The State of Application High Availability

SIOS

> Contents

❯ Executive Summary

An organization's highly available applications (HA applications) are generally the ones that ensure that a business remains in operation. Such systems can range from order-taking systems to CRM databases to anything that keeps employees, customers, and partners working with you. SIOS wanted to understand the general state of HA applications in organizations of all sizes, and we've learned that the news is mixed when it comes to how well HA applications are supported. Here are a few key insights from this report:

- Most (86%), but not all, organizations are operating their HA applications with some kind of clustering or high availability mechanism in place.

- A full 95% of respondents report that they have occasional failure in the underlying HA services that support their applications.

- Ninety-eight (98%) of respondents to our survey indicated that they see either regular or occasional application performance issues.

- When such issues occur, for most organizations, it takes between three and five hours to identify the cause and correct the issue; it also takes using between two and four tools to do so.

- Small companies are leading the way by going all-in on operating their HA applications in the cloud; more than half (54%) of small companies intend to be running 50% or more of their HA applications in the cloud by the end of 2018.

- For companies of all sizes, control of the application environment remains a key reason why workloads remain on-premises, with 60% of respondents indicating that this has played a factor in retaining one or more HA application on-premises rather than moving it into the cloud.

Figure 1: Company size

How many people are in your organization?

(N=390)

22%
5,000 or more

17%
1,000 to 4,999

35%
1 to 249

26%
250 to 999

Figure 2: Respondent role

What is your title?

(N=389)

8%
IT Admin

12%
Senior Management
(VP, CIO, CTO)

13%
Developer

29%
Database
Administrator (DBA)

23%
Architect

14%
Infrastructure
Manager

For this survey, we gathered responses from 390 IT professionals and decision-makers from a broad range of company sizes (**Figure 1**). Respondents consisted of people that manage databases, infrastructure, architecture, systems, and software development as well as those in IT management roles. **Figure 2** provides a look at the respondent role breakdown.

Enterprise Application Landscape

For the purposes of this report, the focus is on tier-1 mission-critical applications, including Oracle, Microsoft SQL Server, and SAP/HANA. They're the lifeblood for most organizations operating these kinds of services; they hold the data that enables the organization to achieve its goals.

Fifty-six percent (56%) of respondents to our survey are operating Oracle workloads, while 49% are running Microsoft SQL Server. Rounding out the survey, 28% have SAP/HANA in production. These are all clearly critical workloads in most organizations, but there are others. For this survey, we provided respondents an opportunity to tell us what, beyond these three big applications, they're operating that can be considered mission critical. Respondents indicate that they're also operating various web databases, primarily from Amazon, as well as MySQL and PostgreSQL databases. To a lesser extent, organizations are also operating some NoSQL services that are considered mission critical.

Figure 3: Mission-critical database systems in use in respondent organizations

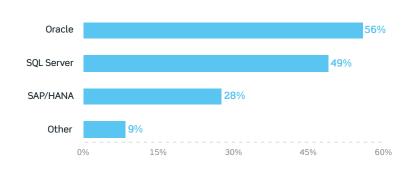**Which of the following applications/databases are you running in your environment?**

(N=390)

Application	Percentage
Oracle	56%
SQL Server	49%
SAP/HANA	28%
Other	9%

Which, if any of the following applications/databases are you running in your environment? (by company size)

- Small Companies
- Medium Companies
- Large Companies

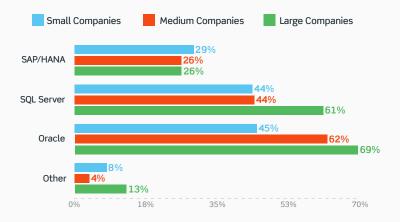|	
SAP/HANA	29% / 26% / 26%
SQL Server	44% / 44% / 61%
Oracle	45% / 62% / 69%
Other	8% / 4% / 13%

0% 18% 35% 53% 70%

Figure 5: SAP/HANA editions in use by respondent organizations

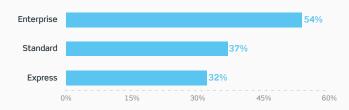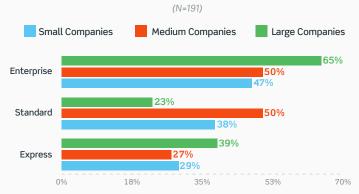## Which SAP/HANA editions are you running?
(N=108; multiple responses allowed)

Enterprise	54%
Standard	37%
Express	32%

0% 15% 30% 45% 60%

Figure 6: SAP/HANA editions in use by respondent organizations (by company size)

What SAP/HANA editions are you running?
(N=191)

- Small Companies
- Medium Companies
- Large Companies

Enterprise	65% / 50% / 47%
Standard	23% / 50% / 38%
Express	39% / 27% / 29%

0% 18% 35% 53% 70%

Breaking this information down by company size reveals that company size, surprisingly, isn't playing a large role in whether or not SAP/HANA is considered a mission-critical workload. However, company size does play a significant role when it comes to SQL Server and Oracle operations, with far more large companies operating these types of environments. As you can see in **Figure 4**, 61% of large companies operate SQL Server for mission-critical applications; just 44% of small and medium-sized companies do so. For Oracle, 69% of large organizations operate this RDBMS platform for mission-critical systems, as opposed to 62% of medium organizations and 45% of small ones.

> ## 61% of large companies operate SQL Server for mission critical applications. Just 44% of small and medium-sized companies do so.

For companies operating SAP/HANA systems, most (54%) are operating the Enterprise Edition. Thirty-seven percent (37%) run the Standard Edition, and 32% run Express (**Figure 5**). Here, company size does appear to be a factor in the chosen edition option; 65% of large SAP-wielding companies operate an Enterprise instance, and 50% and 47% of medium and small companies, respectively. Interestingly, larger organizations are also more likely to be operating an Express Edition for certain purposes; 39% of those running Express are in large organizations (**Figure 6**).

Figure 7: SQL Server editions in use by client organizations

Which SQL Server editions are you running in your environment?

(N=191)

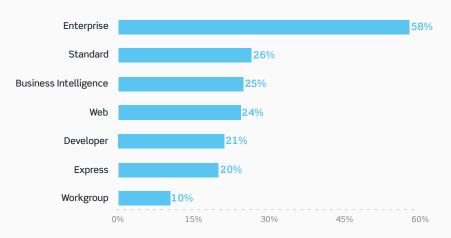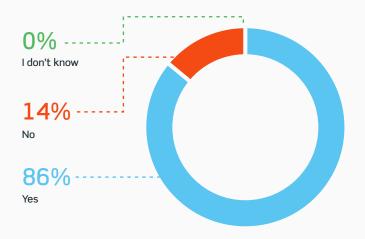| Edition	Percentage
Enterprise	58%
Standard	26%
Business Intelligence	25%
Web	24%
Developer	21%
Express	20%
Workgroup	10%

Figure 8: Organizations with databases operated in a high availability configuration

Are you running your database in a high availability/clustered environment?

0%
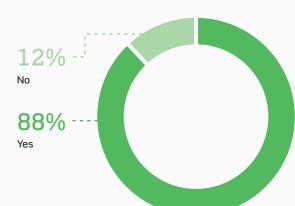I don't know

14%
No

86%
Yes

SQL Server has risen in dominance over the past decade to become a platform on which a significant number of critical business processes operate. Over the years, Microsoft has made available a dizzying array of editions, with feature sets in each edition crafted at particular uses and with a differing set of availability capabilities. Microsoft's SQL Server Enterprise Edition is the company's flagship. Among respondents that indicated that they're running SQL Server, 58% operate at least one Enterprise Edition instance in their environments. Given that the Standard Edition is one of Microsoft's production-grade editions, it's not surprising to see this edition come in second at 26%. Further, until the current latest release of SQL Server, Microsoft made available a comprehensive Business Intelligence Edition, which is still operated by 25% of SQL Server respondents to this survey. **Figure 7** provides a complete breakdown of the various SQL Server flavors in use by respondents.
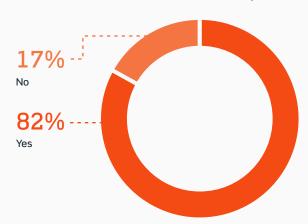
WORKLOAD HIGH AVAILABILITY USAGE

What really indicates whether or not a workload is critical is how much effort is expended on keeping that workload operational. One of the most widespread ways by which database workloads are kept in production is through the use of various high availability and clustering mechanisms. There are a number of ways by which databases can be made highly available, including through features made available from the database vendor as well as through third parties. For this question, we didn't worry about how high availability is being achieved, but whether or not it's being done at all. **Figure 8** demonstrates that 86% of respondents operate at least one of their database environments in a high availability or clustered environment, leaving just 14% without this protection. Company size does not appear to play a role in whether or not high availability is in place (**Figure 9**).

Are you running your database in a high availability/clustered environment? Small Companies

12%
No

88%
Yes

Are you running your database in a high availability/clustered environment? Medium Companies

17%
No

82%
Yes

Are you running your database in a high availability/clustered environment? Large Companies

15%
No

86%
Yes

Application Performance Statistics and Characteristics

Just 2% of respondents say that they never have application performance issues that impact end users in some way. That means that a whopping 98% of organizations do, in fact, have issues with mission-critical applications at various times, ranging from daily (experienced by 18% of respondents) to just one time per year (experienced by 8% of respondents) and everywhere in between. **Figure 10** provides a complete look at the application performance landscape.

This is a critical issue for organizations. Application performance issues lead to customer dissatisfaction and can result in lost revenue and increased expenses. But there appears to be some disagreement around such issues, depending on your perspective in the organization. In **Figure 11**, you can see this broken out by respondent role. It becomes clear that respondents holding decision-maker roles have a more positive view of the performance situation than others. Only 11% of decision makers report daily performance challenges, compared to about 20% of other respondents.

Figure 10: Cloud-based application performance issue frequency

Approximately how often do you have an application performance issue that affects end users?

Daily		1 time/week		1 time/month		1 time/year	
18%	17%	10%	15%	11%	19%	8%	2%
	2 to 3 times/week		2 to 3 times/month		3 to 5 times/year		Never

Approximately how often do you have an application performance issue that affects end users?

(N=390; By Role)

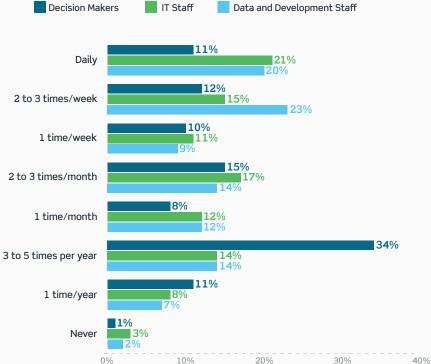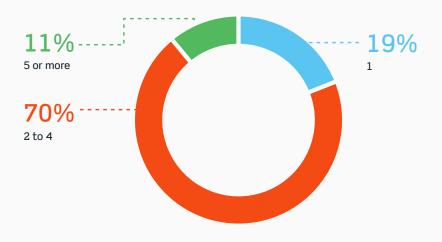■ Decision Makers ■ IT Staff ■ Data and Development Staff

When you have a performance issue with an application running in the cloud, how many tools (monitoring applications, reports, analytics tools) do you typically consult to identify the cause?

(N=382)

11%
5 or more

70%
2 to 4

19%
1

As much as most IT pros using the cloud would like to fully eliminate the potential for application performance issues, the fact is that such situations can and will happen, and IT needs to be ready. There are a variety of tools available in the market to help IT understand and address application performance issues; and IT departments have, over the years, cobbled together troubleshooting toolkits. In general, the fewer tools you need to work with to resolve a problem, the more quickly you can bring services back into full operation. It's disheartening to learn that only 19% of respondents have a single tool to which they can turn to identify cloud application performance issues. This leaves 81% of respondents having to use two or more tools. But it gets worse: 11% of respondents have to use five or more tools to identify performance issues with their cloud applications (**Figure 12**).

How long (on average) does it take to resolve an application performance issue?

(N=383)

3%
5+ hours

23%
Less than an hour

18%
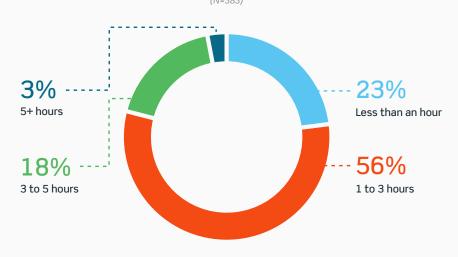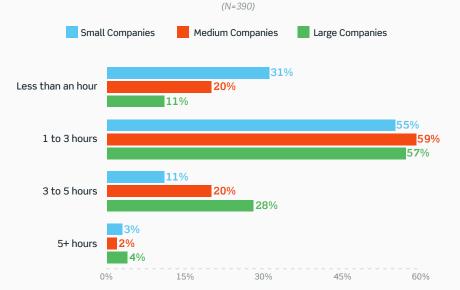3 to 5 hours

56%
1 to 3 hours

Figure 14: Time required to recover from an application performance fault (by company size)

How long (on average) does it take to resolve an application performance issue?

(N=390)

■ Small Companies ■ Medium Companies ■ Large Companies

Less than an hour
31%
20%
11%

1 to 3 hours
55%
59%
57%

3 to 5 hours
11%
20%
28%

5+ hours
3%
2%
4%

0% 15% 30% 45% 60%

Of course, no one wants any application performance challenges; but the real test of an organization's ability to handle such issues comes when measuring the time it takes to recover when something does go awry. In **Figure 13**, you'll note that 23% of respondents can typically recover in less than an hour. Fifty-six percent (56%) of respondents take somewhere between one and three hours to recover. A significant number -- 23% -- take 3 or more hours. This isn't to say that these IT workers are recovering from a complete failure; they're reacting to a performance fault somewhere in the application, one that's serious enough to warrant attention. A goal for most organizations is to reduce the amount of time it takes to troubleshoot problems, which will reduce the amount of time it takes to correct them.

Given how important this information is, we dug a bit deeper into the data, with the results you see in **Figure 14**. Note that smaller companies spend far less time trying to resolve application performance issues than their larger cousins. Thirty-one percent (31%) of small companies can resolve applications performance challenges in less than an hour, compared to just 11% of large companies. In fact, as company size increases, the amount of time it takes to resolve application performance problems increases dramatically. If you look at the statistics in the 3 to 5 hour resolution column, you'll see that just 11% of small companies take this long to correct exceptions, as compared to 28% of larger companies.

To be fair, though, there's a lot more in play here. Larger companies often have far more complexity to deal with in terms of infrastructure and application integration, so it makes sense that they experience longer recovery times. Moreover, as you'll learn later in this report, small companies are also more likely to be operating all of their mission-critical applications

What percent of your highly available (HA) applications are currently in the cloud?

(N=390)

11%
76 to 100%

13%
51 to 75%

32%
25 to 50%

12%
None

32%
Less than 25%

Why have you not made the full move to the cloud?

(N=345)

We prefer to keep our HA applications on-premises where we have more control over the infrastructure	60%
Our organization lacks the internal knowledge/expertise to make the shift	22%
Fear of unexpected costs	16%
We don't think we can provide HA clustering protection in the cloud	14%

in the cloud and/or have assistance from a managed service provider in managing such applications, so they have a "single point of contact" advantage in some ways.

THE ROLE OF THE CLOUD IN HA APPLICATIONS

Bearing in mind that "highly available" applications are those that are particularly important to the organization, we sought to understand what role the cloud plays in respondent organizations. Modern organizations are embracing the hybrid cloud and making strategic decisions around where to operate critical workloads. But not everyone is keen on moving applications into an off-premises environment. In fact, 12% of our respondents have yet to move a single highly available workload to the cloud. This doesn't mean that they're not using cloud at all; it just means that they're not using it for their highly available applications.

In all, just 24% of respondents are running more than half of their highly available workloads in the cloud. There are a variety of reasons for this.

To start with, the cloud isn't for everyone, and not all applications are good candidates for a cloud environment. There are a variety of reasons that respondent organizations haven't made the full jump into the cloud. Control is a huge concern. In this survey, a full 60% of respondents cite infrastructure control concerns as a key reason that they haven't gone all-in on cloud. It's clear that people continue to view their on-premises environments as a preferred location, thanks to the level of control afforded by owning the infrastructure.

Are you planning to move more of your HA applications to the cloud in 2018?

(N=390)

12%
No, we don't plan on moving any HA applications to the cloud in 2018

29%
Yes, we plan to have less than 50% of our HA applications in the cloud by the end of 2018

9%
All of our HA applications are currently in the cloud

50%
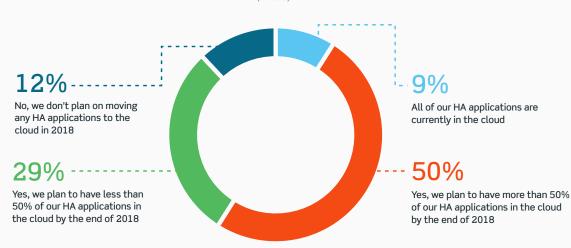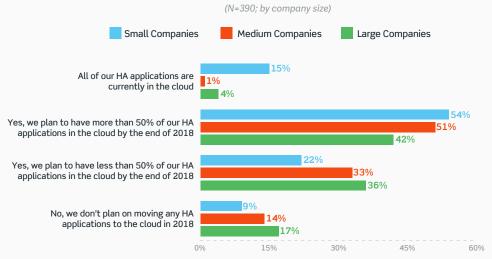Yes, we plan to have more than 50% of our HA applications in the cloud by the end of 2018

Figure 18: Future plans for migrating critical applications to cloud providers (by company size)

Are you planning to move more of your HA applications to the cloud in 2018?

(N=390; by company size)

■ Small Companies   ■ Medium Companies   ■ Large Companies

All of our HA applications are currently in the cloud
- 15%
- 1%
- 4%

Yes, we plan to have more than 50% of our HA applications in the cloud by the end of 2018
- 54%
- 51%
- 42%

Yes, we plan to have less than 50% of our HA applications in the cloud by the end of 2018
- 22%
- 33%
- 36%

No, we don't plan on moving any HA applications to the cloud in 2018
- 9%
- 14%
- 17%

0%   15%   30%   45%   60%

To a lesser, but not unsubstantial, extent, respondents also cite lack of skills (22%) and concerns around cost (16%) as reasons to avoid moving all applications to the cloud. Moreover, 14% of respondents' worries revolve around the perception that they may not be able to provide high availability services to workloads that operate in the cloud. In comments included with the data, some respondents indicate that they also have concerns around security. These worries are pushing them to maintain an on-premises data center presence for key applications.

With these concerns understood, just how well are cloud providers countering them? Are they taking steps to mitigate the challenges that are holding people back from full cloud adoption? The only way to really determine if cloud migration concerns are being addressed is to analyze customer acceptance. We requested information from respondents around their future plans as they pertain to moving additional high availability applications to the cloud. Nine percent (9%) of respondents indicate that all of their most important applications are already in the cloud. By the end of 2018, one-half of respondents expect to have more than 50% of their HA applications migrated to the cloud, while 29% say that they will have less than half of the HA applications in such locations. Finally, 12% of respondents say that they will not be moving any more HA applications to the cloud in 2018.

Figure 18 provides a breakdown of this same information by company size. Small companies are leading the charge when it comes to operating their HA workloads in the cloud, with 15% of such companies indicating that all of their HA applications already reside in the cloud, a statistic that is far ahead of medium and large companies. Of course, smaller companies typically have fewer such
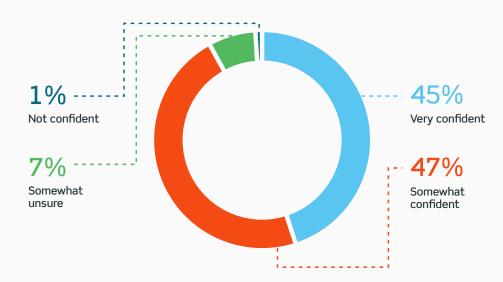
Are you planning to move more of your HA applications to the cloud in 2018?

(N=390; by role)

■ Decision Makers   ■ IT Staff   ■ Data and Development Staff

All of our HA applications are currently in the cloud
- Decision Makers: 15%
- IT Staff: 1%
- Data and Development Staff: 4%

Yes, we plan to have more than 50% of our HA applications in the cloud by the end of 2018
- Decision Makers: 54%
- IT Staff: 51%
- Data and Development Staff: 42%

Yes, we plan to have less than 50% of our HA applications in the cloud by the end of 2018
- Decision Makers: 22%
- IT Staff: 33%
- Data and Development Staff: 36%

No, we don't plan on moving any HA applications to the cloud in 2018
- Decision Makers: 9%
- IT Staff: 14%
- Data and Development Staff: 17%

0%   15%   30%   45%   60%

Figure 20: Confidence in technical team's ability to understand availability objectives and SLAs

How confident are you that your infrastructure team/specialists understand your specific objectives and SLAs for high-availability applications in the cloud?

(N=390; by role)

1% Not confident

7% Somewhat unsure

45% Very confident

47% Somewhat confident

workloads, so it's an easier shift for them. Likewise, small companies are more likely than other companies to have plans to push more than 50% of their remaining HA applications into the cloud by the end of the year.

As is sometimes the case, people in different roles have a different view of where things are headed. This statistic is no exception. It's interesting to note that 60% of those in data or development roles see their company pushing more than 50% of their HA applications to the cloud by the end of the year, compared to about 43% of decision-makers and IT staff. It could be that developers have a different definition of HA applications, or they could be operating based on hope rather than concrete plans.

Not all concerns are around technology. Some organizations experience uncertainty around whether their infrastructure teams have a full grasp of the business's objectives around application usage, and whether their service-level agreements (SLAs) are sufficient, given the importance of specific workloads in the cloud. Fortunately, this is generally a good news situation for respondents. Forty-five percent (45%) say that they're very confident in their team's understanding of cloud-based applications and objectives. However, a full 47% are only somewhat confident, leaving room for doubt. Seven percent (7%) are somewhat unsure about how well their staff are trained to understand these topics. And, for an unfortunate 1% of respondents, they have no confidence that their teams appreciate the criticality of their cloud-based workloads.

ORGANIZATIONAL AGILITY AROUND HA APPLICATIONS

The first step in correcting a problem is understanding what problem you're trying to solve. Gaining knowledge

When performance issues arise in high availability applications (SQL Server, SAP, Oracle, other) in the cloud, it's most often a(n) …

(N=382)

4%
We are never really sure

17%
Cloud platform provider (Azure, AWS, GCP, etc.) issue

15%
Infrastructure issue

64%
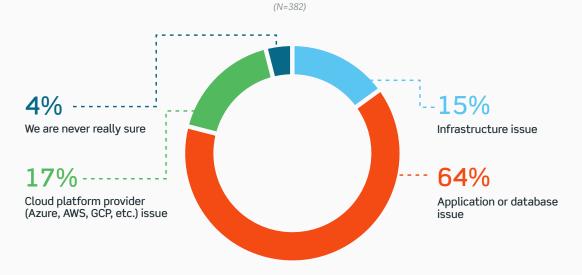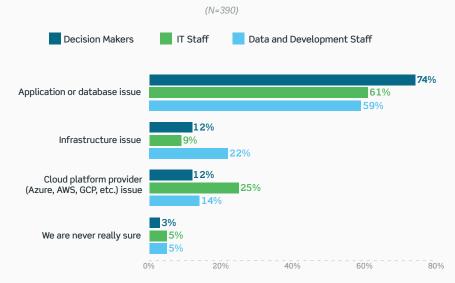Application or database issue

Figure 22: The primary culprit in performance issues in cloud-based highly available applications (by role)

When performance issues arise in high availability applications (SQL Server, SAP, Oracle, other) in the cloud, it's most often a(n) …

(N=390)

- Decision Makers
- IT Staff
- Data and Development Staff

Application or database issue: 74% / 61% / 59%
Infrastructure issue: 12% / 9% / 22%
Cloud platform provider (Azure, AWS, GCP, etc.) issue: 12% / 25% / 14%
We are never really sure: 3% / 5% / 5%

around where performance problems arise in high availability applications means that they can be solved more quickly. Fortunately, just 5% of respondents say that they never really know what causes application performance issues. For 64% of respondents, the most likely culprit is something internal to the application or the database. In what is not a resounding advertisement for cloud providers, 17% believe that their application performance issues typically stem from issues related to the cloud platform in use. Finally, for 15% of respondents, their application performance problems are most often traced back to an infrastructure issue. Bear in mind that this is not a definitive assessment of performance problems. Respondents were asked to provide a response for where they most often see a performance problem arise, and were allowed to choose only a single response.

There has always been some finger-pointing in IT when something goes wrong. Network administrators, for example, have become accustomed to the perception that everything going wrong in the company is the network's fault. **Figure 22** shows that this practice continues today. You'll see that, when application performance suffers, 74% of decision-makers blame the applications team, which blames itself only 59% of the time. In turn, the applications team points the finger at infrastructure 22% of the time, compared to just 9% of the time that it blames the IT operations team. Finally, 25% of the time, the IT operations team is content to blame the issue on a cloud provider, which is about double the tendency of the other two groups of respondents.

Remember: Everyone is coming from their own perspective, and has their own interests to protect.

How do you identify and characterize complex reasons for cluster failover?

(N=390)

Every organization needs a process to determine why an application has failed over, so that preventative measures can be taken to avoid future occurrences. Sometimes, that means looking at reporting tools, a technique used by 65% of respondents. Other times, it's a simple matter of trial-and-error that gets the job done, a technique used by 46% of respondents. And, as is the case for around 5% of respondents, the problem eventually goes away and the cause is never determined. Obviously, this isn't sustainable for these companies. The lack of ability to determine root cause for failover events will ultimately catch up to them.

Figure 24: Methods used to ensure high availability for critical applications in the cloud

How do you ensure high availability of mission-critical applications in the cloud?

(N=390)

Designing a highly available environment from the beginning is the first step in the ongoing success of that application. Administrators use a variety of techniques to do this. They include eliminating single points of failure via technologies like clustering (71%) and over-provisioning resources (24%) to ensure that a resource constraint doesn't bring down a workload. Others -- a full 25% of respondents -- have turned to experts. They outsource their HA activities to a service provider rather than trying to go it alone (**Figure 24**).
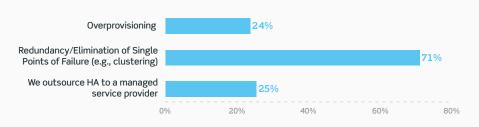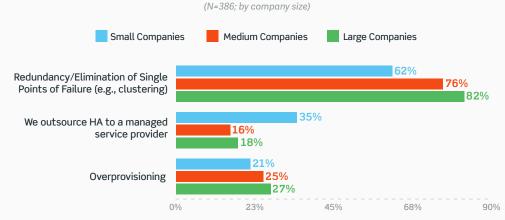
Figure 25: Methods used to ensure high availability for critical applications in the cloud (by company size)

How do you ensure high availability of mission-critical applications in the cloud?

(N=386; by company size)

As shown in **Figure 25**, small companies are far more likely to outsource their HA needs to a managed service provider, with 35% of small companies indicating that they've gone this route. This is about twice the rate of medium and large companies doing the same.

In addition, small companies are far less likely (62%) to deploy redundant configurations and clusters when compared to medium-sized (76%) and large companies (82%). It's unlikely that smaller companies place less importance on their workloads; rather, they may simply be unable to handle the cost and complexity that can be associated with high availability.

The Value of Proactivity

The importance of the ability to quickly understand application performance and other critical issues cannot be overstated. So, the question is, would respondents value the ability to predict challenges in these areas sooner rather than later so that they can take mitigating steps to prevent an issue from happening? Sixty percent (60%) of respondents say that Yes, they would place a high value on this ability, and that they need it. Coming down a step, 35% of our survey respondents indicated that they see such capability as nice to have, but not as a necessity. Rounding out the stat, 5% don't believe that such a tool would have a reasonable value in their organization.

Figure 26: Value placed on being able to identify both application performance and storage utilization using a single tool

How much would you value the ability to predict both *application* performance and storage utilization in your high availability applications in the cloud, and eliminate threats to availability before they happen?

(N=390)

5%
Low value

35%
Medium value. Nice to have, but we can live without it

60%
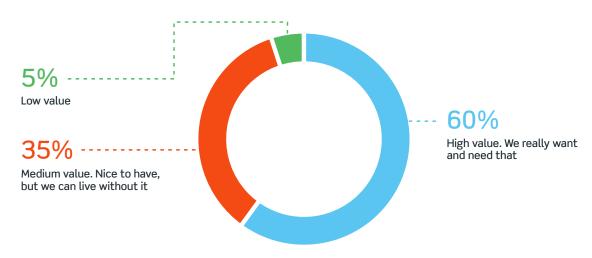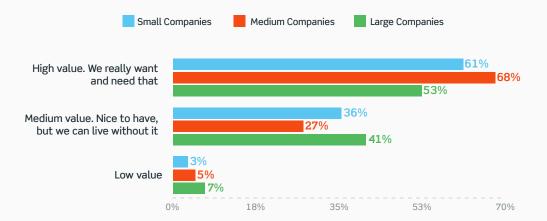High value. We really want and need that

Figure 27: Value placed on being able to identify both application performance and storage utilization using a single tool (by company size)

How much would you value the ability to predict both *application* performance and storage utilization in your high availability applications in the cloud, and eliminate threats to availability before they happen?

(N=390; by company size)

■ Small Companies   ■ Medium Companies   ■ Large Companies

High value. We really want and need that
- 61%
- 68%
- 53%

Medium value. Nice to have, but we can live without it
- 36%
- 27%
- 41%

Low value
- 3%
- 5%
- 7%

0%   18%   35%   53%   70%

Remember that different people interpret the word "value" differently. Some consider it in terms of what a particular tool can do, and some perceive "high value" as "high price," which may not be true. Further, company size plays a role in how people perceive the value of such tools. Those in medium-sized companies are more likely than others to place a high value on them (**Figure 27**).

> ## Just 5% of respondents told us that they never suffer an availability failure. In other words, there is significant room for improvement on this front.

What if you've taken what you believe to be the right steps, but your high availability applications still suffer a failure? Perhaps you've implemented some level of availability, only to find that it doesn't protect you as thoroughly as you'd hope. For 95% of our respondents, this is reality, but to varying degrees. There are a number of respondents – 26%, to be exact – that told us that their availability service fails at least once per month. This is a difficult statistic to grasp, as it would seem that there's a fundamental flaw somewhere that needs to be corrected. Fortunately, not everyone is faring this badly. Twenty-eight percent (28%) of respondents suffer an availability failure every three to six months, and 16% have a similar experience every six to twelve months. Finally, in the "it happens every once in a while" camp, 25% say that they suffer a failure

How often do your availability services that protect your applications degrade or fail?

5%
Never - they always work perfectly

25%
Rarely - once a year or less

16%
Once every 6 to 12 months
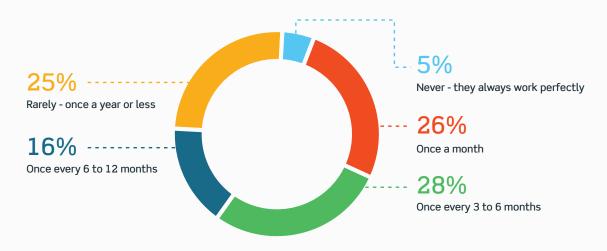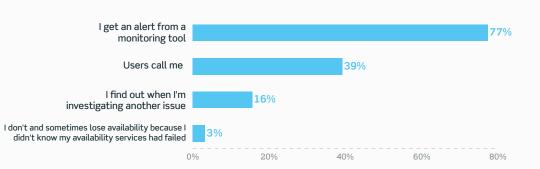
26%
Once a month

28%
Once every 3 to 6 months

Figure 29: Availability service failure discovery methods

How do you discover that your availability services are failing or failed?

(N=372)

I get an alert from a monitoring tool — 77%
Users call me — 39%
I find out when I'm investigating another issue — 16%
I don't and sometimes lose availability because I didn't know my availability services had failed — 3%

only once per year or less often. All that said, just 5% of respondents told us that they never suffer an availability failure. In other words, there is significant room for improvement on this front.

The downside of an availability failure is discovering that one has taken place! Only when you've discovered this situation can you fix it, so having a way to make this discovery is critical. Respondents gave us some good news and some bad news here. On the good news front, 77% are alerted from a monitoring tool that availability has gone bad. But it appears that this isn't always acted on quickly enough, as 39% of respondents also told us that they made the discovery via user feedback. (It's possible that admins check the monitoring console after users call them, hence the discrepancy.) Another 16% say that they basically find out by accident while they're troubleshooting something else. Finally, 3% say that they simply don't find out that availability has failed. This obviously leaves their organizations in a tremendously precarious position.

Although this survey and report represent people's thinking at a single point in time, there are some potentially important trends that emerge: first, it's clear that organizations value their mission-critical applications, as they're protecting them via clustering or other high availability technology. A second takeaway is that even with those safeguards in place, there's more work to be done, as those apps can still suffer failures and performance issues. Companies need to look at the data and ask themselves, therefore, if they're doing everything they can to protect their crucial assets.