



# **HPE Reference Configuration for Komprise with HPE Storage Servers for File and Object**

Solution overview and best practices for data management with Komprise software, HPE servers, HPE StoreEasy, Scality RING, and Qumulo

# Contents

Executive summary.....	3
Introduction.....	3
Solution overview.....	3
Solution components.....	5
Komprise data management software.....	5
HPE StoreEasy.....	6
HPE Apollo 4000 Storage Servers.....	6
HPE Scalable Object Storage with Scality RING.....	6
HPE Apollo 4200 with Qumulo File Fabric.....	7
Best practices and configuration guidance for the solution.....	7
Komprise deployment and configuration.....	7
Komprise configuration for Scality RING.....	11
Komprise configuration for Qumulo File Fabric.....	11
Komprise NAS and target support.....	12
Summary.....	12
Resources and additional links.....	13



## Executive summary

Data is growing very rapidly - much of the world's data was created in the last two years, and enterprise data is growing at a furious pace. There are constant budget pressures, and an increasing need to handle immutable data as well as Big Data. The question of how to retain all this data while keeping within flat budgets is a significant but not insurmountable challenge. Businesses need to do more with less and are looking for affordable scale-out storage options, such as cloud and object storage, to address their growing storage needs.

As data footprints expand rapidly, managing the data with ease becomes very important. Komprise data management software analyzes and manages storage data with no storage agents, no static stubs, no changes to the hot data path, no storage lock-in, and no scaling limits. With Komprise, users can attain cloud-like economics for instant savings of up to 70% or more. Komprise and Hewlett Packard Enterprise are working with customers across industries such as financial services, healthcare, energy, and engineering who are streamlining costs, building a path to more efficient capacity-enhanced storage, and increasing the resiliency of their data in use cases like active archiving, replication, and disaster recovery.

Data management has historically been difficult to adapt and scale to keep pace with today's massive data growth. Komprise addresses this need with a modern architecture and technology that has been built from the ground-up to handle today's massive scale of data with intelligent automation. The Komprise software deploys in under 15 minutes, works across NFS, SMB/CIFS, and S3 object storage without any storage agents, adapts to file system and network loads to run non-intrusively in the background, and scales out on-demand.

Komprise data management software extends the life of primary NAS storage, such as EMC, NetApp, HPE StoreEasy, and HPE Apollo with Qumulo File Fabric, by ensuring it does not fill up with immutable, old, or inactive data. This solution from Komprise and HPE simplifies data management across multiple NAS systems and significantly reduces storage costs.

**Target audience:** This document is intended for solution architects, engineers, sales representatives, end-user IT professionals, and others interested in understanding the features and components of the Komprise solution with HPE StoreEasy, Scalality RING on HPE Apollo 4000, and Qumulo on HPE Apollo 4200.

## Introduction

Komprise research indicates that 60% to 80% of data is infrequently accessed and becomes inactive within months of creation, yet it is often stored and managed in the same way as active business data. Komprise software enables businesses to manage their data intelligently by identifying inactive, immutable, and other targeted data across their NAS storage frames, and transparently moving targeted data with user-defined policies to cost-effective options such as cloud or object storage - without any changes to user or application access. Policies commonly include least recently used (LRU) for moving older and inactive data to lower-cost storage options, as well as moving immutable data, marketing videos, and other audio or video content to a lower tier of storage.

This approach seamlessly extends the capacity of the original on-premises storage virtually, without disrupting users and applications. It ensures that only desired and active data is kept on the tier-1 primary storage by relegating immutable, inactive, and other targeted data to a capacity-enhanced storage system designed to store large quantities of data much more cost effectively. HPE StoreEasy and HPE Apollo with Qumulo are ideal as high performance, scalable tier-1 NAS storage, while HPE Scalable Object Storage with Scalality RING on HPE Apollo is perfectly suited to provide the capacity-enhanced object storage needed to reduce or eliminate the need to increase the capacity of tier-1 NAS storage, thus reducing on-going costs and maintenance.

For redundancy and disaster recovery, many businesses need to keep a copy of their data in an object store that protects its data across multiple geographically dispersed data centers. The Komprise policy-based tools enable the continuous copying of data, for example, from an HPE StoreEasy system or an HPE Apollo 4200 with Qumulo, to HPE Scalable Object Storage with Scalality RING on HPE Apollo 4000 Storage Servers. This provides protection during a network outage between data centers and in cases where one of the data centers becomes inaccessible due to error, equipment failure, or disaster. In a disaster recovery situation, Komprise enables access to data in the remote data center and maintains business continuity.

## Solution overview

The Komprise data analytics software does not require any dedicated hardware and runs as a scale-out grid of VMware® virtual machines. It is an easy-to-deploy, easy-to-scale, distributed architecture that consists of one or more Komprise Observer virtual machines running on the customer's premises, connected to a Komprise Director virtual machine that can run either as a cloud service or on-premises. To scale effectively without centralized bottlenecks, Komprise does not use a traditional SQL database that limits scalability. Instead, the Komprise Observers analyze and aggregate metadata and the Komprise Director presents an accumulation of these aggregates.



Komprise works seamlessly across many data sources using NFS, SMB/CIFS, or S3 object/cloud storage - including file servers such as those from Dell/EMC, HPE, NetApp, and Microsoft® Windows® File Servers. Komprise profiles the data across all storage devices that are visible to it, and provides easy-to-use analytics to answer the following questions:

- How much data is IT managing across all of the storage devices?
- What are the file types and how much of each are currently being managed by IT?
- What is the distribution of file sizes?
- Who is accessing which files?
- How fast is file storage growing?
- What is the associated hosting cost of data at the current locations?

After storage has been analyzed, Komprise software allows specifying various data management objectives and applying different “what-if” scenarios to understand their projected impact on storage capacity and data-hosting costs. When the plan is satisfactory for the data to move or copy, when to move it, and to where, Komprise Observer moves it to the target object stores based on customer-defined policies. The data is moved transparently by replacing the original file with a Komprise dynamic link that looks and operates like the original file. When a user accesses the link, Komprise software is invoked and provides in-situ access to the file as if it were still on the original tier-1 NAS storage. Komprise software can be configured to move the data back to tier-1 NAS storage if it is accessed frequently after having been transferred to an object store.

As an example solution architecture, Figure 1 shows how Komprise moves or copies data from a primary, tier-1 NAS storage and writes it to HPE Scalable Object Storage with Scality RING on HPE 4000 Apollo Storage Servers, optimizing the use of primary storage and capacity-enhanced object stores. Note that the tier-1 NAS could also be HPE StoreEasy or HPE Apollo 4200 with Qumulo.

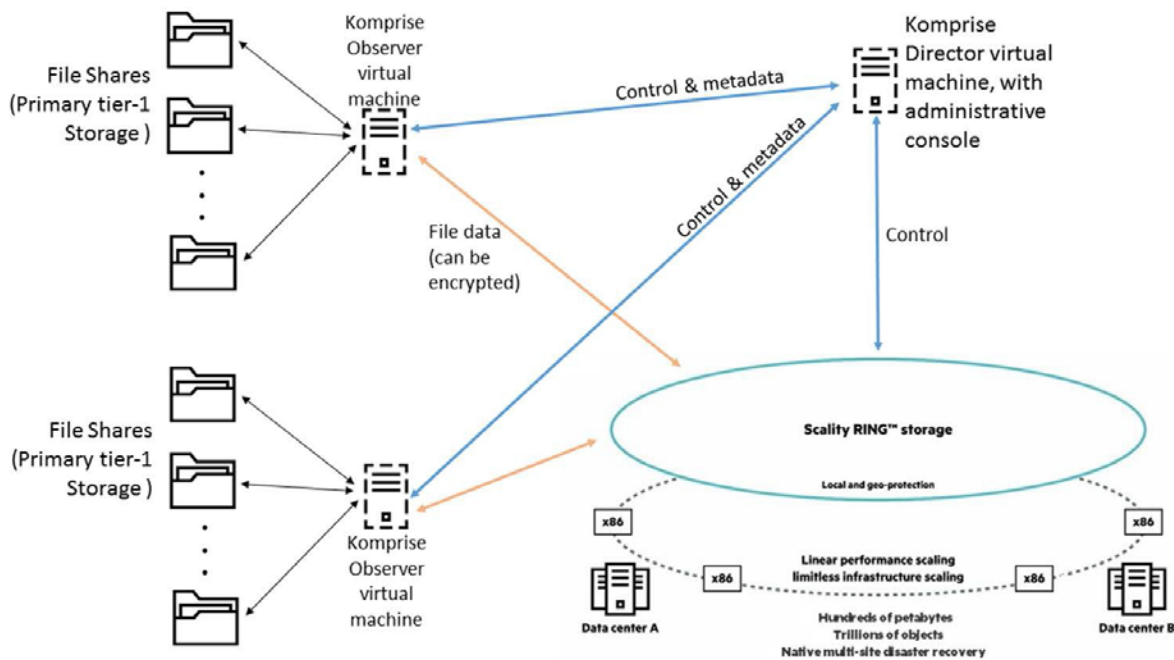


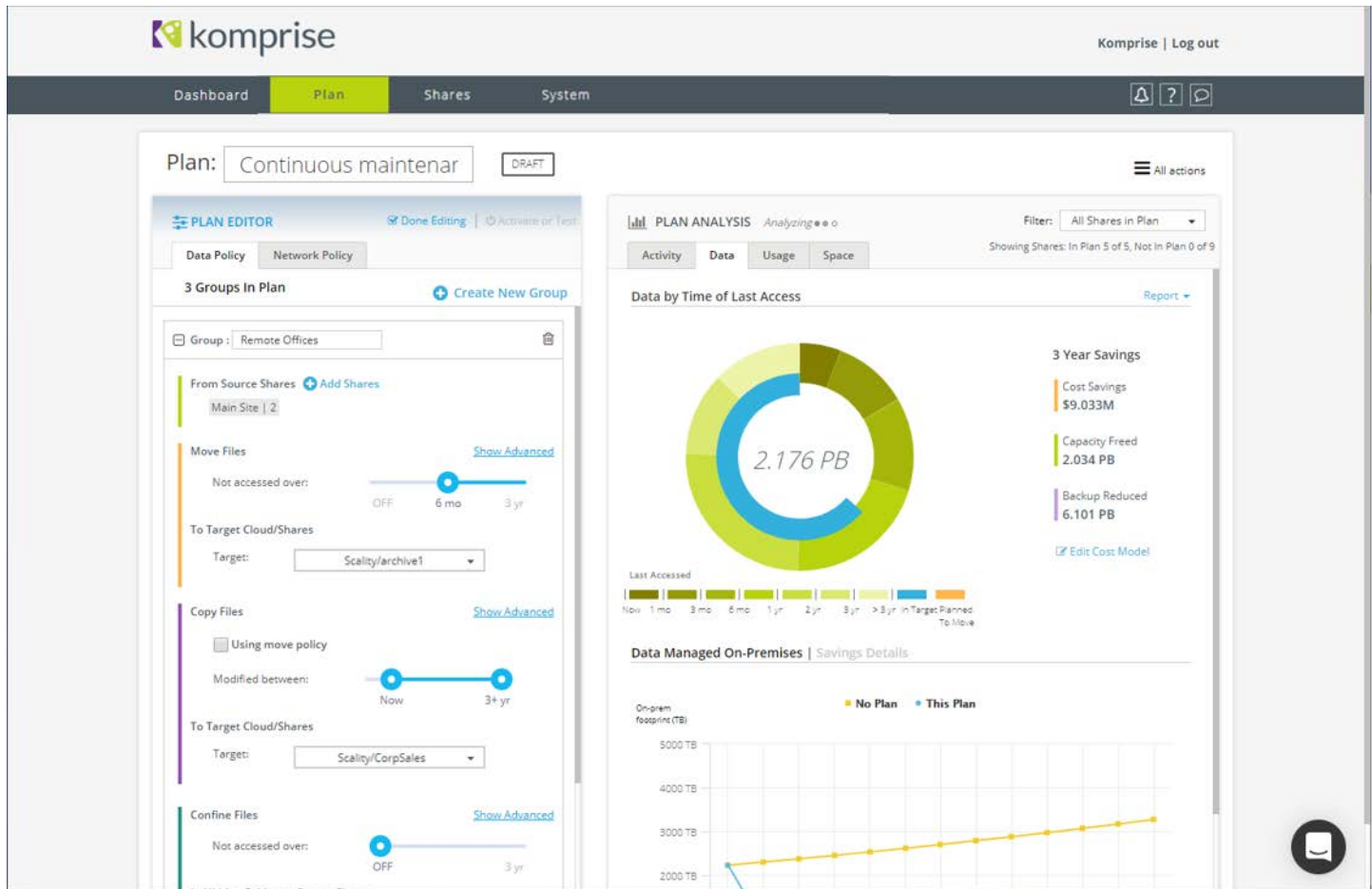
Figure 1. Solution architecture for Komprise with HPE Scalable Object Storage with Scality RING on HPE 4000 Apollo Storage Servers



## Solution components

### Komprise data management software

The management console for Komprise Director, as shown in Figure 2, is used to analyze data usage and growth, plan and create policies to move and manage data, and automate ongoing management across storage sources and targets.



**Figure 2.** Komprise console showing analysis of data usage

After configuring the Komprise data management software to connect to source file shares, Komprise software will analyze the shares and generate several charts showing the time of last access, as shown in Figure 2 above, and space consumed by file type, file size, top users, and top groups.

After storage has been analyzed, a Komprise plan is created and modified for minimizing storage capacity and data-hosting costs. The Komprise management console interactively projects the estimated capacity that will be freed up on the primary storage and the projected cost savings. The cost model created by the Komprise management console can be modified to use your own data-hosting costs, providing a customized ROI calculation. Komprise software also projects the 3-year savings based on historical growth rates of data. This provides a no-risk way to plan the most effective data management approach before actually invoking any data movement.

After the Komprise Plan is finalized, it is activated and the Komprise Observer moves the data transparently from the tier-1 NAS (for example, HPE StoreEasy or HPE Apollo with Qumulo) to the target (for example, HPE Scalable Object Storage with Scality RING), based on user-defined policies specified in the Plan. The transferred data still appears to exist on the source storage system as before. When a user or application accesses the data on the source system, the Komprise Observer transparently returns the data from the target without restoring data to the original source location. This behavior is configurable based on the frequency of requests to access the moved data, and data can either be cached and served

from the Komprise Observer, or restored to the source location. Komprise software delivers file-based access to all the moved data so existing users and applications continue to access the data as before.

Komprise data management software differs from typical storage array-tiering solutions that move inactive data blocks to the cloud and require that data be accessed only by restoring it to the source array, thus limiting customers from leveraging the true power of their object stores as “active archives.” Instead, Komprise moves files as objects and allows object-level access to the moved data from the store, as well as file-level access from the original source storage device using dynamic links.

Komprise data management software 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 due to the increase of data, additional Observers can be added on-demand. All Komprise Observers are managed as a single fault-tolerant grid – making management simple while delivering a robust, efficient, and scalable high-performance solution.

### **Encryption**

By default, the Komprise Observer and software transfer data securely, then uses encryption at rest to further secure the data on the target storage. For additional security, Komprise also supports AES-256 encryption at the source through customer-provided master keys. Komprise software uses the customer-provided keys to encrypt the data before it is moved, then stores the data in its encrypted state on the target object store. When data needs to be retrieved, the encryption keys are required to restore the data.

### **Compression**

If the target storage does not provide data compression capability, Komprise software provides the option to compress data prior to transfer, which can further save storage costs. The compression ratio is dependent on the type of data. Komprise uses an industry standard lossless data compression algorithm that uses a combination of the LZ77 algorithm and Huffman coding.

### **HPE StoreEasy**

HPE StoreEasy is a leading, easy-to-manage NAS product family that provides a centralized space for securely storing documents, images, audio, and video files. With StoreEasy 1000, you can have flexible, efficient, and reliable file storage, right-sized for your organization’s needs and budget, without compromising on quality or capabilities.

### **HPE Apollo 4000 Storage Servers**

HPE Apollo 4000 Storage Servers are third-generation density-optimized platforms, purpose-built to service object storage solutions and Big Data analytics. The power and flexibility of HPE Apollo 4000 platforms provide very robust and scalable object storage solutions that scale out linearly as single protected systems across multiple sites and thousands of servers. HPE Apollo 4000 systems are configured to serve as storage nodes in a Scality RING cluster, providing both storage and connector functionality.

HPE Apollo 4000 platforms are 4U or 2U servers that leverage the modular and efficient HPE Apollo chassis infrastructure to provide storage density and operating efficiency. HPE Apollo 4000 Storage Servers provide configuration flexibility to optimize for capacity, throughput, and responsiveness. The systems are designed to maintain availability, data recovery, and support serviceability. HPE Apollo 4000 systems include the HPE Integrated Lights Out (iLO) management capability, allowing secure management and monitoring of servers from practically anywhere. They also integrate with HPE OneView for automating the management of IT infrastructure, and offer a choice of HPE Smart Array controllers.

Select HPE Smart Array controller cards are capable of Secure Encryption providing enterprise-class encryption. Secure Encryption is FIPS 140-2 certified and has been verified to have a low impact on IOPS for spinning media, in addition to being transparent to the operating system. Hot-plug critical components (disk drives, nodes, fans, and power supplies) support serviceability at every level.

### **HPE Scalable Object Storage with Scality RING**

The HPE Scalable Object Storage with Scality RING is a software-defined storage (SDS), petabyte-scale data storage solution that is designed to interoperate in the modern software-defined data center (SDDC). HPE Scalable Object Storage with Scality RING utilizes the HPE Apollo 4000 storage servers for high-density storage capacity, which enable enterprises to deploy Scality RING with efficiency at the scale their business needs.



### Scality RING

The Scality RING (RING) running on HPE Apollo 4000 storage servers provides an SDS solution for petabyte-scale data storage that is designed to interoperate in the modern software-defined data center (SDDC). The RING software is fashioned to create a scale-out storage system, which is deployed as a distributed system on a minimum cluster of three storage servers. This system can be seamlessly expanded to thousands of physical storage servers as the need for storage capacity grows. To match performance to the deployed capacity, the RING can independently scale out the access nodes to meet a customer's growing I/O throughput requirements.

The RING software is deployed as a distributed system on a minimum cluster of three storage servers. This system can be seamlessly expanded online to thousands of physical storage servers as the need for storage capacity grows. HPE Apollo 4000 Storage Servers provide a combination of HDDs for RING data and solid-state disks (SSDs) for RING metadata. Scality RING connector services can be run directly on the storage nodes for maximum efficiency or on external gateway servers, if desired, for additional flexibility.

### HPE Apollo 4200 with Qumulo File Fabric

Apollo 4200 Gen9 servers running the Qumulo File Fabric (QF2) software is a modern, highly scalable file storage solution that runs in the data center. More economical than legacy storage with leading performance, the Apollo 4200 with QF2 file storage solution provides real-time analytics to let administrators easily manage data no matter how large the footprint or where it's located globally.

### Qumulo File Fabric

Qumulo File Fabric (QF2) is a modern, highly scalable file storage system that runs in the data center and the public cloud. QF2 on-premises instances serve high-performance workloads such as genomic sequencing and analysis, high-resolution video editing, scientific computing and the ingestion of largescale datasets for machine learning applications.

QF2 on HPE Apollo 4200 Servers are more economical than legacy storage appliances from a usable capacity basis, and QF2's up-to-the-minute analytics allow administrators to pinpoint problems and effectively control how storage is used.

## Best practices and configuration guidance for the solution

### Komprise deployment and configuration

The high-level steps to deploy and configure Komprise software are as follows:

1. **If using an on-premises configuration, download the Komprise Director VM and install on a virtual server.** Configure the network settings as needed for the environment. (Skip to step 2 if using the recommended cloud configuration.)
2. **Download the Komprise Observer VM and install on a virtual server.** Set up the network configuration and use the virtual appliance console to connect and authenticate with the on-premises Komprise Director configured above, or with the Komprise Director configured in the cloud.



**Configure the Observer to discover and analyze the shares on the file servers.** After an Observer is authenticated with the Director, the web-based UI is used to add or discover the file shares and enable them for analysis.

Use the Komprise Director management console to discover and connect with file servers and shares where source data resides. In the **Shares** tab, file servers and their shares can be discovered or manually specified. Figure 3 shows the **Shares** tab with several file servers. The **Add File Server** or **Discover Shares** links can be used to add more file servers and shares.

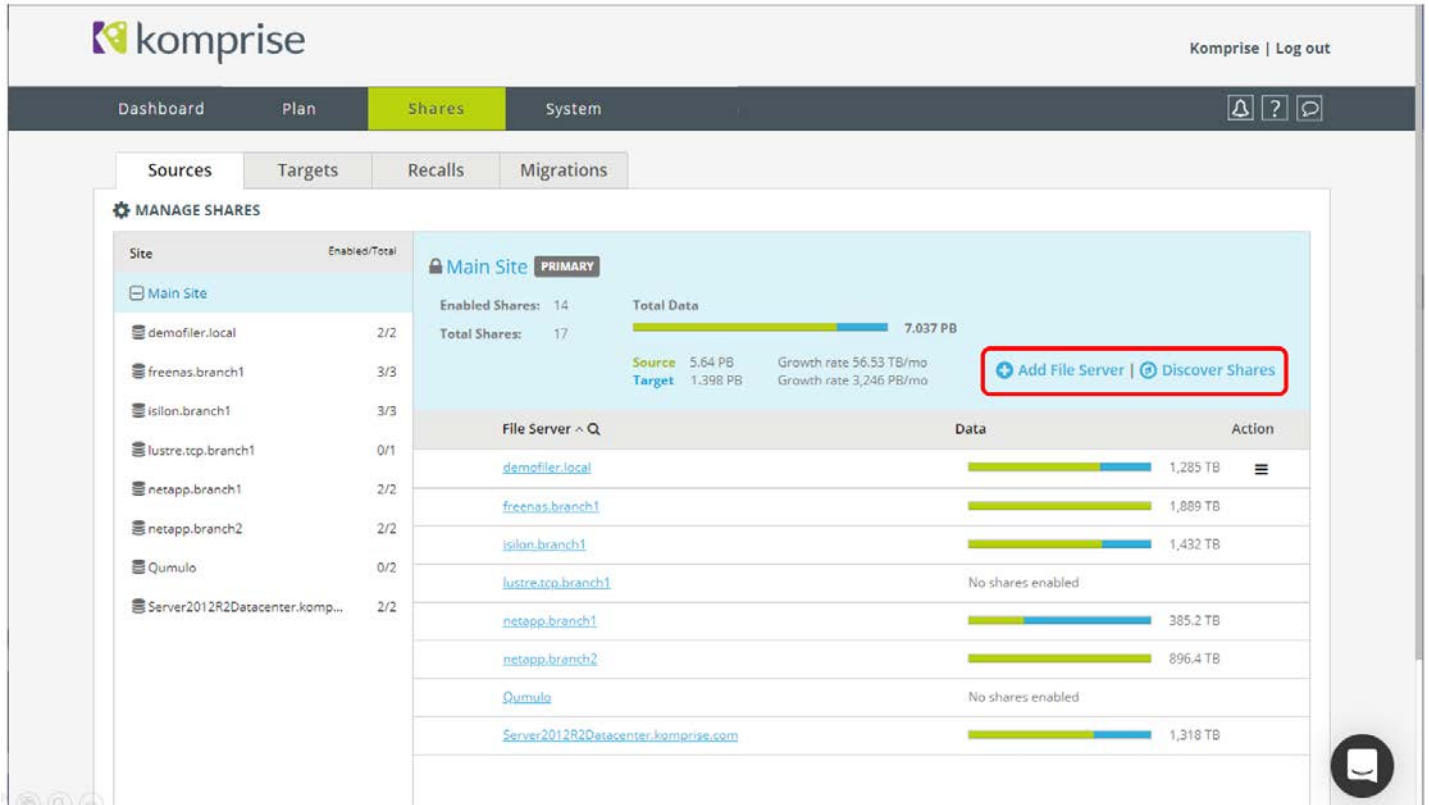
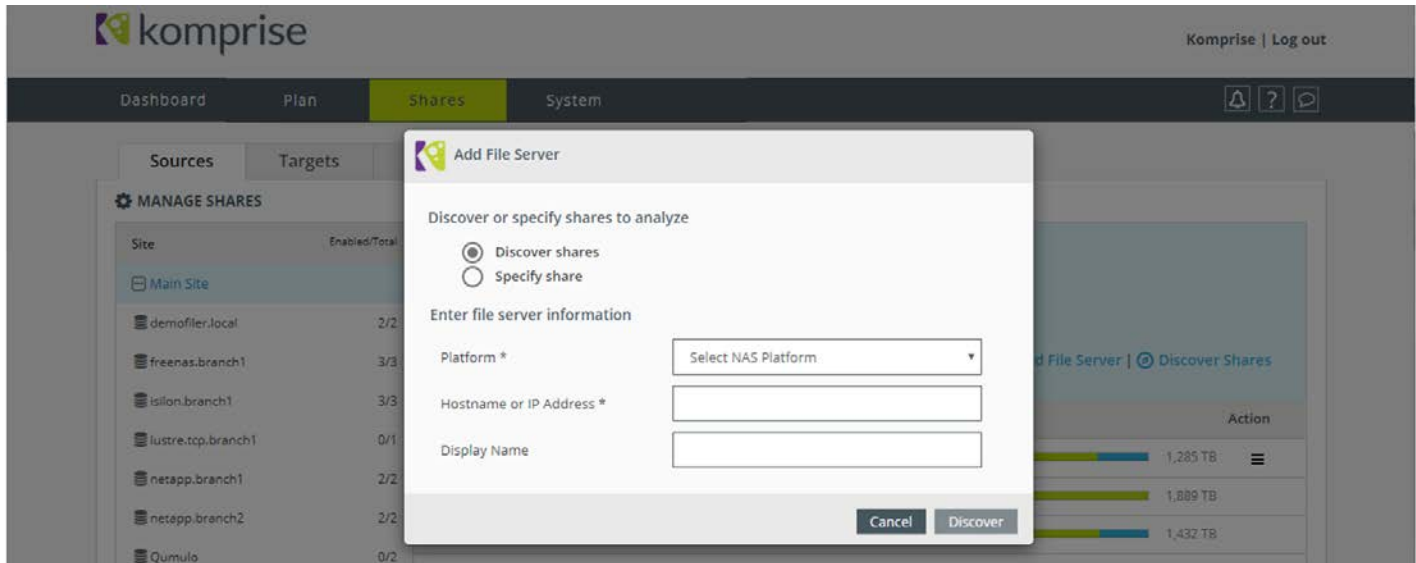


Figure 3. Adding storage via the Komprise Director management console





By clicking **Add File Server** as shown in Figure 4, shares can be discovered on an existing file server, or shares can be manually added by specifying a particular type of file server and the appropriate fully qualified hostname or IP address. File servers such as EMC Isilon, NetApp, Windows File Server, and Linux® are available from the **Select NAS Platform** drop-down list.



**Figure 4.** Adding a file server, via the Komprise Director management console

1. **Enable the shares and analyze them.** A share must be enabled in the Komprise data management software before it is analyzed. After a share is enabled, Komprise software will immediately run an initial analysis, then periodically re-analyze the share to capture changes to its contents. The analysis is designed to provide immediate results, which become more accurate after the first analysis pass of the share is complete. This approach allows the user to receive results immediately and start the planning process without having to wait for the first analysis to complete
2. **Develop a plan and activate it.** After the shares have been added or discovered, enabled, and are being analyzed, the **PLAN EDITOR** on the left pane (See Figure 5) is used to create a plan to manage the data. A Plan consists of one or more Groups. A Group is a collection of shares, a set of policies that operate on those shares to move or copy data, and a target to which data is transferred. As changes are made to the Plan, the **PLAN ANALYSIS** pane changes to show the impact on the Plan including the cost savings over 3 years, the amount of space that will be freed, and the reduction in backup data volume. After a Plan has been developed, it can be activated in test mode or in normal mode. In test mode, the Plan does not transfer any data, but instead creates a list of files that would be affected by the Plan. This allows the administrator to validate the Plan configuration before implementing the Plan. In normal mode, the Plan operates on the files as specified by the policy and will move (archive) and/or copy data to the target or targets specified in the Plan.



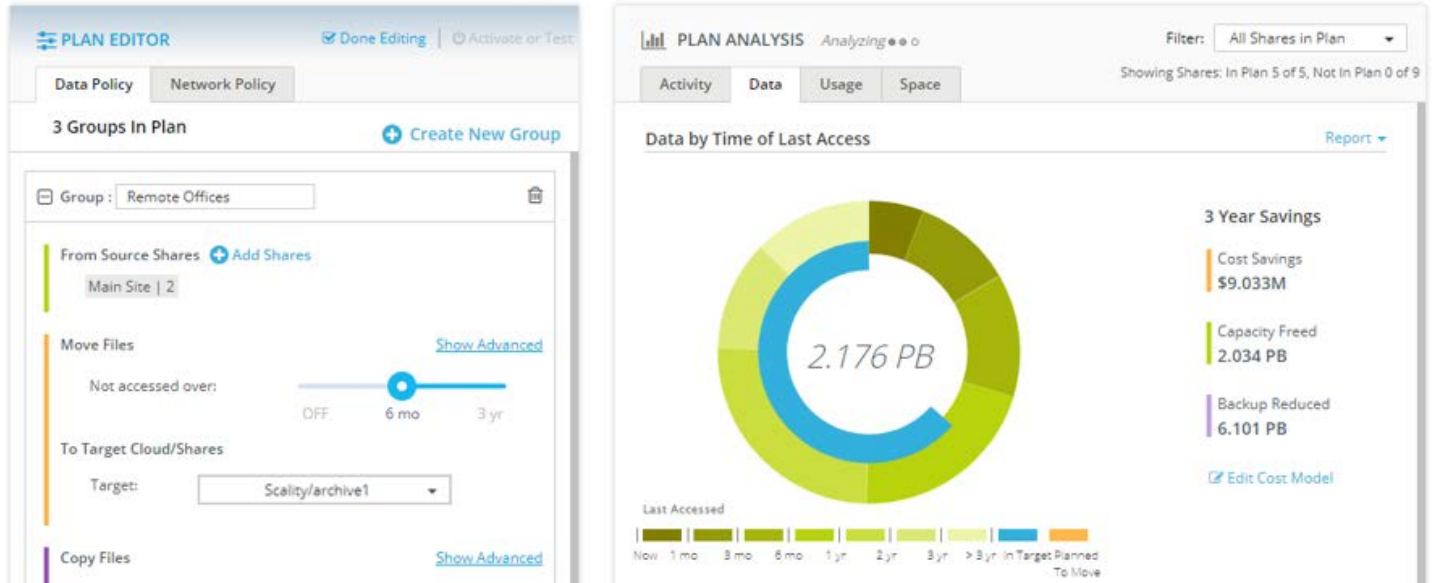


Figure 5. Komprise PLAN EDITOR

The Komprise Observer and software are invisible to the active data path and are not placed inline. Therefore, performance of the active data is unchanged and in fact might improve as the primary storage is less overloaded.



### Komprise configuration for Scality RING

No special configuration of the HPE Scalable Object Storage with Scality RING is needed to use its object storage with the Komprise software. The Scality RING is delivered from HPE as a fully functioning, simple storage service (S3) target, ready to be used immediately with applications such as Komprise.

When configuring the target settings in the Komprise management console, choose Scality RING as the Storage Platform, then provide the Object Store URL, the Bucket Name, the Key ID, and Secret Key, as shown in Figure 6. In addition, encryption of data prior to sending it to the target object store can be turned on or off, as can compression for chunked data format mode.

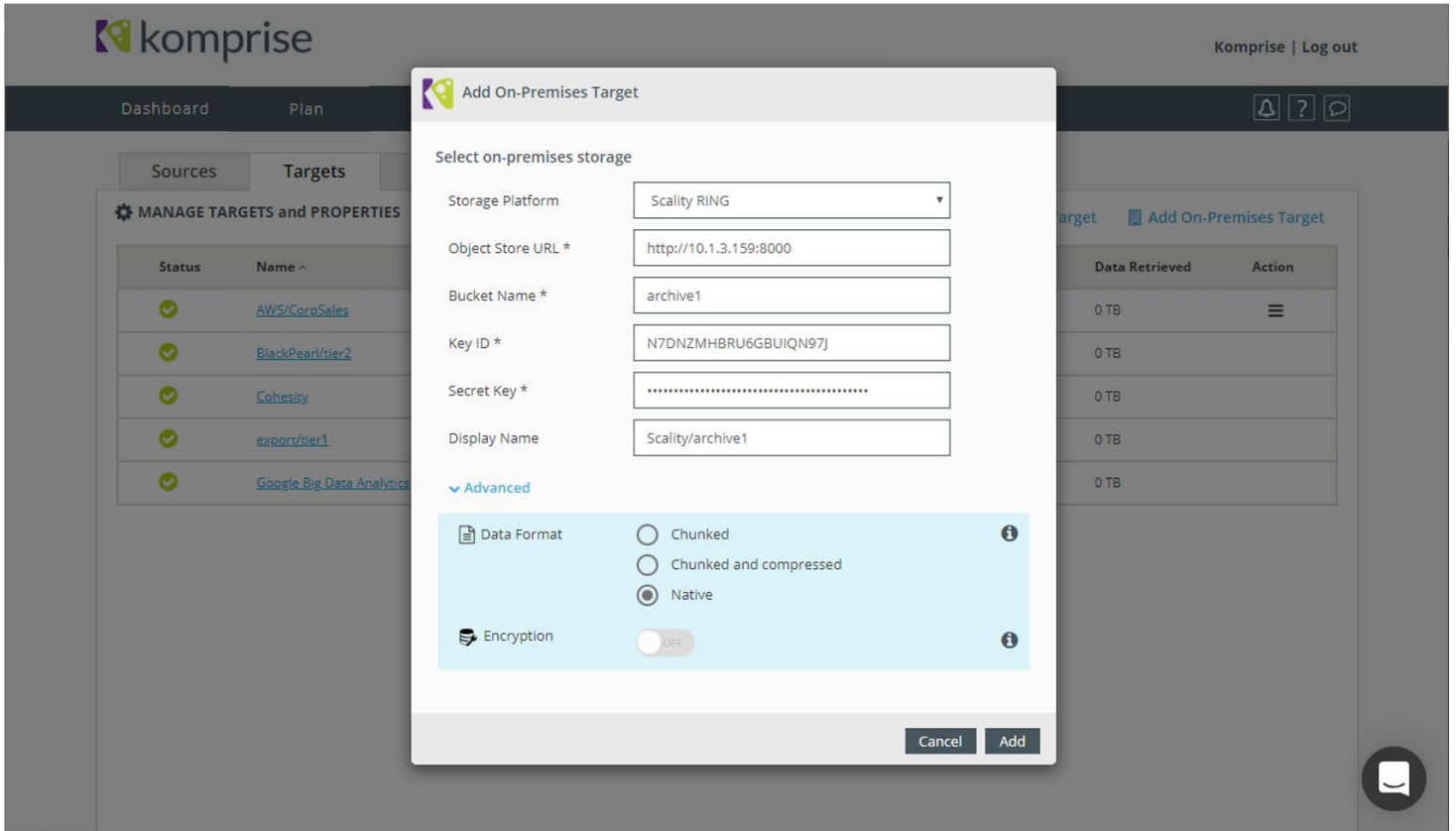


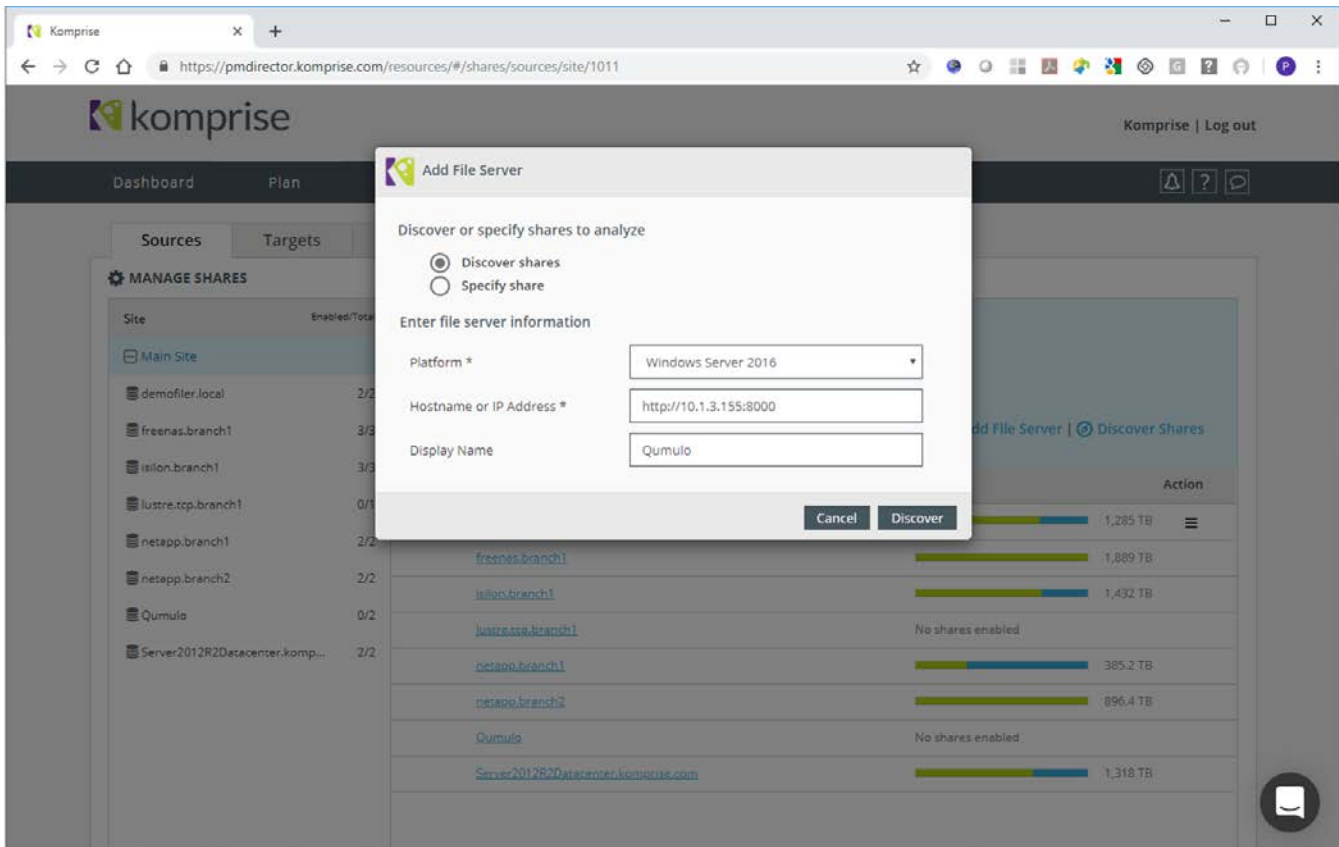
Figure 6. Configuring Scality RING as a target in Komprise

### Komprise configuration for Qumulo File Fabric

No special configuration of Qumulo on HPE Apollo 4200 is needed to enable use of the storage solution as tier-1 NAS storage with the Komprise software.

When configuring Qumulo in the Komprise management console as a File Server source, select Windows Server 2016 as the Platform, then provide the appropriate Hostname or IP Address and any desired Display Name, as shown in Figure 8. You can select “Discover shares” to discover any NFS shares, then you can manually add any SMB shares afterwards, using “Add Shares” in the management console.





**Figure 7.** Specifying Qumulo as a NAS source in Komprise

Once added, the shares can be enabled using the action menu for the new Qumulo file server, and then used as tier-1 NAS sources that can be archived by Komprise.

## Komprise NAS and target support

In addition to supporting HPE StoreEasy and HPE Apollo 4200 with Qumulo File Fabric as tier-1 NAS sources, and Scality RING on HPE Apollo 4000 as archive target, Komprise is vendor-agnostic and supports a wide range of other NAS sources and storage targets: see <http://komprise.com/product/compatibility>.

Wherever your data is stored, with Komprise, you can shrink backups, create low-cost copies, and transparently move cold, inactive files to cheaper storage solutions, which often delivers instant savings of 70%+. There are no storage agents, no static stubs, no hardware, no vendor lock-in, and no changes to the data paths, which means there are no unnecessary costs, delays, or disruptions. With Komprise, you are simply free to make the most of your data and save.

## Summary

Data is growing rapidly and IT budgets are tight. Businesses need a simpler way to efficiently manage data sprawl while cutting costs. Komprise is analytics-driven data management software that identifies inactive data, assesses the ROI of moving it to more cost-effective storage, and then transparently moves it to on-premises or cloud targets of your choice. Unlike legacy data management solutions that are costly, complex, invasive, and hard to scale, Komprise is simple to deploy and operate, seamless to scale, and cost efficient.

Combining Komprise data management software with HPE StoreEasy, HPE Apollo 4200 with Qumulo, or HPE Scalable Object Storage with Scality RING on HPE Apollo 4000 relieves primary storage from the pressure of rapidly growing data by making it easy to move older, inactive, or immutable data to more cost-effective object storage. This extends the life of primary storage by helping ensure it does not fill up with inactive data and achieves cloud economics with on-premises deployments.



## Reference Architecture

### Resources and additional links

HPE Reference Architectures, [hpe.com/info/ra](https://hpe.com/info/ra)

HPE Complete, [hpe.com/storage/hpe-complete](https://hpe.com/storage/hpe-complete)

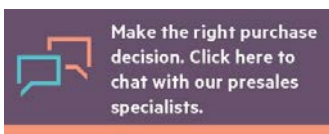
HPE Servers, [hpe.com/servers](https://hpe.com/servers)

HPE Storage, [hpe.com/storage](https://hpe.com/storage)

HPE Scalable Object Storage, [hpe.com/storage/scalableobject](https://hpe.com/storage/scalableobject)

HPE Technology Consulting Services, [hpe.com/us/en/services/consulting.html](https://hpe.com/us/en/services/consulting.html)

To help us improve our documents, please provide feedback at [hpe.com/contact/feedback](https://hpe.com/contact/feedback).



---

**Sign up for updates**

---

---

© Copyright 2019 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

Microsoft and Windows are registered trademarks of Microsoft Corporation in the United States and/or other countries. VMware is a registered trademark of the VMware corporation in the United States. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries. Scality RING is a trademark of the Scality corporation in the United States and other countries.

a00022829enw, June 2019, Rev. 1

