Network Computing

Completed by Fujitsu Softek, Version 11.03.03

Network Computing
Request for Information
Data Protection through Replication

Deadline: November 5, 2003

Vendor response:

Vendor company name: Softek
Vendor contact name: Emory Epperson
Vendor contact telephone number: 408-746-6790
Vendor contact e-mail address: eepperson@softek.fujitsu.com

Date of submission: November 5, 2003

Purpose

This RFI is proprietary to Network Computing and CMP Media, LLC. It is drafted and disseminated for the sole purpose of generating information on data protection/recovery products. Participating vendors must meet the minimum requirements for participation and understand that **all information returned to Network Computing in response to this RFI may be published** in print and in electronic form on our Web site, www.nwc.com. You may be able to redact some information at the discretion of NWC, but we consider it in our readers' best interest to provide as much of the completed RFIs as possible. Please let the author know if this will be a problem.

Effective Dates

Please note: Products proposed in this RFI **MUST** be generally available by December 2003. No beta products please. We reserve the right to examine a test unit (either in our lab or at a customer site) of any product submitted.

Please answer all questions. These will determine the winning bid and our **Editor's Choice Award.**

RFI Issue Data: Oct. 7, 2003

RSVP Deadline: Oct. 15, 2003
RFI Deadline: Postmarked by Nov. 5, 2003
Publication Date: Jan. 22, 2004

Scenario

Darwin's Groceries has a love/hate relationship with the communities it serves. When the company first began its chain of MiniMarts in the 1980s, everyone loved the convenience of a 24-hour grocer whose prices were on par with conventional grocery store chains. However, as the company expanded and began opening its SuperGigantic store chain, it caused an upheaval in the local markets where each store opened by driving "mom and pop" establishments out of business. Demonstrations, occasionally violent, have broken out during groundbreaking ceremonies for new stores. Management is worried that "anti-Darwin malcontents" could eventually target the company's technology infrastructure.

A key to Darwin's success has been the harnessing of information technology to reduce staff overhead. Cash registers forward purchase information back to a central server in each store. In the new SuperGigantic stores, wireless handheld computers with scanning wands are used by staff to update stock reports. A few stores are even opening wireless "hot spots" for use by customers with PDAs for "cartless shopping." As the customer walks down each aisle, his or her PDA shows what's on the shelves; customers click on what they want, and their orders are already billed to their credit or debit card and loaded into bags ready when they are ready to leave the store. Darwin has put a lot of money into this "next evolution in grocery shopping," which it hopes will replace cash register attendants.

Nightly, information collected at each of Darwin's stores is transmitted to a centralized data storage platform at headquarters. Smaller stores use VPN connections over the Internet or direct dial-up modem connections to perform the transfer. Larger stores and the SuperGigantic storefronts have dedicated network high-bandwidth connections to HQ, plus Internet VPN connections as a fallback measure.

Once data is collected at HQ, it is replicated and directed both to a process that updates store management and inventory control systems and to the company's tape backup process. Then, some data from the management and inventory systems is abstracted for use in a data warehouse that helps Darwin spot trends and create new marketing and pricing strategies. These strategies have, thus far, aided in improving company revenues and reducing company costs on an annual basis.

Darwin wants to make sure that its data is well protected during collection, transport and storage. It is seeking to improve its storage infrastructure and management capabilities to enable storage and data protection to scale non-disruptively, and to provide better information on the status of storage-related replication and backup processes. The company is interested in exploring disk-to-disk data replication strategies but has thus far been unable to find a vendor that could support its heterogeneous storage infrastructure. Ultimately, the company would like to use tape for archive and disk-based replication for disaster recovery.

The company would also like to consider cost-effective methods for replicating its headquarters infrastructure at an alternate site so that business would continue uninterrupted in the event of a fire or an interruption in network services. Management is thinking about expanding one of the SuperGigantic stores, located about 80 miles away from headquarters, to serve as a disaster recovery site, either leveraging its existing WAN interconnect or increasing WAN bandwidth to handle remote data replication .

Your job is to give them what they want. Currently, key business processes and their storage infrastructure components are as follows:

1. Individual MiniMarts and SuperGigantic storefronts have centralized servers with NAS arrays for storage. The NAS platforms in individual MiniMarts have a capacity of about half a Terabyte; in the SuperGigantic sites, NAS capacity is about 1 TB. These are databases of store inventory and receipts. Change data is transferred nightly across dedicated networks, Internet-based VPNs, or modem dial-up to headquarters. On weekdays, about 750 GB of data are transferred; on weekends, about 1.5 TB of data pass between the stores and headquarters. Currently, there is no backup for data in the stores themselves. In the event of a disaster, new equipment will be drop-shipped to the stores and data will be reloaded from HQ via network or DVD ROM.

2. Store Accounting Systems at HQ are Oracle databases that contain daily receipt information from all of the company's chain stores. Store information is spread over three large servers with direct Fibre Channel attached arrays, each providing approximately 10 TB of capacity. The company has needed to scale capacity about once per year, usually by deploying a new server and array and moving database components over the expanded platform. The tradeoff for this strategy has been a requirement to purchase all storage from one vendor, currently EMC. However, the vendor has announced that it is changing its platform design, forcing Darwin to migrate all of its platforms to the next generation storage array or find a new strategy. Hitachi Data Systems looks inviting, but the company is open to standards-based alternatives that will prevent a perceived vendor lock-in.

3. HQ's Inventory Management Systems are currently hosted on servers, each with a SCSI-attached 8 TB XIOtech array. This data is also considered critical and needs to be included in the disk-to-disk data replication scheme.

4. Data warehousing and data mining are performed in a workstation cluster sharing a common 10TB HDS array. This data is also deemed critical and Darwin wishes to include it in the burgeoning data protection strategy.

5. Data backup is currently conducted via Gigabit Ethernet and NDMP to three high-end tape libraries that are providing a barely-acceptable 2-4 TB per hour backup speed. Darwin wants to migrate away from tape and into a disk-based data replication solution – preferably one that is platform agnostic. Other attributes sought from the solution include the following:

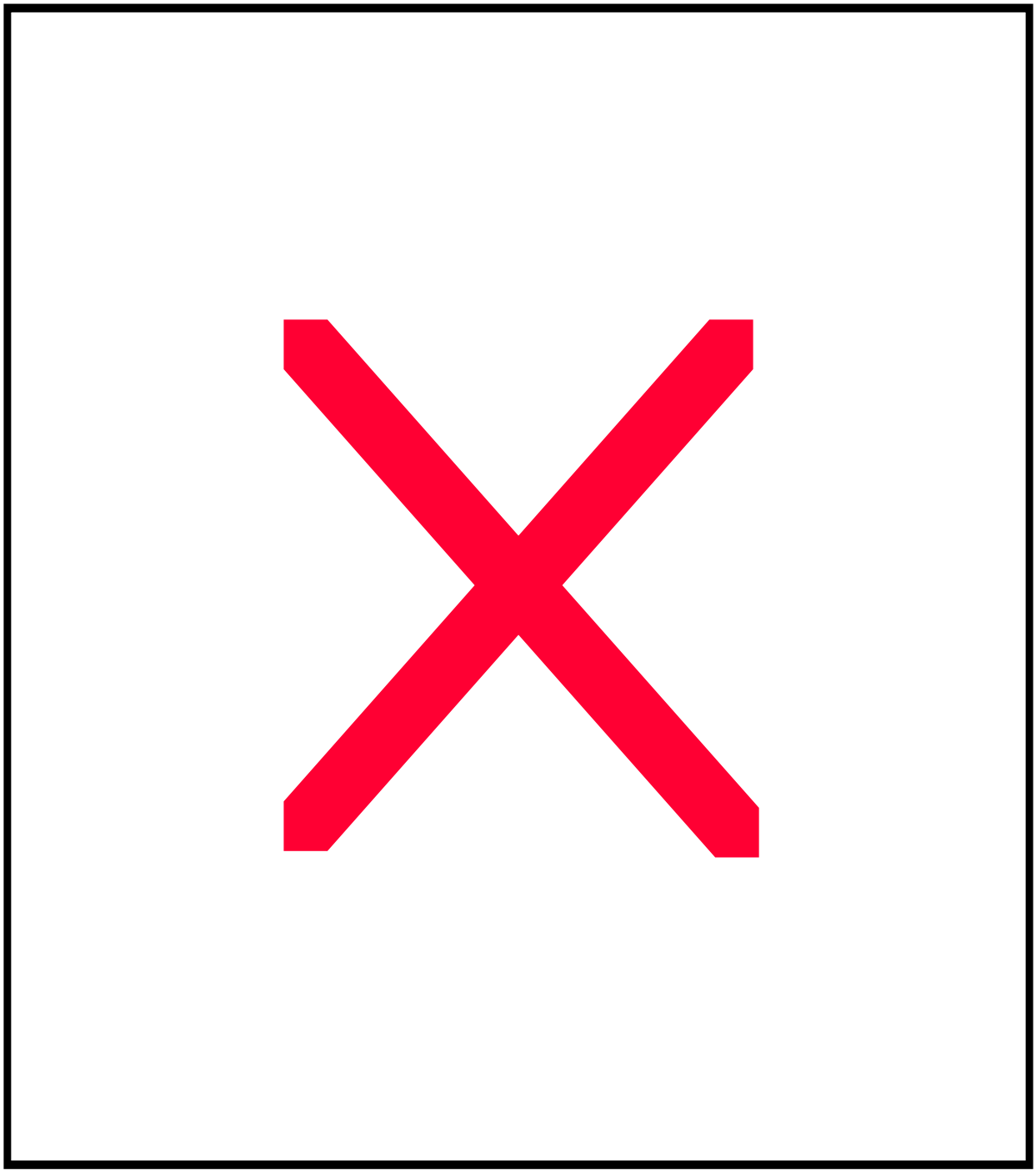
The solution should provide the means for replicating mission-critical data reliably and securely across a WAN so that the remote data copy is synchronized to within five minutes of the original and is available for use by applications within 30 minutes of an interruption of normal processing operations.

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
- Monitoring the on-going performance of the replication strategy
- Testing the replication strategy without disrupting normal application or storage operations
- Securing data from eavesdropping or unauthorized access during the replication process and after "fail over" of application access to the replicated data set
- Scaling readily in response to increases or decreases in the volume of data to be replicated
- 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

DARWIN PHASE 0 – Existing configuration



Softek Response

Overview

Softek offers a unique three-phase solution for Darwin's Groceries headquarters requirements for data protection that is cost-effective and timely. Most immediately, Softek brings Softek Replicator, a host-based data replication product that can easily handle Darwin's data protection and replication challenges and meet its business and technology requirements through its multiple platform OS support, flexible architecture, freedom to work with any storage vendor's hardware, and easy-to-use management interface.

Specifically, we propose in Phase 1:

- Install Softek Replicator on the primary and data warehousing servers at headquarters
- Purchase and install ~40-50TB+ of DAS to cover Darwin's data replication and protection requirements, as well as a number of new servers. Three of the units will be dedicated target replication servers, allowing for Darwin's desired replication operations and planned expansion
- Install Softek Replicator Common Console (on the above-mentioned three servers) to establish replication groups from all primary application and data warehouse servers to newly purchased replication target servers and storage

This will meet Darwin's immediate needs for replication within HQ while laying the groundwork for replication to the offsite DR location. In addition, it will enable Darwin to streamline its existing tape backup process.

Phase 1 assumes Darwin will purchase additional storage capacity to accommodate the newly replicated data. Because of Softek Replicator's heterogeneous support for replication / migration, Darwin can select best-of-breed tier-one or tier-two server/DAS storage combination at its next planned expansion to reduce costs dramatically (lower tier costs allow for purchase of more arrays and a larger pool of storage).

Please note: the proposed three servers used for replication operations will become the target servers, hosting all of the critical HQ replicated data.

During the installation of new storage and servers (outlined above), Softek Replicator can provide an efficient many-to-one configuration. In this phase, Darwin is able to protect all critical accounting, inventory, and sales analysis data by replicating to the new replication targets.

Furthermore, Softek has identified some options available to Darwin that may run contrary to their original assumptions about backup strategy:

- To deliver the best possible solutions with least expense to Darwin, we deliberately avoid the rip-and-replace of Darwin's high-end tape libraries, thereby preserving a still valuable capital asset.
- Until a fully functional DR site is established and operational in Phase 2, Softek recommends maintaining a streamlined tape backup strategy. This new strategy calls for the tape libraries to be connected directly to the target replication (backup) servers via existing network technology.
- Since all critical data can now be replicated to a centralized location, fewer backup server licenses are needed resulting in cost reduction.
- Since the data is new protected with replication, consolidated backups can run against the replicated data on the new target servers without the need for any backup window, and the HQ network load for backup is reduced to zero. This will dramatically improve backup performance while simplifying the over all backup plan.

In Phase 2 we propose extending and enhancing the solution by the following:

- Install Softek Replicator on the dedicated servers at the Super Gigantic Store that will act as the remote disaster recovery (DR) site servers and initiate remote replication of primary application data, as well as data from the data warehouse
- Install Softek Storage Manager agents at the DR site and at all Darwin stores
- Capture key data usage metrics for trending, forecasting, and planning through Storage Manager
- Install Softek Storage Manager on the server at HQ that is hosting the Softek Replicator Common Console
- Initiate consolidated LAN-free tape backup using Replicator at HQ

This is a transition phase to begin data replication at the DR site. The primary objective is to use Storage Manager to scrub data at the Darwin stores by identifying and eliminating redundant, outdated, and unnecessary data. This will significantly reduce the amount of data to be replicated throughout the enterprise, saving disk capacity and bandwidth in the process. This stage also establishes offsite replication although there is no intention to eliminate tape backup. Using the data trends collected and analyzed through Storage Manager, we can then determine when to move on to Phase 3.

In the final stage, Phase 3, we propose that Darwin implement a SAN at HQ.

- Install SAN technology (storage arrays, switches of Darwin's choice)
- Install Softek SANView to provide comprehensive data about the storage infrastructure
- Install Softek Provisioner to perform heterogeneous LUN management
- Use Softek Storage Replicator to migrate data to the SAN
- Install Softek EnView (optional) to manage QoS and SLA compliance

Central to all three phases is Softek Replicator, which replicates any application data on virtually all operating systems, open or proprietary, regardless of which vendor's storage products are used. In the process, it provides an easy way to monitor and take actions for the on-going performance of data replication activities through a common console with an enterprise-wide view. It allows for the testing of the replication strategy with no impact to the application, is highly scalable, and can be implemented to optimize data replication across any WAN bandwidth available. It also allows variable modes of replication configuration. These include many-to-one, one-to-many, chaining of multiple replication servers, symmetric site to site, and cluster configurations.

Since Darwin has declared its desire to keep costs low while replicating a fairly sizable amount of data both locally and remotely, Softek Replicator's many-to-one and chaining configurations are best suited to meet Darwin's needs today and provide an open infrastructure able to accommodate future architectural decisions by Darwin.

Given that Darwin appears to be primarily a UNIX and Windows shop today, Softek Replicator easily supports the replication of data on most UNIX and Windows operating systems, including Linux, providing the ability for storage administrators to use a single product regardless of the operating system. This makes it fast and easy for Darwin to transfer administrator skills to different operating systems when the situation dictates moving the applications to new servers or storage arrays.

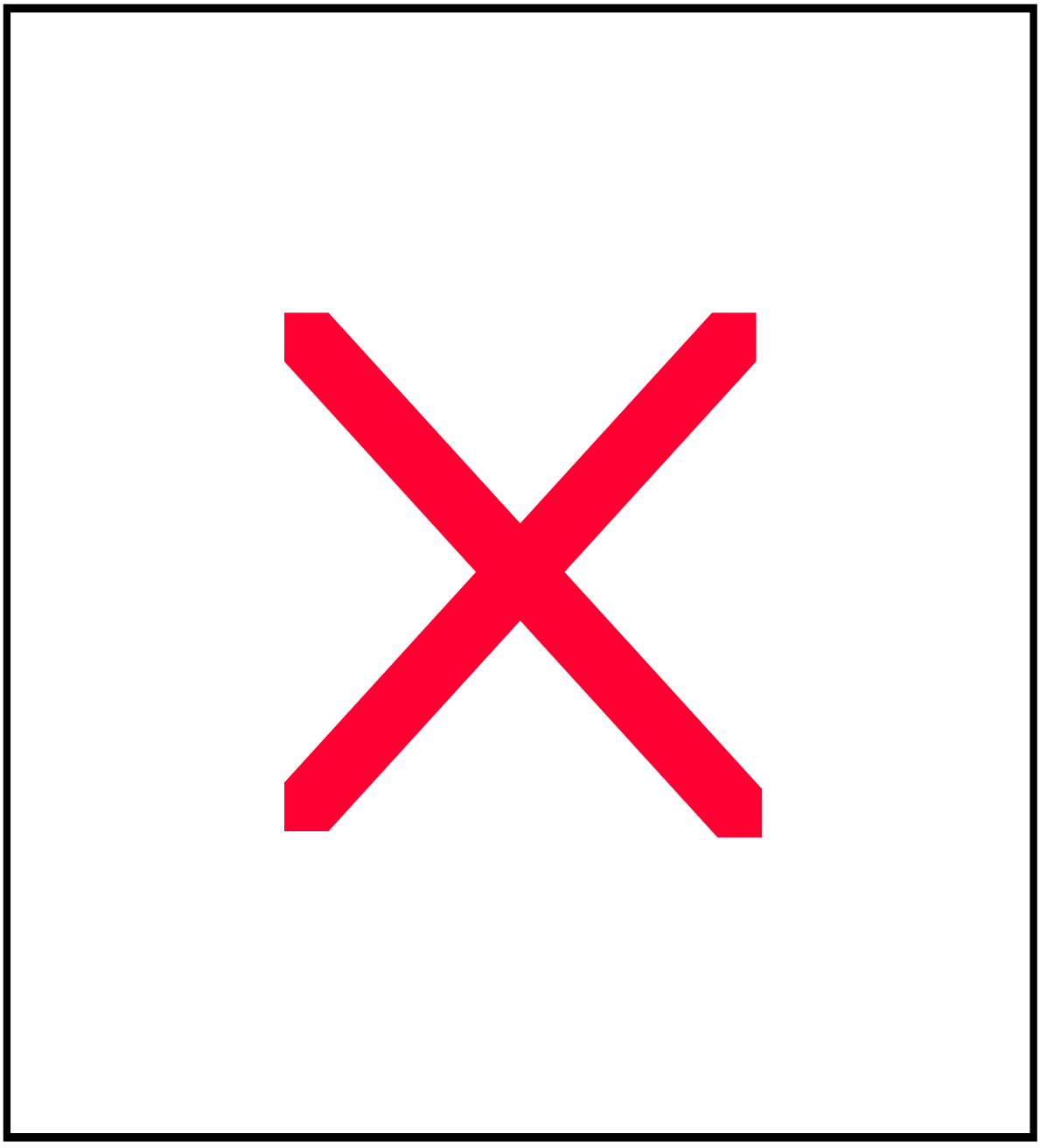
Similarly, Softek Replicator supports the heterogeneous storage environment that Darwin wants to move towards, whether DAS or SAN. Darwin's IT staff will be able to perform on-going data replication activities irregardless of the storage platform being used, and it will be able to deliver the data replication on whatever storage platform Darwin adopts. Softek Replicator runs on any application server and works with whatever storage the server uses. As a result, it can replicate data between storage platforms of different vendors as easily as it can between platforms from the same vendor.

Storage Replicator is built around the use of IP and is therefore uniquely suited to meet Darwin's complex local and wide area data networking needs. It can replicate data between servers and storage at headquarters or remotely over TCP/IP, providing Darwin with a LAN/WAN solution that will be able to deliver fast time-to-recovery objectives for critical application data both locally and remote.

In addition, Softek Replicator provides a management console (Common Console) that Darwin's IT staff can use to perform and monitor the status of any Softek Replicator task from one place. The other Softek products recommended here, as well as the Softek Replicator Common Console, all have the ability to be launched from the Softek Storage Manager console. This eliminates the need to acquire and learn separate management consoles. The ability to view at a glance the state of a given storage process and its associated performance enables administrators to identify and troubleshoot problems easily and quickly.

Finally, the Softek solution will allow Darwin to achieve its data protection goals while preserving its budget. With Softek Replicator Darwin will be able to control hardware, software, and staffing costs. It will be able to shop the market for the best deals on storage and servers. It will experience improved productivity of storage administrators by deploying a single, easily managed solution solving the company's data protection problems through local and remote data replication. Softek has designed its solutions to integrate easily into the existing environment; Darwin will experience immediate benefits without the need to replace any of its existing hardware or software.

DARWIN PHASE 1 – Code Named: Clone Wars



Please note: There is no requirement that the storage at what will become the remote disaster recovery site be of the same make/model or from the same vendor as that at the primary sites. Darwin is free to acquire storage technology from any vendor or combination of vendors.

The Phase 1 Darwin solution is described in detail below. Following this description, we will provide more detailed descriptions of Phase 2 and 3 solutions.

Architecture

The following responses refer to Softek Storage Replicator, the primary product used in our proposed solution. Information on the other Softek products used in the Darwin solution is provided at the end of the document.

1. **Please describe the components or elements of the solution you are describing to Darwin's Groceries. This should include:**

    a. **How you will deliver data replication across an 80-mile distance with, optimally, only five minutes of difference between production and recovery data sets. If a greater delta (difference in data) will be produced by your solution, explain why this is the case and what Darwin might be able to do to address the situation.**

Softek Replicator replicates data using TCP/IP over any distance, including the 80-mile distance facing Darwin. Using asynchronous replication Softek Replicator allows a delta of 5 minutes or less between sites *once* the initial copy of the data is made. To achieve this speed, Softek Replicator allows Darwin to set the chunk size to be sent over the LAN/WAN and replicates multiple logical groups at once. Once the initial copy is made, Softek replicates only the changes or delta by tracking data written to the local or primary disk and mirroring it to the remote disk. This ensures synchronization between the primary copy and the remote replica. Softek Replicator automates the mirroring of data. After the initial full replication, updates can be made quickly and frequently.

    b. **How you will enable the transition (fail-over) of application access from the production storage infrastructure to the remote backup infrastructure (optimally) within 30 minutes of an unplanned interruption in access to, or proper operation of, production storage.**

Softek Replicator's Symmetric Configuration function enables a local and remote server to be failover servers for each other. In the event of a problem that interrupts the operation of the primary server, Softek Replicator can participate in a script that activates the fail over to the remote server. The Symmetric Configuration would also allow a fail back to the original server once the primary site is restored. In order to simplify the process of fail-over when the primary site fails, a simple script is prepared to automate the restart on the secondary system.

    c. **How you will provide security for Darwin Grocery data from the point of creation, during transport, and while stored on production and recovery platforms**.

Storage Replicator operates in a private secure network. This is what we recommend for Darwin. By mirroring data at the block level, Softek Replicator makes it nearly impossible for someone to access and compromise the data during replication.

    d. **How your solution can be subjected to tests without disruption of normal operations.**

Softek Replicator can pause the replication at any point, creating a point-in-time copy at either the local or secondary site. If this is done in conjunction with application consistency points, a full copy of the database will be available for running reports, testing or other activities that cannot be performed against the production database. In most cases, replication can simply be resumed without making a full copy of the data once again.

e. Details of the specific support of your solution for various storage infrastructure components deployed in the Darwin Groceries headquarters data center.

Since Softek Replicator is a host-based replication tool residing on the application server, it can replicate live disk-based data from any volumes on a primary system to disk volumes on a secondary system, regardless of storage vendor.

f. Details of the specific support of your solution for the operating system and application software environments used at Darwin Groceries HQ.

Softek Replicator automatically installs a special device driver, just above the actual disk device drivers or volume management device drivers, but architecturally below filesystems or applications. As a result, any disk-based filesystem supported by AIX, HP/UX, LINUX, Solaris or Windows NT/2000 is compatible with Softek Replicator as well as applications that work directly with raw disk volumes.

g. Details of any topology or hardware changes required (or recommended) to implement or facilitate the benefits of your proposed solution.

The Softek approach assumes Darwin will expand its storage hardware to provide the additional capacity required for replication. Softek does not require any specific vendor's products. Darwin is free to shop the market for the best price and performance.

h. Details of any changes to wide area networks that are required (or recommended) to implement or facilitate the benefits of your proposed solution.

As long as Darwin uses TCP/IP over the WAN and the existing network infrastructure can support the peak I/O update rates of the applications to the remote site, Softek Replicator will require no changes.

i. Details of management capabilities provided as part of your solution, specifically for verifying the proper operation of your solution, alerting Darwin IT managers to error conditions, optimizing the solution for cost-efficient operation especially in terms of WAN costs, and providing audit trails.

Softek Replicator contains powerful management tools that monitor the status of all tasks from a single console. When using this Softek Replicator Common Console, any Softek Replicator command can be issued against any server that has Softek Replicator installed on it. Once a replication is started, each server involved is color-coded to coincide with a particular replication status. Problems and failures are instantly apparent. Servers are assigned to groups as defined by Darwin. All Softek Replicator data is stored in a database, which can be used for historical viewing or reporting.

2. Deployment Issues:

a. Describe the factors that impact the rollout of your solution and discuss the implementation timeframe you anticipate for Darwin Groceries if it selects your solution.

Softek Replicator can be installed and replication groups configured with minimal disruption to the application servers. In fact, Softek Replicator can be installed and configured in advance with no impact on the application servers. During the next planned maintenance, data can be mounted to the Softek Replicator device driver. On UNIX servers, this requires only a brief unmount and mount command. Subsequent phases can follow on a timetable of Darwin's choice.

b. Describe how your solution can be scaled to meet the increasing volume of data generated by Darwin Groceries over time.

Softek Replicator can replicate any size volumes, as well as many volumes defined together as a logical group for DB consistency. It can also replicate many logical volumes at the same time. Softek Replicator has no scalability limitations; the only constraints are those imposed by Darwin's application servers or the network or the storage arrays. Darwin's historic storage growth rate has been 10 TB annually. Softek Replicator can easily handle this level of growth and more.

c. Describe any implementation support services that you offer, including consulting, training, customization, etc. Identify specifically the duration of services (e.g., the length of training) and any additional expense associated with these services.

Softek can assist with deployment of the solution. Typical installations include some post-sales on-site support to get started. If needed, consultation is available at $2,500/day. Darwin's Phase 1 installation would likely require no more than 2-5 additional days.

3. Solution Pricing:

a. Describe your pricing methodology.

Softek offers very flexible licensing. Softek Replicator is priced based upon the number of processors in the application servers. This approach keeps costs to a minimum since companies pay only for the processors they have. Other vendors, by contrast, charge by the amount of data replicated—often a rapidly increasing volume—which results in rapidly increasing software license charges. This is not the case with Softek Replicator.

b. Calculate the cost to Darwin Groceries for your proposed solution including optional components and services.

Darwin has provided no information about the number of processors in its servers. However, the cost for a 4-processor server is $6500. If the Darwin installation required two days of professional services, the additional cost of consulting would be $5000.

c. Identify maintenance or other recurring costs to Darwin after it has implemented your solution.

After an initial 12-month warranty period, annual maintenance is based on 18% of the current list price of the software, which comes to $1,170 annually for a 4-processor server.

d. **Identify any third-party components you have included in your solution and their cost to Darwin Groceries**.

The Softek solution as proposed involves no third-party components other than some additional server hardware and storage capacity as noted above.

4. **Benefits:**

   a. **Describe how your enhanced data protection solution may be differentiated from a tape backup-and-restore solution from the standpoint of**

      i. **Shrinking backup windows**

Softek Replicator allows a complete copy of the data to be made with no disruption to the application. When a full point-in-time copy of the data is desired, you briefly quiet the database for a few seconds for synchronization, which will ensure the logical and physical integrity of the data is maintained. This gives Darwin a complete copy of the application on disk and can be used for immediate recovery if needed without a restore. We do not expect Darwin to discontinue tape backup. With this approach, however, tape backup can be performed against the replicated copy, eliminating the backup window altogether.

      ii. **Reduced time to data (restore) time**

Using Softek Replicator for data replication allows for a complete copy of the data to be maintained on disk if desired, eliminating any need to restore data from a tape. Restoration from disk is far faster than restoration from tape, and more reliable too.

      iii. **Overall solution dependability**

When using Softek Replicator for data replication, the data is replicated remotely in near-real time and is kept in a consistent state at all times. Data is usable at a moment's notice, without the necessity of relying on a backup. Backups can be used solely for recovery from data error if necessary. When replicated to a remote site, the data is protected in the event of a disaster at the primary site.

      iv. **Overall solution cost**

At $11,500 for a 4-processor server, including the likely consulting required, Softek Replicator represents a very competitive solution on the basis of total cost.

   b. **Remote mirroring, one approach for data protection, has always carried with it two deficits in Darwin's view: hardware lock-in and high expense, especially in terms of WAN bandwidth and management. Explain how your solution addresses these concerns.**

Softek Replicator provides data mirroring technology by using the application server as the source of the data replication. The host-based pricing model is very cost competitive and stable, allowing storage volumes to increase without increasing costs. Softek also gives Darwin the benefit of mirroring technology without being locked in to any particular storage vendor and its pricing. Softek Replicator's ability to work with many host/server operating systems and any storage platform also ensures that Darwin can shop the market for the best deal on servers as well as storage. The Common Console lowers the cost of management by making it easy to manage the entire replication process from a single location. The same Common Console also can be used with Storage Manager and other Softek products for even greater cost-efficiency. By replicating deltas after the initial data replication, Softek Replicator greatly cuts down the amount of data crossing over the network for subsequent updates, which reduces the amount of bandwidth Darwin must provide. Similarly, through the data scrubbing enabled by the use of Storage Manager in Phase 2, Darwin will be able to further reduce the amount of data that must be stored and transmitted.

c. Darwin wants a comprehensive business case to offer to management for the solution it selects, one offering not only risk-reduction, but also cost-savings and business-enablement value. Can you describe benefits in each of these categories that derive from your solution?

Not only does Softek Replicator provide continuous data replication to a local or remote site, but the data is also kept in a usable state, ready on a moment's notice, effectively mitigating the risk of a catastrophic failure or disaster. Management is assured that in the event of a primary data center outage, the MiniMarts and SuperGigantic stores could resume operations on the designated DR site with current replicated data ready for use. In addition, Softek Replicator offers the following advantages:

- No vendor lock-in—Softek Replicator is independent of any storage manufacturer, providing Darwin with the ability to negotiate for the most cost effective storage available. Similarly, it works with all the leading operating systems, which allows the company to pick from a multitude of vendors for its servers.
- Leverages existing systems—Softek Replicator can be installed and used in the existing infrastructure, with no additional software or hardware purchases necessary.
- Protects Darwin's technology investment—no matter what server and storage technology Darwin adopts in the future, Softek's flexible, open replication will work with it.

Softek is committed to data replication as a critical IT discipline. With the backing of Fujitsu Ltd. (3rd largest IT supplier in the world and the first software vendor to deliver host-based data replication software on legacy systems) we have the knowledge, experience, and resources to continue to advance the state-of-the-art in data replication.

5. Market Viability

a. Describe how your solution compares with comparable solutions in the market today. (We encourage specific and explicit comparisons to competitive products.)

The Softek Replication architecture gives Darwin the ability to perform multiple data replication tasks with one product, including synchronous mirroring, asynchronous remote mirroring over IP for DR, copies of data for backups or testing or for moving data. By comparison, storage hardware vendors will typically require multiple products to accomplish all of these tasks. In addition, their products will only work on the storage vendor's hardware, making it a propriety solution and effectively locking in the customer. A handful of small, new vendors offer replication products, but these typically are focused on one type of replication task. Also, these vendors are small and lack an established track record. Their long-term viability is questionable.

b. Describe your business model and financial performance to assuage consumer concerns about their investment and your prospects for longevity.

Softek employs a two-tier distribution model. It uses both a direct sales force and a global network of partners to serve customers of any size and at any location. Softek is a private company, a wholly owned subsidiary of Fujitsu Ltd. of Japan (FJ) and does not provide specific financials. However, Softek consistently contributes a positive operating income to FJ and is the global brand for storage management across all Fujitsu operating companies. FJ is a publicly traded company. Its latest financial results can be found at http://pr.fujitsu.com/en/ir/

c. Identify key factors Darwin should consider in its business decision to select and deploy your solution.

The most important factor is ROI. While many products promise a positive ROI, Darwin must examine such promises more closely taking into account additional costs that will delay or reduce the promised payback. For example, what additional hardware and software investments does the solution impose? Will Darwin have to replace existing technology to make the solution work? Will a solution require an extended deployment and services contract? Softek not only delivers the open, flexible replication Darwin requires but does not force the company to replace any of its existing technology and generally requires only a minimal amount of services to become operational and integrated into the existing infrastructure. Furthermore, when Darwin wants to upgrade or expand its systems infrastructure, Softek Replicator does not lock the company into the products of any particular vendor.

Softek offers a complete suite of products to help customers' manager their data and storage, which will be described in Phases 2 and 3. In addition to Softek Replicator, Softek provides Storage Manager, which would allow Darwin to active manage all its storage resources, allowing large amounts of multi-vendor hardware and software to be centrally managed with minimal time and staff resources. Storage Manager eliminates the need for acquiring and integrating numerous point products from different vendors. Other key Softek storage products include SANView, Provisioner, and EnView.

Detail any other information that you consider important to Darwin Groceries' implementation that are not addressed by the questions above.

The above represents Phase 1 of a three-phase solution. The following section describes Phases 2 and 3.

Phase 2

The goal of phase two is to clean up Darwin's stored data at its store locations through the use of Softek Storage Manager as well as extend replication to the remote disaster recovery site. Softek Storage Manager eliminates the need for storage administrators to go from server to server to manually track storage consumption. Instead, Softek software agents at each location collect the data, which is stored centrally at Storage Manager server in Darwin's data center. Softek Storage Manager's reporting capabilities generate customizable reports that can be delivered live or via the web in graphical or tabular format. Through historical and trend, it enables managers to reduce operating costs by intelligently planning future spending.

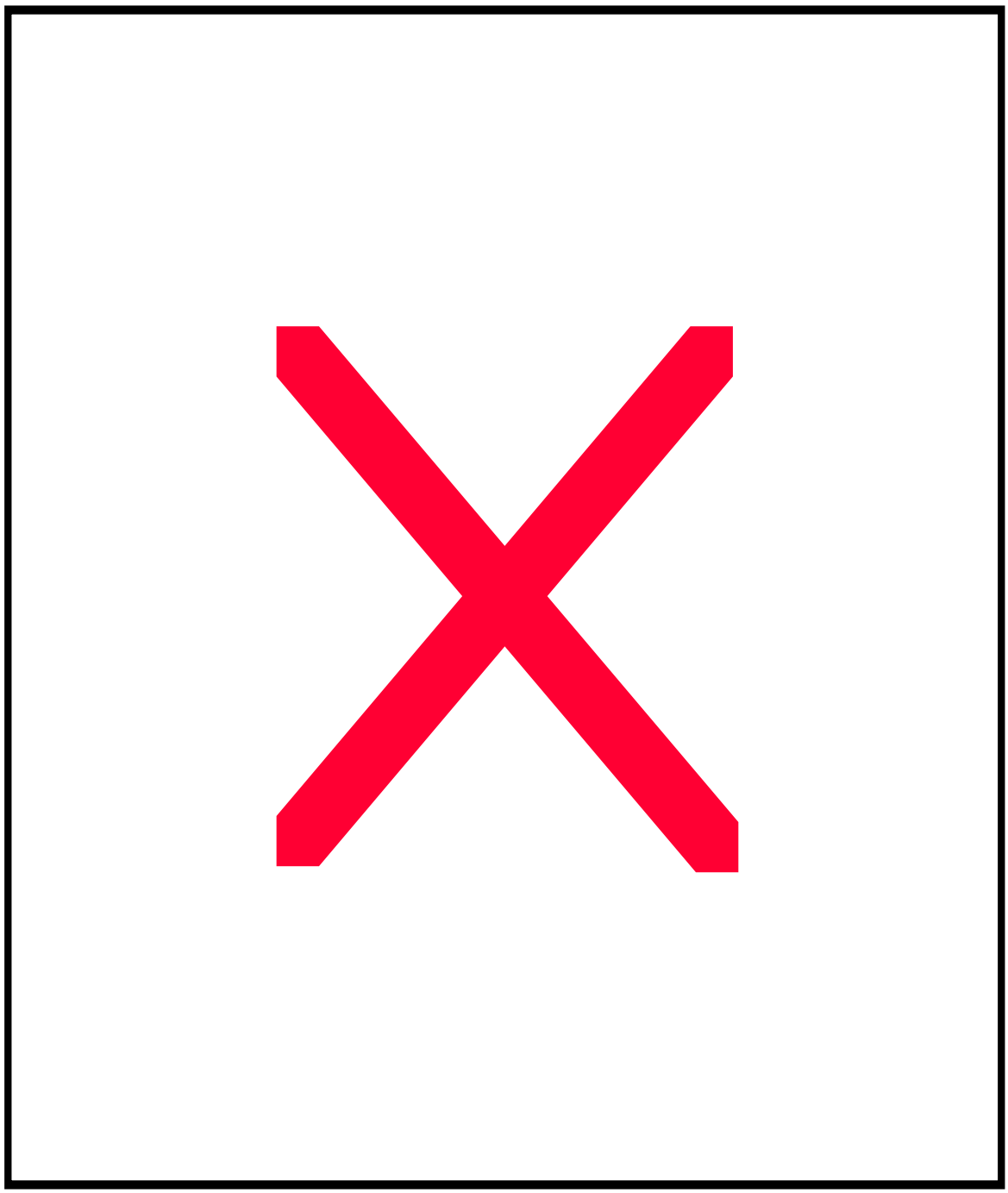
The Softek Storage Manager agent itself is lightweight (about 2K of memory and a maximum 15 percent processor overhead in UNIX environments during full File System scanning). The agent allows for advanced reporting and monitoring of critical file systems within the SAN. Storage Manager agents are non-intrusive and sit dormant during the majority of production operations. They are active only when requested to collect information for file system details, capacity updates, and action execution. In fact, most of the overhead is incurred only during the initial file system scan. Each complete scan of the file system thereafter collects only the changes since the last scan, which greatly reduces the amount of system resources required. These subsequent scans take perhaps 1% of the time to complete. Memory, CPU, and disk resources are utilized with great efficiency and impact is minimal in comparison to operational applications running on the application server.

Storage Manager's ability to capture and analyze usage patterns makes it indispensable during Phase 2. Darwin will use Storage Manager at its store locations to identify data that is no longer needed. In this way, it eliminates redundant, unnecessary, and outdated data. This process will reduce the amount of data Darwin needs to store and transmit, enabling the company to achieve better utilization of the storage capacity and bandwidth it currently has and to defer the acquisition of additional storage and bandwidth.

Architecturally, Storage Manager is simple. The product consists of a server, an optional lightweight agent, and the graphical Common Console for management. The server runs the data repository, hosted by any SQLServer database. Agents, which can be installed on any host you wish to manage, scan the File Systems and transfer the data to the repository. The agent itself is non-intrusive and installs without a reboot on all OS types. The Storage Manager Server data repository is typically striped across two or more disks for performance and requires a Windows 2000 server running SQL / Oracle, IIS and SNMP. The management console uses a lightweight Java client GUI to provide access to Storage Manager and its data repository.

All storage, whether SAN, NAS, or DAS across mainframes and servers with HP-UX Win/NT, Solaris, and AIX can be viewed and managed from a central console with Storage Manager. Although Softek Storage Manager does not technically require installation of agents to view and manage file systems, it does benefit from their deployment and we recommend it for Darwin. The primary functionality of the product's core feature set can be accomplished without the use of any agents by remotely mounting file systems onto proxy management servers, which could be the Storage Manager server itself.

DARWIN PHASE 2 – Code Named: Scrub and Send



All storage, whether SAN, NAS, or DAS across mainframes and servers with HP-UX Win/NT, Solaris, and AIX can be viewed and managed from a central console with Storage Manager. Although Softek Storage Manager does not technically require installation of agents to view and manage file systems, it does benefit from their deployment and we recommend it for Darwin. The primary functionality of the product's core feature set can be accomplished without the use of any agents by remotely mounting file systems onto proxy management servers, which could be the Storage Manager server itself.

Softek software pricing for Phase Two is as follows:

Softek Storage Manager - $39,000 + $795 per managed server

The payback from Phase Two can come very quickly. Storage Manager can immediately free up storage capacity by identifying redundant, outdated, and unnecessary data that can be removed without affecting operations. This will result in immediate savings in storage capacity and allow Darwin to delay additional storage purchases. Its reporting, trending, and analysis capabilities will enable Darwin to forecast future storage needs, eliminating the need to over-provision storage.

Once Darwin has implemented Phase 2, it is then positioned to move its cleaned and scrubbed data to a centralized consolidated SAN should its data volume and capacity trends so indicate. This consolidation is the goal of Phase 3.

Phase 3

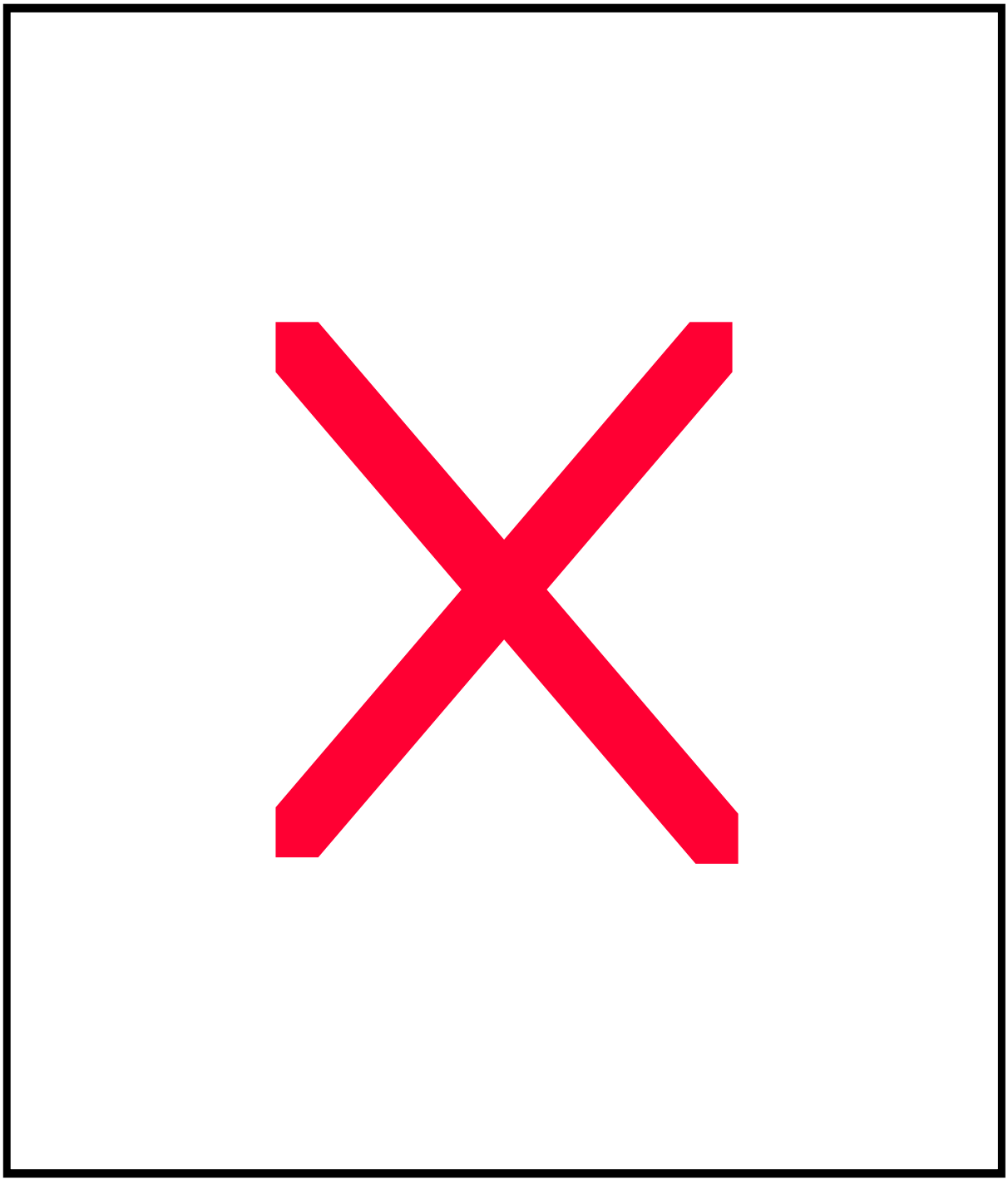
When Darwin's data volumes as collected by Storage Manager in Phase 2 indicate that the company's data volumes are growing sufficiently fast, it is time to move to Phase 3, SAN consolidation. In Phase 3 we recommend that Darwin implement a SAN in its central data center, acquiring the necessary switch and storage hardware. Softek's solutions are hardware vendor independent, which allows Darwin to shop the market for the best SAN components to meet its needs in terms for capacity, performance, and price. Although Softek solutions work equally well in heterogeneous storage environments, we advise at this stage Darwin purchase a single-vendor SAN solution or one with vendor certified compatible components to reduce SAN implementation complications.

Once the SAN is implemented, Darwin can use Softek SANView, Provisioner, and EnView along with Replicator (which it already has, from Phase 1) and Storage Manager (from Phase 2) to capture the benefits of SAN consolidation.

Softek SANView provides both real time and historical traffic and performance monitoring and analysis of connections and devices spanning from storage to the host and from edge through core to edge of the fabric to enable Darwin to monitor and manage its SAN at the storage infrastructure level (Storage Manager manages the storage at the data and file system level). SANView, in conjunction with Storage Manager, enables Darwin to define critical paths and monitor performance through the use of thresholds that identify potential link failure (error monitoring) and potential throughput bottlenecks (blocking or SAN congestion). This allows Darwin to reduce or eliminate downtime due to SAN infrastructure issues. Furthermore, Darwin can use this information to specified proactive automatic alerts and triggers based on defined thresholds or specific configuration changes that alter or stress the SAN. Such triggers can launch applications, scripts, emails, and actions, thereby keeping applications running and speeding problem resolution.

SANView can manipulate and maintain devices in the fabric down to the port level and effect changes in both WWN and port based zoning. Due to integration of CIM / SMIS, SANView will be able to take advantage of storage device-provided controls for common console (single console) manipulation of LUNs and LUN assignment within the SAN fabric. These features are planned for near term releases of this product and will certainly be available before Darwin gets to Phase Three.

DARWIN PHASE 3 – Code Named: Shoe Horn



Architecturally, Softek SANView uses a simple data collector-server relationship with optional agents for collecting additional information. The server resides on a dedicated management system, which can be combined with Storage Manager, Replicator console, and Provisioner console. This platform is Windows 2000 Server. In simplest form SANView deploys on a single machine with client and server residing on the same system. Because of this convenience SANView can be used in the data centers' management network or on a roaming laptop for things like supervising hardware installations.

SANView will export data from its server API for use in other framework applications, such as Tivoli. Softek SANView maintains a database (MySQL comes with SANView although Darwin can continue to use its Oracle database.) that stores not only historical configuration data but performance and detailed traffic information as well. Topology snapshot and image saving allow for comparison of a previous configuration to a later configuration after changes have occurred. SANView allows the SAN configuration to be saved to a file and have the topology recalled automatically across stop/start instances of the client GUI. SAN device configuration data also is stored and restorable. Any SAN device's labeling and description characteristics can be customized and annotated to fit the particular environment as well.

The vast majority devices used in storage networking are fully supported by the Softek solution through one or more modes of discovery and control. Most SAN devices today are FAMIB compliant, and or provide information to our application via other methods such as the Common HBA API, Switch API, common device API, or others. We will be adding the capability of Storage Configuration and LUN assignment for instance, in the next major release of Softek SANView through the implementation of CIM/SMIS (Common Information Model / Storage Management Interface Specification) making use of "Providers" integrated into storage devices.

SANView optionally uses proxy agents to learn more about storage devices and the HBAs in the application servers. This is accomplished by making calls into the common HBA API (multi-vendor collaboration HBA API) built into most HBA drivers today. This is very unobtrusive software residing on the server with no root level access. Softek will use CIM/WBEM and eventually SMIS for remote host (and device) access and control. SANView does not require an agent on the storage devices or additional agents on application servers containing no HBA.

Provisioner is a software product that provides automated storage device management. It consolidates disparate storage arrays into a centralized storage pool and serves capacity on demand to application servers in a heterogeneous environment. Provisioner is integrated with Softek Storage Manager to provide policy-based auto provisioning. This combined solution matches the logical requirements of the application with the physical characteristics of the storage to deliver the right storage automatically without storage administrator intervention. It runs on any Windows/Intel platform Darwin chooses and operates as an in-network, in-band virtualization appliance.

Provisioner provides a Capacity-on-Demand (CoD) capability that establishes a pool of physical disks from which virtual volumes can be created. The created virtual volumes can be of any size up to the largest supported by an application host-operating environment, 2TB in the case of Windows. The physical pool supporting these large capacity virtual volumes can be of a much smaller size. The physical pool backing up these logical volumes is monitored and managed for real usage and capacity automatically added based on crossing a configurable threshold. This means that the application host can be satisfied with a maximum volume size once, and physical storage is added to the underlying pool only as needed. Application hosts do not have to be disrupted to add additional capacity, and no capacity is stranded or wasted. Various multiple pathing options support alternate pathing and cluster pathing to the application hosts. Combined with multiple nodes and mirrors, and possibly dual path to storage, all single points of failure can be eliminated in the storage network. Thus the application host resources and physical resource allocation are decoupled. The result: business is not constrained or disrupted by limited storage capacity yet there is no stranded or wasted capacity and the availability of the data is assured.

Along with consolidating its storage on a SAN, Softek recommends Provisioner be configured for High Availability for critical data. This configuration would involve mirroring each volume between two Storage Provisioner nodes. Each node maintains a full configuration information file to maintain full availability. Virtual volumes can be mirrored once or twice. Snapshot point-in-time copies can be created. Any number of snapshot pairs can be configured and activated or deactivated at any time. Snapshot copies can be full copies or copy on write. Thin provisioning can be accomplished using Capacity on Demand Sets (CoD Sets). CoD Sets are a pool of physical disks from which virtual volumes can be constructed. These virtual volumes can be up to 2TB in size as seen by the application host. The physical pool that uses these volumes can be much smaller than the total of the virtual volumes served.

Finally, Darwin may choose to opt for Softek EnView, an early-warning system that allows organizations to measure the performance, availability, and response of their critical business processes, whether mainframe, client/server, or web-based for purposes of QoS or SLA compliance. EnView provides a proactive means to track and remedy response-time requirements and ensure 24/7 availability of all critical applications. Unlike traditional performance monitors, Softek EnView offers continuous measurement of end-to-end response times—just the way end users see it. The solution is non-intrusive and requires no software to be installed on end-user machines or the sites being monitored. By improving IT service to business users, Softek EnView helps companies reduce costs, improve productivity, and implement proactive service level management.

Softek software Phase Three pricing is as follows:

> Softek Storage Provisioner - $25,000 per storage provisioner engine (2 required for mirroring)
> Softek EnView - $19,500

With Phase 3, Darwin will be able to experience all the benefits inherent in a consolidated SAN including reduced cost of ownership through centralized, storage consolidation. It will be able to take advantage of capacity on demand to ensure business applications have sufficient storage without the need for costly and wasteful over-provisioning. Through centralized, GUI-based management, a small number of administrators can efficiently administer growing volumes of storage. Darwin is assured that the latest data is fully replicated and protected offsite in the event of a disaster while it maintains its tape backup for archiving and other purposes. Finally, the Softek solution does not lock Darwin into any particular storage or server vendor's products.

Review Criteria

1. Solution capabilities and suitability to requirements in the RFI
2. Deployment challenges and services
3. Business benefits of solution
4. Price of solution