SuperDARN Data Distribution Report

Kevin Krieger
SuperDARN Workshop
Moose Jaw - Saskatchewan
May 2013



Data Distribution: Summary

- Hard drives written and distributed
 - 20120101-20120430 Shipped July 9th
 - 20120501-20120831 Shipped November 7th
 - 20120901-20121231 Pending Shipment

 Hard drives are 2TB Seagate FreeAgent GoFlex Desk USB 3.0 compatible

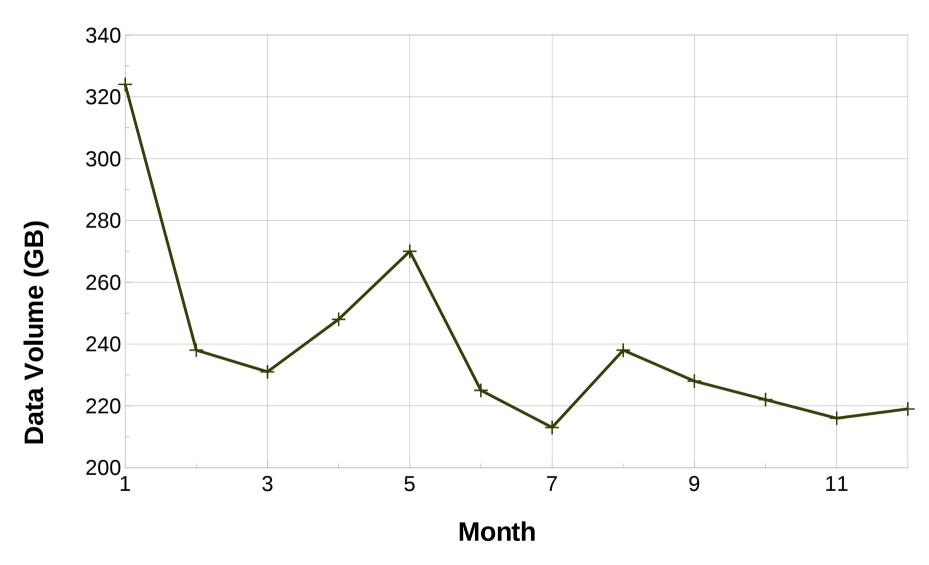


Data Distribution: Gaps

- Both Stu Nylund at APL and I worked to resolve "gaps" in data
- Emails with lists of data gaps sent out to data managers after each 4 month period
- Data managers respond with which gaps are real, which are not – they post data they have for APL to route through U of Saskatchewan



Monthly Data Amount for 2012





Primary Data Mirror: Motivation

- Hard drives suck
- Everyone else is doing it
- Networks
 - Reliable
 - Redundant
 - Faster access to data



Primary Data Mirror: Implementation

- Secure File Transfer Protocol (sftp)
 - ftp-like* transfer protocol
 - Secure
 - Plays well with ssh and chroot
- WestGrid Network File System (NFS) storage
- Synchronization ability via example script
 - Compare bzipped rawacf file hashes on server and user computers



Primary Data Mirror: Implementation

- User accounts setup:
 - Need user computer address
 - Use Public/Private Key Pair authentication (ssh)
 - Read-only access to data files
- Speed factors:
 - Compute local hashes ~0.28s per file
 - Download Variable From < 1s/file to > 3 min/file



Primary Data Mirror: Present

- September 2011 December 2012 available
- January 2013 April 2013 data pending
- 12 users currently set up and tested
- ~7 users actively using mirror



Primary Data Mirror: Future

- Improvements to sync script speed
- Improvements to usability/user-friendliness
- More nodes in SuperDARN Mirror network?

