Tests using the Globus "gsiftp" tool between Caltech and Argonne
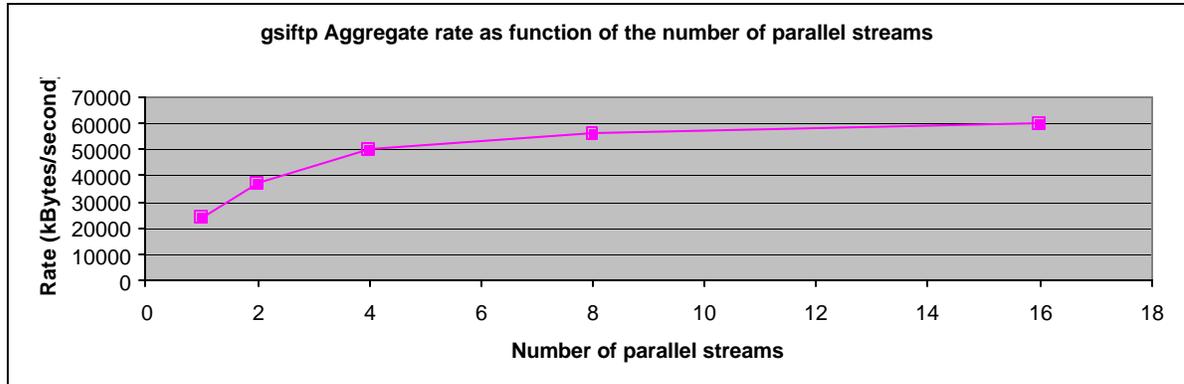
Julian Bunn, 16th December 1999

Introduction

Measurements of throughput on the WAN from Caltech/CACR to Argonne/MCS were made using the "gsiftp" Globus tool. An Objectivity database file of size 240 Mbytes was "put" to /dev/null on the Argonne machine. The CACR machine used was the 256-CPU Exemplar X-class "neptune.cacr.caltech.edu", the Argonne machine an Origin 2000 "denali.mcs.anl.gov". The TCP window size for the tests was set using the "lbufsize" and "rbufsize" gsiftp commands. The values for the local and remote buffer sizes were always set identically. In an attempt to saturate the network, multiple gsiftp *client* streams were used (from one to sixteen streams).

LAN route over HiPPI

As a test of the maximum throughput out of the Exemplar when using the gsiftp server.client combination, I initially measured the transfer speed of a "put" from Neptune out and then into Neptune via the HiPPI switch (theoretically capable of 80MBytes/sec), as a function of the buffer size. The results are shown below:
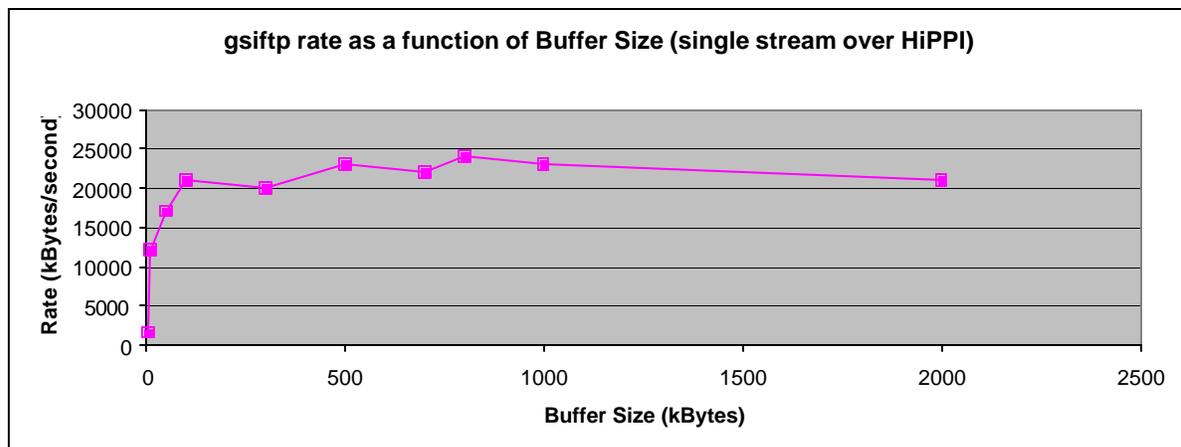
I then measured the aggregate rate achieved when running several parallel gsiftp streams, using a buffer size of 800 kBytes. The results approach the maximum practical throughput for the HiPPI connection. I note that the file used was almost certainly in cache during these tests, otherwise I would see a limit at the disk I/O speed (approx. 20 Mbytes/sec for the RAID device used).

gsiftp Aggregate rate as function of the number of parallel streams

(chart: Rate (kBytes/second) vs Number of parallel streams)
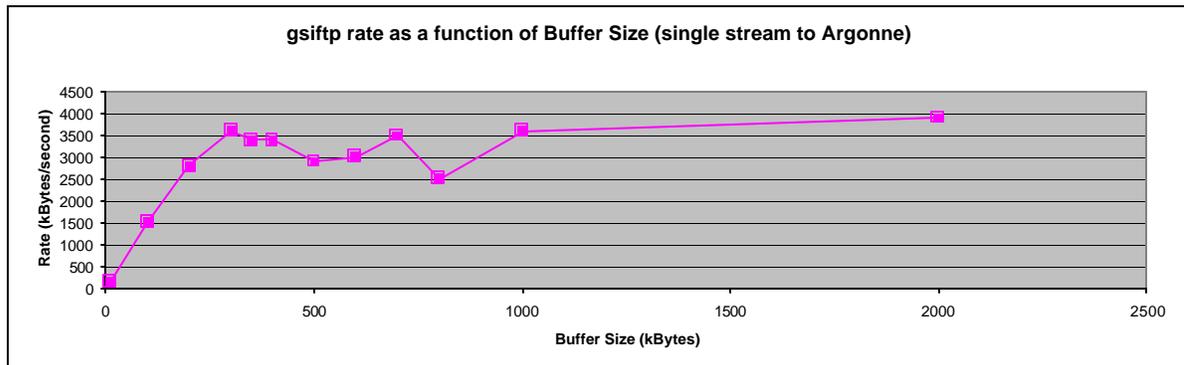
WAN route

The route from Caltech (neptune.cacr.caltech.edu) to Argonne (denali.mcs.anl.gov):

```
traceroute to denali.mcs.anl.gov (140.221.9.116), 30 hops max, 20 byte packets
 1 BBMR-RSM.cacr.caltech.edu (131.215.145.252)    2 ms    1 ms    1 ms
 2 SFL-border.ilan.caltech.edu (131.215.254.252)   1 ms    1 ms    2 ms
 3 192.12.19.249 (192.12.19.249)                   1 ms    1 ms    1 ms
 4 c2-gsr.caltech.edu (192.41.208.49)              1 ms    1 ms    1 ms
 5 UCR--CIT.POS.calren2.net (198.32.248.10)        3 ms    3 ms    3 ms
 6 UCI--UCR.POS.calren2.net (198.32.248.14)        4 ms    4 ms    4 ms
 7 198.32.248.125 (198.32.248.125)                 5 ms    4 ms    5 ms
 8 USC--UCI.POS.calren2.net (198.32.248.18)        6 ms    6 ms    5 ms
 9 abilene--USC.ATM.calren2.net (198.32.248.86)    6 ms    6 ms    6 ms
10 scrm-losa.abilene.ucaid.edu (198.32.8.17)      15 ms   15 ms   15 ms
11 denv-scrm.abilene.ucaid.edu (198.32.8.2)       38 ms   38 ms   38 ms
12 kscy-denv.abilene.ucaid.edu (198.32.8.14)      48 ms   48 ms   49 ms
13 ipls-kscy.abilene.ucaid.edu (198.32.8.6)       58 ms   58 ms   58 ms
14 anl-abilene.anchor.anl.gov (192.5.170.169)     63 ms   62 ms   62 ms
15 stardust-msm-20.mcs.anl.gov (140.221.20.91)    63 ms   63 ms   63 ms
16 denali.mcs.anl.gov (140.221.9.116)             63 ms   62 ms   62 ms
```

Running a single gsiftp stream, and varying the buffer size, I obtained good results (~3000 kBytes/second) around 9am PST on 15th December:

gsiftp rate as a function of Buffer Size (single stream to Argonne)

(chart: Rate (kBytes/second) vs Buffer Size (kBytes))

After this period, something caused the WAN performance to degrade, and I obtained single stream rates of only around 750 kBytes/second, rising to aggregate rates of 1900 kBytes/second with 16 parallel streams (using 800 kBytes buffer size):

gsiftp Aggregate rate to Argonne as function of the number of parallel streams
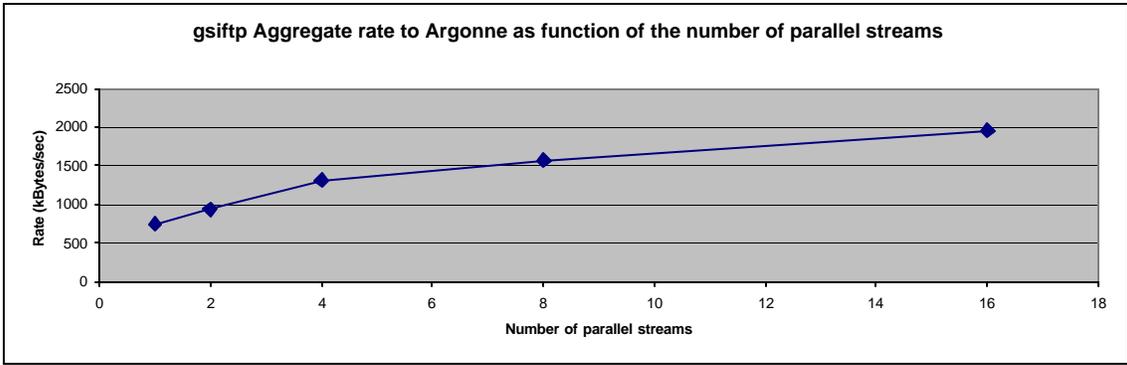
Conclusion

We need more bandwidth!