

The Role of Fault-Tolerant Servers in Protecting Virtualized Applications

System downtime is an issue that still plagues the vast majority of companies. In a survey conducted in June of 2011, only 7% of companies reported that their datacenter had 100% or 99.999% uptime (less than 5 minutes of downtime per year, meaning that 93% of companies still suffer measurable amounts of downtime each year). In February 2012, Aberdeen surveyed 136 organizations of all sizes to learn about their experiences with downtime. This Analyst Insight reports on the types of downtime protection chosen by our survey respondents. In particular this report will focus on the role of fault-tolerant servers in providing the highest levels of uptime for virtualized applications.

Analyst Insight

Aberdeen's Insights provide the analyst perspective of the research as drawn from an aggregated view of the research surveys, interviews, and data analysis

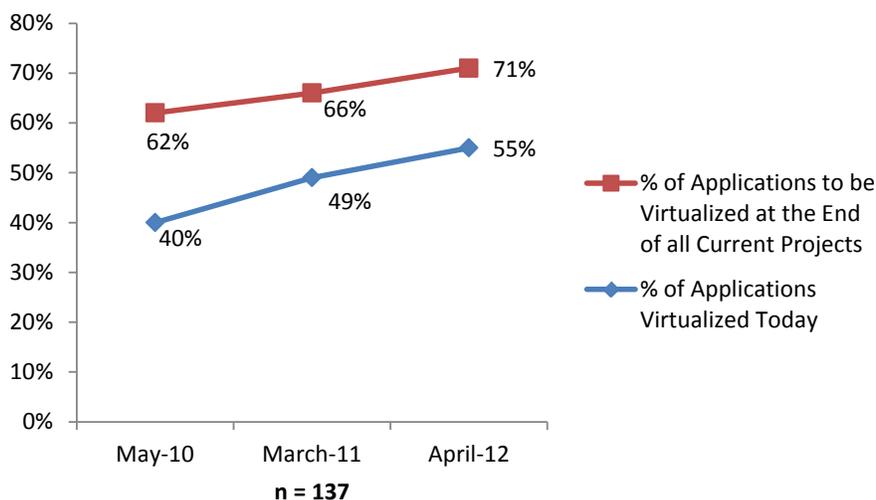
The New Datacenter

Aberdeen has conducted three yearly surveys of datacenter trends. Because of those surveys, Aberdeen has identified three very pronounced ongoing changes that affect application uptime: the increasing rate of server virtualization, the decline in the number of available IT support resources, and the increase in the cost of application downtime.

The Increasing Rate of Server Virtualization

Deploying applications on hypervisor software for server virtualization continues to grow at a steady pace, with April 2012 data showing that just over half (55%) of applications are currently deployed on virtualized servers.

Figure 1: Percentage of Applications Deployed on Virtualized Servers



“Be careful, start slow but realize that almost 100% can be virtualized.”

~ IT Manager, Mid-sized Government Agency, Germany

Source: Aberdeen Group May 2012

The rate of server virtualization will continue to rise to the point of having almost all applications deployed on a virtualized server in the near future. Any form of downtime protection will need to be able to support both non-virtualized (also called physical) application deployments for now, but also virtualized applications as more and more apps are deployed on hypervisor software in the future.

The Decline in the Number of IT Support Resources

From 2010 to 2011 the average company kept IT spending flat and reduced their IT headcount by 3.7%.

Table 1: IT Spending and Headcount Trend

	All Respondents
What is the change in overall IT spending from last year?	0%
What is the change in IT spending for server and app deployment from last year?	- 10.9%
What is the change in overall IT headcount from last year?	- 3.7%
What is the change in IT headcount for server & app deployment from last year?	- 3.1%

Source: Aberdeen Group May 2012

This spending and headcount reduction was particularly felt in the group that supports servers and applications, as they suffered on average an almost 11% reduction in budget. Any form of downtime protection chosen must be easy to install and maintain as there are fewer IT resources to manage the assets. A “set it and forget it” deployment capability with few, if any, false failovers (a system failover for no apparent reason, a common problem with cluster software requiring a complete reset of the downtime protection) is a must as fewer IT professionals will be available to reset the application protection.

Increasing Cost of Downtime

In a March 2012 Aberdeen report, [*Datacenter Downtime: How Much does it Really Cost?*](#), Aberdeen found that the average cost of an hour of downtime is \$138,000 USD. This is the total negative impact of application downtime on an organization. It includes lost productivity, lost data, reduced customer satisfaction, and potentially lost sales from web and order processing systems.

When Aberdeen conducted a survey on downtime in 2010, the average yearly cost of downtime was reported as \$100,000. This current \$138,000 figure reflects about a 15% year-over-year growth on average, across all companies, from what was reported in 2010. This makes sense as more

business processes are now computerized and companies are running leaner, with fewer headcount resources to fall back on for manual processes.

Table 2: Cost of Downtime

	All Respondents
All respondents	\$138,000
Small enterprise (> 100 Employees)	\$6,900
Mid-sized organizations (Between 100 and 1000 employees)	\$74,000
Large enterprise (< 1000 employees)	\$1,130,000

Source: Aberdeen Group February 2012

The remainder of this report will focus on the four classes of tools available to protect applications from downtime. Each has its advantages and costs and provides different levels of protection against downtime and loss or corruption of business data. Figure 2 highlights that different downtime protection technologies offer varying amounts of coverage. Setting the right budget for downtime protection is important, as trying to save money by buying a lower priced solution can end up costing far more in the form of actual downtime costs.

Application Protection Options

In the February 2012 survey Aberdeen used four categories of application downtime protection. These included hypervisor software features such as application mobility and site recovery manager, host-based cluster software, and fault-tolerant hardware. One must also include doing nothing — adding no additional protection to an application other than that offered by general purpose servers — as that might be perfectly adequate for less critical apps.

Today's servers provide features such as error checking memory and redundant disks to keep applications running and provide minimal data protection. Most server vendors describe their servers as having about 99.5% uptime (about 44 hours of unscheduled downtime per year over the life of the server) if all maintenance is performed as directed. Organizations requiring better uptime performance can select from optional software or hardware options that deliver higher uptimes. Some of these technologies can boast levels of up to 99.999% uptime (on average, less than five minutes of downtime per year).

Options for protecting virtualized applications from downtime include:

- **No Protection** — Applications can be run with no additional high availability protection other than that offered by standard hypervisor software and server hardware. This level of protection

must include backing up the data as general purpose servers offer nothing for long term data protection.

- **Hypervisor Software** — One option available to virtualized applications is to take advantage of the high-availability (HA) features offered by leading virtualization vendors. These features protect applications at the hypervisor level, so it does not impact or make unique demands on either the operating system or the application itself.
- **High-Availability Clusters**— Cluster/HA software is host-based software (software that runs on a server, either physical or virtual) that monitors performance, and initiates a failover if it detects a hardware or software issue. Newer versions have become more sophisticated and detect the difference between a server anomaly and a fatal error.
- **Fault-tolerant Hardware** — Several vendors provide hardware-based fault-tolerant servers. These servers have completely duplexed hardware in a single chassis. The single OS and application run in lockstep on both sets of hardware, and if there is an issue, the parallel hardware component keeps on running – without any downtime, or failover. Fault-tolerant systems also protect against loss or corruption of data – even inflight data not yet written to disk is fully protected. Fault-tolerant servers provide the highest level of protection, with companies claiming better than 99.999% uptime.

Choosing the right protection level is important as each option offers different levels of coverage and times to recovery. Figure 2 below highlights the advantages and disadvantages of each protection option.

The chart below uses the following terms:

- **RPO** or Recovery Point Objective – How much data could you expect to lose if there is a failure?
- **RTO** or Recovery Time Objective – How long may it take to recover 100% application performance?
- **Tiers** – Class of applications – Tier 1 apps are the most mission critical with tiers 2 and below having less importance

Figure 2: Availability Protection Options

	Data Snapshots Only	Hypervisor Software	Cluster Software	Fault-Tolerant Servers
RPO	Hours - Day	Minutes	Minutes	0 Min
RTO	Hrs-Days	Hours	Minutes	0 Min
App Tiers	Tiers: All	Tiers: All	Tiers: 1-2	Tiers: 1-2

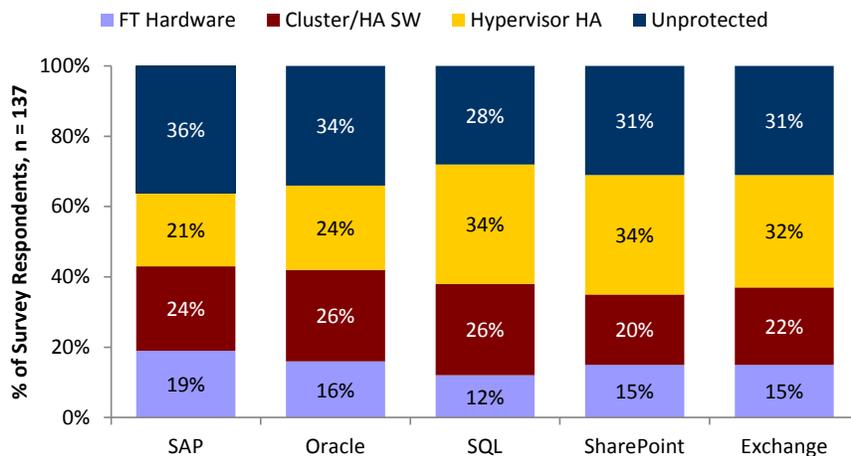
Source: Aberdeen Group May 2012

Clearly the lowest level of protection is chosen by adding no additional application protection options, and the highest level of protection is provided by fault-tolerant servers. We will now look at the results of the survey of 137 worldwide organizations of all sizes and their choice for protection against application downtime.

Tier I Application Protection Choices

For the purpose of this analysis, applications have been divided into just two tiers. Tier I applications are large, enterprise, mission critical applications, such as database, customer relations management, enterprise resource planning, and email. Figure 3 shows the choices survey respondents made in protecting their most important applications.

Figure 3: Tier I Application Protection Choices



Source: Aberdeen Group, May 2012

You can see in the Figure above that the choice of downtime protection varies widely depending on the exact type of application you are discussing. For SAP and Oracle, the most common choice is to implement no

additional downtime protection, while for the Microsoft products of SQL, SharePoint and Exchange, the most common choice is cluster software.

There is a strong segment of users that choose to protect each of their mission-critical applications with fault-tolerant servers. Nineteen percent (19%) of SAP users protect that application with fault tolerant hardware and 15% of SharePoint and Exchange users do the same. The reason for choosing fault-tolerant servers, the highest level of downtime protection will of course vary by organization, but factors include the cost of downtime, the number of people affected by a failed application, and the time and resources required to recover. Clearly there is a sizable group of Tier 1 application users that see fault-tolerant servers as the best solution for any mission critical tier 1 application.

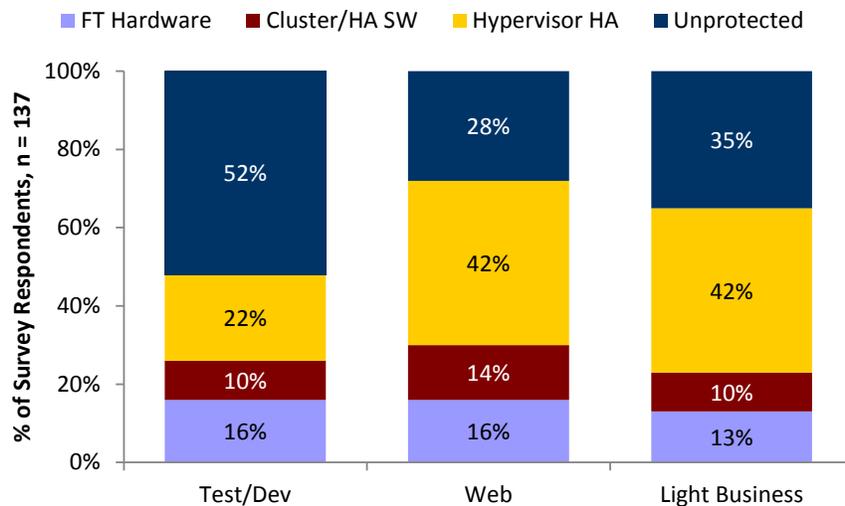
Tier 2 Application Protection Choices

Tier 2 applications are those that are less critical to the organization. They may be apps that are used only by a single department or by a limited number of users. Their cost per hour of downtime will be less as their loss will be felt by fewer employees or customers. Figure 4 shows the application downtime protection chosen by the survey respondents. It is interesting to note that the fault-tolerant server usage rates of this group are very close to Tier 1 levels and in some cases greater. Some organizations seek the highest levels of downtime protection for these Tier 2 applications as well.

“Some smaller applications utilized by the business were labeled unimportant because the business users did not deem them necessary until they lost them during early tests.”

~ CIO, Mid-sized Department Store, US

Figure 4: Tier 2 Application Protection Choice



Source: Aberdeen Group, May 2012

Test and Development applications are very different from any other app. They are used by a limited number of employees (primarily software engineers), they are used during only specific periods of the software life

cycle, and they generally use test or dummy data to see how the application performs. As a result, a majority of users (52%) provide no downtime protection to this class of application.

However, a group of companies of all sizes use fault-tolerant servers to host their Tier 2 applications. Depending on the type of company, the above listed applications can have a large negative impact if they go down. For example, a company that sells primarily through the web would prefer to have their sales systems with zero downtime. Light business applications such as sales processing, employee management or time management applications may also have times of high demand, where a fault-tolerant server and its 99.999% uptime are cost justified.

Summary: Companies are choosing the Highest Levels of Application Protection

Application downtime is still having an impact on organizations today. Companies have multiple choices in how to deal with it, starting at doing nothing more than that offered by general purpose servers to deploying their applications on fault tolerant hardware.

- Fault-tolerant servers provide the highest levels of uptime of any type of downtime protection. This level of uptime can easily be financially justified as the cost per hour of downtime can be as high as \$1M for large enterprises.
- There is a Leading group of organizations who have chosen fault-tolerant servers as their application protection of choice for both Tier 1 and Tier 2 applications. Their reasons for choosing fault tolerant servers may be:
 - They have examined their cost of application downtime and decided that having the highest level of protection makes economic sense. As the cost of these systems has fallen dramatically in the last few years, they are very viable solutions for small to mid-size companies.
 - FT servers operate like single standalone servers – no special skills required. No application changes required. This makes them viable for all sizes of companies, especially those with IT staffing issues.

Talk to organizations that have already done it. Fault-tolerant servers protecting both Tier 1 and Tier 2 applications are a proven and reliable solution. Fault-tolerant servers may make the most practical sense in terms of financial ROI given the high cost of downtime, the limited number of IT resources to deal with downtime recovery and the rising role computers play in even the most basic business process.

Survey Respondents

Individuals answering this survey came from diverse geographies, industries, and corporate roles:

Industries:

- √ IT Services – 21%
- √ Government – 13%
- √ Healthcare/Pharm – 13%
- √ Education – 11%
- √ Finance/Insurance – 10%
- √ Telecomm - 9%
- √ Software - 8%
- √ Manufacturing – 8%
- √ Others – 7%

Roles:

- √ Director and above – 55%
- √ Managers – 21%

Geography:

- √ North America: – 62%
- √ EMEA – 22%
- √ Rest of World – 16%

For more information on this or other research topics, please visit www.aberdeen.com.

Related Research	
<u>Extend your Server Virtualization Program to enable Private Cloud</u> ; May 2012 <u>EMEA SMEs are leading their North American Cousins in Server Virtualization</u> ; February 2012 <u>The State of Server Virtualization in Small and Mid-Sized Organizations</u> ; January 2012	<u>Small vs. Large Enterprise Data Backup: Same Concept, Different Process</u> ; June 2011 <u>High Availability for Virtualized Applications: Protecting Against Unplanned Downtime</u> ; June 2011 <u>Managing Virtualized Applications: Optimizing Dynamic Infrastructures</u> ; April 2011
Author: Dick Csaplar, Senior Research Analyst, Virtualization and Storage Practice (dick.csaplar@aberdeen.com)	

For more than two decades, Aberdeen's research has been helping corporations worldwide become Best-in-Class. Having benchmarked the performance of more than 644,000 companies, Aberdeen is uniquely positioned to provide organizations with the facts that matter — the facts that enable companies to get ahead and drive results. That's why our research is relied on by more than 2.5 million readers in over 40 countries, 90% of the Fortune 1,000, and 93% of the Technology 500. As a Harte-Hanks Company, Aberdeen's research provides insight and analysis to the Harte-Hanks community of local, regional, national and international marketing executives. Combined, we help our customers leverage the power of insight to deliver innovative multichannel marketing programs that drive business-changing results. For additional information, visit Aberdeen <http://www.aberdeen.com> or call (617) 854-5200, or to learn more about Harte-Hanks, call (800) 456-9748 or go to <http://www.harte-hanks.com>. This document is the result of primary research performed by Aberdeen Group. Aberdeen Group's methodologies provide for objective fact-based research and represent the best analysis available at the time of publication. Unless otherwise noted, the entire contents of this publication are copyrighted by Aberdeen Group, Inc. and may not be reproduced, distributed, archived, or transmitted in any form or by any means without prior written consent by Aberdeen Group, Inc.