the Availability Digest

www.availabilitydigest.com
@availabilitydig

Pathway – HPE NonStop's Application Environment

May 2016

At a recent fault-tolerant symposium, I presented an overview of HPE NonStop systems. I stressed the immense scalability of these systems as well as their ability to survive any single fault (and in some cases, multiple faults).

I was struck by the interest shown by the audience in the NonStop process monitor, NonStop Pathway. I came to realize that Pathway is the foundation for application fault tolerance, scalability, and load balancing in NonStop systems. Pathway removes the concerns of these important attributes from the application programmer and implements them 'under-the-covers.'

In this article, we review the architecture of NonStop systems and explain how Pathway provides applications with fault tolerance, scalability, and load balancing with no effort on the part of the application programmer.

The NonStop Architecture

HPE NonStop systems began their life in 1974 as Tandem NonStop systems marketed by Tandem Computers, Inc. Tandem Computers, Inc. was acquired by Compaq in 1997. Compaq was acquired by HP in 2001. In 2015, HP split into two companies, HP and HP Enterprise (HPE). HP retained the PC and printer business, and HPE inherited the systems business. The fault-tolerant systems described herein are now known as HPE NonStop systems.

Right from the beginning, NonStop systems have featured a 'massive parallel processing' (MPP) architecture. In this unique shared-nothing architecture, each processor comes with its own copy of the operating system. There is no shared memory. The system relies heavily on interprocess communication between peer processes.

In contrast to conventional high-availability clusters, the NonStop system provides full hardware/software fault tolerance with immense linear scalability. Each NonStop system can be configured with up to sixteen processors interconnected by a duplex high-speed bus. The original bus was called Dynabus. This was later replaced with a duplexed fabric called ServerNet. With the introduction of the NonStop X, ServerNet has been superseded by the industry-standard Infiniband.

With a high-speed network fabric clustering the system, the limitation of sixteen processors became irrelevant. Much larger configurations with hundreds of processors operating in parallel can now be run very efficiently. Interestingly, after more than four decades of service, the NonStop architecture essentially has remained the same.

The current processors of NonStop systems using ServerNet are implemented with the Intel Itanium microprocessor. These systems are designated NonStop i systems (Figure 1). NonStop i systems are now being superseded by NonStop X systems (Figure 2) based on the industry-standard x86 architecture and interconnected via Infiniband.

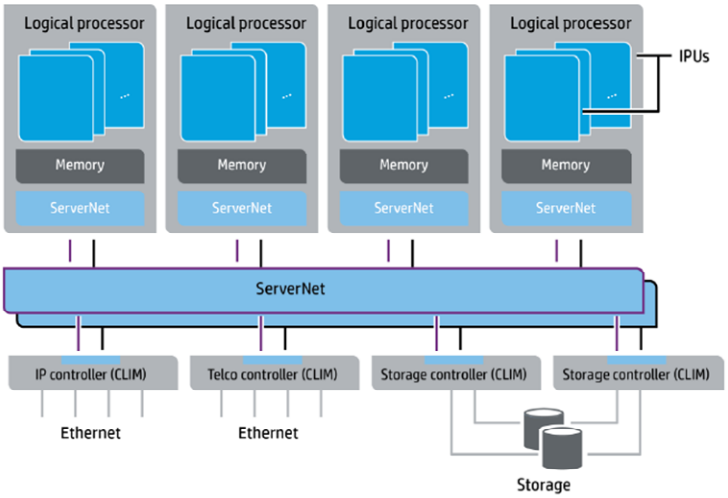

Figure 1: NonStop i

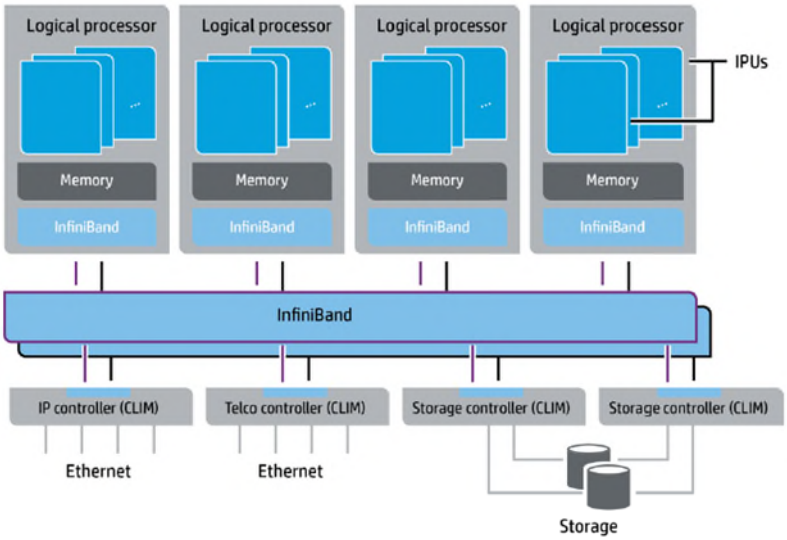

Figure 2: NonStop X

Both product lines leverage multicore operation of the underlying processor chip. The use of Infiniband (which is significantly faster than ServerNet) also opens the opportunity for NonStop X systems to connect directly to other systems such as Linux front ends without having to go through an internal controller (a CLIM). In addition to the greater speed of the x86 microprocessors as compared to the Itanium chips in the NonStop i systems, Infiniband allows NonStop X systems to provide significant additional processing capacity compared to the NonStop i systems.

Architecture Fundamentals

Fault-Tolerance

In NonStop systems, fault tolerance is achieved via checkpointed process pairs. Each critical process is configured with a backup process running in another processor. Whenever the state of the active process changes, it sends a checkpoint message to its backup synchronizing the backup process state to that of the active process. Thus, if the active process fails (typically due to a processor failure), the backup process immediately takes over with no interruption in processing.

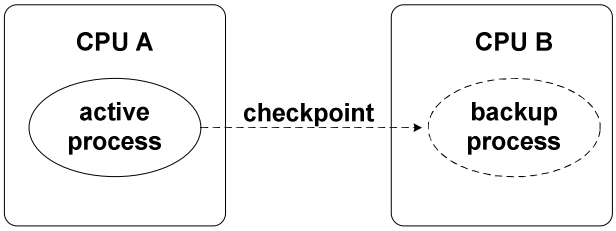

Figure 3: Checkpointed Process Pair

Initially, application programmers were expected to write their own checkpointed process pairs for critical processes. However, this often proved to be too complex for ordinary application programmers. Therefore, the achievement of fault tolerance for application programs was moved to a new system software facility called 'Pathway," as described later. As we shall see, applications written in a Pathway environment can be provided with fault tolerance with little effort on the part of the application programmer. Failover following a fault takes a little longer (seconds) but is fast enough to have little impact on the users of an application. Critical system processes (such as Pathway itself, the NonStop Transaction Monitor TMF, and disk processes) are all implemented as checkpointed process pairs within the NonStop system software.

Expandability

Communication between processes in the processors is via interprocess messages sent over the backbone fabric. Access to external devices is via CLIMs (CLuster I/O Modules) also connected to the ServerNet or Infiniband fabric. There is no common memory used for interprocess communication. Therefore, there is no fundamental limit to the expandability of a system, which is a near-linear function of the number of processors in the system. In fact, up to 255 systems (4,080 processors) can be linked in a cluster, and each additional processor added to the cluster increases the system capacity by 98% of that processor's capacity.

Pathway

The fault-tolerance and immense scalability of NonStop systems are powerful enhancements to bring to an application. But how is an application programmer who is not an expert in NonStop fundamentals expected to take advantage of these features?

The answer is Pathway. The Pathway application manager/provisioning facility provides the runtime execution environment for Pathway applications. As such, it automatically provides fault tolerance, scalability, and load balancing for applications running within its environment. A high-level overview of a Pathway environment is shown in Figure 4.

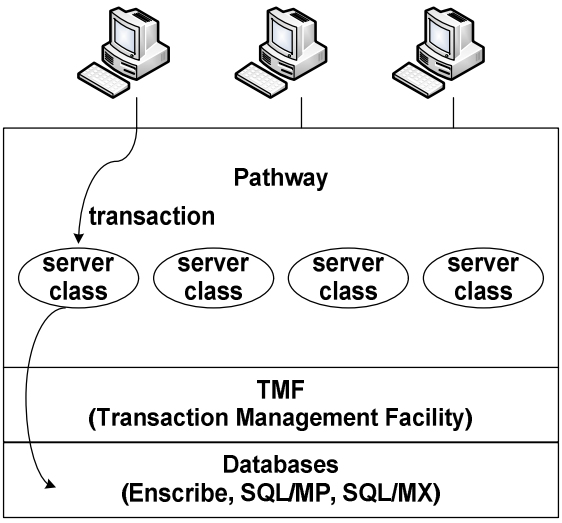

Figure 4: A Pathway Environment

Server Classes

Pathway comprises a collection of *server classes*. A server class contains one or more identical stateless servers designed to process a particular transaction. Typically, a server in a server class will receive a transaction and will route the database updates to the appropriate databases. The NonStop Transaction Management Facility (TMF) is responsible for maintaining the ACID properties (Atomicity, Consistency, Independence, and Durability) of the transaction as it updates the database. NonStop systems support three databases – Enscribe, a relational file system, SQL/MP, a SQL database, and SQL/MX, a newer SQL database that complies more closely with the ANSI SQL standard.

It is the server classes that provide the power of the Pathway environment. The servers in a server class can be spread among several processors in a NonStop system. If a server class becomes overloaded, Pathway can spawn additional servers for the server class. If the load on a server class falls to the point where all of its servers are not needed, it can terminate some of its servers.

Scalability

It is the Pathway server classes that provide the application scalability provided by Pathway. If the load imposed by a transaction starts to overload a server class, Pathway simply spawns more servers for that server class. The servers can be assigned to processors that are carrying less load. Thus, as an application needs more processing power, Pathway automatically assigns more processing power to the application by expanding the server class.

If the load imposed by a transaction begins to recede, Pathway will terminate some of the servers in the server class servicing that transaction. As already described above, the Pathway system ‘breathes’ by swiftly acquiring more system resources during high load situations and by quickly releasing those resources when they are no longer needed.

Fault Tolerance

It is also the server classes that provide fault tolerance to an application. Should a processor in a NonStop system fail, Pathway will route further transactions only to servers running in surviving

processors. In cooperation with TMF, those transactions affected by a failed processor are recovered automatically with no data loss or duplication.

Load Balancing

Pathway balances the transaction processing load across processors by routing transactions to servers that are running in different processors. No separate load balancer facility is required.

Pathway Domains

Multiple identical Pathway environments can be configured as a Pathway domain that behaves as a single, large Pathway application. Any of the Pathway environments within a domain can be taken down for maintenance while the remaining environments within the domain continue processing work with no interruption. This enables online rebalancing of Pathway environments without planned outages. It also relaxes the configuration limits of a Pathway system.

Server classes are replicated across all Pathway environments in the domain. Requests are automatically load-balanced across the domain.

A remote Pathway environment can be configured as part of a domain. Thus, a logical Pathway server class can span multiple NonStop systems. This allows a NonStop node in a domain to be taken down for maintenance while the application remains available on other nodes. Pathway automatically routes requests only to the available nodes in a domain. This also increases scalability because a Pathway application can span multiple NonStop nodes, which can actually be part of a bigger ServerNet or Infiniband cluster containing hundreds of NonStop processors.

Multi-Tenant Operation

Pathway environments can be configured to run the same application concurrently for different customers while reliably separating their access rights and data (this is important for service providers, for instance). Of course, different Pathway environments running entirely different applications can also run concurrently on the same NonStop system. In this regard, the NonStop system resembles the operational characteristics of a traditional mainframe system.

Summary

NonStop systems provide several significant advantages over other processing systems, including immense scalability and fault tolerance. By using the Pathway runtime environment, application programmers can make full use of these characteristics without specialized programming knowledge. Pathway automatically scales applications as needed across multiple processors and even multiple NonStop systems, and it recovers from any processor fault without user involvement.

Acknowledgements

Thanks to our many subscribers who commented on this article and especially to Gerhard Schwartz who provided several corrections and extensions.

Information for this article was taken from the following source:

HP NonStop Pathway with TS/MP Software, *HP Data Sheet.*
http://www8.hp.com/h20195/V2/getpdf.aspx/4AA0-3689ENW.pdf?ver=7