

## Management Switch Fabric Module 3 (MSM-3)

MSM-3 is the third-generation Management Switch-fabric Module (MSM) for the BlackDiamond® 6800 series family. With its implementation of T-Sync enabled hitless failover and hitless software upgrade, MSM-3 sets the standard for resiliency for large, mission-critical Ethernet/IP networks that simply cannot fail. Additional features such as cross module link aggregation, a much greater number of Access Control Lists (ACLs), and faster route computation contribute further to greater system and network resiliency.

In addition to resiliency, MSM-3 also offers enhanced performance and larger BGP routing tables compared to the second-generation MSM64i as shown in Table 1. And given Extreme Networks' long-standing commitment to simplicity and investment protection, MSM-3 delivers the industry's leading resiliency and performance even when deployed in a BlackDiamond 6800 chassis with second-generation "i" series modules.

Category	Feature	MSM-3	MSM64i
<b>General</b>			
	CPU Speed	480 Mhz	240 Mhz
	Memory	512 Mb	256 Mb
<b>Resiliency</b>			
	Hitless Failover	Yes	No
	Hitless Upgrade	Yes	No
	Cross Module Link Aggregation	Yes	No
	Number of ACLs	10,000*	5,000
<b>IP Unicast</b>			
	BGP Routes	850k	330k

\* Future ExtremeWare release

**Table 1: MSM-3 vs MSM-64i Features**

### Moving Toward Hitless Failover and Upgrade

Corporations, governments and other institutions all over the world increasingly depend on their converged Ethernet/IP networks to reliably handle their mission-critical business applications while also carrying other sensitive traffic such as voice calls and video feeds. Consequently, even small periods of network downtime can have a large impact on the productivity and profitability of an enterprise.

Making networks truly unbreakable is a large challenge, but with the release of MSM-3, Extreme Networks® has raised the bar for the entire industry and taken a major step toward the goal of enabling hitless failover.

There is considerable discussion, debate, and hype in the networking industry about hitless failover and what it means. There are various marketing terms such as “non-stop networking” or “hitless protection” that are applied to the concept and each vendor has their own interpretation which can lead to confusion. However, at Extreme Networks we believe there are two simple and related metrics that are crucial to judge how unbreakable a network truly is:

Metric	Definition
Control Plane Restart Time	In the event of a management module failure, time required for a slave management module to come online and fully take over control plane functionality from a failed master management module
Data Plane Restart Time	The time required for a switch to resume forwarding traffic in the event of a control plane restart

If during a control plane restart the data plane restarts instantly (such that no packets are lost), the action can be characterized as hitless.

A BlackDiamond 6800 with MSM-3 achieves hitless failover in a variety of scenarios which are described later and clearly out performs the competition on the two metrics that matter most: control plane restart and data plane restart times (see Tables 2 and 3). The result—BlackDiamond 6800 with MSM-3 can failover to a slave management module with zero packet lost.

### **Hitless Upgrades**

Hitless upgrade is an extension or complement to hitless failover and allows ExtremeWare® software running on a BlackDiamond 6800 to be updated without losing any packets and without taking the switch out of service. Both hitless failover and upgrade are enabled on MSM-3 via innovative T-Sync technology that significantly improves system uptime by interconnecting multiple switch fabrics through a dedicated synchronization bus.

### **Hitless Failover/Upgrade in Layer 2 Networks**

Most Layer 2 Ethernet networks run some resiliency protocol such as Spanning Tree Protocol (STP), Extreme Standby Router Protocol (ESRP) or Ethernet Automatic Protection Switching (EAPS). These protocols ensure there are no loops in the network and that system failures are dealt with appropriately. The challenge however is that each time a switch in the network fails or goes down even momentarily (for example during a management module failover from master to slave) these Layer 2 resiliency protocols must reconverge. This reconvergence results in an unstable network that may drop critical traffic. For example, if a switch carrying a large number of Voice-over-IP (VoIP) calls fails, the calls would all drop, resulting in hundreds or thousands of unhappy end-users complaining about the reliability of their converged network.

In the case of a BlackDiamond 6800 with MSM-3, all STP, ESRP and EAPS protocol state information is preserved in the event of a failure. The preservation of the Layer 2 state (even as the slave MSM takes over from the master MSM) is unique in the industry and consequently, there is no need for reconvergence, and no associated traffic loss. If the BlackDiamond 6800 with MSM-3 were carrying a large number of VoIP calls, all those callers would happily continue talking with no idea that a module had failed in the network.

The scenarios that follow quantify Extreme Networks’ superiority over competitive alternatives in a Layer 2 network.

**Scenario 1:** A network operator upgrades the ExtremeWare software on the slave MSM and forces the master MSM to failover via CLI. This approach is used while upgrading ExtremeWare or when replacing an MSM.

In such a case, the MSM-3 data plane remains operational and established traffic flows continue without interruption. Also, the maintenance of Layer 2 protocol state information eliminates the need for reconvergence. The slave MSM-3 takes over the control plane and comes online within seconds, after which new flows can be established. Contrast this with typical network upgrades that usually require a complete reboot of the switch. This has the effect of completely stopping the network for the period of the reboot. Once the reboot is complete, Layer 2 STP learning must take place. This can result in downtimes that are often measured in minutes.

## Management Switch Fabric Module 3 (MSM-3)

On the other hand, hitless upgrades as enabled by Extreme Networks result in less downtime and greater operator productivity. If several switches need to be upgraded, administrators can save significant network downtime and finish their maintenance much more quickly.

Table 2 clearly shows that MSM-3 greatly outperforms the competition when a network operator initiates a management module failover.

	Hitless	Data Plane Restart Time	Control Plane Restart Time
<b>MSM-3</b>	Yes	0 sec (instant)	12 sec
<b>MSM64i</b>	No	12 sec	12 sec
<b>Cisco Supervisor 2 (with RPR+) - WS-X6K-S2-MSFC2</b>	No	43 sec	43 sec
<b>Foundry J-BxGMR4</b>	No	Feature not available at time of test	N/A

**Test conditions:** Traffic generator sending 1.488 Mpps (64-byte) simultaneously through 16GbE ports.

**Table 2: Performance during user initiated management module failover (upgrade)**

**Scenario 2:** An unexpected error or software exception occurs in the switch control software causing the master MSM-3 to fail. As in Scenario 1, since the data plane remains active established flows continue, the need for reconvergence is eliminated and new flows begin to be established when the slave MSM-3 control plane comes online after some seconds.

Table 2 applies to this scenario as well because from a system performance perspective a user initiated failover and software exception exhibit identical behavior.

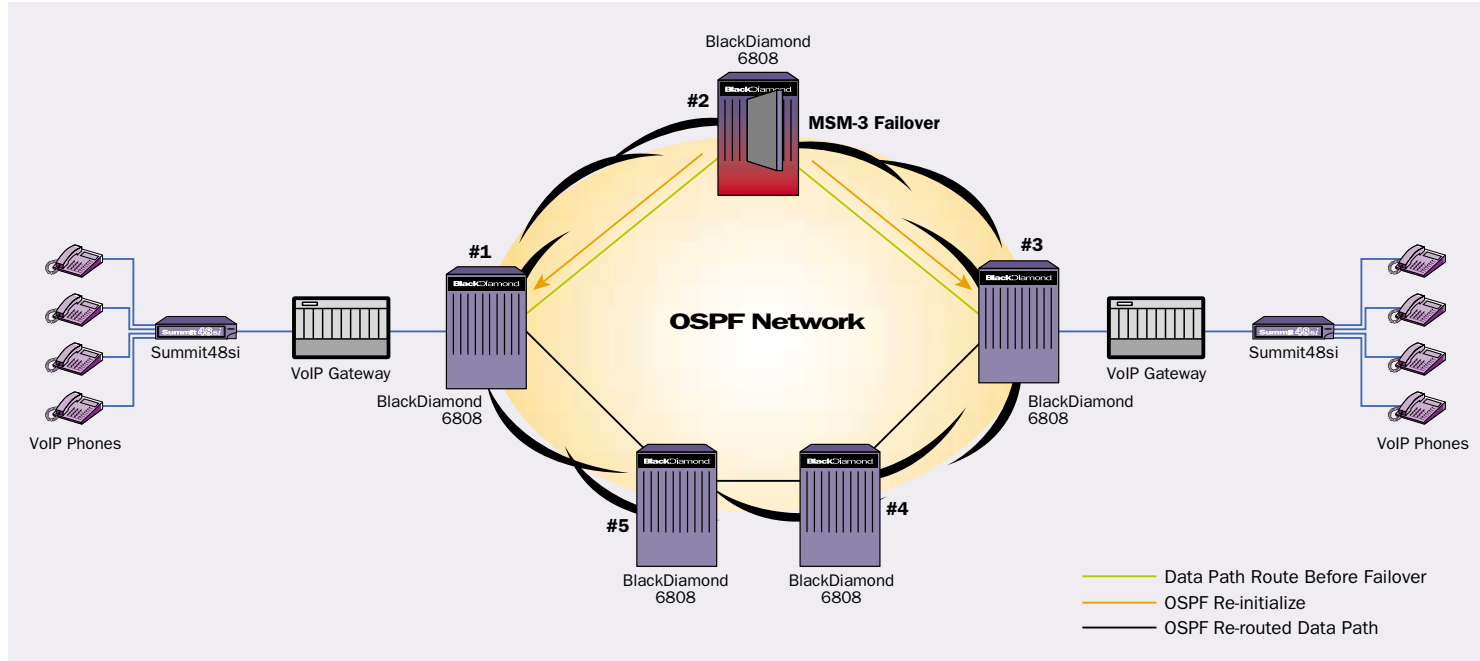
**Scenario 3:** MSM-3 suffers a hardware failure or the MSM-3 is physically removed from the chassis (without first issuing a failover command). Under these circumstances, some traffic may be lost, but the data plane is still capable of restarting even before the slave MSM's control plane is completely active – thereby minimizing packet loss. Such a situation is perhaps more accurately described as near-hitless, but it is worth noting that Extreme Networks significantly outperforms competitive alternatives as shown in Table 3.

	Hitless	Data Plane Restart Time	Control Plane Restart Time
<b>MSM-3</b>	No (near-hitless)	6 sec	12 sec
<b>MSM64i</b>	No	12 sec	12 sec
<b>Cisco Supervisor 2 (with RPR+) - WS-X6K-S2-MSFC2</b>	No	44 sec	44 sec
<b>Foundry J-BxGMR4</b>	No	28 sec	28 sec

**Test conditions:** Traffic generator sending 1.488 Mpps (64-byte) simultaneously through 16GbE ports.

**Table 3: Performance during management module hardware failover or removal**

## Hitless Failover/Upgrade in Layer 3 Networks



**Figure 1: Layer 3 Network**

The hitless behavior in Layer 2 networks described earlier is identical for Layer 3 networks if routes are directly connected or static, because during failover the BlackDiamond data plane continues to forward all Layer 3 flows without interruption.

When dynamic routing protocols such as OSPF are introduced the situation becomes slightly more complex as depicted in Figure 1. When MSM failover occurs in BlackDiamond (BD) #2, established Layer 3 flows such as the VoIP calls shown in green continue to flow uninterrupted through the switch while the slave MSM-3 is initializing. At some point during the slave initialization of BD #2, the OSPF process will restart (shown in orange) and the neighboring routers (e.g. BDs #1 and #3) will realize that BD #1 has restarted OSPF and flush their OSPF database of entries previously received from BD #2 and if an alternate path exists, re-route traffic to that alternate path (shown in black through BDs #4 and #5). Whether this action will be hitless or not (and more importantly whether VoIP calls are dropped or not) depends on how quickly the neighboring router can perform the traffic re-route.

Since the OSPF protocol restarts during hitless failover, the network will have to reconverge, but here again MSM-3 offers a significant advantage. OSPF will begin its convergence only after Layer 2 reconvergence has taken place and after Layer 2 protocols like STP have become stable. So the total period of instability is the sum of the time required for Layer 2 and Layer 3 reconvergence. Because of the MSM-3's ability to maintain Layer 2 state, the process of Layer 3 reconvergence can begin right away unlike a competitor's networks where Layer 2 reconvergence (up to a minute in many cases) must first be established.

There are several draft IETF proposals to modify OSPF ("Hitless OSPF Restart") and BGP ("Graceful restart of BGP") to support Layer 3 routing protocol level hitless failover. These proposals have not yet become standards and therefore are not available on MSM-3. Extreme Networks however participates in these working groups and will implement these extensions in the future as the standards become solidified. With support for such standards, the hitless behavior associated with Layer 2 networks and static Layer 3 networks will also be available in dynamic Layer 3 networks.

## Network Level Resiliency in Layer 3 Networks and Performance Enhancements

In addition to building networks with resilient switches, most network operators also design resiliency at the network level with primary and alternate routing paths. If another switch in the network goes down, the BlackDiamond 6800 with MSM-3 can rapidly failover to an alternate path. Specifically, MSM-3 has an enhanced CPU and memory, relative to MSM64i, thereby improving overall network resiliency through faster computation. The enhanced CPU and memory provide other material performance gains: more than doubling the number of BGP routes that can be processed and also providing a measurable boost to other network protocols and services such as Network Address Translation (NAT), Server Load Balancing (SLB) and others, while also improving basic network control functions such as configuration save and system boot time.

Feature	MSM-3 vs. MSM64i Performance Improvement
Network Address Translation (NAT) Packet Forwarding Rate	3.5 times faster (pps)
Server Load Balancing (SLB) Packet Forwarding Rate	4 times faster (pps)
DHCP Relay Rate	2 times faster (pps)
System Boot Time	4 times faster (pps)
Configuration Save Time	2 times faster (pps)

**Table 4: MSM-3 vs. MSM64i Performance**

## Even More Resiliency

In addition to the significant system resiliency offered by hitless failover and hitless upgrade and greater network resiliency achieved through faster computation, there are several other resiliency enhancements in the MSM-3 such as Cross Module Link Aggregation (CMLA) and a greater number of Access Control Lists (ACLs). CMLA offers the ability to bundle multiple ports on separate line cards in to a single logical link. This greatly improves network resiliency because even in the event of a fiber cut, a port accidentally becoming unplugged, or a module going offline, the network remains 100% operational, all without impacting a single user or application. MSM-3 will also provide up to 10,000 wire-speed ACLs which brings an additional level of resiliency to a customer's network because it allows a network administrator to keep more unwanted traffic - such as from a DOS attack - out of the network.

With its implementation of a comprehensive and mutually reinforcing set of resiliency features, the MSM-3 sets the new standard for system and network resiliency for large, mission-critical Ethernet/IP networks that simply cannot fail. Resiliency for large networks delivered the Extreme Networks way – a commitment to simplicity and investment protection.



3585 Monroe Street Santa Clara, CA 95051-1450 Phone 408.579.2800 Fax 408.579.3000  
Email [info@extremenetworks.com](mailto:info@extremenetworks.com) Web [www.extremenetworks.com](http://www.extremenetworks.com)

© 2003 Extreme Networks, Inc. All rights reserved. Extreme Networks, BlackDiamond, Summit, Summit7i, ExtremeWare, ServiceWatch, Extreme Ethernet Everywhere, Ethernet Everywhere, Extreme Velocity, Extreme Turbodriven and the color purple are registered trademarks of Extreme Networks, Inc. in certain jurisdictions. Alpine, ExtremeWare Vista, Extreme Standby Router Protocol, ESRP, Summit1i, Summit4, Summit4/FX, Summit5i, Summit24, Summit24e2, Summit24e3, Summit48, Summit48i, SummitLink, SummitGbX, SummitRPS, SummitPx1, PxSilicon, EPICenter, vMAN, the BlackDiamond logo, the Alpine logo and the Extreme Networks logo are trademarks of Extreme Networks, Inc., which may be registered or pending registration in certain jurisdictions. ExtremeWorks, the Extreme Turbodriven logo and the Go Purple-Extreme Solution Partner logo are service marks of Extreme Networks, Inc., which may be registered or pending registration in certain jurisdictions. All other registered trademarks, trademarks and service marks are property of their respective owners. Specifications are subject to change without notice.

L-TB-MSM3-309