Individual outages cost an average of $690,204 in 2013, with the most expensive exceeding $1.7 million among enterprises involved in the survey.

Source: Ponemon Report: Sponsored by Emerson Network Power

Increase Uptime

High Availability Clustering

Supporting Business-Critical Operations

High Availability: A Leading Solution

Businesses count on consistent, dependable access to their mission-critical resources and applications. They rely upon systems such as managed file transfer (MFT) solutions to ensure the security and integrity of these assets and to facilitate uninterrupted processes and operations. A gap in this service—even for just a few minutes—can cause losses of data, business opportunities, and revenue.

High Availability (HA) clustering for MFT services seeks to optimize accessibility while reducing the chance that information and applications will be affected by unexpected events such as system failures. By investing in a secure and dependable HA solution, organizations can maximize their uptime and access to mission-critical information.

Why Does it Matter? Downtime Carries Heavy Costs

In this digital age, organizations depend on secure file transfers, such as through MFT solutions, for their day-to-day operations. They need to be able to create, store, process, and access an increasing amount of information, resources, and documents, which are often confidential or sensitive in nature. When an organization's systems go down, it risks losing access to these vital assets until the systems are up and running again. In this situation, time really is money.

Research has pointed to the rising costs of downtime, especially unexpected, total incidents. A December 2013 study by the Ponemon Institute, sponsored by Emerson Network Power, revealed that downtime is becoming progressively more expensive. In 2013, the average cost for unplanned downtime was approximately $7,900 per minute, an increase from $5,600 per minute in 2010.

"As data centers continue to evolve to support businesses and organizations that are becoming more social, mobile, and cloud-based, there is an increasing need for a growing number of companies and organizations to make it a priority to minimize the risk of downtime and commit the necessary investment in infrastructure technology and resources," said Peter Panfil, vice president of global power at Emerson Network Power.

Assessing the Costs

Downtime can burden organizations with the following direct and indirect costs, which the Ponemon report recognized as some of the most significant consequences:

- Harm to mission-critical databases
- Reduced organizational efficiency
- Damaged equipment and other assets
- Investment in detecting and fixing problems in systems and core business processes
- Loss in IT productivity as support teams needs to handle downtime response
- Third-party fees for consultants, contractors, auditors, and other specialists who assist with unplanned outages
- Legal and regulatory issues, including litigation-defense costs
- Lost revenue during the outage itself when transactions are prevented
- Eroded confidence and trust among consumers and stakeholders
- Diminished brand reputation

These expenses and others add up quickly during extended downtime. The study found that individual outages cost an average of $690,204 in 2013, with the most expensive exceeding $1.7 million among enterprises involved in the survey.

Downtime can be caused by a variety of factors. The Ponemon report pointed to the following culprits, listed in order of frequency:

- IT equipment failure
- Cybercrime, such as Distributed Denial-of-Service attacks
- UPS system failure
- Water, heat, or server room air conditioning failures
- Generator failures
- Severe weather and acts of nature
- Accidents or human error

HA and Active-Active Clustering Can Help Maintain Business Continuity

Business continuity is critical in today's digital world. Although it's difficult to anticipate and prevent all of the factors that can interrupt business operations, organizations can take a number of steps to reduce the chances they'll experience unexpected downtime, as well as to mitigate the consequences of system failure.

One of the leading approaches to improve business continuity is implementing HA deployments that make use of active-active clustering. Clustering is how file transfer loads and other demands on the system are distributed among a cluster of nodes, or servers, to draw from multiple resources while providing flexibility in case a problem emerges. With active-active configurations, organizations can take advantage of multiple servers, so if one becomes overloaded or fails, the critical applications and processes can failover to the other nodes. This infrastructure cuts the time it takes to resume full operation to a matter of seconds—or even subseconds.

To give some perspective on the difference that HA solutions can make in terms of reducing downtimes, the Oracle® Fusion Middleware High Availability Guide offers some comparisons between different levels of availability and amount of downtime allowed over the course of a year:

- 95 percent availability has approximately 18 days of downtime
- 99 percent availability: 4 days of downtime
- 99.9 percent availability: 9 hours of downtime
- 99.99 percent availability: 1 hour of downtime
- 99.999 percent availability: 5 minutes of downtime

HA active-active clustering can provide far greater reliability than active-passive systems, such as reaching 99.99999 percent availability, according to Availability Digest. The impact of this on reducing the length of overall downtime—and, therefore, its costs—is quite evident.

How HA strengthens MFT services

MFT solutions that are bolstered by active-active HA deployments provide consistent, reliable data access and transfers. Here's how it works:

- Storage and processes are shared among multiple file servers, making use of clustering to distribute the load.
- When the load balancer detects an issue—such as when one server becomes unresponsive—it stops sending requests to that server.
- Instead, it distributes tasks among the remaining nodes, which have remained active and can accommodate the increased demand.

Because all nodes are always active and current, there is virtually no failover window. The shift happens in near real time as the load balancer monitors and responds to any behavior beyond normal operations.

To the end user, there is no difference: They have a unified, consistent, and streamline experience. That's what clustering is for—it automatically disperses data transfer loads behind the scenes so employees and other stakeholders experience their activities as though everything were running on the same server.

How Are Active-Active Configurations Different from Other Architectures?

Decision-makers may be familiar with traditional arrangements, such as single-server configurations or multi-node active-passive deployments. The active-active HA solution provides a more rapid, robust response to issues than these options. Here's how they work in a failover situation:

- Single-server configurations provide no real failover support. If a system fails, it is out until it can be fixed or until operations are restored using a backup solution. Additionally, this method requires administrators to update the core server using maintenance windows, requiring planning downtime in addition to potentially lengthy unplanned incidents.

- Active-passive clusters improve upon the dependability and availability of single-server solutions. When downtime occurs, the system facilitates automatic failover to use the passive server, which should be unaffected by the problem that disrupted the main system. This can also be used for planned downtimes for maintenance. However, the switch to the secondary node can take a few minutes, causing a delay and possibly interruptions in data transfers.

The failover process within an active-active setup is almost imperceptible. Even for those seconds or subseconds that it takes for the load balancer to make the transition to using responsive nodes, only a subset of users—if any—is affected by the process, Availability Digest explained. In other words, only employees who are using services currently running on the down server will have any possibility of noticing. The other end users experience no interference, because their operations are already being processed on the fully functional nodes.

In contrast, a single-server configuration and an active-passive solution could affect all users for whatever length of time it takes to fix the issue or complete the failover process to the secondary server.

Overall, active-active HA solutions reduce the chances that an unexpected situation–such as an overloaded or failed server–will turn into a catastrophic incident for your business. Even for issues that affect only some of your operations, such as issues with a particular application, active-active clustering can rectify the problem before users even encounter outages or delays. As data resources and computer activities become more central to your enterprise, uninterrupted data availability is critical for your sustained, long-term success.

Additional Benefits

While the primary purpose of active-active HA arrangements may be to prevent lengthy, costly, and damaging periods of downtime caused by system failure, there are a few additional added benefits to this deployment option:

- Eliminates the need for planned downtimes
- Affects fewer users in the event an unlikely interruption occurs
- Improves disaster recovered processes
- Offers scalability and flexibility to adapt to evolving resource needs
- Reduces or prevents data loss when equipment fails

Furthermore, a clustering solution enables the use of multiple, heterogeneous server resources while still offering the simplicity of an integrated system. For both end users and administrators, this means that the complexity of the setup—which is what provides protection in the event of a downtime—doesn't have to make their activities more difficult. Users enjoy a seamless experience and administrators can manage the system from a centralized control, pushing out updates and patches to the entire cluster.

Deciding if HA Active-Active Is the Right Fit

Organizations should consider a variety of factors before deciding which managed file transfer solution is a good fit. Deciding between single-server, active-passive, and active-active configurations depends on the company's unique needs and processes. When considering costs, for example, the initial infrastructure expenses must be evaluated as well as the potential business costs inherent in extended downtimes.

Active-active HA deployments are particularly appropriate for mission-critical processes and data, especially in a growing business environment and for organizations that experience spikes in traffic. Because additional active nodes can easily be added to systems when greater capacity is needed, active-active solutions are incredibly flexible and scalable.

Here are a few typical use cases for active-active configurations:

- Critical databases
- Network file sharing
- Enterprise applications
- Customer-facing programs, like e-commerce sites, especially with a need for "dial tone" reliability
- High or unpredictable volumes of file transfers
- Growing businesses with expanding data needs

In short, if organizations have any resources or processes that require constant, dependable availability, they should think about whether active-active is right for them.

Additional Considerations

Organizations should consider how their programs and processes can take advantage of clustering services. With the right infrastructure in place, administrators can easily set up clustering at installation, such as when they're implementing an MFT solution. Once established, administrators can conveniently manage the active-active deployment through common software such as Microsoft Clustering Services, found in the Windows operating system. This makes it simple to update or make changes to the entire system through centralized management in which changes are pushed through to the entire system.

Although each organization has its own needs, high availability is something that is becoming an increasingly valuable assurance for many companies relying on consistent access to mission-critical resources. When assessing their options, decision-makers should consider their data requirements as well as the potential cost of downtime. Active-active deployments can give them the dependability they need in a simple-to-manage, flexible arrangement.

About Globalscape

Globalscape ensures the reliability of mission-critical operations by securing sensitive data and intellectual property. Globalscape's suite of solutions features Enhanced File Transfer™, the industry-leading enterprise file transfer platform that delivers military-grade security and a customizable solution for achieving best-in-class control and visibility of data in motion or at rest, across multiple locations. Founded in 1996, Globalscape is a leading enterprise solution provider of secure information exchange software and services to thousands of customers, including global enterprises, governments, and small businesses.