

ROBO and regional data centre data protection solution scenarios using HP Data Protector software, HP VTL systems and Low Bandwidth replication

Executive Summary	2
Scenario Summary	4
Introduction: Issues Associated with Remote Office/Branch Office Backup.....	5
Solution Scenarios	7
Scenario 1 –Client based backup over the WAN using HP Data Protector synthetic full functionality.....	7
Scenario 2 –Low Bandwidth Replication using HP D2D and HP Data Protector for up to 24 remote sites	9
Scenario 3 –Low Bandwidth Replication using HP VLS and HP Data Protector for Regional data centers...	11
Summary.....	13
Glossary of Terms	14
Appendix A – HP VTL Family with Deduplication and Low Bandwidth replication	18
Appendix B – HP Data Protector Synthetic Full Backup	19
Appendix C – Deduplication is the enabler for Low Bandwidth replication.	21
Appendix D – HP Data Protector Product Structure & Support Matrix.....	23
Call to action	25

Executive Summary

Protecting data across all corporate locations is critical for business, legal and regulatory compliance purposes in today's information age. Too often, business critical data at remote locations is inadequately protected, exposing the business to the risk of lost data and lost productivity. A study conducted by industry analyst Enterprise Strategy Group found that the top three IT priorities for remote office / branch office (ROBO) locations were driven by business priorities: Improve information security, ensure regulatory compliance, and improve disaster recovery¹.

Some of the biggest challenges in remote office data protection cited by organizations are:

- Limited IT staff and limited backup & recovery skill set
- High capital and operating costs of remote office backup & recovery
- Inability to complete backup jobs within scheduled backup window
- Improving the time it takes to recover data

Centralizing the control, management and policies of remote location backup and recovery

¹ Enterprise Strategy Group Research Report: Branch Office Optimization, 2007

This whitepaper shows how recent developments in the HP data protection portfolio can overcome these obstacles through the use of centralized management with HP Data Protector software and the use of Deduplication and Low bandwidth replication on HP's virtual tape library products. The whitepaper examines three scenarios which cover a wide range of possible implementations.

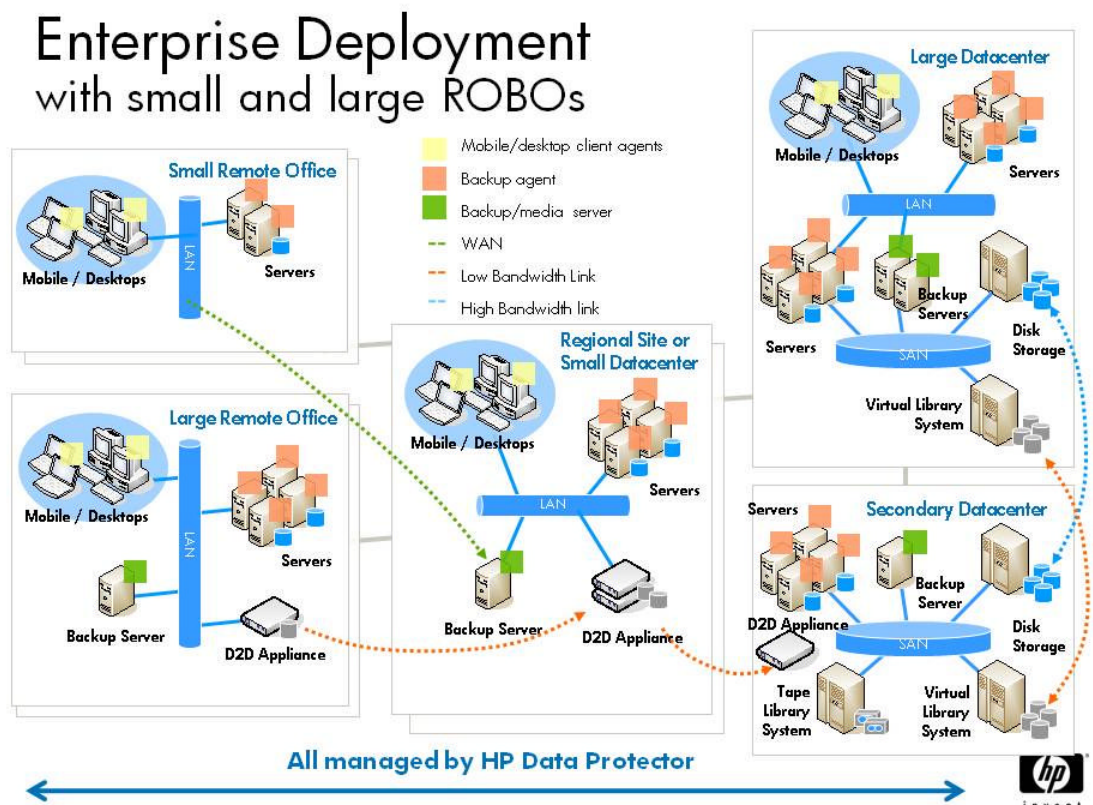


Figure 1: Enterprise deployment with small and large ROBOs

Figure 1 shows the HP Vision for "Protecting data anywhere" using these technologies.

The HP Vision for ROBO based data protection is that backup data should be accessible wherever it is required which in some cases requires it to be in 3 locations as well as on physical tape. This gives maximum flexibility in restore options or DR options. When managed through a single backup application this allows the personnel with relevant expertise to recover the data from multiple locations, be it recovering data at the ROBO site from the ROBO appliance but managing the recovery from the data centre or regional data centre, or even recovering at the ROBO site itself if the relevant expertise is present. The HP vision also realizes that backup appliance demands for ROBO's and data centers are very different both in terms of price and performance which is why the HP vision uses different technologies appropriate to these different customer requirements. Multi-hop replication and further combined appliance storage consolidation in the data centre are possible future enhancements to this vision. As we discuss various scenarios below it is worth noting some unique functionality enhancements between HP data protector and HP D2D & VLS appliances to allow this vision to be delivered today.

Scenario Summary

	Scenario 1	Scenario 2	Scenario 3
Typical sizing	200+ ROBO's with full backup size of 100GB & incremental backups of 10%	100+ ROBO's with 200-500GB full backups	100+ ROBO's with 200-500GB full backups PLUS Up to 4 regional data centres with 7.5TB to 300TB backups requiring DR solution
Problem	Large number of ROBO sites with small backup volumes	Significant backup volume at remote sites and a requirement for local recovery at the ROBO site.	All above data requiring to be consolidated at main data centre
Solution	HP Data Protector Synthetic full backups over WAN to HP D2D with NAS Emulation	Up to 24:1 fan-in from small D2D units in ROBO to larger consolidated single D2D in data centre Offload to Physical tape if required	HP D2D for ROBO's, HP VLS for Regional Data Centres D2D to VLS copies at regional data centre using HP data protector and onward replication to main data center
Why HP	Synthetic full functionality allows object consolidation and ensures only small volumes of data transmitted over WAN. Scalable up to 100s of ROBO sites Recovery to ROBO from data centre under data centre control	Only HP offers a low cost entry point to deduplication and low bandwidth replication for ROBO's (sub \$3000) Target D2D can be tape to tape copied to physical tape using special functions in HP Data Protector End to End control from HP data protector	Only HP offers both a low cost entry point to deduplication and low bandwidth replication(D2D) AND a highly scalable, high performance deduplication appliance (VLS) for Enterprise environments. Target VLS can be tape to tape copied to physical tape using special functions in HP data protector End to End control from HP data protector

The key learnings from a study of these scenarios are:-

- Using physical tape at ROBO's and performing daily off-siting of physical media will become a thing of the past, physical tape will be replaced by WAN based backup or a HP virtual tape libraries so that the whole process is automated and managed from a regional centre or main data centre. The automation will vastly improve the reliability of backup in the most remote offices where there are typically no trained IT staff.
- Deduplication technology is the key enabling technology for Low bandwidth replication which is revolutionizing ROBO backup. Deduplication enables the "unique" pieces of data to be identified and replicated, in turn this means much smaller inter-site links are required to effectively transmit large apparent quantities of data. Up to 95% bandwidth saving is achievable. Low bandwidth replication now ensures we can cost effectively "protect data anywhere". Low bandwidth replication can be "throttled" to use a percentage of an existing link, so as not to affect applications performance running on the same link. One key advantage for HP is its ability to provide "scale down" deduplication devices that are very cost effective and an ideal fit for ROBO's. Consolodation of replicated data from many ROBOs onto a single device at the regional data center or main data center allows for further cost efficiencies.
- Physical tape does still have a role to play in larger regional centres and main data centres, as a long term archive media for customers that have a need to keep data for years. It is relatively simple to integrate physical tape into a virtual tape environment because all backup software supports tape to tape copy. Certain compliance regulations also demand additional copies of data be made to physical tape.
- Close integration of HP D2D & VLS devices with HP data protector software allow for complete end to end solutions to be realized. Features such as synthetic full backup, Object copy and the ability for HP Data protector to detect replicated cartridges and incorporate them into its internal catalog database ensure the solution is always "Disaster recovery ready". Management of the whole process from anywhere in the network ensures a tight integration and expeianced IT staff access to all the relevant components.

Introduction: Issues Associated with Remote Office/Branch Office Backup

For many years remote office/branch office backup and restore has been an issue with IT managers. Some of the challenges they faced were:-

- Having to implement physical tape infrastructure and operators at what could be a large number of sites.
- Paying service providers to pick-up store and retrieve tapes to and from an off-site facility.
- Data in transit (physical tapes) are prone to being lost, and the company relies entirely on a 3rd party for their ability to recover from a disaster.
- Untrained IT personnel at the most remote sites, who may have been responsible for changing physical tapes, and then off-siting the tapes. A very manual but important process which may not always have been executed well.
- Reliability of the process – more prone to errors because it was manual and required regular human intervention. Physical tape drive cleaning and tape rotation added to the complexity.

Errors and failures not always accurately diagnosed. Low restore success rate at a local level was probably a major issue.

- If the backups were large or slow and only a single backup device was available – the backup window could be large and extend into the normal business hours affecting critical business applications.
- Backup over the company WAN to a central site was restricted because of the relatively high volumes of data that needed to be backed up over a relatively low speed link.
- It was very difficult to prove to auditors that a comprehensive site disaster recovery plan was in place and effective.

The solution to these issues is:-

- To automate the process End to End with a single management console across the entire estate.
- Extensive use of virtual tape libraries with deduplication and low bandwidth replication to automate the process. This allows flexible provisioning of backup resources to meet changing customer needs, and offer more effective storage utilization, ability to reduce backup windows and provide faster single file restores.
- Make the best use of existing WAN links from ROBO sites to Regional data centres and from regional data centers into main data centers by using Synthetic full backups and Low Bandwidth replication.
- Perform backup object consolidation by using technologies such as HP data protector synthetic full backups (See Appendix B): Only one initial full backup; no file is stored multiple times, and therefore, can reduce required backup space/media. Only changed files need to be backed up, this enables fast backup of file servers hosting millions of files allowing more frequent backup.
- Spend Less time on physical tape management: There is no tedious tape management with synthetic full backups or virtual tape libraries..
- To be able to maintain and manage Backup copies in multiple locations depending on retention policies and disaster recovery (DR) policies

Solution Scenarios

The scenarios discussed below cover a wide range of implementations that are possible using HP Virtual Tape Libraries and HP Data Protector software, from a small business that has several outlets to a major corporation that has hundreds of remote/branch offices, several regional data centres and multiple main data centers.

Scenario 1 –Client based backup over the WAN using HP Data Protector synthetic full functionality.

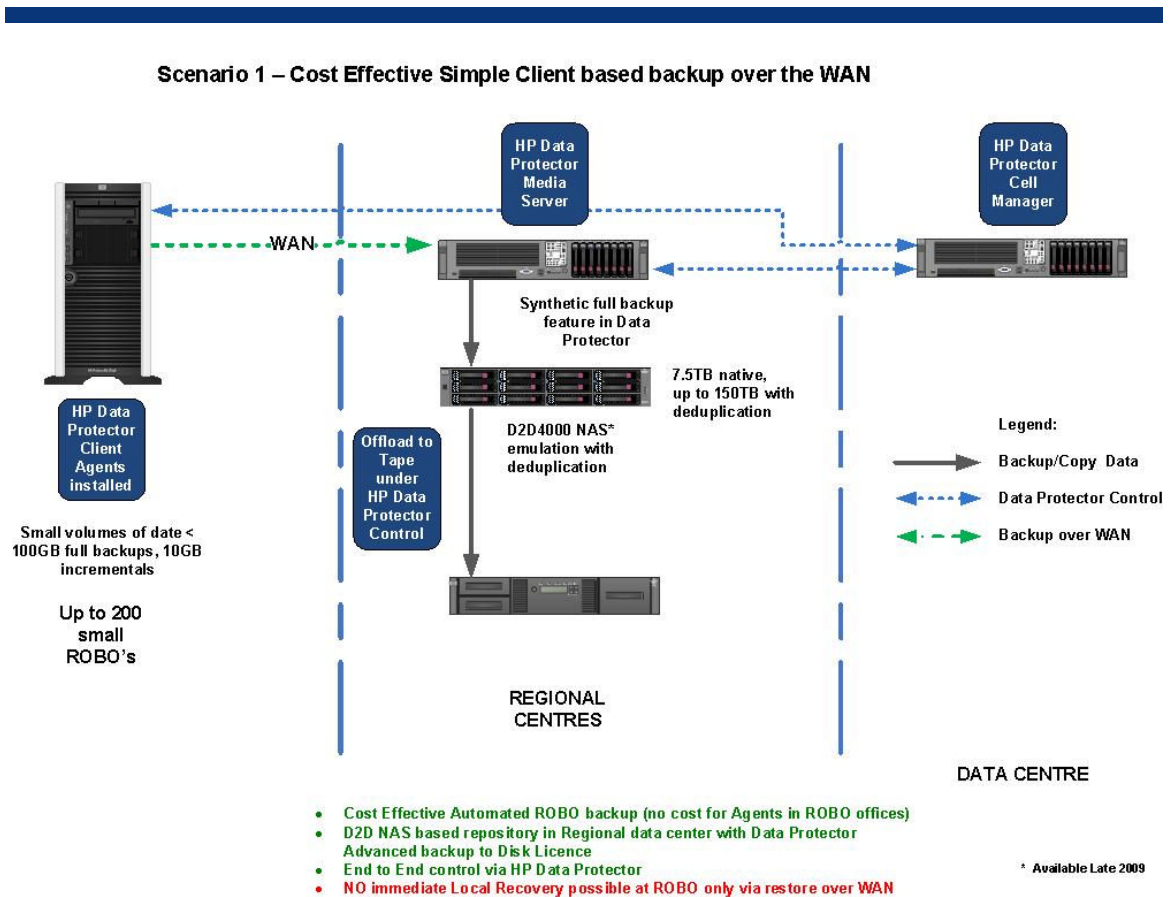


Figure 2: Client based backup over the WAN using HP Data Protector Synthetic Full functionality

Key Message: Cost effective Remote Office Backup with low WAN traffic

In this first scenario we follow a basic client agent backup process, suitable for a large number of ROBO sites where relatively small amounts of data need to be protected. In order to minimize the amount of data that needs to be transmitted over the WAN we can use the “synthetic full” feature of HP Data Protector. After the initial full backup we only need to perform incrementals thereafter. The backup data is sent to a “File Library” device (in this case a NAS share with deduplication on the D2D4000 unit), where a “synthetic full” can be created at any time for a restore. Alternatively there is another feature in HP Data Protector called virtual full that is similar to synthetic full but uses less storage space, but can only be implemented on a file library device such as an HP MSA2000 disk array. Finally again under data protector control the synthetic fulls can be copied to physical tape for long term archiving.

The key benefits to the customer of this solution are:-

- Object consolidation using synthetic full backups to minimize WAN traffic
- Variable Retention policies. Short term on D2D, Long term on tape for archive.
- Fast recovery from D2D4000 with NAS* emulation
- HP Data Protector manages the entire process from the main data centre.
- Reduced WAN utilization (because small incremental backups)

This solution does not provide a local source of the data for recovery. A synthetic full created at the regional centre is required and the necessary files to be restored need to be transferred over the WAN. This could take some time depending on size.

*Available late 2009

Scenario 2 –Low Bandwidth Replication using HP D2D and HP Data Protector for up to 24 remote sites

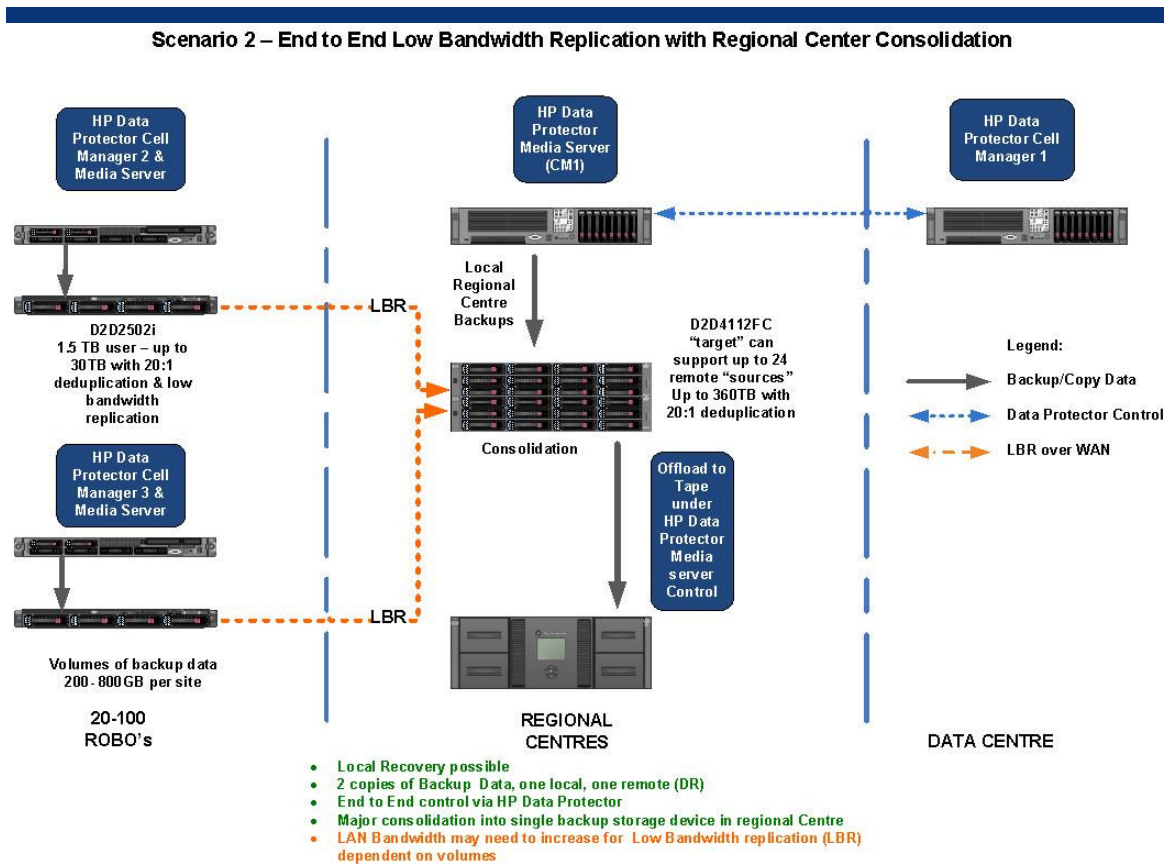


Figure 3 : End to End Low Bandwidth replication solution with regional centre consolidation, all under HP Data Protector Control

Key message: Backup Consolidation & Cost effective Disaster Recovery for larger ROBO's

In this next example the volumes of data on the remote sites are much larger; 500GB perhaps, and the customer really needs to have local restore capability. So in this scenario each remote site has a local D2D2500 with deduplication and the D2D is using virtual tape library emulation mode. A user capacity of 1.5TB can yield up to 30TB of backup storage with a 20:1 deduplication ratio, after several backups have been retained over 2-3 months. If long retention is not required then the D2D2500 can support even larger backups but retained for shorter periods. Direct restore at the ROBO from the D2D unit is always possible because each ROBO site contains a complete Data Protector environment of cell manager and media server and is "self-sufficient" from a restore perspective. The local D2D2500's (sources) can also be configured to replicate only the changed data across a low bandwidth link to the regional centre where a larger D2D4112 device resides. A D2D4112 can support a fan-in of up to 24 D2D sources, and only requires a single licence to enable replication. This allows major storage consolidation benefits at the regional data centre. The replication process replicates the changes in virtual cartridges from the ROBO to the regional centre. On a regular basis the HP data protector cell manager 1 polls the D2D4112 to detect replicated cartridges so it can update its internal database (IDB) with their contents. This ensures that in the event of a disaster at the ROBO the regional centre is immediately aware of which virtual tapes are required for recovery. Finally data from the D2D4112 at the Regional centre can be easily copied to physical tape at regular intervals using the "object copy" functionality within HP Data Protector.

The D2D4112 as well as acting as a replication target for the ROBO sites can also act as a source for backups from the regional centre servers as can be seen in Figure 3.

This solution provides robust disaster recovery, should a total disaster occur at the source (ROBO), the data is safely already offsite at the target (regional data centre). The ROBO site can then be re-constructed by either:-

- a) Re-building a new ROBO server and restoring data at the regional centre using the replicated data
- b) Rebuilding the server at the ROBO site and reverse replicating the critical data from the D2D4112 in the regional centre. Reverse replication involves the total contents of the backup being copied back from the replication target to the source, and may take some time if larger quantities of data are required to be reverse replicated over a low bandwidth link..
- c) Offloading critical data to tape at the regional center and transporting back to the ROBO with the necessary tape hardware for recovery.

The key new area of interest here is the Low Bandwidth link. For the majority of ROBO sites link speeds as low as 2Mbit/sec can be used (dependent on volumes of data to be replicated). The HP suggested usage models for low bandwidth replication are

- Use 100% of the link bandwidth to complete replication in a fixed window (12Hrs). This is for businesses that are not necessarily 24x7 and can afford a window dedicated to replication.
- Within the D2D unit it is possible to configure “throttling” of the link and allow the replication to take place over a 24 hour period. This is for businesses that also use the WAN link for applications, and where giving the whole link over to replication would affect application performance.

To calculate the size of the regional centre D2D device and the size of the link required (either with fixed window or with throttling) HP provides a comprehensive sizing tool available from:-

<http://www.hp.com/go/storageworks/sizer>

As for scalability - In the example above we show 24:1 fan-in, if we had > 24 ROBO sites we would have multiple separate D2D4112 devices at the regional data centre.

The key benefits to the customer of this solution are:-

- Backups performed and controlled locally, reliability of backups not dictated by WAN availability and remote media server as in scenario 1
- Fast local recovery when needed.
- Ability to store several months of backups locally at ROBO with D2D deduplication.
- Significant bandwidth reduction – up to 95% savings in bandwidth compared to all contents being replication rather than just changed data detected by deduplication.
- Cost effective consolidation in the regional centre up to 24 ROBO D2Ds (fan-in) AND D2D4112 can be used for local regional centre backups as well.
- Very Flexible Disaster recovery options.
- Easy offload to tape with HP Data Protector Object copy.
- HP Data Protector manages the entire process from the main data centre.
- Option to manage single ROBO'S from the Regional Centre or Data Centre

Scenario 3 –Low Bandwidth Replication using HP VLS and HP Data Protector for Regional data centers.

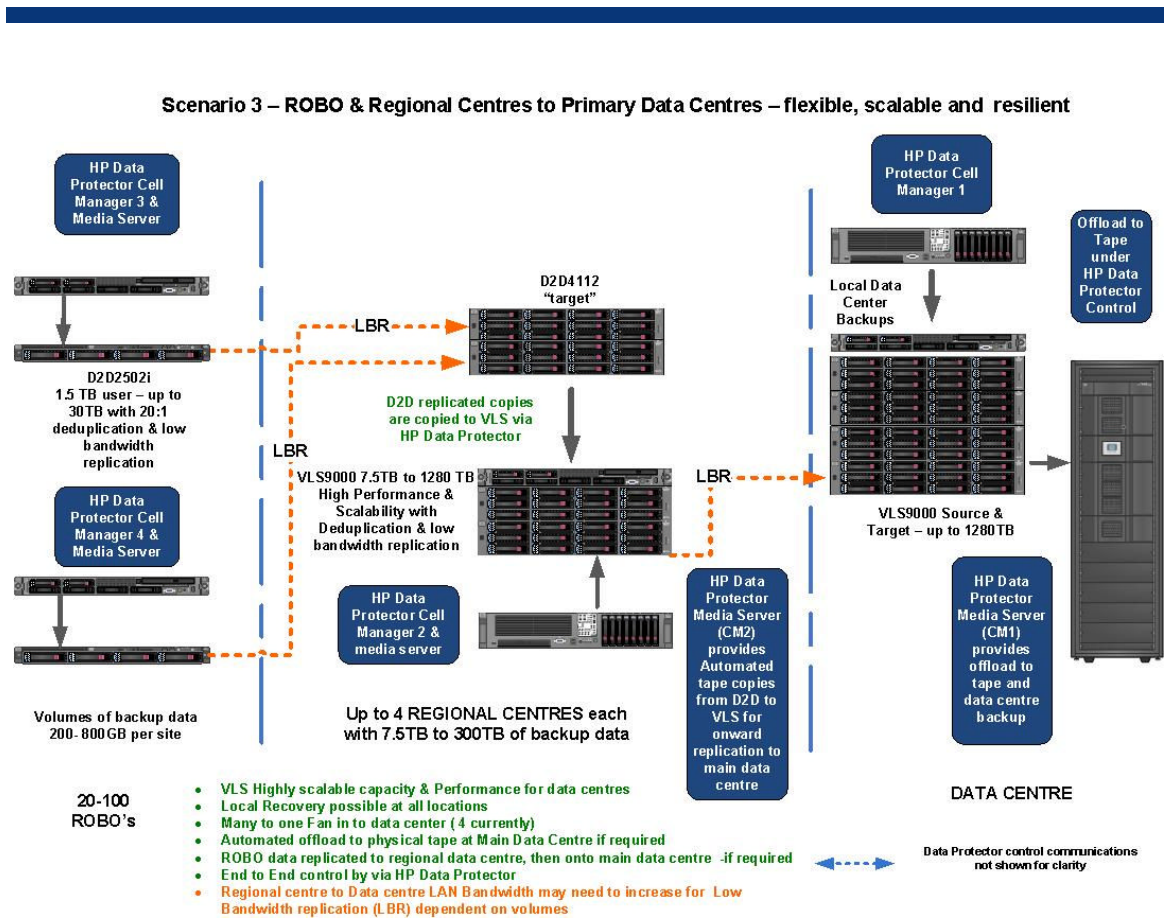


Figure 4: Regional data centre to Main data centre End to End replication with HP VLS all under HP Data Protector control

Key Message: Scalability, Consolidation and Data Centre resilience

In this scenario the volumes of data involved at the regional data centres are significantly high and the need for scalability paramount. HP's enterprise level Virtual Tape Libraries (VLS) are therefore used because in terms of deduplication (the technology enabler for low bandwidth replication) the HP VLS accelerated deduplication is able to scale much higher in terms of capacity and performance than the HP Dynamic deduplication used on HP D2D systems.

The ROBO to regional centre solution is the same as in Scenario 2 but instead of local regional data centre backups being done to D2D because the regional data centre volumes are > 30TB they are now backed up to a local VLS9000 device. VLS systems also support Low Bandwidth replication and currently support up to 4 to 1 fan-in, so we can consolidate replications from several regional data centres in a single VLS9000 device at the main data centre. The VLS9000 uses a multi-node architecture to ensure scalability in terms of backup performance, deduplication performance and replication performance. Because of the higher volumes of replicated data involved however the low bandwidth links for VLS are typically in the area of 100 Mbits/sec upwards. This solution with VLS allows consolidation at the main data centre, the VLS9000 there is used both as a replication target for the 4 regional data centres AND for main data centre backups as well, providing a very good consolidated solution. Deduplication is a licensable feature on the VLS range, and low bandwidth replication licencing is only required at the target site(one licence per node).

Replicated cartridges from the regional centre arriving at the VLS target in the main data centre send an Email on an hourly basis to the Data Protector Cell manager at the main data centre. Data Protector scripts then import these newly replicated tapes into the backup catalog at the main data centre to allow easy access to replicated data in the event of disaster recovery being necessary for the regional data centre.

Should it be required for compliance or archiving HP Data Protector Cell manager 1/Media Server at the main data centre can easily copy any content from the VLS to Physical tape.

A further enhancement to this scenario would be to have Active <-> Active replication where the main data centre backups to the VLS are also replicated back to the regional data center giving the main data centre a very cost effective disaster recovery process using existing infrastructure.

Because of the different nature of D2D & VLS deduplication technologies it is not possible for D2D units to replicate directly into a VLS unit.

The disaster recovery options for the regional data centre using VLS are identical to those used for D2D namely:-

- a) Re-building regional server and restoring data to it at the main data centre using the data replicated to the main data centre VLS.
- b) Rebuilding the server at the regional site and reverse replicating the critical data from the VLS in the main data centre. Although this will be limited to several TB max, because of the replication link being the bottleneck to high capacity restores – for reverse replication the “entire” contents of a virtual tape are reverse replicated.
- c) Offloading critical data to physical tape at the main data centre and transporting physical tapes and associated hardware back to the regional data centre for import and recovery.

Again with the higher volumes of data involved between regional data centres and main data centers HP expect the usage model for VLS to VLS replication to be predominantly

- A dedicated replication link – not shared with application traffic
- 24 Hour replication window available.

VLS replication sizing is also fully supported in the HP StorageWorks sizing Tool

<http://www.hp.com/go/storageworks/sizer>

The key benefits to the customer of this solution are:-

- Fast local recovery when needed at both ROBO, regional data centre and Main data centre.
- Option for ROBO data to be copied to VLS in regional data center for onward replication to main data centre.
- Ability to store several months of backups locally at regional data centre and main data centre using VLS accelerated deduplication - scalable to very high volumes.
- Cost effective consolidation in the main data centre, up to 4 regional VLS units (fan-in) AND VLS at main data centre can be used for local main data centre backups as well.
- Very Flexible Disaster recovery options.
- Easy offload to tape with HP Data Protector Object copy.
- HP Data Protector manages the entire process from the main data centre.
- Main Data centre DR capability if active<->active replication is deployed in the future, using existing infrastructure.

Summary

The three scenarios described above give an insight into the various ROBO & regional data centre DR deployments possible using HP storage components:-

- HP Data Protector product features (synthetic full backup, Object copy, automated tape importing etc)
- HP D2D systems – with Dynamic deduplication as standard and products specifically designed for the ROBO environment
- HP VLS systems with Accelerated deduplication specifically designed for the high performance and high capacity requirements of the enterprise
- Low bandwidth replication support on D2D & VLS with large fan-in ratio support – changing the dynamics and costs of “off-site” data protection
- HP Physical Tape Systems (MSL, EML, ESL) for the ultimate safe repository of data.
- Single point of control from the main data centre using HP data protector management and scheduling tools.

Using the wide range of components above allows flexibility and scalability in the variety of ROBO and data centre backup solutions that can be constructed. With innovative technologies such as synthetic full functionality, deduplication and low bandwidth replication on VTL products HP is able to deliver the optimum solution, with optimum performance at the best cost for each specific location, whilst ensuring central management and control.

Glossary of Terms

HP StorageWorks Sizer Tool

This freely available tool <http://www.hp.com/go/storageworks/sizer> allows users to specify backup capacities, retention schemes, data change rates, replication window, link speed etc and calculates likely backup storage capacity required, VTL model required, deduplication ratios, time to replicate (if link speed specified), link bandwidth required (if replication window specified). The tool then provides a design summary and a priced parts list (list price).

HP D2D

This is HPs virtual tape library designed specifically for ROBO and Small/Medium enterprise, it scales in capacity from 1.5TB user to 18 TB user and uses HP Dynamic Deduplication technology which then enables Low Bandwidth replication. Available with iSCSI Fiber Channel and NAS* interfaces and has inbuilt physical tape offload functionality. See Appendix A * Available Late 2009

HP VLS

This is HPs virtual tape library designed specifically for the Enterprise it scales in capacity from 7.5 TB to 1280 TB user and uses HP Accelerated Deduplication technology which then enables Low Bandwidth replication. A multi-node architecture enables highly scalable performance. See Appendix A

HP Data Protector software

HP Data Protector software automates high-performance backup and recovery from disk or tape to enable 24x7 business continuity at a compelling price-to-performance ratio. Appropriate for medium and large companies, this software helps you reduce IT costs and complexity while delivering the reliability and scalability needed to grow from a single server environment to the largest distributed enterprise infrastructure. With more than 22,000 customers worldwide, HP Data Protector software is used by nearly half of Global 500 corporations.

Target-based Deduplication

This is where the data is deduplicated in a Target device such as a virtual tape library. Deduplication is the ability to detect identical data patterns and save them only once – hence saving on overall storage requirements.

Active->Passive

Replication from a Source to Destination in one direction only.

Active-Active

Replication from a Source device on Site A to a Target Device on Site B and vice versa.

Many-to-One

Replication from multiple sources to a single target device.

Deduplication ratio

The reduction in storage required for a backup (after several other backups have taken place). The ratio is highly dependent on:

- Rate of change of data (for example, 10% of the data in 10% of the files)
- Type of data—files vs database
- Retention period of backups
- Efficiency of deduplication technology implementation

In-Line (HP Dynamic deduplication)

This is where the deduplication process takes place REAL TIME as the backup is actually taking place. This is the technology used on HP D2D devices.

Post Process (HP Accelerated Deduplication)

This is where the deduplication process takes place AFTER the backup has landed on the device to ensure deduplication does not affect backup performance. This is the technology used on HP VLS devices

Source Library

When used as part of a replication pair the source library is the library supplying the data to be replicated.

Target Library

When used as part of a replication pair the target library is the library receiving the data from the source library.

Concurrency

This is the Number of replication jobs that can be active at any one time— this varies depending on if it is a source or target library. Concurrency ensures maximum use of the available link bandwidth.

Fan-In

This is the number of source appliances or libraries that can be channeled into a target appliance or library.

Link Speed

This is speed in Mbits/sec that the Telco provides point to point or Any to Any to allow replication to take place. The speeds can vary from 2 Mb/sec to 1 Gb/sec. Increasing Speed means increasing cost.

Bandwidth Utilization (Throttling)

This is the amount of available bandwidth a device can be configured to use. For example the D2D & VLS devices have network throttling so they can be configured to use for example only 20% of the available bandwidth. This prevents any single device from “hogging” the total bandwidth available, and allows replication to take place over longer periods of time without affecting application responses using the same link.

Synthetic full

Synthetic backup puts a stop to the buildup of incremental backups, and eliminates the need to run lengthy full backups. This technology works by merging all incremental backups into a full ‘synthetic’ backup—a process that can be repeated indefinitely, with no need to run a full backup again. If all the backups, full and incremental, are written to the same HP Data Protector file library which uses a distributed file media format (DFMF) e.g on an MSA2000 disk array, an even more efficient type of synthetic backup is possible. This is called Virtual Full Backup. The solution uses pointers to consolidate data rather than copying the data. As a result, the consolidation takes less time and avoids unnecessary duplication of data. This capability also helps enable continuous data protection for Windows systems. For systems running Windows 2000 and above, HP Data Protector utilizes the built-in Microsoft Windows Change Journal to quickly generate a list of files that have been altered since the last backup was performed. This option avoids the need for timely “tree walks” to identify files that have changed, which can be especially time-consuming on systems with many small files. This allows incremental backups to run much faster and more frequently, providing near-continuous data protection at no additional cost.

HP Data Protector Cell Manager

The Cell Manager is the main system in the cell. The Cell Manager:

- Manages the cell from a central point
- Contains the IDB (Internal Data Base)

The IDB contains information about backup details such as, backup durations, media IDs, and session IDs

- Runs core HP Data Protector software
- Runs Session Managers that start and stop backup and restore sessions and write session information to the IDB.

HP Data Protector Disk Agents

Client systems you want to back up must have the HP Data Protector Disk Agent (DA), also called Backup Agent, installed. To back up online database integrations, install the Application Agent. The Disk Agent reads or writes data from a disk on the system and sends or receives data from a Media Agent. The Disk Agent is also installed on the Cell Manager, thus allowing you to back up data on the Cell Manager, the HP Data Protector configuration, and the IDB.

HP Data Protector Media Server

Client systems with connected backup devices must have a HP Data Protector Media Agent (MA) installed. Such client systems are also called Drive Servers. A backup device can be connected to any system and not only to the Cell Manager. A Media Agent reads or writes data from or to media in the device and sends or receives data from the Disk Agent.

HP Data Protector Object Copy

The Data Protector object copy functionality enables you to copy selected object versions to a specific media set. You can select object versions from one or several backup sessions or object consolidation sessions. During the object copy session, Data Protector reads the backed up data from the source media, transfers the data, and writes it to the target media.

The result of an object copy session is a media set that contains copies of the object versions you specified.

Tape Import scripts

The replication in both the VLS and D2D systems is mirroring the source cartridge to its matching target cartridge so both cartridges have the same barcode, the same tape contents, etc. HP Data Protector currently cannot handle seeing two copies of the same cartridge at the same time (because to the backup application, the cartridge is a single entity in the media management database). With HP Data Protector you can use a second cell manager to import and manage the replicated media/data.

Since replication will result in source and target media having identical barcodes, the source and target device should not be visible in one cell. SAN zoning is a method to assign portions of a virtual library to different backup applications such as HP Data Protector.

For VLS systems, the replication target is a subset or an entire virtual library that is presented on front-end Fibre Channel ports. This library is available in read only mode which allows replicated data to be managed using a second HP Data Protector cell manager. The VLS provides email notification reporting on completion of replication jobs.

For D2D systems, the replication target is an entire D2D library that is presented to a second HP Data Protector cell manager. The D2D provides XML polling notification reporting on completion of replication jobs.

A front-end script will handle the “notification” and feed a “incoming” queue. This frontend is depended on the device (VLS – email processing, D2D XML polling).

The backend processes the incoming queue and distributes the media to import to the available devices. The backend will maintain a log of import jobs and re-queue jobs in case of failures.

Replication Window

A configurable period of time within the D2D & VLS systems during which replication is not allowed to take place. Primarily this is to ensure available WAN bandwidth is not oversubscribed during peak periods of demand which might slow down application performance.

Seeding/Initializing

This term is used to define the process of “synchronizing” the two D2D or VLS appliances with the correct data prior to low bandwidth replication commencing. Because low bandwidth replication only transfers differences the Virtual tape libraries have to have the baseline/reference data established before low bandwidth replication can work.

Importing/Exporting

For Seeding large quantities of data we cannot use a low bandwidth link, so physical tape or co-location is used to allow the source and target devices to be synchronized. If physical tape is used an export and import process is used to get data out of and back into the D2D or VLS devices.

Tape Offload

On D2D devices there is an in-built tape offload utility accessible from the user interface that allows scheduled copies of the replicated data onto physical tape device connected directly to the D2D by SCSI or SAS host bus adapter inside the D2D.

Appendix A – HP VTL Family with Deduplication and Low Bandwidth replication

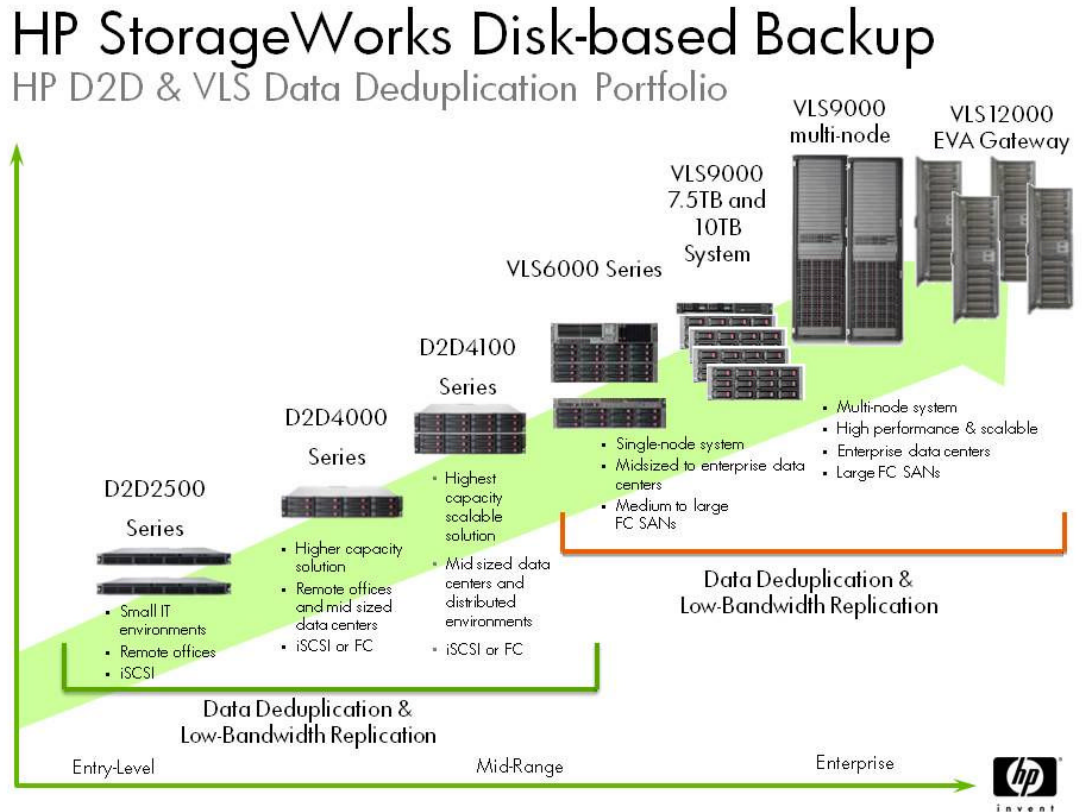


Figure 5: HP Virtual Tape Library portfolio with Deduplication and Low Bandwidth replication

The above figure shows the complete range of HP Virtual Tape Solutions. The D2D2500 is best suited for the most remote offices, the D2D4112 offers major benefits in consolidation by being able to support up to 24 remote D2D2500's. The VLS systems are highly scalable in terms of capacity (7.5 TB to 1280TB) and performance (400MB/sec to 4800MB/sec) and currently support 4 to 1 fan-in. NAS emulation on D2D units to support synthetic full backups will be available Late 2009.

Appendix B – HP Data Protector Synthetic Full Backup

HP Data Protector 6.0 Synthetic Full – Incremental Forever

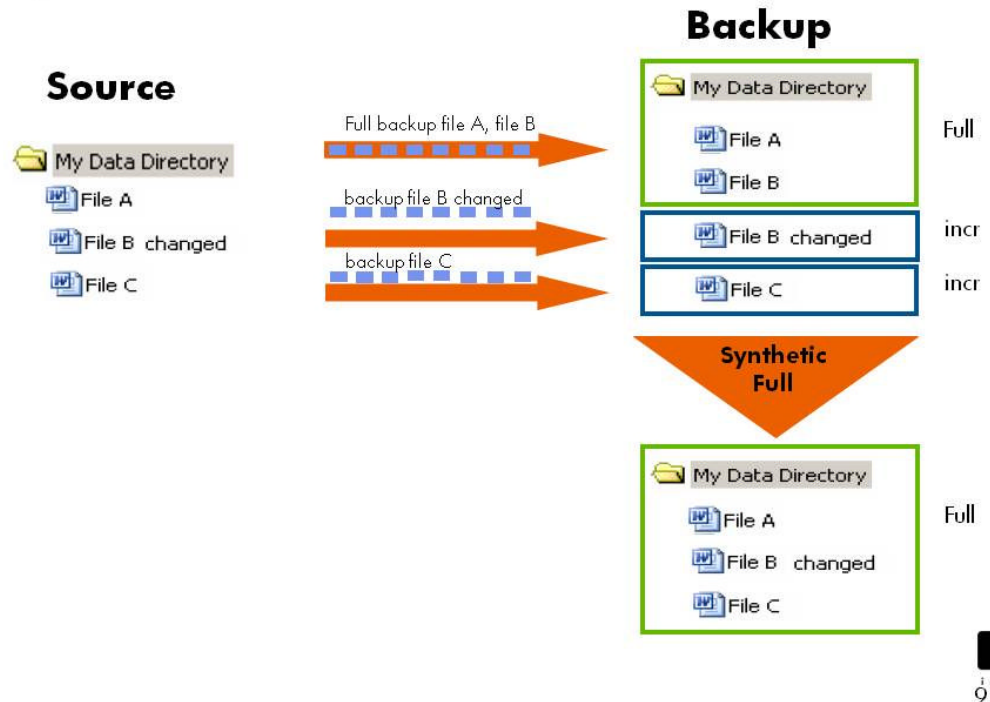


Figure 6: HP Data Protector Synthetic full backup explained

HP Data Protector 6.0

Synthetic Full – Incremental Forever

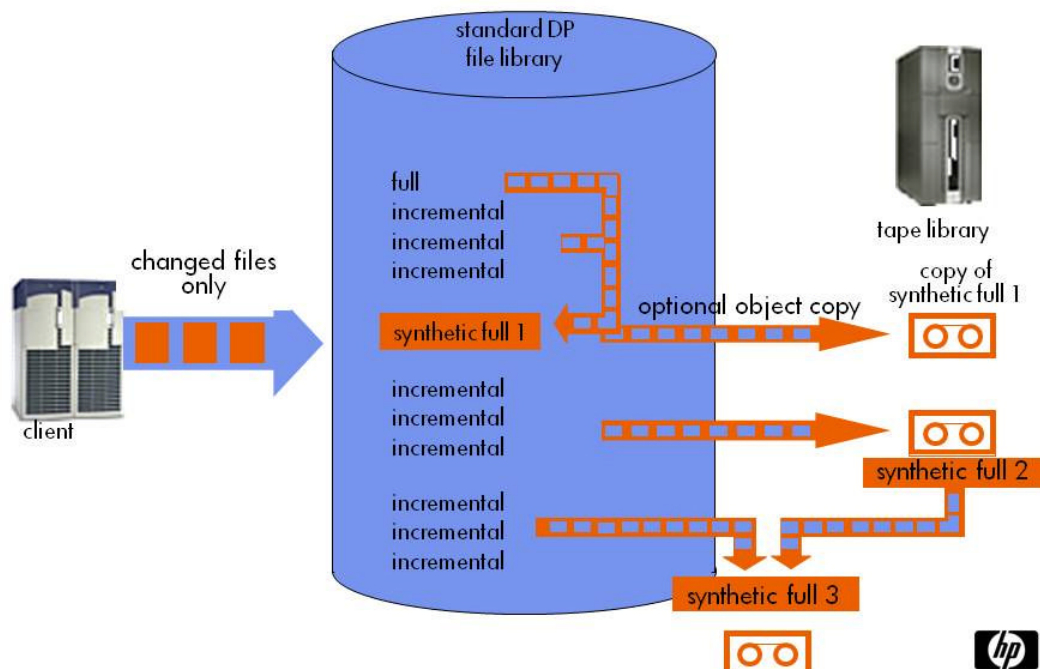


Figure 7: HP Data Protector Synthetic full backup explained

HP Data Protector software offers an advanced backup solution called synthetic full backup. This solution enables you to create synthetic full backups and virtual full backups with an operation called object consolidation allowing implementing an incremental forever backup strategy.

The incremental forever paradigm means that except for the first backup where a full backup is performed, only incremental backups are executed. This concept presents the most efficient way of backups; only the changed data is backed up.

However, without object consolidation, the restore process would last far too long since all backup sessions would have to be restored. Due to this behavior, regular full backups are required and all backups would need to be protected permanently. Object consolidation removes this drawback. During the object consolidation session, HP Data Protector software reads the backed up data from the source media, merges the data, and writes the consolidated version to the target media.

An object consolidation can result either in a synthetic full or a virtual full backup.

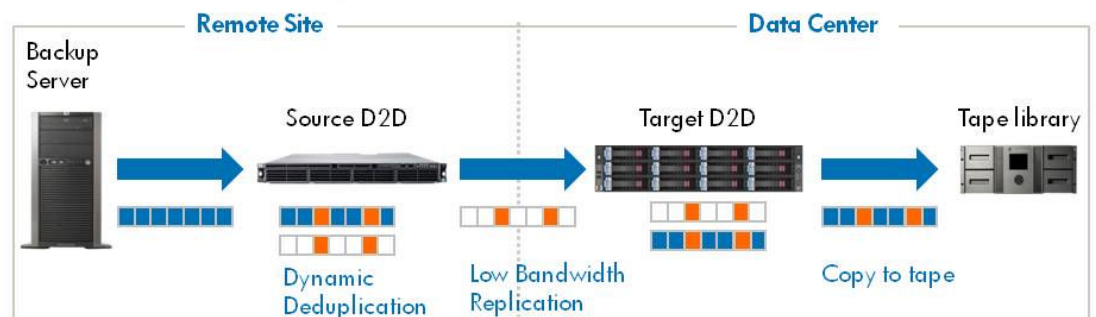
A synthetic full backup is the result of an object consolidation operation where a restore chain of backup objects is merged into a new, synthetic full version of this object. A synthetic full backup is equivalent to a conventional full backup in terms of restore speed. HP D2D NAS emulation supports synthetic full backups

A virtual full backup is an efficient type of synthetic backup where data is consolidated using pointers instead of being copied. It is performed if all the backups (the full backup, incremental backups, and the resulting virtual full backup) are written to a single HP Data Protector software file library using distributed file media format. This requires the File Library to be a disk array such as HP MSA2000

Appendix C – Deduplication is the enabler for Low Bandwidth replication.

Although HP D2D & VLS devices use a different implementation for deduplication both technologies detect unique data and pass that unique data to the replication engine.

Dynamic deduplication enables low bandwidth replication



- Data from the backup server is sent to the source appliance. The source D2D performs dynamic deduplication
- Deduplication compares blocks of data with the previously stored data blocks and removes the redundant blocks thus minimizing the amount of data required to be sent over the network
- If duplicate data is found, a pointer is established to the original data, rather than storing the duplicate data.
- Therefore, only the new or changed data is sent to the target D2D thus reducing the bandwidth required on a WAN
- Up to 6 source appliances can replicate to a single D2D2500 or 16 to the D2D4000
- Physical tape libraries can be directly connected to the D2D appliance for tape copy

Figure 8: Dynamic deduplication is the enabler for Low Bandwidth replication on D2D

Low Bandwidth Replication is enabled by Accelerated Deduplication on VLS

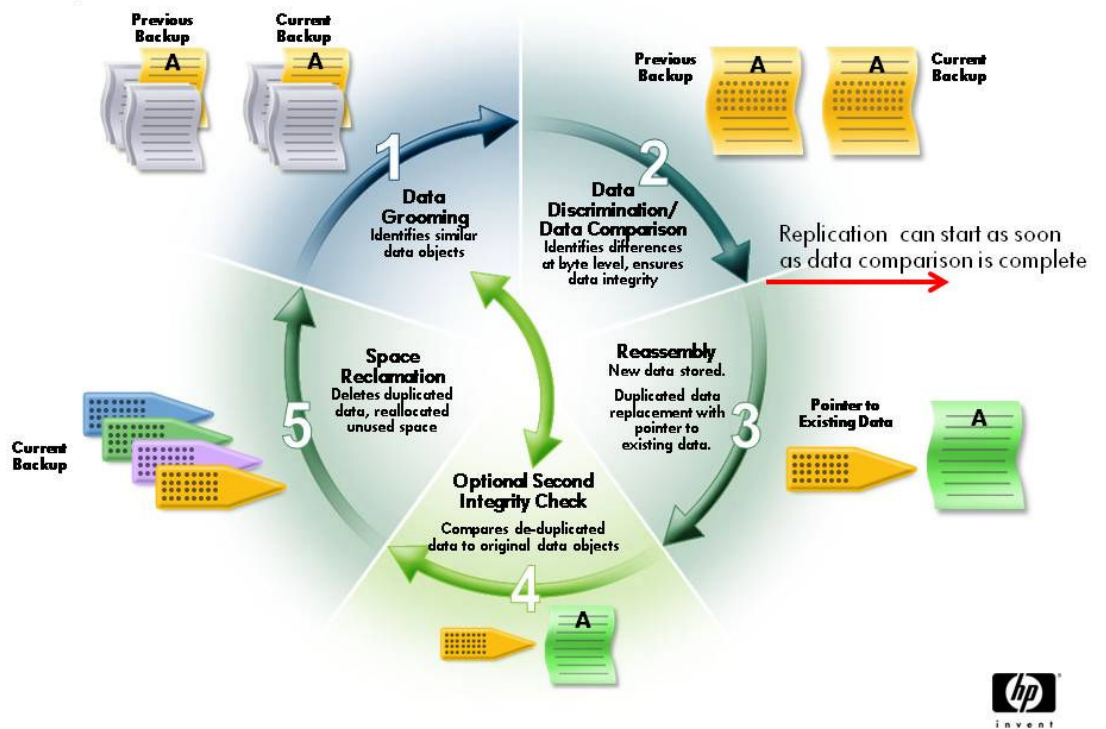


Figure 9: Accelerated deduplication is the enabler for Low Bandwidth replication on VLS

Appendix D – HP Data Protector Product Structure & Support Matrix

HP Data Protector software 6.1

Product Numbers

Single Server Edition			Windows	HP-UX	Solaris		
LTU & media / LTU only migration to starter pack			B7030AA/BA B7031AA	B7020AA/BA B7021AA	B7020DA/CA B7021DA		
1	Starter Packs (required)		All platforms	Windows	Linux	HP-UX	Solaris
	Starter Pack manuals - printed		B6960LA	B6961AA	B6961DA	B6951AA	B6951DA
	LTU & DVDs LTU only DVDs only (2)	1x Cell 1x Cell	B6960MA	B6961BA B6961CA	B6951BA	B6951BA	B6951CA
Drive and library extensions			All platforms	Windows, NetWare, Linux		SAN, UNIX, NAS	
Drive LTU Library LTU			1x drive 1x 6.1-250/unlimited slots 1x upgrade to unlimited slots	B6957BA/B6958BA B6958CA	B6963AA		B6953AA
2	2. Backup to Disk			All platforms			
	Advanced Backup to Disk LTU	1x TB / 10x TB / 100x TB		B7038AA/BA/CA			
3	4. Application Protection			Windows & Linux		UNIX	
	Online Backup LTU	1x system		B6965BA		B6955BA	
	Zero Downtime LTU Instant Recovery LTU	1x TB / 10x TB		HP XP B7023CA/DA B7026CA/DA	HP EVA B7025CA/DA B7028AA/DA	EMC B6959CA/DA	
4	4. Manager of Managers			Windows & Linux		UNIX	
	Manager of Managers LTU	1x system		B6966AA		B6956AA	
	Open File Backup LTU	1x enterprise server/5x workstations 1x 1-server/1x 10-servers	BA155AA/BA154AA BA153AA/BA	CD only		BA152AA	
	Encryption LTU	1x system, all platforms	BB618AA / BA				
	Media Operations LTU	1x 2,000/10,000 media 1x unlimited media	B7100AA/B7101AA B7102AA	CD only/manuals only		B7129AA/B7128AA	
	Direct Backup LTU	1x TB / 10x TB, NDMP	B7022BA/DA	1x TB / 10x TB, HP XP		B7027AA/DA	

Figure 10: HP Data Protector Product Structure

HP Data Protector software 6.1

Broad Interoperability across all major platforms

Java GUI

- Microsoft Windows
- HP-UX (PA-RISC, IA64)
- Solaris (SPARC)
- Linux SuSE, Red Hat

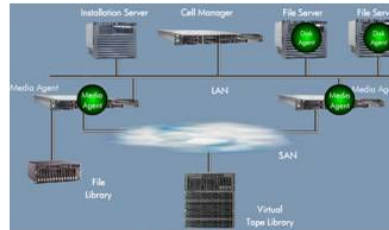
Disk Agent

- Microsoft Windows
- HP-UX
- Linux
- Solaris (SPARC, x64)
- IBM AIX
- HP Tru64 UNIX
- HP MPE/iX
- HP OpenVMS (Alpha, IA64)
- Novell NetWare
- SCO OpenServer; SCO Unixware
- SNI SINIX; SGI IRIX
- Additional platforms via NFS, shared disk (CIFS) or NDMP



Native Windows GUI

- Microsoft Windows



Media Agent

- Microsoft Windows
- HP-UX
- Linux
- Solaris (SPARC)
- IBM AIX
- HP Tru64 UNIX
- HP MPE/iX
- HP OpenVMS (Alpha, IA64)
- Novell NetWare
- SCO OpenServer, Sinix

Cell Manager

- Microsoft Windows
- HP-UX (PA-RISC, IA64)
- Linux SuSE, Red Hat (x64)
- Solaris (SPARC)

On-line Backup

- VMware ESX 3.x
- Microsoft SharePoint Portal Server
- Oracle
- Microsoft SQL Server
- Microsoft Exchange
- Microsoft VSS API
- Microsoft Internet Info Server
- SAP R/3
- SAPDB / MaxDB
- Informix
- Sybase
- Lotus Notes / Domino
- IBM DB2
- HP Network Node Mgr 

Figure 11: HP Data Protector Support Matrix

For more information

www.hp.com/go/VLS

www.hp.com/go/D2D

www.hp.com/go/storageworks/sizer

[HP StorageWorks Deduplication and replication solutions guide for D2D & VLS](#)

(<http://bizsupport2.austin.hp.com/bc/docs/support/SupportManual/c01729131/c01729131.pdf>)

[HP D2D Replication Primer](#)

(<http://h20195.www2.hp.com/V2/GetPDF.aspx/4AA2-4142ENW.pdf>)

[Understanding HP Deduplication strategy](#)

(<http://h20195.www2.hp.com/V2/GetPDF.aspx/4AA1-9796ENW.pdf>)

www.hp.com/go/dataprotector

[Advanced Backup to Disk Integration with Virtual Tape Libraries](#)

(<http://bizsupport1.austin.hp.com/bc/docs/support/SupportManual/c00868706/c00868706.pdf>)

[Integrating HP Data Protector software with HP Data Deduplication solutions](#)

(<http://bizsupport2.austin.hp.com/bc/docs/support/SupportManual/c01601020/c01601020.pdf>)

[Disk-Assisted Backup Whitepaper](#)

(<http://bizsupport1.austin.hp.com/bc/docs/support/SupportManual/c00668397/c00668397.pdf>)

[HP Data Protector software object consolidation: "Best Practices"](#)

(<http://bizsupport1.austin.hp.com/bc/docs/support/SupportManual/c01425312/c01425312.pdf>)

Call to action

www.hp.com/go/robo

© Copyright 2008 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.

Linux is a U.S. registered trademark of Linus Torvalds. Microsoft and Windows are U.S. registered trademarks of Microsoft Corporation. UNIX is a registered trademark of The Open Group.

4AA2-6879ENA, June 2009



i n v e n t