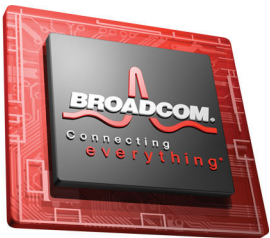# White Paper

Enhancing Scalability Through Network Interface Card Partitioning

Network interface card (NIC) 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 and helping reduce total cost of ownership.

April 2011

Overview

A burgeoning number of applications, rich multimedia content, increasing virtualization, and the growth of cloud computing are all hampering networking efficiency. To handle increased workloads, many organizations have transitioned from Gigabit Ethernet (GbE) networks to 10-Gigabit Ethernet (10 GbE) networks. This bandwidth increase can provide significant performance and cost benefits while supporting efforts to consolidate data centers. However, as organizations increasingly transition from GbE to 10 GbE networks, traffic flow for application workloads may not require the entire bandwidth that is available in a 10 GbE 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 also left excessive pockets of unused server capacity and contributed to network sprawl. While network and storage convergence and server virtualization have helped address these problems, the challenge for 10 GbE networks has been provisioning the appropriate bandwidth and data center resources to meet varying application demands in both physical and virtual environments.

One powerful approach to help eliminate these problems is network interface card (NIC) partitioning, or NPAR, 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 10 GbE NIC partitioning in Dell™ PowerEdge™ blade servers, giving IT organizations enhanced flexibility and capacity as they migrate to 10 GbE networks. For more information on features offered, see "NIC Partitioning Features in Dell Poweredge Blade Servers" on page 6.

Splitting the 10 GbE Pipe

Partitioning begins by configuring a single 10 GbE 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 on page 3).

A partition allows complete physical 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 10 GbE pipe into several network interfaces in whatever increments are appropriate to provide the correct bandwidth for each task. NIC partitioning is switch-agnostic and works with a variety of standard Ethernet switches or passthrough devices.
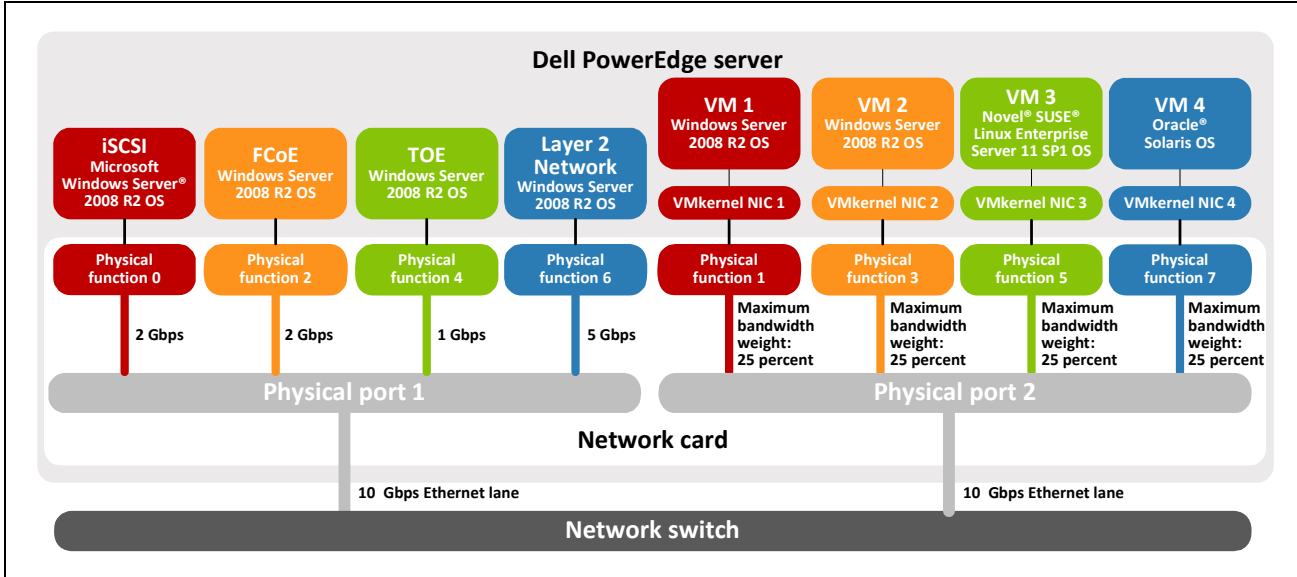
Figure 1: NIC Partitioning Configuration

Dynamic Bandwidth Provisioning

Traditionally, systems administrators may oversubscribe bandwidth for shared connections to handle demand surges by the same applications. Instead, administrators can take advantage of the NIC partitioning capability to handle such demands. Furthermore, systems administrators 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 increases on the other partitions.
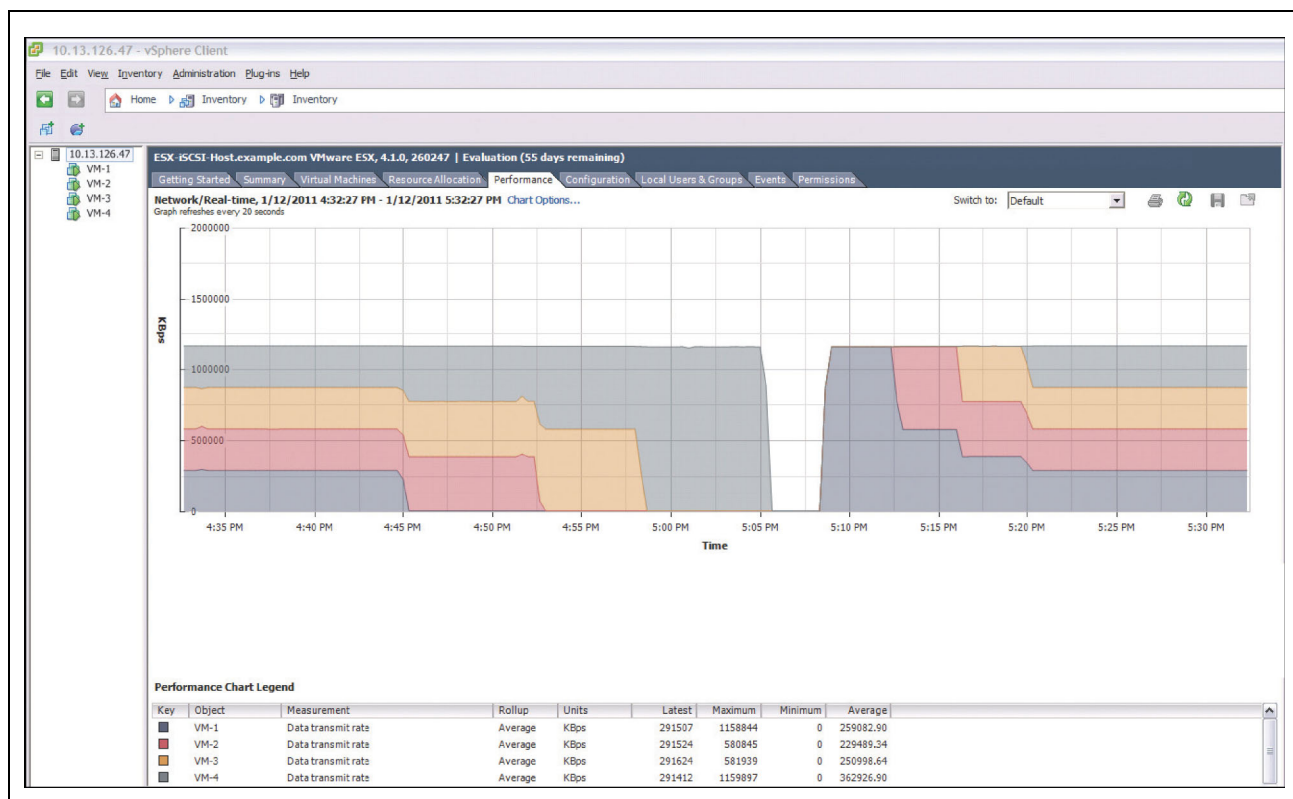
Figure 2: Dynamic Bandwidth Provisioning for Four Partitions Sharing a 10 GbE Connection

NIC Partition 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 multiport PCIe Ethernet devices from a host system and OS perspective—requiring no changes to the OS. After the driver is loaded, each enabled partition is 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, and the VM queue (VMQ) feature of the Microsoft® Hyper-V™ hypervisor, along with storage host bus adapter (HBA) features such as Internet SCSI (iSCSI) HBA and Fibre Channel over Ethernet (FCoE) HBA. Administrators can configure a partition to run iSCSI, FCoE, and TCP/IP Offload Engine (TOE) simultaneously. 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 NIC Partitioning

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

NIC partitioning provides a third alternative — consolidating Ethernet and Fibre Channel connections onto significantly reduced numbers of devices. Like switches, NIC partitioning reduces the number of cables without adding workloads on the network — but NIC partitioning requires fewer devices than when using switches and cables. NIC partitioning 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 10 GbE NIC partitioning approach, the same capacity and flexibility can be provided with just one 10 GbE NIC and converged network adapter (CNA) — consolidating infrastructure onto much less equipment and 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:** NIC partitioning can also save time and labor by helping simplify management 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:** NIC partitioning can play an important role in helping optimize bandwidth allocation and effective utilization of that allocation in both virtualized and non-virtualized environments. Virtualization can reduce the number of physical servers, but can increase I/O demand for each server. Using 10 GbE connectivity with NIC partitioning addresses this problem by providing eight functions on a single integrated network adapter built into the server, so the server can efficiently handle bandwidth requirements but allow for plugging in additional devices to meet growing bandwidth requirements. NIC 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.

  NIC partitioning helps optimize enterprise data centers by providing these benefits for 10 GbE networks.

Using NIC Partitioning to Boost Networking Efficiency

The NIC partitioning approach helps make migrating to today's powerful 10 GbE networks an easy and compelling option for many organizations. 10 GbE networks provide organizations an opportunity to achieve the performance and efficiency capabilities these networks can offer, and to reduce total cost of operation. And 10 GbE networks can deliver the flexibility that enhances the efficiency of Dell servers and storage.

NIC Partitioning Features in Dell Poweredge Blade Servers

Administrators using Dell PowerEdge blade servers that provide Broadcom network interface card (NIC) partitioning can take advantage of several features and functions.

- Support for up to eight partitions per device and up to four partitions per port.
- Support for monolithic and virtualization operating systems—Microsoft Windows®, Linux®, and VMware® operating systems.
- No OS or BIOS changes required.
- Agnostic switch support for industry-standard 10 Gigabit Ethernet (10 GbE) 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, receiveside scaling, and Transparent Packet Aggregation (TPA).
- Support for the TCP/IP Offload Engine (TOE) and Internet SCSI (iSCSI) host bus adapters (HBAs).
- Support for Fibre Channel over Ethernet (FCoE) HBAs—including an expected firmware upgrade.

Authors

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

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