networktest

Virtual Chassis Performance: Juniper Networks EX Series Ethernet Switches

May 2010

Executive Summary

Juniper Networks commissioned Network Test to assess the performance of Virtual Chassis technology on its EX Series Ethernet switches, especially in the areas of latency and high availability. The key objectives were to determine whether Virtual Chassis configurations would provide low latency and fast recovery from network failures.

Among the main findings of this project:

- **Latency is always lower between Juniper EX4200 switches when those switches use a Virtual Chassis configuration rather than a standalone configuration.** Test results clearly show reduced delay, sometimes by better than a 2:1 margin, when using Virtual Chassis configurations. Virtual Chassis latency is lower in both Ethernet switching and OSPF routing configurations.

- **Virtual Chassis configurations recover from hardware and software failures in far less than 1 second.** Tests involved redundancy protocols such as rapid spanning tree and Juniper Redundant Trunk Group as well as various hardware failures. A Virtual Chassis configuration took just 6 microseconds to recover from loss of power to its master switch.

- **Virtual Chassis interconnections operate at 30-Gbit/s rates in each direction between switches for all frame sizes.** Channel capacity between Virtual Chassis member switches was never a bottleneck in this project.

Introducing Virtual Chassis Technology

Up to 10 Juniper EX4200 switches can be interconnected to form a single Virtual Chassis, which appears as a single logical entity.

Virtual Chassis technology offers multiple advantages: It eases network management by using one IP address, one Junos software image and one switch configuration for all switches within a Virtual Chassis configuration. It provides a growth path by allowing more switches to be added as needed. It can reduce latency compared with standalone switch configurations, since no round trip to the network's core or aggregation layers is needed when traffic flows between switches. Perhaps most important of all, Virtual Chassis technology helps boost high availability. Even if a switch member within a Virtual Chassis configuration fails, the Virtual Chassis will continue to operate.

While Network Test observed all these properties in action, testing focused primarily on Virtual Chassis latency, resiliency and channel capacity, as described in the following sections.

Methodology and Results

Figure 1 below shows the basic test bed topology for Virtual Chassis testing. Virtual Chassis 1 (VC1), comprised of four Juniper EX4200 Ethernet switches, served as the primary device under test, while VC2 acted as a traffic sink.

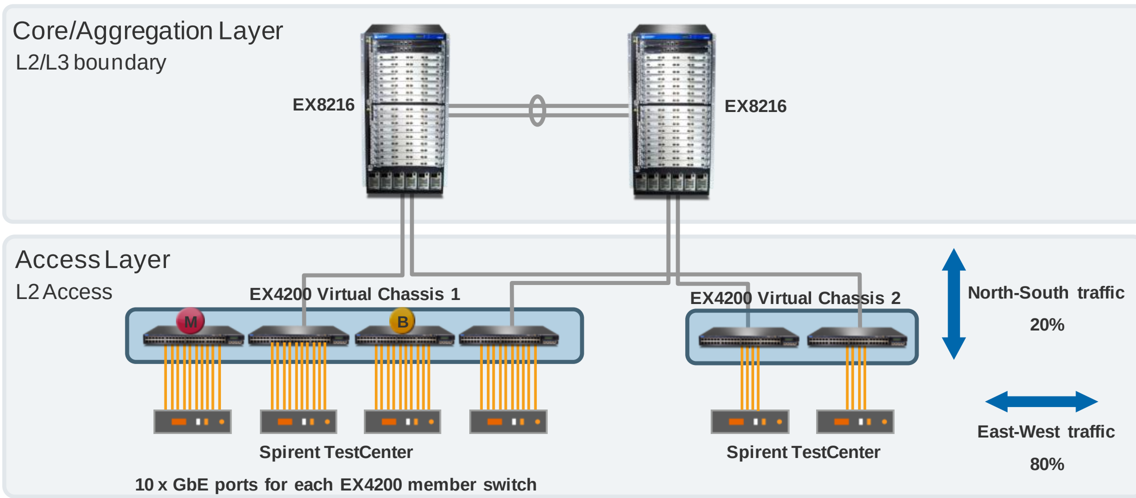

Figure 1: The Virtual Chassis Test Bed

A pair of Juniper EX8216 Ethernet switches connect VC1 and VC2 via 10-gigabit Ethernet uplinks. These core/aggregation switches use standards-based link aggregation to interconnect. The uplinks and interconnections to the core switches are important components in the high-availability tests, where engineers measured failover times for various resiliency mechanisms.

At the edge of the test bed, Spirent TestCenter traffic generator/analyzers offered 64-byte frames (the minimum length in Ethernet, and thus the highest possible frame rate) to each attached interface. Because enterprise traffic often follows an "80/20" pattern, test engineers configured Spirent TestCenter to offer 80 percent of flows local to VC1 in a fully meshed pattern, and to offer the remaining 20 percent between VC1 and VC2 across the core/aggregation switches. Unless otherwise noted, all tests involved traffic offered at line rate for a duration of 60 seconds.

Sharp-eyed readers will note that traffic loads offered to each switch are lower than the 32-Gbit/s bidirectional capacity Juniper claims for Virtual Chassis ports (VCPs) between switches. The traffic patterns used here are solely a function of the number of test ports available, and not due to any limitation of VCP channel capacity. To validate VCP channel capacity, Network Test conducted a separate set of throughput tests described at the end of this report.

Virtual Chassis Latency

A fundamental goal of this project was to compare latency between configurations with standalone Juniper EX Series switches with those involving switches in Virtual Chassis configurations. For the standalone test case, engineers set up redundant connections between six access-layer switches and the core/aggregation layer, as shown in Figure 2 below. In this baseline configuration, switches did not use the Virtual Chassis feature.

To avoid traffic loops on the redundant uplinks, engineers enabled IEEE 802.1w rapid spanning tree protocol on all switches in the standalone configuration.

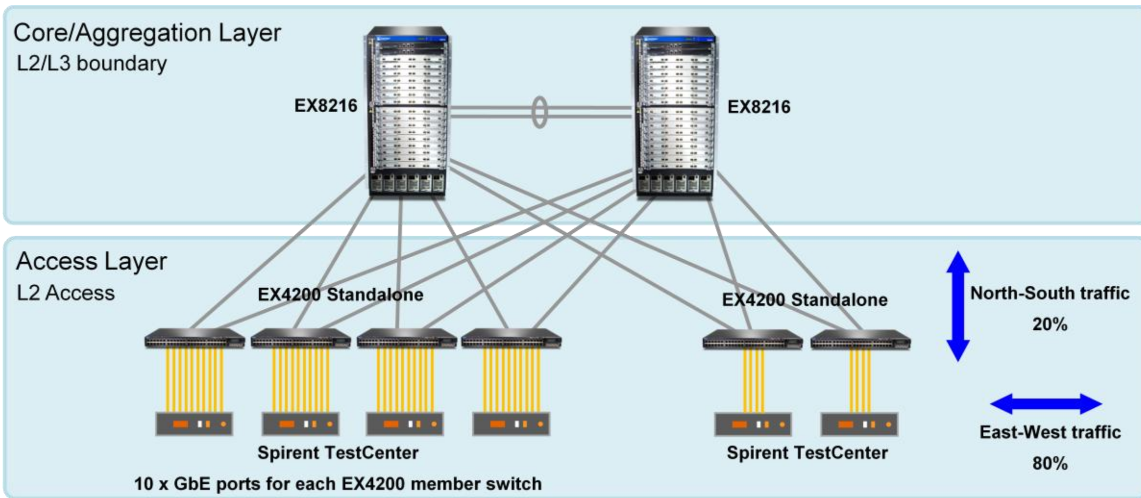

Figure 2: Standalone (non-Virtual Chassis) Test Bed

Network Test offered traffic using the "80/20" pattern described above, and measured latency on the first four switches (those in the lower-left hand corner of Figures 1 and 2). Then engineers reconfigured switches in the Virtual Chassis configuration shown in Figure 1 above. In all, engineers repeated these tests three times to compare standalone and Virtual Chassis latency:

- **Standalone** with all ports in a single VLAN. This is a pure layer-2 configuration, as shown in Figure 2 above.
- **Virtual Chassis** configuration with all ports in a single VLAN. This is a pure layer-2 configuration, as shown in Figure 1 above.
- **Virtual Chassis** configuration with all ports in different VLANs/IP subnets and with OSPF running. While the physical topology is identical to the configuration shown in Figure 1, this is a combination layer-2/layer-3 setup. Here, the boundary between switching and routing is at the access layer; in contrast, that boundary is at the core/aggregation layer in the first two configurations. This final configuration also demonstrates the ability of Virtual Chassis technology to work equally well in layer-2 and layer-3 configurations.

For each set of tests, engineers offered the same number of frames at line rate, and measured the latency of every single frame. This made it possible to make meaningful comparisons between the standalone and Virtual Chassis test cases.

Table 1 below summarizes latency test results. The numbers presented here focus on the first four switches, since they form VC1 in the Virtual Chassis test cases. Flows within VC1 do not traverse the core/aggregation layer, and thus offer the clearest picture of Virtual Chassis latency. One useful way to consider these results is to think of "same switch" numbers as conversations between hosts in a single rack, while connections between switch members in a Virtual Chassis configuration represent conversations between hosts in different racks.

	Average latency (microseconds)	Maximum latency (microseconds)
Standalone, same switch	5.44	9.58
L2 Virtual Chassis, same switch	8.19	11.08
L3 Virtual Chassis, same switch	9.24	12.36
Standalone, switch 0 <-> switch 1	20.81	24.40
L2 Virtual Chassis, switch 0 <-> switch 1	8.96	14.37
L3 Virtual Chassis, switch 0 <-> switch 1	10.22	15.72
Standalone, switch 0 <-> switch 2	22.42	24.40
L2 Virtual Chassis, switch 0 <-> switch 2	9.79	15.09
L3 Virtual Chassis, switch 0 <-> switch 2	11.16	18.44
Standalone, switch 0 <-> switch 3	20.80	24.10
L2 Virtual Chassis, switch 0 <-> switch 3	8.67	14.32
L3 Virtual Chassis, switch 0 <-> switch 3	10.20	15.15

Table 1: Standalone vs. Virtual Chassis Latency

The most notable result is that **latency between switches is always significantly lower in Virtual Chassis configurations, both for average and maximum delay.** The reason for the latency improvement with a Virtual Chassis configuration is simple: Frames do not need to traverse the core/aggregation switches, and thus save the delay added by one or more extra hops through the network.

For traffic within the same switch, latency is slightly higher in the Virtual Chassis test cases. Although both the standalone and Virtual Chassis test cases involve the same fully meshed traffic pattern among all switch ports, the workload is slightly heavier in the Virtual Chassis test case. In that instance, traffic between all ports on other Virtual Chassis member switches traversed each switch in the Virtual Chassis configuration, imposing a slightly heavier load. In the standalone case, there is no such background load: Each switch offloads all non-local traffic to the core/aggregation layer for forwarding, and thus only handles switching of local traffic. However, for any configuration involving multiple switches, average and maximum latency are lower – often by a factor of 2 – when using a Virtual Chassis configuration rather than a standalone configuration.

Virtual Chassis Resiliency

High availability is a greater concern than high performance for most network architects and network managers. After all, the speediest switch in the world is of little help if it's unavailable because of a link or interface failure.

Juniper EX Series switches support multiple redundancy mechanisms to ensure maximum uptime. A key goal of this project was to determine what effect, if any, a Virtual Chassis configuration will have on these failsafe mechanisms.

Network Test measured Virtual Chassis failover times for four test cases in layer-2 mode and three test cases in layer-3 mode. For all tests, Network Test determined recovery time by offering frames at 1,000,000 frames per second (thus, one frame every microsecond) and deriving failover time from frame loss statistics reported by Spirent TestCenter.

The layer-2 test cases included link failures with IEEE 802.1w rapid spanning tree protocol (RSTP) and Juniper's redundant trunk group (RTG) protocol. Other test cases covered the loss of a VCP and Virtual Chassis Routing Engine (RE), which serves as the master switch in a Virtual Chassis configuration.

The layer-3 test cases also involved the loss of a VCP and an RE. In addition, Network Test measured OSPF convergence time after a link failure. As in the layer-3 latency tests, this test involved a configuration in which all switches ran OSPF and each access switch port used a different VLAN and IP subnet.

Figure 3 below shows the traffic patterns involved in the RSTP, RTG, and OSPF failover tests. In these tests, engineers disabled the link between VC1 and the first EX8216 core/aggregation switch, forcing traffic to be rerouted via the second EX8216 switch.

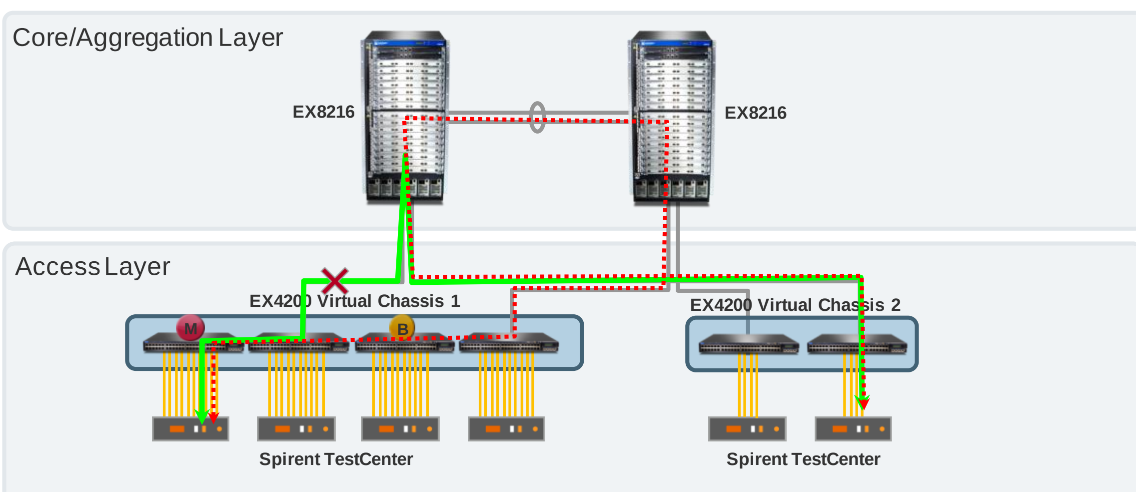

Figure 3: The High-Availability Test Bed

Figure 4 below shows the configuration used in the VCP failover tests. Here, engineers disconnected one of the Virtual Chassis cables linking switches within a Virtual Chassis configuration, forcing the system to reroute traffic across a different Virtual Chassis connection path.

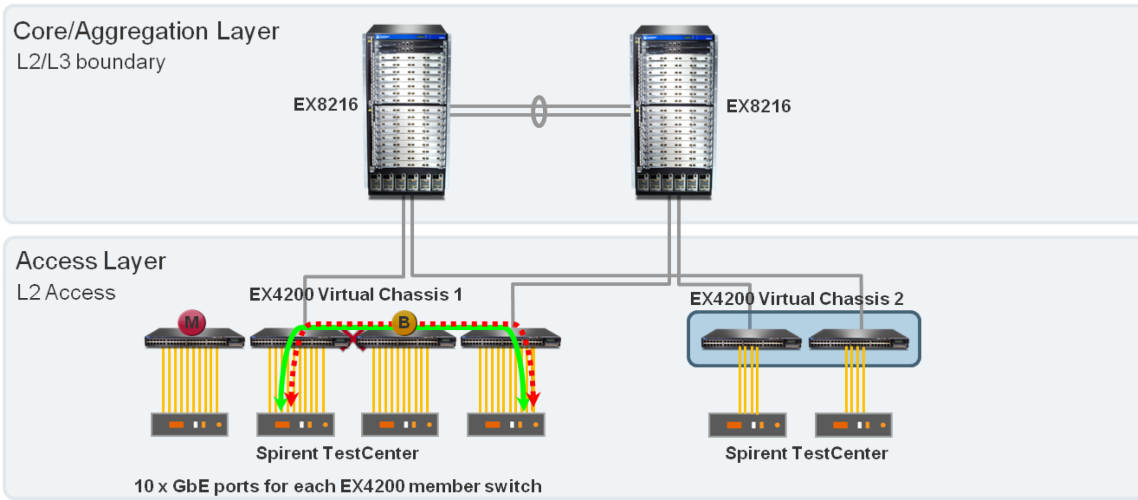

Figure 4: The VCP Failover Test Bed

Figure 5 below shows the topology used for the RE failure test. Here, engineers powered off the master switch in the Virtual Chassis configuration (labeled "M" in the figure), forcing the backup switch (labeled "B") to take over. Virtual Chassis member switches also used RTG in both the VCP and RE failover tests. As in the protocol failure tests, engineers derived failover time from frame loss measured during the cutover period.

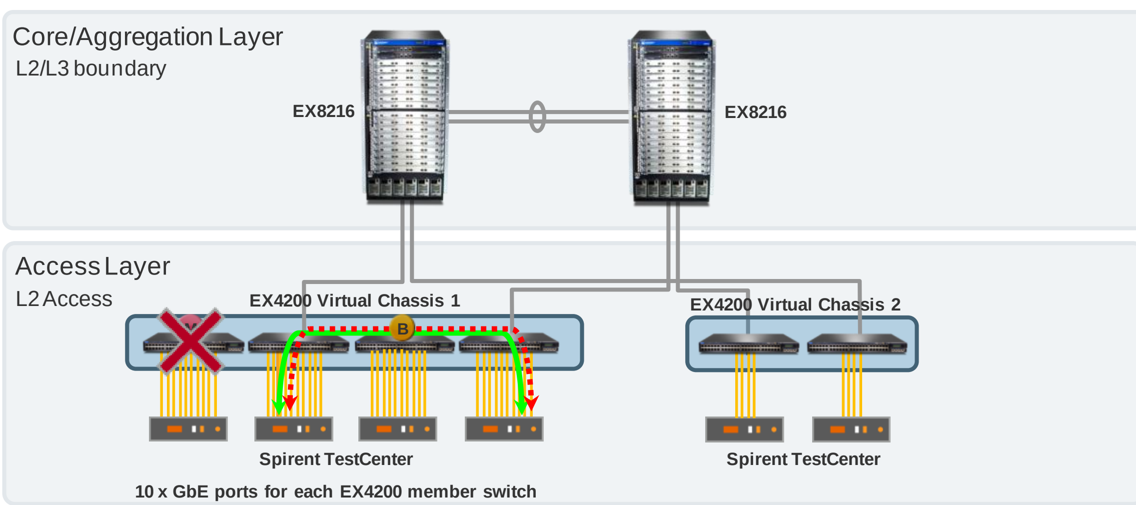

Figure 5: The RE Failover Test Bed

Table 2 below summarizes failover times for the various resiliency test cases.

Test case	Failover time (milliseconds)
Layer 2, rapid spanning tree (RSTP) failover	98.520
Layer 2, redundant trunk group (RTG) failover	96.770
Layer 2, Virtual Chassis port (VCP) failure	0.059
Layer 3, VCP failure	0.116
Layer 2, Routing Engine (RE) failure with RTG	0.006
Layer 3, RE failure	0.213
Layer 3, OSPF failover	285.451

Table 2: Virtual Chassis Failover Times

In all test cases, the Virtual Chassis configuration recovered from failures in far less than 1 second. Even in the case of the longest failover time, for OSPF, it took less than 300 milliseconds for the Virtual Chassis configuration to recover from a link failure and forward traffic along a recomputed path.

Recovery times were lower still for other protocols. The Virtual Chassis configuration converged following RSTP and RTG failures in less than 100 milliseconds.

The fastest recoveries were those following hardware failures. In the fastest failover of all, it took just 6 microseconds to recover from the loss of power to a master switch in the Virtual Chassis configuration in the layer-2 test case. Layer-3 recovery times were somewhat higher due to the greater amount of route processing involved with OSPF, but again were well below the 1-second mark.

Virtual Chassis Channel Capacity Validation

As noted, the number of available test ports – and not any shortage of bandwidth on Virtual Chassis links between switch ports – limited the amount of traffic offered in these tests. To validate that Virtual Chassis links operate at or near Juniper's stated limit of 32 Gbit/s in each direction, Network Test conducted an additional set of tests to determine the throughput rate for Virtual Chassis connections.

As defined in RFC 1242, throughput is the maximum rate at which a system can forward traffic with zero frame loss. Different frame sizes may have different throughput rates due to buffering and other switching effects. To factor for this, Network Test measured the throughput rate for frame sizes ranging between 64 and 1,518 bytes, which are respectively the minimum and maximum sizes allowed by the IEEE 802.3 Ethernet specification. As recommended by RFC 2544, these frame lengths are 64, 128, 256, 512, 1,024, 1,280, and 1,518 bytes.

Figure 6 below illustrates the test bed topology used for Virtual Chassis channel capacity validation. This test bed design achieves two aims: It not only determines the throughput rate for Virtual Chassis connections, but also verifies that this capacity is available when switches operate in a ring topology.

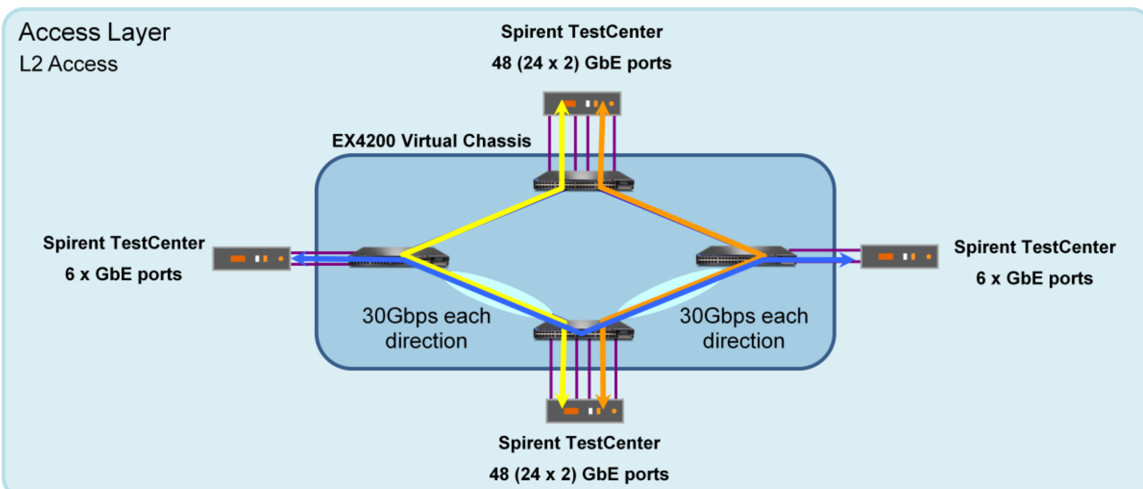

Figure 6: The Virtual Chassis Channel Capacity Validation Test Bed

As shown in Figure 6, this configuration involved a total of 108 gigabit Ethernet test ports. The Spirent TestCenter generators offered a bidirectional aggregate load of up to 30 Gbit/s running between the lower two Virtual Chassis connections. This rate, 30-Gbit/s in each direction on each of two VCPs on the bottom switch in Figure 6, represents a practical upper bound for user data. Although the physical channel capacity is higher (around 32 Gbit/s per VCP in each direction), some bandwidth is used by frame headers added for internal switching.

Table 3 below summarizes throughput rates for the various frame rates tested.

Frame length (bytes)	Theoretical maximum for 30 ports (fps)	Measured throughput on 30 ports (fps)	Percentage of line rate
64	44,642,857	44,642,857	100.00%
128	25,337,838	25,337,838	100.00%
256	13,586,956	13,586,956	100.00%
512	7,048,872	7,048,872	100.00%
1,024	3,591,954	3,591,954	100.00%
1,280	2,884,615	2,884,615	100.00%
1,518	2,438,231	2,438,231	100.00%

Table 3: Virtual Chassis Channel Capacity Throughput

The Virtual Chassis configuration forwarded all frame sizes at line rate, validating that Virtual Chassis connections were not a bottleneck in latency or resiliency testing.

Conclusion

These tests demonstrated several ways in which Virtual Chassis technology can enhance network performance. A Virtual Chassis configuration clearly reduces latency compared with a standalone configuration. Moreover, Virtual Chassis configurations provide sub-second recovery from link and equipment failures (and in some cases, recovery is well into the sub-millisecond range). Throughput testing also demonstrated the ability of Virtual Chassis links to carry user traffic at 30-Gbit/s rates with zero frame loss, regardless of frame length.

Appendix A: Software Versions Tested

This appendix lists the software versions used on Juniper and Spirent equipment on the test bed.

Juniper EX4200: Junos 10.1R2
Juniper EX8216: Junos 10.1R2
Spirent TestCenter: 3.30.5316 (application); 3.30.4496 (firmware)

Appendix B: Disclaimer

Version 2010050400. Network Test Inc. has made every attempt to ensure that all test procedures were conducted with the utmost precision and accuracy, but acknowledges that errors do occur. Network Test Inc. shall not be held liable for damages which may result for the use of information contained in this document.

Network Test Inc.
31324 Via Colinas, Suite 113
Westlake Village, CA 91362-6761
USA
+1-818-889-0011
http://networktest.com