

Data outage last 2 days

Okay, this will attempt to document issues during the past 2 days.

1. ehs USB stick died sometime between the 24 mar and 25 mar 00Z rsyncs. fdisk reports no partitions on this stick. I replaced this with a new USB stick about an hour ago and it is now working.
2. Santiago rebuilt/started eol-rt-data yesterday. Apparently, in the process the ssh configuration was modified to disallow (some) connections to porter2.
3. Yesterday, in an attempt to solve the eol-rt-data issue by myself, I tried restart_process of ssh_tunnel on flux. This killed ssh to flux from the outside. SRS rebooted flux last night at about 1700 and I rebooted it again at about 1200 today (2 separate trips to the tower), which restored the connection from flux to eol-rt-data, but still failed to get all the way through to the eol machines due to eol-rt-data configuration issues. I don't know if this was a red herring and would have fixed itself when eol-rt-data was fixed.
4. Today, we also tried 2 porter2 reboots. At least the second one was justified, since sstat reported most services not running and restart_service didn't bring them back. All was well after the reboot.
5. Even once eol-rt-data was fixed (by Ted restoring an old image of the virtual machine!) (that restored ssh tunnel to flux), connections to bao and ehs were broken. We found that Gordon's check_udp_... only reported errors and didn't restart the udp_tunnel process. Running this by hand finally got data flowing. We had also run this process earlier in the day, so there are a couple of hours of udp data that made it to porter2.
6. Even with everything fixed, rsync_flab_loop.sh failed with a PATH issue (couldn't find rsync_flab.sh). This is really strange, since this script has always worked. I manually added setting of PATH to this script. In the meantime, I also ran the nightly rsync manually (which was hideously slow – about 3 hours to bring 2 days of data back – so I cheated on some by removing the bandwidth limit).

Lessons:

1. Santiago/Ted were unaware that CABL even existed and didn't think to check with Gordon (though he wasn't available) before proceeding with system work. They did ask (vacationing) Gary, who didn't educate them.
2. We still need to fix check_udp_... to restart automatically – I'm guessing that this is just a path issue.
3. Ted thinks that virtual machine "snapshots" were the cause of the config errors and has decreed that they should be avoided in the future.
4. Santiago wants to better document eol-rt-data and is thinking about splitting up some of its services onto other virtual machines.

Impact:

1. BAO tower data should be fine since the DSMs kept running and flux was up to rsync their data. All of these data should be rsynced (and new statsproc files and wwwplots generated) soon.
2. bao data also should be fine since its DSM was up and saving to local storage. It appears that porter2 was able to rsync its data from 0325 and will be able to run tonight. flux wasn't able to get these data and has a gap between 0325_195959 to 0326_202418.
3. ehs data were archived on flux (via udp transmission) until 0325_195959 and are missing until 0326_202418, as with bao. The porter2 files are even worse, with no data from 0325, undoubtedly due to the bad USB stick. More data will not be filled in by the nightly merge. About 28 hours of data were lost. Note that the NetCDF statistics files have 1–2 hours of data for which we don't have raw_data. If these NetCDF files are ever regenerated from scratch, we'll lose this (short) period of data.

Whew!