

A New Market for High-Performance Computing

High Performance Computing (HPC), once the domain of multi-million dollar supercomputers, has undergone fundamental changes. Complex computational problems are now distributed across a compute cluster made up of commodity, 64-bit x86 computers. An entry-level compute cluster can be assembled for a fraction of the cost of a large UNIX® server, bringing high performance computing within the range of smaller companies and individual departments that could not afford these solutions just a few years ago.

New software applications, designed to run on compute clusters and targeted at specific vertical industry problems, now provide a range of industries, including automotive, aerospace, financial services, life sciences and manufacturing, with out-of-the-box solutions. This further reduces the costs and time required to bring high performance computing solutions into operation.

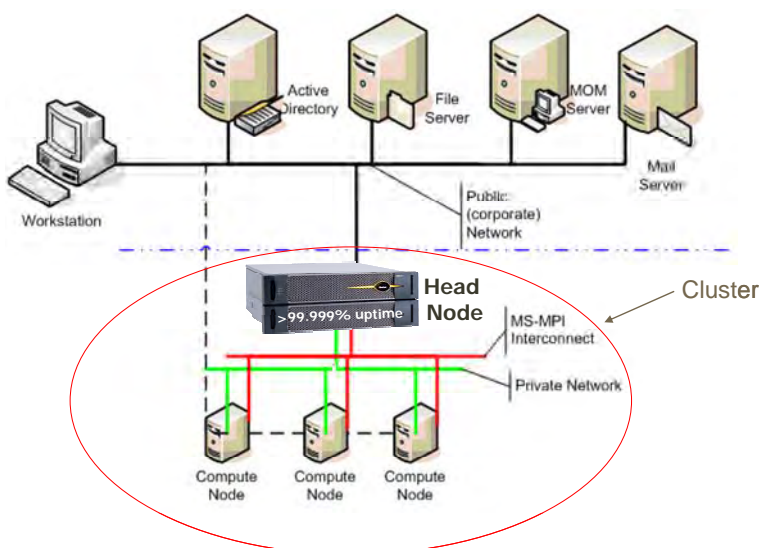
Microsoft now offers a software solution that brings new levels of simplicity and integration to the high-performance computing market. And, with industry-leading uptime that exceeds 99.999% for Microsoft operating environments, Stratus® ftServer® systems are the perfect hardware choice for mission-critical compute cluster components.

Microsoft® Windows Compute Cluster Server

Microsoft Windows Compute Cluster Server (CCS) provides a high-performance computing platform that is simple to deploy, operate and integrate into existing IT infrastructure and tools.

Windows Compute Cluster Server is a software package that can install, configure, operate, and maintain an HPC compute cluster. A typical cluster includes a single head node and one or more compute nodes. The head node is the most critical component of the cluster. It controls and mediates all access to the cluster resources and is the single point of management, deployment, and job scheduling for the compute cluster. Windows Compute Cluster Server uses the existing corporate Active Directory infrastructure for security, account management, and overall operations management using tools such as Microsoft Operations Manager and Microsoft Systems Management Server.

Figure 1. Typical Windows Compute Cluster Server network



In this Windows Compute Cluster Server network, the Stratus fault-tolerant server provides >99.999% uptime for the critical head node and file server functions.

Windows Compute Cluster Server provides many important benefits including:

- Use of commodity hardware and standard Windows operating systems keeps costs low and leverages existing staff knowledge and vendor relationships;
- A complete HPC solution that includes the operating system, job scheduler, message passing interface layer and leading vertical applications;
- Ease of installation, configuration and management using familiar Windows tools such as RIS, MMC, MOM and Perfmon;
- Integration with Windows infrastructure including Active Directory, Exchange, and Windows security;
- Integration with Visual Studio development tools with added parallel debugging and compiling capabilities;
- Microsoft partnerships with leading server, interconnect and application providers.

Ensuring Continuous Availability for Windows Compute Cluster Server

Over the last five years, Microsoft has partnered with Stratus Technologies, a leader in total IT availability, to develop technologies that ensure the support of mission-critical Windows environments. Today, Stratus is both a Microsoft Platinum OEM and Microsoft Gold Certified partner for Enterprise systems. Every Stratus fault-tolerant, ftServer system has passed the rigorous Windows Hardware Compatibility tests and is listed in the Microsoft Server Catalog. As a result, users of Windows Compute Cluster can rest assured that the software will run flawlessly on every ftServer model.

Stratus ftServer Systems with Continuous Processing® Technology

The Stratus ftServer family of Intel®-architecture servers is designed to provide the highest levels of uptime in the industry – 99.999% or better – with measured customer availability results that meet or exceed that goal. Stratus ftServer systems are built from the ground up to avoid downtime altogether. Stratus Continuous Processing technology incorporates lockstep technology, failsafe software and active service architecture. These unique availability innovations all work together to keep ftServer systems in continuous operation.

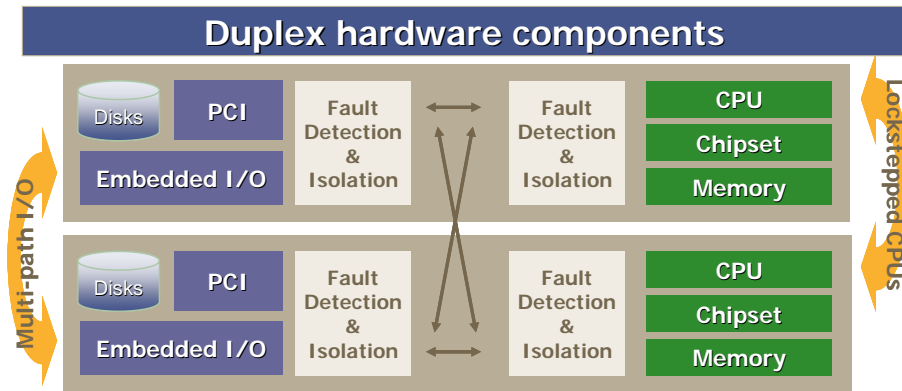
Figure 2: Automatic 99.999% Uptime for Windows Environments



Windows applications benefit from Stratus continuous availability and operational simplicity the minute they are loaded onto the system.

Lockstep technology uses replicated, fault-tolerant hardware components that process the same instructions at the same time. In the event of a component malfunction, the partner component acts as an active spare that continues normal operation and averts system downtime. Fault-tolerant I/O is implemented through use of replicated PCI buses, replicated I/O adapters and replicated devices. Internal SAS and SATA disk storage, along with expansion ftScalable storage, is configured as RAID, connected via two independent storage buses. Any I/O operation failure will result in a retry using an alternate path that ensures successful completion of the I/O operation.

Figure 2: Stratus lockstep architecture diagram



Inside every Stratus ftServer system, replicated fault-tolerant hardware components process the same instructions at precisely the same time. In the event of a component malfunction, the partner component simply continues operating while built-in diagnostics isolate and report the problem.

Failsafe software works in concert with lockstep technology to prevent many software errors from escalating into outages. Unlike standard servers or high-availability software solutions, ftServer handles most errors transparently, shielding the operating system, middleware and application software. Management and diagnostic features capture, analyze and notify Stratus of any software issues. This allows support personnel to take a proactive approach to correction software problems before they recur. In addition, hardened device drivers add considerable reliability to the Windows environment on Stratus ftServer systems. Stratus ftServer systems also offer Active Upgrade™ technology, which allows customers to apply software upgrades and patches to the Windows operating system, system software or applications while applications continue to run. Only a quick application restart is required to bring the upgrades into production.

ActiveService™ architecture is designed to detect and resolve problems before they cause downtime. This architecture combines automatic fault detection, automatic fault isolation, integrated “call-home” remote support, and online component replacement to provide serviceability that is unequalled by other servers.

Microsoft Compute Cluster availability features

Microsoft Windows Compute Cluster Server software incorporates a number of important features that promote availability. CCS management software allows dynamic addition or removal of compute nodes and provides monitoring and recovery features that deal with compute node failures:

- Compute nodes are monitored and notification of compute node failures is provided
- Compute nodes can be dynamically added or removed as workloads change or in response to a failed compute node
- The job scheduler automatically restarts jobs affected by a failed compute node

Protecting the critical CCS head node

The most vulnerable component in the cluster is the head node and its associated services. Any lapse in availability of this critical function will immediately affect the operation of the entire cluster. As such, it is essential to address head node availability in the most effective way possible. Consider the criticality of the following head node functions:

- Manager Service – Cluster management, node discovery, and configuration management
- Scheduler Service – Job queuing, scheduling, resource allocation, and job execution
- SDM Store Service – Read/write operations from the System Definition Model (SDM) data store, integrity management
- SQL Service – Handles database operations for Compute Cluster Server

In addition to the services listed above, CCS job input and output generally use files located on a file server. Loss of these shared files can also disrupt the operation of the entire cluster.

The easiest way to protect the cluster against downtime due to the failure of the head node is to use a fault-tolerant server for this critical function. Using a fault-tolerant head node as the cluster file server will also protect the shared input and output files. By employing a Stratus ftServer system as the Compute Cluster head node, the availability of the head node and of the entire cluster benefit from the five nines track record of availability provided by ftServer systems.

Unlike failover solutions, Stratus ftServer systems avoid downtime due to component failures and avoid any loss of work in process that could require jobs or tasks to be restarted. Stratus ftServer systems also provide built in disk mirroring to protect critical file server data used by the compute cluster. The compatibility and simplicity of ftServer systems means that adding an ftServer head node to a Compute Cluster imposes no additional burdens on the IT staff. The availability of ftServer systems is automatic and transparent. As a result, ftServer systems are installed, configured and managed just like any standard Windows server.

Availability and High-Performance Computing

While not all HPC environments require high levels of availability, environments where application results are time-critical or where compute clusters are large will typically be most sensitive to availability issues. Examples of time-sensitive applications can be found in the financial services industry where applications such as risk management, derivatives trading and portfolio analysis generate results that influence decisions in real time. Any disruption of these high-performance computing applications can result in financial losses or regulatory penalties. Customers in the financial services industry have a long history of critical computer applications and are sensitive to availability issues in all of their computing decisions. Manufacturing applications where calculation results immediately affect operational parameters are another example of time-critical applications that require maximum uptime and data integrity.

Large clusters will often be more vulnerable to lapses in availability because the affects of a cluster outage can be much more significant. Larger clusters usually host larger, longer-running applications. Even if the outage is short, applications running at the time of the outage will be interrupted and require restart of at least some jobs or tasks. Since large, high-performance computing applications can run for hours or even days, the impact of a restart could be severe.

Conclusion

Microsoft Compute Cluster Server brings simple deployment, operation and IT integration to high-performance computing at a price point that allows smaller companies and individual departments to successfully deploy HPC applications. Deploying a Stratus ftServer system as the head node/file server protects the cluster against head node failure and ensures continuous operation of the entire compute cluster. Furthermore, built-in diagnostics and reporting, remote serviceability features and hot-swappable components – integral to every ftServer system – make these servers remarkably easy to install, configure and manage.

Stratus, ftServer, and Continuous Processing are registered trademarks. ActiveService and Active Upgrade, the Stratus Technologies logo are trademarks of Stratus Technologies Bermuda Ltd. Microsoft, Windows and Windows Server are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. Intel is registered trademark of the Intel Corporation in the United States and other countries. All other trademarks and registered trademarks are the property of their respective holders.

Specifications and descriptions are summary in nature and subject to change without notice.

© 2008 Stratus Technologies Bermuda Ltd. All rights reserved.

X970