GPS Integrity Architecture Opportunities

Calvin S. Miles, Federal Aviation Administration
Karl Kovach, The Aerospace Corporation
John Dobyne, Booz Allen & Hamilton
Karen Van Dyke, Department of Transportation
Marc Weiss, National Institute of Standards and Technology

BIOGRAPHIES

Calvin Miles is the GPS Team Lead in FAA Navigation Services. Calvin has over 26 years of experience developing, deploying and approving the design of NAVAIDs with integrity. Those systems include ILS-RMM, VOT, new generation RVR, TLS, three CAT III ILS designs, WAAS, and NCIME. Calvin has worked with GPS III since 1999. He graduated from Harvey Mudd College in 1987 with a degree in engineering.

Karl Kovach is a Senior Project Leader at The Aerospace Corporation in El Segundo, CA. Karl has over 34 years working on various aspects of the GPS program, including 3 years as the Air Force Officer-in-Charge of the GPS Control Segment when it was at Vandenberg AFB, CA (1983-1986). Karl received his BS degree in Mechanical Engineering from UCLA in 1978. He is a recipient of the Institute of Navigation (ION) Captain P.V.H. Weems award and the Satellite Division Johannes Kepler award.

John Dobyne is a Senior Principal Engineer at Booz, Allen, Hamilton (formerly ARINC) Engineering Services in San Diego, CA. John has over 25 years involvement with GPS at Booz, Allen, Hamilton; including support to the GPS program office, support to NATO, GPS receiver development, and avionics integration and certification. John received BS and MEE degrees in Electrical Engineering from Cornell University in 1972 and 1973.

Karen Van Dyke serves as the Director for Positioning, Navigation, and Timing (PNT) in the U.S. Department of Transportation Research and Innovative Technology Administration (RTA). Karen has been involved in navigation-related programs at the Volpe National Transportation Systems Center for over 20 years and currently is responsible for overseeing the navigation program within RTA and development of policy positions on PNT and spectrum. Karen received her BS and MS degrees in Electrical Engineering from the University of Massachusetts at Lowell. She served as the President of the ION and is a recipient of the Award for Meritorious Achievement from the Secretary of Transportation and is a Fellow of the ION.

Dr. Marc Weiss has worked at NIST (the National Institute of Standards and Technology--formerly NBS, the National Bureau of Standards) since 1979, specializing in time transfer techniques. He received the NBS Applied Research Award for a first GPS timing receiver in 1983. He was awarded a patent for the Smart Clock algorithm in 1993, which optimally locks a slave clock to a master. Dr. Weiss won the 2013 NIST William P. Slichter Award, which recognizes achievements by NIST staff in building or strengthening ties with industry. Marc founded and has led WSTS, the Workshop on Sync in Telecom Systems, annually since 1992. Dr. Weiss has also led the NIST program to support the GPS program office in developing clocks and timing systems since 1992, after participating in it since 1980. He received his Ph.D. in Mathematical-Physics from the University of Colorado in 1981.

ABSTRACT

The Global Positioning System (GPS) integrity architecture forms the basis for several GPS integrity augmentation systems, most notably, the Wide Area Augmentation System (WAAS), Receiver-based Autonomous Integrity Monitoring (RAIM) and its proposed extension the Advanced RAIM (ARAIM). This paper surveys several opportunities to increase the integrity of GPS. These opportunities are accompanied by the motivation to use the increased integrity to increase the performance of ARAIM and WAAS. Some of these opportunities already exist today, like taking credit for the GPS a priori failure rates determined by Integrity Failure Modes and Effects Analysis (IFMEA). This paper surveys several opportunities to increase the integrity of GPS. These opportunities are accompanied by the motivation to use the increased integrity to increase the performance of ARAIM and WAAS. Some of these opportunities already exist today, like taking credit for the GPS a priori failure rates determined by Integrity Failure Modes and Effects Analysis (IFMEA). Some are part of the current GPS program evolution, while others could be incorporated into the later stages of the GPS III satellites and the Next Generation Operational Control System (OCX) programs.

The current level of GPS constellation integrity performance is better than $1 \times 10^{-4}$ per hour for the probability of unalerted misleading signal information (UMSI) with the assumption that there are 10 satellites in view. This is based on the U.S. Government commitment to maintain an integrity performance level of better than or equal to $1 \times 10^{-5}$ per hour per satellite UMSI probability.
as documented in the Global Positioning System Standard Positioning Service Performance Standard. Beyond this minimum “threshold” level of integrity performance, there is also an “objective” level of integrity performance specified in requirements documents for the GPS III satellites and OCX. This objective level of integrity performance is three orders of magnitude better than the threshold level of integrity performance. If the objective level of integrity performance of $1 \times 10^{-7}$ per hour probability of UMSI assuming 10 GPS III satellites used in the position solution (i.e., probability of $1 \times 10^{-8}$ per hour per satellite, allocated between the space and control segments) were eventually achieved, then WAAS and RAIM/ARAIM would no longer be needed to provide additional integrity assurance for the aviation user. Our previous paper [1] explored this core GPS III Integrity Concept, where the objective level of integrity performance is provided entirely within the GPS system. The objective level of core GPS III integrity, however, will only come at a substantial cost in technical and programmatic difficulty and requires the full GPS constellation to be repopulated with high integrity GPS III+ satellites. This paper suggests that it may be beneficial to examine the viability of an incremental approach to improving GPS integrity to points in between the threshold and objective levels so as to gain either corresponding improvements in WAAS performance and RAIM/ARAIM capability or – perhaps more importantly – corresponding reductions in WAAS and ARAIM complexity and cost.

INTRODUCTION

Our previous paper [1] explored the core GPS III Integrity Concept, where the objective level of integrity performance is provided entirely within the GPS space and control segments – see Figure 1a. This GPS III objective integrity is equivalent to the level of integrity currently provided by GPS II augmented with WAAS (Figure 1b) and GPS II augmented with RAIM/ARAIM (Figure 1c). Not surprisingly, the objective level of core GPS integrity has a substantial cost in technical and programmatic difficulty and would require the full GPS constellation to be repopulated with high integrity GPS III+ satellites.
Rather than assuming a one-or-the-other dichotomy between the objective integrity and the threshold integrity performance for GPS III (the GPS III threshold level is the same as the current level of integrity performance for GPS II as specified in [2] and [3] and depicted in Figure 1d), it is useful to examine the impacts of an incremental approach to improving GPS integrity over time. There are certain natural breakpoints – which can be thought of as “GPS integrity architecture opportunities” – where a quantum change in GPS integrity can be realized to gain one or more of the following benefits:

- Operationally significant improvements in WAAS and other Satellite-Based Augmentation System (SBAS) performance
- Substantial improvements in both RAIM availability and ARAIM performance/availability
- Possibly large reductions in WAAS/SBAS complexity and decreased sustainment costs
- Major reductions in ARAIM implementation costs that can help make it a reality

Ground-Based Augmentation Systems (GBASs), such as the FAA’s Local Area Augmentation System (LAAS), must support levels of accuracy that are far tighter than will be obtained directly from GPS III. As such, GPS integrity improvements may offer benefit in reduced complexity for LAAS/GBAS.

Four broad categories of GPS integrity architecture opportunities have been identified. These opportunity categories differ in both timeline and in how WAAS/SBAS and RAIM/ARAIM users can take advantage of them for both military and civilian applications. Each of the four broad categories listed below are discussed in separate sections which follow.

I. First are those opportunities that exist today; they simply require engineering and verification to assess.

II. Second are opportunities that are part of the currently planned near-term GPS IIR/IIR-M, GPS IIF, GPS III, and Next Generation Operational Control System (OCX) program evolution; they require additional engineering and verification, but will not take too much time to realize the benefits.

III. Third are opportunities that will take time to realize the benefits from the currently programmed far-term GPS III and OCX programs (e.g., a full constellation of GPS III satellites).

IV. Fourth are additional opportunities that have been identified which are not necessarily tied to a particular point on the timeline but which have particular benefits to WAAS and ARAIM for both civil and military users.

I. OPPORTUNITIES THAT EXIST TODAY

Opportunities that exist today are in two areas:

1. Taking credit for the a priori integrity failure rates in the Integrity Failure Modes and Effects Analysis (IFMEA) [4, 5],

2. Baselining requirements for dual-frequency augmentations and for ARAIM via the Interagency Forum for Operational Requirements (IFOR) process [6].

I.1 Take Credit for Integrity Failure Rates in IFMEA

The GPS Performance Standards [2, 3] commit the GPS positioning services to an overall a priori integrity failure rate (UMSI probability) better than or equal to $1 \times 10^{-5}$ per hour per satellite, but they make no commitments regarding the cause of that UMSI. If a user is sensitive to only one particular type of integrity failure mode, then – absent any other information to the contrary – that user must conservatively assume that the one particular type of integrity failure mode takes up the entire UMSI probability budget of $1 \times 10^{-5}$ per hour per satellite. The user is not justified in independently assuming (for example) there must be at least 10 different types of integrity failure modes and, assuming that all 10 types of integrity failure modes are equally likely, conclude that the one particular type of integrity failure mode they are sensitive to should only take up a tenth of the UMSI probability budget (i.e., improperly assume just $1 \times 10^{-6}$ per hour per satellite).

To avoid unsubstantiated assumptions and conservatively lean towards higher safety margins, the original FAA requirement analysis [7] assigned the full a priori threshold constellation integrity failure rate of $1 \times 10^{-4}$ per hour to each of the following GPS failure modes for each satellite individually: (a) a fault causing a pseudorange step error greater than 3.6 m, (b) a fault causing an abnormally large pseudorange ramp error, (c) a fault causing a pseudorange acceleration error greater than 0.031 m/s², (d) a fault causing an L1 C/A code/carrier divergence exceeding 6.1 m, and (e) a fault causing an abnormally large ephemeris error. Although this resulted in WAAS being able to handle a worst-case satellite taking up the entire constellation integrity failure rate budget by itself with the worst-case integrity failure mode, it also resulted in WAAS being conservative and consequently limiting its own availability.

The GPS integrity failure mode threats that WAAS must guard against can be directly reduced by an order of magnitude across the board by simply taking credit for the UMSI probability in the GPS Performance Standards which is given as $1 \times 10^{-5}$ per hour per satellite; the ‘per satellite’ qualifier being assurance that no single satellite will be allowed to take up more than a tenth of the entire
constellation integrity failure rate budget by itself. However, an even greater reduction in GPS integrity failure mode threats that WAAS must guard against can be achieved by taking credit for the individual integrity failure mode probabilities for each particular satellite block type that are contained in the IFMEA [4, 5]. For example; integrity faults that cause pseudorange step errors are nearly two orders of magnitude less likely to occur from even the worst block of Navstar satellites than what WAAS was originally designed to guard against. Reducing the WAAS design-to integrity failure mode threat probabilities to the IFMEA determined failure mode threat probabilities would enable algorithmic tuning that could significantly improve WAAS performance, but it would require analysis to substantiate the change in safety case and some cost to make the tuning changes, so it is recommended that the potential performance improvement be assessed first. For ARAIM, the a priori failure mode threat probabilities have not yet been determined, but it is recommended that the IFMEA determined failure mode threat probabilities be used instead of the current WAAS a priori failure mode threat probabilities.

1.2 Baselining Dual-Frequency Operation & ARAIM Requirements via the IFOR Process

GPS operational requirements submitted by the FAA to the IFOR [7] were derived to support current aviation operations using GPS. They form the assertion basis (assumptions) that RAIM, GPS/Inertial, WAAS, and LAAS depend upon for their safety cases. For example; RAIM depends upon an “IFOR” requirement which states that “The URA shall be a conservative representation of the root mean square (rms) of the user range error (URE).” The current set of IFOR requirements were originally applied to the L1 C/A positioning signal because that signal is the foundation for current civil aviation operations with GPS and for every one of the existing augmentations.

Multi-frequency augmentation requirements (i.e., L1 C/A and L5I/L5Q) have been developed as part of the GPS Operational Safety Assessment (OSA) [8]. These requirements are now being validated as part of the IFOR process [6]. They do not require anything new in GPS or its operations other than a straightforward extension of today’s IFOR requirements for the L1 C/A positioning signal to also apply to the L5I/L5Q positioning signals.

Further analyses will be required to develop IFOR requirements for ARAIM. For example; a major concern with ARAIM is the threat of common-mode constellation wide failures like erroneous earth orientation parameter predictions (EOPPs) on multiple satellites [9, 10, 11], so an IFOR requirement on the GPS SIS interface could be proposed to reduce the likelihood of common-mode constellation wide failures to near zero. Such a new requirement might well be accepted by the IFOR based on the very reasonable presumption that it would be more cost effective to eliminate the threat by incorporating a set of mitigations in one place (i.e., the GPS Control Segment) than it would be to require every ARAIM capable GPS receiver incorporate its own set of mitigations (e.g., multi-GNSS operations). For example; the way EOPP are uploaded today – sequentially to each satellite individually – procedural safeguards protect against this EOPP threat unless there is a slowly growing EOPP error that does not become significant until after several satellites are uploaded with a common set of erroneous EOPPs. (As will be noted in the following section, the OCX program already intends to substantially mitigate this threat by validating all input parameters, including the EOPPs).

II. OPPORTUNITIES WITHIN THE PLANNED NEAR-TERM GPS PROGRAM

Four improvements are currently planned within the GPS program over the next several years that can be exploited – with additional engineering and verification – to improve the integrity of the GPS positioning services when all of the IIA satellites are retired and GPS consists of a constellation of IIR/IIR-M, IIF, and III satellites being supported by the OCX:

1. User Range Error (URE) performance will continue to improve,
2. WAGE-2 messages will be available for use by ARAIM,
3. Clock reliability is expected to continue to improve, and
4. Ground-based input parameters will be validated by the OCX.

II.1 Improved Signal-In-Space (SIS) URE

Each generation of Navstar satellites has provided an improved signal-in-space (SIS) URE performance largely because of improved clock performance. As the older satellites are retired, the overall GPS constellation URE will continue to improve simply because there will be fewer satellites with larger URE. Advances in the Control Segment (CS), such as better modeling and updated Monitor Stations (MSs), have also contributed to the improvement in SIS URE over time. Figure 2 shows this trend over the past dozen years [12] for the Standard Positioning Service (SPS) SIS URE as an ensemble statistic across all satellites in the constellation compared to the “worst of any” SPS SIS URE performance standard in [2].

Within the next decade, once the IIA satellites are retired and the III satellites are routinely being launched to
sustain the constellation, it is reasonable to expect the civil SPS SIS URE and the military Precise Positioning Service (PPS) SIS URE [3] to operate with performance under 0.6 meter.

GPS SIS URE performance improvements are of little use to differential GPS (DGPS) systems like WAAS/SBAS since those systems provide their own differentially corrected pseudorange accuracy and their integrity information applies to the threats affecting their differentially corrected pseudorange accuracy – not the threats affecting GPS SIS pseudorange accuracy. For example; the classic GPS SIS URE ramp error caused by a runaway satellite clock has no integrity effect on the WAAS/SBAS differentially corrected pseudorange accuracy so long as the size of the ramp error is small enough that it will fit within the numerical range of the WAAS/SBAS differential corrections.

Unlike WAAS/SBAS, GPS SIS URE performance improvements are vitally needed for ARAIM. ARAIM is being designed with the expectation that the URE will be around 1 meter or less. The GNSS Evolutionary Architecture Study (GEAS) Phase II report [13] analyzed ARAIM availability as a function of the URE standard deviation, the broadcast User Range Accuracy (URA), and other parameters. The results in [13] show that ARAIM availability is higher with lower standard deviations of URE and URA. For example; Table 8-4 of [13] shows that with 99.5% availability of LPV-200 using ARAIM can be achieved over most of the earth’s area between 70° S and 70° N latitude with an optimized GPS constellation of 30 satellites for a URE standard deviation of 0.5 m and a URA of 1 m under certain assumptions. Table 4-2 of [13] shows that with two independent GNSS core constellations (e.g., GPS and Galileo), LPV-200 availability of 99.5% is achievable over most of the earth’s surface between 70° S and 70° N if both constellations have a URE standard deviation of 1 m and a URA of 2.4 m, although the coverage area is increased if one constellation has a URE standard deviation of 0.5 m and a URA of 1 m.

The preceding GEAS Phase II results were based on an a priori probability of an integrity fault existing during an approach of $1 \times 10^{-5}$ per hour per satellite for UMSI. If it could be established that the a priori probability of a GPS integrity fault were smaller, ARAIM algorithms could take advantage of the smaller value and provide increased availability. The GPS III satellites present exactly such an opportunity. As described in [14], the GPS III satellites are the first block of Navstar satellites to have specific SIS integrity assurance requirements applied from the very start of their design. Based on the results given in [14], it is reasonable to expect the GPS III satellites to provide an a priori UMSI probability of an integrity fault existing during an aviation approach of better than $1 \times 10^{-6}$ per hour per satellite. The benefits to
ARAIM availability from this order-of-magnitude reduction in UMSI probability are explored in [15] along with similar benefits from improved GPS III URE. The simulation results given in [15] show a promising path forward for achieving better than 99.5% availability of LPV-200 using ARAIM with just a GPS constellation with 27 satellites.

II.2 WAGE-2 Message For Use By ARAIM

Also within the next decade, Wide Area GPS Enhancement-2 (WAGE-2) will come into operation via the new Military Navigation (MNAV) and Civil Navigation (CNAV) messages. WAGE-2 will provide rapidly updated differential correction (DC) values for each satellite in the constellation along with an associated user differential range accuracy (UDRA) value. See the definition of message types (MTs) 34, 13, and 14 in [16] for an explanation of the DC and UDRA parameters to be broadcast by each MNAV/CNAV-capable satellite.

Unlike the current WAGE correction data in the Legacy Navigation (LNAV) messages which have heretofore only been available to PPS users, the WAGE-2 DC and UDRA parameters will be available to both PPS and SPS users alike. Using WAGE-2, the PPS URE and SPS URE performance will be further improved to be on the order of 0.3 meters. This URE improvement will be for every satellite in the constellation, not just the MNAV/CNAV-capable satellites.

The improved URE provided by WAGE-2 will allow new MNAV- and CNAV-capable GPS receivers, which will all also necessarily be dual-frequency capable receivers (i.e., CNAV messages are not available to single-frequency L1 C/A-code receivers), to provide ARAIM protected LPV-200 operations with high availability comparable to the GEAS Phase II results described in the preceding subsection of this paper. GPS receivers for aviation which are limited to L1 and L5 (no L2) will not be able to take maximum advantage of WAGE-2 accuracy for the IIR/IIR-M satellites which do not provide L5 signals, but ARAIM algorithms are being designed to accommodate incoming pseudorange measurements with wide variations in both accuracy and integrity.

In addition to the significantly improved URE provided by rapidly updated DC parameters, WAGE-2 also includes rapidly updated UDRA parameters that provide a measure of integrity for the differentially corrected pseudorange measurements. In some ways, the UDRA parameters can function as part of an Integrity Support Message (ISM) as defined in [17] for ARAIM. The use of WAGE-2 from the new signals has the potential to further improve LPV-200 availability and coverage if the values of the broadcast DC parameters and the broadcast UDRA parameters which describe the resulting differentially corrected URE standard deviations can be “trusted” with high confidence under routine fault-free conditions. This exploitation of WAGE-2 for ARAIM usage will require additional engineering and verification before it is accepted for operational use.

II.3 Improved Satellite Clock Reliability

The largest single integrity failure mode for the SIS from the Block II/IIA satellites has been satellite clock faults. Based on on-orbit service history, the IFMEA [4, 5] failure rate of clocks onboard the II/IIA satellites is equivalent to two ‘soft failures’ per year for a 24 slot constellation comprised entirely of II/IIA satellites. In contrast to the II/IIA satellites, the on-orbit service history for clock faults on the IIR/IIR-M satellites has recorded zero integrity failures to date. Since the cumulative IIR/IIR-M service history spans a total of approximately 150 satellite-years on orbit, this yields an estimated ‘soft failure’ reliability of no less than $1 - 7.6 \times 10^{-7}$ per hour per satellite.

The current IIR/IIR-M satellites – and the yet-to-be-launched III satellites – contain an on-board clock monitor, called a time keeping system (TKS) [18]. The TKS decreases the probability of clock-related anomalies by at least an order of magnitude because it catches many of the anomalies before they become apparent to the user. When the TKS detects a clock-related anomaly, it causes the satellite to switch to broadcasting non-standard code which prevents users from receiving any UMSI. Thus, those clock-related anomalies become continuity outages instead of integrity failures. Combined with inherent reliability improvements in the atomic clocks themselves, this reduces the clock-related integrity failure rate to an estimated UMSI probability of better than $1 \times 10^{-6}$ per hour per satellite [14].

Clock reliability improvements and a TKS can help both SBAS/WAAS and ARAIM by reducing the probability of integrity threats that they must detect and react to. This reduction enables tuning of their detection algorithms to focus more on other anomalies. The case for improving RAIM performance is not as clear cut because many RAIM algorithms have been hard-coded with the current a priori integrity failure rates. There can be substantial costs associated with redesigning, recoding, recertifying, and fielding an updated RAIM algorithm for operational use. The cost/benefit ratio for RAIM improvements may be significantly greater than unity.

II.4 Ground-Based Input Parameter Validation

On Father’s Day 2012, erroneous ephemeris data was broadcast by a Navstar satellite due to the GPS Control Segment using an erroneous set of earth orientation parameters to calculate the ephemeris data [19]. Earth orientation parameters represent the axis and rotation of the earth with respect to inertial space and time. Earth orientation parameters are provided by the National
The use of SBAS can be problematic. Contract air carriers usually take place in areas where the GPS avionics (military missions and logistics support by

Improvements for military aviation operations are limited performance with improved credit for GPS integrity. There are several potential areas of improvement in WAAS users for some time. As has been noted above, there is still room for improvement.

This example illustrates the importance of a rigorous regime of integrity checks for data received from external sources by the Control Segment. While the OCX has a requirement to implement this kind of input parameter checking, it is the detailed design which will determine the thoroughness of the checks. Therefore, it is essential that the OCX design be scrutinized carefully with respect to input parameter validation to ensure the integrity of the input data meets or exceeds the allocated requirements. At the same time, it must be recognized that the external sources of the data also have a responsibility to maintain the integrity of their processes as well. Furthermore, there is a limit to which the OCX can ensure the validity of the incoming data – even comparing against a second set of identical data produced by an independent source (which may not be available) is not foolproof. Hence, there must be an ultimate tradeoff between the rigor and complexity of the checks.

III. INTEGRITY PERFORMANCE OF GPS III AND OCX PROGRAMMED IMPROVEMENTS

The integrity performance of GPS will improve over time. The maturation of the GPS III satellite design and subsequent OCX updates will create opportunities that can be realized once the IIA, IIR/IIR-M, and IIF satellites are retired, OCX is deployed through Phase 2, and a full constellation of III satellites is servicing military and civil users around the globe. Civil aviation users within the North American coverage area of WAAS would likely migrate to use the dual frequency version of the SPS service, although there may be legacy single-frequency SPS users for some time. As has been noted above, there are several potential areas of improvement in WAAS performance with improved credit for GPS integrity. Improvements for military aviation operations are limited unless worldwide ARAIM is implemented in military GPS avionics (military missions and logistics support by contract air carriers usually take place in areas where the use of SBAS can be problematic). Two integrity architecture opportunities will evolve in the later stages of GPS III and OCX as the constellation becomes filled with GPS III satellites and OCX evolves into its mature state:

1. Control Segment command monitoring, and
2. Expanding the operational constellation size beyond 31 satellites.

Direct OCX support of worldwide ARAIM performance via the GPS III satellites will be addressed in the next section as it is an opportunity that is not currently planned for the OCX program. Note that direct support of ARAIM could also be implemented with an ISM delivered by SBAS/WAAS which would be able to take advantage of the currently programmed integrity opportunities, but this would only provide ARAIM capability within SBAS/WAAS coverage regions.

III.1 Control Segment Command Monitoring

In some cases, an erroneous command, improper command sequence, or omitted command from the Control Segment to a Navstar satellite can result in an integrity failure. An example of an improper command sequence would be when an operator fails to set a satellite unhealthy before conducting a satellite station-keeping maneuver. Station-keeping maneuvers are required periodically because satellites gradually drift out of their nominal orbits and need to be maneuvered back onto their defined constellation slots. In addition, satellites are occasionally maneuvered into different constellation slots. If these maneuvers are done while the satellite is set healthy, users can see large range errors resulting in an integrity failure. This kind of event is rare, but has occurred during the operational history of the GPS, even though the operators are well-trained and motivated.

The irony of having human operators is that while they can make mistakes, they also can solve unforeseen problems that pre-programmed machines (computers) can’t. Therefore, humans must ultimately be in command of the GPS Control Segment and must be able to override the machines when necessary. (This is analogous to always having a human pilot being in command of a commercial airliner.)

The solution to most human errors in a system like this is three dimensional, including good situation awareness for the operators, scripting of routine procedures, and machine monitoring of human actions, with operator override capability. The GPS control segment has always had a combination of these features, but since there have been human errors that have affected the user community; there is still room for improvement.

Because the OCX will replace the current system, this is another opportunity for OCX to equal or exceed the capabilities of the current control system. The OCX has
requirements to provide situation awareness, procedure scripting, and machine monitoring. By the time that a full GPS III constellation is realized, routine operations will be automated and non-routine commands will be machine monitored (i.e., double checked). Thus, operator errors will be effectively mitigated.

### III.2 Expanded Constellation Size

The current GPS system, with the current LNAV message, is limited to a maximum of 31 operational satellites in the constellation. In recent years, however, the actual constellation size has been closer to 35 satellites [12, 19]. With a mature OCX, recently added features of the SIS (see [20] for example) will enable the additional satellites (currently known as “residual satellites”) to be brought into service and thereby expand the total constellation size to 35 or 36 satellites, or maybe even more depending on a number of factors.

Having a larger GPS constellation could certainly benefit SBAS/WAAS users, but there are worrisome questions as to how much it would cost to modify the SBAS/WAAS message structure to handle the increased number of Navstar satellites. RAIM/ARAIM users are very sensitive to having enough satellites with good enough geometry to perform fault detection and exclusion, and so having more satellites in the sky to work with will greatly improve RAIM/ARAIM availability. Fortunately for both RAIM and ARAIM receivers, there is no additional cost from the expanded constellation unless the receiver is specifically upgraded to reap the additional benefits of the expanded constellation (i.e., constellation expansion is backward compatible).

### IV. ADDITIONAL INTEGRITY ARCHITECTURE OPPORTUNITIES

Six additional integrity architecture opportunities, above and beyond GPS III and OCX planned improvements, could provide additional integrity protection:

1. Integrity assured Integrity Status Message (ISM) for ARAIM,
2. Integrity assured URA,
3. Enhanced satellite-based clock monitoring and clock ensembling,
4. Satellite-based signal monitoring,
5. Satellite-based monitoring of satellite pointing and ephemeris errors, and
6. Next generation atomic clock.

Each of these opportunities could bring significant improvement in SBAS/WAAS or ARAIM performance. The addition of the original end-state GPS III opportunity (i.e., integrity assured URA with satellite-based signal monitoring and pointing/ephemeris monitoring) or the set of ARAIM opportunities has the possibility of directly attaining the $1 \times 10^{-7}$ level of integrity. They should, however, be analyzed as incremental improvements to determine the most cost and technically viable set of opportunities to implement.

#### IV.1 Integrity Assured ISM For ARAIM

This new concept is intended to support LPV-200 operations worldwide. Generally, it works on the same principle as traditional GPS RAIM. An airborne algorithm calculates multiple position solutions using different sets of satellites in an attempt to isolate faulty satellites or excessive errors in their measurements. ARAIM also relies on external satellite performance data assured to aviation safety of life (SoL) standards. This data would be delivered via an Integrity Support Message (ISM) independently from the standard constellation performance data (e.g. GPS URA). An ISM would be certified by an Air Navigation Service Provider (ANSP) prior to use by the ARAIM algorithm in the aircraft. Table 1 lists the proposed ISM parameters from [17].

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{URA,i}$</td>
<td>standard deviation of the clock and ephemeris error of satellite i used for integrity</td>
<td>ISM</td>
</tr>
<tr>
<td>$\sigma_{URE,i}$</td>
<td>standard deviation of the clock and ephemeris error of satellite i used for accuracy and continuity</td>
<td>ISM</td>
</tr>
<tr>
<td>$b_{nom,i}$</td>
<td>maximum nominal bias for satellite i used for integrity</td>
<td>ISM</td>
</tr>
<tr>
<td>$P_{sat,i}$</td>
<td>prior probability of fault in satellite i per approach</td>
<td>ISM</td>
</tr>
<tr>
<td>$P_{const,j}$</td>
<td>prior probability of a fault affecting more than one satellite in constellation j per approach</td>
<td>ISM</td>
</tr>
</tbody>
</table>
In Table 1, the URA standard deviation (σ_{URA,i}) and URE standard deviation (σ_{URE,i}) are statistical representations of the URA and URE parameters broadcast by GPS. The maximum nominal bias (b_{nom,i}) accounts for any non-Gaussian errors in the satellites, providing a more accurate characterization of its performance. The probability of a single satellite fault (P_{sat,i}) is similar to the GPS probability of a major service failure, and the probability of a constellation wide fault across multiple satellites, (P_{const,i}) is a new probability which is needed to account for common mode faults not historically captured by the GPS IFOR requirements. The values of these parameters would need to be verified using ground monitoring systems. The complexity of the ground monitoring systems would vary depending on the trust in the existing satellite and constellation design performance.

An international working group is evaluating the feasibility of ARAIM using two constellations and notional performance assumptions to support their simulations [17]. ARAIM may be able to be supported by a single constellation given a sufficient, minimum number of satellites and improved performance [15]. This should be examined in further detail as part of future studies.

### IV.2 Integrity Assured URA

This is a concept where each Navstar satellite would broadcast a URA parameter that is the standard deviation of a Gaussian distribution that over-bounds the actual URE distribution of a ranging signal (or a set of ranging signals) from the satellite with a probability of $1 - 1 \times 10^{-8}$ per hour per satellite or better. That would enable a user receiver to use the URAs from up to ten satellites in the position solution to calculate an overbound to the user position solution with an assurance level of $1 - 1 \times 10^{-7}$ per hour or better. This could be done without reliance on an over-determined position solution, as for RAIM/ARAIM, which would maintain availability of integrity with fewer satellites in view than with RAIM/ARAIM. This could also be done without reliance on a GPS augmentation system such as SBAS/WAAS which would allow operations with integrity outside the coverage areas of those systems.

This requires a total GPS Space and Control Segment integrity solution to be able to provide an integrity-assured URA. The principal threats to the integrity of the integrity-assured URA are errors in the URA values themselves or failures of the Space or Control Segments that cause GPS ranging errors which violate the overbound of the URE distribution.

These threats can be mitigated by standard hardware and software assurance processes that assure the URA to the appropriate level of integrity. These processes typically include failure modes effects and criticality analyses, fault tree analyses, rigorous software development processes, and rigorous hardware development processes.

In cases where individual failure modes do not meet the required probability level, monitors can be used to detect those failures and initiate corrective action. In the case of the Control Segment, failures can generally be detected prior to corruption of a satellite command or navigation data upload and thereby be prevented from reaching the satellite and thence the users. In the case of the Space Segment, some failures can be detected and prevented or corrected on board, but the general case is that a failure will result in the satellite terminating broadcast of the signals or issuing some sort of “do not use” indication such as non-standard code, default navigation data, etc. To support aviation users, the satellite detection and correction must be within a few seconds of the occurrence of the failure. The specific time-to-alert requirement is determined by the particular aviation operation supported.

### IV.3 Enhanced Satellite-Based Clock Monitoring and Clock Ensembling

The GPS III clock monitor (TKS) decreases the probability of clock anomalies at least an order of magnitude because it catches many of the anomalies before they become apparent to the user. An improved on-board monitoring system could decrease the probability of clock anomalies by several more orders of magnitude by, essentially eliminating the clock as a significant source of signal anomalies. The TKS detects anomalies by comparing the atomic frequency standard (AFS) to a quartz oscillator (XO), using an on-board phase meter [18, 21]. The current phase meter has a granularity of at least 1.7 ns, which severely limits the ability to see unpredictable behavior in the AFS until it is fairly large. The XO is more stable than the AFS only out to about 100 s [22]. Degraded performance in the AFS can often appear as a slow change in frequency, slower than 100 s [23]. Hence the XO cannot detect this.

There are options to eliminate both of these limitations [23]. A dual-mixer measurement system has been proposed, with accuracy at least in the ps range, as well as the ability to simultaneously compare many devices. A reference for detecting anomalies in the AFS could be one or two other AFSs on-board the satellite, in addition to the XO. Using three AFS would allow not only detection, but also rapid remediation of an anomaly, using majority voting to determine the source of the failure. Measurements could be made at a 10 Hz rate or higher, allowing for detection of measurement error, and eliminating the noise in the measurement system. Combining the measurement of all frequency devices, AFS, XO, and potentially cross-linked time transfers from clocks on other SVs, could be done with an ensembling technique that uses the best properties of all devices.
Clock ensembling is used in the two US national labs, NIST and USNO, to produce their real-time UTC.

IV.4 Satellite-Based Signal Monitoring

Current GPS II satellites perform a variety of on-board monitoring functions to maintain the integrity and quality of the navigation signals, but they do not do end-to-end checks at or near the output of the satellite antenna. This means that the Control Segment must monitor for some satellite navigation signal failure modes. In these cases, the typical response time of the Control Segment to detect a satellite signal failure and return a corrective action to the satellite is significantly longer than the various times to alert for aviation operations (from 1 to 30 seconds). Even with an ideal system design, it is doubtful if the Control Segment could respond in time to meet aviation times to alert, especially if an operator must be in the loop to confirm the detection and corrective action.

Response times to these types of failure modes could be significantly reduced to between fractions of a second and a few seconds, if the monitoring and corrective action functions were hosted on board the satellite. Typical corrective action would be to broadcast non-standard code or shut down a transmitter so GPS user receivers couldn’t track the erroneous signal. One way to accomplish this monitoring is to sample the positioning signals just before they reach the broadcast antenna. Examples of this kind of monitoring could include navigation data read-back, code/cARRIER divergence monitoring, waveform quality monitoring, and output power monitoring. These are all signal monitors that already exist on the ground, for example, in the WAAS, and are in development for OCX, so the monitoring feasibility is already proven. The challenge is to implement them aboard a satellite where payload volume, weight, and power are tight constraints.

For navigation data read-back, what is needed is essentially the front-end of a GPS receiver. The navigation data can be demodulated from the composite signal and then compared bit-by-bit with the data stream at the first point where it was generated. Similarly, the code and carrier can be demodulated and the phase difference measured to ensure that any divergence remains within specifications. Waveform quality can be measured, as on the ground, by correlating with the composite signal and making delay (range) measurements at different points on the correlation peak to measure the shape of the peak and thereby the regularity of the waveform. There may be monitors that are feasible on board the satellite that would not be feasible on the ground, because the signal power is much higher aboard the satellite than received on the ground and the signal is free of atmospheric effects. For example; it might be possible to directly compare a copy of the ideal signal against the demodulated waveform to detect any spurious energy that would represent signal distortion and spurious elements of the signal. Cycle slips in the composite signal might be detected by detecting loss of lock with the demodulated signal or by detecting sudden changes in the phase lock error measurement. Finally, composite and component signal powers can be measured as a byproduct of these other measurements to potentially detect uniform power level drops, energy lost due to distortion and generation of spurious signals, or impedance changes due to a failure in the antenna or signal path.

IV.5 Satellite-Based Monitoring of Satellite Pointing and Ephemeris Errors

Potential threats to GPS integrity include unintended changes in satellite orbits or satellite antenna pointing errors. Both of these errors are difficult to detect in a timely manner from the ground, so they are candidates for on-board monitoring by the satellite. Unintended changes in a satellite orbit or attitude could be caused by uncommanded firing of thrusters, propellant leaks, impacts with micrometeorites or debris, and similar problems. Unintended changes in attitude could also be caused by a failure in an attitude control system. These have historically been rare events, and satellites can already detect extreme events of this kind, but may not be able to detect subtle events to the degree necessary to provide GPS integrity without integrity augmentation.

There are two basic approaches available to detecting these kinds of errors directly. One is by transmitting directional signals between satellites (directional cross-links). The other is by use of a “proof mass” to detect deviations from a pure gravity-driven satellite orbit. By detecting these kinds of errors on-board the satellite, a do-not-use indication (such as non-standard code) can be broadcast for the satellite signals within seconds versus the minutes (at best) that the problem could be detected and responded to by the Control Segment.

Using directional cross-links, range measurements can be taken between satellites to develop a network of relative satellite positions. Deviations from the nominal satellite orbits can be detected as deviations from their predicted relative positions in the network. Satellite attitude errors can be detected as deviations from predicted pointing angles of the satellite cross-link antennas.

The “proof mass” technique consists of a small sphere or cube contained in a shell which protects it from the external drag forces on the satellite such as residual atmosphere and solar radiation pressure. Movement of the proof mass relative to the shell (and therefore the satellite) is a direct measure of the deviation of the satellite from a pure gravity orbit. The proof mass technique has been used in several satellite programs to increase ephemeris predictability (U.S. Navy Transit satellites) or to make precise gravity measurements (NASA Gravity Probe B and ESA Gravity Field and
Steady-State Ocean Circulation Explorer (GOCE)). In these cases, the satellite was steered by use of thrusters to keep the proof mass in the center of the shell and keep the satellite in the equivalent of a drag-free orbit. For Navstar satellites, it is not necessary to maintain a true drag-free orbit. Instead, the ball can be electrostatically maintained in the center of the shell, with the force required used a measure of the satellite deviation from a pure gravity orbit. An unintended change in the satellite orbit would be detectable as a sudden change in the force required to keep the mass centered and would have a magnitude much larger than the nominal force required. Attitude errors can be detected if the proof mass is located along the pointing axis of the satellite, away from the center of mass, where the rotation of the satellite would result in a curved-line motion of the proof mass around the satellite center of mass.

An advantage of the directional cross-link technique is that it is conventional technology. The disadvantage is that most of the satellite constellation has to be equipped with the cross-links to develop a useful network of relative positions with good accuracy. An advantage of the proof-mass technique is that it is independent of whether other satellites are similarly equipped, or not. The disadvantage is that it is a relatively novel technique that is not as widely used as directional antennas and may require some adaptation for the GPS application.

IV.6 Next Generation Atomic Clock.
The next generation atomic clock is in research and development. The current baseline design is a cold atom clock that will be an order of magnitude more stable than the current Block IIR/IIR-M and IIF rubidium atomic frequency standards (RAFS). The expected deviation one day after upload would be 60 ps, or 2 cm, as one standard deviation. With three of these clocks on-board, and a measurement system accurate to the ps region, it would be extremely unlikely to produce an error as large as 1 ns without detection. The stability of the cold-atom clock would take ten days to produce the time error that the best RAFS creates in one day.

Differences in design between the RAFS and the cold-atom clock suggest increased reliability in addition to increased stability. The RAFS uses a rubidium lamp to excite the atoms, which are contained in a glass cell along with a buffer gas. There are various mechanisms for the rubidium lamp light to degrade, as well as for the gas in the cell to migrate through the glass. The cold-atom clock will use a laser system to excite the atoms, and the atoms will be in motion as a beam in vacuum. The laser will much more precisely excite the clock atomic transition, and the atoms will have no collisions or interactions that can cause anomalous behavior in the frequency standard. The life of the clock will be determined by the amount of atoms in the reservoir at launch, which can be chosen appropriately. With these design differences, there will be no “personality” effects among the clocks, as is seen in the current RAFS.

With this new clock, an immediate impact would be that the clock effect on the URE would be less than 1/100 as large as that due to the RAFS. The TKS is also known to have a significant effect on the URE, and that could be reduced also. Other causes of URE would remain, such as ephemeris error and inter-signal biases. The net effect should be to reduce the URE to approximately 15 cm, resulting in an improvement in the availability of integrity for ARAIM users.

CONCLUSION

This paper suggests that it may be beneficial to examine the technical and economic viability of an incremental approach to improving the GPS integrity architecture to benefit improved performance and reduce the complexity of integrity augmentations WAAS and RAIM/ARAIM. The motivation for taking advantage of integrity architecture opportunities is to enable LPV-200 civil and military aviation operations worldwide.

Initially, this incremental approach should take advantage of the opportunities that are available today through IFMEA a priori rates and IFOR requirements for dual frequency augmentation and ARAIM. In the near term, the incremental approach should plan to take advantage of opportunities created by improved URE, WAGE-2 messages available for use by ARAIM, improved clock reliability, and validated ground-based input parameters. In the far-term GPS III and OCX programs, the incremental approach should plan to take advantage of opportunities created by control segment command monitoring and expanding the operational constellation size beyond 31 satellites. Finally, the additional opportunities that are not currently within the GPS program plans should be considered to determine the most cost effective and technically viable set of opportunities to implement.

CAVEAT

The opinions and assertions contained herein are strictly those of the individual authors. These opinions and assertions do not represent the official positions of the authors’ organizations.

REFERENCES


12. Cooley, Col. B., “GPS Program Update to Civil GPS Service Interface Committee (CGSIC)”, 53rd Meeting of the Civil GPS Service Interface Committee, Nashville, TN, 17 September 2013


19. Gruber, Col. B., “GPS Program Update to Civil GPS Service Interface Committee (CGSIC)”, 52nd Meeting of the Civil GPS Service Interface Committee, Nashville, TN, 18 September 2012

