academicJournals

Vol. 3(2), pp. 16-19, September, 2013
DOI: 10.5897/JIIS13.0069
ISSN 1684–5315 © 2013 Academic Journals
http://www.academicjournals.org/JIIS

Journal of Internet and Information Systems

Review

Review on network performance: Meaning, quantification and measurement

Omari J. Kigodi*, Michael Kisangiri and Dina Machuve

School of Computational and Communications Science and Engineering, The Nelson Mandela African Institute of Science and Technology, P. O. Box 447, Tengeru-Arusha, Tanzania.

Accepted 30 August, 2013

This paper surveys various literatures on network performance with main focus on three main issues: its meaning, quantification and measurement. The essence of performance in communication systems is discussed and its meaning explored. Further, we look at relationship between network performance and its characteristics. The weaknesses in literature are disclosed with respect to how performance is conceptualized and applied in networks. The study reveals that there is need to carry out further research to redefine and quantify network performance with intent to get rid of existing misconception and ambiguity.

Key words: Network performance, performance characteristics, quantification, measurement.

INTRODUCTION

The primary function of an ideal network management system is to optimize the operational capability of a network. It implies that the ultimate goal for network management is to keep network operation at peak performance (Cisco, 2008). Performance is also considered during design and re-engineering of communication systems. Although the understanding on network performance does not differ much amongst communications systems, the review is redirected exclusively on computer networks. It is important to note that with newer applications like voice and video in computer networks, performance is the key variable to success and if consistent performance cannot be achieved, the service is considered of low value and fails. In other cases, users simply suffer from variable performance with intermittent application timeouts that degrade productivity and user satisfaction (Haykin, 2001; Lathi, 1998). So maintaining network design integrity, network operational health and final end-node-to-end - node performance is a top priority issue (Nassar, 2000).

Performance is an essential factor in communication links (media), nodes, protocols and traffic on transit (Walrand et al., 2000). Protocols are rules for exchanging messages in or between computing systems and in telecommunications; and they may include signaling, authentication and error detection and correction capabilities. Network traffic or data traffic is data in a network. The kind of protocols in use and traffic on transit in the network pose network performance challenges. Media allow passage of information or traffic amongst the communicating entities and they have an important role to play during data transmission to ensure efficient transmission. Nodes especially network devices must be available, efficient and reliable to guarantee that the network is operationally capable (Keiser, 2002; Lathi, 1998; Haykin, 2001).

The network performance challenges have always been dealt with basing on divide and conquer approach by focusing on performance of those four important aspects of communication networks. Therefore the

*Corresponding author. E-mail: okigodi57@gmail.com, kigodio@nm-aist.ac.tz. Tel: +255 752448545.

characteristic of the overall network performance is an aggregate of the performance characteristics of media, protocols, traffic and network devices (Spohn, 2000). This dividedness has rendered uncountable advantages in terms of simplicity, efficiency, effectiveness and robustness but on the other hand it has posed difficultness in understanding and measurement of the network performance.

MEANING

A lot has been said about network performance because of its essence, variety of networks and complexity of communication process. The eagerness of many studies in this part has been on performance characteristics or data, analysis and evaluation, measurement (Milliken, 2005), degradation and enhancement, baselining (Nassar, 2000),

monitoring or management (Cisco, 2008), optimization (Kouvastos, 2011), and modeling (Keiser, 2002). However, the term has been left hanging somehow in the way it is understood most probably because it is multi-faceted. It has been discussed in terms of network users and owners' expectations, investment goals, (handling of its) challenges, management and design goals. Generally, irrespective of circumstances or type of network or network aspect, performance of networks is conceptualized in terms of its characteristics; therefore performance characteristics are core issues in understanding how networks behave.

Networks must operate at peak performance so as to gain the full return on our investment (Nassar, 2000) because they have become a necessary part of businesses and organizations. Network administrators work tirelessly to ensure that intended performance is achieved. Network performance is the core functional goal of network management, monitoring, evaluation and measurement; the primary aim being to understand how the network is behaving, identifying weak points, and seeing to it that the network is capable enough to save business and organizational purposes.

Seshan et al. (1997) developed a system for discovering network performance called SPAND (Shared passive network performance discovery) and they conceptualized the network performance in terms of characterisctics used to identify performance such as latency, loss rate, available bandwidth, and routing metrics (Seshan et al., 1997). Such orientation is very common and the main approach currently adopted to understand the performance of networks (Gokhale, 2005; Keiser, 2002; Stallings, 1996; Kouvastos, 2011). It is evident from those sources that the network performance has many characteristics which vary greatly depending on a number of factors such as type of network, coverage, and components used.

According to Koendjbiharie et al. (2010), business networks have flourished over the past several decades as many firms facilitated by information and communication technology have clinched the digital ecosystem form of organization (Koendjbiharie et al., 2010). They argue that the thinking of performance at the level of the individual firm has been extended to the level of the ecosystem - the business network as a whole - but still it is not clear how to best conceptualize and measure network performance. Further, the business network performance can be understood by harmonizing three perspectives: the firm perspective, the complex systems perspective and the customer perspective. Basing on the customer perspective, network performance is evaluated in terms of network effectiveness and efficiency; while the complexity perspective focuses on the individual network – its components and functionalities - behavior and how that is reflected onto the whole system to meet firm's goals.

PERFORMANCE CHARACTERISTICS AND MEASUREMENT

According to Mahotra and colleagues (Mahotra et al., 2010) the performance metrics mainly considered in wired and wireless LANs are collision count, data dropped, delay, load, media access delay, queue size and throughput. Collision count is the total number of collisions encountered by a station during packet transmissions. Data dropped implies the number of bits that are sent by wireless node but never received by another node. Delay is a statistic which represents the end to end delay of all packets received by all the stations and forwarded to the higher layer. Load is the total number of bits received from the higher layer. Packets arriving from the higher layer are stored in the higher layer queue. It may be measured in bits/sec or packets/sec. Media access delay means total time (in Seconds) that the packet is in the higher layer queue, from the arrival to the point when it is removed from the queue for transmission. Queue size represents the total number of packets in MAC's transmission queue(s) (in 802.11e capable MACs, there will be a separate transmission queue for each access category). Throughput is the total number of bits sent to the higher layer from the MAC layer. The data packets received at the physical layer are sent to the higher layer if they are destined for this destination.

The basic characteristic of LAN is that its community of users shares the capacity of the transmission links that interconnect them (Keiser, 2002). To be able to utilize this capacity efficiently, protocols are needed for users to access the LAN medium; as such performance of access methods adopted is vital factor in LANs. Keiser argues that the performance of access methods is usually evaluated in terms of the average packet delay or average transfer delay versus throughput.

According to Gokhale (2005), three LAN performance measures that are commonly used are delay, throughput, and the total load on the LAN, including control packets like tokens, and collisions, which are destroyed packets that must be retransmitted. The total load determines network utilization which is the ration of total load to network capacity.

Hong et al. (2009) studied the effect of network information such as network topology, channel state and traffic information on network performance. If more network information is obtained, the network protocol will be more efficient and the network performance can be improved. They determined the lower bound on the amount of information required to achieve certain network performance. The network throughput improves when the scheduler collects more information. However, they did not consider the overhead of collecting network information which also affects network performance because collecting information consumes bandwidth resource, which may in turn affect the network data rate. The communication protocols transmit information of their own which is referred to as overhead (Cavanaugh, 1994).

IP networks are now used to carry everything from best effort internet traffic to mission critical business data, as such Quality of Service (QoS) has become extremely important to provide differentiated treatment of traffic (Barreiros and Lundqvist, 2011). QoS is a performance factor which includes latency, jitter and throughput characteristics; it is especially a significant factor for multimedia data transmission in data networks (Park, 2005). In cellular networks quality of end-user experience (QoE), it is also considered in addition to QoS (Soldani et al., 2006). QoE is a performance element which describes user perceptions of the performance of a service. Usually in order to provide the best QoE to users in a cost-effective, competitive and efficient manner, network and service providers must manage network QoS and service provisioning efficiently and effectively. Although QoE is not usually accounted for computer networks but it is indirectly necessary part-of performance and it can be translated as user feedback to network administrator regarding the network operation behaviour.

Lawniczak et al. (2006) worked on the network performance indicators: throughput, number of packets in transit and average delay time of all packets delivered. They investigated how these indicators are affected by network connection topology and routing. They studied how those indicators capture the phase transition point (that is, the critical load of a network). They investigated how additional links added to the network connection topology affect the network performance indicators and the phase transition point. The same performance parameters were worked on by Ashok and his colleagues (Ashok et al., 2005) for Network on-chip (NOC).

The intermediate devices like routers and switches in a network have buffers where the packets wait in a queue before and after processing (Malhotra et al., 2011). Depending on the packet arrival rate and the packet departure rate which may be higher or lesser than the packet processing rate, the size of input or the output queue may increase or decrease. This increase in queue size may lead to congestion. A key issue in designing any good network is to use congestion control mechanism. The congestion control involves two factors that measure the performance of a network: delay and throughput.

Tobagi and Hunt (1980) analyzed the Carrier Sense Multiple Access (CSMA) throughput in LAN. They did the analysis based on data rate, propagation delay, transmission time, packet size, and link length. The resulting model for the channel throughput indicated how the CSMA protocol is performing.

The works discussed above show that network performance has many elements which are so varied and depends on the type of network or aspect of network considered. It is evident that measurement of network performance has been adopted as measurement of performance characteristics; and that whenever network performance term is used 'eyes are cast on' its elements. Also it is shown that network performance could mean protocol performance, device performance, medium performance, traffic characteristics, or any other term as long as it indicates/influences the operational capability of the network.

Today there are numerous software tools available for analysis, simulation, measurement, evaluation, and diagnosis of network performance (Stanford Linear Accelelator Center (SLAC)). Just a few to mention are: netperf, ns, netmeter, smokeping, AthTek, NetWalk, Cedexis Radar, Enigma, EventSentry, FactFinder Express, Foglight, InfoVista Network Performance Management, Intense School, IP - MAC scanner, MindArray, MonitorHub, NetFort SPAN Port Configurator, Network Timeout, NetXMS, Retrospective, SparrowIQ, Trogon MAC scanner, TruePath Technologies, and QCheck. Those tools work on performance characteristics; and provide a picture of how the network is behaving in terms of (performance) elements.

Following closely the above discussion it is obvious that the approaches for monitoring and evaluation of network performance are varied and depend on performance parameter(s) or aspects being analyzed; but generally the focus usually is to provide a way for network administrators to quickly understand the network health and be able to resolve pertinent problems (Chiu and Sudama, 1992). For example there are several bandwidth measurement techniques for estimating capacity and available bandwidth in individual hops and end-to-end paths. The four major techniques are Variable Packet Size (VPS) probing, Packet Pair/Train Dispersion (PPTD), Self-Loading Periodic Streams (SLoPS), and Trains of Packet Pairs (TOPP). VPS estimates the capacity of individual hops, PPTD estimates end-to-end capacity, and SLoPS and TOPP estimate end-to-end available bandwidth (Prasad et al., 2003).

CONCLUSION

The network performance has been defined and understood in terms of its characteristics or parameters resulting into vague meaning. The basic question is 'What is the set of parameters which best account for overall network performance without imposing ambiguity?' One element or a mini-set of the elements are used to explain performance which is not entirely correct. So, the network performance has been thought as protocol performance, traffic performance characteristics, link/medium performance or device performance which leads to even more confusion.

The network performance relationship with its elements is not explicit. It is an implied relationship and does not indicate the level to which each parameter affects the overall performance. This is a quest which has been overlooked and still lack solid answers to nail it down. For example it is not odd to see the obvious effects of delay on network performance, however, in that understanding scientifically-supported clarity is missing that explain quantitatively to what extent the effect is significant or otherwise to the whole system.

The individual parameters of network performance have been quantified but the overall performance is not yet quantified. Measuring the individual performance parameters does not mean measuring the overall performance and it is inconceivable otherwise due to lack of logical backup. The overall network performance has not been measured to provide its quantitative definition which is important for knowing the level to which the whole network is performing in terms of its essential elements.

Therefore to resolve the challenges, network performance should be redefined. For instance by developing a mathematical formula or model for overall network performance which captures all important characteristics of the network in relation to the network operation. Having a value for network performance also will eliminate the ambiguity. So, further studies are needed to work out the problems either in the suggested directions or other ways possible.

REFERENCES

Barreiros M, Lundqvist P (2011). QoS-Enabled Networks: Tools and Foundations. West Sussex, UK: John Wiley & Sons.

Chiu DM, Sudama R (1992). Network monitoring explained: design and application. Ellis Horwood Series in Computer Communications and Networking.

Cisco (2008). Performance Management Best Practices for Broadband Service Providers. USA: Cisco Systems Inc.

Gokhale AA (2005). Introduction to Telecommunications (2 ed.). New York, United States of America: Thomson Delmer Learning.

Haykin S (2001). Communication Systems (4 ed.). New York, USA: John Wiley & Sons, Inc.

Hong J, Li VOK (2009). Impact of Information on Network Performance – An Information- Theoretic Perspective. IEEE Glob. Telecomm. Conf. Publ. pp.1-6.

Keiser G (2002). Local Area Networks. New York, United States of America: McGraw-Hill Companies.

Koendjbiharie S, Koppius O, Vervest P, van Heck E (2010). Network transparency and the performance of dynamic business networks. 2010 4th IEEE International Conference on Digital Ecosystems and Technologies (DEST). pp.197-202.

Kouvastos DD (2011). Network Performance Engineering: A Handbook on Convergent Multi-service Networks and Next Genereation Internet. German: Springer-Verlag Berlin Heidelberg.

Lathi, B. (1998). Modern Digital and Analog Communication Systems. New York: Oxford University Press Inc.

Lawniczak AT, Tang X (2006). Packet Switching Network Performance Indicators as Function of Network Topology and Routing Algorithms. IEEE Conference (CCECE '06) Publication. pp.1008-1011.

Milliken WC (2005). Patent No. 6978223. USA.

Nassar DJ (2000). Network Performance Baselining (1 ed.). USA: MTP.

Park KI (2005). QoS in Packet Networks. USA: Springer.

Prasad,\ RS, Murray M, Dovrolis C, Claffy K (2003). Bandwidth estimation: metrics, measurement and tools. Networking Journal, IEEE.

Seshan S, Stemm M, Katz RH (1997). SPAND: Shared Passive Network Performance Discovery. USENIX Symposium on Internet Technologies and Systems. California: USENIX.

Soldani D, Li M, Cuny R (2006). QoS and QoE Management in UMTS Cellular Systems. West Sussex, UK: John Wiley & Sons.

Spohn DL (2000). Data Network Design. New York, United States of America: McGraw-Hill Companies Inc.

Stallings W (1996July9). Knowing wiring basics can boost local net performance. Network World , P. 29.

Stanford Linear Accelerator Center (SLAC). (n.d.). Network Monitoring Tools. Retrieved August 29, 2013, from http://www.slac.stanford.edu/xorg/nmtf/nmtf-tools.html.

Walrand J, Varaiya P (2000). High-Performance Communication Networks. USA: Academic Press.