Storage efficiency improvements to help manage data growth while cutting costs

Stop storing so much, move data to the right place, and store more on what you already have

Executive summary

- Rapid growth in data volumes – in particular, unstructured data – is putting significant pressure on capital and operational budgets, storage performance, backup and disaster recovery solutions
- IBM® proposes three simple remedies: stop storing so much, move your data to the right place, and store more with what you have on the floor
- Data compression and de-duplication solutions from IBM can help significantly reduce the total storage capacity you require – cutting costs, improving performance and accelerating backup/recovery
- IBM automated data tiering makes your most frequently used data available at higher speeds, while reducing the cost of storing less valuable data
- With storage virtualisation and thin-provisioning technologies from IBM, you can significantly improve disk utilisation, cut capital and operational costs, and enable non-disruptive growth and migration.

Enterprise storage is under unprecedented pressure. Gartner is projecting 650 percent growth in enterprise data volume over the coming five years, 80 percent of it in the form of unstructured data.[1] Meanwhile, the costs for managing and powering the data storage infrastructure are already rising fast.

The combination of rising costs and data volumes represents an operational time bomb for businesses. To address the challenge and create a sustainable storage infrastructure for the future, businesses will need to get smarter about what they store and how. This will mean identifying the data you really need and storing it as efficiently as possible on high-performance storage, while simultaneously moving lower-value data to less costly storage. It will also mean squeezing the greatest possible value out of existing and future investments in capacity. In plain English, you will need to stop storing so much, move your data to the right place, and store more with what you have on the floor.

Stop storing so much

Even in businesses that have careful controls over which data is stored, there are likely to be multiple redundant copies of the data, and the data itself is likely to take up a larger amount of space than is strictly necessary. IBM offers two key technologies to help you reduce the volume of data stored and claw back capacity for both live and backed-up data: data de-duplication and data compression.

As the volume of unstructured data grows, companies will face increasing difficulty in completing backups rapidly enough to avoid impact on production systems. Larger backups will also mean higher costs and longer restore times, potentially increasing business risk. IBM ProtecTIER® de-duplication technologies use fast, reliable and powerful algorithms to identify and remove duplicate data at a sub-file level. This helps reduce total data volume, offering significant potential savings in disk capacity and enabling faster backup and recovery.

In addition to shrinking and accelerating their backup processes, businesses are looking to accelerate recovery by using remote-site replication to vault recent backups. Up to 12 IBM ProtecTIER Gateways or Appliances can replicate to a ProtecTIER hub at the DR site, creating a rapid disaster recovery solution that helps obviate operational and security issues around tape transportation and storage.

IDC estimates that demand for storage capacity worldwide will increase at a 49.8 percent Compound Annual Growth Rate, and that 65 percent of the total backup will use de-duplication by 2012.[2] As the technologies continue to improve, IDC estimates that businesses will be able to use de-duplication to improve the data reduction ratio from 15:1 in 2008 to 20:1 by the end of 2012.

IBM Real-time Data Compression shrinks data before it is written to disk, reducing total capacity requirements by as much as 80 percent while maintaining existing performance.[3] By freeing up capacity, data compression can help you keep a greater proportion of your data online for high-speed access, potentially making it a key enabler for analytics involving large data sets. Maintaining a smaller total volume of data will accelerate backup and restore, and may help reduce costs for legal discovery in the event of a data audit requirement. And with fewer disks to power and manage, businesses may be able to maintain or even reduce operational costs – even as the volume of data rises.

The benefits of compression and de-duplication will be repeated across all office locations, data centres and offline archives – so the total savings could be substantial.

Move your data to the right place

In most organisations, the vast majority of data is created once, and then rarely or never used again, while a small set of data will be in everyday use. In a conventional landscape, both types of data are stored and managed in the same way. This typically means that the unused data will be disproportionately expensive to store, and will consume budget that could otherwise be used to enable faster and more reliable access to the most commonly used data. Businesses have tried to address this issue by introducing different tiers of disk storage, and by manually archiving data to tape, but this has historically tended to create an additional layer of complexity and administrative cost.

IBM System Storage® Easy Tier™ technology – available on the IBM Storwize® V7000 Midrange Disk System as well as the DS8000® family of enterprise disk arrays, and via IBM SAN Volume Controller for use with any disk system – uses advanced workload monitoring and hotspot analysis to intelligently move only the most frequently used data to ultra-fast solid-state storage. An audited Storage Performance Council benchmark shows that Easy Tier can improve performance up to a factor of three – even if solid-state storage makes up as little as two percent of the total storage capacity.[4]

Live data can be automatically moved between tiers to optimise the price-performance balance – either using automated workload analysis, by manual administrator request, or based on pre-set enterprise policies. Likewise, IBM offers application-aware archiving to move unused or infrequently used data to lower storage tiers, freeing up more costly high-performance capacity to serve the most demanding applications and databases. This could mean faster processing and better response times for key enterprise systems, improved backup and recovery times and a reduction in the overall costs for storage capacity.

IBM offers a broad spectrum of storage technologies that can optimise price-performance for any workload, from ultra-fast solid-state storage through to compact and cost-effective tape. Advanced compression and de-duplication technologies keep capacity requirements to a minimum in every tier – reducing acquisition and energy costs, and enabling faster recovery.

Store more with what you have on the floor

In challenging economic conditions, businesses naturally want to get the greatest possible value from existing and future investments in storage capacity. By shrinking the number of disks required, and by using smaller, more efficient storage devices, businesses can significantly reduce both capital and operational expenditure, as well as extending the useful lifespan of their existing data centre facilities.

With IBM storage virtualisation and thin provisioning technologies, and new higher-density storage, businesses can pack more useful data into a given capacity and more storage systems into a given physical space. In its own data centres, IBM estimates that it has cut block storage costs in half and increased utilisation from 50 percent to 90 percent through the use of storage virtualisation, thin provisioning and storage pooling.[5]

IBM System Storage SAN Volume Controller enables existing and future capacity across different physical arrays to be virtualised, pooled, and treated as a single, flexible resource. New storage requirements can be provisioned at high speed, and the ability to release unused capacity back into the pool for re-provisioning can enable significant improvements in utilisation. Storage capacity allocated to SAN Volume Controller pools is optimised for the available physical disk space, helping prevent the accumulation of small, trapped and unusable portions of space. Online data migration can help eliminate application downtime, and SAN Volume Controller's large cache can significantly improve storage performance. IBM also offers the XIV® Storage System, a fully virtualised grid-architecture disk array that helps protect data and can offer high performance without any need to manage RAID levels or worry about disk hotspots. XIV arrays can be used in combination with SAN Volume Controller to aid in the creation of large aggregations of low-cost storage. Equally, businesses can benefit from Easy Tier's use of Solid State Drives and automated tiering to meet high performance requirements.

Thin provisioning allows administrators to allocate storage capacity to any given system without actually consuming the corresponding physical disk capacity. For example, a new server that will ultimately require 50 GB of storage is allocated the full 50 GB but has no actual disk provisioned to it; only 10 GB is actually written during the server's first year of operation, which is provisioned automatically as needed. The remaining 40 GB that would have been pre-allocated either remains in the central pool for use by other servers that require it, or is not purchased until needed. For a growing business, and in terms of data volume most are, this improves capacity planning and delays the need to invest in additional capacity. Thin provisioning is available as standard on SAN Volume Controller, on the XIV Storage System and on the Storwize V7000, enabling these solutions to support growing requirements in a cost-effective manner. The 'tier-less' XIV arrays promote cost-effectiveness, using high density disks and an innovative grid approach to deliver high performance with a lower administration effort and a ground-breaking intuitive user interface.

In addition to helping shrink the total number of disks required and eliminating 'white space' through the use of server virtualisation and thin provisioning, businesses can take advantage of physically smaller and more efficient disk arrays.

A smarter approach to storage

The accelerating growth in data volumes means that businesses need to find a new, more sustainable approach to data storage.

IBM can help you to reduce the total volume of data that you store and back up, enabling you to store more on current and future capacity. By automating the movement of data between tiers in your storage infrastructure, IBM can also help you to maximise speed of access to the most valuable data while minimising storage costs for the least valuable. Finally, virtualisation and thin provisioning can help you to make the most of what you already have, delay the need to invest in additional capacity, and to make more efficient use of new capacity when you do eventually need it.

Using IBM System Storage technologies, you can reduce capital and operational expenditure on storage capacity, cut implementation time for storage, simplify and accelerate data migration, cut administrative costs and simplify tasks to make administrators more productive, reduce energy consumption, accelerate backup and recovery, and delay the need to invest in new datacentre facilities. What's more, you can also deliver the right data to business users even faster, supporting analytics and turning your data into a business asset.

References

1. Source: **http://my.gartner.com/it/content/1031300/1031312/july1_technology_trends_you_cant_afford_ignore.pdf**
2. Source: IDC, Worldwide Enterprise Storage Systems 2010-2014 Forecast Update: Better Expectations on 2010 Growth, No Changes to Long-Term Outlook, Doc # 224618, August 2010., quoted at **http://www.ibm.com/news/ca/en/2010/10/07/b433705g06402o29.html**
3. Source: **http://www-05.ibm.com/ie/presentations/pdf/storagestory.pdf**
4. Source: Storage Performance Council, April 2010: **http://www.storageperformance.org/results/benchmark_results_spc1#a00092**
5. Source: **http://www-05.ibm.com/ie/presentations/pdf/storagestory.pdf**

IBM United Kingdom Limited
PO Box 41
North Harbour
Portsmouth
Hampshire
PO6 3AU

Produced in the United Kingdom
August 2011