



STORAGE AREA NETWORK

Technical Brief:

Brocade Local Switching with Multi-Stage Backbone and Director Platforms

Maximum networking throughput can be achieved with Brocade Local Switching technology.

BROCADE

CONTENTS

Introduction.....	3
What is Local Switching?	3
Two Director Design Approaches.....	4
Multi-Stage Switching with Interconnected ASICs.....	4
Single-Stage Switching with One or More Crossbars	4
Why One Approach over Another?	5
Port Blades for the DCX Backbone and 48000 Director.....	5
Brocade FC8-16 with One Condor 2 ASIC	5
Brocade FC8-32 with Two Condor 2 ASICs.....	6
Brocade FC8-48 with Two Condor 2 ASICs.....	7
Optimizing with Local Switching.....	8
HA Node Placement: Least Benefit.....	8
Balancing Storage, ISLs, and Hosts by ASIC: Good Benefit.....	8
Mapping Storage Allocations by ASIC: Maximum Benefit.....	9
Summary and Recommendations.....	9
Appendix: Port Blades in the Brocade DCX and 48000	10
FC8-16 or FC4-16 Port Blade.....	10
FC8-32 or FC4-32 Port Blade.....	11
FC8-48 or FC4-48 Port Blade.....	12

INTRODUCTION

This paper covers the following topics:

- Describes Brocade® Local Switching, a high-performance feature of Brocade Fabric OS® (FOS) enterprise-class platforms.
- While some directors can experience oversubscription, Brocade B-Series backbone and director platforms with Local Switching always performs at full speed.
- Using simple guidelines, these platforms can leverage Local Switching to achieve the full, maximum performance on every chassis port.

What is Local Switching?

Brocade Local Switching occurs when two ports are switched within a single Application Specific Integrated Circuits (ASIC). With Local Switching no traffic transverses the backplane of a director so CP4 or CR8 blade switching capacities are unaffected. *Local Switching always occurs at the full speed* negotiated by the switching ports regardless of any backplane oversubscription ratio for the port card. Only Brocade FOS-based director and backbone platforms have Local Switching capability, since they utilize Multi-Stage Switching architectures.

Brocade products with Local Switching include:

- Brocade DCX Backbone
- Brocade 48000 Director
- Brocade 24000 Director (EOL)
- Brocade 12000 Director (EOL)

NOTE: "EOL" stands for "end of life" and means that the product is no longer available from Brocade or its OEM partners, but it may be installed and running at a customer site.

Local Switching always occurs at full speed in an ASIC, although not necessarily a port blade. So, port 0 and port 31 on the FC8-32 blade communicate over the backplane and use the CP4 or CR8 back-end switching ASICs. On the same FC8-32 blade, port 0 can switch locally at full speed with ports 1 to 15 and port 16 can switch locally at full speed with ports 17 to 31. These port groups are called "Local Switching groups," since they each share the same port blade ASIC.

TWO DIRECTOR DESIGN APPROACHES

There are two director architectures currently in the Storage Area Network (SAN) market:

- Multi-stage switching with interconnected ASICs
- Single-stage switching with one or more crossbars

While these architectures have positive and negative aspects, they both perform the basic switching operation for user-facing ports. Neither method is inherently right or wrong, but one method provides much more user benefit: Multi-Stage Switching.

Multi-Stage Switching with Interconnected ASICs

One architecture approach uses ASICs to both process and switch data traffic. Processing could entail measuring performance, enforcing Zoning, and other protocol-related tasks. Switching is simply directing traffic from one port to another.

Only a single ASIC is needed as long as the ASIC has enough ports to accommodate the user-facing ports on the platform. If more user-facing ports are required, two or more ASICs are internally interconnected to provide more user-facing ports than exist on a single ASIC. How these ASICs are interconnected can critically impact the performance and scalability of a multi-stage platform.

Brocade implements a multi-stage approach on the FOS-based Brocade DCX and 48000. With the Brocade DCX, Condor 2 ASICs are on both the port blades and the core blades (CR8), which switch traffic over the backplane. The process is similar on the Brocade 48000 Director with Condor ASICs on the control processor blades (CP4) used to switch data over the backplane. For traffic between ASICs on different port blades, the data is first switched on a port blade, then switched on the CR8 or CP4, and finally switched on the destination port blade.

Brocade minimizes the cumulative latency effect of multiple switching stages by using cut-through routing, in which frames are switched after just the first portion of the frame header is read. The total latency of Brocade FOS switching platforms is measured in microseconds—orders of magnitude smaller than the millisecond metrics with which disk access is measured. The multi-stage architecture in the Brocade DCX, the Brocade 48000 and the EOL Brocade 24000 is analogous to a core-edge fabric with the port blades as edge switches and the CR8s or CP4s as core switches.

Single-Stage Switching with One or More Crossbars

Cisco uses a single-stage architecture for Cisco MDS directors. Brocade also implements single-stage switching with the Brocade Mi10K and M6140 Directors. In this approach, ASICs process data on the port blades, but they do not switch the traffic. Switching for all the ports in a single-stage director is performed by one or more serial crossbars. The crossbar does not process the data; its only function is to switch data between crossbar ports, which are connected to the port blade ASICs.

Why One Approach over Another?

The benefits of the single-stage serial crossbar approach include highly predictable performance across all ports in the chassis and a design that is easier to engineer for higher port capacity. A significant weakness of the single-stage director architecture is performance. All traffic must go through the crossbar—there is no local switching even for ports on the same port blade. The bandwidth maximum for a director chassis using this approach is simply the crossbar's bandwidth.

The multi-stage approach combines switching and processing in a single chip that can be leveraged in the director and switch platforms. This maximizes performance, capabilities, and decreases costs and time to market, since the design and testing effort can be optimized since just one chip is being used. Most importantly for directors, Multi-Stage allows for significantly higher performance since Local Switching is a characteristic of Multi-Stage and not Single-Stage designs. This is the approach Brocade has chosen with the DCX Backbone and 48000 Director and future chassis and blade enterprise switching platforms.

PORT BLADES FOR THE DCX BACKBONE AND 48000 DIRECTOR

The 8 Gbit/sec port blades described in this section are currently available.

Brocade FC8-16 with One Condor 2 ASIC

All 16 external ports make up one Local Switching group, which is made up of 16 x 8 Gbit/sec external ports, and 128 Gbit/sec Brocade DCX or 64 Gbit/sec Brocade 48000 internal bandwidth.

- 1:1 fully subscribed *for all local traffic* on either platform
- 1:1 fully subscribed for non-local traffic on Brocade DCX
- Up to 2:1 oversubscribed for non-local traffic on Brocade 48000

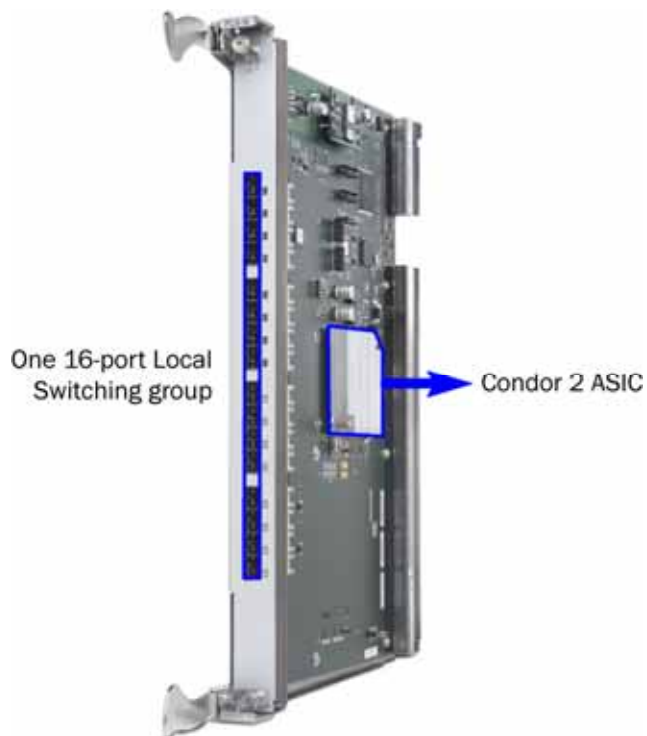


Figure 1. FC8-16

Brocade FC8-32 with Two Condor 2 ASICs

Top 16 and bottom 16 external ports make up two Local Switching groups, each with 16 x 8 Gbit/sec external ports, and 128 Gbit/sec Brocade DCX or 32 Gbit/sec Brocade 48000 internal bandwidth.

- 1:1 fully subscribed *for all local traffic* on either platform
- 1:1 fully subscribed for non-local traffic on Brocade DCX
- Up to 4:1 oversubscribed for non-local traffic on Brocade 48000

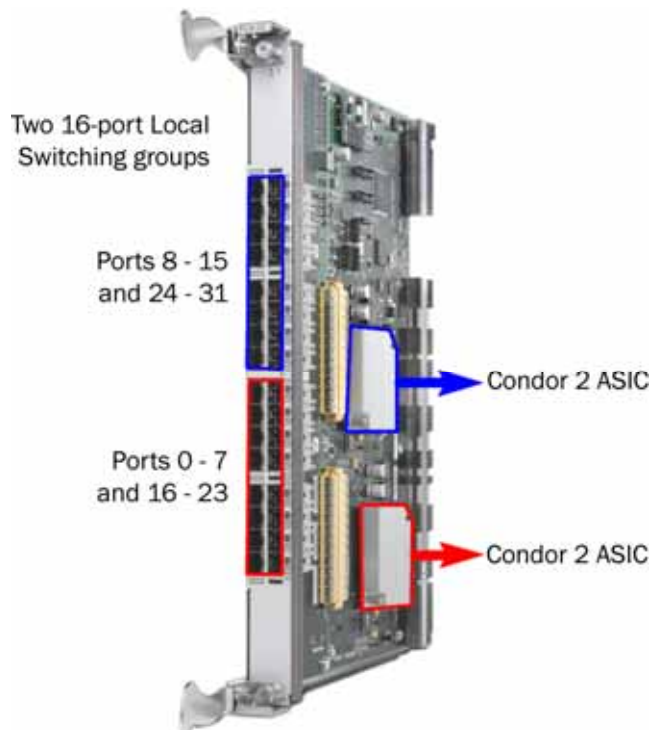


Figure 2. FC8-32

Brocade FC8-48 with Two Condor 2 ASICs

Ports 0 – 7, 24 – 39, and ports 8 – 23, 40 – 47 make up two Local Switching groups, each with 24 x 8 Gbit/sec external ports, and 128 Gbit/sec Brocade DCX or 32 Gbit/sec Brocade 48000 internal bandwidth.

- 1:1 fully subscribed *for all local traffic* on either platform
- Up to 1.5:1 oversubscribed for non-local traffic for Brocade DCX
- Up to 6:1 oversubscribed for non-local traffic for Brocade 48000

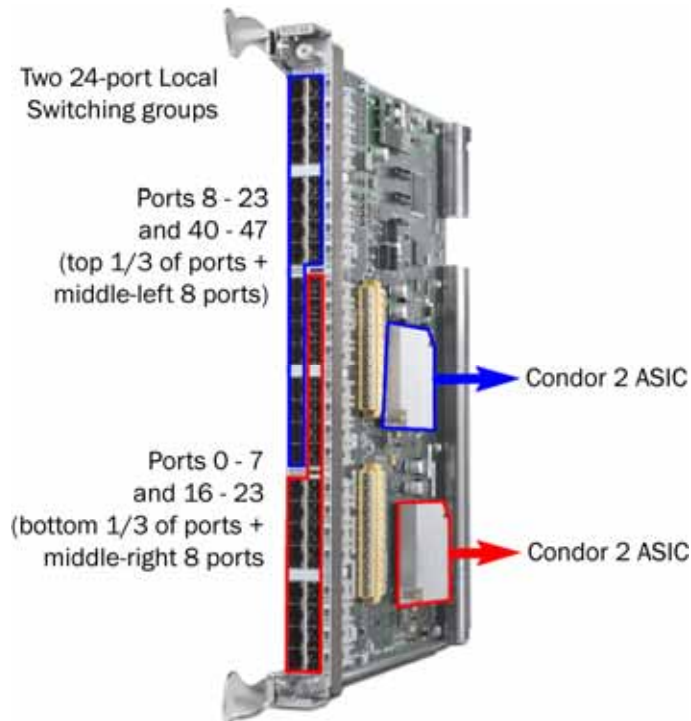


Figure 3. FC8-48

Combinations of local and non-local traffic flows can occur simultaneously. The oversubscription rates listed above are all preceded with the words "up to," since any Local Switching means that fewer ports use the limited bandwidth to the backplane.

- For example, for an FC8-48 in a Brocade DCX Backbone, if one third of the ports in each of the Local Switching groups switch locally, then the remaining two thirds of the ports can use the 256 Gbit/sec bandwidth to the backplane to go at full speed 8 Gbit/sec.
- Another example is an FC8-16 in a Brocade 48000, which can support eight 8 Gbit/sec ports switching non-locally to the backplane when the other eight ports on the FC8-16 are switching locally.
- Lastly, an FC8-48 in a Brocade 48000 Director requires 40 of the 48 ports on the FC8-48 to switch in their respective Local Switching groups, so that the remaining 8 ports can switch at full speed using the 64 Gbit/sec of bandwidth to the backplane. This represents an extreme amount of Local Switching, which would have to be carefully designed, unlike the FC8-48 in the Brocade DCX, which can frequently operate all ports at 8 Gbit/sec with little or no design beyond balancing connections across port blade ASICs.

OPTIMIZING WITH LOCAL SWITCHING

Even with no planning for Local Switching, the bursty, random nature of SAN traffic creates some traffic between ports in Local Switching groups. By performing a modest amount of connection planning, both Local Switching and High Availability (HA) can be maximized. Three methods of making connections to Brocade FOS-based platforms are described in this section.

HA Node Placement: Least Benefit

The easiest method to gain *some* Local Switching benefit is to follow the simple best practice of distributing storage, host, and ISL ports evenly across director port blades. This method minimizes the impact of losing a port blade to a hardware fault or other issue. Just performing this simple HA distribution of port types across blades increases the likelihood of Local Switching occurring.

Balancing Storage, ISLs, and Hosts by ASIC: Good Benefit

A slight modification of the HA method is to distribute storage, host, and ISL ports evenly across port blade ASICs (only for FC8-32 and FC8-48 port blades, which have two Condor 2 ASICs each). Distribution for each connection type should be across the first ASIC on each port blade first (for example, port 0 on each blade) and then the second ASIC on each port card (for example, port 31 or 47). By going across port blades and alternating between the high and low ASICs, Local Switching occurs much more frequently than with the HA node placement method.

Table 1. Balanced storage, ISLs, and hosts by ASIC: Brocade 48000 with three FC4-32 blades in slots 1 – 3

FC4-32	Slot 1		Slot 2		Slot 3	
	Ports 0 - 15	Ports 16 - 31	Ports 0 - 15	Ports 16 - 31	Ports 0 - 15	Ports 16 - 31
6 ISLs	P0, P1		P0, P1		P0, P1	
6 Tape	T8	T31	T8	T31	T8	T31
12 Disk	D9, D10	D29, D30	D9, D10	D29, D30	D9, D10	D29, D30
60 Hosts	H12 – H15	H16 – H21	H12 – H15	H16 – H21	H12 – H15	H16 – H21

In this example, there are 6 ISLs forming three 8Gbit/sec frame trunks, 6 tape ports (1 per ASIC), 12 disk ports (2 per ASIC), and 60 hosts (5 per ASIC). The ISLs are grouped 2 per ASIC over 3 ASICs instead of 1 per ASIC over 6 ASICs to allow the use of Brocade ISL Trunking. This layout, which has ISL trunks, tape, disk, and hosts evenly balanced across ASICs, is optimal for HA and increases the likelihood of Local Switching and increased performance. The ASIC method can be extended to the next method if performance demands require it.

Mapping Storage Allocations by ASIC: Maximum Benefit

The method requiring the most configuration effort pays off with ultra-high performance. First, follow the ASIC method listed above to connect the ports to the director, but purposely group storage, host, and ISL ports that will need to switch with each other within Local Switching groups. Then, ensure that array Logical Units (LUNs) are mapped to a Local Switching group that also has host ports that will require those LUNs. In this way, performance is always maximum, since all traffic is locally switched and no traffic goes over the backplane.

This type of port configuration and storage mapping is very common in high-performance mainframe environments with FICON. Open systems environments executing high-performance computing or other high-bandwidth applications can be configured this way to enable maximum performance of 384 simultaneously switched 8 Gbit/sec ports. Note that some ports can still be switched over the backplane at this maximum performance level—the Brocade DCX requires only one third of ports to be locally switched, while the Brocade 48000 needs five sixths of ports locally switched to support 384 ports at 8 Gbit/sec. This demonstrates the massive performance capabilities of the Brocade DCX Backbone and the excellent investment protection options offered by the Brocade 48000 Director.

SUMMARY AND RECOMMENDATIONS

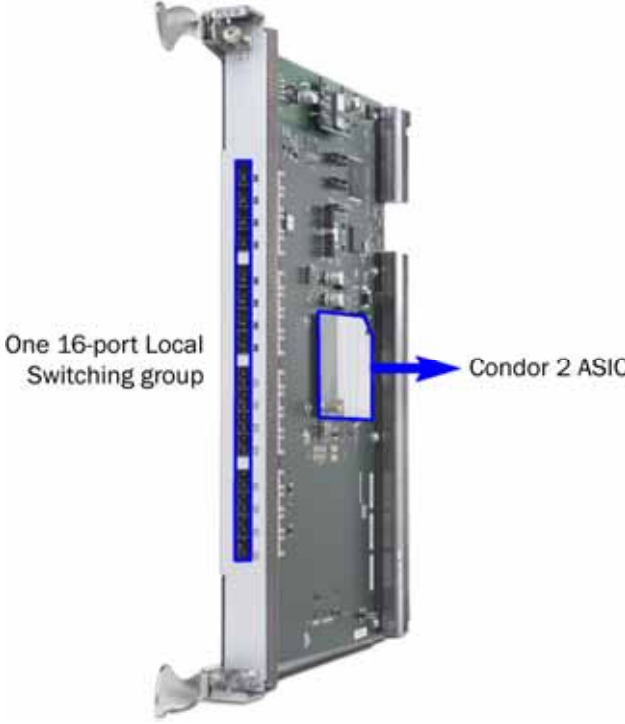
At a minimum, all Brocade FOS-based directors should lay out ports following the ASIC method, shown in Table 1. This is very easy to do, achieves enhanced Local Switching, and has optimal HA to deal with port blade or individual ASIC failures (which occur very rarely).

Using the Brocade DCX Backbone and Brocade 48000 Director with Local Switching, data centers architects can meet their current performance goals, optimize the longevity and Return on Investment (ROI) of their SAN directors, and optimally position the data center infrastructure for future data processing requirements. Brocade Local Switching can be easily implemented to ensure that chassis bandwidth is the highest available for any SAN platform. And following simple connection placement guidelines helps achieve even greater performance from Local Switching.

APPENDIX: PORT BLADES IN THE BROCADE DCX AND 48000

This appendix details how the three port blades discussed in this paper function in a Brocade DCX Backbone and a Brocade 48000 Director.

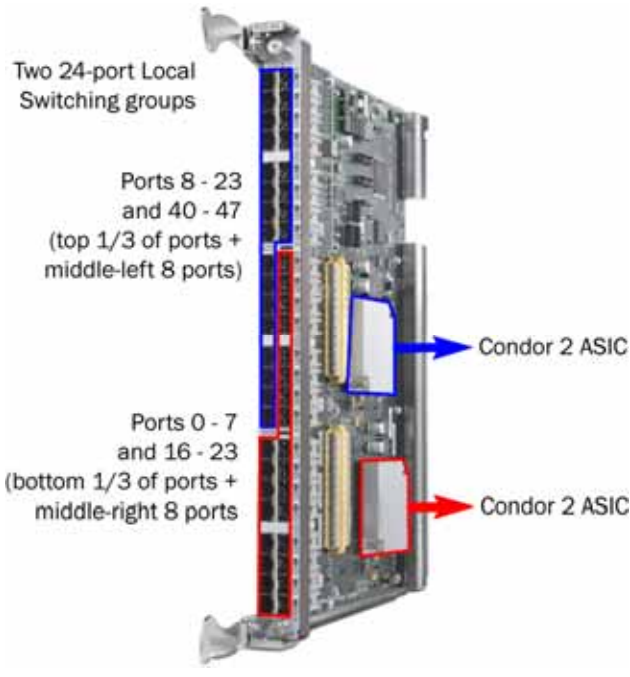
FC8-16 or FC4-16 Port Blade

<p>FC8-16 in Brocade DCX</p>		<p>FC4-16 in Brocade 48000</p>
<p>Port ASIC bandwidth to the backplane: 128 Gbit/sec full duplex (blue arrow)</p>		<p>Port ASIC bandwidth to the backplane: 64 Gbit/sec full duplex (blue arrow)</p>
<p>16 x 8 Gbit/sec user ports per local switching group (blue)</p>		<p>16 x 4 Gbit/sec user ports per Local Switching group (blue)</p>
<p>All ports simultaneously support 8 Gbit/sec for any combination of local or non-local data.</p>		<p>All ports simultaneously support 4 Gbit/sec for any combination of local or non-local data flows.</p>

FC8-32 or FC4-32 Port Blade

<p>FC8-32 in Brocade DCX</p>	<p>Two 16-port Local Switching groups</p> <p>Ports 8 - 15 and 24 - 31</p> <p>Ports 0 - 7 and 16 - 23</p> <p>Condor 2 ASIC</p> <p>Condor 2 ASIC</p>	<p>FC4-32 in Brocade 48000</p>
<p>Port ASIC bandwidth to the backplane: 128 Gbit/sec full duplex (blue and red arrows)</p>		<p>Port ASIC bandwidth to the backplane: 32 Gbit/sec full duplex (blue and red arrows)</p>
<p>16 x 8Gbit/sec user ports per Local Switching group (blue and red)</p>		<p>16 x 4 Gbit/sec user ports per Local Switching group (blue and red)</p>
<p>All 16 ports simultaneously support 8 Gbit/sec for any combination of local or non-local data flows for each Local Switching group.</p>		<p>All ports 16 simultaneously support 4 Gbit/sec if 32 Gbit/sec of traffic is switched in the Local Switching group.</p>
		<p>If no traffic is local (that is, the backplane is used exclusively), all ports in each Local Switching group simultaneously support 2 Gbit/sec.</p>

FC8-48 or FC4-48 Port Blade

FC8-48 in Brocade DCX	 <p>Two 24-port Local Switching groups</p> <p>Ports 8 - 23 and 40 - 47 (top 1/3 of ports + middle-left 8 ports)</p> <p>Ports 0 - 7 and 16 - 23 (bottom 1/3 of ports + middle-right 8 ports)</p> <p>Condor 2 ASIC</p> <p>Condor 2 ASIC</p>	FC4-48 in Brocade 48000
Port ASIC bandwidth to the backplane: 128 Gbit/sec full duplex (blue and red arrows)		Port ASIC bandwidth to the backplane: 32 Gbit/sec full duplex (blue and red arrows)
24 x 8 Gbit/sec user ports per Local Switching group (blue and red)		24 x 4 Gbit/sec user ports per Local Switching group (blue and red)
All 24 ports simultaneously support 8 Gbit/sec if 64 Gbit/sec or more is switched in each Local Switching group.		All 24 ports simultaneously support 4 Gbit/sec if 64 Gbit/sec or more is switched in each Local Switching group.
If no traffic is local (that is, the backplane is exclusively used), all ports in each Local Switching group simultaneously support 5.33 Gbit/sec.		If no traffic is local (that is, the backplane is exclusively used), all ports in each Local Switching group simultaneously support 1.33 Gbit/sec.

© 2008 Brocade Communications Systems, Inc. All Rights Reserved. 07/08 GA-TB-073-00

Brocade, the Brocade B-weave logo, Fabric OS, File Lifecycle Manager, MyView, SilkWorm, and StorageX are registered trademarks and the Brocade B-wing symbol, SAN Health, and Tapestry are trademarks of Brocade Communications Systems, Inc., in the United States and/or in other countries. FICON is a registered trademark of IBM Corporation in the U.S. and other countries. All other brands, products, or service names are or may be trademarks or service marks of, and are used to identify, products or services of their respective owners.

Notice: This document is for informational purposes only and does not set forth any warranty, expressed or implied, concerning any equipment, equipment feature, or service offered or to be offered by Brocade. Brocade reserves the right to make changes to this document at any time, without notice, and assumes no responsibility for its use. This informational document describes features that may not be currently available. Contact a Brocade sales office for information on feature and product availability. Export of technical data contained in this document may require an export license from the United States government.