

# MAXIMIZING YOUR STORAGE COST CUTTING EFFORTS



In 2009 an obvious theme is cutting costs. In IT and storage in particular that cost cutting will be popularized by data archiving and data reduction vendors. Both are very legitimate steps to take.

In the typical data center a good rule of thumb is that 80% of the data on primary storage can be migrated to an archive platform. The user buys the archive solution of their choice, then migrates all of that data, but where is the ROI? Clearly after archiving there are acres of free disk space on the primary storage platforms and some ROI can be derived if this process delays the purchase of additional storage, but the storage administrator can not start turning off empty drives. There aren't any. The way most storage arrays work, the remaining data is still scattered across all the drives in each LUN.

The problem with the archive first and ask questions later is that no analysis was done to determine what data is where and what data can be relocated somewhere else in the system. Tools exist today from companies like Tek-Tools that provide independent third party analysis of the storage environment and how applications interact with that environment. In short to get maximum benefit when driving out costs you need to first drive in efficiency through a proper analysis of the environment.

## Inventory of Data Set

The first step in maximizing your storage cost cutting efforts is to understand what those storage assets are and what they contain. This will not only include a simple file view of data that has not been accessed and can be moved to secondary storage but will also provide an analysis of what that data is and what type of optimization should be applied to it during movement.

For example is that data an ideal candidate for some form of compression or deduplication? If so which method should be used? There are solutions that are focused on providing optimization of standard files like VMware images and then there are those that specialize in optimizing .jpgs and PDF files. An analysis tool will indicate what type of data is in the archival candidate set. Most data reduction technologies cost about the same, the ROI that an analysis tool brings to this step is selecting the technique that is going to provide the maximum reduction for your investment.

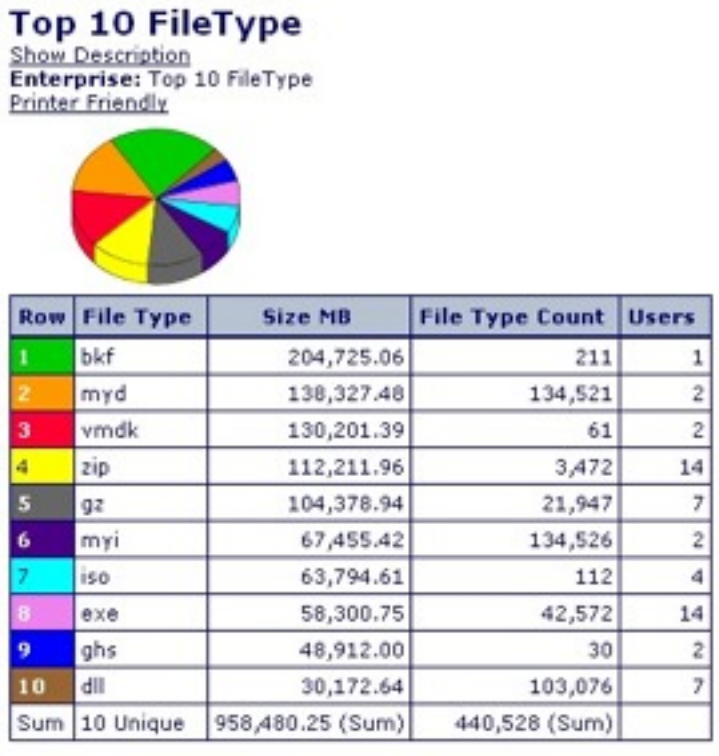
A third and critical parameter is how fast will data be added to the archive. Again with a proper analysis tool those factors can be examined and even trended for future growth. The ROI here is derived from knowing how large of an archive to invest in upfront and to properly budget for not only archive but also primary storage investments in the future.

### C-Level Storage Summary

Row	Time	Resource Name	Raw Capacity	Raw Spares	Raw Used	% Raw Used	Raw Free	Overhead	Usable Capacity	Usable Used	% Usable Used	Usable Free	Free Lbns	Estimated RAID5	Estimated RAID10
1	05/05/2009 00:28:42	\\server2.tes.com	1.06 TB	203.60 GB	679.66 GB	76.90	393.60 GB	203.60 GB	679.66 GB	287.22 GB	40.44	187.84 GB	287.22 GB	336.70 GB	181.86 GB
2	05/05/2009 00:01:49	\\server1.tes.com	1.49 TB	0.00 Bytes	3.36 TB	100.00	0.00 Bytes	736.17 GB	679.66 GB	914.87 GB	79.25	141.08 GB	31.58 GB	0.00 Bytes	0.00 Bytes
3	05/05/2009 00:22:31	\\server2.tes.com	0.29 TB	535.73 GB	3.06 TB	100.00	0.00 Bytes	409.20 GB	2.66 TB	3.67 TB	43.04	5,005.64 GB	20.00 GB	0.00 Bytes	0.00 Bytes
4	05/04/2009 00:21:05	\\server2.tes.com	934.60 GB	67.87 GB	766.52 GB	100.00	0.00 Bytes	67.87 GB	679.66 GB	679.70 GB	99.40	7.93 GB	0.00 Bytes	0.00 Bytes	0.00 Bytes
Sum 4 Unique	4 Unique		6.55 TB (Sum)	497.20 GB (Sum)	5.80 TB (Sum)		293.60 GB (Sum)	5.39 TB (Sum)	4.40 TB (Sum)	3.21 TB (Sum)		5.21 TB (Sum)	313.80 GB (Sum)	336.70 GB (Sum)	181.86 GB (Sum)

**Mapping the Remains**

Cutting storage costs can be as simple as eliminating shelves of unused capacity.



The challenge is that in most data centers this capacity is not represented by racks of storage shelves in an array that have no data on them. More likely it is available capacity that is scattered throughout the environment, none of it in a contiguous block and certainly not on one contiguous shelf.

The next step in maximizing storage cost cutting efforts is to understand the data sets that are left on primary storage. That 20% may be small but it is active, likely critical and as stated earlier, scattered all over the available disk shelves. Ideally the goal is to consolidate this data on to a smaller number of storage platforms and a smaller number of disk shelves being careful to balance cost savings vs. data vulnerability. A data analysis tool

like Tek-Tools Profiler will provide a complete inventory of where the remaining data set actually is. From there the storage administrator can use the tools often

provided by the array vendor to consolidate LUNs and reduce the number of shelves required. These actions will which eventually will result in turning off the shelves, possibly some storage controllers, and will lead to dramatic hard dollar savings in power consumption and the soft dollar savings of reduced management

overhead.

The ability to consolidate this data is going to require not only knowing that there is excess capacity but also where to place the actual data for maximum utilization.

**Inventory of Performance**

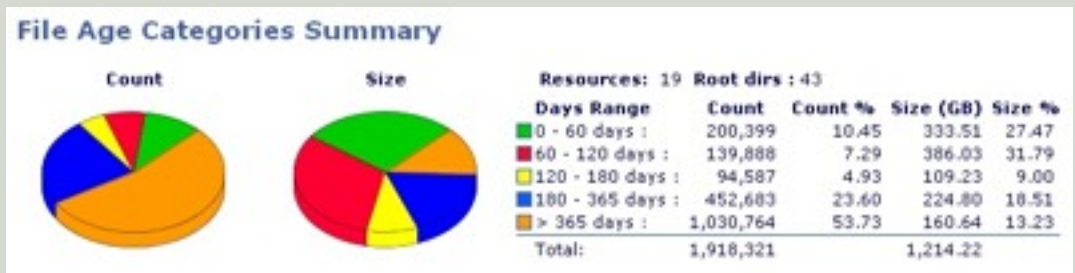
During the process of LUN consolidation it is also valuable to understand what performance the storage systems can deliver and what performance the

applications need. This understanding represents another ROI point for analysis by making sure that applications are operating at peak performance and thereby utilizing the proper amount of storage. An application that can't or does not need to drive data at the full speed of the storage architecture will gain nothing from being on that storage architecture.

Most storage arrays provide similar if not identical data protection like snapshots and replication across all supported performance tiers. The remaining differentiator is the performance those tiers provide and the cost to acquire and power those tiers. For example, faster drives that are smaller in capacity require more physical units, often consume more power on a per unit basis and as a result will require more shelves for the same capacity as larger capacity, slower tiers. Again, acquiring and powering shelves costs money.

A tool that can analyze both the performance aspects of the storage and the performance requirements of the application will allow the storage administrator to correctly match workloads with performance capabilities and optimize the cost reduction.

This step goes hand in hand with the consolidation of LUNs because relocating data onto fewer storage shelves is also going to impact the performance capabilities of the storage that is receiving the relocated data.



## Maximum Resource Utilization Requires Constant Monitoring

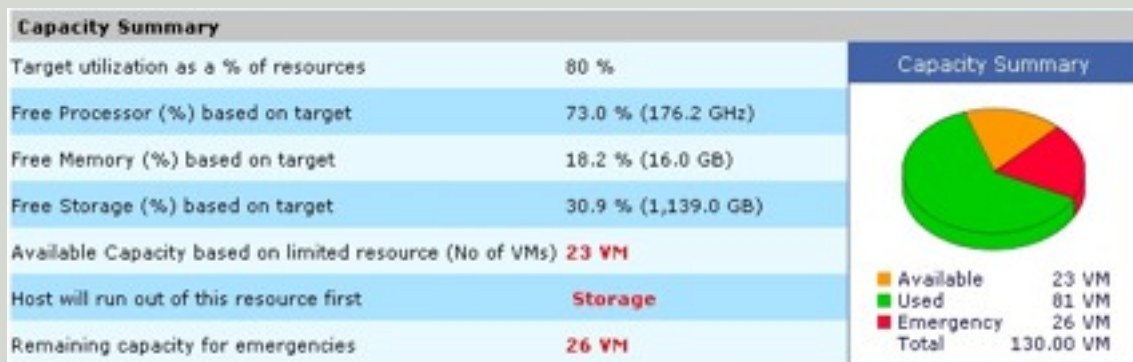
Manual analysis of just the inventory components of this process could be overwhelming to the IT staff, especially one whose levels have more than likely been stretched thinner than ever.

Automating this process is a key requirement.

**Fastest Growing Volumes**  
Show Description  
 Enterprise: Fastest Growing Volumes  
 Node: Node15

Rank	Resource Name	Volume Name	Used MB	Total MB	Current Used %	Rate MB/Day	Rate Percent/Day
1	Tel-Abco	C/S	30,376,277	40,962,338	74.14	1,036.93	2.43
2	Tel-Dame - prod	C/S	17,764,499	49,471,488	35.72	1,004.53	1.45
3	Webserver tek-tools.com	C/S	290,719,175	993,943,422	29.26	252.26	0.33
4	Mailbox tek-tools.com - prod	S/S	17,880,080	30,473,473	58.71	237.68	0.76
5	Index Server	C/S	204,769,127	218,409,861	93.76	197.68	0.89
6	File Server - prod	C/S	214,294,444	214,294,114	99.99	175.37	0.81
7	Tel-DEV 2.05	D/S	91,360,772	79,528,433	114.26	217.00	0.28
8	Media tek-tools.com	S/S	14,677,427	20,473,433	71.71	147.26	0.72
9	Tel-EPD	D/S	94,309,499	94,309,361	100.00	130.23	0.23
10	File Server - prod	S/S	19,236,777	17,486,533	109.67	62.33	0.33
11	Microsoft web tek-tools.com - prod	S/S	1,323,427	7,323,427	18.06	36.04	0.50
12	dev-vm2008-tek-tools.com	C/S	14,429,833	14,382,000	100.34	35.98	0.23
13	Alpha tek-tools.com - prod	S/S	1,174,444	10,473,433	11.22	19.04	0.14
14	Web tek-tools.com - prod	C/S	11,709,533	114,493,099	10.23	17.47	0.03
15	Exchange2007-tek-tools.com	C/S	4,627,777	10,228,866	45.26	26.78	0.26
16	Tel-Dev 3.1	C/S	4,494,333	11,393,144	39.45	18.04	0.13
17	Tel-DEV 1-0	C/S	11,294,899	19,293,544	58.54	15.83	0.04
18	Mailbox tek-tools.com - prod	C/S	4,230,227	10,228,866	41.36	12.73	0.12
19	DBA - prod	App	4,494,233	10,123,000	44.39	10.88	0.08
20	Tel-Exchange	C/S	11,119,333	117,233,111	9.48	8.53	0.01
21	Alpha tek-tools.com - prod	C/S	4,743,866	10,228,866	46.39	7.52	0.07
22	Win 2003 FileServer	C/S	10,369,277	11,393,144	90.93	6.29	0.04
23	Win 2003 FileServer	C/S	4,203,333	11,393,144	36.91	3.08	0.03
24	Netlog	App	4,343,000	4,343,000	100.00	2.79	0.03
25	File Server - prod	C/S	11,804,444	11,294,000	104.44	2.42	0.02
Sum 25 Unique	A Unique		1,490,839,422 (Sum)	1,226,236,488 (Sum)		146.10 (Avg)	0.33 (Avg)

Once the commitment is made to use every IT resource to its safe maximum, the environment needs to be actively monitored and trended. Even if the time and resources were available to manually inventory and analyze the above parameters, which is unlikely since it would take weeks if not months in a data center where all the components are being used to their full potential without excess, a spike in utilization needs to be quickly identified or alerted to (even predicted) so adjustments to the environment can be made.



## Server Virtualization

Of course as part of any cost savings effort the subject of server virtualization (or expanding the server virtualization effort) comes up. The tools used to analyze and provide real time monitoring of the environment, as described above, can also provide real world analysis of servers that are ideal candidates for virtualization. For example Tek-Tools Profiler uses real world statistics and then maps those into a simulator so that performance impact to both the virtual host and the newly virtualized application can now be measured. For more details on this, see our article on Maximizing your Virtualization Project.

In the current economic situation excess capacity, performance, computing power or even archive areas are all expenses that cannot be tolerated. Most likely, even as the economy turns around, data centers will not be allowed to return to wasting budget dollars in excess. The optimized IT environment is a permanent condition. Using realtime analysis tools not only make that an achievable goal, they do so while increasing the overall efficiency of the IT staff, returning to it the most valuable asset of all time, time to do their job and time to go home on time.

## About Storage Switzerland

Storage Switzerland, is an analyst firm focused on the virtualization and storage marketplaces. Storage Switzerland provides strategic consulting and analysis to storage users, suppliers and integrators.