Network Computing Request for Information Data Protection through Replication

Deadline: November 5, 2003

Introduction

Network Computing will publish a feature on data protection in our **January 22, 2004 issue**. A major element of this review will involve our analysis of vendor responses to a Request for Information (RFI) submitted on behalf of mythical retail company Darwin's Groceries.

You have been selected as a vendor that Darwin Groceries would likely consider as a data protection solution provider. To participate, you need to answer the questions and address issues that make up the deliverables portion of this RFI. If a question is not applicable or relevant, please say so.

Products that make up your proposed solution should be available by December 2003.

Schedule and Logistics

Please RSVP to Jon Toigo by October 15, 2003, whether or not you intend to participate. **Completed RFI responses must be received** via e-mail, US mail or delivery service by end of day **November 5, 2003**, to be included in the January 22, 2004 edition of Network Computing magazine. You should respond to this questionnaire by entering information directly into this file. Feel free to attach supporting information, but recognize that delivering a concise response will enable us to better focus on the key capabilities of your solution.

If you have any questions about the overall project or specific RFI questions, please contact the author, Jon William Toigo, at

jtoigo@intnet.net or by phone at 727-736-5367 (O) or 727-504-9311 (M).

Please return the questionnaire to Jon via the above e-mail address. Additional materials that cannot be submitted electronically may be sent to him at the following address:

Jon Toigo
Toigo Partners International LLC
1538 Patricia Avenue
Dunedin, FL 34698

REMEMBER: RFI responses must be received by end of day November 5, 2003, to be included.

Vendor response:

Vendor company name: Tacit Networks
Vendor service name (if any): _____
Vendor contact name: Emilio Dabul, Fusion PR
Vendor contact telephone number: 212-651-4216
Vendor contact e-mail address: emilio.dabul@fusionpr.com
Date of submission: November 14, 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

Architecture

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.

We do not do data replication but we do provide online data coherency. Any writes to remote site are immediately written to the data center. Using advanced streaming, differencing and compression techniques, updates to very large data sets can be transferred online in short time frames.

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.

Network independent. Routes from remote stores to data center are failed over to the backup data center at the network level and operation continues immediately.

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.

Point of creation – secure appliance
During transport – can use existing VPN/network security or appliances use encrypted vpn tunnels
Stored on production and recovery platforms – responsibility of NAS vendor
All file access at all remote sites is authenticated and authorized by the backend. Security requests are always passed through to the backend data center.

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

Pilot testing can be accomplished easily by setting up a separate file server at the data center and placing a remote Tacit appliance at one of the remote sites.

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

Network independent, uses tcp and NFS/CIFS. Talks standard protocols. Server appliance acts as just another client to backend NAS.

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

The Tacit solution works with flat files and, doesn't run application software. It will not work with database files unless those files are transferred to a flat file format. Talks NFS/CIFS. Standard storage protocols

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

Transparent, no changes

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

n/a

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.

SNMP traps for monitoring operation. Statistics for optimizing cost-efficient operation (not avail yet), limited logs – audit trails

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.

Existing and tacit solution can coexist. Phase approach is very easy.

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

Scalable, many remote appliances to server. Remote appliances mount same share – multiple per bin site.

Server is just another client to NAS. Many servers per NAS up to capability of backend.

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.

Tacit offers consulting, training and deployment services; although the solution is extremely easy to implement end to end within one hour.

3. Solution Pricing:

a. Describe you pricing methodology.

Price is based on remote appliance active connections at each remote site(s) plus the cost of a datacenter appliance server.

Data center appliance cost - $ 21,500
Unlimited User remote appliance - $ 22,500
Maximum 30 User remote appliance -$ 17,500
Maximum 20 User remote appliance -$ 12,500
Maximum 10 User remote appliance -$ 7,500

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

The cost to Darwin which would include: a remote appliance at each store, two sets of active/passive failover appliance servers at the datacenter, three years of maintenance and the first full year of 24 x 7 help desk and software maintenance would be $1,245,000

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

3 year hardware maintenance is $250 per appliance. Enhanced Software and help desk support is 18% of purchase price annually

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

n/a

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

Anything that is written on the remote appliance is automatically back in the data center within in 15 seconds without any user intervention

ii. Reduced time to data (restore) time

Since all data is back in the data center there is 0 restore time. The user just has to request the file through the normal interface and it will appear

iii. Overall solution dependability

All the remote appliances are protected by RAID and the datacenter appliance server is protected with a active/passive failover

iv. Overall solution cost

b. Remote disk-to-disk 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.

The Tacit Networks solution reduces the latency over the WAN to near LAN speeds. There is no need to do disk mirroring all the data resides in the datacenter with a cached copy of recently used files at the remote sites. No increase in Bandwidth is needed and management is simplified to a normal LAN model

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?

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.)

Tacit Networks is often compared to replication/copying (most of the storage vendors- EMC, NetApp, HP) and bandwidth optimization (Perabit, etc). Although our solution is a new ground breaking answer to these band aids that are currently applied.

Replication/copying comparison – This is what was also done in the old days when the IBM PC first came out. It was called sneaker-net or floppy copy back then. Then the LAN and Novell came into the pictures and everyone has a LAN now and replication and copying is no longer required locally. Now antiquated replication and copying solutions are being applied to the wide area file sharing between LANs because as in the past there was no solution till now. Tacit Networks allows WAN file sharing in real-time to operate at near LAN speeds eliminating the need for costly replication (one point-in-time) of full data sets. The users are now always working on one data file that is instantaneously updated globally across the entire enterprise in real-time.

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

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

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

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