Resilience in the era of enterprise cloud computing

Design considerations for forward-thinking organizations

IBM

Contents

Introduction

In the beginning, when IT systems and storage were housed internally in a single data center and accessed only by employees, maintaining resilience was far less complicated. Information was managed locally, with nightly backups and offsite storage providing sufficient coverage. Business continuity and security technologies may have been primitive by today's standards, but the risk of exposure was considerably less and the threats far less damaging.

Today's IT environment is altogether different, and it has brought with it a world of new resiliency and continuity challenges. Most organizations rely on an ecosystem of internal and external systems and data to conduct their daily business. Access privileges extend beyond employees to business partners and customers. All of these users are generating unprecedented amounts of data, and their expectation for "always-on, always-available" service is pervasive. They expect to connect whenever and wherever they want via a growing array of devices, from desktops and laptops to tablets and smartphones. This veritable explosion of cloud-based IT and business solutions, combined with the demand for mobility and social connection, has forever changed how, when and from whom technology and services are delivered.

Understandably, keeping such an environment completely protected from adverse events while maintaining continuous processing capability has become increasingly complex. This paper addresses the resiliency challenges facing organizations by examining the elements essential to enabling an effective resiliency program in the era of enterprise cloud computing.

The big picture—extending processing beyond the data center

Cloud computing has already played a major role in the evolution of the data center. While cloud does not account for a large percentage of processing today, it is on a rapid trajectory. Most companies have big plans for cloud and expect to roll it out in great numbers. In its 2013 CloudTrack Survey, IDC found that the percentage of cloud-using organizations more than doubled globally to 52 percent in a single year. Moreover, 77 percent claim to have moved to a "cloud-also" strategy, putting cloud-based solutions on equal footing with traditional IT solutions each time they are looking at implementing new IT functionality or replacing existing functionality.[1]

Cloud offers an extremely scalable, flexible platform—one that extends the availability of compute resources well beyond traditional data center boundaries. It enables organizations to get more from their virtualized systems, storage, networks and infrastructures, increasing both availability and seamless access to the critical IT services required for business growth. And cloud is enabling new user communities to consume these resources. Once limited to internal and corporate users and a few external business partners (B2B), access has now extended outward to all kinds of potential consumers (B2C).

77%

of companies have moved to a "cloud-also" strategy[1]

At the same time, the consumerization of IT and pervasive mobility have driven the demand for always-on service and instantaneous access to information. Users want information to be available around the clock, and they want it easily consumable from any device. The challenge is that digital data is being created at a dizzying pace in a wide variety of structured and unstructured formats. The escalating volume of "big data" has made it more difficult to assure the continuous availability of data in daily processing as well as in the event of a disaster. It has also made high availability solutions and comprehensive disaster recovery strategies mandatory.

Certainly, in this new IT environment where traditional enterprise computing and advanced cloud technologies interact continuously, delivering resilience and business continuity is more complicated. Figure 1 provides a high-level view of the current topology. The infrastructure must be capable of providing the resiliency needed for daily operations while safely accommodating explosive data growth and expanded access to systems and data. Processing can take place on premises using traditional enterprise computing resources and off premises using public, private or hybrid clouds to augment internal resource capacity and availability. Organizations can accommodate sporadic processing needs without an increased capital investment, including non-production development and testing, analytics and disaster recovery workloads.

When designing an always-on, always-available IT environment, there is no doubt that resilience should be front and center. As cloud technologies continue to mature and as consumer expectations continue to rise, it is IBM's belief that the success of a resiliency program will depend on having a solid strategy that addresses the evolving nature of IT. Toward that end, special consideration must be given to:

- Addressing the advances brought about by cloud
- Managing the big data explosion
- Enabling systems of engagement

All three are pertinent to the design of a resilient architecture. Understanding these implications and using that knowledge to build resilience into the business operation are vital. At stake are organizations' readiness and ability to respond quickly to threats and disruptions in the era of enterprise cloud computing.

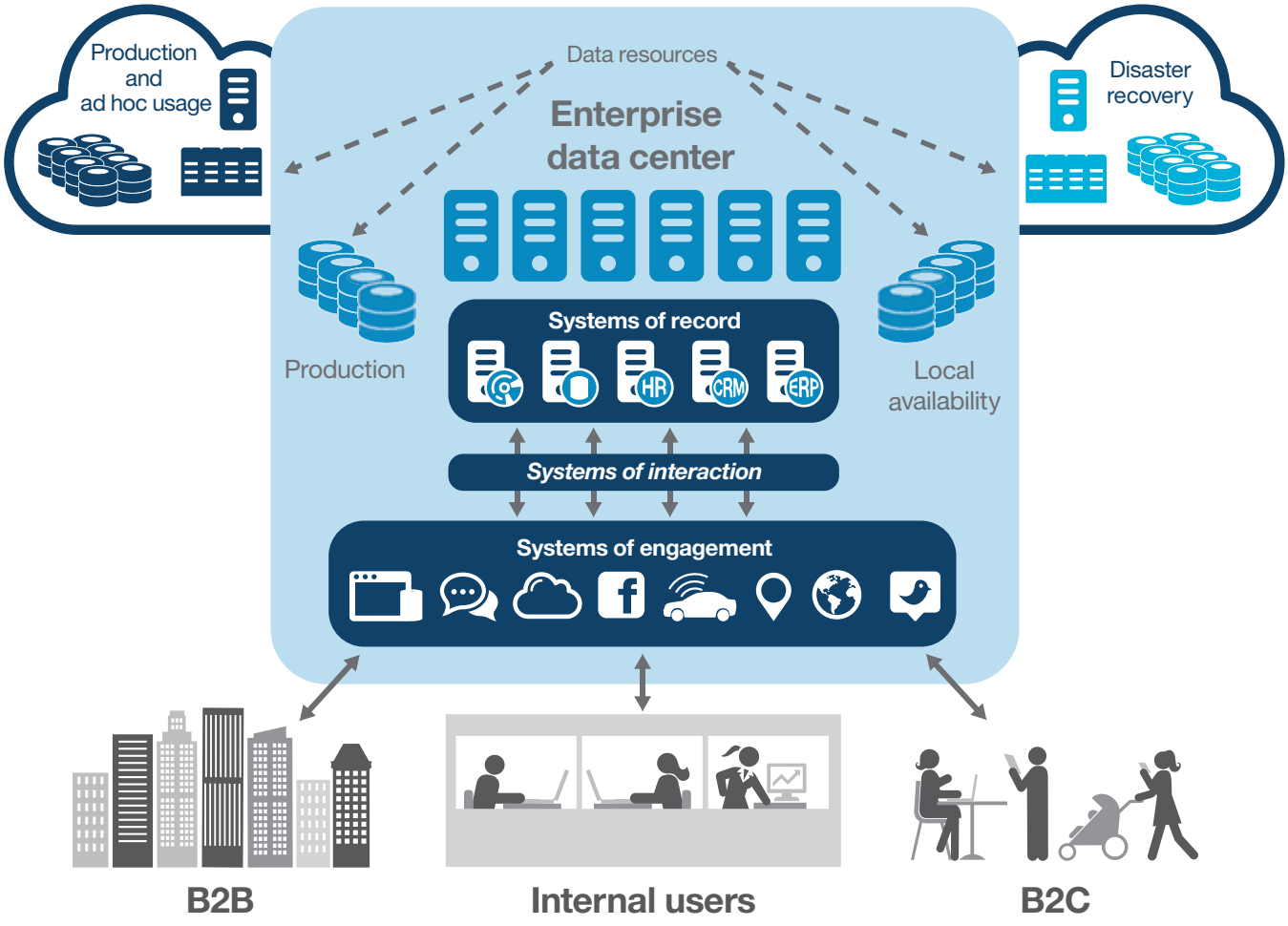


Figure 1. **High-level view of today's IT environment.** With traditional enterprise computing and advanced cloud technologies interacting continuously, organizations are increasingly challenged to deliver resilience and business continuity cost-effectively.

Addressing the advances brought about by cloud

IT operations have evolved from purely physical, dedicated resource allocation to a more technologically advanced, automated approach that leverages shared resources for all kinds of workloads, from production to disaster recovery. Most companies have put a virtualization strategy in place that enables server, storage and network resources to be configured so that a large number of images can be run via a single physical footprint. From a technology standpoint, advances like these have laid the groundwork for cloud computing and led to an array of options for extending processing to the cloud.

With its roots in mainframe computing, server virtualization has perhaps seen the greatest advances. In the late 1960s, multiple guest operating systems were resident on a single physical processor. Today, with advances in server hardware and virtualization software, the ability to share pooled resources across numerous physical systems is more the norm than the exception.

Storage architectures have also seen significant progress, providing denser, faster access and making increased volumes of data available for processing. The ability to create "on demand" instances of data across multiple sites in a relatively short timeframe has been key. It has advanced storage utilization in a manner that reduces the need to maintain duplicate copies and additional hardware, providing organizations with increased availability and recovery at a more efficient price point.

Network advancements have increased availability too, and they have expedited recovery by improving organizations' ability to access systems, move data and dynamically divert and balance user traffic. By combining multiple isolated physical network resources into a single, software-based virtual network, organizations are able to enhance performance and throughput while improving the efficiency of system and application communications.

The evolution of the server, storage and network virtualization technologies underlying cloud is far from over, as is their impact on resiliency. As new capabilities roll out in the cloud, their impact on business resilience must be addressed. This necessitates some additional areas of focus in the design of an effective resiliency program.

The effects of technology advances on resilience should be a front-and-center consideration when designing a cloud-based, always-on, always-available IT infrastructure.

Techniques for building resiliency in cloud deployments

Complete, enterprise-level recovery requires the integration of all cloud and traditional IT platforms. It also requires that their associated server, storage and networking designs comprehensively address the resumption of business operations. The overall cloud design must be taken into consideration, be it public, private or hybrid. This will help determine what resiliency measures are taken and how they are implemented. Additionally, from a logistics standpoint, it is important to understand whether cloud capabilities are being delivered locally or remotely.

Once the requirements for integrating cloud into the enterprise computing environment have been identified, many of the very same techniques used to enhance the production environment can be used to enhance resiliency as well, with certain areas requiring increased focus and attention:

- Virtualization techniques enable organizations to improve their recovery time objectives (RTO) and recovery point objectives (RPO) by allowing system images, associated configuration specifications, and all applications and data to be encapsulated and transmitted to a remote site for use in recovery.
- Automation is essential to rapid recovery. Automated server provisioning and system orchestration will need to be implemented and fine-tuned to meet business expectations. This will require detailed knowledge of the organization's server restoration priorities, along with a blueprint for connecting to critical storage and networks during the recovery process.
- Organizations should have the capability for cross-platform, multivendor storage replication to address data synchronization concerns associated with managing storage instances that are replicated to remote sites.
- As with traditional storage replication, isolated copies of storage should be utilized for recovery validation, with the primary replica remaining intact during testing to help assure business viability.
- Consistent LAN/WAN connectivity for data replication and user access that meets both RTO and RPO objectives as well as user (B2B and B2C) expectations for access and performance must be designed and validated.

Key resiliency metrics

Recovery time objective (RTO). The duration of time and the service level within which a system must be restored after a disruption to avoid a break in business continuity. It refers to the point in time when complete business operations have resumed.

Recovery point objective (RPO). The maximum tolerable period in which data might be lost. Focused on backup frequency, it refers to the point in time in the past to which a system is recovered.

Managing the big data explosion

Data has become a natural resource, and its growth, availability and protection have become vital to meeting business objectives. Increased scrutiny over how data is managed and made available to users is needed. Organizations must understand how data will be protected, archived and recovered to maintain continuous availability. A resiliency strategy should address:

- How data will be accessed—and by whom. As user expectations increase, organizations will need to provide an always-available IT design that is capable of providing real-time access to data from a multitude of devices that serve a wide variety of users.
- How application dependencies and user performance and throughput will affect the integration of enterprise and cloud storage. Storage needs to be designed and delivered relative to the actual location of data and overall access requirements to make sure that information accuracy and accessibility meet business expectations.
- How cross-platform data synchronization will be provided. Identifying and architecting data synchronization is especially important when considering the movement of workloads from the enterprise data center to external cloud-based processing sites.
- The role that each data type (file, database, graphics, and so on) will play. This enables organizations to determine how and where data should be positioned and stored, and how it should be managed.

Techniques for building resiliency with big data

As the volume of data grows, the amount of storage required to support the production environment—and the business—increases too. Such rapid and enormous growth puts increased strain on the network for data creation and information sharing, but also for archiving and disaster recovery. Meeting user expectations for data availability and access requires a redesign of the storage and networking infrastructure.

It is likely that organizations are already leveraging some of the best methods for protecting data. For example, most have implemented various forms of server or storage based snapshots or replication for enhanced information availability. The real advantage comes in using a combination of these traditional methods, along with new and emerging methods that are designed to accommodate such expansive data growth. The following techniques are essential to maintaining data availability and access in the era of big data:

- A consistent data capture strategy should be developed for cloud and enterprise computing environments to address data synchronization concerns and facilitate cross-platform workload processing.
- Detailed storage hierarchies and tiering should be used to enable data to be automatically assigned to the most appropriate storage system, based upon its value to the business at a particular point in time and on the needs of applications like replication, archiving and resilience. Establishing a hierarchy and tiering system enables organizations to effectively identify and manage the number of data copies required to maintain resilience.

- Data types should be defined, and data should be categorized based on business value, then mapped back to the storage hierarchy for inclusion in the organization's data capture and replication strategies. Subsequent data protection designs will need to be put in place to make certain that data accuracy, availability and integrity are in alignment with the structured, hierarchical approach for business protection.
- The methods employed by users to manage and protect data—backup, snapshot and replication—should be evaluated to be sure they are easily accessible and aligned with user expectations. This will require that an integrated approach to server- and storage-based replication be identified and validated based upon stated recovery objectives. Such an approach helps the organization verify that overall system, application and data integrity are in place to deliver the desired user experience.

The dizzying pace of data creation requires new strategies and tools to assure accuracy, availability and integrity of information

Enabling systems of engagement

As the demand for continuous access to systems and information rises, so do user expectations for self-service capability. Users increasingly expect to access resources and perform other routine functions on their own, with limited external interaction or support. They expect to be able to perform transactions quickly, conveniently and easily. The online experience has become a high priority for users, and it is driving the requirement for systems of engagement.

Systems of engagement are customer-facing systems like email, social, mobile and collaboration systems. They engage users with intuitive design, incorporating technologies like cloud to encourage B2B and B2C interaction. In short, they reflect how today's users want to work.

IT is steadily evolving to systems of engagement from more traditional systems of record, which for the most part provide information to internal employees passively via back-office processes like HR, accounting and payroll. These core business systems typically live behind the corporate firewall, making it easier to verify the accuracy and integrity of data and processes that support business transactions. As they give way to systems of engagement with infinitely more user touchpoints and a virtual avalanche of data, the potential for risk increases. This shift in how, when and by whom today's systems are being accessed is dramatically changing the way resiliency strategies should be designed and implemented.

Techniques for building resiliency with systems of engagement

As the user experience changes and becomes more interactive, any type of IT disruption becomes far more apparent. Organizations must have the ability to resume business operations immediately following an outage. But providing continuous service in this environment is more challenging. Today's resiliency measures must be scalable and adaptable, with the ability to speed recovery as the demand for alway-on, always-available service escalates.

59%

of firms with an in-house disaster recovery solution were only somewhat successful in meeting recovery objectives during testing[2]

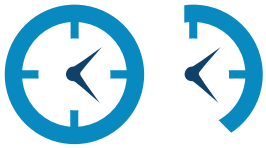
Most organizations will need to increase current service levels for availability, and this is likely to require technological or process changes to day-to-day operations. Resiliency strategies must address these changes to provide continuous availability and consistent interaction. They need to specify the technologies and services required to protect not only back-office systems of record, but also the increasing number of front-end systems of engagement. Moreover, systems of interaction will need to be developed to see to it that front-end applications are integrated into the back-office systems and processes.

For today's mobile and social users, their first contact with an organization is often through a system of engagement.

Regulatory requirements are increasing pressure on organizations to deliver seamless processing functionality for mobile and social business systems. This is driving organizations to focus new attention on the audit and compliance aspects of resiliency. While access to traditional systems is regulated and contained, the openness and spontaneity associated with systems of engagement pose new content management and access challenges for IT. Longstanding practices for control and governance must be adapted for these new systems.

New resiliency programs should be built with the capability to validate their own integrity, to prove that their response to a wide-scale event meets the needs of users and, ultimately, the needs of the business. Resiliency testing should re-create the actual user experience to accurately assess the value of the program, especially for systems of engagement. This moves away from some of the more traditional, scripted approaches to system restoration and data validation, and also from the application verification metrics that have been in place to validate the IT function, but not necessarily the user experience.

Systems of record will continue to leverage traditional methods for system, application, data and database resiliency to resume complete business functionality in the event of a disruption. Speed of recovery (RTO) with minimal data loss exposure (RPO) will remain key measures of system recovery.



Less than 1.3 hours
is the recovery time objective of best-in-class companies[3]

Systems of engagement require extended resiliency capabilities to make sure that users have continuous access to all business applications (and their associated data) with minimal latency. Fully functional communications—that is, complete access with equivalent performance and response time—will be required in order for end-to-end transactions to be validated. Organizations need to be able to mimic the user experience by creating scenarios that simulate the user's online business environment. For example, in order for application designers to validate the user's perspective of access, performance and response time, organizations should consider leveraging scripts to conduct user acceptance testing in the development environment. This would help re-create more accurate, end-to-end transaction scenarios. The objective is to verify that availability, performance and responsiveness have met agreed-upon service levels in accordance with user expectations.

Systems of interaction (see Figure 1) will require more detailed design for integrating the back-office systems with the front-end applications to make sure that when workloads are re-established following an outage, all components are communicating in a timely manner. Increased focus will be on data integrity (in the form of RPO) and cross-platform application synchronization to help assure that the total environment resumes operations with optimum accuracy, efficiency and readiness.

Conclusion

The growing ubiquity of cloud infrastructures, big data and systems of engagement is driving significant change in how enterprise computing environments are defined and ultimately delivered. As enterprise computing continues to evolve, organizations will need to reassess how they approach resiliency, from system, storage and network design to recovery processing.

The role and relative importance of resiliency will continue to parallel that of cloud. As the evolution of cloud continues, resiliency will take on a greater role within the enterprise. Likewise, as IT delivery models change and become more complex, the amount of risk associated with delivering uninterruptible service will increase accordingly. This more agile, versatile environment will be further complicated by the increasing amounts of new data and the desire for continuous access to it.

Keeping pace with these new, more fluid designs for IT delivery requires resiliency measures that are capable of addressing the demand for more complete systems and data protection. Accompanied by more stringent recovery expectations, resiliency will need to be delivered in the form of increased service levels for availability, performance and responsiveness. The methods for verifying and validating resiliency capabilities will also need to be modified to replicate more of the user experience, capturing core business processes and data being accessed by the very systems of engagement that deliver business results. The best resiliency measures will be developed based on this experience.

For more information

To learn how IBM is helping organizations deliver resilience in the era of enterprise cloud computing, please contact your IBM representative or IBM Business Partner, or visit

ibm.com/services/continuity

IBM Global Services
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
July 2014

[1] IDC, *"Findings from IDC's 2013 Global CloudTrack Survey: An Evolving Set of Cloud Drivers,"* IDC #244624, November 2013.

[2] "The Risks of 'Do It Yourself Disaster Recovery," a commissioned study conducted by Forrester Consulting on behalf of IBM, January 2103.

[3] "Virtualization: Gateway to Business Continuity,"Aberdeen Group, April 2013.

RLW03030-USEN-00