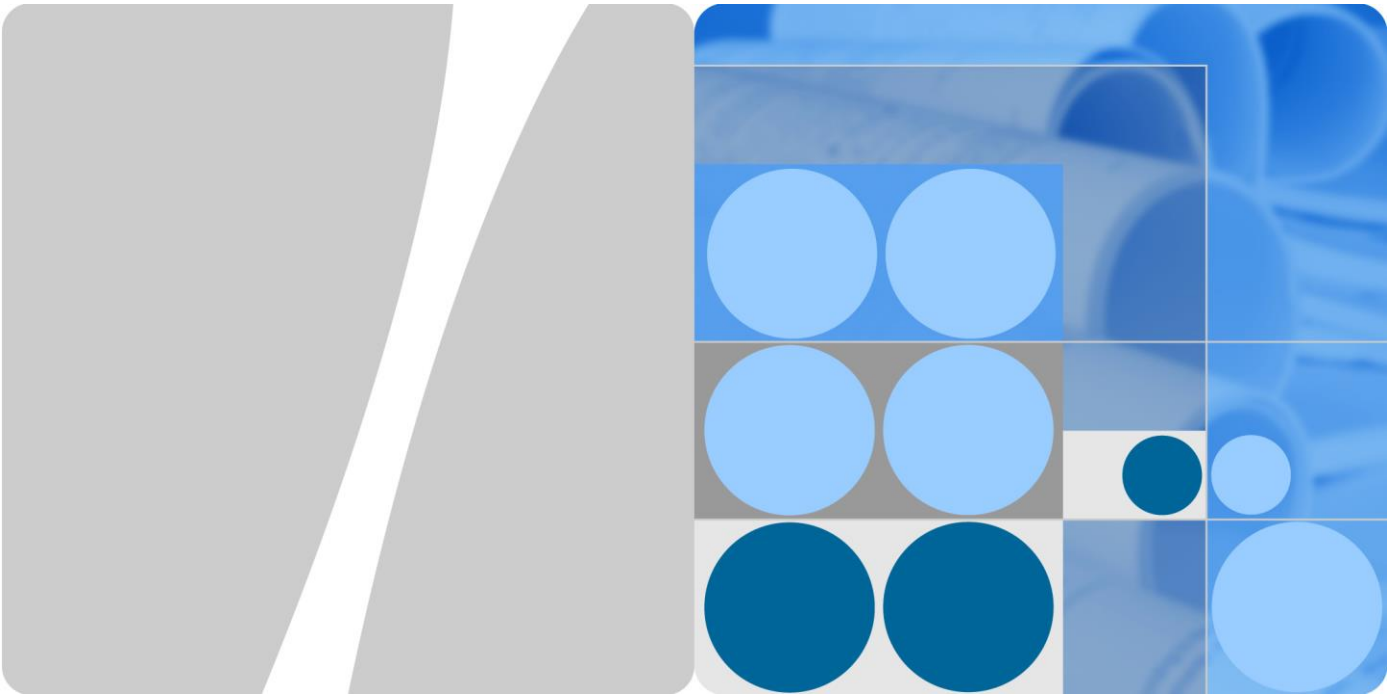


Huawei OceanStor SNS2624, SNS3664, SNS3696E, SNS5604, SNS5608 V200R001

Technical White Paper

Issue 1.0

Date 2018-07-05

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions


and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: http://www.huawei.com

Email: support@huawei.com

Contents

1 Introduction

Huawei-produced OceanStor SNS V200R001 series Fiber Channel switches (the SNS V2R1 series for short) are composed of 5 models: OceanStor SNS2624, OceanStor SNS3664, OceanStor SNS3696E, OceanStor SNS5604, and OceanStor SNS5608. OceanStor SNS2624, OceanStor SNS3664 and OceanStor SNS3696E are fixed switches while OceanStor SNS5604 and OceanStor SNS5608 are modular guides.

The SNS V200R001 series adopt the Gen 6 Fibre Channel technology to deliver integrated network sensors with industry-leading performance and enhanced service flexibility, which quickens data access. The SNS V200R001 series are capable of meeting ever-changing requirements on service continuity, hyperscale virtualization, at-scale cloud infrastructure, and flash storage environment. The Gen 6 ASIC adopted by the SNS V200R001 series is the most powerful and efficient Fibre Channel switch technology in the industry that provides 32 Gbit/s and 128 Gbit/s speeds. The SNS V200R001 series also feature higher performance, excellent I/O performance, and various functions such as Fabric Vision, IO Insight, VM Insight, ClearLink Diagnostics, online data compression, and forward error correction (FEC).

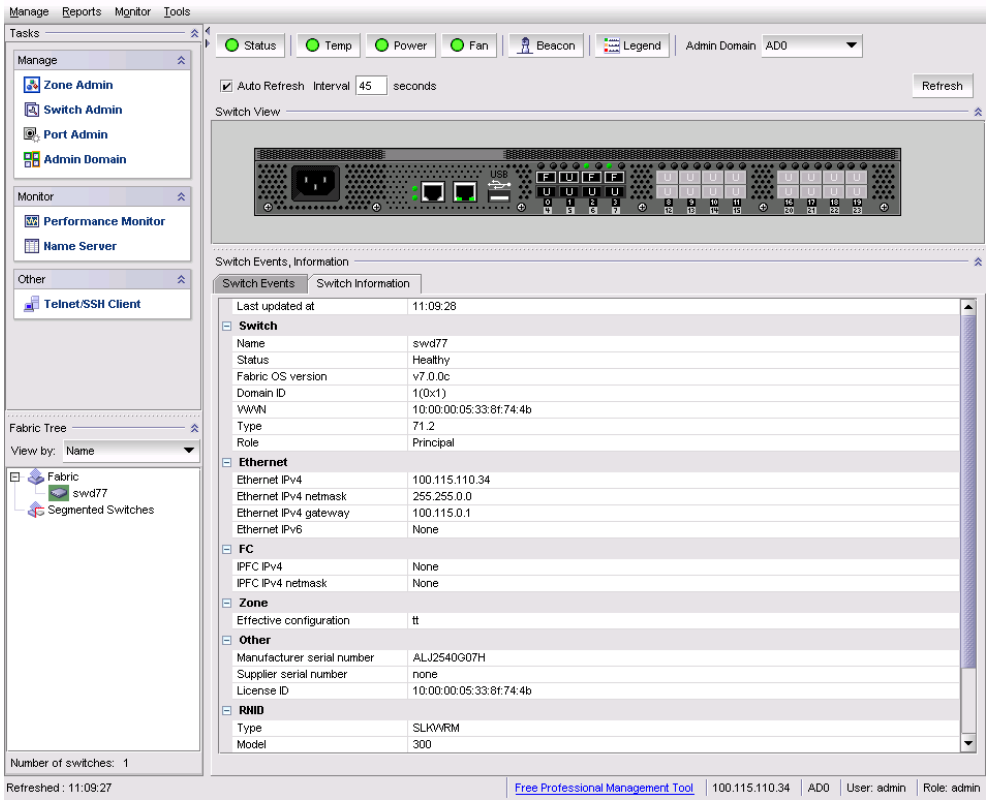
The SNS V200R001 series integrate reliability, availability, and serviceability (RAS) features, comply with scalable requirements of enterprise switches, and are interoperable and easy to use not only for small storage scenarios, but also for medium-and large-sized enterprise data centers. The storage environment is facing challenges brought by digital transformation because customers desire to access data anytime, anywhere. The increasing data volume requires evolution of networks to facilitate enterprise development. The SNS V200R001 series featuring excellent availability, optimum performance, and good scalability ensure data consistency and predictability, and provide high processing performance.

2 Solution

2.1 Enhanced Management Software

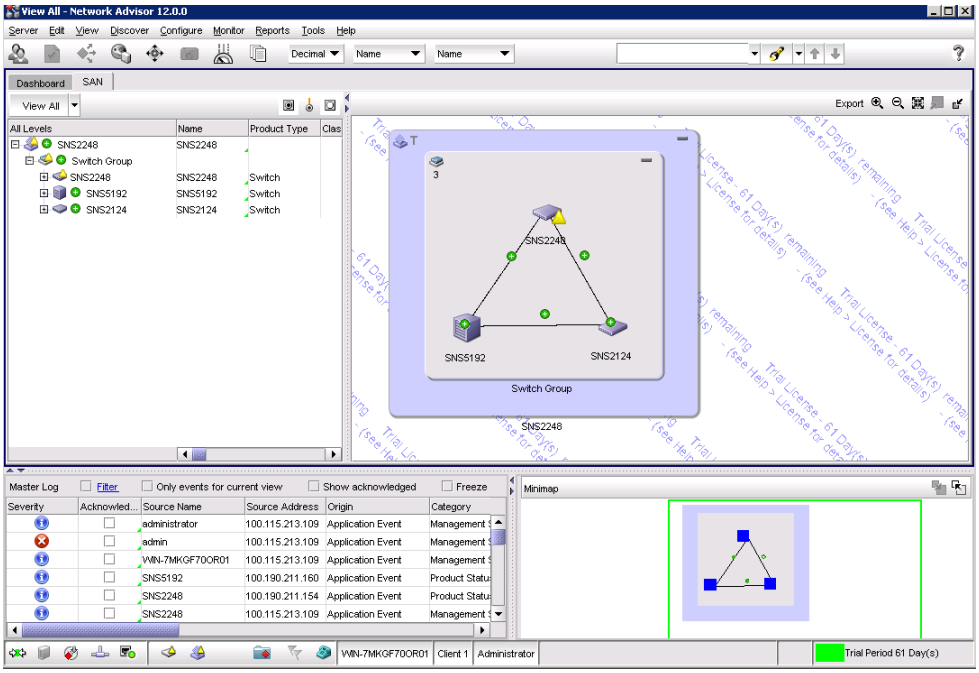
1. You can monitor and manage Fabric OS (FOS) switches using the web-based Web Tools without installing it. Web Tools provides a series of functions for switch task management, including switch firmware upgrade, license management, route policy selection, ISL trunking management, and long-distance transmission management.

Figure 2-1 Management interface of Web Tools



2. You can manage the FOS switch using the CLI.
3. Brocade Network Advisor (BNA) that provied by Brocade helps organizations manage the entire network lifecycle, including monitoring, diagnostics, change management, troubleshooting, and remediation. BNA provides a smart control panel to quickly detect and troubleshoot faults and delivers capabilities such as real-time and historical performance data reporting, visible VM management, and advanced Call Home, etc.

Figure 2-2 Management interface of BNA

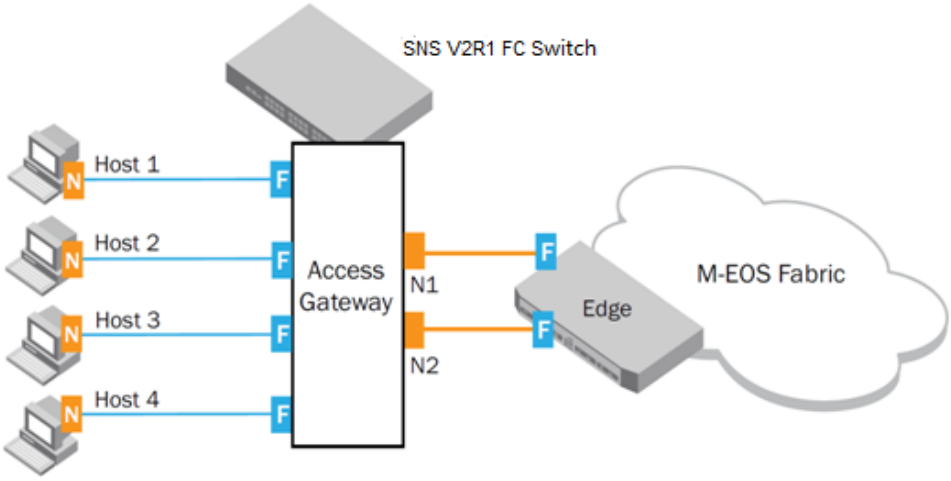


2.2 Access Gateway Mode (SNS2624 and SNS3664)

The Access Gateway (AG) mode of the SNS V200R001 series is implemented by the N_Port ID Virtualization (NPIV) technology. Switches produced by different vendors cannot be cascaded through the Expansion port (E_Port). This is why the AG mode appears. The SNS V200R001 series can be connected to other switches in AG mode, which technically avoids E_Port cascading. The AG mode ensures connection flexibility of SAN storage devices.

1. After configured to the AG mode, a switch is not connected to other switches in the fabric through the E_Port, but accesses the fabric through the N_Port. In AG mode, the switch connects to a device or director that supports the NPIV technology. The AG mode ensures connection flexibility of the SAN.
2. In AG mode, a switch presents an N_Port not an E_Port to the fabric. Thus other switches in the fabric simply deem the AG-mode switch as a node device not a switch. Additionally, adding a new switch in AG mode means that several ports are added. No domain ID assignment needs to be considered, which simplifies fabric management.

Figure 2-3 SNS switches in AG mode



2.3 ClearLink Diagnostic Port (D_Port)

OceanStor SNS V200R001 series Fiber Channel switches supply D_Port which fulfills ClearLink Diagnostics to ensure optical and signal integrity for Gen 5 and Gen 6 Fibre Channel optics and cables, simplifying deployment and support of high-performance fabrics. By actively verifying the integrity of the key transceiver, enterprises are capable of quickly addressing any physical layer problem without employing special optical test instruments.

ClearLink Diagnostics of OceanStor SNS V200R001 series enables users to conduct automatic tests before deployment or when any physical layer problem occurs to check and verify switch link latency and distance, thereby verifying the integrity of 32 Gbit/s or 16 Gbit/s transceiver. ClearLink Diagnostics frees administrators from repeated and time-consuming tests for fault locating and is not in demand of any access optical module or cable. In addition, you only need to disable the ports that connect to the link to be tested while other ports work properly online.

In addition to switch link verification, the SNS V200R001 series also collaborate with the Gen 5 and Gen 6 Fibre Channel adapters of QLogic and Emulex to enable ClearLink Diagnostics for switches of the FOS 8.1.0b and later version.

D_Port test consists of four steps:

1. Electrical loopback test (E-WRAP)
2. Optical loopback test (O-WRAP)
3. Link traffic test
4. Link latency, distance, and power measurement

Figure 2-4 D_Port

2.4 Virtual Fabric (Excluding SNS2624)

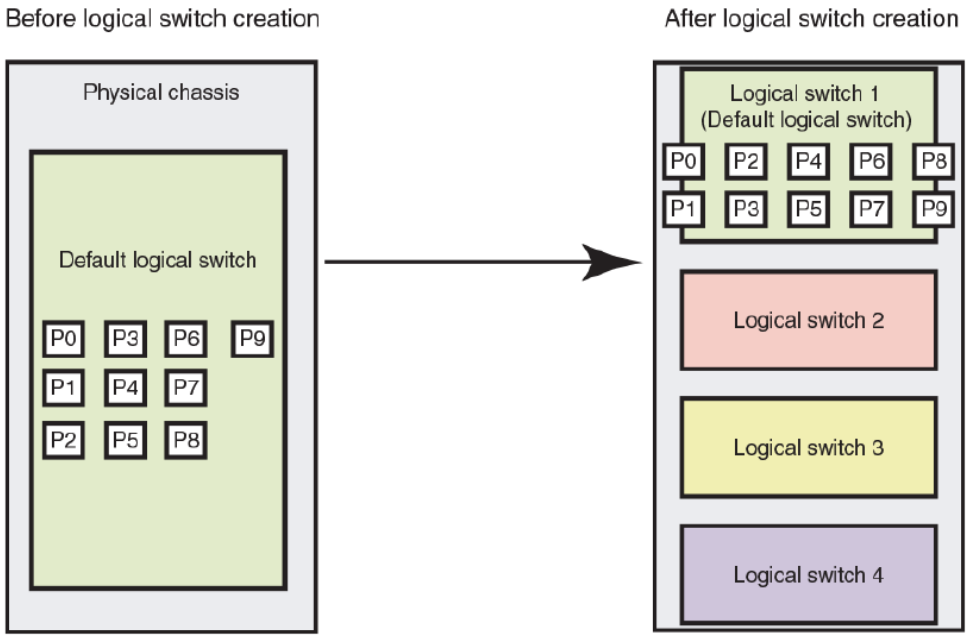
Virtual Fabric is a hardware virtualization architecture. Conventionally, storage networks are designed and managed on the physical switch layer. However, Virtual Fabric allows storage network design and management to be implemented on the port layer. The SNS3664, SNS3696E, SNS5604, and SNS5608 switches support the Virtual Fabric function.

Specifically, Virtual Fabric consists of three functions: logical switch, logical fabric, and device sharing, which can be customized on demand.

1. Logical switch

   Logical switches are basic components of Virtual Fabric. After enabling the Virtual Fabric function, a physical switch can be divided into multiple logical switches. Furthermore, ports on the physical switch can be dynamically assigned to any logical switches. Before logical switches are created, the physical switch with the Virtual Fabric function enabled is deemed as a logical switch by default, as shown in Figure 2-5. A default logical switch cannot be deleted.

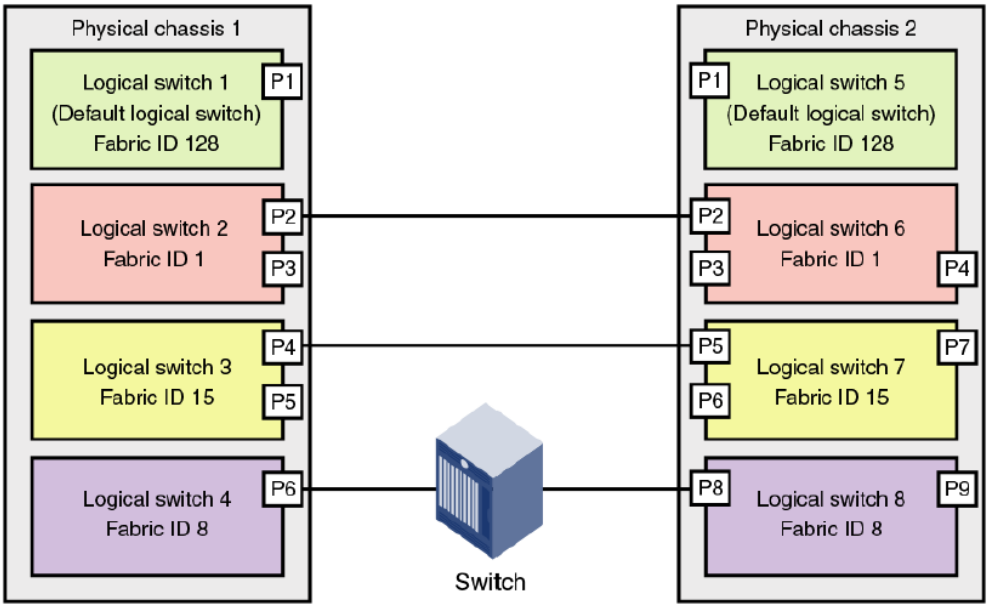
Figure 2-5 Physical switches and logical switches



2. Logical fabric

   A logical fabric contains at least one logical switch. The fabric ID (FID) is an ID of a logical switch in a specified logical fabric. If the logical switches of different physical switches have the same FID, they can be assigned to the same logical fabric. Logical switches in the same logical fabric support ISL cascading, frame aggregation, and dynamic path selection (DPS).Figure 2-6 shows that logical switches whose FIDs are 15 on physical chassis 1 and physical chassis 2 constitute a logical fabric through the physical ISL link. Logical switches whose FIDs are 8 and a physical switch constitute a logical fabric through the physical ISL link.

Figure 2-6 Logical fabrics and logical switches

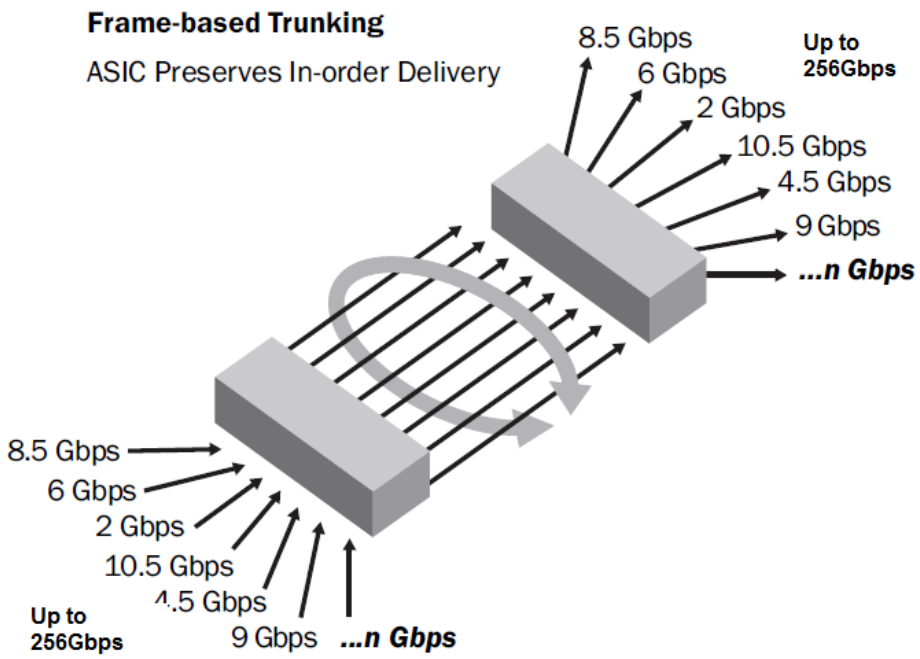


Virtual Fabric can specify physical switches as independent logical switches based on ports, optimize data transmission and management, and enable fault isolation. It also supports enhanced FOS features such as frame aggregation, DPS, Fibre Channel routing, AG, and FCIP.

2.5 Frame-based Trunking

The trunking function of the SNS V200R001 series combines a group of physical links into a single logical link to optimize bandwidth usage. Traffic is dynamically and sequentially assigned in a trunking group so that fewer links can deliver higher performance, as shown in Figure 2-7. In a trunking group, multiple physical ports serve as a single port to simplify management. The trunking function ensures system reliability by transmitting data in order and avoiding I/O retries after a link breakdown. The trunking technology is based on frame rather than switch. The frame-based trunking technology of the SNS V200R001 series is more precise and balanced than the switch-based trunking technology, and ensures full link utilization.

Figure 2-7 Frame-based trunking



The trunking function of the SNS V200R001 series consists of ISL trunking (E_Port trunking), EX_Port trunking, F_Port trunking, and N_Port trunking, which all require a trunking license.

ISL trunking aggregating cascading links between switches are generally used.

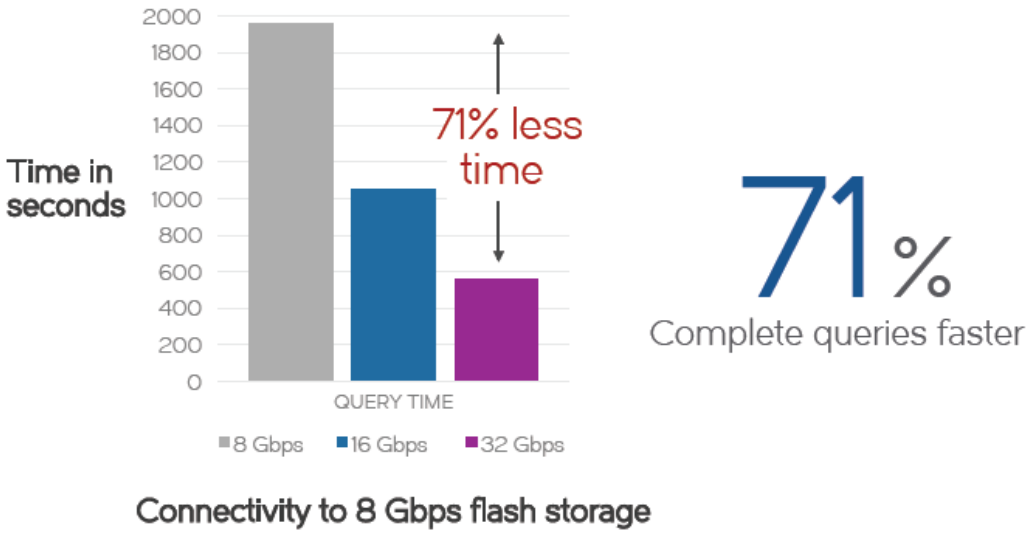
ICL trunking does not require a trunking license.

2.6 32 Gbit/s

The OceanStor SNS V200R001 series provide the 32 Gbit/s or 128 Gbit/s Fibre Channel to meet the demand of high performance and exploit the potential of all-flash technology. Enterprises require flash storage technologies that can fulfill their demands of balancing the increasing workload. The conventional infrastructure cannot meet the dynamic and versatile workload requirements and the aged network blocks the performance of the all-flash data center. According to Demartek test, when you conduct database application tests using the all-flash array at the speed of 8 Gbit/s, you do not need to change the storage system but only upgrade the network to the Gen 6 Fibre Channel featuring a speed of 32 Gbit/s, as shown in Figure 2-8. Compared with the 8 Gbit/s network, the 32 Gbit/s network saves 71% of time.

Thus the 32 Gbit/s switch not only increases the transmission speed, but also optimizes the all-flash array performance.

Figure 2-8 32 Gbit/s quickening the 8 Gbit/s all-flash storage



2.7 Long-Distance Transmission

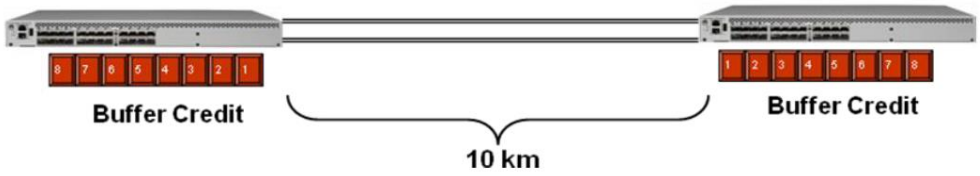
According to Fiber Channel protocol standard definitions, Fiber Channel ports on any end of the extended connection must support high-level buffer credits for long-distance Fiber Channel signal transmission.

A buffer credit allows the transmit (TX) to send a data frame. The TX does not send the next data frame until receiving an acknowledgement from the receive (RX) indicating that the data frame has arrived. There is an acknowledgement latency during this process. The latency increases as the distance elongates, and the bandwidth available for users decreases. Therefore, the key to Fibre Channel transmission over a long distance is to provide sufficient buffer credits for the TX. This approach enables the TX to continue sending data frames over a Fibre Channel while waiting for acknowledgement information.

The optimal number of buffer credits is determined by the distance (frame transmission latency), frame sending latency, link signaling frequency, and frame size. As the link speed increases, the frame transmission latency is reduced and the number of buffer credits must be increased to obtain full link utilization. Excessive buffer credits also easily cause a data frame congestion at the RX. Buffer credits are allocated from a common buffer pool to a group of ports on a switch. The buffer credit value can be changed for specific applications and service scenarios.

The OceanStor SNS V200R001 series support four long-distance connection modes, including Normal Mode (L0), Extended Mode (LE), Dynamic Mode (LD), and Static Mode (LS). If the distance is over 10 km, you need to configure Extend Fabric.

Figure 2-9 Buffer Credit



2.8 Fabric Vision

Fabric Vision technology, an integration of Gen 5 and Gen 6 Fibre Channel ASIC, FOS, and BNA, provides unprecedented visibility and insight across the storage network. Offering innovative diagnostic, monitoring, and management capabilities, Fabric Vision technology helps administrators avoid risks, facilitates new application deployment, and drastically reduces operational costs.

Fabric Vision technology, an extension of Gen 6 Fibre Channel, provides the following leading technologies in the industry:

Monitoring and Alerting Policy Suite (MAPS): provides a policy-based monitoring and alerting suite to actively monitor the health status and performance of SAN infrastructure, ensuring normal running and high availability of applications.

Flow Vision: provides tools with powerful functions, which help administrators identify, monitor, and analyze specific applications and data flows to maximize performance, avoid congestion, and improve resource utilization. Flow Vision provides three features, including Flow Monitor, Flow Generator, and Flow Mirror.

IO Insight: provides proactive, smooth, and real-time I/O monitoring, helping users understand storage performance and behaviors more comprehensively to ensure stable system running and quality service level.

Fabric Performance Impact (FPI) monitoring: detects and reports severity levels of latencies based on the preset MAPS policies, and finds out "slow drain" devices that affect network performance. This feature enables users to precisely locate devices that have bottlenecks or are affected by bottleneck ports, and automatically isolates slow drain devices to prevent buffer credits from being used up.

Configuration and Operational Monitoring Policy Automation Services Suite (COMPASS): exploits automated switches and fabric architecture to simplify deployment, ensure consistency, and improve running performance of at-scale scenarios. Administrators can configure a template or make a template with existing configurations to seamlessly deploy services on the fabric architecture. In addition, with the COMPASS setting and policy on BNA, configurations will not change as the time goes.

Control panel: displays switch status and causes that affect switch running status, enabling administrators to quickly locate the fault on any switch hot spot and troubleshoot the fault.

ClearLink Diagnostics: ensures the integrity of optical signal of the Gen 5 and Gen 6 Fibre Channel optical modules and cables, and simplifies the deployment and support of the high-performance fabric architecture. Non-SNS devices require the Fabric Vision technical license.

Forward Error Correction (FEC): recovers Gen 5 link errors to improve transmission reliability and performance. FEC is a must for the Gen 6 link to support the 32 Gbit/s speed.

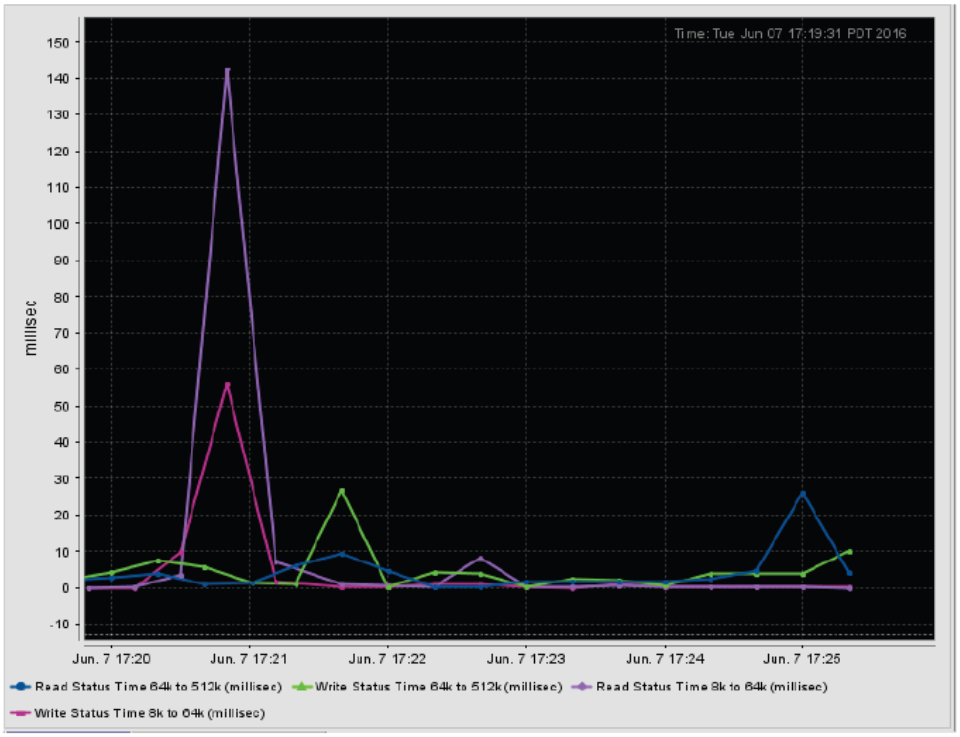
Credit loss recovery: detects and recovers from virtual channel (VC)–level buffer credit loss to prevent performance deterioration and improve application availability.

2.9 IO Insight(Excluding SNS2624) and VM Insight

IO Insight integrates the Gen 6 Fibre Channel ASIC and network sensor. It provides enterprises with full control and insight in fault locating, shortens the troubleshooting duration, and meets the requirements of the service level agreement (SLA). IO Insight can smoothly collect I/O statistics through any device port on the Gen 6 Fibre Channel platform, and use these statistics to configure thresholds and alarms on the policy-based monitoring and warning suite. Monitoring for application and device latency and IOPS helps detect performance deterioration, which enables administrators to actively optimize the performance and availability. The following functions are provided:

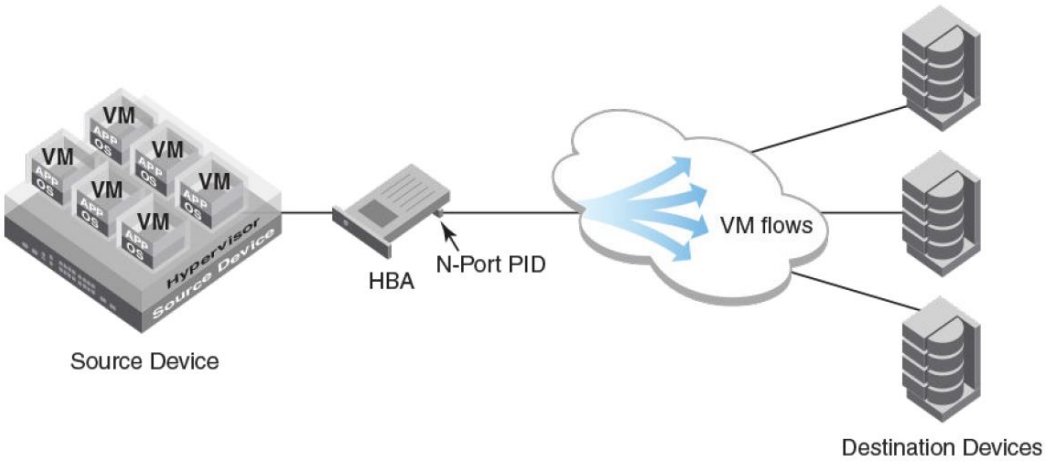
1. Provides proactive, smooth, and real-time I/O monitoring and alerting, and visible health status and performance of storage I/Os.
2. Provides monitoring for storage devices, helping users understand network performance comprehensively and meeting SLA requirements.
3. Collects statistics such as the total I/Os of a specific host or storage device, the maximum or average first response time, the maximum or average I/O latency (ECT), the maximum or average I/O to be processed, and the process performance, to diagnose I/O running issues.

Figure 2-10 Real-time performance data curve of IO Insight on BNA



VM Insight helps collect performance statistics of Fibre Channel I/O and SCSI I/O and I/O latency from VMs using a single N_Port ID (PID) to destination devices. Collecting VM flows enables the isolation of VMs whose performance is deteriorated, helps migrate VMs to implement load balancing, or provides reference for other purposes.

Figure 2-11 Monitoring VM flows



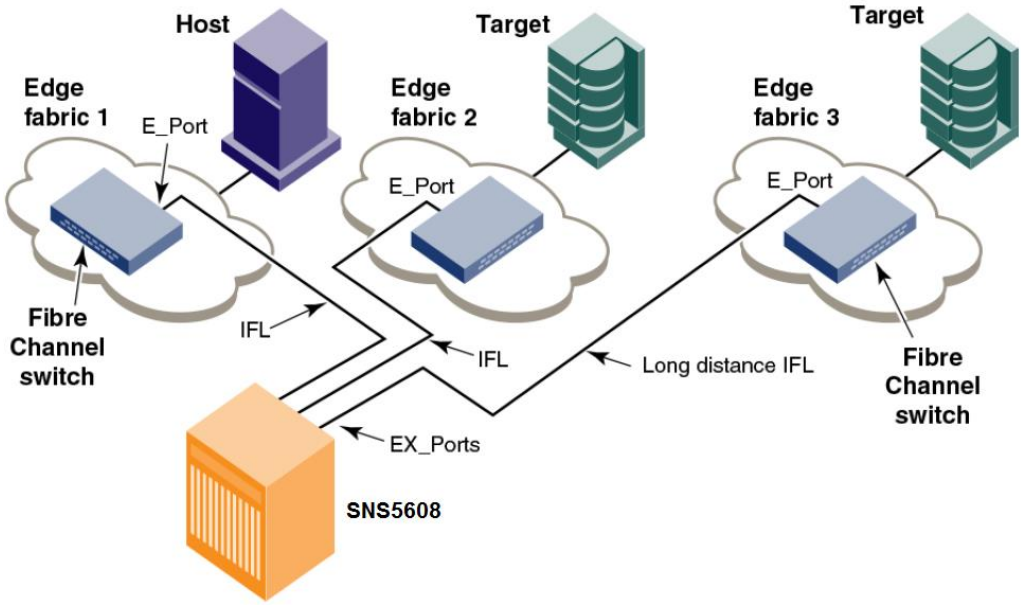
2.10 FC-FC Routing (Excluding SNS2624)

The OceanStor SNS V200R001 series provide the FC-FC Routing service to enable Fibre Channel routes between two or multiple fabrics, in which the fabrics are not merged. FC-FC Routing is layer 3 routing.

With FC-FC Routing, tape drives can be shared among fabrics without any management problems, such as change management, network management, scalability, reliability, availability, and maintainability brought by fabric merging. Figure 2-12 shows a metaSAN architecture (a network that uses FC-FC Routing), in which three edge fabrics and SNS5608 are connected by Inter-fabric links (IFL). The three edge fabrics are independent from each other, but their storage space can be shared by using the FC-FC Routing function of SNS5608.

FC-FC Routing applies to scenarios, in which multiple SANs need to be integrated, the remote disaster recovery network needs to be isolated, or the limited access to independent SANs needs to be implemented.

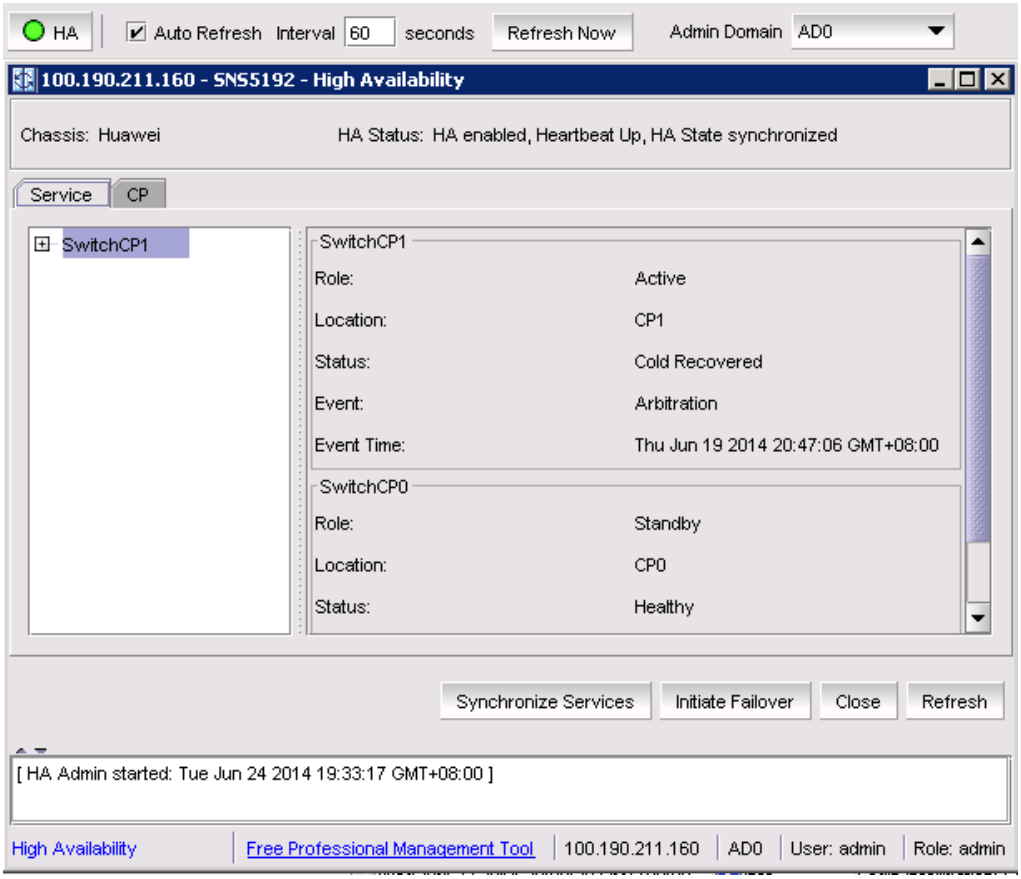
Figure 2-12 A metaSAN with IFL



2.11 High Availability (Only Support SNS5604 and SNS5608)

The High Availability (HA) feature provides maximum reliability and non-disruptive replacement of key hardware and software modules. HA is available only on SNS5604 and SNS56084. HA supplies redundant active/passive control processor, redundant active/active core switching blades, and redundant WWN cards, which are designed to provide 99.999% uptime capabilities.

Figure 2-13 HA



2.12 UltraScale ICL (Only Support SNS5604 and SNS5608)

Gen 5 director SNS5384 and SNS5192 and Gen 6 director SNS5608 and SNS5604 can provide a maximum of 32 quad small form-factor pluggable (QSFP) ports in chassis with 8 slots and a maximum of 16 QSFP ports in chassis with 4 slots to help save switch ports on terminal devices. Each QSFP port has four independent 16 Gbit/s or 32 Gbit/s links that can provide a total bandwidth of 64 Gbit/s or 128 Gbit/s.

The OceanStor SNS V200R001 series provide a high-density solution with the UltraScale ICL technology which uses the 128 Gbit/s link to double the bandwidth. As a result, a flatter, faster, and easier fabric architecture is built to enhance convergence while reducing network complexity and costs.

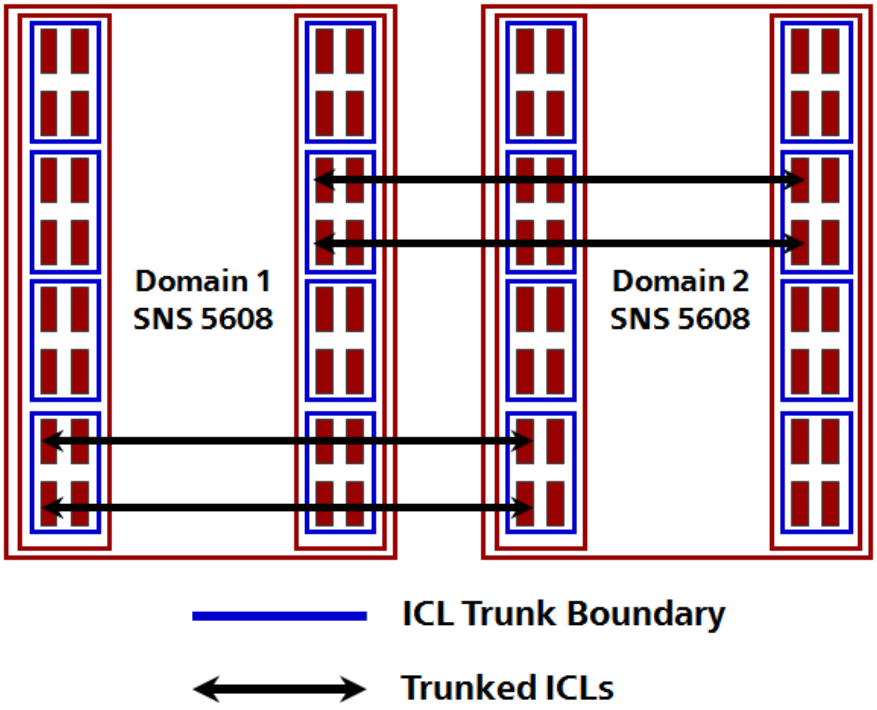
The OceanStor SNS V200R001 series provide UltraScale ICL ports that support standard optical cables of 100 m at the Gen 6 speed (4 x 32 Gbit/s) and optical cables of 2 km at the Gen 5 speed (4 x Gbit/s), enabling long-distance data transmission. Gen 6 UltraScale ICL ports are compatible with Gen 5 ICL ports, which helps enterprise fully use their investments. UltraScale ICL allows a maximum of 12 chassis in a core-edge topology and a maximum of 9 chassis in an active-active mesh topology. These high-density chassis topologies reduce 75% of cables between switches and save 33% of ports that can be used for connecting servers and storage devices. As a result, the port density in a minimum chassis is maximized and the ports being saved can be used to connect to servers and storage devices. To ensure high reliability

of ICL, each core blade must have two ICL links at least. In this case, two directors which are connected to each other have four ICL links at least, as shown in Figure 2-14.

ICL links can connect chassis of Gen 6 director SNS5604 and SNS5608 and Gen 5 director SNS5192 and SNS5384.

ICL link supports the trunking function. Thus a trunking license is not required.

Figure 2-14 Minimum configuration of 128 Gbit/s ICL link (SNS5608)



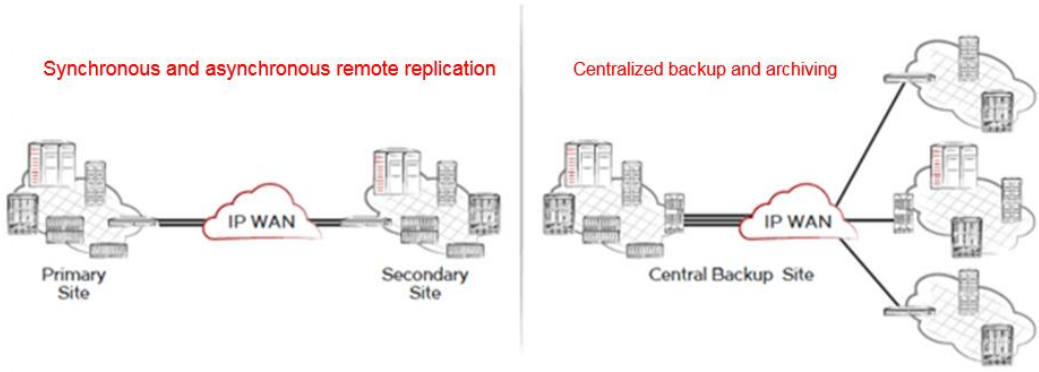
2.13 FCIP/SX6 Extended Blades (Only Support SNS5604 and SNS5608)

SNS SX6 extended blade is a multi-protocol extension solution that can move more data securely and support the most severe environment in the world. With Gen 6 Fibre Channel, IP expansion, Fabric Vision, the platform can provide SNS5604 and SNS5608 with optimum performance, high security, continuous availability, high scalability, and simplified operations. The extension solution can meet the requirements of growing data flows in data centers of Fibre Channel-, FICON-, and IP-based storage environments.

Each chassis of SNS5604 or SNS5608 supports a maximum of four SX6 extended blades. Each SX6 extended blade provides 16 32 Gbit/s Fibre Channel or FICON ports, 16 GE or GbE ports, and two 40 GbE ports. As a result, a high bandwidth, port density, and throughput are provided for WAN connection to obtain optimal performance and meet the requirements of the most demanding disaster recovery.

SX6 extended blades can interconnect with Brocade7840 FCIP switches.

Figure 2-15 SX extended blades extending Fibre Channel at a long distance and conducting IP-based storage replication for disaster recovery and data protection

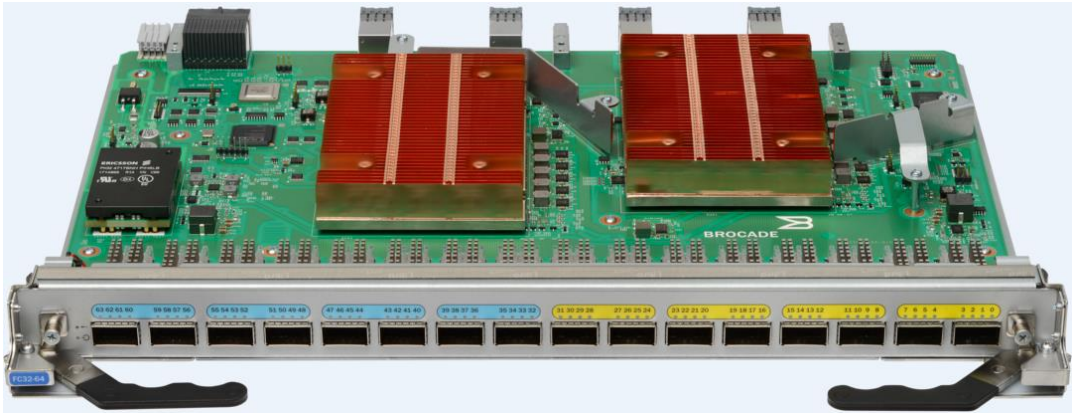


2.14 FC32-64 Port blade(SNS5604/SNS5608)

With the Huawei FC32-64 high-density port blade, the Huawei SNS5608 can now scale even further up to 512 ports and SNS5604 can now scale even further up to 256 ports, both of two models could offering 33 percent more device ports, while increasing total system bandwidth up to 20 Tbps and 10 Tbps. This blade provides 64 Fibre Channel ports in an elegant, highdensity form factor designed with Q-Flex connections, enabling administrators to simplify cabling infrastructure. As a result, organizations can effectively scale to build high-density fabrics to meet data growth demands, handle more workloads, and drive efficiency by maximizing rack space.

Combining Huawei's robust Gen 6 capabilities with multiprotocol flexibility and enhanced latency monitoring of NVMe over Fibre Channel traffic, the Brocade FC32-64 Port Blade enables a seamless transition to the all-flash data center and drives IT innovation.

Figure1 FC32-64 Port blade



3 Acronyms and Abbreviations

Table 3-1 OceanStor SNS acronyms and abbreviations

Acronym and Abbreviation	Definition
Activity LED	A port LED that indicates when frames are entering or leaving the port
AL_PA	Arbitrated Loop Physical Address, a unique one-byte value assigned during loop initialization to each NL_Port on a loop
Alarm	A message generated by the switch that specifically requests attention. Severity of some alarms can be configured.
Alias	A named set of ports or devices. An alias is not a zone, and cannot have a zone or another alias as a member.
Arbitrated Loop	A fabric topology where ports use arbitration algorithm to establish a point-to-point loop
Arbitrated Loop Physical Address (AL_PA)	A unique one-byte value assigned during loop initialization to each NL_Port on a loop
ASIC	Application Specific Integrated Circuit, a chip designed for a specific application
Auto Save	Parameter that determines whether changes to the active zone set that a switch receives from other switches in the fabric will be saved to the non-volatile storage medium on that switch
BootP	A type of network server
Buffer Credit	Flow control on the FC protocol layer. The source and destination set the number of unacknowledged frames (buffer credits) allowed to accumulate before the source

Acronym and Abbreviation	Definition
	stops sending data.
Cascade	A serial connection
Cascade Topology	A Fibre Channel network in which switches are connected in series. If you connect the last switch back to the first switch, you create a cascade-with-a-loop topology.
Class 2 Service	A unicast transmission service
Class 3 Service	A unicast transmission service
CLI	Command Line Interface
Default Visibility	Zoning parameter that determines the level of communication among ports/devices when there is no active zone set
Domain ID	User-defined number that identifies the switch in the fabric
Effective configuration	Zone set that defines the current zoning for the fabric
E_Port	Cascading port that is used to connect to other switches through ISL
Event Log	Logs that describe events occurring in the fabric
F_Port	Port on a switch being used to connect two N_Ports
Fabric Management Switch	Switch through which the fabric is managed
Fabric Name	User-defined name associated with the file that contains user list of the fabric
Failover	Automatically switching control from one CPU to another due to an error condition. The alternative is switchover.
FDMI	Fabric Device Management Interface
FL_Port	Fabric Loop Port
Flash Memory	Memory on the switch that contains the chassis control firmware
Frame	Minimum data transmission unit on the Fibre Channel protocol layer which consists of a start-of-frame (SOF) delimiter, header, data payload, CRC, and an end-of-frame (EOF) delimiter
FRU	Field Replaceable Unit

Acronym and Abbreviation	Definition
FT feature	Fault Tolerant feature, which implements failover between two CPU blades in one SNS5120
GUI	Graphical user interface
Hard Zone	Hard zoning divides the fabric for purposes of controlling discovery and inbound traffic.
Hop	Transmission latency
ICL	Inter Chassis Link, the link between directors or chassis
ISL	In Switch Link, the cascading link between switches
Inactive Firmware	Firmware image that is not used by the switch
In-band Management	The ability to manage a switch through another switch over an inter-switch link
Initiator	The device that initiates a data exchange with a target device
In-Order-Delivery	A data transmission mode that requires that frames be received in the same order in which they were sent
Input Power LED	A chassis LED that indicates that the switch is receiving proper DC voltages
Inter-Switch Link	The connection between two switches using E_Ports
LIP	Loop initialization primitive
Logged-in LED	A port LED that indicates device login or loop initialization status
Management Information Base	Parameters used for enabling or disabling SNMP functions
N_Port	Node Port. The node can be a storage device, disk drive, or disk library.
NPIV	N_port ID virtualization parameter that determines whether a N_Port has virtualized ports
Pending Firmware	Firmware image that will be activated upon the next switch reset
Port Binding	List of device WWNs that can log in to a switch port.
POST	Power-On Self Test Diagnostics

Acronym and Abbreviation	Definition
Principal Switch	The switch in the fabric that manages domain ID assignments
RSCN	Registered State Change Notification. Devices of a fabric will be notified of the status change of any device in the fabric.
SFP	Small Form-Factor Pluggable, an optical module
SNMP	Simple Network Management Protocol, an application protocol that manages and monitors network communications and functions
Soft Zone	Soft zoning divides the fabric into independent zones that contains WWNs and ports.
Target	A storage device that responds to an initiator device
trunk	A batch of port which can transmit data concurrently
WWN	Worldwide Name, a unique 64-bit address assigned to a device by the device manufacturer
zone	A group of ports or devices