Performance and Fairness Issues in Big Data Transfers

[Extended Abstract]

Se-young Yu, Nevil Brownlee, and Aniket Mahanti
University of Auckland, New Zealand
{se-young.yu,n.brownlee,a.mahanti}@auckland.ac.nz

ABSTRACT

We present performance and fairness analysis of two TCP-based (GridFTP and FDT) and one UDP-based (UDT) *big data* transfer protocols. We perform long-haul performance experiments using a 10 Gb/s national network, and conduct fairness tests in our 10 Gb/s local network. Our results show that GridFTP with jumbo frames provides fast data transfers. GridFTP is also fair in sharing bandwidth with competing background TCP flows.

Categories and Subject Descriptors

H.3.4 [**Systems and Software**]: Performance evaluation

Keywords

Big data transfer protocols; fairness; performance

1. INTRODUCTION

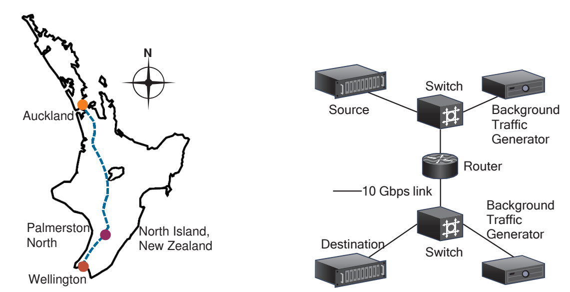
Large-scale scientific installations such as the Large Hadron Collider and Sloan Digital Sky Survey generate huge amounts of scientific data (or 'big data') every day. There are several teams of researchers distributed around the world who utilize this data to answer fundamental questions about the origins of the universe.

The time-sensitive nature of this scientific work necessitates fast and effecient delivery of the data. Many research sites are equipped with high capacity links (10 Gb/s) to transfer their big data, however, generic TCP or UDP implementations do not sufficiently utilize these links. Modified protocol implementations (e.g., GridFDT, FDT, and UDT) encompassing the application and transport layers exist that do a better job than the generic implementations.

There have been limited practical studies [1, 2] on fairness between protocols or discussion of definitions of fairness as they apply to big data transfer. In this paper, we present a study of three big data transfer protocols, namely, GridFTP, FDT, and UDT. We extend upon [3] by conducting experiments to understand the *performance* and *fairness* of these big data transfer protocols under *controlled* and *real-world* situations.

CoNEXT Student Workshop'14, December 2, 2014, Sydney, Australia.

ACM 978-1-4503-3282-8/14/12 ...$15.00.
http://dx.doi.org/10.1145/2680821.2680824.



(a) National network map (b) Local network testbed

Figure 1: Experimental setup

2. METHODOLOGY

Figure 1 shows our experimental setup that consists of a *national research network* and a *local in-lab* testbed. We designed our experiments to emulate real-world situations to help users understand the behaviour of the big data transfer protocols, and help choose one that suits their need.

The national network setup consisted of two machines located in Auckland and Wellington, respectively. These two sites were separated by a physical distance of about 650 km. The machines communicated using the 10 Gb/s national network[1]. The baseline round trip time (RTT) between the two machines was 9.59 ms, and `traceroute` showed that there were three hops on the path. We measured each protocol under varying environmental settings such as using *jumbo frames* and increasing numbers of parallel flows. Our big data simulates were zero-filled 30 GB files for disk-to-disk transfers, and `/dev/zero`[2] for memory-to-memory transfer tests. We did not measure congestion-awareness behaviour of the protocol because we did not want to inconvenience other national network users. We used *goodput* as the performance metric.

The in-lab testbed was a controlled dumbell network consisting of a source and destination each connected to an OpenFlow switch through a 10 Gb DAC cable. There was one background TCP flow generator on each end connected by 10 Gb DAC cable to the OpenFlow switches. The two OpenFlow switches were connected by a router using 10 Gb

[1]`http://reannz.co.nz/`
[2]UDT did not support this feature. Hence, we did perform memory-to-memory tests for UDT.

Table 1: Protocol goodput for a single flow

Protocol	Jumbo	Disk-to-Disk (Gb/s)	Memory-to-Memory (Gb/s)
GridFTP	No	1.3	5.0
	Yes	1.3	4.5
FDT	No	1.2	2.8
	Yes	1.2	4.5
UDT	No	1.3	N/A
	Yes	1.3	

SFP optical cable. We used `nuttcp` for background traffic generation. We set up `nuttcp` to generate a maximum of 1 Gb/s to 10 Gb/s of background TCP traffic, and measured the actual throughput on the wire against our big data transfer throughput. We used *throughput* of the actual big data transfers as well as competing background traffic to investigate *fairness* issues.

3. RESULTS

3.1 National Network Testbed

We start by discussing results on performance of the protocols in the national network. Our experiments were performed when the national network link utilization was minimal. While discussing the results, we consider the national network testbed to be *uncongested*.

Table 1 shows the single flow performance of each protocol for disk and memory transfers with and without jumbo frames. We observe that disk I/O throttles the performance of the transfer protocols. To remove that limitation, we measured memory-to-memory transfer performance for GridFTP and FDT. Although using jumbo frames slightly reduced GridFTP goodput, they improved FDT goodput by about 40%. FDT (is implemented in Java) has more per-packet overhead than GridFTP, which is implemented in C.

Figure 2 shows the goodput of the two protocols (with no background traffic) when we incremented the number of parallel flows. Note that we performed memory-to-memory transfers in this case. We observe that multiple parallel flows did not improve GridFTP or FDT goodput significantly.

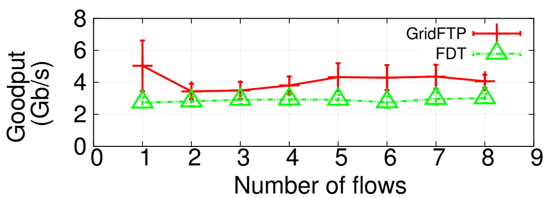


Figure 2: Protocol goodput with incrementing parallel flows on the national testbed

Without a fast disk system, there may be no benefit in choosing one protocol over another in a 10 Gb/s network. With an efficient disk system, using GridFTP may give better performance than FDT and UDT. *We recommend GridFTP as our choice on an uncongested network.*

3.2 Local Network Testbed

We next discuss the fairness issues related to the big data transfer protocols. In particular, we look at how well the protocols share the bandwidth with competing background TCP flows in a *congested* network. We sent our big data simulates into the link at a fixed rate of 10 Gb/s, while

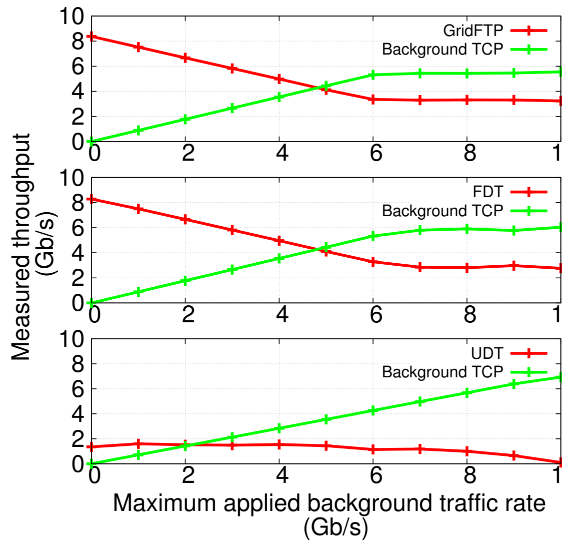


Figure 3: Protocol throughput compared to background traffic throughput on the local testbed

varying the maximum applied background TCP traffic rate from 1 Gb/s through 10 Gb/s.

Figure 3 shows the change in throughput for each of the protocols with increasing applied rate of background traffic. The average throughput of GridFTP and FDT decreased at an equal rate as background traffic increased, respectively. As applied background traffic rate was ramped up to 6 Gb/s, we observe an equilibrium where the big data transfer and background traffic throughputs flatten out. UDT did not utilize the maximum link capacity even in the absence of background traffic.

On a congested 10 Gb/s link, GridFTP and FDT will reach an equilibrium where they share the bandwidth with competing traffic while trying to maximize their own utilization. In contrast, UDT does not effectively utilize the link capacity and is unable to share the link with competing flows. *We recommend using GridFTP because of its fair capacity utilization in congested link.*

4. CONCLUSION AND FUTURE WORK

Our goal was to make recommendations for a choice of big data transfer protocol in a 10 Gb/s network. We measured the performance of GridFTP, FDT and UDT in terms of their performance and fairness. We observed that GridFTP and FDT performed similarly but UDT had some utilization problem with our testbed. We found that using GridFTP with jumbo frames for transferring big data is well suited in both congested and uncongested links.

In future, we will look into measuring performance of UDP-based big data transfer protocols with high RTT links. Measuring performance of UDP-based transfer protocols will give us more insight into utilization of long-haul links and fairness among the protocols. We will also look at the effect of varying frame sizes on the big data transfer protocols. Jumbo frames may make significant difference in high-speed data transfer protocol. We will quantify the improvement in future experiments. Lastly, we will endeavour to discover better metrics for measuring fairness of the protocols.

5. REFERENCES

[1] R. Cottrell, S. Ansari, P. Khandpur, R. Gupta, R. Hughes-Jones, M. Chen, L. McIntosh, and F. Leers. Characterization and evaluation of TCP and UDP-based transport on real networks. *Annales Des Telecommunications*, 61(1-2):5–20, 2006.

[2] S. Ha, L. Le, I. Rhee, and L. Xu. Impact of background traffic on performance of high-speed TCP variant protocols. *Computer Networks*, 51(7):1748–1762, May 2007.

[3] Se-young Yu, N. Brownlee, and A. Mahanti. Comparative performance analysis of high-speed transfer protocols for big data. In *Proc. IEEE LCN*, pages 292–295, Oct 2013.