



Science DMZ network architecture deployment and performance evaluation to large scaled data transfer efficiency

Jong-Seon Park¹, Seung-Hae Kim¹, Min-Ki Noh¹, Bu-Seung Cho^{1*}

¹ Div. of Supercomputing/KREONET Center, Korea Institute of Science and Technology Information, 245 Daehangno, Yuseong, Daejeon, 34141, KOREA

*Corresponding author E-mail: bscho@kisti.re.kr

Abstract

Background/Objectives: Recently, it has been great issue to transfer large-scale science data such as scientific field of high energy physic, astronomical space, super-computing simulation. To solve the transfer issue and to increase transfer efficiency, it needs a multi-dimensional approaches.

Methods/Statistical analysis: To improve the transfer performance, approaches from the perspective of components such as network equipment, transmission protocol, and transmission application have been suggested. Effort to TCP congestion control algorithm and parallelism of data transfer channel are representative example to improve performance. However, the solution through the each component has a limitation in maximizing the transmission efficiency.

Findings: Science DMZ is a new network architecture that can maximize transfer performance. It maximizes transfer efficiency through approach to all components, such as network equipment, dedicated network path, transfer applications, and local institute firewall policies. With these complicated components, science DMZ network architecture can greatly improve the transfer efficiency. In this paper, we design and construct a science DMZ network architecture between two organizations that utilize supercomputing resources based on KREONET and evaluate the performance.

Improvements/Applications: After configuring the experiment environment, we measured network performance through iperf and file transfer performance test through SCP. Experiment result showed around 388% Improvement than that of existing method.

Keywords: Large-Scale Science Data; Transfer Performance; Science DMZ; TCP; UDP Based Transfer; Parallel Data Transfer.

1. Introduction

With the development of network transfer system technology and the expansion of network infrastructure, the demand for user transfer speed has been also increasing. However, user performance is less than expected throughput. Especially, the size of data produced through observation equipment or simulation is very large in scientific field of high energy physic, astronomical space, super-computing simulation. Therefore, increasing the transfer efficiency to scientific large scaled data is a very important and critical issue.

The main reason for this can be found from a multidimensional perspective. In other words, to maximize the end-to-end transmission efficiency, approach to all components of the network is required. For these kind of components that determine the performance of the user, data transmission protocol, transmission application, transmission path, and firewall are represented example. Although TCP (Transmission Control Protocol) is a typical transmission protocol for data transfer, it shows low transfer efficiency in an environment named as LFN(Long Fat Network) where the network bandwidth is large and the data transfer distance is long2-3. And TCP shows very low transfer efficiency from network congestion or packet loss due to any kind of reasons. Therefore, existing studies for TCP performance are mainly focused on reducing packet loss by predicting network congestion. The transfer

efficiency due to the packet loss is very sensitive in the LFN environment. To increase the performance in LFN environment, parallel transfer method has been used. Parallel transfer can maximize performance with a very aggressive transfer mechanism through multiple streams4-5. UDP (User Datagram Protocol)-based transfer application is proposed to replace TCP6-7. In the case of UDP-based transfer applications, additional data transfer reliability and stability functions such as TCP are implemented in the user layer based on the highly aggressive UDP characteristics.

Recently, there have been active studies on the science network model (science DMZ) to increase and maximize transfer efficiency by the research institutes and universities that generate large scaled scientific data8-10. Science DMZ is proposed by the US energy research network (ESNET) and presents an innovative network model based on a data transfer system, high-speed transfer tool, minimized security policy and performance monitoring system.

In this paper, we design and construct a science network architecture between two organizations that utilize supercomputing resources based on KREONET. And we suggests method to maximize efficiency of scientific large scaled data transfer. The proposed science network architecture can increase the transfer efficiency by reuse the existing network environment with minimum cost and effort.

The composition of this paper is as follows. In Section 2, we discuss the high speed transfer method of scientific data. And de-

scribe proposed network architecture of science DMZ in Section 3. In Section 4, we measure the network performance and file transfer performance and analyze the measured results. In Section 5, we conclude the paper.

2. High speed data transfer to user performance

So far, many researches to transfer efficiency of end to end users have been carried out. The main network components of these researches are summarized as transfer protocols, data transfer systems, and transfer applications. Especially, research for TCP transfer efficiency is the main issue because most of user data transfer applications are based on TCP. In the case of TCP, reducing the packet loss is the best way to increase the performance since the transfer efficiency due to the packet loss is drastically reduced. Estimating network conditions using RTT (Round Trip Time) or OWD (One Way Delay) can significantly reduce this packet loss rate. A sudden change in RTT can prevent packet loss by reducing the amount of packets determined by congestion in the bottleneck region. One of the most advanced methods of TCP research is to predict the available bandwidth while predicting network conditions. That is, the available bandwidth of the current network is calculated using RTT or OWD and the amount of transfer packet can be accurately set to the value of the available bandwidth calculated even if packet loss occurs. This can dramatically increase the transfer efficiency than the artificially reducing the amount of packets at the congestion avoidance stage after the packet loss. Transport protocols such as high speed TCP (HSTCP), fast TCP (FTCP), and TCP-Illinois are typical transfer protocols that improve performance through available bandwidth prediction.

However, transmission performance is limited in the LFN environment even with this improved transfer protocols. Main method for improving the transfer performance in the LFN environment is a parallel transfer technique of using multi-streams. Data transfer through multiple streams can be very aggressive by reducing the latency of response waiting time of TCP. Figure 1 and Figure 2 show the comparison of data transfer through single stream and multi stream.

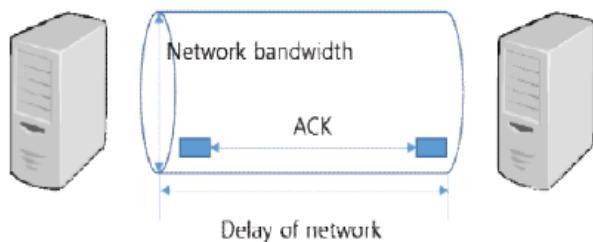


Fig. 1: Data Transfer with Single Stream.



Fig. 2: Data Transfer with Multi Stream.

Another alternative for high speed transfer in the LFN environment is to use UDP. In the UDP-based transfer application, continuous data transfer is possible. But it is difficult to guarantee data transfer reliability. Therefore, a separate function such as TCP reliability is additionally implemented in the user layer. Ta-

ble 1 compares the characteristics of TCP-based and UDP-based applications.

Table 1: Comparison of Characteristics According to Transfer Application

	data transfer channel	amount of packet	congestion control
TCP based application	data transfer and control message with single channel	amount of packet according to available bandwidth	waiting ACK from receiver after transfer data-congestion avoidance after packet loss
UDP based application	distinction of channels for data and control message	packet sending period according to available bandwidth	continuous data transfer for certain period-control packet loss at the same time after certain period

3. Design and deployment of science network architecture

Recently, a community that generate scientific large-scaled data can utilize from 10Gb/s to 100Gb/s network bandwidth for global data transfer and data sharing among researchers. In order to increase the transfer efficiency of user performance in the scientific LFN environment, it is necessary to access not only the transfer application but also the complicate components constituting the network. In case of United States, it has been deployed and operated science network model for large data transfer starting with the CC*IIIE program. ESnet (Energy Sciences network) proposed a science DMZ network architecture optimized for transfer of large scientific data. Science DMZ is considered as a network model considering end to end user performance through comprehensive approach to the components that constitute the network.

The science DMZ network model proposed by ESNET includes a dedicated network for large-scaled data transfer, a data transfer system, a data transfer application and an internal firewall policy of organization. In the case of a dedicated network, it is possible to utilize the available bandwidth by isolating large-scaled traffic from general traffic. In the case of user transfer applications, high-speed data transfer tools such as globus-online or GridFTP are used. In addition, complicated security policies should be eliminated in order to eliminate bottlenecks from the WAN (Wide Area Network) which are dedicated networks to the local area. Therefore, in order to maximize the efficiency of large-capacity data transfer, approaches to all the components for end-to-end data transfer is required. Figure 3 shows the network components for end-to-end data transmission.

The science DMZ component of figure 3 can be summarized as follows.

- Separation of networks for large-scaled data traffic and general data traffic-Dedicated networks for scientific data transfer from existing network environments should be separated to quickly transmit large-scaled observations and simulation data generated by the research community.
- Routing for wide area network (WAN) connections Optical path (light-path / layer1), virtual LAN (vlan and trunk / layer2), virtual circuit and software-defined networking (SDN) technology should be applied.
- Application of lightweight network security technology without performance degradation Security technology that does not cause degradation such as network traffic access control (ACL), threat traffic monitoring should be applied.
- Data transfer node (DTN)-High-performance data transfer system with high-capacity specification should be equipped.
- Network monitoring system-Monitoring of network segments and data transfer nodes should be performed in real time.

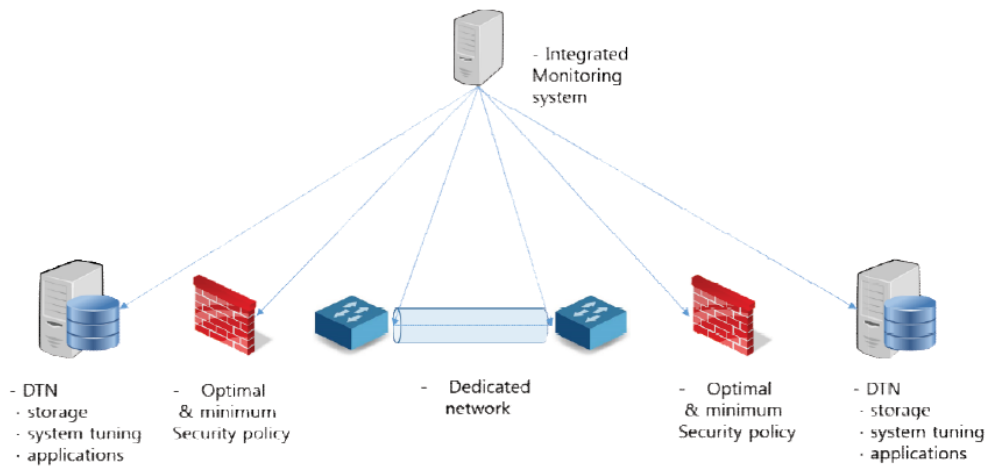


Fig. 3: Integrated Network Components for Scientific Data.

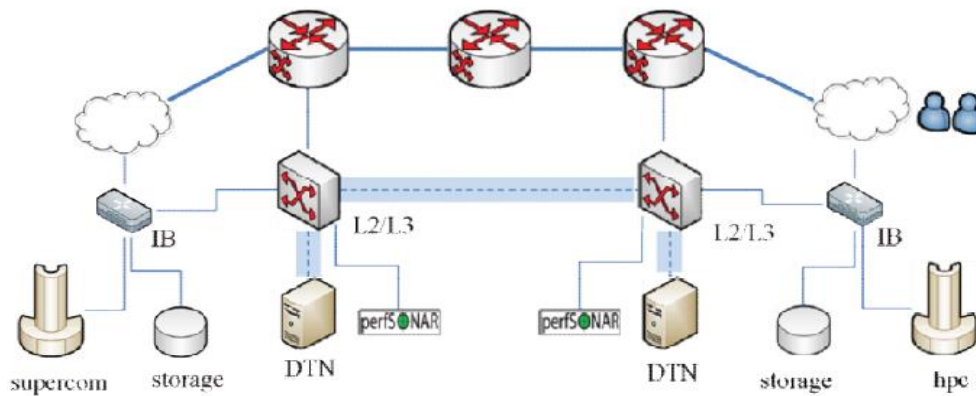


Fig. 4: Deployment of the Science DMZ between Two Organizations through KREONET.

Figure 4 illustrates a science DMZ network architecture based on KREONET between two organizations of super-computing center and APEC climate center. The data transmission between the two organizations is downloaded from super-computing center to the APEC Climate Center. Without deployment of science DMZ, it is difficult to utilize available bandwidth efficiently due to congestion resulting in packet loss. In addition, complex firewall policies in both organizations cause packet loss resulting in performance degradation due to the bottleneck between KREONET and local organization network. In the case of TCP-based transfer tools such as SCP or FTP, the bottleneck must be removed because the transfer efficiency is significantly degraded due to small packet losses. The components of the newly deployed scientific data transfer network are DTN, a performance measurement tool (perfSONAR), an L2 / L3 switch for internal interworking with KREONET, a minimum security policy. The dedicated path for data transfer can be configured via the virtual LAN of the KREONET backbone router. It is meaningful that the existing network transfer system is reallocated and the scientific data transfer path is established with minimum cost. As shown in figure 4, the infiniband (IB) switch for processing large-scaled simulation results is connected with the scientific data transfer switch in the supercomputing center interworking. In addition, the scientific data transfer switch is connected with a large-capacity data transfer node. Using the dedicated network and dedicated transfer nodes deployed from figure 4, the available bandwidth from the KREONET backbone network can be fully utilized.

4. Experiment result

This section evaluates and analyzes the performance of the deployed science network. Performance is measured using a performance monitoring server connected to the organization switch. Table 2 shows the specifications of the performance monitoring server.

Table 2: Specification of Performance Monitoring Server

Specification	Description
operating system	CentOS release 6.8 (Final)
cpu	Intel(R) Core(TM)2 Quad CPU Q9300 @ 2.50GHz × 4
memory	2GiB DIMM DDR2 Synchronous 800 MHz × 2
hdd	HDD/500G/SATA 800Mbps
network interface card	Intel Corporation 82573E Gigabit Ethernet Controller 1Gbps

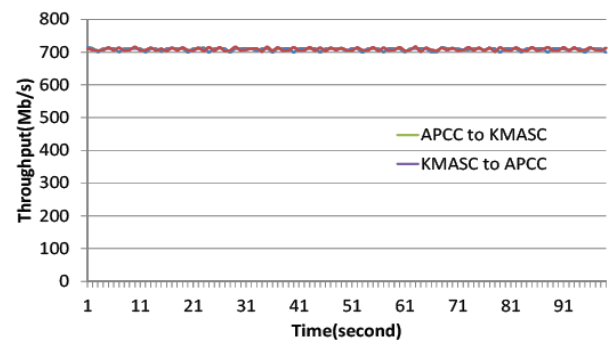


Fig. 5: Bi-Directional Performance Measurement Result with Science Network.

Figure 5 shows the performance results measured after deploying the science network. In the newly deployed network, it can be confirmed that the 700Mbps available bandwidth provided by the testbed is fully utilized since the minimum firewall policy of the organization is applied and the dedicated path for data transfer is used.

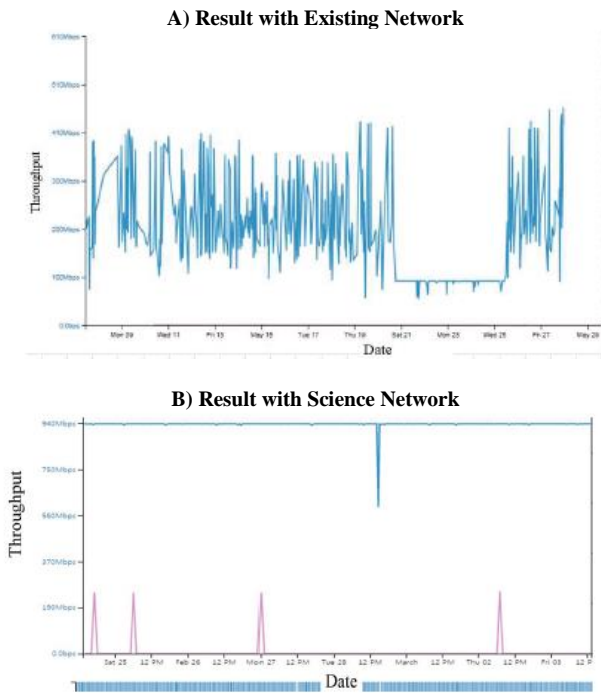


Fig. 6: Performance Results from Super-Computing Center to APEC Climate Center with Perfsonar.

Figure 6 shows the results of one week monitoring result with perfSONAR. Performance measurements of perfSONAR are measured every 20 minutes for 30 seconds. The left side of figure 6 shows the performance result with existing network resulting in fluctuated throughput. The right side of figure 6 shows the performance result with science network. The network performance shows that the available bandwidth of 1 Gb/s is utilized efficiently enough.

5. Conclusion

Transfer efficiency is a very critical issue to transfer large-scaled data such as scientific data. End-to-end transfer performance is determined by various network components such as transfer paths, data transfer systems, storage, transfer application tools, and fire-wall policies of the organization. Most of the existing researches have focused on the transfer protocols that are based on transfer applications. However, a comprehensive components must be considered in order to maximize performance. In this paper, we designed and deployed a scientific data transfer network based on KREONET. To do this, KREONET backbone link is directly connected to data transfer system. And minimum security policy is applied in each organizations. This is a method to increase transfer efficiency with minimal cost while making maximum use of existing network. The network performance from super-computing center to the APEC climate shows the improvement of 363% compared with that of existing network.

References

- [1] C. Jayashri, P. Abitha, S. Subburaj, S. Yamuna Devi, S. Suthir, S. Janakiraman, Big Data Transfers through Dynamic and Load Balanced Flow on Cloud Networks, 2017 Third International Conference on AEEICB.
- [2] T. Kelly, Scalable TCP: improving performance in high speed wide area networks, Computer Communications Review, 2003, 32(2).
- [3] D. Kliazovich, F. Granelli, and D. Miorandi, Logarithmic Window Increase for TCP Westwood+ for Improvement in High Speed, Long Distance Networks, Computer Networks, 2008, 52(12), pp. 2395-2410.

- [4] M. A. Alrshah, M. Othman, Performance Evaluation of Parallel TCP, and Its Impact on Bandwidth Utilization and Fairness in High-BDP Networks based on Test-bed," in Proc. 2013 Malaysia International Conference on MICC, pp. 23-28.
- [5] J. Park, D. An, G. Cho, an Adaptive Channel Number Tuning Mechanism on Parallel Transfer with UDT, 2013 International Conference on ICOIN.
- [6] Y. Gu, R. L. Grossman, UDT: UDP-based Data Transfer for High Speed Wide Area Networks, Computer Networks, 2007. 51(7), pp. 1777-1799.
- [7] M. Masirap, M. H. Amaran, Y. M. Yussoff, R. Rahman, H. Hashim, Evaluation of Reliable UDP-based Transport Protocols for Internet of Things (IoT), 2016 IEEE Symposium on ISCAIE.
- [8] E. Dart, L. Rotman, B. Tierney, the Science DMZ: a Network Design Pattern for Data-Intensive Science, Supercomputing Conference, 2013.
- [9] I. Monga, E. Pouyoul, C. Guok, Software-Defined Networking for Big-Data Science-Architectural Models from Campus to the WAN, 2012 SC Companion on SCC.
- [10] H. Neeman, TheOneOklahoma Friction Free Network: Towards a Multi-Institutional Science DMZ in an EPSCoR State, 2014 Annual Conference on ESEDE.