



ITSC Meeting Chicago 10/6/2006 8:00 AM

Attendees: Jon Lea, Lelli Van Den Einde, Shannon Whitmore, Dan Wilson (via phone), Jason Hanley, Jacobo Bielak, Peter Couvares, Jerry Hajjar, Ahmed Elgamal, Phil Maechling (New Member), Cliff Roblee

Absent: Laura Lowes, Andrei Reinhorn, John Bobbitt, Sri Sritharan, Cherri Pancake, Christopher Stanton

Purpose of Meeting: To review v1.0 of NEESit's FY07 Work Plan

Side Note: This was the first ITSC meeting for Phil Maechling.

1. Jacobo asked if there were any comments on the minutes from the last two meetings. There were none and they were approved.
2. Since this was the first meeting for Phil Maechling, Jacobo asked that everyone give a brief overview of their background. Phil gave a more detailed description of his background:
 - a. He describes himself as a software developer/systems integrator.
 - b. He has a formal education in applied physics and computer science.
 - c. He has worked for small companies, in aerospace, for the military (command and control), 8 years at CalTech (seismology: Tri-Net, shake maps), in commercial work (Kaiser hospitals) and currently at SCEC (Southern California Earthquake Center) on an ITR grant as System Architect doing earthquake simulations, grid-based efforts and workflow technologies, database, SRB storage, system integration.
3. Lelli began with a summary presentation on the status of the work outlined in FY06 statement of work. She said that new versions of NEEScentral (v1.6) and RDV (v1.4) will be released soon. She gave a PowerPoint presentation detailing the status of work items in the FY06 work plan.
4. Lelli noted that the data model is getting relatively large and that NEESit would like to release v1.0 soon and then concentrate more on data usage in FY07 since it is consuming a lot of FTE time to continually add to the data model and add new GUI elements for the new data items.

Phil mentioned that metadata has been a problem at SCEC since it takes time to get consensus on what to formally have in the database and the associated development and maintenance issues. It was restated that the more metadata in the database the more GUI forms are needed. It was asked



if in fact we do have a lot of metadata. Jon asked Phil approximately how many fields there are in the SCEC data model and he said around 1500-2000. It was noted that there are a couple of hundred fields in the NEES database. It was also noted that we could minimize user's time and effort in specifying metadata by having good data ingestion tools. Lelli noted that NEESit is delivering some basic ingestion tools as part of the FY06 work and they will be enhancing and adding to those tools in FY07. There was also a short discussion on webservices (NEEScentral uses a SOA) and Phil mentioned that SCEC does not support webservices – they do support some reading of their data. It was mentioned that the idea of sharing data is relatively new.

Ahmed noted that NEESit is using the services of several data experts (e.g., Chaitan Baru of SDSC and Kincho Law of Stanford) and they have told NEESit that the NEES data model is at a point that it can be released as v1.0 and see how the community responds to it. It was discussed that as data usage grows the data model can be extended to support needed functionality.

Jon asked Phil how SCEC manages their data model and he noted that they don't have a formal process in place but rather users send in requests for new metadata; these are evaluated for general use and added if appropriate.

It was noted that NEESit is getting many, many requests from many, many different sources for additions, enhancements, etc., and it is reaching a critical mass. It was discussed that perhaps requests be channeled through a single person in each Use Case Project group (e.g., Bruce in the data use case project) in order to reduce the number of people making requests. Jon noted at his previous job they too had many requests for enhancements, etc., and that the all requests were put into a database and later evaluated by Product Engineers. Several people acknowledged that receiving many requests for new items (especially from many sources) is time consuming but is very good to have and should not be discouraged. Decisions should then be made by a few designated people.

5. Lelli continued with the status of the FY06 work and noted that a lot of good work has been done in telepresence.
6. Jon then presented the status of Phase I and Phase II work efforts for hybrid simulation. He presented the status information obtained from UCB and UIUC for Phase I (both schools are progressing well and will complete their deliverables very soon) and then moved onto the status of Phase II. He reported that a draft work plan has been produced and is under review by UCB and UIUC stakeholders. He stated that he anticipates the work plan to be finished in the next couple of weeks so it can be included in the overall NEESit FY07 work plan.



7. There was another discussion on trying to reduce the number of communication channels to NEESit. It was stated that individual BOD members are involved in using and testing the tools and they also are providing suggestions. It was noted that during the use case project WebEx meetings that there needs to be decisions made on topics before moving onto new topics. Lelli noted that it would be good if people used it.support.org for their suggestions. Ahmed reiterated that his staff is struggling with the constant stream of communication and that he wanted to acknowledge their extraordinary efforts, that he wanted to thank Jason Handley for his good work, that he thanked the ITSC for their exceptional interests and oversight and that the data model has really grown during FY06 and that we need to capitalize on it by really using the data.
8. Ahmed presented his PowerPoint presentation. He presented a slide listing several categories of work that will be present in FY07 and noted that NEESit will be similar to other NSF funded programs at SDSC. He said that NEESit (with the help of Kincho Law) will be designing and implementing a flexible data model so that they can reduce the amount of developer time on the data model; yet still provide users a mechanism to “extend” it to capture their personal metadata. He also mentioned that NEESit will have an end-to-end data usage presentation by the BOD meeting in November.
9. Jon asked Phil if SCEC has a formal data dictionary and Phil said that they don't.
10. It was discussed that NEESit is repairing their miniMOST-I and II Hybrid Simulation test beds (it was noted that a three-site test, UCB, UIUC and SDSC was successfully performed today) and that SDSC will not drive the EOT effort but they will provide the framework for it. It was also noted that NEESit will provide a collaboration web portal (they are looking at gridSphere which is used by Bern and GEON) which will start out with the basics and grow with use and time. Ahmed noted that SDSC has a lot of experience with gridSphere (it is JSR 168 compliant) so it make sense for NEESit to take a good look at what it can provide for NEES. He noted that perhaps NEEScentral will be “consumed” by gridSphere.

It was asked who develops and maintains gridSphere: students, professionals? Ahmed said that three of four of the initial developers maintain it and a commercial version of it is marketed by IBM under the name “WebSphere”.

11. Workflows then were discussed. Ahmed stated that NEESit is looking at Kepler, which is a scientific workflow application. Jon noted that NEEScentral needs to have a higher-level process workflow so that data sets can be



tracked and managed through the data conduit. Ahmed noted that if NEESit ends up using Kepler that they will need to provide some support to the Kepler program.

Phil discussed SCEC's involvement with scientific workflows. He noted that they depend on the underlying architecture of what they are addressing, e.g., grid, job submission, error handling and recovery, etc. It was stated that it is necessary to separate the workflow from the tools. It was also noted that the need for scientific workflows at this point in time for NEES is not clear. And it was noted that it needs to be used for some specific applications as opposed to being in the abstract. It was asked if Kepler is the only solution.

Ahmed responded by stating that NSF has recommended it and that they are trying to develop a demo similar to nanohub. Peter noted that nanohub helps consultants to put content on their website. Ahmed said that he and Lelli recently visited nanohub. It was noted that there is a danger that users will be confused if too many different types of sites are available. Also, it was stated that we should not separate portals, data, telepresence, hybrid simulation but rather have a complete integration of these activities. Ahmed noted that a portal is a longer-term activity but NEES will need to have one.

12. Jacobo noted that simulation was not part of the presentation and Ahmed pointed out where in the presentation it was addressed.
13. It was noted that NEESinc should drive an effort to develop and produce an e-Journal. Jacobo noted that the BOD had recommended it and Cliff stated that if NSF allows the use FY06 Category-4 dollars then it will be addressed. Jerry mentioned that we could link EQ Spectra with NEEScentral.
14. Jacobo asked about the support for experimental specimens in the current version of the NEES data model. There was a discussion on the topic and it was noted that currently there isn't a definition of a specimen in the data model and that NEES needs to have one to really reach its goals. He noted that he wants to have a description of a specimen to a degree that a FE model can be generated. He emphasized that it is very important to NEES that the database have a detailed description of the model. Peter suggested that it is ok to fix the data model at v1.0 now but continue to work towards v1.1.

Jon asked Phil how many FTEs are assigned to enhance and maintain SCEC's data model; Phil said they have just one person with that responsibility.

It was noted that there needs to be a good mechanism to gather data requirements, from any and all sources, and then Ahmed and NEESinc can



work together to prioritize activities. Ahmed reiterated that he firmly believes that data usage is the key for FY07.

15. Ahmed indicated his desire that he really would like to have a favorable response by the ITSC on NEESit's plans for FY07.
16. Phil did a quick show-and-tell on SCEC's website and portal.
17. Ahmed did quick show-and-tell on GEON's website and portal.
18. It was asked when DataTurbine will become integrated with NEEScentral. It was stated that the current priorities for DataTurbine are stability and performance and Create is helping with those activities.
19. The end-to-end data use was discussed and Lelli showed a slide depicting NEESit's proposed scheme. After some discussion it was concluded that several simpler versions of the figure shown in the slide would be good for users to see what was happening and which applications (and who is responsible for them) are involved. It was noted that someone needs to take overall responsibility for the E2E process to make sure everything gets finished by the BOD meeting. Ahmed commented that if any part of the E2E process is successful it will be a big win.
20. NEESit's proposed use of FY06 carryforward dollars was discussed. Ahmed said that they have been working with the cluster group at SDSC and it is a good opportunity to make good use of the budget. SDSC would maintain it and it would provide users "on demand" access to a dedicated super computer. Jacobo discussed that NSF really wants the super computers used by all communities and the ITSC is not in favor of NEESit pursuing a cluster at this point in time.

The ITSC members entered into executive session and NEESit personnel were recused from the discussion.

NEES personnel were brought back into the discussion. Jacobo presented the conclusions and recommendations from the executive session.

Jacobo began his presentation by commending the fine, hard work and extreme dedication by all NEESit staff.

He then continued with the recommendations.

1. The ITSC has not yet seen v1.0 of the Executive Summary version of the work plan; it needs to have it before it can make any formal recommendation



to the BOD. (It was noted that NEESit was originally asked to have this ready by 10/17/2006).

2. Concurrs with NEESit plans that highest priority is data usage in an end-to-end demonstration.
3. Next highest priority is providing users good tools to ingest metadata and data into NEEScentral.
4. Once the E2E process is demonstrated for the current Use Case Projects, it is to be done at the remaining equipment sites.
5. A parallel effort, concurrent with the above activities, is to be carried out to investigate the complexity of adding a specimen description to the data model.

Note: There was not enough time to discuss any topics other than data in the executive session.