Securing the Future of Government Data in the Cloud: a case study

In collaboration with AWS and Govloop, this special cast study discusses how the City of Oakland in California shifted to a comprehensive backup and recovery tool to guard against data loss.

CITY OF OAKLAND

BEST PRACTICES FOR DATA BACKUP IN THE CLOUD:

1. Don't assume that if you have resources in the cloud, they are fully backed up.
2. Make sure your cloud-based backup solution prioritizes application consistency.
3. Automation is critical when backing up cloud workloads.
4. Choose a backup or recovery solution that positions you for the future.

WATCH THE VIDEO

Executive summary

Agencies continue to move workloads to the cloud, citing increased operational efficiency and agility, reduced costs and improved citizen services. While the cloud provides many benefits, it also delivers new challenges. For example, over time, most agencies have accumulated many different, often disconnected, cloud repositories. Accumulation makes it difficult to understand where all data is located and to recover quickly in the event of an emergency.

To learn more about how agencies can provide fast backup and recovery of data while maintaining and increasing manageability, GovLoop teamed with N2W, a cloud-native backup, disaster recovery and archiving solution.

The challenge: data sprawl in the cloud

Backing up your cloud environment is an important first step to making sure that your data and infrastructure is always available, but it doesn't solve everything. Because agencies tend to deploy workloads across dozens of cloud platforms, for example, it can be difficult to quickly find the specific server or files required. Instead, administrators can spend hours examining the backups of each cloud platform to find what they need.

Speed of recovery is another challenge. With backups distributed across multiple clouds, it can be difficult to recover quickly in the event of an outage – what's known as Recovery Time Objective (RTO). Closely tied to RTO is Recovery Point Objective (RPO) — the time between backups. **"If you back up once a day and have a failure after 23 hours, you have lost 23 hours' worth of data. Not too many agencies have that kind of tolerance for data loss,"** said Sebastian Straub, a principal solutions architect with N2W, which specializes in data protection for cloud-based workloads.

With so many workloads spread across so many clouds, it also can be difficult to control who has access to your backups. And then there are concerns about data sovereignty — where the data actually lives — and whether vendors or other parties might be able to access that data.

These security and compliance concerns constitute very real roadblocks for agencies storing workloads in the cloud. It's also important to have confidence that backups are occurring consistently, and that the right people get notified when failures or other issues occur. For example, do you have the ability to carefully audit your environment to determine whether there are workloads that are not being backed up but should be?

The solution: automated, policy-driven backup and recovery

To protect against data loss, security breaches and slow recovery times while providing real-time data access, agencies need an automated, policy-driven, comprehensive backup and recovery plan that includes all data stores and infrastructure. To provide fast availability and recovery time, choose a solution that prioritizes speed and automation. For example, the N2W virtual appliance uses fast snapshot-based block storage, allowing it to recover an entire environment — all data, network configurations and servers — within about 30 seconds. That's particularly important when an entire region or data center fails.

If a disaster occurs, you should be able to recover everything at the same time instead of machine by machine. If a region fails, your solution also should support cross-region recovery. N2W, for example, supports disaster recovery between AWS GovCloud (US) regions, so if one fails, everything is immediately available on the other site.

Providing full security and compliance is important with all backup scenarios. One way to do this in cloud environments is by using a solution that does not actually see or access any data it is backing up. Instead of filtering the data through a solution, for example, look for a solution that simply applies the instructions set by the organization to perform specific data manipulations. No third party, including the vendor, should have access or visibility into anything.

A comprehensive backup solution also should be able to automate how data is moved throughout tiers to save money. For example, data backed up to expensive Amazon Elastic Block Store (Amazon EBS) should be able to be moved to less expensive Amazon Simple Storage Service (Amazon S3) or even Amazon S3 Glacier Deep Archive, depending on its importance, relevance and how quickly it needs to be recovered.

> **"There was a level of granularity in the backup scheduling that exceeds our ability internally, and we could do both incremental and full backups, which we couldn't easily do before."**
>
> Julian Ware, Spatial Data Administrator

Recovery Time, Peace of Mind Prove Critical for City of Oakland

Known for its trendy neighborhoods, street festivals and major sports teams, Oakland, California, is as busy a city as you'll find. To keep up with maintenance, safety and citizen services, the city's IT staff began running a major geographic information system (GIS) in an AWS Cloud in 2012.

As the mapping system and associated data grew, the IT department realized it needed more comprehensive backup and recovery to guard against potential data loss. Access to mapping data is critical for many city services, including police and fire. "AWS allows you to do a lot of things, but we were just too busy. We didn't have time to script out backup routines and make sure that everything was being consistently snapshotted and backed up," said Julian Ware, a spatial data analyst for the city of Oakland.

After evaluating the options, the team settled on N2W, a backup and recovery tool designed for AWS workloads. When the team was a few days into the two-week pilot, Ware said he knew he had found the answer. "There was a level of granularity in the backup scheduling that exceeds our ability internally, and we could do both incremental and full backups, which we couldn't easily do before," he said. Ware also noticed that recovery time improved dramatically, and that performing backups didn't significantly impact IT resources.

More than anything, Ware appreciates the peace of mind. "Being able to look at my email in the morning and know that backups ran at the time I scheduled them for and that there were no problems means I don't have to worry about it for the rest of the day," he said.

Conclusion

As agencies move more workloads to the cloud, they must find ways to make sure that their data and network configurations are fully backed up and can be quickly recovered. At the same time, they must make certain that security and compliance are not compromised, and that budgets stay under control.

An automated backup and recovery solution built specifically for the cloud will help agencies meet all of these goals. A "born in the cloud" approach to backup and recovery is better equipped to deal with workloads in multiple clouds, protect against data loss and security breaches, ensure compliance and enable realtime data access. As mission requirements evolve, backup and recovery processes must change to keep pace. Using an automated, modern approach to backup, recovery and disaster recovery is the key to accomplishing these goals.