Switch Independent Partitioning and Data Center Bridging

Dell and QLogic Deliver iSCSI over Data Center Bridging with Type-Length-Value (TLV) Support

**Switch Independent Partitioning**

Switch Independent Partitioning, also referred to as NIC Partitioning (NPAR), is a method of dividing a single physical 10GbE Ethernet port into multiple PCI physical functions or partitions with flexible bandwidth capacity allocation. This approach enables administrators to apply QoS to the virtual ports, which improves I/O performance while maintaining a low total cost of ownership (TCO). Dell's implementation maps four PCI functions to each physical port on a dual-port 10GbE device. Switch Independent Partitioning presents the eight PCI functions per device using standard PCI configuration space. Each function or partition is assigned a unique MAC address. Each partition can support concurrent networking and storage protocols, enabling flexible bandwidth provisioning to applications.

The administrator can assign these partitions to run both networking and storage protocols, with an implementation that is agnostic to the external Ethernet switch. For certain configurations, the QLogic unique eSwitch implementation enables the switching of VM-to-VM traffic via the embedded switch located in the adapter, thereby reducing the traffic and latency of flow through an external switch port. This results in lower CPU utilization while preserving I/O performance, providing sufficient capacity to run additional applications. Switch Independent Partitioning with the unique eSwitch implementation ensures the highest levels of virtualization with assured interoperability in heterogeneous data centers.

**Introduction**

Data Center Bridging (DCB) is a series of enhancements to the Ethernet protocol that enable the use of traditional Ethernet-based LAN switching to support Fibre Channel SAN data center traffic.

Converged networking requires enhanced Ethernet capability with Quality of Service (QoS) support to deliver the reliable solutions that enterprise data center deployments require. Administrators can configure guaranteed granular bandwidth allocation for storage and networking traffic over 10Gb Ethernet by setting the QoS attributes. QoS settings allow bandwidth allocation between traffic classes with assigned priority attributes. This provides efficient I/O utilization and management of networked traffic and eliminates the need for discrete 1Gb Ethernet ports—enhancing the value of network virtualization.

This technology brief examines the use of QLogic® Converged Network Adapters with Switch Independent Partitioning to enhance the user flexibility of managing virtual data paths at the I/O level in Dell® 12G servers in a DCB environment with iSCSI TLV.

Dell® and QLogic® are driving next-generation server I/O virtualization with Switch Independent Partitioning, which is based on QLogic’s VMflex® technology.
Switch Independent Partitioning and
Data Center Bridging

SWITCH INDEPENDENT PARTITIONING – HARDWARE-BASED QoS

Dell and QLogic are driving next-generation server I/O virtualization with Switch Independent Partitioning, which is based on QLogic’s VMflex technology. Switch Independent Partitioning, also referred to as NIC Partitioning (NPAR), is a method of dividing a single physical 10GbE Ethernet port into multiple PCI physical functions or partitions with flexible bandwidth capacity allocation.

This approach enables the administrator to apply QoS to the virtual ports, improving I/O performance while maintaining a low total cost of ownership (TCO). Dell’s implementation maps four PCI functions to each physical port on a dual-port 10GbE device. Switch Independent Partitioning presents the eight PCI functions per device using standard PCI configuration space. Each function or partition is assigned a unique MAC address. Each partition can support concurrent networking and storage protocols, enabling flexible bandwidth provisioning to applications.

Unlike competitive solutions, these partitions can be assigned to run both networking and storage protocols, with an implementation that is agnostic to the external Ethernet switch. For certain configurations, the QLogic unique eSwitch implementation enables the switching of VM-to-VM traffic via the embedded switch located in the adapter, thereby reducing the traffic and latency of flow through an external switch port.

This results in lower CPU utilization while preserving I/O performance, providing sufficient capacity to run additional applications. Switch Independent Partitioning coupled with the unique eSwitch implementation provides a non-captive (switch and OS agnostic) solution to customers, ensuring the highest levels of virtualization with assured interoperability in heterogeneous data centers.

DATA CENTER BRIDGING

Lossless Ethernet is specifically designed to meet the needs of storage applications by managing the flow of traffic onto the network to avoid congestion and the risk of packet loss for the storage traffic. DCB encompasses four new capabilities that were added to the Ethernet 802.1Q specifications to support new requirements to allow storage traffic to be transported over Ethernet: Priority-based Flow Control (PFC: 802.1Qbb) for controlling the amount of traffic that is allowed onto the network based on the defined traffic class, Enhanced Transmission Selection (ETS: 802.1Qaz) for classifying traffic and requesting lossless services, Congestion Notification (CN: 802.1Qau) for monitoring the congestion in the network and triggering flow control events between devices, and an exchange protocol called Data Center Bridging exchange (DCBx: 802.1Qaz) for devices that support DCB to communicate capabilities and configuration information to each other and manage the network.

This protocol leverages functionality provided by IEEE® 802.1AB (LLDP) and is included in the 802.1az standard. DCB also leverages existing capabilities of standard Ethernet devices to dedicate specific amounts of bandwidth to specific traffic classes and enforce class-based priority access to the network. It uses DCBx to exchange configuration information with the directly connected peer. DCBx is expected to operate only over a point-to-point link and if multiple peers are discovered, the peers’ TLVs should be ignored until the multiple peers’ condition is resolved.

QLogic was the first to introduce the Converged Network Adapter class of converged networking products to the market and maintains a leadership position in that segment. The company’s Converged Network Adapters support concurrent LAN (TCP/IP) and SAN (FCoE and iSCSI) traffic over a shared 10Gb Ethernet link and are ideal for beginning the evolution to converged networking at the SAN edge. The adapters deliver industry-leading LAN networking performance. They are the only adapter offering in the market with full hardware offload for FCoE and iSCSI protocol processing and an extensive set of features that enhance system virtualization. In addition, these adapters leverage QLogic’s existing Fibre Channel and iSCSI drivers that have gone through more than 10,000 man-months of reliability testing and are field hardened with deployments of more than 11 million ports in enterprise data centers. This makes the adapters the most dependable and field-hardened Converged Network Adapters on the market.

---

Figure 1. Switch Independent Partitioning
Supporting DCB for iSCSI allows QLogic’s 10Gb iSCSI to participate in Enhanced Transmission Selection (ETS), which allocates to iSCSI network bandwidth, and Priority Flow Control (PFC), which prevents dropped frames on a per-protocol basis. This implementation allows iSCSI and other traffic (NIC or FCoE) to converge on the same cable, each having an allocated share of the bandwidth and a PFC priority with which it can be independently throttled without affecting the flow of the other protocols’ traffic.

**DCB AND SWITCH INDEPENDENT PARTITIONING**
ETS enables a converged network switch to allocate bandwidth to the converged protocols (NIC, FCoE, and iSCSI) configured on an attached port. Traffic transmitted from the Converged Network Adapter to the switch is examined by the switch that passes the traffic to the fabric proportionally to the protocols’ allocations established by DCBx. If a Converged Network Adapter port supports Switch Independent Partitioning and is configured for multiple NIC functions, ETS makes no distinction between the traffic flows from the different NIC functions.

QLogic provides Switch Independent Partitioning QoS to divide the NIC bandwidth among the configured NIC functions. The DCBx specifies the application protocol TLV that allows the upper layer protocols and related priority information to be advertised and exchanged over the DCB link. The iSCSI TLV application protocol parameters are used to set the priority flow control and any other required parameters. A traffic agreement is set between the two nodes, and all traffic subsequent to the TLV agreement follows the settings established in the exchange.

**LEARN MORE**
QLogic offers complete solutions to some of the most complex issues facing the data center. To visit the QLogic solutions Web page, click “QLogic Solutions” under the Solutions tab at www.qlogic.com.

QLogic has created demonstrations that showcase QLogic technologies, such as NIC Partitioning and the application of QoS in a VMware® ESXi v5.x environment using a QLogic QLE8242 Converged Network Adapter. In addition, demonstrations of technology performance or proof of concepts can be configured for your specific environment. To register for a demonstration of QLogic product capabilities, visit http://www.qlogic.com/solutions/Pages/qos.aspx and click “Schedule a future ‘live’ demo.”
Switch Independent Partitioning and Data Center Bridging

KEY BENEFITS
Here are some of the ways Switch Independent Partitioning offers value to an organization:

Reduced Capital and Operational Expenditures
- Enables hardware consolidation of multiple 1GbE ports into a single 10GbE port
- Allows the creation of multiple partitions for flexible bandwidth provisioning on 10GbE adapters
- Provides lower adapter, cabling, switch port, and management costs

Eliminates Dependency on a Captive Switch
- Removes dependency on an external Ethernet switch to regulate and manage bandwidth
- Offers customers the freedom to choose an Ethernet switch of their choice

Simplified Deployment and Management
- Provides concurrent support for Ethernet, FCoE, and iSCSI protocols, which appear as discrete functions similar to native OS, minimizing deployment disruptions
- Eliminates the need for OS changes to implement flexible bandwidth provisioning

Improved Ability to Scale the Business
- Reduces I/O emulation overhead
- Assigns I/O hardware directly to virtual machines (VMs)
- Allows the dedication of bandwidth (QoS) for VMs and associated applications
- Conserves PCIe® slots in all server environments
- Enables high I/O performance in virtualized environments

Advantages of Switch Independent Partitioning
- Available today