Rapid Prelaunch Testing of the JPSS VIIRS Sensor

Carl H. Fischer

MIT Lincoln Laboratory has an extensive history helping develop hardware and communications systems for previous environmental monitoring satellites and is assisting, to varying degrees, with the design and development of all major sensors for the next-generation Joint Polar Satellite System (JPSS). Recently, Lincoln Laboratory has leveraged this experience to develop a Data Analysis and Decision Support cluster that allows scientists to extract, process, and plot statistics expeditiously for tens of gigabytes of sensor data contained across hundreds of data files. These "quick looks," or rapid assessments of the health and sufficiency of test data, have been implemented for the thermal-vacuum test phase of one JPSS sensor, the visible/infrared imager/radiometer suite (VIIRS), which will provide important environmental data, including most of the atmospheric cloud and surface information needed by weather forecasters.

» **For more than 40 years, polar-orbiting** satellites have been providing data for global environmental monitoring, benefitting both civilian and military communities through timely weather forecasting and improved situational awareness. Satellite data also have supported aviation and search-and-rescue missions and have advanced the understanding and prediction of changes in the Earth's environment. Two independent U.S. polar satellite systems that have historically provided these data are the Polar-Orbiting Environmental Satellite (POES), operated by the National Oceanic and Atmospheric Administration (NOAA), and the Defense Meteorological Satellite Program (DMSP), operated by the Department of Defense.

The next generation of satellites, previously known as the National Polar-orbiting Operational Environmental Satellite System (NPOESS) and now reorganized as JPSS, will extend the system's capabilities in several mission areas [1]. Updated existing systems will be combined with five new instruments illustrated in Figure 1: a hyperspectral cross-track infrared sounder (CrIS), an advanced technology microwave sounder (ATMS), a microwave imager/sounder (MIS), an ozone mapping and profiling suite (OMPS), as well as a visible/infrared imager/radiometer suite (VIIRS). These instruments will yield information about soil moisture, sea surface temperature, ozone depletion, and ocean color, among other mission products. In addition, higher spatial resolution measurements of atmospheric temperature, water vapor, and pressure will be possible using smaller, lighter, less power-demanding sensors.

Prior to launch of the first JPSS satellite, which is scheduled for 2014, careful testing and calibration of satellite sensors in space-like conditions are essential to achieving mission objectives once the satellite is in orbit. The NPOESS Preparatory Project (NPP) is a precursor satellite scheduled to be launched in fall 2011. The NPP is a risk-reduction effort designed to test several of the new sensors in space and to ensure that the ground control systems work properly before launch of the first operational JPSS satellite [2]. This ambitious satellite sensor suite has been referred to as potentially “the most complex environmental satellite program ever developed,” and presents numerous technical, developmental, integration, and managerial challenges [2].

Design and manufacturing problems have threatened the overall project cost and schedule baseline. A principal challenge of typical prelaunch sensor test programs is to carry out a thorough and accurate characterization of complex hardware within a constrained schedule and budget. The necessary array of vacuum chambers, test fixtures, and measurement equipment is expensive and requires operating teams of five to 10 experts working 24 hours a day. The total costs often exceed $100,000 per day. Faulty sensors or errors in data acquisition can impact greatly the budget and timeline of the overall project.

MIT Lincoln Laboratory has helped develop hardware and communications systems for previous environmental monitoring satellites, including the POES and the Geostationary Operational Environmental Satellite systems. Lincoln Laboratory has leveraged this knowledge to assist, to varying degrees, with the design and development of all of the major JPSS sensors. Furthermore, Lincoln Laboratory has recently applied its extensive experience in sidecar systems to develop a system to allow on-site science team members to browse several different levels of summarized instrument data in real time through a simple web interface. The Data Analysis and Decision Support (DADS) cluster allows scientists to extract, process and plot statistics expeditiously for tens of gigabytes of sensor data contained across hundreds of data files. These “quick looks,” or rapid assessments of the health and sufficiency of test data, have been implemented for the thermal-vacuum test phase of the VIIRS sensor. VIIRS will provide important environmental data, including most of the atmospheric cloud and surface information needed by weather forecasters. By permitting sensor test data to be quickly inspected, DADS has delivered a substantial cost savings for this project.

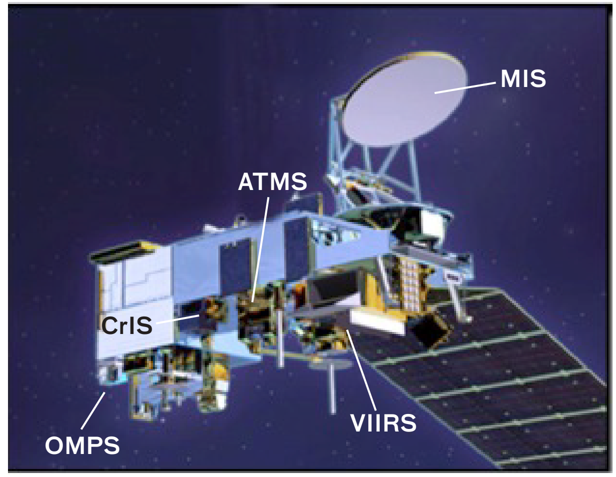

FIGURE 1. This artist’s illustration of the National Polar-orbiting Operational Environmental Satellite System shows the location of the new instruments: a hyperspectral cross-track infrared sounder (CrIS), an advanced technology microwave sounder (ATMS), a microwave imager/sounder (MIS), an ozone mapping and profiling suite (OMPS), as well as a visible/infrared imager/radiometer suite (VIIRS).

Four key attributes of the DADS system that highlight key design decisions are critical to its success. First, its easy-to-use web-based interface minimizes the learning curve for users. Second, the system precomputes commonly used statistics so they can be displayed on interactive time scales. Third, queries to extract data for the plotting engine are built from administratively defined configuration tables, thereby enabling rapid configuration of new data visualizations as they are requested. Finally, the reporting engine is based on re-usable, user-defined templates that can be quickly run against sensor data as they arrive. These attributes encouraged use of the system by scientists from several organizations and substantially reduced the level of effort necessary to accomplish the most frequently occurring test data review tasks.

Data Distribution

The vacuum testing for VIIRS was performed at the site of the VIIRS contractor, Raytheon Space and Airborne Systems, at El Segundo, California. Each day of testing, VIIRS typically generates 100 gigabytes (GB) of data, which should be distributed to analysts around the country in less than 24 hours to ensure an adequate review by various science teams, including Science Systems and Applications, Inc., Lincoln Laboratory, and Aerospace

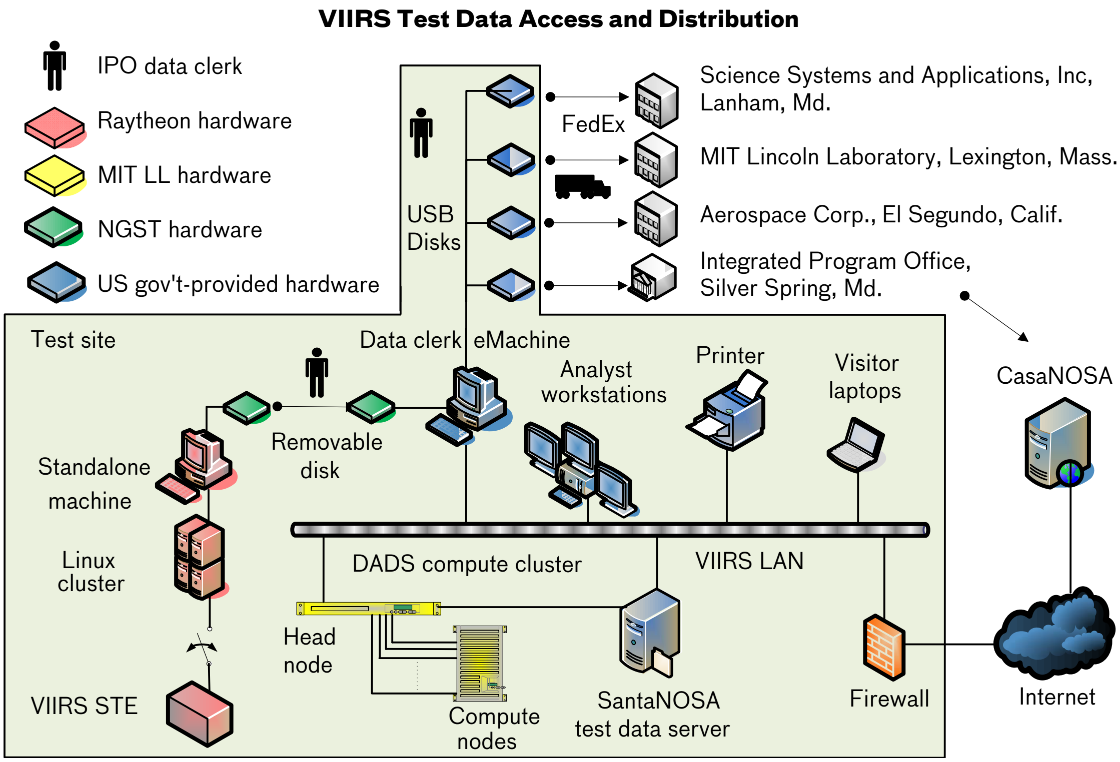

FIGURE 2. About 100 gigabytes of data are collected during each day of testing for the VIIRS sensor. Raw data is transferred to science teams by three methods: on-site access is given to the SantaNOSA test data server over a secure virtual LAN; off-site scientists received daily shipments of USB disks containing the previous day's test data; and the IPO program NOSASync pushed data on SantaNOSA to CasaNOSA, which is accessible via an Internet connection to registered users of the calibration/validation teams.

Corporation [3, 4]. During previous phases of the test program, these organizations thoroughly assessed sensor performance by performing test data analysis over a two-week period. Such a lengthy response window was no longer practical for prelaunch tests, as irregularities in test data should be identified before the configuration of the sensors or test equipment changes. Thus, the Lincoln Laboratory DADS system balanced turnaround speed with depth of analysis and reworked existing methods for distributing and reviewing data. An overview of the data-transfer process is illustrated in Figure 2.

The Data Products Division at the Integrated Program Office (IPO) designed a low-cost data distribution operation that directed test data in nearly real time from the high-bay to a data server owned and operated by Raytheon. Analysts were given access to test data on this server via a read-only workstation. For security reasons, an air gap was formed between this read-only workstation and government equipment. An on-site "data clerk team" was responsible for many aspects of data capture for the 100-plus days' duration of the thermal vacuum test. The data clerk transferred test data (typically several gigabytes of data at 30-minute intervals) via a removable external hard disk from the read-only workstation to an on-site IPO-provided test data server called SantaNOSA. Data typically became available on SantaNOSA between two and ten hours after collection.

Raw sensor data was distributed using one of three methods. On-site sensor scientists had access to SantaNOSA and the DADS cluster data over a secure VIIRS local area network (VLAN) and could commence analysis immediately. Off-site scientists received daily overnight shipments of USB hard disks containing any test data collected by 4:00 p.m. the previous day. Data were also automatically transferred to the NOAA headquarters by using NOSASync, an IPO application installed on SantaNOSA that pushes data to CasaNOSA, an Internet-accessible server. CasaNOSA serves as an official test data repository for analysts and calibration/validation teams around the country.

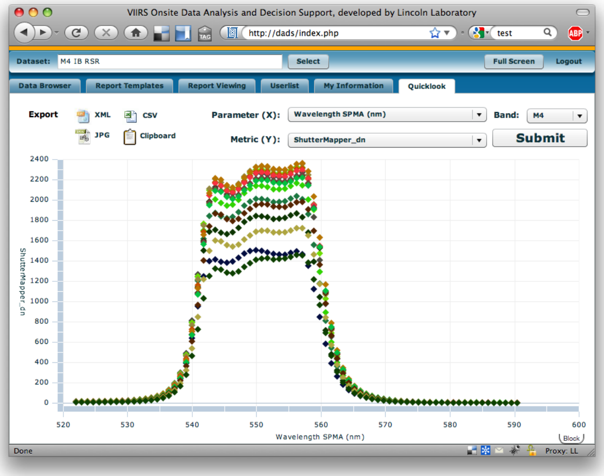

FIGURE 3. A screen shot of the DADS user interface illustrates its utility in plotting large data sets to quickly monitor sensor performance during testing. Each x coordinate represents a single data file of tens of megabytes. All detectors are shown for the selected band.

The DADS system operates by computing a standard set of statistics from every data file as they are delivered to SantaNOSA. The results of these data analyses are cataloged in a MySQL database. The database delivers these statistics on demand to a web interface that allows users to browse data and build test data reports.

Data Extraction and Analysis

To provide a system capable of many basic data health metrics, Lincoln Laboratory leveraged its previously developed tools for the parallel extraction and summarization of data sets. The new system allows on-site science team members to browse several different levels of summarized instrument data in real time.

VIIRS test data files, ranging in size between 10 and 600 megabytes, are discovered and queued by a set of bash and perl scripts. The processing queue is maintained as a table in the database. In addition to the raw sensor data files, the satellite vendor delivers ancillary information, such as test equipment parameters, in text format such as comma- or tab-separated values. Perl scripts parse and load these data into tables that can be efficiently joined to the appropriate sensor-data statistics tables. This critical capability allowed the DADS system to deliver plots of sensor response versus test equipment parameters, such as monochromator wavelength (as shown in Figure 3) or blackbody temperature, upon request of the science team representative.

Once the data are ingested and queued, an array of 14 worker nodes (Dell Power Edge M600 blade servers) begins to process the files. Each node has eight gigabytes of random-access memory and two quad-core Intel Xeon processors. The nodes are managed using Rocks cluster management software [5, 6]. Two Matlab threads run on each node for a total of 28 simultaneous processing tasks. All 28 nodes perform the same data-reduction task on every file in the queue, and a single-instruction, multiple-data approach is used to ensure that no two nodes process the same file.

The DADS system computes a standard set of statistics from every data file as it is delivered to the file server. The results are cataloged in a query database and the commonly used statistics may be displayed to the user within interactive time scales. Users may build test data reports by browsing data, building queries from administratively defined configuration tables, and extracting data for the plotting engine. The reports may be saved as a presentation (e.g., in Microsoft PowerPoint) or as a portable document (e.g., Adobe PDF).

Data Reduction

VIIRS test data sets often contain hundreds to thousands of data files, each of which contains data from a stable set of test conditions. Most analysis routines begin by reducing the volume of test data to a manageable size—often one or two metrics are necessary per detector from each data file. The method for calculating these metrics varies widely on the basis of the specific parameter being tested. For example, to assess the quality of relative spectral response testing, an analyst would typically plot the mean difference between shutter-open and shutter-closed data versus monochromator wavelength. On the other hand, to assess emissive band linearity, the analyst might review plots of mean sensor digital number response divided by blackbody radiance versus blackbody radiance. To facilitate the wide array of potential user requests, an extensible processing system was developed in Matlab to handle the known processing requirements as well as the unexpected needs that may arise during the test.

After each data file is claimed, extracted, and loaded by a worker node, it is stored on a network share site, allowing other analysts to skip the time-consuming extraction process. Next, the node runs through a set of Matlab processing scripts on the file. The required parameters

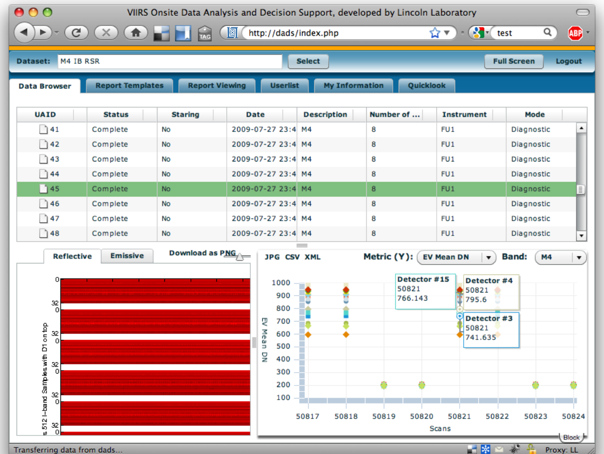

FIGURE 4. A screen shot of the DADS user interface includes a table at the top that lists all data files associated with the selected sensor test and two panels below that show data from the selected file. Clicking a new file refreshes the bottom panels. Drop-down menus on the bottom right panel allow the user to select plots and figures that are relevant to the specific test.

and results from the scripts are documented with examples, allowing other teams to develop custom statistics for future expansion. These scripts populate statistics tables on the central database server with simple metrics, such as calculations of mean value, standard deviation, maximum and minimum values, and counts of saturated samples. Statistics are stored with the file from which they originated, the date and time as reported by the sensor, and the absolute scan number. Once the processing scripts finish, the node marks the file complete in the processing queue table and repeats the recipe with another file in the queue.

The end result of these steps is a comprehensive table of relevant statistics for each file processed. These statistics are organized by a relational database for efficient access and represent a data reduction of several orders of magnitude. Multiple instances of this process can run simultaneously across a compute cluster without conflicts because the processing queue ensures only one instance works on each file.

User Interface

DADS includes a custom web interface to display selections from the statistics database to end users. Because the government on-site science team comprised users from various organizations, it was important that the user interface be compatible with a variety of different operating systems and security postures. Adobe® Flash was selected for its extensive library of visualization components, rapid development cycle, and the wide adoption of Flash Player on end users' web browsers. The interface went through four major revisions in the short three-month Agile Scrum software development cycle [7].

Using the interface, shown in Figure 4, scientists and engineers can browse a table of data files and, with a few clicks, perform the following tasks: retrieve plots and images that are updated nearly instantaneously, generate plots of test parameters versus sensor response from hundreds of data files in just a few seconds, design report templates to tabulate and plot relevant information from a set of data files, request interim reports from existing report templates, annotate those reports on the basis of their interpretation of the results, and finalize annotated reports to be included in the program archive.

The server-side processing necessary to deliver data to the flash components was implemented in PHP, a scripting language especially suited for web development, by using an object-relational model. The use of a relational database allows the efficient joining of information from sensor statistics, test logs, and test equipment settings.

DADS Performance

An on-site data clerk team proved to be a low-cost, effective method for VIIRS test data distribution. Unlike a strictly electronic distribution path, the human data clerk was always on hand to obtain and organize documentation for test procedures and various other shift logs, test logs, and test reports. The data clerks served a program librarian role, locating data, finding documents, managing the data flow process, and serving any requests from the sensor scientist teams. This approach allowed government sensor scientists to concentrate on data analysis rather than on how or where to find the necessary data and documentation.

The off-site data distribution method by USB disks proved to be a low-cost, easy, and reliable method to distribute data and was often preferred by analysts over electronic retrieval. While fully electronic transfer was always a viable option, the USB data distribution eliminated many technical barriers, lowered total costs, and gave scientists more time to focus on building a quality VIIRS sensor.

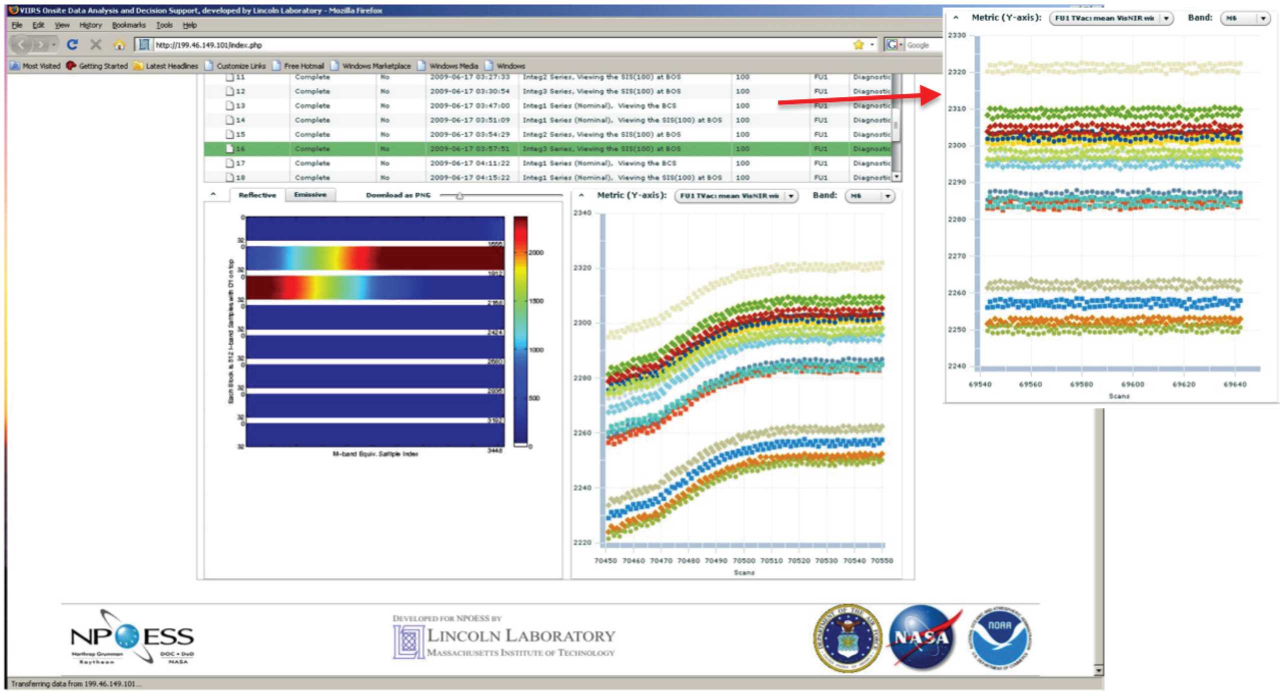

FIGURE 5. Sequential data collections can be presented by clicking on the appropriate scan in the main screen. The plots of data sets 12 and 16 are shown. Between scans, bus voltages were changed. The data for collect 12 (upper right) suggest that the sensor had stabilized, but the variation in response magnitude suggests that the instrument had not yet stabilized after reducing the bus voltage before collect 16 (center).

Despite its rather simple collection process of automated statistics, the DADS system often delivered the first insight into several potential issues that warranted a review of raw data. Screen captures from the DADS system became commonplace in the daily shift reports from scientists serving as the on-site science team representatives and played prominently in the early detection of test issues.

The success of the DADS system is attributed to the dramatic reduction in time necessary to inspect data and the corresponding increase in the amount of data that was independently reviewed. Simplifying the tasks of plotting and browsing commonly needed statistics enabled the scientists to review more data and investigate anomalies more closely. For example, the DADS tool was applied to investigate a reported sensitivity of measurements to bus voltage variation. Plots of average digital numbers were generated for a series of 100 scans for several bands in some collects. Between scans, bus voltages were changed, and two example plots using the shortest integration time are shown in Figure 5. While the data in the top plot suggest that the sensor had stabilized for collect 12, the variation in response magnitude suggests that the instrument had not yet stabilized after reducing the bus voltage before collect 16. A DADS user was able to detect this problem early and avoid delays and a cost of ~$1,000,000 in lost testing time.

In another example, the DADS tools were used to flag suspicious data (e.g., measured points that were orders of magnitude larger than neighboring points). Anomalies, such as bit flips or multiple bit flips and the bands in which they occurred, were more easily identified and catalogued by reviewing the statistics around these flagged points.

At the conclusion of thermal vacuum testing, the DADS statistics database contained over 1.2 billion records. Despite this large volume, the interface remained responsive to user interaction, and lags in user-interface response were dominated mostly by the Flash Player's ability to generate plots, not the database response times.

REFERENCES

1. NOAA Satellite and Information Service Integrated Program Office: http://www.ipo.noaa.gov
2. Testimony of the Assistant Administrator for Satellite and Information Services, National Environmental Satellite, Data, and Information Service Gregory Withee on the National Polar Orbiting Operational Environment Satellite System (NPOESS) program before the Senate Subcommittee on Disaster Prevention and Prediction, Committee on Commerce, Science, and Technology, March 30, 2006.
3. C. Welsch, H. Swenson, S.A. Cota, F. DeLuccia, J.M. Haas, C. Schueler, R.M. Durham, J.E. Clement, and P.E. Ardanuy, "VIIRS (Visible Infrared Imager Radiometer Suite): A next-generation operational environmental sensor for NPOESS," *Proc. Intl. Geoscience and Remote Sensing Symposium*, vol. 3, 1994, pp. 1020–1022.
4. C. Schueler, J.E. Clement, L. Darnton, F. DeLuccia, T. Scalione, and H. Swenson, "VIIRS sensor performance," *Proc. IEEE Intl. Geoscience and Remote Sensing Symp.*, vol. 1, 2003, pp. 369–372.
5. F.D. Sacerdoti, S. Chandra, and K. Bhatia, "Grid systems deployment and management using rocks," *IEEE Intl. Conf. Cluster Computing*, 2004, pp. 337–345.
6. P.M. Papadopoulos, M.J. Katz, and G. Bruno, "NPACI rocks: Tools and techniques for easily deploying manageable Linux clusters," *Concurrency and Computation: Practice & Experience*, vol. 15, no. 7, 2003, pp. 707–725.
7. K. Schwaber and M. Beedle, *Agile Software Development with Scrum,* Upper Saddle River, N.J.: Prentice Hall, 2001.

ABOUT THE AUTHOR

Carl H. Fischer was a technical staff member in the Sensor Technology and System Applications Group of Lincoln Laboratory between 2004 and 2010. Fischer was the 2009 recipient of the MIT Excellence Award for Innovative Solutions. He earned both a bachelor's degree in engineering-physics and a doctorate in electrical engineering from the University of Michigan. He is currently a principal engineer at Ball Aerospace.