"Achieving Century Uptimes"
An Informational Series on Enterprise Computing

As Seen in *The Connection*, A Connect Publication
December 2006 – Present

About the Authors:

Dr. Bill Highleyman, Paul J. Holenstein, and Dr. Bruce Holenstein, have a combined experience of over 90 years in the implementation of fault-tolerant, highly available computing systems. This experience ranges from the early days of custom redundant systems to today's fault-tolerant offerings from HP (NonStop) and Stratus.

Gravic, Inc.
Shadowbase Products Group
17 General Warren Blvd.
Malvern, PA 19355
610-647-6250
www.ShadowbaseSoftware.com

Achieving Century Uptimes
Part 21: Active/Active NonStop Blades
March/April 2010

Dr. Bill Highleyman
Paul J. Holenstein
Dr. Bruce D. Holenstein

As we move into the second decade of the 21st century, we have pretty much left behind the primary issue on which we in data processing have previously focused – performance. Our systems have become blazingly fast, achieving thousands of transactions per second at a per-transaction cost measured in millicents.[1]

Our focus must now shift to another goal – availability. With today's demands for 24x7 service, three and four 9s of availability (up to eight hours of downtime per year) are simply not acceptable. Today's systems must operate flawlessly and continuously to avoid massive downtime costs, compliance issues, and – yes – even worldwide negative publicity in the press.

Continuous availability is achievable with active/active systems.[2] HP's NonStop Blade Systems are an excellent choice to implement active/active systems. In this article, we review some of the many advantages of these systems to achieve continuous availability.

What is an Active/Active System?

In an active/active system, multiple geographically-distributed processing nodes cooperate in a common application using distributed copies of the application database. Should there be a failure anywhere in the application network, all that needs to be done is to reroute transactions to a surviving node.

Recovery from a fault can be so rapid that users are unaware of the fault – in effect, there has been no outage, and continuous availability is achieved. There are many examples of active/active systems that have been in production for over a decade that have never had an outage.

The NonStop Blade System

Let us first review the attributes of NonStop Blade Systems that make them so applicable to active/active systems.[3]

[1] A $10,000,000 system amortized over two years and processing on average 1,000 transactions per second has a transaction cost of 0.016 cents per transaction – 16 millicents per transaction.

[2] What is Active/Active?, *Availability Digest*, October 2006.

[3] W. H. Highleyman, HP's NonStop Blades, *The Connection*; May/June 2009.
HP's NonStop Blades, *Availability Digest*; August, 2008

From a configuration viewpoint, an HP NB50000c NonStop Blade System is similar in many respects to a NonStop S-series or Integrity system. A fully-configured blade system delivers twice the power of the HP Integrity NS16000, until recently HP's largest NonStop server, in half the footprint. Existing applications can be ported seamlessly to the new bladed system.

The Processors

Based on dual-core Itanium processors, the new multicore architecture is called NSMA, the NonStop Multicore Architecture. The NSMA system uses HP's c7000 blade processors running the NonStop operating system. An NSMA system comprises two to sixteen blade processors. Each blade is driven by an Intel dual-core (dual instruction-processing-unit, or IPU) Itanium microprocessor with up to 48 gigabytes of memory.

In addition to the standard blade configuration, an NSMA blade carries one additional card – a ServerNet card for connecting to the dual ServerNet fabric There are two redundant ServerNet fabrics for a NonStop Blade System. A full NSMA system can provide up to 48 ServerNet I/O connections.

Performance tests of NonStop Blade Systems using HP's Order Entry benchmark show that the processing power of an IPU is substantially that of an NS16000 processor. Therefore, since there are two IPUs, the throughput of a NonStop blade is roughly twice that of an NS16000 processor. Order-Entry benchmark tests show the following per-logical processor capacities:

NS16000	165 tps
NonStop Blade	349 tps

Since an IPU is equivalent to an NS16000 processor, the response times are roughly equivalent.

The Enclosure

NSMA uses the standard HP c7000 enclosure to house the blade processors. This is a 10U chassis that can hold up to eight blades. It comes prepopulated with six power supplies and 10 fans. Though only some of these are required to power and cool the enclosure, this provides spares and, more importantly, room for growth as quad- and eight-core chips become available.

Only one processor enclosure may be put into a cabinet because of power, cooling, and weight restrictions. Therefore, a fully-configured NonStop blade system with sixteen processors requires two cabinets, half the space of an NS16000 but with twice the processing power.

The I/O Subsystem

The CLIM (Cluster I/O Module) is the storage and communication interface for NSMA. It supports SAS (serial-attached SCSI) disks and IP interconnects. The CLIM platform is an HP Proliant DL385 rack-mounted server with a 2U height. A NonStop Blade System can support up to 44 CLIMs in addition to two required storage CLIMs.

The storage CLIM supports dual-ported SAS disks. 25 disks can fit into a 2U rack-mounted enclosure and can either be 72 gigabyte, 15,000 rpm disks or 146 gigabyte, 10,000 rpm disks. A fully-configured NonStop Blade System can contain over 160 terabytes of mirrored storage and can support data transfer rates up to twelve gigabytes per second.

The communication CLIM supports both IPv4 and IPv6 with IP Security (IPSec). A fully-configured NonStop Blade System can support up to 220 Gigabit Ethernet channels.

System Management

The management of NonStop blades is integrated into the wider HP system-management products, especially HP's Systems Insight Manager (SIM), which provides management of heterogeneous systems across the entire enterprise environment. All NonStop management tools and third-party products continue to provide their existing system management functionalities.

Going Active/Active with NonStop Blade Systems

NonStop Blade Systems indeed pack a lot of power into a small package. But what does this have to do with building active/active systems? As we shall see, NonStop Blade Systems bring many advantages to the deployment of active/active systems, including greater availability, smaller footprint, and lower cost.

Availability

The use of fault-tolerant nodes in an active/active system can reduce the number of nodes required to achieve a desired availability. There are three reasons a user may be denied service by a redundant system:

- One node has failed, and users are being failed over to another node (this can take hours in some cases).
- The failover to a backup node fails.
- Enough nodes fail so that the system can no longer handle the application load.

In active/active systems, failover is virtually instantaneous (it can be done in subseconds), and it can be so fast that users are not aware that there has been an outage. Furthermore, successful failover is assured because the other nodes in the application network are known to be operational – after all, they are currently processing transactions. Therefore, the only reason that users will be denied service is if enough nodes in the network fail so that there is not enough capacity to handle the application load.

Let us assume that we have the choice of using fault-tolerant nodes with an availability of four nines or industry-standard (ISS) x86 nodes with an availability of three nines (typical availabilities for these platforms). Furthermore, we want to achieve an availability of seven nines. It can be shown that two nodes in an active/active configuration double the nines of one

node.[4] Thus, a two-node active/active system using fault-tolerant nodes will have an availability of eight nines, more than satisfying our requirement. However, a two-node active/active system using ISS nodes will only have an availability of six nines, an order of magnitude less than we require. In order to achieve our desired availability with ISS nodes, we need to add a second spare node, resulting in a three-node active/active ISS system that can withstand the loss of any two nodes. Each of these three nodes must be configured to handle the entire application load since it may be the only surviving node.

This leads to a general rule in availability, which we have called the *NonStop Maxim: An active/active system will generally require fewer nodes if fault-tolerant nodes are used rather than industry-standard nodes.*[5]

Node Count

Consider a very large active/active system – say one that requires two sixteen-processor Integrity systems to handle the application load. To survive a processor failure, a third system must be added, creating a three-node active/active system. A sixteen-processor NonStop Blade System has twice the capacity of its Integrity equivalent. Therefore, the application load can be handled with only two NonStop Blade System nodes while surviving a node failure.

There is an additional benefit here, and that is availability. Let us assume that the availability of an Integrity system and a Blade System are the same. There are three ways that two nodes in a three-node Integrity system can fail and bring down the application network – nodes 0 and 1, 0 and 2, and 1 and 2. There is only one way that a two-node Blade System can fail. Therefore, the NonStop Blade System will inherently have *three times* the reliability of the Integrity system.

Smaller Nodes, More Reliability

We have just seen how the number of nodes in an active/active system can affect its availability – the more nodes there are, the less is its availability because there are more failure modes. This same rule holds for processors in a NonStop system – the more processors there are, the more ways the system can fail because of a dual-processor failure.[6]

Thus, in our previous example, if we instead compared a four-node Integrity active/active system to a four-node Blade System active/active system with the same capacity, we would be looking at four sixteen-processor Integrity systems versus four eight-processor Blade Systems. Because the Blade Systems in this example have half the number of processors than the Integrity systems, we would expect them to be substantially more reliable.[7]

---

[4] See Rule 2, "Providing a backup doubles the nines," Chapter 1, The 9s Game, *Breaking the Availability Barrier: Survivable Systems for Enterprise Computing*; AuthorHouse; 2004.

[5] See Rule 39, Chapter 1, Achieving Century Uptimes, *Breaking the Availability Barrier II: Achieving Century Uptimes with Active/Active Systems*, AuthorHouse; 2007.

[6] See Rule 3, "System reliability is inversely proportional to the number of failure modes," Chapter 1, The 9s Game, *Breaking the Availability Barrier: Survivable Systems for Enterprise Computing*; AuthorHouse; 2004.

[7] More specifically, a sixteen-processor system has 120 ways that two processors can fail, and an eight-processor system has 28 ways that two processors can fail – over a 4:1 availability advantage for the smaller system. Things aren't quite this bad (or good) because not all cases of a dual-processor failure will take a system down. For

Total Cost of Ownership (TCO)

There are several factors that speak well for a reduced acquisition (basic) cost and a reduced TCO when using NonStop Blade Systems for active/active systems instead of other equivalent systems. Research by The Standish Group shows the following comparisons:[8]

NonStop Blade System versus	Basic Cost Savings	TCO Savings
NonStop Integrity System	26%	11%
NonStop S-Series	55%	42%
IBM z10	30%	24%
IBM z9	51%	36%

HP's software licensing polices support the use of active/active systems. No longer is software licensed by system – this had always been an impediment to breaking up a single system into multiple active/active nodes. Software is now licensed by processor. Therefore, a four-node active/active system with four processors per node has substantially the same software license fees as a single sixteen-processor node.

NonStop Blade Systems leverage HP's continuing cost/performance improvements in their broader Integrity product line. As performance improvements and cost reductions are made in the blades that are used across the board in HP's products, these improvements come seamlessly to NonStop Blades. No longer must special hardware development be supported for NonStop. NonStop efforts can now be focused on fault-tolerant software, which is the differentiating factor for NonStop systems.

A major factor recognized in the industry is that NonStop systems are especially easy to manage. In particular, the number of system administrators needed to manage a large NonStop system is substantially less than that required for any other system.[9] This has a major impact on TCO, especially when support for multiple nodes in diverse locations is required.

Data Center Footprint

In addition to the higher availability that comes with the use of NonStop Blade Systems, the resultant system will have one-quarter the footprint of an Integrity system in the data center (twice the capacity at half the size). This can substantially impact data center facilities costs.

---

example, if both the primary and backup disk processes for a critical file are located in the cpu's that fail, that will bring down the application at that node, whereas the failure of two processors with non-related disk processes will not. Furthermore, with today's highly reliable hardware, system faults are more likely to be caused by software bugs, operator errors, and environmental faults (power, cooling, etc.). Nevertheless, the reliability advantage of smaller systems can still be significant.

[8] Trends in IT Value, *The Standish Group*; 2008.
Megaplex: An Odyssey of Innovation; *The Standish Group*; 2009.

[9] The Standish Group reports that one customer is managing 19 NonStop systems comprising more than 200 processors with five full-time system administrators. See Megaplex: An Odyssey of Innovation, *The Standish Group*; 2009.

HP support

Last, but certainly not least, is the level of support that is available from HP and its third-party partners for implementing and managing active/active systems. Active/active technology started with NonStop over a decade ago, and HP and its partners remain the primary source for active/active systems.

Summary

Utilizing the increased computing density of NonStop Blade Systems, active/active systems can now be implemented that provide even higher availability, greater capacity, and lower cost of ownership than ever before. As HP blades move in the near future from dual-core to quad-core and eight-core processors, this exciting trend will continue. Now is an opportune time to consider moving your most critical applications into an active/active environment – the technology is proven with many production deployments in the field. The elimination of downtime costs may more than justify the costs of your move.