High-Availability Blade Server Clustering with the Dell PowerEdge Cluster FE555W

Dell high-availability clusters such as the Dell™ PowerEdge™ Cluster FE555W are based on cost-effective, modular Dell PowerEdge servers. The flexible, rack-mountable PowerEdge 1855 blade server supports 10 removable server blades, fabric switches, and Ethernet switches—all efficiently stored within the 7U Dell Modular Server Enclosure. By incorporating the dense hardware configuration and redundant components of the PowerEdge 1855 blade server, the PowerEdge Cluster FE555W is designed to save data center space and enable high availability.

BY GREG BENSON, BRYANT VO, AND FARRUKH NOMAN

Dell high-availability (HA) clusters are designed to provide maximum uptime for business-critical environments with continued efficiency and effectiveness. The Dell PowerEdge Cluster FE555W incorporates the flexible, industry-standard Dell PowerEdge 1855 blade server—a rack-mountable server system that supports 10 removable server blades, fabric switches, and Ethernet switches within the 7U Dell Modular Server Enclosure. The PowerEdge Cluster FE555W supports two server modules running the Microsoft® Windows® 2000 Advanced Server OS and two to eight server modules running Microsoft Windows Server™ 2003, Enterprise Edition or Enterprise x64 Edition. For storage, the cluster uses Fibre Channel–based Dell/EMC CX300, CX500, or CX700 arrays.

Understanding the building blocks of the Dell PowerEdge Cluster FE555W

By incorporating redundant hardware features, Dell HA clusters are designed to ensure that resources remain available to network clients and the applications can recover from failure. The PowerEdge Cluster FE555W comprises cluster nodes, Fibre Channel modules, Ethernet modules, and storage systems.

Cluster nodes. The individual server blades of the Dell PowerEdge 1855 blade server, which use dual Intel® Xeon™ processors, function as the cluster nodes. Multiple clusters can be configured depending upon the number of available blades. Supported configurations may use a maximum of two nodes per cluster running Windows 2000 Advanced
Each server blade is configured with two integrated Gigabit Ethernet network interface cards (NICs) to provide cluster heartbeat communication and public client access on two separate networks. These NICs are internally connected to a corresponding port on either an Ethernet pass-through module or a switch module. The Ethernet pass-through module provides a non-switched, one-to-one connection between the server blade and the external Gigabit Ethernet device or another server blade. These ports are preset to communicate only at 1 Gbps and will not auto-negotiate to a slower speed. In contrast, the Gigabit Ethernet switch module—the Dell PowerConnect 5316M switch—provides a switched connection with 6 uplink ports and 10 downlink ports. The option of using any one of these modules for the private and public networks provides flexibility to meet different user requirements. In a configuration with a large number of cluster nodes or with multiple clusters, best practices recommend employing switch modules. Figure 1 shows the private network cabling for pass-through and switch modules in various cluster configurations.

**Storage systems.** Server nodes within a cluster can share one or more external storage systems connected through Fibre Channel switch fabric modules or pass-through modules. The PowerEdge Cluster FE555W solution supports Dell/EMC CX300, CX500, and CX700 storage arrays, which can be configured through a management station using EMC’s Navisphere Manager software.

### Choosing the appropriate configuration

Direct attach storage (DAS) and SAN-based configurations are widely used for HA clusters. Choosing an appropriate configuration depends on the load of the cluster nodes and the application servers that need to participate in the cluster. A DAS-based cluster, suitable for light loads, requires the use of Fibre Channel pass-through modules to communicate directly with the storage system; whereas a SAN-based cluster uses an embedded fabric switch to open the scope for multiple nodes in the cluster to share heavy loads. Using the PowerEdge Cluster FE555W in a DAS environment (where the cluster is typically limited to two nodes) would not be an efficient use of the large number of resources, unless multiple clusters were employed to make use of all the server blades.

The current release of the Dell PowerEdge Cluster FE555W supports 2 Gbps end-to-end cluster environments in a direct attach configuration and both 2 Gbps and 4 Gbps cluster environments in a SAN-based configuration. In the future, Dell plans to provide support for 4 Gbps end-to-end cluster environments in both direct attach and SAN configurations. Based on the enterprise structure, specific storage requirements, and the number of clusters, IT organization can implement various SAN configurations.

### Single-server, single–storage system configuration

In this configuration, a maximum of four ports are available from the fabric switches to provide connectivity to a dual-storage processor (SP) storage system with four ports. Based on the available ports, any Dell/EMC CX storage array can be connected to the

---

1 This term does not connote an actual operating speed of 1 Gbps. For high-speed transmission, connection to a Gigabit Ethernet server and network infrastructure is required.
PowerEdge 1855 blade server. Figure 2 shows the redundant connections to a Dell/EMC CX700 array.

**Single-server, dual–storage system configuration.** Cluster storage capacity can be increased by connecting to two storage systems. Only two dual-SP storage systems with four ports (two from each SP) can be connected to the fabric ports on the PowerEdge 1855 blade server, unless an external fabric is linked to the ports via Inter-Switch Links (ISLs). Figure 3 depicts the cluster configuration with two storage systems.

**Dual-server, single–storage system configuration.** Multiple nodes or clusters can share a single storage system. In this scenario, if multiple PowerEdge 1855 blade servers are employed, then each server requires a direct path to the storage system to help ensure high availability for applications. At least two ISLs are required to provide connectivity between each Fibre Channel switch of the PowerEdge 1855 blade servers. A link to connect the private networks is also required if a cluster is using server blades from both PowerEdge blade servers. Figure 4 shows the connections of two PowerEdge 1855 blade servers to a single storage system.

**Dual-server, dual–storage system configuration.** A second storage system can be added to a dual-server cluster to increase the storage capacity and help ensure high availability by providing direct paths from Fibre Channel modules of both PowerEdge 1855 blade servers. Figure 5 shows the connections between two PowerEdge 1855 blade servers and two storage systems.

**Setting up the SAN**

Setting up the SAN involves configuring the Fibre Channel connections from the HBAs to the storage targets through the fabric. Logistically, this setup requires administrators to configure the Fibre Channel daughter card (HBA) on each blade, set up the storage access, and create fabric zones to secure host-to-storage connectivity. A direct attach configuration employs the same set of rules except for fabric zoning.

**Fibre Channel HBA setup.** The Fibre Channel daughter card on each server blade in a PowerEdge 1855 blade server must be configured for the SAN topology. Administrators can verify or change these settings by using the QLogic SANsurfer utility or the Fast!UTIL option during the boot-up process. Through either tool, the connection and topology settings can be configured for a fabric or direct attach environment.

**Storage setup.** The storage system should be updated with the core software, and EMC Access Logix™ software should be enabled.
to help prevent data corruption. Access Logix software is highly recommended on a Dell/EMC CX storage system if one of the following conditions is true:

- Two or more stand-alone servers or clusters are configured to access the same storage system.
- EMC MirrorView™, SnapView™, and SAN Copy™ software are installed on the storage system and being used by the cluster.

Access Logix restricts access to logical units (LUNs) by organizing them into storage groups for particular hosts. Administrators can activate these access restrictions by enabling Access Control (a feature of Access Logix) in the storage system properties and placing the hosts and their assigned LUNs in a storage group. A set of hosts residing in one storage group will not be allowed to join any other storage group.

Fabric zones setup. Zoning is implemented on Fibre Channel switches to isolate servers and storage systems from each other. In an environment that uses zoning and Access Logix software, multiple PowerEdge clusters can share Dell/EMC storage systems in a switched fabric. Dell supports only single-initiator zoning, in which each Fibre Channel daughtercard port resides in a separate zone with the storage ports—to help ensure that Fibre Channel communication between the cards and their target storage systems do not affect one another. Each Fibre Channel port can be connected to a maximum of two storage systems.

Configuring the cluster

Once all the cluster components are configured properly and the EMC Navigisphere agents are active on the cluster nodes, the Fibre Channel HBAs should be able to make logical connections to the storage. The Navigisphere agent allows each host (cluster node) to register with the storage system. EMC PowerPath™ software also should be installed on the hosts to provide automatic rerouting features for the Fibre Channel I/O traffic, to help offset primary-path failures and address load-balancing needs.

The participating cluster nodes are configured in a storage group with shared LUNs. These LUNs, which can now be accessed from the disk management feature of each node, are prepared with unique disk signatures to actively participate in a cluster. The public and private networks are also verified to be uniquely identifiable in separate domains. Microsoft Cluster Service can be configured with the available resources on one node at a time. If the cluster is formed properly, resources can fail over between different nodes.

Gaining high availability and modularity

Business-critical applications that require maximum uptime are excellent candidates for clustering. The Dell PowerEdge Cluster FE555W can help deliver continuous infrastructure availability for critical applications by leveraging the flexibility and scalability of a SAN. This approach provides enterprise-class features while enabling the PowerEdge Cluster FE555W to help organizations keep acquisition costs low, maximize rack density, and enhance power efficiency.