WHITE PAPER

Lean storage

A phased approach to information lifecycle management can improve the productivity of a storage environment — today, not years from now

The principles of lean manufacturing offer valuable lessons for companies wrestling with rapid data growth and rising storage costs. As manufacturers have found, the solutions to systemic problems lie neither in "big bang" approaches that solve all problems at once nor in quick-fix tactics that address hot issues. Instead, companies need a strategic framework for solving large-scale problems in an incremental manner. For data centers, information lifecycle management provides this framework.

JUNE 2004

1 Executive summary

- The achievements of lean manufacturing — such as driving out waste and inefficiencies — offer valuable lessons for enterprises struggling with data storage issues.
- As a steel manufacturer learned, well-intended quick fixes can actually exacerbate larger business problems. The same holds true for storage.
- Information lifecycle management is not a big-bang solution that is years away. It provides a strategic framework for addressing today's storage problems in an incremental manner.
- An organization can begin its ILM journey by attacking the inefficient use of primary disk space, implementing tiered archiving solutions and addressing systemic backup and recovery issues.

2 The challenge

2.1 The case for beginning the information lifecycle management journey — today

Much of the data storage industry is thinking about information lifecycle management in very limited terms. Vendors are promising, and their customers are awaiting, the big-bang emergence of automated tools to deliver the right information to the right tier of storage, at the right cost and at the right time.

And though it's true that tools are going to mature to enable this movement of data to capitalize on the economic arbitrage between tiers, there will be no big-bang product metamorphosis. The sooner the IT industry accepts that reality, the sooner they can begin to realize higher efficiencies and investment returns in their data storage environments.

So, why isn't the industry moving rapidly in this direction today? And more importantly, how do customers, vendors and service providers get started? Is the answer in a new, magical, software product as valuations in the storage software space imply? In a word, no. Software is only a small part of the story, and it's certainly no panacea.

The problem is that organizations are looking too far ahead to the arrival of all-encompassing solutions. While they wait for the big fix, they miss the opportunities to make very meaningful gains that can be enabled by taking incremental steps in the longer information lifecycle management journey.

For perspective, data storage users and vendors would do well to look to the industrial sector and consider some of its lean-manufacturing lessons.

3 Eliminating waste

3.1 The lessons of lean manufacturing

The principles of lean manufacturing were first introduced in the United States in the automobile industry in the mid-1980s. Dr. Edward Deming and others determined that Japanese producers like Toyota had a distinct competitive advantage in the global arena, and it was rapidly manifesting itself as higher profitability and market share.

That underlying advantage was the extremely self-critical recognition of seven types of muda, or waste, within operational environments, and the process of kaizen, or continuous improvement, to rigorously and ruthlessly attack that waste. The seven types of muda are overproduction, overprocessing, inventory, motion, waiting time, transportation and correction.

The kaizen process involves measurement of the current state, benchmarks against peers and world-class companies, aspiration setting, problem identification, root-cause investigation for the problems' key drivers, action plans for countering root causes, and continuous iteration of all the steps.

Kaizen is a never-ending journey and a belief that while perfection is unachievable, its relentless pursuit is rewarded by excellence. Toyota's, Honda's and Nissan's dramatic share gains versus global competitors — including General Motors, Ford and Chrysler in the United States — during the 1980s and 1990s are proof of the power of kaizen.

3.2 Parallels between steel and storage

The lean manufacturing journey has twisted through U.S. automotive OEMs and many industries. Consider a producer of flat rolled stainless steel in 1999. This customer looked a lot like the U.S. automotive competitors who had suffered losses to competitors in the late 1980s.

A quick scan of internal and external indicators of success told the sorry tale:

- Long and unpredictable cycle times through its manufacturing lines
- Substandard delivery reliability against customer commitments
- Low asset utilization levels
- High inventory carrying costs
- Unacceptable delivered quality performance
- Low customer satisfaction with both delivery reliability and product quality
- Declining profitability and market share

Let's pause to consider the similarities to today's data storage issues. Although steel manufacturing is a world apart from data storage, there are some notable similarities between the issues listed above and the issues that IT departments face today.

- Long and unpredictable cycle times through steel manufacturing lines are analogous to data center managers' challenge to fulfill uncertain but long regulatory retention requirements and to comply with corporate policies for the destruction of data.
- Substandard steel delivery reliability against customer commitments is analogous to data center managers' challenge to meet recovery time objectives (RTO) on data that is often unreliably backed up. For 25 percent of companies, at least one in twenty backups fails.[1]
- Low steel asset utilization levels are similar to the poor utilization rates for the most expensive primary storage assets in today's data centers. Primary storage — the $15–30 per gigabyte storage that accounts for a $12.4 billion market today[2] — currently has net utilization of less than 50 percent.[3]
- Steel's high inventory carrying costs are similar to the petabytes of warm or cold copies of data that are unnecessarily residing on primary storage, instead of ATA disk or tape. These are the copies for which fast access (sub 200 milliseconds) is not required to meet business requirements.
- Steel's unacceptable delivered quality performance is, again, similar to backup's problems. Data center managers struggle to deliver data back to customers within the desired recovery point objective (RPO) because of poor backup reliability, poor recovery reliability, and the window between the last backup copy and the point at which the original data was lost.
- Low customer satisfaction with both delivery reliability and product quality abound in storage, mostly in the backup, archival and recovery spaces.
- Declining profitability and market share in the steel industry are analogous to the storage industry's problems in meeting the business requirement to deliver increasingly higher levels of storage services at decreasing costs — despite ever-increasing data volumes and retention requirements. Again, the current trajectory for many storage environments is sub-optimal.

3.3 Applying lean manufacturing in steel

In the case of the steel producer, an executive-level steering team was formed to attack the long list of waste-related issues. Its first step was to compose a team of external operational consultants and key internal employees from across levels, plants and functions. The company recognized that it needed a blend of experts, change agents and the old guard in the plant, who would be instrumental in making initial gains stick while beginning to move the culture in the direction of kaizen.

The team's mission was great, but it recognized that other companies had been down the same path before. If disaggregated into its logical steps, the mission could be achieved.

To begin, the team needed to:

- Define the problem.
- Set benchmarks versus best-in-class companies.
- Set internal aspirations and gain the steering team's buy-in to those aspirations.
- Identify root causes behind the problem indicators.
- Launch the kaizen journey.

After initial benchmarking and aspiration-setting exercises, the team was ready to embark on the hardest part of its journey — identifying root causes.

[1] Peripheral Concepts, 2003.
[2] IDC, 2003.
[3] "Improving Storage Utilization: The Benefits Are Often A Myth," January 23, 2003, Gartner, Inc.

3.4 Root causes behind the steel failures

When the team investigated the root causes of problems, they found many committed people taking targeted actions. Production planning and manufacturing departments were capitalizing on flexible machinery by batching multiple jobs on single machines in an effort to decrease changeover times and increase asset productivity. Operations and sales teams were escalating hot orders and pushing them through the plant on a fast track in an effort to improve customer satisfaction. Yet despite such efforts, asset productivity, delivery reliability and customer satisfaction spiraled down. But why?

The team members asked why each problem was really happening. They determined that the efforts of many qualified people in this production system, although individually well intended, were treating symptoms of problems and not the root causes themselves. Many of the well-meaning business practices being implemented had actually become multipliers of the problem. For instance:

- The practice of batching jobs on single machines created a complex flow of in-process inventory, making customer orders virtually untraceable as they moved through the plant. Machine operators were capable of working with any of the product lines but were masters of none, causing quality problems.
- Adding inventory to the system to increase asset productivity only compounded problems. Longer runs resulted in larger batches of failed coils that had to be rerun.
- The practice of escalating hot customer orders through the plant further compounded problems. Large batches ready to run were pushed aside to accommodate the few urgent coils required that week.

3.5 The quick-fix problem in storage

The experiences of the steel manufacturer parallel the experiences of many data centers. Storage problems are often addressed with narrowly targeted, short-range fixes. If a particular application is running out of storage capacity, an administrator does a quick fix — buys more disks. If users are running into the storage limits of an e-mail system, a company mandates that individual users retain no more than one gigabyte of e-mail. If an operations team can't keep up with daily backup tasks, an organization hires an extra staff member.

The storage industry has many examples of the pitfalls of quick fixes. Here's a look at a few:

3.5.1 Quick fix: Throw disk at the problem

There is a temptation to use primary or ATA disk in all places because of falling prices and a perception of flexibility. Primary disk is like the flexible machinery in the steel plant. It can be used everywhere. ATA is the second choice.

But disk of any type is not a panacea. Performance disk is not economically feasible for warm or cold data. ATA is not economically feasible for cold data. Neither type is removable, long lived or reliably recoverable. Many data centers lack the personnel, capacity and tools to classify and move data from performance disk to lower-cost storage tiers. So this quick fix results in data piling up on primary disk, compounding overall storage problems.

3.5.2 Quick fix: Keep it on disk longer

Similarly, there is a tendency to leave more data copies, snaps and mirrors on disk for longer periods of time. That's because it seems safer and easier to leave data where it is. But this is the storage equivalent of shoving more product through a manufacturing line. It doesn't address the root cause of the problem. Ultimately, this approach results in more data on higher-cost storage, more waste and a lower return on storage.

3.5.3 Quick fix: Focus on the hot issues

Life in data centers can be extremely reactive. IT staff can spend a great deal of time responding to one-off customer demands to expand capacity, recover backup copies or recall archived files. As in the steel plant, focusing on hot issues takes time, delays other operations and exacerbates broader problems. Uncovering root causes, one sees that a quick-fix approach isn't addressing the real issues — a low reliability for backup/recovery and a reliance on manual processes.

As with the steps taken in the steel plant, quick fixes like these are all well-intentioned actions. And they all solve immediate problems. But in so doing, they can actually exacerbate larger issues. Cold data piles up on expensive primary disk. Management burdens rise as more disk devices come online. Storage costs grow. And data center staff members spend their time putting out fires rather than managing the forest.

4 Addressing root causes

4.1 Solution design in steel

For the steel producer, solution design ironically was pretty simple after it investigated and gained an understanding of root causes. The design had three primary facets to it, each designed to counter one of the root causes:

1. To stop the use of flexible machinery to serve any product line, the company dedicated four primary "flowpaths" through the mill, each specializing in a product family with like operational characteristics and customer needs. In this way, a flowpath line could function as a team to target metrics against its product family, including profitability and share. The team could also become specialized enough to increase process repeatability as well as to identify and solve process abnormalities commonly occurring.
2. The process of pushing excess inventory into the system was replaced by a pull system, in which a coil of steel would be released into the system only when it had a customer order attached. Every inventory coil in the mill then became important, and overall inventory could be greatly reduced, beginning with finished goods.
3. Finally, and perhaps most importantly, a system of controls was put in place between sales and operations to halt the practice of rushing most hot orders. And because excess inventory and quality problems would be corrected with the first two practices, the overall number of hot orders would be reduced to begin with. This was difficult because it involved changing the balance of power for a time from sales to operations, at least until the operations could be brought more under control.

4.2 Phasing in the solution

The results are less important than the journey. Once the root causes and solution design were well understood and agreed to by the steel manufacturer's steering team, it seemed obvious to source a new, automated scheduling system and develop a turnkey changeover plan for the entire mill, right? Wrong, the team decided.

Although the business was suffering, there was still much risk to transition all at once. Such a big-bang approach also offered no migratory process to bring the rest of the culture along for the ride. Even if the design were absolutely correct, not having the hearts and minds of the individuals on board would doom it to failure.

So the team created a phased implementation approach.

1. They selected one pilot flowpath to do first. And they selected a family of products that had significant enough problems, but offered enough upside to the company if it succeeded that it could generate momentum for the broader transformation.
2. They migrated to a second flowpath, and then a third. They leveraged success and momentum along the way, taking on even more critical orders.
3. Finally, the fourth, fifth and sixth lines were cut over to the new dedicated flowpath scheme.

The journey was not without its bumps. But success became irresistible, and the urge to win began to outstrip the resistance to change. People in the plant changed because they could see the win.

4.3 Applying lean manufacturing to storage

So how can the principles of lean manufacturing, the elimination of muda and the goals of kaizen be applied to information management? How can the complex process of storing and serving up data for demanding internal application owners be anything like that of rolling steel?

The answer is that they are both operational processes with demanding customers and talented personnel driving flawed processes with inadequate tools, and with insufficient metrics to define success. In both cases, operational processes are extremely difficult, but the associated challenges can be overcome through a kaizen-like process.

In data storage environments, a total systems approach to information lifecycle management brings the kaizen concept to life. This approach eliminates waste by balancing the costs and performance of storage options with the purpose of data at different points in its lifecycle. It distinguishes the value of a piece of information based on such characteristics as its age, its importance to day-to-day business operations, user access requirements, and legal and regulatory guidelines. By better matching the purpose of information with the cost of storing it, this approach responds directly to the challenge of rapid data growth in a time of falling IT budgets.

Information lifecycle management does not hinge on the availability of new technology, nor does it require the purchase of costly hardware or software solutions. It can be implemented today by making strategic improvements to a current storage infrastructure.

4.4 The information lifecycle management opportunity

Primary disk storage, despite sharp recent declines, still costs roughly $15–30 per gigabyte — and up to $60 per gigabyte for highest performance enterprise disk.[4] Tape storage systems, also declining sharply, cost roughly $0.50–3.00 per gigabyte.[5] So an information lifecycle management strategy is simple: manage storage costs and efficiencies through effective data management.

Less than 5 percent of users' total data is truly "hot" at any moment in time[6], so moving vast amounts of the "cold" data to automated tape libraries offers the potential for tremendous cost savings. For cold data, leading automated tape technology offers reasonable access times — 20–30 seconds.

Enter ATA disk drives and the story gets more interesting. ATA cost points are in the $5–10 per gigabyte range.[7] The performance of ATA disk is close enough to primary disk for many primary applications, and more than adequate for the data protection of many other applications. Now there is a place for "warm" data in between traditional disk and tape.

IT organizations only need to move warm data off primary storage and onto ATA to unlock savings of up to 68 percent. This can be accomplished with little to no impact on users' access times, which should stay in the sub-second range. If cold data is moved from primary and ATA storage and onto tape storage, they can unlock savings as high as 90 percent. When taking into account the massive data storage budgets for many corporations, the potential savings are staggering. And these are just the hardware savings. There is also potential to save on management costs as well.

Information lifecycle management is like lean manufacturing for data storage. It is an IT-based process improvement journey. It is a new way of thinking. The first step for enterprises is to adopt this new way of thinking instead of waiting for the perfect tool. IT organizations can and should start on this journey today, and the IT industry should support them.

5 Recommended actions

Embark on three process journeys:

1. Attack inefficient use of primary space head-on:

- Use SRM and services to measure and optimize partitioning and capacity utilization today. Defer buy decisions — now.
- Introduce modular or ATA storage wherever application requirements allow. Challenge what business requirements really exist by application.
- Commoditize high-cost primary disk software by raising its functionality to the network level and virtualizing it.

2. Implement tiered data movement archive solutions:

- Measure and monitor what copies you are making by application and by tier of storage. Set the baseline.
- Pick application-specific pilots in which to get started.
- Roll those successes out to rest of your unstructured/structured and regulated/unregulated data.

3. Address systemic backup and recovery processes:

- Measure and monitor your backup and recovery success rates, by application, by location, by time of day, week, month and year. Understand your performance and your volatilities. Find which ones you need to target.
- Apply backup management information and management tools. Use disk buffering in front of tape for some applications. Use disk-based backup entirely for other applications.
- Break the paradigm. Enact continuous point-in-time replication. Replace backup process entirely with better archival practices.

[4] Horizon Information Strategies, 2003.
[5] Ibid.
[6] Ibid.
[7] Ibid.

6 Planning considerations

- Don't wait for a high-level CIO/CFO/COO-led steering group and a full-time working team to make information lifecycle management happen. Define the problem, measure the status quo, create benchmarks, set aspirations, design the solution, pilot and do kaizen.
- Start where your pain is the greatest. Is primary spending out of control? Are backup and recoveries unacceptable? Are archival and compliance with business and regulatory needs your biggest concern?
- Begin by addressing a single area of concern, unless you're staffed for more.
- Do your own analysis. Don't skip the measurement and assessment steps.
- Design your solution architecture, without anticipating a big-bang product. (It probably doesn't exist.) Pilot and go — find wins today.
- Attack the muda. It's everywhere. The tools will come along. Build momentum inside your organization.

ABOUT STORAGETEK®

Storage Technology Corporation (NYSE: STK) is a $2 billion global company that enables businesses, through its information lifecycle management strategy, to align the cost of storage with the value of information. The company's innovative storage solutions manage the complexity and growth of information, lower costs, improve efficiency and protect investments. For more information, visit www.storagetek.com, or call 1.800.275.4785 or 01.303.673.2800.

WORLD HEADQUARTERS

Storage Technology Corporation
One StorageTek Drive
Louisville, Colorado 80028 USA
1.800.877.9220 or 01.303.673.5151

MC 1091 A 06/04