

the *Availability Digest*

www.availabilitydigest.com

Stratus' ftServer Flexes Its Recovery Muscle

August 2010

The ftServer from Stratus Technologies is a hardware-based, fault-tolerant server for running Windows and Linux applications. Its scalability and recoverability were recently measured in June, 2010, by Principled Technologies, Inc.



Under a test commissioned by Stratus and NEC¹ to assess ftServer's scalability and its resilience to catastrophic events, Principled Technologies stressed the ftServer incrementally by adding virtual CPUs (vCPUs) to a single VMware virtual machine (VM). At peak load, one of the two redundant servers - memory, processor, I/O subsystem, disks, and all – was pulled from the chassis to measure the ftServer's recovery time from such a catastrophic failure.

The results showed performance measured in tens of thousands of orders per minute. Scalability was reasonably linear up to four vCPUs. With eight vCPUs running, the induced massive fault caused by removing one of the redundant processors resulted in virtually no performance degradation and no loss of application data or its integrity.

The ftServer

The ftServer architecture is described in some detail in our earlier article, Fault Tolerant Windows and Linux from Stratus, which was published in the September 2007 issue of the *Availability Digest* (http://www.availabilitydigest.com/public_articles/0209/stratus.pdf). ftServers from Stratus Technologies (www.stratus.com) provide plug-and-play fault tolerance for Windows and Red Hat Linux applications. Using Intel Xeon chips in a dual modular redundancy architecture, ftServers bring extremely high availability – five 9s and beyond – to the industry standard marketplace at affordable prices.

Dual Processors



The high availability obtained by the Stratus ftServer product line is achieved by running all applications on dual processors that are lockstepped at the memory-access level. The dual physical processors appear to the outside world as a single logical processor.

Should there be a disagreement between the processors, one of the processors has suffered a fault and is taken out of service. The faulty processor can be replaced and synchronized with the operational processor while the system continues to run.

¹ Stratus and NEC have worked together for ten years to bring fault tolerance to the Intel Server Architecture. See the NEC Express 5800/R320 at http://www.nec.com/global/prod/express/product/f_tolerant/index.html.

Each processor is packaged as a 2U module. The processors are interconnected via a small, passive backplane. A complete standalone ftServer, therefore, has a 4U form factor.

I/O Subsystem

Each 2U module contains its own I/O subsystem. Both I/O subsystems also have their own fault detection and isolation logic. In normal operation, peripheral devices are driven by both I/O subsystems via a multipath connection. However, if an I/O subsystem detects a malfunction, it will remove itself from service.

Each logical processor is connected to both I/O subsystems. In this way, any combination of one logical processor failure and one I/O subsystem failure will not render the system inoperable.

Integrated Disks

Each processor contains not only its microprocessor cores, memory, and I/O subsystem but also three 500 gigabyte integrated disks. Thus, both logical processors have direct access to 1.5 terabytes of local storage. This storage is mirrored between the processors, providing 1.5 terabytes of fault-tolerant storage within the ftServer itself.

Operating Systems

The Windows and Red Hat Linux operating systems that Stratus supports are those that are commercially available to anyone. Therefore, the ftServers are application binary interface (ABI) compatible with Windows and Linux applications. Any application that can run under Windows or Red Hat Linux on an industry-standard server can run without modification on an ftServer. The installation and administration procedures are identical. The user should see no difference except for downtime – and that is the big difference.

The Test Configuration

System Configuration

The Principled Technologies' test was run on a Stratus ftServer 6300. Each of the redundant processors provided dual sockets for quad-core microprocessor chips. Therefore, the logical ftServer had access to eight cores.

Both physical processors were configured with 48 GB of memory, giving the logical processor 48 GB of memory.

The ftServer was connected via fibre channel to a NEC D3-10 SAN that contained the virtual machine.

VMware ESX 4.0 was configured to be the hypervisor for the test. The test did not use the VMware Fault Tolerance or the VMware High Availability products to provide availability. The availability characteristics tested were strictly those that were native to the ftServer.

A single VM running Microsoft Server 2008 R2 as its guest operating system was configured for the application. The number of virtual CPUs running in that VM was varied from one to eight, the maximum number of cores available.

The benchmark test was database-intensive. Microsoft SQL Server was used as the database manager for the application. Three virtual disks were configured – one for the operating system, one for the database, and one for the SQL logs. A 5 GB application database was built.

The Test Driver

The open-source “DVD Store Version 2” (DS2) benchmark² was used to exercise the system. DS2 simulates an online ecommerce DVD store. Simulated customers log in, browse movies by actor, title, and category, and purchase movies. The title and actor searches use full-text search. DS2 provides other administrative functions, such as adding new customers. Each simulated customer results in an order.

A single instance of DS2 was configured with 32 threads to run on a control server. DS2 is structured to use a web interface to submit requests to a database server. However, in this test, the web interface was bypassed so as to maximize the load that could be placed on the ftServer.

DS2 measures *orders per minute (OPM)* as its metric. Each test was run for 30 minutes with OPM measurements taken every second. The last OPM value at the end of the test was reported as the test result.

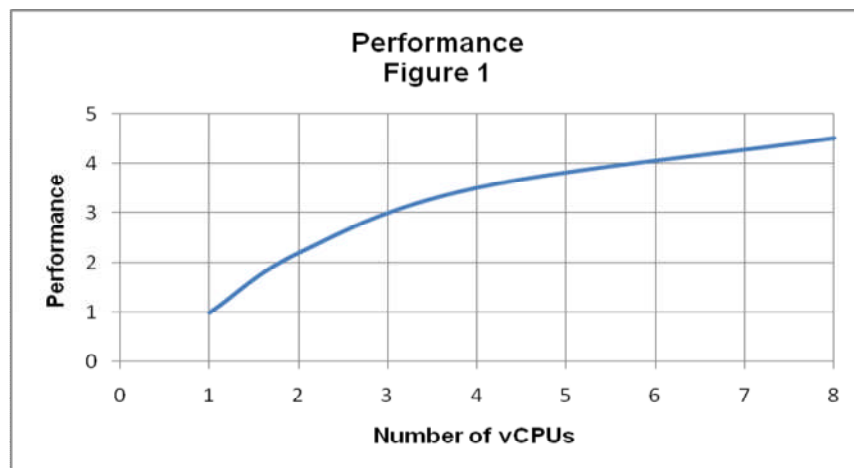
Test Results

Scalability

The scalability test measured scaling in a single VM as vCPUs were added. This test is particularly important for database applications that must scale beyond one vCPU. Note that the test system used a 5 GB database in a processor with 48 GB of memory. Therefore, the database was likely to be cached in memory; and the test measured raw processor power. Test results are summarized in Table 1 and Figure 1.

No. of vCPUs	OPM	Orders Per Second	Orders Per Sec. per vCPU	Relative Capacity
1	16,771	280	280	1.0
2	36,833	614	307	2.2
4	59,409	990	248	3.5
8	76,129	1269	159	4.5

ftServer Scalability
Table 1



² <http://www.delltechcenter.com/page/DVD+Store>

The ftServer scaled fairly linearly up to four vCPUs, as shown in Table 1 and in Figure 1. Doubling the number of vCPUs from one to two resulted in more than twice the capacity. This was perhaps due to SQL Server optimization for multiple CPUs. Moving to 8 vCPUs showed saturation of the system and was used to test recoverability.

Recoverability

The recoverability test measured the ability of the ftServer to recover from a catastrophic failure. Running at maximum capacity with eight vCPUs, one of the redundant lockstepped servers was pulled from the chassis - processor, memory, I/O channels, disks, and all. The processing capacity of the remaining processor was then monitored every second to determine the impact on the ftServer.

The test showed an initial drop in capacity that was barely measurable. The transaction processing rate dropped from 76,129 OPM to 76,010 OPM, a drop of 0.16%.

The transaction-processing rate began to recover twelve seconds after the fault, and it returned to its pre-event level in 26 seconds. No application data was lost nor its integrity compromised as a result of the fault.

The recoverability test demonstrated two important attributes of the ftServer's fault tolerance:

1. A catastrophic fault has virtually no visible impact on the functioning of the ftServer.
2. There is no measurable lockstepping overhead that detracts from the capacity of the ftServer while it is running in dual-processor mode.³

Summary

The independent testing by Principled Technologies showed linear scalability over a 4:1 range for a virtual machine running on a Stratus ftServer as virtual CPUs were added. It also demonstrated that there was virtually no measurable impact on the operation of an ftServer as it recovered from a catastrophic fault under heavy load.

The test further showed that the lockstep, hardware-based, fault-tolerant approach used by the ftServer imposed no overhead on the system during normal redundant operation.

Principled Technology Inc.

Principled Technology, Inc.,⁴ is an independent technology-assessment firm located in Durham, North Carolina, USA. The detailed test procedure used by Principled Technology may be found at http://www.principledtechnologies.com/clients/reports/NEC/ftServer_Stratus_NEC.pdf.

³ As opposed to software fault-tolerance techniques that impose a processing overhead during fault-tolerant operation.

⁴ <http://www.principledtechnologies.com/>