GSN Strategy for the Next Decade

Rhett Butler – GSN Program Manager
White paper presented to GSN Standing Committee
4/5/2010

The Global Seismographic Network is beginning the decade of quality. The network is in place, and has captured with full-fidelity the third and fifth largest earthquakes ever measured since the dawn of instrumental seismology. GSN real-time data are used at more than ten-fold their acquisition rate, flowing to the worldwide scientific community and to tsunami warning, earthquake hazard response, and nuclear treaty monitoring facilities. Augmentation funding to GSN from NSF and USGS has enabled the GSN to capitalize the standardization and modernization of its data acquisition and sensor systems across the network, and to accelerate their deployment. Operationally, the GSN envisioned in the mid-1980s is now in place. The new dimension for growth is quality.

The GSN started with quality. New equipment, evaluated and tested, met GSN design goals. Sustaining our fleet of equipment and infrastructure with its spectrum of ages requires a continuous engagement in monitoring the state of the network. To make the existing network in place even better requires a new approach to data quality. The framework exists. GSN already performs absolute calibrate in the field, and calibration via telemetry—an essential feature of the new, standard data acquisition systems now being deployed. Information about the state of GSN data—quality metrics—must be on a par with the data themselves. This information must be as routinely, easily available to the user community as it is for the network operators themselves. When the GSN started, written notes about its sensors were stored in file cabinets and data were shipped weekly on magnetic tapes; today, with the Internet we can distribute data in real-time, and should post all knowledge about the state of sensor data on GSN websites. Confidence in GSN data requires knowledge of their quality.

Growth in quality will build on the framework in place for the network, including ASL and IDA Network Operators, USGS and IDA Data Collection Centers, the IRIS Data Management System, and GSN management, together with the help of the user community, as exemplified by the Lamont Waveform Quality Center, and GSN and DMS Standing Committees. Everyone in this framework already realizes the importance of data quality. Progress will be made through working together for this common goal, and increasing efforts to improve data quality and its communication to the user community.

Focusing on quality as the primary goal naturally includes prior efforts emphasizing data availability, for when a sensor is down all of its data quality metrics may be considered as zero. However, the focus on quality permeates the approach to the network and to data flow. Practices that improve data quality are encouraged and adopted, whereas those that don't are reevaluated or discontinued. Such an approach creates natural momentum toward quality.

In building quality as a high priority dimension for the GSN, the foundation is based upon the extensive network infrastructure today. A comprehensive strategy forward starts with this foundation. Increasing efforts toward a sustainable high-quality network is the clear objective.

Efficiency and the NGS

With the Federal stimulus funding of 2009, the GSN has been able to leap from the operation and maintenance of an established, aging network into a new paradigm where upgrading and sustaining a new generation of standard instrumentation is now possible. The next generation data acquisition system (NGS) now being installed throughout the GSN (25 sites in 2009, 58 total to date) is the cornerstone for achieving new efficiencies in network operations and improved data quality.

The NGS brings four major core improvements: standardization between ASL and IDA, low-power for improved reliability in remote locations, remote calibration, and lower cost. Each old data acquisition system replaced by the NGS frees both personnel and equipment resources previously tied down to obsolete equipment, unsupported by manufacturers, repaired with cannibalized parts. By utilizing common Q330HRs, power and sensor interface equipment, and sensors, ASL and IDA can now develop and share system knowledge, work together, and provide each other with assistance in GSN maintenance. Already, common equipment has been swapped between network operators to assure continuity of operations during delays in vendor shipments. Repairing NGSs may be outsourced to Quanterra for the coming decade, relieving the burden on inhouse engineering staff. The low-power utilization of the NGS will improve data availability at remote installations where back-up power is challenging. Remote calibration capability via the telemetry link is essential for routine monitoring of sensor state-of-health, documenting the sensor response, and tracking data quality. As each NGS upgrade (and subsequent site visit) is associated with an absolute field calibration, location, and azimuth analysis, the remote calibration capability provides for regular, routine assessment of the data quality.

In order to realize the efficiencies and concomitant improvements in data quality from the NGS standardization across the GSN, the most important activity is to complete the NGS upgrade of the whole GSN as soon as possible. Therefore, the GSN must not decelerate from its current upgrade plan with IDA completing the II sites by 2013, and the ASL completing the IU/IC sites by late 2014 (with some contingency for working with Russia and China). In order to sustain this upgrade rate, it is important that IDA and ASL upgrade their respective, current suite of stations. Switching stations between the network operators would require negotiations with station hosts and transfer of station-specific knowledge, and would dilute the reservoir of goodwill built-up over the years between the network operator and station host. Therefore, in order expedite the efficiencies to be gained from the NGS, there should be no change in the relative station configuration between the II and IU networks of GSN at this point. This view is fully supported in discussions by USGS, IDA, and IRIS GSN management.

In the time frame following the completion of the NGS upgrades and network standardization (2015-2020), the opportunity for re-structuring the network will naturally begin. Bolstered by several years in collaboration with standard equipment, ASL and IDA will have developed confidence in coordinating together on maintenance at select remote sites. Re-structuring can take advantage of logistics (grouping sites in a region), politics (university versus government MoUs), and funding (both NSF/IRIS and USGS).

Such re-structuring may also take place in the context of closing/relocating stations, or adding new stations to the GSN (as unique opportunities arise). Further, large national partners (e.g., Russia and Australia) may then be amenable to change when GSN equipment is standard and up to date across their territory. On a single station, case-bycase basis,

international hosts may wish to take greater responsibilities, or even full responsibility, for local GSN stations. These arrangements and discussions will be constructively met, and encouraged insofar as the quality of the GSN data may be assured.
To sustain the current NGS roll-out rate, GSN will continue to support base personnel levels for IRIS/IDA and USGS/ASL staff through the 27-month proposal. By reducing the burden for maintaining obsolete equipment, the productivity of our GSN field staff with the new, standard equipment will allow for a shift in emphasis towards an improved focus on data quality for the whole network. This personnel-efficiency gain further underscores the fundamental importance for completing NGS roll-out expeditiously. Staff support added in 2009 for the NGS roll-out will be reviewed following completion (2013-14), and may be released or re-engaged to quality assurance roles.

Sensors

Good-working sensors are the crux of data quality. The Federal stimulus funding of 2009 enabled the GSN to purchase the sensors available for immediate problem areas. These include replacement E300 electronics for the aging STS-1, and secondary sensors (STS-2, T240) to ensure working, redundant broadband sensors at all GSN sites, microbarographs, and upgraded strong-motion sensors. There is no replacement yet for the STS-1 mechanical assembly, nor is there a reliable borehole sensor available to replace the obsolete KS54000.

Therefore in the coming years to sustain the network and maintain high data quality, the GSN must keep abreast of sensor repairs and replacements—falling behind leads to situations where for want a sensor, stations are down or data quality is poor. As part of its O&M review in 2008, the GSN determined mean-time-to-replacement (MTBR) rates for all GSN sensors, based on more than 6,000 sensor-years of actual GSN field experience. Consequently, the GSN can predict with good accuracy the necessary numbers of sensors that must be procured or repaired per year to meet replacement needs. These MTBRs will be updated regularly to maintain adequate projections forward, especially for those sensors or components where we have limited field experience (T-240s, Metrozet E-300s, new microbarographs, etc).

Until replacement primary vault or borehole sensors are available, the GSN will replace defective, un-repairable units with the STS-2/3 or T240. Funding for repair of KS54000s must be continued. In the coming years the GSN must invest in prototype-replacements for the primary sensors, both procurement and testing/evaluation. This is necessary to engage the manufacturing/university community, and to provide for options forward. It is unlikely that this process will converge to successful candidates until the next IRIS proposal in 2013. At that time, IRIS/USGS could review opportunities for a large funding increment for a bulk purchase. Without such an increment, the GSN will need to replace units based upon the projected MTBR rate of the KS54000 and STS-1/E300, plus units necessary for sites operating only the secondary broadband sensor.

Through the coming decade and beyond, sustaining the data quality of the GSN will require continuous repair and replacement of sensors as they age, are damaged, or become obsolete. GSN measured MTBR rates enable the network to accurately project these needs.

NGS Future

For the NGS the situation is different. The manufacturing lifetime of a DAS is about ten years, after which time the manufacturer discontinues the product line (the original components become impossible to obtain) and no longer supports repairs. This has been observed in the past in the GSN and was quoted by the vendor as the expected manufacturing life span of the NGS DAS. In this transition period the GSN must maintain and repair units internally, and may resort to cannibalization (as seen before the NGS). Sufficient spares need to be available for this transition period. Assuming a life-cycle of about 12 years, our initial installation of NGS in 2008 will be followed by a renewal in about 2020. Therefore in the 2016-17 time frame, the GSN will need to review its design goals, draft new specs, and begin the procurement process for the "next NGS." This would also be the time frame to build a foundation for a large funding increment to assure that a large, standardized purchase can be made. Looking forward, given trends observed to date in the DAS equipment and peripheral interface equipment, it is likely that the "next NGS" will cost substantially less than the current system, and have even better performance. Inasmuch as the GSN data acquisition system is the heart of the station's body of sensors, keeping this essential component efficiently functioning is the sine qua non of a station's data quality.

Infrastructure

The GSN station infrastructure affects data quality. A state-of-the-art, well-calibrated sensor on a noisy pier yields noisy data. Problematic piers and boreholes, aging buildings, etc—all must be maintained. During the NGS roll-out the network operators are performing a systematic assay of station infrastructure issues. Some are being addressed immediately and as necessary to complete the NGS upgrade. For instance, applying new, state-of-the-art installation techniques subtantially improves the quality of the secondary sensors. However, most infrastructure work is being deferred until after the upgrade cycle is completed. Therefore, starting in about 2014, the GSN will need to commence a systematic approach toward improving/upgrading station infrastructure in order to assure long-term station viability and data quality. Sensors and stations that cannot be maintained to quality standards may need to be re-located or replaced or closed altogether.

Data Quality

The transition in the GSN from performance focused primarily on data availability to data quality began several years ago. This is evidenced by the development of absolute field calibration methods, improved secondary sensory installation techniques to enhance their noise performance (to serve as a "true" back-up for the primary sensor) and augmentation funding for a large procurement of Metrozet E300s (developed with IRIS GSN funding) for the STS-1 problems.

Repairing/replacing sensors with data quality problems are an essential part of the equation for making the GSN a high-quality network. This work is being conducted as part of the NGS roll-out, with trips of opportunity for emergency repair as can be expedited. GSN has in place a program for station calibration both in absolute terms in the field, and relative terms via telemetry—in 2009 there were 27 absolute and 66 relative calibration events, at all sites with telemetry links. Calibration events include measurements immediately before and after all major instrumentation changes. This yearly program is being rigorously followed, with additional focus on sites with apparent temporal stability issues. GSN is also reviewing its calibration policy and will coordinate with the WQC to weigh the advantages of more frequent, lower-intrusive step calibrations versus the potential impact to continuous, ultra-long-period, time-series data. Sensor metadata have been routinely updated by IDA, and this practice is now being initiated by ASL, along with a full review of prior calibration history.

The Lamont Waveform Quality Center (WQC) has noted significant lapses in past calibration practice (sites not visited often enough for good calibration, sensors operating out of specification, metadata not updated). The GSN Operations Group has prepared specific responses to these reports, and is prepared to continue to work with WQC in answering historical data quality issues. WQC has brought to bear on GSN data a number of data quality measures and methods that will be adopted by GSN and brought into its quality-control regime. This will require additional personnel, both at IDA and ASL to fully implement these methods. These plans are now being made. An additional QC staff position is being requested this year and for the coming proposal for pan-GSN data quality control, which will be a coordinated position between GSN and DMS at the IDA DCC. ASL is currently re-bidding its Honeywell services contract, but it is anticipated that additional USGS QC support will be forthcoming, both at ASL and in collaboration with QC efforts taking place at the USGS data collection center at NEIC. These additional QC staff will also review historical GSN data quality, to assure that the data and metadata in the archive are correct, appropriately flagging anomalies, and openly publishing the status on a GSN/DMS website. Data problem reports, which have lapsed at the DCCs recently, will be revised and renewed with end-to-end tracking to resolution, in conjunction with network field engineering staff.

Discussions have already initiated in the Network Operations group for developing GSN data quality metrics, which assess the variance of sensor data from our published Design Goals, and include timing, noise level, linearity, calibration accuracy, and orientation. In designing the metrics, care must be taken to balance the effects such that reasonable changes in a given measure correspond to another measure in a meaningful way (normalization). Thinking of the process as a traffic light, the design goals clearly represent "green" but setting the levels for "yellow" and "red" require care and definition with respect to the use of the data for science and for operations. For the metrics to be routine, they must be automated. To be of general use for the GSN they must also be uniform across the GSN.

Metrics need to be defined and evaluated as soon as possible, and then automated. Web-published metrics not only offer a clear status and history of sensor data quality for the scientist using the data, but also better enable GSN network operators to monitor quality, to bring engineering expertise to problems identified, and for making decisions on the allocation of resources for field trips. In principle, like the DMS BUD webpage showing real-time telemetry status, a uniform GSN QC page could display color-coded quality status information for stations, which could be drilled down for channel information and detail. This information would also be stored in the DMS as a historical record of the state of data quality. Whatever the eventual form and presentation of the information, it is important to present to the scientific and user communities the status of GSN data and sensors as clearly as possible. Close coordination with and help from DMS web-personnel, in collaboration with ASL and IDA, will enable the GSN to publish a wide suite of essential information related to GSN data quality on the web.

Ongoing QC efforts at the IRIS DMS (data gaps and overlap, availability, RMS signal, PDF spectra, latency, STA/LTA ratio) provide a framework for adding new GSN QC processes being developed. Mining the existing DMS QC/PDF dataset potentially offers additional information on historical GSN data quality.

Data quality is essential to the GSN and is a primary concern of the IRIS Data Management System. The IRIS GSN and DMS Program Managers will work together to coordinate their programs to insure high quality GSN data.Within the IRIS GSN Management structure, the GSN Operations Manager will coordinate implementation with DMS efforts, and as chair of the GSN Operations group will coordinate implementation with USGS and IDA efforts. GSN will actively engage the participation of WQC in these efforts.

The Data Collections Centers were originally under the GSN program in IRIS, but were moved under DMS in 1990 in order to improve the flow of data into the DMS. As the DCCs serve a central role in data quality control for the GSN, it may be appropriate to restructure the oversight of this resource to include GSN more directly. The quality control analysts at each of the DCCs are clearly as essential to efficient field operations as they are to defining the metadata and assuring the effective quality data flow through to the DMS. In this context, it would be appropriate for the GSN Operations Manager to share oversight responsibilities with the DMS for the DCCs regarding quality control aspects, and participate in DMS/DCC meetings.

GSN is on the quality path—this is where the GSN started. Instrumentation and methods now available are already being installed and implemented to improve quality. With funding proposed for additional GSN QC personnel in collaboration with current DMS/DCC efforts this year, GSN data quality will advance. A fundamental step is developing and implementing automated data quality metrics. The GSN is a champion of open data, and must also embody the principle of open information regarding its data quality. Here GSN leadership can potentially lead to improved data quality practices beyond the GSN, a worthy goal in itself.