HPE GreenLake

Technical white paper

HPE GreenLake for Backup and Recovery—Backup-as-a-Service built for the hybrid cloud

A data protection service for on-premises VMware and AWS workloads

Contents

Executive summary

As businesses face unrelenting data growth, stringent recovery and restore service-level agreements (SLAs), and increasingly virtualized environments, protecting and recovering VMware® VMs and Amazon EBS Volumes and EC2 instances become even more challenging. Traditional agent-based and proxy server-based processes provide reliable backup, recovery, and retention, but they can also impact application performance and add cost and complexity. While snapshots offer fast, non-disruptive point-in-time copies, snapshots alone cannot deliver comprehensive protection of the virtual workload. They have retention limitations, corruption vulnerabilities, and dependence on the underlying storage system. In fact, snapshots are at risk if the storage system fails.

Traditional backup approaches are not designed for today's dynamic and distributed business environment. They are time-consuming and costly to manage. They cannot easily scale to meet progressive, changing needs. They require upfront CAPEX, expensive software licensing, and costly over-provisioning. Yesterday's approaches fall short when it comes to neutralizing modern threats such as ransomware and other sophisticated malware.

Delivered through HPE GreenLake cloud data services, HPE GreenLake for Backup and Recovery provides a comprehensive data management service consolidating all the data silos to deploy compute resources, provision storage, and protect workloads with a single unified access and cloud operational experience. This service protects on-premises and cloud native workloads in a simple and efficient manner, with global protection policies for consistent protection on-premises or in the cloud via a single SaaS console. It brings policy-based automation to protect your VMware VMs on any storage and Amazon EBS volumes and EC2 instances in a few simple steps—within minutes—eliminating the complexities of managing your backup and recovery operations on-premises or in the cloud. There are no additional cloud gateways, agents, backup software, proxies, media servers, or backup targets to manage, and cloud storage is fully managed and scaled automatically by the service. Threats like ransomware and malware are neutralized with built-in encryption, data immutability, dual authorization, and flexibility to store backup copies in an air-gapped manner making them inaccessible to cybercriminals. Organizations lower the cost of protecting data on-premises or in the cloud with consumption-based pricing and ultra-efficient data reduction technologies.

Target audience: Pre-sales consultants, solution architects, IT and storage administrators responsible for the protection of virtual workloads.

Document purpose: This paper describes how HPE GreenLake for Backup and Recovery is differentiated from other solutions offering data protection, with unmatched storage efficiency and ease of use.

Service overview

HPE GreenLake for Backup and Recovery is a cloud native Backup-as-a-Service (BaaS) designed for the hybrid cloud and delivered through HPE GreenLake edge-to-cloud platform. The service provides a solution for protecting on-premises VMware and AWS workloads with SaaS simplicity and best-in-class storage efficiency. The data flow is shown in Figure 1.

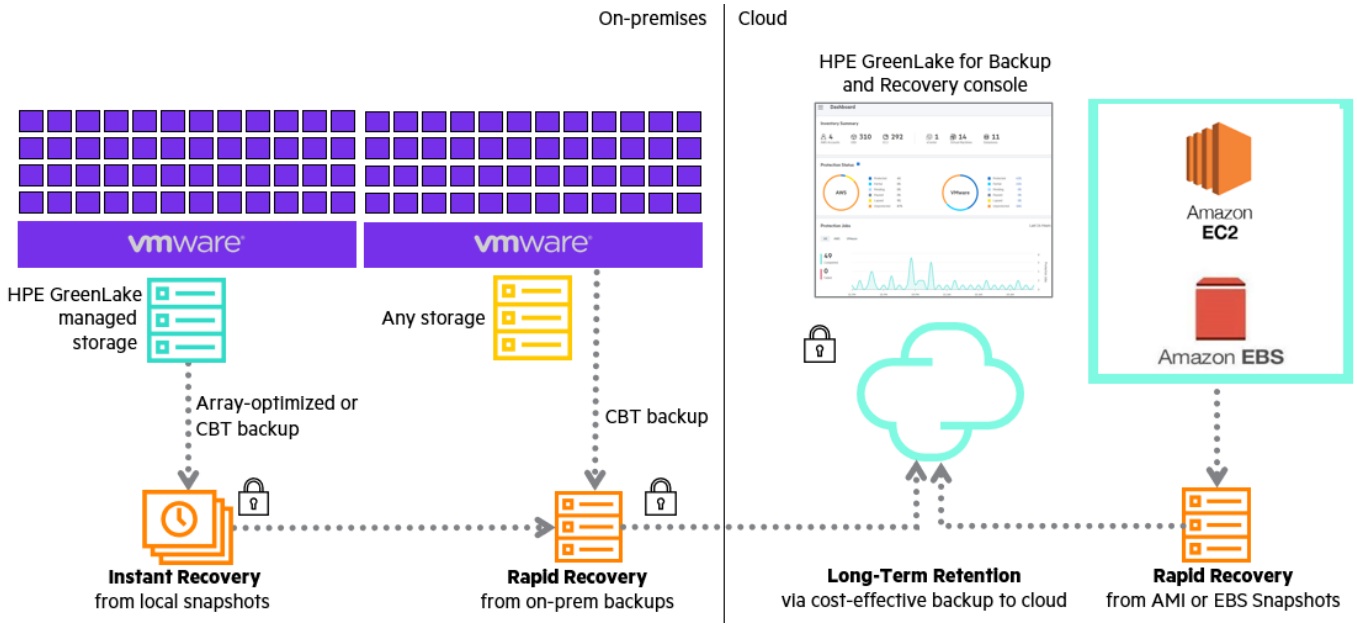

Figure 1. HPE GreenLake for Backup and Recovery provides hybrid data protection for on-premises VMware and AWS workloads

For on-premises storage and backup types, see Storage (HPE GreenLake managed or any other)

Key benefits of HPE GreenLake for Backup and Recovery

Effortless protection to meet your SLAs

The HPE GreenLake for Backup and Recovery console makes it easy to manage backup and recovery operations from anywhere. This cloud native console is designed for simplicity of operation. There is no requirement to be a backup and recovery expert or require training to set up and use the service. At its core is the capability for flexible protection that covers on-premises and cloud resources. On-premises backups include array snapshots for instant recovery, on-premises backups for rapid recovery, and cloud backups for long-term retention. Cloud backups include AWS snapshots and AMI (Amazon Machine Images) for instant recovery and cloud backups for long-term retention.

Unlike traditional backup and recovery software, you do not have to build a solution based on media servers, catalog servers, backup servers, and integration with secondary storage devices. HPE GreenLake for Backup and Recovery runs the service in the cloud and includes the integrated management of backup targets, which significantly reduces the hassle of managing backup infrastructure. This means you can focus on the protection delivered by the service—not the maintenance and operation of the backup infrastructure. Meeting the protection and recovery commitments to your organization and achieving compliance with data governance regulations is assisted with global Protection Policies and Protection Groups that provide auto-protection of existing and newly created resources. APIs are also available for scripted and automated orchestration of protection.

HPE GreenLake for Backup and Recovery is managed through a cloud native console designed to be easy to use. It can be launched from anywhere with internet access via console.greenlake.hpe.com, see Figure 2 as an example.

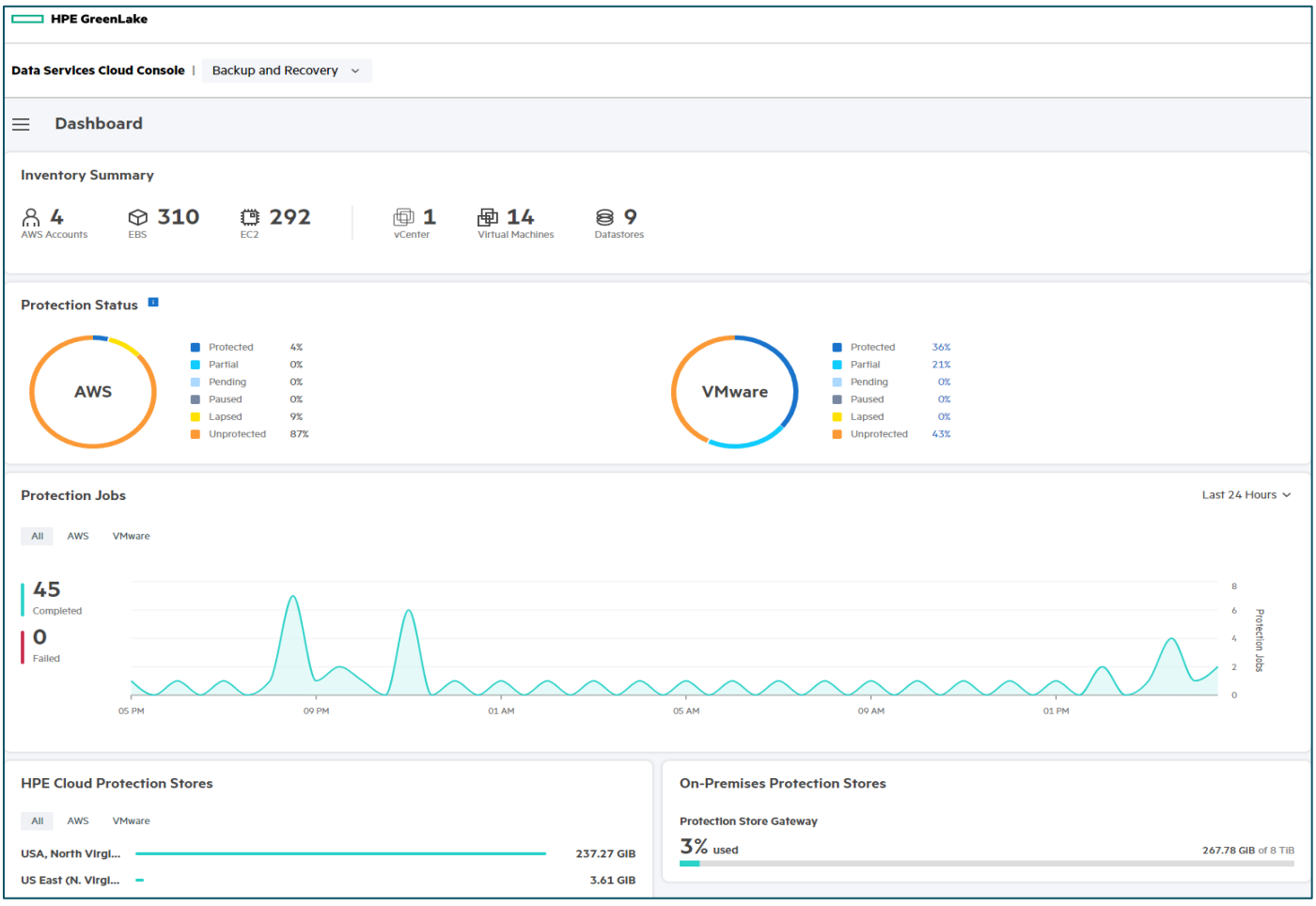

Figure 2. HPE GreenLake for Backup and Recovery dashboard provides a rich view of the overall protection status

Flexible billing, based on usage

HPE GreenLake for Backup and Recovery offers pay-as-you-go[1] usage and billing. Alternatively, the service can be reserved for a committed duration and usage for lower per-unit prices. There are no licenses to manage, just the subscription you choose. Usage-based billing has two billing metrics: the number of protected resources and the capacity of the cloud storage consumed. Elastic scaling provides more capacity when you need it so that resources always match the demand. There is no separate public cloud subscription required for long-term retention and there are no egress charges for recovery.

Unmatched storage efficiency

HPE GreenLake for Backup and Recovery is highly space efficient. It delivers unmatched storage efficiency by using HPE Catalyst technology with small deduplication chunk sizes. Testing has shown that backups consume up to 8x less space than similar solutions so there is less storage to manage, pay for, and move to the cloud for protection. Figure 3 shows the storage efficiency compared to four other solutions that offer similar protection. The relative advantage of HPE GreenLake for Backup and Recovery is clear—enabled by the superior HPE Catalyst deduplication technology.

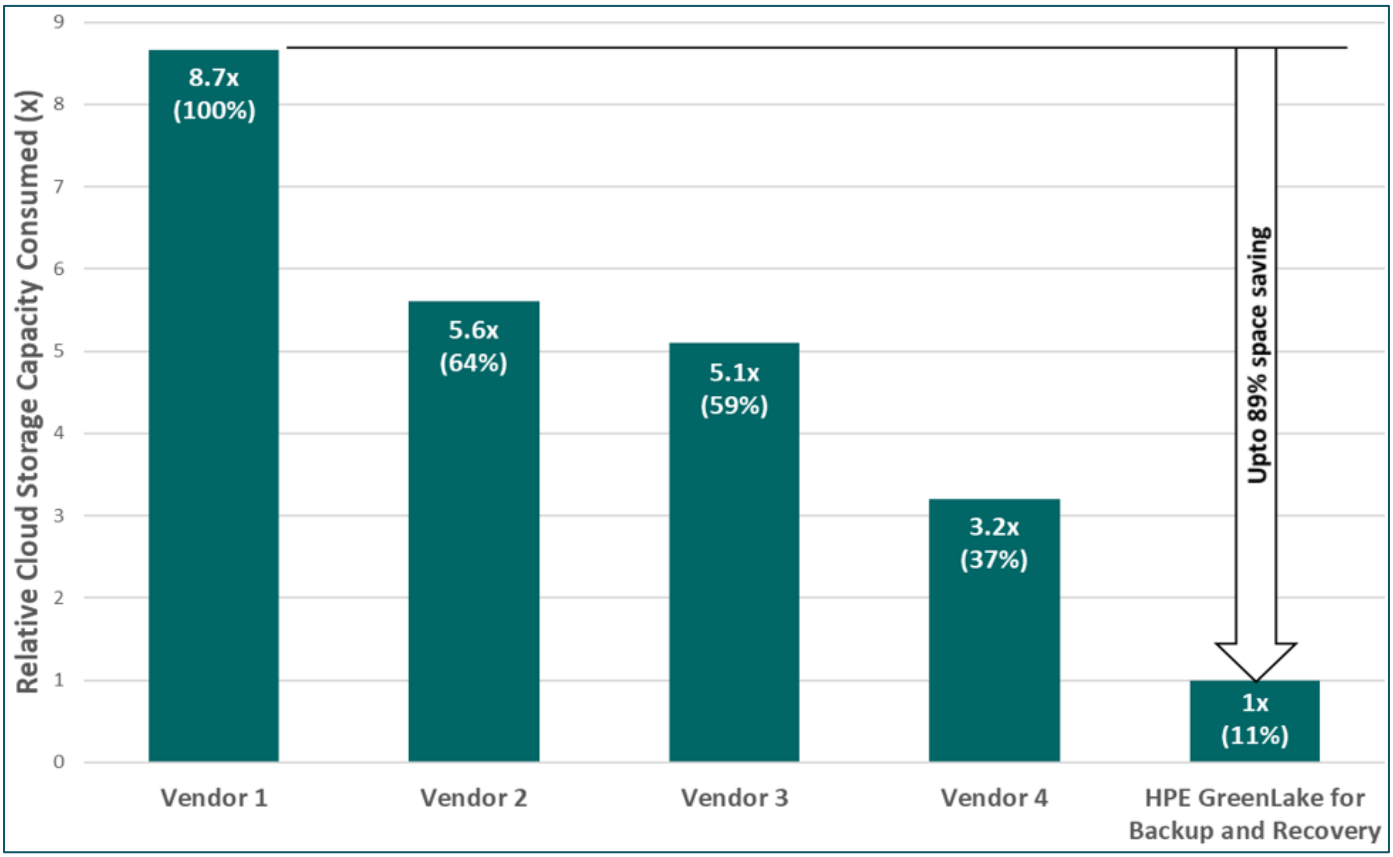

Figure 3. Internal testing showed that HPE GreenLake for Backup and Recovery uses up to 8x less storage capacity than four similar solutions

Note

These test results are based on HPE internal testing conducted in Q2–Q3, 2022 comparing HPE GreenLake for Backup and Recovery with other competing backup solutions. For details, see Appendix A: Storage efficiency test environment.

Secure by design

Ransomware protection is achieved with encrypted and immutable backups. All backups are encrypted for storage and transmission, which make backups unreadable to cyberattacks. Configurable backup data immutability and dual authorization prevent backup data from being modified or deleted by such hackers or bad actors. With dual authorization in place, destructive operations require escalation and approval.

User identity and access for all services are managed centrally by HPE GreenLake edge-to-cloud platform. Roles and permissions are enforced by using role-based access control (RBAC) to ensure that the correct level of access is given to each user. Roles are used to set

[1] May be subject to minimums or reserve capacity may apply.

user permissions, and scopes are used to limit the resources that each role can access. The account administrator has the option to assign predefined (or custom) roles and scopes provided by HPE GreenLake edge-to-cloud platform for each service. The predefined roles include a Backup and Recovery Administrator with full control, or a Backup and Recovery Operator with limited control.

Security is built into HPE GreenLake data services architecture. The Data Services Cloud Console (DSCC) runs in the cloud but is only a management control plane. The data collection is limited strictly to configuration and performance-related data. User data that resides in the protected resources is not exposed to or accessible from the DSCC.

For security details, see Data Services Cloud Console Security Guide.

HPE GreenLake for Backup and Recovery architecture

HPE GreenLake for Backup and Recovery is a cloud native service used for hybrid data protection of resources hosted on-premises or in the cloud. The high-level architecture is shown in Figure 4. For AWS data protection, no on-premises components are required. For VMware data protection, the service deploys and manages components in the cloud and on-premises.

For deployment details, see HPE GreenLake for Backup and Recovery Deployment Considerations Guide.

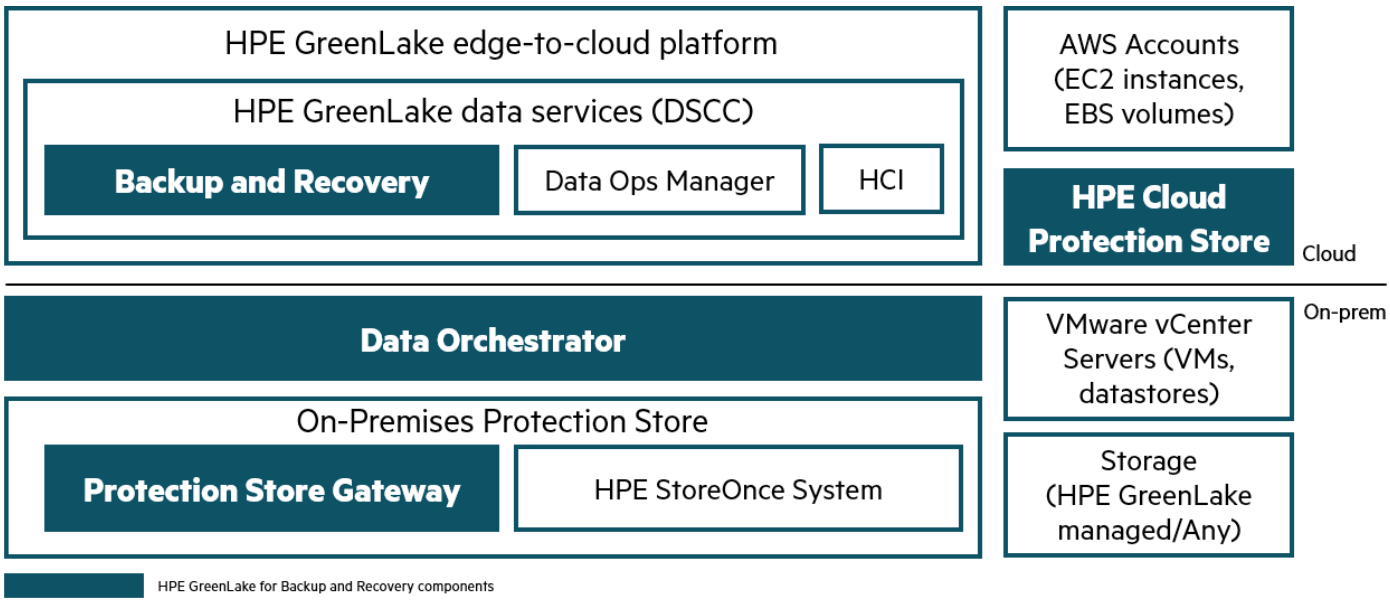

Figure 4. HPE GreenLake for Backup and Recovery provides hybrid data protection with components in the cloud or on-premises (VMware protection only)

Cloud components (AWS and VMware protection)

HPE GreenLake for Backup and Recovery is a cloud data service hosted in the HPE GreenLake edge-to-cloud platform. The service removes the management costs of traditional backup and recovery applications and associated storage devices. Multiple sites can be managed from a single cloud console and integration with other HPE cloud services provides a common usage and support experience.

HPE GreenLake for Backup and Recovery components (cloud)

Backup and Recovery

Backup and Recovery is the HPE GreenLake data service, which is the focus of this paper. The service is flexibly built using a micro-services architecture that enables a high frequency of new features added into the service with zero effort by the user. This is a huge benefit relative to the planning and downtime required to update traditional backup and recovery software.

HPE Cloud Protection Store

HPE Cloud Protection Store is fully managed by Hewlett Packard Enterprise and provides the destination and capacity for cloud backups (per region, selected by the user). Usage is billed based on the capacity used after the cloud backups have been deduplicated and compressed. There are no additional (egress) charges for data recovery.

Other components (cloud)

AWS Accounts

AWS Accounts host the customer's EC2 and EBS resources being protected.

Data Ops Manager

Data Ops Manager is an HPE GreenLake data service that simplifies storage management at scale, with fleet-wide management and monitoring. HPE GreenLake-managed storage devices can be managed from this service without needing to use a separate device UI.

HCI

HCI is an HPE GreenLake data service that simplifies global lifecycle management of HCI infrastructure and virtualization resources. A tight integration with HPE GreenLake for Backup and Recovery provides consistent protection SLAs across the entire global infrastructure. This is achieved by adding Backup and Recovery global protection policies to HCI VM provisioning policies.

On-premises components (VMware protection only)

HPE GreenLake for Backup and Recovery provides two on-premises components running as VMs: Data Orchestrator (DO) and Protection Store Gateway (PSG). These components are entirely managed by the service and can be recovered using a guided recovery should they be corrupted or deleted. To further secure HPE GreenLake for Backup and Recovery, only the on-premises Data Orchestrator establishes a connection to the DSCC in the cloud. The Protection Store Gateway never connects to the DSCC, and for cloud protection, it only sends data over an encrypted link to the HPE Cloud Protection Store. The DO and PSG run a hardened Linux® OS that reduces the attack surface for malware.

HPE GreenLake for Backup and Recovery components (on-premises)

Data Orchestrator

Data Orchestrator manages service operations defined in the cloud. It establishes a mutual Transport Layer Security (mTLS) tunnel to the DSCC Backup and Recovery component in the cloud. This secure tunnel remains active to send northbound management-event data and southbound management requests. The on-premises network proxy or firewall must permit outbound connections only.

Protection Store Gateway (On-Premises Protection Store)

Protection Store Gateway processes backup data before it is sent to the On-Premises Protection Store and/or HPE Cloud Protection Store. It hosts a service-defined On-Premises Protection Store. Communication between the Protection Store Gateway and the console is through the secure Data Orchestrator. The Protection Store Gateway uses the HPE Catalyst engine to deduplicate, compress, and encrypt backup data before it is sent to the On-Premises Protection Store and the Cloud Protection Store. For security, the Protection Store Gateway is deployed on a hardened Linux virtual machine. Multiple Protection Store Gateways can be deployed to add scale and performance. The Protection Store Gateway deployment is flexible to accommodate different sized On-Premises Protection Stores and performance requirements. Sizing guidance is provided.

Other components (on-premises)

HPE StoreOnce system (On-Premises Protection Store)

An HPE StoreOnce system is an alternative to a service-defined On-Premises Protection Store hosted on a Protection Store Gateway. HPE StoreOnce systems are dedicated, high-efficiency backup appliances. The physical separation of the HPE StoreOnce appliance from the protected VMware environment reduces resource contention and allows scaling on-premises storage without the need to add more VMware storage. Backups to HPE StoreOnce appliances can run over iSCSI or Fibre Channel.

Using an HPE StoreOnce system for on-premises backup storage delivers the same unmatched space saving, encryption, and immutability features as the service-defined On-Premises Protection Stores. Any HPE StoreOnce appliance running HPE StoreOnce software 4.3.4 or newer can be connected to HPE GreenLake for Backup and Recovery. Once connected, the HPE StoreOnce appliance is used in Protection Policies the same as service-defined On-premises Protection Stores. An HPE StoreOnce appliance can be used by HPE GreenLake for Backup and Recovery and other backup software concurrently.

VMware vCenter Servers

VMware vCenter Servers® host the customer's VM and datastore resources being protected.

Storage (HPE GreenLake managed or any other)

HPE GreenLake-managed storage offers the best performance, using **Array-optimized** backups. This backup type uses fast and efficient array snapshots and has the least impact on production.

Any storage includes third-party and HPE storage systems not managed by HPE GreenLake, using **VMware CBT** backups. This backup type uses more overhead but supports more storage systems.

For more details and a list of supported storage, see
HPE GreenLake for Backup and Recovery QuickSpecs → "VMware Array-optimized and VMware CBT Backups" and "Storage System support for high performance array-optimized protection" sections.

HPE GreenLake for Backup and Recovery in use

This section illustrates the ease of use of HPE GreenLake for Backup and Recovery and how simple it is to configure Protection Groups and global Protection Policies to automate protection.

The example use case shown in this section is for a customer named "XYZ" that has empowered their developers and business units outside of IT to be able to deploy workloads on-premises and in AWS on-demand, while also backing-up these new workloads automatically. A global Protection Policy is used for consistent protection of both on-premises and cloud workloads (databases).

Note
HPE GreenLake for Backup and Recovery is a cloud native service that gets updated on a regular basis. There might be occasions when the latest user interface screens and names do not match the screenshots taken at the publication time of this paper.

HPE GreenLake for Backup and Recovery is launched from the Data Services Cloud Console, shown in Figure 5

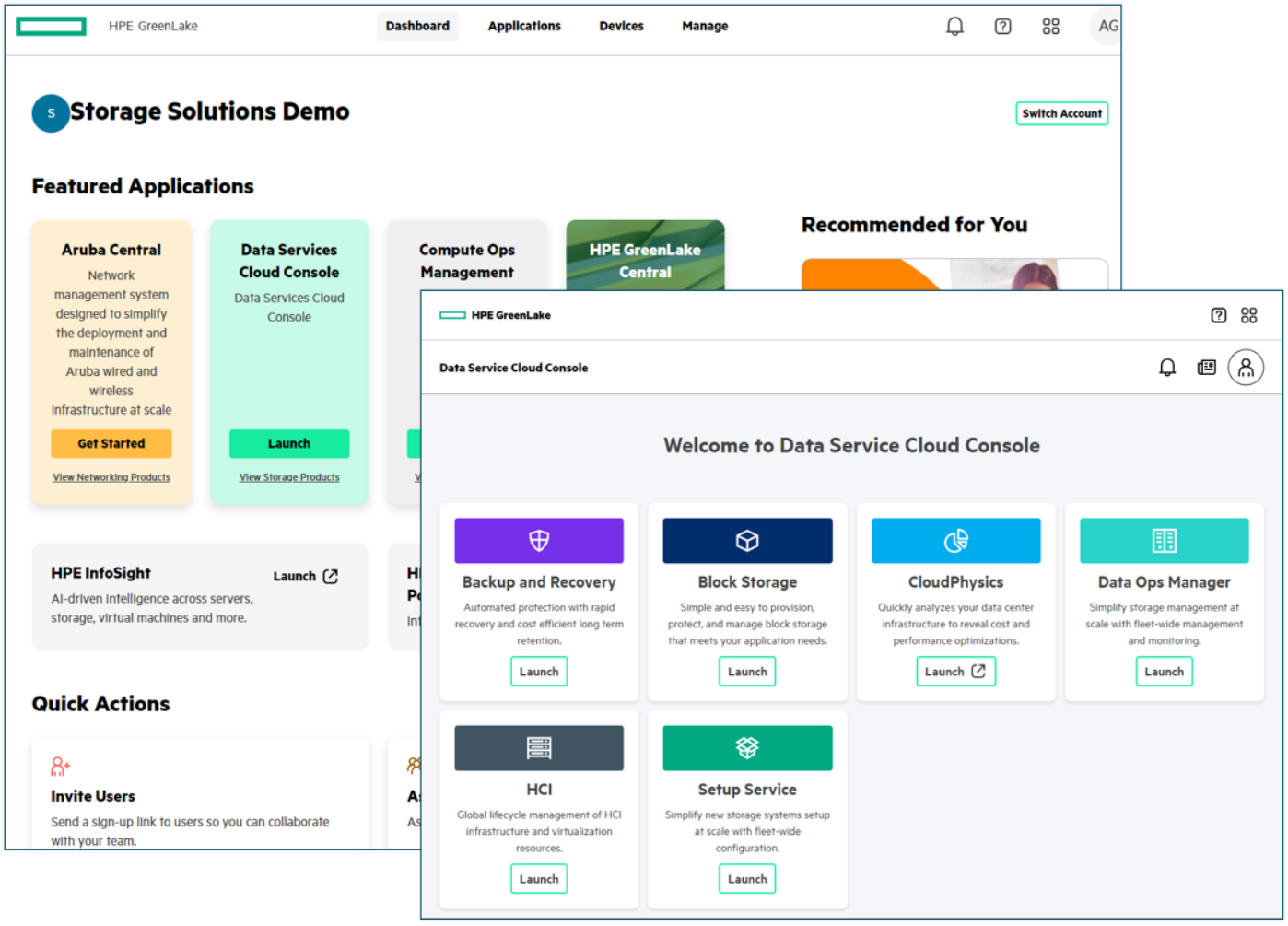

Figure 5. HPE GreenLake console streamlines user, device, and subscription management across all cloud services and is the launching point for data services

Operations

HPE GreenLake for Backup and Recovery makes it easy and reliable to meet recovery SLAs for on-premises VMware and AWS resources. The required recovery points and times are used to define Global Protection Policies. Automatic protection is achieved by defining Protection Groups for resources (VMware virtual machines/datastores, AWS instances/volumes) that require similar protection. The resources are then protected by assigning a policy to an individual resource or group.

It is as easy as: define a policy, then group, protect, and recover data.

HPE Catalyst backup data deduplication

Regardless of the backup data type chosen, the backups will be stored space-efficiently after being processed by the HPE Catalyst deduplication and compression technology. This technology is owned and developed by Hewlett Packard Enterprise. The initial application was originally developed for HPE StoreOnce systems. As a key component of HPE StoreOnce systems, the HPE Catalyst technology delivers significant backup data space savings for tens of thousands of users. This proven technology is built into HPE GreenLake for Backup and Recovery and delivers unmatched space efficiency.

Central to HPE Catalyst is the deduplication technology that efficiently stores the backup data as chunks. A chunk is stored only once. Duplicate chunks are replaced by a much smaller reference to the existing chunk. The process of chunking and matching is a highly optimized in-line process that has been refined over many releases. While this is conceptually like other deduplication processes, what sets it apart is the efficiency of the process to deliver high-performance backup—and the 4 KB chunk sizes (the smallest of any vendor) that provide unmatched space savings. These small chunk sizes used with HPE GreenLake for Backup and Recovery incremental-forever backups enable fast, efficient on-premises and cloud backups.

The use of HPE Catalyst adds to the security of HPE GreenLake for Backup and Recovery. As well as providing encryption and data immutability, the HPE Catalyst application programming interface (API) is obfuscated from malware, which provides additional protection against ransomware attacks. Malware is incapable of activating within an HPE Catalyst-based protection store because HPE Catalyst does not use standard operating system commands for its operations.

Architecturally, HPE Catalyst is built using an HPE Catalyst client and server. The client executes the chunking and communicates with the server to identify duplicate chunks. The client can be run close to the data and the server at the most appropriate location for the backup data storage. Only the unique chunks are transmitted to the server to optimize use of network bandwidth for high performance and to control costs. For VMware protection, the HPE Catalyst server runs in both the On-Premises Protection Store and the HPE Cloud Protection Store. For AWS protection, the HPE Catalyst server runs in the HPE Cloud Protection Store.

Consolidated inventory views

The protection status for all registered resources is shown in the HPE Backup and Recovery inventory views. These views of resource types for all locations provide global views that are not available from individual vCenter Server or AWS consoles. Global search and filtering makes it simple to find resources.

Figure 6 is an example of viewing the protection status for customer XYZ's cloud instances

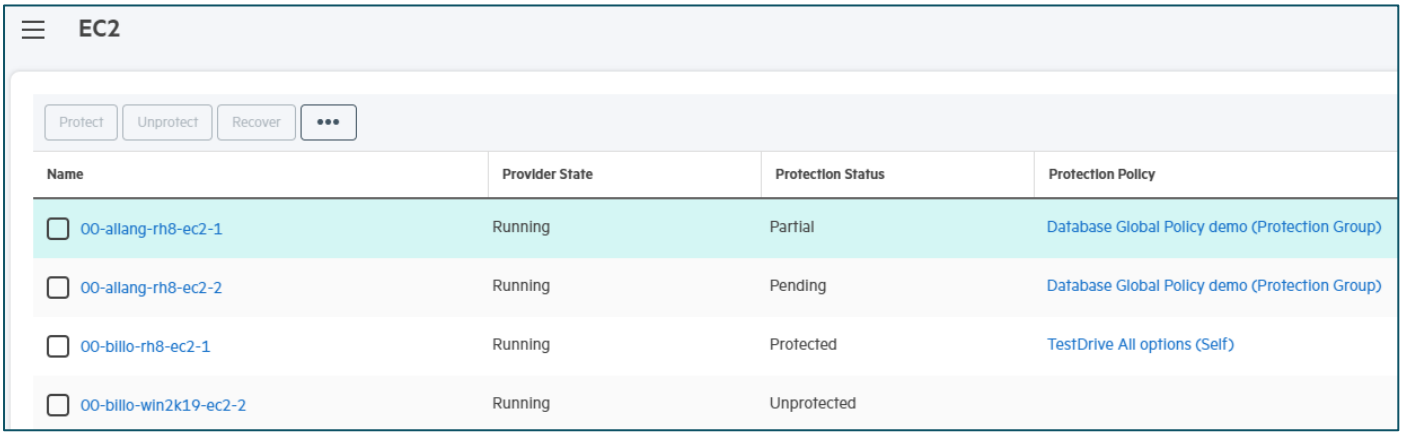

Figure 6. Inventory views based on resource type show a consolidated view of protection status for all resources

Protection Policies

Protection Policies allow you to set up global rules for defining the creation and lifecycle management of the recovery points, which can be applied to both on-premises and cloud resources. They include where to store the backups (Array Snapshot, On-Premises Protection Store, AWS Snapshot, or HPE Cloud Protection Store), how often backups are captured, how long backups are retained, and backup data immutability.

Note
The immutable option is only available for array snapshots if the array supports immutability. The immutability time is defined by the backup retention period. During this time, an immutable backup cannot be modified or deleted by anyone—not even by users who otherwise would have the appropriate permissions to delete backups.

The protection targets are flexible and can be configured with the following tiers:

Table 1. Protection tier options

Data source	Protection targets	Recovery Time Objectives (RTO)
On-Premises - HPE GreenLake managed storage	• Array Snapshot only • On-Premises Protection Store only • HPE Cloud Protection Store only • Array Snapshot, On-Premises Protection Store • Array Snapshot, On-Premises Protection Store, HPE Cloud Protection Store	• Instant Recovery • Rapid Recovery • Long-Term Retention • Instant Recovery, Rapid Recovery • Instant Recovery, Rapid Recovery, Long-Term Retention
On-Premises - Any storage (not managed by HPE GreenLake)	• On-Premises Protection Store only • HPE Cloud Protection Store only • On-Premises Protection Store, HPE Cloud Protection Store	• Rapid Recovery • Long-Term Retention • Rapid Recovery, Long-Term Retention
AWS	• AWS Snapshot only • HPE Cloud Protection Store only • AWS Snapshot, HPE Cloud Protection Store	• Rapid Recovery • Long-Term Retention • Rapid Recovery, Long-Term Retention

Figure 7 is an example of applying global rules for protecting customer XYZ's cloud and on-premises resources.

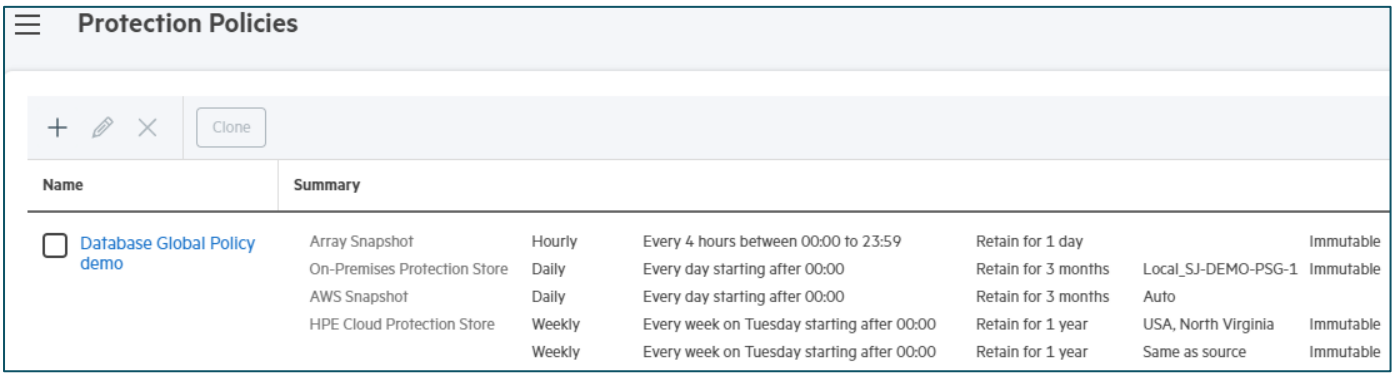

Figure 7. Protection Policies provide the protection service levels the organization needs

Protection Groups

Protection Groups define what is to be protected; these allow you to group a set of resources that require a similar level of protection. Automatic Protection Groups are used to automatically protect resources added to underlying entities (VMware folders/containers, AWS tags). Custom Protection Groups are used to individually select resources for protection.

Figure 8 is an example of grouping customer XYZ's cloud resources for automatic protection.

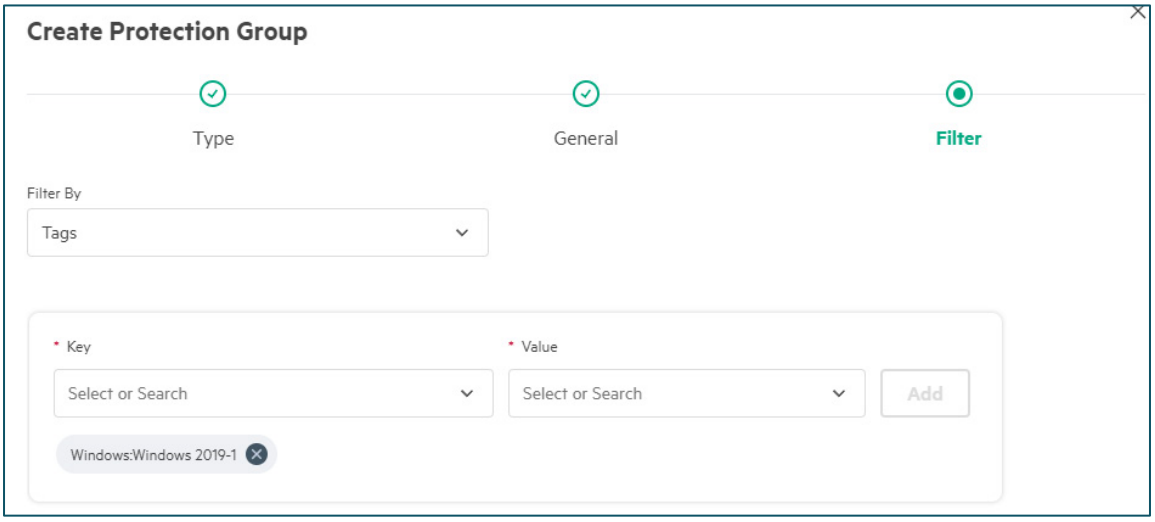

Figure 8. Automatic Protection Groups protect all resources, existing and new, assigned to an underlying entity (AWS tags in this example)

Protecting data

After defining Protection Policies and Protection Groups, you can easily start protection (at scale) by assigning a Protection Policy to a Protection Group or individual resource. Advanced Options, such as application consistency, can be applied here.

Figure 9 is an example of protecting customer XYZ's cloud and on-premises resources.

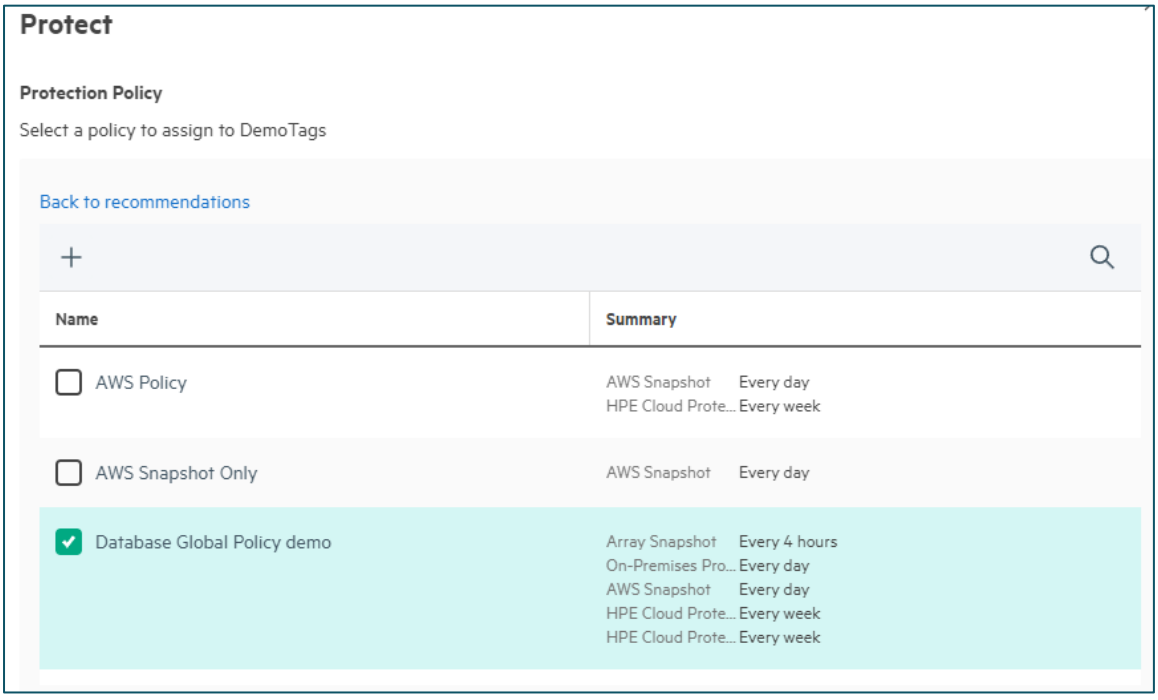

Figure 9. Applying a Protection Policy to a Protection Group or individual resource starts automated protection and provides consistent protection at scale

Recovering data

After data is protected, you can easily view and recover from recovery points. A recovery point is a snapshot or backup of data at a point in time. Data can be recovered to the original or a new resource using any recovery point. For VMware VMs, there is also an option to recover individual virtual disks. The recovery points continue to exist even after a resource is deleted from the inventory. A consolidated view of all

recovery points per-protected resource allows you to easily choose where and when to recover from. VMware resources are recoverable from an Array Snapshot, On-premises Protection Store, or HPE Cloud Protection Store. AWS resources are recoverable from an AWS Snapshot, Staging Snapshot, or HPE Cloud Protection Store.

Figure 10 is an example of recovering customer XYZ's cloud instances

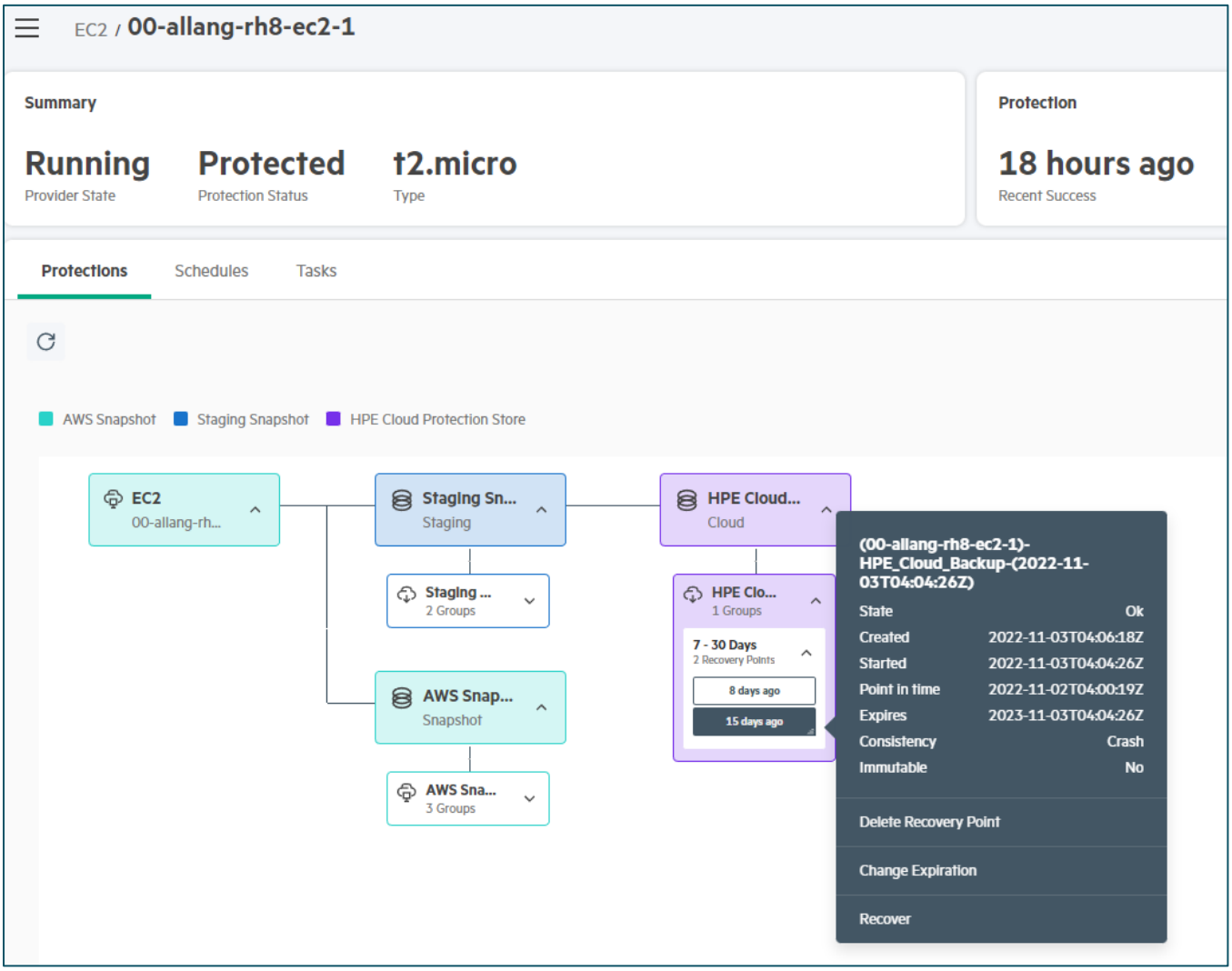

Figure 10. Choose a recovery point to recover from the desired backup location and point in time

Summary

HPE GreenLake for Backup and Recovery is the modern way to protect your on-premises VMware and AWS workloads. This cloud native service, available from the HPE GreenLake edge-to-cloud platform, removes the complexity and costs of deploying traditional backup and recovery software. HPE GreenLake for Backup and Recovery redefines backup and recovery with the simplicity and flexibility of software delivered and managed as a service. Global policy-based protection—configured from a single cloud console—enables you to set up and automate the protection of your virtual machines in a few simple steps. In the event of a data loss or error, all restore points are easily visualized so the data from the required recovery point can be quickly brought back online. The service is secure and efficient, delivering protection from ransomware and lower TCO. Security is built into the HPE GreenLake edge-to-cloud platform, plus the backup data is encrypted and optionally set as immutable, along with dual authorization required for backup deletion. The superior storage efficiency provided by HPE Catalyst technology reduces the amount of backup storage capacity needed by up to 8x compared to similar solutions.

The service can co-exist with other backup software for evaluations, and you can try the service—without any commitment for 90 days—at connect.hpe.com/HPE_Backup_and_Recovery_Trial.

Appendix A: Storage efficiency test environment

Testing was conducted at Hewlett Packard Enterprise in Q2–Q3, 2022 by protecting virtual machines with HPE GreenLake for Backup and Recovery and four other vendors. The testing was designed to represent a typical virtual user environment and protection schedules. The testing used capacity information as reported by the vendor and/or object storage provider.

The testing simulated 50 weekly backups using a typical production workload:

- Five Windows virtual machines
- 100 GB of different file data per virtual machine
- Virtual machines backed up using HPE GreenLake for Backup and Recovery and each of the four protection vendors
- File data in each virtual machine changed by a script between each backup iteration
- Process repeated 50 times with used cloud capacity measured for each vendor after each backup iteration

Resources, contacts, or additional links

HPE GreenLake for Backup and Recovery QuickSpecs
hpe.com/psnow/doc/a50004269enw?section=Product%20Documentation

HPE GreenLake for Backup and Recovery Deployment Considerations Guide
hpe.com/psnow/doc/a00128576enw

Application page for 90-day HPE GreenLake for Backup and Recovery evaluation
connect.hpe.com/HPE_Backup_and_Recovery_Trial

Data Services Cloud Console Security Guide
hpe.com/psnow/doc/a00113337enw

HPE GreenLake for Backup and Recovery
hpe.com/us/en/storage/data-protection-solutions/backup-recovery.html

Data Services on HPE GreenLake webpage
hpe.com/us/en/storage/data-services-cloud-console.html

HPE GreenLake for hyperconverged
hpe.com/us/en/greenlake/hyperconverged-infrastructure.html

HPE GreenLake for Backup and Recovery protecting dHCI
community.hpe.com/t5/around-the-storage-block/unleash-the-power-of-hpe-backup-and-recovery-service-to/ba-p/7177406#.Y5oGlnbML-g

HPE StoreOnce website
hpe.com/us/en/storage/storeonce.html

Learn more at

hpe.com/us/en/storage/data-protection-solutions/backup-recovery.html

a00123325enw, Rev. 1

Hewlett Packard Enterprise