Building a disaster-proof data center with HP OpenVMS

Executive summary

HP OpenVMS has long been known as the "gold standard" for disaster-tolerance[1] and has long been hailed in the industry for its security and reliability.[2] Over the years it has helped many commercial users avert business tragedies by protecting their critical operations and keeping their business processes up and running *despite and during* major disasters. These disasters range from localized human error to catastrophes such as that which occurred on September 11, 2001. Yet because enterprises generally do not like to speak publicly about any type of disaster, it has been difficult to demonstrate the capabilities of OpenVMS to those who remain skeptical.

That has now changed. In a successful and dramatic demonstration of the disaster-tolerant capabilities of OpenVMS, Hewlett-Packard set up multi-site disaster-tolerant cluster configurations, including an OpenVMS cluster, and then destroyed the equivalent of one entire data center, with all of its servers, SAN, storage, and network equipment. In the wake of this event on May 17, 2007, the applications on OpenVMS paused only briefly before continuing unaffected and with no loss of data or transactions, thus proving HP's claim that OpenVMS environments can, indeed, provide up to 100% application availability.

Full disaster tolerance with OpenVMS is qualified at distances up to 500 miles. However, customers do take advantage of the functionality at greater distances. At least one OpenVMS customer utilizes this capability at an inter-site distance of greater than 3,000 miles.[3] Synchronous data replication with OpenVMS host-based volume shadowing software has been tested at distances up to 60,000 miles (keeping in mind that the speed of light does limit the efficiency at greater distances).

With the costs of application downtime capable of reaching tens of millions of dollars per minute or more, it is easy to see the bottom-line benefit and short payback period of an OpenVMS clustered environment.

Moreover, in addition to the benefits of increased application uptime and reduced data loss, the fact that OpenVMS clusters are able to function as a single, load-balancing system also delivers reduced cost of management. Because of this characteristic, the cost and effort of management does not rise in proportion to the number of nodes in a cluster. In fact, it rises very little as more nodes are added.[4]

[1] Illuminata, Inc. Research note: "The 'I' is for Integrity," 20 March 2006, page 3.

[2] Ideas International, Inc. Technology Trend: "HP Extends Integrity portfolio with integrated solutions," 9 February 2005, page 5; and D.H. Brown Associates: "The United States User View of Business Continuity and Recovery," June 2004, page 49.

[3] Note that this level of separation exceeds the officially supported distance. However, the customer's application is able to tolerate the inherent latency associated with distance and the speed of light.

[4] For more information on management efficiency, see www.hp.com/go/openvms.

What was done in the test

The purpose of the test was to demonstrate how an OpenVMS disaster-tolerant cluster configuration can survive the destruction of an entire data center. A cluster was set up in a multiple data-center configuration separated by a safe distance. The cluster nodes included both HP Integrity servers and HP AlphaServer systems, demonstrating both the mixed-architecture and disaster-tolerant capabilities of OpenVMS.

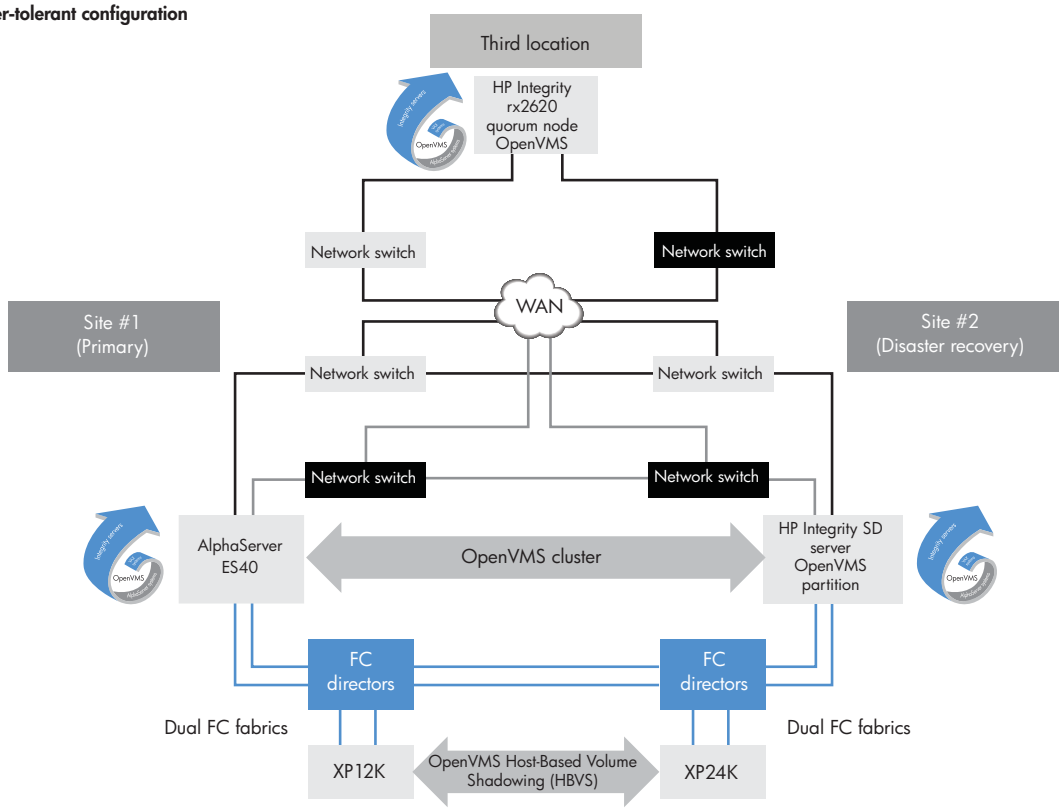
While applications were running, a controlled explosion blew up one of the data centers.[5] The remaining OpenVMS systems in the cluster automatically reconfigured, assuming full responsibility for all of the operations at the obliterated data center while maintaining the entire cluster's full functionality prior to the explosion.

One aspect of the test involved running a simple application that constantly updated a shared data file. In this simulation, the success of each transaction depended on updated data from the previous transaction—in much the same way a successful automated teller machine (ATM) withdrawal may depend on the correct results of a previous deposit transaction. In the test, the application provided an easy way to verify that no data was lost in conjunction with the disaster.

Another simple application running simultaneously wrote a record to a shared data file every 10 milliseconds, recording the time of the current write operation and the actual elapsed (response) time from the previous write operation. Following the simulated disaster, it recorded any delays in writing data associated with the cluster's recovery.

[5] See a video of the entire test and explosion at www.hp.com/go/DisasterProof.

Figure 1: Tested disaster-tolerant configuration



Configuration

The configuration was designed the same way an OpenVMS customer who demands maximum application uptime and maximum data protection would deploy OpenVMS. It simulated two main data centers, each with its own server and storage, plus a quorum node at a third site to provide quick, automatic decision-making during a disaster. With the quorum node in place, the surviving cluster sites were able to continue processing as rapidly as possible.

HP StorageWorks XP Disk Arrays were incorporated. XP disk arrays are large enterprise-class storage systems designed for organizations that simply cannot afford any downtime or any data loss. They are designed to deliver uninterrupted availability. Integrated with the OpenVMS environment, the XP disk arrays are architecturally compatible with both HP Integrity server and AlphaServer system platforms. When you integrate OpenVMS, OpenVMS cluster software, and XP disk arrays into your environments, such as was done in this test, your enterprise functions as a cohesive information-management engine that drives the most critical demands for data and applications.

Result: 100% application uptime

No data was lost as a result of the disaster, and the applications continued to run successfully on the surviving nodes. During a few seconds of pause, the OpenVMS cluster evaluated the situation and took the appropriate recovery actions.

The first application continued to run without interruption and without disruption in its data sequence.

The record of write operations recorded by the second application, under conservatively set parameters, showed a maximum delay of 13.7 seconds[6] in writing data to the shared data file. This delay was the result of the cluster identifying the lost node and reconfiguring itself to spread that node's workload over the surviving nodes. While this was going on in the background, the applications continued to run on the surviving node.

Not only was no data lost, but no transactions were missed—and users of the applications would not have noticed this delay since it was a background function.

[6] The background reconfiguration time is variable, depending on how key tuning parameters are set. While the application continues to run uninterrupted on surviving nodes, it is possible to configure clusters with background reconfiguration times as low as 5 seconds.

Implications for your business

CIOs cite business continuity as a top priority whenever they are surveyed about their critical needs. Mitigating business and operational risks—and reducing the business impact and costs of disruptions and outages—is an ongoing concern for business and IT executives alike. These concerns have been heightened by significant technology, social, and business changes over the past decade. In that time, more business and customer services have moved online. Indeed, business pressures have driven more and more businesses to deliver their services without interruption, around the clock, every day of the year.

For many enterprises, the need to deliver IT services around the clock, without interruption, has become a requirement and a challenge. This is true not just for the obvious industries such as financial services, public safety, and telecommunications. But it is becoming an imperative almost across the board. These needs are driven by changes such as globalization, online employees, and supply-chain IT services as well as by online customer and customer support services. Taken together, all these drive the need for full-time 24x365 access to IT services. This need specifically means that long recovery times jeopardize the health of the business. In these circumstances, enterprises must move from a recovery orientation to a disaster-tolerant orientation. They make this move when the potential impact on the business is great enough to affect the top line, productivity, customer loyalty, or brand reputation. Rather than just recovering their IT services, they must *keep their business in operation* during and despite all outages, whether they are caused by disasters, technology failures, or human error.

OpenVMS was the first environment to deliver these capabilities. HP has continued to improve on them ever since. As a result, OpenVMS is deployed to protect the most sensitive, critical, and vital operations in such industries as financial services, transportation, manufacturing, healthcare, telecommunications, government and defense agencies, and utilities. In short, OpenVMS is found in environments in which downtime is simply not an option. Furthermore, the capabilities of OpenVMS go a long way to reduce costs. The cost of downtime, whether caused by a disaster, system maintenance, or anything else, can range into tens of millions of dollars or more. Yet at the same time that OpenVMS delivers the gold standard of disaster tolerance, these very same capabilities also help to increase management efficiency and thus reduce costs even when a disaster may not be on the horizon. As a result of the way OpenVMS is designed, engineered, and configured:

- Maintenance and upgrades can be managed with no application downtime.
- System expansion, whether by scaling out or scaling up, can be achieved with no downtime, and little or no increase in system management effort or personnel.
- Entire data centers can be moved to new locations without application downtime. See 10 years of 100% application uptime, including data center relocation, at **http://www.openvms.org/stories.php?story=03/11/28/7758863**

How OpenVMS achieves up to 100% application uptime

At the most fundamental level, OpenVMS achieves these results through its unique and elegant clustering capabilities.

Typical non-OpenVMS clustering

Clusters in other environments generally function in a failover model, which means that the systems are totally independent and constantly checking on each other to see if there is a problem. When a problem is detected, the backup server must assume responsibility as though it had not been associated with the lost functionality at all. It must load the application, connect to the database, and then begin processing—this is what is meant by failover. As a result, significant delays, lost data, lost transactions, and possible data corruption can occur, especially when the database versions at the two locations are eventually re-synchronized.

OpenVMS clustering

An OpenVMS cluster—whether in a simple two-node configuration or the maximum supported 96-node configuration—functions as a single virtual computing environment that constantly balances the workload and priorities among the nodes; this is called a shared-*everything* environment in which all processing power, storage, memory, I/O, and users are available to all nodes in the cluster. It also means that a cluster of any size can be managed as though it were a single system, thus maintaining simplicity and significantly lowering costs.

Therefore, if a node fails for any reason—all the way up to a catastrophic disaster—the OpenVMS cluster (virtual computing environment) sees this as just another load-balancing task. As long as the cluster and data replication configurations are set up properly, an OpenVMS cluster provides up to 100% application uptime with no data loss, no corruption, and with only minimal delays.

The following are among many key OpenVMS cluster characteristics:

- Applications run simultaneously at multiple sites, rather than running at one and requiring failover after a disaster occurs. OpenVMS continues to operate in surviving site(s), providing the best response times by spreading users evenly across computers at all sites.
- Automatic load-balancing mechanisms such as IP cluster alias (Failsafe IP) and batch queues do not waste computing capacity as compared with a hot/standby (failover) cluster.
- A quorum node can be located at an additional site for an automatic decision to re-allocate resources within seconds of a node loss.
- Bridging of LANs between sites for the cluster interconnect is required at this time; however, the option of using an ordinary IP network as the cluster interconnect is presently under development, as noted in the OpenVMS roadmap plans (www.hp.com/go/openvms/roadmap).
- Host-Based Volume Shadowing (HBVS) delivers automatic storage replication, enabling decisions within seconds because all disks are in simultaneous use. This synchronous shadowing achieves zero data loss. HBVS has been successfully tested up to an inter-site distance of more than 60,000 miles (keeping in mind that the limitations of the speed of light will limit the efficiency at such distances). Moreover, HBVS has special algorithms and on-disk metadata designed to prevent an application from ever accessing out-of-date data.
- Many disaster-tolerant solutions can be configured to recover from the loss of an entire data center. Some OpenVMS customers have configured their disaster-tolerant environments to continue with zero data loss and automatic unattended continuation despite the loss of two out of three entire data centers.
- OpenVMS disaster-tolerant environments offer a flexible choice in storage, including Fibre Channel SAN, SAN extension (e.g., FCIP), or locally accessible storage at each site plus OpenVMS MSCP-serving for remote storage access (requiring no separate storage interconnect).

- OpenVMS clusters are easy to set up and operate in a disaster-tolerant configuration. Many customers have been doing just that after Ethernet was first supported as an OpenVMS cluster interconnect more than two decades ago.
- The OpenVMS file system never leaves any disk unreadable after a system crash, thanks to a technique called "careful write ordering" with write-through caching.
- Even managing a cluster with many nodes requires only slightly more effort than managing a single system because the cluster can be managed as if it were a single system.
- OpenVMS can boot multiple systems from the same system disk (even across multiple sites). So software and patches need to be installed only once for the entire cluster, reducing system administration workload.
- The same copy of OpenVMS can be used to boot systems from the smallest to the largest. It is unnecessary re-install the operating system for a change in server hardware. There is never a need to "re-gen" the kernel for new hardware either. Thus, a server can be instantly reprovisioned and its power applied to a different workload in a different cluster simply by booting it from a different system disk on the SAN.
- Different applications work well together, and OpenVMS keeps them from interfering with one another. A separate server for each different application is not needed, thus further reducing data-center complexity and cost.
- In addition to OpenVMS support for scale-up strategies with the largest multiprocessor systems, OpenVMS clusters readily and painlessly scale out by adding servers to the cluster at any time and having them pick up a share of the workload. To applications, the environment on each node in the cluster appears identical, so they do not know (or care) what node(s) they are running on. All file systems are visible and usable in a coordinated, shared fashion from all nodes simultaneously.
- They are designed to preclude viruses and intrusions—not merely to respond to them. So constant operating system patches to fix virus vulnerabilities are not necessary because proactive intrusion-protection updates keep hackers out.
- A specialized set of services known as Disaster Tolerant Cluster Solution (DTCS) services is staffed by experts in designing, deploying, and managing OpenVMS disaster-tolerant environments.

Conclusion

On May 17, 2007, HP demonstrated beyond doubt the disaster-tolerant, security, and reliability capabilities of HP OpenVMS. Following a controlled explosion of the equivalent of an entire data center, with all of its servers, SAN, storage, and network equipment, one server of which was running OpenVMS in a disaster-tolerant cluster, the applications running on OpenVMS continued to function, thus proving HP's claim that OpenVMS environments can, indeed, provide up to 100% application availability.

Full disaster tolerance in OpenVMS environments can be achieved at distances up to and beyond 500 miles between cluster nodes. Synchronous data replication with OpenVMS has been tested and proven at distances up to 60,000 miles. Additionally, ease of deployment and simplicity of management add value to OpenVMS by significantly reducing bottom-line costs.

To see the actual disaster-tolerance test, please visit: **www.hp.com/go/DisasterProof**

See 10 years of 100% application uptime, including data center relocation, at **http://www.openvms.org/stories.php?story=03/11/28/7758863**

Additional information on HP OpenVMS disaster tolerance can be found at: **http://www2.openvms.org/kparris/** and

http://h71000.www7.hp.com/openvms/journal/v1/disastertol.pdf

For more information on HP OpenVMS, visit: **www.hp.com/go/openvms**

To learn more, visit www.hp.com

4AA1-3405ENW, June 2007

invent