Alaska Department of Fish and Game Solves Its Alaska-Sized Backup Challenges with EMC Data Domain

Alaska is more than twice the size of Texas with a population less than that of Hawaii, which leaves it with little room for inefficiency in any of its operations. So when Corey Kos, Infrastructure Manager for the Alaska Department of Fish & Game recognized his agency’s huge backup and disaster recovery challenges, he turned to EMC to help transform their backup and recovery processes into tasks so simple that any size state could manage them. Utilizing EMC Data Domain deduplication storage systems has eliminated the agency’s use of tape and further its VMware server virtualization initiatives by centrally protecting and consolidating backup management.

November 2010
By Jerome Wendt, DCIG Lead Analyst

Company
Alaska Department of Fish & Game (ADF&G)
P.O Box 115526
1255 W. 8th Street
Juneau, AK 99811-5526
907.465.4100
www.adfg.state.ak.us

Industry
Government

Challenges
• Decentralized backup management put integrity of backup data in question
• Disparate tape technologies make creation of reliable DR plan impossible
• Decision to virtualize servers not feasible with legacy backup infrastructure
• Low bandwidth network limited the ability to architect a centralized solution

Solution
• Four EMC Data Domain deduplication storage systems
• EMC Data Domain Replicator software

Benefits
• Lowered operational costs through centralized procurement and management of backup and DR environment
• Reduced administration through consolidated management of all backup data at four primary ADF&G data centers
• State-wide backup and DR plans enable sites across the state to meet backup windows and recovery service level agreements
• Flexibility to dynamically reconfigure Data Domain appliances to meet changing demands associated with virtualization and replication from remote offices
• Reliable replication over state’s finicky WAN connections without additional bandwidth

Overview
Alaska Department of Fish & Game’s Infrastructure Manager, Corey Kos, recognized that his agency’s use of a hodge-podge mix of tape products for data protection at multiple sites would not meet the demands of the virtualized environment that was planned and that a move to a centralized disk-based backup solution was in order. Deploying EMC Data Domain systems enabled ADF&G to continue doing backups locally using their existing backup software. Further, Data Domain appliances could efficiently replicate data using ADF&G’s existing network infrastructure across its remote locations and then centrally store deduplicated data at the agency’s primary data centers.

Alaska Department of Fish & Game (ADF&G) Profile
ADF&G’s goals are to optimize the economic benefits that Alaska derives from its fish and wildlife resources; optimize public participation in Alaskan fish and wildlife pursuits; and, increase the public’s knowledge that fish and wildlife in Alaska are responsibly managed. ADF&G currently has about 1,500 employees with primary data centers in Anchorage and Juneau and two (2) remote data centers in Fairbanks and Kodiak.

ADF&G Data Protection Strategy Needed an Overhaul
Corey Kos, ADF&G’s Infrastructure Manager, recognized almost from the moment he took his post in 2007 that ADF&G’s data protection strategy needed an overhaul.

The hodge-podge mix of tape drives and formats as well as different backup software products in use at the individual sites were only the tip of the iceberg of the problems that he encountered. Other challenges that ADF&G was facing included:

• Inefficient data protection supply procurement processes.
  Obtaining backup software, tape cartridges and drives, etc. was an ad hoc process and done by each office such that Kos had no sense of who had what or how much money was being collectively spent on backup across the organization.
Management of backups was highly distributed. IT staff in each of the central and remote data centers performed their own backups without any centralized, collaborative effort. As a result, backups were not done consistently and there was no formal mechanism in place to check the integrity of these backups or that data could be recovered.

Backup infrastructure would not support ADF&G’s planned virtualization initiative. While remote offices would continue to use physical servers, the servers across the four (4) data centers were to be virtualized on VMware with 15 ESX Servers running 200-300 virtual machines (Windows 2003/2008 & Linux). In this virtualized environment, the heterogeneous tape-based data protection approach would quickly become a bottleneck.

ADF&G’s inability to effectively manage its existing backup environment coupled with the pending introduction of a new virtualized environment led ADF&G to mandate the introduction of a common, scalable state-wide platform and consolidated management for backup.

New Backup Solution Criteria
Part of the criteria that the new solution for ADF&G needed to satisfy included centralizing all backups, implementing disk, and supporting this highly virtual environment that ADF&G was creating.

Unfortunately in this situation Kos could not really compare the cost of a new backup solution to the cost of ADF&G’s old tape system since the backup infrastructure was in such disarray. However as a government agency, the cost of new solutions and equipment was still an obstacle so the following criteria were established in order to select the best solution for ADF&G. It had to:

• Allow for centralized management by the new ADF&G infrastructure services team
• Easily drop into any ADF&G office and work with existing backup software
• Keep costs under control by reducing the amount of storage required
• Lessen and ideally eliminate the administrative burdens of backup at each office
• Meet the individual data protection and retention needs for each office
• Scale to meet future data growth

Based upon this set of requirements, it soon became clear to Kos that EMC Data Domain systems were the best overall solution for ADF&G. It worked seamlessly with any backup software so ADF&G could continue to use its existing Backup Exec, Oracle RMAN and Windows Server backup software while laying the framework for ADF&G to backup its emerging virtual environment using a disk-based target.

Using EMC Data Domain ADF&G’s backup and recovery related tasks have become so small that we can support them with minimal effort even though we need to support a geographic area the size of Alaska.”

―Corey Kos, Infrastructure Manager
Alaska Department of Fish & Game

The EMC Data Domain appliances replicate data so efficiently that even over our worst WAN links we can get all of our data from our two most remote data centers by noon of the next day.”

―Corey Kos, Infrastructure Manager
Alaska Department of Fish & Game

Hybrid Approach Used for Remote Office Backup
Kos determined a hybrid approach would best fit the backup and retention requirements across remote offices and data centers. To protect his dozens of remote offices (some of which are only operational during the summer months), he opted to deploy Windows file servers at these sites. Since they only stored minimal amounts of data for most of the year, he did not really need to back them up.
All data in these offices was then stored to this local Windows file server and, using the replication feature within Windows DFS (Distributed File System), replicated to a Windows file server in one of ADF&G’s data centers.

Now that the data from his remote offices was on Windows file servers in one of his four primary data centers, he then backed up the Windows file servers to the EMC Data Domain appliance located at that site.

At the smaller Fairbanks and Kodiak data centers, Kos deployed Data Domain DD510 appliances that each hold up to 5.7 TBs of disk capacity prior to deduplication benefit. At the larger Anchorage and Juneau data centers, he deployed higher capacity Data Domain DD565 systems that each hold up to 16.2 TBs of physical disk capacity. Using this hybrid approach, he was now able to protect the data of all of ADF&G’s offices on disk and still stay within budget.

**Efficiency of EMC Data Domain Replication Seals the Deal**

However Kos had not yet solved all of his problems. While his remote offices had two copies of backup data (one at their site and one to the primary site to which they were replicating), his four data centers still needed to replicate data offsite in the event of a disaster.

This piece of the plan initially concerned Kos. WAN connectivity in the State of Alaska is dicey at best and while he saw really good throughput between the Juneau and Anchorage sites, the WAN connections in from Fairbanks and Kodiak were limited in bandwidth.

These limitations in bandwidth were further aggravated by how the State of Alaska manages its WAN links. Alaska has an enterprise technology group that acts as a centralized IT team and which runs its state WAN. While these shared WAN links help save Alaska money, it provides ADF&G almost no visibility into how saturated these links are. This results in unpredictable performance when doing replication.

It was this lack of visibility and limited bandwidth that in part drove Kos to select EMC Data Domain systems. He says, “The lack of visibility into our WAN connections between our four main data centers made it difficult to architect this part of our solution. But the EMC Data Domain appliances replicate data so efficiently that even over our worst WAN links we can get all of our data from our two most remote data centers by noon of the next day.”
ADF&G Solves Alaska Sized Backup Challenges with EMC Data Domain

Two years have passed since ADF&G first implemented disk-based backup and Kos could not be more pleased with how well the EMC Data Domain systems have performed over that period of time. He has now completely eliminated tape in all of his remote offices and data centers and can recover data in any of those offices. The data in all of those offices is automatically and daily replicated from his remote offices and data centers back to one of his two primary data centers using a combination of Windows DFS and Data Domain.

The introduction of Data Domain deduplication storage also contributed to increasing the number of virtual machines he could host on each server. While he initially was using Backup Exec on his virtual machines (VMs), using Backup Exec was creating too much overhead on the individual VMs.

So once the Data Domain solution was fully deployed he moved to a backup software product (PHD Virtual) that was optimized for backing up VMs to disk. Kos says, “Using this new backup approach, we were able to virtualize 90% of our servers using our existing VMware ESX servers.”

Having so many other tasks on his plate as ADF&G’s Infrastructure Manager, he is relieved that data protection and disaster recovery are tasks that he no longer has to think about or manage. Kos adds, “Using EMC Data Domain ADF&G’s backup and recovery related tasks have become so small that we can support them with minimal effort even though we need to support a geographic area the size of Alaska.”

About EMC Corporation
EMC Corporation (NYSE: EMC) is the world’s leading developer and provider of information infrastructure technology and solutions that enable organizations of all sizes to transform the way they compete and create value from their information. Information about EMC’s products and services can be found at www.EMC.com.

About DCIG
DCIG analyzes software, hardware and services companies within the storage and ESI industries. DCIG distributes industry, company and product analysis by way of viral marketing and community building using the burgeoning BLOG infrastructures created worldwide.