16.1 OBSTACLES AND SOLUTIONS IN NEAR-REAL-TIME PROCESSING OF TERRA AND AQUA MODIS DATA

Paul Haggerty* and Kristina Sprietzer
Science and Technology Corporation, Suitland, Maryland

Gene Legg, National Oceanic and Atmospheric Administration, Suitland, Maryland
Ray Luczak, Computer Sciences Corporation, Suitland, Maryland

1. INTRODUCTION

NANOOK, the near-real-time MODIS processing system was conceived as a proof-of-concept program and implemented under the Computer Sciences Corporation (CSC) Central Satellite Data Processing (CSDP) contract. It was designed to provide NOAA (National Oceanic and Atmospheric Administration) with rapid access to MODIS (Moderate Resolution Imaging Spectroradiometer) environmental data from the Terra and Aqua satellites. Additionally, NANOOK provides risk reduction for future high data rate satellite systems by identifying the risks and problems associated with high data volumes in near-real-time operations.

2. THE MODIS INSTRUMENT

The MODIS instrument provides a data stream divided into 36 bands across the spectrum, ranging from the visible to the thermal infrared, with spatial resolutions of 250 meters, 500 meters, and 1 KM, depending on the particular band. The band combinations were chosen to provide high-quality science data for research into Land, Ocean, Atmosphere, and Climate disciplines. A MODIS instrument was launched aboard the Terra Spacecraft in December 1999 and the Aqua Spacecraft in May 2002.

3. THE PREHISTORY OF NANOOK

The NOAA/NESDIS near-real-time project began in 1997 with an original task order describing a project to develop an extremely limited processing system designed to provide data to the National Weather Service inside a mandated time limit of three hours from observation. At that time, the launch of Terra was scheduled for June 1998, and agreements were put in place between NOAA and NASA (National Aeronautics and Space Administration) that would allow NOAA to receive raw data from EDOS (EOS Data and Operations System) and all necessary processing code from MEBS (MODIS Emergency Backup System). MEBS was originally formed by NASA as a parallel group which could handle overflow data from the main MODIS program, and to handle other requirements which might come up and need to be handled without interrupting the official system. NASA soon released that the data load from Terra would be larger than the primary system could take. MEBS was later renamed MODAPS (MODIS Adaptive Processing System) and folded into the primary production environment.

By November 1997, the agreements were in place, floor space at the Goddard Space Flight Center (GSFC) had been allocated, and the initial hardware had been delivered. In order to keep the porting of NASA code to a minimum, NOAA purchased the same architecture that NASA would be using; in this case an Origin 2000 computer from SGI, with 2 CPUs, 1 Gigabyte of RAM, and 5 9GB internal hard drives. Even at this point, upgrades were already seen as a necessity. But with the existing system, preliminary work could begin.

Due to the size of the incoming data stream (estimated at 6GB every 100 minutes), the NOAA system was required to be co-located at GSFC with EDOS so that a high-speed network could be used to transfer the raw data. As EDOS was operating a 100-megabit FDDI (Fiber Distributed Data Interface) ring, an agreement was reached to hook the NOAA server, now named SOAP, to the ring. This turned out to be the first obstacle to be overcome. In a situation faced by computer purchasers everywhere, the FDDI network card procured, with a PCI bus, was incompatible with the Origin 2000, which used an XIO bus. The solution in this case was to put the FDDI card into the NOAA router, which would be used for communications with other NOAA computers and as the outgoing network junction for communicating back to the NOAA offices in Suitland. This communications link, in turn, became vital when a tentative agreement with NASA to base NOAA personnel at GSFC fell through, requiring SOAP to be managed remotely, with only occasional maintenance visits by NOAA personnel.

In order to minimize processing, the original task called for processing only continental United States (CONUS) granules (a five-minute segment of data), yielding a total processing data set of twelve to twenty-four granules/day depending on the given orbital tracks. Each orbit would consist of two or three granules, with three to four orbits in the day, and a like number at night. The data delivery requirement called for delivery to customers no more than three hours after observation, 90 percent of the time. As it took NASA

* Corresponding author address: Paul Haggerty, STC, NOAA/NESDIS/IPD Federal Center FB-4, 4401 Suitland Road, Suitland, MD 20746;
e-mail: Paul.Haggerty@noaa.gov

two hours to transfer the data through their own networks before arriving on the NOAA processor, only one hour remained for all additional processing to be completed and the data sent back to the Suitland facility for distribution. In the original plan, the raw RBD (Rate-Buffered Files) would be received directly from NASA, converted to Level 0 files, and then to Level 1A and Level 1B using the NOAA server at GSFC, and the Level 1B products sent back to Suitland for additional Level 2 processing on existing hardware before final distribution.

A look at the available network speeds quickly showed that this was not a viable plan. Evaluating the existing test files produced by NASA indicated that the Level 1B files would be approximately one gigabyte in size and require 40 minutes of processing to reach Level 1B. Coupled with a 10mbit/sec FNS network line, having an estimated real throughput of five Mbits/sec, granules would take 30 minutes or more to be transferred from GSFC to Suitland. In the case of a one-granule orbit, this would mean that the final granule would not arrive at Suitland until just after the three-hour limit had been reached. A two-granule or three-granule orbit would be more than thirty minutes and sixty minutes late, respectively, before Level 2 processing could even begin. It was decided that as much Level 2 processing as possible should also be performed at GSFC, with only the smaller Level 2 products being sent back to Suitland for distribution via the existing CEMSCS (Central Environmental Satellite Computer System) system.

The scheduler system software was to be provided by NASA. Since it would be the same system that NASA itself was using to process MODIS data, it was expected that this would be one of the more robust segments of the system. This turned out not to be the case. The scheduler system was a mixture of Fortran, C, and shell scripts, all held together with a database implemented with SYBASE. After SYBASE was procured and installed, work turned to porting the scheduler code over to the NOAA hardware. Although NASA's original intention was to create a scheduler that would be machine independent, a great deal of work was still necessary locating and changing all the hard-coded computer names, IP addresses, and directory paths that had crept into the code over the years. Several months of effort were spent with occasional consultations with MEBS engineers in order to work out misunderstandings, subtle bugs, and a large number of configuration issues.

4. THE FIRST INCARNATION

The hardware side of the NRT system has been upgraded several times throughout its life. In late 1999 and early 2000, a RAID (Redundant Array of Inexpensive Disks) system was procured to double the storage resources and to provide a more robust environment. In what appears to be a comedy of errors, it took several months to accomplish. First the RAID disks arrived, but there was no rack to install them in. When the rack was delivered, it was an Origin 2000 rack, which was incompatible with the RAID components, and when the RAID rack arrived, the power supply system proved to be incapable of supporting the RAID configuration. Finally, though in early 2000, the entire disk package was completed and installed. The SOAP server was also upgraded to eight CPUs and eight gigabytes of internal RAM. These upgrades were deemed essential based on the new vision of producing the Level 2 data on site rather than having Level 1B data sent back to the Suitland Facility. Work progressed on obtaining reliable data access to ancillary data that would be required once the real data began. Version 1 of the NASA scheduler was installed and functioning, but delivery of the actual operations version was delayed several times, causing a severe difference of opinion to develop. While NASA was reluctant to release science software unless they were certain of its reliability, NOAA was watching the count-down to launch and becoming desperate to begin installation of anything. In December of 1999, just before the launch of Terra, the code was finally delivered and integrated, and everyone involved could breathe again.

In early 2000, while Terra was still being checked out in orbit and the data calibrated, the last software components were being completed to satisfy the original requirements. Geolocation of MODIS data is performed as part of the Level 1A step. Waiting for this step would require all 24 granules per orbit to be processed through this stage before the majority, if not all, of them could be discarded. As an alternative, orbit prediction software was located that would allow the system to know ahead of time which granules were wanted and which were not. By specifying a location, a circle is plotted around it and the orbit of the satellite is traced relative to this circle. In this way a granule whose start time or end time lay within the circle was designated for processing. All other granules were immediately deleted from the system. Two overlapping circles, centered over Pennsylvania and Colorado, were used to fully define the Continental US.

Terra processed its first-light data in February of 2000, and while the near real time (NRT) system was capable of a limited manual processing capacity, the system was still unable to keep up with anything approaching operations. At this time, the system was upgraded again to 12 CPUs and 250GB of RAID storage, drastically increasing the computing power; although with an input rate of 72 GB/day and product creation of 24 GB/day, it was still necessary to immediately delete all data as soon as generation was completed and confirmed.

As full processing began, differences in the processing environment between NASA and NOAA became even more prominent. Chief among these were problems with the ancillary data sets used by the various products. Since NASA is not a NRT processing environment, the standard methodology is to wait until all required data sets are available before beginning. Unfortunately, in the case of ancillary data, this could cause a delay of hours and sometimes days. NOAA's NRT system could not tolerate these delays. Many of the ancillary data sets are first-guess, (or analysis) data, created by summarizing yesterday's observed data, and used as a preliminary reference point as to what todays data will be. The problem using this method is that today's data will not be available until tomorrow, while for NRT processing, it is needed immediately, and sometimes sooner. As a solution, the NRT system took the position that yesterday's forecast data was better than no data at all, and simply programmed the system to use the very latest data. In early test comparisons to NASA-produced Level 1B files, less than a 0.25% difference was found to result.

Problems were also observed in the raw Rate Buffered data stream. The RBD-to-Level 0 converter program was updated numerous times to work out subtle and not-so-subtle problems in the data stream. The largest of which was the data communications problem between Terra and the groundstation. Due to a timing problem, when the satellite attempts to lock-up for data download, the first part of the MODIS data is missed. So after the entire stream was transmitted, the recorder is reset to the beginning and the first part re-transmitted. These so-called rewinds cause timing discontinuities to appear in the data so that temporally contigous data appears in separate regions of the download, and sometimes in multiple locations in the same download. A subtler problem was in simply defining exactly when the granule began and ended. It seemed obvious enough at first. The 12:00 granule would start at 12:00:00 and end at 12:05:59. In reality, however, a given packet of data, which can basically be related to a single data point of output data, consists of 1.477 seconds of data, which means that at granule could begin at 12:00:01 or could end at 12:04:58 (after rounding). A large number of granules refused to process until the converter program could be recoded to adjust for these details.

In April of 2000, the NRT system truly was born. On April 6th, a granule containing eastern China, Taiwan, and Korea became the first granule to be processed totally from RBD to Level 1B and then to be extracted in image form. Even though image generation has never been an official requirement, the NRT team felt that no project could be considered complete if it was impossible to visibly inspect the results.

5. NANOOK IS BORN

Work during the summer of 2000 consisted of mostly upgrading the various PGEs (Product Generation Executables) provided by NASA and slowly integrating them into the system to provide more of the required products. Updates to the scheduler system were also being made on a regular basis, and this, more than anything, led directly to the creation of the separate NOAA scheduler system, NANOOK. All the old problems of hard-coded names still lurked in the code and needed to be routed out with each new version. Additionally, a good deal of configuration updates needed to be made for parts of the scheduler that would never be used. For example, while NASA is tasked with producing Level 3 and Level 4 products that required archiving data products for weeks and months, NOAA is strictly an NRT shop, and has no need for a large and complex archival system. On several occasions through the summer, NASA would release new versions of the scheduler before the old version had been successfully integrated.

During the weeks around Labor Day 2000, widespread wildfires burned through Montana and Idaho. Flights by reconnaissance planes were hazardous due to the heavy smoke from the fires, causing the Fire Lab of the National Forest Service to rely heavily on satellite imagery. As the MODIS sensor is sensitive to both visible and infrared radiation, it has the capacity to detect hot flames on the surface even through smoke and clouds. Near real time access to satellite data could help the firefighting command center monitor smoke dispersion, which is a critical issue for the health of those living in the area. The production of NRT MODIS data proved to be extremely useful to the Fire Response Team in reevaluating wildfire management on a daily basis. However, manually pushing the appropriate granules every day (including weekends) indicated a real need for an automated delivery system based on the production server. However, integrating such a system with the NASA scheduler would heavily complicate an already unwieldy system. A decision was therefore made to abandon the NASA scheduler and create a simplified version that would accomplish only the tasks needed by the NRT project.

In September of 2000, a bare bones replacement scheduler was designed and implemented. The new scheduler, nicknamed NANOOK replaced the SYBASE database with three small text files to track:

1) What products had been create for each granule.
2) What granuless required processing.
3) The status of the available processing resources.

The amalgamation of shell scripts, Perl, C, and Fortran code was recreated in a set of Perl scripts which:

1) Detected products waiting,
2) Allocated resources and loaded necessary data,
3) Processed the data using the appropriate PGE,
4) Queued follow-on products.
5) Deleted products after all processing.

By the end of September, NANOOK was handling full automated processing of CONUS data through level 1B and by mid-October was producing all Level 2 products.

With added capabilities come added requirements. Alaska and Hawaii were added to the area of interest, nearly doubling the number of granules to process per day and increasing the output data volume to nearly 90GB per day. Plans to distribute data through existing NOAA channels were scrapped due to the mounting data volumes. Of course, network capacities were already a major choke point. Even if CEMSCS were capable of handling the data load, there was no way to get the data from GSFC to Suitland within the required times. Distribution was placed on the back burner while plans to install a Fiber-Optic network between GSFC and Suitland were formed, and the role of producing Level 2 products was official shifted to SOAP.

Although working well, a number of necessary additions were immediately obvious. The new version of NANOOK would need to have parallel development and production systems so that new PGEs and new code could be tested on one side without impacting on-going operations. And although distribution had been officially sidelined, with the fresh experience of the Labor Day fires in mind, an automated distribution system would definitely be needed. With this in mind the Version 2 release was coded, tested, and became operational in early December of 2000.

6. NANOOK GROWS UP

In 2006, the NPP (NPOESS Preparatory Project Satellite) will be launched as a test bed to validate the new sensors that will be launched as part of the NPOESS (National Polar-orbiting Operational Environmental Satellite System) constellation beginning in 2009. NPOESS combines NASA, NOAA, and Department of Defense meteorological programs into a single program for the entire nation. Among the instruments carried aboard NPP and NPOESS will be VIRS (Visible Infrared Imaging Radiometer Suite), the follow-on instrument to MODIS, which will have a data load approximately three times larger than MODIS.

NANOOK was therefore tasked as a risk reduction system in preparation for these new sensors. NANOOKs purpose in this regard is to identify the hardware, software, and network configurations that will allow NOAA to handle these extremely large datasets in near real time.

NOAA will use MODIS data to simulate these high volumes and multi-spectral instruments. NANOOK will then provide real benchmarks for processor performance, data communications, and storage, thus allowing NOAA to generate findings and recommendations for NPP and NPOESS planning. These findings will then allow NPP and NPOESS to use real-world experiences when making decisions for ground system and instrument science designs.

To support this risk reduction effort, NASA provided funding for the acquisition of new computing hardware designed to handle global processing of two MODIS instruments. In June of 2001, the order was placed for an Origin 3800 with sixty-four CPUs, an Origin 3400 with thirty-two CPUs (which would be used by another project for the AIRS sensor aboard AQUA), and an Origin 3200 with two CPUs. RAID storage of two terabytes were provided for each of the 3800 and 3400 for on-line processing, and a RAID with five terabytes was procured for the 3200, which was designated as the file distribution server for both MODIS and AIRS. NOAA provided a new communications hub with gigabit capacity for communicating between the servers, and ultimately for the fiber-optic network, which was still under construction. The old Origin 2000 (SOAP) was reclassified as a dedicated development system for code changes. SGI delivered the new systems to GSFC in early July of 2001, and the next set of problems began.

In 1997 when the first boxes of computer equipment for the inital configuration of SOAP arrived, NASA provided floor space and power for the machines in an area close to the EDOS cluster. The allocated area was now far too small to hold the new computers, and the room itself had been fully allocated to other projects. Over the next several months, NOAA worked with NASA, trying to find a location for the new machines that had space, power, and network capacity. By the end of 2001, NASA decided to abandon its search and simply find a completely new area. A room with appropriate free space was allocated, and work orders were generated to install power and network connections. Unfortunately, the work order was at the end of a rather long list of work orders that the electricians were already working on, and it was not until February of 2002 before all the components were in place. (See Figures 1 and 2)

Figure 1. Little Crates All n a Row

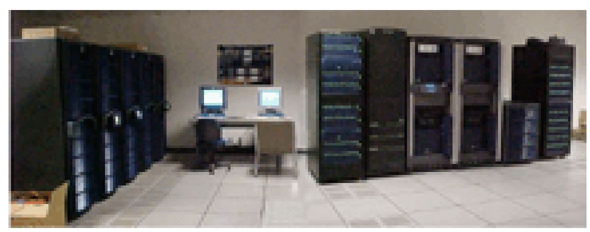

Figure 2. Unpacked and Ready for Action

Based on the experience gained when splitting the Version 1 code into separate development and production systems, and the experience gained while trying to remove all the hard-coded information from the original NASA scheduler, the Version 2 code of the NANOOK scheduler was designed for maximum portability within a SGI environment. The success of this effort was demonstrated in March when a total of three days was required to totally port and configure the processing environment to the new production server. Between the beginning of the week and the end, NANOOK moved from processing 12-24 granules per day to processing 288; from CONUS to global.

Growth happened quickly once the new server was in place. Aqua was launched in May of 2002, and NANOOK was ready for it, instantly doubling production volumes to 576 granules per day; 175 gigabytes of data being processed into 4.3 terabytes every day.

The Air Force and Navy both became very interested in acquiring MODIS data as NRT inputs into their weather models. To facilitate access, the NANOOK servers were connected to a private network allowing 100 Mbit service. Finally, NANOOK had a pipe big enough to reliably deliver data to a customer. In June of 2002, the distribution system that had been waiting since December of 2000 was activated. It promptly choked while attempting to send 20 gigabytes of products to the first customer. A couple of days of debugging intervening networks, and reconfiguring the customer's disks soon smoothed out the bumps, and data began flowing automatically. Slowly, more customers have been added and more data requested until NANOOK is now pushing nearly 500 gigabytes of data per day to various customers.

7. NANOOK TODAY

Today NANOOK continues to process in near real time, but several external factors have helped to accelerate production. First, NASA enabled the Tracking and Data Relay Satellite System (TDRSS) mode on Terra, allowing half orbits to be relayed to the ground via that TDRSS constellation. This allows for three gigabytes of data to be delivered every fifty minutes, and therefore NANOOK can get a head start on the beginning of an orbit while the end of the orbit is still being acquired. NASA also performed extensive analysis of the ground system and data networks and managed to cut nearly an hour from the time needed to send data from the ground station to NOAA. In conjunction with faster processors and faster disks, NANOOK can now distribute data to customers within two to three hours of observation (averaging 2.5), well within the requirements range.

8. SUMMARY

NANOOK has come a long way in five years. It has moved from a single crate of parts and a requirements document, to a 64 CPU mainframe with Terabytes of storage capacity. It has grown from an area of interest of 24 granules over the US, to global coverage with two satellites, processing 576 granules per day. And most importantly, it has progressed from a single operator checking logs for an error status before deleting the data that constantly threatened to overfill the data storage resources, to a small team overseeing a nearly operational production system serving hundreds of gigabytes of data to a dozen customers for real time applications. And with NPP coming up in 2006, and NPOESS in 2009, the data volumes and application uses continue to grow.

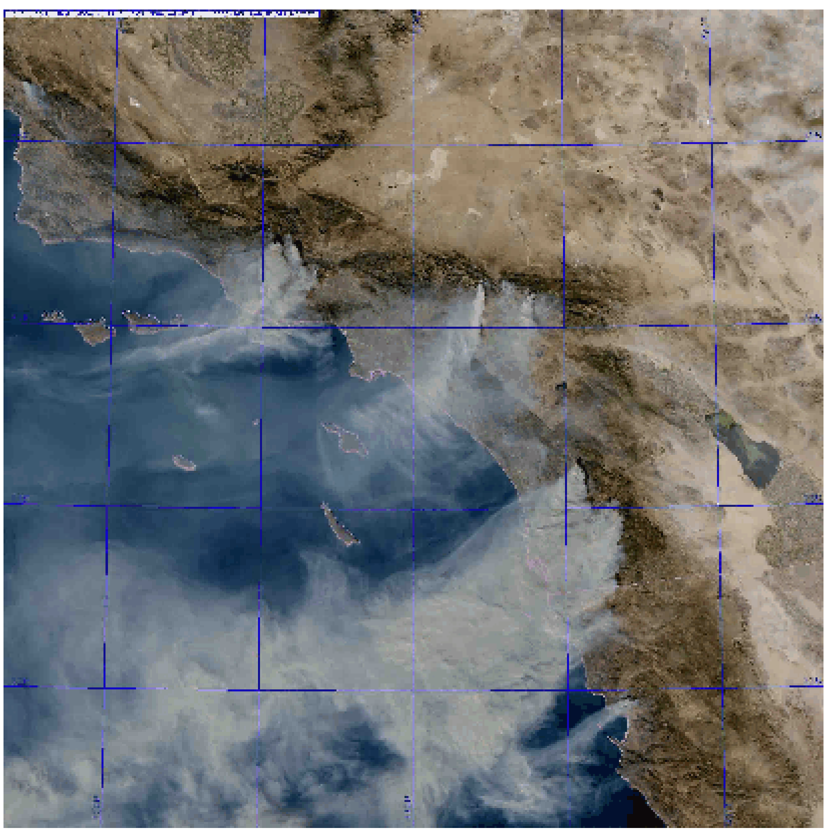

Figure 3. Smoke Plumes from California Wild Fires October 26th 2003.
Produced from NOAA/NESDIS NRT data by the Naval Research Lab in Monterey California.