Komprise and Spectra Logic: Archiving Data with Insight

August 2017

800-833-1132 • 303-449-6400 • (Fax) 303-939-8844 • 6285 Lookout Road • Boulder, CO 80301 USA SPECTRALOGIC.COM

Contents

Abstract **2**

The Data Growth Challenge **2**

How to Squeeze Massive Data Growth within Flat Budgets 2

How to Manage Data Based on ROI, and not Based on Where Users Put It 2

Legacy Solutions too Costly, Complex and Disruptive 2

Intelligent Data Archiving with Komprise and Spectra Logic **3**

Overview 3

Manage with Insight 3

Simple to Operate Service that Scales on-Demand 3

Up and Running in Under 15 minutes 3

Data Analysis and Interactive ROI Visualization 4

Move Data with No Changes to User and Application Access 5

Scale-Out On-Demand to Handle Data Growth 6

The Komprise Difference: Raises the Bar on Simplicity and Efficiency 6

Spectra Logic Deep Storage **7**

Spectra Archive Storage Options 7

Options for Archive Storage – Disk, Tape, Cloud 7

BlackPearl with Komprise 9

Overall Costs and ROI Examples 11

Summary **12**

Abstract

Enterprise organizations often seek ways to easily identify and move inactive data off of their primary storage and onto more cost-effective storage. Komprise data management software combined with Spectra Logic archive storage solutions provide easy, reliable, and cost effective data management by continuously and automatically moving inactive data off of primary storage and onto archive storage without changing user or application access paths.

The Data Growth Challenge

How to Squeeze Massive Data Growth within Flat Budgets

Data is growing fast – nearly 90 percent of the world's data was created in the last two years, and enterprise data is doubling every two years. The challenge is how to retain all this data (most of which is valuable) while keeping within flat budgets. As businesses look to do more with less budget, they are turning to affordable, scalable storage options such as cloud and object storage.

How to Manage Data Based on ROI, and not Based on Where Users Put It

Over 60 to 90 percent of data is infrequently accessed within months of creation, and yet is often stored and managed in the same way as active data. This is due to a lack of easy approaches to identifying and moving inactive data to more affordable storage without disrupting users.

Legacy Solutions too Costly, Complex and Disruptive

Legacy data archival management and storage solutions are costly, complex, and disruptive. Legacy solutions are:

- Costly – requiring expensive enterprise licenses and storage, not allowing users to migrate to more affordable archive disk, tape and cloud storage
- Complex – with multiple moving parts such as storage agents, hardware, software and databases to manage
- Disruptive – creating performance slowdowns due to the management overhead they generate, and user disruption by not maintaining transparent access to moved data
- Missing ROI Analysis – no way to identify the savings potential or Return on Investment (ROI) with the archive solution
- Brittle – with static stubs that can be corrupted or orphaned, agents that need to be kept up-to-date as the storage evolves and detailed rules that need to be specified and managed
- Lacking Data Protection – no way to easily archive data to more than one storage type for "genetic diversity" data protection

Data management has been hard to adopt and scale to meet today's massive data growth. Komprise and Spectra Logic address this need with a modern archiving architecture that has been built from the ground up to handle today's massive amounts of data with intelligent automation.

Intelligent Data Archiving with Komprise and Spectra Logic

Overview

With the Komprise/Spectra solution, Komprise software is used to identify inactive data that resides on primary storage. Komprise then automatically and continuously moves that data to Spectra archive storage.

Figure 1. Basic Solution Flow Diagram

Manage with Insight

Komprise enables businesses to manage their data intelligently by identifying inactive data across a customer's storage environment and transparently moving infrequently accessed data to Spectra archive storage – without any changes to user or application access. This cuts over 75 percent of costs and solves the business challenges of managing data growth within flat budgets without any changes to user or application access.

Simple to Operate Service that Scales on-Demand

Komprise uses an easy-to-deploy, easy-to-scale distributed architecture that consists of one or more Komprise Observer virtual machines running at the customer site connected to a Komprise Director that runs either as a cloud service or on-premise. Komprise works seamlessly across any on-premise NFS, SMB/CIFS, and Lustre file system storage.

Up and Running in Under 15 minutes

Deploying Komprise is simple – simply download the Komprise Observer and point the Observer at existing primary storage via NFS, SMB/CIFS or Lustre. Within 15 minutes, Komprise provides analytics on how much data you have across your storage, how it's being used, who is using it, and how fast it is growing. Even on petabytes of data, Komprise provides nearly instantaneous results using a patent-pending dynamic sampling technique.

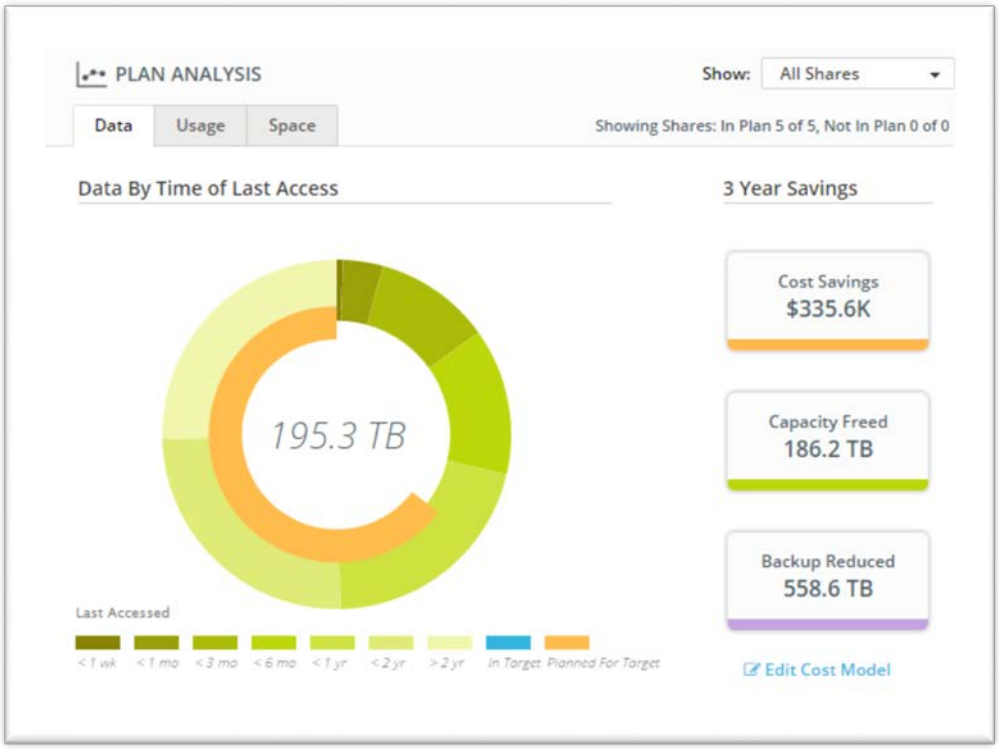

Figure 2 Komprise Analytics of Data Usage

Data Analysis and Interactive ROI Visualization

Users set various data management objectives and play with "what-if" scenarios to understand their projected impact on footprint and costs. Komprise interactively projects 1) how the data footprint at the source will change based on the chosen objectives, 2) the cost impact, and 3) the projected Return on Investment (ROI). This provides a no-risk way to plan capacity and the most-effective data management approach before actually moving any data.

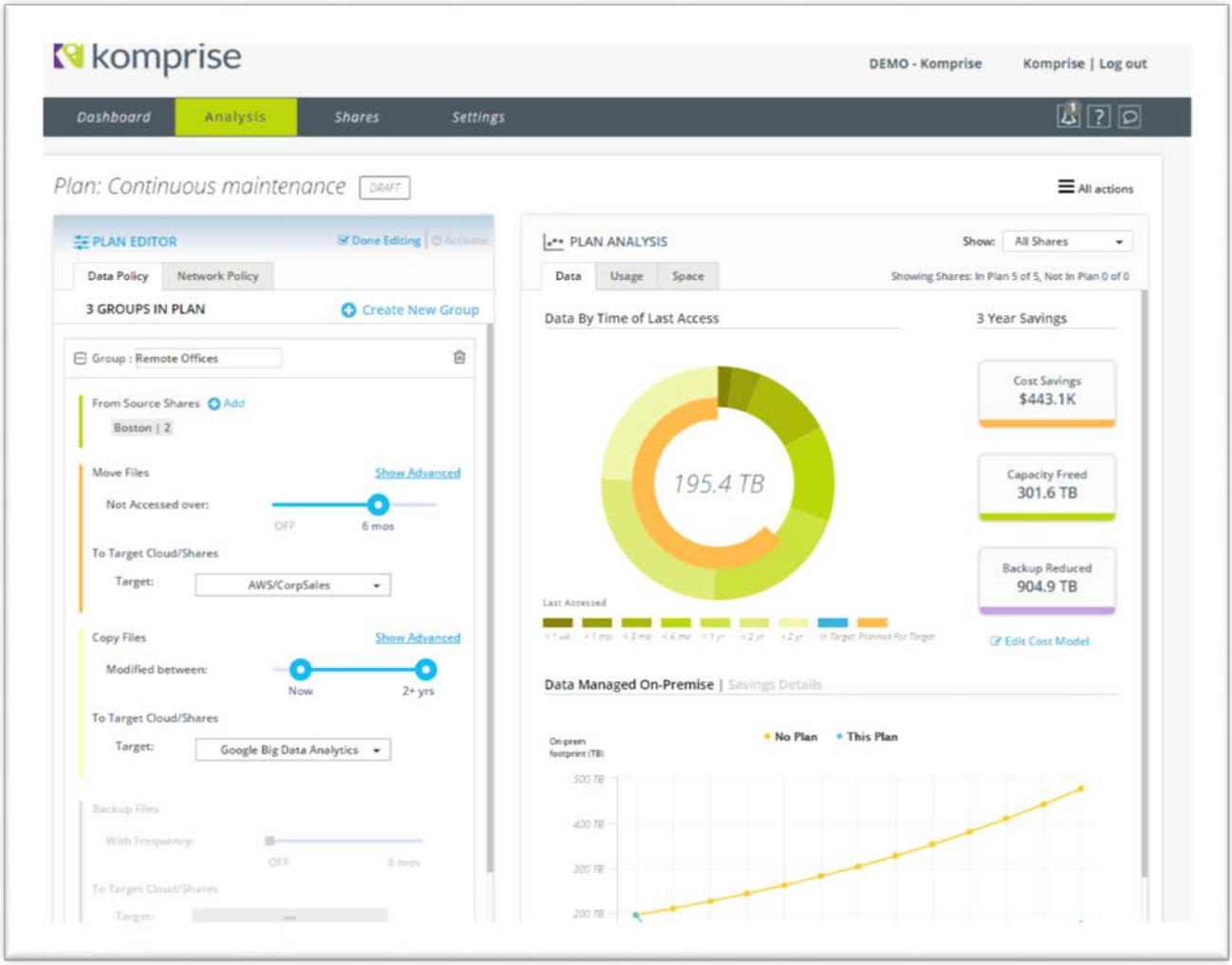

Figure 3 Interactive "What-if" analysis of Plan Objectives

Move Data with No Changes to User and Application Access

Once satisfied, simply activate the plan and Komprise moves the data transparently to Spectra Logic archive storage – without using any proprietary agents or static links on the storage systems (mechanisms that have traditionally caused problems). There is no change to manage for users and applications – they continue to see and access the data as they did before.

Scale-Out On-Demand to Handle Data Growth

Komprise is designed to deploy in minutes and yet scale out to handle massive data growth. The architecture is a distributed fault-tolerant scale-out design. As the load on the Komprise Observer grows, additional Observers can be added on-demand and Komprise manages the observers as a single fault-tolerant grid – making management simple while delivering a robust, efficient and scalable solution.

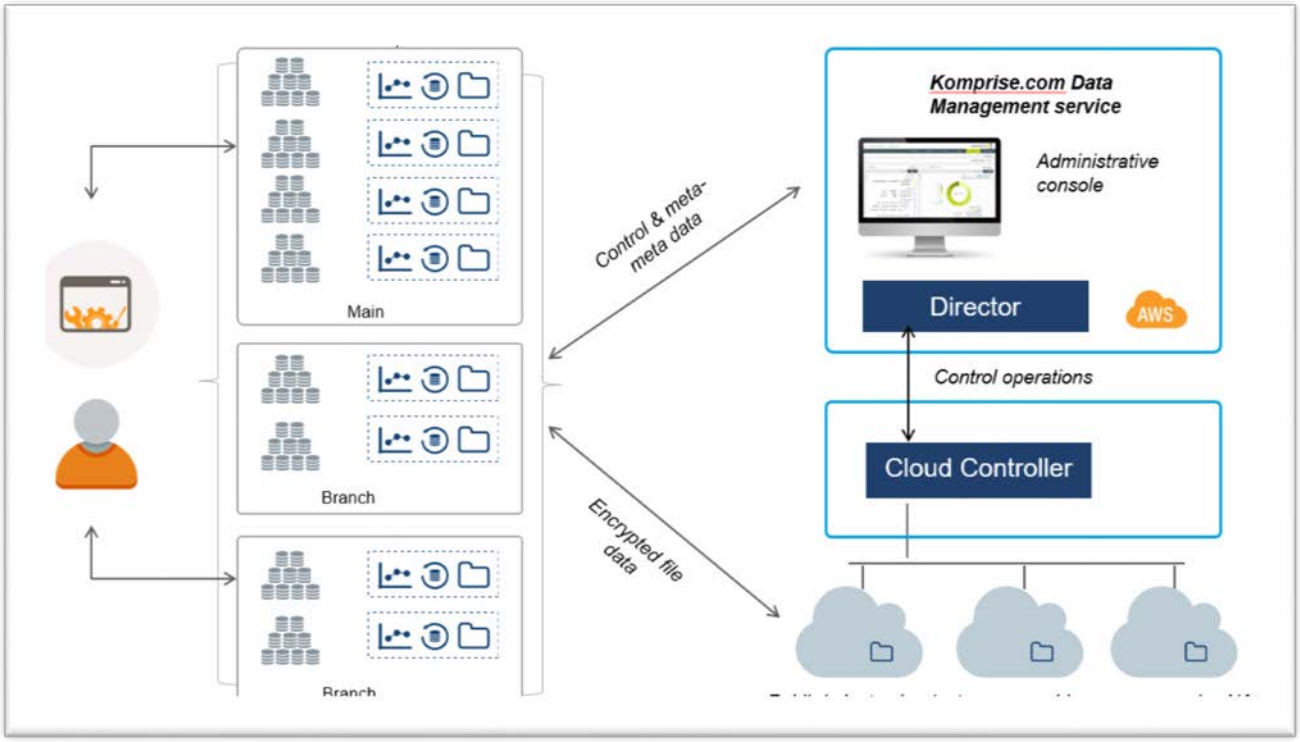

Figure 4 Komprise Architecture

The Komprise Difference: Raises the Bar on Simplicity and Efficiency

Komprise raises the bar on simplicity and efficiency by eliminating the cost, complexity, limited scalability and disruption of prior approaches

- **No hardware to deploy**: Komprise does not require dedicated hardware or upfront infrastructure investments. Deploying Komprise is a 15-minute setup of the Observer virtual machine.
- **No storage agents:** Komprise does not require any agents on any of the storage systems. Komprise works seamlessly across NFS, SMB/CIFS, and Lustre storage using standard protocols without requiring any proprietary interfaces.
- **No stubs:** Legacy data management solutions use stub files – static pointers to redirect users to the new location of files. But, these static stubs have lots of issues. They can be accidentally deleted, leaving the moved files orphaned. They can be corrupted –which also leaves the moved files orphaned. Or, they can become invalid if the moved file gets moved again. Komprise uses a patent-pending dynamic link technique that does not have any of these issues. Users and applications continue to access moved files as they did before without any changes.

- **No impact on hot data path:** Komprise sits outside the active data path so access to hot data remains unchanged. In fact, performance of hot data may improve since Komprise reduces clutter on the primary storage.
- **No scaling limits:** Komprise and Spectra Logic have a scalable architecture that seamlessly scales on demand. As capacity grows, simply add more Observer virtual machines and more Spectra archive storage.
- **No silos:** Komprise delivers a unified view of your data across all your storage and enables you to move data across storage based on your requirements.
- **No costly licenses:** With sub-cloud pricing, Komprise and Spectra Logic are designed to instantly cut 70 percent or more of costs on every terabyte managed.
- **No complex configurations:** Komprise and Spectra Logic products are simple to deploy and operate. The combined solution requires no proprietary interfaces or complex infrastructure setup.
- **No disruption:** Komprise runs seamlessly in the background and does not interfere with file system activity. Komprise is designed to adaptively throttle itself back when file systems are in active use, so it runs invisibly without any noticeable impact on performance.
- **No changes to user or application access:** Komprise moves data transparently to Spectra Logic storage without any changes to user or application access. Moved files are just as accessible as before from the primary storage without any static stubs.

Spectra Logic Deep Storage

Spectra Archive Storage Options

For nearly 40 years Spectra has focused on innovation in storage systems and solutions. The leaders in data-intensive industries, government entities, and researchers rely on Spectra solutions that are optimized to support their specific workflows. Spectra has a rich ecosystem of leading solution providers that, combined with Spectra® BlackPearl® with Storage System, provide a single interface to the industry's leading tape and disk storage using cloud protocols.

Options for Archive Storage – Disk, Tape, Cloud

Spectra Logic is the only vendor to provide archive options for all of the different modern secondary storage types:

- Affordable disk storage on archive drives for fast access to archived data.
- Digital tape, the most affordable and reliable storage type, when fast retrieval times are not critical. Spectra Logic is the only vendor to support LTO, IBM® TS11xx, and Oracle® T10000 tape technology in a single tape library unit.
- Archive to public cloud and public "cold cloud" for offsite storage and disaster recovery copies.
- Spectra Verde® NAS Solution with Komprise
- Spectra Logic's Verde network-attached disk archive storage is the perfect complement to Komprise's data management software. Verde is affordable NAS storage based on archive disk technology that delivers high density bulk storage. With the ability to store up to 7.1PB in a single rack, Verde delivers an amazing 180TB per rack unit density. Key features of Verde with Archive Drives are:

- **Digital Preservation** - Multiple layers of protection ensure data is protected, from replication to triple parity to a unique check summing strategy.
- **Simplicity** – Verde's simple user interface makes setup and management easy. And with easy expansion, Verde grows with your needs.
- **Broad Application Support** – Verde is certified with the industry leading backup and archive software packages – most importantly Komprise – to extend your storage infrastructure and capacity at an affordable price.
- **Affordable NAS** - Verde offers you the lowest cost enterprise disk hardware at 7.5¢ per gigabyte without sacrificing reliability and data protection features.
- **Tier to Tape** – Easily make a copy of the data to the most reliable and lowest cost storage type, including Spectra tape libraries, for true "genetic diversity" in data protection.
- **Supporting You is Our Number One Priority** - SpectraGuard® support ensures that you are completely satisfied with your Verde NAS solution, delivering several support options.

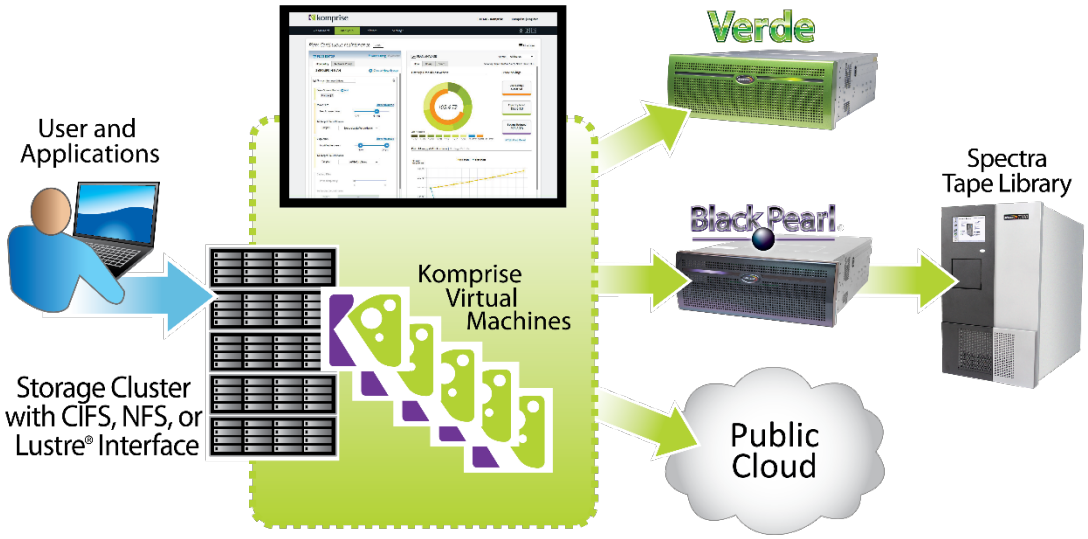

Figure 5. Archiving to Spectra Verde via Komprise

Data archived to Verde via Komprise can be recalled immediately and transparently by the user via the original primary storage path. Komprise's dynamic link technology couple with the fast retrieval times of Verde make the user unaware that the file has been archived and removed from its original location.

To further protect data that resides on Verde, users have the option to make an additional "disaster recovery" copy with one of the following methods:

- Make a copy on tape via Verde's optional connection to BlackPearl and tape (see below)
- Replicate the data to another Verde at another site via Verde's replication feature
- Use Komprise to make a copy a second location:
  - In the public cloud (Amazon, Azure, etc.) - see Figure 5. Above
  - On a BlackPearl

BlackPearl with Komprise

Komprise can easily archive to all of Spectra Logic's various storage tiers – disk, tape and public cloud - with the Spectra BlackPearl. BlackPearl allows data to be seamlessly archived to disk, tape and public cloud storage through a single private cloud object storage interface. BlackPearl eliminates middleware and enables users to store their data on the most cost-effective tiers of storage for their unique environments. Key features of BlackPearl include:

- **The lowest cost storage available** - BlackPearl takes advantage of tape's low cost, economically scaling to exabyte capacities. It also promotes reductions in OPEX–related expenditures by leveraging the environmental benefits of Spectra's tape libraries, including high-density footprints and low power consumption.
- **Minimal software costs** - Integrating BlackPearl requires minimal software acquisition costs to attach archive storage into your environment by leveraging Komprise or other software clients based upon RESTful interfaces. Users can develop their own clients for the Spectra S3 interface or access a library of existing software clients available through the Spectra Software Developer program.
- **Reduce time spent managing long-term storage** – Archive storage is the ultimate resting place for all data, and with BlackPearl there is no need to continually move data to another storage target. BlackPearl minimizes the time spent moving data, allowing users to target objects for this type of storage. With BlackPearl, users do not need any previous experience with tape library systems. BlackPearl handles how data is written to tape, the tape library inventory, robotic commands, and other aspects of how the library stores the data.
- **Simple approach to interface and manage tape** - By leveraging the Spectra S3 interface, BlackPearl unlocks tape's true calling. BlackPearl takes user–defined buckets, caches the data on a combination of high speed SSD and SAS HDD and determines the best method to write it to tape for greatest system performance. To ensure perpetuity and ease of data access, BlackPearl writes data on tape in the open LTFS format to ensure you will always have access to it.
- **Deep Storage Interface modernizes the use of tape** - The first RESTful interface to tape, Spectra S3 provides BlackPearl and client interaction via REST, allowing you to move large data sets easily to and from archive storage. The user only moves data through the client while BlackPearl handles all interaction with the tape library.
- **Ensures data security, reliability and integrity** - Spectra tape libraries provide a powerful set of tools that allow for data integrity verification and notifications of media, drive, library status and health. BlackPearl also offers extended data security through Spectra's AES 256-encryption offerings.
- **Fast transfer of large storage objects** - BlackPearl offers high-performance transfer of bulk data to archive storage, allowing tape to sequentially read and write. BlackPearl optimizes how buckets are written to tape for best performance. Additionally, staging data to BlackPearl's tape cache allows for up to 3 GB per second reads and writes to easily offload data to tape.
- **Easily expands to exabytes and beyond** - BlackPearl unlocks tape's economies of scale, allowing you to drive the total cost of your data storage down as your system grows. Standard LTO and IBM TS tape technology roadmaps also span through future generations with significant capacity improvements to realize even greater capacity and storage density benefits.

- **Designed to keep data forever** - BlackPearl enables you to realize the long-term benefits of data storage, helping you retain your data indefinitely. BlackPearl includes future support for built-in media migration for seamless transfer of data to future archive storage media formats. It also includes and supports different drive generations to ensure a seamless transition to newer tape media formats.
- **Advanced Bucket Management (ABM) for BlackPearl** - BlackPearl uses internal management policies for different storage devices, enabling the system to write data intelligently to different tape libraries, disk storage, the public cloud, and other BlackPearls (replication). This allows BlackPearl to control where and how data is being written. Within a data policy, BlackPearl can make multiple copies of data on tape or disk. This provides a very flexible platform and storage management system, while protecting the organization's data with genetic diversity of media and storage devices. With ABM it's possible to have a redundant copies of a file on IBM® TS1155 tape, LTO–7 tape, SAS HDD, and SATA HDD.

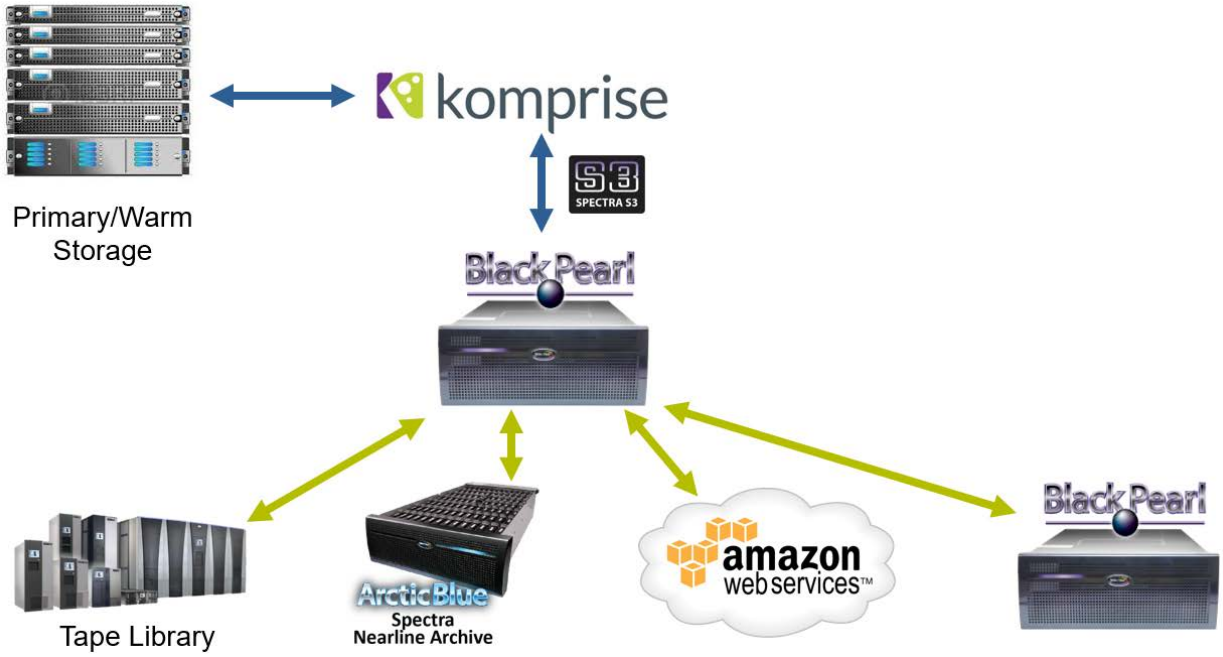

Figure 6. Archiving to Spectra BlackPearl via Komprise

Files archived to BlackPearl by Komprise can take varying amounts of time to restore depending on storage type being used behind BlackPearl. In the case of disk storage being used behind BlackPearl, restore will be almost immediately, similar to Verde as explained above. When data is stored on tape behind BlackPearl, restore times can take longer. To deal with the restore time from tape, users can sign up to receive email alerts when the files are restored, as shown in the image below.

With BlackPearl's "Advanced Bucket Management" data policy feature, users have the option to make a "disaster recovery" copy of the data archived via BlackPearl with one of the following methods:

- Make two copies of the data on two different tapes
- Make one copy of the data on disk and one copy of the data on tape
- Make one copy of the data on tape or disk, and one copy of the data on public cloud
- Make one copy of the data on tape or disk, and replicate the data to another BlackPearl (which itself has tape or disk behind it)

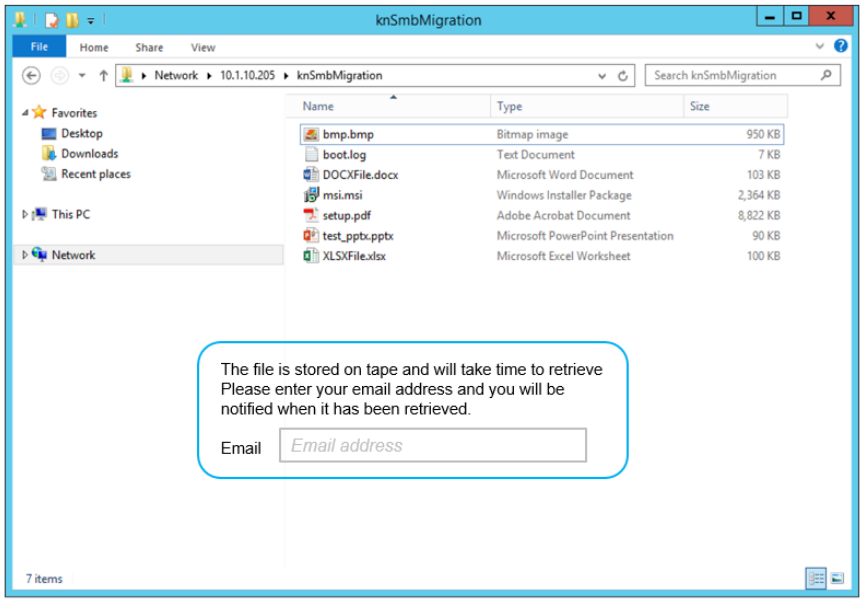

Figure 7. Accessing Files Archived to BlackPearl/Tape

Overall Costs and ROI Examples

By moving data off of primary storage using the Komprise and Spectra solution, customers can save money by creating additional space on their primary storage, eliminating the need to purchase additional primary storage. When considering the cost of primary storage, the investment to back up the data and store the backups must also be considered.

In the table below, the typical costs of primary storage per terabyte are shown. This is compared to the cost of the Komprise/Spectra solution. In this case, it is assumed that the data is being migrated to Spectra Verde with a Disaster Recovery (DR) copy being made to a Spectra tape library via the BlackPearl. Cost savings can be very significant, as shown in this example which shows an 86 percent savings.

CURRENT COSTS	PER TERABYTE	WITH KOMPRISE/SPECTRA	PER TERABYTE
Primary Storage	$1,800	Spectra Storage	$124
Backup/DR Storage	$350	DR Storage	$96
Backup Software	$486	Komprise Software	$150
Total Current Costs	$2,636	Total New Costs	$370
		SAVINGS	**$2,266 (86%)**

Figure 8. Cost Savings with Komprise/Spectra Solution

Summary

Data is growing fast and budgets are tight. Businesses need a simple way to efficiently manage data sprawl while cutting costs. The Komprise/Spectra solution is analytics-driven data management software coupled with archive storage that enables customers to identify inactive data, assess the ROI of moving it, and transparently move it to archive storage. Unlike legacy data management solutions that are costly, complex, invasive, and hard to scale, the Komprise/Spectra solution is simple to deploy and operate, seamless to scale, and very cost-efficient.

About Spectra Logic

Spectra Logic develops data storage solutions that solve the problem of short- and long- term digital preservation for business and technology professionals dealing with exponential data growth.

Dedicated solely to storage innovation for nearly 40 years, Spectra Logic's uncompromising product and customer focus is proven by the adoption of its solutions by industry leaders in multiple vertical markets globally.

Spectra enables affordable, multi-decade data storage and access by creating new methods of managing information in all forms of storage—including archive, backup, cold storage, private and public cloud. To learn more, visit www.SpectraLogic.com.

303-449-6400 • 800-833-1132 • (Fax) 303-939-8844 • 6285 Lookout Road • Boulder, CO 80301 USA • spectralogic.com

V1-081517