Hub Gigabit Ethernet Test Report

Table of Contents

1 FEX Hub Module Switch Testing Report

The goal of this testing was to observe the effects on bandwidth of sending data to/from multiple ports of the FEX Hub Module's three switch chips. Results are organized according to the number of ports, referred to as clients, sending data to another port, referred to as a server.

1.1 Overview

First the layout of the FEX Hub Module as it relates to the switch chips is described, as well as the software used for the bandwidth testing and the actual setup of the machines used to perform the testing. Then the data gathered from the testing is shown, first the results of sending data from one client to one server, then the results of sending data from two clients to one server, and finally the results of sending data from three clients to one server. Comparisons are made to a commercially available gigabit switch.

It is seen that using the FEX Hub Module connections between a single client and a single server maintain a bandwidth of 935 Mbits/sec, which is the same as that of the commercial switch. It is further seen that when sending data from multiple clients to a single server, an overall bandwidth of 940 Mbits/sec is maintained. The bandwidth of each individual connection between client and server is about what one would expect based on the bottleneck of traffic. The only difference is that when there is competition between connections, the connection whose data travels through more switch chips has more available bandwidth. The difference between the expected bandwidth based on the bottleneck of traffic and the actual bandwidth resulting from the topography of the connections is noticeable, but overall not significant.

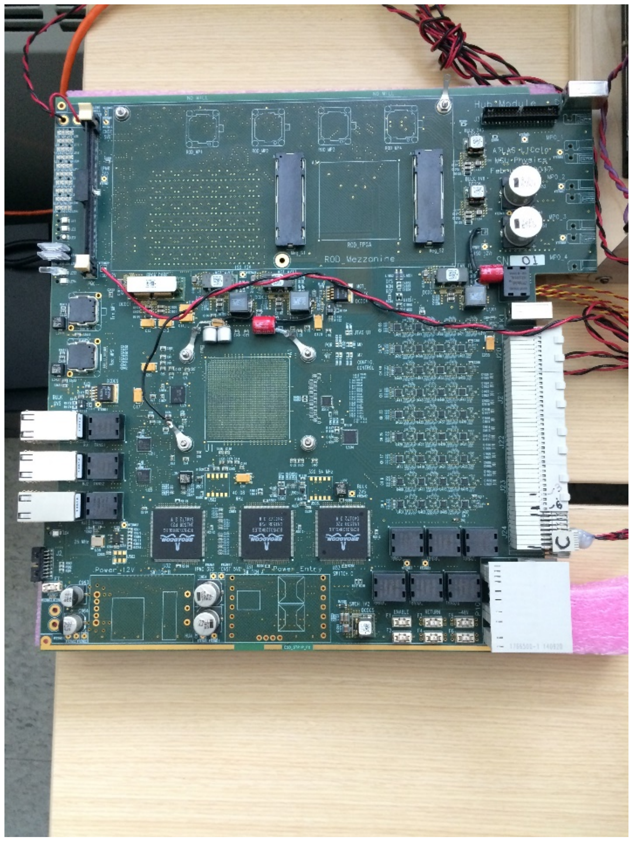

Figure 1: Hub Module GbE Test Setup (1/2)

1.2 Description of Hub Hardware

The FEX Hub Module possesses 3 switch chips, labeled A, B, and C, which each have 8 ports. In the FEX Hub Module used for these tests, chips A and C each use one of their ports to connect to chip B, while chip B uses two of its ports to connect to the other two chips (refer to diagram). These inter-chip connections can be enabled/disabled depending on whether the connecting circuit is capacitor coupled. The FEX Hub front panel has six RJ45 Connectors/Ethernet Ports, four of which connect to the Hub's switch chips. The remaining two connect to the FEX Hub Module's ROD FPGA and IPMC. The four ports connected to the switch chips are referred to SW-A6, SW-B6, SW-B7, and SW-C6. Port SW-A6 is connected to chip A, port SW-B6 is connected to chip B, and so forth. Backplane ports were also tested, but these go without individual names.

1.3 Iperf Software

Iperf is a network management tool designed to measure maximum achievable bandwidth on IP networks. The National Laboratory for Applied Networking Research: Distributed Applications Support Team (NLANR/DAST) originally developed the software, which can be found and downloaded at https://iperf.fr. The user specifies a client to generate traffic, and a server to discard traffic (multiple clients can and will be assigned to a single server). Once a connection between a client and a server is established the maximum bandwidth of the connection is measured, and a report of the test is created (in CSV format). For these tests, each client generated traffic for a 60 second period, and the results of each five-second interval of the test were reported. When sending traffic from multiple clients, each connection was initiated within five seconds of the others, and as such only the results of the intervals between seconds 5.0 and 55.0 of each test are fully representative of the competing traffic among all the clients.

1.4 Test Setup

Tests were conducted using five machines. One machine ran CentOS7 and used Secure Shell to remotely operate the other four, which ran perfSONAR Toolkit, a custom distribution of CentOS containing several tools and services for monitoring network performance, including iperf. PerfSONAR Toolkit can be found and downloaded at https://www.perfsonar.net. The machine running CentOS7 was connected by an ethernet cable to port SW-B7 of the FEX Hub Module, while the other machines were connected to the other ports of the FEX Hub Module as needed. Furthermore, a commercial HP 1410-8G Gigabit Switch is also tested for comparison. The commercial switch has 8 ports, numbered 1 through 8. When testing the commercial switch, the machine running CentOS7 connected to port 1, while the other four machines connected to ports 2-5.

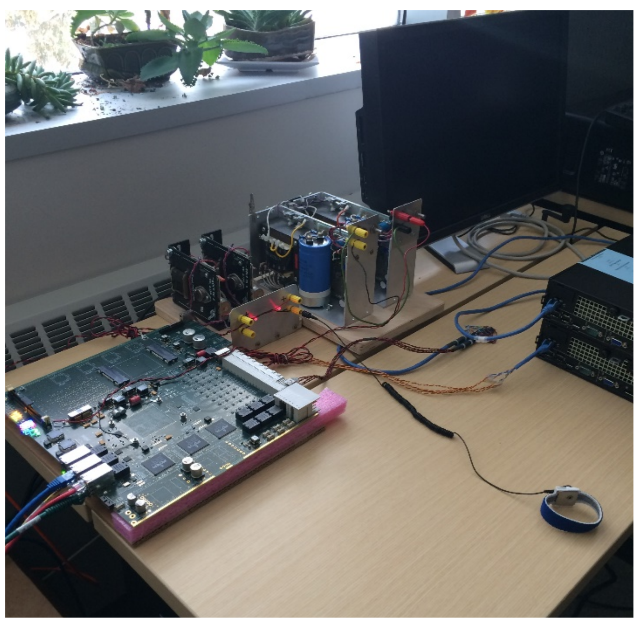

Figure 2: Hub Module GbE test setup (2/2)

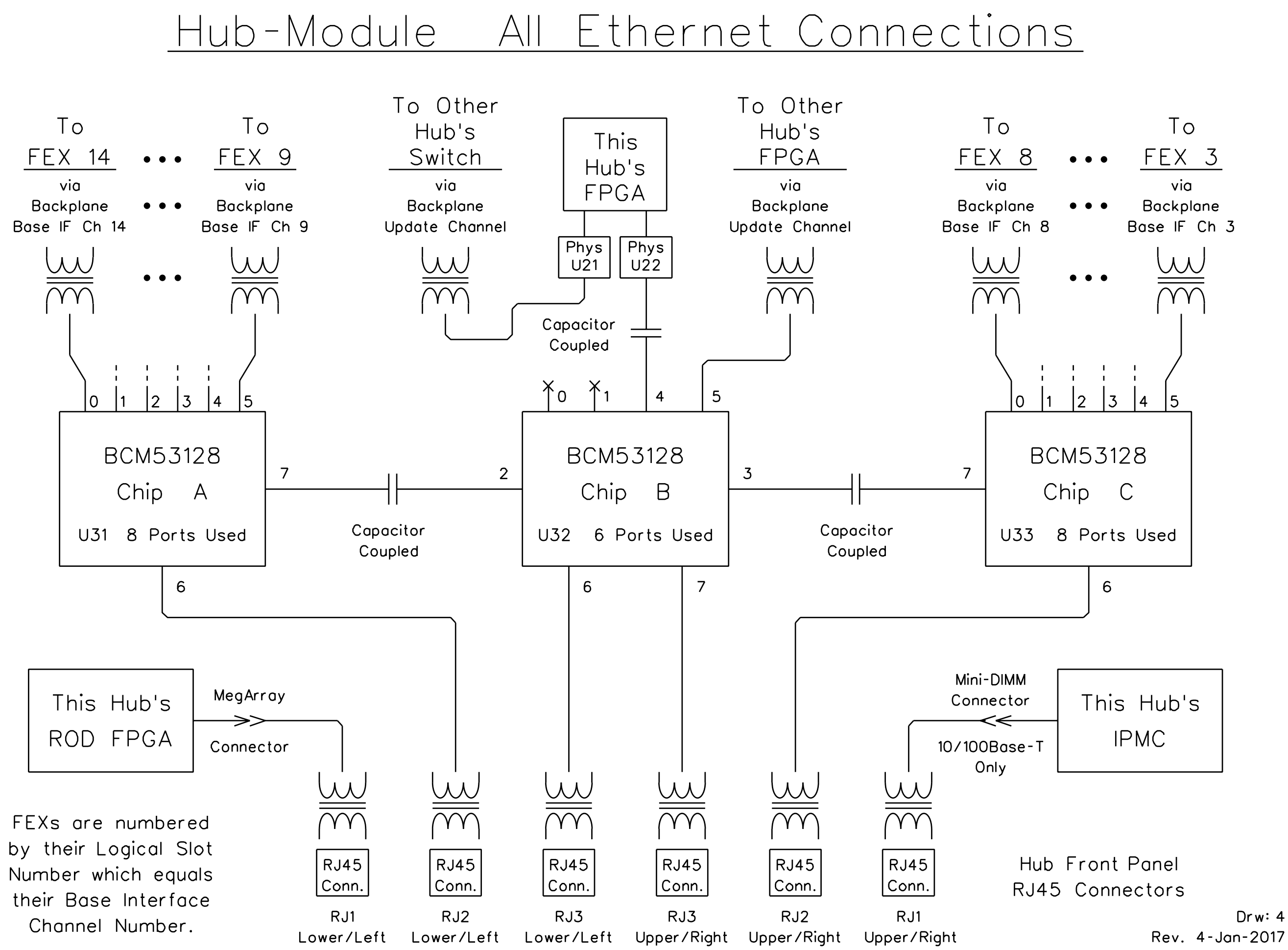

Figure 3: Hub GbE connection diagram.

1.4.1 Single Client, Single Server Connections

All combinations of single client, single server connections using the ports SW-A6, SW-B6, and SW-C6 of the FEX Hub Module were tested. All connections maintained a bandwidth of 935 Mbits/sec, regardless of which or how many switch chips the data was transferred through. This is the same bandwidth observed when using the commercial gigabit switch. The results are given in the following table.

Single Client, Single Server Test Results		
Server Port	Client Ports / Average Connection Bandwidth (in Mb/sec)	
	SW-B6	SW-C6
SW-A6	935	
SW-A6		935
	SW-A6	SW-C6
SW-B6	935	
SW-B6		935
	SW-A6	SW-B6
SW-C6	935	
SW-C6		935
Commercial Gigabit Ethernet Switch		
	Port 2	
Port 1	935	

1.4.2 2 Clients, Single Server Connections

In all 2 clients, single server connection configurations using the FEX Hub Module, an overall bandwidth of 940 Mbits/sec is maintained with a reduced bandwidth for each individual connection due to competing traffic. The same overall bandwidth is observed when using the commercial switch. Where the Hub and the commercial switch differ is in the distribution of bandwidth between each client/server connection.

When port SW-B6 acts as the server, we see an even distribution of bandwidth between each connection, as is the case with the commercial switch, but the distribution is uneven when port SW-A6 or SW-C6 is used as the server. In these cases, the "longer" pathway – where length is measured in terms of the number of chips the data being transferred in a connection must pass through – receives ~15 Mbits/sec more bandwidth than what would normally be expected based on the bottleneck of traffic (an even split of bandwitdth). The "shorter" pathway, on the other hand, receives ~15 Mbits/sec less bandwidth than what would normally be expected. Further testing suggests that it is indeed the "length" of a connection that affects the bandwidth distribution, rather than the distribution resulting from the workings of chip B specifically.

The inclusion of a chip C backplane port in these tests is intended both to show that the backplane ports are working and functionally equivalent to the front panel ports, and to further help demonstrate that the "length" of connection impacts its available bandwidth. The 2 clients, 1 server connection tests were repeated multiple times using different combinations of machines plugged into each port in order to test whether the machines themselves had an impact on bandwidth. The overall bandwidth and distribution remained the same, regardless of machine configuration.

Two Client, Single Server Test Results			
Server Port	Client Ports / Average Connection Bandwidth (in Mb/sec)		
	SW-B6	SW-C6	Chip C Backplane
SW-A6	454	487	
SW-A6	454		487 Mbits/sec
	SW-A6	SW-C6	Chip C Backplane
SW-B6	470	470	
SW-B6	470		470
	SW-A6	SW-B6	
SW-C6	487	454	
Chip C Backplane	487	454	
Commercial Gigabit Ethernet Switch			
	Port 2	Port 3	
Port 1	470	470	

1.4.3 3 Clients, Single Server Connections

In all 3 clients, single server connections using the FEX Hub Module, an overall bandwidth of 940 Mbits/sec is maintained. This meets the same standard as the commercial switch. But once again, the distribution of bandwidth is different when using the Hub than when using the commercial switch. The results of the 3 clients, single server tests mostly adhere to what one would expect based on the bottleneck of traffic and the previously observed trend of the "length" of a connection how much bandwidth it receives.

Consider the results of using port SW-A6 as the server and ports SW-B6, SW-C6, and a chip A backplane port as clients. The "shortest" connection, that between ports SW-A6 and the chip A backplane port, maintains a bandwidth of 453 Mbits/sec. The two "longer" connections have a combined bandwidth of 487 Mbits/sec. This is the same distribution of bandwidth seen in the 2 clients, 1 server test when two connections of different length compete for bandwidth. This makes sense given that the two "longer" connections enter switch chip A through the same port. In terms of competition for bandwidth in switch chip A, the two "longer" connections are treated as if they were a single connection, and so the distribution of bandwidth is reduced to a 2 clients, 1 server distribution. Furthermore, between the two "longer" connections the "longest" one, that between ports SW-A6 and SW-C6, has the greatest bandwidth. Indeed, we can see that the ratio of the port SW-C6 connection's 255 Mbits/sec bandwidth to the overall 487 Mbits/sec bandwidth of the two "long" connections is about the same as that between a 487 Mbits/sec bandwidth and a 940 Mbits/sec bandwidth. In other words, the port SW-

C6 connection receives ~52% of the bandwidth available in switch chip B, where it competes for bandwidth with the SW-B6 connections, which receives the other ~48% of chip B's bandwidth. Then the SW-C6 connection receives ~52% of the 487 Mbits/sec of bandwidth available to the SW-C6 and SW-B6 connections in switch chip A, while the SW-B6 connection receives the other ~48% of that available bandwidth. The same results are witnessed in the analogous test using port SW-C6 as the server.

Three Client, Single Server Test Results					
Server Port	Client Ports / Average Connection Bandwidth (in Mb/sec)				
	SW-B6	SW-C6	A Backplane	B Backplane	C Backplane
SW-A6	232	255	453		
SW-A6	313	313		313	
SW-A6	454	243			243
	SW-A6	SW-C6	A Backplane	B Backplane	C Backplane
SW-B6	235	470	235		
SW-B6	327	327		286	
SW-B6	470	235			235
	SW-A6	SW-B6	A Backplane	B Backplane	C Backplane
SW-C6	243	454	243		
SW-C6	313	313		313	
SW-C6	255	232			453
Commercial Gigabit Ethernet Switch					
	Port 2	Port 3		Port 4	
Port 1	313	313		313	

The other results can be explained in the same way, examining both the bottleneck of traffic and the "length" of each connection competing for bandwidth at an individual bottleneck. A somewhat notable exception occurs when using port SW-A6 as a server and ports SW-B6, SW-C6, and a chip B backplane port as clients. Based on the previous results, one might expect the SW-C6 connection to have more bandwidth than the other two connections, since it is "longer", but instead we see an even 3-way distribution of bandwidth. The same result is witnessed in the analogous test using port SW-C6 as the server. This distribution would be expected based solely on the bottleneck of traffic at switch chip B, but does not adhere to the trend of "longer" paths receiving more bandwidth. This warrants further investigation into how/whether the "length" of connections affects the distribution of bandwidth.

1.5 Conclusion and Areas of Further Study

From these results it can be seen that the FEX Hub Module's three switch chips are functional and support a bandwidth of 935 Mbits/sec when sending data from one port to another, or 940 Mbits/sec when data is sent to a port from multiple other ports. This meets the same standards as a commercial gigabit switch. This bandwidth is maintained when traffic is generated by up to three ports. Further testing should examine whether this bandwidth is maintained under higher-stress conditions, such as when a single port receives traffic from all other ports, or when several ports receive traffic from many other ports.

Distribution of bandwidth occurs as one would expect based on the bottleneck of traffic in each switch chip, only with the condition that traffic which travels over a longer distance, in terms of the number of switch chips it passes through, receives more bandwidth than traffic traveling over a shorter distance (when the traffic is being generated simultaneously). This condition generally appears to hold true, but as stated previously, the distribution of bandwidth in certain client/server configurations warrant further investigation under exactly what circumstances the condition holds. Nevertheless, the results gathered suggest that the "length" of connection only amounts to an at most ~15 Mbits/sec difference between the actual bandwidth for each client/server connection and the expected bandwidth based only on traffic bottleneck. This difference is minor compared to the actual bandwidth of each connection, and as such does not pose an issue, at least not when few ports are communicating.

It is possible that this phenomenon may be an issue when sending data between many different ports across multiple switch chips, as the bandwidth of each connection becomes more divided. For example, if the bandwidth of the switch chips is so divided among connections that each connection has a bandwidth of at most 50 Mbits/sec, then a 15 Mbits/sec difference in bandwidth is much more significant than when we are dealing with bandwidths in the realm of 470 Mbits/sec. However, some results of the 3 clients, 1 server tests suggest that the difference "length" makes in the bandwidth of a connection may be reduced as the number of competing connections with different "lengths" increases. Further testing should examine this and whether the effects of communicating between multiple switch chips while many ports send and receive traffic are significant enough to warrant concern. However, if switch chips A, B, and C are not connected (their connecting circuitry is not capacitor coupled) then there are no "shorter" and "longer" paths to consider, and as such bandwidth would be distributed equally between connections running through a single switch chip, just as it is in a commercial gigabit Ethernet switch.