Tactical approaches for making a successful satellite passive microwave ESDR

M. A. Hardman[1] (mhardman@nsidc.org), M. J. Brodzik[1], J. Gotberg[2], D. G. Long[2], A. Paget[2]
[1]National Snow & Ice Data Center/CIRES, University of Colorado at Boulder, [2]Microwave Earth Remote Sensing Laboratory, BYU

Project Objectives

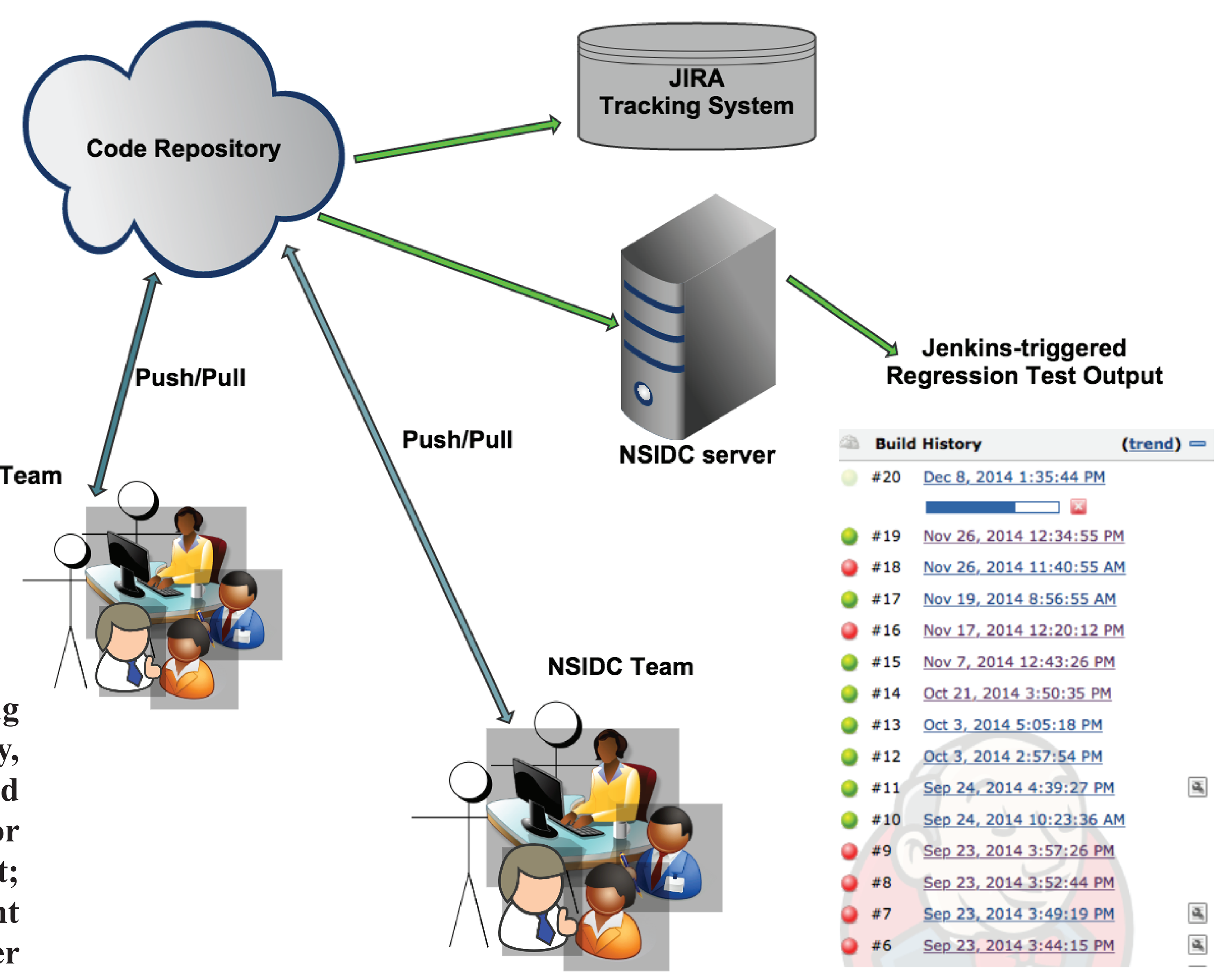
We are producing a NASA MEaSUREs Earth System Data Record (ESDR) with enhanced-resolution gridded brightness temperatures from the 36-year historical satellite passive microwave (SMMR, SSM/I-SSMIS and AMSR-E) time series.
Our project goals:

1) to produce a well-documented, consistently processed, high-quality historical record at higher spatial resolutions than have previously been available, and
2) to transition our production software to the NSIDC DAAC for ongoing processing after our project completion.

We describe our tactical approaches to some of the challenges we are facing.

Challenge: Port the legacy system from BYU to NSIDC

Our baseline production system was originally developed at BYU for research purposes. We had to port it to NSIDC to modify it to work in a production environment with additional sensors.



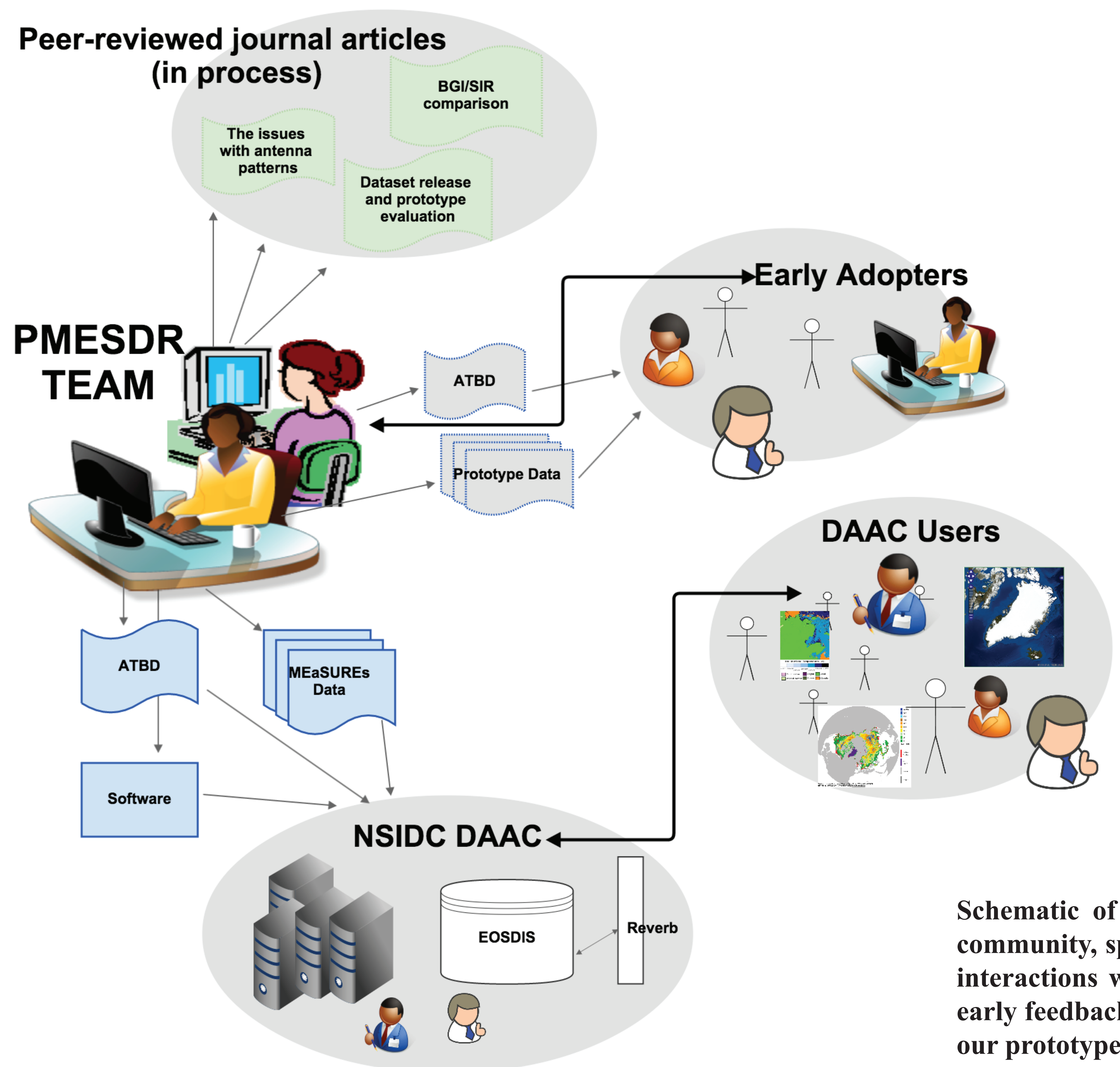
Project development environment, showing distributed teams using shared code repository, with automated regression testing that is triggered whenever code is pushed to the repository. Red or green regression status is indicated in screenshot; email to team members is automatically sent whenever test status changes. Our issue tracker also links to repository push actions.

What we have done	Why we did this
Checked software system into shared "git" code repository	• To allow all team members to have access to code history and revisions • To manage software version control
Defined baseline regression output at BYU and matched this output at NSIDC	• To ensure integrity of software results as we modify software
Automated regression testing upon any code commits to repository	• Automatic testing is more likely to succeed, because it requires no extra action by developers • Every software change is automatically tested as soon as possible to guard against unintended changes • Any unintended changes to regression output can be corrected as soon as possible after being introduced
Moved git code repository from BYU to BitBucket	• private BYU repo was difficult to administer; access for NSIDC team members was intermittent and complicated by firewalls • BitBucket is reliable and open; team members can focus on project-specific tasks
Adopted JIRA issue tracker, linked to git repository	• All project issues are documented in JIRA issues • Each git "push" action that mentions specific JIRA issues is automatically catalogued in JIRA, providing project history and rationale for each software change

2014 AGU Fall Meeting: IN31A-3711

http://nsidc.org/pmesdr

Challenge: Engage our user community

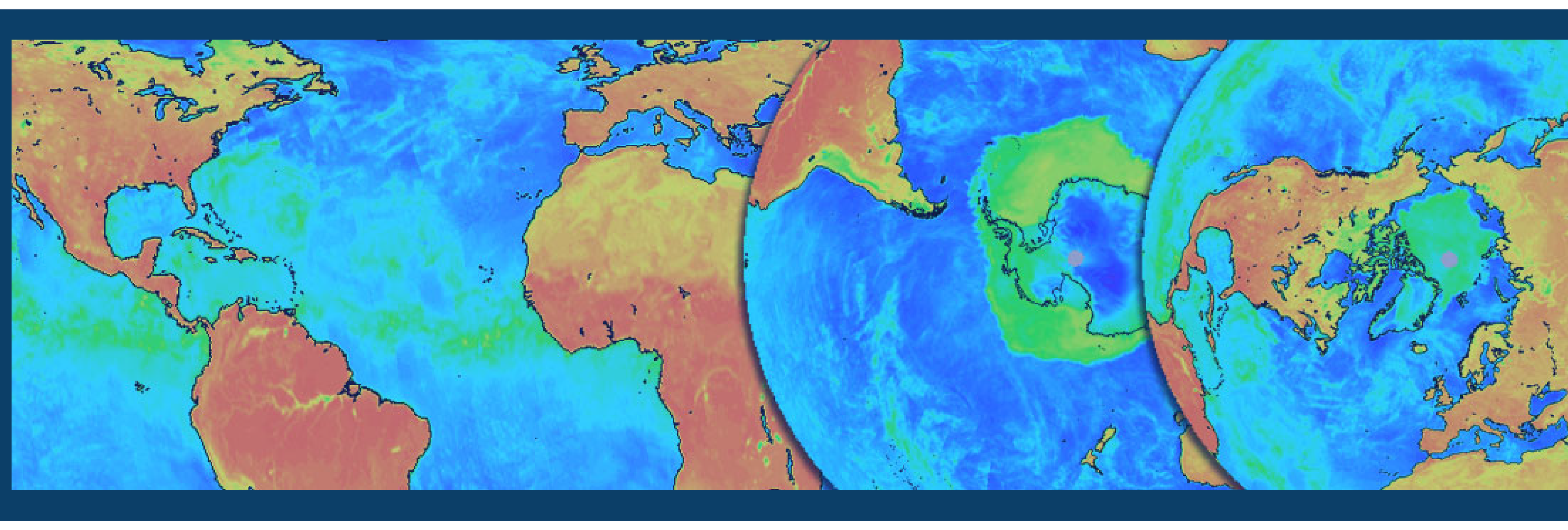


To ensure a successful ESDR, our user community needs to engage with our approach and agree that our data is the highest-quality possible. We intend for our ESDR to replace the current collection of related historical data sets; we must demonstrate why the replacement is justified.

Schematic of interactions with our user community, specifically showing expected interactions with our Early Adopters for early feedback, comment and iteration on our prototype data set.

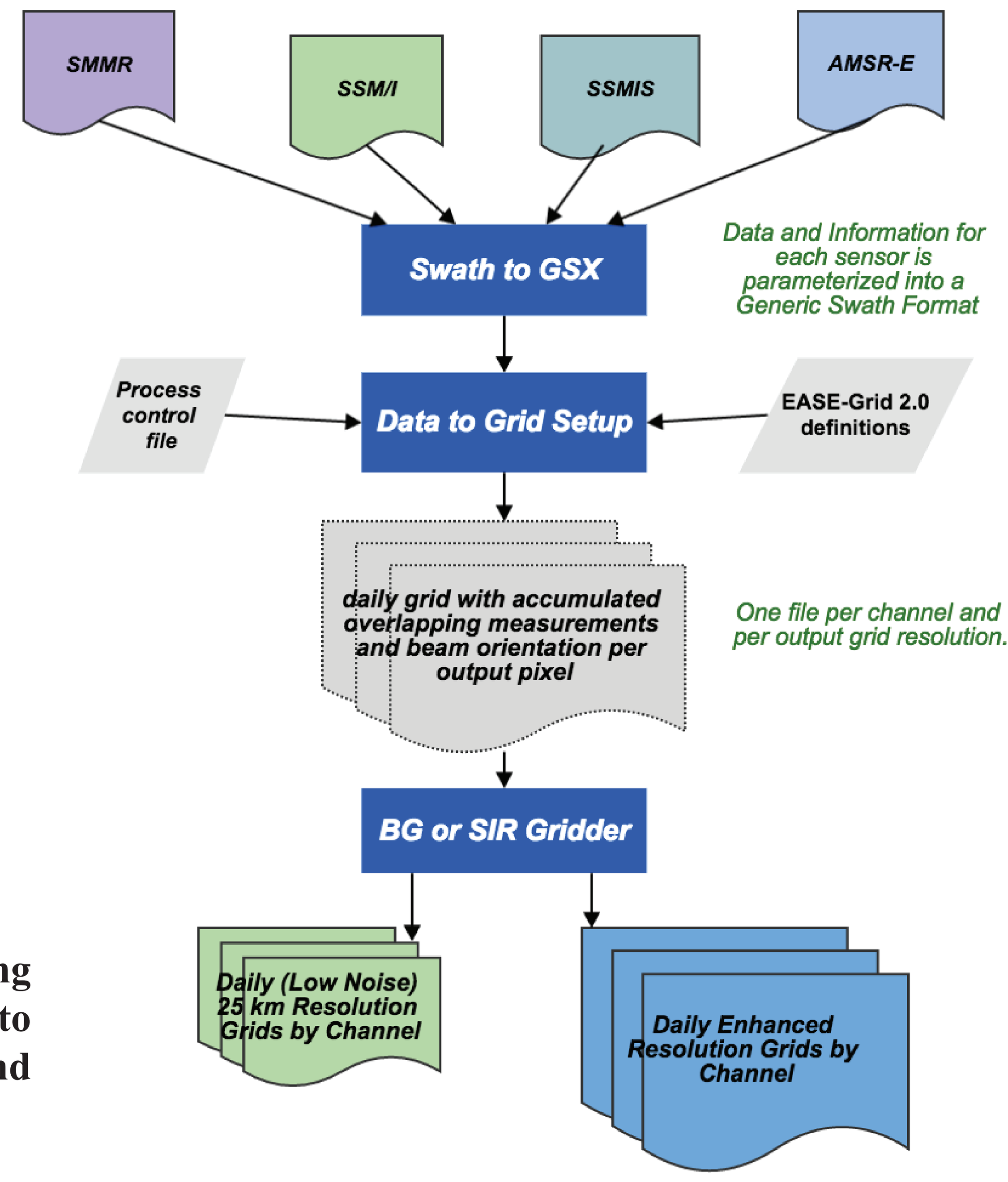
What we are doing	Why we are doing this
Inviting representative members of our user community to be Early Adopter volunteers for feedback and evaluation	• Open avenues for feedback from user community • Invite them to review and evaluate our prototype data and documentation midway through the project • Leverage their influence with the rest of the community • Hold ourselves accountable to actual user requirements • Improve quality of final data and documentation by incorporating user comments
Writing Algorithm Theoretical Basis Document	• Clarify our objectives and goals • Provide accountability and transparency for the ESDR • Document rationale for all technical decisions • Archive our technical decisions in a document that will be retained after our funding is complete
Planning peer-reviewed publications on key project decisions (image reconstruction, antenna-pattern sensitivity)	• Establish transparency • Encourage review and evaluation by experts • Engender confidence in our methods
Maintaining close relationship with NSIDC DAAC liaison	• Talk to the DAAC early and often to ensure mutually understood expectations for format specifications, data and metadata requirements • Leverage experience of DAAC staff who have long history with similar gridded products and systems • Prepare for smooth delivery of final ESDR product to DAAC • Prepare for smooth transition of production system to DAAC at end of MEaSUREs project
Planning Prototype ESDR data set release in Year 2 (early 2015)	• Provide subset of selected data to Early Adopters for feedback on data, metadata and documentation • Give Early Adopters access to new preliminary data • Plan time for user feedback to influence and improve final product • Prepare user community for changes between new ESDR and historical legacy data

This work has been funded by NASA MEaSUREs 2012.
We use the Janus supercomputer, which is supported by the National Science Foundation (award number CNS-0821794) and the University of Colorado Boulder.
The Janus supercomputer is a joint effort of the University of Colorado Boulder, the University of Colorado Denver and the National Center for Atmospheric Research.



Challenge: Modify the system for new sensor inputs

The legacy system was originally designed for scatterometry data that had been modified to process SSM/I radiometer input data; many grids were output that our project did not need. We are removing unwanted/unused software and preparing for inputs from new sensors and swath data sources. We began with a manual code review of the legacy system, to identify sections to eliminate and to redesign system input for generic swath format.



Proposed system architecture showing modifications to input interfaces to handle multiple data sources and formats.

What we are doing	Why we are doing this
Identifying hardcoded, sensor-specific values	• Make use of generic sensor abstraction and consolidate locations where sensor-specific parameters are required • Prepare for extensibility to multiple sensor input and formats • Improve maintainability of sensor-specific parameters • Determine what to parameterize in generic swath inputs from different sensors
Replacing conditional compilation in setup step with extended generic swath input model	• Replace sensor-specific hardcoded values with parameterized inputs • Replace conditional compilation in setup step, which is fragile and not easily extended to new sensors • Modularize and encapsulate the handling of input data differences with a generic swath design pattern (multiple input data sets will be converted to a generic format for input to system) • Reduce system logical complexity by separating sensor differences from image reconstruction logic
Eliminating references to output grids and/or logic we do not need	• Reduce complexity by retaining only the logic that is needed for our product • Improve maintainability

Plans

Our project faces challenges in providing the highest-quality ESDR possible and in transforming a research-quality production system into an extensible, maintainable system that will continue to produce data after our project funding is complete. These dual goals are key to organizing our project priorities.

In early 2015, we will release a prototype data set with one year of data (2003) from AMSR-E and two sources of SSM/I F13 data (both are MEaSUREs Level 2 FCDRs). We will include prototype data on the 25 km EASE-Grid 2.0 as low-noise GRD images and in enhanced-resolution images from SIR and BGI. We will ask our Early Adopters for feedback on the two reconstruction methods. Early Adopter feedback will be important in choosing an image reconstruction method for full data set production.

See also related poster by Paget et al. on our image reconstruction methods this Friday, A51I-3158.

If you are interested in volunteering to be an Early Adopter, please contact us (brodzik@nsidc.org).