USE OF THE CLOUD FOR MT

Replacing pesky humans with robots!

National Geoelectromagnetic Facility

net.oregonstate.edu
Driving Question

How can the “Cloud” improve process flow and minimize impacts on field operations due to failures in data transportation?
Data Workflow – Prior to Cloud storage

Collect data, metadata field QA/QC daily production reports

FTP to OSU MT server

FedEx Memory Cards to OSU

Data QA/QC

Data processing MT response functions mSEED, SPUD

Upload to IRISDMC
Data Workflow – Using the Cloud

1. Collect data, metadata field QA/QC daily production reports
2. Auto sync to Dropbox cloud
3. Upload to OSU NGF NAS
4. Autosync from NGF NAS 1 to NGF NAS 2
5. Data QA/QC
6. Data processing MT response functions mSEED, SPUD
7. Manual upload to AWS Amazon Glacier
8. Upload to IRISDMC
Data Transmission and Quality Control Process – KISS

- Our field crews are heterogeneous – so we keep things simple and automated
- Instrument control is through easy to use LabView interface

It is nearly impossible for field crews to start a data collection run without entering complete metadata and without confirming that all sensors are operating correctly.
During site visits, the field process the MT data in the field using a single station robust processing algorithm. This lets the field crew know, at the field site, if the data quality has met requirements, or if the site needs reconfiguration and if an additional data run is required.
Why Dropbox? (we’re agnostic about which Cloud service; for now this one works well, but we can easily switch)

- Robust servers
- File deletion protection
- Versioning controls
- Dropbox software verifies file integrity
- Robust against often slow hotel/MiFi internet connections
NGF MT Repository

2 x Synology RX1214 NAS + expansion

- 2 x 24 Enterprise Class 4 TB HD (~2 x 96 TB of capacity)
- RAID 6, 1 hot swap drive per rack, 4 racks
- Automatic sync across NAS’s triggered by any file updates
- Compilation of software available to run based on our needs.
  - Service tickets
  - Inventory control
  - Cloud server
Real life example – recovery from total disaster

- ~6 Enterprise class 4 TB hard drive failures
- Complete control hardware failure
Real life example – recovery from total disaster

- ~6 Enterprise class 4 TB hard drive failures
- Complete control hardware failure

- 100% data recovery
Standard Database Schema Across Projects

Can this be generalized so all projects can be provided pre-populated templates + Cloud so manual “uploads” To DMC can be eliminated?

- Field Data
- Pictures
- Site visit reports

- Daily Production Reports
- Permits
- Maps
Local Disaster Protection

- Amazon Glacier
  - Low-cost cloud archiving
  - Built for large storage
  - Optimized for infrequently accessed data
  - Retrieval time is several hours
Summary and Future Work

• Data management: move from obtuse/time consuming to robust procedures
• Remove human error portion, automation!
• Standardization of procedures across projects

• Future Work:
  • Move away from Dropbox, onto alternative clouds or Synology cloud software
  • Develop systems to upload directly to cloud storage
  • Develop cloud based software: inventory control, service tickets, etc.
ANY questions?