OPUS

All Services, All the Time:
High Availability Integrated Services Modules for the Catalyst 6500/Cisco 7600

prepared for Cisco Systems
September 2004

CONTENTS

ILLUSTRATIONS

Executive Summary

In data-center environments, high availability is at least as important a design goal as high performance. When millions of transactions are involved, even a small amount of downtime can lead to significant revenue loss. Complicating this problem for Web and SSL services is the fact that these services are *stateful*. Typically, the loss of one state-tracking component leads to the loss of all connection state in all components.

Cisco Systems has introduced redundancy features for Catalyst 6500/Cisco 7600 integrated services modules that preserve connection state even if one or more modules fails. By replicating connection state tables, the Cisco solution ensures that end-users' transactions continue uninterrupted. Furthermore, the integrated services modules all work together in a way that ensures that loss of one has no impact on the others.

Cisco commissioned Opus One, an independent networking consultancy, to assess the high availability capabilities of integrated services modules for Catalyst 6500 series switches and Cisco 7600 series routers[1]. For this report, Opus One conducted failover tests in a variety of scenarios, all intended to determine whether the Cisco solution would preserve session state despite the loss of one or more components.

Cisco also asked Opus One to conduct performance benchmarks to determine scalability of the integrated services modules when used together in a typical data-center deployment.

As to whether the integrated services modules are highly scalable and highly available, the answer to both questions is a resounding "yes." Among the highlights of this test:

- Support for nearly 1 million concurrent TCP connections through the Firewall Services and Content Switching Modules

- Support for establishment of more than 3,300 new SSL sessions per second using the SSL Services Module, exceeding Cisco's own published specification and representing a new high in a published benchmark.

- Support for nearly 76,000 HTTP transactions per second through the Firewall Services and Content Switching Modules, and nearly 9,000 HTTPS transactions per second through the SSL Services Module

- Zero session loss in the event of a failure of a Firewall Services Module

- Zero session loss in the event of a failure of a Content Switching Module

[1] Integrated services modules work in both the Catalyst 6500 series switches and Cisco 7600 series routers. This report describes tests conducted on Cisco 6509 switches, but all results apply to Cisco 7600 routers as well.

- Zero session loss in the event of a Supervisor card failure, thanks to Cisco's Non-Stop Forwarding/Stateful Switchover capabilities

- Zero session loss in the event of the loss of an entire switch with HTTP traffic

- Minimal session loss in the event of a failure of the SSL Services Module, with dropped sessions representing 0.25 percent or less of all active sessions

This report is organized as follows. We begin by introducing the three integrated services modules tested – the Firewall Services Module (FWSM), the Content Switching Module (CSM), and SSL Services Module. Then we discuss the configuration of the test bed and test traffic. Finally, we present results of scalability and failover tests.

Integrated Services Modules

Integrated services modules extend the functionality of Catalyst 6500 series switches and Cisco 7600 series routers by performing numerous upper-layer tasks. In this project, we examined the use of three such modules for firewall, content switching, and SSL tunneling.

The **Firewall Services Module** provides stateful multilayer access control. In addition to multilayer traffic inspection, the FWSM offers dynamic, per-user authentication and authorization; safeguards against denial-of-service (DoS) attacks; content filtering capabilities; support for multiple DMZs on logical or physical interfaces; remote management; and virtualization, a means of partitioning one FWSM into multiple virtual firewalls, or "security contexts." A single FWSM supports up to 256 security contexts.

The FWSM keeps track of connection state for virtually any flow, even those using stateless UDP. Once it adds an entry to its state table, the FWSM sends all packets from a given flow through its hardware-assisted "fast path." Our tests demonstrated the ability to scale to nearly 1 million concurrent connections. We present more performance data on the FWSM later in this report.

More data on the FWSM is available at http://www.cisco.com/en/US/products/hw/switches/ps708/products_data_sheet09186a00801daa53.html.

The **Content Switching Module** provides load-balancing capabilities. When placed in front of a group of servers, caches, or firewalls, the CSM improves scalability and protects investment in existing back-end infrastructure.

The CSM uses numerous criteria for load-balancing, including TCP or UDP port numbers, URLs, cookies and HTTP headers. The CSM also offers a choice of load-balancing algorithms.

To help ensure high availability, the CSM continuously monitors servers and applications, and offers rate-limiting safeguards against denial-of-service attacks.

Like the FWSM, the CSM scales to support nearly 1 million concurrent connections. We provide more performance data on the CSM later in this report.

More data on the CSM is available at http://www.cisco.com/en/US/products/hw/modules/ps2706/products_data_sheet09186a00800887f3.html.

The **SSL Services Module** offloads key security functions onto dedicated hardware, freeing up server resources and reducing transaction times. Hardware offload is especially important with SSL, given that authentication and especially encryption are highly processor-intensive tasks.

The SSL Services Module and CSM work together to use a “sticky” cookie that links particular users with particular servers. Session persistence between the modules helps ensure a seamless experience for clients, even across multiple changes in source IP addresses.

The SSL Services Module’s dedicated security hardware speeds transactions to unprecedented rates. Our tests demonstrated the ability to establish more than 3,000 new SSL sessions per second and handle more than 50,000 concurrent clients. Each new SSL “session” represents the combination of a TCP three-way handshake; an SSL handshake; and the transmission of an HTTPS GET message. We give more SSL performance data later in this report.

In addition to speeding transaction times, the SSL Services Module protects investment in existing data-center infrastructure. The SSL Services Module extends the lifetime of existing servers by offloading key security tasks; thus, no additional hardware is needed to scale SSL transaction capacity. In addition, the SSL Services Module obviates the need to buy additional server certificates from third-party certificate authorities for each back-end device; instead, the SSL Services Module presents one certificate for all client requests.

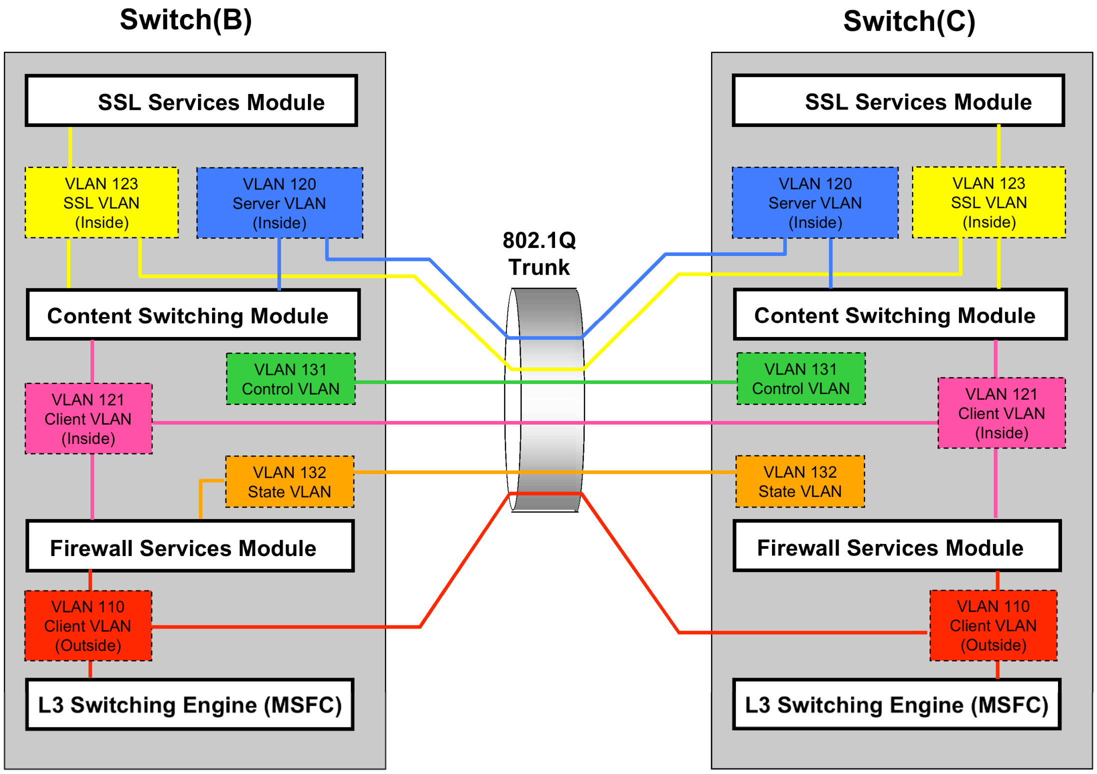
More data on the SSL Services Module is available at http://www.cisco.com/en/US/products/hw/modules/ps2706/products_data_sheet09186a00800c4fe9.html.

The Test Bed

Test Bed Configuration

Figure 1 below illustrates the logical configuration used in all tests. A pair of Catalyst 6509 switches, designated Switches B and C, are each equipped with integrated services modules providing SSL tunneling, content switching, and firewall functionality. Each switch also houses a Multilayer Switching Feature Card (MSFC), which runs routing protocols and populates Cisco Express Forwarding (CEF) tables in switching ASICs.

Figure 1: High Availability Logical Test Bed


Switch(B)
Switch(C)
SSL Services Module
VLAN 123 SSL VLAN (Inside)
VLAN 120 Server VLAN (Inside)
802.1Q Trunk
Content Switching Module
VLAN 131 Control VLAN
VLAN 121 Client VLAN (Inside)
VLAN 132 State VLAN
Firewall Services Module
VLAN 110 Client VLAN (Outside)
L3 Switching Engine (MSFC)

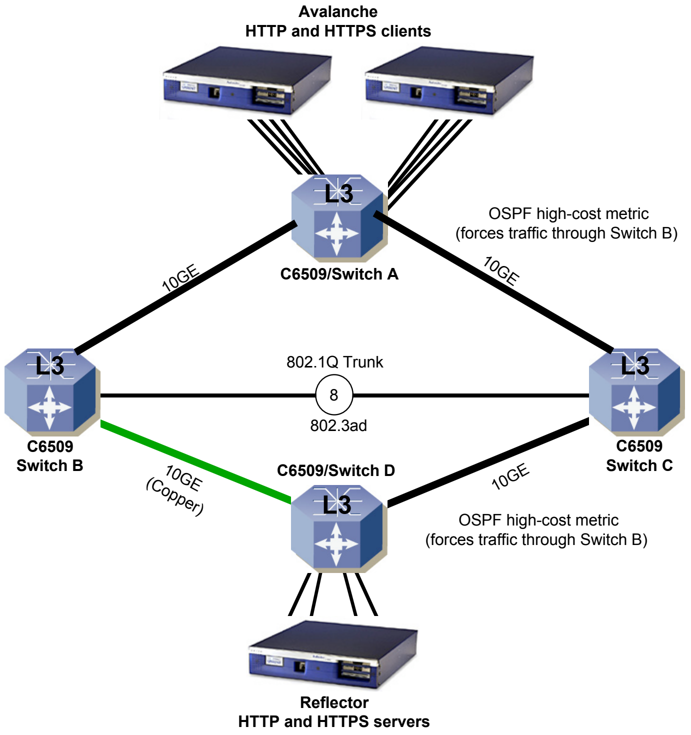
Each switch has a single set of integrated services modules, configured so that modules in Switch B provide stateful backup for those in Switch C and vice-versa.

The switches replicate state tables and other information through the use of VLANs. Separate VLANs are created for various tasks, such as sharing of state tables, forwarding of client traffic, device control, and so on. The two switches share VLAN traffic through a trunk port defined on each switch.

Cisco demonstrated investment protection in these tests by equipping Switch C with legacy Supervisor 2 management cards. Switch B used redundant Supervisor 720 management cards.

Figure 2 below shows the physical configuration of the test bed. The Spirent Avalanche and Reflector HTTP and HTTPS traffic generators attach to Switches A and D, respectively.

Figure 2: High Availability Physical Test Bed



To force traffic over a particular path, the OSPF configuration in Switches A and D assigned a lower-cost metric through Switch B than Switch C. In most tests, the only traffic to Switch C was replication of state information sent over the trunk port between Switches B and C.

Thus, there was no load-balancing of traffic between the two switches. For each failover event, 100 percent of the connection state information was transferred from one integrated services module to another.

Traffic Types

All tests generated a large amount of L4-L7 state that required synchronization between online and standby services modules. These tests stressed the ability of the services modules to maintain synchronization, and to seamlessly migrate connection state in the event of a module failure.

We used two traffic models to build up connection state: HTTP alone, and HTTP plus secure HTTP (HTTPS) transmitted through tunnels set up with Secure Sockets Layer (SSL). For both traffic types, we used Spirent Communications' Avalanche and Reflector test instruments. Avalanche emulates HTTP and/or HTTPS clients, while Reflector emulates servers.

We chose HTTP because of its continued prevalence on IP backbone links.[2] We used SSL because of that protocol's rapid growth in popularity as a means of setting up virtual private networks (VPNs). Using Cisco's SSL Services Module, our tests set up secure tunnels with client stations.

[2] See http://ipmon.sprint.com and http://caida.org for analysis of Internet backbone traffic by application.

Test Results

Integrated Services Module Scalability

Although this report's primary focus is high availability, we begin with performance baseline measurements to demonstrate the scalability of the Firewall Services, Content Switching, and SSL Services Modules.

These tests measure scalability four ways: concurrent connections; HTTP transaction rate; HTTPS transaction rate over SSL; and SSL tunnel establishment rate.

Concurrent Connections

Cisco claims support for nearly 1 million concurrent TCP connections in its Firewall Services and Content Switching Modules. Our tests verified that claim. With the integrated services modules working together, we sustained more than 900,000 concurrent TCP connections in failover scenarios.

Cisco also claims that the SSL Services Module supports more than 50,000 concurrent SSL tunnels. Once again, our tests validated that claim.

It is important to note that our scalability tests did not determine the absolute maximum number of connections supported. The original goal was to begin with 900,000 HTTP concurrent sessions and then add 50,000 concurrent SSL tunnels. As discussed later in this report, we were unable to do so because of a limitation with our test equipment. Nonetheless, the integrated services modules did demonstrate the ability to scale up to or beyond the performance levels claimed by Cisco.

HTTP Transaction Rate

For HTTP transaction rate tests, we configured the Spirent Avalanche and Reflector test instruments to process HTTP transactions as quickly as possible. To maximize transaction count, we configured Avalanche's emulated clients to use HTTP 1.1 with persistence (so that new requests would re-use existing TCP connections). To minimize transaction time, we configured Avalanche to request small (1-kbyte) objects from Reflector's emulated servers.

Each "transaction" represents the interval starting when a client sends an HTTP GET request and ending when the client receives the entire requested object.

These tests fully exercised the FWSM and CSM working together. Without the firewall handling this heavy load, packet loss or even session timeouts would have occurred. Without the CSM sending off requests to servers, transactions would have failed. We only considered a test run successful if the FWSM and CSM worked together to process every transaction with zero failures. The two integrated services modules did so in all cases.

In the HTTP transaction rate tests, the FWSM and CSM handled an average of 75,902 transactions per second. That is at least one order of magnitude greater than even many large enterprise data centers handle.

Table 1 below presents real-time results from the HTTP transaction rate tests.

Table 1: HTTP Transaction Rate

Elapsed time (seconds)	HTTP transactions per second
4	75,977
8	75,951
12	75,521
16	76,805
20	75,213
24	76,611
28	74,936
32	76,893
36	75,469
40	75,928
44	75,716
48	76,592
52	74,919
56	76,879
60	75,114

HTTPS Transaction Rate

We then conducted the same test to determine the transaction rate for HTTPS over SSL. Given that SSL is highly processor-intensive, transaction rates are inevitably lower. Typically, software-only SSL devices can handle, at most, a few hundred requests per second. In contrast, Cisco's SSL Services Module scales much higher: It handled an average of 9,149 HTTPS requests per second during our test. Clearly, the SSL Services Module's offloading of compute-intensive authentication and encryption tasks has a beneficial effect on scalability.

Each "transaction" represents the interval starting when a client sends an HTTPS GET request and ending when the client receives the entire requested object.

As in the previous tests with HTTP only, we maximized transaction count by using HTTP 1.1 with persistence (so that new requests would re-use existing TCP connections). We also configured Avalanche to request small (1-kbyte) objects from Reflector's emulated servers to minimize response time.

Table 2 below presents real-time results from the HTTPS transaction rate tests.

Table 2: HTTPS Transaction Rate

Elapsed time (seconds)	HTTPS transactions per second
4	9,087
8	9,030
12	9,060
16	9,138
20	9,610
24	8,987
28	9,285
32	8,792
36	9,416
40	9,153
44	8,849
48	9,459
52	9,050
56	9,094
60	9,222

SSL Session Establishment Rate

The final scalability test measured SSL session establishment rate. This is arguably the most stressful of the scalability tests, since each connection establishment request required not only the processing of a TCP three-way handshake, but also the compute-intensive task of SSL session establishment, followed by an HTTPS request.

For the concurrent SSL tunnel test, we configured Avalanche to make HTTPS requests, and configured Reflector to wait 60 seconds before returning each object; this ensured long-lived connections, and allowed us to build up a large amount of connection state. Note that each "tunnel" represents one three-way TCP handshake, one SSL session handshake, and one HTTPS GET request.

In previously published data of SSL VPN performance, the maximum connection establishment rate observed was 1,100 SSL sessions per second.[3]

Cisco's SSL Services Module went well beyond that level in this test, and even exceeded Cisco's published claim of 3,000 connections established per second. In these tests, the SSL Services Module set up an average of 3,398 SSL tunnels per second.

Multiplying that figure over time, the SSL Services Module can set up more than 12 million tunnels per hour, or more than 293 million tunnels per day.

[3] Newman, David. "SSL VPNS: Access Anywhere, Anytime," Light Reading, December 2003. http://www.lightreading.com/document.asp?doc_id=44442

To put these numbers in perspective, consider that eBay, one of the world's busiest e-commerce sites, handles around 533 million transactions per day[4]. This entire load could be processed through just two Catalyst 6500 or Cisco 7600 devices equipped with SSL Services Modules. (Of course, sensible network design dictates a somewhat more distributed approach).

Note that the level of concurrent HTTPS transactions in our tests – around 50,000 – is far higher than either the HTTPS transaction or setup rates. That is because the units of measurement are different. Concurrency is a measure of capacity, while the transaction and setup tests are measures of rate.

Table 3 below presents real-time results from the HTTPS connection establishment rate tests.

Table 3: HTTPS Connection Setup Rate

Elapsed time (seconds)	HTTPS sessions established per second
4	3,213
8	3,567
12	3,419
16	3,282
20	3,349
24	3,508
28	3,361
32	3,306
36	3,518
40	3,435
44	3,408
48	3,406
52	3,428
56	3,414
60	3,358

[4] Marshak, David S. "eBay Creates Technology Architecture for the Future." Patricia Seybold Group for Sun Microsystems, 2003. p. 12. http://www.sun.com/service/about/success/recent/Sun_eBay6-2_forWeb.pdf

Firewall Services Module Failover

All firewalls keep track of the connections flowing through them, but what happens if a firewall fails? With conventional firewalls, failure results in the loss of all connection state, leaving users unable to re-establish access with enterprise computing resources.

Cisco’s integrated firewall services module (FWSM) provides high availability by replicating TCP connection state with a redundant FWSM in another switch. This redundancy ensures there is no loss in connectivity for end-users, even if one firewall ceases to function.

We verified FWSM high availability with failover tests involving nearly 1 million concurrent TCP connections.

As shown in Figure 2, our test bed connected clients and servers through four Catalyst 6500 switches. Two of these were with a Firewall Services Module; a Content Switching Module; and an SSL Services Module. One of these chassis used redundant Supervisor 720 management modules; we designated this system Switch B. The other Catalyst 6509 used Supervisor 2 modules; we designated this Switch C.

As noted, the switches’ OSPF configuration assigned unequal metrics to force traffic over a particular path. There was no load-balancing between the two paths on the test bed, and we verified that all connection state was transferred between integrated service modules in the event of a failover.

We began with a baseline test to verify the FWSM’s ability to handle a large number of concurrent HTTP sessions. The FWSM and CSM are rated at 1 million concurrent connections.

In the baseline test, we loaded up the system with approximately 900,000 concurrent HTTP sessions using the Spirent Avalanche and Reflector test instruments. For each HTTP GET request received by an emulated server on Reflector, it would wait 60 seconds before returning the object, thus ensuring long-lived connections. We deliberately set the load at 900,000 rather than the rated capacity of 1 million to allow additional headroom for the switches to handle any failed connections.

The baseline tests showed the FWSMs capable of handling 900,008 concurrent connections, each involving long-lived HTTP transactions.

We then repeated the same test while forcing a failure of a FWSM. About 30 seconds into the test duration, we physically removed the FWSM from Switch B. The failure of the primary FWSM forced the system to reroute connections and connection state through the FWSM in Switch C.

We repeated this test twice, first with Switch B acting as the primary switch, and again with Switch C as the primary.

In both cases, the system maintained all sessions with no loss in connectivity. In the case of a firewall failure in Switch B, the redundant FWSM maintained state for all 900,010 concurrent connections. When we tested a firewall failure in Switch C, once again the failure led to zero loss in connectivity for all 900,006 concurrent TCP sessions established in the test.[5]

There was no loss of connectivity from the end-users' perspective in these tests. Equally important, there was no period during the test when an attacker could have penetrated the network, since firewall rules were always in force on both the primary and secondary FWSMs.

Table 4 below summarizes results from the firewall failover tests with HTTP traffic. Entries in *green italic type* show the number of concurrent connections after loss of the primary FWSM.

Table 4: Firewall Services Module Failover With Long-Lived HTTP Sessions

Elapsed time (seconds)	Established TCP connections, baseline case	Established TCP connections, loss of Switch B FWSM	Established TCP connections, loss of Switch C FWSM
4	900,008	900,010	900,006
8	900,008	900,010	900,006
12	900,008	900,010	900,006
16	900,008	900,010	900,006
20	900,008	900,010	900,006
24	900,008	900,010	900,006
28	900,008	900,010	900,006
32 (post-failover)	900,008	*900,010*	*900,006*
36 (post-failover)	900,008	*900,010*	*900,006*
40 (post-failover)	900,008	*900,010*	*900,006*
44 (post-failover)	900,008	*900,010*	*900,006*
48 (post-failover)	900,008	*900,010*	*900,006*
52 (post-failover)	900,008	*900,010*	*900,006*
56 (post-failover)	900,008	*900,010*	*900,006*
60 (post-failover)	900,008	*900,010*	*900,006*

[5] The slight differences between TCP connection counts in the baseline and failover cases reflect minor variations in connection establishment time, and are dependent to some degree on the mechanics of TCP itself.

For the SSL tests we used a total of 200,000 concurrent connections, 50,000 of which were HTTPS sessions designed to exercise the stateful failover capabilities of the SSL Services Module in addition to those of the FWSM and CSM.

The difference in connection counts – 200,000 here, versus 900,000 when HTTP alone was used – serves to illustrate the heavy processing burden imposed by SSL encryption and authentication not only on the device under test but also on the test equipment itself. It was a processing limit of the available test equipment that ultimately determined the overall number of sessions in the test, not any deficiency in the Catalyst 6500 or Cisco 7600 integrated services modules.

In our baseline tests, Switch B maintained an aggregate total of 200,007 concurrent sessions. When we removed the FWSM from Switch B, we observed a slight increase in concurrent connection count.

The explanation for this is that a small number of connection attempts actually did fail during the FWSM failover. We configured the Avalanche traffic generators to maintain 200,000 active connections at all times. To keep the connection count constant, Avalanche attempted to set up new connections as old ones close. In this case, approximately 500 new connection attempts – or about 0.25 percent of the 200,000 total – failed during the FWSM failover.

We draw two conclusions from this result. First, there is a slight but nonzero cost to failover when SSL traffic is involved. Second, the cost is minimal. Note that overall connection count actually increased post-failover since Avalanche kept trying to establish more sessions.

We observed similar behavior after removing the FWSM from Switch C. Once again, the total TCP connection count increased by approximately 500 new connection attempts after the failover.

Table 5 below summarizes results from the firewall tests with HTTP and HTTPS. Entries in *green italic type* show the number of concurrent connections after loss of one FWSM.

Table 5: Firewall Services Module Failover With Long-Lived HTTP and HTTPS Sessions

Elapsed time (seconds)	Established TCP connections, baseline case	Established TCP connections, loss of Switch B FWSM	Established TCP connections, loss of Switch C FWSM
4	200,007	200,000	200,003
8	200,007	200,000	200,003
12	200,007	200,000	200,003
16	200,007	200,000	200,003
20	200,007	200,000	200,003
24	200,007	200,000	200,003
28 (post-failover)	200,007	*200,026*	*200,036*
32 (post-failover)	200,007	*200,027*	*200,036*
36 (post-failover)	200,007	*200,027*	*200,036*
40 (post-failover)	200,007	*200,027*	*200,036*
44 (post-failover)	200,007	*200,027*	*200,036*
48 (post-failover)	200,007	*200,027*	*200,036*
52 (post-failover)	200,007	*200,027*	*200,036*
56 (post-failover)	200,007	*200,027*	*200,036*
60 (post-failover)	200,007	*200,027*	*200,036*

Content Switching Module Failover

In data-center environments, the content switch – the device that parcels out client requests to multiple back-end servers – is among the most critical elements of any network design.

While a data center can continue to function if one or more servers fails, the loss of a single content switch essentially means the loss of *all* servers – and with them, availability of all users' connections. Thus, ensuring maximum uptime for content switches is a key requirement of data-center design.

Cisco offers high availability for its Content Switching Modules by replicating data about connection state between redundant CSMs. Stateful failover is virtually instantaneous, ensuring no loss in client connectivity.

As in the firewall services module tests, we verified the CSM's high availability with failover tests involving nearly 1 million concurrent TCP connections. Our test bed connected clients and servers via a pair of Catalyst 6509 switches, each equipped with a CSM; a Firewall Services Module; and an SSL Services Module – a typical deployment in many data-center environments. Once again, Switch B used redundant Supervisor 720 management modules, while Switch C used Supervisor 2 modules.

We began with a baseline test to verify the CSM's ability to handle a large number of concurrent sessions. The CSM and FWSM are rated at 1 million concurrent connections. In the baseline test, we loaded up Switch B with approximately 900,000 concurrent TCP connections. As in the firewall tests, we chose a number below the rated limit to allow headroom for any failed connections.

The clients sent HTTP requests to a virtual IP address set up by the CSM; the CSM, in turn, redirected these requests to "servers" emulated by the Reflector test instrument. The baseline tests showed the CSMs capable of handling 900,008 concurrent connections, each involving long-lived HTTP transactions.

We then repeated the tests twice, each time causing one CSM to fail. In the first case, we designated Switch B as the primary device. Approximately 30 seconds into the steady-state phase of the test, we physically removed the CSM from Switch B. The loss of a primary CSM caused the system to reroute all traffic through the CSM in Switch C. In this test, we observed a constant level of 900,010 concurrent sessions with no loss in client connectivity.[6]

[6] The slight differences between TCP connection counts in the baseline and failover cases reflect minor variations in connection establishment time, and are dependent to some degree on the mechanics of TCP itself.

In the second case, we physically removed the CSM from Switch C. This time, we observed a total of 900,006 concurrent TCP sessions, again with no loss in client connectivity.

In both failover cases, clients continued to retrieve Web pages as if no failure had occurred. Further, by replicating connection state, the redundant CSMs ensured that existing TCP connections remained active despite the failover.

Sharp-eyed readers may note that results in the CSM failover tests are identical to those in the firewall services module tests. This demonstrates that failure of one module has no effect of on the other. Indeed, the results show that failure of either the CSM or FWSM produces exactly the same result: Preservation of connection state for all client sessions.

Table 6 below summarizes results from the CSM failover tests. Entries in *green italic type* show the number of concurrent connections after loss of one content switching module.

Table 6: Content Switching Module Failover With Long-Lived HTTP Sessions

Elapsed time (seconds)	Established TCP connections, baseline case	Established TCP connections, loss of Switch B CSM	Established TCP connections, loss of Switch C CSM
4	900,008	900,010	900,006
8	900,008	900,010	900,006
12	900,008	900,010	900,006
16	900,008	900,010	900,006
20	900,008	900,010	900,006
24	900,008	900,010	900,006
28	900,008	900,010	900,006
32 (post-failover)	900,008	*900,010*	*900,006*
36 (post-failover)	900,008	*900,010*	*900,006*
40 (post-failover)	900,008	*900,010*	*900,006*
44 (post-failover)	900,008	*900,010*	*900,006*
48 (post-failover)	900,008	*900,010*	*900,006*
52 (post-failover)	900,008	*900,010*	*900,006*
56 (post-failover)	900,008	*900,010*	*900,006*
60 (post-failover)	900,008	*900,010*	*900,006*

We reran the CSM failover with 150,000 HTTP and 50,000 HTTPS sessions, all concurrent.

As in the FWSM failover tests, we noticed a very slight loss in connectivity when offering HTTPS traffic through the CSM. The same issues apply as in the FWSM tests: Avalanche strove to maintain 200,000 active connections at all times. To keep the connection count constant, Avalanche attempted to set up new connections as old ones close. A very small number of these new connection attempts failed during the CSM failover.

For example, in failing over from Switch B to Switch C, we observed the loss of five TCP sessions – or 0.0025 percent of the total 200,000 connections – during the failover. As Avalanche continued to bring up new connections, the ultimate post-failover total was 200,002, or two connections more than the original state.

After failing the CSM in Switch C, we observed the loss of a single TCP connection, representing 0.0005 percent of all connections.

Table 7 below summarizes results from the CSM failover tests involving HTTP and HTTPS traffic. Entries in *green italic type* show the number of concurrent connections after loss of one Content Switching Module.

Table 7: Content Switching Module Failover With Long-Lived HTTP and HTTPS Sessions

Elapsed time (seconds)	Established TCP connections, baseline case	Established TCP connections, loss of Switch B CSM	Established TCP connections, loss of Switch C CSM
4	200,007	200,000	200,005
8	200,007	200,000	200,005
12	200,007	200,000	200,005
16	200,007	200,000	200,005
20	200,007	200,000	200,005
24	200,007	200,000	200,005
28	200,007	200,000	200,005
32	200,007	200,000	200,005
36 (post-failover)	200,007	*199,995*	*200,004*
40 (post-failover)	200,007	*200,002*	*200,004*
44 (post-failover)	200,007	*200,002*	*200,004*
48 (post-failover)	200,007	*200,002*	*200,004*
52 (post-failover)	200,007	*200,002*	*200,004*
56 (post-failover)	200,007	*200,002*	*200,004*
60 (post-failover)	200,007	*200,002*	*200,004*

SSL Services Module Failover

In addition to tests demonstrating the high availability of other integrated services modules when handling SSL tunnels, we also measured the effects of loss of the SSL Services Module itself.

We physically removed an SSL Services Module from one switch while offering a combination of 150,000 HTTP and 50,000 HTTPS connections. Half the total – or 25,000 HTTPS sessions – were affected by the failover because the Content Switching Module load-balanced connections between the SSL Services Modules in each switch.

In migrating from the Switch B (equipped with Supervisor 720 modules) to Switch C, we observed a small net gain in concurrent connections after the failover. This increase represents new connections initiated by the Avalanche test instrument after the loss of a few existing connections.

The change in connection count occurred because the SSL Services Modules do not synchronize SSL connection state. When a connection fails (as it did when we removed an SSL Services Module), the client will attempt to establish a new connection, but using a new and different SSL session ID.

In our tests, removing an active SSL Services Module caused the failure of all 25,000 SSL sessions associated with that module. The Avalanche clients then attempted to establish new connections as quickly as possible. The vast majority of these new connection attempts succeeded on the first try. Those that did not tried retransmitting (a normal part of TCP behavior) but ultimately a handful of the retransmissions failed.

We observed similar behavior when removing the SSL Services Module from Switch C (equipped with Supervisor 2 modules). The number of concurrent connections increased, and actually varied slightly, after the failover. The same mechanics were at work here as in the Switch B failover, with dropped connections replaced by (mostly) re-established new connections.

The number of connections lost was small – at most, around 0.009 percent of the total of 200,000 concurrent connections.

In no event did we observe the Catalyst 6509 forwarding unencrypted traffic during the failover. The Avalanche client emulator did not report any anomalies in SSL behavior.

Table 8 below summarizes results from the SSL Services Module failover tests. Entries in *green italic type* show the number of concurrent connections after loss of one SSL Services Module.

Table 8: SSL Services Module Failover With Long-Lived HTTP and HTTPS Sessions

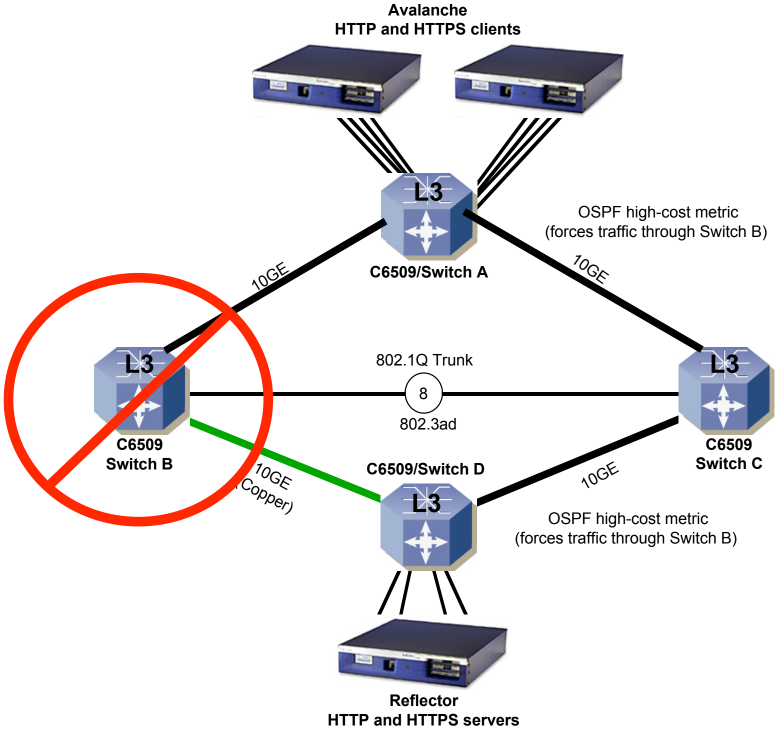
Elapsed time (seconds)	Established TCP connections, baseline case	Established TCP connections, loss of Switch B SSL Services Module	Established TCP connections, loss of Switch C SSL Services Module
4	200,007	200,004	200,002
8	200,007	200,004	200,002
12	200,007	200,004	200,002
16	200,007	200,004	200,002
20	200,007	200,004	200,002
24	200,007	200,004	200,002
28	200,007	200,004	200,002
32	200,007	200,004	200,002
36	200,007	200,004	200,002
40	200,007	200,004	200,002
44 (post-failover)	200,007	*200,022*	*200,020*
48 (post-failover)	200,007	*200,022*	*200,020*
52 (post-failover)	200,007	*200,022*	*200,018*
56 (post-failover)	200,007	*200,022*	*200,017*
60 (post-failover)	200,007	*200,019*	*200,017*

Power-Off Tests

Besides verifying the availability of individual integrated services modules, we also simulated a more catastrophic event: the loss of an entire switch, and along with it all integrated services modules. The objective was to determine whether multiple redundant services modules in another switch would maintain all existing sessions with no loss in connection state.

Figure 3 below shows the physical test bed we used for the first of two power-off tests. As illustrated, we shut off power to Switch B (equipped with a Supervisor 720 module) approximately 30 seconds into the steady-state phase of the test. The backup Catalyst, Switch C, maintained all 900,012 active TCP sessions with no loss in connectivity.

Figure 3: Power-Off Physical Test Bed



The second test, a power loss on Switch C (equipped with a Supervisor 2 module), also resulted in no loss of connection state. This time, Switch B picked up all 900,007 sessions, again with no loss in connectivity.

Despite the loss of entire switches, end-users' TCP connections were never affected.

Table 9 below summarizes results from the power-off tests. Entries in *green italic type* show the number of concurrent connections after loss of one switch.

Table 9: Power-Off Tests With Long-Lived HTTP Sessions

Elapsed time (seconds)	Established TCP connections, loss of Switch B	Established TCP connections, loss of Switch C
4	900,012	900,007
8	900,012	900,007
12	900,012	900,007
16	900,012	900,007
20	900,012	900,007
24	900,012	900,007
28	900,012	900,007
32 (post-failover)	*900,012*	*900,007*
36 (post-failover)	*900,012*	*900,007*
40 (post-failover)	*900,012*	*900,007*
44 (post-failover)	*900,012*	*900,007*
48 (post-failover)	*900,012*	*900,007*
52 (post-failover)	*900,012*	*900,007*
56 (post-failover)	*900,012*	*900,007*
60 (post-failover)	*900,012*	*900,007*

NSF/SSO Supervisor Failover

The previous tests all exercised the stateful failover capabilities of various integrated services modules. But what happens if an active Supervisor card fails? How much disruption to service can users expect under these circumstances?

In situations where redundancy of integrated services modules is deemed insufficient, Cisco recommends the use of its Non-Stop Forwarding (NSF) and Stateful Switchover (SSO) mechanisms. NSF/SSO adds even more resilience by protecting against the failure of a Supervisor card in the device housing the integrated services modules.

With NSF/SSO in place, there is no single point of failure – not in the integrated services modules and not in the Supervisor cards.

Although this report focuses primarily on redundancy in upper-layer services modules, we also explored the use of NSF/SSO to increase availability. Our objective was to determine what effect, if any, the loss of a Supervisor card would have on TCP connection state in the FWSM, CSM, and SSL Services Modules. A forthcoming report in this series will focus on NSF/SSO failover scenarios.

NSF makes use of the graceful restart mechanisms being developed by the IETF. It preserves layer-3 forwarding state during the loss and restart of a routing session. Cisco's NSF works with BGP, OSPF, and IS-IS. We used OSPF in these enterprise-focused tests.

SSO synchronizes layer-2 forwarding tables between Supervisor modules, ensuring that forwarding will continue even after the loss of a Supervisor module.

Conceptually, the NSF/SSO tests were similar to those for the various services modules: After establishing approximately 900,000 HTTP sessions with the Avalanche and Reflector test instruments, we physically removed the active Supervisor card from the switch under test. The only difference in this event was that NSF/SSO maintained L2 and L3 forwarding during failover to a standby Supervisor card in the same switch.

We designated a Catalyst 6509 equipped with Supervisor 720 cards as Switch B. The other system, which we dubbed Switch C, was equipped with dual Supervisor 2 cards. The Supervisor 720 not only establishes and maintains routing protocol and spanning-tree state, but also integrates a 720-Gbit/s switch fabric. Catalysts equipped with a Supervisor 2 use a separate switch fabric card; thus, when the Supervisor 2 is pulled, the data path through the switch fabric remains intact, whereas in the case of the Supervisor 720 it disappears for a short interval.

In the Switch B Supervisor failover tests, a Catalyst 6509 maintained 900,003 concurrent HTTP sessions, with no loss of client TCP connections. In the Switch C failover tests, a Catalyst 6509 maintained 900,008 concurrent sessions, again with no loss in connectivity for clients.

In both tests, traffic continued to flow uninterrupted through the same FWSM, CSM, and SSL Services Modules on each switch. In no case was traffic failed over to the integrated services modules on the other switch, nor did the services modules lose connection state.

Although NSF/SSO preserves L2 and L3 forwarding in the event of a Supervisor card failure, note that it does not preserve upper-layer connection state. For example, NSF/SSO will not save TCP connection state if a single (non-redundant) integrated services module fails. For this reason, Cisco recommends the use of NSF/SSO in conjunction with redundant services modules for maximum reliability: The services modules provide high availability for layer 4-7 connections, while NSF/SSO adds resiliency with non-stop forwarding at layers 2 and 3.

Table 10 below summarizes results from the NSF/SSO failover tests with HTTP traffic. Entries in *green italic type* show the number of concurrent connections after loss of a Supervisor module.

Table 10: NSF/SSO Supervisor Failover With Long-Lived HTTP Sessions

Elapsed time (seconds)	Established TCP connections, loss of Switch B Supervisor 720	Established TCP connections, loss of Switch C Supervisor 2 module
4	900,003	900,008
8	900,003	900,008
12	900,003	900,008
16	900,003	900,008
20	900,003	900,008
24	900,003	900,008
28	900,003	900,008
32 (post-failover)	*900,003*	*900,008*
36 (post-failover)	*900,003*	*900,008*
40 (post-failover)	*900,003*	*900,008*
44 (post-failover)	*900,003*	*900,008*
48 (post-failover)	*900,003*	*900,008*
52 (post-failover)	*900,003*	*900,008*
56 (post-failover)	*900,003*	*900,008*
60 (post-failover)	*900,003*	*900,008*

In tests involving both HTTP and HTTPS, we observed behavior similar to that of other events involving SSL: A small loss in connectivity during the failover, followed by modest gain in concurrent connections.

Table 11 below summarizes results from the NSF/SSO failover tests. Entries in *green italic type* show the number of concurrent connections after the loss of a Supervisor module.

Table 11: NSF/SSO Supervisor Failover With Long-Lived HTTP and HTTPS Sessions

Elapsed time (seconds)	Established TCP connections, loss of Switch B Supervisor 720	Established TCP connections, loss of Switch C Supervisor 2 module
4	200,000	200,000
8	200,000	200,000
12	200,000	200,000
16	200,000	200,000
20	200,000	200,000
24	200,000	200,000
28	200,000	200,000
32 (post-failover)	*200,017*	*200,015*
36 (post-failover)	*200,021*	*200,019*
40 (post-failover)	*200,021*	*200,019*
44 (post-failover)	*200,021*	*200,019*
48 (post-failover)	*200,021*	*200,019*
52 (post-failover)	*200,021*	*200,019*
56 (post-failover)	*200,021*	*200,019*
60 (post-failover)	*200,021*	*200,019*

Conclusion

Although the tests described in this document explore many different failure scenarios, there is one common thread among them: Existing connection state is preserved during and after component failure, even when nearly 1 million concurrent sessions are involved.

The tests also demonstrate the ability of multiple integrated services modules to work together in a single chassis. Further, the tests show that the loss of one module has no impact on others; for example, a Firewall Services Module will continue to function even if a Content Switching Module fails over to a redundant CSM.

Where even greater reliability is required, Cisco also demonstrated NSF/SSO protection against the loss of a Supervisor card. These tests stressed the NSF-aware capabilities of the integrated services modules: During Supervisor failover in one switch, we verified that integrated services modules continued to operate normally, and did not fail over to the standby modules in the other switch. This combination of upper- and lower-layer redundancy ensures maximum reliability for enterprise data centers.

Acknowledgements

Opus One gratefully acknowledges the support of Spirent Communications, which supplied engineering assistance for this project. Spirent service delivery manager Philip Joung and test engineer Brooks Hickman assisted with configuration and troubleshooting of Spirent’s Avalanche and Reflector test instruments.

TESTED with SPIRENT

About Opus One®

Opus One® is a consulting and information technology firm based in Tucson, AZ. Founded in 1989, Opus One's corporate goal is to help our clients make the best use of information technology. We focus on efficient and effective solutions in the areas of data networking, electronic mail, and security. For more information, see http://opus1.com or contact us at:

Opus One
1404 East Lind Road
Tucson, AZ 85719
+1-520-324-0494