

Integrating data protection for VMware Infrastructure



Considerations for efficient virtual environment protection

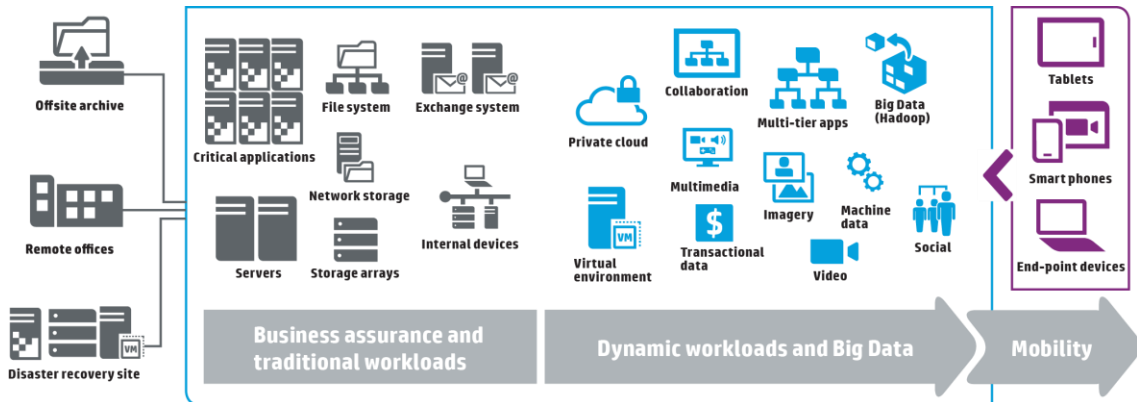
Table of contents

Executive summary	2
Data protection considerations in increasingly virtualized data centers	2
Consideration 1: agent, agentless, or hardware-assisted virtual machine backups	2
Consideration 2: a separate or point-based solution for virtual machine protection or a unified data protection solution	4
Consideration 3: squeezing every last bit to maximize capacity without taking a byte out of performance using deduplication	5
Consideration 4: creating a backup and recovery strategy that is efficiently tiered	7
Consideration 5: building for the future based on knowledge of the present with analytics	11
About HP Data Protector	14
Conclusion	14

Executive summary

The reality of today’s modern data center is that it’s peppered with physical and virtual deployments, technology silos, aging applications that should be retired but can’t for various reasons, and often supported by minimal staff (figure 1). Add to these challenges the marginally growing, flat, or declining IT budgets, and it stands to reason that greater adoption of virtualization allows organizations to lower capital, optimize the data center with the economies of scale that virtualization delivers, and build an automation foundation that results in greater business agility.

Figure 1. Traditional data centers conjoined with dynamic workloads and edge devices



As the perspective of IT changes from one of supporting technology assets to one of increasing service, the demands placed on IT will increase. Business managers will inevitably request that new services and the value of virtualization will start to be measured in terms of time to market and responsiveness. If services can be delivered faster, the business can respond faster to market and competitive pressures. The result is increased revenue and lower costs. While these new approaches help IT efficiently meet the accelerated demands of their business and respond quicker to external pressures, it does increase the challenges to IT beyond service delivery: ensuring the right protection of data and applications in these mixed and/or highly virtualized environments.

Regardless of whether organizations define data protection based on service-level agreements (SLAs) or service-level objectives (SLOs), balancing the agreed upon service level with the right data protection infrastructure and software is one of the most difficult challenges that IT organizations face today. This challenge becomes amplified as more and more data center and mission-critical applications move to virtual environments. This results in a number of considerations that must be addressed in order to achieve the economies of scale that virtualization provides, including identifying the right degree of protection to ensure the supportability of the data, the applications, and the business.

This technical white paper discusses how HP Data Protector enables you to address these data protection considerations with integrated backup and recovery for VMware® Infrastructure.

Data protection considerations in increasingly virtualized data centers

While the accumulated value organizations can achieve by adopting virtualization can be substantial, blindly adopting technology without considering its impacts to your existing infrastructure capabilities can result in long-term headaches and difficulties. Regardless of your existing data protection infrastructure and software choices, when investigating the right approaches to take for protecting virtual environments, the following considerations should be addressed:

Consideration 1: agent, agentless, or hardware-assisted virtual machine backups

Successful virtual infrastructure protection is a constant balance between the criticality of the application or workload running on it and the resources it requires for optimum operation. While traditional server protection approaches can work in virtual environments, the right approach is to select a solution that uniquely address the needs of both virtual and physical server environments. However, knowing whether to run an agent in the virtual machine, use an agentless approach, or some other data protection strategy can be a daunting task for administrators. The following discusses each option, per figure 2., and how HP Data Protector provides organizations with all three backup operation approaches to meet the most common use cases when protecting virtual machines.

- **Agent-based virtual machine backup software:** This is the traditional way of performing file and full system backups. In this scenario, a backup agent (a software service that handles all processes that include handling I/O, deduplication, cataloging backup sets, etc.) is installed in the VM’s guest operating system. A backup specification or policy is defined in HP Data Protector, and during the defined backup window, the agent copies files to the desired

backup server/backup target. This solution offers simple file-level restore functionality, and provides application-level backup and restore operations not otherwise available via file system backup operations. Although this method is proven to be very reliable, by default it does not benefit from any of the advanced features offered by an enterprise-level virtualization platform. Hypervisors are architected to utilize the full capabilities of the physical server to accommodate multiple virtual instances running in tandem. An implementation like this is the same as for a physical server environment and can result in adverse VM performance due to resource competition with peer-level VMs running on the same host. There are instances where an agent-based approach is the right one, for example an OS or application that does not support the Microsoft® Volume Shadow Copy Service (VSS), point-in-time recovery goals, backing up multi-node applications such as exchange, etc. However, this use-case should be carefully considered as it complicates both the management of VMs and backup operations and may impact the licensing costs as VMs proliferate.

– **Benefits:**

- Proven technology and approach.
- Protects investment in current backup software solution and administrator knowledge.
- Agents are normally tailored to specific applications.

– **Risks:**

- Extended backup windows or operations can affect peer-level VMs.
- Agents are not specifically geared to the virtual environment or the sharing of VM resources and must be actively monitored.

- **Agentless virtual machine backup software:** This is the most common way of performing virtual machine backups. In this scenario, the backup software solution coordinates the backup and recovery options with the hypervisor via the hypervisor's specified API. For VMware environments, VADP is the interface that allows an HP Data Protector server (physical or virtual) to perform VM backups without requiring agents in the VM. As a virtualization-aware backup application, HP Data Protector coordinates the backup operation by coordinating VM preparation and the creation of snapshot images to be used as the source of the backup. The use of software-based snapshots enables the backup operation to be performed without requiring downtime of the VMs running on the hypervisor host. As a result, agentless backups are non-disruptive to running VMs, allowing the operation to be run at any time without the need for extended backup windows and adverse impacts to running applications. Additional resource savings can be gained when offloading backup-related operations (such as deduplication, encryption, etc.) to the backup server and/or backup target devices.

– **Benefits:**

- Uses hypervisor supplied APIs to coordinate the backup and recovery operation.
- Uses space-efficient software-based snapshots as the source for the backup operation.
- Backup can be done across a SAN to minimize network traffic.
- VM performance impacts are greatly reduced.

– **Risks:**

- Relies on Microsoft Volume Shadow Copy Service for OS and application consistency.
- Requires a mount point to establish backup traffic connection to a proxy server.
- Consumes hypervisor CPU, memory, and networking to complete backup operation.

- **Agentless, hardware assisted, virtual machine backup:** With this approach, the benefits of the agentless backup methodology, previously described, is augmented to remove the compute, networking, and memory consumption levied on the hypervisor host to complete the backup operation. In this scenario, HP Data Protector coordinates the backup operation via VADP, but alleviates the necessity to involve the hypervisor in the transfer of the backup set to the backup target. Instead, HP Data Protector locates the snapshots on disk and brokers the connection between the primary storage array and the backup target to complete the backup operation (similar to the SAN transport method). This approach offloads the I/O processing, and data movement from the hypervisor host allows it to continue to provide maximum physical resource support to the VMs it is hosting.

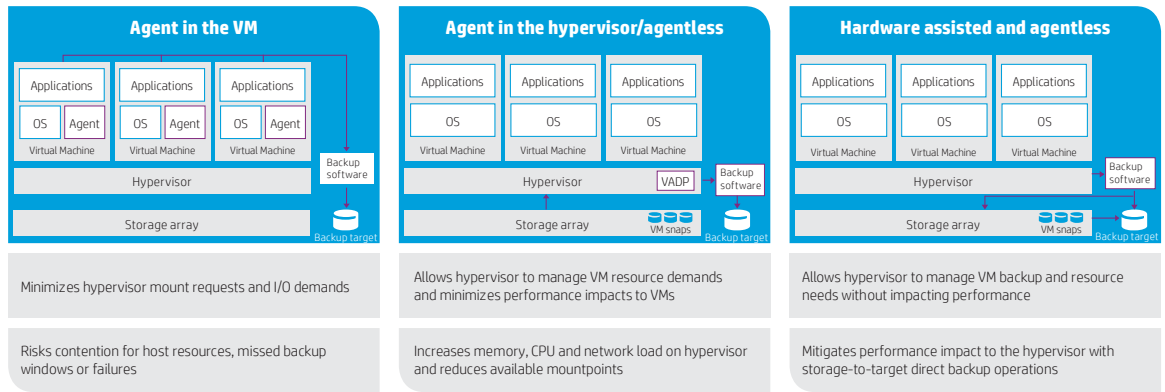
– **Benefits:**

- Offloads the processing and movement of backup data from the hypervisor.
- Uses hardware-based snapshots (storage array) as the source for the backup operation.
- VM and hypervisor performance impacts are minimized.

– **Risks:**

- Relies on specifically supported storage arrays (see the appropriate [hardware compatibility list](#) for more details).

Figure 2. Modern data protection solutions must offer the choice on where backup operations occur



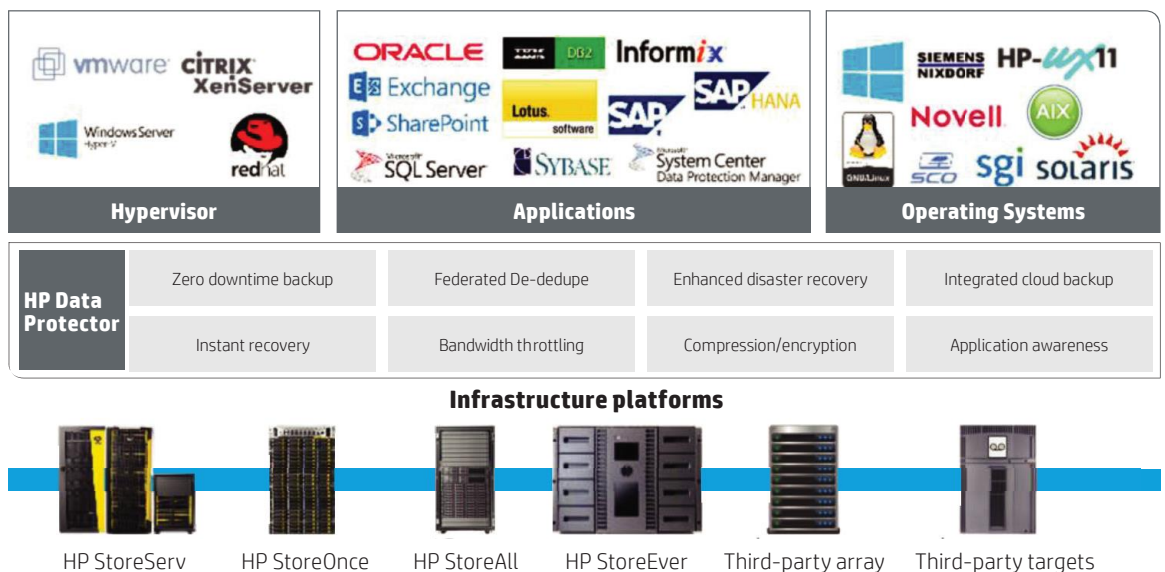
Consideration 2: a separate or point-based solution for virtual machine protection or a unified data protection solution

Data center growth can bring with it organically grown processes, point-based solutions, and purpose-built appliances—these factors result in technology silos that often result in long-term costs related to their management and specialized knowledge required by the administrative staff. However, data center maturity is the goal of most organizations. As such, the deployment of a single, full-featured solution that can address the data protection needs of varying workloads, physical and virtual environments, remote sites and is centrally managed, is the ideal approach.

When organizations operate both physical and virtual environments, the responsibilities for how data is protected is often split between IT and VM administrative teams. It is not uncommon for each team to have their desired backup software solution, and too often, these solutions are never the same product. A dual data protection software approach adds cost and complexity from both a CAPEX and OPEX perspective. Capital expense is required in the form of duplications in software licensing, support costs, annual maintenance fees, duplicate storage costs (space allocation and/or infrastructure), and separate administrative staff. Operational expenses are tied into process complexity, scripting, and the need to maintain specialized knowledge and training within the IT staff.

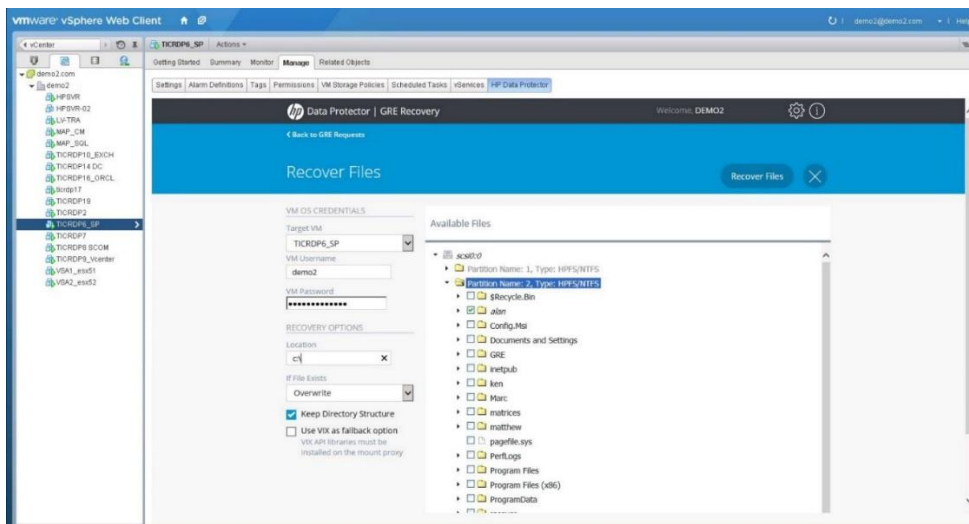
HP Data Protector is a unified and comprehensive backup and recovery solution for physical and virtual environments (figure 3). With HP Data Protector, administrative staff benefit from a solution that provides end-to-end visibility and control over the data protection process with centralized management, alerting, reporting, and monitoring. With support for the four major hypervisors, 14 of the most common mission-critical applications, 11 of the industry-leading operating systems, and HP Data Protector integrated array-based storage features, organizations benefit from a single backup software solution that is capable of addressing the entire data center, not just a single use case. With respect to VMware, the capabilities of HP Data Protector can be extended into the VMware management tools with its Granular Recovery Extension.

Figure 3. HP Data Protector—a single enterprise solution for virtual and physical server protection



The HP Data Protector GRE (figure 4) is a VMware vCenter plug-in, or extension, that enables VMware administrators to perform the various recovery operations using a data protection infrastructure that supports the entire data center workloads, not just VM protection. Using the GRE, VMware administrators can recover VMs, virtual disks, and single items within the same management console they use to manage their daily VM-related tasks, all with the need to access HP Data Protector.

Figure 4. HP Data Protector Granular Recovery Extension for VMware



Backup and recovery is not limited to VMs—HP Data Protector also supports those organizations using VMware vCloud Director. With HP Data Protector, VM administrators can backup vCloud Director vApps and VMs. Additionally, as new VMs are added to both vCloud Director and the virtual environment, they benefit from backup policy inheritance or one-touch-protection. This means that VMs inherit the backup and recovery policy of their parent container, and for organizations that create performance tiers within the virtual infrastructure, these same tiers can extend beyond the performance parameters (CPU, networking, and storage) to include how the VM is being protected.

Consideration 3: squeezing every last bit to maximize capacity without taking a byte out of performance using deduplication




Data redundancy is a leading contributor to the unabated data growth that plagues data centers and eats away IT budgets. Virtualization compounds this problem as more and more VMs are deployed across physical hosts; efficiently storing the huge volumes of digital data becomes a challenge, especially as those VMs are filled with duplicated system and user files, operating systems, and application instances. Primary storage and backup software vendors use deduplication technologies to identify unique blocks of data within the target data set(s) to reduce the storage space required, I/O transfers, and networking congestion within the data center, at remote sites and at disaster recovery locations.

Regardless of the protected workload, HP Data Protector provides deduplication options that can be deployed at the application source, the backup server, and/or at the backup target (figure 5). While each of these deployment options have their tradeoffs, deduplication should be a given—the question should simply be where should one execute the deduplication process.

- **Client/application/source-side deduplication**—identifies the unique blocks on the machine running the application and/or containing the information that is being protected. It is an ideal backup software solution for environments having a low daily rate of change in the data being protected at remote sites, on servers with enough physical resources to meet the needs of both the application(s) and the deduplication process, and when running the backup agent on the hypervisor host (resources permitting).
 - **Benefits:**
 - Backup windows across high latency and/or low bandwidth networks more likely to complete on time.
 - Reduced storage required for backups.
 - Backup and replication SLAs are more achievable due to faster backup operations.
 - **Risks:**
 - Resource contention on the client is between deduplication process and application(s).
 - May require a client-side agent (see agent-based VM deployment consideration previously mentioned).

- **Backup server deduplication**—offloads the identification and processing of unique blocks to the backup server that resides between the client and the target. It is an ideal intermediary solution for virtual machines that are resource constrained but have sufficient network bandwidth to support the full backup set transfers, and for backup servers with predictable periods of low utilization or gaps in the backup window to support the deduplication process.
 - **Benefits:**
 - Reduces the resource load on the client.
 - Network bandwidth requirements are reduced during the backup operation between the backup server and the backup target.
 - Beneficial to bandwidth constrained environments.
 - **Risks:**
 - Resource contention on the backup server may occur.
 - Does not reduce the bandwidth requirements between the client and backup server.
- **Target-side deduplication**—relies on a purpose-built backup appliance to identify the unique blocks, and when using an integrated backup infrastructure with the primary storage array, backup software, and target, it often results in a highly performant backup and recovery approach for virtual environments. It is an ideal solution for virtual environments as they are generally easy to deploy and use, and is often backup software agnostic.
 - **Benefits:**
 - Offloads the deduplication processing to a purpose-built backup target.
 - Reduces resource overhead on the client/VM and the backup server.
 - **Risks:**
 - Does not reduce the backup window or bandwidth requirements between the source and the target.
 - Full backups can take longer to complete due to the lack of deduplication at earlier points in the architecture (incremental backups or performing the deduplication on the client will improve).

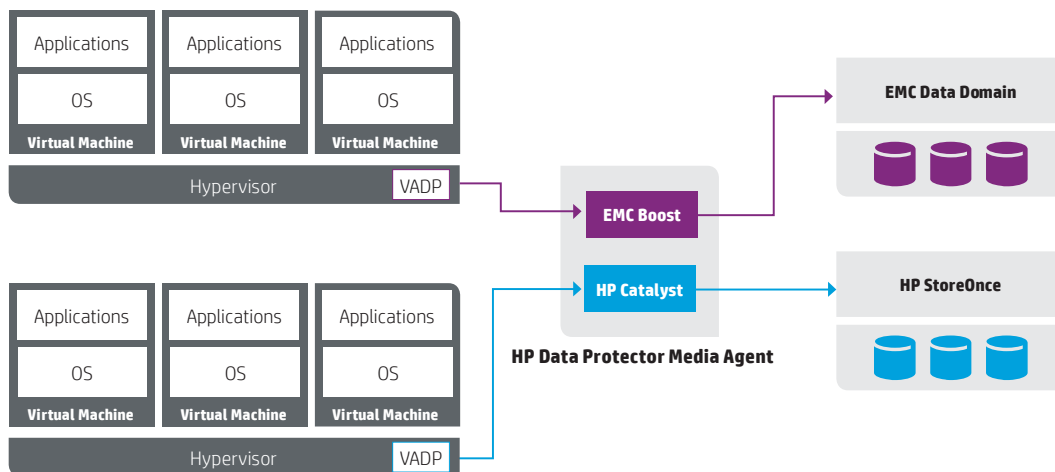
Figure 5. HP Data Protector deduplication options

Deduplicate at the application source	Deduplicate at the backup server	Deduplicate at the backup target
		
Minimizes network traffic by processing and sending only unique blocks	Minimizes load on application server	Highly performant solution responsible for both backup set storage and deduplication
Risks contention for application host resources and can effect other workloads on the host	Can impact backup performance; requires full backup set transfer from the source	Requires full backup payload transmission from source through backup server
Best used when minimal application loads is predictable and network congestion is a concern	Best used with resource intensive applications, bandwidth and backup operations can support it	Best used when application and backup workloads are resource intensive

Today, only HP can deliver the same block-level deduplication algorithm that scales from a software-only footprint to a scale-out, purpose-built-appliance, enabling efficient data replication and storage. The combined HP Data Protector software and HP StoreOnce Backup appliance solution enables federated deduplication, which gives you the flexibility of tri-location deduplication described above. Deduplication is federated, meaning that the deduplication occurs once—at the location specified—and the space saving ratios achieved are consistent wherever the deduplication occurs. By integrating with HP StoreOnce Catalyst, HP Data Protector can centrally manage both low-bandwidth data replication across multiple virtual data centers for disaster recovery and the deduplication process. Additionally, because HP StoreOnce is application-agnostic, it provides deduplication support for a wide array of mission-critical applications.

HP Data Protector is the only backup and recovery solution in the market today that places real-time operational analytics at the core of the data protection process—a value discussed later in this paper. With respect to deduplication, HP Data Protector media agent (responsible for read/write actions to back up media) analyzes the backup path between the client and the target to identify the most optimum deduplication algorithm to use: HP StoreOnce Catalyst or EMC Data Domain Boost (figure 6) based on the configured destination backup target.

Figure 6. HP Data Protector target-aware deduplication



The understanding of the source-to-target backup and recovery relationship enables organizations, with heterogeneous backup targets, to benefit from the same deduplication placement options (client, server, and target) without running multiple agents. In addition to the integration with HP StoreOnce, HP Data Protector media agent is integrated with EMC’s Data Domain Boost API. As shown in figure 6, the HP Data Protector media agent (the component responsible for reading and writing the backup set to the target) is aware of the capabilities of the target device and the source-to-target relationship, which means it will use the appropriate deduplication algorithm during the data backup and recovery process.

Because the deduplication deployment options are similar between HP StoreOnce and EMC Data Domain, it is important to note that the pan-HP relationship between HP Data Protector and HP StoreOnce results in a tighter integration, which means that wherever deduplication is applied in the backup process, no additional downstream deduplication processing, or rehydration is required. In contrast, with the HP Data Protector and EMC Data Domain integration, additional processing is required once the backup set is sent to the backup appliance to best accommodate the backup set.

It stands to reason that the question is not whether to deduplicate VMs, but where the deduplication should occur. Because deduplication reduces the disk space required to store backup data sets longer without impacting backup performance, retaining more backup data on disk and at various points of recovery enables greater data accessibility for rapid restores. Remember that deduplication ratios are strongly influenced by two factors—data change rate and retention periods of the data on the deduplication appliance. Low data change rates and data retained for longer periods of time often yield greater deduplication ratios. Key recommendations for VMware VM backups include:

- For optimal storage use: perform sequential VM backups where possible to improve and maintain the highest deduplication ratios and achieve the most optimal storage utilization over time
- For optimal network throughput: back up multiple VMs simultaneously to improve backup throughput performance
- Configure VMware Change Block Tracking to reduce the amount of data in incremental backups
- Use larger backup block sizes for faster backup throughput performance and better deduplication ratios
- Configure a weekly full and daily incremental backup schedule to reduce the amount of end-to-end data and decrease the time required to run daily backups
- Configure disk-to-disk replication to seamlessly replicate all VM backup images to an appliance in a remote facility for easier disaster recovery

Consideration 4: creating a backup and recovery strategy that is efficiently tiered

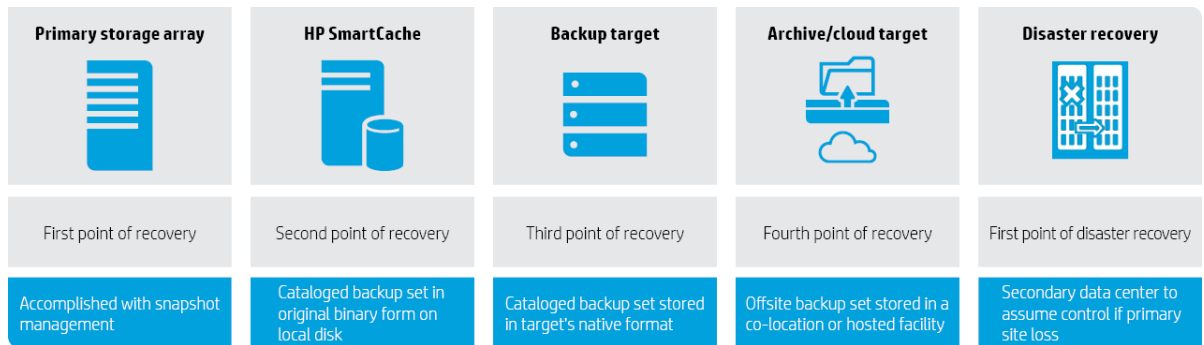
The dynamic nature of virtual environments is a cause to rethink the way IT architectures are designed and used when considering the storage and data protection needs of virtual machines. Equally challenging is the consistent request to balance performance with reducing costs. Of course emphasis is always placed on cost reduction and forgiveness never offered for poor performance. Balancing this does not have to be a tradeoff, especially if you design your backup and recovery strategy around data that has been classified and aligned to a corresponding SLA, and the physical storage infrastructure is tiered to meeting the recovery point objectives for the business and/or each VM. Designing a tiered backup and recovery infrastructure begins with classifying the workload and aligning it to the business needs that drive the performance requirements, which in turn guides the data protection tier configurations.

Data, application, and VM classification does not require expensive tools or services, and most organizations should be able to characterize their data and map it to the appropriate storage tier that matches the performance needs within desired cost windows. At a minimum, organizations should use a backup and recovery classification for their VMs with a four-tiered approach:

- **Mission critical**—VMs and VM workloads that are the most valuable to the business with high rates of access and change. These workloads are highly performant, require high availability, and are bound to SLAs that tolerate little to no downtime.
- **Operational**—VMs and VM workloads that are used daily are important to the business, but the viability of the business is not at risk if some downtime is experienced. These workloads tend to require reasonable performance and availability, have balanced data change rates, and can tolerate no less than six- to eight-hour recovery windows.
- **Archival**—VM and VM workloads that are not accessed regularly and often tied to compliance rules/regulations and may consist of fixed content with very low rates of change.
- **Offline**—VM and VM workloads that are no longer accessed, but fall under specific retention policies that require they be kept for a specific period of time should they be needed.

Classifying your VM and VM workloads allows for better alignment with the infrastructure tier based on the recovery-point and recovery-time objectives of the workload and the capabilities of the tier. In addition to improving the speed of backup or recovery, having multiple data protection tiers, as shown in figure 7, results in a balanced approach to mitigating the risk of data loss. But data backups should not just be in different locations, they should also be stored on different types of media to avoid system or media processing errors that may delay recovery operations.

Figure 7. HP Data Protector tiered protection options for VMware environments



With respect to VM and VM workloads, the backup solution must support the following points of recovery:

- **Primary storage array:** This approach relies on the storage array’s snapshot (a space-efficient point-in-time copy of the source data object or container) feature. This is a common scenario in backup and recovery because the snapshots are internal operations, fast, allow for quick restores, and easy to manage.
 - **Benefits:**
 - Low impact/non-disruptive to the production servers at the time of the backup.
 - Supports quick restores and easy management options.
 - Enterprise-class storage arrays offer higher snapshot frequencies before performance impacts are felt (generally speaking).
 - **Risks:**
 - Snapshots remain on the same array as the production volumes—loss of the array results in a loss of both the production volumes and their snapshots.
 - Not all snapshots are created equally—vendor implementation of snapshots varies and is effected differently with their use and degrees of integration.

HP Data Protector integrates with primary storage arrays to coordinate the creation, management, retention and deletion of snapshots used in the backup and recovery process—also known as Zero Downtime Backup (ZDB). This removes the need to take the production workload offline to perform the backup operation and uses associative APIs to address the integrity of the workload and perform backups that are application and operating system consistent. Additionally, ZDB enables administrators to use the Instant Recovery feature to perform granular restores from snapshots when they are available.

For VMware VMs, HP Data Protector snapshot integration with HP 3PAR results in fast, direct, single item restores, VM power on, and VM live migrate from the snapshot repository that is easily managed through the HP Data Protector application or using the HP Data Protector GRE vCenter plug-in.

- **Backup target/appliance:** This approach relies on the use of an “off array” appliance, often purpose built for backup and recovery, where the backup and recovery software transfers the backup set to an appliance-defined repository. Backup sets are cataloged and stored in formats that are native to the appliance and restores are normally managed via the backup and recovery software.

– **Benefits:**

- Integrated appliances remove capacity demands on the primary array and are easy to implement.
- Enterprise-class appliances offer storage efficiency features such as deduplication, compression, and optimization.
- Organizational information can be retained longer, and within the data center, allowing for longer data windows to recover information from.

– **Risks:**

- Without the level of integration to the backup software to manage backup and recovery centrally, appliances can create management and deployment complexities.
- Appliances, like primary arrays, are capacity bound and can result in scalability limitations without a properly defined data retention policy in place.
- Infrastructure performance is at risk because data is being copied from the source to the appliance (a risk that can be minimized with deduplication, compression, and WAN optimization features, if available).

HP Data Protector supports a myriad of appliances as backup targets (see the [HP Hardware Compatibility List](#) for details) with varying levels of integration. Integration with HP StoreOnce results in centralized management, federated deduplication, compression, and WAN optimization with the ability to provide HP StoreOnce Catalyst, NAS, and Virtual Tape Library (VTL) repositories via IP, FC, iSCSI, and FCoE network protocols. Additionally, HP Data Protector provides integrations with EMC Data Domain appliances via their Boost API to centrally manage the backup process with deduplication technologies and the replication of backup sets between Data Domain Boost appliances.

With respect to VMware VMs, HP Data Protector integration with HP StoreOnce and EMC Data Domain appliances delivers a non-cached solution (backup objects must be restored from the backup repository before the recovery operation can be completed) for VM backups based on snapshots. The result is a centrally managed backup and recovery process for VMs with granular restore functionality using the HP Data Protector GRE vCenter plug-in.

- **Archive targets:** This approach relies on the use of a tape media appliance or with a disk-to-disk appliance that provides virtual tape library support, as the destination for a backup set. As with the backup appliance, backup sets are cataloged and stored in formats that are native to the archive target and often are compressed to reclaim as much physical media capacity possible.

– **Benefits:**

- Low energy consumption and high capacity/data density results in cost containment and improved ROI for long-term retention.
- Reduces capacity demands on more expensive storage tiers while still providing managed access to the data.
- Risk reduction is possible by allowing organizations to retain access to information for longer periods of time on a low-cost storage tier.
- Portable backups resulting in easier restore operations to another environment for DR purposes.

– **Risks:**

- Data loss due to physical media wear (an issue with traditional archiving media).
- Manual processes are often involved in the management of archive media, processes that should be automated to reduce error and improve speed.
- Typically, a longer recovery window results in restoring from archive.

Similar to the disk-based backup appliances, HP Data Protector supports a myriad of archive targets and media formats (see the [HP Hardware Compatibility List](#) for details) with varying levels of integrations. Integration with HP StoreOnce results in a centralized mechanism to coordinate the backup process and copying the backup set to the archive target. For organizations wanting to eliminate the use of tape, HP StoreOnce offers Virtual Tape Library support that can be configured within HP Data Protector for on-disk archiving of information.

- **Cloud targets:** This approach relies on a hosted, co-hosted, private, or hybrid infrastructure using industry-standard operating environments such as OpenStack. Hosted solutions often provide a low-cost/low-service level approach, whereas private clouds tend to be an extension of the on-premises data protection operation.

- **Benefits:**

- Outsource the physical infrastructure and management without impacting human resources.
- Adds an additional backup tier that is commonly off-premise, thereby broadening the backup and recovery strategy and addressing disaster recovery considerations.
- Flexible payment options often reduces impacts to the IT budget when considering purchases of additional hardware, and the OPEX costs to house and operate them.

- **Risks:**

- Troubleshooting may be complicated depending upon the number of vendors and technology layers that must be negotiated.
- Hosting vendors are subject to acquisition (possibly by a competitor) and the financial state of the market (i.e., resulting in a permanent shutdown).
- Data access likely abstracted through gateways making flexibility difficult, non-integrated with existing backup and recovery tools, and potentially troublesome to reclaim the backup data and its history if reclaiming backup sets to host within the data center.

HP Data Protector provides integrated support with HP Helion and Amazon Web Services (AWS). Integrations with HP Cloud (public), or HP Helion (private cloud), provide VM administrators with a fast and reliable cloud device, or target, to back up and restore VMs. Backup sets can be copied to the cloud target from HP Data Protector's file library or the HP StoreOnce library with flexible policies to enable hybrid recovery models (on-premise and/or remote repositories). Also, HP Data Protector integrates with user-managed keys and cloud authentication services to ensure secure and compliant transfers of information to/from defined cloud repositories. With respect to AWS, HP Data Protector supports the use of this offsite cloud target through Amazon's Gateway-Stored Volumes and Gateway-Cached volumes for backup and recovery operations.

With respect to VMware, organizations using VMware's vCloud Director (vCD) solution benefit from HP Data Protector's ability to define protection policies (full, incremental, and differential) for VMware Virtual Data Centers, independent of the underlying physical setup, with automatic protection policy inheritance for new vApps and VMs. Additionally, HP Data Protector can be used to back up and recover the vCD setup and configuration.

- **Disaster recovery sites:** This approach relies on the use of an offsite location to support ongoing operations in the event of an unplanned outage or disaster that effects the primary data center. Disaster recovery sites are often one of the following:

- *Cold sites:* Data center space with power and networking that is ready whenever it is needed, but does require the relocation of hardware from the primary site to address the outage.
- *Warm sites:* Similar to a cold site with the exception that the organization pre-installs the hardware and configures connectivity before a disaster. Should a disaster occur, the organization only has to install the software and data necessary to support continued operation.
- *Hot sites:* An off-premises replication of the production data center (or some specific parts) to support an immediate failover in the event of a disaster with limited manual intervention to maintain the operational state of the business.

Organizations choose the type of disaster recovery site based on budget, sensitivity of the data, and amount of risk/downtime that can be tolerated when responding to an unplanned event.

- **Benefits:**

- Offsite location (especially hot sites) improves business resiliency.
- Reduces primary data center infrastructure demands for storing backup sets.
- Disaster recovery site can be used as a secondary primary data center for lower priority workloads.

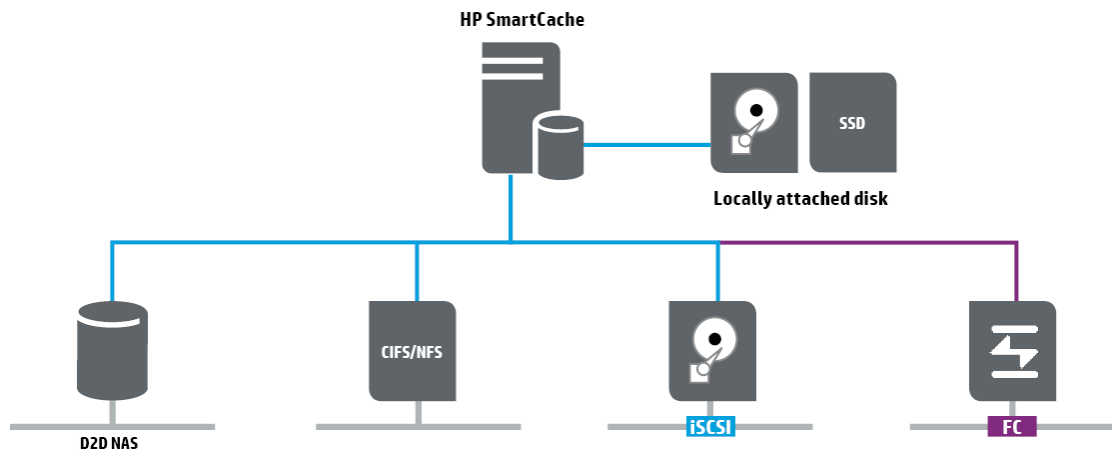
- **Risks:**

- Different site options can have impacts to the IT budget, OPEX overhead, and the length of time to restart normal operations.
- Continuous data transfer can impact network performance.
- Manual intervention/scripting may be necessary to complete the failover/failback process.

To address the criticality of VM and VM workloads and strike a balance between the need for performance in the backup and recovery operation with the sensitivity to the costs of higher tiered storage, HP Data Protector offers an intermediary backup target—SmartCache (see figure 8). This backup-to-disk target provides a new backup target option with all of the benefits of snapshots stored on the backup server cataloged and ready to be moved to the backup target, or kept ready for rapid recovery operations. Moreover, it is designed specifically to handle VMware workloads enabling direct access to VM data for whole-system or granular object restores and integrated within the HP Data Protector GRE. The benefit to VM administrators is that restore operations can occur directly from the SmartCache, or select a non-staged backup set on the associated disk-to-disk appliance to restore from. Lastly, administrators can make use of locally attached, or internal, disks on the backup server, or remotely connect NAS, CIFS, NFS, iSCSI, or FC storage to configure as a SmartCache solution.

- **Benefits:**
 - Improved recovery speed without consuming expensive primary storage.
 - Integrated into the recovery chain and controlled by the VM administrator.
 - VM data is browsable with direct access for restore operations.
 - VM power on and live migration from the SmartCache.
- **Risks (as of the publication of this document):**
 - Does not support compression/deduplication.
 - Only supports VMware workloads.

Figure 8. HP Data Protector SmartCache disk options



Consideration 5: building for the future based on knowledge of the present with analytics

Data centers often struggle with two common facts: unabated data growth and flat/declining IT budgets. These two facts require a smarter approach to the backup and recovery process, but with so many more backup choices than organizations have had historically, balancing the tradeoffs between faster backups and faster recoveries in an intelligent and well-informed way is difficult, but not impossible. A smarter approach is based on backup and recovery software’s ability to deliver integrated analytics, optimization, and architectural efficiencies.

Analytics is more than just static/structured reporting. It is an integral part of the backup and recovery software design, where by the process, from source to target and all points in between are continually monitored, data is collected, and is readily accessible, analyzed, visualized, and can be acted upon. At HP, this is referred to as Adaptive Backup and Recovery.

HP’s innovative approach to backup and recovery is based on the use of operational analytics targeting the day-to-day use of the backup infrastructure. More importantly, this approach adds trending capabilities and predictive algorithms enabling IT teams to make decisions about the backup and recovery process before problems surface. At a high level, Adaptive Backup and Recovery provides the following capabilities that address the shortcomings associated with traditional/commoditized backup and recovery models:

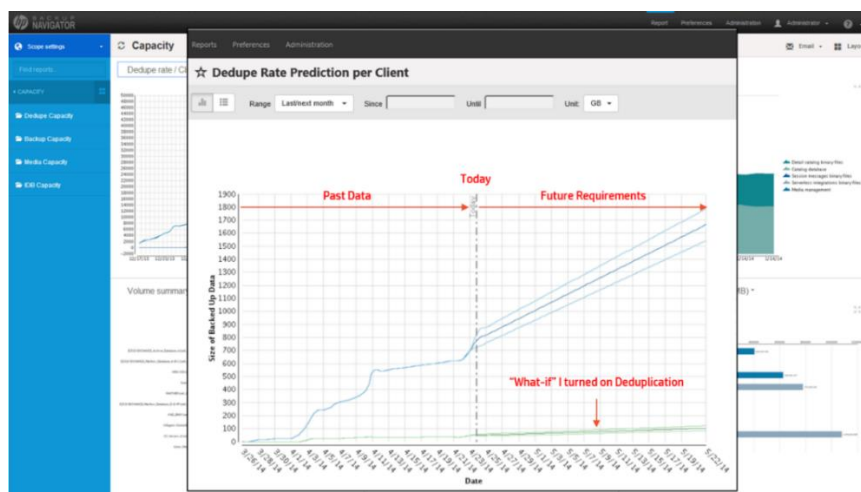
- Application and infrastructure backup analysis to prioritize backup operations effectively.
- Infrastructure performance monitoring to identify potential conflicts associated with interconnects between the source and target media, and session throughput rates.
- “What-if” scenarios to characterize backup data sets that use a common federated deduplication model to curb backup capacity requirements potentially before the space is depleted.

- Capacity trending to expose current infrastructure utilization rates in conjunction with forward-looking, or predictions, gaps in the infrastructure where the backup and recovery needs are not going to be met.
- Root-cause analysis with resolution recommendations.
- Recommend solutions to optimize backup environments, and self-tune and self-correct backup policies to improve business continuity within your IT environment.

These capabilities are delivered through a suite of inter-related software products (HP Data Protector, HP Backup Navigator, and HP Data Protector Management Packs) that, when combined, enable the operational analytics features described above.

HP Backup Navigator (see figure 9) is designed to provide this kind of process and infrastructure intelligence. Using an interactive interface, the IT staff is no longer left to isolate data protection problems at a point when the business needs are centered on the need to recover vital information. Instead, having end-to-end insight into the physical and logical data protection infrastructure enables the staff to make smarter decisions concerning how the backup and recovery process is implemented and uncover the root causes of issues before the business has to rely on that process. The scope and use of primary/secondary storage, disk-to-disk, disk-to-tape, target media, and the logical constructs of backup and recovery process are used to support business units, operating environments, and applications that are vast and often includes both physical and virtual infrastructures making use of general and purpose-built devices, cloud targets, and archival systems. The same kind of data characteristics discussed earlier are complicating the backup and recovery process as it is based on a variety of infrastructure that becomes more complex as it expands and requires more capacity to address the speed at which data is growing. The future of backup and recovery is in enabling the IT staff to make smarter decisions, develop future-proof protection plans and resolve issues with a clear end-to-end view over a diverse non-centralized infrastructure—and this is what the Adaptive Backup and Recovery initiative with HP targets.

Figure 9. HP Backup Navigator analyzes and predicts future needs based on historical trends

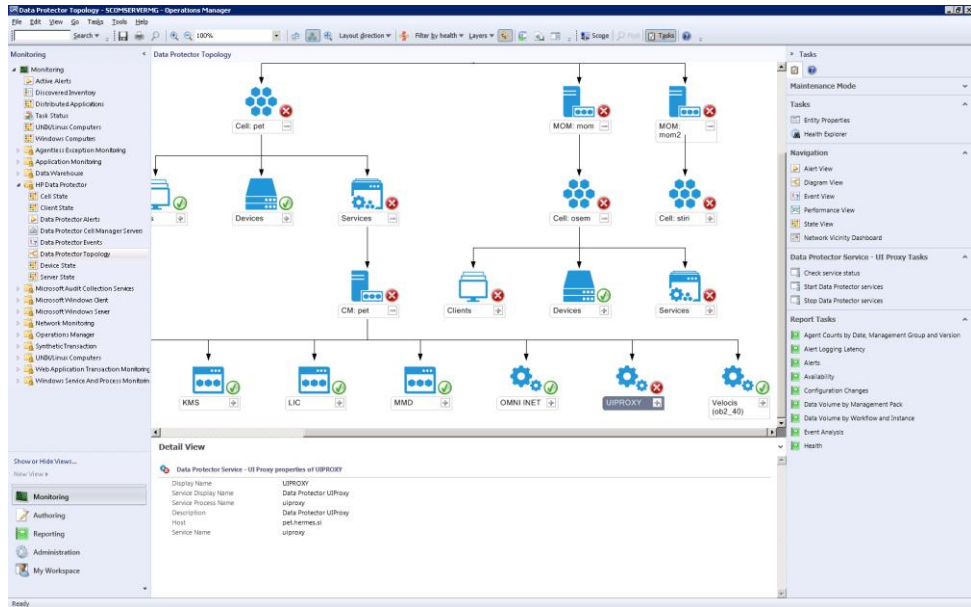


HP Backup Navigator provides the IT staff with intuitive and interactive dashboard and analytical reports based on 75+ key performance indicators concerning backup and recovery operations. With this, the IT staff can immediately identify inefficiencies within the backup operations, unbalanced use of backup resources, and uncover failures before they are exposed in the recovery process. More importantly, HP Backup Navigator monitors the real-time backup and recovery operations analyzing the use of the physical and logical resources identifying trends in the data sets that lead to gaps in the infrastructure you rely on. This intelligence means that you can make proactive decisions about performance and capacity needs before they are exhausted. In addition, this kind of intelligent insight provides a degree of foresight enabling the IT staff to evaluate the overall process capabilities before taking on new data protection workloads and applying a strategy in an ad hoc manner. Instead, the protection strategy is based on an assessment supported by real-time analytics and trending to ensure that maximum infrastructure utilization can be achieved without sacrificing the success of the operation. With so much information to be analyzed, HP Backup Navigator delivers the ability to create flexible and customized visualizations that can be scheduled and shared with key stakeholders or used as a form of collaboration between the members of the IT staff. In addition, the value of the reports can be exported in a variety of forms and formats for ingestion into other system such as broader business analytics, billing systems, regulatory reporting, data center health, etc.

With **HP Data Protector Management Pack** (see figure 10), the IT staff is enabled with detailed backup and recovery health information viewable from within Microsoft System Center Operations Manager and HP Operations Manager relative to the physical, virtual, and logical backup and recovery components of HP Data Protector infrastructure.

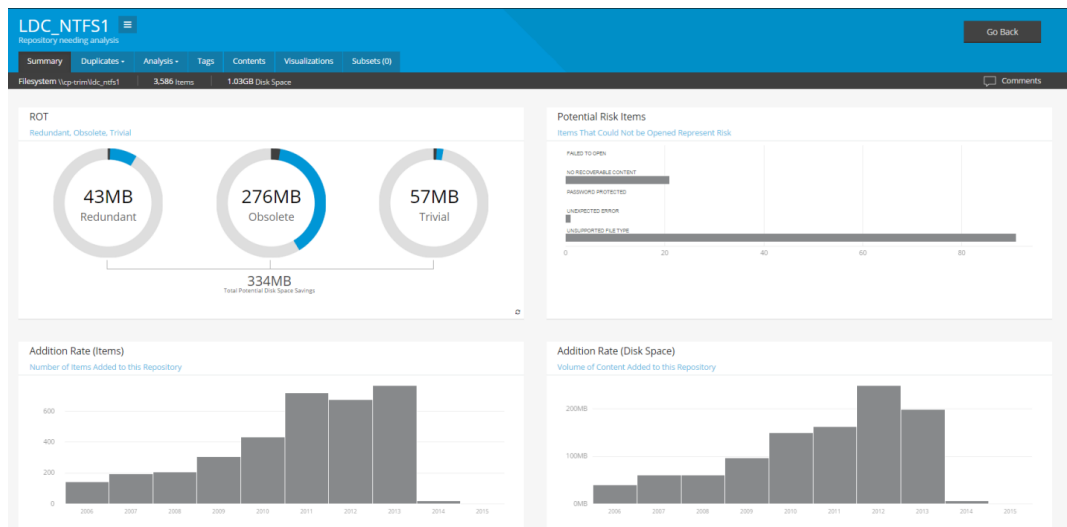
The result is unprecedented visibility into the health of your backup and recovery infrastructure from within the System Center. IT administrators can use Data Protector’s SCOM or HP Operations Manager interface for integrated end-to-end monitoring, automated diagnosis, and intuitive reporting across their complete backup and recovery infrastructure (physical and/or virtual).

Figure 10. HP Data Protector Management Packs delivers real-time operational analytics with assistive remediation



Lastly, the health and status of the virtual and physical infrastructure/workloads is not the only beneficiary of analytics. The relevance of information is just as dynamic as the virtual environment making the need to constantly prune primary data repositories difficult. HP addresses the need to profile and identify information assets that are candidates for archiving with its **HP Storage Optimizer** solution (see figure 11). With HP Storage Optimizer, IT staff can schedule storage repository and file analysis to identify specific information assets that have relevance to the business, but should be offloaded to an appropriate archive tier without severing access to the information. Once the assets are identified and archived, HP Storage Optimizer replaces the file with a link, or stub, that allows the user or administrator to access the information anytime it is needed (HP Storage Optimizer manages the access process and commits the data to disk should it become active—i.e., the user begins editing the document). This operation will not affect the backup and recovery process because HP Data Protector understands when it encounters an HP Storage Optimizer stub and handles it accordingly (i.e., does not recall the archived file during a backup operation).

Figure 11. HP Storage Optimizer analyzes storage, identifies inactive data, and archives it based on defined policies



About HP Data Protector

HP Data Protector is an enterprise backup solution that provides reliable data protection and high availability for your fast-growing business data. Data Protector offers comprehensive backup and restore functionality specifically tailored for enterprise-wide and distributed environments, and offers the following benefits:

- **Real-time operational intelligence:** Personalized/customizable dashboards and reports, intelligent scheduling, root-cause analysis, scenario-based modeling and predictive analytics for resource balancing, and identification and resolution of potential conflicts and contentions.
- **Zero Downtime Backup/Instant recovery:** Array-based snapshot integrations provide rapid protection and recovery while removing the burden that traditional backup technologies have on the production environment.
- **Storage optimization:** With compression, federated deduplication, storage management, and analytics, organizations achieve increased scalability and cost efficiency, and better utilization of the IT infrastructure.
- **Automated disaster recovery:** Centralized bare metal recovery from physical to physical, physical to virtual, virtual to virtual, and virtual to physical from any backup set at no additional cost.
- **Application consistent recovery:** Leading business application integrations extend backup, automated point-in-time recovery, and granular restores to application owners enabling them to manage, drive, and service their own backup and recovery requirements based on the backup infrastructure defined by IT.
- **Advanced virtual server protection:** Hypervisor integrations, and support, offer virtual machine protection inheritance, tiered recovery options, process automation, analytics, and visualization for virtual environments.
- **Standardized protection:** A unified and flexible architecture enables centralized protection across heterogeneous environments, disparate operating systems, and critical applications from core data centers to remote sites.
- **Information retention:** Automated retention and replication management across different backup media, storage tiers, and locations for compliance and efficient long-term data retention.

Conclusion

When you consider the compelling benefits of server virtualization (such as consolidation, improved management, and greater levels of reliability/standardization), it's no surprise that organizations are turning to virtualization to transform the data center and increase infrastructure agility. However, the flexibility and agility that virtualization brings can often directly contribute to both data growth and data complexity. As a result, organizations can end up building overly complex backup and recovery infrastructure to address the different characteristics of the data being protected using a range of unique features available through the IT infrastructure.

HP Data Protector allows organizations to make the most out of their virtual as well as physical infrastructure investment by delivering a data protection strategy that is as dynamic and agile as the virtual data center itself. By integrating with major hypervisors, HP Data Protector delivers protection inheritance, tiered recovery options, process automation, and analytics and visualization for virtual environments.

Learn more at
hp.com/go/dataprotector



Share with colleagues



Rate this document

© Copyright 2015 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP 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. HP shall not be liable for technical or editorial errors or omissions contained herein.

Microsoft is either registered trademark or trademark of Microsoft Corporation in the United States and/or other countries. VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions. Oracle is a registered trademark of Oracle and/or its affiliates. Red Hat is a registered trademark of Red Hat, Inc. in the United States and other countries. SAP is the trademark or registered trademark of SAP SE in Germany and in several other countries. Citrix is a registered trademark of Citrix Systems, Inc. and/or one more of its subsidiaries and may be registered in the United States Patent and Trademark Office and in other countries.

