# Maximizing Visibility and Flexibility for 10GbE/FCoE/iSCSI Analysis

The decision whether to use a hardware-based analyzer versus a software-only utility has been an ongoing debate for almost as long as Ethernet has been around. On the surface, the discussion seems to focus on price, especially with the availability of several packages based upon the open-source library pcap. While it is difficult to argue with free, these software-based analyzers still need to provide the capability to reliably analyze and troubleshoot networks if they are to be worth their price. A tool that fails to uncover problems will end up requiring significant investment in terms of additional development effort and lost time-to-market. Even worse, if problems make it undetected out the door, the damage to brand and the cost of updating/recalling product can be substantial. As the old adage goes, "You get what you pay for."

While there are appropriate uses for software-based analyzers, there are many reasons these tools simply fall short when analyzing emerging network technologies. Design, development, and troubleshooting processes are changing as the market shifts to 10 Gigabit Ethernet and Fibre Channel over Ethernet (FCoE). This shift is exacerbated with the simultaneous development of Data Center Bridging (DCB), formerly known as Converged Enhanced Ethernet or Data Center Ethernet.

The challenge of determining where errors are actually occurring, with the increasing blend of storage and networking, is spilling over into the domain of mid-tier vendors. Even SAN administrators, with the evolution to "lossless Ethernet" for FCoE, now must deal with networking issues that until now have always been the domain of "those network guys". In order to succeed, developers and administrators alike will need full access to data, including 100% traffic capture at line-rate, full-duplex visibility across protocol domains, and flexibility in analyzer placement in order to guarantee that network equipment performs as required.

## Analyzing Ethernet

Part of the challenge for developers making the decision about which analyzer to use is their familiarity with the maturity of Ethernet. Traditional Ethernet analysis, in most cases, is well-served by software-based tools, and there are several good reasons for this. For example, just about every 10/100 Ethernet network adapter (i.e. Intel, 3Com, etc.) supports a promiscuous mode to analyze networks. When set in this mode, these adapters will capture, to the best of their ability, streams of incoming and outgoing data into buffers for user analysis. Many also calculate basic link statistics – such as CRC errors, fragments, collisions, broadcast frames, multicast frames, and other anomalous/troublesome network behavior – which can be accessed through programs such as the Microsoft SMS Network Monitoring Agent, Wireshark or tcpdump.

Developers of Ethernet products also can rely heavily upon the fact that the IP and TCP Ethernet layers have been implemented in software. This facilitates "live" debugging with meaningful information extracted directly from drivers and microcode that can be tuned for individual data streams. Furthermore, a range of analysis tools are available for many levels/layers of the different software stacks that are handled by the network application layers. Programs such as Microsoft's Perfmon, as well as a multitude of SNMP and HP OpenView snap-in software packages, can monitor performance and behavioral traits in the network.

## Lossless Ethernet

The IT industry, however, never stands still for long. With its constant shift into a new world of protocols and transports that are more than just evolutions of previous standards, what is transmitted over the wire is becoming more than just the worry of NIC and switch engineers. Many developers believe that since Ethernet is a mature technology, errors are automatically handled and recovered by the TCP and IP layers, making hardware-based analysis unnecessary.

The rise of 10 GE and so-called Converged Enhanced Ethernet for next-generation data center networking is proving that Ethernet still has work to be done. With FCoE being ratified as a standard, the concept of "lossless Ethernet" continues to gain popularity. As stated by FC-BB-5, Clause 7.2:

*"Although a generic Ethernet network may lose frames due to congestion, a proper implementation of appropriate Ethernet extensions (i.e., the Pause mechanism defined in 802.3-2005, Part 3) allows a full-duplex Ethernet link to provide a lossless behavior similar to the one provided by the buffer to buffer credit mechanism in native Fibre Channel. The protocol mapping defined by FC-BB_E is referred to as Fibre Channel over Ethernet (FCoE) and requires the underlying Ethernet layer to be full-duplex and lossless."*

An Ethernet network that is lossless is a much different protocol to analyze than traditional Ethernet. For example, the Pause mechanism used to help achieve lossless performance is implemented through the use of MAC Control Pause Frames. Pause frames are valid at the MAC layer only and are never passed up to the receiving network stack. As a consequence, software-based analyzers are hard-pressed to report on the use, or misuse, of Pause frames. When analyzing an FCoE implementation, this introduces a very large blind spot as regards accurately assessing network performance. With the introduction of Priority Flow Control, knowing whether your equipment is using the newer PFC instead of traditional MAC PAUSE becomes an even bigger issue.

This reliance of software-based analyzers on the network stack is the source of a wide range of potential problems. Many software analyzers depend on the Ethernet hardware to only pass up valid frames to the networking stack. However, if a frame is dropped due to an incorrect CRC, the analyzer has no visibility into why the frame was dropped, or even that it was dropped. Developers have only incomplete information as to why performance is not meeting expectations or, as might be the case for FCoE, why there is an Abort Sequence suddenly happening.

When analyzing TCP-based traffic, such as iSCSI, ignorance of missing frames can cause even greater troubleshooting headaches. Was the missing frame a simple duplicate ACK or a large run of missing segments? Or were any segments really missing at all? This problem can be exacerbated by the use of TCP Offload Engines (TOEs). For example, it may not be clear whether the software capture is seeing frames before or after they have been processed by the TOE. Many software captures are rife with supposed TCP checksum errors merely because the frame has already traversed an external stack. In addition, the increasing use of TCP Offload Engines has begun to blur the line as to whether performance data is available or not.

Even 10GbE is too much for software-based analyzers at this time. 10 Gb/s throughput stresses system resources. Because software-based analysis competes with the OS (and by extension, the networking stack) for resources, it can easily run out of buffers with which to capture frames. As a result, the capture will be missing frames which may contain important clues as to the performance of the network, frames that need to be seen in order to make intelligent decisions about the design and stability of a particular component or technology. For example, if the capture is missing any of the multiple TCP segments across which iSCSI Protocol Data Units (PDU) are spread, reconstruction of the iSCSI conversation may be impossible. Note that as the 10 Gigabit Ethernet standard only allows for full-duplex transmission and that such conversations can easily run into gigabytes of trace data for even a short capture, the problems arising from incomplete captures are not isolated or even rare.

Despite its ubiquity and perceived maturity, Gigabit Ethernet is also not without its ambiguous grey areas and corresponding interoperability issues. Some of these issues are masked by the fact that a frame must traverse the card prior to being analyzed by the software capture utility. Consider the use of Jumbo Frames between two different vendor's GbE cards. Using just a software-based analyzer, developers may be unable to determine that a card sent the packet size expected or that the receiving card even received a frame. If the received frame is not of a size the receiving card expects, it may silently drop it or allow the upper layers to respond with a Please Fragment message. With only a software-based analyzer, a dropped frame would not even be recognized in the analysis.

## Hardware-based Analysis

Hardware-based analyzers provide substantially more information and more options for developers than software-based analysis tools. They allow access to traffic at all layers and provide a richness of information simply not possible with software-based analyzers:

- Hardware-based analysis is able to capture frames as they are transmitted, compute their correct CRC, and flag them as being a bad frame if the CRC is wrong. Software analyzers may not even be aware that a bad frame was ever sent or received.
- Repeated testing has shown that software analyzers catch as little as 15–20% of the total traffic passing over high speed links. Hardware analyzers, with dedicated ASICs and capture buffers, capture 100% of traffic.
- It is important to bear in mind that iSCSI and FCoE require full-duplex data to be analyzed – SCSI and TCP are bi-directional protocols that cannot be analyzed using only half of a conversation! Most modern Ethernet switches do provide a way to monitor full-duplex traffic through the use of a mirror port, after which a software analyzer attached to that port can monitor the conversation. A mirror port, however, may not mirror all data when the switch is congested since the forwarding of real frames is a higher priority than mirroring. As such, there is no guarantee that all of the frames will ever be delivered out of a mirrored port.
- Accurate decoding is critical to fast identification and resolution of network issues. However, while a software-based analyzer may have the FCoE portion of the frame correct, most consider the embedded FC frame in the same way they consider embedded TCP traffic: it's just "data". Decoding this "data", however, can have a huge impact on one's ability to understand the operation (or lack of operation) of a converged FCoE network. To this end, most vendors of hardware-based analyzers have Fibre Channel backgrounds and have a vast brain trust to rely upon in the decoding of what is on the wire.
- Many initial deployments of FCoE will rely on bridges into an existing Fibre Channel infrastructure. The mixed protocol capabilities of hardware analyzers provide vital visibility into the Fibre Channel portion of FCoE packets, enabling developers to analyze both sides of a conversation to determine whether the sending adapter or the bridge is causing an issue. No software packages exist that will allow visibility into the Fibre Channel portion of transactions.
- The addition of several new pieces to the Ethernet family to create Data Center Bridging has created an interoperability challenge not seen since the early days of the protocol. Priority Flow Control (PFC), Enhanced Transmission Selection (ETS) and Congestion Notification are all Layer 2 conversations, and as such require the use of a hardware analyzer to understand what is being done.
- Hardware analyzer vendors are very active in the standards committees. Many times, these vendors are aware of changes or new standards almost from conception. The analyzers are used throughout the development of the protocols, and as such allow the vendors a chance to 'fine-tune' the decode information so when the standard is finalized the hardware analyzer is ready to be used in development and production environments.

## Connecting to the Network

The most common implementation method used by software analyzers is end-point analysis where the analyzer software is installed onto one of the end-points of the desired conversation and set to capture traffic from one or more of the interfaces available. This many times involves the use of a capture library that places the interface into "promiscuous" mode, allowing it to capture all data it sees on the wire, even if it is not addressed to that interface. With the almost ubiquitous use of switched networks, this will still result in traffic bound only for that adapter being captured. The installation of the capture utility has also now placed an extra load on the host processor, which can result in the misidentification and time-consuming resolution of false performance issues. Hardware analyzers can provide a better end-point analysis by being placed in-line directly before the end device under test.

In-line analysis is almost exclusively the domain of the hardware analyzer. Simple to set up, this method provides 100% capture as well. Given that the analyzer is placed in-line, it does disrupt the link being monitored when the analyzer is inserted and removed. While analyzing in-line does add two more connections and a cable, digital retiming modes on the analyzer can be used to regenerate signals, thus avoiding adding any jitter on the link. Regeneration also has the desirable effect of making it easier to locate problems that arise from signal degradation.

It is easy to forget the vast amounts of research that go into bringing a new protocol to life. Once this work is completed and a protocol matures, then it is reasonable to assume that the major incompatibilities and ambiguities have been worked out and that testing can be more relaxed. However, during the early stages of a fledgling protocol's rise, a hardware-based analyzer is often the only way to verify adherence and compliance to specifications. Because of their early involvement, hardware-based analyzer vendors are in a unique position to gain in-depth knowledge of a protocol and enable packages to be much more robust and reliable at the release of a technology. In contrast, software-based analyzer developers many times must wait to see new technologies already deployed before they can begin considering supporting a new protocol or transport.

There are circumstances where the use of a software-based analyzer offers good value. For those times when developers only need to make a quick check that traffic is getting through and that it is nominally formatted correctly, then a software-based tool may be the best option. However, as communications protocols continue to increase in data rate and protocol complexity, software based analysis tools lose ground on providing them full visibility into the network. Having a hardware based protocol analyzer that captures 100% of traffic at line rate becomes necessary. Moreover, only a hardware analyzer is capable of verifying next generation enhanced Ethernet technology, confirming that data integrity and "lossless" congestion management mechanisms operate as expected. While software-based analyzers have been and will be a useful tool, it is clear that the hardware-based analyzer will continue to be at the forefront of compliance and accuracy testing for the foreseeable future.

Contact Us **+1 844 GO VIAVI**
(+1 844 468 4284)

To reach the Viavi office nearest you,
visit viavisolutions.com/contacts.


10gfcoeiscsianalysis-wp-san-tm-ae
30162830 900 0117