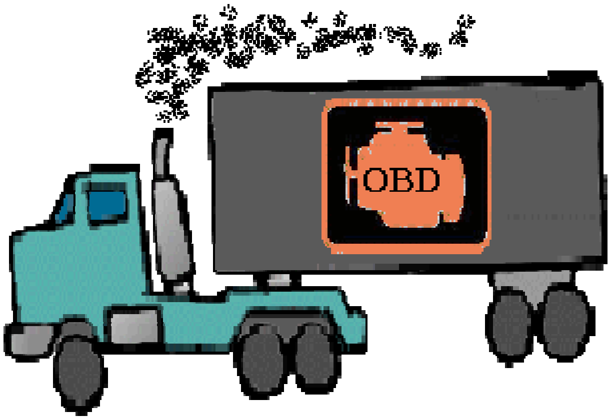
State of California
AIR RESOURCES BOARD

DRAFT PRELIMINARY STAFF REPORT: INITIAL STATEMENT OF REASONS

Technical Status and Malfunction and Diagnostic System Requirements for 2007 and Subsequent Model Year Heavy-Duty Vehicles and Engines (HD OBD)

Date of Release:	July 25, 2003
Scheduled for Public Workshop:	August 13, 2003



Mobile Source Control Division
9528 Telstar Avenue
El Monte, California 91731
www.arb.ca.gov

Table of Contents

I. INTRODUCTION AND BACKGROUND INFORMATION

Introduction

On-board diagnostics (OBD) systems are comprised mainly of software designed into the vehicle's on-board computer to detect emission-control system malfunctions as they occur. This is done by monitoring virtually every component and system that can cause increases in emissions. With a couple of exceptions, no additional hardware is required to perform the monitoring; rather, the powertrain control computer is designed to better evaluate the electronic component signals that are already available, thereby minimizing any added complexity. When an emission-related malfunction is detected, the OBD system alerts the vehicle operator by illuminating the malfunction indicator light (MIL) on the instrument panel. By alerting the operator of malfunctions as they occur, repairs can be sought promptly, which results in fewer emissions over the life of the vehicle. Additionally, the OBD system stores important information, including identifying the faulty component or system and the nature of the fault, which would allow for quick diagnosis and proper repair of the problem by technicians. This helps vehicle owners achieve less expensive repairs and promotes repairs done correctly the first time.

Currently, California regulations require all 1996 and newer passenger cars, light-duty trucks, and medium-duty vehicles and engines to be equipped with OBD systems (referred to as OBD II systems). The Air Resources Board (ARB) first adopted the OBD II regulation (title 13, California Code of Regulations (CCR) section 1968.1) in 1989 and subsequently adopted modifications to this regulation in regular updates to the Board in subsequent years to address manufacturers' implementation concerns, strengthen specific monitoring requirements, and add new monitoring requirements, among other reasons. In 2002, ARB adopted title 13, CCR sections 1968.2 and 1968.5, which established OBD II requirements and an OBD II-specific in-use enforcement protocol, respectively, for 2004 and subsequent model year passenger cars, light-duty trucks, and medium-duty vehicles and engines.

The OBD II requirements serve an important role in helping vehicle manufacturers achieve and maintain low vehicle emissions. Manufacturers are required to improve their emission control system performance and durability in order to meet the very low and near-zero emission standards of the Low Emission Vehicle II program. Since the OBD II program is designed to ensure maximum emission control system performance for the entire life of the vehicles (regardless of mileage), it is able to monitor the low-emission performance of vehicles and ensure that they are performing as required throughout their useful lives and beyond. This is important, since most emission problems occur as vehicles age and accumulate high mileage.

Input from manufacturers, service technicians, pilot Inspection and Maintenance programs, and in-use evaluation programs indicate that the OBD II program is very effective in finding emission problems and facilitating repairs. The United States Environmental Protection Agency (U.S. EPA), in fact, issued a final rule that indicates its confidence in the performance of OBD II systems by requiring states to perform OBD II checks for these newer vehicles and allowing them to be used in lieu of current tailpipe tests in Inspection and Maintenance programs. Overall, ARB staff is pleased with the significant and effective efforts of the automotive industry in implementing the program requirements.

Why Require OBD Systems on Heavy-Duty Vehicles and Engines?

Heavy-duty vehicles are important, and thus prevalent, throughout the country, being used for on-road interstate and intrastate transportation. Given some economic advantages from fuel efficiency, maintenance costs, and durability, diesel engines are employed on the vast majority of the heavy-duty trucks in lieu of gasoline engines. Unfortunately, the emissions emitted from these heavy-duty trucks, especially diesel trucks, are of great concern. Currently, diesel truck emissions account for about 28 percent and 16 percent of the total statewide mobile source oxides of nitrogen (NOx) and particulate matter (PM) emissions, respectively. NOx is a precursor to ozone as well as a lung irritant, while diesel PM is carcinogenic and has been identified as a toxic air contaminant by ARB. While emissions from heavy-duty diesels are of great concern, emissions from heavy-duty gasoline vehicles are also a concern.

As stated previously, OBD systems are required on all 1996 and newer passenger cars, light-duty trucks, and medium-duty vehicles and engines. Presently, however, there are no regulations in California requiring OBD systems on heavy-duty vehicles (i.e., vehicles with a gross vehicle weight rating greater than 14,000 pounds). ARB staff is thus proposing the adoption of a separate OBD regulation (proposed title 13, CCR section 1971) to apply to all heavy-duty Otto-cycle (gasoline) and diesel engines and vehicles.

The reasons for requiring OBD systems on heavy-duty vehicles and engines are analogous to those for requiring OBD II systems on light- and medium-duty vehicles. Like the light- and medium-duty vehicles, the emission standards for heavy-duty vehicles have become increasingly stringent over the years. By 2004, the heavy-duty diesel emission standards for oxides of nitrogen (NOx) and particulate matter (PM) have been reduced by over 60 to 80 percent compared to the standards in 1990. In 2007, both emission standards would be reduced further by 90 percent compared to the 2004 standards. Emission standards for heavy-duty gasoline vehicles and engines are also similarly reduced in 2008. While the adoption of increasingly stringent standards are a step towards meeting California's air quality goals, there must be some assurance that these standards continue to be met in-use, since emission-related malfunctions can cause vehicle emissions to increase well beyond the standards that they are intended to meet. To meet these stringent standards, manufacturers must improve existing emission control technologies as well as utilize new technologies. The technologies include combinations of electronic powertrain and emission controls as well as exhaust aftertreatment components. Accordingly, in order to maintain low emissions throughout the vehicle's life, the durability and performance of these components and systems must be monitored. Additionally, with these changes comes the development of more complex electronic emission control systems, which increases the reliance on computer-based control. Therefore, the diagnosing of malfunctions related to emission-related components and systems becomes more complicated as well. OBD systems would ensure that emission-related malfunctions are quickly detected as well as properly identified and repaired by providing repair technicians with enough information concerning the malfunctioning component and the type of failure present.

Having adopted the first set of OBD II requirements for light- and medium-duty vehicles in 1989, ARB staff has had extensive experience with OBD systems and in developing diagnostic requirements. ARB staff has determined that many of the diagnostic requirements for heavy-duty vehicles would be similar to those of light- and medium-duty vehicles. Specifically, several of the major system and component monitors would be directly calibrated to an emission level correlated to the emission standards (i.e., require a fault to be detected before emissions exceed the standards by a certain amount) while other component monitors (e.g., comprehensive components) would require individual components on the vehicle to be checked for circuit faults and rationality or functionality. Thus, ARB staff has used the OBD II regulation (title 13, CCR section 1968.2) as a base for developing the proposed heavy-duty OBD regulation (title 13, CCR section 1971). For manufacturers concerned about the technical feasibility of meeting the proposed requirements, the staff and industry have identified methods that are expected to be effective in monitoring various emission-related components and systems. In many cases, the staff has identified only one or two potential monitoring strategies for a particular component even though many other equally effective strategies may exist. Further, as past history has usually shown, the staff expects that manufacturers will be quite innovative and may develop even better techniques as the underlying emission control technology evolves.

To provide manufacturers with sufficient leadtime to comply with the new heavy-duty OBD regulation, a phase-in is proposed beginning with the 2007 model year. Manufacturers would be required to phase-in OBD systems on 30% of 2007 model year vehicles, 60% of 2008 model year vehicles, and 100% of 2009 model year vehicles. Staff believes this approach would allow manufacturers to effectively implement comprehensive OBD systems in a timely manner within available resources. Requiring only 30% of the 2007 model year vehicles to meet the OBD standards would allow manufacturers to focus on a limited number of engines and gain in-use experience before applying the OBD system to all engines. That said, staff is aware that it will take a significant amount of resources to upgrade from the current diagnostics typically implemented on engines to an OBD-compliant system and is seeking feedback on staff's proposed implementation schedule or other alternatives that would be equally effective in getting to fully implemented and comprehensive OBD requirements in the same time frame as staff's proposal.

Additionally, staff expects that manufacturers will likely be required to substantially revise the emission control systems on all engines in the 2007 model year to meet the 2007 PM standards as well as some revisions to meet intermediate NOx emission reduction levels. Typically, these modifications will include hardware changes (such as the addition of PM traps) and software modifications (such as EGR flow rates and fuel injection parameters). As such, it would typically be more cost effective and a more efficient use of engineering resources to implement modifications such as an OBD system at the same time. Given the comprehensiveness of the OBD system requirements, though, staff believes it most appropriate to phase-in these systems. Staff is concerned, however, that manufacturers will be implementing new emission control technologies such as PM traps without implementing adequate diagnostics to alert vehicle operators to a problem and provide repair technicians with sufficient information to find and fix emission-related malfunctions. In the past, manufacturers have typically voluntarily implemented some level of diagnostic capability to aid repair technicians, although such diagnostics are often limited to circuit faults of various input and output components. To ensure that manufacturers are planning on providing sufficient diagnostic capability on all 2007 and newer model year vehicles, staff is seeking feedback from engine manufacturers as to the level of diagnostics that they are anticipating to implement on vehicles not subject to OBD requirements in the 2007 model year (e.g., the 70% not required to have OBD under staff's proposal). If manufacturers are not likely to have sufficient diagnostics in place for alerting the vehicle operator to failures of the newer emission control components such as PM traps, staff may also propose some minimum monitoring requirements for vehicles not included in the phase-in beginning in 2007 model year.

What Would the Heavy-Duty OBD Regulation Require?

As with the light- and medium-duty OBD II requirements, for some emission control systems and components, the proposed heavy-duty OBD regulation would require malfunctions to be identified before any problem becomes serious enough to cause vehicle emissions to exceed the standards by a certain amount. This would require manufacturers to correlate component and system performance with emission levels to determine when deterioration of the system or component will cause emissions to exceed 1.5 or 1.75 times the engine emission standards. When this occurs, the proposed regulation would require the diagnostic system to alert the operator to the problem by illuminating the MIL. The malfunction thresholds will be based on the emission standards that the particular engine is certified to, be it an established engine emission standard or a manufacturer-specific family emission limit (FEL) used in accordance with the averaging, banking, and trading program. Components and/or systems whose monitors would be calibrated to the 1.5 or 1.75 times the standard criterion include the fuel system, the exhaust gas recirculation (EGR) system, the catalyst system, and other aftertreatment devices (e.g., particulate matter traps, etc.). It should be noted that staff will consider the use of alternative thresholds during the phase-in period to allow industry to gain experience in properly calibrating them and to limit compliance liability in the early years of implementation.

For the components and systems in which the 1.5 or 1.75 times the standard criterion is not sufficient or cannot easily be applied, the proposed regulation would establish different malfunction criteria to identify emission problems. For example, in addition to having to detect engine misfire before emissions exceed 1.5 or times the standards on gasoline engines, the proposed regulation would require that misfire levels be detected that will cause catalyst damage due to overheating.

Given that diesel and gasoline applications often utilize different emission control technologies or strategies, the proposed regulation would contain several separate monitoring requirements for diesel and gasoline applications. For example, diesel applications would be required to monitor diesel-related emission control technologies such as particulate traps and NOx absorbers/traps, while gasoline applications would be required to monitor gasoline-related technologies such as evaporative emission systems. Additionally, for emission controls common to both diesel and gasoline engines, the proposed regulation would include a section that details monitoring requirements that apply to both diesel and gasoline applications. These include variable valve timing and/or control system monitoring and comprehensive component monitoring.

Regarding evaporative system monitoring for gasoline applications, the 1.5 times the emission standard criterion would not be applicable. The proposed regulation would require the OBD system to detect leaks equivalent or greater in magnitude to a 0.090 inch diameter hole and a 0.030 inch diameter hole. While data from passenger car evaporative system designs show that leaks approaching a 0.020 inch hole begin to rapidly generate excess evaporative emissions (up to 15 times the standard), current monitoring technology for the large tanks typically found on heavy-duty vehicles and serviceability issues do not appear to permit detecting and repairing smaller leaks.

The 1.5 times the emission standard criterion would also not be applicable to the monitoring of electronic powertrain components that can cause emissions to increase when malfunctioning, but generally to less than 1.5 times the standard. The proposed regulation would require such components to be monitored for proper function on both diesel and gasoline applications. For example, for components that provide input to the on-board computer, the OBD system would be required to monitor for out-of-range values (generally open or short circuit malfunctions) and input values that are not reasonable based on other information available to the computer (e.g., sensor readings that are stuck at a particular value or biased significantly from the correct value). For output components that receive commands from the on-board computer, the OBD system would be required to monitor for proper function in response to these commands (e.g., the system verifies that a valve actually opens and closes when commanded to do so). Monitoring of all such components is important because, while a single malfunction of one of these components may not cause an exceedance of the emission standards, multiple failures could synergistically cause high in-use emissions.[1] Further, the OBD system relies on many of these components to perform monitoring of the more critical emission control devices. Therefore, a malfunction of one of these input or output

---

[1] The proposed regulation would only require detection of any single component failure that can affect emissions rather than detection of every combination of multiple component degradations that can cause emissions to exceed the standards, due to the overwhelming time and cost resources that would be required to evaluate the latter.

components, if undetected, could lead to incorrect diagnosis of emission malfunctions or even prevent the OBD system from checking for malfunctions.

In addition to malfunction detection requirements, the proposed regulation would require that diagnostic repair information be provided to aid service technicians in isolating and fixing detected malfunctions. For each malfunction detected, a specific fault code would be stored identifying the area and nature of the malfunction (e.g., a mass air flow sensor with an inappropriately high reading). The OBD system would also provide technicians with access to current engine operating conditions such as engine speed, engine load, coolant temperature, etc. The OBD system would even store the operating conditions that exist at the time a malfunction is detected. All of this information would be accessed with the use of a generic scan tool (i.e., one tool that can access all makes and models of vehicles), and would help assist the technician in accurately diagnosing and repairing problems.

What Do the Federal OBD Regulations Require?

Currently, the U.S. EPA only has OBD requirements for light-duty vehicles and trucks and for federally defined "heavy-duty" vehicles and engines with a gross vehicle weight rating (GVWR) between 8,500 to 14,000 pounds. These are the same categories of vehicles covered by ARB's OBD II regulations which apply to light- and medium-duty vehicles (where medium-duty is defined in California as the 8,500 to 14,000 pound GVWR range). Presently, like ARB, the U.S. EPA does not have OBD requirements for vehicles and engines above 14,000 pounds, which is the weight range for California’s “heavy-duty” class. ARB staff and the U.S. EPA staff have been discussing the heavy-duty OBD requirements and the U.S. EPA staff has indicated its intent to propose and adopt an OBD regulation for heavy-duty vehicles and engines over 14,000 pounds. While the timing and specifics of such a federal OBD program have not yet been proposed, U.S. EPA staff have indicated a strong interest in working with ARB, the heavy-duty industry, and other stakeholders to develop harmonized ARB and federal OBD programs.

Heavy-Duty OBD and Inspection and Maintenance

As stated before, one of the main purposes of OBD is to keep emissions low for the entire life of the vehicle. In order to achieve this, an inspection and maintenance (I/M) program is needed to ensure that emission-related malfunctions that are being detected are repaired in a reasonable timeframe. Before the OBD II system check was recently incorporated into the I/M program for light- and medium-duty vehicles, the current California I/M program (i.e., “Smog Check”) relied primarily on tailpipe testing to identify vehicles with emission-related malfunctions. When these vehicles were identified, repair technicians then were required to diagnose the cause of the emission failure and performed the necessary repairs. The effectiveness of the repairs in bringing the vehicle back into compliance can be known with certainty only when the vehicle again undergoes a tailpipe test. The incorporation of OBD system checks greatly simplifies and improves this process. Instead of measuring tailpipe emissions directly once every two years, the technician would only have to check the OBD system. If the MIL were not illuminated, nor any fault codes stored, there would be considerable assurance that the vehicle is not emitting excessive emissions (i.e., virtually all the potential sources for an emission problem are operating without defect). In addition, an OBD-I/M check can catch faults of emission-related components and systems that cannot otherwise be checked during a tailpipe-only I/M test, such as cold start emission reduction devices or fuel system malfunctions that occur exclusively outside of the I/M driving conditions.

Currently, ARB has two enforcement programs that target excessive smoke emissions from heavy-duty trucks and buses. The first program, the Heavy-Duty Vehicle Inspection Program (HDVIP), consists of ARB inspectors conducting smoke opacity snap-acceleration tests and visual tamper inspections (where inspectors look under the hood for visible signs of tampering) on vehicles at various roadside locations, such as California Highway Patrol (CHP) weigh stations. The second program, the Periodic Smoke Inspection Program (PSIP), requires owners of heavy-duty truck and bus fleets to perform and maintain records of annual self-inspections of their own vehicles. These also consist of smoke opacity snap-acceleration tests and tamper inspections of the vehicles. These current programs, however, focus mostly on reductions of hydrocarbons and particulate matter (which smoke is mostly composed of) and reflect how the vehicle is performing only at the moment of inspection (as opposed to continuously on the road) and under the conditions tested (i.e., snap acceleration). ARB is currently considering several approaches to improving the inspection program, including the development of a NOx screening test, which incorporates a dynamometer test, and a manufacturer-run in-use compliance program which tests the engines under the Not-to-Exceed (NTE) test. Whatever the final program turns out to be, the incorporation of OBD checks into this program would enable a more thorough inspection by continuously monitoring the entire emission control system while the vehicle is in-use and providing emission-related information at the time of inspection. Further, a heavy-duty vehicle operator will know before the inspection whether the vehicle will pass or fail based on the presence or absence of the MIL warning light. This can eliminate uncertainty on the vehicle operator’s part in wondering whether or not the truck will fail the inspection and can lead to reduced risk of citations or NOVs.

Enforcement for Heavy-Duty OBD

Under the OBD II requirements for light- and medium-duty vehicles, ARB has adopted a separate, stand-alone enforcement regulation for OBD II systems. For heavy-duty OBD, staff anticipates doing the same but does not have a staff proposal at this time. Staff anticipates adopting enforcement regulations specific to heavy-duty OBD compliance under a separate rulemaking (or during a biennial review of this regulation) prior to implementation of OBD systems in the 2007 and subsequent model years. Accordingly, the staff report and proposed regulation do not contain specific enforcement provisions.

Staff has spent some time considering the uniqueness of the heavy-duty industry with separate engine and component suppliers and the difficulties that can present in enforcement. The heavy-duty industry is similar in some aspects to other regulated industries or products such as marine engines, off-road engines, and incomplete vehicles, and ARB has experience in dealing with complicated supplier, manufacturer, importer, and dealer relationships both in certification as well as in enforcement. With OBD being fairly complicated and sensitive to interaction from the various components installed with an engine into an end vehicle, staff expects these relationships may become even further complicated. In the end, however, the vast majority of the proposed OBD requirements would apply directly to the engine or its associated emission controls, and the engine manufacturer would have complete responsibility to ensure those requirements are met. Given the central role the engine and engine control unit would play in the OBD system, the staff anticipates proposing that the party certifying the engine and OBD system (typically, the engine manufacturer) also be the responsible party for in-use compliance and enforcement actions. In this role, the certifying party would be ARB’s sole point of contact for noncompliances identified during in-use or enforcement testing. ARB would not take on the role of going beyond identifying the noncompliance to determine the ultimate party responsible for the noncompliance (e.g., engine manufacturer, coach builder, other supplier). In cases where remedial action would be required (e.g., recall), the certifying party would take on the responsibility for arranging to bring the vehicles back into compliance. To protect themselves, it is expected that engine manufacturers will require engine purchasers to sign indemnity clauses or other agreements to abide by the build specifications applicable to the engine and to bear ultimate financial responsibility for noncompliances caused by the engine purchaser. Given that heavy-duty engines are already subject to various emission requirements including engine emission standards, labels, certification, etc., engine manufacturers currently do impose restrictions on engine purchasers to ensure the engines do not deviate from their certified configuration when installed. As such, it is likely that the engine manufacturers already require such agreements from engine purchasers to protect themselves. Further, if not done for emission certification purposes, certainly the engine manufacturer would have similar protections in place for items that cause premature engine component failure or warranty cost. Engine manufacturers would certainly have provisions in place to safeguard them from warranty costs or problems caused by the engine purchaser (e.g., insufficient engine cooling system installed resulting in overheating and premature engine damage).

II. GENERAL MONITORING REQUIREMENTS

A. Monitoring Conditions

As stated previously, the intent of the OBD system is to detect malfunctions of the emission control system while the vehicle is being operated. To best achieve this, the OBD monitors would have to be designed to run during conditions routinely encountered by drivers of heavy-duty vehicles. If OBD monitors were designed to run only during extreme (i.e., rarely encountered) conditions, emission-related malfunctions would rarely, if ever, be detected, which could lead to unnecessary excess emissions, defeating the purpose of OBD. While manufacturers may limit the conditions under which certain monitors would run to ensure effective monitoring of the component or system, it is important that these conditions are not so restrictive such that monitoring would rarely occur during real-world driving. Given the wide variety of operating patterns used by various aspects of the heavy-duty industry such as refuse trucks and transit buses to line-haul applications, it is especially imperative that heavy-duty manufacturers design monitors to run under as broad a range of driving conditions as possible.

The staff is proposing some guidelines that manufacturers would have to follow when developing their OBD monitors. The proposed regulation would require the monitors to run during conditions that (1) are technically necessary to ensure robust detection of malfunctions, (2) ensure monitoring will occur during normal vehicle operation, and (3) ensure monitoring will occur during the heavy-duty Federal Test Procedure (FTP) transient cycle. ARB would determine if the monitoring conditions proposed by the manufacturer for each monitor abide by these requirements. The staff is also proposing requirements that would measure the real world monitoring performance of many OBD monitors (see section VI. of the Staff Report for more details). These proposed requirements would assist the staff in determining if the monitoring conditions are sufficiently broad for frequent monitoring during normal operation.

The proposed regulation would require each monitor to run at least once per driving cycle in which the applicable monitoring conditions are met. The proposal would also require certain monitors to run continuously throughout the driving cycle. These include a few major monitors (e.g., fuel system monitor) and most circuit monitors. While a basic definition of a driving cycle (e.g., from ignition key on and engine start up to engine shut-off) has been sufficient for passenger cars, the driving habits of many types of vehicles in the heavy-duty industry dictate an alternate definition. Specifically, many heavy-duty operators will start the engine and leave it running for an entire day or, in some cases, several days or weeks, continuously. As such, the staff is proposing a modification to the definition to define any period of continuous engine-on operation of four hours to be considered a complete driving cycle and to trigger the start of a new driving cycle. Thus, monitors that are required to run once per driving cycle would be reset to run again (in the same key-on engine start or trip) once the engine has been operated for over four hours continuously. This will avoid an unnecessary delay in detection of malfunctions simply because the heavy-duty vehicle operator has elected to leave the vehicle running continuously for an entire day or days at a time.

B. MIL and Fault Code Requirements

When an emission-related malfunction is detected by the OBD system, there must be some indication to the driver of the presence of this fault so that it can be repaired as soon as possible. In the event of a malfunction, the proposed regulation would require the manufacturer to store a fault code identifying the nature of the malfunction and illuminate the MIL to alert the driver of the presence of the fault.

The staff is proposing to standardize the location and image of the MIL. Generally, the MIL would be required to be located on the driver's side instrument panel and, when illuminated, to display the phrase "Check Powertrain." The proposed regulation would not allow manufacturers to use the MIL for any other purpose other than those related to OBD (i.e., those purposes specified in the proposed regulation). Manufacturers have expressed their desire to utilize existing engine and transmission-specific lights on the dashboard to indicate both emission-related and non-emission-related malfunctions. While the proposed regulation would not prohibit the illumination of the current lights when engine or transmission-related problems occur, the staff believes that a separate, OBD-specific light must also be illuminated in conjunction with the other light when the problem is an emission-related fault. This would significantly help the future incorporation of OBD checks into a heavy-duty I/M program, where vehicles would "fail" due to the presence of an emission-related fault. If a vehicle did not have an OBD-specific light, technicians would individually have to determine whether the illumination of an engine or transmission-related light was emission-related or not. Past experience in California with warning lights that combine emission-related and non emission-related faults has shown great discrepancies in interpretation by individual technicians and has resulted in unnecessary confusion and difficulty.

Generally, a manufacturer would be allowed sufficient time to be certain that a fault truly exists before illuminating the MIL. It is to the advantage of neither the manufacturer, the vehicle operator, the service technician, nor ARB for the MIL to be illuminated spuriously or when a repairable malfunction is not truly present. Thus, for most OBD monitoring strategies, manufacturers would be expected to illuminate the MIL only after the same malfunction has occurred on two separate driving cycles. The first time a malfunction is detected, a "pending" fault code identifying the suspected failing component or system would be stored in the on-board computer. If the same malfunction is again detected the next time the vehicle is operated, the MIL would be illuminated and a "confirmed" fault code would be stored. A technician would use the "confirmed" fault code to determine what system or component has failed, what the exact problem is, and how to fix the problem.

In order to minimize the possibility of the MIL cycling on and off, the staff is proposing specific requirements to prevent the MIL from extinguishing too easily, thereby improving owner confidence in the diagnostic system. Specifically, once the MIL is illuminated, the MIL would not be extinguished unless the monitor related to the malfunction runs on three subsequent successive driving cycles (or trips) and no longer detects a malfunction present. Thus, in the case of an intermittent fault, the malfunction would need to be present for "two-trips-in-a-row" to illuminate the MIL and subsequently, it would have to be gone for "three-trips-in-a-row" to extinguish the MIL.

The staff is also proposing specific requirements to retain fault code information for a longer period of time to aid repair technicians. A confirmed fault code would be erased only if the identified malfunction has not been again detected in at least 40 engine warm-up cycles and the MIL is not presently illuminated for that malfunction. This provides added benefit to the vehicle operator and repair technicians by allowing access to fault information even if the MIL is not currently illuminated.

There may be malfunctions of the MIL itself that would prevent the illumination of the MIL. While a technician or inspector can still determine the status of the MIL (i.e., commanded "on" or "off") by reading electronic information available through a scan tool, if the MIL malfunctions, there would be no indication to the driver of any emission-related faults should they occur. Unidentified malfunctions may cause excess emissions to be emitted from the vehicle and may even cause subsequent deterioration or failure of other components or systems without the driver's knowledge. In order to prevent this, the manufacturer must ensure that the MIL is functioning properly. Thus, the staff is proposing two requirements to check the functionality of the MIL. First, the MIL would be required to illuminate for a minimum of 15 to 20 seconds when the vehicle is in the key-on, engine-off position. This allows an inspector, technician, or vehicle operator to ensure the MIL is capable of illuminating by simply cycling the key on. While the MIL would be physically illuminated during this functional check, the MIL command status would be required to indicate “off” during this check (unless the MIL was currently being commanded “on” for a detected malfunction).

The second functional check requirement would be for the manufacturer to perform a circuit continuity check of the electrical circuit that is used to illuminate the MIL to verify the circuit is not shorted or open (e.g., burnt out bulb). While there will not be an ability to illuminate the MIL when such a malfunction is detected, the electronically readable MIL command status in the on-board computer would be changed from commanded "off" to commanded "on". This again greatly simplifies the I/M program and allows the inspection to be completely automated instead of a combination of pass/fail criteria based on electronic information obtained through a scan tool plus manually inputted visual results entered by the inspector. Feedback from passenger car I/M programs has indicated that the current visual bulb check performed by inspectors is excessively subject to error and has resulted in numerous vehicles being falsely failed or passed. By requiring monitoring of the circuit itself, the entire pass/fail criteria of an I/M program could be determined by the electronic information available through a scan tool, thus better facilitating quick and effective inspections and minimizing the chance for manually-entered errors.

While most monitors are expected to be designed as “two-in-a-row" driving cycle monitors (i.e., illuminate the MIL and store a confirmed fault code in two driving cycles), the proposed regulation would allow manufacturers to seek ARB approval to use “statistical algorithms” in their monitoring strategies, which generally analyze diagnostic information collected over more than two driving cycles. For ARB approval of the alternate statistical MIL illumination and fault code storage protocol, the manufacturer would have to submit information demonstrating that the alternate protocol is able to evaluate the system performance and detect malfunctions in an effective and timely manner equivalent to the standard “two-in-a-row” protocol. Additionally, the staff is proposing to limit the "run length" of these alternate strategies to six on average. With alternate strategies, even with a limit of six on average, some malfunctions would not be detected until 10 or more trips due to the variation associated with the algorithm. Should the limit be increased, the variation would also increase, causing malfunction detections to be delayed until 20 or more trips in some cases, which would not be reasonably timely nor equivalent to the standard MIL illumination protocol.

The proposed regulation would also require manufacturers to illuminate the MIL when the vehicle enters a default mode of operation (e.g., over-temperature management strategies) that can affect emissions or the performance of the OBD system. However, manufacturers would be exempt from illuminating the MIL if either of the following occurs: (1) the strategy causes an overt indication (e.g., illumination of a warning light such as a "hot light") such that the driver is certain to respond and have the problem corrected, or (2) the default strategy is an auxiliary emission control device (AECD)[2] strategy that is properly activated due to the occurrence of conditions that have been approved by the Executive Officer. The manufacturer would be required to submit documentation supporting the exemption for ARB approval.

Additional detailed technical requirements pertaining to fault codes are provided in section VII. (Standardization Requirements) of the Staff Report.

III. TECHNICAL STATUS AND PROPOSED MONITORING SYSTEM REQUIREMENTS FOR DIESEL/COMPRESSION-IGNITION ENGINES

A. FUEL SYSTEM MONITORING

Background

An important component in emission control is the fuel system. Proper delivery of fuel (in both quantity and injection timing) can play a crucial role in maintaining low engine-out emissions. The performance of the fuel system is also critical for aftertreatment device control strategies. As such, thorough monitoring of the fuel system is an essential element in an OBD system. The fuel system is primarily comprised of a fuel pump, fuel pressure control device, and fuel injectors. Additionally, the fuel system generally has sophisticated control strategies that utilize one or more feedback sensors to ensure the proper amount of fuel is being delivered to the cylinders. While gasoline engines have undergone relatively minor hardware changes (but substantial fine-tuning in the control strategy and feedback inputs), diesel engines have more recently undergone substantial changes to the fuel system hardware and now incorporate more refined control strategies and feedback inputs.

For diesel engines, a substantial change has occurred in recent years as manufacturers have transitioned to new high-pressure fuel systems. One of the most widely used is a "common-rail" fuel injection system, which is generally comprised of a high-pressure fuel pump, a fuel rail pressure sensor, a common fuel rail that feeds all

---

[2] AECDs are discussed in more detail in section V.E. of the Staff Report.

injectors, individual fuel injectors that directly inject fuel into each cylinder, and a closed-loop feedback system that uses the fuel rail pressure sensor to achieve the commanded fuel rail pressure. Unlike older style fuel systems where fuel pressure was mechanically linked to engine speed (and thus, varied from low to high as engine speed increased), common-rail systems are capable of controlling to any desired fuel pressure independent of engine speed. This increase in fuel pressure control allows greater precision relative to fuel quantity and fuel injection timing that provides engine manufacturers with tremendous flexibility in optimizing the performance and emission characteristics of the engine. The ability of the system to generate high pressure independent of engine speed also improves fuel delivery at low engine speeds.

While most diesel engine manufacturers use common-rail systems, some use improved unit injector systems, which are a result of continuous design improvements to a conventional unit injector system. In this system, fuel pressure is generated within the injector itself rather than via an electric fuel pump in a common-rail system. Typically, the injector is a combination electrically and hydraulically-controlled unit. A high-pressure oil pump is used to deliver oil to the injector, which in turn activates a plunger in the injector to increase the fuel pressure to the desired level. Earlier versions of unit injector systems were able to achieve some of the advantages of common-rail systems (e.g., high fuel pressures) but still had limitations on the pressure that they could build based on engine speed. Further, the fuel pressure was a function of engine speed and could not be modified to a lower or higher pressure at a given engine speed. Newer design iterations have created an injector with extra valves that allow the system to deliver higher or lower pressures at a given engine speed. Thus, while there is still some dependence on engine speed for the fuel pressure, it is largely adjustable and can achieve much of the same fuel pressure range a common-rail system is capable of achieving.

Precise control of the fuel injection timing is also crucial for optimal engine and emission performance. As injection timing is advanced (i.e., fuel injection occurs earlier), hydrocarbon (HC) emissions and fuel consumption are minimized but oxides of nitrogen (NOx) emissions are increased. As injection timing is retarded (i.e., fuel injection occurs later), NOx emissions can be dramatically reduced but HC emissions, particulate matter (PM) emissions, and fuel consumption increase.[3] Engine manufacturers must continually optimize the system to deliver the desired fuel quantity precisely at the right time.

The common-rail system or improved unit injector system also provides engine manufacturers with the ability to separate a single fuel injection event into discrete events such as pilot (or pre) injection, main injection, and post injection. A system using a pilot injection and a main injection instead of a single injection event has been shown to generate a 16 percent reduction in NOx emissions[4] in addition to providing a

---

[3] Tschoeke, H., 1999. “The Fuel Injection System”, SAE International-Diesel Engine Design Academy, Ypsilanti, MI, July 1999 (www.dieselnet.com, Diesel Fuel Injection, Advanced Injection Systems and Their Control).

[4] Tullis, S., Greeves G., 1996. “Improving NOx Versus BSFC with EUI 200 Using EGR and Pilot

substantial reduction in engine noise. Another study has shown that the use of pilot injection versus no pilot injection can lead to a 20 percent reduction in PM emissions and a five percent reduction in fuel usage at a similar NOx level.[5]

Lastly, the high pressures and near infinite control in a common-rail or improved unit injector system begin to open the door for manufacturers to modify the fuel injection pressure during a fuel injection event which results in different fuel quantity injection rate profiles or "shapes." "Rate-shaping," as it is commonly known, allows manufacturers to begin a fuel injection event with a set injection rate and end the injection at a different injection rate. This could be used to progressively increase the fuel quantity during the injection event and has been shown to lower NOx emissions in laboratory settings.[6]

Given these various aspects of common-rail systems and improved unit injector systems, malfunctions or deterioration that would affect the fuel pressure control, injection timing, pilot/main/post injection timing or quantity, or ability to accurately perform rate-shaping could lead to substantial increases in emissions (primarily NOx or PM), often times with an associated change in fuel consumption.

Proposed Monitoring Requirements

For diesel engines, the staff is proposing several monitoring requirements to verify the overall fuel system's ability to meet the emission standards and to verify that individual aspects or capabilities of the system are properly functioning.

Fuel Pressure Monitoring

The staff is proposing monitoring requirements that continuously verify the system is able to control to the desired fuel pressure. Manufacturers would be required to indicate a malfunction when the system can no longer control the fuel system pressure such that emissions remain below the 1.5 times the applicable standards.

Fuel Injection Quantity Monitoring

The staff is proposing monitoring requirements that verify the fuel system is able to accurately deliver the proper quantity of fuel required for each injection. Manufacturers would be required to detect a fault when the system is unable to accurately deliver the desired fuel quantity such that emissions exceed 1.5 times the applicable standards. If no failure can cause emissions to exceed 1.5 times the applicable standards, than the OBD system shall detect a fault when the fuel injection system has reached its control authority limits and can no longer increase or decrease

---

Injection for Heavy-Duty Diesel Engines", SAE 960843 (www.dieselnet.com, Diesel Fuel Injection, Common-Rail Fuel Injection).

[5] Greeves, G., Tullis, S., and Barker, B., 2003, "Advanced Two-Actuator EUI and Emission Reduction for Heavy-Duty Diesel Engines", SAE 2003-01-0698.

[6] Erlach, H., Chmela, F., Cartellieri, W., and Herzog, P., 1996. "Pressure-Modulated Injection of an HD Diesel Engine", Off-Highway Engineering, SAE International, April 1996 (www.dieselnet.com, Advanced Diesel Engine Technologies, Advanced Technologies: Fuel Injection and Combustion).

the commanded injection quantity to achieve the desired fuel injection quantity. Malfunctions or deterioration of the system such as injector deposits that restrict flow can result in individual cylinder variations that alter the injection quantity or injection profile and lead to increases in emissions. Unlike gasoline engines, however, there is no feedback system on diesels that directly verifies the proper fuel quantity. While large decreases in the fuel injection quantity can be noticed by the vehicle operator (e.g., reduction in power output of the engine), small changes may substantially impact emissions by reducing the ability of the system to accurately deliver fuel (through separate pilot, main, or post injections or timing) before being noticeable to the driver. As an example, pilot injections typically represent only a few percent (e.g., four to five percent) of the total fuel injected for an individual cylinder fueling event but can have a disproportional impact on increases in NOx emissions (e.g., +16 percent). Deterioration or other malfunctions could affect the ability of the system to accurately deliver the pilot injection yet still achieve acceptable performance to the vehicle operator.

Multiple Fuel Injection Performance Monitoring

The staff is also proposing requirements to verify the ability of the fuel injection system to deliver multiple injections during a single cylinder firing event. With the ability of these systems to divide the injection into several discrete events such as a pilot, main, and even post injection, engine manufacturers are able to optimize injection to reduce various emissions while still maintaining good fuel economy and minimizing engine noise. If the system were no longer able to accurately deliver multiple injections in a very short time period, the emission benefits of the multiple injection strategy would be lost. In some cases the time between the end of the first injection and the start of the second injection could be as little as two to three degrees crank angle or approximately 0.0002 seconds at 2000 revolutions-per-minute (rpm). Malfunctions or deterioration that cause an injector to open or close slower or deposits which decrease the rate of fuel being delivered could result in unacceptable multiple injection performance.

Fuel Injection Timing Monitoring

Lastly, the staff is proposing that manufacturers implement monitoring to verify that fuel injection timing is correct; that is, that fuel is injected at the precise time that it is commanded to happen. Small changes in fuel timing (advance or retard) can have significant impacts on emissions. If the injector were to open too soon (due to a deteriorated needle lift return spring, etc.), fuel would be injected too soon and potentially at a lower than desired fuel pressure. If the injector were to be delayed in opening (due to restrictions in the injector body passages, etc.), fuel would be injected later than desired and potentially at a higher fuel pressure than desired. As such, manufacturers would be required to verify that the fuel injection occurs within a manufacturer-specified tolerance of the commanded fuel timing point and indicate a malfunction prior to emissions exceeding 1.5 times any of the applicable standards.

Technical Feasibility of Proposed Monitoring Requirements

For diesel engines, under the light- and medium-duty OBD II requirements, a few passenger cars and several medium-duty applications utilizing diesel engines have been monitoring the fuel system components since the 1997 model year. Recently, this has included vehicles using common-rail fuel injection and improved unit injector systems, the same new technology expected to be used throughout the heavy-duty industry. For some aspects of these high-pressure fuel systems, however, the monitoring proposed by the staff for heavy-duty diesel engines does extend beyond existing medium-duty applications.

Fuel Pressure Monitoring

The first monitoring requirement proposed by the staff is to identify malfunctions that prevent the system from controlling to the desired fuel pressure. Manufacturers control fuel pressure by using a closed-loop feedback algorithm that allows them to increase or decrease fuel pressure until the fuel pressure sensor indicates they have achieved the desired fuel pressure. For the common-rail systems currently certified as medium-duty vehicles, the manufacturers are indeed continuously monitoring the actual fuel system pressure sensed by a fuel rail pressure sensor, comparing it to the target fuel system pressure stored in a software table or calculated by an algorithm inside the on-board computer, and indicating a fault if too large of an error exists between the two. The error limits are established by engine dynamometer emission tests to ensure a malfunction will be detected before emissions exceed 1.5 times the applicable emission standards. In some cases, manufacturers have developed separate strategies that can identify small errors over a long period of time versus large errors over a short period of time. In other cases, one strategy is capable of detecting both types of malfunctions at the appropriate level. In cases where no fuel pressure error can generate a large enough emission increase to exceed 1.5 times any of the applicable standards, manufacturers are required to set the threshold at their control limits (e.g., when they reach a point where they can no longer increase or decrease fuel pressure to achieve the desired fuel pressure). This monitoring requirement has been demonstrated as technically feasible by the several medium-duty applications that already meet this requirement. Further, the nature of a closed-loop algorithm is that such a system is inherently capable of being monitored because it simply requires analysis of the same closed-loop feedback parameter being used by the system for control purposes.

Fuel Injection Quantity Monitoring

The second diesel fuel system monitoring requirement being proposed is that the system verify that the proper quantity of fuel is being injected. Again, manufacturers would be required to establish the malfunction criteria by engine dynamometer emission tests to ensure a malfunction will be detected before emissions exceed 1.5 times the applicable emission standards. In cases where no fuel quantity error can generate a large enough emission increase to exceed 1.5 times any of the applicable standards, manufacturers are required to set the threshold at their control limits (e.g., when they reach a point where they can no longer increase or decrease fuel quantity to achieve the desired fuel quantity).

As there is no overall feedback sensor indicating that the proper mass of fuel was injected, this monitoring is more difficult. One manufacturer, however, has identified a strategy currently being used that verifies the injection quantity under very specific engine operating conditions and appears to be capable of determining that the system is accurately delivering the desired fuel quantity. This strategy entails intrusive operation of the fuel injection system during a deceleration event where fuel injection is normally shut off (e.g., coasting or braking from a higher vehicle speed down to a low speed or a stop). During the deceleration, fuel injection to a single cylinder is turned back on to deliver a very small amount of fuel. Typically, the amount of fuel would be smaller than, or perhaps comparable to, the amount of fuel injected during a pilot or pre injection. If the fuel injection system is working correctly, that known injected fuel quantity will generate a known increase in fluctuations (accelerations) of the crankshaft that can be measured by the crankshaft position sensor. If too little fuel is delivered, the measured crankshaft acceleration will be smaller than expected. If too much fuel is delivered, the measured crankshaft acceleration will be larger than expected. This process can even be used to “balance” out each cylinder or correct for system tolerances or deterioration by modifying the commanded injection quantity until it produces the desired crankshaft acceleration and applying a correction or adaptive term to that cylinder to compensate future injections of that cylinder to the desired nominal amount. Each cylinder can, in turn, be cycled through this process and a separate analysis can be made for the performance of the fuel injection system for each cylinder. Even if this procedure requires only one cylinder be tested per revolution (to eliminate any change in engine operation or output that would be noticeable to the driver) and requires each cylinder to be tested on four separate revolutions, this process would only take two seconds for a six cylinder engine decelerating through 1500 rpm.

The crankshaft position sensor is commonly used to identify the precise position of the piston relative to the intake and exhaust valves to allow for very accurate fuel injection timing control and, as such, has sufficient resolution and data sampling within the on-board computer to be able to measure such crankshaft accelerations. Further, in addition to the current use of this strategy by a medium-duty diesel engine manufacturer, a nearly identical crankshaft fluctuation technique has been commonly used on medium-duty diesel engines during idle conditions to determine if individual cylinders are misfiring since the 1997 model year.

Another technique that may be used to achieve the same monitoring capability is some variation on the current cylinder balance tests used by many manufacturers to improve idle quality. In such strategies, fueling to individual cylinders is increased, decreased, or shut off to determine if the cylinder is contributing an equal share to the output of the engine. This strategy again relies on changes in crankshaft/engine speed to measure the individual cylinder’s contribution relative to known good values and/or the other cylinders. Such an approach seems viable to effectively determine the fuel injection quantity is correct for each cylinder but has the disadvantage of not necessarily being able to verify the system is able to deliver small amounts of fuel precisely (such as those commanded during a pilot injection).

Multiple Fuel Injection Performance Monitoring

A similar, or even the same, technique could potentially be used to meet the last two proposed requirements for fuel system monitoring. Specifically, the staff is proposing that the OBD system verify the system is capable of delivering multiple injections during a single cylinder event (if such a strategy is used by the engine) and that fuel was injected at the proper time (usually a function of crank angle). For the first part, the ability of the fuel system to deliver the fuel in multiple injections rather than one continuous injection is crucial to proper emission performance. In following the same technique identified above where fuel is intrusively injected during a deceleration event, a manufacturer could add a function where a “double shot” of fuel is injected after the cylinder balance algorithm is carried out. That is, after the specific injector has been compensated to deliver the amount of fuel necessary to generate the desired crankshaft acceleration, the injector could be commanded to deliver two injections of the same size during a single firing event on a subsequent revolution. If the system is working properly, the two injections should generate a higher crankshaft acceleration than a single injector and can be compared to a pre-determined threshold to determine if the system truly delivered the correct amount of fuel over the multiple injections. If the system is unable to either end the first injection fast enough or begin the second injection soon enough, the measured crankshaft acceleration would be different. In this manner, the ability of the system to deliver rapid multiple injections during a single cylinder firing event would be verified.

Fuel Injection Timing Monitoring

In the same manner, fuel injection timing could likely be verified. By monitoring the crankshaft speed fluctuation and, most notably, the time at which such fluctuation begins, ends, or reaches a peak, the OBD system could compare the time to the commanded fuel injection timing point and verify the fluctuation occurred within an acceptable time delay from the commanded fuel injection. If the system was working improperly and actual fuel injection was delayed relative to when it was commanded, the corresponding crankshaft speed fluctuation would also be delayed and result in a longer than acceptable time period between commanded fuel injection timing and crankshaft speed fluctuation.

B. MISFIRE MONITORING

Background

Misfire, the lack of combustion in the cylinder, causes increased engine-out hydrocarbon emissions. On gasoline engines, misfire is due to absence of spark, poor fuel metering, and poor compression. Further, misfire can be intermittent on gasoline engines (e.g., the misfire only occurs under certain engine speeds or loads).

Consequently, the existing light- and medium-duty OBD II regulation requires continuous monitoring for misfire malfunctions on gasoline engines. However, for diesel engines, manufacturers have maintained that misfire only occurs due to poor compression (e.g., worn valves or piston rings, improper injector or glow plug seating), and when poor compression results in a misfiring cylinder, the cylinder will misfire under all operating conditions. Accordingly, the existing light- and medium-duty OBD II regulation does not require continuous monitoring for misfire malfunctions on diesel engines.

Proposed Monitoring Requirements

For diesel engines, the staff is proposing to require manufacturers to monitor for engine misfire that occur continuously in one or more cylinders during idle conditions. Additionally, to the extent possible, manufacturers would be required to identify the misfiring cylinder or indicate if multiple cylinder misfiring is occurring (through the storage of the appropriate fault codes). The proposed regulation would require misfire monitoring to occur at least once per drive cycle in which the monitoring conditions (i.e., idle conditions) are met. The proposed regulation would not allow the idle period under which misfire monitoring is to occur to require more than 15 seconds continuously to collect data, nor would it allow more than 1000 continuous engine revolutions of data to make a decision.

The proposed monitoring requirements for misfire monitoring are identical to the requirement for light- and medium-duty diesel vehicles and are based on the premise that a misfiring diesel engine always misfires, as the engine manufacturers have asserted. However, the staff is concerned that real world malfunctions that cause misfires on diesel engines can occur intermittently or only during off-idle conditions, contrary to manufacturers' assessment. The staff will continue to investigate the possibility of these misfires but currently does not have sufficient information or data at this time to thoroughly validate these concerns. As additional information becomes available for future Board reviews of the HD OBD regulation, the staff may propose a more comprehensive requirement.

Technical Feasibility of Proposed Monitoring Requirements

Diesel engines certified under the light- and medium-duty OBD II requirements have been monitoring for misfire since the 1998 model year. The monitoring requirements proposed by staff for heavy-duty diesel engines are identical to those of current medium-duty diesel applications. The technical feasibility has clearly been demonstrated for these packages.

C. EXHAUST GAS RECIRCULATION (EGR) SYSTEM MONITORING

Background

Since the 1980's, diesel engine NOx emissions have dropped from an uncontrolled level of 15 grams per horsepower-hour (g/hp-hr) to less than four g/hp-hr through the application of advanced technologies. These include turbocharging, charge air cooling, and electronic fuel injection (replacing mechanical systems). In addition, advanced turbocharger systems now provide quick boost response, variable boost pressure, and variable exhaust back pressure to minimize emissions while maximizing fuel economy.

Exhaust gas recirculation (EGR) systems are currently being used to complement these advanced fuel injection and turbocharger systems to meet NOx levels of approximately two g/hp-hr (the 2004 standard is 2.5 g/hp-hr NMHC+NOx with a 0.5 g/hp-hr NMHC cap). Some systems also use an EGR cooler to further reduce NOx emissions. While NOx control technologies have evolved and been refined on gasoline engines over the last 30 years, they have not been readily adapted to diesel engines. However, as light- and medium-duty diesel engines have been subject to increasingly more stringent emission standards, EGR systems are becoming more commonplace and will likely be a key emission control component on future heavy-duty diesel engines. In fact, most heavy-duty diesel engines certified for the 2002 model year are equipped with EGR. The staff anticipates that EGR usage will continue as even more stringent heavy-duty diesel standards are phased-in in the near future.

NOx emissions are formed under high combustion chamber temperature and pressure conditions. EGR reduces NOx emissions through two mechanisms. First, recirculated exhaust gas dilutes the intake air (i.e., oxygen and nitrogen are displaced with relatively non-reactive exhaust gases). Dilution of the fresh air provides less reactants to form NOx. Second, EGR absorbs heat from the combustion process, thereby reducing combustion chamber temperatures with an attendant reduction in NOx formation. Heat absorption capacity is in turn a function of EGR flow rate and its temperature, both of which are commonly controlled to minimize NOx emissions. EGR coolers can be added to the EGR system to lower the EGR temperature.

While in theory the EGR system simply routes some exhaust gas back to the intake, production systems can be complex and involve many components to ensure accurate control of EGR flow and maintain acceptable PM and NOx emissions while minimizing effects on fuel economy. To determine the necessary EGR flow rates and control EGR flow, EGR systems normally use the following components: an EGR valve, valve position sensor, boost pressure sensor, intake temperature sensor, intake (fresh) airflow sensor, and tubing or piping to connect the various components of the system. EGR temperature sensors and exhaust backpressure sensors are also commonly used. Additionally, some systems use a variable geometry turbocharger to provide the backpressure necessary to drive the EGR flow. Therefore, EGR is not a stand alone emission control device. Rather, it is carefully integrated with the air handling system (supercharging and intake cooling) to control NOx while not adversely affecting PM emissions and fuel economy.

The staff anticipates manufacturers will need to design EGR systems that accurately and continuously control EGR flow under both transient and steady state load conditions to meet the certification standards applicable for the 2007 and subsequent model years. Further, EGR will have to be accurately controlled under the range of ambient conditions represented by the Not-to-Exceed, or NTE, test to maintain emissions while maximizing in-use fuel economy (refer to section VII. G. of the Staff Report for more details of the NTE zone). The staff believes all of the components used for control (including alternate emission control device or "AECD" operation) purposes can also be used for monitoring. The staff projects that manufacturers would not have to add any components specifically for EGR monitoring.

Proposed Monitoring Requirements

A common phrase in diesel emission control discussions is the "NOx/PM trade-off." Typically, as air-fuel ratio, fuel injection (e.g., start of injection) and EGR parameters are varied, changes that improve NOx emissions tend to increase PM emissions, and changes that improve PM emissions tend to increase NOx emissions. Specifically for EGR system design, excessive EGR flow causes increased PM emissions, and insufficient EGR flow causes increased NOx emissions. When manufacturers design engines and emission control systems, they have to balance this trade-off to achieve both the NOx and PM emission standards.

Given the need to accurately control EGR to maintain acceptable emission levels, the staff is proposing monitoring requirements for flow rate and response rate malfunctions. Additionally, on vehicles equipped with EGR coolers, the OBD system would be required to monitor the cooler for insufficient cooling malfunctions.

EGR Flow Rate Monitoring

Under the staff's proposal, the OBD system would be required to detect an EGR system malfunction before the change (i.e., decrease or increase) in flow from the manufacturer's specified EGR flow rate causes vehicle emissions to exceed 1.5 times any of the applicable emission standards. In situations where no failure or deterioration of the EGR system that causes a decrease in flow could result in vehicle emissions exceeding 1.5 times any of the applicable standards, the OBD system would be required to detect a malfunction when the system has no detectable amount of EGR flow. Similarly, if high flow malfunctions do not cause emissions to exceed 1.5 times any of the applicable standards, the OBD system would be required to detect a malfunction when the EGR system has no ability to reduce EGR flow when it commands a reduction in EGR. Since the EGR system may experience flow rate malfunctions only under some conditions (e.g., a "sticking" EGR valve may not fully open to achieve a desired high flow EGR condition but may still be able to open enough to achieve lower flow rates), the EGR system would be continuously monitored for low and high flow malfunctions.

Under the high flow rate monitor, the OBD system would also monitor for leaking EGR valves. A leaking EGR valve can cause increased PM emissions under conditions where EGR flow is commanded off (i.e., during aggressive engine transients). While a leaking valve may be characterized as a high flow malfunction, it might not necessarily be detected by the high flow diagnostic discussed above. A leaking valve is likely to be caused by a failure of the valve to seat properly when commanded closed, and only has an emission impact under conditions where the valve is commanded closed or turned off. Functional failures for valve opening and valve control would be detected by the flow and response diagnostics discussed above, but these diagnostics may not detect proper valve closing/seating (e.g., if the EGR control system is in an "open loop" mode when it is commanded closed, the flow and response diagnostics would likely be disabled and would not detect the leaking valve).

EGR Response Rate Monitoring

Manufacturers will likely use transient EGR control to meet the emissions standards. EGR rates will be varied with transient engine operating conditions to maintain the balance between NOx and PM emissions. The staff is therefore proposing a response rate diagnostic to verify that the system has sufficient response. This monitor would detect the inability of the EGR system to modulate EGR flow rates under transient engine conditions. Specifically, the OBD system would be required to detect a response malfunction of the EGR system if it is unable to achieve the commanded flow rate within a manufacturer-specified time that would cause a vehicle's emissions to exceed 1.5 times any of the applicable standards. Similar to the monitoring requirements for flow rate malfunctions, if response rate malfunctions do not cause emissions to exceed 1.5 times the standards, the staff is proposing that the diagnostic system be required to simply verify that the EGR system has some response to computer commands.

The manufacturer would be required to monitor response rate during both increasing and decreasing EGR flow rate conditions. Considering the NOx/PM trade-off discussed above, slow response while trying to increase EGR rates may result in increased NOx emissions. Similarly, slow response while trying to decrease EGR rates may yield in increased PM emissions. Manufacturers would have to account for these trends when determining their malfunction thresholds. Further, it is necessary to monitor response rate under both increasing and decreasing conditions because some malfunctions may only affect response under one (i.e., increasing or decreasing) condition. For example, some EGR valves are held in the closed position with a spring. As the spring deteriorates, it may still properly hold the valve in the closed position, but the valve would close at a slower rate (and might even open at a faster rate). Such a malfunction would only be detected by monitoring the response rate under decreasing EGR conditions.

EGR Cooling System Monitoring

Insufficient EGR cooling can result in higher NOx emissions and can lead to default operation where EGR is shutoff. Accordingly, the staff is proposing monitoring requirements for proper EGR cooling system performance. Specifically, the OBD system would be required to indicate an EGR cooling system malfunction when the reduction in cooling of the exhaust gas causes emissions to exceed 1.5 times any of the applicable standards. For vehicles in which no failure or deterioration of the EGR system cooler could result in a vehicle's emissions exceeding 1.5 times any of the applicable standards, the OBD system would be required to detect a malfunction when the system has no detectable amount of EGR cooling. Some manufacturers using EGR coolers have indicated that the cooler is not used for emission reductions but rather for EGR valve and system durability. These manufacturers have also requested to forego monitoring of the EGR cooler. If a manufacturer demonstrates that emissions will not be affected under any reasonable driving condition due to a complete lack of EGR cooling, the manufacturer would not be required to monitor the EGR cooler.

At this time, the staff is not proposing monitoring requirements for malfunctions that result in EGR overcooling. While overcooling can lead to accelerated deterioration of the EGR system and engine components due to formation and condensation of corrosive gases, the staff has not reviewed any data indicating emissions are affected due to overcooling of EGR gases. However, to address the condensation issue, manufacturers may employ bypass designs that do not cool the exhaust gas under conditions that can result in condensation. Manufacturers would be required to monitor the bypass system to verify that bypass does not occur when cooling is needed.

Other Monitoring Requirements

Manufacturers would be required to monitor all electronic components of the EGR system (e.g., temperature sensors, valves) for proper function and rationality under the comprehensive component monitoring requirements.

Technical Feasibility of Proposed Monitoring Requirements

EGR Flow Rate Monitoring

The EGR control system has to determine and control the EGR flow. While the system designs from different manufacturers will vary, they will employ a similar closed loop control strategy. First, the control system determines a desired EGR flow rate based on the engine operating conditions. Manufacturers will likely store the desired flow rate/valve position in a lookup table in the engine control module (ECM) (e.g., the desired EGR values, which are based on engine operating conditions such as engine speed and engine load, are established when the manufacturer designs and calibrates the EGR system). The ECM commands the valve to the position necessary to achieve the desired flow. EGR flow rate and/or valve position is feedback-controlled. The ECM calculates or directly measures both fresh air charge and total intake charge. The difference between the total intake charge and fresh airflow is the actual EGR flow. The closed-loop control system continuously adjusts the EGR valve position until the actual EGR flow equals the desired EGR flow.

These closed-loop control strategies could be readily monitored and are the basis for many existing monitors on both gasoline and diesel light- and medium-duty vehicles. The OBD system could evaluate the difference (i.e., error) between the look-up value and the final commanded value to achieve the desired flow rate. When the error exceeds a specific threshold, a malfunction would be indicated. Typically, as the feedback parameter or learned offset increases, there is an attendant increase in emissions, and a correlation could be made between feedback adjustment and emissions. This type of monitoring strategy could be used to detect both high and low flow malfunctions, and is currently in production on a medium-duty vehicle.[7]

While the closed-loop control strategy described above is effective in measuring and controlling EGR flow, some manufacturers are currently investigating the use of a second control loop based on a wide-range exhaust gas sensor to further improve EGR control and emissions. With this second control loop the desired air-fuel ratio is calculated based on engine operating conditions (i.e., intake airflow, commanded EGR flow and commanded fuel). The calculated air-fuel ratio is compared to the air-fuel ratio from the exhaust gas sensor and refinements can be made to the EGR and airflow rates (i.e., the control can be “trimmed”) to actually achieve the desired rates. On systems that use the second control loop, flow rate malfunctions could also be detected using the feedback information from the wide-range oxygen sensor. Manufacturers would be required to monitor this feedback control loop, and the OBD system would be required to detect a malfunction when the EGR system is unable to maintain emissions below 1.5 times any of the standards or when the feedback control based on the wide-range sensor has used up all the adjustment allowed by the manufacturer.

Two types of leaking valves are required to be detected. One type is the failure of the valve to seal when in the closed position (e.g., if the valve or seating surface is eroded, the valve could close and seat, yet still allow some flow across the valve). A flow check is necessary to detect a malfunctioning valve that closes properly but still leaks. EGR flow (total intake charge minus fresh air charge) could be calculated with the valve closed using the monitoring strategy described above for high and low malfunctions, and when flow exceeds unacceptable levels, a malfunction would be indicated. Some cooled EGR systems will incorporate an EGR temperature sensor, which could also be used to detect a leaking EGR valve. For a properly functioning EGR valve, EGR temperature should be a minimum when the EGR valve is closed. An elevated EGR temperature when the valve is closed would indicate a malfunctioning valve. A leaking valve can also be caused by failure of the valve to close/seat (e.g., carbon deposits on the valve or seat that prevent the valve from fully closing). The flow check described above would detect failure of the valve to close/seat but would require a repair technician to further diagnose whether the problem is a sealing or seating problem. Failure of the valve to close/seat could be specifically monitored by checking

---

[7] “2003 MY OBD System Operation Summary for 6.0L Engine” at website http://www.motorcraftservice.com/vdirs/diagnostics/pdf/Dobdsm304.pdf.

the zero position of the valve with the position sensor when the valve is closed. If the valve position is out of the acceptable range for a closed valve, a malfunction would be indicated. This type of zero position sensor check is commonly used to verify the closed position of valves/actuators used in gasoline OBD II systems (e.g., gasoline EGR valves, electronic throttle) and would be feasible for diesel EGR valves.

EGR Response Rate Monitoring

The EGR response rate diagnostic is similar to the flow rate diagnostic. While the flow rate diagnostic would evaluate the ability of the EGR system to achieve a commanded flow rate under relatively steady state conditions, the response diagnostic would evaluate the ability of the EGR system to modulate (i.e., increase and decrease) EGR flow as engine operating conditions and, consequently, commanded EGR rates change. Specifically, as engine operating conditions and commanded EGR flow rates change, the monitor would evaluate the time it takes for the EGR control system to achieve the commanded change in EGR flow. This monitor could evaluate EGR response passively during transient engine operating conditions encountered during in-use operation. The monitor could also intrusively evaluate EGR response by commanding a change in EGR flow under a steady state engine operating condition and measuring the time it takes to achieve the new EGR flow rate.

EGR Cooling System Monitoring

Some diesel engine manufacturers are currently using exhaust gas temperature sensors as an input to their EGR control systems. On these systems, EGR temperature, which is measured downstream of the EGR cooler, could be used to monitor the effectiveness of the EGR cooler. For a given engine operating condition (e.g., a steady speed/load that generates a known exhaust mass flow and exhaust temperature to the EGR cooler), EGR temperature will increase as the performance of the EGR cooling system decreases. During the OBD calibration process, manufacturers could develop a correlation between increased EGR temperatures and cooling system performance (i.e., increased emissions). The EGR cooling monitor would use such a correlation and indicate a malfunction when the EGR temperature increases to the level that causes emissions to exceed 1.5 times the emission standards.

While the staff anticipates that most, if not all, manufacturers will use EGR temperature sensors to meet future standards, EGR cooler monitoring may also be feasible without an EGR temperature sensor by using the intake manifold temperature (IMT) sensor. EGR cooler performance could be evaluated by looking at the change in IMT (i.e., “delta” IMT) with EGR turned on and EGR turned off (IMT would be higher with EGR turned-on). If there is significant cooling capacity with a normally functioning cooling system, there could be a significant difference in intake manifold temperature with EGR turned on and off. As cooling system performance decreases, the change in IMT would increase. Delta IMT could be correlated to decreased cooling system performance and increased emissions.

D. BOOST PRESSURE CONTROL SYSTEM MONITORING

Background

Turbochargers are used on internal combustion engines to enhance performance by increasing the mass and density of the intake air. Some of the benefits of turbocharging include increased horsepower, improved fuel economy, and decreased exhaust smoke density.[8] Most modern diesel engines take advantage of these benefits and are equipped with turbocharging systems. The power increase associated with turbocharging also brings higher engine stresses, so the robust design of the diesel engine makes the addition of a turbocharger less problematic compared to gasoline engines. Since turbochargers increase the efficiency of the diesel engine, exhaust emissions are also improved. Moreover, smaller turbocharged diesel engines can be used in place of larger non-turbocharged engines to achieve the desired engine performance characteristics.

The most widely used turbochargers utilize exhaust gas to spin a turbine at speeds from 10,000 to over 150,000 rpm. The turbine is mounted on the same rotating shaft as an adjacent centrifugal pump. The energy that would otherwise be exhausted as waste heat is used to drive the turbine, which in turn drives the centrifugal pump. This pump draws in fresh air and compresses it to increase the density of the air charge to the cylinders, thereby increasing power.

A boost pressure sensor is typically located in the intake manifold to provide a feedback signal of the current turbo boost. As turbo speed (boost) increases, the pressure in the intake manifold also increases. Hence, engine designers may compare the boost pressure signal to a target boost for the given engine speed and load conditions. Target boost pressure is then obtained by either modulating a wastegate valve or turbo vanes.

Proper boost control is essential to optimize emission levels. Even short periods of over- or under-boost can result in undesired air-fuel ratio excursions and corresponding emission increases. Additionally, the boost control system directly affects exhaust and intake manifold pressures. Another critical emission control system, EGR, is very dependent on these two pressures and generally uses the differential between them to force exhaust gas into the intake manifold. If the boost control system is not operating correctly, the exhaust or intake pressures may not be as expected and EGR may not function as designed. In high-pressure EGR systems, higher exhaust pressures will generate more EGR flow and, conversely, lower pressures will reduce EGR flow. A malfunction that causes excessive exhaust pressures (e.g., wastegate stuck closed at high engine speed) can produce higher EGR flowrates at high load conditions and have a negative impact on emissions.[9]

---

[8] Ecopoint Inc., 2000. "Turbochargers for Diesel Engines", DieselNet Technology Guide.

[9] Ecopoint Inc., 2000. "Effects of EGR on Engine and Emissions", DieselNet Technology Guide.

Manufacturers commonly use charge air coolers to maximize the benefits of turbocharging. As the turbocharger compresses the intake air, the temperature of the intake air charge increases. This increasing air temperature causes the air to expand, which is directionally opposite of what turbocharging is attempting to accomplish. Charge air coolers are used to exchange heat between the compressed air and ambient air (or coolant) and cool the compressed air. Accordingly, a decrease in charge air cooler performance can affect emissions by causing higher intake air temperatures that can lead to increased NOx emissions from higher combustion temperatures.

One drawback of turbocharging is known as turbo lag. Turbo lag occurs when the driver attempts to accelerate quickly from a low engine speed. Since the turbocharger is a mechanical device, a delay exists from the driver demand for more boost until the exhaust flow can physically speed up the turbocharger. In addition to a negative effect on driveability and performance, improper fueling (e.g., over-fueling) during this lag can cause emission increases (typically PM).

To decrease the effects of turbo lag, manufacturers design turbos that spool up quickly at low engine speeds and low exhaust flowrates. However, designing a turbo that will accelerate quickly from a low engine speed but will not result in an over-speed/over-boost condition at higher engine speeds is difficult. That is, as the engine speed and exhaust flowrates near their maximum, the turbo speed increases to levels that cause excessive boost pressures and heat that could lead to engine or turbo damage. To prevent excessive turbine speeds and boost pressures at higher engine speeds, a wastegate is often used to bypass part of the exhaust stream around the turbocharger. The wastegate valve is typically closed at lower engine speeds so that all exhaust is directed through the turbocharger, thus providing quick response from the turbocharger when the driver accelerates quickly from low engine speeds. The wastegate is then opened at higher engine speeds to prevent engine or turbo damage from an over-speed/over-boost condition.

An alternative to using a wastegate is to use an improved turbocharger design commonly referred to as a variable geometry turbo (VGT). To prevent over-boost conditions and to decrease turbo lag, VGTs are designed such that the geometry of the turbocharger changes with engine speed. While various physical mechanisms are used to achieve the variable geometry, the overall result is essentially the same. At low engine speeds, the exhaust gas into the turbo is restricted in a manner that maximizes the use of the available energy to spin the turbo. This allows the turbo to spool up quickly and provide good acceleration response. At higher engine speeds, the turbo geometry changes such that exhaust gas flow into the turbo is not as restricted. In this configuration, more exhaust can flow through the turbocharger without causing an over-boost condition. The advantage that VGTs offer compared to a waste-gated turbocharger is that all exhaust flow is directed through the turbocharger under all operating conditions. This can be viewed as maximizing the use of the available exhaust energy.

Proposed Monitoring Requirements

The staff is proposing manufacturers be required to monitor boost control systems for proper operation. Manufacturers would be required to continuously monitor for appropriate boost to verify that the turbocharger is operating as designed and conditions of over-boost or under-boost are not occurring. The boost control monitor would need to be calibrated to indicate a malfunction before emissions exceed 1.5 times the emission standards.

The staff is also proposing that manufacturers be required to monitor for slow response malfunctions. That is, the time required to reach the desired boost, whether transitioning from high to low boost or low to high, would be measured and a malfunction detected when the response time increases to a level where emissions are increased. The boost response monitor would also need to be calibrated to 1.5 times the emission standards.

The electronic components of the boost control system (e.g., actuators, pressure sensors, position sensors) that provide or receive a signal from the engine control module (ECM) would need to be monitored under the comprehensive component requirements for malfunctions such as circuit failures, rationality faults, and functional response to computer commands.

Lastly, the staff is proposing that charge air coolers be monitored for proper cooling of the intake air. A malfunction would need to be detected before emissions exceed 1.5 times the emission standards. If no charge air cooler malfunction can cause emissions to exceed 1.5 times the emission standards, then the cooler would need to be monitored for proper functionality (e.g., verify that some detectable level of cooling is occurring).

Technical Feasibility of Proposed Monitoring Requirements

To monitor boost control systems, manufacturers are expected to look at the difference between the actual pressure sensor reading (or calculation thereof) and the desired/target boost pressure. If the error between the two is too large or persists for too long, a malfunction would be detected. Manufacturers would need to calibrate the length of time and size of error to ensure robust detection of a fault occurs before the emission malfunction threshold is exceeded. Given the purpose of a closed-loop control system with a feedback sensor is to continually measure the difference between actual and desired boost pressure, the control system is already continually monitoring the difference and attempting to minimize it. As such, a diagnostic requirement to indicate a fault when the difference gets too large and the system can no longer properly achieve the desired boost is essentially an extension of the existing control strategy. Additionally, three different diesel medium-duty vehicle engines are currently certified to the light- and medium-duty OBD II regulation requirements with OBD II systems that meet these proposed requirements.

To monitor for malfunction or deterioration of pressure sensors, manufacturers could validate sensor readings against other sensors present on the vehicle or against ambient conditions. For example, at initial key-on before the engine is running, the boost pressure sensor should read ambient pressure. If the vehicle is equipped with a barometric pressure sensor, the two sensors could be compared and a malfunction indicated when the two readings differ beyond the specific tolerances. A more crude rationality check of the boost pressure sensor may be accomplished by verifying that the pressure reading is within reasonable atmospheric limits for the conditions the vehicle will be subjected to.

Rationality monitoring of VGT position sensors may be accomplished by comparing the measured sensor value to expected values for the given engine speed and load conditions. For example, at high engine speed and loads, the position sensor should indicate that the VGT position is opened more than would be expected at low engine speed and loads. These rationality checks would need to be two-sided. That is, position sensors would be checked for appropriate reading at both high and low engine operating conditions.

E. CATALYST MONITORING

Diesel Oxidation Catalyst

Background

Diesel oxidation catalysts have been used with off-road diesel engines since the 1960s and on trucks and buses in the U.S. since the early 1990s. Oxidation catalysts are generally used for reducing HC and carbon monoxide (CO) emissions via an oxidation process. Current diesel oxidation catalysts, however, are also optimized to reduce PM emissions. Specifically, while promoting the chemical oxidation of HC and CO, diesel oxidation catalysts also oxidize the soluble organic fraction (SOF) of diesel particulates. The SOF consists of hydrocarbons adsorbed to the carbonaceous solid particles and may also include hydrocarbons that have condensed into droplets of liquid. At sufficiently high temperatures diesel oxidation catalysts can convert up to 90 percent of HC and CO emissions and 30 percent of PM emissions. Considering the low engine-out concentrations of HC and CO in diesel exhaust, diesel manufacturers are projected to primarily use these catalysts to reduce diesel particulate emissions. Oxidation catalysts may also be used in conjunction with other aftertreatment emission controls such as NOx adsorber systems, selective catalytic reduction (SCR) systems, and PM traps to improve their performance.

Unfortunately, oxidation catalysts also promote the oxidation of sulfur dioxide (the exhaust concentration of which is highly dependent on the fuel sulfur content). This generates more sulfate particulates, and thus may increase total PM emissions. As such, much of the current focus in improvements to the diesel oxidation catalyst has been to make them more selective (i.e., to formulate catalysts that exhibit sufficiently high HC and SOF oxidation activity and acceptably low sulfate particulate formation).

Current research has been concentrating on designing catalysts that extend PM control into the higher exhaust gas temperatures.

Proposed Monitoring Requirements

The staff is proposing that manufacturers monitor the oxidation catalyst for proper performance. Specifically, the staff is proposing that manufacturers be required to indicate a malfunction when the conversion efficiency decreases to a point that emissions exceed 1.75 times the applicable non-methane hydrocarbon (NMHC) or PM (or if applicable, NMHC+NOx) standards. At a minimum, manufacturers would be required to monitor the catalyst once per driving cycle in which the monitoring conditions are met. If a malfunctioning catalyst cannot cause emissions to exceed the emission threshold of 1.75 times the applicable standards, a manufacturer would only be required to functionally monitor the system and indicate a malfunction when no conversion efficiency of the emission of concern could be detected. Additionally, through the 2009 model year, no monitoring would be required if the conversion efficiency of the catalyst system is less than 30 percent.

In order to determine the proper OBD malfunction threshold for the oxidation catalyst, manufacturers would be required to progressively deteriorate or “age” the catalyst(s) to the point where emissions exceed 1.75 times the standard. The method used to age the catalyst(s) must be representative of real world catalyst deterioration (e.g., thermal and/or poisoning degradation) under normal and malfunctioning operating conditions. For engines with aftertreatment systems that only utilize diesel oxidation catalysts, the catalyst(s) can be aged as a system to the emission threshold for determining the malfunction threshold. However, for engines with aftertreatment systems that utilize multiple different catalyst technologies (e.g., an aftertreatment system that includes an oxidation catalyst, catalyzed NOx trap, catalyzed PM trap, and lean NOx catalyst), determining the OBD malfunction threshold for the diesel oxidation catalyst becomes more complex since the aging affects on the catalyst is dependent on many factors, including the location of the oxidation catalyst relative to the other aftertreatment technologies and the synergism between each component in the system. Given that each component in the system are dependent on each other, are part of an overall catalyst system, and deteriorate in-use as a system, it would not be appropriate to treat each component in the system completely independent of each other. Since it is uncertain what exhaust configurations and aftertreatment systems manufacturers will use to comply with the future emission standards for the 2007 and later model years, it is important for the staff to develop and specify a “one-size-fits-all” aging process that accurately represents every possible future aftertreatment configuration. Once diesel aftertreament system designs have stabilized to a level similar to gasoline aftertreatment systems (i.e., the variation of aftertreatment systems is limited) defining a definitive but generic catalyst aging plan will be more simple and practical. Until then, the staff would require manufacturers to submit a monitoring plan to the Executive Officer for review and approval of the monitoring strategy, malfunction criteria, and monitoring conditions prior to introduction on a production engine. Executive Officer approval would be based on the representativeness of the catalyst system aging to real world catalyst deterioration under normal and malfunctioning operating conditions, the effectiveness of the monitor to pinpoint the likely area of malfunction, and verification that each catalyst component is functioning as designed.

Technical Feasibility of Proposed Monitoring Requirements

Monitoring of the oxidation catalysts could be performed similar to three-way catalyst monitoring, which uses the concept that oxygen storage correlates well with hydrocarbon and NOx conversion efficiency. Thus, oxygen sensors located upstream and downstream of the catalyst can be used to determine when the oxygen storage capability of the catalyst deteriorates below a predetermined threshold. Determining the oxygen storage capacity would require lean air-fuel (A/F) operation followed by rich A/F operation or vice-versa during catalyst monitoring. Since a diesel engine normally operates lean of stoichiometry, the lean A/F operation portion will be a normal event. However, the rich A/F operation would have to be commanded intrusively when the catalyst monitor is active. The rich A/F operation could be achieved with the engine fuel injectors through late fuel injection or with a dedicated injector in the exhaust upstream of the catalyst. With lean operation, the catalyst will be saturated with stored oxygen. As a result, both the front and rear oxygen sensors should be reading lean. However, when rich A/F operation initiates, the front oxygen sensor would switch immediately to a “rich” indication while the rear oxygen sensor should stay reading “lean” until the stored oxygen in the catalyst is all consumed by the rich fuel mixture in the exhaust. As the catalyst deteriorates, the delay time between the front and rear oxygen sensors reading lean would become progressively smaller. Thus, by comparing the time difference between the responses of the front and rear oxygen sensors to the lean-to-rich or rich-to-lean A/F changes, the performance of the catalyst could be determined. Although conventional oxygen sensors are utilized to illustrate the monitoring method above, these sensors could be substituted with wide-range oxygen sensors for additional engine control benefits such as EGR trimming and fuel trimming.

Alternatively, if only a functional monitor of the catalyst is required (e.g., a malfunctioning catalyst cannot cause emissions to exceed 1.75 times the emission standard), temperature sensors could be used for monitoring. A functioning oxidation catalyst is expected to provide a significant exotherm when it oxidizes HC and CO. By placing one or more temperature sensors at or near the catalyst, the temperature of the catalyst could be measured. Depending upon the efficiency of the catalyst and the duty cycle of the vehicle, the exotherm may be difficult to discern from the inlet exhaust temperatures. To add robustness to the monitor, the functional diagnostic would need to be conducted during predetermined operating conditions where the amount of HC and CO entering the catalyst are known. This may require an intrusive diagnostic that actively forces the fueling strategy richer (e.g., through late or post injection) than normal for a short period of time. If the measured exotherm exceeds a predetermined amount that only a working catalyst can achieve, the diagnostic would pass.

Lean NOx Catalyst

Background

Lean NOx catalysts are essentially reduction catalysts (i.e., catalysts primarily involved in reducing NOx emissions via reduction processes with hydrocarbons) specifically aimed at reducing NOx emissions in the presence of oxygen-rich exhaust gases (i.e., lean conditions) characteristic of diesel engines. Lean NOx catalysts are relatively simple systems that can utilize hydrocarbons from diesel exhaust (a process known as passive lean NOx reduction) to reduce NOx emissions. In general, lean NOx catalysts show increasing NOx conversion rates with increasing HC concentrations. Since the concentration of HC in diesel exhaust is normally low, enrichment of the exhaust with added HC (a process known as active lean NOx reduction) has been pursued as an approach to improve NOx conversion rates. Enrichment of the diesel exhaust can be done by injecting diesel fuel through a dedicated injector into the exhaust system upstream of the catalyst or through late fuel injection into the cylinder. However, even with the addition of HC into the exhaust stream, the average NOx conversion efficiency of lean NOx catalysts remains generally low (less than 30 percent). These catalysts also tend to possess a less favorable efficiency/fuel penalty tradeoff and are most effective in a limited temperature-operating window that does not always correspond to the exhaust temperature at which most NOx emissions are generated. Additionally, catalyst efficiency is affected by HC/NOx ratios and oxygen content in the exhaust.[10] Due to these problems, further improvements need to be made for lean NOx catalysts to achieve widespread commercialization. Currently, lean NOx catalyst technology is primarily aimed at providing small NOx reduction functionality in other technologies, such as diesel oxidation catalysts.

Proposed Monitoring Requirements

The proposed monitoring requirements would require monitoring of the lean NOx catalyst (i.e., catalysts primarily involved in reducing NOx emissions via reduction processes) for proper NOx conversion performance. Specifically, the staff is proposing that manufacturers be required to indicate a malfunction when the conversion efficiency decreases to a point that emissions exceed 1.75 times the applicable NOx (or if applicable, NMHC+NOx) emission standards. If a malfunctioning catalyst cannot cause emissions to exceed the emission threshold of 1.75 times the applicable standards, a manufacturer would only be required to functionally monitor the system and indicate a malfunction when no conversion efficiency of the emissions of concern could be detected. At a minimum, manufacturers would be required to monitor the catalyst once per driving cycle in which the monitoring conditions are met. For lean NOx catalysts that utilize an active/intrusive diesel injection strategy (i.e., active lean NOx catalysts), monitoring must be conducted continuously since precise control of reductant addition throughout the engine's operation range is essential for good NOx performance from the system . Additionally, through the 2009 model year, no monitoring would be required if the conversion efficiency of the catalyst system is less than 30 percent. Further, if an active lean NOx catalyst is utilized, the mechanism for adding the fuel reductant must be monitored for proper function.

---

[10] "Lean NOx Catalyst," www.dieselnet.com.

Technical Feasibility of Proposed Monitoring Requirements

In order to monitor the lean NOx catalyst, manufacturers are projected to use NOx sensors. NOx sensors placed upstream and downstream of the lean NOx catalyst could be used to determine the NOx conversion efficiency directly. Alternatively, manufacturers could potentially use a single NOx sensor placed downstream of the catalyst to measure catalyst-out NOx emissions during engine operation within a controlled window where NOx engine-out (i.e., catalyst inlet) emission performance is relatively stable and can be reliably estimated. Within this engine operation window, NOx catalyst-out measurements could be compared with a calibrated emission threshold for determining a malfunctioning or deteriorated lean NOx catalyst system. If both an upstream and downstream NOx sensor are used for monitoring, the upstream sensor could be used to improve the overall effectiveness of the catalyst by controlling the air-fuel ratio in the exhaust precisely to the levels where the catalyst is most effective.

If an active lean NOx catalyst is utilized, manufacturers would be required to monitor the mechanism for adding the fuel reductant for proper function. This could be done by using a temperature sensor located near or at the catalyst to determine if an exotherm resulting from the injection has occurred. A temperature sensor placed near or at the catalyst is projected to be needed for control purposes on these catalysts to determine when the catalyst is active. As previously described, lean NOx catalysts tend to have a narrow temperature range where they are most effective. Adding reductant when the catalyst is not sufficiently active would adversely affect fuel economy without a reduction in emission levels. Therefore, a temperature sensor placed in the exhaust could help determine when reductant injection should occur. This same sensor can also be used to monitor the injection. Alternatively, the NOx sensors that are used to monitor the lean NOx catalyst can be utilized to determine if the injection has occurred. Since NOx sensors also have the capability to determine the air-fuel ratio in the exhaust stream, the diesel fuel injection into the exhaust can also be verified with this sensor.

Selective Catalytic Reduction (SCR) Catalyst

Background

The SCR catalyst has been used on power plants and stationary engines since the 1970s and is now being developed for use on on-road diesel engines. SCR catalysts are considered one of the most promising exhaust aftertreatment technologies for NOx control. While lean NOx catalysts use hydrocarbons as reductants to reduce NOx, SCR systems use nitrogen-containing compounds such as ammonia or urea, which is injected from a separate reservoir into the gas stream before the catalyst. Currently, the SCR system, with NOx reduction rates of over 80 percent achieved on heavy-duty engines, is one of the more promising catalyst technologies capable of achieving the most stringent future low NOx emission standards.

SCR catalyst systems require an accurate ammonia control system to inject precise amounts of reductant. Currently, urea is considered the best reductant for providing ammonia on heavy-duty applications due to its non-toxicity, ease of transport and handling, and wide availability. At temperatures above 160 degrees Celsius, urea thermally decomposes to ammonia in the exhaust, thereby providing ammonia to the SCR catalyst. Concerning ammonia, an injection rate that is too low may result in lower NOx conversions while an injection that is too high may release unwanted ammonia emissions (referred to as ammonia slip) to the atmosphere. In general, ammonia to NOx ratios of around 1:1 are used to provide the highest NOx conversion rates with minimal ammonia slip. Therefore, it is important to inject just the right amount of ammonia appropriate for the amount of NOx in the exhaust. For stationary source engines, estimating the exhaust NOx levels is fairly easy since the engine usually operates at a constant speed and load and the NOx emission rate is generally stable. However, on-road diesel engines operate over a variety of speeds and loads, thereby making NOx exhaust estimates difficult without a dedicated NOx sensor in the exhaust. With an accurate fast response NOx sensor, closed-loop control of the ammonia injection can be used to achieve and maintain the desired ammonia/NOx ratios in the SCR catalyst for high NOx conversion efficiency (i.e., greater than 90 percent) necessary to achieve the 2007 emission levels under various engine operating conditions. Currently, however, such an accurate fast response NOx sensor is not yet available. It has been estimated that achieving the 2007 NOx emission standards with SCR systems will require NOx sensors that can measure NOx levels accurately around the 20 to 40 ppm range with little cross sensitivity to ammonia.[11] Current NOx sensors do not yet meet these specifications, but sensor technology is improving quickly such that zero to 500 ppm resolution sensors have been achieved[12] and zero to 100 ppm sensors are being developed.[13] With further development, sensors are expected to achieve the 20 to 40 ppm NOx sensitivity in time for the 2007 emission standards. Further, recent production plans indicate most manufacturers intend to meet the 2007-2009 phase-in of lower standards by making all engines meet intermediate standards. Such a plan could avoid the need for significant NOx aftertreatment and the associated monitors and sensor technology until the 2010 model year. Regarding cross-sensitivity to ammonia, work has been done that indicates ammonia and NOx measurements can be independently measured by conditioning the output signal.[14] This signal conditioning method resulted in a linear output for both ammonia and NOx from the NOx sensor downstream of the catalyst.

For SCR systems, closed-loop control of the reductant injection will likely require the use of two NOx sensors. The first NOx sensor would be located upstream of the catalyst and the reductant injection point would be used for measuring the engine-out

---

[11] Song, Q. and Zhu, G., “Model-based Closed-loop Control of Urea SCR Exhaust Aftertreatment System for Diesel Engine,” SAE Paper 2002-01-0287.

[12] Kato, N., Kokune, N., Lemire, B., and Walde, T., “Long Term Stable NOx Sensor with Integrated In-Connector Control Electronics,” SAE Paper 1999-01-0202.

[13] Kobayashi, N., et al., “Development of Simultaneous NOx/NH3 Sensor in Exhaust Gas,” Mitsubishi Heavy Industries, Ltd., Technical Review Vol.38 No.3 (Oct. 2001).

[14] Schaer, C. M., Onder, C. H., Geering, H. P., and Elsener, M., “Control of a Urea SCR Catalytic Converter System for a Mobile Heavy Duty Diesel Engine,” SAE Paper 2003-01-0776.

NOx emissions and determining the amount of reductant injection needed to reduce emissions. The second NOx sensor located downstream of the catalyst would be used for measuring the amount of ammonia and NOx emissions exiting the catalyst and providing feedback to the reductant injection control system. If the downstream NOx sensor detects too much NOx emissions exiting the catalyst, the control system can inject higher quantities of reductant. Conversely, if the downstream NOx sensor detects too much ammonia slip exiting the catalyst, the control system can decrease the amount of reductant injection.

In addition to exhaust NOx levels, another important parameter for achieving high NOx conversion rates with minimum ammonia slip is catalyst temperature. SCR catalysts have a defined temperature range where they are most effective. For example, platinum catalysts are effective between 175 and 250 degrees Celsius, vanadium catalysts are effective between 300 and 450 degrees Celsius, and zeolite catalysts are most effective between 350 and 600 degrees Celsius. Injecting urea into the SCR catalyst outside the effective temperature band could lead to deactivation through poisoning or collapse of the crystal structure of the catalyst.[15] Furthermore, the reaction kinetics between ammonia and NOx are sensitive to temperature. In general, at higher catalyst temperatures, more ammonia needs to be added to the exhaust to achieve the desired NOx conversion rates while at lower temperatures, ammonia injection rates need to be limited to prevent ammonia slip.[16] To determine exhaust catalyst temperature for reductant control purposes, manufacturers are likely to use temperature sensors placed in the exhaust system. It is projected that only one temperature sensor positioned just downstream of the SCR system will be utilized for reductant injection control purposes.

Production SCR catalyst systems may also contain auxiliary catalysts to improve the overall NOx conversion rate of the system. An oxidation catalyst is often positioned downstream of the SCR catalyst to help control ammonia slip on systems without closed-loop control of ammonia injection. The use of a “guard” catalyst could allow higher ammonia injection levels, thereby increasing the NOx conversion efficiency without releasing un-reacted ammonia into the exhaust. The guard catalyst can also reduce HC and CO emission levels and diesel odors. However, increased N_2O emissions may occur and NOx emission levels may actually increase if too much ammonia is oxidized in the catalyst. Some SCR systems may also include an oxidation catalyst upstream of the SCR catalyst and urea injection point to generate N_2O for reducing the operating temperature range and/or volume of the SCR catalyst. Studies have indicated that increasing the N_2O content in the exhaust stream can reduce the SCR temperature requirements by about 100 degrees Celsius.[17] This “pre-oxidation” catalyst also has the added benefit of reducing HC emissions. However, additional sulfate PM emissions can occur when high sulfur fuel is used.[16]

---

[15] “Selective Catalyst Reduction,” www.dieselnet.com.

[16] Van Helden, R., van Genderen, M., van Aken, M., et al., “Engine Dynamometer and Vehicle Performance of a Urea SCR-System for Heavy-Duty Truck Engines,” SAE Paper 2002-01-0286.

[17] Walker, A. P., Chandler, G. R., Cooper, B. J., et al., “An Integrated SCR and Continuously Regenerating Trap System to Meet Future NOx and PM Legislation,” SAE Paper 2000-01-0188.

Despite its high NOx conversion efficiency, there are several concerns in applying SCR systems to mobile applications. First, proper injection control is difficult under transient conditions. Second, design modifications to accommodate the necessarily large SCR catalysts may be difficult and costly. Further, there are many as yet unresolved issues regarding infrastructure changes that would be necessary to address the storage and refilling of the reductant supply on vehicles. Nonetheless, there is extensive research going on in the development and improvement of applying SCR to heavy-duty vehicles.

Proposed Monitoring Requirements

The proposed regulation would require monitoring of SCR catalyst systems for proper NOx conversion performance. Specifically, the staff is proposing that manufacturers be required to indicate a malfunction when the conversion efficiency decreases to a point that emissions exceed 1.75 times the applicable NOx (or if applicable, NMHC+NOx) emission standard. If a malfunctioning catalyst cannot cause emissions to exceed the emission threshold of 1.75 times the applicable standards, a manufacturer would only be required to functionally monitor the system and indicate a malfunction when no conversion efficiency of the emission(s) of concern could be detected. The mechanism for adding the fuel reductant must also be monitored for proper function. Since precise control of reductant addition is essential for good NOx performance from the SCR system, manufacturers would be required to continuously monitor the SCR system while the system is in operation. Additionally, through the 2009 model year, no monitoring would be required if the conversion efficiency of the catalyst system is less than 30 percent.

Technical Feasibility of Proposed Monitoring Requirements

As mentioned earlier, current NOx sensor technology tends to have a cross-sensitivity to ammonia (i.e., as much as 65 percent of ammonia can be read as NOx).[14] Although this cross-sensitivity can be detrimental to SCR controls (i.e., reductant injection/NOx reduction efficiencies), it is actually beneficial for monitoring purposes. Monitoring of the catalyst can be done by using the same NOx sensors that are used for SCR control. When the SCR catalyst is functioning properly, the upstream sensor should read high (for high NOx levels) while the downstream sensor should read low (for low NOx and low ammonia levels). With a deteriorated SCR catalyst, the downstream sensor should read similar values as the upstream sensor or higher (i.e., high NOx and high ammonia levels) since the NOx reduction capability of the catalyst has diminished. Therefore, a malfunctioning SCR catalyst could be detected when the downstream sensor output is equal to or greater than the upstream sensor output.

Monitoring of the fuel reductant injection functionality could be done in a manner similar to that for lean NOx catalyst monitoring. The same temperature sensor that is used for control purposes could also be used for monitoring the injection. With proper injection, the catalyst should see a temperature increase afterwards. In addition, the

NOx sensors that are used for control purposes could be used to monitor the reductant injection. With a properly functioning injector, the downstream NOx sensor should see a change from high NOx levels to low NOx levels. In contrast, a lack of reductant injection would result in continuously high NOx levels at the downstream NOx sensor. Therefore, a malfunctioning injector could be found when the downstream NOx sensor continues to measure high NOx after an injection event has been commanded.

F. NOx ADSORBER/TRAP MONITORING

Background

NOx adsorbers (also called NOx traps) are another NOx control technology that has been experiencing significant progress in development and optimization. This is one of the newer technologies being optimized for use in diesel vehicles as well as lean-burn gasoline vehicles. NOx adsorber systems generally consist of a conventional three-way catalyst (e.g., platinum) with NOx storage components (i.e., adsorbents) incorporated into the washcoat. The concept of the NOx adsorber involves the trapping, release, and reduction of NOx from the exhaust stream in the catalyst washcoat. The adsorbers chemically bind (i.e., “trap”) the oxides of nitrogen during lean engine operation. Generally, when the storage capacity of the adsorbers is saturated, regeneration occurs and the stored NOx is released and converted. This occurs under rich engine operation and includes the chemical reduction of the released NOx to nitrogen by carbon monoxide, hydrogen, and hydrocarbons on a precious metal site. The rich running conditions, which generally last for several seconds, are typically achieved using a combination of intake air throttling (to reduce the amount of intake air), exhaust gas recirculation, and post-combustion fuel injection.

NOx adsorber systems have demonstrated NOx reduction efficiencies from 50 percent to in excess of 80 to 90 percent. This efficiency has been found to be highly dependent on the fuel sulfur content because NOx adsorbers are extremely sensitive to sulfur. The NOx adsorption material has a greater affinity for sulfur compounds than NOx. Thus, sulfur compounds can saturate the adsorber and limit the number of active sites for NOx adsorption, thereby lowering the NOx reduction efficiency. Accordingly, low sulfur fuel is required to achieve the greatest NOx reduction efficiencies. Although new adsorber washcoat materials are being developed with a higher resistance to sulfur poisoning and ultra-low sulfur fuel will be required in the future, it is projected that NOx adsorber systems will still be subject to sulfur poisoning and will require a sulfur regeneration mechanism.[18] Sulfur poisoning, however, is generally reversible through a desulfurization process, which requires high temperatures (i.e., 500 to 700 degrees Celsius) accompanied by a rich fuel mixture that can be achieved with post-injection and installation of a light-off catalyst upstream of the NOx adsorber. Because the sulfur regeneration process takes much longer (e.g., several minutes) and requires more fuel and heat than the NOx regeneration step, permanent thermal degradation of the NOx

---

[18] Bailey, O., H., Dou, D., and Molinier, M., “Sulfur Traps for NOx Adsorbers: Materials Development and Maintenance Strategies for Their Application,” SAE Paper 2000-01-1205; “NOx Adsorbers,” www.dieselnet.com.

adsorber and fuel economy penalties may result from too frequent sulfur regeneration. However, if regeneration is not done frequently enough, NOx conversion efficiency is compromised and fuel economy penalties will also be incurred from excessive purging of the NOx adsorber.[19]

Installation of sulfur traps upstream of the NOx adsorber can help in alleviating sulfur poisoning problems. The sulfur trap is essentially an adsorber catalyst aimed at trapping sulfur compounds. Similar to the NOx adsorber, once the sulfur trap becomes saturated, the trap must undergo sulfur regeneration. Unfortunately, depending on the temperatures, this regenerated sulfur may be re-adsorbed downstream in the NOx adsorber, so strategies must be carefully developed to minimize this effect (e.g., allowing sulfur trap regeneration to occur less frequently than NOx adsorber regeneration or using bypass valves).

In order to achieve and maintain high NOx conversion efficiencies while limiting negative impacts on fuel economy and driveability, vehicles with NOx adsorption systems will require precise air/fuel control in the engine and in the exhaust stream. Many of these control strategies are still undergoing rapid development. However, diesel manufacturers are expected to utilize NOx sensors and temperature sensors to provide the most precise closed-loop control for the NOx adsorber system.[20] These sensors will provide the adsorber control system with valuable information regarding the NOx levels, oxygen levels/air-fuel ratio, and adsorber temperatures that are needed to achieve and maintain the highest NOx conversion efficiencies possible with minimum fuel consumption penalties during all types of operating conditions. Further, these same sensors can also be used to monitor the adsorber system as will be described later.

Alternatively, if NOx sensors are not used to control the NOx adsorber system, it is projected that wide-range air-fuel (A/F) sensors (located upstream and downstream of the adsorber) can be used effectively as a substitute. A/F sensors are currently used by one manufacturer on a gasoline-fueled vehicle equipped with a NOx adsorber system to control and monitor the system, and at least one other gasoline-fueled vehicle manufacturer plans to introduce a similar system soon. Although manufacturers have previously expressed concerns regarding the durability of A/F sensors in diesel applications, these concerns apparently have been sufficiently addressed since at least one diesel manufacturer has indicated plans to introduce wide-range A/F sensors on a 2004 model year vehicle for EGR control. On diesel applications, A/F sensors have several advantages over NOx sensors including lower cost, wide availability, and a mature technology. However, A/F sensors cannot provide an instantaneous indication of tailpipe NOx levels, which would allow the control system to precisely determine when the trap system is filled to capacity and regeneration should be initiated. If A/F sensors are used in lieu of NOx sensors, an estimation of NOx engine-out emissions and their subsequent storage in the NOx adsorber can be achieved indirectly through modeling. However, this may require significant development work.

---

[19] Ingram, G. A. and Surnilla, G., “On-Line Estimation of Sulfation Levels in a Lean NOx Trap,” SAE Paper 2002-01-0731.

[20] “NOx Adsorbers,” www.dieselnet.com.

Proposed Monitoring Requirements

To ensure the desired NOx emission levels are achieved throughout the engine's useful life, the NOx adsorber must maintain a high conversion efficiency. Therefore, the staff is proposing that manufacturers monitor the NOx adsorber for proper performance. The staff is proposing that manufacturers be required to indicate a malfunction when the conversion efficiency decreases to a point such that emissions exceed 1.5 times any of the applicable standards. If a malfunctioning NOx adsorber cannot cause emissions to exceed the emission threshold of 1.5 times the applicable standards, a manufacturer would only be required to functionally monitor the system and indicate a malfunction when no NOx conversion efficiency could be detected. Since the performance of NOx adsorbers is critical to complying with stringent emission standards, manufacturers would be required to continuously monitor the NOx adsorber during a driving cycle in which the monitoring conditions are met.

Additionally, due to the importance of desulfurization on the performance of the NOx adsorber, the NOx adsorber system diagnostic must be sufficiently robust to distinguish poor NOx conversion performance from temporary/reversible sulfur poisoning. Although manufacturers would not be required to separately monitor for proper desulfurization, manufacturers would be required to design their NOx adsorber diagnostic to be able to rule out temporary sulfur poisoning as the source of poor NOx conversion performance. If the NOx adsorber diagnostic continues to indicate poor performance after temporary sulfur poisoning has been ruled out (e.g., immediately after desulfurization), the adsorber system would be considered malfunctioning and the MIL would be illuminated.

Technical Feasibility of Proposed Monitoring Requirements

As mentioned earlier, either NOx sensors or A/F sensors along with a temperature sensor are projected to be used for controlling the NOx adsorber system. These same sensors could also be used to monitor the adsorber system. The use of NOx sensors placed upstream and downstream of the adsorber system would allow the system's NOx reduction performance to be continuously monitored. For example, the upstream NOx sensor on a properly functioning adsorber system operating with lean fuel mixtures, will read high NOx levels while the downstream NOx sensor should read low NOx levels. With a deteriorated NOx adsorber system, the upstream NOx levels will continue to be high while the downstream NOx levels will also be high. Therefore, a malfunction of the system can be detected by comparing the NOx levels measured by the downstream NOx sensor versus the upstream sensor. With further development, there is a possibility that manufacturers will be able to model the upstream NOx levels (based on other engine operating parameters such as engine speed, fuel injection quantity and timing, EGR flow rate), thereby eliminating the need for the front NOx sensor for both control and monitoring purposes.

Alternatively, if NOx sensors are not used by the adsorber system for control purposes, monitoring of the system could be conducted by using wide-range A/F sensors to replace one or both of the NOx sensors.[19] Under lean engine operation conditions with a properly operating NOx adsorber system, both the upstream and downstream A/F sensors will indicate lean mixtures. However, if the exhaust gas is intrusively commanded rich, the upstream A/F sensor will quickly indicate a rich mixture while the downstream O2 sensor should continue to see a lean mixture in the exhaust due to the release and reduction of NO_2 in the adsorber. Once all of the stored NO_2 has been reduced, the downstream A/F sensor will indicate a rich reading. The more NOx that is stored in the adsorber, the longer the delay before the downstream A/F sensor indicates a rich exhaust gas. Thus, the time differential between the upstream and downstream A/F sensors' lean-to-rich indication is a gauge of the NOx adsorption capability of the adsorber and can be calibrated to indicate different levels of performance. Fresh NOx adsorber systems will have the highest NOx adsorption capability and consequently the longest "lean-to-rich switch" time differential while deteriorated adsorbers with no adsorption capability will have the shortest time differential. Therefore, the NOx adsorber system could be monitored by calibrating the lean-to-rich time differential to indicate a fault when the NOx adsorber system has deteriorated to a level such that 1.5 times the emission standards would be exceeded. Honda currently utilizes A/F sensors in a similar manner as described above to monitor the NOx adsorber on a 2003 model year gasoline vehicle.

Since sulfur poisoning reversibly diminishes the performance of the NOx adsorber system, it is imperative that sulfur poisoning be distinguished from a true deteriorated system. Otherwise, perfectly good NOx adsorber systems could erroneously be identified as being bad (i.e., false MILs could occur). Manufacturers of gasoline vehicles with NOx adsorber systems are aware of this issue and are taking various measures to account for adsorber sulfation. These approaches should also work on diesel vehicles. Basically, the monitoring method works on several phenomena. As sulfation of the adsorber increases, the NOx adsorption capacity of the system progressively decreases. When the NOx adsorption capacity decreases past a predetermined threshold, a desulfation event is intrusively commanded (e.g., with an external heat source or rich fuel mixture) to sufficiently heat up the adsorber for sulfur removal. After desulfation, the adsorber system's NOx capacity is again reevaluated. If the NOx capacity is now below the predetermined threshold, the NOx adsorber is judged good and the previous deteriorated result was due to sulfur poisoning. However, if the NOx capacity is still below the threshold, the NOx adsorber is truly bad and the MIL should be commanded on and a fault code identifying the deteriorated adsorber stored.

G. PARTICULATE MATTER (PM) TRAP MONITORING

Background

As indicated earlier, the particulate matter (PM) emission standards for the 2007 model year will be reduced by 90 percent from the 2004 model year standards. In order to meet the increasingly stringent standards, manufacturers will likely use aftertreatment devices such as PM traps to achieve the necessary emission levels. PM traps are considered the most effective control technology for the reduction of particulate emissions and can typically achieve PM reductions in excess of 90 percent. In general, a PM trap consists of a filter material that permits exhaust gases to pass through but traps the PM emissions. In order to maintain the performance of the PM trap and the vehicle, the trapped PM must be periodically removed before too much particulate is accumulated and exhaust backpressure reaches unacceptable levels. The process of periodically removing accumulated PM from the trap is known as regeneration and is very important for maintaining low PM emission levels. PM trap regeneration can be passive (i.e., occur continuously during regular operation of the filter), active (i.e., occur periodically after a predetermined quantity of particulates have been accumulated), or a combination of the two. With passive regeneration, oxidation catalyst material is typically placed on the PM trap system to lower the temperature for oxidizing PM. This allows the trap to continuously oxidize trapped PM material during normal driving. In contrast, active systems utilize an external heat source such as an electric heater or fuel burner to facilitate PM trap regeneration. It is projected that virtually all PM trap systems will have some sort of active regeneration mechanism.

One of the key factors that needs to be taken into account for a trap regeneration control system is the amount of soot quantity that is stored in the PM trap (often called soot loading).[21] If too much soot is stored in the PM trap when regeneration is activated, the soot can burn uncontrollably and damage the filter. However, activating regeneration when there is too little trapped soot is also undesirable since there is a minimum amount of soot quantity needed to ensure good combustion propagation. Another important factor to be considered in the control system design is the fuel economy penalty involved with trap regeneration. Prolonged operation with high backpressures in the exhaust and too frequent regenerations are both detrimental to fuel economy and durability. Therefore, trap designers will need to carefully balance the regeneration frequency with various conflicting factors. In order to optimize the trap regeneration for these design factors, the control system for the regeneration system is projected to utilize both pressure sensors and temperature sensors to model soot loading among other phenomena.[21] Through the information provided by these sensors, designers can optimize the PM trap for high effectiveness and maximum durability while minimizing fuel economy and performance penalties.

Proposed Monitoring Requirements

Regardless of the regeneration method, regeneration must be monitored by the OBD system since this process is vital in maintaining the performance of the PM trap. Thus, the staff is proposing to require manufacturers to monitor PM traps for proper performance of the regeneration process and proper performance of the trap itself. Manufacturers would be required to indicate a PM trap malfunction when the trapping capability decreases to a point such that emissions exceed 1.5 times any of the applicable standards. Malfunction modes that are expected to be detected by the PM

---

[21] Salvat, O., Marez, P., and Belot, G., “Passenger Car Serial Application of a Particulate Filter System on a Common Rail Direct Injection Diesel Engine,” SAE Paper 2000-01-0473.

trap monitor include clogged filters, cracked filters, empty filter cans, and any other malfunctions that can affect emissions to the prescribed levels. If a malfunctioning PM trap cannot cause emissions to exceed the emission threshold of 1.5 times the standards, a manufacturer would only be required to perform functional monitoring of the system and indicate a malfunction when no PM trap capability could be detected. Since the proper performance of the PM trap is integral to complying with stringent emission standards, manufacturers would be required to monitor the PM trap continuously for the presence of a malfunction. For PM trap regeneration, manufacturers would be required to verify that regeneration has actually occurred when driving conditions allowing for its occurrence have been achieved or when commanded on.

Technological Feasibility of Proposed Monitoring Requirements

It is anticipated that manufacturers will not need additional hardware to meet the PM trap monitoring requirements. The same pressure and temperature sensors that are used to control trap regeneration are projected to be used for monitoring. In general, a pressure sensor placed upstream of the trap and at least one temperature sensor located near the PM trap are used for the control system. As mentioned earlier, pressure sensors are expected to be used on PM trap systems to prevent damage due to delayed or incomplete regeneration that could lead to excess temperatures. When a pressure sensor placed upstream of the trap senses excessively high pressures, active regeneration can be activated. The same pressure sensor could also be used to determine if regeneration is functioning correctly and to evaluate the suitability of the trap for controlling particulate emissions. For example, after a regeneration event, the backpressure should drop significantly since the trapped soot and particles are removed. If backpressure does not drop within the range expected after a regeneration event, the regeneration did not function correctly and the OBD system would alert the vehicle operator of a problem. Also, backpressure on a normal PM trap should progressively increase as the mass of soot and trapped particles increase. In general, the mass of soot and trapped particles should increase as the mileage traveled or time of operation increase. However, a cracked filter or missing filter may not increase backpressure as expected. Therefore, a cracked or missing filter can be detected if the backpressure fails to increase at the rate projected by the soot-loading model. One European vehicle manufacturer has incorporated PM trap monitoring on their PM trap-equipped vehicles since 2000.

As mentioned earlier, manufacturers are projected to also use temperature sensors for regeneration control purposes. As an additional benefit, this same sensor could also be used on these systems to monitor active regeneration of the trap. If excess temperatures are seen by the temperature sensor during active regeneration, the regeneration process can be stopped or slowed down to protect the trap. If active regeneration is commanded on and there isn’t a sufficient temperature rise in the PM trap system for the amount of soot stored in the trap, the regeneration system is malfunctioning and the OBD system would alert the driver of a problem.

IV. TECHNICAL STATUS AND PROPOSED MONITORING SYSTEM

REQUIREMENTS FOR GASOLINE/SPARK-IGNITED ENGINES

A. FUEL SYSTEM MONITORING

Background

An important component in emission control on gasoline engines is the fuel system. Proper delivery of fuel is essential to maintain stoichiometric operation and minimize engine out emissions. Proper stoichiometric control is also critical to maximize catalyst conversion efficiency and reach low tailpipe emission levels. As such, thorough monitoring of the fuel system is an essential element in an OBD system.

For gasoline engines, the fuel system generally includes a fuel pump, fuel pressure regulator, fuel rail, individual injectors for each cylinder, and a closed-loop feedback control system using oxygen sensor(s) or wide-range lambda sensor(s). The feedback sensors are located in the exhaust system and are used to regulate the fuel injection quantity to achieve a stoichiometric mixture in the exhaust. If the sensor indicates a rich (or lean) mixture, the system reduces (or increases) the amount of fuel being injected by applying a short term correction to the fuel injection quantity calculated for the current engine operation condition. To account for aging or deterioration in the system such as reduced injector flow or vacuum leaks that introduce excess air, more permanent long term corrections are also learned and applied to the fuel injection quantity for more precise fueling.

Proposed Monitoring Requirements

For gasoline engines, fuel system monitoring has been implemented on light- and medium-duty vehicles from the 1996 model year under the OBD II regulations. For heavy-duty gasoline engines (many of which are the same engine used in lighter medium-duty applications), the system components and control strategies are identical to those used in the light- and medium-duty categories. As such, the monitoring requirements established for light- and medium-duty engines can be directly applied to heavy-duty gasoline engines.

The staff is proposing that the fuel system be continuously monitored for its ability to maintain engine emissions below the standards. Manufacturers would be required to detect a malfunction when the system can no longer achieve this. Since the systems are essentially “self-correcting” and adapt for deterioration, monitoring of the system is accomplished by looking at the adaptive terms (e.g., short term and long term fuel trim) and indicating a fault when the corrections get so large (or reach their adaptive limits) that emissions cannot be maintained below the emission standard. Manufacturers would also be required to verify that the fuel system is in closed-loop operation (e.g., is using the oxygen sensor for feedback and can make changes to the adaptive correction values). Manufacturers have a pre-defined set of criteria that must be satisfied to begin closed-loop operation which typically include a minimum time after engine start, a minimum engine coolant temperature, and some indication that the oxygen sensor is warmed-up and ready. Manufacturers would typically meet this requirement with separate diagnostics that verify each individual criterion is satisfied (which also provides valuable diagnostic information to help repair technicians pinpoint the root cause of the malfunction).

The individual components of the fuel system would also be covered by separate monitoring requirements for oxygen sensors, misfire (for the fuel injectors), and comprehensive components (in systems such as those with electronically-controlled variable speed fuel pumps or electronically-controlled fuel pressure regulators).

Technical Feasibility of Proposed Monitoring Requirements

For gasoline engines, the light- and medium-duty OBD II regulations have required identical fuel system monitoring since the 1996 model year. Over 84 million cars have been built and sold in the U.S. to these fuel system monitoring requirements including medium-duty vehicles which utilize the exact same gasoline engines that are also used in some heavy-duty vehicle applications. The technical feasibility has clearly been demonstrated for these packages.

B. MISFIRE MONITORING

Background

One of the primary causes of catalyst degradation is engine misfire, which is the lack of combustion due to the absence of spark or poor fuel metering, among other causes. When misfire occurs, unburned fuel and air are pumped into the catalyst, greatly increasing its operating temperature (where the temperature can soar to above 900 degrees Celsius). This problem is usually most severe under high load, high speed engine operating conditions, causing irreversible damage to the catalyst. Though the durability of catalysts has been improving, most are unable to sustain continuous operation at such high temperatures. Engine misfire also contributes to excess emissions, especially when the misfire is present during engine warm-up and the catalyst has not reached its operating temperature.

Proposed Monitoring Requirements

Accordingly, for gasoline engines, the staff is proposing continuously monitoring for engine misfire at all positive torque engine speeds and load conditions. Additionally, manufacturers would be required to identify a misfiring cylinder or indicate if multiple cylinder misfiring is occurring (through the storage of the appropriate fault codes). With regards to catalyst-damaging misfire, manufacturers would be required to determine the level (i.e., percentage) of misfire per 200 revolution increments (e.g., two seconds at 6000 rpm) for each engine speed and load condition that would result in a temperature that causes catalyst damage. The proposed regulation would establish a specific means of determining the temperature at which catalyst damage occurs. With regards to misfire that can cause excess emissions, manufacturers would be required to determine the level of misfire per 1000 revolution increments that would result in emissions exceeding 1.5 times the applicable standards. To establish this percentage of misfire, manufacturers would utilize misfire events occurring at equally spaced, complete engine cycle intervals, across randomly selected cylinders throughout each 1000-revolution increment. The staff is also proposing to set a lower limit on the level of misfire that is required to be detected (i.e., five percent for misfire causing catalyst damage, and one percent for misfire causing emissions to exceed 1.5 times the standards), due to increased difficulty in diagnosing misfire at such low percentages.

Although the proposal would require misfire monitoring to occur continuously for gasoline engines, the proposed regulation would allow manufacturers to temporarily disable misfire monitoring during certain operating conditions where misfire cannot be reliably detected. These conditions include driving on rough roads, during manual transmission gear changes, and during extremely rapid throttle changes. Manufacturers that want to disable misfire monitoring during conditions not specifically stated in the proposed regulation would be required to request Executive Officer approval of such disablement. Some manufacturers may request disablement during a certain amount of time from engine start-up (end of crank), since they may contend that such conditions may cause unreliable misfire detection. The staff, however, is concerned that misfire could occur during start-up (i.e., during cold start when the engine can run rough) and then cease once warming of the engine has occurred. Such misfire problems would significantly impact emissions, since the catalyst would not have reached its operating temperature. Thus, the proposed regulation would require misfire monitoring to occur no later than the end of the second crankshaft revolution after engine start-up.

Technical Feasibility of Proposed Monitoring Requirements

For gasoline engines, the light- and medium-duty OBD II regulations have required identical misfire monitoring requirements since the 1996 model year. One of the most reliable methods for detecting misfire that has been demonstrated is the use of a crankshaft position sensor, which would measure the fluctuations in engine angular velocity and determine if misfire exists, and a camshaft position sensor, which can be used to identify the misfiring cylinder. This method has been shown to be technically feasible for misfire monitoring on light- and medium-duty vehicles.

C. EXHAUST GAS RECIRCULATION (EGR) SYSTEM MONITORING

Background

Exhaust gas recirculation (EGR) is one of the most effective emission control technologies for reducing NOx emissions in vehicles today. Generally, NOx emissions are formed under high combustion chamber temperature and pressure conditions. EGR systems redirect spent combustion gases from the exhaust stream to the intake system to dilute the oxygen concentration and increase the heat capacity of the air/fuel charge. This effectively reduces the combustion temperature, which results in lower levels of

NOx emissions. EGR systems can involve many components to ensure accurate control of EGR flow, including valves, valve position sensors, and actuators.

Proposed Monitoring Requirements

The EGR system would need to be monitored to ensure that the appropriate amount of EGR flow reaches the intake system. The staff is proposing that manufacturers be required to indicate an EGR system malfunction when the EGR flow rate increases or decreases to a point where emissions exceed 1.5 times the application standards. While decreased EGR flow can cause increased emissions, excessive EGR flow can also cause increased emissions and driveability problems. Manufacturers would be required to monitor the EGR flow rate at least once per driving cycle in which the monitoring conditions are met. If a malfunctioning EGR system (with a reduced flow or excessive flow fault) cannot cause emissions to exceed the emission threshold of 1.5 times the applicable standards, a manufacturer would only be required to perform functional monitoring of the malfunction of concern (e.g., indicate a malfunction when no detectable amount of EGR flow is detected). The individual electronic components utilized by the EGR system would be monitored under the comprehensive components monitoring requirements.

Technical Feasibility of Proposed Monitoring Requirements

The light- and medium-duty OBD II regulations have required identical EGR system monitoring since the 1996 model year. Manufacturers have been detecting malfunctions of EGR flow rate generally by looking at the change in fuel trim or manifold pressure under conditions when the EGR system is active. The technical feasibility of EGR monitoring has already been demonstrated for these applications.

D. COLD START EMISSION REDUCTION STRATEGY MONITORING

Background

The largest portion of exhaust emissions from gasoline vehicles is generated during the brief period following a cold start before the engine and catalyst have warmed up. In order to meet increasingly stringent emission standards, manufacturers are developing hardware and associated control strategies to reduce these emissions. Most efforts are centering around reducing catalyst warm-up time. A cold catalyst is heated mainly by two mechanisms - heat transferred from the exhaust gases and heat that is generated in the catalyst as a result of the catalytic reactions.

Manufacturers are implementing various hardware and control strategies to quickly light off the catalyst (i.e., reach the catalyst temperature at which 50 percent conversion efficiency is achieved). Most manufacturers use substantial spark retard and/or increased idle speed to maximize the heat available in the exhaust following a cold start to quickly light off the catalyst. However, customer satisfaction and safety (i.e., vehicle driveability and engine idle quality) limit the amount of spark retard or increased idle speed that a manufacturer will use to accelerate catalyst light off. On a normally functioning vehicle, engine speed drops when the ignition timing is retarded, therefore causing the idle speed control system to compensate and allow more airflow (with a corresponding increase in fuel) to the engine in order to maintain idle speed stability during spark retard. Since idle quality is given a high priority, spark retard is typically limited to an extent that the idle control system can quickly respond to and maintain idle quality. Conversely, a deteriorated or poorly responding idle control system would reduce the capability of the engine to compensate and may cause the on-board computer to command less spark retard than would normally be achieved for a properly functioning system, thereby causing delayed catalyst light off and higher emissions. Though the proposed regulation would require monitoring of the idle control system and monitoring of the ignition system by the misfire monitor, the idle control system is normally monitored only after the engine has warmed up, and malfunctions that occur during cold start may not be detected by the OBD system, yet have significant emission consequences.

Additionally, given the escalating cost of precious metals, there is an industry trend to minimize their use in catalysts. To compensate for the reduction in catalyst performance, manufacturers will likely employ increasingly more aggressive cold start emission reduction strategies. It is crucial that these strategies be successful and properly monitored in order to meet the new, more stringent emission standards and to maintain low emissions in-use.

Proposed Monitoring Requirements

Considering the issues outlined above, the staff is proposing a requirement to monitor the individual components used to implement cold start emission reduction strategies. This would ensure that the target conditions necessary to reduce emissions or catalyst light-off time are indeed achieved and emissions do not exceed 1.5 times the emission standard. These components would need to be monitored while the strategy is active. For example, if the target idle speed for catalyst light-off could not be achieved or maintained adequately to maintain emissions below 1.5 times the standard, a malfunction would need to be indicated. Similarly, if the target spark retard necessary for catalyst light-off could not be achieved due to an idle control system malfunction, a fuel system malfunction, or any other malfunction, a fault would need to be indicated.

Technical Feasibility of Proposed Monitoring Requirements

Monitoring techniques that are projected to be used for cold start monitoring strategies would be similar to those already outlined during the light- and medium-duty OBD rulemaking, which mainly involve software modifications. For example, if ignition retard is used during cold starts, the commanded amount of ignition retard would have to be monitored if the amount of timing retard can be restricted by external factors such as idle quality or driveability. This can be done with software algorithms that compare the actual overall commanded final ignition timing with the threshold timing that would result in emissions that exceed 1.5 times the standard. Cold start strategies that always command a predetermined amount of ignition retard independent of all other factors and do not allow idle quality or other factors to override the desired ignition retard do not require monitoring of the commanded timing. Other methods to ensure the actual timing has been reached include verifying other factors such as corresponding increases in mass air flow and idle speed indicative of retarded spark combustion. Since mass air flow and idle speed are both currently used by the engine control system and the OBD system, only minor software modifications should be required to further analyze these signals while the cold start strategy is invoked.

As required for other OBD monitors, the stored fault code would, to the fullest extent possible, be required to pinpoint the likely cause of the malfunction to assist technicians in diagnosing and repairing these malfunctions. The proposal would also allow a manufacturer to develop calibrations on representative vehicles and apply the calibrations to the remainder of the product line.

E. SECONDARY AIR SYSTEM MONITORING

Background

Secondary air systems, which are expected to be utilized only on gasoline vehicles, are used to reduce cold start exhaust emissions of hydrocarbons and carbon monoxide. Although many of today's vehicles operate near stoichiometric (where the amount of air is just sufficient to completely combust all of the fuel) after a cold engine start, more stringent emission standards may require secondary air systems, generally in combination with a richer than stoichiometric cold start mixture, to quickly warm up the catalyst for improved cold start emission performance. Secondary air systems typically consist of an electric air pump, various hoses, and check valves to deliver outside air to the exhaust system upstream of the catalytic converters. This system usually operates only after a cold engine start for a brief period of time. When the electric air pump is operating, fresh air is delivered to the exhaust system and mixes with the unburned fuel at the catalyst, so that the fuel can burn and rapidly heat up the catalyst. Problems with the secondary air systems that may be found in the field include corroded check valves, damaged tubing and hoses, and malfunctioning air switching valves. Given the importance of properly functioning secondary air systems to emission performance, monitoring is needed.

Proposed Monitoring Requirements

The secondary air system would have to be monitored to verify secondary air delivery to the exhaust system during cold engine starts when it is normally active. Thus, the staff is proposing that manufacturers be required to monitor the proper functioning of the secondary air delivery system including all air switching valves. Specifically, a manufacturer would be required to indicate a malfunction prior to a decrease from the manufacturer's specified air flow during normal operation (e.g., during vehicle warm-up following engine start) that would cause a vehicle's emissions to exceed 1.5 times the applicable standards. Manufacturers would be required to monitor the secondary air system at least once per driving cycle in which the monitoring conditions are met. If a malfunctioning secondary air system cannot cause emissions to exceed the emission threshold of 1.5 times the applicable standards, a manufacturer would only be required to perform functional monitoring of the system by indicating a malfunction when no detectable amount of air flow is delivered during normal operation. The individual electronic components utilized by the secondary air system would be monitored under the comprehensive components monitoring requirements.

Technical Feasibility of Proposed Monitoring Requirements

In order for the OBD system to effectively monitor the secondary air system when it is normally active, linear oxygen sensors (often referred to as wide-range oxygen sensors or air-fuel ratio sensors) would most likely be required. These sensors are currently installed on many new cars and their implementation is projected to increase in the future as more stringent emission standards are phased in. Linear oxygen sensors are useful in determining air-fuel ratio over a broader range than conventional oxygen sensors and are especially valuable for controlling fueling in lean-burn engines and other engine designs that require very precise fuel control. Since linear oxygen sensors are able to determine air-fuel ratio accurately, the amount of secondary airflow needed to keep emissions below 1.5 times the tailpipe emission standard could be correlated to the air-fuel ratio, making linear oxygen sensors useful for secondary air system monitoring.

F. CATALYST MONITORING

Background

Three-way catalysts are one of the most important emission-control components utilized by gasoline engines. They consist of ceramic or metal honeycomb structures (i.e., “substrates”) coated with precious metals such as platinum, palladium, or rhodium. These precious metals are dispersed within an alumina washcoat containing ceria, and the substrates are mounted in a stainless steel container in the vehicle exhaust system. Three-way catalysts are so-designated because they are capable of simultaneously oxidizing HC and CO emissions into water and carbon dioxide, and of reducing NOx emissions (by reacting with CO and hydrogen) into elemental nitrogen, carbon dioxide, and water.

This three-way conversion activity only takes place efficiently, however, when the fuel system operates at stoichiometric (i.e., the air/fuel ratio where there is just the required amount of air to completely burn all of the fuel in the engine). Manufacturers achieve and maintain stoichiometric fuel delivery by incorporating closed-loop fuel control systems that utilize an exhaust gas oxygen sensor to provide feedback on the status of the air-fuel ratio being achieved. Most closed-loop fuel control systems actively cycle the air-fuel ratio slightly above and below the stoichiometric point to maximize three-way catalyst conversion efficiency. The precious metals are used to temporarily retain the HC, CO, and NOx molecules in the catalyst and promote the chemical reactions while the ceria in the washcoat is used to store and release oxygen that is needed to complete the reactions. Oxygen is stored in the catalyst during the lean portion of the fuel system's cycling (i.e., when the air-fuel ratio is slightly higher than stoichiometric) and is released during the rich excursion.

While improvements to catalysts over the years have increased their durability, they are still subject to high temperature deterioration that occurs when excess air and fuel enter the catalyst. This can be caused by misfire (i.e., unburned fuel and air that are pumped into the catalyst) among other factors, and will result in reduced catalyst conversion efficiency. Catalyst performance can also deteriorate due to catalyst deactivation from poisoning (e.g., lead, phosphorus). Additionally, catalysts can also fail due to mechanical problems, such as excessive vibration or damage to the catalyst itself.

Proposed Monitoring Requirements

Due to the importance of the catalyst system in a vehicle's emission control system, the staff is proposing monitoring for proper catalyst system performance. Specifically, manufacturers would be required to indicate a catalyst malfunction when the catalyst system's conversion capability decreases to a point that emissions exceed 1.75 times the applicable HC or NOx standards. The staff is proposing that the catalyst monitor run at least once per driving cycle. Manufacturers that utilize multiple catalyst systems would only be required to conduct catalyst OBD monitoring on catalysts exposed to untreated exhaust gas (except for bypass catalysts). These catalysts are most likely to be damaged and would provide the earliest indication of a catalyst system problem. Replacement of these catalysts alone would also restore a high conversion efficiency to the system since the majority of emissions occur during a cold start and the forward catalysts are the most important for controlling cold start emissions.

When determining the proper OBD malfunction threshold for catalysts, manufacturers would progressively deteriorate or "age" catalysts (by replicating excessive temperature conditions via oven aging or misfire aging) to the point where emissions exceed 1.75 times the standard. Thus, the staff is also proposing specific requirements for catalyst aging and determining the malfunction thresholds for the catalyst monitor. Specifically, manufacturers would be required to use deterioration methods that more closely represent real world deterioration, thereby ensuring that the MIL would illuminate at the appropriate emission level during real world operation. The proposal would further require that the catalyst system be aged as a whole (i.e., manufacturers would simultaneously age the monitored and unmonitored catalysts) to the malfunction criteria. This accounts for the fact that the unmonitored catalysts could also experience some real world deterioration. However, manufacturers that use fuel shutoff to misfiring cylinders in order to minimize catalyst over-temperature would be allowed to age the monitored catalyst to the malfunction criteria and the unmonitored catalysts to the end of the useful life. Such systems are less likely to be subjected to extreme temperatures, so they would likely age with the monitored catalyst experiencing most of the deterioration.

Technical Feasibility of Proposed Monitoring Requirements

A common method used for estimating catalyst efficiency is to measure the catalyst's oxygen storage capacity. This monitoring method is utilized by all current light- and medium-duty gasoline vehicles since the OBD II regulation was first fully implemented in the 1996 model year. Generally, as the catalyst's oxygen storage capacity decreases, its conversion efficiency of HC and NOx also decreases. With this strategy, a catalyst malfunction would be detected when its oxygen storage capacity has deteriorated to a predetermined level. Manufacturers could determine this by utilizing the information from the upstream oxygen sensor and a second oxygen sensor located downstream of the monitored portion of the catalyst (this second sensor is also used for trimming the front sensor to maintain precise fuel control). By comparing the level of oxygen measured by the second sensor with that measured by the primary sensor located upstream of the catalyst, manufacturers determine the oxygen storage capacity of the catalyst and thus, estimate the conversion efficiency. With a properly functioning catalyst, the second oxygen sensor signal will be fairly steady since the fluctuating oxygen concentration (due to the fuel system cycling about stoichiometric) at the inlet of the catalyst is damped by the storage and release of oxygen in the catalyst. When a catalyst is deteriorated, such damping is reduced, causing the frequency and peak-to-peak voltage of the second oxygen sensor to simulate the signal from the front oxygen sensor because the catalyst is no longer capable of storing and releasing oxygen.

G. EVAPORATIVE SYSTEM MONITORING

Background

In addition to emissions from a vehicle's tailpipe, ARB is concerned about emissions from a vehicle's evaporative system. Emissions that vent to the atmosphere through leaks in the evaporative system (e.g., disconnected evaporative system hoses) can be many times the evaporative emission standards. Additionally, evaporative purge system defects such as deteriorated vacuum lines, damaged canisters, and non-functioning purge control valves may occur, also resulting in high evaporative emissions.

Proposed monitoring requirements

Thus, the staff is proposing to require manufacturers to monitor the evaporative system separately for (1) small leaks equal to or greater than a 0.030 inch diameter hole; and (2) gross leaks equal to or greater than a 0.090 inch diameter hole. The 0.090 inch leak monitoring requirement is intended to quickly detect larger leaks such as split or disconnected evaporative system hoses or loose/missing gas caps, and thus would generally have much less restrictive monitoring conditions than the small leak (i.e., 0.030 inch) monitor. With regards to the orifice shape and length, the staff proposes the use of a specific orifice supplied by O'Keefe Controls Corporation, a manufacturer and supplier of precision orifices used by many in the industry. Orifices with equivalent specifications from other suppliers would also be acceptable. Additionally, the proposed regulation would require manufacturers to verify the purge flow from the vehicle canister system (i.e., to verify that the purge flow is actually reaching the engine and not venting into the atmosphere).

While the OBD II regulations have required leak detection for 0.020 inch leaks beginning with 2000 model year, light- and medium-duty manufacturers have found that fuel tanks larger than 25 gallons are extremely difficult to monitor to the leak sizes required by the OBD II regulation. To address this issue, the OBD II regulation contained a provision that allowed manufacturers to revise the leak size requirements for vehicles equipped with larger fuel tanks provided the manufacturer demonstrate the need for this allowance. Given that the vast majority, if not all, of the gasoline tanks in the heavy-duty industry are likely larger than 25 gallons, the staff evaluated the capability of the medium-duty manufacturers with large tanks and has accordingly proposed heavy-duty OBD monitoring only to 0.030 inch leaks in lieu of 0.020 inch leaks. The 0.030 inch detection level is consistent with the performance of most of the medium-duty applications and would ensure that all manufacturers are consistently designing and calibrating the system to an equivalent performance level.

Technical feasibility of proposed monitoring requirements

As mentioned above, the OBD II regulation has required monitoring of evaporative system leaks as small as 0.020 inches on light- and medium-duty vehicles for several years. These include medium-duty applications such as incomplete trucks and engine dynamometer certified configurations similar (and in many cases, identical) to the configurations used on heavy-duty applications. Applications successfully meeting the OBD II requirements have also included dual tank configurations as well as applications with up to 55 gallon tanks. Manufacturers have successfully implemented these requirements by utilizing monitoring techniques that create either a vacuum or pressurized condition in the fuel tank and evaporative system and check the change in vacuum/pressure over time. In general, these systems require the addition of an evaporative system pressure sensor and a canister vent valve capable of closing the vent line. In some cases, manufacturers have elected to add pressure pumps to generate a positive pressure in lieu of using the engine as a vacuum source. Further, in a few cases, manufacturers have implemented changes to the on-board computer to allow a portion of the control module to remain "on" even while the engine is off and monitor the natural vacuum and pressure fluctuations that occur in the system due to heating and cooling of the gasoline in the tank. Evaporative systems that have too large of a leak will be unable to build or hold pressure or vacuum for a sufficient amount of time and can be distinguished from systems without a leak.

Heavy-duty gasoline applications are expected to use near identical, if not identical, evaporative system components and the staff is not aware of any reason the existing monitoring techniques would not continue to work on heavy-duty applications. Thus, the technical feasibility of achieving the staff's proposed monitoring requirements has clearly been demonstrated by the successful use on medium-duty applications over the last several years.

V. TECHNICAL STATUS AND PROPOSED MONITORING SYSTEM REQUIREMENTS FOR ALL VEHICLES

A. VARIABLE VALVE TIMING AND/OR CONTROL (VVT) SYSTEM MONITORING

Background

Variable valve timing (VVT) and/or control systems are used primarily to optimize engine performance and have many advantages over conventional valve control. Instead of opening and closing the valves by fixed amounts, VVT controls can vary the valve opening and closing timing (as well as lift amount in some systems) depending on the driving conditions (e.g., high engine speed and load). This feature permits a better compromise between performance, driveability, and emissions than conventional systems. With more stringent NOx emission standards being phased in, more vehicles are anticipated to utilize VVT. By utilizing VVT to retain some exhaust gas in the combustion chamber to reduce peak combustion temperatures, NOx emissions are reduced. Manufacturers utilizing VVT are often able to remove external exhaust gas recirculation (EGR) valves and controls from their vehicles, offsetting the cost increase for the system.

Proposed monitoring requirements

Since valve timing can directly affect exhaust emissions, the staff is proposing specific requirements for monitoring VVT and/or control systems. In addition to monitoring the individual electronic components used in the VVT system, manufacturers would be responsible for detecting target errors and slow response malfunctions of these systems. For target error and slow response malfunctions, the diagnostic system would be required to detect malfunctions when the actual valve timing and/or lift deviates from the commanded valve timing and/or lift such that 1.5 times the applicable emission standard would be exceeded. For VVT and/or control systems that cannot cause emissions to exceed 1.5 times the standard, manufacturers would still be required to monitor the system for proper functional response under the comprehensive component requirements.

Technical feasibility of proposed monitoring requirements

VVT systems are already in general use in light- and some medium-duty applications. Further, under the OBD II requirements, such systems have been monitored for proper function on the applications that have used VVT systems since the 1996 model year. While monitoring under the OBD II regulation does not require monitoring to an emission threshold of 1.5 times the applicable standards until the 2006 model year, the same monitoring strategies that have generally been used to verify proper functionality are also expected to be used to monitor to the emission threshold. Such strategies include the use of the crank angle sensor and camshaft position sensor to confirm that the valve opening and closing occurs within an allowable tolerance of the commanded crank angle. By calculating the difference between the commanded valve opening crank angle and the achieved valve opening crank angle, a diagnostic algorithm could differentiate between a malfunctioning system with too large of an error and a properly functioning system with very little to no error. By calibrating the size of this error (or integrating it over time), manufacturers could design the system to indicate a malfunction prior to the required emission threshold. In the same manner, system response can be measured by monitoring the length of time necessary to achieve the commanded valve timing. To ensure adequate resolution between properly functioning systems and malfunctioning systems, most manufacturers only perform this type of check when a large enough "step change" in commanded valve timing occurs.

B. EXHAUST GAS SENSOR MONITORING

Background

Exhaust gas sensors (e.g., oxygen sensors, wide-range air-fuel (A/F) sensors, NOx sensors) are important to the emission control system of vehicles. In addition to maintaining the air-fuel ratio at stoichiometric, which helps achieve the lowest engine emissions, these sensors are also used for enhancing the performance of several emission control technologies (e.g., catalysts, EGR systems). Many modern vehicles traditionally perform fuel control with an oxygen sensor feedback system. In order for the emission control system to operate most efficiently, the air-fuel ratio must remain within a very narrow range (less than one percent deviation) around the stoichiometric ratio. Oxygen sensors are typically located in the exhaust system upstream and downstream of catalytic converters. The front or upstream oxygen sensor is generally used for fuel control, while the rear or downstream oxygen sensor is generally used for adjusting the front oxygen sensor as it ages and for monitoring the catalyst system. Many vehicles use wide-range A/F sensors, which provide a precise reading of fuel trim, in lieu of conventional oxygen sensors for fuel control and catalyst monitoring. Both of these sensors are expected to be used by the heavy-duty manufacturers to optimize their emission control technologies as well as satisfy many of the proposed heavy-duty OBD monitoring requirements, such as fuel system monitoring, catalyst monitoring, and EGR system monitoring. NOx sensors are also anticipated to be used for optimization of several diesel emission control technologies, such as lean NOx catalysts and selective catalytic reduction (SCR) systems. Since an exhaust gas sensor can be a critical component of a vehicle's fuel and emission control system, the proper performance of this component needs to be assured in order to maintain low emissions. Thus, it is important that any malfunction that adversely affects the performance of any of these exhaust gas sensors is detected by the OBD system.

Proposed monitoring requirements

The staff is proposing that a manufacturer be required to monitor the output voltage, resistance, impedance, response rate, and any other characteristic of an exhaust gas sensor that can affect emissions and/or other diagnostics. This requirement applies to both primary sensors (which are used for fuel control) and secondary sensors (which are used for control/feedback and monitoring of certain emission control technologies). Since proper fuel control and emission control system performance is essential in meeting the emission standards and maintaining low emissions, malfunctions where the system is unable to optimize these functions should be detected. Thus, manufacturers would also be required to indicate a malfunction when a sensor fault occurs such that the fuel system or an emission control system stops using the sensor as a feedback input. Additionally, for heated exhaust gas sensors, manufacturers would be required to monitor the heater for proper performance as well as circuit continuity faults.

Most of the exhaust gas sensor monitors (e.g., response rate) would be required to operate at least once per driving cycle. However, the staff is proposing that for circuit continuity faults, out-of-range values, and faults that prevent the sensor from being used as a feedback input, monitoring would be required to be continuous. While fuel system monitors may already be able to identify some of the oxygen and A/F sensor malfunctions, fuel system faults are generally one of the most difficult faults to diagnose and repair due to the substantial number of possible causes. As such, these requirements would help to pinpoint the oxygen or A/F sensor as the malfunctioning component if a circuit problem is occurring. A manufacturer may request Executive Officer approval to disable the continuous exhaust gas sensor monitoring when a sensor malfunction cannot be distinguished from other effects (e.g., disable out-of-range low oxygen sensor monitoring during fuel cut conditions).

Technical feasibility of proposed monitoring requirements

The light- and medium-duty OBD II regulations have required similar oxygen sensor monitoring since the 1996 model year. The technical feasibility has clearly been demonstrated for these packages. Additionally, A/F sensor monitoring has also been required and demonstrated on these vehicles for many years.

NOx sensors are a recent technology and currently still being developed and improved. However, the staff is expecting manufacturers would design their NOx sensor monitors to be similar to those of A/F sensors.

C. ENGINE COOLING SYSTEM MONITORING

Thermostat

Manufacturers typically use a thermostat to block the flow of coolant within the engine block during cold starts to promote rapid warming of the engine. As the coolant approaches a specific temperature, the thermostat begins to open and allows circulation of coolant through the radiator. The thermostat then acts to regulate the coolant to the specified temperature. If the temperature rises above the regulated temperature, the thermostat opens further to allow more coolant to circulate, thus reducing the temperature. If the temperature drops below the regulated temperature, the thermostat partially closes to reduce the amount of coolant circulating, thereby increasing the temperature. If a thermostat malfunctions in such a manner that it does not adequately restrict coolant flow during vehicle warm-up, an increase in emissions could occur do to the prolonged operation of the vehicle at temperatures below the stabilized, warmed-up value (i.e., due to cold start engine control strategies). The emission impact may vary considerably from one manufacturer to another based on cooling system design and air-fuel control strategies; however, it is generally acknowledged that the component can impact emissions significantly, particularly at lower ambient temperatures (e.g., 50 degrees Fahrenheit). Further, since the engine coolant temperature would potentially be used as an enable criterion for other OBD diagnostics, if the vehicle's coolant temperature does not reach a manufacturer-specified warmed-up value, several diagnostics may effectively be permanently disabled from identifying other emission-related malfunctions.

The staff is proposing that manufacturers be required to monitor the thermostat for proper performance. Manufacturers would be required to detect malfunctions if, within a certain time period after engine start, the engine coolant temperature does not achieve the highest temperature required to enable other OBD monitors or warm up to within 20 degrees Fahrenheit of the manufacturer-specified thermostat regulating temperature. The time period threshold(s) (i.e., the time after engine start when the thermostat would be considered malfunctioning) would be a function of starting engine coolant temperature and vehicle operating conditions that contribute to coolant temperature warm-up. Regarding the latter requirement (i.e., malfunction detection when the coolant temperature does not warm up to within 20 degrees Fahrenheit of the thermostat regulating temperature), subject to Executive Officer approval, a manufacturer would be permitted to monitor the thermostat for a larger deviation from the nominal warmed-up temperature if it adequately demonstrates that a thermostat operating at the lower temperature will not cause an emission increase of 50 or more percent of any of the applicable standards (e.g., a 50 degree Fahrenheit emission test). Manufacturers would be required to submit test data and/or an engineering analysis of the coolant temperature-based modifications to the engine control strategies to support their request. The thermostat monitoring requirement could be satisfied by verifying that the coolant temperature reaches a stabilized value after a period of engine operation, taking into account engine load and coolant temperature at engine start.

Some of the manufacturers' largest vehicles require a high capacity passenger compartment heating system. In cold weather, use of the heaters may not allow sufficient coolant temperature to be achieved in order to avoid illumination of the malfunction light, even when the thermostat is functioning normally. As a result, manufacturers have been forced to select very restrictive monitoring conditions that may not be frequently encountered in-use to ensure an accurate decision.

Therefore, the staff is proposing that vehicles that do not reach the temperatures specified by the malfunction criteria would be allowed to use alternate malfunction criteria and/or temperatures that are a function of coolant temperature at engine start. This provision would apply only for engine starts below 50 degrees Fahrenheit and would require the manufacturer to demonstrate why the standard malfunction criteria are not sufficient. Above 50 degrees Fahrenheit, the monitor would need to meet the standard malfunction criteria.

Engine Coolant Temperature Sensor

Manufacturers generally utilize engine coolant temperature (ECT) as an input for many of the emission-related engine control systems. For gasoline engines, the ECT is often one of the most important factors in determining if closed-loop fuel control will be allowed by the engine's powertrain computer. If the engine coolant does not warm up sufficiently, closed-loop fuel control is usually not allowed and the vehicle remains in open-loop fuel control. Since open-loop fuel control does not provide precise fuel control, this results in increased emission levels. Diesel engines generally use ECT to initiate closed-loop control of some emission control systems, such as EGR systems. Similar to closed-loop fuel control on gasoline engines, if the coolant temperature does not warm up, closed-loop control of these emission control systems will usually not begin, which will also result in increased emissions. For both gasoline and diesel engines, ECT would potentially be used to enable many of the diagnostics that are required by the heavy-duty OBD regulation (e.g., an OBD monitor would not run until the coolant temperature is above or below a certain temperature to ensure accurate detection capability). If the ECT sensor malfunctions and remains at a low or high reading, many diagnostics would not be enabled.

The staff is proposing that manufacturers be required to monitor the ECT sensor for proper performance. Manufacturers would be required to monitor the sensor to ensure that the vehicle achieved the highest minimum temperature needed for closed-loop control of all emission control systems (e.g., fuel system, EGR system) on gasoline and diesel vehicles within an Executive Officer-approved time after start-up, which would be based on ECT at start-up and/or intake air temperature. The Executive Officer would approve the time interval upon determining that the data and/or engineering evaluation submitted by the manufacturer supports the specified times. Vehicles that do not utilize engine coolant temperature to enable closed-loop control of any emission control system would be exempted from this monitoring requirement.

Additionally, manufacturers would be required to monitor the coolant temperature sensor for rationality, electrical, and out-of-range failures. Since the ECT sensor is essential for both fuel and spark timing control as well as for other OBD monitors, the rationality monitor needs to be more capable in detecting sensor faults than rationality monitors of non-temperature sensors (which follow the comprehensive component monitoring requirements). Accordingly, the proposed regulation would require that rationality monitoring for ECT sensors identify ones that read inappropriately low or high (and thus, disable or delay operation of other monitors). Generally, however, manufacturers may be exempt from rationality monitoring of low sensor readings that disable other OBD monitors, since the OBD monitor for the thermostat (described below) would generally be designed to detect this fault. Additionally, manufacturers may be exempt from monitoring ECT sensors stuck at high temperature regions: (1) where the MIL would be illuminated for default mode operation (e.g., overtemperature protection strategies), or (2) that fall within the red zone of the temperature gauge in cases where the ECT sensor is used for both the OBD system and the temperature gauge.

Technical Feasibility of Proposed Monitoring Requirements

The light- and medium-duty OBD II regulations have required identical ECT sensor and thermostat monitoring since the 1996 model year. The technical feasibility has clearly been demonstrated for these packages.

D. POSITIVE CRANKCASE VENTILATION (PCV) SYSTEM MONITORING

Background

Combustion in each cylinder is achieved by drawing air and fuel into the cylinder, compressing the mixture with a piston, and then igniting the mixture. After the combustion event, the mixture is exhausted from the cylinder with another stroke of the piston. However, during the combustion process, exhaust gases can escape past the piston into the crankcase and subsequently to the atmosphere. The PCV system is used to remove these gases (known as “blow-by”) from the crankcase and direct them to the intake manifold to be burned by the engine. The PCV system generally consists of a fresh air inlet hose, a crankcase vapor outlet hose, and a PCV valve to control the flow through the system. Fresh air is introduced to the crankcase via the inlet (typically a connection from the intake air cleaner assembly). On the opposite side of the crankcase, vapors are vented from the crankcase through the valve by way of the outlet hose to the intake manifold. The intake manifold provides the vacuum that is needed to accomplish the circulation while the engine is running.

The valve is used to regulate the amount of flow based on engine speed. During low engine load operation (e.g., idle), the valve is nearly closed allowing only a small portion of air to flow through the system. With open throttle conditions, the valve opens to allow more air into the system. At high engine load operation (i.e., hard accelerations), the valve begins to close again, limiting air flow to a small amount. For most systems, a mechanical valve is all that is necessary to adequately regulate PCV system air flow.

Problems may occur such that the PCV system does not function properly and emissions are vented into the atmosphere. The hoses utilized by the PCV system may be subject to cracks or deterioration. However, the staff does not believe that such failures have a significant impact on emissions because vapors are drawn by intake manifold vacuum into the engine. Therefore, air is likely to be drawn into the hose through the crack as opposed to crankcase vapor being forced out. The more likely cause of PCV system malfunctions and excess emissions is improper service or tampering of the PCV system. These failures include misrouted or disconnected hoses, and missing valves. Of these failures, hose disconnections on the vapor vent side of the systems and/or missing valves can cause emissions to be vented to the atmosphere.

Proposed Monitoring Requirements

Thus, the staff is proposing that manufacturers be required to monitor the PCV system for malfunctions. Specifically, staff proposes that manufacturers be required to monitor the PCV system for disconnections between the crankcase and the PCV valve and between the PCV valve and the intake manifold. Because disconnections between the valve and the intake manifold will result in a significant intake air leak, effective monitoring should be readily achievable through the existing monitoring strategies for the idle air control system or the fuel system. Additionally, if the leak is sufficiently large, the disconnection will render the vehicle inoperable by causing the engine to stall. The staff’s proposal does not require the stored fault code to specifically identify the disconnection if additional hardware would be required for this purpose, and provided service information generated by the manufacturer directs technicians to examine the connection as a possible cause of the indicated fault.

Regarding disconnection between the PCV valve and the crankcase, detection would be significantly more difficult with existing monitors, and would likely require additional hardware such as a pressure switch to ensure flow between the crankcase and the PCV valve. However, in order to facilitate cost-effective compliance, the staff proposes to exempt manufacturers from detecting this type of disconnection if the PCV valve is fastened directly to the crankcase in a manner that makes technicians more likely to disconnect the intake manifold hose from the valve rather than disconnect the valve itself from the crankcase during service. Staff believes that this would eliminate most of the disconnected hose and valve events because technicians who do not reconnect the hose when the service procedure is completed will be alerted to a diagnostic fault as explained in the previous paragraph that will lead the technician back to the disconnected hose.

For PCV system designs that utilize tubing between the crankcase and the valve or any additional tubing or hoses used to equalize pressure or to provide a ventilation path between various areas of the engine (e.g., crankcase and valve cover), the proposed regulation would allow for an exemption from detecting disconnection in this area. This exemption would be obtained if it is demonstrated that all of these connections are resistant to deterioration or accidental disconnection, are significantly more difficult to remove than the connections between the intake manifold and the valve, and are not subject to disconnection during any of the manufacturer’s repair procedures for non-PCV system repair work. Again, the staff believes these safeguards will eliminate most of the disconnected hose and valve failures previously observed in the field while still providing manufacturers with adequate design flexibility to meet the requirement.

The staff is not proposing to require monitoring of the identified PCV valve failures that generally do not have a significant impact on emissions such as disconnected fresh air lines and plugged valves. As stated previously, the emission impact is generally minimal (if any effect at all) due to the fact that vapors are not directly vented to the atmosphere. Further, detection of these additional failure modes would almost certainly require additional vehicle hardware. Considering the small emission benefit expected, monitoring would not be cost-effective.

Lastly, manufacturers that utilize PCV systems that do not have any external hoses or tubing would be exempted from these monitoring requirements completely. These systems typically use internally machined passageways or other similar arrangements which are not subject to failure modes causing emissions to be vented to the atmosphere.

For vehicles with diesel engines, the staff is proposing that prior to introduction on a production vehicle, manufacturers would be required to submit a plan for Executive Officer approval of the monitoring strategy, malfunction criteria, and monitoring conditions. Executive Officer approval shall be based on the effectiveness of the monitoring strategy to monitor the performance of the PCV system to the extent feasible with respect to the proposed malfunction criteria detailed above.

Technical Feasibility of Proposed Monitoring Requirements

The light- and medium-duty OBD II regulations have required identical PCV monitoring since the 1996 model year. The technical feasibility has clearly been demonstrated for these packages.

E. COMPREHENSIVE COMPONENT MONITORING

Background

Similar to the OBD II requirements for light- and medium-duty vehicles, the staff is proposing that manufacturers monitor for malfunctions of comprehensive components on heavy-duty vehicles, which covers all other electronic powertrain components or systems not mentioned above that either can affect vehicle emissions or are used as part of the OBD diagnostic strategy for another monitored component or system. Comprehensive components are generally identified as input components, which provide input directly or indirectly to the on-board computer, or as output components/systems, which receive commands from the on-board computer. Typical examples of input components include temperature sensors and pressure sensors, while examples of output components/systems include the idle speed control system, glow plugs, wait-to-start lamps, and automatic transmission solenoid or controls.

While the emission impact of a malfunctioning comprehensive component may not be as high as the major emission-related components, they still could result in a measurable increase in emissions. With the heavy-duty emission standards becoming increasingly stringent in the near future, manufacturers need to ensure that their emission-control systems are working properly in order to meet these standards. Furthermore, the proper performance of these components can be critical to the monitoring strategies of other components or systems. Malfunctions of comprehensive components that go undetected by the OBD system may disable or adversely affect the robustness of other OBD monitors without any indication. This could potentially result in the failure to detect other faulty emission-related components or systems. Due to the vital role these components play, it is important that these components are properly monitored.

Proposed Monitoring Requirements

The staff is proposing that manufacturers monitor for malfunctions of comprehensive components. The staff is proposing that input components be monitored continuously for out-of-range and circuit continuity faults (shorts, opens, etc.). Additionally, they would be monitored for rationality faults (e.g., where a sensor reads inappropriately high or low but, unlike out-of-range faults, still within the valid operating range of the sensor) whenever the monitoring conditions are met. Regarding rationality checks, the monitors would be “two-sided” (i.e., detect both inappropriately high and low readings) to the extent feasible and would have reasonable malfunction thresholds and operating conditions (not extreme operating conditions) so that faults are detected efficiently. For example, a reasonable diagnostic for a mass air flow sensor would look for a signal indicating moderate or moderate-to-high engine load, not extremely high engine load (i.e., a near out-of-range value) while the engine is operating at or near idle. Rationality monitoring would be required to use all available information and would generally be accomplished by comparing the output characteristics of multiple sensors that read the same metric during certain engine operating conditions. For example, the output characteristics of the barometric pressure sensor and manifold absolute pressure sensor could be compared during certain conditions to verify either sensor.

The staff is proposing that output components be monitored for proper functional response (i.e., that the component has properly carried out a command from the on-board computer) at least once per driving cycle. If functional monitoring is not feasible, then circuit continuity monitoring would be required. The proposed regulation would contain more specific monitoring requirements for the idle speed control system, glow plug, and intake air heater system monitors.

In contrast with other monitors, the proposed regulation would not require illumination of the MIL for all comprehensive component malfunctions. The staff is proposing that a manufacturer illuminate the MIL for comprehensive component failure only if it meets two requirements: (1) a malfunction of the component causes emissions to exceed 15 percent or more of the standard, and (2) the component is used as part of the diagnostic strategy for any other monitored component or system. Even if the MIL is not required to be illuminated, the manufacturer would still be required to store the associated confirmed fault code.

Auxiliary Emission Control Devices

Heavy-duty engine manufacturers are currently allowed to implement auxiliary emission control device (AECD) strategies that activate an alternate engine/fuel/emissions control strategy in order to protect the engine or emission control system. An AECD generally refers to any device or element of design that (1) senses temperature, engine speed, vehicle speed, manifold vacuum, or any other parameter for the purpose of activating, modulating, delaying, or deactivating the operation of the emission control system; and (2) reduces the effectiveness of the emission control system under conditions that may reasonably be expected to be encountered in normal urban vehicle operation and use. Consequently, when an AECD strategy is active, the engine usually emits more emissions into the atmosphere due to the nature of the engine control changes. For the goal of minimizing in-use emissions, it is important to limit manufacturers' use of AECDs to only when they are absolutely necessary. From the perspective of OBD and the more specific goal of minimizing in-use emissions due to emission-related malfunctions, it is important to verify that manufacturers invoke AECDs only when the vehicle is actually operated in conditions that warrant the use of the AECD.

AECDs are usually activated when input parameters reach specific values or other combinations of sensed values meet certain criteria. An overly simplified example is an AECD device that shuts off the exhaust gas recirculation (EGR) system for engine protection if the engine reaches an over-temperature condition. The over-temperature condition may be identified by the engine coolant temperature (or the engine oil temperature) sensor output exceeding a specific temperature. Currently, manufacturers are required to submit their AECD descriptions to ARB for review and approval. When everything is working correctly, most AECDs are generally activated only under “extreme” conditions.

However, when a faulty input component or sensed parameter outputs an incorrect reading, the AECDs can be erroneously activated. For example, if the engine coolant temperature sensor outputs a temperature reading that is much higher than the actual temperature and causes the engine control module to falsely think that the engine is overheating, the AECD will erroneously be activated. The staff is concerned that malfunctions may occur that cause the AECD to activate even during normal driving without any indication to the driver that there is a problem. During such occurrences, vehicle emissions may likely increase substantially.

Accordingly, the staff is proposing that manufacturers be required to monitor any input component, sensed/calculated value, or other parameter that is used to activate an AECD (which, by definition, is emission-related). Specifically, the OBD system would be required to detect a failure of a component, sensed value, or other parameter that would cause the system to falsely activate an AECD. This monitoring requirement would be included as part of the comprehensive component monitoring requirements in the proposed regulation which requires monitoring of any electronic powertrain component that can affect emissions or is used as part of the monitoring strategy for any other emission-related component. Under the proposed comprehensive component monitoring requirements, manufacturers would be required to monitor input comprehensive components for circuit, out-of-range, and rationality faults. To the extent technically feasible, the staff is expecting manufacturers to design the input comprehensive component rationality monitor to catch the AECD-related faults described above. A typical rationality monitor uses all available information to identify components that are operating within their normal range but no longer accurate due to sensor drift or deterioration, and are usually “two-sided” (i.e., look for inappropriately high or low readings). The staff wants to ensure that the rationality monitor is able to detect faults at a level that would trigger inappropriate activation of an AECD. Manufacturers whose typical rationality monitors would not be able to identify AECD-related faults for an input component would likely need to either improve their existing monitor, add another monitor, or modify their AECD strategy to achieve this.

Additionally, to enable the staff to verify that the monitoring strategies used by the manufacturer cover malfunctions that would falsely trigger AECD activation, manufacturers would be required to submit detailed descriptions of all the AECDs used as part of their OBD certification application (refer to section IX of the Staff Report). This description would include the purpose of the AECD, the actions taken when the AECD is activated, and the exact criteria used to decide when the AECD is activated. While this information is currently submitted as part of the engine emission certification application, it is anticipated that manufacturers may follow the path of light-duty manufacturers and submit their OBD certification application for review and approval in advance of the engine emission certification application. As such, the description of the AECDs will need to be included in the OBD application. However, the description required with the OBD application is identical to that required for engine emission certification, so the manufacturer will simply be required to submit the same information at the time of OBD certification (should it occur at a different time than the engine emission certification review).

Technical Feasibility of Proposed Monitoring Requirements

The light- and medium-duty OBD II regulations have required identical comprehensive component monitoring since the 1996 model year. The technical feasibility has clearly been demonstrated for these packages.

F. OTHER EMISSION CONTROL OR SOURCE SYSTEM MONITORING

While the heavy-duty OBD regulation would list very specific requirements for most emission controls commonly used today, manufacturers are continually innovating new emission control technologies in addition to refining existing ones. In cases where the technology simply reflects refinements over current technology, the heavy-duty OBD monitoring requirements described above would generally be sufficient to ensure the improved devices are properly monitored. However, in cases where the new technology represents a completely different type of emission control device, the monitoring requirements for existing emission controls may not be easily applied. Typical devices that fall under this category include hydrocarbon traps and thermal storage devices.

Given that the purpose of OBD is to monitor all emission-related and emission control devices, the staff is proposing to require manufacturers to submit a monitoring plan for ARB's review and approval for any new emission control technology prior to introduction on any future model year vehicles. This policy has worked effectively for the light- and medium-duty OBD II regulation, allowing manufacturers and ARB staff to evaluate the new technology and determine an appropriate level of monitoring that was both feasible and consistent with the monitoring requirements for conventional emission control devices.

Within the proposed requirement, the staff would provide guidance as to what type of components would fall under the requirements of this section instead of under the comprehensive component section. Specifically, the staff is concerned that uncertainty may arise for emission control components or systems that can also be defined as electronic powertrain components because they fit the definitions of both sections. As such, the proposal would delineate the two by requiring components/systems that fit both definitions but are not corrected or compensated for by the adaptive fuel control system to be monitored under the provisions of the "other emission control devices" requirements rather than under the comprehensive component requirements. A typical device that would fall under this category instead of the comprehensive components category because of this delineation is a swirl control valve system. Such delineation is necessary because emission control components generally require more thorough monitoring than comprehensive components to ensure low emission levels throughout a vehicle's life. Further, emission control components that are not compensated for by the fuel control system as they age or deteriorate can have a larger impact on tailpipe emissions relative to comprehensive components that are corrected for by the fuel control system as they deteriorate.

Also, to ensure that all devices that can generate emissions on hybrids and other advanced vehicle propulsion technology vehicles are properly monitored, the proposal would expand the requirement to require monitoring of "emission source devices" in addition to emission control devices. For purposes of the proposed regulation, "emission source devices" would be defined as components or systems that emit pollutants that are subject to vehicle evaporative and exhaust emission standards (e.g., NMOG, NOx, PM). These may include non-electronic components and non-powertrain components such as fuel-fired passenger compartment heaters and on-board reformers. For these devices, manufacturers would be required to submit a plan for Executive Officer approval of the OBD II monitoring strategy, malfunction criteria, and monitoring conditions in the same manner used for emission control devices.

G. EXCEPTIONS TO MONITORING REQUIREMENTS

Under certain conditions, the reliability of certain monitors may be significantly diminished. Accordingly, ARB is proposing to allow manufacturers to disable the affected monitors when these conditions are encountered in-use. These include situations of extreme conditions (e.g., very low ambient temperatures, high altitudes) and of periods where default modes of operation are active (e.g., when a tire pressure problem is detected). In some of these cases, ARB may allow manufacturers to revise the emission threshold to ensure the most reliable monitoring performance. More details of the exceptions to the proposed monitoring requirements are specified in the proposed regulation.

VI. A STANDARDIZED METHOD TO MEASURE REAL WORLD MONITORING PERFORMANCE

A. Background

In designing an OBD monitor, manufacturers must define enable conditions that bound the vehicle operating conditions where the monitor will execute and make a judgment as to whether a component or system is malfunctioning. Manufacturers would be required to design these enable conditions so that the monitor is: (a) robust (i.e., accurately making pass/fail decisions), (b) running frequently in the real world, and, (c) in general, also running during the FTP heavy-duty transient cycle. If designed incorrectly, these enable conditions may be either too broad and result in inaccurate monitors, or overly restrictive and prevent the monitor from executing frequently in the real world.

Since the primary purpose of an OBD system is to continuously monitor for and detect emission-related malfunctions while the vehicle is operating in the real world, a standardized methodology for quantifying real world performance would be beneficial to both ARB and vehicle manufacturers. Generally, in determining whether a manufacturer’s monitoring conditions are sufficient, a manufacturer would discuss the proposed monitoring conditions with ARB staff. The finalized conditions would be included in the certification applications and submitted to ARB staff, who would review the conditions and make determinations on a case-by-case basis based on the expert judgment of the staff. In cases where the staff is concerned that the documented conditions may not be met during reasonable in-use driving conditions, the staff would most likely ask the manufacturer for data or other engineering analysis used by the manufacturer to determine that the conditions will occur in-use. In proposing a standardized methodology for quantifying real world performance, the staff believes this review process would be made easier and faster. Furthermore, it would better ensure that all manufacturers are held to the same standard for real world performance. Additionally, the staff believes it is necessary to propose procedures that will ensure that monitors operate properly and frequently in the field.

The staff is therefore proposing that all manufacturers be required to use a standardized method for determining real world monitoring performance and hold manufacturers liable if monitoring occurs less frequently than a minimum acceptable level, expressed as minimum acceptable in-use performance ratio. The proposed regulation would also require manufacturers to implement software in the on-board computers to track how often several of the major monitors (e.g., catalyst, EGR, PM trap, other diesel aftertreatment devices) execute during real world driving. The on-board computer would keep track of how many times each of these monitors has executed as well as how often the vehicle has been driven. By measuring both these values, the ratio of monitor operation relative to vehicle operation can be calculated to determine monitoring frequency. The proposed requirements would also establish a minimum acceptable monitoring frequency that manufacturers must meet for each monitor. The proposal would likely make it easier for ARB to identify problematic monitors.

The proposed minimum acceptable frequency requirement would apply to many of the OBD system monitors. In the proposed OBD regulation, monitors would be required to operate either continuously (i.e., all the time), “once-per-driving-cycle” (i.e., once per driving event), or in a few cases, "multiple-times-per-driving-cycle" (but only when the proper monitoring conditions are present, not continuously). For components or systems that are more likely to experience intermittent failures or failures that can routinely happen in distinct portions of a vehicle’s operating range (e.g., only at high engine speed and load, only when the engine is cold or hot), monitors would be required to be continuous. Examples of continuous monitors include the fuel system monitor and most electrical/circuit continuity monitors. For components or systems that are less likely to experience intermittent failures or failures that only occur in specific vehicle operating regions or for components or systems where accurate monitoring can only be performed under limited operating conditions, monitors would be required to run “once per driving cycle.” Examples of “once-per-driving-cycle” monitors typically include gasoline catalyst monitors, evaporative system leak detection monitors, and output comprehensive component functional monitors. For components or systems that are routinely used and perform functions that are crucial to maintaining low emissions but may still require monitoring under fairly limited conditions, monitors would be required to run each and every time the manufacturer-defined enable conditions are present. Examples of "multiple-times-per-driving-cycle" monitors typically include input comprehensive component rationality monitors and some diesel exhaust aftertreatment monitors.

Monitors that would be required to run continuously, by definition, would always be running and a minimum frequency requirement is unnecessary. The new frequency requirement would essentially apply only to those monitors that are designated as “once-per-driving-cycle” or "multiple-times-per-driving-cycle." For all of these monitors, manufacturers would be required to define monitoring conditions that ensure adequate frequency in-use. Specifically, the monitors would need to run often enough so the measured monitor frequency on in-use vehicles would exceed the minimum acceptable frequency. However, even though the minimum frequency requirement would apply to nearly all “once-per-driving-cycle” and "multiple-times-per-driving-cycle" monitors, manufacturers would only be required to implement software to track and report the in-use frequency for a few of the major monitors. These few monitors generally represent the most critical emission control components and the most difficult monitors to run. Standardized tracking and reporting of only these monitors should, therefore, provide sufficient indication of monitoring performance.

B. Why frequent monitoring is important

It is important that OBD monitors run frequently to ensure early detection of emission-related malfunctions and, consequently, maintain low emissions. Allowing malfunctions to continue undetected, and thus go without repair, for long periods of time allows emissions to increase unnecessarily. In other words, the sooner the emission-related malfunction is detected and fixed, the fewer the excess emissions that are generated from the vehicle.

Frequent monitoring can also help assure that intermittent emission-related faults (i.e., faults that are not continuously present, but occur for days and even weeks at a time) are detected. The nature of mechanical and electrical systems is that intermittent faults can and do occur, and the less frequent the monitoring, the less likely these faults will be detected and repaired. Additionally, for both intermittent and continuous faults, earlier detection is equivalent to preventative maintenance in that the original malfunction can be detected and repaired prior to it causing subsequent damage to other components. This can help vehicle operators avoid more costly repairs that would have resulted had the first fault gone undetected.

Infrequent monitoring can also have an impact on the service and repair industry. Specifically, monitors that have unreasonable or overly restrictive enable conditions could hinder vehicle repair services. In general, upon completing an OBD-related repair to a vehicle, a technician will attempt to verify that the repair has indeed fixed the problem. Specifically, a technician will ideally operate the vehicle in a manner that will exercise the appropriate OBD monitor and allow the OBD system to confirm that a malfunction is no longer present. This affords a technician the highest level of assurance that the repair was indeed successful.

However, if OBD monitors operate infrequently and are therefore difficult to exercise, technicians may not be able (or may not be likely) to perform such testing. Despite the future, proposed ARB service information regulation amendments that would require manufacturers to make all of their service and repair information available to all technicians, including the information necessary to exercise OBD monitors, technicians would still have difficulty in exercising monitors that require infrequently encountered vehicle operating conditions (e.g., abnormally steady constant speed operation for an extended period of time). Furthermore, this information and the time required by the technician to perform this verification would not be free. Ultimately, vehicle owners would pay for this information and labor time through their repair bills. Additionally, to execute OBD monitors in an expeditious manner or to execute monitors that would require unusual or infrequently encountered conditions, technicians may be required to operate the vehicle in an unsafe manner (e.g., at freeway speeds on residential streets or during heavy traffic). If unsuccessful in executing these monitors, technicians may even take shortcuts in attempting to validate the repair while maintaining a reasonable cost for consumers. These shortcuts, however, would likely not be as thorough in verifying repairs and could increase the chance for improperly repaired vehicles being returned to the vehicle owner or additional repairs being performed just to ensure the problem is fixed. In the end, monitors that operate less frequently can result in unnecessary increased costs and inconvenience to both vehicle owners and technicians.

While technicians (and/or consumers) may elect not to spend the additional time and money to validate a routine repair, repairs made in the context of passing an I/M (Smog Check) test would require this validation. For an OBD-based I/M inspection, the driver or technician would be required to exercise the OBD monitors and verify that the repairs are successful before the inspection can be performed. This is because this inspection would require specific internal flags in the OBD system known as readiness flags to be set before the vehicle can pass the inspection. These flags would only set upon each of the major OBD monitors executing and completing at least once since the last time fault codes were erased. Vehicles failed during an OBD-based I/M inspection (due to the presence of a malfunction) would be required to have malfunctions repaired (and thus, fault codes cleared) before returning for re-testing to verify the repairs. If OBD monitors are incapable of executing frequently and verifying repairs in a timely manner, technicians would have a difficult time preparing a vehicle for re-inspection or would be able to do so only with considerable effort, and thus, at considerable cost to the vehicle owner. With especially troublesome monitors, vehicle owners may have to wait several weeks or months before the repair is verified, the readiness flag is set by the OBD system, and the vehicle can be re-inspected at the I/M station. In contrast, monitors that function frequently would be easier for technicians and even vehicle owners to exercise. Clearly, monitors that function infrequently would subject vehicle owners to unnecessary delays and/or increased repair costs that would hinder the effectiveness and efficiency of the I/M program.

C. Detailed description of software counters to track real world performance

As stated above, manufacturers would be required to track monitor performance by counting the number of monitoring events (i.e., how often each diagnostic has run) and the number of vehicle driving events (i.e., how often has the vehicle been operated). The ratio of the two would give an indication of how often the monitor is operating relative to vehicle operation. Thus:

$$\text{In - Use Performance (Ratio)} = \frac{\text{Number of Monitoring Events (Numerator)}}{\text{Number of Driving Events (Denominator)}}$$

To ensure all manufacturers are tracking performance in the same manner, the proposed regulation would include very detailed requirements for defining and incrementing both the numerator and denominator of this ratio. Manufacturers would be required to keep track of separate numerators and denominators for each of the major monitors, and to ensure that the data are saved every time the vehicle is turned off. The numerators and denominators would be reset to zero only in extreme circumstances when the non-volatile memory has been cleared (e.g., when the on-board computer has been reprogrammed in the field, when the on-board computer memory has been corrupted). The values would not be reset to zero during normal occurrences such as when fault codes have been cleared or when routine service or maintenance has been performed.

Further, the numerator and denominator would be structured such that the maximum value each can obtain is 65,535, the maximum number that can be stored in a 2-byte location, to ensure manufacturers allocate sufficient memory space in the on-board computer. If either the numerator or denominator for a particular monitor reaches the maximum value, both values for that particular monitor will be divided by two before counting resumes. In general, the numerator and denominator would only be allowed to increment a maximum of once per driving cycle because most of the major monitors are designed to operate only once per driving cycle. Additionally, incrementing of both the numerator and denominator for a particular monitor would be disabled (i.e., paused but the stored values would not be erased or reset) only when a fault has been detected (i.e., a pending or confirmed code has been stored) that prevents the monitor from executing. Once the fault is no longer detected and the pending fault code is erased, either through the allowable self-clearing process or upon command by a technician via a scan tool, incrementing of both values would resume.

To handle many of these issues, staff has worked with industry and SAE to develop standards for storing and reporting the data to a generic scan tool. This would also help ensure that all manufacturers report the data in an identical manner and thus help facilitate data collection in the field.

1. Number of monitoring events (“numerator”)

For the numerator, manufacturers would be required to keep a separate numeric count of how often each of the particular monitors has operated. However, this is not as simple as it may seem. More specifically, manufacturers would have to implement a software counter that increments by one every time the particular monitor meets all of the enable/monitoring conditions for a long enough period of time such that a malfunctioning component would have been detected. For example, if a manufacturer requires a vehicle to be warmed-up and at idle for 20 seconds continuously to detect a malfunctioning catalyst, the catalyst monitor numerator could only be incremented if the vehicle has actually operated in all of those conditions simultaneously. If the vehicle is operated in some but not all of the conditions (e.g., at idle but not warmed-up), the numerator would not be allowed to increment because the monitor would not have been able to detect a malfunctioning catalyst unless all of the conditions were simultaneously satisfied.

Another complication is the difference between a monitor reaching a “pass” or “fail” decision. At first glance, it would appear that a manufacturer should simply increment the numerator anytime the particular monitor reaches a decision, be it “pass” or “fail”. However, monitoring strategies may have a different set of criteria that must be met to reach a “pass” decision versus a “fail” decision. As a simple example, a manufacturer may appropriately require only 10 seconds of operation at idle to reach a “pass” decision but require 30 seconds of operation at idle to reach a “fail” decision. Manufacturers would only be allowed to increment the numerator if the vehicle was at idle for 30 seconds even if the monitor actually executed and reached a “pass” decision after 10 seconds. This is necessary because the primary function of OBD systems is to detect malfunctions (i.e., to correctly reach “fail” decisions, not “pass” decisions), and thus, the real world ability of the monitors to detect malfunctions is the parameter that needs to be measured. Therefore, monitors with different criteria to reach a “pass” decision versus a “fail” decision would not be able to increment the numerator solely on the “pass” criteria being satisfied.

It is imperative that manufacturers implement the numerators correctly to ensure a reliable measure for determining real world performance. “Overcounting” would falsely indicate the monitor is executing more often than it really is, while “undercounting” would make it appear as if the monitor is not running as often as it really is. Manufacturers would be required to demonstrate the proper function of the numerator incrementing strategy to ARB prior to certification, and to verify the proper performance during production vehicle evaluation testing.

2. Number of driving events (“denominator”)

The proposed amendments would also require manufacturers to separately track how often the vehicle is operated. In the simplest of terms, the denominator would be a counter that increments by one each time the vehicle is operated. The issue of how to best count or measure vehicle operation was the subject of considerable discussion during the light-duty OBD II regulatory revisions adopted in April of 2002. Several proposals were considered, including very simple measures such as the number of key starts as well as more complex measures that require several individual criteria to be met on a single driving cycle before it would increment the denominator counter. At this time, the staff is proposing to use the same methodology for heavy-duty OBD as currently adopted in the OBD II regulation. That is, the denominator counter would only be incremented if several criteria were satisfied on a single driving cycle. This method allows very short trips or trips during extreme conditions such as very cold temperatures or very high altitude to be filtered out and excluded from the count. This is appropriate because these are also conditions where most OBD monitors are neither expected nor required to operate.

Specifically, the denominator would be incremented if on a single key start, the following criteria were satisfied:

(1) minimum engine run time of 10 minutes;
(2) minimum of 5 minutes, cumulatively, of vehicle operation at vehicle speeds greater than 25 miles-per-hour; and
(3) at least one continuous idle for a minimum of 30 seconds encountered; and the above three conditions met while:
(4) ambient temperature between 20 and 100 degrees Fahrenheit;
(5) altitude of </= 8000 feet.

As with the light-duty OBD II regulations, the staff will work with industry to collect data during the first few years of implementation and make any adjustments, if necessary, to the criteria used to increment the denominator to ensure the ratio provides a meaningful measure of in-use monitoring performance.

D. Proposed standard for the minimum acceptable in-use performance (“ratio”)

Determining how frequent is “frequent enough” for monitors to operate is a complex task that requires consideration of several different factors, including the technical capability of OBD systems, the severity of the malfunction, the consequences of delayed detection and repair of the malfunction, and expected driving patterns and habits. When considering all of these factors, the staff has established a target frequency of malfunction detection (and MIL illumination) within two weeks from occurrence of the fault for 90 percent of the vehicle population. The vast differences in vehicle operation over a two-week period, however, make it difficult to objectively ascertain whether or not this criterion is satisfied. The proposed regulation would attempt to simplify this task by specifying a minimum acceptable monitoring frequency in a quantifiable format, known as the minimum acceptable in-use performance ratio.

In order to determine the appropriate minimum acceptable in-use performance ratio that correlates with the target frequency of two weeks, an analysis of in-use driving patterns of heavy-duty vehicles would need to be conducted. This would take into account the real world variability in driving habits, which would help ensure that the vast majority of heavy-duty vehicles are capable of detecting malfunctions in a timely manner. This analysis requires a fairly large data set of real world driving cycles from all types of vehicles in the heavy-duty industry. While the staff did indeed perform such an analysis for the light-duty OBD II regulation, the staff has not yet identified a suitable database that contains the information necessary to perform such an analysis for the heavy-duty industry. Nevertheless, the staff believes that a minimum ratio must be set to ensure that OBD monitors are indeed running and detecting emission-related malfunctions. Therefore, the staff is proposing a minimum ratio of 0.100 for all monitors required to meet the in-use performance requirement. Based on the analysis done during the OBD II regulatory development, a ratio of 0.100 will generally translate to much less frequent monitoring than the target of two weeks. However, this ratio will still ensure some monitoring is occurring in-use on some portion of the heavy-duty vehicles and will provide manufacturers with considerable flexibility to gain experience during the first few years OBD is required on heavy-duty vehicles. As more data becomes available, staff will perform a more accurate analysis targeting the two-week standard and modify the proposed minimum acceptable ratio(s) during future rulemaking reviews.

VII. STANDARDIZATION REQUIREMENTS

Similar to the light- and medium-duty OBD II regulation, the heavy-duty OBD regulation would include requirements for manufacturers to standardize certain features of the OBD system. Effective standardization assists all repair technicians in diagnosing and repairing malfunctions by providing equal access to essential repair information, and requires structuring the information in a common format from manufacturer to manufacturer. Additionally, the standardization would help facilitate the potential incorporation of OBD checks into a future heavy-duty I/M program.

Among the features that would be standardized under the proposed heavy-duty OBD regulation include the diagnostic connector, hardware and software specifications for tools used by service technicians, the information made available by the on-board computer, the methods for accessing the information, the numeric fault codes stored when a malfunction is detected, and the terminology used by the manufacturer in service manuals.

At this time, the staff has not made a final decision as to which standards heavy-duty manufacturers should be required to conform to. As discussed further in the following sections, there are two sets of standards that seem to be the leading choices for the heavy-duty industry. They are SAE J1939 and ISO 15765/15031 (“light-duty CAN”). As can be expected, both options have their advantages and disadvantages and both have ardent supporters and opponents. From staff’s past experience with standardization under light-duty, however, ultimately it is desirable to have a single set of standards used by all heavy-duty vehicles. Staff has found this is generally beneficial for the service and repair industry, I/M or roadside-type inspections, diagnostic equipment and tool manufacturers, and the regulatory agencies in terms of verifying all vehicles are built in conformance with the standards. Further, a single protocol is also beneficial for fleet operators that utilize add-on equipment such as data loggers and for vehicle/coach builders that integrate various engine and component suppliers that eventually must all work together. Thus, staff’s goal is indeed to end up with a single set of standards for all heavy-duty vehicles.

With that in mind, staff is leaning towards a phase-in of the OBD system but a single set of standards for all vehicles included in the phase-in rather than an approach allowing multiple protocols before finally harmonizing on a single protocol. However, staff is still seeking feedback on the best approach to reach that goal in the shortest time possible as well as what set of standards should be used for that final goal. Further, staff is seeking feedback on what approaches can or should be used prior to reaching the final goal. Please note, however, that the current draft regulatory language for the standardization requirements does not reflect staff’s position for heavy-duty vehicles. The draft regulatory language is primarily a carry-over from the light-duty regulatory language and since staff has not reached any decision on the standardization issue for heavy-duty at this time, it was left largely unchanged and will remain that way until a decision is made based on feedback from industry.

One other important aspect to keep in mind is that the proposal by staff would only require that a certain minimum set of emission-related information be made available through the standardized format, protocol, and connector selected by staff. It does not limit engine or vehicle manufacturers as to what protocol they use for engine or vehicle control, communication between on-board computers, or communication to manufacturer-specific scan tools or test equipment. Further, it does not prohibit engine or vehicle manufacturers from equipping the vehicle with additional diagnostic connectors or protocols as required by other suppliers or purchasers. For example, fleets that use data logging or other equipment that requires the use of SAE J1587 communication and connectors could still be installed and supported by the engine and vehicle manufacturers. The OBD rules would only require that manufacturers also equip their vehicles with a specific connector and communication protocol that meets the standardized requirements to communicate a minimum set of emission-related inspection and diagnostic information.

A. Communication Protocol

During the initial years of implementation of the light- and medium-duty OBD II regulation, ARB allowed manufacturers to use one of four protocols for communication between a generic scan tool and the vehicle's on-board computer. A generic scan would automatically cycle through each of the allowable protocols to establish communication with the on-board computer. While this has generally worked successfully in the field, some communication problems have arisen in the field due, in part, to the use of multiple protocols. Recent amendments to the OBD II regulation now require all manufacturers to use only one protocol by the 2008 model year to help address this issue. To avoid this problem for heavy-duty vehicles, the staff is leaning towards solutions that require the use of a single protocol for all vehicles.

While the industry has been very sharply divided over the issue of which protocol will best serve the needs of industry and ARB, at this time, the staff is leaning towards ISO 15765 as the protocol best suited to meet the needs in the 2007 timeframe. This is the same standard starting to be used in the light-duty industry in the 2003 model year and required on all light- and medium-duty vehicles by the 2008 model year. By harmonizing with the light-duty protocol, equipment and tool manufacturers will be able to adapt existing tools very easily to work on heavy-duty vehicles and will provide even more diagnostic equipment choices for heavy-duty repair and maintenance personnel. Further, the ISO 15765 and associated ISO 15031 standards have already been updated to accommodate nearly every standardized requirement proposed for heavy-duty vehicles. The use of the 15765 protocol and 15031 messages will also provide a consistent format for technicians and inspectors on all types of vehicles. Lastly, the use of the same protocol used in current medium-duty applications provides vehicle and engine manufacturers (as well as other suppliers) that currently produce product for both the medium-duty and the heavy-duty sectors the ability to use a common software set for all products.

One of the biggest drawbacks of the ISO protocols, however, is that the diagnostic information made available through the OBD connector is limited to a subset of the emission-related information available on the vehicle. While the content is suitable for performing inspections, in some cases it can be marginally acceptable for purposes of performing emission-related diagnostics. As heavy-duty vehicles incorporate more advanced emission controls and increased reliance on sensors and computer communications, it may be become more difficult for technicians to quickly and effectively isolate malfunctions with a limited subset of the available information.

Several of the engine manufacturers have requested the allowance of another protocol, SAE J1939, for the heavy-duty industry. However, surveys of the existing heavy-duty engine manufacturers, transmission manufacturers, equipment manufacturers, and other suppliers indicate that the a significant portion of the interested parties do not favor SAE J1939 as the best protocol choice for a long term solution. Further, discussions with various heavy-duty manufacturers indicate that SAE J1939 is not currently being used by the majority of heavy-duty manufacturers nor is it used on the majority of engines sold in the heavy-duty market. In some cases, it appears that several manufacturers continue to use J1587 as the primary communication network on the vehicle.

That said, there are some distinct advantages that SAE J1939 could have over the ISO 15765 protocol. One such advantage could be the opportunity to access not only the minimum parameter set required by the OBD regulation but to access all parameters available on the vehicle through the same protocol and message structure. This would be a clear advantage for repair technicians by providing a more powerful repair tool if all of these additional parameters are standardized and can be automatically translated by the scan tool without any additional manufacturer-specific software. In the same manner, SAE J1939 could offer the ability to access enhanced emission-related (and potentially non-emission-related) diagnostic information other than just parameters with a single tool and without manufacturer-specific software/cartridges/adapters to translate the information. However, discussions with some in the heavy-duty industry have indicated that the majority of “enhanced” (e.g., beyond the minimum required by the OBD regulation) diagnostic information, while accessed through the J1939 connector on the J1939 network, is not accessed using defined and standardized J1939 messages (nor is it required to be by SAE J1939) in a manner that would automatically translate the results to useable information for a repair technician. As such, manufacturer-specific scan tool software is still required to access and use the enhanced information for a particular engine model and make. If this is the case, then SAE J1939 offers little advantage in this aspect relative to the ISO 15765 protocol as repair technicians would still be required to purchase additional scan tool software every year for each specific make and model. If SAE J1939 really can offer access to all enhanced diagnostic information without the use of manufacturer-specific software, J1939 would have a distinct advantage over the ISO protocols. Staff is seeking additional feedback from heavy-duty suppliers (e.g., engine and transmission manufacturers) as well as repair tool equipment manufacturers and the repair industry on this issue.

Others in the industry had requested that ARB allow manufacturers to choose one of several communication protocols including SAE J1939, ISO 15765, as well as various hybrid combinations of the two protocols. However, as stated previously, the staff's past experience is that multiple protocols will result in some of the same communication problems that the light- and medium-duty vehicles experienced. Additionally, with the potential incorporation of an OBD check into the future heavy-duty I/M program, the use of only one communication protocol on all heavy-duty vehicles would be beneficial. A single protocol also offers a tremendous benefit to scan tool designers as well as technicians. Scan tool designers can focus on added feature content and can expend much less time and money validating basic functionality of their product on all the various permutations of protocol interpretations that are implemented. As such, technicians will likely get a scan tool that works properly on all vehicles without the need for repeated software updates that incorporate "work-arounds" or other patches to fix bugs or adapt the tool to accommodate slight variances in how the multiple protocols interact with each other or are implemented by various manufacturers. Staff’s goal is to eventually have a single protocol used by all heavy-duty vehicles. As mentioned earlier though, staff is seeking feedback on the most effective way to get to that end point in the shortest time possible.

B. Diagnostic Connector

All vehicles would be required to incorporate a diagnostic connector conforming to the specifications contained in the standards ultimately selected. The diagnostic connector would be required to be located in the driver’s side foot-well region of the vehicle interior and would need to be easily identified and reachable by a technician or inspector crouched or standing on the ground on the driver's side of the vehicle with the vehicle driver's door open. Additionally, if a manufacturer wished to utilize a cover over the connector, the manufacturer would be required to label the cover with the text "OBD" to assist technicians in identifying its location and would be required to make the cover easily removable by hand (without the use of tools). The manufacturer would be required to submit the label to ARB for approval. The staff's experience from the light-duty industry has been that connectors that are difficult to locate cause unnecessary but substantial problems both in the repair community and the I/M community. Further, feedback from ARB heavy-duty inspectors has indicated that a location that would be easily accessible without entering the vehicle and while standing on the ground provides the most efficient means for inspection and would be preferred by most vehicle owner/operators.

C. Readiness Status

Manufacturers would be required to incorporate readiness status indications of several major emission control systems and components into their vehicles, which would determine if the OBD monitors have performed their system evaluations. When the vehicle is scanned, the monitor would report a readiness status of either “complete” (if the monitor has run a sufficient number of times to detect a malfunction since the memory was last cleared), “incomplete” (if the monitor has not yet had the chance to run since the memory was last cleared), or “not applicable” (if the monitored component in question is not equipped or monitored on the vehicle). The readiness status of monitors that are required to run continuously would always indicate “complete.” The proposed heavy-duty OBD regulation would detail the process of setting readiness status for each monitor. The readiness status would be set to “incomplete” whenever the fault memory is cleared either by a battery disconnect or by a scan tool, but not after a normal vehicle shutdown (i.e., key-off).

For the light- and medium-duty OBD II regulation, the main intent of the readiness status is to ensure a vehicle is ready for I/M testing (i.e., that monitors have run) and to prevent fraudulent testing. In general, for OBD-based I/M tests, technicians “fail” a vehicle if the MIL is illuminated, which indicates a fault is currently present. Without readiness status, drivers (or even technicians) could possibly avoid “fail” designations by disconnecting the battery and clearing the computer memory prior to an I/M inspection, which erases any pre-existing fault codes and extinguishes the MIL. The readiness status information allows a technician to determine if the memory in the on-board computer has been recently cleared (e.g., by a technician clearing fault codes or disconnecting the battery). The presence of unset readiness flags will cause the vehicle to be rejected from testing and required to return for a re-test at a later date. With the potential incorporation of OBD checks into the revised heavy-duty I/M program in the future, the staff anticipates that the readiness status would serve the same function.

Technicians would also potentially use the readiness status to verify OBD-related repairs. Specifically, technicians would clear the computer memory after repairing an OBD-detected fault in order to erase the fault code, extinguish the MIL, and reset the readiness status to “incomplete.” Then the vehicle would be operated in such a manner that the monitor of the repaired component would be exercised (i.e., the readiness status of the monitor is set to “complete”). The absence of any fault codes or MIL illumination would indicate a successful repair.

Unfortunately, the presence of unset readiness flags may be due to circumstances beyond the driver’s control (i.e., the car was not driven under the conditions necessary to run some of the monitors) and these drivers would be rejected from I/M testing. For example, vehicle operation in extreme ambient conditions would prohibit monitors from running and setting readiness status to “complete”. To address the issue of extreme ambient conditions, the proposed regulation would allow, subject to Executive Officer approval, that in situations where monitors have been disabled for multiple driving cycles due to extreme ambient conditions, the readiness status for the subject monitors would be set to “complete,” even if monitoring has not been completed. As another example, if a vehicle with the MIL illuminated was repaired shortly before an I/M test, there may be instances where the vehicle has not had sufficient time to operate (i.e., exercise the monitors) after the repair services so that it may have unset readiness flags. These vehicles may consequently be rejected from I/M testing.

Originally, ARB staff envisioned that all readiness flags on a vehicle would be required to be set to "complete" prior to I/M testing. However, given the situations cited above and trying to balance consumer inconvenience with fraud detection, the U.S. EPA recommends allowing vehicles to pass the light- and medium-duty I/M inspection as long as there are two or fewer readiness flags set to "incomplete" (most vehicles have a total of four readiness flags). However, a substantial amount of feedback regarding readiness flags and clearing of codes prior to I/M inspection has been gathered in the last few years as 17 states across the nation, including California, have implemented some form of OBD II inspection into the I/M program. Specifically, there is now more evidence that the "two or fewer" criterion that knowingly created a potential loophole for vehicles to fraudulently get through an I/M inspection is indeed being exploited by vehicle owners, technicians, and inspectors. As such, the staff is looking for additional improvements to the readiness flag logic that will better differentiate between vehicles that are attempting to fraudulently get through an I/M inspection prior to re-detection of a fault and those that have been correctly repaired recently or otherwise have unset readiness flags through no fault of the vehicle operator. The staff's additional proposed requirements to help address this issue are detailed below, however, the staff is open to further suggestions and/or modifications to the staff's proposals to better address this issue.

Distance and Number of Warm-up Cycles Since Code Clear

The staff's proposal would require all vehicles to make available data on the distance elapsed and the number of warm-up cycles since the fault memory was last cleared. By combining these data with the readiness data, technicians (or I/M programs) would better be able to determine if unset readiness flags or an extinguished MIL are due to recent clearing of the memory or circumstances beyond the driver’s control. For example, a vehicle with several "incomplete" readiness flags but with a high number of miles traveled and of warm-up cycles since code clear would be less likely to have undergone a recent clearing event solely to extinguish the MIL prior to I/M repair. On the other hand, a vehicle even with only one or two "incomplete" readiness codes and a very low number of miles traveled and warm-up cycles since code clear would be a more likely candidate to be rejected or failed at the I/M inspection and required to return once the readiness status is "complete" for all monitors. This would better allow an I/M program to be set up to reject only those vehicles with recently cleared memories from the I/M inspection while minimizing the chance to reject vehicles that have monitors that are difficult to execute or possess monitoring conditions that are not frequently encountered due to the specific vehicle owner's driving habits.

Permanent Diagnostic Trouble Code Storage

The staff is also proposing a requirement to make it much more difficult for a vehicle owner or technician to clear the fault memory and erase all traces of a previously detected fault. Currently for light- and medium-duty vehicles, a technician or vehicle owner can erase all diagnostic trouble codes (DTCs) and extinguish the MIL by issuing a command from a generic scan tool plugged into the vehicle or, in many cases, simply by disconnecting the vehicle battery. While this does reset the readiness status for all monitors to "incomplete" and would reset the two counters described in the previous paragraph to zero, it also removes all trace of the previous fault that was detected on the vehicle.

The staff's proposal would require manufacturers to store up to four confirmed DTCs that are presently commanding the MIL on to non-volatile memory (NVRAM) at the end of every key cycle. By requiring these permanent DTCs to be stored in NVRAM, vehicle owners would not be able to erase them simply by disconnecting the battery. Further, manufacturers would not be allowed to clear or erase these "permanent" DTCs by any generic or manufacturer-specific scan tool command. Instead, these DTCs would only be allowed to be self-cleared by the OBD system itself, once the monitor responsible for setting that DTC has indeed run and passed enough times that is has confirmed that the fault is no longer present. Once this has occurred, the specific DTC stored in NVRAM would be erased. Thus, if more than one emission-related fault existed, to erase all the permanent DTCs stored in NVRAM, each monitor related to each permanent DTC would have to run and pass.

This approach provides several benefits to an I/M program. First, it would allow an I/M program to very specifically target and reject/fail those vehicles that have recently had the MIL illuminated and have not subsequently been driven enough to exercise the specific monitor previously responsible for illuminating the MIL on that vehicle. With readiness status, I/M stations are forced to either require that all monitors have run and passed since the last code clear or allow some monitors to remain incomplete and gamble that the incomplete monitors are not the ones that were previously responsible for illuminating the MIL on that particular vehicle. For example, a vehicle could show up at an I/M inspection with the catalyst monitor incomplete and the EGR monitor complete. If that particular car recently had a MIL on for a catalyst fault, it could still have the fault present and ideally, it would fail I/M until the catalyst monitor was complete. However, if that particular car recently had a MIL on for an EGR fault, it is highly likely that the EGR fault has been confirmed to no longer be present because the readiness status for EGR is complete and there is less likelihood that the vehicle is sneaking through the I/M test with a fault still present, even though the catalyst monitor is still incomplete. Unfortunately, with only the readiness status to make a decision on, there is no way for a technician or I/M program to know which of the above two cases applies to the vehicle. With the permanent DTC method, however, an I/M program could essentially create a custom readiness status for every car and only reject/fail those cars that indeed have recently had the MIL on and have not had an opportunity to re-run that same monitor. For the first case in the above example, a permanent DTC for the catalyst would be present if the car indeed had recently had a catalyst MIL-on fault and had not yet had a chance to re-run the monitor. The lack of a permanent DTC for the catalyst would provide a high degree of confidence that the car does not need to be failed because, even though the catalyst monitor has not run since code clear to reset the readiness status, this particular car has not recently had the MIL on for a catalyst fault. In this manner, I/M programs could reject/fail any vehicle that has a permanent DTC stored in it while it could potentially pass any vehicle that had zero permanent DTCs stored in it.[22]

The permanent DTC method also has advantages for a technician attempting to repair a vehicle and then prepare it for I/M testing. The permanent DTC would identify the specific diagnostic that would need to be exercised after repair and prior to inspection to remove the permanent DTC. By combining this information with the vehicle manufacturer's service information, technicians could identify the exact conditions necessary to operate a particular monitor. As such, technicians could more effectively target after repair verification and would be able to verify that the specific monitor that previously illuminated the MIL has run and confirmed the repair has been made correctly. This also provides added incentive for the technician to "fix it right the first time" and reduces vehicle owner "come-backs" for incomplete or ineffective repairs.

Real Time Indication of Monitor Status

Provisions are also proposed to make it easier for technicians to prepare the vehicle for an I/M inspection following a repair by providing real time data which indicates whether certain conditions necessary to set all the readiness flags to "complete" are currently present. These data would indicate whether a particular monitor still has an opportunity to run on this driving cycle or whether a condition has been encountered that has disabled the monitor for the rest of the driving cycle. While these data would not provide technicians with the exact conditions necessary to exercise the monitors (only service information will do that), this information in combination with the service information should facilitate technicians in verifying repairs and/or preparing a vehicle for inspection. Technicians would be able to use this information to identify when specific monitors have indeed completed or to identify situations where they have overlooked one or more of the enable criteria and need to check the service information and try again.

Communicating Readiness Status to Vehicle Operator

As mentioned above, substantial feedback has been received through the roll-out of OBD II-I/M programs throughout the U.S. and much of this feedback has to do with the issues regarding the effect on consumers because of possible rejection from I/M testing due to unset readiness flags. To address this, some light-duty manufacturers requested the option to communicate the vehicle’s readiness status directly to the vehicle owner without the use of a scan tool. This would allow the vehicle owner to be

---

[22] An I/M program would likely still want to require some or all of the readiness flags to be complete at the time of inspection instead of relying solely on the presence of permanent DTCs. This is due to the structure of most OBD systems, which may disable relevant monitors upon detection of a fault with one or more related components. If the vehicle owner ignored the detected fault for a substantial period of time, other components could have subsequently malfunctioned but will not be monitored until the first malfunction has been repaired. Requiring some or all readiness complete will increase the likelihood that the vehicle is not in a condition to trigger a "chain" of successive faults.

sure that the vehicle is ready for inspection prior to taking the vehicle to an I/M station. Such a provision was recently adopted in the OBD II regulation. The staff is also proposing to allow heavy-duty manufacturers to do the same. If manufacturers choose to implement this option, though, they would be required to do so in the standardized manner prescribed in the proposed regulation. On vehicles equipped with this option, the vehicle owner would be able to initiate a self-check of the readiness status, thereby greatly reducing the possibility of being rejected at the I/M inspection.

D. Fault Codes

Fault codes are the means by which malfunctions are reported by the OBD system and displayed on a scan tool for service technicians. The proposed heavy-duty OBD regulation would require manufacturers to report all emission-related fault codes using a standardized format and to make them accessible to all service technicians, including the independent service industry. The standards selected would define many generic fault codes to be used by all manufacturers. In the rare circumstances that a manufacturer cannot find a suitable fault code already standardized, a unique “manufacturer-specific” fault code could be used. However, these manufacturer-specific codes are not as easily interpreted by the independent service industry. Increased usage of manufacturer-specific codes may increase the time and cost for vehicle repairs. Thus, the proposed regulation would restrict the use of manufacturer-specific fault codes. If a generic fault code suitable for a given malfunction cannot be found, the regulation would require the manufacturer to pursue approval of additional generic fault codes to be added. This proposal would affirm the intent of the OBD regulation to standardize as much information as possible.

Additionally, the staff is proposing that the OBD system store fault codes that are as specific as possible to identify the nature of the fault, which would provide technicians with detailed information necessary to diagnose and repair vehicles in an efficient manner. In other words, manufacturers should use separate fault codes for every diagnostic where the diagnostic and repair procedure or likely cause of the failure is different. Generally, a manufacturer would design an OBD monitor that detects different root causes (e.g., sensor shorted to ground or battery) for a malfunctioning component or systems. The staff expects manufacturers to store a specific fault code such as “sensor circuit high input” or “sensor circuit low input” rather than a general code such as “sensor circuit malfunction.” The staff further expects manufacturers to store different fault codes distinguishing circuit faults from rationality and functional checks, since the root cause for each problem is different, and thus the repair procedures may be different.

For most OBD strategies, manufacturers would be expected to illuminate the MIL only after the same malfunction has occurred on two separate driving events. This “double” detection would ensure that a malfunction truly exists before alerting the owner. The first time a malfunction is detected, a “pending” fault code identifying the suspected failing component or system would be stored in the on-board computer. If the same malfunction is again detected the next time the vehicle is operated, the MIL would be illuminated and a “confirmed” fault code would be stored. A technician would use the “confirmed” fault code to determine what system or component has failed. A “pending” fault code, however, could be used by service technicians to help diagnose intermittent problems as well as to verify that repairs were successful. In these instances, a technician could use the “pending” fault code as a quicker, earlier warning of a suspected (but as yet unconfirmed) problem. The staff is proposing that manufacturers store and make available a “pending” fault code for each currently malfunctioning monitored component or system, regardless of the MIL status or the presence of a “confirmed” fault code. Descriptions of the proposed fault code storage and erasure requirements are described in section II. B. of the Staff Report.

The staff is also proposing requirements that would help distinguish between fault codes stored for present faults and fault codes stored for past faults. As described in section II. B., a manufacturer would generally be allowed to extinguish the MIL if the malfunction responsible for the MIL illumination is not detected (i.e., the monitor runs and determines that the fault no longer exists) on three subsequent sequential driving cycles. However, a manufacturer would not be allowed to erase a confirmed fault code unless the identified malfunction associated with the code is not detected in at least 40 engine warm-up cycles and the MIL is not presently illuminated for the malfunction. So even though the malfunction may no longer be present and the MIL not illuminated, the fault code would still remain as a “history” code. Consequently, if another unrelated fault occurs and the MIL illuminates for this new fault, another fault code would be stored in addition to the “history” code. When trying to diagnose the OBD problem, technicians accessing fault code information may have trouble distinguishing which fault code is responsible for illuminating the MIL (i.e., which fault actually exists), and thus would have problems determining what exactly must be repaired. Therefore, the staff is proposing requirements that would help distinguish a fault code that illuminates the MIL and a “history” code. Specifically, manufacturers would be required to separately identify "history" codes, confirmed codes (that are currently commanding the MIL "on"), and pending codes. The staff is working with the standards setting committees to determine the best method to distinguish these codes (e.g., readable by different scan tool commands, distinguished by some sort of add-on identifier bit).

"Permanent" fault codes (described above in section VII. C.) would also need to be separately identified from the other types of fault codes. The staff is also working with the standards setting committees to best determine the method for doing this, but it will likely be done in a similar manner to that used to distinguish the other types of codes. Additionally, as mentioned above, manufacturers would be required to develop additional software routines to properly store and erase permanent fault codes in NVRAM and prevent erasure from any battery disconnect or scan tool command.

E. Data Stream/Freeze Frame/Test Results

An important aspect of OBD is the ability of technicians to access critical information from the on-board computer in order to diagnose and repair emission-related malfunctions. ARB believes there are certain emission critical components and systems for which electronic information access through the data link connection would provide invaluable assistance in properly repairing vehicles. The availability of real-time information would also greatly assist technicians in responding to driveability complaints because the vehicle could be driven under the problem conditions and the technician would be able to know what the fuel system was delivering at that time. Fuel economy complaints, loss of performance complaints, intermittent problems, and others could also be addressed.

The proposed regulation defines a number of data parameters that manufacturers would be required to report to generic scan tools. These parameters, which would include information such as engine speed and exhaust gas sensor readings, would allow technicians to understand how the vehicle engine control system is functioning, either as the vehicle operates in a service bay or during actual driving. They would also help technicians diagnose and repair emission-related malfunctions by allowing them to watch instantaneous changes in the values while operating the vehicle.

In the event an emission-related malfunction is detected by the OBD system, the proposed regulation would also require manufacturers to make available “freeze frame” information, which displays the operating conditions of the vehicle at the time of a malfunction detection, in addition to the fault code associated with the data. The required freeze frame data would include the absolute load value, engine speed, vehicle speed, and engine coolant temperature. Further, the required freeze frame data would be required to include all other standardized data parameters available in the on-board computer that detected and stored the fault. For the purposes of this requirement, "available" means any other data parameter that is input to (directly wired or sent via other modules or network messages) or calculated within the on-board computer. This would allow the freeze frame data to assist the technician in two ways. First, the technician should be able to identify how the vehicle was being operated by the driver at the time of the fault should he or she need to duplicate the driving conditions to find an intermittent malfunction or verify a repair under the same conditions where it was originally detected. Second, the inclusion of all other available data provides the technician with the ability to "see" some of what the on-board computer was seeing when it set the malfunction. This can be particularly useful when a specific fault is indeterminate (e.g., could have been caused by more than one root cause or more than one malfunctioning sensor).

The proposed regulation would also require manufacturers to store the most recent monitoring results for most of the major monitors. Manufacturers would be required to store and make available to the scan tool certain test information (i.e., the minimum and maximum values test limits as well as the actual test value) of the most recent monitoring event. “Passing” systems would store test results that are within the test limits, while “failing” systems would store results that are outside the test limits. The storage of test results would greatly assist technicians in diagnosing and repairing malfunctions and would help distinguish between components that are performing well below the malfunction thresholds from those that are potentially marginally passing the malfunction thresholds.

F. Identification Numbers (Cal ID, VIN, CVN)

OBD diagnostics are comprised of software routines and calibrated limits and values to determine if a component or system is malfunctioning. When dealing with the OBD II systems on light- and medium-duty vehicles, manufacturers often released updates to the software in the on-board computer to add new features and improvements or to correct errors or "bugs" found in the system. The staff is anticipating the same situation would occur for heavy-duty vehicles. Thus, to determine if the correct software has been installed, the staff is proposing that manufacturers be required to report two identification numbers. The first item, Calibration Identification Number (CAL ID), would identify the version of software installed in the vehicle. Subsequent releases of software by the manufacturer that make changes to the emission controls or OBD system would require a new CAL ID. The second item, Calibration Verification Number (CVN), would help ensure that the software has not been inappropriately corrupted, modified, or tampered with. Both CAL ID and CVN requirements were adopted for the light- and medium-duty OBD II regulation to ensure the integrity of the OBD II system during I/M inspections. CVN would require manufacturers to develop sophisticated software algorithms that can verify the integrity of the emission-related software and ensure that the diagnostic routines and calibration values have not been modified inappropriately. The CVN would essentially be a self-check calculation of all of the emission-related software and calibration values in the on-board computer and would return the result of the calculation to a scan tool. If the calculated result did not equal the expected result for that CAL ID, the software would be known to be corrupted or otherwise modified. The proposed regulation would require that the CVN result be made available at all times to a generic scan tool.

The proposed regulation would also require manufacturers to make available an additional identification number, the Vehicle Identification Number (VIN), in a standardized format. The VIN would be a unique number assigned by the vehicle manufacturer to every vehicle built. The VIN is commonly used for purposes of ownership and registration to uniquely identify every vehicle. For the light- and medium-duty OBD II regulation, the VIN is used during an I/M inspection to identify the exact vehicle being tested. For the heavy-duty industry, the VIN is used to identify the vehicle on citations or notice-of-violations (NOVs) issued at roadside inspections under the HDVIP. Without the adoption of a VIN requirement, I/M inspectors could fraudulently pass failing vehicles by manually entering the VIN of one vehicle and performing an emissions test on a known "clean" vehicle (a practice known as "clean-piping" or "clean-scanning"). By requiring the VIN to be stored in the vehicle and available electronically to a generic scan tool, the possibility of a technician performing a fraudulent inspection would be smaller. Electronic access to this number would also greatly simplify the inspection process and reduce transcription errors from manual entry. The VIN requirement would also serve the same purpose for an OBD-based heavy-duty I/M program.

The proposed heavy-duty OBD regulation would require the VIN to be electronically stored in a control module, not necessarily the engine control module, in the vehicle. As long as the VIN is correctly reported according to the standards selected, it is irrelevant as to which vehicle module (e.g., engine controller, instrument cluster controller) contains the information. And, while the ultimate responsibility would lie with the engine manufacturer to ensure that every vehicle manufactured with one of its engines satisfied this requirement by having the VIN available, the physical task of implementing this requirement would likely be passed from the engine manufacturer to the vehicle manufacturer via an additional build specification. Thus, analogous to how the engine manufacturer currently provides engine purchasers with detailed specifications regarding engine cooling requirements, additional sensor inputs, physical mounting specifications, weight limitations, etc., the engine manufacturer would likely include an additional specification dictating the need for the VIN to be made available electronically. It would be left to each engine manufacturer to determine the most effective method to achieve this, as long as the VIN requirement is met. Some manufacturers may find it most effective to provide the capability in the engine control module delivered with the engine coupled with a mechanism for the vehicle manufacturer to program the module with the VIN upon installation of the engine into an actual vehicle. Others may find it more effective to require the vehicle manufacturer to have the capability built into other modules installed on the vehicle such as instrument cluster modules, etc. It should also be noted that staff has observed several current vehicles with engines from three different engine manufacturers that already have the VIN available through engine-manufacturer specific scan tools so achieving this requirement for all vehicles should not be infeasible.

G. Tracking Requirements

In-use Performance Ratio Tracking Requirements

The tracking requirements for the in-use performance ratios are discussed in section VI of the Staff Report and listed in the proposed regulation.

Engine Run Time Tracking Requirements

Numerous new emission requirements for heavy-duty engines are taking effect over the next five years. One of those requirements is commonly referred to as the "not-to-exceed", or NTE, standards. These standards define a wide range of engine operating points where a manufacturer must design the engine to be below a maximum emission level. In theory, whenever the engine is operated within the speed and load region defined as the NTE zone, emissions will be below the required standards. Additionally, the NTE zone was defined to try and cover the portion of the engine speed and load region most commonly encountered during in-use driving. Thus, the engine operation conditions most widely used by vehicle operators during in-use driving should be covered and result in the engine operating at or below the required emission standards.

However, heavy-duty engines are used in a variety of vehicle applications and there is no guarantee that all vehicle applications will be designed to operate the engine in the same speed and load regions or even at speed and load regions that are within the NTE zone. Additionally, manufacturers make use of a provision that allows them to use auxiliary emission control devices (AECDs), even while operating within the NTE zone. Typically, AECDs consist of alternate control strategies or actions taken by the engine controller for purposes of engine, engine component, or emission control component protection or durability. AECDs usually involve the engine controller utilizing various sensors or combinations of them to determine that action must be taken to protect the engine or emission control components. An example of a common AECD is when the engine enters into over-temperature conditions, in which case the engine controller will take actions such as modifying fuel injection timing or EGR flow to attempt to reduce engine temperature. Manufacturers are currently required to disclose details of all such AECD strategies at the time of certification, including what conditions are necessary to activate the AECD, what action is taken when the strategy is invoked, what the emission impact is when the strategy is invoked, and how often the AECD is expected to be invoked during in-use driving. ARB staff reviews the strategies and approves them upon finding the manufacturer has validly demonstrated the technical need for such AECDs and that they will be encountered as infrequently as possible in-use such that they should not have a substantial in-use emission impact.

However, estimates of how often these AECDs are expected to be activated in-use is generally based on some generic model data or a couple of hours worth of data from one or two test vehicles. Given the wide variety of in-use driving patterns throughout the heavy-duty industry, it is highly unlikely that these few data points adequately represent all of the various in-use applications. Further, since engines are commonly sold to various coach builders and each coach builder installs the engine in different vehicles with different aerodynamic factors, expected duty cycles, and engine cooling capacity, the performance of the in-use vehicles is likely to be significantly different from the few pre-certification test points. Accordingly, the expected emission impact from activation of these AECDs could be grossly over or under estimated. Thus, since ARB's approval of such strategies is generally partially predicated on the small emission impact due to the expected infrequent operation of the AECDs in-use, a significant difference in actual operation versus the numbers estimated at the time of certification would be material to the staff's determination to grant approval.

Accordingly, the staff is proposing a requirement for manufacturers to track both vehicle operation inside and outside of the NTE zone as well as frequency of AECD operation on all in-use vehicles. This would provide a valuable feedback mechanism to validate both the currently defined speed and load window of the NTE zone as well as the original estimates by the manufacturer as to how often the AECDs will be invoked. Such feedback could be incorporated into subsequent model year certifications to better adjust the estimates to historical in-use performance or even into future regulatory modifications or new standards. Further, such data could provide the staff with valuable information to be used in enforcement cases where actual heavy-duty implementations (either inside or outside of the engine manufacturer's build specifications) have resulted in non-compliant systems (e.g., an application built with insufficient engine cooling such that the vehicle routinely encounters an engine overheat condition under normal operation and relies on the activation of one or more AECDs to maintain acceptable vehicle performance).

While the staff is still open to feedback from industry as to the best method and metric to track in-use performance, staff's proposal would be to have several cumulative engine operating hour counters to track the various states of operation. Thus, the proposed regulation would require a manufacturer to store counters that would log engine run <u>time</u> in the following conditions:

(a) Engine running (i.e., total engine run time);
(b) Engine running and in NTE zone as defined by engine speed and load;
(c) Engine running and operating at idle;
(d) Engine running and operating with power take-off (PTO) active;
(e) Engine running and operating with AECD#1 active
(f) “ “ “ “ “ AECD#2 active
(…) “ “ “ “ “ AECD#n active (for engines with “n” total AECDs).

The proposed regulation would require all these counters to be stored in non-volatile memory (NVRAM) so that vehicle owners would not be able to erase them simply by disconnecting the battery nor would the values be able to be erased by using a scan tool.

The tracking of total engine run time, engine run time in the NTE zone, and engine run time while an AECD is active (conditions (a), (b), and (e)-(…) above) could give ARB the ability to sample a fleet of trucks and determine if the manufacturer's claims about estimated frequency are accurate or not. They could also verify which AECDs are being activated routinely and how often a vehicle is operating in the NTE zone versus outside the NTE zone. Additionally, this information could reduce the certification review burden on ARB and engine manufacturers, since it would allow manufacturers to provide more accurate data documenting the in-use frequency of various strategies. Regarding the tracking of AECD operation (condition (e) above), manufacturers would be required to store separate counters for each individual AECD strategy the engine employs. Each counter would log the engine operating time while each of the AECD strategies is active. Each individual AECD would be identified with a manufacturer-specified number (i.e., #1 through #n), and engine manufacturers would be responsible for submitting information identifying the manufacturer-specific numbered AECD during certification so that ARB could easily reference these documents when attempting to interpret which AECD strategy each counter is referring to.

Regarding the tracking of idle operation (condition (c) above), engine idle operation is a unique operating condition that is outside of the NTE zone and often has its own AECD strategies. In some truck applications such as long-haulers with sleeper cabs, considerable time can be spent operating at idle. By requiring manufacturers to implement a separate counter identifying "engine operating at idle," the staff would be better able to separate out engine run time while at idle operation from non-idle. Further, under a separate rulemaking, ARB is proposing idle-off requirements to minimize time spent at idle and require engine manufacturers to implement strategies that forcibly turn off the engine after a specified amount of idle operation. By tracking idle operation, staff would be able to better quantify how well such strategies are working. Future heavy-duty I/M inspections could also use this parameter to help identify vehicles that have been tampered or otherwise modified to bypass the idle-off strategies.

The tracking of engine running time while PTO is active (condition (d) above) would also be helpful data for ARB staff. It is anticipated that engine manufacturers would generally disable some OBD monitors when PTO is active. By requiring this counter to be tracked, the staff would be better able to interpret in-use monitoring frequency data (as detailed in section VI. of the Staff Report). Specifically, for monitors that seem to demonstrate very low monitoring frequency, the staff could determine if this was due to frequent PTO activation (if the PTO active counter was really high) or due to other conditions.

H. Service Information

Once a malfunction has been detected by the OBD system, the emission reduction benefits are obtained only when the problem is corrected. When repairing an OBD-related problem, a repair technician generally accesses the available information from the on-board computer to determine the component or system that failed. After repairing the malfunction, the vehicle would then be driven in a manner such that the monitor for the malfunctioning component runs and determines that the fault no longer exists. In order to do this, the repair technician would need information that would help pinpoint the malfunctioning component, determine the cause of the malfunction, and ensure that the problem has indeed been corrected. Therefore, access to adequate service information is an important part of the OBD program. Specifically, all emission-related vehicle service information necessary to make use of the OBD system and to perform emission-related repairs should be made available to all service technicians, including independent and aftermarket service technicians, and in a format for easy accessibility of the information.

For the light- and medium-duty vehicles, the service information requirements are detailed in a stand-alone regulation, section 1969 of title 13, California Code of Regulations, which requires this information to be made available on the internet. The required information includes OBD monitor descriptions, information necessary to execute each monitor (e.g., enable conditions), information on how to interpret the test data accessed from the on-board computer, and other information. ARB is currently revising section 1969 to include service information requirements for heavy-duty vehicles. Manufacturers would be required to provide the same OBD-related information that is currently being required for the light- and medium-duty vehicles in the same format. Specifically, heavy-duty manufacturers would be required to create Internet websites that meet the requirements of section 1969 and contain all of the required information. Under the provisions of section 1969, manufacturers are allowed to charge for access to the website and typically offer multiple subscription options including daily, monthly, and yearly subscription prices.

However, because the proposed amendments to section 1969 (which is scheduled to go before the Board in December 2003) will not be adopted before the proposed heavy-duty OBD regulation is adopted, the proposed heavy-duty OBD regulation would include language detailing basic service information requirements. Additionally, the staff is including language in the proposed OBD regulation that would clarify that, to the extent the service information regulation is effective and operative, it would supercede any redundant service information requirement in the proposed OBD regulation.

VIII. DEMONSTRATION TESTING REQUIREMENTS

As stated previously, the OBD system is designed to detect malfunctions of the emission control system to help prevent increases in emission levels. The proposed OBD regulation would require manufacturers to design OBD monitors for each emission-related component or system to indicate a malfunction before emissions exceeded a proposed emission malfunction threshold (generally 1.5 times the applicable standards for most monitors). While the proposed certification requirements (discussed in section IX of the Staff Report) would require manufacturers to submit technical details of each monitor (e.g., how each monitor worked, when the monitor would run), ARB staff would still need some assurance that the manufacturers' monitors are indeed able to detect a malfunction before the emission threshold is exceeded. Thus, in order to verify that the OBD malfunction threshold values set by manufacturers are appropriate, the staff is proposing that manufacturers conduct demonstration testing on its major monitors validating their malfunction threshold values. The proposed heavy-duty regulation would require manufacturers to submit documentation and emission data demonstrating that the major monitors are able to detect a malfunction before emissions exceed the emission threshold (e.g., 1.5 times the applicable standards) as part of the proposed certification requirements. In addition to testing the system with "threshold" components (e.g., components that are deteriorated or malfunctioning right at the threshold required for MIL illumination), manufacturers would also be required to test the system with "worst case" components. By testing both the threshold, or best performing failing system, and the worst case, or worst performing failing system, the staff would be better able to verify that the OBD system should perform as expected regardless of the level of deterioration of the component. This could become increasingly important with new technology aftertreatment devices that could be subject to complete failure (such as PM traps) or even to tampering by vehicle operators looking to improve fuel economy or vehicle performance. The demonstration test procedures are detailed in the proposed regulation.

The number of test vehicles manufacturers would be required to conduct demonstration testing on would be based on the total number of engine families the manufacturer would be certifying for that model year. Specifically, a manufacturer certifying one to five engine families in a model year would be required to conduct demonstration testing and submit emission data from a test vehicle from one engine family. A manufacturer certifying six to ten engine families would be required to submit data from test vehicles from two engine families. Lastly, a manufacturer certifying more than ten engine families would be required to submit data from test vehicles from three engine families. Given the difficulty and expense in removing an in-use engine from a vehicle for engine dynamometer testing, this demonstration testing would likely represent nearly all of the OBD emission testing that would likely ever be done on these engines. Requiring a manufacturer, who is fully equipped to do such testing and already has the engines on engine dynamometers for emission testing, to test one to three engines per year would be a minimal testing burden that provides invaluable (and in a practical sense, nearly otherwise unobtainable) proof of compliance with the OBD malfunction thresholds.

Regarding the selection of which engine families would be demonstrated, manufacturers would be required to submit descriptions of all engine families planned for the upcoming model year and the Executive Officer would review the information and make the selection(s). For each engine family, the information submitted by the manufacturer would need to identify engine model(s), power ratings, emission standards, emission controls used by the engine, and projected engine sales volume (including a sales breakdown by engine purchaser/coach builder and vehicle application and transmission usage). Factors that would be used by the Executive Officer in selecting the one to three vehicle configurations/variants within the engine families for testing include, but are not limited to, new engines, transition to more stringent emission thresholds, sales volume, and highest sales volume vehicle applications/configurations.

Manufacturers required to submit data from more than one engine family would be granted some flexibility by allowing the data to be collected under less rigorous testing requirements than the official FTP certification test. That is, for the second and third engine families required for testing, manufacturers would be allowed to submit data using internal sign-off test procedures that are representative of the official FTP test in lieu of running the official test. Commonly used procedures include the use of engine test cells with less rigorous quality control procedures than those required for the FTP or the use of forced cool-downs to minimize time between tests. Manufacturers would still be liable for meeting the malfunction thresholds on official tests run according to the FTP procedure. However, this latitude would allow them to potentially use some short-cut methods that they have developed to ensure themselves that the system is calibrated to the correct level without incurring the additional testing cost and burden of running the official FTP test procedure on every application.

IX. CERTIFICATION REQUIREMENTS

The OBD system certification requirements would require manufacturers to submit diagnostic system documentation representative of each engine family. The certification documentation would contain all the information needed for ARB to determine if the OBD system meets the proposed requirements of the heavy-duty OBD regulation. The proposed regulation would list all the information that is required to be in the certification package. If any of the information in the certification package is standardized for all of a manufacturer's engine families (e.g., the OBD system general description), the manufacturer would only be required to submit one set of documents covering the standardized items for all of its engine families per model year.

As is done with medium-duty applications that use engines certified on an engine dynamometer, engines would be required to be certified to the OBD requirements applicable to the designated vehicle model year, not the engine model year. Thus, all designations of model years in the proposed regulation are applicable to vehicle model year. For instance, an early introduction 2007 model year vehicle (built during the middle or end of 2006 calendar year) may actually contain a 2006 model year engine. However, under the OBD requirements, the engine would be required to meet all OBD requirements applicable to the 2007 model year if it was installed in a 2007 model year vehicle. This would help avoid confusion during certification and during an OBD-based I/M inspection in cases where the vehicle is labeled as a particular model year but the engine is actually a previous model year. Most motor vehicle registration and I/M inspection requirements are based on vehicle model year, not engine model year, and accordingly, the I/M requirements will likely be applied on a vehicle model year basis. Allowing a particular vehicle model year to have both previous model year OBD systems and current model year OBD systems could result in an inappropriate I/M test being applied. This model year designation is only used for purposes of determining compliance with the heavy-duty OBD monitoring requirements, so all other certification requirements for engines (e.g., emission standards) would remain unaffected and would continue to be applied as they are currently. Thus, a 2007 model year truck could be built with an engine that meets 2006 model year engine dynamometer emission standards and 2007 model year OBD standards.

Further, as is also done with medium-duty applications that use engines certified on an engine dynamometer, manufacturers would be required to get OBD approval of the entire vehicle package, not just the portions of the proposed OBD requirements that are incorporated into engines. While the majority of the proposed OBD requirements would apply to the engine and be incorporated by design into the engine control module, a significant portion of the proposed OBD requirements would apply to the vehicle and not be self-contained within the engine. Examples include the proposed requirements to have a MIL in the instrument cluster, a diagnostic connector in the cab compartment, and comprehensive component monitors on all electronic powertrain components that affect emissions such as electronic transmission components or hybrid powertrain components. As such, on medium-duty applications, OBD II approval is given for a particular combination of engine, transmission, and vehicle. The staff is expecting to follow a similar method for the heavy-duty industry because it is the only practical method that the staff is aware of to verify that the entire complement of OBD requirements are indeed satisfied prior to making the determination that the system complies with the proposed OBD requirements.

The staff does, however, see opportunities to simplify the process. As is currently done by the engine manufacturers, a build specification is provided to vehicle manufacturers detailing mechanical and electrical specifications that must be adhered to for proper installation and use of the engine (and to maintain compliance with emission standards). The staff expects engine manufacturers will continue to follow this model in providing detailed specifications for those items that the vehicle manufacturer will need to be aware of or responsible for to maintain compliance with the proposed OBD regulation. These would include specifications regarding the location, color, and wording of the MIL (as well as electrical connections to ensure proper illumination), location and type of diagnostic connector, electronic VIN access, etc. Rather than requiring the engine manufacturer to provide details of every single variation of vehicle that will use a particular engine to ensure these requirements are satisfied, the staff is anticipating that OBD approval for these types of items could be based on the engine manufacturer's build specification. During the certification process, in addition to submitting the details of all of the diagnostic strategies and other information required, engine manufacturers could submit a copy of the OBD-relevant build specifications provided to vehicle manufacturers and a description of the method used by the engine manufacturer to ensure vehicle manufacturers adhere to the provided build specifications (e.g., required audit procedures or signed agreements to adhere to the requirements). This would greatly simplify the certification process while still providing the staff with a reasonable level of verification that the proposed OBD requirements are indeed satisfied.

Engine manufacturers would thus be responsible for submitting a certification package that includes description of all OBD diagnostics performed by the engine control unit (including diagnostics on signals or messages coming from other modules that the engine control unit relies on to perform other OBD diagnostics) as well as a copy of the OBD-relevant build specifications provided to chassis builders and the method used to reasonably ensure compliance with those build specifications.

Regarding electronic powertrain components other than those used by the engine control unit that require monitoring under the comprehensive component monitoring requirements such as electronic automatic transmissions or hybrid drivetrain system components, the staff would need to review and approve the details of monitoring strategies, etc. with the same level of documentation as required for engine component diagnostics. To accomplish this, it is anticipated that most transmission manufacturers would develop their own self-contained diagnostics and develop certification documentation describing those diagnostics (including diagnostics on signals or messages received from other control modules and relied upon for other OBD diagnostics). These transmission suppliers would then need to separately apply to ARB for OBD approval of their transmission diagnostics. If the diagnostics meet the proposed requirements, the transmission would be approved by the staff as OBD compliant. Further, the engine manufacturer would then need only to reference the use of a previously approved transmission (if applicable) or incorporate the requirement to equip the vehicle with an OBD-approved transmission into the build specification. Given the relatively small number of transmission suppliers in the heavy-duty industry and the fairly limited number of transmission models, the staff believes this would be the most effective way to ensure compliance with the requirements on automatic transmissions.

For components that are part of systems other than the engine or transmission such as a hybrid powertrain, the engine manufacturer would be required to submit the details along with the engine description at the time of certification. Given the interaction between a hybrid powertrain and the engine itself, it is expected that engine manufacturers will be substantially involved in the system before allowing a different module on the vehicle to take more direct control of the engine (e.g., turning on or off the engine during vehicle operation or augmenting engine power) to ensure adequate driveability and performance of the complete system. Further, the impact of malfunctions within the hybrid system on the engine (e.g., requiring the engine to work harder or more often) will likely dictate which hybrid components necessitate monitoring and cannot be completely isolated or handled separately from the engine. As such, it is appropriate that engine manufacturers are involved in certification of the complete system and take the responsibility to coordinate the entire diagnostic package at the time of certification.

The proposal would allow manufacturers (engine and transmission) to establish OBD groups consisting of test groups with similar OBD systems and submit only one set of representative OBD information from each OBD group. The staff anticipates the representative information will normally consist of an application from a single representative test group. In selecting the representative test group, the manufacturer would need to consider tailpipe and evaporative emission standards, OBD phase-in requirements (i.e., if a representative test group meets the most stringent monitoring requirements), and the exhaust emission control components for all the test groups within an OBD group. For example, if one test group within an OBD group has additional emission control devices, that test group should be selected as the representative test group. If one test group does not adequately represent the entire OBD group, the manufacturer may need to provide information from several test groups within a single OBD group to ensure the submitted information is representative. Manufacturers wishing to consolidate several test groups into an OBD group would be required to get ARB approval of the grouping prior to submitting the information for certification.

Two of the most important parts of the certification package would be the OBD system description and summary table. The OBD system description would include a complete written description for each monitoring strategy outlining every step in the decision-making process of the monitor, including a general explanation of the monitoring conditions and fault criteria. This section may include graphs, diagrams, and/or other data that would help the staff in understanding each monitor. Specific parameter values would be included in the OBD summary table. This table would provide a summary of the OBD system specifications, including: the component/system, the fault code identifying each related malfunction, the monitor strategy, the parameter used to detect a fault and the fault criteria limits to evaluate the parameter (the malfunction criteria and threshold value), secondary parameter values and conditions needed to run the monitor, the time required to execute a monitoring event, and the criteria or procedure for illuminating the MIL. In these tables, manufacturers would be required to use a common set of engineering units to simplify and expedite the review process by ARB staff.

Among the other items that would be required for submittal include: a logic flowchart for each monitor illustrating the step-by-step decision process for determining malfunctions, data supporting the criteria used to detect faults that cause emissions to exceed the specified malfunction thresholds (e.g., 1.5 times the FTP standards) for fuel system, EGR, boost pressure, catalyst, NOx adsorber/trap, PM trap, cold start strategy, secondary air, evaporative system, VVT system, and exhaust gas sensor monitors, data demonstrating the probability of misfire detection by the misfire monitor over the full engine speed and load operating range (for gasoline engines only) or the capability of the misfire monitor to correctly identify a one cylinder out misfire for each cylinder (for diesel engines only), a description of all the parameters and conditions necessary to begin closed-loop fuel control operation (for gasoline engines only), closed-loop EGR control (for diesel engines only), closed-loop fuel pressure control (for diesel engines only), and closed-loop boost control (for diesel engines only), a listing of all electronic powertrain input and output signals (including those not monitored by OBD) that identifies which signals are monitored by the OBD system, detailed descriptions of all the auxiliary emission control device (AECD) strategies used by the manufacturer, and the emission data from the demonstration testing (as described in section VI). The proposed regulation lists the rest of the information that is required to be in the certification package.

X. PRODUCTION VEHICLE EVALUATION TESTING REQUIREMENTS

A. Verification of Standardized Requirements

An essential part of OBD systems is the numerous standardized requirements that manufacturers would have to design to. The proposed standardized requirements include items as simple as the location and shape of the diagnostic connector (where technicians can "plug in" to the on-board computer) to more complex subjects concerning the manner and format in which fault information is accessed by technicians via a “generic” scan tool. The importance of manufacturers meeting these standardized requirements is essential to the success of the heavy-duty OBD program, since it would ensure access for all technicians to the stored information in the on-board computer in a consistent manner. The need for consistency is even higher with the potential incorporation of OBD into a future heavy-duty I/M (which relies on access to the information via a “generic” scan tool). In order for I/M inspections to work effectively and efficiently, it is essential that all vehicles are designed *and built* to meet all of the applicable standardized requirements.

While it is anticipated that the vast majority of vehicles would comply with all of the necessary requirements, some problems involving the communication between vehicles and “generic” scan tools may occur in the field as it did for the light- and medium-duty vehicles. The cause of the problem could range from differing interpretations of the existing standardized requirements to oversights by the design engineers to hardware inconsistencies or last minute production changes on the assembly line. To try and minimize the chance for such problems on future vehicles, the staff is proposing that manufacturers be required to test a sample of production vehicles from the assembly line to verify that the vehicles have indeed been designed and built to the required specifications for communication with a “generic” scan tool.

Under the proposal, manufacturers would be required to test one vehicle per software "version" released by the manufacturer to ensure it complies with some of the basic “generic” scan tool standardized requirements, including those that are essential for proper I/M inspection. With proper demonstration, manufacturers would be allowed to group different calibrations together to be demonstrated by a common vehicle. Prior to acquiring these data, the proposed regulation would require engine manufacturers to submit for ARB review and approval a test plan that verifies the vehicles tested would be representative of all vehicle configurations (e.g., each ECM variant coupled with and without the other available vehicle components that could affect scan tool communication such as automatic transmission or hybrid powertrain control modules). The plan would include details on all the different applications and configurations that would be tested. Additionally, manufacturers would be required to conduct this testing on actual production vehicles, not stand-alone engines. In the past, the staff found that light-duty vehicles that do not properly communicate with a scan tool or I/M equipment cause huge problems at repair facilities and I/M stations, since technicians are unable to access all of the necessary emission-related information from the vehicle’s on-board computer. In fact, it is such a big issue that under the light- and medium-duty OBD II enforcement regulation (section 1968.5), this specific problem has been identified as one that would result in mandatory recall. Thus, to avoid this problem with heavy-duty vehicles, it is imperative that the proposed testing be representative of all applications. Further, the staff has also had numerous issues in the past with light-duty vehicles where, despite each controller independently working properly, interaction problems between two controllers (e.g., ECM and TCM) have caused communication problems with scan tools, such as lack of communication or communication with only one module. In this case, separate testing of the controllers would be blind to this problem. There have even been cases where interaction problems between emission-related controllers and non-emission-related controllers (e.g., ABS, airbag) have caused scan tool communication problems. Since heavy-duty engine manufacturers are expected to sell the same engine (with the same calibration) to various coach builders who would put them in different final products (e.g., with different TCMs), the same communication problem would be expected to occur. Furthermore, on some occasions, the staff has found applications that communicated properly with generic scan tools during development but last minute production changes (including component supplier changes, etc.) have caused actual production vehicles to differ from pre-production development vehicles and to not properly communicate. Thus, for heavy-duty vehicles, it would be necessary to have proposed testing done on the end vehicle product, not just the engine, and to have the proposed testing be representative of all possible configurations of controllers.

Verification testing of standardized requirements should occur early enough to provide manufacturers with early feedback of the existence of any problems and time to resolve the problem prior to the introduction of the entire model year of vehicles being introduced into the field. The proposed regulation would require that testing be done and data submitted to ARB within six months of the start of normal engine production. However, understanding that a substantial delay may occur from the start of engine production to the start of vehicle production (and thus, availability of an end vehicle product), upon good cause, the Executive Officer may extend this time period for testing.

To verify that all manufacturers are testing vehicles to the same level of stringency, the proposed regulation would require the vehicle manufacturers to get ARB approval of the testing equipment used by the manufacturer to perform this testing. ARB approval of the testing equipment would be based upon whether the equipment can verify that the OBD system complies with the standardized requirements and will likely communicate properly with any off-board test equipment (e.g., generic scan tools) that is also designed to meet the standardized requirements. The staff anticipates that the vehicle manufacturers and scan tool manufacturers will likely develop a common piece of hardware and software which could be used by all vehicle manufacturers at the end of the assembly line to meet this requirement. Two different projects (SAE J1699 and LOC3T) are already far along in the process for developing such equipment under the light-duty OBD II requirements. It is anticipated that this equipment will be ready for the 2005 model year for light-duty and similar type equipment could be developed in time for the 2007 model year for the heavy-duty industry and communication standards selected. Ideally, this test procedure will verify each and every requirement of the communication specifications including the various physical layers, message structure, response times, message content, etc.

It is important to note, however, that this verification equipment would not replace the function of existing “generic” scan tools used by technicians or I/M inspection stations. This equipment would be custom designed and used expressly for the purposes of this assembly line testing and would not include all of the necessary features for technicians or I/M inspectors.

B. Verification of Monitoring Requirements

The proposed OBD regulation would require comprehensive monitoring of virtually every component on the vehicle that can cause an increase in emissions. To accomplish this task, manufacturers would need to develop sophisticated diagnostic routines and algorithms that are programmed into software in the on-board computer and calibrated by vehicle engineers. This would translate into thousands of lines of software programmed to meet the diagnostic requirements but not interfere with the normal operation of the vehicle. While most manufacturers would likely develop extensive verification or "sign-off" test procedures to ensure that the diagnostics function correctly, problems could and will probably happen. Moreover, the majority of the validation testing done by the manufacturer would probably focus on finding problems that would be noticed by the vehicle operator such as those that will cause the MIL to falsely illuminate when no malfunction really exists rather than verifying that the MIL will indeed illuminate when a malfunction does exist.

The problems that occur could vary greatly in severity from essentially trivial mistakes that have no noticeable impact on the OBD system to situations where significant portions of the OBD system and normal vehicle fuel and emission control system are disabled. Furthermore, it is often very difficult to assess the impact the problem may or may not have on vehicles that will be on the road for the next 10-30 years. The cause of the problems could also vary from simple typing errors in the software to carelessness to unanticipated interactions with other systems or production or component supplier hardware changes.

In an attempt to minimize the chance for significant problems going undetected and to ensure that all manufacturers are devoting sufficient resources to verifying the performance of the system, the staff is proposing that engine manufacturers be required to perform a thorough level of validation testing on two to six actual production vehicles per model year and submit the results to ARB. As with the proposed standardization requirement testing (section X. A. above), engine manufacturers would be required to conduct this testing on real production vehicles, not stand-alone engines. This would give a better idea of how monitors would run on actual vehicles in the real world, since some OBD monitors would most likely rely on data or information from other modules on the vehicle. Additionally, similar to the demonstration testing requirement (section VIII. of the Staff Report), the number of vehicles engine manufacturers would be required to test would be based on the total number of engine families the manufacturer would be certifying for that model year. Specifically, an engine manufacturer certifying one to five engine families in a model year would be required to conduct testing on vehicles from two engine families. An engine manufacturer certifying six to ten engine families would be required to conduct testing on vehicles from four engine families. Lastly, an engine manufacturer certifying more than ten engine families would be required to conduct testing on vehicles from six engine families. As with durability demonstration vehicle testing, the Executive Officer would select the two to six vehicle applications/configurations/variants to be tested by the manufacturer from the information submitted by the manufacturer.

For the testing, engine manufacturers would be required to individually implant or simulate malfunctions to verify that virtually every single engine-related OBD diagnostic on the vehicle correctly identifies the malfunction. Prior to testing, manufacturers would be required to submit a test plan for review and approval by the Executive Officer detailing the method used to implant each fault and verify proper diagnostic operation. The Executive Officer would exempt manufacturers from testing that could not be done without causing physical damage to the production vehicle. The testing would be required to be completed and reported to ARB within six months after a manufacturer begins normal engine production to provide early feedback on the performance of every diagnostic on the vehicle. As with the other production vehicle evaluation testing requirements, upon good cause, the Executive Officer may extend this time period for testing.

The staff is also proposing that electronically-controlled transmission manufacturers be required to perform validation testing of their OBD transmission monitors on production vehicles and submit the results to ARB. Similar to engine manufacturers, transmission manufacturers would be required to conduct this testing on real production vehicles, not stand-alone transmissions. The number of vehicles transmission manufacturers would be required to test would be based on the number of electronically-controlled transmission models that are sold for that model year. Specifically, manufacturers would be required to test one production vehicle for every three transmission models offered by the transmission manufacturer for use in OBD-equipped heavy-duty vehicles. For the testing, transmission manufacturers would be required to individually implant or simulate malfunctions to verify that virtually every single transmission-related OBD diagnostic on the vehicle correctly identifies the malfunction. Transmission manufacturers would be required to conduct the testing and submit the results to ARB within six months after transmission production starts. Further, upon good cause, the Executive Officer may extend this time period for testing.

As an incentive to perform this thorough validation testing, a manufacturer could request that any problem discovered during this self-testing be evaluated as a deficiency and take effect retroactively to the start of production of the engine. If the other factors necessary to qualify for a deficiency are indeed satisfied, the Executive Officer would amend the certification to retro-actively assign the deficiency to the start of production of the affected engines. In contrast, problems discovered later by ARB staff during in-use testing would become noncompliance issues and handled in accordance with OBD-specific enforcement regulations.[23]

C. Verification and Reporting of In-use Monitoring Performance

The staff is proposing that manufacturers track the performance of several of the most important monitors on the vehicle to determine how often they are executing during in-use operation. These requirements are discussed in more detail in section VI of the Staff Report. Essentially, the proposed regulation would standardize a method for measuring and determining how often monitors are executing in the real world and set a minimum acceptable performance level. Monitors that perform below the acceptable levels would be subject to remedial action including potential recall.

---

[23] While the regulatory package being considered for adoption does not currently include a separate OBD-specific enforcement regulation due to time and resource constraints, the staff intends to come back to the Board with a proposed enforcement regulation prior to the introduction of OBD systems on heavy-duty vehicles. It is the staff's intention to have a stand-alone OBD enforcement regulation, analogous to the separate OBD II enforcement regulation for light-duty vehicles, title 13 CCR section 1968.5. See section I of the Staff Report for more details.

In conjunction with the proposal to measure in-use monitoring frequency, the staff is also proposing that manufacturers be required to collect this in-use data within the first six months after engine production begins. This information would provide ARB with early indication as to whether or not the system is performing adequately as well as provide valuable feedback as to the appropriateness of the minimum ratio. As discussed extensively in section VI, the staff is proposing a ratio of 0.100 for the first few years of implementation primarily because a sufficient database does not currently exist that would allow the staff to develop a more accurate estimate of fault detection in a reasonable time period such as two weeks. The requirement for manufacturers to collect and report some of these data in the early years would provide an invaluable source of real world data and allow the staff to revise the regulatory requirements as necessary to establish a ratio that more closely correlates with the desired in-use monitoring frequency.

Prior to acquiring these data, engine manufacturers would be require to submit for ARB review and approval a sampling plan that verifies the data collected would be representative of California driving for all applications (e.g., buses, long-haul trucks) the engine families are used for. The plan would detail all applications that employ the engines, the number of engines per application group that would be tested and the method in which the data would be collected. Discussing the plan with ARB would allow each manufacturer to identify the most cost-effective way to obtain the data. Some manufacturers may find it easiest to collect data from vehicles that come in to its authorized repair facilities for routine maintenance or warranty work during the time period required, while others may find it more advantageous to hire a contractor to collect the data. Further, upon good cause, the Executive Officer may extend the six-month time period for the collection of data to cover situations where manufacturers have difficulty in gathering the required data within the six-month time period.

As stated before, the data collected under this program are primarily intended to provide an early indication that the systems are working as intended in the field, to provide information to "fine-tune" the proposed requirements for tracking the performance of monitors, and to provide data to be used to develop a more appropriate minimum ratio for future regulatory revisions. The data are not intended to substitute for testing that would be performed by ARB under the future heavy-duty OBD-specific enforcement regulation to determine if a manufacturer is complying with the minimum acceptable performance levels established in the OBD regulation. In fact, the data collected would not likely meet all the required elements for testing by ARB to make an official determination that the system is noncompliant.

XI. DEFICIENCIES

As discussed in the introduction, the proposed OBD regulation would require monitoring of virtually all components and systems that can affect vehicle emissions. Most components and systems would be monitored for more than one type of failure. Therefore, OBD systems would contain many diagnostic algorithms. During the early stages of OBD implementation for light- and medium-duty vehicles, some manufacturers encountered unforeseen and generally last minute problems with some monitoring strategies despite a good faith effort to comply with the requirements in full. The staff anticipates the same problems to occur during heavy-duty OBD implementation.

Thus, like the light- and medium-duty OBD regulation, the staff is proposing a provision that would permit certification of heavy-duty OBD systems with “deficiencies” in cases where a good faith effort to fully comply has been demonstrated. Specifically, in granting deficiencies, the Executive Officer would consider the following factors: the extent to which the proposed requirements of the OBD regulation are satisfied overall based on the application review, the relative performance of the resultant OBD system compared to systems fully compliant with the proposed requirements of the OBD regulation, and a demonstrated good-faith effort on the part of the manufacturer to: (1) meet the proposed requirements in full by evaluating and considering the best available monitoring technology; and (2) come into compliance as expeditiously as possible.

The deficiency provisions would facilitate OBD implementation by mitigating the danger of manufacturers not being able to certify vehicles with relatively minor implementation problems. However, to prevent misuse of the provision and ensure equity for manufacturers able to meet the proposed requirements in full, the staff is proposing that manufacturers be subject to fines for deficiencies in excess of two for a particular model. The fines would be in the amount of $25 or $50 per deficiency per vehicle depending on the significance of the monitoring strategy in question. Given the leadtimes proposed for the monitoring requirements and the experience of light- and medium-duty OBD compliance, the staff is anticipating very few vehicles that would be subject to fines.