CONDUIT Update, December 1, 2005

Sites Receiving CONDUIT

As of December 1, the CONDUIT data feed is currently being received by at least 80 unique hosts in 50 domains (Fig 1.), including sites in the United States, Brazil and Portugal. Sites reporting reception include research and education universities, government facilities and laboratories, as well as both commercial and non-profit organizations. Recipients can obtain the data without providing information back to Unidata, so the known hosts are a lower bound of the total.

Changes since January AMS Meeting

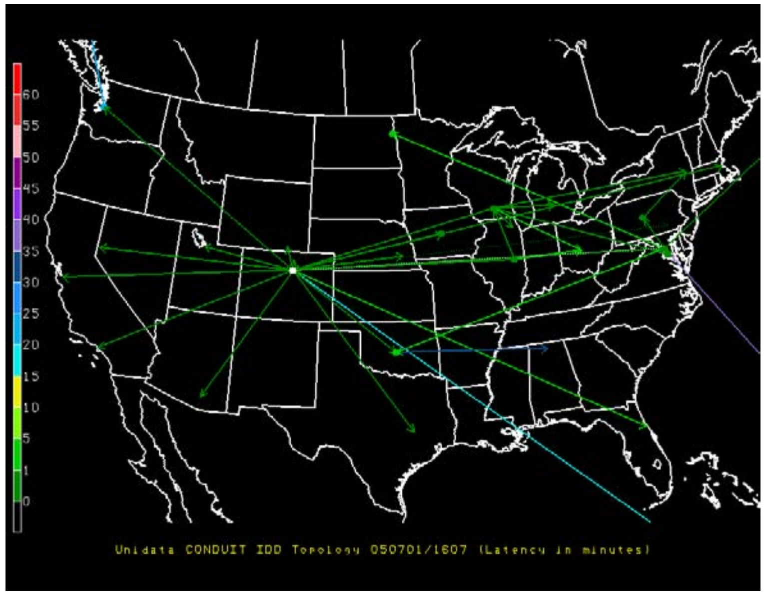

Figure 1. CONDUIT IDD topology with latency values.

The addition of 0.5 Degree GFS data to CONDUIT in November, 2004 stressed the infrastructure which inserts the data into the top level LDM host on TGSV32.NWS.NOAA.GOV to the point where data sets were not being delivered in a timely fashion due to the large number of files being processed as well as the increase data volume. Several attempts to improve the rate at which the data could be inserted into the LDM data queue involving both the insertion software on TGSV32, as well as the remote DBNet hosts, TGSV10/11, which trigger the insertion processes on TGSV32 were tested. The results of software and operating system monitoring revealed that the primary bottleneck was the amount of physical memory available for the LDM product queue on TGSV32. On February 7, 2005 the product queue on TGSV32 was decreased in size from 1.5GB to 800MB. The resulting change provided for a return to acceptable reliability in the data feed. The decrease in product queue size was a tradeoff which limits the time period (approximately 20 minutes can now be retained) for which a relay site connected to TGSV32 can lose connectivity and still be able to obtain the data posted during the time of disconnection. Typical time to deliver data from TGSV32 to downstream LDM relays is less than 30 seconds, as shown in Figure 2.

On March 16, 2005, DBNet enhancements provided by Brent Gordon's group were added which now allow concurrent insertion from the remote connections from TGSV10/11 to TGSV32 which further improved timeliness in the time between when a file is available via the AFS network and when it can be inserted into the LDM product queue on TGSV32. At present, up to 4 concurrent insertion processes are running on TGSV32. The number of processes that can simultaneously insert data files into the LDM queue on TGSV32 will also be affected by memory limitations on that machine. Unrelated changes at NCEP to the ensemble file posting has since made the data available on the AFS server much sooner (ensemble files now post nearly

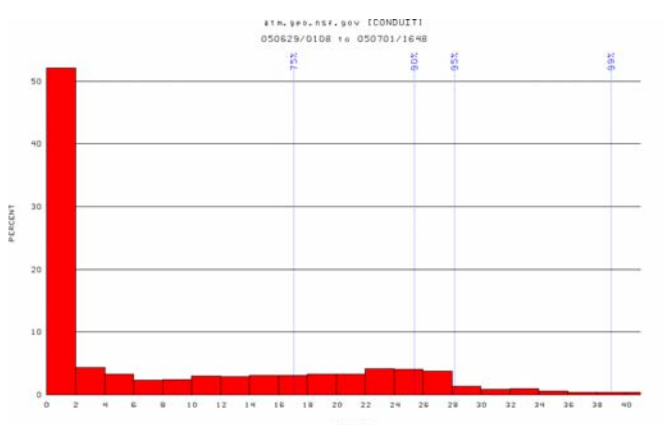

Figure 2. Histogram showing time to receive data via CONDUIT feed from TGSV32 to ATM IDD sites.

simultaneously, where previously had taken up to 1 hour), which has led to greater injection delay as seen in CONDUIT posting, although the data is available sooner to the LDM relay sites than it had been previously.

CONDUIT Reliability and Timeliness

Delivery of data sets via CONDUIT has been very reliable since the modifications to the LDM data queue on TGSV32 and the addition of multiple insertion processes for DBNet on TGSV10/11. As shown in Figure 3, the daily volume of data received at the LDM relay site ATM.GEO.NSF.GOV consistently totals nearly 30GB per day.

CONDUIT data sets are generally inserted into the LDM product queue on TGSV32 within a few minutes of availability via the AFS file system, with the primary exception of the extended forecast hours of the GFS run (available 4 times per day) and the global ensemble files (available twice per day) as depicted in Figure 4.

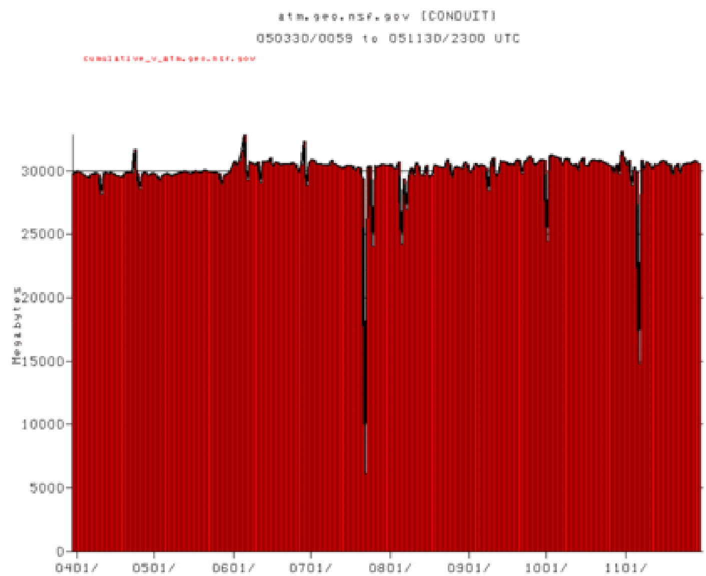

Figure 3. Daily volume of CONDUIT data received at ATM.GEO.NSF.GOV (March 30-December 1, 2005).

Changes to CONDUIT Data Stream

The time series depicted in Figure 4 shows that the NAM/ETA insertion into the LDM queue on TGSV32 is currently able to be accomplished in good time. As a result, we have requested to:

1) Add the 12km GRIB2 format NAM files which contain surface, boundary layer and pressure level data;
2) Remove the 20km GRIB1 format surface and boundary layer grids.

This change is in response to pending action items from the CONDUIT meetings, with the overall effect of maintaining the same number of files, while replacing 7MB per forecast hour files with 21 MB per forecast hour files (roughly 406MB more data). The expectation is that the existing TGSV32 hardware and software configuration can accommodate this modification.

Changes to the data stream which would increase the number of files or volume of data being posted during the GFS and Global ensemble cycles would be problematic given the current hardware limitations of TGSV32, and would best be accomplished through comparable deletions from the data stream where data set duplications occur (for example, if the 0.5 degree GFS output were extended past the

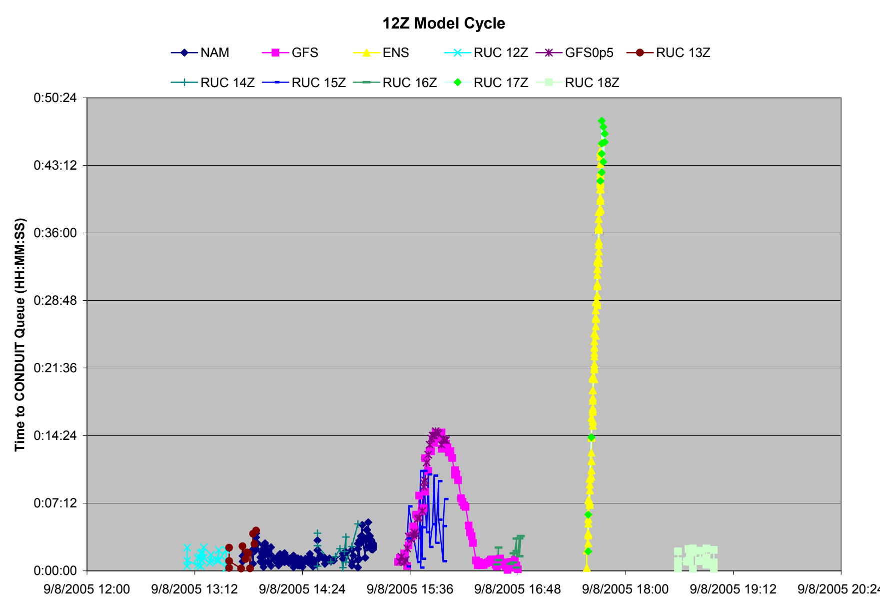

Figure 4. Typical duration of time between when data files are available via FTP server, and when inserted into the LDM queue on TGSV32. 12Z model cycle shown for June 20, 2005.

current 84 hour availability to 180 hours, then adding these files should be accompanied by a compensating removal of 1.0 degree (Grid 3) GFS files for forecast period 0 through 180 hours. The addition of the GDAS analysis fields may be acceptable given the expected size and availability. Coordination with NCEP for delivery of the data set will be needed following the moratorium on changes within the NWS.

Accommodating Future Changes to CONDUIT Data Stream

As the volume of data being added to CONDUIT grows, it is believed that the existing TGSV32 hardware will be incapable of maintaining a data buffer that is sufficient to ensure full receipt by downstream recipients in the face of transient network outages.

As part of its continued testing regimen, the UPC has recently successfully explored a cluster approach to IDD data relay. The cluster built in-house is able to relay the full set of IDD data, including CONDUIT, to over two hundred downstream connections without introduction of product latency and while maintaining a backing store of over three hours of data.

The cluster data server backend was implemented using modestly priced (approx. $5500), multiple Sun SunFire V20Z PCs that feature dual 64-bit AMD Opteron processors, 2x 36 GB 10K RPM SCSI hard disks, redundant power supplies, and a minimum of 12 GB of RAM in a 1U rackmount chassis. These machines were purchased through a Sun educational discount program. Each data server was configured to run the 64-bit version of Fedora Core 3 Linux after comparative testing of three operating systems (Sun Solaris x86 Version 10, FreeBSD 5.3, and Fedora Core 3). Notably, Fedora Core 3 Linux is the free distribution of RedHat Enterprise WS 4 Linux.

The most robustly configured of the data servers tested (configured with the maximum possible system 16 GB of RAM, and costing $6500) was able to IDD relay data to approximately 230 downstream connections with no introduction of product latency while essentially idling. Given the excellent performance of this machine, we strongly recommend that the NWS consider it as a replacement for the current TGSV32.