TIC Conference Call
Tuesday, February 25, 2014
11 a.m. PT

Attendees
DW Dave Wilson      PC Phillip Crotwell
DN Doug Neuhauser   RB Relu Berlacu
DO Dave Oppenheimer  RH Renate Hartog
EY Ellen Yu          SK Sue Kientz
FC Faria Chowdhury   SZ Stephane Zuzlewski
PA Prabha Acharya    VT Valerie Thomas

Meeting Discussion
Meeting topics are summarized below, with some issues and suggestions explained in more detail.

Developer Updates

• **Anticipated Server Outage** previously reported in the distributed Agenda has changed.
  - All SIS operational machines will be up the week of March 3-7. Some issues with the A/C contractor for the server room may necessitate a brief loss of power only.
  - The Maui Trac wiki will be down for a day.
  - People will be able to upload to and view the new SIS. Development will continue.

• **PA has added a wiki page to help with loading**, which includes tips and covers known issues. See http://maui.gps.caltech.edu/SIStrac/wiki/SIS/Dataloading/ExtendedStationXML

• **A Validator is now available** for validating the XML code before it’s loaded into the database. If the code does not pass the validation test, an email is sent to the file’s submitter.

• **SZ has successfully loaded several hundred files into SIS**. PC has tried a few times, and the developers are continuing to work with him to find out what parameters or elements need to be corrected.

• **A Batch Upload Script is now available** that can load hundreds of files at one time, rather than the initial process which entailed clicking “upload” for each one.
  - The script is run on the client side and uploads the files to the loader, which then loads them into the SIS database.
  - Requires that curl Base64 be installed.
  - When you’re ready to upload, email sis-help@gps.caltech.edu and we’ll send you the script.
  - If any file does not validate or load, it sends an e-mail (per file) to the submitter. So be careful; don’t try to upload 100 invalid files, as you will get 100 error emails.
  - A summary e-mail is also sent to the submitter at the end of the day, summarizing the upload activity.

• **The XML Writer is almost done.** In approximately a week, SZ will test it with the AQMS loader. SZ reports that they have not started on the Reader (AQMS Loader).

• **We are working on the Extended XML at this time.** After they finish with this, then they will start working on
the Standard or FDSN Station XML. The Extended version has all the equipment information, and the FDSN will also be offered.

• **We will be relying on IRIS’s XML Converter to Dataless SEED.** That’s where the Standard FDSN XML will be needed. The IRIS converter can be found at:
  - https://seiscode.iris.washington.edu/projects/stationxml-converter
  - https://seiscode.iris.washington.edu/projects/stationxml-converter/wiki (there are some critical caveats in the conversion, documented on this page)

• **Read-only IRIS-NRL Response Lookups, as loaded into the SIS database, are now available.** They are found in the Equipment section at http://anss-sis.scsn.org/sistest/equipment/; scroll to Master Table Views for links.
  - There was a human-driven process to map our model names to the IRIS model names, where we couldn’t just use the IRIS model names, e.g., Q330-S. It has the same response as the Q330 so we mapped it to the existing Q330 response in the NRL, so this is a case where there are more models in SIS than there are in the NRL, and in some cases they map to the same response as another model. This is something to keep in mind as we add more models to SIS and the NRL.

• **Some Admin-Level Forms now available.** They will make more sense when the non-Admin level files become available.
  - There are three classes of users: Super User, Admin, and Regular User
  - Certain things are available to Admin users. When you become a user, state whether you are the Admin user for your RSN.
  - We can decide what actions should be available for Admin Users, e.g., Creating a new Model is an admin action, but Creating New Equipment will not be an admin action.

• **SZ: Did you solve the problem with the concurrency issue when loading multiple XML files?**
  - FC: I set the concurrency of the job to 1 and restarted the process. It’s using Rabbit MQ, that’s the job queue software that’s running that. There are other options I must research, like rate limit. That way it doesn’t overload the database when there are a lot of uploads. SZ’s mass upload of 800+ files didn’t impact the Development SIS interface, but that’s because no one else was using this version. We need to address that possible issue of an upload causing performance issues for other users of the UI.

• **SZ: Can I remove the sleep statement?**
  - FC: Keep the sleep 3 for now. We have to figure out whether we need the client to do rate limiting, or if this can be done purely on the server side.

**How Do You Anticipate Using the Project and Owner Fields?**

Some questions have arisen on how one might identify stations that might be part of a particular temporary deployment, versus stations or equipment involved in a particular project from a fiscal point of view. Perhaps there are additional requirements we have not considered, and this might lead RSNs to use these fields in ways that temporarily help their needs but long-term might cause problems. We may need additional fields to keep this potential “abuse” of these fields from occurring.

That said, whatever solution or plan we come up with would not be included in the upcoming version of SIS. With this group, however, we want to figure out how “Project” and “Owner” are planned to be used. It may be that we don’t know the answer to this now, and RSNs will have to start using SIS to discover any new needs or issues.

**Owner:** The Entity that owns the site/station/equipment
**Project:** Certain contracts want their equipment want to be able to track it into the future. On the government side, there was the ARRA project, the MHDP project through the USGS; on the Caltech side, we have the UASI project coming through. In some sense, it’s a desire to keep track of which equipment was purchased with which funds, so that 10 years from now, when whoever bought that equipment wants to collect it all or find out where it is, they can look it up.

Some people might see “Project” as a place to identify equipment/stations that are deployed in an area that just had an earthquake, for an “aftershock project.” In this case, it’s not a funding name; it’s rather tied to the name of the earthquake, or tied to the locality where a temporary deployment is going in.

So this appears to be two ways of using “Project,” one a scientific way, and the other a fiscal way. What other projects might be out there? Do we want to possibly split this field into two, give them different names? Are there any other applications or ways you might want to track your equipment? Fiscally, it makes a lot of sense to track equipment under a fiscal project; I don’t know whether it makes sense to track a station under a fiscal project. Maybe temporary deployments, the project-ness of them, is a quality of the station itself and not that of its equipment. Or do we care about this issue?

**Discussion**

EY: At Caltech people sometimes want to track things by a field project, e.g., we’re converting a number of our analog stations to Basalts, and people want to see the progress of that.

RH: We do something like that. We make station tables on our wiki of stations that need to be upgraded. Maybe this is one of those settings.

DO: Are we talking about just one attribute? This sounds like tags to me. DN: But tags could potentially be used for multiple things, and you don’t know what the category of the tag is, e.g., is this a fiscal tag or a project tag.

VT: Would project tag be at the station level, and fiscal tag be at the equipment level? DO: That’s the nice thing about tags – you can do whatever you want with them.

PA: Currently we can tag a site and tag equipment. The site tags have categories, but for equipment the tags are at the model level. We would have to add another level there. There are also sub-categories, e.g., for sites, what kind of site is it, seismic or GPS or waypoint. Those values are for “site-type.” In the equipment models, one tag type is “measurement type,” and takes values like Velocity Transducer, Pressure Transducer. The tag values are not restricted; anyone can create new tags. Those categories we would like to keep limited, so reporting becomes easier. We can add new categories, but it has to be with consensus, so we don’t inadvertently duplicate tags.

FC: While we don’t have tagging at the equipment level, we do have settings. Like “Firmware revision,” which people like to track for loggers.

DN: Something I’ve been asked to track that I haven’t been able to do very well is, when my boss asks me that he want to see how much data are people collecting using our Broadbunds, and how much data are people using from our bore holes. Sometimes we have broadbands down in bore holes. That’s a categorization problem I don’t know how to solve. It’s a tagging on a per-channel basis. Suggestions?

VT: So it has to be on a per-channel basis, because that’s how people request the data.
DN explains how some stations are mixtures of several sensors and it’s difficult to separate the data streams to provide such requested information. VT points out that you can’t really do this at the sensor level because we track our data based on the SNCL that has been fed out of the database. It’s a SNCL tagging problem. DN: I have to be able to differentiate SNCLs at a site. By a channel level, I mean at the SNCL level. FC: Couldn’t this be derived from the channel code and the placement depth? DN explains why he can’t necessarily do that.

DN: If we’re talking about tagging sites and tagging equipment, I’m suggesting we also consider tagging channels as it could be useful.

FC: The difference between Equipment settings and tags is equipment settings is a tag plus ondate and offdate, because there was a need to track when a setting went into effect. If you want to tag at the channel level, did you want every channel to have an ondate and offdate, or are we going to keep it simple, just key and value?

PC wonders if Comments are different from Tags. PA: Comments have a date range but tags do not support a date feature, just a key-value pair. You can have as many tags as you want. A station can be tagged seismic, GPS and telemetry. FC: Tags are also text. Don’t expect to store numbers and expect a numeric sorting.

DN: My channel tagging request would have to include a time range. PA: You could tag the channel epoch? DN: I might want to maintain that separately. That seems like a lot of work. PA says we can think about this, whether we need this channel tagging involving just the SNCL or at the channel epoch level.

PA: As for comments, they can be added at the channel level and have their own date. Comments can be added at the station level, and there are comments available at each channel epoch, but then there are no separate dates – one comment per channel epoch. At the equipment level, for each equipment epoch, you can add one comment. This is separate from the logs that can be added to equipment, that have their own date range. So we have two types of comments: one that’s like a log, and you can attach more than one, at whatever level (station, channel, or equipment), and we have another type of comment that has one row in the database. So it’s just another comment. FC: If people want to search on comments, it’s going to be slow.

EY attempts a summary of the discussion and apparent consensus, but DN does not agree that Project might be reserved for specifically fiscal project designations. VT offers some examples as to why we might limit Project to fiscal identification. She proposes that “Project at the Equipment level is tied to money, and Projects that aren’t money-related get indicated at the station level.” Discussion continued with many examples of exceptions offered, but no clear-cut consensus was reached on a definition of Project.

**In Summary**

The Project field as currently implemented may be too narrow. The question is, do we broaden Project itself or do we find another place to handle the additional attributes. Currently, Project contains fiscal information.

While currently we have tagging for models and stations, some TIC members suggest it may be useful to tag equipment and channels, e.g., this is a bore-hole channel, and this is a state-of-health channel. These are not features that will show up immediately, but they are certainly worth considering. It’s agreed that such a scheme would be completely human driven. Some business rules or guidelines should be generated to recommend how such tags could be used. Every network approaches things in their own unique way, but there are some things we can all agree on for basic tagging. Then within networks, there could be free-form tagging, in a way that makes sense just to that network.
Business Rules for SIS Naming Conventions

The group agreed that the current naming convention business rules look fine. The basic idea is to look at what the manufacturer named the item, and then to adhere as much as possible to that name.

What to do when new items go into SIS before the NRL?

We’ve run into a few instances where pieces of equipment have gone into SIS but have yet to be added into the NRL. Mary seems to be amenable to the idea that if SIS gets there first, as long as we’re all following the same exact convention that the NRL is, that the NRL will adopt the name used in SIS.

DN doesn’t believe we can rely on that in the future, in that there will be too many people involved. VT agrees that that can become a problem, but perhaps only one person per RSN should be authorized to add new model names. Anyone can add new equipment, as long as the model already exists in SIS. VT suggests that when a new model comes along, it might be wise for the RSN involved to send e-mail to the others, to say “I’ve got a new thingie here and am going to call it a Fritz-Natz Thingie.” We are open to ideas on how to control the growth of model names.

DN wonders if a large network comes into SIS, which has its own variations of model names, it will be hard to deny them using those model names when they need to upload their data. VT figures we should have some protection if we are using manufacturer’s official names; it will be perhaps with items like thermometers and micro-barographs that we would have problems.

Portable Instrumentation and How to Track

Per earlier discussion, VT assumes that “portable” could become a tag that would allow searching for portable stations or instrumentation to be effected. It’s unsure how quickly this could be put in place, as tagging at the equipment level does not exist yet.

PC: How to you know what might already be deployed? VT: The big pieces would be easy to track, if it’s in inventory versus installed. “Station groupings” is a new idea, that a full stations-worth would be a ‘yes’ on the portable list. It’s easy to identify individual pieces of equipment as portable, but we’re going to have to think about the ‘groupings’ idea.

PC: Maybe a script could be created that could match together disconnected items tagged for portable deployment, into a grouped set of instrumentation. DN: But SIS doesn’t monitor cables.

DW: Here at ASL, we put together fake stations, give it a name like “Box #12,” and then give it a real name and location when we deploy it. But before deployment, the equipment is put together in a box, zip-tied, and ready to ship. DN: What if equipment in there needs to be powered? DW: You can also prebuild the metadata in there as well.

The next meeting is March 18, 2014, at 11 a.m., same phone information as always.

If you have something you’d like to discuss at next meeting, send Valerie Thomas a note by March 13.

Accessing the Meeting Recording

If you wish, you can listen to the recording of the meeting by calling 1-703-648-4848, conf. code 38103. After you enter into an empty meeting, press *3. You will be prompted for the meeting recording code, which is 0019.
When prompted, jump 1 minute into the recording to get to the very start of the meeting. To stop playback, press *3.

You cannot pause or reverse/fast forward the playback. You can only stop playback and start it again (*3), then jump to a later minute.