What can Ethernet provide for Big Data systems? Ideas for getting more out of an Ethernet Fabric

Casey Miles, Brocade
cmiles@brocade.com

1. Introduction

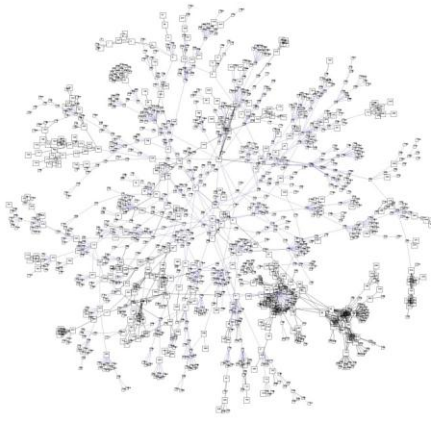
Ethernet is rarely the focus of a Big Data conversation. Usually the conversations tend towards scalability, new applications that leverage massively parallel file systems, or trying to put a definition to the term "Big Data". When Ethernet is mentioned it is quickly waved off with, "Yes, we understand that Ethernet will be the network" and not given another thought. What I would like to spark in this paper is the idea that the network itself can provide information about data and has the potential to turn a good solution into a wow solution.

2. Identifying Networks

Let's start by first identifying where you most commonly would find Ethernet in a Big Data system. We'll start from the very core of the system and work our way outwards to all the attached and dependent systems that feed into a Big Data cluster. Keep in mind, this is a generic example and there will be modifications for different applications. Many Hadoop-like file systems have settled on gigabit Ethernet as the standard interconnects for small to medium size clusters. It is the natural selection as these days, many server boards come standard with multiple gigabit Ethernet ports built conveniently into the board. Some larger systems or systems who's applications require a bit more throughput are looking to 10 gigabit Ethernet as the interconnect and 1 gigabit as a management or monitoring network. These systems are either using data at rest or data in motion for their analysis, so this requires a data center network attached to a corporate network, a storage network, or straight to the Internet depending on function. All of these networks are mostly Ethernet networks. There are many Ethernet switches and routers in the walls and down the hall quietly humming away supplying volumes of data to these systems. What if they could take a more active role? What if they could be more than the canvas of this painting?

3. Sensor Networks

I'm sure most of you who have enough interest to read this brief paper have also heard that data is multiplying exponentially year over year and that 2009 signified a change in who or what produces the most data. The cross happened quietly in the night some time during 2009 that sensors now create more data than people do. This leads me to one area where I believe Ethernet has an opportunity to provide a lot value to Big Data: sensor networks. To build on a common example, consider the thousands of miles of railroad track in the US. Every few feet there is a sensor reporting if the track is broken, weight of the train going by, if there is a train going by at all, temperature, and a host of other readings I am not aware of. These sensors connect back to a



server and that server over an Ethernet network back to a server at some central site where the data is either processed as it comes in or stored for future use. There will be an application, if there isn't already one, that will go through these logs and pick out bits of information to help the railroad run a more efficient operation, reduce maintenance downtime, improve shipping speeds, and anything else you could think of to ask a Big Data machine armed with endless logs of sensor information. What if the switches that connected all those sensors were a little smarter? What if the switches reported back trends, spikes, and collated information about all the servers plugged into that switch or network of switches? Would it be important to see seismic activity hitting one sensor then another then another on a single pane of glass display as it is happening? The individual sensors will tell you what they are recording but are they aware of what the sensors next to them are recording? Are they putting that data together in a way that means more than just the individual readings? What if that data could be packaged and sent to a Big Data machine in a data format that could add additional insight to what is unfolding in the field? What if it was global financial markets or a battlefield instead of trains? Would that extra second be worth something? Would the collated metadata in short provide additional insights about why all-of-a-sudden a few sensors in one area are producing a lot more data than the rest or they had in the past?

4. Controlled Interconnects

Ethernet is the great untapped resource in Big Data clusters. There are many opportunities to incorporate the fabric that touches every Big Data machine as well as the data that it will process whether that data is at rest or in motion. Another area of untapped potential is dynamic Ethernet fabric control. Consider for a moment that every application of Big Data has a few phases. At times, a parallel file system may rebalance itself, multiple jobs could be running at the same time, or a hero machine could go through a few different phases of a large reduce function. In each situation different aspects of Ethernet become important. For example: A data reduction function starts and queries every node for interesting traffic. Some nodes will respond and move a portion of data to intermediary nodes. In this scenario, there have already been two vastly different requirements on the network. In the first portion, there were many small frames passed to all nodes. A full mesh, capitalizing on low latency would be the best design for this function. However, in the second stage, when a few nodes moved data to intermediary nodes for further processing, a fat-tree would be the best topology to capitalize on aggregate bandwidth over many equal-cost paths; a mesh would be a nightmare of latency and dropped traffic! Imagine if you could have both optimal solutions automatically form when you need them without user interaction. Imagine if your application or your file system could speak to the network and prioritize certain traffic, change topologies with OpenFlow technology from a mesh to a fat-tree, or reserve cache on a granular level for your most important flows on the fly. Building a fabric for a Big Data cluster without knowing what the application traffic patterns are is like making a suit without having any measurements. Odds are, some pieces might fit but it won't be a good fit. Imagine if the fabric itself was self-forming and changed based on what the application needed at any given time. These are just a few areas where the canvas can change to add more to the painting. The network is more than the canvas. The Ethernet fabric is the nervous system of Big Data and can respond and change to the environment. The network has a lot to offer.