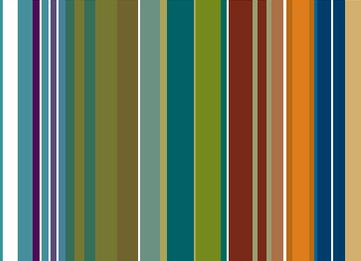




THE NEED FOR HIGH AVAILABILITY AND UPTIME



All Clouds Are Not Created Equal

INTRODUCTION

Companies increasingly are looking to the cloud to help deliver IT services. Many have already moved email, sales force management, and other applications to Software-as-a-Service (SaaS) cloud service providers. And now there is growing interest in infrastructure services to cut costs and to offer companies the agility to be more responsive when business conditions change.

However, companies have been hesitant to move production and business-critical applications to public clouds due to concerns about performance, reliability, and security. Recent outages of some cloud providers' services, which knocked out major sites such as Instagram, Netflix, and Pinterest for hours or longer, confirmed their greatest fears. A most telling comment about the recent outages came from the *Wall Street Journal*. It noted¹ that "[what we're seeing] is the rapid expansion of a big new industry that is still in its shakedown phase—finding and fixing problems."

This should be a cause for concern for any company that wants to run production applications in the cloud or use cloud services for other vital business operations such as backup or disaster recovery.

However, the point to keep in mind is that not all cloud services are created equal. To run production applications, organizations must select a provider that offers high availability and uptime, backed up with service level agreements.

One recent uptime study³ found that businesses lose an average of about \$5,000 per minute (\$300,000 per hour) in an outage.

THE NEED FOR HIGH AVAILABILITY/UPTIME

The use of public cloud services for SaaS is widely embraced and for infrastructure needs it is slowly gaining momentum. A survey² of 600 global enterprise and mid-market companies published in early 2012 found that 27 percent were using public cloud Infrastructure-as-a-Service (IaaS) solutions. That was 10 percent higher than what was found in a similar survey published in early 2011.

To date, the main uses of SaaS and IaaS have been for everything but mission-critical applications.

Organizations are typically using SaaS for email, calendaring, and internal collaboration. They use IaaS for test and development, to build and deploy new Web-based applications, and to move non-critical applications that are not subject to regulatory requirements off of expensive to maintain on-premises data center equipment.

Another area where both SaaS and IaaS are finding acceptance is when a business unit needs to respond to a new opportunity and doesn't want to wait for IT to provision internal resources.

However, most organizations are taking a cautious approach when it comes to using the cloud for key business applications.

The reason: Businesses run 24/7. Employees want to check mail, schedules, and share information around the clock. The global nature of organizations, their supply chains, and customer base means systems must be available at all times.

While downtime might be acceptable to the general web surfer, businesses cannot afford that luxury. If a customer tries to check his account or place an order and finds a site or application down, he can easily take his business to a competitor. Workers who cannot access key applications lose valuable time and the company incurs mounting lost productivity for each lost minute or hour.

The costs can quickly add up.

Regardless the source of an interruption in availability, the end result is the same. Namely, business suffers. One recent uptime study³ found that businesses lose an average of about \$5,000 per minute (\$300,000 per hour) in an outage. In several industries, downtime can be even more costly. Past studies have pegged the cost of an hour of downtime at \$1.1 million for a retailer and up to \$6.48 million per hour for a brokerage firm.

Making these numbers particularly worrisome is that many businesses routinely experience outages. In 2009, Dunn & Bradstreet found⁴ that 49 percent of Fortune 500 companies experience at least 1.6 hours of downtime per week. That translates into more than 80 hours annually.

Businesses obviously need to take steps to avoid costly downtime, lost productivity, and regulatory problems. And if an outage does occur, service recovery must be rapid. All of these points are amplified when production and mission-critical applications are involved.

A properly architected provider infrastructure can help ensure minimal service disruptions.

EVALUATING A PROVIDER

For production applications used to run a business, not all clouds are created equal. Organizations need to select a cloud service that offers availability and uptime characteristics that match its applications' tolerances for downtime.

Immediately, organizations will find that there are not only great technology differences between providers, but there are also variations in operational procedures, responses to problems, and the way security is handled.

So what are the key characteristics to look for in order to run production applications in the cloud with the assurance they will be highly available?

Essential infrastructure elements: First, look at the cloud provider's architecture. Does the provider employ an enterprise architecture with built-in resiliency and security designed for production workflows? Are there multiple connectivity paths to the provider's site? If a site goes down due to a power failure, natural or man-made disaster, how is it restored and how fast?

Is there more than one site? If one site becomes inaccessible or its services are unavailable, do workloads automatically failover to a second site without disruption? Can you load balance between sites?

All of these points are important. A properly architected provider infrastructure can help ensure minimal service disruptions.

Dig a layer deeper in examining the infrastructure. Are the services based on open source solutions or best-in-class solutions? Who are the provider's technology partners? These issues can become important when problems arise. If the provider has strung together services based on open source solutions, its technical staff will have to go it alone when trying to troubleshoot problems and restore services. On the other hand, a provider that partners with technology leaders will have the expertise and help of those companies when trouble occurs.

Next, find out about the provider's operational procedures. What happens if something fails? Suppose a blade server crashes. Is restoration automatic? How long does it take? Are the OS, application, drivers, and data restored instantly or does the server have to be rebuilt from scratch?

What if a site goes down? How does the provider handle routine problems like cut cables or power outages? Are there redundant line feeds from different telecom providers into the data center?

Major provider sites would naturally have onsite backup power generators. But how often is that generator and its ability to automatically kick into action tested? This may seem trivial; one might assume providers routinely test their backup power solutions in real-world scenarios. Don't assume. A recent widespread and prolonged outage of a major provider's services was due to the failure of a data center to switch over to its backup generators.⁵ This eventually drained the center's uninterruptable power supplies resulting in shutdown of all site hardware.

How are the data centers staffed? Is there someone on site 24/7? What are the provisions in place to bring staff in when problems arise?

Hosting key business applications and their associated data in the cloud means organizations also must take a deep look at the provider's approach to security. How does the provider handle system and physical security?

When evaluating cloud service providers, check to be sure its application and infrastructure service level agreements (SLAs) match the characteristics of the production applications that will be using the service.

On the system security front, does the provider follow best practices for protecting systems from malware and cyber attacks? Many government and industry groups (such as ISO and NIST) have developed security guidelines and frameworks. Which ones does the provider follow?

How does the provider handle physical security? How is the data center secured? Who is allowed access? What mechanisms are in place to ensure only authorized staff have access to the server room? Are the server racks locked down?

Procedural aspects: Many organizations need to abide by regulations on availability, data protection, and data privacy. How does the provider handle these issues?

Look at the provider's governance, risk, and compliance strategies. What security practices are followed to protect production workflows?

For organizations that operate in regulated industries, how does the provider address and assist with compliance with regulations such as PCI-DSS (Payment Card Industry Data Security Standard), HIPAA (Health Insurance Portability and Accounting Act), Sarbanes-Oxley, and many others? And to help ensure compliance, what auditing practices does the provider follow?

Do the provider and its staff have the needed certifications to carry out the security procedures and guarantee that regulatory standards are met? Increasingly, one measure of a provider's security and regulatory fitness is having SSAE (Statement on Standards for Attestation Engagements) 16 Type II certification. This certification, related to auditing, is typically requested if a service provider's services affect the financial statements of another company. Most publically traded companies require that their service providers have this certification.

Availability issues: Production applications have little tolerance for downtime. To run such applications on a public cloud service, organizations must have contractual guarantees on availability and uptime.

When evaluating cloud service providers, check to be sure its application and infrastructure service level agreements (SLAs) match the characteristics of the production applications that will be using the service. And make sure you understand exactly what the service level agreements cover and what points you are responsible for.

Take a deeper look at the provider's infrastructure and procedures. How does the provider address resiliency in its infrastructure? Are there multi-site options for high uptime of mission-critical applications? Can the sites be configured so both are primary, load-balancing when all systems are fine and failing over with no downtime if services at one site are lost?

For applications that can accommodate some risk and downtime, what are the availability options based on the recovery time objectives (RTO) and recovery point objectives?

SunGard's Enterprise Cloud Services is fully managed and built to meet the requirements of an organization's most critical production applications.

Basically, what it comes down to is that providers need to offer multiple layers of data and application protection and availability to run production applications.

SUNGARD AS YOUR TECHNOLOGY PARTNER

Many companies today are challenged by lack of IT resources and technology infrastructure to support their critical business applications. Having reaped the benefits of public cloud services for many of their other applications, there is growing interest in leveraging the cloud for production applications.

But because these applications are so vital to the health of the business, the cloud service must have enterprise-class availability and uptime. That's where SunGard can help.

SunGard's Enterprise Cloud Services is fully managed and built to meet the requirements of an organization's most critical production applications. To deliver the services, SunGard uses a highly secure, enterprise-grade virtual data center infrastructure that delivers the high availability and scalability needed by businesses today. These services are backed with the security and experience SunGard is known for. To that end, SunGard offers multiple layers of data protection, with a wide range of recovery options mapped to the needs of each application.

SunGard's Enterprise Cloud Services is built on best-in-class Vblock architecture, which is comprised of Cisco, EMC, and VMware technology. Availability is backed by service level agreements for both production and recovery environments.

For applications that require the highest level of uptime, SunGard offers Managed Multi-site Availability. This includes automated failover to a replicated cloud environment or fast recovery of a cloud infrastructure.

SunGard's Enterprise Cloud Services is delivered in data centers built to the ITIL v3 framework, audited under SSAE 16 Type II, and certified to the ISO 20000-1 standard. Its cloud platform is built to support the regulatory requirements of PCI-DSS and HIPAA.

Organizations exploring the idea of moving their production applications to the cloud can use optional SunGard consulting services including a Cloud Readiness Assessment and Cloud Migration Plan to help assess and migrate an environment to the cloud and ensure a strong return on investment.

Additionally, SunGard's Infrastructure-as-a-Service (IaaS) solution offers the expertise and services needed to move to key applications such as SAP, Citrix, MS SQL, MS Exchange, Active Directory, intrusion detection systems, and geographic load balancing to the cloud.

For more information about SunGard's high availability cloud services, visit: www.sungardas.com.

¹ <http://online.wsj.com/article/SB10001424052702303644004577523134171172096.html>

² <http://www.networkworld.com/supp/2012/enterprise2/040912-ecs-iaas-257610.html>

³ <http://www.eweek.com/c/a/IT-Infrastructure/Unplanned-IT-Downtime-Can-Cost-5K-Per-Minute-Report-549007/>

⁴ http://www.information-management.com/infodirect/2009_133/downtime_cost-10015855-1.html

⁵ <http://www.zdnet.com/amazon-web-services-the-hidden-bugs-that-made-aws-outage-worse-7000000186/>