



Network Computing

Network Computing Request for Information Data Protection through Replication

Darwin's Groceries

Vendor response:

Vendor company name: Quantum Storage Solutions Group

Vendor service name (if any): _____

Vendor contact name: Kim Nelson

Vendor contact telephone number: (949) 725-1831

Vendor contact e-mail address: kim.nelson@quantum.com

Date of submission: November 5, 2003

The RFI did not mention several aspects of the customer environment that would significantly affect the design and cost of the proposed solution(s) in the RFI response. Specifically:

- No mention was made as to what backup software is currently being used.
- The RFI mentions that NAS servers are in use at the remote stores, and that changed data is transferred nightly to the data center. No mention is made of what technology is being used to transfer the data. There is no mention of whether or not there is a NAS server in the main data center, and no mention is made of brand or model (which can significantly impact the design of the solution).

Assumptions:

For the purposes of this RFI response, we assumed the following:

- VERITAS Netbackup 4.5 was already in use in the environment. Given the fact that Darwin Groceries is already backing up to three large tape libraries, and doing it via NDMP, it will be assumed that they are using enterprise class backup software.
- The NAS filers in question (at the stores and in the data center) are Network Appliance 700 and 800 series Filers. This means they can support NetApp's Snapshot technologies over the TCP/IP links between the stores and the main data center. Although there is no specific mention of any NAS servers at the main data center, it can be assumed that there are NAS servers because of the mention of NDMP backups over Gigabit Ethernet (GbE).
- The drawing in the RFI shows a tape library directly connected to the mainframe. Because it is directly connected to the mainframe and is not mentioned in the text of the RFI, it will be treated as an "island" and won't be considered in this response.
- The RFI did not state what operating systems Darwin is using on its database servers and other systems. Given the highly transactional nature of Darwin's operations, and the need for high performance systems, we would recommend moving to an all Sun Solaris environment (if that were not already the case). This would obviously be limited if any specific in-house applications required the use of Windows or a specific flavor of UNIX (i.e. HP-UX or AIX).

However, the use of other operating systems will still allow the use of the recommended solution.

Solution

Our proposed solution consists of the following parts, discussed in order of implementation. Please use the attached PowerPoint slide as a visual reference.

1. Replace the Store Accounting Systems with a Hitachi Lightning 9900 array and a Brocade Fibre Channel switch. The current method that Darwin is using to expand their database is very inflexible and requires separating the components of the database. By using a single high performance disk array, the entire database can be kept in one array which will make this part of the solution much more scalable in the future. Scaling is a simple matter of adding more disk capacity to the array system.
2. Setup a switched Fibre Channel fabric in the data center, using either Brocade 3800 or 12000 switches. Initially, only the Store Accounting System Hitachi array, database servers, and tape backup devices will be attached to the fabric. Eventually, both the Inventory Management Systems data and the Data Warehousing DAS data can be moved to the fabric. This can be done by either expanding the Hitachi 9900 array, physically moving the data over, or attaching the existing DAS arrays via Fibre Channel (SCSI to FC bridges will be likely be required). Moving these two applications to a switched fabric will increase backup speed and will provide additional options for future replication strategies (i.e. replication over Fibre Channel instead of replication over IP).
3. Establish the remote data center and purchase a leased, high-speed WAN line between the remote data center and the main data center. The remote data center will require servers to house replication and possibly clustering software, and will also house disk storage devices (i.e. disk and tape). Because this replication strategy includes software-based replication (see next step), Darwin can reuse older disk arrays at the remote site for storage of replicated data. Security was a major concern for Darwin. They can choose a dedicated leased line as it is relatively secure, but if Darwin desires an extra measure of security, a VPN encryption tunnel could be established between the sites. The VPN could be done

via existing firewalls, or with dedicated VPN/encryption devices. Third party bandwidth management tools can be implemented to help manage the use of the bandwidth to this remote site (to ensure that the replication between primary and remote sites has sufficient bandwidth, given that other applications or processes could be using the same network link). Due to the performance and security requirements mentioned in the RFI, the use of public Internet connection is not recommended (unless it was to be made available as a secondary or tertiary failover link).

4. Install VERITAS Volume Manager and VERITAS Volume Replication (VVR) for UNIX on the Store Accounting Systems servers, the Inventory Management Systems servers, and the Data Warehousing servers. Adequate server capacity and disk space would need to be setup and configured at the remote data center. VVR allows the use of either synchronous or asynchronous replication over a TCP/IP network. Given the latency of an 80-mile TCP/IP connection, we recommend the use of the asynchronous feature of VVR. The asynchronous connection minimizes the impact on the performance of the application, and meets the RFI requirement of having the remote data synchronized within five minutes of the original data. Disaster recovery capabilities could be enhanced if Darwin exercised their option to use clustering technology (such as VERITAS Cluster Server, or VCS). This would ensure that the application server at the remote site was ready for business in the event of a failure or disaster at the main data center, and would fail over well within the 30-minute requirement given in the RFI (This assumes that all applications are cluster aware). If the cost of the clustering technology were too high, then the Darwin IT staff would need to ensure that the application servers at the remote site were tested and ready in the event of a failure at the main data center.
5. Replace the existing three tape libraries with a Quantum DX100 enhanced backup system. This would greatly reduce the backup window and significantly reduce the amount of floor space required in the main data center. Additionally, long-term maintenance costs on the three libraries would be reduced or eliminated, since the maintenance costs of the DX100 may be significantly less. Given the large amount of data needing to be archived, a 64 TB DX100 is recommended. Since Darwin is relying on the VVR software for primary data replication, the DX100 would be acting solely as a secondary restore device (given data corruption or potential

problems with the VVR software) and as a staging device for the archive of data to tape. In a crisis, data from the DX100 could be restored very quickly back to the primary arrays (much faster than from tape). It will also reduce the backup window and the time it takes to create archive tapes (the NAS filer(s) would backup directly to tape, since they essentially already make backups of themselves internally via snapshot technology). Also, the high performance of the DX100 allows a relatively fast restore of data, especially when used with VVR's checkpoint restore feature, which allows the data to be "rolled forward" to a consistent state in conjunction with a restore from DX100 (which might be necessary in certain types of data loss or database corruption).

6. Install a Quantum MAKO PX720 library with twelve SuperDLT600 drives in the main data center. This library would be used for two purposes:

1. Backing up the NAS servers (via NDMP over GbE through DinoStor GbE/SCSI bridges). This would require the use of VERITAS' NDMP option.
2. Archiving data to tape for the purpose of taking tapes offsite for disaster recovery purposes.

While we eliminated the three existing libraries in the previous step, a Quantum PX720 with SDLT 600 drives provides a very dense library with high performance tape drives. This allows Darwin to reduce costs of buying and storing (offsite) media, and reduces maintenance costs (vs. their old libraries). Four tape drives would be connected to the GbE network via the DinoStor bridges, and would be used to backup the NAS filers in the main data center. The remaining eight drives would be Fibre Channel attached to the Brocade SAN and would be used with VERITAS Netbackup's SSO (Shared Storage Option) to share the drives dynamically via fabric. The MAKO offers native Fibre Channel drives and bridged drives, so Darwin could choose which technology they prefer. Optionally, a smaller Quantum library with SDLT600 drives could be setup at the remote site for disaster recovery from offsite tape (in the event of a failure in the WAN links or VVR replication technology) or one of the three libraries that had been taken off-line could be re-deployed. Tapes can be driven off site by a third party DR provider (ie. Iron Mountain) to the remote data center for future restore needs.

Requirements Addressed in RFI: (answers in italics)

Your solution should include mechanisms or functionality for:

- Hosting replicated data on storage platforms or topologies that do not replicate on a one-for-one basis storage platforms located in the production environment, thereby enabling greater flexibility and lower cost for the overall recovery strategy
 - o *ANS: the use of VVR (and not proprietary solutions provided by disk vendors) allows the replication of data on a non one-to-one basis. For example, the older Xiotech and HDS arrays could be redeployed to the remote site to store the replicated data.*
- Monitoring the on-going performance of the replication strategy
 - o *ANS: provided by VVR's VRA advisor software.*
- Testing the replication strategy without disrupting normal application or storage operations
 - o *ANS: application can be tested at remote site with synchronized data (clustering software will improve this).*
- Securing data from eavesdropping or unauthorized access during the replication process and after "fail over" of application access to the replicated data set.
 - o *ANS: provided via 3rd party VPN/encryption to remote site*
- Scaling readily in response to increases or decreases in the volume of data to be replicated
 - o *ANS: VVR software architecture allows easy scaling as requirements change, especially since it's not tied to any particular hardware product.*
- Culling from replicated data duplicate and/or non-critical data as well as data or files containing virus signatures or other malicious software code
- Automated techniques for optimizing data transfers across WAN interconnects of varying bandwidth and for optimizing WAN interconnects for best possible cost-efficiency
 - o *ANS: use 3rd party "bandwidth grooming" tools, which can be integrated with firewalls and/or encryption products.*

The asynchronous option within VVR will allow the data to be synchronized within the five minute window stated in the RFI.

Products used:

Quantum PX720 MAKO tape library with twelve SDLT600 tape drives
Quantum DX100 w/64 TB disk
DinoStor Gigabit Ethernet to SCSI bridges (2)
VERITAS Netbackup 4.5 w/Shared Storage Option and NDMP Option
VERITAS Volume Manager
VERITAS Volume Replicator for Unix
Brocade 3800 switches or single Brocade 12000 switch (at Darwin's discretion)
Hitachi Data Systems Lightning 9900 array with 64 TB of disk
Third party encryption and bandwidth monitoring products
Optionally: VERITAS Cluster Server for some or all applications

1. Deployment Issues:

- The project will likely take approximately six months to implement in monthly phases. This time-frame will be heavily influenced by the decision to use an existing store facility for DR versus building a dedicated DR site. Three months are likely required on the front end for planning and procurement purposes.
- Quantum has designed its solutions to meet customers' increasing data storage needs. The MAKO PX720 offers performance and capacity that is easily configured to meet any customer requirements today and in the future. MAKO's Performance on Demand (PoD) feature enables customers to increase library performance as needed by adding modular drive clusters containing up to four tape drives, redundant fans and power with failover capabilities and built in management. Additionally, Capacity on Demand (CoD) is also available to support growth of library capacity without unnecessary up-front investment. When requirements grow to exceed the capacity of a single MAKO frame, the CrossLink™ mechanism allows the library to expand up to five linked MAKO frames that are centrally managed to support up to 100 tape drives and up to 3,660 LTO-2 tapes for a total capacity of 732 TB (native). Additionally, customers can scale the DX100. Starting at 8 TB, customers can expand the DX100 capacity in scalable increments of

4 TB, up to 64 TB, while the system continues to be managed by the backup software as a single tape library.

- Quantum has an extensive list of service and support capabilities. However, for a project of this magnitude, we would obviously rely on the service delivery capabilities of our VAR and ISV partners to deliver key planning and integration services as part of the overall solution.

2. Solution Pricing:

As discussed in the previous paragraph, Quantum would deliver this solution in conjunction with key VAR and ISV partners. It would require a highly specific infrastructure analysis, needs assessment and budget review to determine an exact cost for the implementation described above. The list price for the Quantum specific portions of this solution are as follows:

DX100 – 64TB \$615,500

MAKO Library – 12 SDLT 600 drives \$290,205

Service and support costs will vary depending upon length of term and service level agreement.

3. Benefits:

- The DX100 has significant advantages over tape-based backup and restore architectures as well as other disk-based backup methods. Backup windows become much more manageable because data can be backed up at disk-based speeds and migrated to tape during “daylight” hours. The use of the DX100 also allows the administrator much more flexibility in the way that retention periods can be managed. As a result, data will often be restored from disk rather than tape, again taking advantage of the performance capabilities of disk-based media. Overall solution dependability in the DX100 is enhanced by the use of RAID protected disk arrays, hot spare capabilities (ideal for lights-out environments), and redundant and hot-swappable components throughout (i.e. fans and power supplies). DX100 has a significant impact on overall solution costs since it gives the IT department the flexibility to eliminate or re-allocate older devices that are more difficult to manage and more expensive to maintain. In addition, using the DX100 as the backup target and the recommended tape devices for the archive solution allow the IT department the

flexibility to re-architect where data is stored and for how long which can yield significant savings in hard costs or allow the administrator the flexibility to backup more data and keep it for longer periods of time.

- Quantum's solution utilizing the functionality of the VERITAS software eliminates vendor lock-in. Software-based replication also prevents vendor lock-in and allows redeployment of existing storage resources.
- Darwin Groceries will recognize risk reduction benefits from remote replication, the MAKO PX720 tape library with reliability features of redundant cooling and power systems, advanced library management software that performs component-level system monitoring to address potential failures before they occur as well as enhanced robotics engineered to execute more than three million swaps before failure. Additional risk reduction benefits are realized by the DX100 which has high availability features and RAID 5 with hot spare disk options. By using the DX100 for backup and restore, customers can lower overall costs through more efficient use of media. If Darwin chooses to implement clustering technology, then further cost benefits would be realized by reusing existing disk. All of this ensures more system uptime and network security.

4. Market Viability

There are many players in the disk-based backup market today. The majority of them are either start-ups with a new twist on the same technology or established market players that have repurposed existing devices for backup and are now playing in a somewhat "foreign" environment. As a leader in the storage and data protection market for more than a decade, Quantum is uniquely qualified to deliver useful solutions to our customers with a very high degree of expertise. Quantum understands that successful backup and restore capabilities are more about process than they are about the implementation of a specific technology. So the ability of the DX100 to preserve the existing infrastructure and process (including classes, scripts, etc.), coupled with the market leading characteristics of the MAKO tape library (density, capacity, performance, reliability, etc.) resonates with our customers and will provide a useful and scalable solution to Darwin Groceries. Few other suppliers will be able to offer the financial stability of a

market leader coupled with the key relationships required to offer and support a solution of this magnitude. Key factors Darwin should consider in its business solution to deploy the DX100 and MAK0 are:

- Preservation of the existing environment (no forklift upgrade required)
- Competing solutions that offer the comparable levels of performance and technology advantage require much larger commitments in terms of process change, financial investment and people costs (re-training, time spent managing secondary storage).
- Investment protection is maximized due to the highly scalable nature of the solution. Choosing this path does not result in a dead-end for Darwin Groceries at some point in the future where they are no longer able to expand the solution to meet capacity and performance requirements.
- Quantum is a recognized leader in the data protection market and is offering technology that has been proven in the field (DX100 is a second generation product and MAK0 is Quantum's fifth generation library). This is extremely important in the relatively young market for disk-based backup.