SCI StorInt™ Dispatch – EMC080519
EMC World Announcements: DL Deduplication and New Avamar® DataStore

EMC announced in Las Vegas the latest iteration of their Disk Library product line which now includes partner supplied target deduplication technology as well as a new generation of Avamar® DataStore

Disk Library (DL) Target Deduplication

EMC's next version of their high-end disk library (DL 4X00) and the newly introduced low-end DL 3D 1500 and DL 3D 3000 will support target deduplication. DL 3D 1500 and DL 3D 3000 support software oriented deduplication while the DL 3D 4000 will support added hardware to provide deduplication services.

Deduplication technology can be provided at the backup source or target. EMC Avamar® and Symantec PureDisk both provide source deduplication technology, integrated with special backup software that is ideally suited for remote office backup to central sites. Last year EMC announced (see SCI's EMC070904 StorInt™ Dispatch) support for an Avamar client connected to a NetWorker backup server that could provide both remote and central site source deduplication for backup data.

Target deduplication technology supports any software, appliance, or service that writes to either disk or tape as a backup target. Disk targets support NAS interfaces (NFS or CIFS) for backup to disk. VTL technology can attach to any software that writes tape data, which includes all backup systems and has much wider applicability. The low-end DL 3D 1500 and 3000 support both NAS and VTL operations and the DL 3D 4000 only supports VTL. Target deduplication technology is currently available from DataDomain, FalconStor, NetApp, Quantum DXI, Sepaton, and now EMC.

One ongoing concern with this technology has been the processing overhead required to deduplicate data. EMC's offering, supplied by their partner Quantum, can be configured to support offline deduplication for higher data ingest performance or conversely can be configured to support parallel deduplication while data is being ingested. Ingestion throughput may suffer by using parallel deduplication but the capacity required to buffer data while waiting for offline deduplication is minimized.

The DL 3D 1500 and DL 3D 3000 are available immediately and support a maximum of 36TB and 148TB of deduped backup data storage respectively. The DL 3D 4000 deduplication engine will be available in late July and although EMC has not released maximum capacities yet for this product, the current version supports over 670 TB using 750GB SATA drives. EMC has announced support for 1TB SATA drives on the new DL 3D 4000.

Also announced was a new spin down capability for the DL4000 and DL 3D 4000. Particularly useful for backup targets when not active, drives can be spun down to

Silverton Consulting
Strategy, Storage & Systems
P.O. Box 492|Broomfield, CO 80038-0492|
720-221-7270|info@silvertonconsulting.com

minimize power consumption. Combining this with the new 1TB SATA drives will improve power consumption considerably.

Avamar® Enhancements

Avamar's DataStore size has been doubled to 2TB per node. With a 16-node configuration and 500X dedupe compression ratio, a fully configured DataStore could support a PB of backup data. Compression rates are highly dependant on backup cycle and the amount of data duplicated throughout your storage, so your results may depart significantly from projected results.

Announcement significance

Deduplication technology is breaking out all over what with IBM purchasing Diligent, FalconStor solving their deduplication problems, DataDomain still leading the market and the rest of the players trying to keep up. EMC's move to support target deduplication levels the playing field somewhat, at least until the next round.

Silverton Consulting, Inc. is a Storage, Strategy & Systems consulting services company, based in the USA offering products and services to the data storage community

Silverton Consulting
Strategy, Storage & Systems
P.O. Box 492|Broomfield, CO 80038-0492|
720-221-7270|info@silvertonconsulting.com