Trust within Open, High-End Networking and Data Centers for Understanding and Improving Cybersecurity

Author: Trung Tran, Altera Corporation

Problem statement:

Scientific computing has grown in size because the amount of data that can be collected from sensors has improved significantly and because the internet has given us the ability to include a variety of data sources. The size of this data presents a problem in compute resources because it becomes impractical to store and maintain petabytes and even exabytes of data in a single location. Increasingly the scientific community is looking towards private/public cloud data centers to provide a flexible way to maximize the amount of computing that needs to be done. These trends has made security a concern as the community faces the reality that they can no longer rely on physical security to ensure the integrity of their results. Previously, it was impossible to be able to even run virus scans on large amounts of data because of time it takes to run the software. Any attempt to evaluate the data integrity of the information seems unlikely. Add to that the complexity of doing that scan across distributed networks and the task becomes even more complicated.

FPGA torus systems can be used to accelerate real time data analytics. The new Stratix 10 family of products have both the bandwidth (96X28G XCVRs) and the processing capability (10 TFLOPs of single precision floating point). These are critical elements needed to be able to evaluate data in real time. This technology can easily be scaled and sized. The following approach discusses the type of IPs which can be utilized in a FPGA Torus to ensure data integrity.

Approach:

Step 1: Encryption

1a) Partnering with Mellanox, we are able to provide encryption down to the file level with open routing on a standard infiniband network which is used in most Storage Area Networks. The key is to MACSEC encryption and doing the encryption with very little added latency in the network. This means understanding the difference between control packets and the data packets which will actually be written into the hard drive. Each file can be given its own separate key and public key certificate exchange can be used to ensure that only authenticated users can have access/rights to view and or modify the data. Physical theft of drives can be defeated because the data is never stored in clear text. Also network rules can be set-up to prevent local storage of the data on client machines. This also clears the issue of public/private cloud because the data in the cloud is never stored or routed unencrypted.

1b) Altera has demonstrated the ability to create a light weight cypher which can be incorporated into remote senor equipment without the cost of a heavy weight cypher such as AES. Remote sensor equipment are constrained by the amount of power which can be provided to them and the amount of circuitry they can support. The cypher we demonstrated is a tenth the power of AES and 5% the logic required by AES. It also adds only once cycle of latency which is important when collecting real time data.

Step 2: IDS/IPS

Standard SNORT Intrusion Detection System and Intrusion Prevention System (IDS/IPS), allow systems to monitor traffic for malicious code and determine the risk that it presents to the system. Often times though the implementation of SNORT require tap aside networks which effectively 10x the amount equipment needed to secure a 100G link. The size of data sets need a solution which scales to 600G and is a bump in the wire approach to reduce the power and space footprint that this type of security system can provide. Working with Clemson University, Altera can provide a scalable version of SNORT with a standard Linux interface which scales to 400G and is a bump in the wire system without the need for a tapped network,.

Step 3: Traffic Analytics

Altera is working on IP which can create large scale graphs of relational data which can be used to track the access to data inside the network. Using network logs, we can graph the IP interactions between multiple devices in the network to see which location and who in those locations are accessing critical data. A sense of normal operation can then be established in the network. Additionally a machine learning algorithm can use Breath First Search and Depth First Search, to look for changes in the edges and vertices of the graph and be able to on a sliding window basis be able to alert the network of any activity which deviates from the normal.

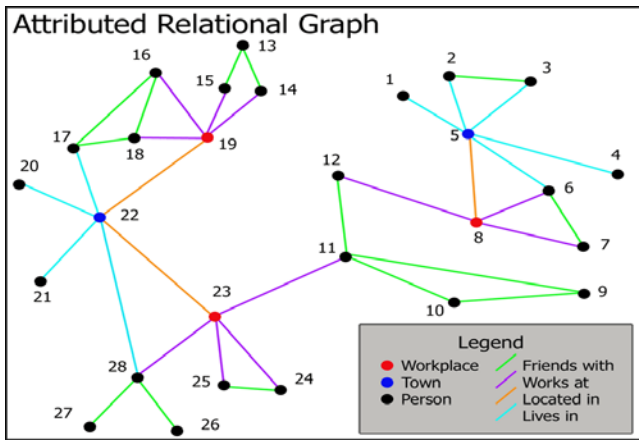

Figure #1: An attributed relational graph

Step 4: Streaming Data Analysis

Altera and IBM are developing a non- Von Neumann compute architecture specifically to evaluate the relevance of streaming data and do initial analytics on those datasets in order to filter out or correct data. The key is real time feature extraction of data characteristics and the evaluation by a machine learning algorithm of the relevance of the data based on a predefined criteria. Examples of where this is used includes internet search engines, object detection of you tube video, or large scale radio astronomy. A criteria can be created so that data can be validated before its computed. Data that does not meet the criteria can be stored

or deleted. The criteria does not need to be a direct match but can be base on a variety of factors with a sliding window approach allowing for a confidence value being assigned to the data. Figure #1 shows a model which takes streaming information and makes a decision of the integrity of that data.

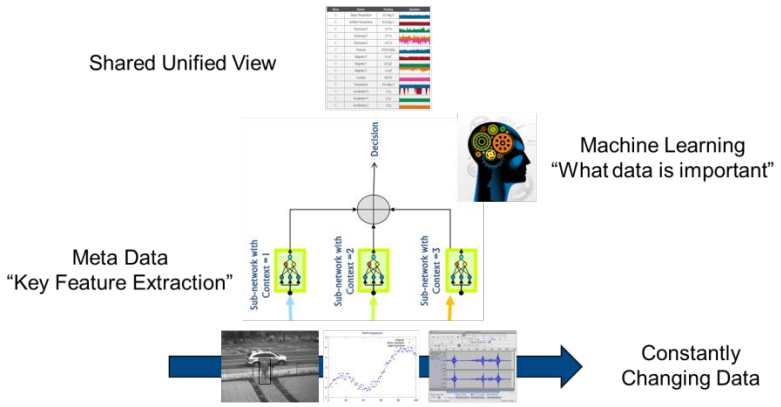

Figure #1: Streaming Analytics Steps

The criteria which can be used can be as simple as source of the information, authentication tags to determine access, appropriate size of the data. Or it can be as complicated as searching through the structure of the data and the data types inside the data set and checking to see if matches the type of data that is expected in the system. This filtering out of non-authorized sources in band and in real time, will ensure that the data that is being processed is from a valid source and can be considered reliable. The result of the computation can then be considered valid. The process for this can be done in a scaled out manner using one FPGA for 10G to thousands of FPGA for exabytes of data.

Conclusions:

Ensuring confidence in the results of scientific data is critically important especially when the results can affect decisions at the government level as is the case with global warming. The open nature of cloud data centers and collaborative computing makes ensuring the integrity of that data very difficult. A multi-tiered approach to security is needed to ensure that not only is access to the data restricted but also that any attempt to access is monitored with possible tampering flagged. The key is using the FPGA as an accelerator which allows for the streaming of large amounts of data and which allows for the analysis of that data in band and in real time.