Concurrent File Transfers into the Cambridge HPC

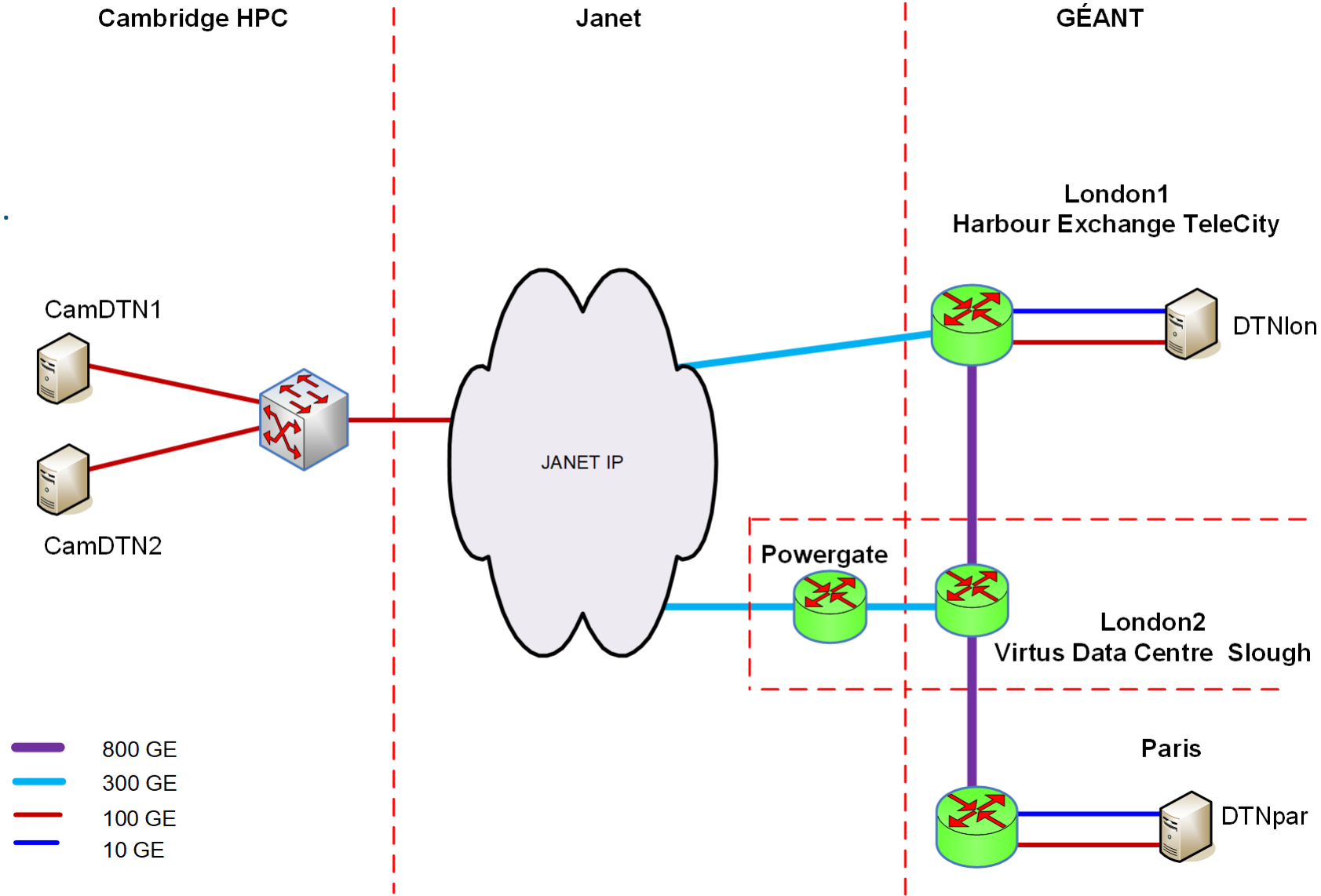
Richard Hughes-Jones

Network Topology Connecting the Data Transfer Nodes

Storage

- GÉANT DTNs
  Six INTEL DC P3700
  400GB, NVME SSD 8 lane PCIe3.

- Cambridge
  Lustre file system stripe over 2
  Object Storage Targets



Instrument davix-put and davix-get

- Measure times:
  - From start of program to start of transfers
  - Measure times during the transfer
    using the CPU cycle counter which counts CPU cycles (Time Stamp Counter).
  - Total time the program runs.
- Report transfer rate and number of TCP re-transmits at intervals during the flow
- Ability to set fixed size chunk reads from or writes to the network and disk.
- Measure and histogram times for “getChunk” & “putChunk”
  - socket read & disk write for davix-get
  - disk read & socket write for davix-put
- Note the TCP performance parameters using tcp_info struct for each chunk
- Record “getChunk” & “putChunk” times and tcp_info as time series in memory
- Option not to write a file (davix-get)

davix-put (http): DTNlon→CamDTN1 Disk-Disk Concurrent Flows

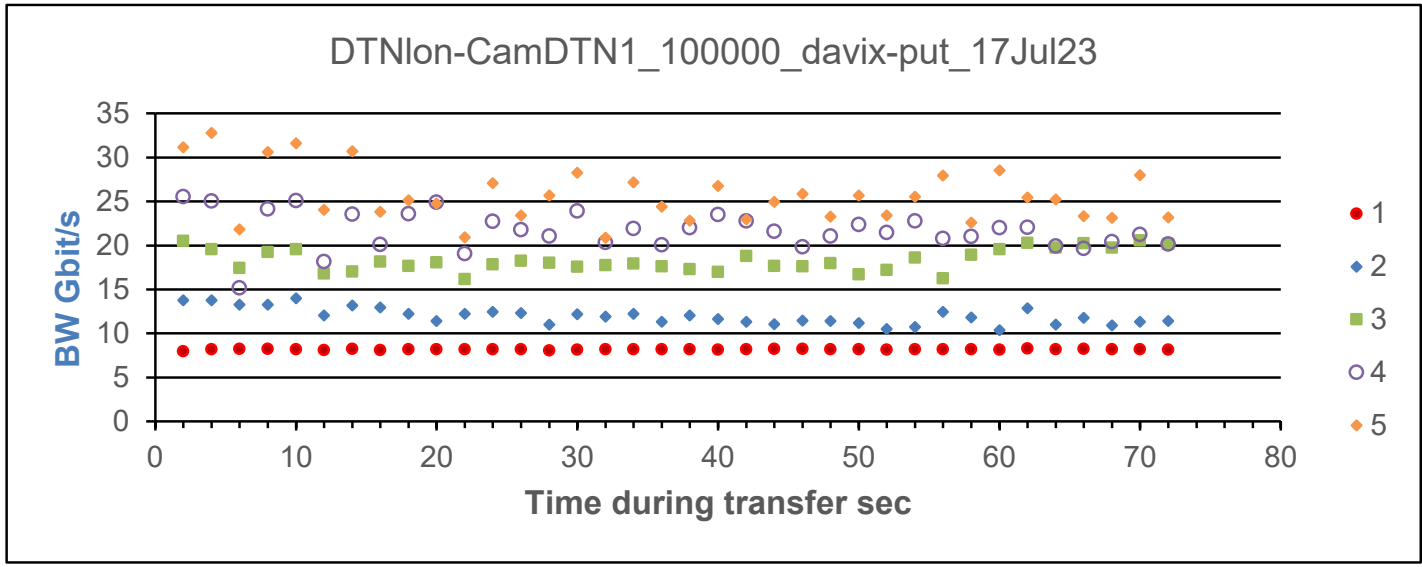
- CamDTN1 xrootd v 5.6.1

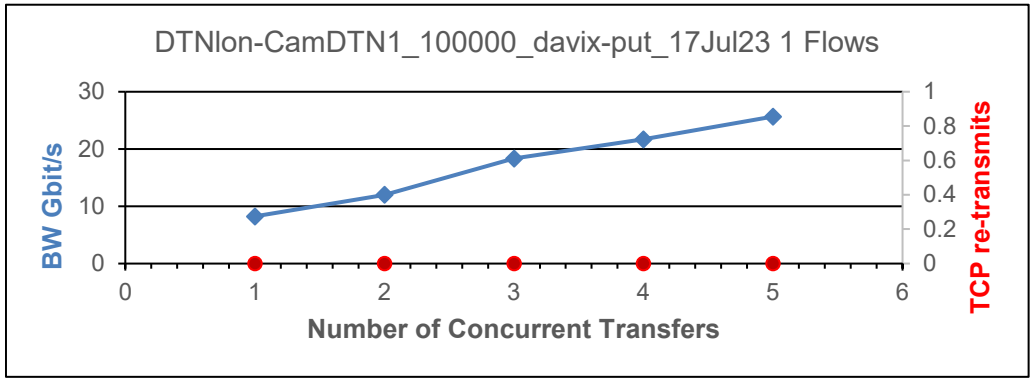
```
xrootd-http.cfg
xrd.protocol XrdHttp:5201 /home/dc-hugh3/Xrootd/build-5.6.1/src/libXrdHttp-5.so

[dc-hugh3@cpu-p-167 build-5.6.1]$ src/xrootd  -c ../xrootd-http-5.6.1.cfg
```

DTNlon davix-put

```
[richard@DTNlon davix_tests]$ ./cmd_run_davix-put_multiflow.py  --bufsize 1048576 --srcfile /mnt/dtn/DTNFILE100000 --dstfile /rds-d7/project/rds-
bRdYdViqoGA/rhj/stripe2/davix-put100000 -A 6 -d 192.84.5.1 -p 5201 -o  DTNlon-CamDTN1_100000 -n 5
```

- Each davix-put client on a different core
- Test with 1 File, 2 Files … n Files Concurrent
- Plot the total transfer rates
- Flows generally smooth as function of time.
- Not TCP re-transmits





- Scales well from 1 flow 8 Gbit/s to 5 flows 25 Gbit/s

davix-put (http): DTNlon →CamDTN1 10 Disk-Disk Concurrent Flows

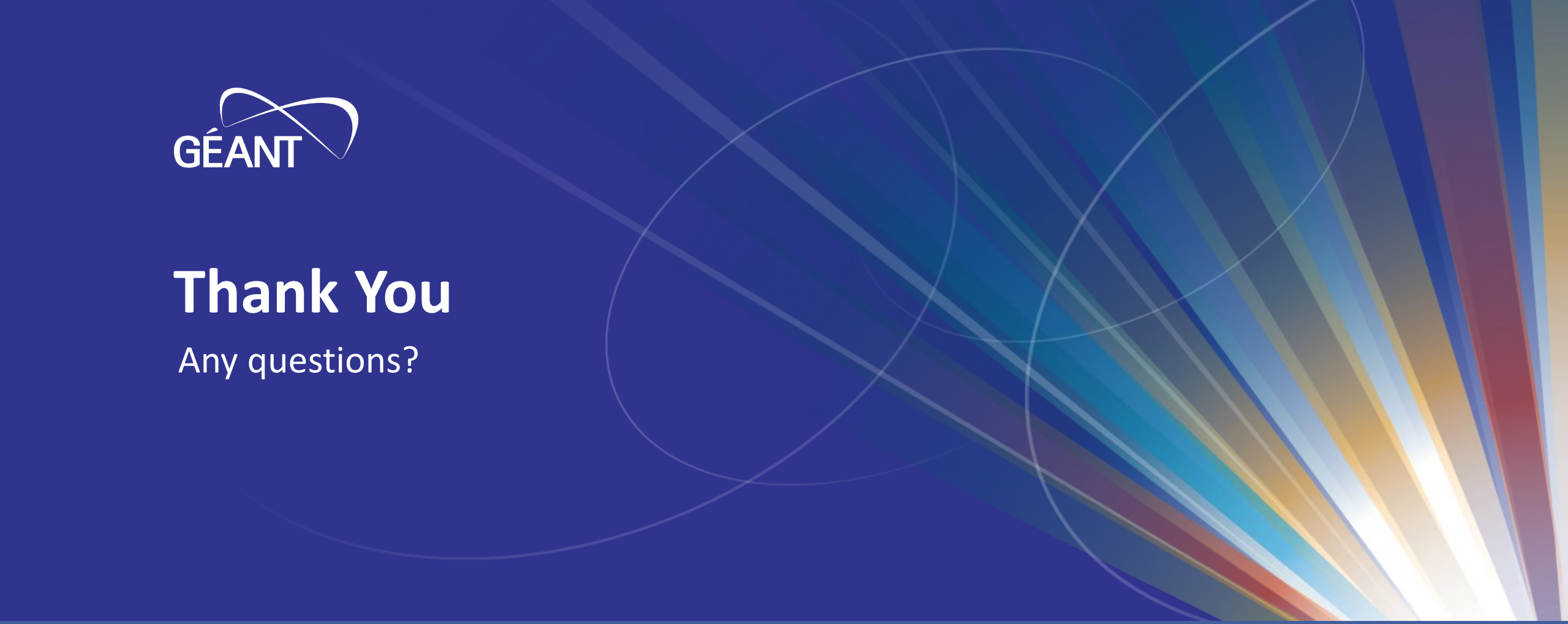
- DTNlon → CamDTN1 10 flows
- 100 GByte files 1 TB total data transferred
- On average a file took 2 min 36 sec
- 1 T Byte in 3 min 30 sec
- Some transfers took 10s to start moving data.
- Modest number of TCP re-transmits
- Scaling good 53 Gbit/s with 10 flows

Time_startup_s	Time_sending_s	Total_Time_s	Data_rate_GBytes/s	Data_rate_Gbit/s	Total TCP retrans
0.88	201.64	202.52	0.50	3.97	0
0.88	151.29	152.16	0.66	5.29	0
0.88	136.71	137.58	0.73	5.85	0
0.87	144.38	145.25	0.69	5.54	14
10.27	140.96	151.23	0.71	5.68	7
10.27	140.96	151.23	0.71	5.68	7
0.87	129.96	130.83	0.77	6.16	0
0.87	209.17	210.04	0.48	3.82	0
0.87	140.66	141.53	0.71	5.69	7
0.87	143.60	144.47	0.70	5.57	7
sum				53.24	42
ave				5.323889	

davix-put (http): DTNlon →CamDTN1 DTNpar →CamDTN1 10 Disk-Disk Concurrent Flows

- DTNlon → CamDTN1 5 flows
  DTNpar → CamDTN1 5 flows
- 100 GByte files 1 TB total data transferred
- On average a file took 3 min
- 1 T Byte in 8 min 20 sec
- One flow took a long time
- Scaling good 54 Gbit/s with 10 flows

Time_startup_s	Time_sending_s	Total_Time_s	Data_rate_GBytes/s	Data_rate_Gbit/s	total_TCP retrans
0.03	141.38	141.40	0.71	5.66	0
0.03	138.28	138.30	0.72	5.79	0
0.02	134.62	134.65	0.74	5.94	6
0.02	130.21	130.24	0.77	6.14	0
0.02	139.98	140.01	0.71	5.72	0
0.05	134.81	134.86	0.74	5.93	0
0.04	130.54	130.58	0.77	6.13	0
0.03	139.64	139.67	0.72	5.73	0
0.04	501.19	501.22	0.20	1.60	0
0.03	139.05	139.08	0.72	5.75	0
sum				54.39	6
ave				5.44	0



Thank You

Any questions?

www.geant.org

Co-funded by
the European Union