Managing Distributed Storage System through Network Redesign

Tahani H. Hussain
System and Software Development
Kuwait Institute for Scientific Research
P.O. Box 24885 Safat 13109 Kuwait
thussain@kisr.edu.kw

Paulvanna N. Marimuth and Sami J. Habib
Computer Engineering Department
Kuwait University
P. O. Box 5969 Safat 13060 Kuwait
sami.habib@ku.edu.kw

***Abstract*—In this paper, we have assessed the performance of an existing enterprise network before and after deploying distributed storage systems through a co-simulation environment integrating NS-2 with Simulink. The existing enterprise networks can benefit from replacing the expensive monolithic disk array systems with low-cost and scalable distributed storage systems. The storage replacement task is a challenging redesign problem, where we need to ensure the suitability and performance sustainability of existing network to support distributed storage systems. We have chosen three performance parameters such as average response time, packet loss rate and storage system throughput as the outcomes of the co-simulation, and these parameters can guide the redesign process by maximizing the storage system throughput with the intuition to reduce average response time and packet loss rate. The simulation of an enterprise network with 680 clients and 54 servers followed by redesigning shows improvements in the storage system throughput by 13.9% with 24.4% decreased average response time and 38.3% reduced packet loss rate.**

Keywords—redesign; storage systems; simulation; NS-2; Simulink; performance; network

I. INTRODUCTION

The exploitation of information technology (IT) in businesses has grown to a larger extent and the success of these businesses depends directly on their ability to store, process and retrieve data. Moreover, the growth of distributed systems is a major concern in many organizations as the standards of support in a devolving environment are not always adequate. Consequently, the storage growth is increasing at an alarming rate when combined with more servers to support storage and services, such situation is unmanageable as far as the storage management is concerned. On top of that, increasing in storage cost leverages high demands for efficient and cost-effective storage systems.

The traditional approach to storage along with the local area network (LAN) and wide area network (WAN) can't handle data-storage load or provide data reliability, availability, scalability or management required today. Distributed storage solutions, such as storage area networks (SAN), can offer flexibility for connecting storage, ensure much greater utilization of disk storage space and support for server consolidation. In our previous effort [1], we revealed that consolidation of servers and storages in an existing distributed network could be solved by considering SAN as an alternative. In addition, with SAN, the operational cost of the enterprise network drops. Nevertheless, the success of implementing SAN mainly depends on the existing network characteristics and also, the performance of the underlying network. We have extended our earlier work [1] on introducing SAN into an existing enterprise, by introducing a co-simulation model to study the impact of SAN on the existing network performance and also the impact of topology redesign on the performance of SAN and the underlying network. Exploring alternative network redesign solutions can provide a better understanding of the impact on the underlying network by the performance of the storage system.

In this paper, we have proposed a simple methodology for redesigning existing network topology to enhance the performance of SAN and the underlying network. First, we assess the performance of existing network before implanting SAN system. Then, we assess the performance of the existing network after deploying SAN. We evaluate the performance within a proposed co-simulation environment integrating network simulator (NS-2) and Matlab/Simulink simulation tool. NS-2 works as a master simulator, and it is utilized to model and to simulate the data transfer over the existing network, whereas Simulink is utilized to model the data process within the storage systems. The performance of the network is analyzed by investigating the average response time and packet loss rate of existing network before and after implementing SAN. Finally, the combined results are utilized to propose a redesign solution intended to enhance the network and SAN performances by evaluating the network performance along with SAN throughput.

The rest of this paper is organized as follows: section II contains the related works, focused mainly on storage systems and network redesign. Section III provides the background of SAN and simulation tools. Section IV describes the proposed redesign network topology considering the storage system methodology. Section V elaborates the selected case study along with the simulation results of proposed methodology. Finally, Section VI presents the summary of the proposed work and future research directions.

II. Related Work

Many researchers had attempted to address various issues in replacing traditional storage system with distributed storage system [2-4]. SAN started with the traditional small computer system interface (SCSI) technology and it moved to fiber channel technology [2]. The earlier stages of research were focused on the implementation of SAN within the enterprise networks [3] and analyzed the advantages of fiber channel technology over traditional SCSI technology [4].

Several approaches [5-9] had been proposed to carry out simulation studies for analyzing the performance of storage systems. Additionally, performance models were also proposed by few researchers [10-13] to evaluate the performance of distributed storage systems. The authors in [5] developed a SAN simulator, which has considered several interconnection topologies, real-world I/O traces and synthetic I/O traffic, failures in links, different routing algorithms and switch architectures. While, the modeling and simulation of storage systems were done through N-SPEK [6], ParlOSim [7], and NS-2 [8] simulators; Hatnik and Altmann [9] introduced the concept of simulator coupling and they integrated Matlab/Simulink, ModelSim and network simulator (NS-2) to provide a co-simulation environment, which was used to simulate complex heterogeneous systems and to study the performance of the storage systems. Aizikowitz et al. [10] presented a component based performance modeling focusing on the performance of I/O interactions of host servers and storage subsystems via the SAN fabric. Authors in [11] and [12] discussed the root cause analysis of performance problems in SAN. Similarly, authors in [13] discussed strategies to optimize storage area network for high performance.

Network redesign operations have mainly aimed to redistribute the resources or to add/delete new/existing resources [14] [15] [16], whereby the authors have attempted to analyze the performance of the redesigned network through optimization algorithms. In another work, the authors have adapted the redesign methodology based on the concept of keeping virtual machines for hosting services only and deploying SAN for storage unification [1].

However, our previous works attempted to study the performance of the storage system only. Hereby, first, we proposed a framework for assessing enterprise network performance along with distributed storage system performance using a co-simulation model integrating NS-2 with Matlab/Simulink. Second, we redesign the network topology not only to enhance the performance, but also to ensure its suitability and performance sustainability to support distributed storage systems. Third, we use the simulation results as guides for the redesign process.

III. Background-Storage and Simulation Tools

A. Storage Area Network

A storage area network is a dedicated storage network intended to optimize the entire communications between network servers and storage disk arrays. SAN permits servers of different types and platforms to access the same storage simultaneously. Also, additional storage is independent of existing servers and the amount of storage that can be scaled up is virtually unlimited. With SAN, all storage systems can be monitored and managed by a single management console. Finally, by separating the servers from the storage, the problem of underutilized disk space is eliminated, since the storage devices are now consolidated.

B. Network Simulator

Network simulator NS-2 has the capabilities in modeling network topologies, network protocols, routing algorithms, packet scheduling schemes, as well as traffic generation methods [17]. It has been broadly used in the field of computer network design and simulation.

C. Matlab/Simulink

Simulink is a graphical programming language tool for modeling, simulating, and analyzing dynamical systems [18]. It supports linear and nonlinear systems, either in a continuous time, sampled time, or hybrid of the two. The most important feature of Matlab is its flexible programming capability, which allows to add user-developed functions and to access C language codes using external interfaces.

IV. Redesign Methodology

The first objective of the proposed redesign methodology is to address issues in the performance of the underlying network while deploying SAN. Also our methodology analyzes the impact of replacing traditional storage system with SAN through simulation study. The second objective is aimed to study, whether redesigning of underlying network topology will enhance the performance of enterprise network and SAN. The third objective determines the set of modifications necessary to carry out within the existing network topology to optimize its performance along with the storage system performance.

A. Performance Measurements

We have carefully chosen the performance measurements pertinent to network and storage traffics to evaluate the performances before and after deploying SAN. We measure the average response time and packet loss rate for evaluating the network performance and we measure the storage system throughput to assess the storage system performance.

The average response time ($\boldsymbol{\alpha_t}$) is defined as the average time taken by the enterprise network from sending request by the client till receiving the reply. In our work, the response time $\boldsymbol{\alpha_t}$ is calculated using Equation (1), whereby $\boldsymbol{S_t}$ is the total time taken by the distributed network to fulfill all the clients' requests and $\boldsymbol{R_t}$ is the total clients' requests during the simulation time $\boldsymbol{t}$.

$$\alpha_t = (S_t / R_t) \tag{1}$$

We calculate the packet loss rate ($\boldsymbol{\rho_t}$) using Equation (2), which is defined as the ratio of the total dropped packets $\boldsymbol{D_t}$ to the total generated packets $\boldsymbol{G_t}$ during the simulation time $\boldsymbol{t}$. We measure ρ_t by considering the entire traffic generated during the simulation time regardless of the traffic type and directions. The packets, which are initially dropped and successfully retransmitted, do not count as lost packets. However, these packets increase the average response time.

$$\rho_t = (D_t / G_t) \times 100\% \quad (2)$$

The storage system throughput (β) is defined as the amount of data processed in a given time interval, and it is represented in megabytes per second (MB/s). The storage system with better response time is resulted from better storage processing. Nowadays it's slightly more complex with system having highly sequential data transfer of workloads, such as database queries, video or audio, throughput, wherein higher data transfer rate can lead to much faster workloads. Let B_s be a constant integer number that represents the data request block size and B_T be the total number of blocks to be sent during T data processing seconds. Since, the amount of processed data by the data storage system is calculated by multiplying B_s by B_T, β can be calculated using Equation (3).

$$\beta = (B_s \times B_T) / T \quad (3)$$

B. Co-simulation Model

The contribution in this paper is utilizing the co-simulation environment for a simple distributed network considering traditional storage system and SAN, as illustrated in Fig. 1 and 2. The co-simulation model integrates NS-2 with Matlab\Simulink, whereby NS-2 is used to model and simulate the communication protocols and also the traffic related to the data transfer within the distributed network. The data handling and processing within the storage system is modeled and simulated using Matlab/Simulink simulator. NS-2 works as a master simulator and as an interface to the co-simulation model. The exchange of simulations' data and events between the two simulators and a program script written in C programming language are integrated.

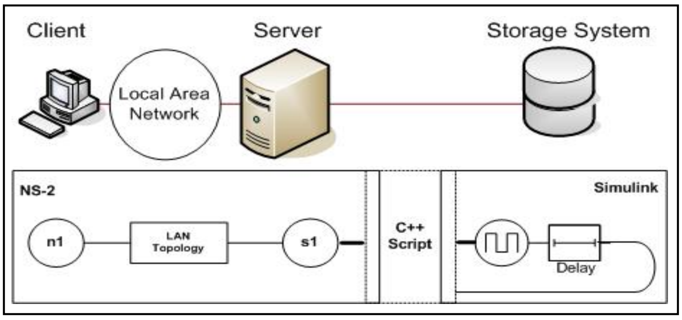

Fig. 1. Co-simulation model for distributed network supporting traditional storage system.

The distributed network consists of a set of clients, servers, and storage system along with the interconnecting network topology. The network components are modeled using NS-2. The simulation procedure includes building the nodes, the interconnecting links and the related communication protocols. The data process is formulated and modeled using Matlab/Simulink blocks. The processing model consists of clock and time delay block to closely resemble the real environment. The client node (n1) sends a request to a server (s1) through LAN topology using UDP/TCP communication protocols, which is forwarded by the server node (s1) to the storage system model. The node s1 in Fig. 1 acts as an interconnecting node between the network traffic in NS-2 and the data process in Simulink, whereas in Fig. 2, s1 is connected to a private network with SAN and a virtual node (v) is created to act as the interconnecting node. In both figures, the s1 is the node that it is responsible to satisfy the request by sending the response back to n1.

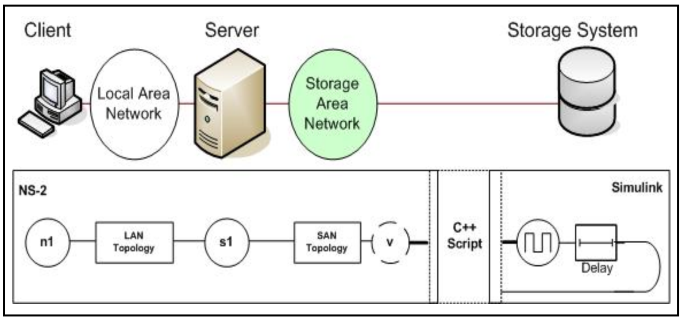

Fig. 2. Co-simulation model for distributed network supporting SAN.

C. Redesign Process

Redesign process is aimed to maximize the storage system throughput by reducing the average response time and the packet loss rate through reconfiguring the network topology. The proposed redesign process includes: addition of one or more network device or links, geographical relocation of one or more nodes (clients or servers) and change in the flow path over the network. The decision on redesign operation depends on the outcome of the co-simulation model and also the performance analysis.

The response time is defined as the round-trip time to send a request and to receive a response (i.e. the latency plus the processing time), is used to evaluate the impact of deploying distributed storage system and to identify the source of creating delay. The loss rate is the ratio of packets failed to reach a destination to the total number of packets sent to that destination, which is used to define the bottleneck of the network. The network redesign decisions are based on the performance measurements as described below:

1) Network devices with long response time are considered as the sources of delay and they need to be redesigned

2) Network devices with high packet loss rate are considered as the bottlenecks and they need to be redesigned.

The proposed redesign process is illustrated in Fig. 3, wherein, it starts by identifying the specifications of existing network before (ENB) deploying SAN. It proceeds to identify the specifications of existing network after (ENA) deploying SAN. Further, ENB and ENA are simulated within the co-simulation environment to assess the performance. Subsequently, we compare and analyze the simulation outcomes and based on these outcomes, a redesign process is performed on ENA to enhance the performance. This model iterates of the redesign procedure; at the end of the iteration the design with the highest performance measurements is selected as the model optimum solution.

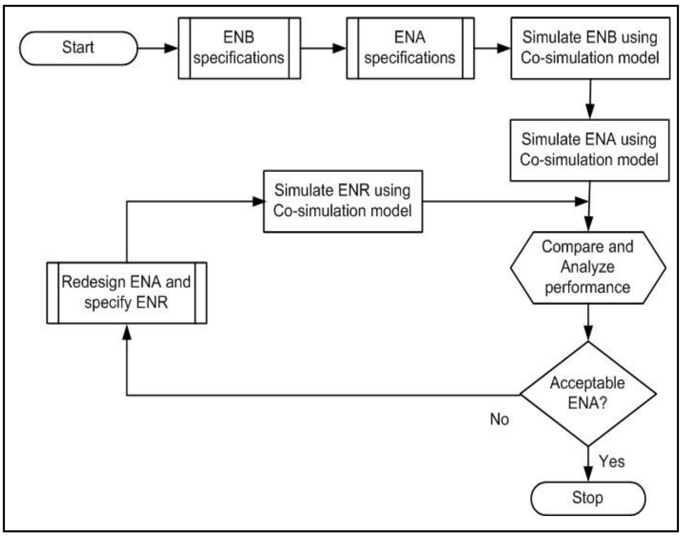

Fig. 3. The proposed network redesign methodology.

V. Experimental Results

The efficiency of our methodology has been evaluated on a medium size existing network (ENB) consisting of 680 clients, 54 physical servers hosting 54 services, 13 Layer-2 switches (L2), 5 Layer-3 switches (L3) and a router (R), as illustrated in Fig. 4.

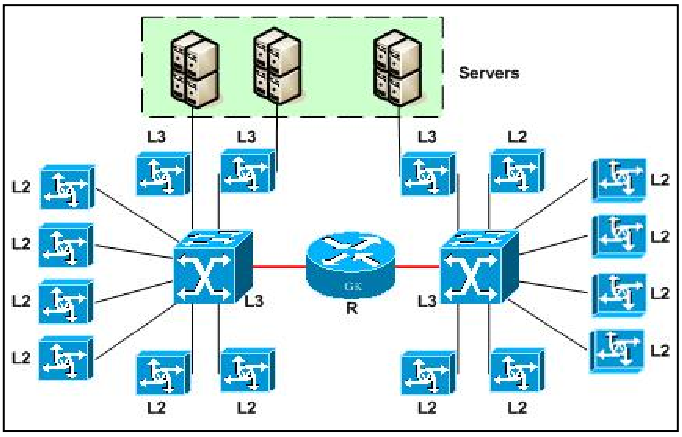

Fig. 4. Existing Network Before (ENB).

Two of the Layer-3 switches are used to connect the switches of each side and the others (3 Layer-3 switches) are used to connect the servers. Full-duplex links of 100Mbps capacity with queuing delay of 10 ms were employed to establish the connections between the clients, servers, Layer-2 and Layer-3 switches. The links between Layer-3 switches and the router are configured with 1 Gbps link using OSPF (Open Short Path First) routing protocol. Each switch is configured with a single virtual local area network (VLAN) such that the devices connected to the switch are part of its particular VLAN.

To apply SAN in ENB, previous work [1] is used to optimize the server and storage consolidation considering SAN for ENB. The solution from [1], which is ENA, consists of 25 physical servers hosting 54 services serving the 680 clients with a 10 Terabyte of SAN storage capacity are deployed. The SAN LAN has been configured with new 3 Layer-3 switches and a router as shown in Fig. 5.

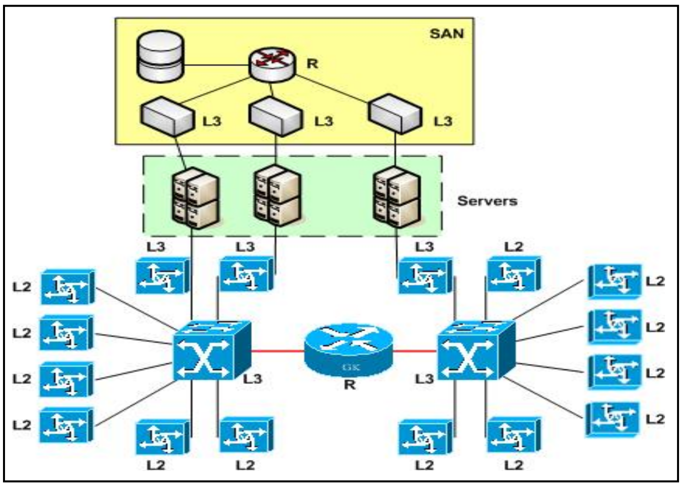

Fig. 5. Existing network after (ENA).

We have considered three network scenarios: existing distributed network (ENB), distributed network with SAN (ENA) and redesigned network after deploying SAN (ENR). ENB and ENA are compared against storage throughput as shown in Fig. 6. The simulation of ENB demonstrates that the maximum storage throughput is 98 MB/s and the minimum is 7 MB/s. It decreases below 75 MB/s at an early stage of the simulation time (45). On average, the storage throughput in ENB is 40 MB/s. However, in ENA, the maximum, minimum and average throughputs are 129, 64 and 95 MB/s respectively. Thus, the network with SAN shows that on average the storage throughput is enhanced by 58%.

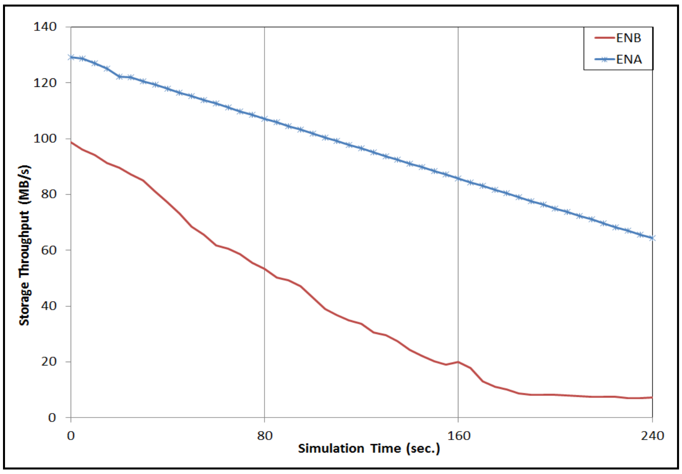

Fig. 6. Storage throughput of ENB and ENA.

ENB and ENA are compared against the average response time in Fig. 7. It is observed that the average response time is affected undesirably by deploying SAN, whereby it is increased up to 20.9%. The average response time exceeded

60 sec. in ENB at simulation time 220 seconds and in ENA at simulation time 195 seconds. It is seen that the average response time reaches a maximum of 77seconds and 97 seconds in ENB and ENA respectively.

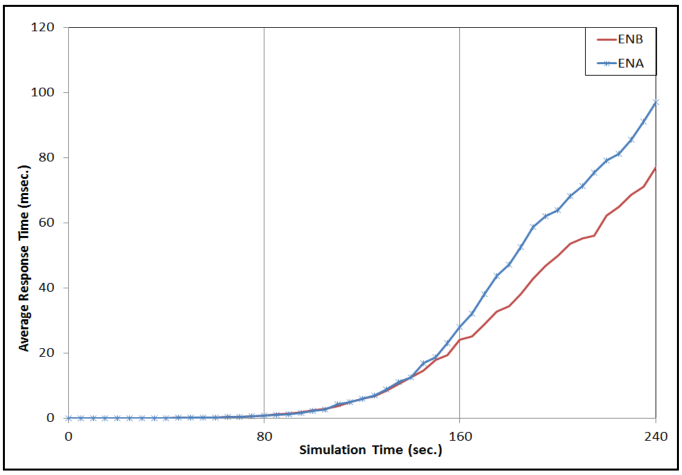

Fig. 7. Average resposne time of ENB and ENA.

The loss rate is evaluated on each network devices at the end of the simulation time. The minimum, average and maximum recorded packet loss rate are tabulated in Table I. Although the minimum loss rate has been reduced by deploying SAN, the average and maximum loss rate values seriously affected. ENA has increased the average loss rate in ENB by 14.2%. The packet loss rate greater than 50% is recorded in Layer-3 switches and in the main router.

TABLE I. PACKET LOSS RATE FOR ENB AND ENA.

Value	**Network**	
	ENB	***ENA***
Minimum	2.8%	0.0%
Average	25.5%	29.7%
Maximum	63.4%	71.2%

From the above results, we observed that although deploying SAN has improved the storage throughput, the ENA topology has failed to meet other network performances. Moreover, we noticed that the router and the Layer-3 switches are the main bottlenecks and they are the sources of increased response time. In addition, the number of connections in the two Layer-3 switches connecting the main router in ENB is not balanced (i.e. one switch is with 402 connections and the other is with 332 connections).

Thus, we applied load balancing and redundancy solution to ENA topology to eliminate the bottlenecks and sources of response time degradation. Also, allocation to the nominated servers and clients are applied to provide some balancing in the number of connections in each Layer-3 switch side (one switch is with 348 connections and the other is with 357 connections). Finally, Layer-3 switch along with dedicated VLAN are employed in parallel to the main router to serve client-to-client traffics without prorogating through the main router. Fig. 8 illustrates the proposed redesign solution (ENR).

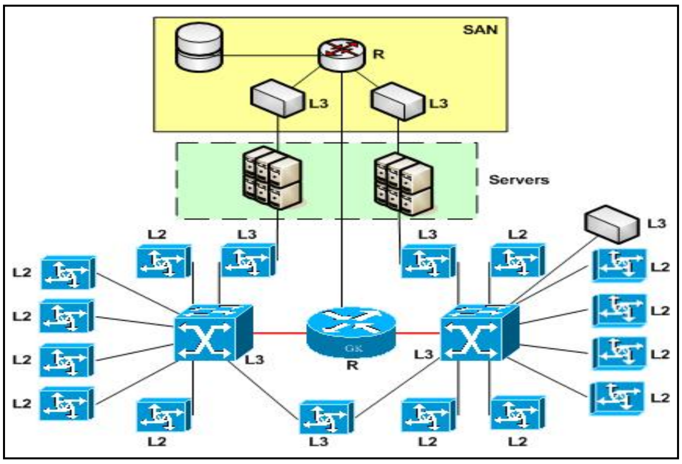

Fig. 8. Existing network redesined (ENR).

The simulation of each network scenario is ran for 240 simulation seconds, and the networks ENB, ENA and ENR are compared against the storage throughput, average response time and packet loss rate (Fig. 9-10 and Table II).

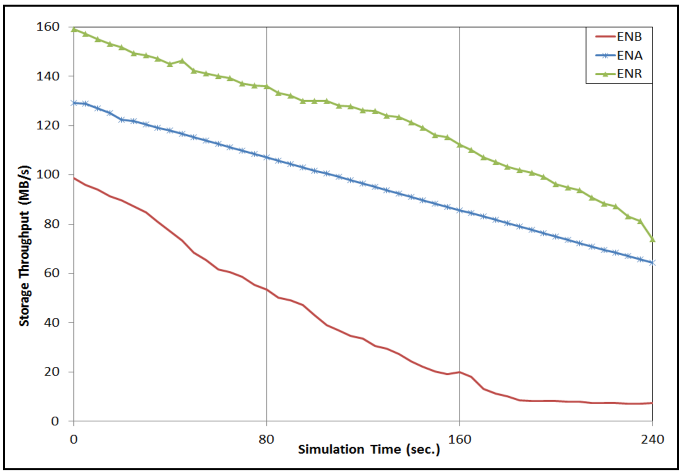

Fig. 9. Storage throughput of ENB and ENA.

From Fig. 9, we noticed that the storage throughput starts decreasing below 75 MB/s at an early stage of the simulation in ENB, whereas ENR maintains higher level of storage throughputs within the stipulated period of simulation. The maximum, minimum and average storage throughputs of ENR are 137, 74 and 112 MB/s. Thus, on average ENR had enhanced the throughput of ENA by 13.9%.

From Fig. 10, it is shown that the proposed redesign solution reduces the average response time of ENA by 24.4% and it is closer to ENB with marginal difference of 1.28%. The average response time exceeded 60 seconds. in ENR at simulation time 210 compared to 195 in ENA and 220 in ENB.

Remarkably, from Table II, we noticed that the maximum loss rate on the redesigned network is reduced by 50% compared to ENA. Nevertheless, on average, the redesign network reduced the packet loss rate by 38.3%.

Finally, we can conclude that redesigning the existing network after deploying SAN has greater impact on the network performance and storage throughput.

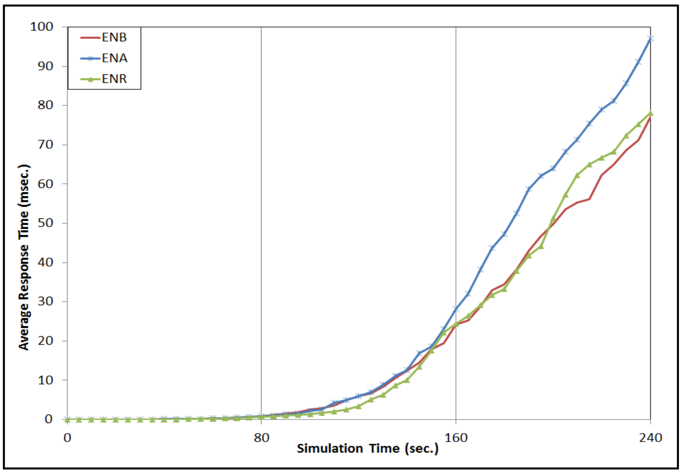

Fig. 10. Average resposne time of ENB and ENA.

TABLE II. PACKET LOSS RATE FOR ENB AND ENA

Value	**Network**		
	ENB	***ENA***	***ENR***
Minimum	2.8%	0.0%	0.1%
Average	25.5%	29.7%	21.4%
Maximum	63.4%	71.2%	31.1%

VI. CONCLUSIONS AND FUTURE WORKS

In this paper, we have proposed a methodology to assess the performance of the existing enterprise network before and after deploying distributed storage systems through co-simulation environment. Existing network is redesigned based on the performance measurements to ensure the suitability and performance sustainability of the existing network to support distributed storage systems.The co-simulation environment integrates NS-2 with Simulink, and it measures the average response time, the packet loss rate and storage system throughput pertain to the network to ensure the performance sustainability of existing network. The outcome of simulation of an enterprise network with 680 clients and 54 servers followed by redesigning demonstrates improvements in the storage system throughput by 13.9% with 24.4% reduced average response time and 38.3% reduced packet loss rate.

Our future work includes testing the efficiency of the redesign model on more sophisticated distributed network having variant workloads, incorporating the storage system communication protocols in the co-simulation model and considering other related network and storage systems performance criteria.

ACKNOWLEDGMENT

This work was supported by Kuwait University, Research Grant No. EO 05/12.

REFERENCES

[1] T. Hussain and S. Habib, "A Redesign Methodology for Storage Management Virtualization," In the Proceedings of IFIP/IEEE 18th International Symposium on Integrated Network Management, Ghent, Belgium, 27-31 May 2013.

[2] B. Phillips, "Have Storage Area Networks Come of Age?," Computer, vol. 31, no. 7, pp.10-12, Jul 1998.

[3] S. Milanovic and Z. Petrovic, "Building the Enterprise-wide Storage Area Network," IEEE International Conference on Trends in Communications, Bratislava, Slovakia, vol. 1, pp. 136-139, 4-7 July 2001.

[4] X. Molero, F. Silla, V. Santonja and J. Duato, "A Tool for the Design and Evaluation of Fibre Channel Storage Area Networks," In the Proceedings of 34th Annual ACM/IEE International Symposium on Simulation, Austin, TX, USA, pp. 133-140, 2001.

[5] R. Buyya, T. Cortes, and H. Jin, "Modeling and Evaluation of Fibre Channel Storage Area Networks," High Performance Mass Storage and Parallel I/O:Technologies and Applications, Wiley-IEEE Press, 2002, pp. 464 -473.

[6] M. Zhang, Q. Yang and X. He, "Performability Evaluation of Networked Storage Systems using N-SPEK," In the Proceedings of 3rd IEEE/ACM International Symposium on Cluster Computing and the Grid, Tokyo, Japan, pp. 736-741, 12-15 May 2003.

[7] NASA Advanced Supercomputing Division: NAS Parallel Benchmarks, http://www.nas.nasa.gov/publications/npb.html, Accessed 4 March 2013.

[8] J. Liu, "Approaches to Network Simulation Based on NS," Application Research of Computers, vol. 9, pp. 54-57, 2002.

[9] U. Hatnik and S. Altmann, "Using ModelSim, Matlab/Simulink and NS for Simulation of Distributed Storage Systems," In the Proceedings of IEEE International Conference on Parallel Computing in Electrical Engineering, Dresden, Germany, pp. 114-119, 7-10 September 2004.

[10] N. Aizikowitz, A. Glikson, A. Landau, B. Mendelson and T. Sandbank, "Component-based Performance Modeling of a Storage Area Network," In the Proceedings of the 37th Winter Simulation Conference, Orlando, FL, USA, pp. 2417-2426, 4-7 December 2005.

[11] D. Breitgand, E. Henis, E. Ladan-Mozes, O. Shehory and E. Yerushalmi, "Root-cause Analysis of SAN Performance Problems: an I/O path Affine Search Approach," In the Proceedings of IFIP/IEEE 9th International Symposium on Integrated Network Management, Nice, France, pp. 251-264, 15-19 May 2005

[12] R. Routray, S. Gopisetty, P. Galgali, A. Modi and S. Nadgowda, "iSAN: Storage Area Network Management Modeling Simulation," In the Proceedings of IEEE International Conference on Networking, Architecture, and Storage, Guilin, China, pp. 19-208, 29-31 July 2007.

[13] J. Choi, Y. Yu, H. Eom, H. Yeom and D. Shin, "SAN Optimization for High Performance Storage with RDMA Data Transfer," In the Proceedings of High Performance Computing, Networking, Storage and Analysis, Honolulu, Hawaii, USA, pp. 24-29, 10-16 November 2012.

[14] T. Hussain and S.J. Habib, "Capacity Planning of Network Redesign – a Case Study," In the Proceedings of International Symposium on Performance Evaluation of Computer and Telecommunication Systems, Ottawa, Canada, pp. 52-57, 11-14 July 2010.

[15] S.J. Habib, P. N. Marimuthu, and N. Al-Awadi, "Network Redesign through Clusters Consolidation," In the Proceeding of the International Symposium on Performance Evaluation of Computer & Telecommunication Systems, Istanbul, Turkey, pp. 333-338, 13-16 July 2009.

[16] S. Saleem, P. N. Marimuthu and S.J. Habib, "Capacity Planning of Enterprise Information System Through Simulation," In the Proceeding of the International Symposium on Performance Evaluation of Computer & Telecommunication Systems, pp. 1-6, 8-11 July 2012.

[17] The Network Simulator NS-2: Virtual Internetwork Test bed Collaboration, http://www.isi.edu/nsnam/vint/, Accessed 25 March 2013.

[18] The Matrix Laboratory Matlab, "Simulink: Simulation and Model-base Design," http://www.mathworks.com/products/simulink/, September 2012.