IBM PROTECTIER AND SAP: CRITICAL DATA PROTECTION WITHOUT DATA DISRUPTION

OCTOBER 2011

SAP is a business-critical application for enterprise and mid-range business. SAP products generate huge volumes of critical data that must be fully protected, the majority of it stored in databases like DB2, Oracle and SQL. These databases are fully backed up once a day or more, which makes for a huge load on traditional tape-based backup infrastructures.

We believe in physical tape as a long-term data retention medium and as a perfectly good solution in some backup environments. But backing up SAP to tape is an entirely different story. No matter how fast a given tape drive may be, this level of SAP database backup is simply overwhelming for physical tape.

The most common solution for these environments is virtual tape libraries (VTLs). They replace physical tape with disk-based virtual tape cartridges that do not require significant changes to backup application settings and scripts. In simple backup environments a commodity VTL may serve, but not in SAP backup environments with their strict service level agreements (SLAs) for backup and recovery. VTLs serving the SAP market must meet stringent requirements:

- **The VTL must be high performance.** Meeting backup windows and recovery time objectives (RTO) is a constant challenge in SAP backup environments. Both backup and recovery must be high performance to meet attendant SLAs. The VTL should ingest backup data at a very fast rate while maintaining complete data integrity and should also restore quickly and reliably.

- **The VTL must be high capacity.** Disk-based backup is faster than tape but it is considerably harder to slot in new disk than new tape cartridges. The VTL must be able to store rapidly growing data without the need for constant reprovisioning. This requires high physical capacity appliances and/or VTL gateways with scalable backend storage, and deduplication for much higher virtual capacity.

- **The VTL must offer excellent data protection.** Restoring large databases from tape takes a day or more but only an hour from the virtual cartridge. DR for critical SAP data also requires that the VTL replicate data to a remote site. In this case restoring a replicated data set will make all the difference between a small business setback and a major data disaster.

This Solution Profile will describe IBM ProtecTIER's fast performance, scalable capacity and replication for SAP backup environments. We will also describe two real-life SAP and ProtecTIER deployments to see how IBM is meeting SAP backup needs in even the most challenging environments.

SAP and the Storage Impact

SAP offers customers a wide range of business operations products. The flagship system is SAP R/3, an online transaction processing system for Enterprise Resource Planning (ERP). SAP has also invested heavily in additional applications that offload intensive data workloads and has built out Internet-based communications and data delivery networks. Databases are an integral component of most SAP applications, data and operations. With business processes numbering in the hundreds and thousands, these processes generate a huge amount of business data that is stored in databases.

Industry professionals tend to throw around the "critical data" buzzword but in SAP's case it absolutely applies. The data stored in SAP databases is essential to the business and must be fully protected. This data is growing and it is growing fast: large SAP users are commonly experiencing year-over-year growth rates of 30 to 60 percent of their SAP data. Much of this data growth is occurring in the core records retention and business intelligence programs. But SAP is seeking to retain and grow market share by expanding from their core ERP suite to include a host of management and collaborative platforms. This level of data generation is producing hundreds of TB of disk storage and is reaching into PB levels.

SAP STORAGE PAIN POINTS

Pain Point	Why It Hurts
High storage costs	Storage costs are radically increasing across the enterprise and in SAP environments. Poor storage practices result in infrastructure strains, higher management overhead, budgetary pressure and the risk of inflated backup windows and poor recoverability.
Failing SLAs	Storage administrators frequently backup large SAP databases; many of these are full backups. Tape-based backup and restore tape threatens SAP backup windows and RTOs.
Extreme data growth	SAP offerings and adoptions are increasing, leading to high rates of data growth. Traditional tape-based backup is toppling under the load of increasingly large database backups.

IBM Storage and SAP

IBM meets these difficulties with a large selection of storage options for SAP environments. Let's take a brief look at IBM and SAP's long-term partnership and then move to a discussion of ProtecTIER in these SAP data protection.

SAP was founded nearly 40 years ago by former IBM engineers and the two companies have partnered ever since. IBM offers a variety of system storage solutions across the SAP Business Suite applications. Storage offerings include IBM System Storage DS8000 series for data-intensive enterprise SAP environments, IBM XIV Storage System with its open disk technology and economy, and the IBM System Storage DS6000 series. Mid-range sees the DS5000 and entry-level DS3000 series. IBM also provides NAS with System Storage N series N6000 and N7000 systems and the mid-sized IBM System Storage N3000. Tape storage systems include LTO with LTO-5 optimized for capacity and performance. All of these storage types have successfully served in SAP storage environments, including LTO-5 for tertiary long-term backup retention.

PROTECTIER AND SAP

ProtecTIER is the essential storage step that optimizes SAP and storage resources. The VTL dramatically reduces backup windows by replacing tape with high performance backup disk and improves capacity usage with high deduplication ratios for SAP databases.

ProtecTIER accepts incoming backup directly or as a secondary storage target from disk-based storage and applies inline deduplication and compression. Processing is extremely fast using an efficient index that resides in memory. Only byte-level changes are stored. Its processing performance achieves throughput rates of up to 2000 MB/second performance and data recovery is even faster at 2800 MB/second sustained performance.

ProtecTIER uses HyperFactor data deduplication technology. HyperFactor algorithms work by comparing incoming data to previously stored data on the disk repository. HyperFactor does not use hashing so does not require exact data chunk matches, which results in greater accuracy and efficiencies for incoming database data. HyperFactor deduplication can result in 25:1 compression ratios for databases, which are well suited to deduplication. Disaster recovery is essential in SAP environments. ProtecTIER replication operates one-to-one, many-to-one and many-to-many to provide complete flexibility for data protection and disaster recovery.

Replicating deduplicated data relieves the intensive load on LAN and WAN bandwidth and saves on secondary site storage capacity. Replication occurs simultaneously with backup which allows IT to fully protect the deduplicated data before it even enters primary ProtecTIER storage. Users have the option to clone replicated data sets to physical tape at the secondary location. The deduplicated data cuts down dramatically on the number of physical cartridges required.

ProtecTIER is available in a wide variety of gateway and appliance models for open systems, IBM System *I*, and IBM System *z* mainframe environments. The appliances offer 4 to 36TB useable capacity for mid-sized environments and are available in single or two-node cluster configurations. The gateway's backend storage can scale to support up to 1PB physical storage capacity, which represents 25PB or more backup data behind a single ProtecTIER system. The gateways can be clustered for additional performance and high availability.

For example, a typical configuration is a ProtecTIER appliance with 36TB internal disk storage. The appliance acts as the direct storage target for the SAP database backup application. A large-scale environment might deploy the 100TB ProtecTIER gateway with backend IBM ESS storage units. In this case TSM backs up the databases to a DS8000, which in turn backs up to the ProtecTIER. The ProtecTIER can also be set to save backups to LTO tape for long-term data retention and regulatory compliance.

BEST PRACTICES FOR PROTECTIER AND SAP

When backing up SAP databases to the ProtecTIER, disable any compression, multiplexing and encryption operations that may be running in the database backup tool such as Oracle RMAN or MS SQL LiteSpeed. This is why:

- **Compression**. Compressed database backup data has fewer duplicate blocks than uncompressed data, making it difficult for HyperFactor to match patterns for deduplication. ProtecTIER will compress the data itself after deduplication for further space savings. Compressing data a second time can also can expand data, not exactly what the storage administrator wants to see. Note that this only refers to backup database compression, not to space-saving database server compression.

- **Encryption.** Encryption renders backup data unique and thus useless for deduplication. Disable backup server encryption features for ProtecTIER.

- **Multiplexing.** Do not use database backup tool's multiplexing feature with ProtecTIER storage. Multiplexing works by opening channels for multiple storage streams to physical cartridges but interferes with deduplication in the ProtecTIER VTL. Multiplexing changes the backup data by threading blocks from different files so ProtecTIER can store but not deduplicate the multiplexed data.

OPTIMIZING DATABASES FOR PROTECTIER

Database	Recommended Settings
DB2	Turn off backup multiplexing, compression and encryption. If performance suffers, tune backups for maximum buffers, minimum parallelism and minimum sessions.
SQL	Disable compression, multiplexing and encryption in SQL Server and any SQL tools such as SQL LiteSpeed. Also check the following for these settings: SQL Safe, Ultra Base and SQL Backup 3.
Sybase	Disable Sybase settings for backup compression, multiplexing and encryption.
Oracle	Disable compression, multiplexing and encryption in Oracle RMAN. Increase MAXOPENFILES and PARALLELISM if performance suffers.
Informix	Disable backup agent compression.
MaxDB (SAP DB)	Disable backup multiplexing, compression and encryption as you would for SQL.

ProtecTIER and SAP in the Real-World

We've talked about technologies and best practices. Let's take a closer look at real deployments and how ProtecTIER is optimizing critical SAP backup.

#1: BALDOR ELECTRIC

Baldor Electric Company designs, manufactures and markets industrial electric motors, drives, generators and power transmission products all over the world. The 7000-person company, a division of Swiss-based ABB, is headquartered in Arkansas. Its global activities and large customer base require Baldor's IT department to protect a vast store of critical data, most of it in SAP run on an IBM mainframe.

The company originally backed up to physical tape but wanted to switch to virtual tape for improved backup and restore. The crucial factor was that the VTL must support the IBM System z10 mainframe running a large SAP deployment. The company made a huge investment in SAP software with many modules supporting manufacturing, marketing, engineering, accounting, and more. They divided the z10 into multiple logical partitions running different operating systems. Each of these partitions was further divided into production, testing and data sharing and failover instances. The storage infrastructure includes multiple System Storage DS8000 systems for primary storage. The DS8000 backed up to physical tape, which had become a serious bottleneck to the backup process.

Baldor adopted an IBM TS7680 ProtecTIER Deduplication Gateway for IBM System z with a back-end IBM XIV Storage System containing 36.4 terabytes of physical storage space. Since ProtecTIER returns a 14:1 database deduplication ratio on the SAP's DB2 databases, the virtual space is much larger. The new deployment also reduced backup times by 40% and restore times by 50%. Following the ProtecTIER installation, Baldor phased out its physical tape within weeks.

Baldor puts the ProtecTIER through its paces by backing up five production databases twice a day and retaining the data for 14 days. Space savings are significant: on a backup size of 152 TB only 10 physical TB have been used on the gateway's backend XIV. With the speed of inline deduplication, general backup times went down by 40% and restore times went from 8 hours to 5. Backup times on the critical SAP production databases reduced from 4.5 hours to 2.5 – a huge time savings. The backed up data simultaneously replicates to the remote site and finishes within 30 minutes of the completed backup task at the data center. In addition to all this, the company also freed up about 100 square feet of valuable data center space.

The SAP core system is on a single DB2 database. Eight additional modules include Process Integration, Business Intelligence, Integrated Planning, and more. Each has its own DB2 production database and its own test/dev sandbox databases. The DB2 databases total about 40, all of which backup through the ProtecTIER gateway.

Baldor dumps the production databases twice a day using FICON channels for backup. When they first went to the ProtecTIER Gateway they used IBM TotalStorage Enterprise Storage Server (ESS) 8300 as the backend storage. But the ProtecTIER was accessing the backup data so fast that it degraded the ESS cache I/O. The company upgraded their backend storage to the ESS 8700 which solved the problem. Now Baldor is upgrading again to the IBM DS8800 with SSD for the production database backup which will prove even faster. All production systems are flash copied using ESS' internal copy mechanism then the full copied volume is dumped to the ProtecTIER. Test systems alternate between flash copies and dumping data straight to ProtecTIER.

Baldor is now adding an IBM System z196 to the z10 in order to build a dual footprint for better redundancy coverage. IBM DB2 data sharing will accomplish the data sharing between the two mainframes. Baldor plans to keep the production load on the newer z196 and the test dev/sandbox systems on the z10. The new z196 will come with 8Gb FICON. Once Baldor hooks up the z196 channels to ProtecTIER's 8gb FICON port the already excellent throughput will improve significantly.

#2: ITELLIGENCE OUTSOURCING INC.

Cincinnati-based itelligence Inc. hosts SAP environment for mid-market and large enterprise customers. itelligence Inc. is part of the SAP services service provider itelligence AG that provides software, consulting, outsourcing and application support for SAP customers on a global Basis. itelligence is one of only a few SAP certified global hosting partners.

itelligence hosts customer deployments in their secure data centers. As a hosting provider for SAP, itelligence must provide strict service level agreements around availability, optimized infrastructure and strict business continuity. The company found that its physical tape backup was no longer feasible in light of 200TB hosted data and daily full database backups. Backup windows were pushing their boundaries even with added tape drives and fast tape libraries. And storing tapes in off-site vaults threatened RTOs in case of a disaster. It was even hard to make immediate restorations from the main tape library since tape drives were often engaged in backup throughout the day.

itelligence replaced their tape infrastructure with an IBM ProtecTIER 7650 appliance containing 36TB physical capacity. The SAP databases – primarily DB2, Oracle and SQL Server – mostly back up directly to the ProtecTIER. Some specialized instances use TSM to backup to ESS storage first and then to ProtecTIER. ProtecTIER's deduplication provides up to 900TB or more of storage capacity.

itelligence has since added two more ProtecTIER appliances: two of them in the Cincinnati data center and another in a second itelligence data center. This appliance serves as the replication target for full redundancy and fast recovery. In fact, itelligence set out to use ProtecTIER's native replication for DR as a competitive differentiator. This sets them apart from their competition that is still using physical tape backup storage.

itelligence experienced an immediate 40% backup window time savings. There has also been a dramatic improvement in recovery times. What took physical tape a day to restore now takes an hour with the ProtecTIER. itelligence can do disk-to-disk copy instead of locating tape and then starting the slower restore process. ProtecTIER also replicates backup data to the remote data center; deduplicated data lowers WAN bandwidth requirements and IBM's non-hashing technology provides the highest levels of data integrity. Itelligence plans to add more ProtecTIER appliances as hosted SAP data grows.

ProtecTIER Benefits for SAP Environments

- **Provide data protection for critical systems.** SAP applications generate critical data that must remain fully protected and highly available. Service level agreements support fast backup for large database protection and very fast and granular restore rates for disaster recovery and information governance. ProtecTIER enables backup rates up to 2000 MB/s or higher and restore rates of up to 2800 MB/s or higher from individual files to entire backup sets. This level of backup and restore speeds goes far beyond what physical tape can provide.

- **Scale storage capacity to petabyte levels.** With the amount of data that SAP generates, even differential backups can be very large. ProtecTIER deduplication allows SAP databases to write to virtual tape located on high capacity disk-based storage.

- **Consolidate to reduce energy usage.** Energy costs from power and cooling are growing and some data centers cannot expand because of a limited power grid. The less energy that storage takes, the better the data center can operate. ProtecTIER deduplication reduces the amount of stored data which in turn significantly lowers energy usage. ProtecTIER's deduplication ratio is particularly effective with SAP databases such as DB2 or Oracle, which can realize up to 25:1 deduplication ratios – a dramatic savings in energy usage, costs and risk.

- **Maintain data integrity.** Mission-critical SAP requires full-blown data protection at primary and secondary storage systems. ProtecTIER does not use hashing for deduplication, which eliminates the problem of hash collisions and protects data integrity. And ProtecTIER replicates data using highly customizable options so customers can optimize their SAP data protection plans.

Taneja Group Opinion

SAP products are powerful applications but they can have a strong negative impact on tape-based backup infrastructure. Tape remains attractive for long-term data retention but even fast tape libraries cannot scale to meet immediate SAP backup and recovery requirements. This situation will get progressively worse as SAP databases grow.

However, when companies consider SAP purchases they calculate TCO for purchase and server hardware expenses but will often forget the large cost and risk of SAP storage in a tape-based environment. Critical SAP data requires extremely strong SLAs for backup and recovery, and IT must provide those guarantees without making radical and expensive changes to the backup infrastructure.

ProtecTIER solves the problem by providing high speed, capacity and data protection to SAP environments -- without an expensive overhaul of the backup storage system. IBM ProtecTIER is an immensely attractive solution for protecting SAP data and its meeting crucial SLAs. SAP owners cannot take chances with their data, and should not wait until broken backup windows cause a disaster. Take a close look now at replacing tape for your backup and recovery needs with the ProtecTIER deduplication suite of solutions.