Data Transfer in a Science DMZ using SDN with Applications for Precision Medicine in Cloud and High-performance Computing

Nam Pho[*]
Harvard Medical School
Research Computing
Boston, USA

Dino R. C. Magri
University of São Paulo
LARC - PCS - EPUSP
São Paulo, Brazil

Fernando F. Redigolo
University of São Paulo
LARC - PCS - EPUSP
São Paulo, Brazil

Byoung-Do Kim
University of Virginia
Research Computing
Charlottesville, USA

Timothy Feeney
Tufts Medical Center
Department of Surgery
Boston, USA

Heidi L. Morgan
Florida International University
CIARA
Miami, USA

Chirag J. Patel
Harvard Medical School
Department of Biomedical
Informatics
Boston, USA

Chris Botka
Harvard Medical School
Research Computing
Boston, USA

Tereza Cristina M. B.
Carvalho
University of São Paulo
LARC - PCS - EPUSP
São Paulo, Brazil

ABSTRACT

As the difficulty and cost of generating data has decreased, there have been emerging problems of data movement between where it was collected to where it can most effectively be processed. The Science DMZ model is an established solution to help maximize the throughput from any given network infrastructure between sites but there still exists room for exploration and improvement, including the use of Software-Defined Networking (SDN). We aim to explore how the use of SDN-enabled flows in a Science DMZ may impact its main trait: performance. In this paper we have successfully demonstrated improvements in data transfer between certain sites and under certain conditions. Though we have also seen sites where there is no difference or even a performance penalty, all the results are consistent and reproducible over hundreds of test transfers suggesting there are unique properties of each site configuration leading to real and sustained performance increases to a Science DMZ.

Keywords

Science DMZ; SDN, Software Defined Networking; Brocade;

---

[*]The author can be reached at: nam_pho@hms.harvard.edu

OF, OpenFlow; Genomics; Precision Medicine; Personalized Medicine; Cloud Computing; Top500; HPC, High-performance Computing

1. INTRODUCTION

In early 2015, a little over a decade after the first draft of the human genome was published in 2003, the United States announced a new $215-million precision medicine initiative [10]. Though daunting in scale at its goal to collect high-throughput biological (e.g., genomic) data from millions of individuals, it follows as a natural progression of increasingly larger-scale sequencing studies such as the 1,000 genomes [9], The Cancer Genome Atlas (TCGA) among others. The exponentially decreasing cost of sequencing an individual's genome has dwarfed the rate of Moore's law and turned medicine into a big data discipline. There are even predictions in the scientific community that it will exceed the storage growth rate of established big data fields such as online video (YouTube), social media (Twitter), and astronomy (Australian Square Kilometre Array Pathfinder project). Conservative estimates put genomics on par with these other fields at two exabytes of genomic data by 2025 to more extreme predictions of up to 20 orders of magnitude more [8]. The computational resources required to process this data will also grow in tandem. The second sizable challenge is security. Precision medicine is defined as human-subjects research and must comply to regulation that ensure the privacy and confidentiality of research subjects (e.g., HIPAA). Therefore, the transfer of data must also be secure.

Regardless of the discipline, moving large volumes of data between sites has been increasingly stressed by the ease of data accumulation; pockets of data rich and data poor sites
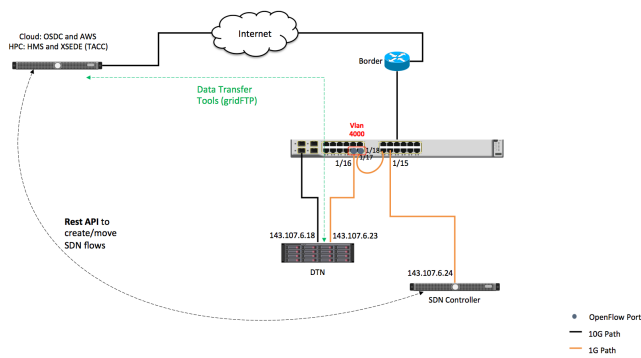
Figure 1: A network diagram of data as it transits the internet from the DTC to the DTN through a Science DMZ

naturally form. Particularly in the genomics field, there is a tension between data production centers, such as those producing next-generation genomic sequencing information, and supercomputing or cloud data centers that have computational resources required to process and analyze data. Through domestic initiatives such as ESnet and Internet2 and international ones between North and South America such as AmLight [4], improved bandwidth interconnectivity and other data transfer methods were introduced to address increasing data sizes. However, even with sufficient 10 Gigabit bandwidth infrastructure between locales, there are blocks to the natural flow of data such as stateful firewalls that the Science DMZ was created to address [2]. The concept of a Science DMZ has also been adopted internationally, particularly in Brasil, with a similar nation-wide initiative from Brazil's National Research and Educational Network (RNP - Rede Nacional de Ensino e Pesquisa) led by the University of São Paulo (USP) researchers [1].

2. SDN-ENABLED SCIENCE DMZ

Software-Defined Networking (SDN) has been a maturing technology that has seen measured rates of adoption in the real-world [5]. The use of a SDN-enabled Science DMZ was proposed by Dart, et al. in 2013 but we describe the implementation and benchmarking of it in practice for genomics.

At USP we are using a Brocade hybrid Openflow switch (CES-2024C-4X) with the latest firmware update (v5.8.0) for the Science DMZ. The University's border router and the SDN controller are connected to a non-OpenFlow vlan (vlan1), whereas the Data Transfer Node (DTN) has ports both in vlan1 and in an OpenFlow-enabled network segment (vlan4000). A physical loop connects both segments.

As shown in Figure 1, Data Transfer Clients (DTCs) running on different external sites initiate a request for file transfer to the DTN at USP via a REST API call to the SDN controller, which will then insert OpenFlow rules into the switch in order to enable traffic from each DTC to reach the DTN. Upon completion of transfer the flows will subsequently be removed.

Using the ovs-ofctl OpenFlow command we create four distinct flows on the switch via a REST API on the SDN controller when the request is initiated. Two of them handle ARP packets to and from the DTN. The third directs IP packets destined for the DTN to the OpenFlow-enabled

port through the physical internal cable loop, while the last one handles packets leaving the DTN in the reverse direction through the internally looped port bridge. When a transfer is finished another call to the REST API on the SDN controller will destroy the OpenFlow table and prevent packets from reaching the DTN.

3. DATA TRANSFER

As the goal of any Science DMZ is to optimize data transfers, we aim to show that the use of additional features provided by SDN come at no expense to existing performance. We use one of the first available full genome sequences from an individual as our transfer data set [6] where the goal is to process and map the genome from raw data at a compute-rich cloud or HPC site then transfer the assembled sequence via USP's Science DMZ. All machines involved are running Linux (CentOS 6.6) using GridFTP 5.2.5 for the transfer agent. As disk latency is known to also be a factor in data transfer we will try different storage options on the DTC to a local RAID-1 disk on the DTN at the Science DMZ. In each condition we will perform 100 file transfers and record their average speed in Mbps.

3.1 Personalized Medicine

The chemical DNA of an individual is comprised of approximately 2.8 billion base pairs (bps) divided into 23 pairs of chromosomes of which each base is represented by a single IUPAC code as ASCII on disk with accompanying metadata. The most common method of sequencing a genome involves shearing chemical DNA into fragments of several hundred bps and sequencing each fragment from one or both ends, using the sequence piece to map to a reference genome. The mapping process is computationally intensive though embarrassingly parallel in nature. The genome to be assembled and transferred was sequenced at 7.5-fold coverage and is 180 Gb of raw data and 2.8 Gb of processed data on disk though the standard now is commonly 30-fold coverage or greater. Details on the raw data and processing protocol are beyond the scope of this paper and can be described in detail in Levy, et al.

3.2 Academic and Commercial Clouds

Cloud computing is an attractive option for processing data as it allows for elasticity; the cloud can dynamically scale to fit the task at hand and costs are only commensurate with the resources required of a project. Despite the availability of cloud solutions, adoption of it has been slow for research projects though in recent years regulatory and funding support has caught up to the changing research paradigm [7].

The Open Science Data Cloud (OSDC) is an academic cloud and data commons running OpenStack hosted at the University of Chicago [7]. Academic clouds are useful for being specialized to individual research domains and more forgiving as a private research environment. For example, Bionimbus is an academic cloud for biomedical research providing a central data commons in conjunction with the elastic compute services [3]. In our example we are running a m1.large instance on the OSDC with virtual and GlusterFS storage options as a DTC.

In the commercial cloud computing space players such as Google, Amazon, Microsoft, and Rackspace have established solutions. The benefit of a commercial solution is the pro-
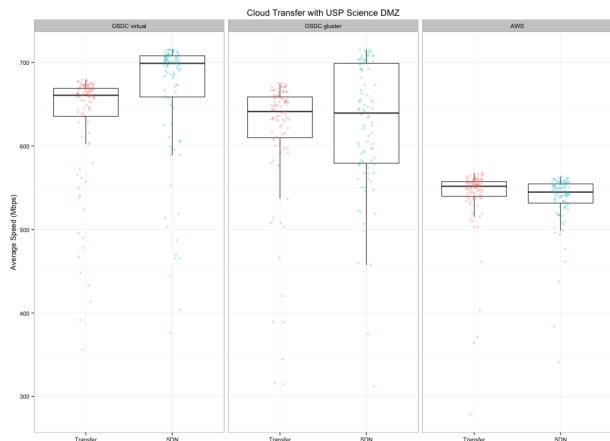
Figure 2: Genome transfer using GridFTP with and without OpenFlow enabled between academic (OSDC) and commercial (AWS) clouds with the USP Science DMZ

Figure 3: Genome transfer using GridFTP with and without OpenFlow enabled between Top500 (TACC) and academic medical (HMS) supercomputing centers with the USP Science DMZ.

prietary engineering behind each cloud, which is often derived from the company's experience delivering other services through their proven global infrastructure. Using the Amazon Web Services (AWS) cloud a DTC was provisioned on an m3.xlarge instance with a SSD backed EBS in the Northern Virginia (US) zone.

When transferring from a virtual filesystem on OSDC we experienced statistically significantly faster data transfers (p=2.45e-3), an increase from 630 to 663 Mbps on average at 5% improved throughput (Figure 2, left panel). Using the same DTC on OSDC with the GlusterFS volume we experienced no change in data transfer performance (p=0.379) from the normal transfer at 615 Mbps on average though the variability in speeds increased 2.4-fold from an IQR of 48.9 to 120 Mbps under SDN transfers (Figure 2, middle panel).

Using AWS we saw slower average file transfer speeds compared to the OSDC and no difference in speed using normal or SDN-enabled flows (p=0.538). However, there was much greater consistency for each transfer at 539 Mbps on average for normal conditions and only a 5.44 Mbps difference in IQR between normal and SDN flows in AWS (Figure 2, right panel).

3.3 Top500 and Academic Medical HPC Centers

Supercomputing sites such as those found in the Top500 are an attractive solution to help process data; a network of HPC centers form the backbone of the Extreme Science and Engineering Discovery Environment (XSEDE) consortium in the US [11]. Among one of the sites is the Stampede supercomputer at the Texas Advanced Computing Center (TACC), among the top 10 clusters in the world as measured by more than 5,000 Teraflops of processing power. At TACC the data was staged for file transfer on two separate production Lustre-based storage devices functioning as home and scratch directories.

In the area of precision medicine, genomes are processed locally to where the samples are collected at academic medical centers and supporting HPC facilities are the first to
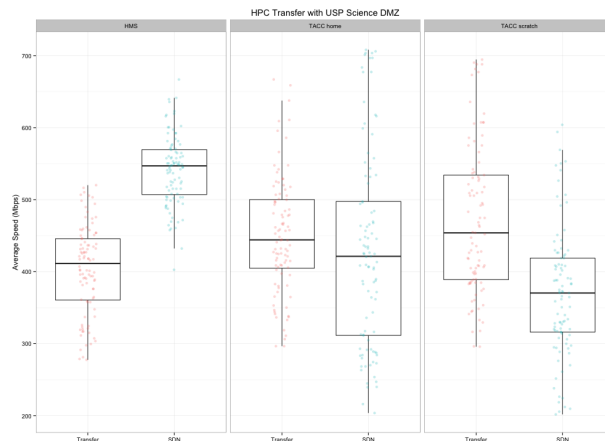
analyze the data. For example, the Orchestra cluster has over 7,000 CPU cores to support Harvard Medical School (HMS) and its 12 affiliated hospitals in all fields of biomedical research. At HMS the transfer data was staged on a local SSD storage volume.

Using the DTC at HMS we demonstrated a statistically significant increase in file transfer performance using SDN (p = 2.2e-16). Speeds to USP increased from 403.2 Mbps (sd=61.6) to 541.0 Mbps (sd=48.2) on average, a 34.2% increase in performance throughput (Figure 3, left panel).

From the TACC site there were mixed results with no change in average file transfer performance speeds using home (p = 0.199) but a performance penalty under scratch (p=3.55e-11) designated Lustre devices. Using the home device we averaged 450.3 Mbps of transfer while under scratch the average speed decreased from 470.8 to 371.0 Mbps, a performance degradation of 21% (Figure 3, middle and right panels).

4. DISCUSSION

Though Science DMZs have been an established technology to improve data transfer performance, we sought to explore methods of further optimizing it or extending its functionality by means of the SDN paradigm. Through our exploratory survey we received mixed results including performance increases at the OSDC (virtual) and HMS, indifferent results through AWS and OSDC (GlusterFS) and TACC (home), and even performance decreases at TACC (scratch). Given that each of these transfer sites is a production environment and tests were done over multiple days it is difficult to control with 100% certainty all potential confounding variables. For example, internal TACC benchmarks have the scratch device capable of 160 GBps of max throughput versus 12 GBps for home. Neither of these should be rate limiting properties for our tests yet file transfers from TACC using scratch had worse performance under SDN flows.

SDN-enabled Science DMZs remain a viable option for optimizing the secure transfer of potentially private genomics data on many thousands of individuals, however we have

shown that performance is context dependent and confounded. The results presented are promising and we recommend further study and consideration. In particular, it is our intent to execute increased testing under more controlled conditions to better understand enabling and penalizing factors for SDN-enabled data transfer. Future work includes extending the protocol between the DTC and SDN controllers for supporting additional features, such as improved security methods through key-based authentication prior to flow modification and using dynamically provisioned Layer-2 circuits (such as the ones provided by Internet2 ION service) for lower latency data transfers to the DTN.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] D. Cristofoleti Magri, T. Melo de Brito Carvalho, F. Frota Redigolo, M. Torrez Rojas, M. Simplicio Junior, L. Neumann Ciuffo, G. Dias, A. Soares de Moura, and F. Vetter. Science dmz: Support for e-science in brazil. In *e-Science (e-Science), 2014 IEEE 10th International Conference on*, volume 2, pages 75–78, Oct 2014.

[2] E. Dart, L. Rotman, B. Tierney, M. Hester, and J. Zurawski. The science dmz: A network design pattern for data-intensive science. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, SC '13, pages 85:1–85:10, New York, NY, USA, 2013. ACM.

[3] A. P. Heath, M. Greenway, R. Powell, J. Spring, R. Suarez, D. Hanley, C. Bandlamudi, M. E. McNerney, K. P. White, and R. L. Grossman. Bionimbus: a cloud for managing, analyzing and sharing large genomics datasets. *Journal of the American Medical Informatics Association*, pages amiajnl–2013–002155+, Jan. 2014.

[4] J. Ibarra, J. Bezerra, H. Morgan, L. F. Lopez, D. A. Cox, M. Stanton, I. Machado, E. Grizendi, and L. F. Lopez. Benefits brought by the use of openflow/sdn on the amlight intercontinental research and education network. In R. Badonnel, J. Xiao, S. Ata, F. D. Turck, V. Groza, and C. R. P. dos Santos, editors, *IM*, pages 942–947. IEEE, 2015.

[5] D. Kreutz, F. M. V. Ramos, P. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig. Software-defined networking: A comprehensive survey. *CoRR*, abs/1406.0440, 2014.

[6] S. Levy, G. Sutton, P. C. Ng, L. Feuk, A. L. Halpern, B. P. Walenz, N. Axelrod, J. Huang, E. F. Kirkness, G. Denisov, Y. Lin, J. R. MacDonald, A. W. C. Pang, M. Shago, T. B. Stockwell, A. Tsiamouri, V. Bafna, V. Bansal, S. A. Kravitz, D. A. Busam, K. Y. Beeson, T. C. McIntosh, K. A. Remington, J. F. Abril, J. Gill, J. Borman, Y.-H. Rogers, M. E. Frazier, S. W. Scherer, R. L. Strausberg, and J. C. Venter. The diploid genome sequence of an individual human. *PLoS Biol*, 5(10):e254, 09 2007.

[7] L. D. Stein, B. M. Knoppers, P. Campbell, G. Getz, and J. O. Korbel. Data analysis: Create a cloud commons. *Nature*, 523(7559):149–151, July 2015.

[8] Z. D. Stephens, S. Y. Lee, F. Faghri, R. H. Campbell, C. Zhai, M. J. Efron, R. Iyer, M. C. Schatz, S. Sinha, and G. E. Robinson. Big data: Astronomical or genomical? *PLoS Biol*, 13(7):e1002195+, July 2015.

[9] The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature*, 467:1061–1073, 2010.

[10] The White House. The precision medicine initiative, 2015. Available at: https://www.whitehouse.gov/precision-medicine. Accessed: 2015-08-03.

[11] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. Peterson, R. Roskies, J. Scott, and N. Wilkins-Diehr. Xsede: Accelerating scientific discovery. *Computing in Science Engineering*, 16(5):62–74, Sept 2014.