



*Measuring and Minimizing Tap Switchover Times
Using VSS Monitoring vAssure™ Technology*
Technical Note

Before the prevalence of Gigabit Ethernet over twisted pair, network taps consisted of fiber-optic splitters or passive electronic amplifiers connected to 10/100 twisted-pair wiring.

Passive network inputs present a problem for multi-input devices such as port aggregators which must also service single-output links such as SPAN ports. A passive tap will not work with a SPAN port input, since there must be a link partner on the other end of the cable. Thus, the SPAN port will never receive any incoming link pulses and will not send any data to the aggregator device.

The no-link problem is addressed by designing the tap such that each network port links directly to the tap, rather than being passively connected to another port. The tap then acts as a Layer 1 bridge to pass packets between the two network ports, while duplicating the packets to the monitoring ports. This allows inputs to also be connected to single-output devices such as SPAN ports, packet generators and the monitoring ports of other taps.

Active network-port architecture is also required for all copper Gigabit-capable taps. Passive monitoring of 1000BASE-T is not possible due to both sides of a link transmitting simultaneously on each of the four wire pairs in a cable. As each end of a 1000BASE-T link knows what it is transmitting, it can subtract this signal from what is seen at its own receiver, and thus recover the signal arriving from the other end. However, a tap between two Gigabit network elements has no way to separate the two mixed signals as it does not know what either side is sending.

Two copper Gigabit Ethernet capable devices normally use the standard auto-negotiation procedure to establish a link between each other. Each end device sends fast link pulses on its Tx pair (pins 1 and 2) to the other device. These link pulses are received by the other device on its Rx pair (Pins 3 and 6) and are used to initiate a conversation between the two devices.

This fast link pulse connectivity is used by both devices to specify the capabilities they support, such as Speed (10, 100 and 1000), Duplex (Half and Full), Device Type (Hub/Switch or End Point) and PAUSE (Symmetric and Asymmetric). Once each device receives the other device's capabilities, it attempts to negotiate the best possible combination of settings both devices can support. Additionally, Gigabit-capable devices must determine which end will be the master clock source. As Gigabit

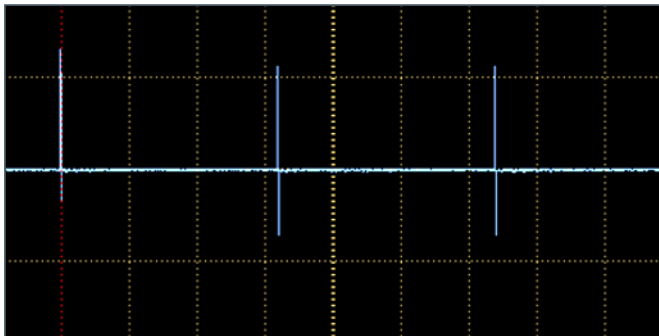


Figure 1. Auto-negotiation Fast Link Pulses

links must be synchronous with each other, only one side can use its internal clock to send data. The other "slave" device must use the incoming data from the "master" device to clock its transmitted data. This designation of which device is master and which device

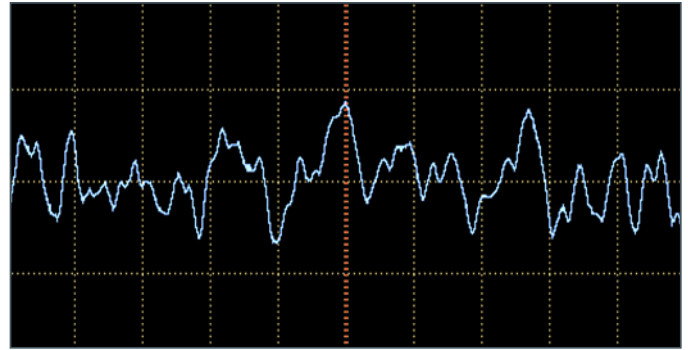


Figure 2. Gigabit data signal

is slave is decided during the auto-negotiation process.

Note that most devices also have auto-MDIX, which enables internal swapping of the Rx and Tx connections automatically. Thus a device will send fast link pulses on both Tx and Rx pairs until a conversation is started with another device.

Once all parameters are negotiated, the devices bring the link up in the correct state. In the case of Gigabit devices, the master side sends an idle data pattern to the slave, which the slave uses to phase lock its clock. At that point the slave sends the same idle pattern back to the master and the link is marked by both sides as up.

The scope picture in Figure 2 shows what appears to be a random signal. This is due to the fact that the signal is a composite of both ends transmitting an idle pattern on the same wire simultaneously. Each end subtracts its own transmitted signal from what is seen on the pair, thus determining what the other end is sending. Because of this mixture of signals, it is impossible to use passive sniffing to watch Gigabit Ethernet traffic. A device viewing the signal has no way to know what either end is sending, and thus cannot separate the two mixed signals from each other.

The solution is to have the tap device link with each network device, decode the data, transfer the data to the output of the other network port, and then re-encode the data to pass it on.

Gigabit taps can be viewed as a specialized Layer 1 hardware bridge; accordingly, each network element establishes a link with the tap. To provide a link between network ports during power failures, network port pairs must be physically transferred from the tap input ports to their matching network port. This causes a very short interruption of data flow upon power loss and restoration. While the physical transfer of the ports is very fast (typically < 5 milliseconds), link speed, configuration and re-starting of auto-negotiation may cause short loss-of-link events. These loss-of-link events can introduce delays of up to 2-3 seconds before the link recovers. If upper link-layer protocols such as Spanning Tree Protocol are in use, this short loss of link event can cause links to be held inactive for many seconds and initiate

unnecessary network topology reconfigurations.

For network configurations that cannot tolerate interruptions of several seconds, port and network device configurations can be adjusted to reduce such delays to tens of milliseconds with no apparent link loss to the connected devices.

For 10 and 100 Mb/s links, auto-negotiation can be disabled for faster link switchover times. Note that most devices, which have auto-MDIX, disable this capability when auto-negotiation is turned off. Thus, it may be necessary to use cross-over cables after turning off auto-negotiation or to manually configure tap and network device ports for MDI or MDIX on ports using 10 and 100 Mb/s.

For Gigabit links, auto-negotiation must remain enabled per the 802.3 standard. If auto-negotiation is disabled on a Gigabit copper link, the link will typically fail to start.

The most common reason for long link transfer times is the use of auto-negotiation on the link. VSS Monitoring taps implement vAssure, a proprietary Gigabit link recovery technology that allows tap power up/down without loss of the link, even with auto-negotiation enabled.

The second most common cause of long transfer times is having Spanning Tree or Fast Spanning Tree enabled or not properly configured for the port. Spanning Tree is a protocol that enables two or more switches to be interconnected with multiple ports without causing routing loops. The default mode for most switches that support this protocol is to start up in a listen-only mode for up to 45 seconds after the link comes up. This results in an extended interruption of packet flow even if the link is up. This can be addressed in several ways. One way is for Spanning Tree to be disabled if multiple parallel links are not used for this switch port. If Spanning Tree is required on the port, most switches will allow the port to be configured to immediately start routing packets upon the link up state. An example of this is the "portfast" option in Cisco® switches. The other way is where some switches will allow setting the timeout value for how long a link must be down before it is declared disconnected or unroutable. If this timeout is made 3 seconds or more, the link dropping and recovering during a switchover event will not cause a spanning tree re-route and data flow will continue as soon as the link is re-established. However, neither of these are recommended or desirable.

VSS taps with vAssure allow Spanning Tree to remain enabled and configured as desired by the user, as the tap switchover does not cause loss of link.

Typical switchover times with different link configurations

Auto-negotiation 10/100	500-1000ms	Link resets
Auto-neg. 10/100/1000	1-3sec.	Link resets
vAssure 10/100/1000	30-60ms	Link stays up
100 Forced	5-200ms	Link stays up
10 Forced	5-200ms	Link stays up

Table 1. Link Switchover Test Data

Lab Test

All testing was performed by inserting two taps in-line with a pair of Spirent® GX-1420B Gigabit Ethernet cards. Each card was configured to transmit packets of random byte length with minimum inter-frame gap (96 ns). The two test streams were started and one of the in-line taps was powered on or off.

The second in-line tap had an oscilloscope connected to the Data Valid (DV) line of the GMII interface of the Physical Layer chip for the network port. This line is held high during receipt of valid data; thus, any data interruption will display as a reading of 0V on the scope. Note that the 96ns inter-packet gap will also be visible as a low on the DV line. The traces below show intermittent low glitches in the DV lines due to the inter-packet gap being aliased because of the slow scope speed required to capture the entire data loss event.

Traces are identified as follows: 1) Reset line of the powered on/ off tap, used as trigger for oscilloscope; 2) Network port DV line.

This is typical behavior with Gigabit taps that do not have vAssure technology. Note that during the loss of link, most devices such as a switch or NIC will no longer transmit packets. Given the large number of packets that may arrive during this link interruption, most will be lost at the device, because required buffer size to hold this many packets is several orders of magnitude larger than a typical port buffer on a switch.

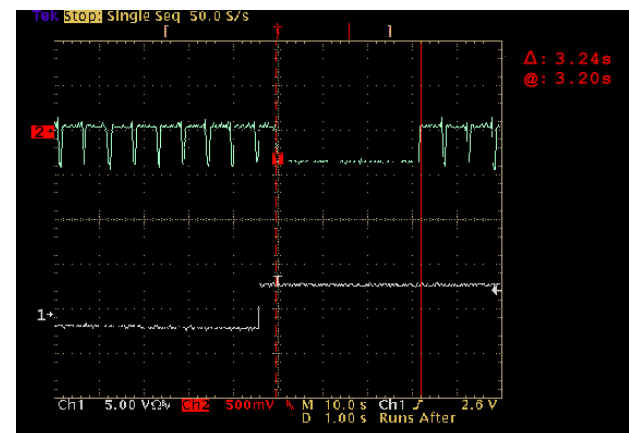


Figure 3. Power-Up without vAssure

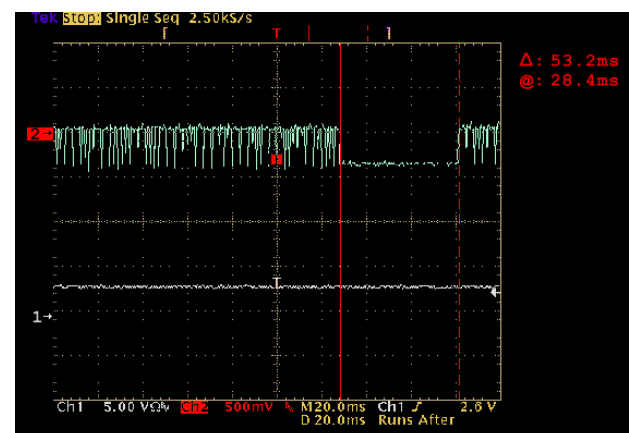


Figure 4. Power-Up with vAssure

Figure 3 depicts a data stream interruption during the power-up on a Gigabit link with auto-negotiation enabled. The data shows that a data-flow interruption lasts approximately 3.24 seconds.

Whereas, Figure 4 shows the same event with VSS Monitoring's vAssure enabled. The data flow has dropped from seconds to a mere 53.2 milliseconds. With vAssure, connected devices do not see a loss of link and do not have to re-negotiate the link on the physical connection.

During the same lab test, VSS captured data from a test platform saturating the Gigabit link at 100 percent. The same event as shown above is depicted in Figures 5 and 6. Figure 5 shows the number of random packets lost during power-up with auto-negotiation enabled without vAssure. Approximately 500,000 packets are lost in each direction during this 5-second network interruption. The same event with VSS Monitoring's vAssure, depicted in Figure 6, shows a 60x improvement with only 8,441 packets lost at 100 percent saturation.

VSS Monitoring is leading the way in innovation for high-availability Ethernet, Fast Ethernet, Gigabit, and 10-Gigabit Ethernet monitoring. For more information, please visit <http://www.vssmonitoring.com>.

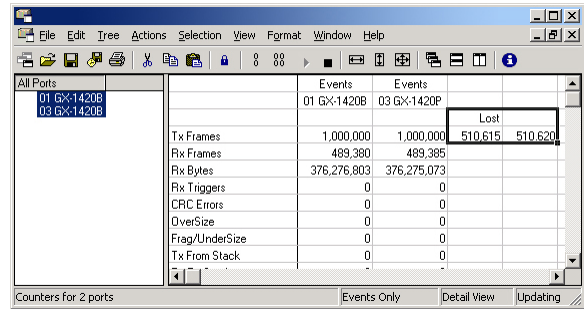


Figure 5. Test tool saturating a Gigabit link without vAssure

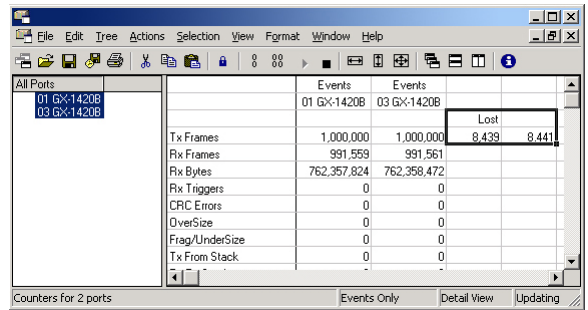


Figure 6 - Test tool saturating a Gigabit link with vAssure

Real-World Test

Usage:

```
fping 192.168.1.2 + 10 -n 2000 -w 1
```

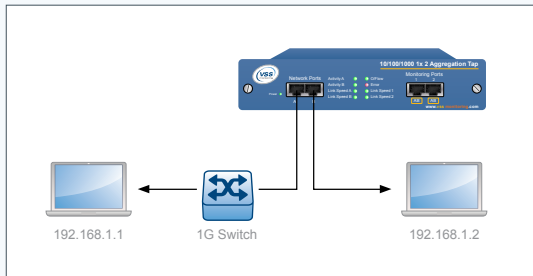


Figure 7 - vAssure Test Case

Results:

Normal Case: Power Off

Ping statistics for 192.168.1.2:
 Packets: Sent = 2000, Received = 1973, Lost = 27 (1% loss)
 Approximate round trip times in milli-seconds:
 Minimum = 0.1 ms, Maximum = 0.5 ms, Average 0.1 ms

With vAssure: Power On

Ping statistics for 192.168.1.2:
 Packets: Sent = 2000, Received = 1972, Lost = 28 (1% loss)
 Approximate round trip times in milli-seconds:
 Minimum = 0.1 ms, Maximum = 1.5 ms, Average 0.1 ms

Without vAssure: Power-On

Ping statistics for 192.168.1.2:
 Packets: Sent = 2000, Received = 396, Lost = 1604 (80% loss)
 Approximate round trip times in milli-seconds:
 Minimum = 0.1 ms, Maximum = 0.5 ms, Average 0.2 ms



Network Visibility. Optimized.

USA
 (Corporate HQ)
 + 1 650 697 8770 phone
 + 1 650 697 8779 fax
 38 Adrian Court
 Burlingame, CA 94010
 USA
www.vssmonitoring.com

Japan
 + 81 422 26-8831 phone
 + 81 422 26-8832 fax
 T's Loft 3F, 1-1-9,
 Nishikubo, Musashino,
 Tokyo, 180-0013
 Japan
www.vssmonitoring.co.jp

China
 + 86 10 6563-7771 phone
 + 86 10 6563-7775 fax
 C519, 5 Floor,
 CBD International Tower
 16 Yong'An Dong Li,
 Beijing, China 100022
www.vssmonitoring.com.cn

VSS Monitoring, Inc. is the world's leading innovator of Distributed Traffic Capture Systems™ and network taps, focused on meeting the rapidly evolving requirements of security and performance conscious network professionals. Distributed Traffic Capture Systems herald a new architecture of network monitoring, one which fundamentally improves its capability and price-performance.

VSS, Distributed Traffic Capture System, vAssure, LinkSafe, 12x4 Distributed Tap and 8x8 Distributed Tap are trademarks or registered trademarks of VSS Monitoring, Inc. in the United States and other countries. Any other trademarks contained herein are the property of their respective owners.