



White Paper

Enhancing Scalability Through Network Interface Card Switch Independent Partitioning

Network interface card (NIC) Switch Independent Partitioning can simplify end-to-end networking by dividing a network controller port into as many as four partitions, enabling dynamic allocation of bandwidth as needed while reducing the total cost of ownership.

May 2012



Third-party information
brought to you
courtesy of Dell®.

Overview

A burgeoning number of applications, rich multimedia content, increasing virtualization, and the growth of cloud computing hamper networking efficiency. To handle these increased workloads, many organizations have transitioned from Gigabit Ethernet (GbE) networks to 10-Gigabit Ethernet (10GbE) networks. This bandwidth increase can provide significant performance and cost efficiencies while supporting efforts to consolidate data centers. However, as organizations increasingly transition from GbE to 10GbE networks, traffic flow for application workloads may not require the entire bandwidth that is available in a 10GbE pipe.

Many organizations have built their IT infrastructures by adding separate servers, networks, and storage capacity for each application workload. These silo application deployments helped simplify management in some ways but left excessive pockets of unused server capacity that contributed to network sprawl. While network and storage convergence as well as server virtualization have helped address these problems, the challenge for 10GbE networks has been provisioning the appropriate bandwidth and data center resources to meet varying application workload demands in both physical and virtual environments.

One powerful approach to help eliminate these problems is Switch Independent Partitioning, in both physical and virtual environments. This approach enables administrators to split up the pipe to divide and reallocate bandwidth and resources as needed. In collaboration with Broadcom®, Dell® provides 10GbE Switch Independent Partitioning in Dell PowerEdge™ rack, tower and blade servers, giving IT organizations enhanced flexibility and capacity as they migrate to 10GbE networks. For more information on features offered, see [“Switch Independent Partitioning Features in Dell PowerEdge Servers” on page 6.](#)

Splitting the 10GbE Pipe

Partitioning begins by configuring a single 10GbE port to represent as many as four separate partitions or physical functions. Each partition is an actual PCI Express® (PCIe™) function that appears to the system ROM, OS, or virtualization OS as a discrete physical NIC with its own driver software, and each partition behaves as an independent NIC port (see [Figure 1](#)).

A partition allows complete logical isolation and division of networking and storage resources, and administrators can dynamically assign these resources on a per-partition basis. By provisioning each of these physical functions with the desired bandwidth in 100 Mbps units, administrators can achieve high granularity in partitioning. In essence, the technology allows administrators to split the 10GbE pipe into several network interfaces in whatever increments are appropriate to provide the correct bandwidth for each task. Switch Independent Partitioning works with a variety of standard Ethernet switches and pass-through devices.

Each partition is configured with Relative Bandwidth Weight and with Maximum Bandwidth. The Maximum Bandwidth is used to limit the bandwidth given to a partition, even if there is available bandwidth that no other partition uses. The relative weight is used to determine the relative bandwidth a partition will get with regard to the other partitions on the same port when they compete on the same available bandwidth. A partition will get its relative weight out of the available bandwidth but never more than its configured Maximum Bandwidth. If a partition is not using its configured weight, its bandwidth will be available for the use of other partitions.

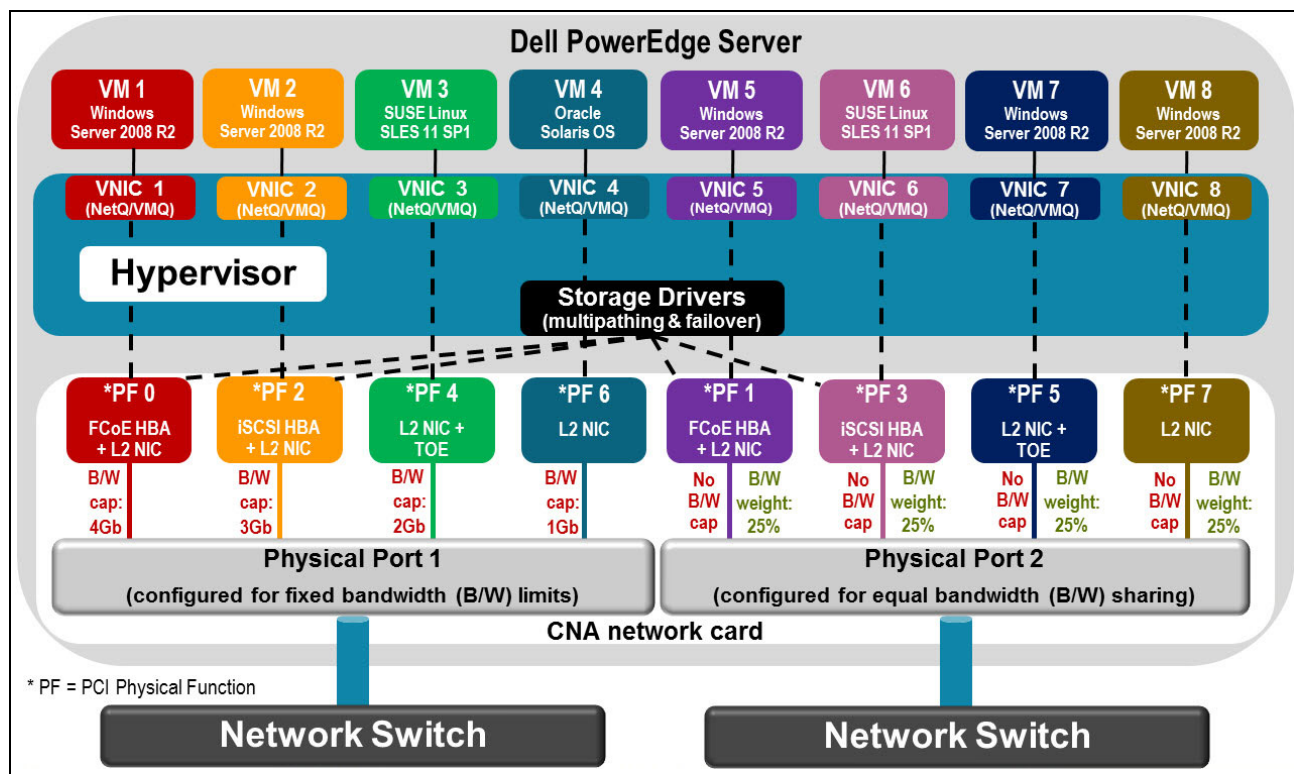


Figure 1: Switch Independent Partitioning Configuration

Dynamic Bandwidth Provisioning

Traditionally, systems administrators would oversubscribe bandwidth for shared connections to handle demand surges by the same applications. Instead, administrators can take advantage of the Switch Independent Partitioning capability to handle such demands and can configure the weighting of each partition to provide increased bandwidth presence when an application requires it (see [Figure 2](#)). Each virtual machine (VM) assigned to a specific partition with equal weights can acquire as much free bandwidth as is available, while incrementally yielding bandwidth back as demand from the other partitions increases.

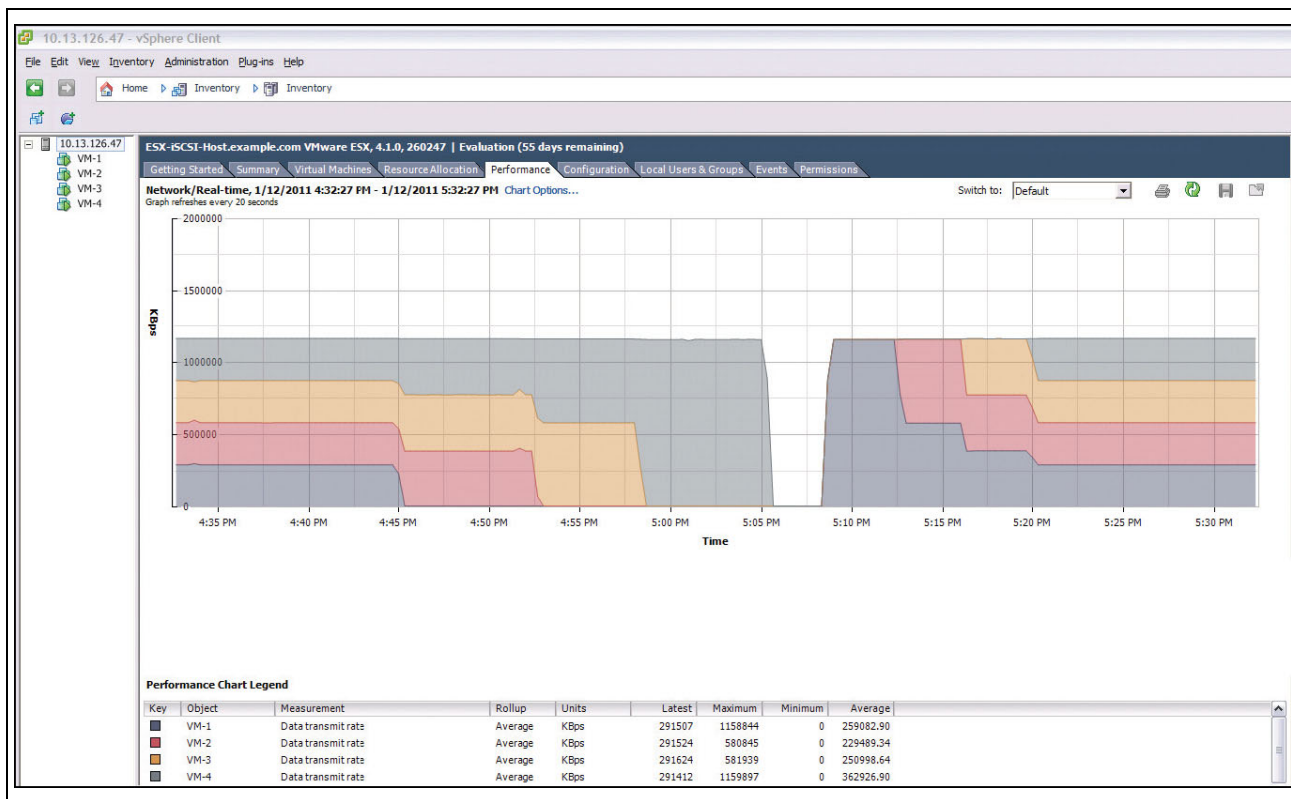


Figure 2: Dynamic Bandwidth Provisioning for Four Partitions Sharing a 10GbE Connection

Switch Independent Partitioning Configurations

Each NIC partition is presented with a dedicated PCI function that can be enabled or disabled by administrators. With this PCI function, partitions behave as normal PCIe Ethernet devices from a host system and OS perspective — requiring no changes to the OS. After the driver is loaded, each enabled partition becomes essentially a dedicated network connection or Ethernet and storage controller device. Standard device configurations can then be applied to the partition.

Each partition can support networking functions such as TCP checksum offload, large send offload, Transparent Packet Aggregation (TPA), multiqueue receive-side scaling, VMware® NetQueue and Microsoft® VMQ, along with storage host bus adapter (HBA) features such as Internet SCSI (iSCSI) and Fibre Channel over Ethernet (FCoE). Administrators can configure a partition to run iSCSI, FCoE, and TCP/IP Offload Engine (TOE). All features can be enabled or disabled on each partition as necessary.

Administrators can provision NICs quickly and easily using the Dell Unified Server Configurator (USC) and the Broadcom Advanced Control Suite (BACS, an integrated graphical user interface management application that can be used to monitor statistics and track configurations across devices).

Streamlining Networking with Switch Independent Partitioning

Switch Independent Partitioning helps simplify the data center and the network and storage infrastructure in several ways. For example, when connecting servers to LANs and storage area networks (SANs), administrators may have to use many cables, sometimes adding switches to reduce cable proliferation.

Switch Independent Partitioning provides a third alternative — consolidating Ethernet and Fibre Channel connections onto significantly reduced numbers of devices. Like switches, Switch Independent Partitioning reduces the number of cables without adding workloads on the network — but Switch Independent Partitioning requires fewer devices versus using switches and cables, and is designed to offer the following additional benefits:

- **Reduced network sprawl:** With a two-port GbE device, a data center can usually deploy 10 NICs, two HBAs, and many cables and switch ports to meet bandwidth needs. With the 10GbE Switch Independent Partitioning approach, the same capacity and flexibility can be provided with a 10GbE converged network adapter (CNA) — consolidating the infrastructure onto much less equipment while using less facility space than when deploying GbE devices.
- **Maximized network scalability:** Having a reduced number of network devices and cables allows IT organizations and data centers to easily scale their networks and add servers and network devices to meet growing IT demands.
- **Simplified administration:** Switch Independent Partitioning can also save time and labor by helping to simplify the management tasks for IT administrators and enhancing their self-sufficiency. They can add or replace network cards or move workloads from one partition to another within minutes.
- **Optimized resource allocation:** Switch Independent Partitioning can play an important role in helping to optimize bandwidth allocation and the effective utilization of that allocation in both virtualized and non-virtualized environments. Virtualization can reduce the number of physical servers and can increase I/O demand for each server. Using 10GbE connectivity with Switch Independent Partitioning addresses this problem by providing eight functions on a single integrated network adapter built into the server, so the server can efficiently handle current bandwidth requirements and still allow additional devices to be plugged in to meet future bandwidth needs. Switch Independent Partitioning also allows systems administrators to fine-tune the amount of bandwidth for each adapter port, as required by the application, and to set the bandwidth for each physical function. Switch Independent Partitioning helps optimize enterprise data centers by providing these benefits for 10GbE networks.

Using Switch Independent Partitioning to Boost Networking Efficiency

The Switch Independent Partitioning approach helps make the migration to today's powerful 10GbE networks an easy and compelling option for many organizations. 10GbE networks provide organizations an opportunity to achieve the performance and efficiency capabilities these networks can offer, while reducing total cost of operation. Also, 10GbE networks can deliver the flexibility that enhances the efficiency of Dell servers and storage.

Switch Independent Partitioning Features in Dell PowerEdge Servers

Administrators using Dell PowerEdge servers enabled with Broadcom Switch Independent Partitioning can take advantage of several features and functions, as follows:

- Support for up to four partitions per 10GbE port.
- Support for monolithic and virtualization operating systems—Microsoft Windows®, Linux®, and VMware® operating systems.
- No OS or BIOS changes required.
- Switch support for industry-standard 10GbE switches.
- NIC control of the transmit flow rate from the server.
- Flexible and dynamic bandwidth allocation.
- Comprehensive support for standard network offload technologies including large send offload, TCP/IP and TCP/UDP, TCP checksum offload, receive-side scaling, and TPA.
- Support for the TOE and iSCSI host bus adapters.
- Support for FCoE host bus adapters.

Authors

Abhijit Aswath is a senior product line manager for Ethernet controller software at Broadcom.

Srinivas Thodati is a product marketing director for the Dell PowerEdge M-Series blade servers at Dell.