Why is High System Availability Important ?

- in " Cloud Computing " and with " Platforms "

How many " 9's" does your application need to - meet or exceed user expectations, address increasing business demands, and avoid outage impact ?

When it comes to system availability, whether for " cloud computing " or on a " platform " , people frequently hear about " 9's " of uptime. Many service level agreements (SLAs) specify at least 99.9 % availability (3 9's - or - < 9 hours of downtime / year), or, 99.99 % (4 9's – or – < 1 hour of downtime / year). In conjunction with this, to further define system performance, the SLA Availability Metrics agreed on need to be from the End User perspective and apply from the time the application initially goes live.

When people talk about a service being highly available, the following chart is what they're referring to. The major cloud and platform providers can perform to the level requested with traditional systems typically with an army of technical personnel, many tools, as well as decades of established processes and experience. Alternatively, with Fault-tolerant and newer networked systems, the burden to achieve high availability level is less – either a little less or a lot less – depending on the availability objective as well as the sophistication of the cloud service or platform. As you can see in the chart, the more demanding the application or service importance to the organization, the more 9's the better -

High availability percentages of SLAs

PERCENTAGE	YEARLY DOWNTIME*
99.9	8hr 45m 57s
99.99	52m 35.7s
99.999	5m 15.6s
99.9999	31.6s
99.99999	3.2s
99.999999	0.3s
99.9999999	31.6 ms

Because achieving very high service availability is very challenging and costly with popular systems, is why only very large organizations with extensive funding and resources can have these information services in-house. However with costs and complexity increasing, to reduce the burden and get better value, is why Cloud Computing and/or using Fault-Tolerant Platforms is appealing. This is important in an **increasingly on-line, real-time, all-the-time world** to meet expanding business needs, rising User expectations, improve agility and responsiveness, deliver greater value, etc.

For example, a simple static website can easily expect to achieve 4 9's or more of uptime - because there are few potential points of failure. *Yes, there's the web server and the machine it is running on, but with the addition of a floating IP address, a load balancer and a redundant server - it would be rare if people experienced any preventable downtime.* For a more complex, monolithic web application, while 4 9's may be possible, the pressure to achieve it increases as you add components to the mix – ie: new services / applications, many Users, numerous devices, a database, caching servers, object storage, etc. Further, segment the application into microservices, and the number of potential failure points grows !

Understanding this is important since as the complexity of an application increases so too does the risk of losing a 9 in availability. With popular systems, while you can add more redundancies to improve availability, this increases costs and complexity. For example, it's challenging to keep multiple copies of a database in sync, avoid data corruption when a failure occurs, etc.). In addition, because IT operations demands have grown and become more complex, it's to the point where there are cases where there too many alerts for human operators to process and little to no visibility into which alerts are impacting business. And it's only getting worse with the addition of new devices and the growing list of enterprise digital services — which further slows detection and resolution times. This is a huge issue that is becoming more problematic because of the many siloed system management tools, the periodic need for manual intervention, and the challenges to find the exact root cause of problems – despite extensive work and investment to automate processes. The fundamental issue is trying to fix the plane in flight !

With there being much information to achieve the needed levels of availability, the next step is to identify the consequences of losing a nine in your SLA. For example, how will your customers react if you have 54 minutes of downtime versus 540 minutes or 5,400 minutes in a year ? How many customers will you lose at each of those levels ? See **Aligning Applications and Platforms** for insights on improving outcomes by mapping application importance / business impact – with – the appropriate cloud service / platform. Extending on the 9's requirement, given the importance of success with digital services to the future of every organization, have penalties for SLA non-compliance (ie: $1m / minute for excessive downtime / system outage from when the application goes live). This is important to respect the consequences of not performing (ie: financial loss, business devaluation, lost Customers, inability to attract new business / Customers, challenges to attract and retain talent, etc.)

These are important considerations to address in an SLA – whether for a Cloud service or with Platforms to support information services essential to the business – today, and even more so going forward. High availability is important – and the more important the application, the more important it runs in a very high availability computing environment. Examples of important applications where the need is for 7 9's or high availability are - ATMs, POS and Payment Systems, communication services, e-commerce platforms, numerous enterprise on-line services, tele-health, etc. From this, it's clear these digital capabilities are essential for an " always on organization ". In addition, to prudently manage the downside, 7 9's or higher availability are also important to mitigate risk and avoid brand damage from unplanned system outages or poor performance during peak activity periods. For confirmation of the high cost of not understanding risk, ask the folks at Wells Fargo, NASDAQ, Lloyds, Robinhood Markets, Google, etc. - who experienced severe consequences from system outages with their on-line services !

For additional insights on better positioning your organization for success with high availability information services, and how NonStop contributes in the " Cloud " or as a " Platform ", see –

Improving Business Outcomes with NonStop

https://www.nonstopinsider.com/uncategorised/the-abcs-of-an-affordable-business-critical-compute/

Oct 30, 2020 By Ronald Thompson , CAIL   www.cail.com   info@cail.com