



NEC Computers White Paper

Fault Tolerant Servers: The Choice for Continuous Availability

NEC Computers Inc.
80 Central Street
Boxborough, MA 01719
www.necftservers.com

Introduction

Computer systems are a critical component of a company’s competitive and financial well-being. A single downed server can cause disruption of business and loss of revenue.

IT staffs are faced with an assortment of availability alternatives. Servers with integrated redundant features, high-availability configurations of two or more systems, and fault tolerant (FT) solutions offer a continuum of uptime ranging from better than average to superior.

The choice of one solution versus another typically involves tradeoffs between an acceptable level of availability, initial purchase price, and the time and effort involved to implement and maintain the solution. Some solutions are more expensive than most organizations can afford, so the organization spends less on initial purchase price and accepts reduced availability.

This paper discusses today’s options for achieving continuous availability and how NEC’s Express5800/ft servers can provide every company with an affordable way to own a true fault tolerant server.

High Availability

High-availability systems are defined as having 99% or more uptime. The most expensive high-availability server configurations offer solutions approaching only 99.9% uptime. This translates into an average of 8.7 hours of downtime per year, which is unacceptable in some business environments.

Availability (%)	Downtime per year
99	87 hours, 36 minutes
99.5	43 hours, 48 minutes
99.9	8 hours, 44 minutes
99.95	4 hours, 23 minutes
99.99	53 minutes
99.999	5 minutes

Clusters and Redundancy

Most common high-availability network configurations are based on clustering and/or redundancy.

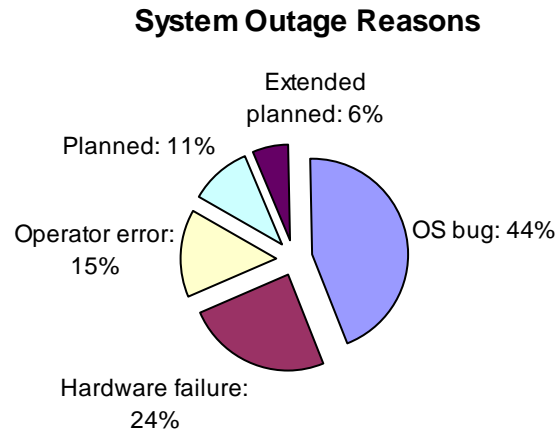
Clustering involves setting up multiple systems as nodes functioning in a single network. Network software recognizes the nodes, provides common resources, and distributes applications and services across them to reduce the chance of a single point of failure. If one node fails, the OS moves the processing to a different node.

Clustering provides only high availability, rather than continuous availability, due primarily to the time involved for the network to notice that a node is not functioning properly and begin a failover routine. This

routine involves restarting applications and databases, which is time-consuming and involves the risk of transactions being lost due to the failure.

Redundancy involves duplicate hardware and applications, with mirroring to maintain duplicate data. If one system fails, the other can be brought up to take its place.

The chart below describes the percentage breakdown for outage given in a survey of IT professionals.



Source: The Standish Group International, Inc. ©2001

Fault Tolerance

Fault tolerant systems like the Express5800/ft series go beyond high availability, into the realm of “continuous availability.” A continuously available server is designed to deliver 99.999% uptime, averaging less than 5 minutes of unplanned downtime per year—including time spent repairing failures, installing upgrades, and performing general maintenance.

NEC Computers Express5800/ft servers, engineered with Stratus[®] fault tolerant technology, address the drawbacks of clusters and redundancy. These servers offer

- one of the highest levels of availability in the industry
- zero failover time should a hardware failure occur
- protection of memory and disk data
- the simplest high-availability solution to install and maintain
- service with no experienced IT staff onsite

Replicated fault tolerant subsystems—CPU, PCI, memory, hard drives, and power supply—virtually eliminate any single point of failure. The replicated subsystems operate in lockstep, processing the same instructions at the same time, with automatic failover to the redundant subsystem should a fault occur. In

the event of a hardware failure, users can replace subsystems while the server and application continues to run.

The merits of fault tolerant solutions servers can be determined most efficiently by analyzing them in light of four major criteria:

- performance
- reliability
- serviceability
- total cost of ownership

Performance

Reconfiguring two or more conventional computers on an Ethernet LAN and relying on that LAN to send availability messages back and forth between systems can cause unexpected performance bottlenecks from excessive LAN traffic. Another potential performance bottleneck occurs when the application allocating CPU cycles performs availability messaging. CPU overhead is required to keep the two servers synchronized, and LAN overhead can be required to exchange synchronization messages. Failover routines take time to run—up to 30 minutes for failovers involving database recovery. If both of the linked systems are being used, a system's work capability can be greatly reduced by a failure.

High-availability configurations of conventional computers are not appropriate for all applications. They are designed for availability rather than speed or scaling. When the nature of an online application is critical, such as insuring network reliability and integrity, continuous availability provides a higher level of assurance that the application will be kept running smoothly 24 hours a day, 365 days a year.

Reliability

Conventional high-availability networks are not fault tolerant; they have many single points of failure that can crash the system. Their goal is to recover as quickly as possible rather than to prevent a crash in the first place. Recovery times can vary from seconds to many minutes, depending on the complexity of the application and the communications involved. Recovery can take much longer if the application must be returned to the point where it was before the failure occurred. Users generally have to re-log in, and disks and databases may have to be resynchronized. If a crash occurs in the middle of a transaction, corrupt data may enter the system, or transaction data may be permanently lost.

Cluster networks rely on layered systems software or custom application code residing above the operating system to recover from system crashes. These configurations have limited ability to identify hardware failures and cannot detect transient hardware failures. As a result, although the hardware

platform may continue to run, the mission-critical software application can be rendered useless by bad data.

Express5800/ft servers' hardware-based fault tolerance addresses these failings through a unique "pair and partner" technology. Hardware is engineered to perform logic self-checking, and its main subsystems (CPU, PCI, memory, hard drives, and power supply) are physically replicated. Self-checking logic is resident on each major circuit board, to detect and immediately isolate failures. Because of this replication, normal application processing continues, even if one of the components should fail.

In this manner, fault tolerance is transparent to the software application, and no performance degradation occurs if a component fails. Also, logic self-checking allows data errors to be isolated at each clock tick, assuring that erroneous data never enters the bus.

Serviceability

Conventional high-availability networks have limited or no remote service notification functionality. As a result, a problem typically reveals itself to end users only through a system crash. All support must be supplied onsite by the vendor's field personnel. This can cause considerable delays in repairing the conventional system's supporting mission-critical applications.

Express5800/ft servers are equipped to constantly monitor their own operation. Because existing components can be replaced while the system is running, the change is transparent to the application and users. Likewise, components can be upgraded, and maintenance and backups can be performed online.

Total Cost of Ownership (TCO)

Total cost of ownership of an availability solution is not only the initial hardware purchase price but also the software cost and the administration and downtime costs over the life of the system. Considering just the hardware cost may result in a long-term total cost of ownership that is higher for a lower-end solution than a high-end solution. This is why NEC recommends evaluating the total cost of ownership (TCO) instead of focusing on the purchase price.

Hardware Cost. Creating cluster networks requires reconfiguring existing systems or purchasing new equipment. Reconfiguring conventional servers for high availability requires custom network configurations. These are usually Ethernet or token ring networks that must be redundantly configured with duplex LAN interface cards and custom LAN failure software scripts.

Express5800/ft servers are single units with a low initial cost. They contain standard Intel® Pentium® III processors with a fault tolerant chipset that provides synchronization control, fault detection, disconnection, and recovery. Express5800/ft servers are designed to continue processing if a hardware fault occurs.

Software Cost. Linking conventional systems into an active/active cluster network typically incurs significant additional costs. Just as a user must purchase duplicate hardware components, multiple copies of the operating system and applications are required. Cluster networks also require custom software applications designed to initiate recovery from a crash and handle shared processing. This custom code can become complex and expensive, depending on the sophistication of the application. Furthermore, it is virtually impossible to code for every possible failover scenario; programmers must select the most likely scenarios and remain on standby to handle unanticipated failures.

In contrast, Express5800/ft servers require only a single copy of the operating system and applications. They run standard Microsoft® Windows® 2000 applications, eliminating the need for customized or modified software.

Administration Cost. In an age of increasingly expensive and scarce IT resources, the cost of personnel can easily eclipse the purchase price of the hardware itself. Clustered high-availability systems typically depend on the programming and administrative skills of the IT staff. Businesses contemplating the use of high-availability systems must consider the costs of personnel required for constant monitoring and fine-tuning performance. IT staff are also needed to synchronize files and memory and administer two separate systems so that they look identical.

Because Express5800/ft server architecture has no such dependencies, organizations can deploy revenue-generating applications much more quickly without overburdening IT resources.

Downtime Cost. Depending on how a server is used, the cost of an hour of downtime can reach hundreds of thousands of dollars. In the financial services industry, applications related to brokerage operations or online credit card sales authorization can incur millions of dollars per hour in downtime.

Industry Downtime Costs

Application	Cost per minute
Call location	\$27,000
Number portability	\$14,400
ERP	\$13,000
Supply chain management	\$11,000
Electronic commerce	\$10,000
Internet banking	\$7,000
Universal personal services	\$6,000
Customer service center	\$3,700
ATM/POS/EFT	\$3,500
Messaging	\$1,000

Source: The Standish Group International, Inc. ©2001

The NEC Solution

Fault tolerant solutions once carried a premium price tag and, when evaluated solely on the basis of initial purchase price, appeared expensive compared to high-availability alternatives. NEC has designed Express5800/ft servers to be priced competitively with an equivalently configured two-node, high-availability cluster.

Major subsystems in Express5800/ft servers are commodity hardware. The server uses processors, memory, disks, and power subsystems identical to those used in conventional servers from first-tier vendors. NEC Computers then adds value by enhancing the server's availability features. This continuously available server provides one of the highest levels of security against downtime for a single system/service price. That price is all-inclusive, in that it includes all of the elements for continuously available computing.

Express5800/ft server	Clustered system
<ul style="list-style-type: none"> • A single server 	<ul style="list-style-type: none"> • Multiple computers function as a single server
<ul style="list-style-type: none"> • A single copy of the OS 	<ul style="list-style-type: none"> • Multiple copies of the OS
<ul style="list-style-type: none"> • A single off-the-shelf copy of each application 	<ul style="list-style-type: none"> • Multiple copies of programs, either off-the-shelf or custom programmed to handle cluster-based failure recovery

Fault tolerant Express5800/ft servers are the only computing platform for Windows 2000 applications in which every major aspect of the system is designed for uninterrupted uptime from the start.

Express5800/ft servers offer one of the industry's highest levels of system availability in a cost-effective platform that is faster to implement, easier to operate, and simpler to manage than comparable alternatives.

This document is for informational purposes only. NEC Computers Inc. shall not be liable for technical or editorial errors or omissions contained herein. The information in this document is subject to change without notice.

NEC and ESMPRO are registered trademarks of NEC Corporation and/or one or more of its subsidiaries. Simplified Availability is a service mark of NEC Computers Inc. Stratus and the Stratus logo are registered trademarks and the Stratus 24x7 logo is a trademark of Stratus Technologies International, S.à.r.l. Microsoft and Windows are registered trademarks of Microsoft Corporation. Intel and Pentium are registered trademarks of Intel Corporation. All other trademarks and trade names are the property of their respective owners. ©2001 NEC Computers Inc. All rights reserved.