Mitch's Thoughts on a Satellite Algorithm Testbed

- Environment where algorithms worthy of operational consideration can be:
  - » Tested in a repeatable manner
  - » Routinely validated against similar observations from other observing systems or satellite systems
  - » Ultimately demonstrated to meet operational performance requirements
  - » Executed routinely in near-real time with their output made available to the user community
- Establish guidelines and processes throughout NESDIS.

Algorithm Standards for Implementation

Walter Wolf[1]
[1]NOAA/NESDIS/STAR

Satellite Algorithm Test Bed Meeting
February 26, 2008

Testbed

- Possible Testbed parts:
  - » Hardware
  - » ***Processing System***
  - » Validation System
  - » Data Distribution
  - » User Readiness

Transition to Operations

- OSD, OSDPD and STAR have been working on improving the process of transitioning algorithms/systems to operations.

- A number of groups are working together and are moving toward this common goal.

SPSRB Activities

- PSDI, G-PSDI, and GIMPAP Project Plans
  - » Lead: Tom Schott
  - » This effort is to improve the project plans
  - » Possible additions are: EVM, Gantt charts, work breakdown structures, and monthly reports

- The Standards Working Group
  - » Lead: Maurice McHugh
  - » The SPSRB common standards group recommends standards to the SPSRB for adoption among the participating organizations
  - » Leads the effort to bring standards and consistency between development efforts
  - » Topics includes documentation standards, coding standards, life cycyle development (SPSRB and CMMI), and delivery standards

STAR Activities

- STAR IT Advisory Committee
  - » Lead: Ingrid Guch
  - » Organizing the efforts within STAR associated with IT Security, Data Management, Common Standards, and the Collaborative Environment
- The Data Management Group
  - » Lead: Celso Barrientos
  - » Organizes the data within STAR such that research and validation efforts have the data required. Data formats and metadata standards are addressed by this group.
- SMCD Integration Team
  - » Lead: Hank Drahos
  - » Address the issues associated with the actual transition of algorithms to operations. This team looks at the software development side and the day to day issues of transition.

STAR Collaborative Environment (CE)

- Host Algorithms In A Near Operational Environment
  - Same compilers, operating systems
  - One set of standardized documentation
  - Extensive test input & output data sets
  - Same ancillary data sets as the operational environment
  - Common software
    - Example: Community Radiative Transfer Model
  - Define benchmarks for the algorithms
- Provide Access To CE For prospective transition projects
- Data Server(s) associated with the CE where research/experimental products are made available to users

STAR Internal Activities: Software Improvements

Flexible Systems

- Microwave Integrated Retrieval System (MIRS)
  - » This system has being developed with an end to end capabability of calibrating and characterizing the radiances measured from satellite microwave instruments, and retrieving the environmental data records with the state-of-the-art algorithm science
- Operational Satellite Winds Processing System
  - » Enables winds to be generated from GOES-I/N, MODIS, MTSAT-1R, and AVHRR instruments
- CrIS/ATMS Processing System
  - » A system designed to produce sounding products using the same algorithm for all hyperspectral instruments. The data processed through this system are AIRS/AMSU/HSB, IASI/AMSU/MHS and CrIS/ATMS.
- GOES-R AWG Framework
  - » A framework where all the GOES-R AWG algorithms can be run within one program.

Improving the Process

- System developers, system integrators and scientists working together at the start of, and throughout, each of the projects:
  - The scientists develop the algorithms.
  - The developers and integrators design the system, work with operations, lead the design reviews and eventually build simulated near-real time product processing systems.

- IASI Project
  - Fully functional processing system that was running a year before launch.
  - This system was delivering data within a week of when the IASI data was released and was transitioned to operations 9 months after NOAA first received the data (IASI was a new instrument).
  - CMMI Level 3 pathfinder project within STAR:
    - Test Readiness Review
    - System Readiness Review
    - Expanded suite of documentation

Improving the System

- GOES-R Algorithm Working Group (AWG) Integration Team
- Designed a program/framework that can be used to process any type of satellite data
- New algorithms and data sources can be added easily
- Designed with the concept that it will be a plug and play system

Improving the Software

- Algorithm acceptance procedures
- Implemented software standards
- Standardized the inputs
- Fully integrated C/C++ and Fortran program
- Created code generators
- Compile and test the code on multiple platforms with different compilers
- System is under configuration management

Algorithm Standards

- Coding Standards
  - » SPSRB software standards
- Security
  - » Check for memory leaks
  - » Compilation errors
  - » Linking errors
  - » Code profiling
- Software Configuration Management
  - » ClearCase and ClearQuest

Common Tools

- Code generators
  - » Allocation
  - » Deallocation
  - » Reading
  - » Writing
  - » Copying
- Code checkers
  - » Check to meet some of the coding standards
- Code Separators
  - » A way to maintain research code within the software that is not run in the test bed or eventually in operations.

More Commons Tools

- Common Libraries
  - » Reduce development time
  - » Consistency is maintained between algorithms
  - » Contain mathematical and scientific functions
- Common Data Formats
  - » Standardized the input and output data formats
  - » Works on both little endian and big endian machines

Specifics: Interfaces

- Software Interfaces
  - » Input control
  - » Product precedence
  - » Algorithm interface into the system
- *Hardware Interface*
  - » *Scheduling*
  - » *System sizing*

Specifics: Input Control

- Standardized the input handling for the GOES-R Framework
- We have organized all the inputs into Production Control Files (PCF)
- Designed to handle product precedence

Specifics: Product Precedence

- The framework has to understand the precedence for each algorithm
- When the inputs are read in, the product information is stored
- Once the product information is stored, then the software sorts the precedence information
- This sorted precedence information is used to define the calling sequence for each product
- If a precedence is not run, then the algorithm is run using default information or a climatology

Specifics: Algorithm Interface

- The AWG Integration Team has developed a common way to implement algorithms into the framework using an interface subroutine
- The interface subroutine will either pass the data structure directly through to the algorithm, align the variables between data structures, or pass variables.
- If the variables are aligned, then the algorithm variables are pointers that point to the framework data structure variables

Specifics: Interface Subroutine Results

- This alignment of variables results in minimal impact to the product teams software development
- Algorithms can be plugged into the framework with minimum modifications
- Only issue is that if work is done on the algorithm outside the framework, then the researchers will have to maintain any interface modifications to enable a simple plug in of the algorithm into the framework on any subsequent deliveries

Algorithm Issues Addressed

- Use data from multiple time steps
- Use data from multiple data blocks/sections (swaths, granules, orbits)
- Advantage occurs if the algorithm is written to process data from multiple instruments

Summary

- OSD, OSDPD and STAR have been working towards improving the transition to operations process.
- Internally, STAR is taking steps to improve the software development part of the scientific algorithms
- Within the next couple of months, STAR will have a framework where algorithms can be plugged in to process any type of satellite data.

Extra Slides for Discussion

Testbed Definition

- Wikipedia
  - » A **testbed** is a platform for experimentation for large development projects. Testbeds allow for rigorous, transparent and replicable testing of scientific theories, computational tools, and other new technologies.
  - » The term is used across many disciplines to describe a development environment that is shielded from the hazards of testing in a *live* or production environment.
  - » In software, the hardware and software requirements are known as the testbed. This is also known as the test environment.
- Websters.com
  - » Any device, facility, or means for testing something in development

Testbed Questions

- How many types of satellite data do you want to process?
- What operating systems should it run on?
- What type of languages should be allowed?
- How do you want to run the test bed? Scheduler?

Testbed Questions

- Will there be one program on the testbed or will each algorithm have it own main program?
- If one program, who will integrate the algorithm into the test bed?
- How will product validation be organized?

Testbed Questions

- Will the hardware be sized?
- How much data will be required?
- How many algorithms will be run on the testbed?
- Will the testbed address data distribution?
- Will there be user participation with the research products?