



ESG REPORT

The Far Reaching Benefits of IT Efficiency

By Steve Duplessie

February, 2009

Table of Contents

Table of Contents	i
Introduction	1
Overview	1
Data Growth is the Common Factor Driving Inefficient IT	1
Infrastructure Consolidation	2
Data Virtualization: Attacking the Problem at the Source	3
The Value Multiplier: Concurrently Addressing the Infrastructure <i>and</i> the Data Optimizes ROI	4
Summary	5

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508)482-0188.

Introduction

Recently, I had a discussion with a very senior operations manager at a major European telecommunications company who was under pressure to cut costs. We were discussing his experience with a backup consolidation effort (hundreds of remote data sites) which leveraged VTL (virtual tape library) systems with data deduplication technologies he had deployed approximately six months earlier. What I learned was entirely unexpected and became the impetus for this paper.

It seems that not only was the company able to realize the assumed and obvious benefits of consolidation that one would expect, but because the new process was so much more efficient, it altered the way the company fundamentally viewed *all* of the applications, processes, and services IT delivered to the business. He stated that there were substantial gains in far reaching areas—from IT’s ability to significantly improve ITSM delivery capabilities across regulatory bodies and borders, to security, privacy, and beyond. Because of the success of what began as a fairly simple “consolidation” exercise in the area of backup and recovery, the company is now aggressively investigating other opportunities to improve efficiency and gain even greater benefits, despite the current economy. All this happened as a result of something as seemingly simple and benign as consolidation; I became interested in discovering what other value could be derived by pushing overall efficiency improvements throughout the business and IT.

Overview

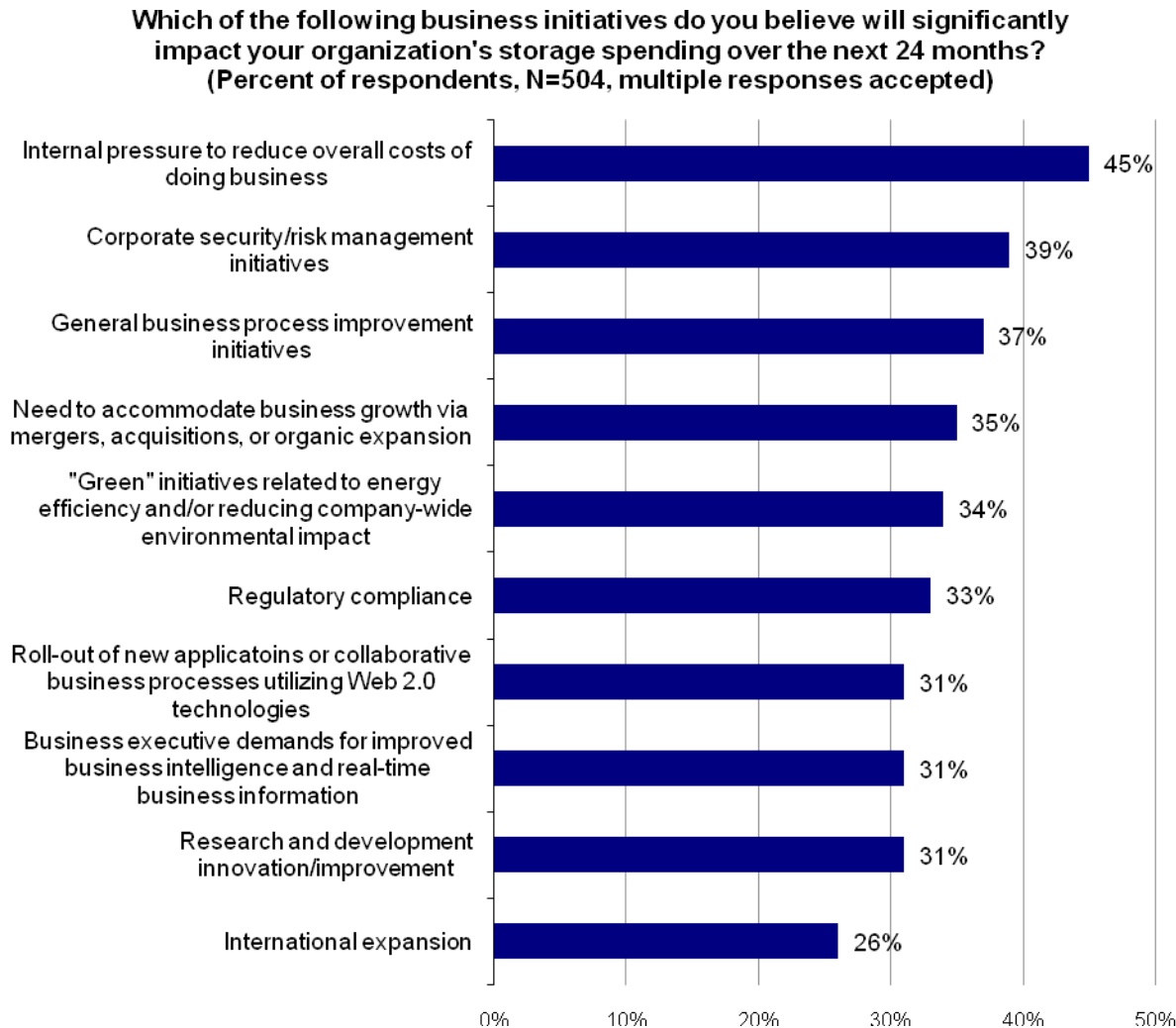
Data Growth is the Common Factor Driving Inefficient IT

Data growth is the only absolute constant in IT and, arguably, is the inevitable cause of most operating problems. New data growth rates never abate; they only accelerate. Data growth occurs whether the economy is up or down.

As a society, we are constantly increasing the capabilities and opportunities for people to create even more data at an even faster pace. Its growth naturally causes “sprawl” throughout the IT ecosystem. We need more devices to store it, more servers to execute upon it, more applications to manipulate it, and more networking to connect to it. More data generates more new use cases, which in turn generates even more data, and so on. We perpetuate the problem further as we duplicate and replicate data—normally for good reason. We must protect our data and our ability to execute on it in the event of local and site failures. We push data to the far edges of our remote organizations and throughout our supply chains to extend the value of that data over the entire organization.

The primary means of gaining efficiency across the business and IT is realized by consolidating infrastructure and data, both physically and virtually. In addition, we must begin to look outside of the data center—to the business processes and applications driving much of the data growth that will be impacting our spending over the coming months and years (see Figure 1).

FIGURE 1. BUSINESS INITIATIVES IMPACTING STORAGE SPENDING OVER THE NEXT 24 MONTHS



Source: Enterprise Strategy Group, 2008

Infrastructure Consolidation

There are three primary “layers” of data center infrastructure: server, network, and storage. Consolidating physical servers mean less network connections, less “boxes” to manage and support, and so on. Furthermore, virtualization technologies enable IT operational and administrative personnel to respond to business unit requirements significantly faster due to enhanced provisioning time. Secondary benefits, such as “liquidity”—or motion—have positive effects ranging from dynamic system migrations to high-availability and improved disaster recovery; all without the normal disruptions traditionally experienced by the business.

Server and storage layer consolidation has the same obvious effects. Consolidating more physical components into fewer physical or virtual management elements simplifies the entire infrastructure stack, which in turn simplifies everything operationally connected with that stack. *Less is easier to manage and operate than more.*

Leveraging virtualization technologies from within and across infrastructure elements can enable dynamic movement for tactical, error prone, and time consuming functions that normally require significant planning and operating attention. Data and system migrations that can be done virtually mitigate a significant amount of risk and business disruption. “Thin provisioning” in storage and server virtualization—the concept of showing an application a virtual (large) representation of “allocated” resources, but in reality only supplying that application what it actually requires—provides huge utilization improvements and clear cost advantages. ESG conducted a

survey of storage administrators focused on the limitations and challenges of traditional storage provisioning methods.¹ More than half reported that between 31% and 50% of their purchased capacity was stranded and unused, while 80% felt that storage provisioning had a significant negative impact on IT time and resources. In other words, traditional storage provisioning methods were placing a strain on capital and operating budgets.

Consolidation for operational benefit is obvious, but the incremental value gained by doing so with secondary enabling technologies, such as virtualization, can magnify those efficiency gains well beyond the initial cost savings. Less equipment means less power and cooling, less people to support and manage it, and higher levels of utilization—which leads directly to lower operating costs and higher operating efficiency.

Data Virtualization: Attacking the Problem at the Source

While attending a “Green IT” conference in China recently, I spoke with a senior official responsible for China’s global economic analysis. He asked, “Why do you focus on the symptom (meaning infrastructure) before the cause (meaning data)?” It was a great question. Most commercial Chinese IT operations are relatively new—unburdened by many of the long term historical issues European and North American environments face. China’s relatively new venture into the world of IT is such that the infrastructure “sprawl” issue others have spent 50 years dealing with is something that Chinese IT has yet to face and, as such, can stay focused on the cause: the data itself. Most of us are not so lucky.

This realization helped crystallize the situation: if IT can put itself back into a position to treat every decision from the perspective of the data itself—and not simply the infrastructure—our efficiencies and effectiveness could be optimized.

If data growth is the inevitable cause of our problems, then logically, we do ourselves a disservice by simply treating the symptoms—or the results—of those problems. By treating the causes or sources of the problem directly, we can dramatically enhance or eliminate the efficiency issues that arise “downstream.” People often overlook this basic fact.

Data virtualization—a virtual instance of data instead of a physical one—is not often categorized, but is perhaps the most interesting means of gaining efficiency throughout the entire spectrum of IT operations. Deduplication, thin provisioning, and snapshots are examples of techniques that dramatically reduce or eliminate the true physical issues associated with data itself while continuing to provide unfettered “virtual” access.

Consider data deduplication. In short, data that is duplicated time and time again (pushing the limits of capacity of all infrastructure and taxing all of our operating processes), often for perfectly valid reasons, has a multiplying ***negative effect*** on the backup, recovery, and disaster recovery processes we conduct. By default, most backups create full images, at regular intervals, of data that is *already duplicated*—sometimes dozens of times. The positive gains achieved by eliminating duplicate data inefficiencies are felt throughout an entire organization.

For example, by backing up to a disk target with deduplication, IT gets an immediate primary benefit of speed as disk enables a tremendous performance gain. When the data is deduplicated at the source, it takes up significantly less space at rest and during transmission. Therefore, data that has been deduplicated at point A will require significantly less infrastructure (bandwidth, capacity, etc.) to store and transmit that “virtual” instance of the data to point B, and so on. Secondly, because typical backup deduplication rates can be 20:1 or higher, the economics of keeping the virtual (or deduplicated) data on disk are far more practical, which in turn dramatically improves the efficiency of recovery operations themselves to bring even higher service levels and value back to the business. This new efficiency gain is multiplied again by improving overall IT operations staff utilization—allowing an already-taxed group to gain perhaps the most valuable of all operating assets: time.

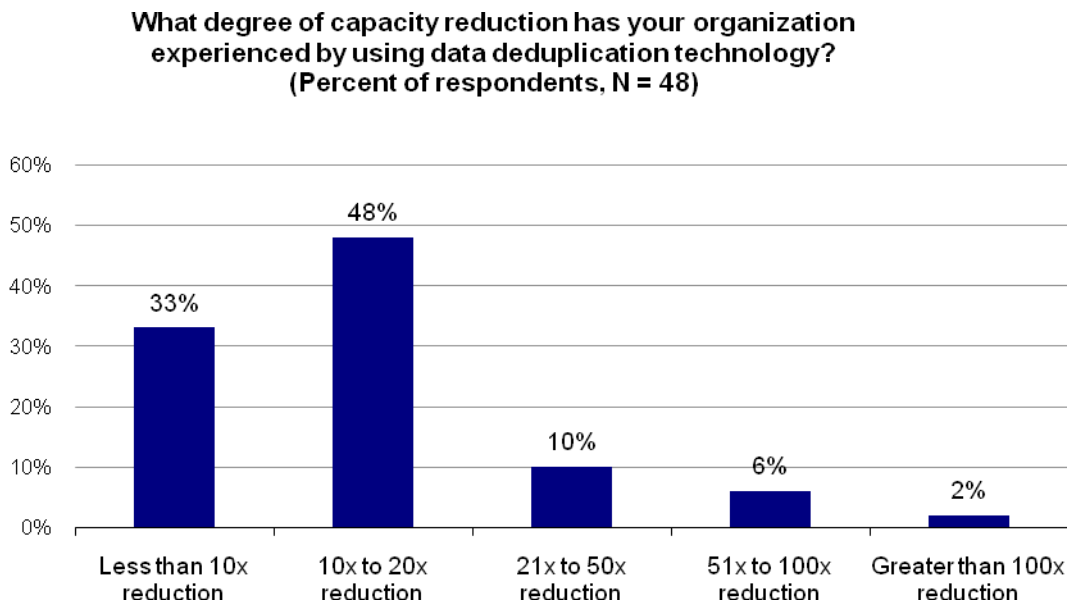
Reducing the actual amount of data that traverses our corporate networks enables even further consolidation and improved utilization, multiplying the benefits derived yet again. By implementing these technologies at remote offices, companies will find they can now easily centralize all corporate backup/recovery and disaster recovery operations to a primary, centralized location—eliminating the need for much of the remote IT support and associated costs presently required.

¹ Source: ESG Report, *Thin Provisioning*, April 2006.

One of the key measures of any capacity reduction solution is the degree of capacity reduction—or “reduction ratio”—it enables. As an example, a “10x” or “10 times” reduction would indicate that an organization was able to reduce the size of a 500 GB backup to just 50 GB. Among ESG research respondents currently using data deduplication technology, approximately one-third (33%) say they have experienced a less than 10x reduction in capacity requirements, 48% report a 10x-20x reduction, and 18% report reductions ranging from 21x to more than 100x (see Figure 2).

In general, these results are consistent with the results of hands-on product tests conducted by ESG Lab. While data deduplication ratios will vary based on the types of data involved and the frequency of full backups and retention, ESG’s testing has found a 20:1 ratio—when combined with data compression—to be broadly achievable. The fact that some percentage of ESG’s survey base has experienced much higher degrees of capacity reduction is also not surprising as ESG Lab has seen data deduplication ratios of up to 89:1. While it is easy for IT vendors and users alike to get caught up in these eye-popping numbers, it should not be overlooked that reductions of 4:1 or 5:1 can still be extremely valuable.

FIGURE 2. CAPACITY REDUCTION RESULTING FROM DATA DEDUPLICATION TECHNOLOGY USE



Source: Enterprise Strategy Group, 2008

Security capabilities are dramatically enhanced because less data physically moving around the enterprise reduces the risk of intellectual property falling into the wrong hands. Collaboration, discovery, and recovery operations required throughout the entire enterprise become suddenly attainable. Centralizing or consolidating backup and DR functions, enabled by consolidating and virtualizing the data itself, improves consistency, security, and efficiency throughout the entire enterprise—all while lowering capital costs, operating costs, and risk.

The Value Multiplier: Concurrently Addressing the Infrastructure *and* the Data Optimizes ROI

Data itself lives—and as such, can be further optimized by enabling efficiency gains at different times in its lifecycle. For example, mandated or regulatory retention requirements of an e-mail record are fairly easy to understand: “you must keep each record on immutable storage for 7 years.” We understand exactly what we must do to comply with the mandate. Where it gets less obvious is when we need to determine the most effective and economical means of complying. The easiest means is to apply the “rule” to the data, put it somewhere, and never move it so you can always find it if and when you need to. That is often at odds with gaining efficiency and optimization.

However, e-mail is a representation of data and during its early life, it may require high performance or availability that it simply will not require after a certain period of time. When it becomes a fixed (persistent) data object that will never change, it will only be accessed in the event of a recovery or discovery operation. As such, it makes more sense to gain economical advantages by housing that e-mail on the most cost effective infrastructure platform attainable—which most likely is NOT the originating platform. Just because retention and immutability are mandated, it does not mean that object is relegated to inefficient treatment forever.

The same benefits are realized—and magnified—as we apply this logic to “non-mandated” data. Eventually, every data object, regardless of form, will become a persistent, non-changing asset that will be infrequently accessed. Data in this stage, which represents the overwhelming majority of corporate data being managed, has radically different attribute requirements from when it was active and dynamic. Generally, it is safe to say that whether considered as data within an archive or simply data in a lower “tier” of infrastructure, once it stops changing and being heavily accessed, we can apply the same efficiency logic to it. By deduplicating this data, we can more easily and efficiently protect, access, secure, and store it without requiring many of the superhuman efforts our IT staff is currently forced to provide. Imagine the overall business efficiency gains (and value) if 90% of all of the problems and issues associated with IT inefficiencies went away because we were able to address those issues at the source?

Efficiency improvements implemented closer to the point of data origination—in the business processes and applications themselves—provide the greatest potential value throughout the data lifecycle. Content management systems can be leveraged more efficiently to enforce and automate policies based on the data objects themselves, enabling efficiency optimization at the object itself—often automatically.

Summary

The simple, well understood benefits of implementing efficiencies across the business and IT have profound ripple effects—we only need to look for them. Less physical does not imply less logical. As long as we do not inhibit our ability to use and access data as we require, eliminating or consolidating physical assets can only provide value. In fact, less physical data living and moving over our “plumbing” and through our process engines means we can move *more virtual instances of that data into even more areas of our business, which in turn drives even more value*. This is a positive cycle that feeds and leverages a basic truth: less is better than more.

Business, operations, and IT professionals can start the process of gaining additional efficiency leverage by asking simple questions as they are making common decisions, such as:

- Does the solution I am considering have any additional benefits upstream or downstream beyond the immediate problem I am trying to solve? If so, what are they? Recognizing incremental benefits can help you to further justify your decision or to put new light on alternative offerings.
- Does the solution I am considering have any multiplying effects when implemented? If I am solving one problem, in what other areas am I able to gain leverage? If I solve a backup problem by leveraging deduplication technology, am I also driving other benefits such as remote office consolidation or security?
- Am I solving an immediate problem only to create yet another problem for myself? Even if that is what you are forced to do, knowing can be helpful.
- Does the solution I am considering position us to achieve our long-term growth strategies? Or does it simply band-aid a problem without delivering leverage?
- Am I using technology as an enabler to the upstream drivers and the downstream infrastructure? Have I changed the processes around the technology?

By fully considering all of the positive and negative potential issues created or helped by a single decision, you will find yourself more prepared to handle the issues of today—and those that will arise tomorrow.



20 Asylum Street

Milford, MA 01757

Tel: 508-482-0188

Fax: 508-482-0218

www.enterprisestrategygroup.com