rubrik

A CHECKLIST OF WHAT YOU NEED TO LOOK FOR IN A MODERN DATABASE BACKUP AND RECOVERY

Enterprises are quickly onboarding next-genera-tion, mission-critical applications, such as e-com-merce, artificial intelligence, machine learning, and internet of things on non-relational databases and distributed local storage. Organizations are moving these applications to scalable databases because they are easier to deploy, always avail-able, and cost effective. However, a robust and scalable backup and recovery solution is needed to protect these modern databases. The challenge is that although enterprise adoption of next-gen-eration applications has been rising, the risk of data loss is prevalent. This is because non-rela-tional databases lack a viable data protection and recovery solution. What this means:

- **Currently, there is no solution that allows for corrupted data to be removed,** replayed, and propagated with minimal downtime to customer-facing applications.

- **Reputable studies have concluded that as much as 75% of data disasters** ultimately leading to downtime is the result of some sort of human error. Because human errors have a big impact on business operations and occur randomly, they cannot be solved by solutions that are offered today, such as native database replication or node-level snapshots.

- **Database replication does not provide point-in-time backup and any point-in-time recovery** so enterprises cannot go back and fix these errors. In fact, if errors are introduced, the databases' redundant-node replication can lead to almost immediate corruption across all nodes of the database.

Here's what to keep in mind as you begin to explore backup and recovery solution for your next-generation applications.

1. Storage Is No Longer Unified or Shared Storage

A key architectural transition is occurring where monolithic pools of shared storage are no longer used for next-gen applications and databases. Next-generation databases are scale-out in nature and use commodity storage at a node level, creating a highly distributed storage pool. This distributed storage pool provides both low cost and high performance, but is not conducive to using traditional database backup and recovery tools, which rely on database durability logs (from a shared storage LUN) for a point-in-time backup.

2. Eventual Consistency Has a Big Impact on Backup and Recovery

Next-generation databases are not built around the traditional ACID (Atomicity, Consistency, Isolation, Durability) transaction model. Transac-tions are a mechanism for guaranteeing these properties; they are a way of grouping related actions together so that as a whole, a group of operations can be atomic, produce consistent results, be isolated from other opera-tions, and be durably recorded. Eventual consistency, however, brings forth new challenges to generate a consistent point-in-time view of the database for operational recovery or test/dev needs. Therefore, novel and scale-out data protection technology to address eventual consistency is needed.

3. Deduplication Matters at Scale

With traditional databases, unified or shared database storage with block level snapshots made image-based recovery and rollback relatively efficient. Distributed storage combined with a scale-out database that commonly involves replication of data rather than joins across tables lead to a rapid proliferation of redundant data and increasing storage requirements (3-4 copies of every write operation). Effective data reduction at scale is critical to success and economics for effective data protection in this environment.

4. Single Points of Failure Do Fail

Legacy data protection approaches have commonly involved a media serv-er, which presents a choke point for a secondary data path and single point of failure. Given the very high ingest rates, coupled with the real-time nature of applications deployed on modern scale-out databases, next-generation data protection infrastructure has to be distributed to handle performance (RPO and RTO windows) and resiliency in always-on environments.

rubrik

CHECKLIST FOR MODERN DATABASE DATA PROTECTION:

☑ How do you protect your next-generation scale-out databases today?

☑ What scenarios of data loss are you worried about?

☑ What downtime can your applications or business handle in the event of data corruption?

☑ Have you validated backup and recovery for your next-generation databases?

☑ Can you prove compliance with regulatory or line-of-business requirements for backup of your critical next-generation infrastructure?

☑ Can you recover to a desired or any point in time quickly? What manual effort is required to do this?

☑ How much storage is required to store your desired pool of recovery versions?

☑ Does your data protection approach scale out with your application and data growth?

☑ Can you efficiently refresh your test/dev environment to a smaller or larger cluster size as needed?

☑ What happens to your backup solution when a database node fails? Can you still perform backups? How about intermittent network failures?

☑ Native formats free you from vendor lock-in and allow for data management services, such as search and analytics on secondary data. Can you recover your data in native format?

If you are you considering a scale-out data protection solution, let us know how we can help!

rubrik