there's a smart way to run your website
up to 150 times faster in SSL mode.

Infrastructure **FEATURE**

Gigabit Over Copper: Bandwidth To Burn?

September 18, 2000
By Joel Conover

They said Gigabit Ethernet over standard copper wiring could not be done. But on June 23, 1999, IEEE members passed the 802.3ab Gigabit Ethernet Over Copper standard. Now, more than a year later, the industry is delivering its first gigabit-over copper-solutions to the mass market.

Gigabit over copper holds two strengths over its fiber brethren: lower cost and the ability to run over existing copper wiring. But the technology's reliability may not be enough for its target applications in data centers, ISP collocation sites and high-performance workstations and servers, where fault-tolerance is absolutely critical.

At Schneider National, one of our partner labs, in Green Bay, Wis., we tested the first wave of products. Although they deliver wire-speed throughput in the lab, we found gigabit over copper may deliver only four-nine reliability when pushed to its limits. The small chance that the physical layer could be to blame may be reason enough to cost-justify fiber.

What Do Readers Think?

Check out our e-poll on Gigabit Over Copper!

The Gigabit Ethernet market is also surfing through an incredible commoditization phase. While the price rift between 1000BASE-SX optical and 1000BASE-TX gigabit over copper is slowly closing, 1000BASE-SX fiber ports still cost nearly 50 percent more per port than their copper equivalents. Network managers must decide if a 50 percent cost savings is worth a few thousandths of a percent risk.

The Manifest

Network Computing built a test plan and asked vendors to submit a complete end-to-end solution for a high-density ISP collocation site using gigabit-over-copper technology. We dubbed our test the ISP Rack Collocation Challenge. Each of the vendors had to deliver an end-to-end solution suitable for deployment in one 84-inch tall, 19-inch wide ISP collocation rack, using gigabit-over-copper technology and focusing on fault-tolerance, server load-balancing and space optimization.

Why the end-to-end approach? Because the vendors pitch such solutions to their customers and had begged us to test them this way. We invited 15 vendors to stop their huffing and to deliver on their bluster, yet only four vendors stepped up. We were not concerned with an apples-to-apples comparison; rather, we simply required each vendor to present its best gigabit-over-copper solution, keeping in mind collocation and server-room consolidation as primary targets for its technology.

In our test plan, we told vendors we'd be judging the solutions' physical fault-tolerance, fault management and availability. Furthermore, we asked each vendor to show us three features that significantly differentiate its solution from the competition.

We solicited participation from Alcatel, Alteon WebSystems, Asante Technologies, Cisco Systems, Compaq Computer Corp., Enterasys Networks, Extreme Networks, Foundry Networks, Hewlett-Packard Co., Intel Corp., Lucent Technologies, Nortel Networks, SysKonnect, Sun Microsystems and 3Com Corp. Alcatel, Compaq and HP claimed they had no products that fit the nature of our tests (gigabit over copper and server load-balancing). Lucent and Sun, though they had products, chose not to participate. Enterasys and 3Com sent no response.

Although Alteon, Asante and Nortel initially agreed to participate, we never got to test their products. Alteon informed us it felt a competitor's not-yet-shipping product was not a fair comparison with Alteon's already-shipping product. (It has always been Network Computing's policy to insist that products we test be shipping by the date of publication.) Nortel said its Passport 8600 gigabit-over-copper blades were not yet available for tests. Finally, we determined that the Asante FriendlyNet GX4 was not an appropriate contender among these large, enterprise-class products, and we removed it from our roundup.

Ultimately, we tested four complete solutions and two NICs. Representatives from Cisco, Extreme, Foundry and Intel trekked to our Wisconsin lab with a full rack of collocation gear. Products included server load-balancing hardware, firewalls, SSL accelerators, core switches and content-switching solutions--all gigabit-over-copper attached. SysKonnect and Intel sent NICs, which we tested separately (see "Gigabit-Over-Copper NICs.")

Analyzing the Report Card

Our report card is broken down into five areas: price, management, fault-tolerance, density metric and advanced features. The vendor's price score is based on the cost of the solution in the rack configuration. The price metric includes all ancillary equipment. The management score is a measure of the overall ability to manage faults, performance monitoring and troubleshooting of the solution. Fault-tolerance is a function of the total solution. We asked each vendor to demonstrate that there was no single point of failure in the network and to show us the levels of fault management available in the solution. We weighted this metric equally with the level of connectivity and server density each vendor provided in the rack.

We also asked each vendor to show off how it manages its solution in terms of statistics, device management and overall server-farm health monitoring. These results, along with our analysis of the unique features, are reflected in the advanced-features score.

Final Analysis

In addition to evaluating each vendor's solution based on technical merit, feature set and fault-tolerance options, we contrived several metrics to quantify how fast and dense the vendor's overall solution was. Our primary metrics were bandwidth per inch (the higher the better, as it indicates more potential throughput), cost per inch (lower is better, providing quicker return on investment) and load per server (the lower the better, delivering faster response times) assuming an OC-12 connection to the upstream Internet provider. Based on the size and complexity of the products involved, typical server load never exceeded 20 Mbps, a load that almost any modern server can handle with ease. Of course, if larger (than 1U) servers are used, the load per server goes up accordingly. Several vendors also pointed out that a typical large ISP collocation effort would involve multiple sites connected at OC-3 speeds or below, to enhance fault-tolerance and response time on the global Internet. If that is true, the typical server will barely push more than 100 Mbps. Gigabit over copper leaves you with plenty of bandwidth to burn.

We chose Cisco's ArrowPoint solution as our Editor's Choice. Its content-aware networking gave it a specific and valuable advantage over the competition. Foundry and Intel tied for a close second. The small edge Cisco held over its competitors was largely due to its extremely useful reporting features and incredible content-aware switching. Extreme came in last, showing the green nature of its first-generation server load-balancing solution.

Reviews

Cisco's Solution Fits Rack to a 'T'

Yes, it's expensive, but Cisco's Catalyst-ArrowPoint-PIX combo delivers all the fault-tolerance, flexibility and features an ISP could ask for.
By By Joel Conover

Cisco Systems

A-

Cisco's answer to our ISP Rack Collocation Challenge features hardware from its recent ArrowPoint Communications acquisition. With its unique ability to replicate content based on popularity and its superior fault-tolerance, Cisco's solution represents the best use of gigabit-over-copper technology in a collocation setting. Although expensive, the Cisco solution offers incredible flexibility and unmatched features--its bang makes up for the bucks.

The core of the Cisco solution consists of a Catalyst 6506 with dual Supervisor Engine 1A management modules and 32 ports of gigabit-over-copper connectivity. Eight ports of fiber gigabit connectivity provide additional links to the other equipment. Cisco also brought along a pair of Cisco ArrowPoint CSS-11154 Content Smart Switches and a pair of fault-tolerant PIX 515 Gigabit Ethernet-enabled firewalls. The total solution is 28.85 inches tall, leaving room for 31 servers on our rack. Cisco's core bandwidth capacity is 32 Gbps, or 381 Mbps for every inch of rack space. The solutions from Foundry and Extreme deliver four times the bandwidth per inch at the core. While higher is better, we felt that because the solution was delivering only 20 Mbps of load to each server, the long-term scalability was well within reason. The Cisco solution provides a load of 20.06 Mbps to each of the 31 servers in the rack, assuming an OC-12 connection to the upstream provider. Rack-space cost for the entire solution is $2,488 per inch. One slot is available on the Catalyst 6506 for future expansion, to increase port density; as an option, Cisco offers an intrusion-detection blade that can be outfitted as part of the solution.

Cisco's solution addresses the need for fault-tolerance elegantly. At the lowest layer, the core Catalyst 6506 performed admirably in our fault-tolerance tests. Cisco demonstrated its nonstop-code-upgrade availability, as well as extensive fault-tolerance through dual redundant Supervisor modules. Two content-switch load-balancers provide a second level of fault-tolerance, and a pair of fault-tolerant PIX firewalls with firewall load-balancing provide a third level.

The ArrowPoint demand-based replication option is Cisco's ace in the hole when it comes to differentiating features. Demand-based replication is unique to the ArrowPoint architecture. Essentially, the ArrowPoint load-balancer CSS-11154 Content Smart Switch tracks "hot" URLs based on a list of the 10 most heavily visited sites. When a predefined set of criteria (based on URL hits, bandwidth and/or number of hits over time) is met, the hardest-hit content can be mirrored, via FTP, to additional backup servers to provide extra capacity. As "hot flashes" on your collocated content change, the switch can move source material dynamically to accommodate the additional load. All this is done transparently, without user intervention. Demand-based replication also is aware of changes to the source material after replication; thus, if content is changed on the origin server, those changes will be propagated to the backup servers, too. This advanced feature makes it possible to collocate a large number of sites with different traffic patterns in a relatively small space. We tested Cisco's demand-based replication by setting up several Web servers in the lab and generating sufficient load to trigger the demand-based replication. It worked simply, quickly and elegantly.

Cisco Systems

Catalyst 6506 chassis; two Catalyst 6000 Supervisor Engine 1A modules; Catalyst 6000 16-port 1000TX Gigabit Ethernet module; Catalyst 6000 eight-port Gigabit Ethernet module; two CSS-11154 switches; two PIX 515 Firewalls with two Gigabit Ethernet interfaces each; space for 31 servers.

Total cost: $208,992

Cisco also touted its NAT (Network Address Translation) peering function, which it claims allows the CSS-11154 Web load-balancer to redirect content to the server intelligently with the lowest overall round-trip latency (when compared with traditional translational NAT implementations). This feature transparently redirects a URL request to the fastest server in a geographically distributed collocation deployment, in which you deploy multiple collocation racks such as the one we tested.

Finally, Cisco's firewall load-balancing feature lets it gain maximum performance from a firewalled environment by distributing load across multiple active PIX firewalls. The trick here is to track and return packets back through the firewall from which they came. The overall benefit is increased performance without having to worry about the firewall being the bottleneck.

By the time we finished testing Cisco's network, we were left with one solid impression: The Cisco solution is uniquely content-aware. The ArrowPoint content switch can do much more than just balance traffic based on a particular URL. It is suited to track the content of your Web site down to the URL level, and provide billing, statistics, troubleshooting and other vital information based on the content visited. The other vendors left us with a sense of content-enabled switching, but only Cisco convinced us that its product is truly content-aware.

Catalyst 6506 chassis, $11,990; Catalyst 6000 Supervisor Engine 1A, $29,995; Catalyst 6000 16-port 1000TX

Gigabit Ethernet module, $15,995; Catalyst 6000 eight-port Gigabit Ethernet module, $9,995; PIX 515 Firewall with four NICs, $12,400; CSS-11154 12 pt FE (TX), 2 pt GE (SX), 128-MB, HD, AC, $24,995; Cisco Systems, (800) 553-6387, (408) 526-4000; fax (408) 526-4100, www.cisco.com.

Foundry Networks

B+

Foundry stepped up to our challenge with its BigIron 8000 and ServerIronXL platforms. Armed to the teeth for fault-tolerance, performance and reliability, the Foundry solution was a first-class experience.

Foundry brought us a BigIron 8000 chassis that was equipped with four eight-port gigabit-over-copper modules, dual Management IV modules (with four fiber gigabit ports each), and two ServerIronXL eight-port 10/100 switches with dual gigabit uplinks. The entire solution occupies 26.25 inches of rack space, leaving room for up to 33 servers. Foundry specified 32 servers in its design (leaving 1U of open space in the design). The solution's total available bandwidth is 128 Gbps, or 1.52 Gbps per inch. That's an incredibly large amount of bandwidth, offering almost infinite scalability for this application. Foundry balances the big iron of a chassis switch with the density of a stackable, delivering 19.44 Mbps of load to each server in our 622-Mbps-attached environment. The overall cost per inch is $1,724 per inch of rack space, $764 per inch less than Cisco's offering. Foundry has room for up to two more eight-port modules in its solution.

Foundry matches Cisco nearly feature for feature in terms of fault-tolerance. With only three components to the system, there isn't much that could go wrong. However, we penalized Foundry in our report card's management category because the vendor didn't offer a nonstop-code-upgrade feature, and the solution required several reboots when different features were enabled throughout the course of our tests. The ServerIronXL has two modes of operation: active-active and active-standby. In active-active configurations, you can gain extra performance from the redundant server load-balancing switch. Foundry's big-ticket item in fault-tolerance is stateful failover. If a switch fails during operation, existing sessions are not disrupted but are instead gracefully failed over to the backup unit. This ensures that shopping-cart applications and the like are not affected by the loss of a mission-critical network component in the active-active configuration. In active-standby, only one ServerIronXL is active; the other waits silently until a failure occurs. Stateful failover was unique to the Foundry solution.

At the core, the BigIron 8000 matches Cisco's configuration with dual management modules. While Foundry offers stateful failover at the load-balancing level, it does not keep state at the core level. Thus, if one management module fails, there is a slight interruption in service (less than 30 seconds). Likewise, code upgrades require a system reboot. However, the Foundry switch's reboot time was almost four times as fast as Cisco's reload time. We were slightly frustrated to find that some command-line options--such as enabling ACLs (access control lists) for the first time--required a switch reboot. This may seem trivial, but if you choose not to deploy ACLs at installation time and later decide to enable them--for example, in response to a DOS (denial of service) attack--you would incur a switch reboot.

Foundry Networks

BigIron 8000 chassis; two Management IV modules with four 1000BASE-SX GBICs each; four eight-port 1000BASE-TX gigabit-over-copper modules; one 24-port 10/100BASE-TX module; ServerIronXL; two-port 1000BASE-SX uplink module; space for 33 servers.

Total cost: $144,816

Foundry's hot-spots are its URL switching, slow-start recovery mechanism, switchback technology and firewall load-balancing. Similar to Cisco's CSS-11154 switch, Foundry's ServerIronXL is URL-aware. We tested this feature and found that it works flawlessly, load-balancing to servers on a round-robin URL-aware basis. However, the statistics, tracking, and overall flexibility of Foundry's URL-based load-balancing weren't as dynamic as those of the CSS-11154.

Foundry's other strengths come from its architecture. Like Cisco's CSS-11154, the ServerIronXL can track and deliver traffic through the proper source firewall, if one exists. Foundry's slow-start recovery algorithm is also unique and quite valuable. Typically, when a server is added to a server load-balancing pool or rebooted, its hit counters will be much lower than other servers in the pool. As a result, that server may become inundated with traffic because of the disparity in counters. To counteract and prevent this, Foundry provides slow-start pacing, where the newly added server is ramped up to full production over a series of smaller traffic load increments. This unique feature is good for pools with constantly changing servers or servers with frequent downtimes.

Foundry's switchback technology--the third feature the vendor chose to highlight--impressed us, but every vendor offers some variety with similar performance and features. In Foundry's design, the Web load-balancers operate similarly to one-arm routers, hanging off to the side of the primary data path. Switchback technology allows packets being returned to the client to bypass the SLB on the way back.

Our impression of the Foundry solution is that it's fast. Foundry provides ample and superior bandwidth throughout its architecture, from dual gigabit links to its server load-balancers, down to an incredibly fast and dense core. The solution's URL load-balancing is highly advanced, and its features compete with those of Cisco's CSS-11154; however, accounting and management--particularly the breakout of utilization, hit counts and load-balancing stats--are not as robust.

ServerIronXL eight-port Layer 4-7 switch, $7,995; two-port 1000BASE-SX uplink modules, $3,695; BigIron 8000 chassis, $9,995; Management IV module, $25,995; 1000BASE-SX GBICs, $495; eight-port 1000BASE-TX modules, $12,995; Foundry Networks, (888) TURBOLAN, (408) 586-1700; fax (408) 586-1900, www.foundrynetworks.com.

Intel Corp.

B+

Intel delivered a unique mix of products that included four Intel NetStructure 480T routing switches and two 7180 e-Commerce Directors. The 7180 e-Commerce Director is a powerhouse, but the 480T doesn't provide enough expandability to clinch a win for Intel.

Each component of Intel's solution is 2U tall, for a total of 12U, or 21 inches, leaving room for up to 36 servers, making it the highest capacity solution we tested. Each 480T switch provides 12 ports of gigabit-over-copper connectivity and four GBIC (gigabit interface converter) ports for interswitch connectivity and auxiliary device support. The 7180 e-Commerce Directors connect via 10/100 connections to the solution. Intel's total density is 36 servers, with some being dual homed to two switches and some being single homed. The solution's total switch bandwidth is 32 Gbps, offering the same 381 Mbps per inch as Cisco, more than sufficient for the load per server it delivered. With 36 servers in the rack, Intel's total load per server at OC-12 speed is the best, at 17.3 Mbps per server. Intel's solution comes in at $1,951 per inch of rack space, slightly more than Foundry's solution.

Intel uses a combination of dual-homed servers, OSPF equal-cost load-balancing and Layer 2 trunking to ensure end-to-end fault-tolerance. Up to 31 of the 36 available servers can be dual homed for maximum fault-tolerance. At the core of the network, the 480T routing switches provide wire-speed switching at Layers 2, 3 and 4. However, Intel uses the 480T primarily as the interconnect for servers and its 7180 e-Commerce Director, which takes on the brunt of the load-balancing work.

The 7180 e-Commerce Director, the shining star of Intel's solution, shows off Intel's insight into issues that plague a typical SLB (server load-balancing) installation. At the top of its list of unique features are Intel's SSL (Secure Sockets Layer) termination and acceleration capabilities. The 7180 e-Commerce Director can terminate SSL connections at the box and then pass secure transactions in clear text on to the hosting servers.

Intel Corp.

Four Intel NetStructure 480T routing switches; two Intel NetStructure 7180 e-Commerce Directors; space for 36 servers.

Total cost: $163,884

In the lab, we tested the device with and without this feature. With the feature turned off, our server hosted about 100 connections per second and ran at 100 percent CPU utilization. When we tried to push the server beyond 100 connections per second, latency rose from milliseconds to several seconds per connection. In the real world, such poor response time is bound to lose customers. But when we turned on the 7180 e-Commerce Director's SSL termination capabilities, CPU utilization dropped from 100 percent to 5 percent to 10 percent, and latency fell to submillisecond values. The server farm could handle up to 600 connections per second without any impact on performance. No other vendor offered SSL acceleration.

The 7180 e-Commerce Director also has extensive server-monitoring and fault-management options. At the top of this list is content-aware switching. Like Foundry and Cisco, Intel can switch based on URL information; however, the 7180 takes the process one step further. If one or more servers are down, or content is missing from those servers, the 7180 intelligently deals with the resulting 400/500/600-series error codes. Rather than terminating the SSL session on a failure, the load-balancer redirects the traffic to the next server until it finds one with the appropriate content--never showing an error message to the client, unless no server has the appropriate content. Only Intel has this extremely useful feature, which prevents users from getting 404 errors because of incorrectly replicated content. The switch also can deliver customized error messages based on 400/500/600-level error codes, letting you customize the messages a user might otherwise receive because of server overload.

As a final feature, Intel demonstrated its intelligent, load-based SLB functionality. Rather than use round-robin traffic delivery, the 7180 monitors response time and server load, and uses this information to distribute HTTP and SSL requests to the least-loaded server. In an environment with dissimilar servers, this feature ensures that larger servers receive their fair share of traffic without requiring the administrator to manually determine weights for reach resource in the SLB pool. When combined with content-based switching, the feature also lets the 7180 give certain users priority access to specific Web farms, based on availability and source IP address.

Because the 7180 e-Commerce Director can be dropped into an operational network, adding intelligent SLB and SSL acceleration with this product can be done without any interruption to the network service. Once you've configured the 7180, simply update your DNS entries, and in 24 to 48 hours, the device will start picking up and balancing server traffic, all without network downtime.

The Intel solution is highly e-commerce-enabled. Its unique features for handling server-content mismatch and SSL termination make it an ideal solution for any collocation environment. Overall, the solution made a strong showing, but its 16-port core configuration (the NetStructure 480T switch) didn't leave much room for expansion. In an ideal world, the 7180 e-Commerce Director could fit into any vendor's load-balancing solution, enhancing SSL performance and delivering a best-of-breed SLB solution.

Intel NetStructure 480T routing switch, $12,295; Intel NetStructure 7180 e-Commerce Director, $39,995; Intel Corp., (800) 628-8686, (408) 765-8080; fax (408) 765-9904, www.intel.com.

Extreme Networks

B+

Extreme Networks' solution stands apart in its strong management and use of partner products to fill the bill. Extreme's solution is a good choice for basic load-balancing at a great price, but it lacks a number of features important to advanced e-commerce sites, features that the other vendors' solutions demonstrated. Extreme includes components from several of its partner vendors, including F5 Networks' 3-DNS global load-balancer, a pair of NetScreen-100 firewalls and a pair of Extreme Summit 7i switches. Also part of the solution is a Sun Microsystems Netra t1 Model 105 server, which is responsible for monitoring server health using Extreme's ServiceWatch platform.

The Extreme solution measures in at just 13U, leaving room for up to 35 servers. Extreme also specified a partially fault-tolerant server farm with dual-homed servers for 18 of the servers and an additional 17 single-attached servers. The total solution provides 1.31 Gbps per inch of bandwidth, and a fully loaded server farm would receive about 17.77 Mbps per server on average. Extreme's solution prices in at $1,369 per rack inch, the lowest cost of the solutions provided.

Extreme's solution focuses on site monitoring, high availability and performance. Compared with the other products, however, the server load-balancing features of Extreme's solution are a bit basic. The standard SLB features are there, but the Extreme Summit 7i doesn't do cookie persistence or URL-based load-balancing. Like the products from Foundry and Cisco, Extreme's supports a direct return path technology the vendor calls transparent server load-balancing. With transparent SLB, the sessions are distributed to the server farm, and the traffic is returned directly to the querying host without passing through the server load-balancing code path again. This significantly improves performance over Extreme's translation mode, in which the packets are fully translated in both directions.

Extreme's top selling point is device consolidation. Rather than having a separate box for SLB, which introduces artificial bottlenecks and additional points of failure, all the server load-balancing is done inside the Summit 7i switch.

Unlike our first look at the ExtremeWare 6.0 code (which we first tested on the Alpine platform; see "Extreme Networks Alpine 3808: Broadband Access for MANs," this time we were able to see the full F5-based load-balancing scheme at work. Extreme has licensed and ported F5's full load-balancing code to its i-Series switches, including the Alpine and the Summit 7i. The Summit 7i has support for health checks--from simple ICMP (Internet Control Message Protocol) pings to more complex service-response checking, multiple load-balancing mechanisms and SSL session persistence. Although the Summit 7i cannot track cookies, it can properly load-balance SSL traffic based on the SSL

Extreme Networks

Two Summit 7i 100/1000BASE-T switches; two NetScreen Technologies NetScreen-100 firewalls; Sun Microsystems Netra t1 Model 105 running Extreme Networks ServiceWatch; F5 Networks 3-DNS controller; space for 35 servers.

Total cost: $114,996

ID bytes of the packet. When persistent SSL sessions based on cookies are required, Extreme partners with F5, using F5's Big-IP solution to augment Extreme's own products. In this solution, Extreme uses F5's 3-DNS global load-balancer to provide architectural scalability and performance. If additional servers were to be added, 3-DNS would provide a method for distributing load across multiple server farms.

The biggest strength of Extreme's solution is its service-management offering. Extreme's solution includes a useful hardware service monitoring platform, called ServiceWatch. ServiceWatch runs on the Sun Netra t1 105 server Extreme included as part of its solution. ServiceWatch provides out-of-band management for server farms, tracking response time and availability. It can be configured to keep track of individual server processes, such as the Web server, mail server and other mission-critical applications, running on the server farm. Over time, ServiceWatch enables you to better plan capacity, thanks to its trending and analysis features. Extreme is the only vendor to offer a management component for capacity planning and server monitoring. Competing products can do health checks but can't approach the level of trending and analysis ServiceWatch provides.

Overall, Extreme's solution is good but doesn't have as many bells and whistles as some of the competition. Extreme doesn't skimp on performance or fault-tolerance; this solution is great as a bread-and-butter load-balancing solution. But the lack of certain critical features like cookie persistence makes it unsuitable for some applications. Although Extreme can add this via third-party hardware, it is not an integrated part of the vendor's solution. For basic load-balancing at a damn good price, this solution is hard to beat.

Summit 7i 100/1000BASE-T, $19,995; Summit 7i 1000BASE-SX, $28,495; Extreme Networks, (888) 257-3000, (408) 579-2800; fax (408) 579-3000, www.extremenetworks.com, or info@extremenetworks.com.

Testing configuration information: 3-DNS, $34,990, F5 Networks, (888) 882-4447, (206) 505-0800. www.F5.com; NetScreen-100, $9,995, NetScreen Technologies, (408) 330-7800; fax (408) 330-7850, www.netscreen.com, or info@netscreen.com.

Netra t1 Model 105 running Extreme Networks ServiceWatch, ($2,995 for ServiceWatch) Sun Microsystems, (800) 555-9786, (650) 960-1300, www.sun.com.

Executive Summary

Gigabit-Over-Copper Solutions

Gigabit Ethernet over copper is finally appearing en masse in the market. This technology typically offers up to 50 percent cost savings over traditional fiber-based networks, and can make a low-cost interconnect medium for data centers, collocation facilities and high-performance clusters. But if 99.999 percent reliability is your priority, you'll want to think twice. Our tests showed gigabit over copper can be finicky--particularly sensitive to cabling-plant issues and external noise.

We invited vendors to display their gigabit-over-copper solutions as they would recommend them for a typical collocation environment. Cisco Systems, Extreme Networks, Foundry Networks and Intel showed up in our labs, touting complete collocation solutions. Cisco took top honors, winning our Editor's Choice award for its expensive but enticing solution based on its Catalyst 6500 series and ArrowPoint Communications content-switching products.

Because gigabit-over-copper solutions also require specialized NICs, we invited vendors to submit their copper-based cards for testing. Intel and SysKonnect chose to participate in these tests, with SysKonnect showing a production card that ran rings around Intel's beta effort.

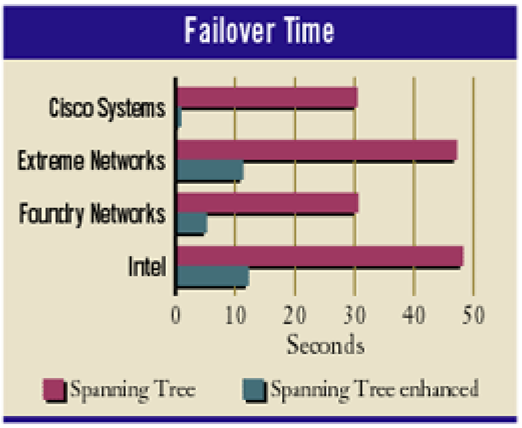
By the Numbers

How We Tested Gigabit Over Copper

Our ISP Rack Collocation Challenge called for end-to-end solutions; we stress-tested the contestants' fault-tolerance and server load-balancing capabilities. By Joel Conover

In our first ISP Rack Collocation Challenge, we invited each of the vendors to submit a complete, end-to-end gigabit-over-copper solution, designed to fit in one 84-inch tall, 19-inch wide ISP collocation rack.

Our test plan had two sections. Physical fault-tolerance, fault management and availability testing made up the first part. We tested standard and enhanced Spanning Tree failover, switch reboot time, power-supply redundancy, SNMP trap generation, code-upgrade behavior, QoS (Quality of Service) and gigabit-over-copper physical reliability.



In addition to these common tests, we asked each vendor to devise up to three unique tests that demonstrated its most valuable and differentiating features. We discuss these features in the reviews section and present the common benchmarks here.

Spanning Tree Failover

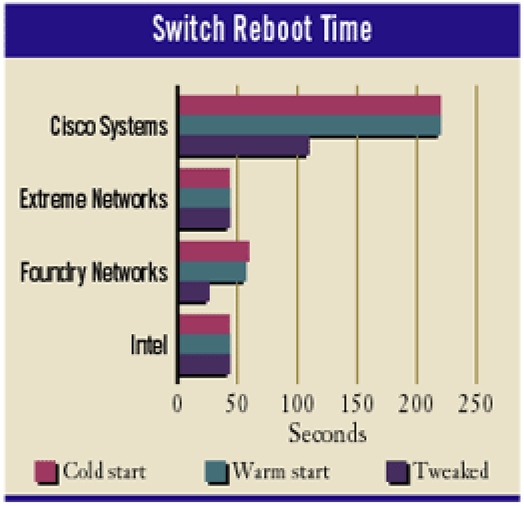
We began by testing the Spanning Tree protocol, which all vendors but Extreme specified as one level of fault-tolerance in their solutions; Extreme specified trunking instead. Spanning Tree works by detecting loops in networks and disabling one port on the network so that the loop doesn't generate broadcast storms. Vendors design loops into their solution to increase fault-tolerance, and rely on Spanning Tree to detect network failures and route around them on the redundant loops.

However, Spanning Tree also serves to provide fault-tolerance; should the primary link on the network go down, the backup link becomes active automatically and starts forwarding packets. The default implementation of Spanning Tree takes about 30 seconds to discover a downed port and fail over to the backup. However, many vendors allow you to "tweak" the Spanning Tree protocol by adjusting timers; some have even implemented advanced Spanning Tree algorithms that go beyond just tweaking timers.

We measured each core switch's default Spanning Tree behavior, then reran the test after each one minimized Spanning Tree failover time. Cisco's and Foundry's solutions had default times of 30.5 seconds, and Intel's and Extreme's times clocked in between 47 and 48 seconds. Optimized, however, the Cisco solution had by far the best failover time: just 0.97 seconds. Its unique UplinkFast advanced Spanning Tree algorithm also provides a mechanism to exchange bridging tables between switches when failover occurs, a feature that guarantees minimal flooding and retrain time after a Spanning Tree event. Foundry's enhanced Spanning Tree support helped reduce its solution's failover times to 5.3

seconds; Extreme and Intel reduced their products' failover times to 11.5 and 12.2 seconds, respectively. (See "Failover Time," above.)

Switch Reboot Time



Regardless of how mission-critical your network is, sooner or later you're going to need to reboot your infrastructure to upgrade, repair or replace some part of the solution. Many enterprises schedule weekly or monthly maintenance windows to handle hardware and firmware upgrades. When the need to reboot presents itself, you want the transition to be as seamless as possible. In testing reboot times, we found that most of the tested products took an extra 30 seconds to begin forwarding packets after reboot.

We concluded that this extra time was the result of Spanning Tree algorithms being initialized at reboot time. This problem plagues any Ethernet solution, not just gigabit over copper. (See "Switch Reboot Time," right.)

No mission-critical installation would be complete without redundant power supplies. Extreme and Intel use switches with integrated dual power supplies. If a power supply fails, the switch must be powered off, removed from the rack and serviced by an authorized technician (on-site). The Cisco Catalyst 6506 uses dual hot-swappable power supplies. Foundry's BigIron 8000 has room for four power supplies, of which at least two must be operational for proper function.

To determine how the products behave when one or more power sources fail, we took all but one power supply offline for each vendor. All the devices handled the power failure gracefully. Even Foundry's BigIron 8000, whose manual says the device requires two power supplies, survived our test with only a single power supply online.

Power-Failure Notification

We also tested each core switch's ability to notify a management station that power had failed. The Foundry, Extreme and Cisco switches passed this test with no problem, delivering traps to our HP OpenView management workstation and notifying the station of power supply failure and reinsertion. Intel's private MIB does not throw traps to HP OpenView, however; thus there was no notification of power supply failure for the Intel NetStructure 480T switch. Intel claimed it was a bug in its code and that a release to correct the problem would be available by press time.

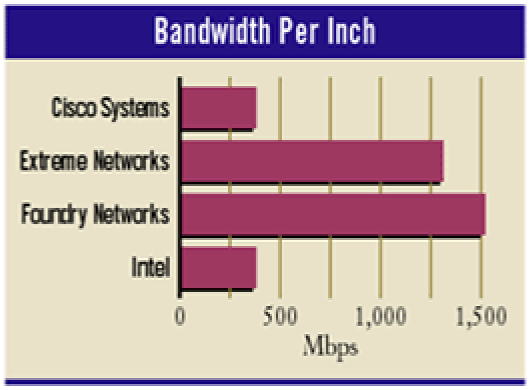
To minimize network downtime, it is desirable to be able to upgrade code without switch reboots. We asked vendors to demonstrate their switches' behavior during a code upgrade and found significant differences. The Intel and Extreme switches offer incremental configuration changes and scheduled reboots. Scheduled reboots are primarily done to implement a code upgrade (a reboot is required for a code upgrade). The switch administrator also can schedule how often incremental upgrades are applied. In this fashion, a form of version control can be introduced on the Extreme and Intel platforms.

Foundry's BigIron 8000 also requires at least one reboot after upgrading the primary management module. Once the primary module is upgraded, the secondary syncs up with the primary in the same amount of time a reboot takes. Only Cisco offered a nonstop-code-upgrade facility. By enabling the high-availability mode on the Catalyst 6506 Cisco provided, we were able to upgrade one supervisor module, reboot it and bring it back online without interrupting network service.

Once the primary image is upgraded, the secondary image can be cycled in a similar manner. With this facility, you can upgrade switch code without a reboot, increasing your overall uptime. Cisco representatives spent considerable time demonstrating that not only does this feature work, but that there is less than half a second of downtime between failovers.

Quality of Service

Guaranteed delivery of certain types of traffic is a requirement for top-notch network performance. To characterize the QoS classification and delivery abilities of the devices under test, we used a pair of VBrick video encoder/decoder devices from Vbrick Systems. These devices stream live MPEG-1 video across the network at up to 3,000 KB per second and audio at 384 KB per second. With the video running across the network, we oversubscribed the 1-Gbps backbone by 2 to 1 using a Netcom Systems SmartBits traffic generator. Our goal was to have each vendor demonstrate that it could classify traffic by Layer 3 or Layer 4 protocol information and to observe the quality of the video with QoS enabled.



Extreme's and Intel's devices passed the test with no notable picture degradation. However, on the Cisco Catalyst 6506 we saw choppy, less-than-perfect behavior. The Catalyst uses a weighted round-robin queuing mechanism, and it shows in this particular application. Whereas other vendors' switches delivered crystal-clear audio and video, Cisco's core devices struggled with the load we provided, delivering some quality of service, but not a perfect picture. This is partly due to Cisco's weighted-fair-queuing mechanism, which cannot have queue weights of zero (thus some packet loss will occur in the high-priority stream).

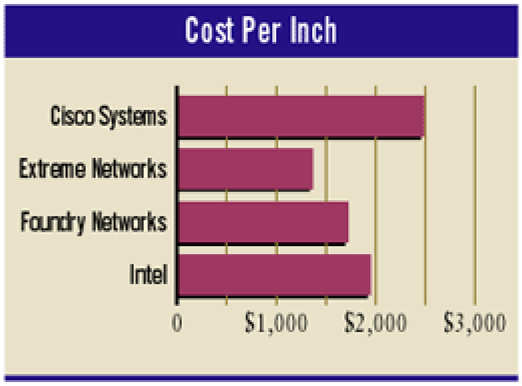
Foundry ran into a configuration problem in the lab, such that the vendor could not complete our QoS tests. In past tests, Foundry has demonstrated the ability to deliver priority-queuing services, but we couldn't use this video test to measure the vendor's QoS delivery mechanism. After the fact, Foundry discovered that it must have an IP address assigned to the router interface for port-based priority queuing to work properly, even if the traffic isn't crossing a routed boundary.

Gigabit-Over-Copper Performance

We also designed a vigorous test to stress the gigabit-over-copper physical medium. According to the IEEE standard, gigabit over copper is designed to run over standard Category 5 wiring. We used Category 5E wiring (rated slightly better than standard Category 5) and patch panels provided by Krone Americas as our test bed in the lab. Our test configuration consisted of a patch panel and a wall jack, with 300 feet of Category 5E-certified cabling between the two switches.

At one end, a patch cable connected one DUT (device under test) to the cabling infrastructure via a wall jack connector. At the other end, a patch cable attached to a panel provided the connection to the other DUT. A Netcom SmartBits device generated load traffic. The goal was to determine whether the physical interface could drive 1 Gbps of traffic in each direction without errors or packet loss. Intel's, Extreme's and Cisco's solutions passed the test with little or no difficulty. Foundry's switches showed a few packets coming in with FCS Alignment errors from the Gigabit Ethernet run. After a few test trials, we observed a flawless run.

While the solutions performed admirably, it is difficult to say whether or not other solutions would have exhibited problems had we tested the link more times. Foundry took several attempts to pass the test; other solutions hit it right the first time, but we didn't do exhaustive testing to determine how reliable each vendor's product was. Given that our test bed was well in excess of IEEE 802.3ab gigabit-over-copper specs for quality of copper cabling, and was within the gigabit-over-copper distance and interconnect specifications, we were a little concerned that we saw any corruption, regardless of the vendor.



The results scream a word of caution to shops considering gigabit over copper: This technology is really pushing the limits of conventional copper wiring, and a less-than-perfect network may generate significantly higher error counts than our tests revealed. Of course, these errors don't halt operation; they merely decrease throughput and increase network errors. But such errors are highly undesirable, and you should think twice before blindly dropping gigabit over copper into an unknown environment.

During our tests, we saw several phenomena that we can only attribute to gigabit-over-copper interference. During several of our distance tests, we had to recable, repatch and double-check the consistency of our links. Our tests were done on a cabling plant with brand new, end-to-end Category 5E certified runs. We tested each run with a Fluke Corp. DSP-4000 cable tester and verified at least 3 dB to 4 dB of headroom over Category 5E on each run. Furthermore, when we ran into errors, we retested our runs to ensure that no patch cables or patch ports had been damaged.

Of course, the errors and difficulties we saw were only present in the most extreme cases. We encountered no errors when running gigabit over copper at distances of 50 or 100 feet. It was only when we pushed the spec to the limit that we began to see strange phenomena cropping up on our network. Specifically, when we used a 98-meter run, we saw intermittent errors that were nonexistent at shorter distances. The problem became more pronounced as we tested additional runs simultaneously. In our worst-case experience, we lost more than 1,000 packets over 15 million sent (99.993 percent reliability). And our tests were run on Category 5E, whereas the 1000BASE-T specification calls only for Category 5 wiring, which is more commonly used. Why should you have to sacrifice any reliability to the physical layer?

Essentials

Gigabit-Over-Copper NICs

To take advantage of gigabit over copper, your server needs gigabit-over-copper NICs. SysKonnect's SK-9822 card is an ace in the hole. By Joel Conover

When we began our tests, we solicited gigabit-over-copper NICs from Alteon WebSystems, Intel, Sun Microsystems, SysKonnect and 3Com. Only Intel and SysKonnect could pony up a product, with Alteon backing out because it didn't want to compete against Intel's beta product. As of press time, the 3Com product is an Alteon OEM card. Sun Microsystems simply said it didn't want its products to be part of any competitive tests.

In our two-vendor NIC shoot-out, SysKonnect's card was the clear winner. Intel's early beta card showed promise, but some outstanding driver issues caused the card to behave improperly under high receive-traffic loads.

We tested both vendors' cards in a Compaq Computer Corp. ML-530 enterprise workgroup server with dual 800-MHz Pentium III Xeon processors and 640 MB of RAM. Using Ganymede Software's Chariot 3.2 testing software, we connected a dozen PCs as clients and hammered the server with TCP-based file copy traffic (simulated using Chariot's filesndl script) to determine maximum server throughput. We made no changes to the operating system other than to install the latest NIC drivers and service packs.

SysKonnect SK-9822 1000Base-T Gigabit Ethernet

SysKonnect has a reputation for making high-quality server-class products for enterprise applications, and the SK-9822 is no exception. SysKonnect's gigabit-over-copper adapter is available in single- and dual-port configurations. We tested the single-port version and took a quick look at the dual-port adapter. Both adapters feature the same hardware and performance figures. The dual-port adapter offers additional hardware failover fault-tolerance. SysKonnect claims the product has hardware failover times of roughly 100 milliseconds.

The SysKonnect 9822 adapter is a 64-bit, 66-MHz PCI card with fallback to 64/33, 32/66 and 32/33 operation modes. We installed the NIC in a 64-bit, 66-MHz slot on our Compaq ML-530 server and ran it through its paces. We used the card's DOS-based diagnostics tool to verify that the card was running in 64/66 mode; SysKonnect doesn't offer a Microsoft Windows-based diagnostics tool.

The SysKonnect NIC performed admirably in our tests. The card had a maximum throughput of 860 Mbps, which was measured in our Microsoft Windows NT 4 receive throughput test. In transmit mode, the card sustained throughput of 432 Mbps (Linux) to 649 Mbps (Novell NetWare), and 492 Mbps (Linux) to 860 Mbps (Windows NT 4) in receive mode. Intel's offering was slightly faster, with a top speed of 938 Mbps, and generally slightly faster throughput, but also had slightly higher CPU utilization. We tested the SysKonnect card using a 100-meter Category 5E cable run through a pair of patch panels and found that it performed flawlessly even at maximum distance.

The SysKonnect 9822 features 1 MB of on-board memory, and has quite a few features that enhance performance and fault-tolerance. These include dual-voltage (3.3 volt or 5 volt) support; driver and hardware support for TCP, UDP (User Datagram Protocol) and IP Checksum off-loading; support for jumbo frames; 802.1p and 802.1Q class of service and VLAN tagging; and on-board temperature and voltage sensors. The card also supports PCI hot-plug specifications. SysKonnect offers a huge range of driver support, including drivers for Apple Computer MacOS; FreeBSD 3.0 and later; Hewlett-Packard HP-UX; IBM's AIX; Linux 2.x and later; Microsoft Windows NT 4 (alpha and x86), Windows 2000, Windows 95 and Windows 98; Novell NetWare Client32 for DOS and NetWare Server versions 3.12 through 5.x; SCO UnixWare; and Sun Solaris (x86 and SPARC architectures).

As we were testing, SysKonnect was close to releasing a new driver that will enable both ports on the dual-port card to function simultaneously. Link aggregation and trunking will be an integral part of the new driver suite. Currently, you must run third-party software to handle link aggregation and failover across multiple NICs. The hardware failover (dual-port NIC option) requires no special drivers.

SK-9822 1000Base-T Dual Port Gigabit Ethernet, $1,595, SysKonnect, (800) 716-3334 or (800) 752-3334, www.syskonnect.com or sales@syskonnect.com.

Intel Corp. Pro/1000 T Server Adapter

Intel's Pro/1000 T Server Adapter is Intel's second-generation Gigabit Ethernet NIC. Like SysKonnect's SK-9822, Intel's Pro/1000 T is a 64-bit, 66-MHz PCI product.

We tested an early beta version of Intel's Pro/1000 adapter. This beta version had drivers only for Windows 2000, NetWare 5.1 and Windows NT 4. Intel's previous-generation card also offered support for Linux, Intel Architecture Solaris and SCO UnixWare. These drivers are expected to be available on the Pro/1000 T Server Adapter later this year. However, with full link aggregation and fault-tolerance for multiple adapters as an integral part of the driver, Intel's driver suite has a leg up on SysKonnect's.

During our tests, we discovered some autonegotiation issues with the

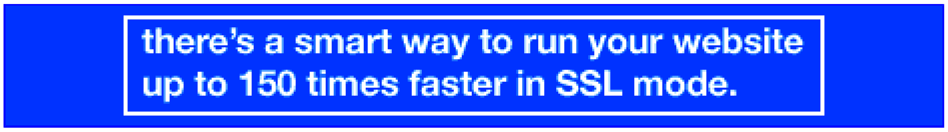
Pro/1000 T adapter and our Extreme Summit 7i and Cisco Catalyst 6506 switches. In Extreme's case, the fault was due to prerelease Broadcom PHY components in the Summit 7i we tested. However, we could not explain the problem with the Cisco switch. Some of these problems also may be because Intel doesn't use the Broadcom PHY chip in its design. Perhaps another good reason to stick with fiber.

In the lab, Intel's card was a real screamer, though we ran into some beta driver issues with certain traffic types. After a quick workaround from Intel, we obtained reliable performance data on the Pro/1000 T Server Adapter. In our throughput tests, the Intel card benchmarked throughput of 751, 938 and 752 Mbps for transmit, receive and bidirectional throughput (respectively) under Windows NT. Performance under Windows 2000 was lower by 75 to 225 Mbps. NetWare throughput also was lower, posting a top speed of 711 Mbps.

Like the SysKonnect card, Intel's offering supports all the features you need in a gigabit environment, including hot-plug PCI, jumbo frame support, VLAN tagging, IP and 802.1p QoS, and 802.1Q VLAN tagging. Intel's product passed our 100-meter distance test with no problems.

Intel Pro/1000 T Server Adapter, pricing not yet announced, Intel Corp., (800) 628-8686, (408) 765-8080; fax (408) 765-9904, www.intel.com.

Send your comments on this article to Joel Conover at jconover@nwc.com.



Copyright 1999, CMP Media Inc.