International Journal of Trend in Research and Development, Volume 4(5), ISSN: 2394-9333
www.ijtrd.com

Analysis of Link Aggregation Protocols Applied on Efficient Data Center Networks

[1]Sarhan M. Musa, [2]Nader F. Mir, [2]Avanish Gupta, [2]Shakun Rajan Yawatkar, [2]Anam Khan and [2]Sravani Chavali,
[1]Department of Engineering Technology, Prairie View A&M University, Prairie View, Texas 77446, U.S.A
[2]Department of Electrical Engineering, San Jose State University, San Jose, California 95192, U.S.A

Abstract—Widespread deployment of network virtualization technology in ever-growing Enterprise, Data center and service provider network, brings about its own set of problems. The concept of LAG aims to overcome some of the problems such as IP address exhaustion, limited link bandwidth, bottleneck links, Level-3 routing and lack of resilience. There are two main protocols for LAG, Link Aggregation Control Protocol (LACP) and Port Aggregation Protocol (PAgP). Both these protocols are used widely in networking industry. LACP is an IEEE standard while PAgP is Cisco Proprietary protocol. This paper presents implementation, simulation and testing the performance of Link Aggregation (LAG) protocols applied on data center networks. Also, we show the performance analysis of each protocol separately for LAG considering various factors such as file transfer rate, number of parallel connections and link overloading. Apart from individual analysis of LACP and PAgP, a brief comparative analysis of these two protocols is presented.

Index Terms—*Link Aggregation, Link Virtualization, Port Aggregation,*

I. INTRODUCTION

In high speed networks, Link Aggregation (LAG) is used to improve speed and throughput and reduce the cost to transmit data. LAG increases transmission speeds and network capacity. LAG also reduces the requirement for new hardware, thus saves the cost of new hardware devices. Combining several physical links which are on different devices into one logical or virtual path is the main goal of link aggregation. Also, the multiple links which are combined by LAG act as back up for each other. It increases reliability by introducing redundancy [3]. Advantages of Link Aggregation can be as follows:

- Increased Network Reliability: Link Aggregation achieves redundancy of the several links it aggregates. If one of the aggregated link fails, another one carries the network traffic keeping the communication uninterrupted.
- The data flow continues. So overall network reliability increases.
- Load Balancing: Various Network Load Balancing algorithms are implemented so as to distribute the load on the network.
- Increased Speed: LAG increases the speed of the network by small amounts. It saves cost and resources. There are two ways to configure LAG: Static and Dynamic.

Link Aggregation can be deployed in three kind of networks: Switch to Switch Networks, Switch to Station Networks and Station to Station Networks. Limitations of only one connection can be overcome by utilizing parallel networks with Link Aggregation Deployed [4][6][8]. Physical Ports on the same switch can be aggregated using this. Split Multi Link Trunking (SMLT) and Routed SMLT can overcome limitation of only one connection and ports from different switches can be aggregated using this, popularly called - Link Bundling, Network Interface Card Teaming. Fast Ethernet and Gigabit Ethernet with Link Aggregation can further enhance accessibility and limit between different devices [5][11].

II. LINK Aggregation Protocols

A. *Link Aggregation Control Protocol*

Link Aggregation Control Protocol (LACP) is a standard protocol used for link aggregations. Various misconfigurations are avoided in network using LACP. Thus, it ensures that aggregation of links is done only if they are connected and configured consistently. Aggregation of links gives greater bandwidth by merging of links among two or more nodes [10]. Ports on which LACP is configured can be grouped and LACP can be automatically configured on them. LACP enables another link to act as back up for a primary link in case it fails. The adding or deleting of link automatically happens when a link fails among the group. LACP can keep a track of links and their appropriate groups [6].

In a nutshell, to automatically configure and maintain link aggregation, Link Aggregation Control Protocol is used as an extension of LAG. Being dynamic, LACP can be configured among switches of various make. Initial configurations are needed on ports being involved in aggregated links. The additional links present can be auto activated after a certain link failure.

The two LACP Configurational modes are:

- Active mode – When a port is up, it sends LACP PDUs immediately.
- Passive mode – If the port is configured as passive, it just responds to LACP PDUs received. But, negotiation is not started by itself.

If both the sides are configured for active negotiation of LACP, link aggregation can be formed easily as both the sides will negotiate actively. If one side is configured as active and the other side is configured as passive, then also the LAG can be configured as other side will respond to LACP PDUs sent by the opposite side. If both the sides are configured as passive for LACP negotiation, it will fail to negotiate link aggregation between the two sides.

B. *Port Aggregation Protocol*

Port Aggregation Protocol (PAgP) is a Cisco proprietary protocol. Ethernet switch ports called as Ether Channels are logically aggregated using PAgP. Cisco switches can function in three modes when enabled with PAgP namely Auto, Desirable and On.

In Auto Mode, channel is negotiated passively. While in Desirable Mode, the channel is negotiated actively. In On Mode, it is assumed that other side of the network is already having active PAgP and no protocol is used to negotiate.

III. IMPLEMENTATION FOR TOP OF RACK NETWORKING

Both LACP and PAgP protocols are implemented in experiments on different topologies including a basic topology applied on top of rack (ToR) switching of data center networks. To start with, a square topology and a fully connected mesh topology are introduced as follows.

A. Basic Topology

To start the experiment, it is decided to have a minimal topology which can be used for setup and testing of basic features of an aggregated link. To maintain simplicity, a topology with only 2 switches and two aggregated links is chosen. The topology is as shown Figure 1.

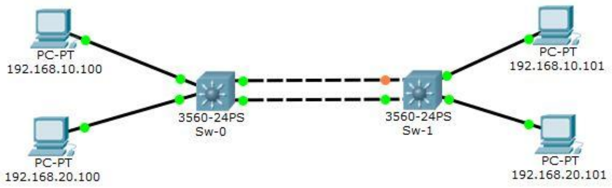

Figure 1: Basic Topology

The next step in the experiment is to configure the Link aggregation for the two links. Once both the routers are configured for link aggregation for both links, the network undergoes a stabilization phase. During this time the aggregated links first go down, then start negotiation with each other. Within a few minutes the negotiations are complete and the links become active. The various stages of this equalization and stabilization process can be seen, where the color of link interfaces changes in the negotiation stage.

B. Square Topology

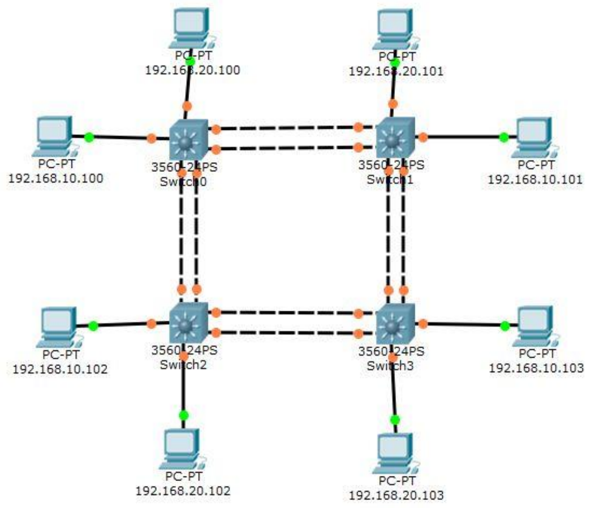

Figure 2: Square Topology

The next step in the process is to gradually upgrade the size of the network simulated, and verify the functioning of the multiple aggregated links. To this end, a network with 4 switches arranged in a square fashion with two hosts connected to each switch, is chosen. Figure 2 shows the network as simulated in square toloplgy.

It can be observed that the above topology has four branches that have aggregated links. Each of these branches has two links each. It should be noted that while Spanning Tree Protocol (STP) does not block any of the interfaces on the switches to prevent loops between any two switches after Link aggregation has been achieved, it still blocks some interfaces to prevent the larger loop consisting of all the four switches.

The four branches have been aggregated independently. Thus each switch has two groups of aggregated links. Once all the links have been aggregated, connectivity is tested between the various hosts to verify successful configuration.

C. Fully Connected Mesh Topology

To further build upon the network created in the previous stage and more closely mimic practical network designs, the square topology is converted into a fully connected mesh network. Though the number of switches and the hosts in the network remains the same, addition of extra aggregated links helps reduce the number of hops. Figure 3 below shows the simulated network of fully connected mesh. As observed in the previous topology, here too STP disables several interfaces to prevent loops. Since this topology is more complicated than the previous one, more numbers of interfaces are disabled. Six interfaces are disabled compared to two in the previous case.

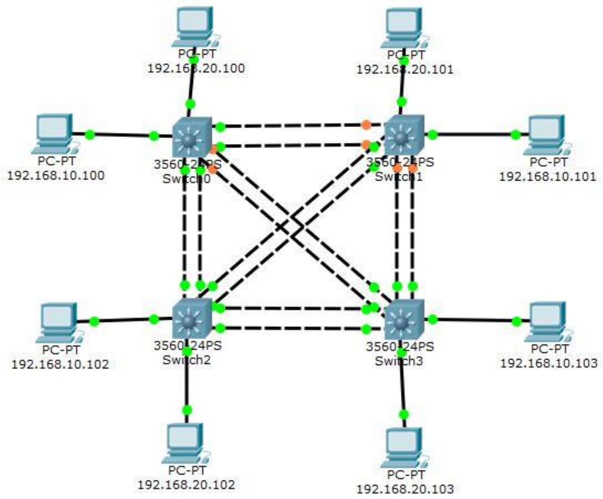

Figure 3: Fully Connected Mesh

It is observed in the three topology simulations as described above that a test of connectivity is a necessary but not a sufficient condition for successful implementation of link aggregation.

D. Performance Parameters

Following performance parameters are varied while doing experiments for LAG.

1) Link Aggregation protocol being used (No LAG / LACP/ PAgP)

This variation in the LAG protocol being used allows us to observe the performance of different LAG protocols and compare them with each other. We can also observe the impact of LAG by comparing it with performance when no LAG is incorporated. This can help us choose the protocol that best suits our needs. All experiment performed here use two most common LAG protocols: Link Aggregation Control Protocol (LACP) and Port Aggregation Protocol (PAgP), as well as without using any form of LAG.

2) Network Topology

By studying the performance of LAG in different network topologies, we can get a better understanding of the impact of variation in topology on the LAG performance. This can help us optimize our network to better suit the LAG configuration being used. The three network topologies used for the experiments are Basic topology, Square topology and fully connected mesh. These three topologies and their respective configurations have been detailed in the previous section.

3) File Size Used for File Transfer

To more extensively understand the impact of LAG on the performance of a typical network, it becomes imperative to observe the effect of file size on different LAG protocols. To this effect, all experiments have been repeated for four different file sizes. The file sizes chosen are 5571584 Bytes, 15522644 Bytes, 33591768 Bytes and 50938004 Bytes. The choice of file sizes is made to include a wide range (5 to 50 MB).

4) Number of Links Aggregated

Link Aggregation can be performed by aggregating any number of links available. This is usually limited by the number of physical links available for use since links and ports are limited and hence valuable. Hence the number of links being aggregated must be optimized to best suit the network requirement. By comparing the performance with number of links aggregated under different conditions, an optimized number of links to be aggregated can be approximated for a particular network.

5) Number of Simultaneous File Transfers

To effectively test the aggregated link to its limiting conditions, link overloading is being used. This is done by running multiple simultaneous file transfers over the aggregated link. Tests are conducted with one, two and three file transfer running simultaneously. This ensures that the aggregated link behaves as the bottleneck link. All other links used in the file transfer (for example the link from the router to the end host where file transfer application is running) are subjected to lesser traffic than the aggregated link.

IV. RESULTS

The performance evaluation results for all the three topologies are tabulated and analyzed. The detailed results and analysis graphs are included in this work and can be referred for detailed reading.

It is observed from these results that minor changes in the network topology do not have much effect on the performance of LAG results. This is in line with the expectations as link aggregation protocols are designed to affect only a particular link and remain virtually invisible to other aspects of the network. Hence changes in the topology should not have any effect.

It is also noted that as the file size used in data transfer between two ends increases, the overall transfer rate decreases slightly. This is observed to be true for bots LACP and PAgP. However, PAgP performance is slightly better than LACP protocol when the file sizes are more.

One way to test the limiting condition on a link is by using Link Overloading. In this setup, initially done for a single non-aggregated link, the link is tested under a progressively increasing load, so as to observe its saturation conditions as illustrated in Figure 4. The traffic on the link is increased by running one, two, and three simultaneous file transfers as shown in Figure 5. As the number of simultaneous file transfers increase, the link became the bottleneck in the data transfer path. It is observed that the transfer rate decreases as the file size increases for all the three cases. The total data transfer rate between two ends including all the file transfers going on between two ends slightly decreases as the number of simultaneous file transfers increase. This can be due to overhead increase, when the number of simultaneous file transfers increases. As expected, individual data rates fell down proportionally as the number of file transfers increased.

While comparing the performance of LACP and PAgP protocol, it is observed that under most conditions, the performance of PAgP protocol is slightly higher than LACP protocol. Even though the difference is not much, when large amount of traffic is under consideration, such minor differences cannot be ignored. It must, however, be noted that PAgP protocol is a Cisco proprietary protocol, and can be used on only supporting Cisco devices. Also since the Testing tool used in this research is a Cisco product, its performance may have been optimized for a Cisco proprietary protocol (PAgP), thus giving it a slight edge over the open source LACP protocol.

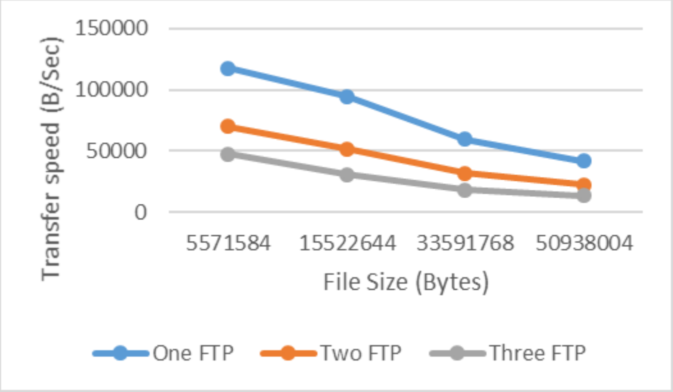

Figure 4: Link Overloading Results

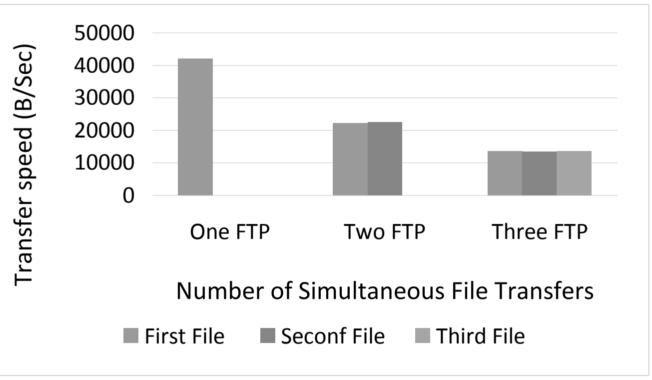

Figure 5: Effect on Different files

Until the testing is done using a third party testing platform, the test results will remain unverified. But, since PAgP protocol is not available as an open source, this becomes an unlikely possibility. Due to the lack of detailed information on the packet structure and frame format of PAgP protocol, its slight edge over LACP cannot be justified either.

This observation is performed on Network Topology II where first, overall Delay is observed and then the throughput. Delay in packets received in case of four switches aggregated with four links is measured and compared with the Delay in receiving packets in case of single link without any aggregation. Figure 6 shows comparisons while using FTP servers with different file sizes. The blue bar represents single link output while orange represents LACP output.

Figure 6: Delay comparison graph

The graph clearly shows a significant reduction in Delay while using aggregation for increasing number of FTP servers. As observed the variation is minimum while using single FTP server and is quiet significant while using two and three servers. This observation proves that in large Data Centre Networks where delay can result is critical data loss LACP proves to be of significant importance. Similarly, next throughput is observed on same topology and the following graphs in Figure 7 are observed in case of single link and link aggregation.

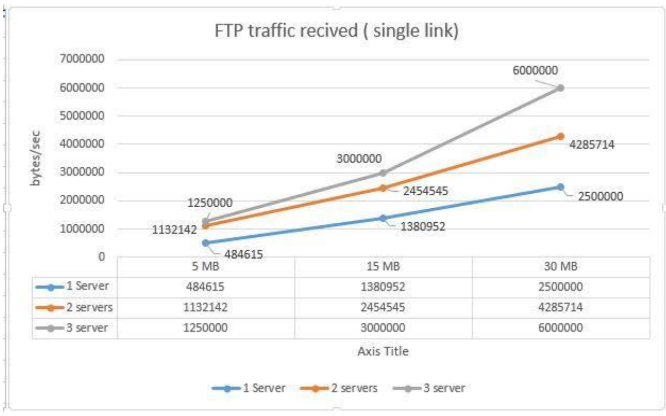

Figure 7: Throughput graph for FTP traffic (single link)

Figure 7 shows throughput results for various file sizes while using FTP servers. The graph line in blue shows the throughput while using 1 FTP server, similarly orange and grey represents two and three FTP servers all using a single link and without aggregation. Also the file sizes are varying on the y-axis. Considering first 5MB file size, there is an increase in throughput as the number of servers' increases. Similarly, for 15MB and 30MB, throughput increases when the number of FTP servers connected is increased. Figure 8 summarizes the same result in case of link aggregation using LACP on four links.

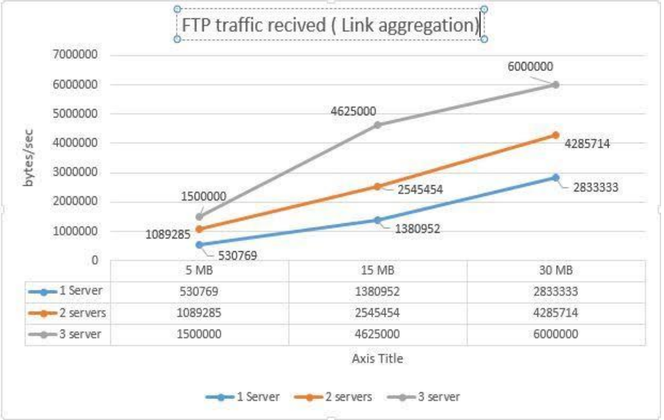

Figure 8: Throughput graph for FTP traffic (aggregated links)

Figure 8 summarizes the throughput result by increasing the number of FTP servers and keeping different file sizes in each case. Considering the file size of 5MB throughput is increased linearly by increasing the number of servers. However, in case of 15MB file size the throughput is considerably increased for three FTP servers implying that link aggregation becomes more effective in case of large amount of data being transferred. Similarly, for file size of 30MB it shows wide range of variation with increasing the number of FTP servers. Also the maximum throughput of 6 megabytes can be observed in case of 50 MB file size with three FTP servers. Figure 8 effectively shows link aggregation is more prominent for greater file sizes. Comparing the above two graphs, i.e. the one with single link and the other with aggregation and individually considering the file sizes. For 5MB, the throughput in case of single link is around 11.5MB and with link aggregation is 13MB. For 15MB file size the single link throughput is around 3MB and aggregated links have throughput of 3.5MB.

CONCLUSION

We implemented, simulated and tested the performance of Link Aggregation (LAG) protocols applied on data center networks. We showed the performance analysis of each protocol separately for LAG considering various factors such as file transfer rate, number of parallel connections and link overloading. In addition, we made a brief comparative analysis of LACP and PAgP.

References

[1] "*IEEE 802.3ad Link Aggregation Task Force*" available at http://www.ieee802.org/3/ad/

[2] IEEE Std 802.1AX-2008 IEEE Standard for Local and Metropolitan Area Networks available at http://standards.ieee.org/getieee802/download/802.1AX-2008.pdf

[3] "Link aggregation" available at https://en.wikipedia.org/wiki/Link_aggregation

[4] "IEEE 802.3ad Link Aggregation (LAG)" available at http://www.ieee802.org/3/hssg/public/apr07/frazier_01_0407.pdf

[5] "Understanding IEEE 802.3ad Link Aggregation" available at http://www.juniper.net/techpubs/en_US/junose14.2/topics/concept/802.3ad-link-aggregation-understanding.html

[6] "Configuring Link Aggregation" available at http://www.cisco.com/c/en/us/td/docs/wireless/controller/

7-4/configuration/guides/consolidated/b_cg74_CONSOLIDATED/b_cg74_CONSOLIDATED_chapter_0100011.html

[7] "LACP (802.3ad) for Gigabit Interfaces" available at http://www.cisco.com/c/en/us/td/docs/ios/12_2sb/feature/guide/gigeth.html

[8] "ns-3 Tutorial" available at https://www.nsnam.org/docs/release/3.24/tutorial/singlehtml/index.html

[9] "RIVERBED MODELER network simulator" available at http://www.riverbed.com/products/steelcentral/steelcentral-riverbed-modeler.html

[10] "802.1ad - Provider Bridges" available at http://www.ieee802.org/1/pages/802.1ad.html

[11] "IEEE 802.1ad" available at https://en.wikipedia.org/wiki/IEEE_802.1ad

[12] "Understanding Q-in-Q Tunneling and VLAN Translation" available at https://www.juniper.net/techpubs/en_US/junos13.2/topics/concept/qinq-tunneling-qfx-series.html

[13] "IEEE 802.1Q-in-Q VLAN Tag Termination" available at http://www.cisco.com/en/US/docs/ios/lanswitch/configuration/guide/lsw_ieee_802.1q.html

[14] "An Overview of Link Aggregation and LACP" available at https://thenetworkway.wordpress.com/2015/05/01/an-overview-of-link-aggregation-and-lacp/

[15] "Method and system for VLAN aggregation" US patent 6914905 B1 available at http://patft.uspto.gov/netacgi/nph-Parser?Sect2=PTO1&Sect2=HITOFF&p=1&u=/netahtml/PTO/search-bool.html&r=1&f=G&l=50&d=PALL&RefSrch=yes&Query=PN/6914905

[16] "VLAN Aggregation for Efficient IP Address Allocation" RFC 3069 available at https://tools.ietf.org/html/rfc3069.

[17] IEEE standard for local and metropolitan area networks--link aggregation, 1st ed, 2014,

[18] S.-E. Jeon, "Topology aggregation method for multiple link parameters," IEEE Global Telecommunications Conference, 2015. GLOBECOM '15