Virtual Update Notes 7/21 and 7/22 2008 Report on NISAC Activities 2008

Mon, 07/07/2008 - 2:43pm — cgries [1]

Wade Sheldon presented an update on NISAC CI Implementation Plan progress.

A short intro powerpoint may be found here [2]

And the NISAC spring report is here [3]

John Porter's notes from 7/21/2008 session:

IM Converence Call - 07/21/2008

A. In attendance
1. Balsinger
2. Benson
3. one more from NTL
4. Boose
5. Vanderbilt
6. Remillard
7. Sheldon
8. San Gil
9. Walsh
10. Porter
B. Etiquette for call
1. mute when not speaking
a) but turn ON when speaking! :)
2. raise hand for recognition
3. Focus - help take care of presentations prior to IM meeting
4. contact IMEXEC if you see problems with this VTC update approach, etc.
C. Stream Chem Database
1. would be good to have Don Henshaw but he is in France until
the end of the week.....
D. Summary report of NISAC - CI Implementation
1. presentation posted on IM web site -
News/Committees/NISAC/July 2008 briefing PDF
2. some areas may be good candidates for discussion, working groups at IM meeting
3. Porter and Boose will take notes on this meeting.....
a) paste into news meeting updates on IM website

4. see presentation - most discussion here....
a) http:/intranet.lternet.edu/imcommittees/NISAC/Reports/vtc_july2008.pdf
b) summary of past process on coming up with CI cost numbers and needs
1) 1 FTE + 80K CI infrastructure
2) uncertainties in decadal plan....
c) bulleted list of tasks to advance initiatives
d) operational and governance issues
1) clearly defined framework needed for coordination of CI development
a> LNO
b> Committee framework
c> individual sites
2) much more formal way of doing network-level CI development
with defined endpoints, goals etc.
3) fund and reward site IMs for working on projects for LTERNET
4) need site commitment to NET-level CI and IM
e) basic idea - need framework and matrix of tasks
f) upshot - progress has been slow on pushing this forward
1) time constraints
2) turnover in NISAC membership
g) next steps
1) fleshing out document
2) writing up task details
3) good time for input on specific issues
h) 40+ specific tasks broken out
1) see document (initiative 1 etc.)
2) Initiative 1
a> questions
1> auditing?
a: needed so sites and data contributors get feedback for data
served via distributed servers
2> virtualization for point of presence nodes?
a: promising technology
b: need to see what technologies are best
c: may not need if everyone has own suite of services
E. Discussion
1. impressive job in drafting implementation plan
a) concern that much of the development needs to be SCIENCE DRIVEN
1) build around project
2) frustration of committee as well.....
b) pure CI budget may be easier to pursue....
1) infrastructure in absense of science questions raises
problems with LTER researchers
c) can sell general capabilities - integrating data - clear
goals palitable to users
1) "scenario" approach may be useful
2) TRENDS also useful model - would require lots of CI investment/testing
a> have working group proposal to look at architecture
b> good working scenario for cross site synthesis
2. report well done - agree with need for science drivers
a) could talk to scientists involved in TRENDS and NCEAS
workshops to see what problems they encountered
3. thanks for good work on this.... addressed important issues
a) network level interaction
b) but what are next steps?

1) perhaps more work with site scientists to frame CI need
2) site to site, cross site
3) benefits to sites for network participation
c) Those ARE the open questions!
1) multiple approaches will need to be used
2) new money would help!
a> obvious gesture to tell site leadership it is important...
b> may need to route through network rather than sites themselves
c> may need some statutory limits imposed to guarantee FTEs go to CI tasks...
3) Site review criteria was a big step forward
a> need to keep expanding its roles
4) EML provides a good example
a> sites stimulated to create EML
b> LNO played very helpful role....
4. any specifics to convey to IMEXEC?
a) group looking at governance, priority setting, implementation
b) ideas of how to fund IM at site for sustained network effort not seen before!
1) development projects have lacked formal recognition and rewards.....
2) can't do what we need to do without sustained, high level input from sites
c) different sites have different levels of commitment to NET projects.....
d) might help to send out site PI letter of commendation for people that have been helpful....
1) NISAC or IMEXEC provide site PI and NSF letters of support for work etc.
e) let PIs of each site know where there IM team stands - what is their effort towards network-level IM....
1) comparisions with other sites are valuable.....
2) IMs themselves are excellent (but some are overloaded)
3) try to spot voids and communicate those to the PIs....
f) that is a ticklish issue.... much like site reviews
1) had lots of pushback from quantitative criteria in site reviews from domain scientists on NISAC
5. suggest homework for group
a) bring to the table one need they have at the site and how it connects to the network level
b) what is your bottleneck in terms of IM?
c) what are limits on network level activities? Site activities?
d) what is science issue that drives IM/CI issue?
e) would be good to circulate straw-man list so we know what you are looking for
f) survey might be helpful at one point
g) so much already in CI Implementation Plan... need to set PRIORITIES
1) bottom up good
2) but need to decide on priorities!
6. would like to see some good prototype SOA applications
a) e.g., Personnel directory via web service
b) get something concrete for people to see....
c) LNO could do personnel database

1) need both some clients and servers
2) could share client for NBII thesaurus
d) need client as well as web service
1) something simple that works for demo!
e) service for synchronizing personnel would be valuable
f) personnel has not been accessible as data source.... but I'd love to be able to access data for personnel for inclusion in EML documents etc.
g) Drupal could also be extended to software repostitories etc. to provide real collaboration technologies
1) shared code, stylesheets - shake out better, practical system.....
2) prototype efforts would help here....

Emery Boose's notes from the 7/21/2008 session:

Background

* NISAC asked by EB last winter to draft specific components of CI plan. Specific tasks, priorities, funding sources.

* Short-term request last winter from NSF (Henry, Dan) for initial estimates of implementation costs. Corinna, Barbara, James, & co. assembled figures. Initial estimate - 1 FTE, 80K in equipment per site

* NISAC began with optimal site functionality document (Kristen, Will)

* CI plan focused on large scale, cross-site projects

* CI implementation plan should focus on general site functionality to support large projects and inter-site collaboration

* NISAC meeting at LNO, March 2008. Reduced attendance. Developed matrix of tasks for initiatives identified in decadal plan. Identified specific tasks for each initiative. Included near, mid, and long-term tasks.

* NISAC identified significant operational & governance issues that must be resolved

* Need clearly defined framework for groups, tasks, endpoints, etc. More formal setting for network-level CI development.

* Need more efficient process for developing and adopting IT standards, esp. in web services.

* Need to clarify roles of different LTER groups: IMC, NISAC, EB, etc

* Need new ways to fund and reward site IMs for working on network-level CI projects. E.g. product-oriented meetings, tech transfer, site leadership support.

* These issues need to be resolved early in the process

* How to organize CI plan to make it palatable to EB, useful as reference doc, etc?

* Progress to date has been slow. Transfer of leadership in NISAC. Conflicts with other responsibilities.

* Currently fleshing out document from March meeting. To be vetted by IMExec. Good time for input & discussion.

* See presentation slides for enumeration of specific tasks

* Initiative - service-oriented architecture

* Near-term tasks include: Web services interfaces for existing network databases. Migrate climate & hydro DB to LNO. Develop & prototype auditing systems to track use of network resources (sites need to be able to report data use, etc).

* Mid-term tasks include: Develop federated services. Develop web services schema (based on EML) for information transfer. Develop wrapping technologies that work with web services. Other middle ware resource & discovery services.

* Long-term tasks include: Develop integrated applications using web services. Develop POP (point of presence) nodes at sites. How to manage & pay for? Virtual machine technologies may facilitate theses tasks.

* Initiative - build CI technology at sites

* Near-term tasks: Improve standardization & quality of EML documents. How can we use the EML we are producing? Implement standardized keywords & units from recent efforts. Standardize access to data from individual sites. Standardize setup & management of sensor networks.

* Mid-term tasks: Develop standardized attributes. Develop automated QC procedures for high volume data streams. Develop standards for missing data and QA/QC. Develop common data models and best practices.

* Long-term tasks. Identify technology for automated metadata & data capture in field. Identify key common measurements that would support synthesis.

Discussion

* CI development needs to be science-driven. What are the driving questions?

* CI funds may be easier to obtain than science funds

* How to improve the science focus? Scientists are more interested in results (e.g., seamless integration of data) than details of CI implementation.

* Developing scenarios can be useful in communicating with scientists. E.g. the "dream information system" from years ago is still a useful target (when placed in proper context).

* Trends project is another source of information on synthesis projects. One of few synthesis projects mentioned by name in decadal plan. Good example of what synthesis can be.

* An infusion of new money (e.g. site augmentations across network) earmarked for CI could make a real difference. In the absence of new, dedicated funding, current trends will continue.

* Alternatively NSF could rewrite the LTER RFPs to require a greater allocation of funding for CI. But NSF has been reluctant to go there.

* EML was a good model for site & LNO cooperation

* Develop better cooperation between domain scientists & IT personnel

* IMExec governance group is looking at decision making process in IMC

* How to fund site IMs for sustained efforts on network-level CI projects? Little formal recognition or reward in the past. But LNO staff can not do all of this work without substantial support from the sites.

* E.g. letters of support to site PI from EB, NISAC, etc

* Let site PI know where their IM team stands (not as criticism, but as information). Site PIs react when they see site comparisons.

* Ask sites to identify an important need for their site. Are their commonalities? Are these needs science-driven? What are bottlenecks for site or network efforts? Straw man list might facilitate. More formal survey might follow.

* CIIP plan is quite full. Need to set priorities. Encourage bottom-up process, e.g. from IMC meeting

* Prototypes to demonstrate web service architecture will be critical. Difficult to conceptualize without a working prototype. E.g. network personnel database.

* Transparent business to business application is what we need. Web service with a couple of simple clients. Need to get larger group to buy in.

* E.g. a service to synchronize databases would be very useful. IM working group could create a list of candidates.

* E.g. ability to extract entries in XML format from personnel database for use in EML. Couple with simple style sheet for web display.

* How can we contribute to community technology development? Need architecture for collaborative work. Drupal website is a first step. Shared style sheets, code, etc.

Attachment	Size
Porter_notes_IM080721.txt [4]	7.27 KB
Boose_notes_IM080721.doc [5]	33.5 KB

Virtual Updates [6]

This material is based upon work supported by the National Science Foundation under Cooperative Agreement #DEB-0236154. Any opinions, findings, conclusions, or recommendations expressed in

Source URL: http://im.lternet.edu/node/199

Links:
[1] http://im.lternet.edu/user/24
[2] http://im.lternet.edu/im/news/committees/NISAC/reports/vtc_july2008
[3] http://im.lternet.edu/im/news/committees/NISAC/reports/may2008
[4] http://im.lternet.edu/sites/im.lternet.edu/files/porter_notes_IM080721.TXT
[5] http://im.lternet.edu/sites/im.lternet.edu/files/boose_notes.doc
[6] http://im.lternet.edu/taxonomy/term/169