Sponsored by:

This story appeared on Network World Fusion at
http://www.nwfusion.com/news/2004/0426cluster.html

Server clusters offer speed, savings

By Jennifer Mears
Network World, 04/26/04

When retail services firm Datavantage acquired the code last year to roll out its gift-card offering that would provide retailers a transaction platform to store and manage retail credits, it knew its back-end system couldn't stand any downtime. It also knew it didn't want to shell out loads of money to keep the system running on the expensive Unix infrastructure on which it was built.

So the Cleveland firm scrapped the legacy platform, a Sun E450 running Oracle. Instead, Ian Amit, development manager at Datavantage, deployed a cluster of Intel-based HP ProLiant servers running Linux and Oracle's 9i Real Application Cluster database.

Increasingly, companies are looking at ways to cluster smaller, low-end servers to achieve performance and reliability that is equal to or better than expensive, high-end boxes. Clustering servers is nothing new. For years, mainframes have been strung together in what are called Parallel Sysplex Clusters to let workloads be shared across all available resources. Unix systems also provide clustering capabilities with vendors providing proprietary software, such as Sun's Sun Cluster and IBM's high-availability cluster multiprocessing technology.

But the improved performance and reliability of low-end servers has IT managers looking at clustering in a new light, analysts say. Systems vendors and software makers are pushing the trend. For example, last year Dell and Oracle unveiled efforts to push clusters of low-cost, standards-based systems to provide business customers with processing power previously only available on expensive, high-end machines. Finding applications that can run in such distributed environments is one hurdle, but software vendors are beginning to introduce offerings. Oracle's 9i RAC, for example, was designed specifically to run on clustered servers. Analysts say business customers can expect more applications to follow.

Part of the reason for the shift is that IT managers, forced to do more with less in recent years, have begun buying more low-end servers, which at the same time have become more powerful, analysts say. While server revenue has dragged in the midrange and high end during the past few

years, server sales on the low-end (servers priced less than $25,000) continued to grow throughout the downturn, according to IDC.

"IT managers got so accustomed to buying the volume servers during the downturn, I think that has kind of predisposed them to say, 'Gee, let me see how I can use this type of computing going forward,'" says Jean Bozman, vice president of global enterprise server solutions at IDC.

For Amit, the cluster of low-end Linux servers has resulted in a substantial cost savings and improved performance. Whereas the Sun infrastructure cost about $1 million, Amit set up a new infrastructure that includes the cluster of two two-processor ProLiant DL580 servers running Oracle for about $250,000.

"We didn't like the old platform. It had down time that exceeded the [service-level agreement] that we were supposed to provide customers. We had problems managing the Oracle instance, and the data itself was becoming a nightmare because you literally cannot take the system down. For every little hiccup that the system had, that was downtime," he says.

With the cluster and Oracle 9i RAC, servers work together to pick up the load so downtime becomes insignificant, Amit says. If one server goes down or becomes overloaded, the other server automatically picks up the slack. And with the cluster configuration, Amit easily can expand capacity, an important consideration with the company expecting its gift-card application to double its workload in the next year or two.

"I can expand horizontally and vertically. I can add more horsepower to the current servers, and I can add more servers to the configuration," he says. "The whole design had to keep everything completely open to expansion without incurring a single second of downtime. Anything you wanted to do in terms of expanding the old setup was downtime and cost."

Bozman says an IDC study of 325 IT managers running clusters last year found that about 80% of Windows and Unix clusters are being deployed in high-availability configurations. However, the move toward running workload balancing clusters, such as the one Amit deployed, is increasing, especially in the Linux world, where about 80% of those clusters are focused on resource sharing, she says.

Cluster time

Corporations are considering clusters of low-end servers to get the power and reliability typically associated with bigger boxes. Three hundred and twenty-five IT managers surveyed by IDC last year listed benefits and challenges inherent in cluster deployments:

Reasons for clustering:
- Improved availability
- Reliability/dependability
- Improved scalability
- Improved workload management

Challenges in implementing clusters:
- Lack of understanding
- Getting it operational
- Configurations

GlobeXplorer, which provides satellite images and aerial photography via the Internet, made its first foray with Linux when it rolled out a cluster of more than 40 Dell servers running Red Hat Linux earlier this year. The servers help GlobeXplorer deliver and manage images that must be located and decompressed.

Rob Shanks, CEO of the Walnut Creek, Calif., company, says GlobeXplorer handles more than a million images per day, so processing power and reliability is invaluable. Without the cluster, the company would have had to spend millions of dollars to deploy bigger iron, he says.

"The alternative would have been something beyond Sun Fires, which are multi-, multi-million dollar machines [for example, a high-end Sun Fire E25K starts at about $1 million for a four-processor configuration]," he says. "But since we built the software from the ground up, we built it around this clustered technology."

Porting software to the clustered environment is a challenge for companies looking to deploy clusters. Jim Knight, manager for infrastructure services at outdoor clothing and gear retailer Recreational Equipment Inc. (REI) in Kent, Wash., says that was the biggest hurdle in deploying an Oracle 9i RAC cluster on IBM Unix

Cluster time

Corporations are considering clusters of low-end servers to get the power and reliability typically associated with bigger boxes. Three hundred and twenty-five IT managers surveyed by IDC last year listed benefits and challenges inherent in cluster deployments:

Reasons for clustering:
- Improved availability
- Reliability/dependability
- Improved scalability
- Improved workload management

Challenges in implementing clusters:
- Lack of understanding
- Getting it operational
- Configurations

servers two years ago.

REI had been running an Oracle database on one Unix server with hot backup, but realized it wasn't getting the reliability it needed - and was paying too much to have idle hardware standing by - as it watched sales on its online retail site jump from about $25,000 per hour to as much as $95,000 per hour during peak seasons.

With the Oracle 9i RAC cluster, REI avoids downtime, but it took time to get its application developers used to the idea of writing code for a distributed environment, Knight says.

"Our code was originally designed against a single database server with a hot backup, so everybody coded to a specific server," he says. "With RAC you don't need to do that. You code to a database and you let the servers talk to each other. So the challenge that we had was re-educating our developers. Running a database across multiple boxes at that time was unheard of."

Another challenge in deploying clusters can be architecting storage and back-up systems to ensure data is shared across multiple servers.

"With Microsoft clusters, the backups we've run have proven to be not really an obstacle, but more of a challenge to get the back-up system to work with the cluster - to recognize when a failover has taken place in the cluster and move over," says Jim Hammelef, a senior systems programmer at Oakwood Healthcare in Dearborn, Mich.

RELATED LINKS