Bay Area Rapid Transit District
Advance Automated Train Control System
Case Study Description

By Victor Winter (2615), Raymond Berg (2615) and Jim Ringland (8112)
All from Sandia National Laboratories[1]

Objective

This document contains an informal description of a portion of the Advanced Automatic Train Control (AATC) system being developed for the Bay Area Rapid Transit (BART) system. BART provides commuter rail service for part of California's San Francisco bay area. Specifically, the informal specification given below focuses on those aspects of BART that are necessary to control the speed and acceleration for the trains in the system. Other aspects of BART control such as (1) communication error recovery, (2) routing (via switches) and (3) right-of-way signaling (via "gates") are largely ignored. The scope of this case study is narrower than the AATC project as a whole, but within this narrowed scope, enough detail has been supplied to give a sense of the level of complexity involved.

The overall objective of this case study is to construct a system (software or otherwise) within the infrastructure given, that can control the speed and acceleration of trains in the system subject to the various constraints that are described in the specification. In particular, it is not the purpose of the case study to criticize the infrastructure.

You are asked to present your research in the context of this case study. Specifically, how would your work positively impact the construction of the speed and acceleration control system? Also, since the purpose of this case study is to provide a context for presenting your research, it is perfectly acceptable to make simplifications when necessary. And finally, the informal specification given will no doubt contain ambiguities. Again, given that this is an academic exercise, feel free to resolve any ambiguities in whatever manner seems most reasonable to you.

General Background on the BART Train System

It is not assumed that those participating in this case study have an extensive knowledge of train systems and terminology. This section gives a general overview the BART train system.

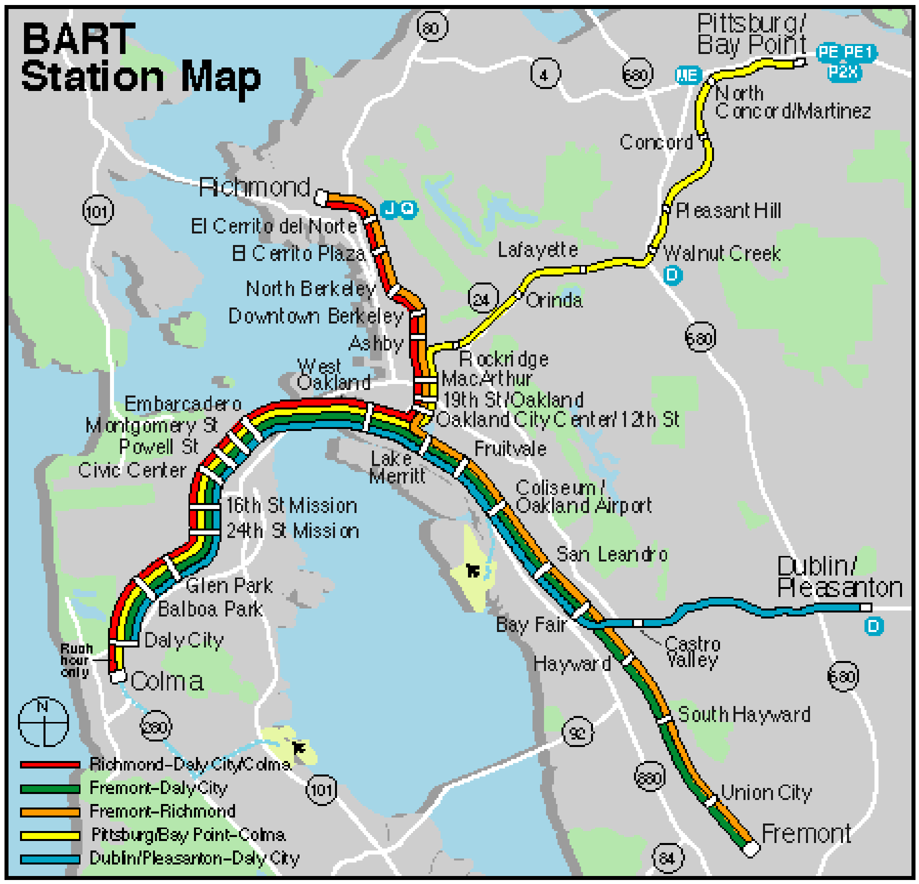
BART provides heavy commuter rail service in the San Francisco Bay Area. On a typical work day it serves around 250,000 passengers. During commute hours over 50 trains, most consisting of 10 cars, will be in service. Cars are driven by electric motors powered by a 1000 VDC, "third rail." Cars use both regenerative and friction brakes. The system is controlled automatically. On-board operators have a limited role in normal operations. Operators signal the system when

[1] This work was supported by the United States Department of Energy under Contract DE-AC04-94AL85000. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy.

the platforms are clear so a train can depart a station. More importantly, operators trouble-shoot problems, and can operate trains manually (at low speeds) when there is a problem.

The system operates most of the day, but there is some maintenance time available at night. Trains run on the BART system from 4:00 AM, when the first trains leave the yards to position themselves for the day's service, until about 1:30 AM, when the last trains return to the yards. On Sunday mornings, trains leave later: around 7:00 AM.

With a few minor exceptions, the BART system consists of double track: one track going one direction and one track going the other. The trains go from a starting point to an ending point (i.e., the track is not a loop.) As discussed later, there is an AATC controller at the front and back of each train (consist). At the end of the line, the front and back controllers are redefined and the train goes in the other direction.



As the map shows, the lines feed from points north, south and east into San Francisco, with the critical link being the section from Oakland to points west where four lines share one pair of tracks. To serve more passengers, BART needs to utilize this section more efficiently. Adding new tracks to this segment — in a tube under the bay and underground in the heart of San Francisco — would be prohibitively expensive. BART needs to run trains more closely spaced. AATC will allow them to do that.

A track is partitioned into *track segments*. Track segments may be bounded by gates. A gate can be viewed as a traffic light of sorts, where a *closed gate* corresponds to a red light and an *open gate* corresponds to a green light. Gates establish the right-of-way where tracks join or merge at switches. The purpose of gates is to keep trains from passing through switches before they are appropriately positioned. Gates not associated with switches can be used to control traffic flow.

Informal Specification for the AATC System

The AATC system replaces part but not all of BART's current control system. AATC consists of new station computers, a radio communications network that links the stations with the trains (communications currently are through the running rails, with very low bandwidth), and software modifications to the front and back controllers on-board the trains. The communications system provides ranging information (from wayside radios to train radios and back) that allows the system to track train positions. On the trains, the control computer is located in the lead car. This computer controls the operations of the brakes and motors on all the cars in the consist.

Most of the control computation is done at the stations; the trains respond to speed and acceleration commands. Each station controls trains only in its immediate area. A station must thus communicate with its neighbors to receive and hand-off trains. BART has asked its contractors to size the system to be able to handle 20 trains in each station's control zone. This would correspond to an extreme back-up situation. In locations where stations are closely spaced, such as in downtown Oakland or San Francisco, one computer may manage a zone that spans several stations.

This case study will concentrate on one of the most critical functions of the new station computers – calculation of the speed and acceleration commands that are sent to the trains. For this study we will assume that the communications link, the on-board train control system, and station computer functions other than speed and acceleration selection work as intended. The latter include conversion of ranging data into train position estimates, managing entry and exit of trains into the system, and hand-offs between stations.

The other major aspect of train control is *interlocking* – the management of track switches and associated signals to enter or not enter whole blocks of track. (In the 19-th century, track switches and flags to signal train operators were mechanically interlocked. The name has stuck.) At BART interlocking is handled by another system that is not being replaced as part of the AATC development. The AATC system will simply see "go" or "stop" indicators at various track locations in the system. A train should enter such a location (a "gate") only if allowed. Otherwise it should stop before the gate if it can. (It is the responsibility of the interlocking system not to move a switch if a signal change occurs when it is too late for an approaching train to stop. The system interlocking may, however, "close" a gate at any time as a signal to following trains.) Stopping at a station platform is much like stopping in front of a gate, although there are some additional controls (not discussed here) to assure the final stop is at precisely (within about a foot) of the desired location.

The responsibility of the speed and acceleration selection process is to get trains from one point to another as fast and smoothly as possible, subject to various constraints. These constraints include:

- A train should not enter a closed gate.
- A train should never get so close to a train in front that if the train in front stopped suddenly (e.g., derailed) the (following) train would hit it.

- A train should stay below the maximum speed that segment of track can handle.

This is a simplification – this case study is omitting many details with which system designers must deal.

The AATC system operates on 1/2-second cycles. Each half-second the station computer receives ranging and - speed information (derived from tachometer data) from the trains. It uses that information to compute an uncertainty envelope for the location of each train (mean and standard deviation). For this study, assume this function works as intended and provides fully reliable inputs to the speed/acceleration selection problem. This information, along with track signal and track layout information, is used to compute speed and acceleration commands.

Commands are time stamped and become invalid 2 seconds after the identified time. This time stamping (or time tagging) is done by what is called a Message Origination Time Tag (MOTT). When a train sends performance data back to the station, it attaches the time that it sends (originates) the message. When that information is used to update the position estimate, the MOTT is associated with that position estimate. The time stamp provides a measure of the currency of a position estimate. When that position estimate is used to compute a speed/acceleration command, that MOTT is attached to the command. The train then checks the MOTT before exercising a command. A train will continue to exercise that command until a new one arrives or until that command expires, 2 seconds after the originating time.

If the train does not have a currently valid command, it goes into maximum braking. The control algorithms thus have to be designed so that if all communications are lost, then when commands expire and trains come to a stop, no safety violations will have occurred. That is, the stopping location of any train after lost communications, has timed out, and has come to a stop will be (1) behind any closed gates and (2) behind the rear end of any trains ahead. This sequence of events, along with a very pessimistic definition of how long it physically takes to stop a train once braking begins, defines what is called the "Worst Case Stopping Profile." Similarly, the control algorithms have to be designed so that whatever happens, the train must not exceed any track speed limits.

The station computer itself is divided into two systems. First is the Non-Vital Station Computer (NVSC). ("Vital" in the context of railroad safety can be interpreted to mean that the function is critical to the safety of the system.) The NVSC is a fast, flexible computer. The hardware is reliable enough to meet performance-driven Mean Time Between Failure (MTBF) and Mean Time Between System Shutdown (MTBSSD) goals but not safety-related Mean Time Between Hazard (MTBH) goals. These goals are defined precisely below. Second is the Vital Station Computer (VSC), which is slower but reliable enough to meet safety-related MTBH goals.

The NVSC proposes speed and acceleration commands, considering both safety and performance goals. The VSC checks that they do not exceed maximum bounds for safety. For smooth operations, the NVSC generally will have to propose speeds and accelerations lower than the absolute maximums against which the VSC checks. For this case study, you may partition functions among the two machines as you wish, subject to meeting overall safety goals, the fact that the NVSC hardware alone cannot support these goals, and the VSC cannot do more than essentially one computation of a bounding speed and acceleration.

(Note. This description of the NVSC and VSC does not precisely represent the actual interaction between the two machines being designed for the AATC system.   The discussion above captures the basic concepts, but omits some real [and proprietary] system design complexities.)

The following sections provide some additional details about the control system.

Inputs and Outputs to the Control Algorithm

The control algorithm receives the following information, updated every 1/2 second.

- *The outputs of the position algorithm -- mean and standard deviation on both position and velocity* – for all trains in the area. Nominal performance for the system is for the standard deviation to be 2 to 3 feet. Note that both the mean and standard deviation will vary a little (e.g., a few feet) from each ½ -second-time step to the next.
- *The Message Origination Time Tag (MOTT).*   This is the time a given train sent its most recent report.   Information from this report is folded into the position tracking system. Thus the MOTT is a measure of information currency.   After receiving a MOTT from a train, the control system then attaches the same MOTT to acceleration and velocity commands that are then sent back out to the train, and the commands are only valid until MOTT + 2 seconds.
- *Gate information (open, closed) from the Interlocking system.*
- *Any special speed restrictions on either the whole system or individual track segments.*

It is the responsibility of the control system to separate trains sufficiently so that if a train in front instantaneously stopped (e.g. it derailed) then the following train could stop without hitting it (assuming worst-case conditions). Additionally, if a "leading train" (one in front of the train for which a speed/acceleration command) stops sending reports back to the station, the algorithm has to assume it is in the location associated with its last report (e.g., the train could have just derailed). And finally, it is the responsibility of the control system to prevent a train from entering a closed gate if it can. As noted earlier, it is the responsibility of the interlocking system not to move a switch if a signal change occurs when it is too late for an approaching train to stop. The system interlocking may, however, "close" a gate at any time as a signal to following trains.)

The following static data is also available for segments of the track. *Segment location* (end points and length)

- *Grade* (less than 4% system-wide). Each track segment has only one defined grade. Grades can either be constant or be part of a parabolic change from one grade to another
- *Maximum allowable speed*. (Maximum speeds typically range from 36 to 80 mph, with some lower numbers on crossovers and siding.) It is the responsibility of the

control systems to slow trains down before entering segments with low allowable speeds.)

- *Locations of gates* (at ends of some segments)

The table at the end of this informal specification gives a sample of the sort of information in the track database. This sample covers part of western San Francisco, from the Daly City station to the 24-th Street Mission stations. Between these are the Balboa Park and Glen Park stations. (See the map earlier in this write-up.) In this particular example, one station computer controls the entire area from a point about half way between Daly City and Balboa Part to a point about half way between Glen Park and 24-th Street. That is, one station computer is actually controlling the area around two stations.

The control algorithm then must produce for each train, a command message. All processing for all trains must be complete at the end of the 1/2-second cycle. That message contains the following safety critical data:

- *ID of the train* for which the command is intended
- *Commanded speed (*between 0 and 80 mph)
- *Commanded acceleration (*-2 to -0.45 mphps in (closed loop) braking, 0 to 3 mphps in propulsion)
- *Message Origination Time Tag*
- *A fixed four bit code* identifying this as a command message

Some non-safety-critical information is also sent.

<u>Physical Performance of the Train in Response to Commands</u>

When given a speed and acceleration command, the train will accelerate to 2 mph under the commanded speed at the indicated acceleration. Maximum propulsion rates depend on speed. At low speeds, up to 31 mph, accelerations as high as 3 miles per hour per second (mphps) on flat track are possible. At higher speeds, only lesser accelerations are possible. In practice, BART uses a look-up table to get exact values, but for reference, at 80 mph, accelerations of only 0.78 mph are possible. 80 mph is the maximum speed that the control system can send to the train. When the actual speed is within 7 mph of the commanded speed, the train controller will limit maximum allowable accelerations to smoothly reach a speed 2 mph below the commanded speed. (Note that the train controller is a system that is separate from the control system.) Once a speed of 2 mph below the commanded speed is achieved, the train will attempt to maintain that speed, although there may be some small fluctuations (on the order of +- 1 mph). The train can coast while in propulsion mode. Rolling friction and wind resistance can then reduce train speed, except perhaps on down-grades.

The intent of maintaining an actual speed 2 mph below the commanded speed is to ensure, given these fluctuations, that the actual speed not exceed the commanded speed. The train will initiate braking to prevent speeds in excess of the commanded speed, as might be the case on a downgrade. As will be discussed below, it takes a little time for the train to move from propulsion mode into braking mode. Doing so frequently results in a jerky ride for the passengers and excess wear on the equipment. Seeking to maintain a speed 2 mph below the command speed limits but allowing for small fluctuations minimizes mode changes.

Trains have two different braking modes. In *open-loop* (maximum) braking the train simply attempts to brake at its maximum rate, 3 mphps. The trains are designed to revert to open-loop braking unless a DC signal is continuously applied to inhibit it. Thus, in the case of control failure, the trains should stop safely. In *closed-loop* braking, less than maximum rates can be applied and those rates can be continuously varied.

The maximum brake rate may not be achievable in practice. If wheel slippage is detected, the controls to that set of wheels reduces the braking force so that wheel adhesion can be regained. Wheel adhesion and slippage limits, not the capabilities of the braking system, often determine how fast a train can stop. The worst case profile to be discussed below assumes only decelerations of 1.2 mphps on wet track and 1.6 mphps on dry track, although these are very pessimistic bounds.

(Adhesion can also limit acceleration performance, although this is not critical for safety. Note also that when the wheels are slipping, tachometer data will not be representative of real speeds.

The minimum brake rate is 0.45 mphps.

There are delays in moving from propulsion into braking and in changing brake rates. It takes 1.5 seconds to move from maximum propulsion to coasting and 1 second to reconfigure while coasting from propulsion mode to braking mode (the same timings are true when going in the reverse direction). Brake rates (d^2x/dt^2) then ramp up with a jerk rate (d^3x/dt^3) of –1.5 miles per hour per second per second.

Likewise, there are delays in changing brake rates or moving from braking into propulsion. Again brake rates can be changed subject to a jerk rate of 1.5 mphpsps. There is a 1 second mode change. Propulsion acceleration can change with jerk rates of 2 mphpsps.

Grades will change actual accelerations. For the purposes here, this case study will assume that all potential energy is converted to kinetic energy (or vice versa) on grades. This will affect the acceleration limits due to motor or brake performance.

Worst Case Stopping Profile

One fundamental safety requirement is that train speeds and accelerations be selected so that (1) the train does not hit a train in front of it or (2) enter a closed gate. These safety requirements must be satisfied even in the case of very poor stopping conditions.

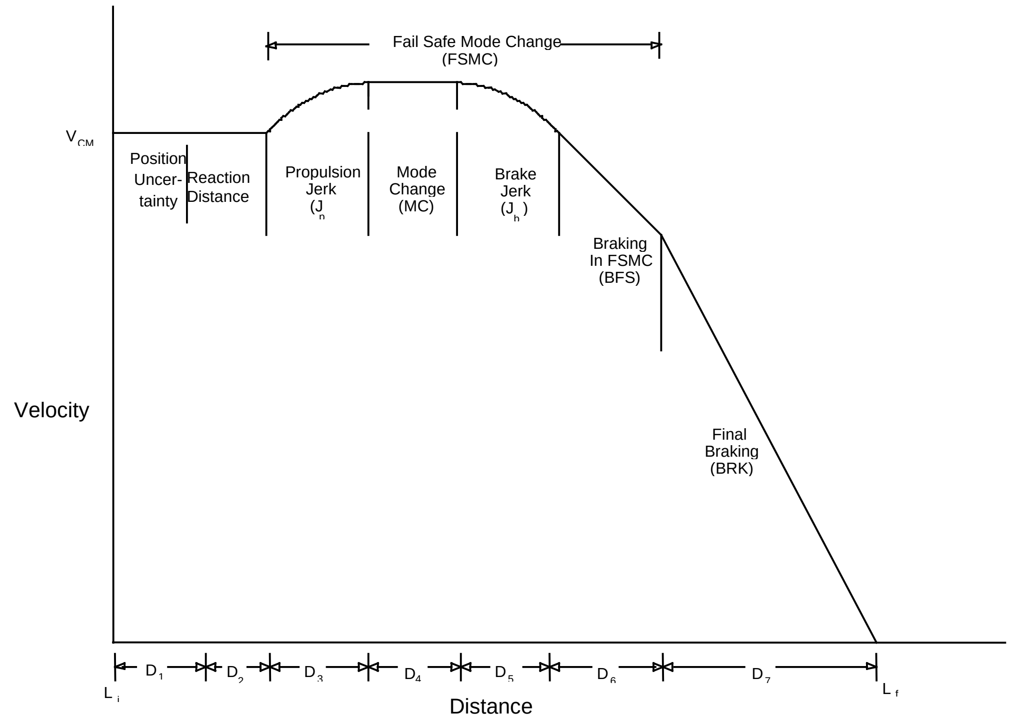
A profile defining a "worst-case" stopping distance of trains has been agreed upon. While this profile does not represent the absolute limit of what could happen, it does represent a very conservative bound. The use of this profile as a worst-case bound was approved by the California Public Utilities Commission Interim Order of June 3, 1980. Thus, performance of a train that is worse than this profile – which is theoretically possible but extremely unlikely -- represents a publicly agreed upon, accepted risk.

Train stopping performance is limited by the time it takes to move into full braking, and then, once in braking, the adhesion of the train wheels to the track.

The Worst Case Model assumes the velocity distance profile shown in the figure on page 9. This figure assumes the train is configured for propulsion or speed maintaining and that the track grade is constant throughout the stop. The stopping distance is the sum of D_1 through D_7 as shown at the bottom of the figure. The length of each segment is based on worst case assumptions.

The figure on page 9 starts with a train speed maintaining at the commanded speed (V_{CM}) with a command that is time-tagged at the time when the train is located at position L_i. The control algorithms cannot know exactly where L_i is since train position is reported only with an estimated mean and standard deviation. The first segment of the worst case profile, D_1 accounts for this uncertainty. That command is valid for two seconds after that tagged time. If the train receives no further commands, then after two seconds, the train will enter maximum braking. If a new command is received, it will act on that command. If the new command calls for braking, then the train will initiate the braking sequence at that time. Thus, the latest response to an obstacle occurs if the train waits two seconds and then "times out". D_2 is the distance the train would travel during those two seconds.

When the train reacts to the need to stop at the beginning of D_3, the agreed-upon worst case model assumes that the train suddenly is in full propulsion (a non-physical bounding assumption). The train then has to stop propulsion (D_3 - it takes 1.5 seconds to move from full propulsion to no propulsion), reconfigure for braking (D_4 - it takes 1 seconds to reconfigure), and then apply the brakes (D_5 – it takes up to 1.5 seconds to apply the brakes, although the actual time depends on the brake rate). Because of details in the operation of the train brakes, the worst-case model assumes that initially a train has only 60 percent braking capability (e.g., the brakes are not responding in 40% of the cars in the train). With respect to the non-braking cars (i.e., 40% initially) it is assumed that half of these cars (20% of the total train) will recover 8.5 seconds after the beginning of D_3. (This accounts for the possibility that some cars may not respond to the command to reconfigure for braking, but will successfully go through a back-up "fail-safe mode change." This will force the cars in question into braking mode, but there is some additional delay. The worst case model assumes that the other half of the malfunctioning cars (20% of the total train) never contribute to braking. (This worst-case bound covers total brake failures, locked up (skidding) wheels, and transient brake drop-outs.)



Worst Case Model – Train accelerating or maintaining speed
Constant Grade

- **D_1:** This distance accounts for the uncertainty (e.g., 6 sigma) of the train position from the perspective of the control system.

- **D_2:** In the case where no command is received (e.g., a time out = 2 seconds maximum), D2 denotes the worst case distance that could be traveled before a time out.

- **D_3:** This is the distance that the train would travel while the train switches from full acceleration mode to coasting mode. At this point it is assumed that the train (suddenly) is in full acceleration mode. This is a non-physical worst case bound.

- **D_4:** This is the distance that is traveled while the train switches from coasting in propulsion mode to braking mode.

- **D_5:** The distance traveled before the brakes fully ramp on.

- **D_6:** This is the distance that is traveled while 60% of the cars are braking and 40% are not. 8.5 seconds after the beginning of D_3, the train achieves 80% percent braking. The model assumes 20% of the cars go through a slower fail-safe mode change to reconfigure for braking.

- **D_7:** At this point we have 80% of the cars braking normally and the remaining 20% of the cars do not contribute to the braking. D_7 represents the distance required for the train to come to a complete stop.

The equations for calculating D_1 through D_7 are shown in the shaded tables on the next pages.

An item of interest is that distance D_3 is non-linear with velocity because the acceleration rate A_p decreases with increasing velocity. (The actual values involve a lookup table, but the values range from 3 mphps at train speeds under 31 mph to .78 mphps at 80 mph. However, the jerk time does not change.

The design brake rate in D_7, depends on whether the track is exposed (open to the elements) or covered (in tunnels). Exposed track may be wetter. The wheels may begin to slip at lower speeds on exposed track than on covered track. About 1/3 of the BART system is underground.

In practice, trains may be commanded to decelerate at less than the maximum rate. The worst-case stopping profile once a train is in braking (but at less than full rate) omits intervals D_3-D_6. Interval D_1 applies unchanged. Interval D_2 applies, but one can assume the train is decelerating at the minimum brake rate of 0.45 mphps. (For the worst-case calculation, we assume pessimistically that only minimum braking is applied.)

In handling stops of varying grades, BART has developed some rather lengthy-to-describe heuristics. The physical facts are relatively simple. Track grades can be represented either as being of constant grade : x feet at one grade, y feet at another grade, or as part of a parabolic transition from one grade to another. On grades, potential energy is exchanged for kinetic energy: $mgh = 1/2\ mv^2$. The tables give the formulas appropriate for constant grades. One cannot, however, just easily subdivide the stopping profile into smaller segments where the grade is constant. Trains are not point masses. A full train is 710 feet long. The potential energy of a train is constantly varying. (For the purposes of the case study, you may either assume constant grades, accept that a reasonable approximation exists, or work out and solve the actual, moderately complicated, differential equations. However, BART has found that their heuristics, for the actual track at hand, are fully adequate.)

For these worst-case safety-related calculations, the effects of rolling friction and wind resistance (which are real) are not included.

WORST CASE TRAIN CALCULATIONS

$D_1 = NOSE + (PUF \times PU)$

$D_2 = V_{CM} \times AD$

$J_P = A_P/T_{JP}$

$D_3 = V_{CM} \times T_{JP} + 1/2\ A_P \times T_{JP}^2 + 1/6\ J_P \times T_{JP}^3 + 1/2\ A \times T_{JP}^2$

$V_3 = V_{CM} + A_P \times T_{JP} + 1/2\ J_P \times T_{JP}^2 + A \times T_{JP}$

$D_4 = V_3 \times MC + 1/2\ A \times MC^2$

$V_4 = V_3 + A \times MC$

$Q_{FSMC} = (NCAR - NFAIL - NFSMC)/NCAR$

$T_{JB} = BRK/J_B$

$D_5 = V_4 \times T_{JB} + 1/6\ J_B \times Q_{FSMC} \times T_{JB}^3 + 1/2\ A \times T_{JB}^2$

$V_5 = V_4 + 1/2\ J_B \times Q_{FSMC} \times T_{JB}^2 + A \times T_{JB}$

$T_6 = FSMC - T_{JP} - MC - T_{JB}$

$BFS = BRK \times Q_{FSMC}$

$D_6 = V_5 \times T_6 + 1/2\ BFS \times T_6^2 + 1/2\ A \times T_6^2$

$V_6 = V_5 + BFS \times T_6 + A \times T_6$

$Q = (NCAR - NFAIL)/NCAR$

$D_7 = (V_F^2 - V_6^2) / 2\ (BRK \times Q + A)$

$$D_{WC} = \sum_{i=1}^{7} D_i$$

WORST CASE TRAIN PARAMETERS

NOSE - Estimated train nose location (ft) (from Update Position Algorithm)

PUF - Uncertainty Factor = 6

PU – Position Uncertainty reported as one standard deviation (ft)

V_{CM} - Commanded Speed (mph)

AD – AATC Delay = 2 seconds

J_P - Jerk Limit in Propulsion (Function of Speed and Grade) (mphps ps)

A_P - Acceleration in Propulsion (Function of Speed and Grade) = mphps

T_{JP} - Jerk Time in Propulsion = 1.5 Seconds

A - Acceleration Due to Grade = 21.9 mphps * $\frac{\text{(Grade in \%)}}{100}$

MC - Mode Change = 1 Second

NCAR - Number of Cars in Consist = 10 Cars

NFAIL - Number of Failed Cars = 2 Cars

NFSMC - Number of Cars In FSMC = 2 Cars

J_B - Jerk Limit in Braking = -1.5 mphps ps

BRK - Design Brake Rate = -1.5 mphps for exposed track;
-2.0 mphps for covered track

FSMC - Fail Safe Mode Change Time = 8.5 Seconds

Considerations with Acceleration and Speed Commands

As noted in the introduction, actual speeds and accelerations must be consistent with worst-case safety bounds and must never allow a train to exceed an allowable maximum speed. Once these are met, there are secondary performance objectives. These include the following

- Minimize the time travelling from station to station. (Schedules can be maintained by dispatching trains appropriately and holding in stations if need be.)
- Avoid large and frequent changes in speed and acceleration. This maximizes passenger comfort, provides minimal stress on equipment, and minimizes power usage.
- Avoid mode changes. (Note coasting can be done in propulsion mode.)

There are two concerns when accelerating a train from a lower speed to a higher speed in an area where trains are close to one another.

- Spacing trains farther apart than necessary because the commanded velocity (which is used in the worst case stopping distance calculation) is much greater than the actual velocity.
- Accelerating a train so fast that it nears infringing on the worst case stopping distance and must go into braking.

Regarding the first, note that worst case stopping distances are calculated using the commanded speed, not the actual speed, so commanded speeds should not be too much greater than the actual speed. Conversely, since the train automatically limits accelerations when within 7 mph of the commanded speed, the commanded speeds should not be too close to estimated speeds.

Regarding the second, when a train is accelerating behind another train which is also accelerating and is located close to the worst case stopping distance of the following train, the speed of the following train will tend to oscillate. The oscillation occurs because the stopping distance increases with the square of the velocity while the velocity of the following train increases linearly with acceleration. The following train will accelerate to nearly within the worst case stopping distance, brake, and then repeat the cycle. This is undesirable. In order to avoid this oscillatory behavior, it is necessary to maintain a separation distance between the two trains that is greater than the worst case stopping distance of the rear train throughout the acceleration profile and to choose acceleration rates judiciously.

Uncertainties in position estimates can lead to some irregularities in the worst case stopping distance from one1/2-second cycle to the next. If the position uncertainty envelope grows a little bit or if the best location estimate varies a little bit within the uncertainty bound, a train that was a little bit outside the worst case stopping distance of an obstacle could suddenly (and artificially) find itself within the worst case stopping distance. The appropriate response at this point is to go into braking. And then perhaps at the next cycle the train may find itself a little bit outside the worst case distance, so braking is no longer needed. This too produces an undesirable oscillation.

The acceleration from gravity on a non-zero grade adds to the acceleration / deceleration rates. The vehicle controllers use data from accelerometers as feedback when controlling train acceleration. If the train is accelerating or decelerating at a constant rate due to gravity, the accelerometers will not sense that acceleration. They will sense only additional acceleration supplied by brakes or motors. Thus, if the train is asked to accelerate at 1 mphps on a downgrade, it will actually accelerate at 1 mphps + (Grade in % /100)*(21.9 mphps [Note 21.9 mphps = 32.1 fpsps = g = gravitational acceleration.). The train does not know the grade it is on. Any corrections for this effect have to be done by the station computer.

Similarly, velocity commands to a following train should not lead to oscillatory behavior. In a simple case with a following train within the stopping distance plus a some margin of another train which is speed maintaining, the following train should not be given a velocity command that is much above the speed of the lead train as the margin narrows.

Quantitative Quality and Safety Metrics to be Demonstrated

In addition to the system performance issues discussed above, there are several qualitative quality and safety metrics to be demonstrated. The metrics are to be computed system-wide, assuming 26 control stations, 72 miles of double track, and 160 AATC-equipped vehicles (80 trains in service).

The *Mean Time Between Hazard* must be more than 1,000,000,000 hours (i.e., 10^9 hours) system-wide. For this case study, a "hazard" is one of the these three conditions:

- A train being physically within the worst-case bounds of any obstacle such as a closed gate or the rear of a leading train. (Note this condition is stated in terms of physical reality, not in terms of the estimates and uncertainties available to the control system.) This cannot be directly measured and must therefore be bounded (e.g., the probability of a train being more than 6 sigma ahead of the nominal position is on the order of 10^{-9}.)
- A train operating a speed higher than either the allowable maximum on a track segment. (Again, this is stated in terms of physical speeds, not estimates from the train tracking system or readings from tachometers.)
- A train operating at a speed higher than an externally commanded reduced speed.

The actual contract has a few other conditions associated with other elements of the system (e.g. problems with the locating and tracking system). Also, in the actual BART contract, this MTBH is limited to hardware faults. For this case study, we invite comment on how such a goal might be demonstrated for the design and implementation of the control algorithms.

The *Mean Time Between System Shut Down* should be more than 26,000 hours system-wide. A "system shutdown" is any hardware failure or software fault that prevents the AATC from sending a speed command to the train control system on one or more operational control cars for a period of five seconds or more.

The *Mean Time Between Failure* for a control station AATC must be more than 2000 hours. A "failure" is any event that impairs subsystem performance or requires manual intervention. This is primarily being used to provide a performance metric for hardware failures, including ones that involve no loss of functionality due to redundancy, that require repair.

The *Service Reliability Requirement* states that system wide all first failures are to be repaired within 24 hours. Total service recovery time, which includes the time for maintenance crews to reach the equipment plus the mean-time-to-restore-service, shall average 12 hours.

The *Cold Start Boot-Up Time* for a control station AATC, as would be required after a system shutdown shall be no more than seven minutes.

Vital Station Computer (VSC) Issues

As noted in the introduction, the vital station computer system provides for a high degree of internal checking that drives any computational error rates down to safety-acceptable levels. Moreover, elements of system design, implementation, and testing assure acceptable robustness against operating system and compiler faults. For the purpose of this case study, assume this system performs as desired. This specially-designed system is, however, relatively slow. (In the actual application, this machine is programmed in a relatively restricted subset of C++. We will not apply any particular language limitations in this case study.)

Note that if VSC does flag a non-vital speed as unsafe, the command is not sent. If no commands reach the train for two seconds, there is a time out. This does not count against the MTBF number , but if the outage continues for five seconds it counts against the MTBSSD number. If the algorithm does not flag a non-vital unsafe speed as unsafe, it is a safety failure and if the trains actually enter a worst-case stopping distance of each other or of a closed gate, it counts against the MTBH number.

You may assume the following. Safety critical tables of train parameters used by the train control algorithm are always available, i.e., they are stored with enough redundancy to meet all safety and reliability requirements.

Miscellaneous Questions and Answers

- Q: Can a train go backwards?
  A: No.

- Q: Does the control algorithm need to consider the case of a runaway train?
  A: No. This is managed by the controls built into the train.

- Q: Is a specific track always "one-way"?
  A: Yes, for the purposes of this case study. In actual operations there are special protocols (reduced speeds, special routing rules) for operating a train in the reverse direction on a specific track.

- Q: Does the speed and acceleration control system need to consider routing of trains?
  A: No. Routing is done by the interlocking system and other external systems.

- Q: Is it true that, in general, a train must begin deceleration more than one segment prior to a closed gate?
  A: Yes. The worst case is spectacular. On a continuous 4% downgrade, a train at 80 mph would take over 3 miles to stop.

- Q: Although a train can be in either a propulsion or a braking mode, can it also be in a neutral mode?
  A: No, neutral, or coasting, is done in propulsion mode (with acceleration = 0). When in braking, the brake rate is at least -0.45 mphps.

- Q: Are all trains the same length?
  A: No, but for the purposes of this study, you can assume so. Nominally, during rush hours, trains are made up of 10 cars, giving the total length of 710 feet introduced earlier. At off-peak times, shorter trains are used. And even at rush hours, some lines may have shorter trains.

- Q: Why are gates closed/opened?
  A: Gates are like stoplights on a highway and they serve a similar purpose. A stoplight prevents traffic from entering an intersection where the traffic is going in the other direction. The gates define who has the right of way coming up to a switch where two tracks join or diverge. A gate prevents a train from entering a switch when it is set for traffic in another direction. A gate (and a stoplight) can also control or limit flow.

- Q: If a gate wants/needs to be closed and the position of a train prevents this, why is this acceptable?

A: This gets at the logic of train routing, with is beyond the scope of this case study, but some background might be useful. Gates being opened and closed essentially define who has right of way when tracks come together. The situation where a gate wants/needs to be opened or closed occurs when different trains want/need the switches positioned differently. The interlocking system is set up so that "wants" can be handled as expeditiously as possible, but trains that do not have the right of way may have to wait. The speed selection system has to make sure that a train moves into a position where it "needs" to have a gate in front of it changed. The interlocking system has to make sure that it never changes a switch so that a train is put at risk. (If the right-of-way is going to be changed, it may be the case that all the gates are closed -- the stoplights are red in all directions -- as a train finishes passing through a switch.)

What actually happens is that as a train approaches a gate that is not open (but before it has to slow down), a separate system sends a "request" to the interlocking system asking for changes in switches and gate positions (i.e., that ask for changes in who has the right-of-way). A request is checked against the positions of all other trains near the switch. Acting on a requested action may occur in stages. All gates approaching the switch in question may be closed while a train is crossing a switch or is too close to stop. No switches will be moved until all trains are clear of the area. And only after a switch is physically moved will the appropriate gates be opened. When a routing request is not granted, a train that does not have the right-of-way will have to either slow down or stop until the gate opens. It is a functionality issue, but not a safety issue, to make sure that train schedules do not lead to too many conflicting requests and that routing requests come in early enough that they can be granted without having the trains slow down excessively.

- Q. Why are we worried about going faster than the maximum speed on a piece of track?
  A. The train may derail.

- Q. Why are we worried about entering a closed gate?
  A: If the gate is in front of an incorrectly aligned switch, the train may derail. If the switch directs the train onto an inappropriate track, there may be collision.

- Q. Why can't the train accelerometers detect the component of train acceleration due to gravity?
  A. An accelerometer is, in gross simplification, a mass at the end of a spring. Acceleration is measured by how much the spring is compressed. When the motors accelerate the train, the only way the mass gets accelerated is for the spring to push on it, so the spring is compressed. When the gravity accelerates the train, it also accelerates the mass on the spring. There is no difference in the forces applied at the two ends of the spring, so no change (no differential acceleration) is measured.

Sample Track Layout Information

Segment						
Begin	End	Length	Comment	Civil Speed	Grade	Exposure
(feet)	(feet)	(feet)		(mph)		
5940	6640	700	Daly City Station Platform	36	-0.80%	Open
6640	6741	101		36	-0.80%	Open
6741			Gate			
6741	7588	847	Switches to Crossover and Spur	36	-0.80%	Open
7588			Gate			
7588	8522	934		36	-0.80%	Open
8522	8722	200	Parabolic grade transition. Midpoint at 8622	80	transition	Open
8722	9003	281		80	2.75%	Open
9003			Control Station Boundary			
9003	9509	506		50	2.75%	Open
9509	11355	1846	Parabolic grade transition. Midpoint at 10428	50	transition	Open
11355	11900	545		50	-3.50%	Open
11900			Gate			
11900	12300	400	Switches to Crossover	80	-3.50%	Open
12300	12369	69		80	transition	Open
12369			Gate			
12969	13100	131	Parabolic grade transition. Midpoint at 12800	80	transition	Open
13100	13400	300		80	-2.11%	Open
13400	14190	790	Parabolic grade transition. Midpoint at 13950	80	transition	Open
14190			Tunnel Portal			
14190	14500	310	Parabolic grade transition (continued)	70	transition	Tunnel
14500	14850	350		70	-3.19%	Tunnel
14850	15150	300	Parabolic grade transition. Midpoint at 15000	70	transition	Tunnel
15150	15410	260		70	-1.00%	Tunnel
15410	16110	700	Balboa Park Station Platform	36	-1.00%	Tunnel
16110	16140	30		36	-1.00%	Tunnel
16140	16500	360		50	-1.00%	Tunnel
16500	17000	500	Parabolic grade transition. Midpoint at 16750	50	transition	Tunnel
17000	17025	25		50	2.11%	Tunnel
17025			Tunnel Portal			
17000	17250	250		50	2.11%	Open
17250	18075	825	Parabolic grade transition. Midpoint at 17650	50	transition	Open
18075	18218	143		50	-3.89%	Open
18218	19245	1027		70	-3.89%	Open
19245			Tunnel Portal			
19245	19550	305		70	-3.89%	Tunnel
19550	20550	1000	Parabolic grade transition. Midpoint at 20110	70	transition	Tunnel
20550	20795	245		70	2.14%	Tunnel
20795	21295	500	Parabolic grade transition. Midpoint at 21045	70	transition	Tunnel
21295	21485	190		70	0.30%	Tunnel
21485	22185	700	Glen Park Station Platform	36	0.30%	Tunnel
22185	22420	235		80	0.30%	Tunnel
22420	22780	360	Parabolic grade transition. Midpoint at 22600	80	transition	Tunnel
22780	24950	2170		80	-0.68%	Tunnel
24950	25635	685	Parabolic grade transition. Midpoint at 25300	80	transition	Tunnel

25635	28115	2480		80	-3.12%	Tunnel
28025	28115	90		80	transition	Tunnel
28115			Control Station Boundary			
28115	28374	259	Parabolic grade transition. Midpoint in next segment	80	transition	Tunnel
28374	29025	651	Parabolic grade transition. Midpoint at 28525	63	transition	Tunnel
29025	29535	510		63	0.55%	Tunnel
29835	30065	230	Parabolic grade transition. Midpoint at 29950	63	transition	Tunnel
30065	30080	15		63	-0.30%	Tunnel
30080			Gate			
30080	32281	2201		63	-0.30%	Tunnel
32281	32981	700	24-th Street Mission Station Platform	36	-0.30%	Tunnel
Notes	The area controlled by one station computer is from location 9003 to location 28115 (total of 19107 feet = 3.6 miles)					
	Note one station computer actually will control an area spanning two train stations in this example.					
	Low numbered locations are to the west; numbers increase to the east					
	Civil speeds represent maximum allowable speed. Actual speeds can be lower.					
	On these grades, the current system operates trains at speeds much lower than the civil speeds					
	This is one of the most hilly sections of in the BART system -- much of the track is on flat ground.					
	Negative grades are down hill travelling west to east; postive grades are up hill west to east					
	For grade transitions, the plans give a mid-point, where the grade is halfway between the two grades at the ends. These three points (location and grade) define a parabola.					