Dell PowerEdge Select Network Adapters-

…the freedom to choose

This feature overview describes Dell PowerEdge Select Network Adapters Family, also known as Network Daughter Cards (NDCs).

Rich Hernandez

Pankaj Gupta

Rahul Deshmukh

Nick Harris

Contents

This document is for informational purposes only, and may contain typographical errors and technical inaccuracies. The content is provided as is, without express or implied warranties of any kind.

November 2011 | Rev 1.0

Abstract

The twelfth-generation Dell™ Power Edge™ server portfolio uses a Dell custom daughter card to house the complete LAN on Motherboard (LOM) subsystem. In these systems, the LOM is provided on the Network Daughter Card (NDC) as part of Dell PowerEdge Select Network Adapters family. There are two form factors of Select Network Adapters,one for blade servers and one for rack servers. Blade Select Network Adapter provides dual port 10GbE from various suppliers. Rack Select Network Adapter provides a selection of 1GbE and 10GbE port options such as 1000Base-T, 10GBASE-T and 10Gb SFP+. The Select Network Adapter form factor continues to deliver the value of LOM integration with the system, including BIOS integration and shared port for manageability while providing flexibility.

Changes in server interconnectivity

Historically, servers have provided a fixed LAN on Motherboard (LOM) subsystem integrated directly on the main system planar as the primary network access. This provided a fixed physical layer interconnect, typically One Gigabit Ethernet, and a fixed number of ports. This has served customers well until the advent of high-speed 10GbE technologies, which now provide significant benefits and physical-layer interface choices. As server designs have become more dense, it has become inevitable that the built in LOM would give way to offer more flexibility. Mainstream servers are now requiring 10Gigabit per second interfaces with added demand for more flexibility from the LOM slot. Dell has listened to our customers and is now offering a flexible approach to the LOM on our rack and blade mainstream Power Edge servers.

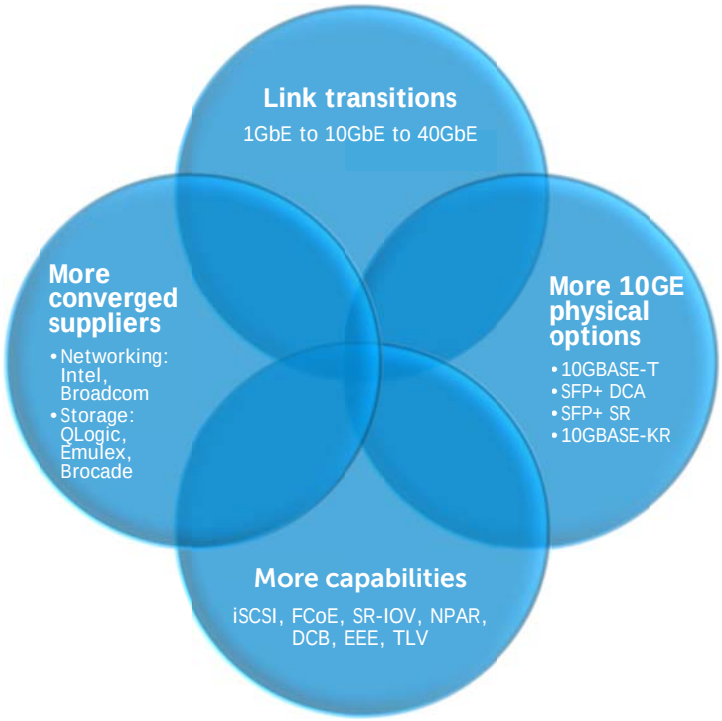
Server virtualization, network convergence, and space efficiency are three primary factors driving the need for flexibility on the choice of LOM subsystem. In addition, each factor drives the need for higher bandwidth interconnects, such as 10GbE and 40GbE.

The server network interconnect landscape has changed significantly. The following points, along with Figure 1, help to explain these changes.

- The transition to 10GbE has begun with various physical interface choices. Both copper and optical choices are needed, including 10GBASE-T and SFP+ optical and Direct Copper Attached. 10GBASE-T is the physical interconnect of choice for installations requiring long cable runs (for example, >10 m). SFP+ and Direct Copper Attached is the current preference for top-of-rack 10GbE connectivity, but we expect many installations to shift to 10GBASE-T in the future. In addition, XAUI-to-KR transition is driven by blade infrastructure to provide 10GbE on all fabrics.
- Network convergence (iSCSI and FCoE [Fibre Channel over Ethernet]) and virtualization is driving an ever richer feature set that offer choice of capabilities and performance.
- The need for higher space efficiency, or the ability to do more with the same or smaller space, means the network fabric taken by the LOM must be able to do more than just networking via Ethernet.

- Traditional networking and storage server I/O is converging, broadening the supplier choice. The Select Network Adapter provides link speed choice, network interface choice, and vendor choice.

Figure 1. The server interconnect landscape



The need for LOM flexibility

Dell unique PowerEdge Select Network Adapters family have a lot of benefits. You can respond to a network infrastructure upgrade in mid-server cycle with full Dell management integration. It is also possible to transition from 1GbE to 10GbE and to converged 10GbE without having to replace servers The transition from optical to copper network connectivity can be done easily as well. With customers starting to ask for convergence options, it is now possible to deploy full end-to-end convergence without having to rip and replace servers. Networking solutions can be selected depending upon the requirements, and not limited by existing infrastructure, which has not been possible with traditional LOM designs in the current market. This unique feature gives control back to you to choose the best of breed technology which best fits your requirements, and it does not require you to pay for something that you will not use.

Dell has chosen the Select Network Adapter to deliver a choice of LOMs. This includes traditional LOM networking capabilities, as well as a new CNA storage functionality. Given the broad spectrum of vendors and capabilities, there is no a single solution that addresses all customer needs and priorities. Some solutions are better at core networking functionality at lower power, while others deliver best converged performance. The Select Network Adapter provide a set of flexible vehicles to transition from 1GbE to 10GbE. By offering 10Gb port devices as well as 1Gb and 10Gb devices on the same NIC you can maintain legacy 1Gb connectivity while providing 10Gb availability for new

functions and workloads. It allows adaptation to technology changes during server generations, such as 40GbE and Multi-port 10GbE.

Dell's rack Select Network Adapter options include quad port 1GbE and 10GbE options. 10GbE options include two ports of 10GbE and two ports of 1GbE to support a variety of use cases for server interconnect. For example, the 1GbE ports on the rack server NDC could be used for remote management or for Wake on LAN (WoL)**.** Dell's blade Select Network Adapter provides dual-port 10GbE interfaces. Dell chose to offer only 10GbE Network Daughter cards for blades due to faster adoption of 10Gb , convergence and its cost advantages. The cost advantages are due to no dependence of optics transceivers and aggregation via chasis switches.

Dell provides a choice of LOMs that can increase the operational efficiency and longevity of servers and thereby proving investment protection. Servers do not need to be replaced if IO requirements change, but the servers can be better utilized, allowing you to react faster to a changing networking technology landscape. An IO device can be replaced with a new NDC, which provides more benefits than the traditional LOM. These adapters are fully managed by the Dell Universal Server Configurator (USC) for device configuration, updates, and real-time monitoring. All Select Network Adapters can be monitored in real time, utilizing Dell Embedded Management through iDRAC.

Choosing which NDC to use

Choosing a blade server NDC depends on which features are required:

- Offloads
- Convergence
- Virtualization
- Switch Independent Partitioning
- Bandwidth requirements

Choosing a rack server NDC also depends on a number of factors, such as:

- Link bandwidth requirements: 1 Gbps or 10 Gbps
- If 10GE is required, then SFP+ or 10GBASE-T is needed
- Feature requirements such as:
    - Offloads
    - Convergence
    - Virtualization
    - Switch Independent Partitioning

Dell has also provided information on key strengths and features of these NDCs when customer is configuring a server for purchase. This will help customer choose the right NDC for them.

Available NDC options

Table 1 and Table 2 detail the NDC options for PowerEdge rack servers and blade servers, respectively. In addition to these, Dell plans to offer more options in the future as well. Keep in mind that an NDC must be present in the system at all times.

Table 1. Supported Select Network Adapters options for PowerEdge rack servers

Features	Broadcom 5720 Base-T	Intel® I350 Base-T	Broadcom 57800 Base-T and SFP+	Intel x540 Base-T
Ports x link speed	4 x 1Gb	4 x 1Gb	2 x 1Gb + 2 x 10Gb	2 x 1Gb + 2 x 10Gb
Supported Speed	1Gb	1Gb	1Gb and 10Gb	1Gb and 10Gb
SR-IOV	Not Supported	Not Supported	Not Supported	Supported (10GE Only)
ISCSI HBA	Not supported	Not supported	Supported[1]	Not supported
EEE	Supported	Supported	Not supported*	Not supported
FCoE	Not supported	Not supported	Not supported	Supported[2]
Switch Independent Partitioning	Not supported	Not supported	Supported[3]	Not supported
DCB	Not supported	Not supported	Not supported	Supported[4]

Features	Broadcom 5720 Base-T	Intel® I350 Base-T	Broadcom 57800 Base-T and SFP+	Intel x540 Base-T
iSCSI TLV	Not supported	Not supported	Not supported	Supported[5]
Supported PowerEdge servers	R620, R720, R720XD	R620, R720, R720XD	R620, R720, R720XD	R620, R720, R720XD

[1]Only 10GbE ports have iSCSI HBA support

[2]Only 10GbE ports have FCoE support

[3]Only 10GbE ports have Switch Independent Partitoning support. The maximum number of partitions supported is 4 (2 partitions per 10Gb port). 1Gb ports don't support Switch Independent Partitioning.

[4]Only 10GbE ports have DCB support

5 Only 10GbE ports have ISCSI TLV support

Note: Dell will provide a software update for above listed unsupported features in a future release.

Table 2. Supported Select Network Adapters options for PowerEdge blade servers

Features	Broadcom 57810S KR	Intel x520 KR	QLogic QMD8262 KR
Ports x link speed	2x10Gb	2x10Gb	2x10Gb
Supported speed	1Gb and 10Gb	1Gb and 10Gb	10Gb only
SR-IOV	Not supported	Supported	Not supported
ISCSI HBA	Supported[1]	Not supported	Supported1
FCoE	Not supported	Supported[2]	Supported[2]
Switch Independent Partitioning	Supported[3]	Not supported	Supported[3]
DCB	Not supported	Supported[4]	Supported[4]
DCB with iSCSI TLV	Not supported	Supported[5]	Supported[5]

Features	Broadcom 57810S KR	Intel x520 KR	QLogic QMD8262 KR
Supported PowerEdge servers	M620	M620	M620

[1]Only 10GbE ports have iSCSI HBA support

[2]Only 10GbE ports have FCoE support

[3]Only 10GbE ports have Switch Independent Partitoning support. The maximum number of partitions supported is 8 (4 partitions per 10GE port).

[4]Only 10GbE ports have DCB support

5 Only 10GbE ports have ISCSI TLV support

Typically, the system board will have one blade NDC connector for 2S half-height blades. An NDC is always required for the system to POST without error.

1GbE port features

Dell's 1GbE rack server NDCs support triple link speeds—10M/100M/1G—over unshielded twisted pair. Features such as iSCSI boot and EEE are supported as well.

ISCSI boot

The 1GbE NDC ports support the ability to boot the server from a remote iSCSI target over the network. This is fully managed by Dell USC for device configuration, updates, and real-time monitoring.

Energy Efficient Ethernet (EEE)

Dell Broadcom and Intel NDC GE ports support Energy Efficient Ethernet (EEE). EEE is based on the IEEE 802.3az specification. EEE defines a mechanism to reduce power consumption on the port during periods of low-link utilization. EEE reduces power consumption on an operational link by transitioning the link to Low Power Idle state when the transmitter is idle.

The EEE implementation ensures the following behavior:

- The link status does not change as a result of the transition to Low Power Idle
- No frames in transit are dropped or corrupted during the transition
- The transition is transparent to upper layer protocols and applications

10GE port features

The 10GBASE-T rack server NDC supports triple link speeds—100M/1G/10G—over unshielded twisted pair. This allows an easy way to provision for 10GE on the server side while allowing connectivity to the legacy GE infrastructure.

The 10GE NDCs offer a broad set of capabilities, including the following:

- iSCSI HBA
- FCoE
- Data center bridging
- Multiple Receive Queues
- Switch Independent Partitioning -
- SR-IOV

ISCSI boot

The 10GbE NDC ports support the ability to boot the server from a remote iSCSI target over the network. This is fully managed by Dell USC for device configuration, updates, and real-time monitoring.

iSCSI HBA

The iSCSI offload adapter supports full offload of the data path for iSCSI IO, including offload for iSCSI header and data digests. Session initiator and teardown is managed by host driver components, but once an iSCSI session is established, all iSCSI protocol involved in initiating and completing SCSI IOs are offloaded to the network controller hardware.

Fiber Channel over Ethernet (FCoE)

The FCoE offload adapter supports FCoE and FIP per the T11 American National Standard for Information Technology- Fibre Channel—Fibre Channel Backbone - 5 (FC-BB-5). Full offload of the data path for SCSI IO including offload for FC and FCoE is supported. Session initiation and teardown is managed by host driver components, but once an FC session is established, all FC protocol involved in initiating and completing SCSI IOs are offloaded in hardware.

Data Center Bridging (DCB)

Data Center Bridging (DCB) includes support for the following IEEE standards:

- **P802.1Qaz:** IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks - Amendment: Enhanced Transmission Selection.
- **P802.1Qbb:** IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks - Amendment: Priority-based Flow Control.
- **DCBx iSCSI Application Type-Length-Value (TLV):** Dell has recently added this application TLV, which provides the primary mechanism for communicating capabilities from switch-to-switch or switch-to-edge device node within the DCBx protocol. This is an extension to the 802.1AB LLDP protocol. Of particular need is the support for an application TLV that supports the TCP well-known port 3260 (iSCSI).

Enhanced Transmission Selection

Enhanced Transmission Selection (ETS) enables transmission bandwidth management and rate limiting per traffic class. ETS allows traffic classes to use spare bandwidth above the guaranteed minimum. ETS provides common management of bandwidth allocation on a per priority group basis. Switches that claim support for DCB, should allow configuration of priority groups, assignment of individual classes of service to each group and the allotment of per-port bandwidth to each group.

Priority-based Flow Control

Priority-based Flow Control (PFC) enables multiple traffic types to share a common Ethernet physical link without interfering with each other. PFC allows link flow control to be performed on a per-priority basis. Priority-based Flow Control (PFC) provides a link level flow control mechanism that can be controlled independently for each priority queue. PFC provides a link level flow control mechanism that can be controlled independently for each frame priority class of service. Switches that claim support for DCB should provide mechanisms for assigning a class of service to different "streams" of traffic.

Data Center Bridging Capability Exchange Protocol

Data Center Bridging Capability Exchange Protocol (DCBCXP) is an LLDP-based protocol used to exchange link configuration parameters for PFC, and ETS. It is a discovery- and capability-exchange protocol that is used to convey link capabilities and configuration. Link partners can choose the following supported features and parameters for PFC and ETS using LLDP, as defined by IEEE 802.1AB:

- Protocol to exchange DCB parameters
- Set local operational parameters based on exchange
- Resolve conflicting parameters

Multiple Receive Queues

Multiple Receive Queues is a hypervisor-enabled technology that offloads the software vswitch from sorting and forwarding the received packets to specific VMs. The NDC controller does the sorting and forwarding of received traffic into queues that are mapped to a given VM using destination MAC Address and VLAN ID, if applicable. This eliminates bottlenecks in the vswitch implementation, increasing the total throughput. VMware calls this technology NetQueue, and Microsoft calls it VMQ.

Switch Independent Partitioning

Switch Independent Partitioning is also referred to as NIC Partitioning (NPAR). Switch Independent Partitioning provides the capability to divide a 10GE NIC port into multiple PCI functions with flexible bandwidth capacity allocation that looks to the OS and network infrastructure as separate NIC interfaces. On the host OS side, Switch Independent Partitioning presents up to eight PCI functions per device using standard PCI configuration space. Dell's implementation maps four PCI functions to a physical port on a dual-port 10GE device. Each function or partition is assigned a unique MAC Address.

Switch Independent Partitioning enables allows you to replace multiple 1GE NICs with partitioning-enabled 10GE NICs. This allows consolidation of multiple GE ports into fewer 10GE ports, reducing switch port and cabling complexity while maintaining the network segmentation and isolation. In addition, flexible-bandwidth allocation per partition allows for efficient use of the 10GE link.

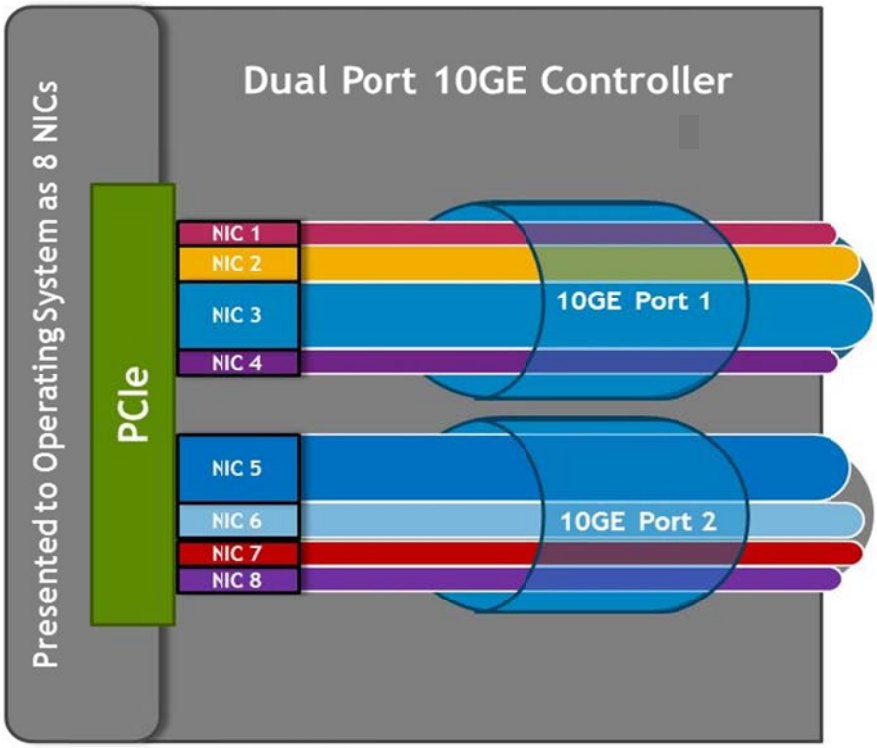
A Switch Independent Partitioning enabled 10GE link supports the following server use cases:

- **Server segmentation:** The partitions can be on separate subnets or VLANs.
- **Server high availability:** The partitions support NIC teaming, including switch dependent link failover and load balancing.
- **Physical and virtual server:** NIC Partitioning is supported in both native OS and hypervisor-based OS. In a virtual server, emulated and direct assignment I/O of partitions to VMs are supported.

Partitions are available for VM assignment (direct or emulated) and for application segmentation via VLAN or IP subnets.

Note: The maximum number of supported partitions by Broadcom 2x1Gb+2x10Gb Rack NDC is 4 (2 partitions per 10Gb port). The 1Gb ports don't support switch independent partitioning.

Figure 2. 10Gb Switch Independent Partitioning



SR-IOV ready

Single Root I/O Virtualization (SR-IOV) enables a natively shareable device that can bypass the hypervisor on a virtualized server for the main data movement. It enables a VM to have direct access to a PCI I/O device while sharing the device among multiple VMs or GOSes. The PCI-SIG Single Root I/O Virtualization and Sharing specification defines a standard to implement a natively shareable device in a virtualized host. Virtual functions provide independent DMA streams that can be assigned to a VM. A Physical Function includes the resource capabilities to create the VFs.

A single Ethernet port appears as multiple devices showing in PCI Configuration space as multiple functions. SRIOV enabled hypervisors initialize and assign VFs to virtual machines.

Table 3. Dell PowerEdge Select Network Adapters software features

Vendors	Chipset	Supported speed	FCoE (offload, boot)	ISCSI (boot, offload, software)	Switch Independent Partitioning	SR-IOV	DCB with iSCSI	System management with real-time monitoring
Broadcom®	5720	1GbE	Not supported	Software iSCSI, iSCSI boot	Not supported	Not supported	Not supported	Supported