Technologyforecast

A quarterly journal
Summer 2009

In this issue

PricewaterhouseCoopers

Contents

Features

Interviews

Departments

Message from the editor

Cloud computing, cargo transshipment, and fruit flies. Three topics you wouldn't expect to find in the same article. But our focus on cloud computing and its role in the enterprise has unearthed interesting parallels between them.

The history of logistics since the Industrial Revolution and the rise of the railroads is one of fits and starts around standardization. Any system of cargo transshipment seeks to minimize the handling of individual items across the modalities of transport. But it has been a slow, laborious process to remove unnecessary handling, because legacy infrastructure is costly to replace.

Consider railway gauges—the distance between the inner sides of the two parallel rails that compose a railway line. Initially, decisions about gauges were determined locally, and many different gauges emerged. The downside? Break of gauge—the term used to describe the process of moving all people and cargo between trains running on different gauges.

An example of break of gauge is on the Trans-Mongolian Railway, where Mongolia uses broad gauge and China uses standard gauge. According to Wikipedia, at the border, "each carriage has to be lifted in turn to have its bogies changed and the whole operation, combined with passport and customs control, can take several hours."

Broad standardization of railway gauges has greatly reduced such inefficiencies around the world. But cargo often uses multiple transport modalities, including trains, trucks, and ships. Changing modalities was once equivalent to break of gauge. Fortunately, the rise of standards for intermodal containers has continued the trend of more efficient transport that began with gauge standardization. Movement of cargo is now "loosely coupled" with the underlying mode of transport.

Enterprise computing today is much like the early days of railways. Compute environments are purpose-built collections of technology dedicated to individual workloads. Individual pieces might be "standard," but the overall environment is one of a kind. Cloud computing offers relief. But not as you might expect, given all the press devoted to external cloud services. Enterprises need to learn the lessons of logistics standardization and modularization. Doing so will empower IT to deliver much higher responsiveness and greater financial flexibility to the business.

Who are some of the early adopters of this approach? Here's where fruit flies come in. Drosophila melanogaster, commonly known as the fruit fly, transformed the study of genetics back in 1910. Thomas Hunt Morgan's research in the "fly rooms" at Columbia University elucidated many basic

principles of heredity, including sex-linked inheritance, epistasis, multiple alleles, and gene mapping. What was the fruit fly's contribution? Short, 10-day intergenerational time periods and high birthrates (females lay 100 eggs per day) meant researchers could track gene expressions over many generations in a matter of a few months.

Bechtel, the engineering and construction giant, is to data centers as the field of genetics is to fruit flies. Bechtel's initiation of large, multibillion-dollar, multiyear projects requires the frequent provisioning of project data centers and application environments. Each effort is a "learning opportunity."

Bechtel's big takeaway? Traditional best practices for deploying and managing data center environments always impact cost, complexity, and business agility. Geir Ramleth, Bechtel's CIO, recognized that cloud service providers such as Amazon.com and Google offered a new set of best practices that most enterprises weren't adopting. With a new generation of projects always stacked up in front of them, Ramleth realized that Bechtel had an opportunity to design a new pattern of DNA for data centers that would deliver much higher business value.

This issue of the *Technology Forecast* covers the emerging trends associated with cloud computing. The first article describes Bechtel's journey and introduces the Evergreen IT concept—the idea that the goal isn't cloud computing but a new approach to the provisioning and management of IT that avoids the creation of legacy complexities and cost.

The second article takes the Evergreen IT concept down into the details of the technology that can support such an approach. We examine two broad technology domains, virtualization and data center automation, and explain their key roles in delivering Evergreen IT.

The third article describes the way forward from the CIO's perspective. Although there will be no single path, we suggest five broad stages of transformation most companies must accomplish to reach the goal of Evergreen IT.

As always, our articles are supported by in-depth interviews with leading executives and thought leaders defining the future of IT. Geir Ramleth of Bechtel shares his vision and experience of transforming the Bechtel IT organization from its legacy roots and toward the vision of Evergreen IT. Erich Clementi and Irving Wladawsky-Berger of IBM emphasize how standardization and mass customization principles will reduce complexity and industrialize the IT function. Simon Crosby of Citrix describes how virtualization creates the separation between the layers of the IT stack necessary for Evergreen IT. Kirill Sheynkman of Elastra explains how intelligent software he is building will model and automate data center operations. Russ Daniels of EDS, an HP company, shares his insights about how IT needs to move from a build-to-order culture to a configure-to-order culture. Doug Hauger of Microsoft notes how cloud computing represents a technology and business model shift.

Please visit pwc.com/techforecast to find these articles and other issues of the *Technology Forecast*. If you would like to receive future issues of the *Technology Forecast* as a PDF attachment in your e-mail box, you can sign up at pwc.com/techforecast/subscribe.

And as always, we welcome your feedback on this issue of the *Technology Forecast* and your ideas for where we should focus our research and analysis in the future.

Paul Horowitz
Principal
Technology Leader

paul.j.horowitz@us.pwc.com

The real promise of cloud computing

Evergreen IT delivers flexibility and lower cost. Finally, enterprises can shake off the burdens of their legacy systems.

Several years ago, the IT staff at Bechtel Corporation needed at least 60 days to respond to requests for new or changed application workloads. This requirement is hardly agile in an enterprise where each multiyear, multibillion-dollar construction project is much like building a new business that has unique staffing needs, geographic challenges, and a large IT component. Today, the IT staff can implement most changes overnight, a vast improvement that illustrates how IT agility can support business agility.

How did Bechtel do it? Central to the transformation is the on-demand availability of IT resources shared across a pool of users. This availability is achieved through some of the emerging practices and technologies collectively known as cloud computing. At present, cloud computing services are best illustrated by Amazon.com, Google, salesforce.com, and other service providers. These companies adopted patterns of IT architectures and data center provisioning and management that are disruptively more cost efficient, agile, and reliable than most enterprise IT environments. By adopting the best practices of cloud service providers for its internal IT infrastructure, Bechtel transformed its systems and processes. (See "Evergreening IT at Bechtel" on page 16 for more information.)

PricewaterhouseCoopers (PwC) believes Bechtel has made an impressive leap toward a vision we call Evergreen IT. An Evergreen IT environment is the end state that frees enterprises of costly, inefficient, legacy IT by enabling enterprises to build a more flexible and agile environment. Evergreen IT requires technology architecture and IT infrastructure management strategies consistent with the strategies used in cloud computing. It creates the potential for continually refreshing IT capabilities without distracting the business and for staying in alignment with business needs. Although this vision is inspired by external cloud computing providers, the goal is Evergreen IT, not cloud computing per se.

This issue of the *Technology Forecast* explores the concepts, the technology, and the road map to Evergreen IT. This article begins with a close look at Bechtel. It then describes how legacy IT has become an obstacle to IT agility, and ends with a prescription for resolving the issues through Evergreen IT.

An Evergreen IT environment is the end state that frees enterprises of costly, inefficient, legacy IT by enabling enterprises to build a more flexible and agile environment.

An IT transformation inspired by cloud computing

Agility in the face of rapidly changing opportunities and risks is the effective strategy for business sustainability at any company. At Bechtel (2008 revenue of $31.4 billion), big changes come on two dimensions: the geographies where it operates, and an elastic workforce of 44,000 employees, plus contractors, subcontractors, and supply chain partners. The company undertakes some of the world's most daunting construction projects, including the Channel Tunnel linking England and France, the Hong Kong International Airport, and the recently opened Tacoma Narrows Bridge near Seattle. A flexible and responsive IT infrastructure is imperative to managing these mega-projects with agility.

Like most large enterprises, Bechtel previously had an IT infrastructure defined by a provisioning pattern that allocated dedicated servers and storage to specific applications with little concern for utilization or technology standards. The IT organization managed and maintained the infrastructure primarily by manual processes, which did not support agility. To reduce complexity and cost, CIO Geir Ramleth initiated in 2002 a series of IT rationalization efforts, including the consolidation of 27 data centers into 7 and the removal of unneeded duplication and variations in the IT environment. Ramleth calls this the Bechtel Alignment Stage. (See Table 1.)

Milestones	Phase I: Rationalization		Phase II: Transformation	
Motivation	Address complexity and costs		Pursue opportunity for flexibility and competitiveness	
Stages (time frame)	**Disruption** (up to 2005)	**Alignment** (2005 to 2007)	**Transformation** (2007 onward)	**Expansion** (future)
Model	IT centric	Company centric	Collaboration centric	Partner centric (ecosystem)
IT Project, services deployment time	60–90 days	30 days	Overnight	Policy-driven, real-time, self-service
IT workforce	2,000	1,250	1,100	Balanced to needs
Data center size	20+ centers; 35,000 sq ft	7 centers; 20,000 sq ft	3 centers; 1,000 sq ft	TBD
Applications	More than 1,600, with several current versions of each	230, with an average of 4 to 5 versions of each	230, with 3.5 versions of each	~200, ideally with only 1 version of each
Standardization	None	General standardization	Strict standards for optimization (goal of one solution, one version)	Strict guidelines with flexibility for rapid change (Evergreen IT)
Server utilization	2–3%	30–40%	60–70%	On demand
Resource/services provisioning	Manual	Somewhat automated	Policy-driven provisioning of access	Policy-driven access to external IT resources (move toward end-user self-service)

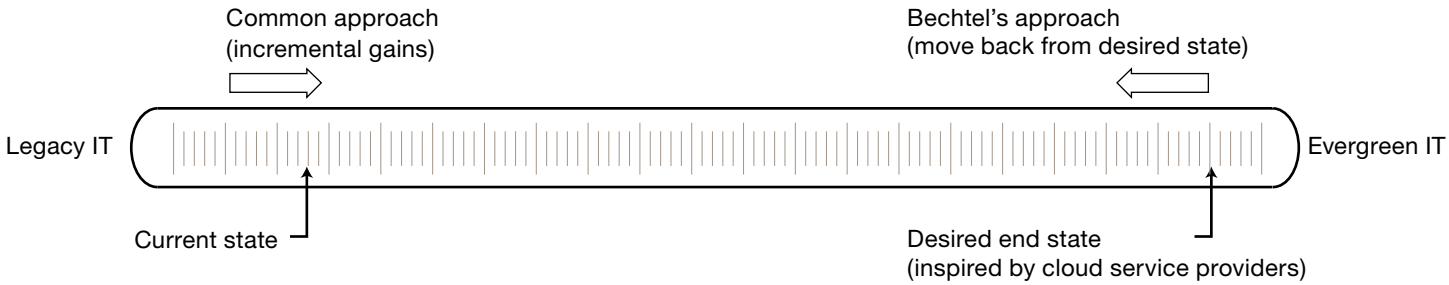
Table 1: **The decade-long transformation of Bechtel IT spanned two key phases and four stages**

Source: Bechtel Corporation

The improvements were significant, but they did not address the fundamental drivers of Bechtel's business that described where IT could perform even better—flexibility and responsiveness to geographic and labor force diversity. Alignment allowed Bechtel IT to reset the base of its problems, but in time the legacy approaches would simply re-create the complexity. "We were not in pain with the old state, but we did see that we would have to take a new approach to have the flexibility for what we needed to do going forward. We had already completed a comprehensive rationalization process and brought about other needed efficiencies. What we needed going forward was a global, flexible collaboration platform," Ramleth says.

Ramleth was convinced that IT needed a new strategy. His staff had collected benchmark data from cloud computing service providers, including Amazon.com, Google, salesforce.com, and YouTube. Even after the alignment stage, Bechtel IT was still orders of magnitude more costly and less flexible than IT at these well-known Internet brands. The external cloud computing providers became Ramleth's inspiration.

In 2007, Ramleth embarked on Phase II to transform IT.[1] (See Table 1.) Rather than focus on incremental improvement, Ramleth saw an opportunity to learn from cloud providers and deliver radically greater flexibility. He also decided to approach the problem as a startup would: If there were no legacy, how would they do things? Ramleth instructed his team to define the end state that would deliver the targeted benchmarks and to compromise on end-state design only where absolutely necessary. (See Figure 1.)



Legacy IT resulting from build-to-order culture		Evergreen IT leveraging configure-to-order culture
$3.75/GB/mo at Bechtel	Storage benchmark	$0.15/GB/mo at Amazon.com
$0.55/server/hour at Bechtel	Compute benchmark	$0.10/virtual machine/hr at Amazon.com
1 admin/100 servers at Bechtel	Server admin benchmark	1 admin/20,000 servers at Google
$500/Mb at Bechtel	Network benchmark	$10-15/Mb at YouTube
230 apps x 5 versions x many upgrades at Bechtel	Application benchmark	1 application image x 4 upgrades/year at salesforce.com

Figure 1: **Bechtel's approach to transformation: Backtrack from a desired end state rather than making incremental gains from the existing state to overcome perceived risks**

Source: Bechtel Corporation

1. For more details on the transformation at Bechtel, please see the interview with Geir Ramleth in this issue on page 16 as well as the following articles:
http://www.networkworld.com/news/2008/102908-bechtel.html
http://www.cio.com/article/453214/Cloud_Computing_to_the_Max_at_Bechtel
http://cloudstoragestrategy.com/2009/03/bechtel-harnesses-the-cloud-a-case-study-in-service-delivery.html

Although the transformation affected all areas of IT, the journey began by radically simplifying the infrastructure of servers, storage, and networking. This simplified infrastructure created the foundation onto which application workloads are now being migrated. In most cases, the applications must be modernized. In all cases, they must be certified to operate in an infrastructure more like the high-performing environments at Amazon.com and Google. Bechtel now has an IT infrastructure markedly different from one aimed at sustaining legacy IT.

The specific benefits of radical simplification are many. Data center space, which had been rationalized in Phase I from more than 35,000 square feet to about 20,000 square feet, was reduced to 1,000 square feet in each of three identical data centers. Standardization, consolidation, and virtualization of servers increased utilization from less than 3 percent to about 70 percent, resulting in a 30 percent savings on operations and management costs. By judiciously choosing data center locations and opting to manage its own bandwidth, Bechtel IT moves 10 times the former amount of data traffic for the same networking expense.

And the simplification drive is extending into the application domain as well. Bechtel's IT staff rationalized the number of applications from more than 1,600 to about 230. Of those 230 applications, 60 percent are custom-built legacy solutions, typically for a specific business unit or project. To support the varied needs of projects on a single application code base, Bechtel is on a path to convert many of them to the software-as-a-service (SaaS) model. The design pattern in this case is to allow individual projects to configure the functionality without changing the underlying code—a capability found with many commercial SaaS offerings.

The Bechtel transformation represents improvement in two key dimensions. It creates a more loosely coupled vertical stack of IT defined by applications, middleware, systems, and storage to eliminate dependencies that cause complexity and inefficiency. And it automates manual IT infrastructure management processes to promote self-service and responsiveness. In both cases, it standardizes assets and processes to reduce complexity. The result is Evergreen IT, a key characteristic of which is the ability to manage different layers of IT independently of each other.

Bechtel proved that by radically simplifying its technology infrastructure, it could more closely achieve the infrastructure characteristics in place a Amazon.com, Google, or salesforce.com. The business directly benefits because IT responds to business needs with greater agility. IT is especially flexible financially and in service delivery. The remainder of this article delves deeper into the IT issues that enterprises face as they work to establish Evergreen IT, and the traits and benefits of this vision.

Legacy IT: Old habits die hard

Time and again, IT has proved its value in individual projects for process automation, data management, and decision support. The results are often stunning, but IT consumes increasingly more of its budget just to run existing applications. Operational expense is typically more than 80 percent of the IT budget. And everyone has a story about IT not fulfilling requests quickly enough. Business agility is definitely constrained by a lack of IT agility.

In simple terms, today's IT solution is a vertical stack of three principal layers: infrastructure, including compute, storage, and networking assets; application workloads that enable business processes; and end-user devices, which provide the interfaces to access the applications. (See Figure 2.) For a long time, management has focused its attention (and investment) on building each individual solution. Few IT staffs look at the entire collection of systems and ask, "Do the IT functions and associated architectures as they have grown over the years still make sense?"

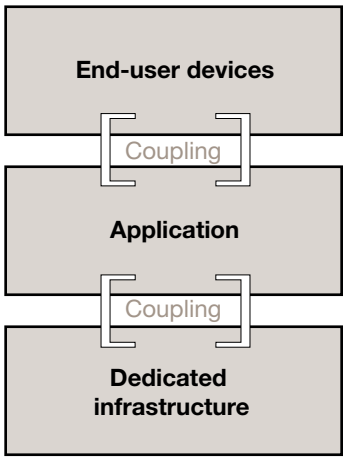


Figure 2: **A simplified view of today's IT stack**

Despite best intentions, most IT organizations have fallen into dysfunctional standard operating procedures in their IT infrastructures.[2] The main culprits for this decline are from two broad areas: vertical stacks and manual processes.

IT solutions are built in standalone, vertical stacks

Most IT projects are individually funded, and the needs of a particular business activity define the IT provisioning requirements.[3] The primary consideration is that the vertical technology stack delivers a solution with acceptable performance. The hardware, application software, and any middleware tools are built to order for the application requirements. This creates an instant legacy problem of interdependencies among process logic, data logic, integration logic, compute capacity, storage capacity, and networking functionality.

Over time, many standalone stacks are built and integrated, creating interdependencies that vastly increase the cost and complexity of making changes to individual layers. (See Figure 3.) As technology advances, the components of the stack become less cost-effective and maintenance knowledge becomes scarce, even for compute elements that may have once been standard.

"The vertical stack of knowledge [the interdependencies between and among applications and infrastructure] about how to make a particular function work in the IT infrastructure is a direct enemy of deploying a new thing, because new things change every one of those silos," explains Simon Crosby, CTO of the Virtualization and Management Division of Citrix Systems. The challenge is not limited to the technology stack alone, but is also mirrored in the IT organization. "The enemy of agility is also the existing vertical integration of expertise and function in the organization structures of IT," Crosby adds.

Too many manual processes, too little automation

While IT delivers ever more capable solutions to enhance business unit processes, the IT unit itself operates using too many inefficient, error-prone, manual processes. The routine functions of finance, sales,

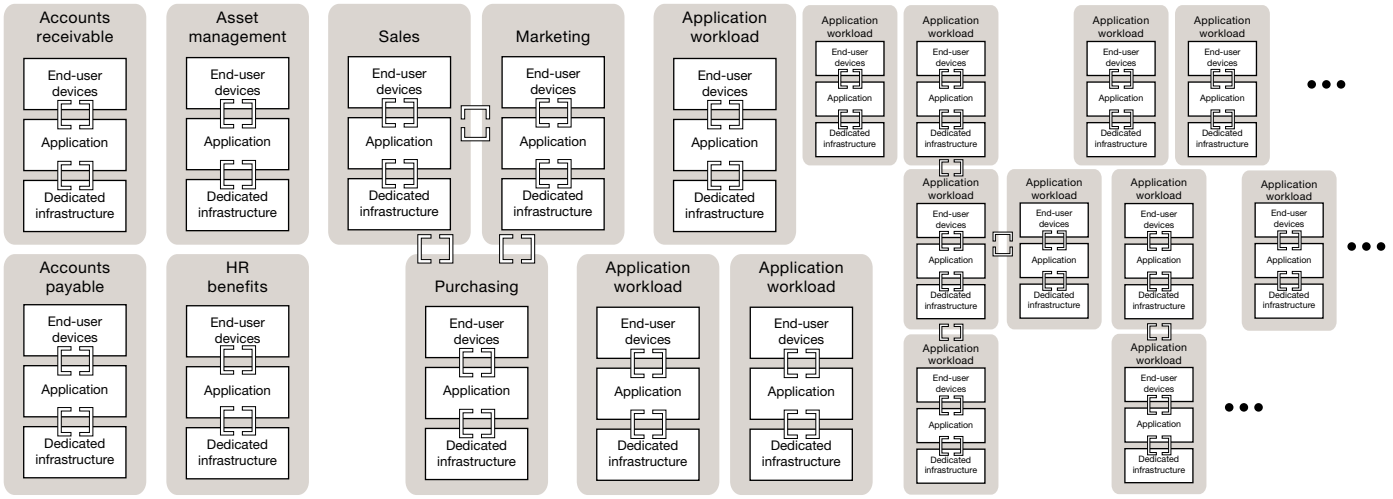


Figure 3: **The typical IT environment is a collection of hundreds or even thousands of tightly coupled vertical stacks integrated with each other**

2. In IT, much of the infrastructure of servers, storage, and networking can be found dispersed across the enterprise and data centers. However, there is a growing trend toward consolidating the infrastructure inside data centers. Therefore in this article, the terms "IT infrastructure" and "data centers" are being used interchangeably.

3. In IT, provisioning is the process of preparing the IT infrastructure with all the requirements (of servers, storage, memory, network, operating system, and so on) of a particular application that needs to be run.

supply chain, and other enterprise activities are highly automated, but not those of an IT organization. Many IT automation tools today are designed to manage one part of the infrastructure. Therefore, they suffer from the same legacy integration issues as everything else in the IT environment.

Due to manual processes, on a percentage basis the growth in infrastructure support costs at most enterprises outpaces the growth in costs for application support, help desk, and other labor. Based on data from Gartner, PwC estimates that between 2002 and 2008, the personnel costs for applications rose 45 percent while personnel costs in the data center rose 95 percent. (See Figure 4.)

IT staff spending and percentage change of IT total spending

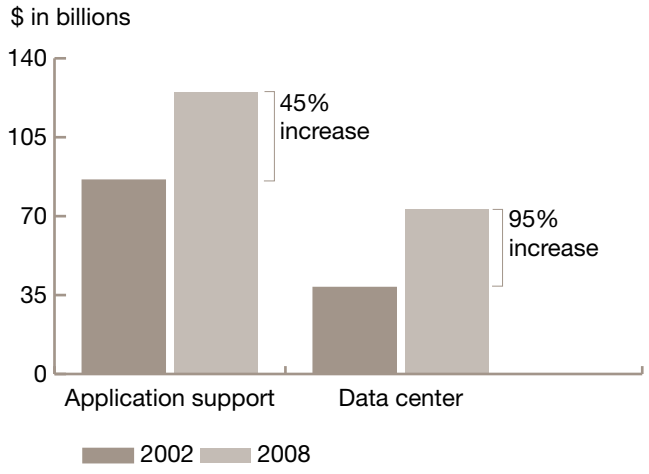


Figure 4: **Between 2002 and 2008, the labor cost to support the infrastructure in the data center rose faster than the labor cost to support applications**

Source: Industry sources and PwC estimates

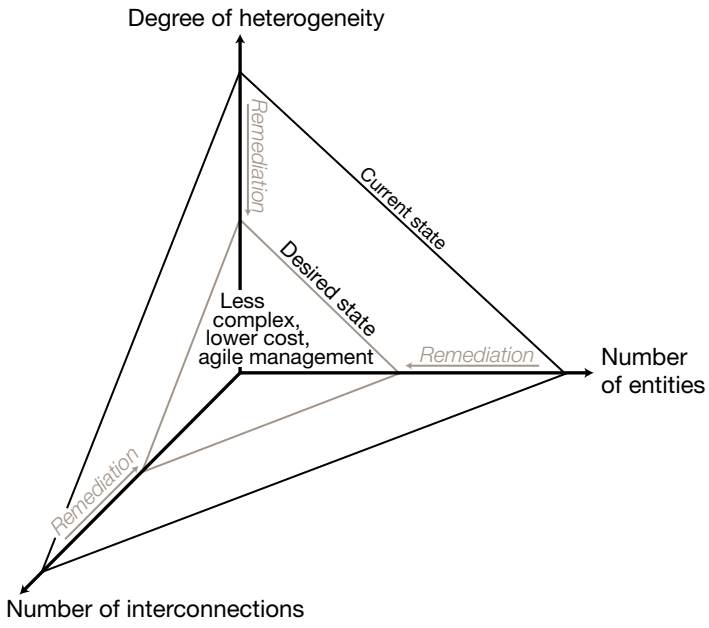


Figure 5: **Three dimensions of complexity**

This increase in support costs is more evidence that IT has not focused on process improvement or using intelligent software in the data center. Scripts passed from one systems administrator to another usually support the process of starting up a server. IT organizations usually maintain little corporate knowledge that describes or controls how changes in one IT asset (for example, a version of Linux) will affect another (for example, a database). The result is a huge focus on break/fix expediencies and much higher costs.

These common practices became an issue over time. “What is ultimately driving cloud computing is a crisis of complexity, driven by the distributed computing model,” says Erich Clementi, general manager of Enterprise Initiatives at IBM. As discussed in the Summer 2008 *Technology Forecast*, complexity grows in relation to three drivers: the total number of entities, their degree of

heterogeneity, and their degree of interconnectedness. (See Figure 5.)

Remediating complexity inevitably requires an approach that reduces one or more of these drivers. In the data center, prevalent IT practices have allowed all three to grow. This is true of the IT components themselves and the practices that define the human aspects of data center management.

Conventional wisdom on how to fix IT

Many enterprises have decided the only solution to dysfunctional enterprise IT is to outsource it. But in most cases, IT outsourcers cannot deal with the silos of built-to-order solutions found in a typical data center. The legacy solutions greatly limit the efficiencies and agility an outsourcer can deliver. Outsourcing transfers the problem from one place to another rather than creating any fundamental change.

More recently, vendors have begun offering cloud computing services as a provisioning option for IT infrastructure. Rather than managing and maintaining built-to-order IT solutions, companies could move them to an external cloud provider. Standard workloads such as corporate e-mail probably can be well managed this way. But IT solutions supporting mission-critical activities are unlikely to be moved to an external cloud provider at present because of reliability and performance concerns. And regulatory issues such as data privacy laws create other constraints. Also, most of today's legacy applications must be assessed and adapted in a time-consuming and potentially expensive "de-verticalization process" informed by knowledge of interdependencies within and among siloed solutions. This is detailed knowledge that most IT organizations have never collected, much less actively managed.

But cloud vendors have adopted IT architectures as well as provisioning and management approaches that are more cost efficient, more agile, and more reliable than most enterprise IT environments. And enterprise IT could benefit—as Bechtel has—by transforming IT systems and processes around new norms based on the best practices of cloud computing providers. By doing so, internal IT can reclaim the mantle of cost-efficient, high-performance computing, and move toward Evergreen IT.

Evergreen IT: Undoing legacy IT

Evergreen IT is a pattern of provisioning, architecture, and operational management that addresses the two drivers of dysfunctional enterprise IT—silos of tightly coupled application and IT infrastructure, and limited use of intelligent software to manage the IT infrastructure. Evergreen IT is an infrastructure architecture designed to deliver loose coupling between distinct layers of the IT stack, and a systematic transition of IT operations from predominantly manual to predominantly automated through intelligent software. In this approach, incremental investment in IT does not create an ongoing burden from legacy systems. All aspects of IT have the potential to be continually refreshed as needed by changes in technology or business, thereby staying current and aligned with the needs of the business.

Figure 6 shows the various technologies in cloud computing and how they support Evergreen IT. The technologies direct attention to legacy IT assets and their vertical and horizontal interdependencies. By capturing and modeling these interdependencies, IT can take steps to reduce complexity in the data center and to transition application workloads to cloud like infrastructures—internal/private clouds or external/public clouds.

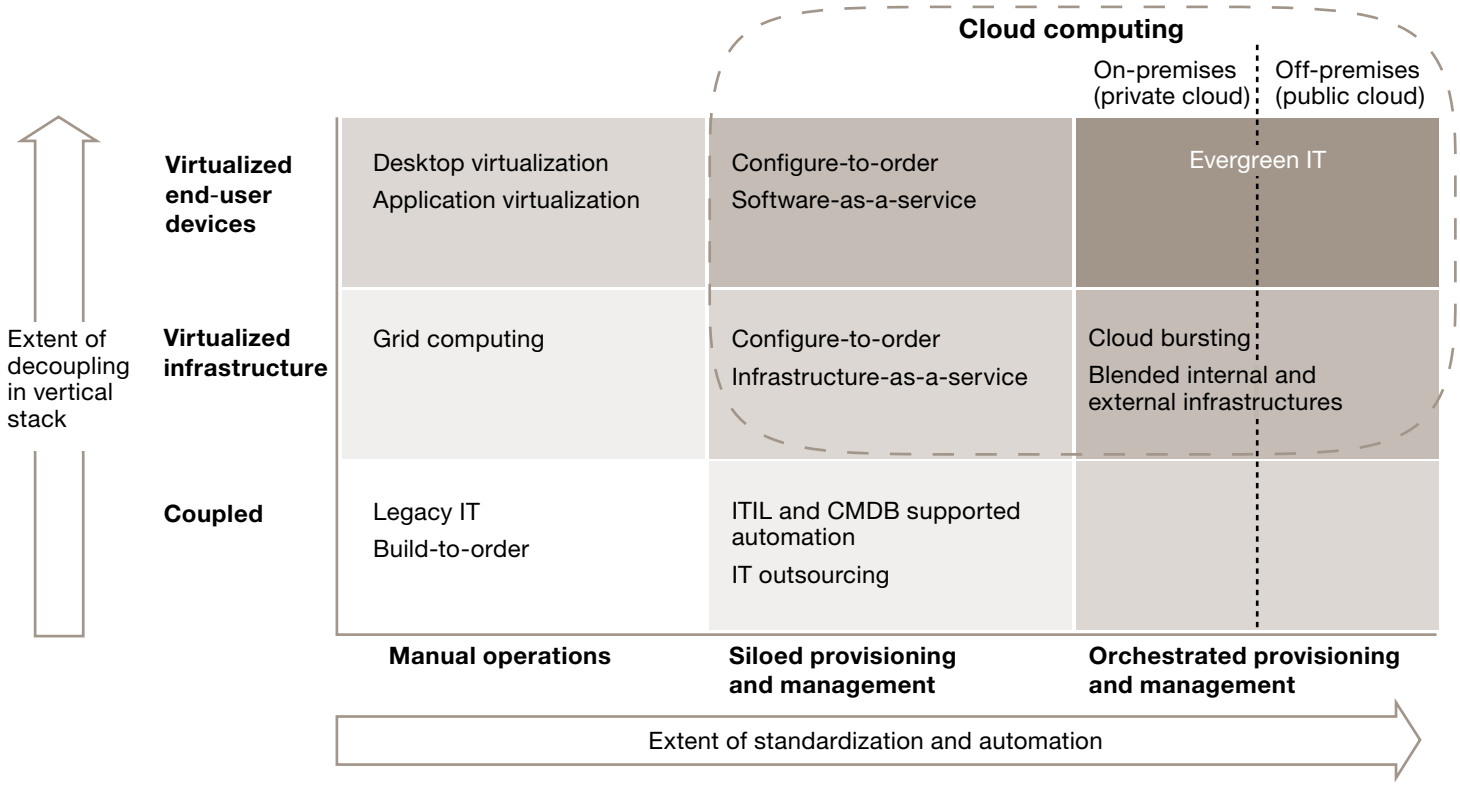


Figure 6: **Principles from cloud computing, when applied to decouple the IT stack and automate the operations, lead to Evergreen IT**

The term "evergreen" suggests that applications, data warehouses, business intelligence, decision support, and other workloads will have non-interrupted access to nearly unlimited compute, storage, and network resources. Application workloads declare their infrastructure requirements programmatically and receive support. At the infrastructure level, IT manages compute, storage, and networking capacity as a procurement flow defined by generic capacity requirements. Reusable and interchangeable components supplement newer, more advanced components on refresh cycles defined by genuine technological improvements and budgetary considerations. In Evergreen IT, new applications are provisioned from a pool of shared resources whether the scope is for one business unit or the entire enterprise.

This shift moves IT away from provisioning applications as standalone solutions, avoiding the creation of legacy technology. The goal is to be legacy free, which means the only reason to swap out a component is because it is broken or less cost efficient than an alternative. The only reason to keep a component is because it efficiently delivers commodity compute, storage, or network capacity.

Characteristics of Evergreen IT

Evergreen IT consists of two main attributes: a loosely coupled IT architecture, and software-driven provisioning and management. They deliver three categories of benefits: lower cost, legacy free, and increased agility.

Loosely coupled IT architecture

Loose coupling between hardware or middleware and application workloads is the critical ingredient in the Evergreen IT architecture. Evergreen IT requires a software layer that sits between the raw compute, storage, and network environment and the applications, shown as the software mediation layer in Figure 7. This software layer mediates differences between what the infrastructure provides in its native form and what the workloads expect. In practice, this means the mediation layer should be configurable to emulate whatever application programming interfaces (APIs) the applications were written to, be they for Windows, various versions of UNIX, or a mainframe operating system.

Early manifestations of this capability are evident in virtualization software. The current practice is to use virtualization for one specific operating environment. For example, data center managers use it to consolidate multiple Windows-compatible workloads running on a plethora of standalone machines onto a single platform running multiple virtual Windows environments.

Technologies are also advancing to virtualize desktop applications and create separation between application workloads on servers and the various client devices—laptops, smart phones, and so forth—to which they are delivered. The article, “Enabling Evergreen IT,” on page 24 provides more detail about the software mediation layer.

Software-driven provisioning and management

The second attribute of Evergreen IT is process standardization and modernization of data center operations, which can be accomplished with or without the IT architecture described previously. Evergreen IT must incorporate intelligent software based on advanced models of provisioning and IT infrastructure management. Given the crucial role IT plays in business today, industrialization of IT operations is long overdue.

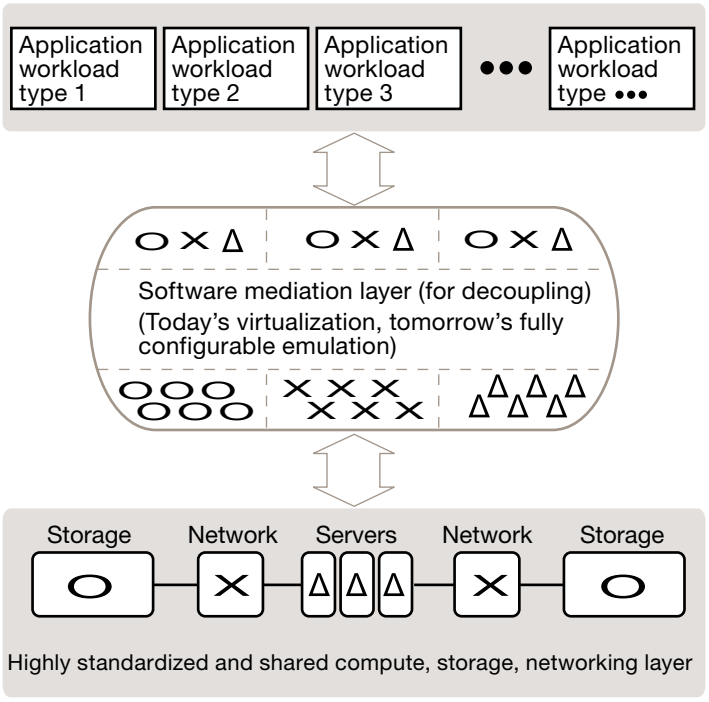


Figure 7: **Evergreen IT depends on a software mediation layer to loosely couple infrastructure and application workloads and to support the automation of the IT processes of provisioning, management, and orchestration**

Automation through intelligent software creates much of the economic benefit that cloud service providers achieve. (See the benchmarks in Figure 1.) By applying the discipline of process formalization and standardization to IT, cloud computing changes the way enterprises consume IT, says IBM's Clementi. "So far, IT has been something like an art," he explains. "Cloud computing is doing to the IT supply chain what Henry Ford did with the conveyor belt. Otherwise you cannot explain the economics."

IT organizations can achieve a new level of data center management by externalizing and formalizing three things: details of a company's IT assets, deep knowledge of asset interdependencies, and management interfaces that extend and integrate with existing systems management tools.

Automation is also the basis for changing IT processes from the build-to-order model to the configure-to-order model. "Automation ultimately ends up being more important, because you can automate both physical as well as virtual environments and support configure to order of more and more complex application architectures," says Russ Daniels, CTO of EDS, an HP company.

Automation shifts where IT can create value. As routine provisioning and management are automated, value moves to how IT orchestrates internal or external resources to meet the demands of the business almost in real time. (See Figure 6.) For example, emerging concepts such as cloud bursting—transferring overflow processing during peak times to other clouds—create hybrid environments combining internal and external IT resources.

Evergreen IT must incorporate intelligent software based on advanced models of provisioning and IT infrastructure management.

Three benefits from Evergreen IT

As Evergreen IT takes shape, enterprises that have significant IT investments will redefine IT complexity to a more manageable, sustainable level. The data center will deliver higher performance in cost efficiency and agility and will be legacy free.

Cost characteristics

Managing technology is expensive at most enterprises. "I think businesses are spending way too much money and time on things that don't bring them value, but there are things that they need to do, such as managing technology environments," explains Irving Wladawsky-Berger, chairman emeritus of the IBM Academy of Technology. "As a result, they don't have enough time to focus on the things that do bring them value," he continues.

By reducing complexity, Evergreen IT can have a profound impact on costs. Among the results, the capital cost of IT infrastructure for any new application will be much less. Bechtel reports closing more than a dozen data centers and retiring hundreds of servers. (See Table 1.) Bechtel also reports that its version of Evergreen IT reduced IT operations costs by 30 percent. More important, with shared resources, the cost of infrastructure management grows much slower than the rate of IT infrastructure growth. And by loosely coupling the pool of raw compute, storage, and networking capacity, IT can buy commodity technology that most closely tracks Moore's Law in price/performance. Advances that would be available only to applications written to a specific environment can be claimed by all applications.

Evergreen IT returns economies of scale to IT, something the creeping complexity of dysfunctional IT has prevented.

Legacy-free IT

Legacy IT has gotten a bad name due to the high costs of maintaining old applications and hardware, and deservedly so. The tight coupling of hardware and software typically seen in today's IT solutions often means that IT organizations must maintain

Evergreen IT shifts the dependency in the IT stack from hardware to software by decoupling the infrastructure from the application workload via a software mediation layer.

non-economical servers and storage devices to support legacy applications that manage key processes. And to keep these systems running, data centers must retain staff who have idiosyncratic, non-transferable knowledge.

Evergreen IT shifts the dependency in the IT stack from hardware to software by decoupling the infrastructure from the application workload via a software mediation layer. Companies can select software mediation solutions on the basis of the legacy environments they support and retire old hardware. Staffing and skill sets can be standardized around a small core set of technologies.

In effect, this decoupling makes the data center legacy free. Hardware's useful life (the replacement return on investment [ROI] logic) will be defined by cost per millions of instructions per second (MIPS), per bytes, or per bandwidth—not by ties to particular application workloads.

Improved IT agility

PwC has stated in past issues of this journal (see the *Technology Forecast,* Summer 2008) that business agility is created when businesses leverage both standards and flexibility, and decide when each applies on the basis of customer value. Sustainable agility occurs by synchronizing that which is stable (standard) and that which is flexible through a mediation layer rich in information semantics. See "Crossing the big divide between strategy and operations" in the *Technology Forecast*, Winter 2009.

Evergreen IT replicates this architecture exactly. (See Figure 7.) The IT infrastructure defines highly standardized and efficient IT resources. Normally these workloads would need a variety of IT operating environments. But the software mediation layer in Figure 7 can configure the infrastructure to the needs of the workloads. By being standard and configurable to a wide range of workload requirements, the infrastructure can flexibly support many workloads that vary in function and scale—creating an agile IT infrastructure.

Conclusion

The history of IT has been dominated by build-to-order systems that result in legacy environments characterized by vertically integrated stacks of infrastructure and applications, and managed by nonstandard, manual operations. Legacy IT soaks up much of the available IT budget and is a primary barrier to IT responsiveness and overall business agility.

Benchmarks from external cloud providers show that radically better efficiencies and flexibility are possible. Most cloud service providers started without any legacy burden, so the challenge for a typical enterprise is to modernize in a way that moves it away from re-creating more legacy complexity and cost.

To accomplish this modernization, PwC urges enterprises to pursue the vision of Evergreen IT—IT that is increasingly free of costly legacy infrastructures. To do this, IT must adopt an architecture that creates loose coupling between the IT infrastructure and application workloads. It also must modernize and automate IT's own internal business processes for provisioning, managing, and orchestrating infrastructure resources.

Evergreen IT does not need to happen in one big rip and replace of all applications. Rather, the process starts with the creation of a standardized infrastructure onto which compliant workloads can be migrated sooner rather than later. Over time, IT organizations can expect to move more complex application workloads to this standardized infrastructure as software mediation (virtualization) technology matures. Most important now is to stop adding to the legacy solutions. Every workload moved to Evergreen IT will free up time and budget for IT to create greater business value.

For more information on the topics discussed in this article, contact Brian Butte at +1 704 344 7600 in the US or Gary Delooze at +44 (0) 161 245 2582 in the UK.

Evergreening IT at Bechtel

Geir Ramleth of Bechtel describes how he used cloud computing principles to transform IT and make Bechtel's computing environment more agile.

Interview conducted by Vinod Baya, Bo Parker, and Gregg Agens

Geir Ramleth is senior vice president and CIO of Bechtel Group, where he leads the global Information Systems and Technology organization. Prior to rejoining Bechtel in January 2002, Ramleth held leadership positions for a variety of companies in the technology and communications industry. Among his responsibilities in those roles, he defined strategy and market position and secured multiple forms of financial resources.

Previously, Ramleth served as chairman and CEO of DigiPlex S.A., based in Zurich, Switzerland, which owned and operated large communications and networking facilities throughout Europe. Prior to that, he served as president and CEO of Genuity, a Bechtel Enterprises company formed to deliver Internet and data services. Ramleth has held management positions with Oracle, PageMart, and PacTel Personal Communications.

In this interview, Ramleth shares his insights on how he is transforming IT at Bechtel to move away from managing, maintaining, and building legacy solutions to an environment that leverages virtualization, standardization, and automation to be flexible and deliver the right services anytime, anyplace, and to any device that has a browser and an Internet connection.

PwC: You have been leading a major transformation of the IT function at Bechtel. How did you get over the hurdles common to such transformations?

GR: Many of the hurdles that people think are there are really perception hurdles. They're not necessarily true hurdles. I believe that the technical side is only 20 percent of the problem. Change management is 80 percent of the challenge, and getting your own IT people to step outside their comfort boundaries is a key part of this. When you get them outside their comfort zones, you can get them to do things in a different way.

At one of our planning meetings, I laid out about a dozen and a half benchmarks from Internet services and consumer-based providers. I said, "If they [cloud service providers, Amazon.com, Google, salesforce.com] can achieve totally different operating models, think what we could do." I went through all these different benchmarks and said, "You know, if people can deliver it for free, why does it cost so much when we do it?" After that first day, we concluded that we can't directly use any of these 15 or 20 companies' services. We would have to go out and learn how they [deliver at these price targets, at this speed, or with these delivery dynamics]. We decided that we needed to build it ourselves.

PwC: Such a transformation is quite ambitious. What made you think you would succeed?

GR: If your approach is to modify your current state, you will probably end up incrementally better. But if you go to the desired end state and move back only where it's absolutely necessary, you will probably end up with a much more profound transformation. So we said, "Assume we have no heritage. Assume we're starting from scratch. Assume that we actually are a startup that doesn't have over a hundred years of experience and suboptimized IT legacy." They started doing that, backing off only when needed, and that's how we got to where we are now rather than just an incrementally modified state from the past.

We then carried out the plan by designing the desired end state and building it. We made no modifications to existing infrastructure—meaning we built new data centers, we built new networks, we built everything new. We used a rigorous application certification process, and nothing was allowed to come from the old state to the new one unless it passed that certification. Without it, you would just end up diluting a new set of services.

PwC: What was moved from the past state to the new state?

GR: No hardware was moved at all. The only thing that was moved was software. From a fundamental base level, all software was new. But when you look at specific applications, that can't necessarily be the case. Now we're moving SAP, for example. SAP has been certified, so we're going to move it over.

PwC: What were the problems with the old state that you wanted to fix?

GR: We were not in pain with the old state, but we did see that we would need to take a new approach to have the flexibility for what we needed to do going forward. We had already completed a comprehensive rationalization process and brought about other needed efficiencies. What we needed going forward was a global, flexible collaboration platform.

When you look at our business, our concerns incorporate far more than those of Bechtel itself. Our world is really much broader than that. We operate in more and more locations for shorter durations than we used to. We don't have many large permanent offices that last 10 or 20 years—some offices might last only for a project's duration or maybe even just a portion of a project's life cycle.

And the people who participate with us are not all Bechtel employees. They are contractors; they come from the customer side; they come from the supply chain side; and sometimes they are our competitors. So if you take the geographical diversity and the project staff diversity as the real challenge, what's your strategy for IT? How do you ensure intellectual property protection, for example? You want to open it up to be available to anybody anytime anywhere, but

"We had already completed a comprehensive rationalization process and brought about other needed efficiencies. What we needed going forward was a global, flexible collaboration platform."

how do you then protect yourself? How do you move from everything sitting behind a firewall to doing everything outside the firewall?

We asked ourselves, "Now, wait a moment. Nobody sent you to school to learn how to buy a book from Amazon.com. Why should you have to learn how to do business with us? Why shouldn't we be just as easy to do business with as when you do a Google search?" To get there, we found that you need to accept a new security paradigm. That's really the biggest challenge on the technical side.

Milestones	Phase I: Rationalization		Phase II: Transformation	
Motivation	Address complexity and costs		Pursue opportunity for flexibility and competitiveness	
Stages (time frame)	**Disruption** (up to 2005)	**Alignment** (2005 to 2007)	**Transformation** (2007 onward)	**Expansion** (future)
Model	IT centric	Company centric	Collaboration centric	Partner centric (ecosystem)
IT Project, services deployment time	60–90 days	30 days	Overnight	Policy-driven, real-time, self-service
IT workforce	2,000	1,250	1,100	Balanced to needs
Data center size	20+ centers; 35,000 sq ft	7 centers; 20,000 sq ft	3 centers; 1,000 sq ft	TBD
Applications	More than 1,600, with several current versions of each	230, with an average of 4 to 5 versions of each	230, with 3.5 versions of each	~200, ideally with only 1 version of each
Standardization	None	General standardization	Strict standards for optimization (goal of one solution, one version)	Strict guidelines with flexibility for rapid change (Evergreen IT)
Server utilization	2–3%	30–40%	60–70%	On demand
Resource/services provisioning	Manual	Somewhat automated	Policy-driven provisioning of access	Policy-driven access to external IT resources (move toward end-user self-service)

Table 1: **The decade-long transformation of Bechtel IT spanned two key phases and four stages**

Source: Bechtel Corporation

PwC: Can you provide a high-level overview of your journey?

GR: We see our current cycle, a decade of IT, in four stages. [See Table 1.] The first one was disruption. That was when the business said, "You know, now IT is messing with me again. IT costs too much." It was when ERP [enterprise resource planning] was installed. The vendor said that the world is going to be so great, but all we had was pain, and it's slow, and it doesn't really change the business the way the vendor said it would. So that was the first phase, the disruption phase. It was very IT centric.

The second stage, alignment, was when we heard from the business, "Hey, at least they are working on what they should be doing, and we understand the cost structure." So you're aligned with the business and where it is going. During alignment, we went from 27 to 7 data centers, and we made the Internet our friend. But we were still very focused just on what was happening in the company, and the IT group was very focused on what they do from a technology standpoint. This stage was very company centric.

The third stage was transformation, motivated by our need to address the nature of our business, which I talked about earlier—how business opportunities result in frequent changes in geographies where we are present, and how the people we deal with shouldn't have to learn how to do business with us. How do you transform yourself to just be in the Internet rather than attached to the Internet? This is our current stage and we define it as global-collaboration centric.

The last stage, which is in the future, is business expansion, where we become partner centric (with direct and indirect participants).

PwC: How has IT become more agile through this transition?

GR: In the disruption stage, when we wanted to fire up a new project, it took us about 60 to 90 days to get a project ready from an IT standpoint. In the alignment stage, we got that down to about 30 days, which is really good. Now that we've reached the transformation stage, we're basically saying, "Why shouldn't it be available overnight? Why shouldn't it all just be open in the same way that Amazon.com is open?" We've come to realize that we needed to build a security environment driven by policies and not one that's driven by topology.

So instead of saying, "I'm going to get you in here behind this firewall," you're saying, "No, I just want to know who you are and what you need, and then by policies I will create the rest for you."

We think of this as our virtual foundation. You don't have to have a Bechtel login. You use whatever your e-mail address is, just like when you go to any modern Internet site. As long as we know that and we can authenticate you with that, we can give you what you need.

This is a huge improvement over how bad it was previously. At one time, when using our former authentication approach of assigning internal user IDs and such, one-third of the people on our core network were not Bechtel employees. We don't have the same right to do background checks for people who come on our network from the outside. The partner or the customer can say, "Add these 200 people. They are now a part of our project." We have to add them without the same rigor that we apply to our own users. We don't know who they are other than a name. With our new model, it doesn't matter. Our trust and security model is set by policy. When you seek more confidential information or try to access more secure IT elements, we put on stricter and stricter authentication, but it's all done by policy.

> "The important thing is that we standardized. Now, we very much live by the power of one. Do it one way, in one place, using one mechanism, using one approach, and all operated by one group."

PwC: Complexity is a big concern in many IT organizations. How did you deal with that in your transformation?

GR: We said, "Let us make this as simple as possible," so we went for a single vendor. I say half-jokingly that our old data centers were like a hardware hall of fame. We had one of every vendor's products. For example, now our computing and storage is from HP, and other solutions follow the same model. The important thing is that we standardized. Now, we very much live by the power of one. Do it one way, in one place, using one mechanism, using one approach, and all operated by one group.

Initially, we even tried to run one data center in one location. That didn't work. Our applications are not designed for a high degree of latency. So we designed three data centers: one in the UK, one in Singapore, and one in the US. But we said they would have the same hardware, the same setup, and they would be wired the same way. Then you can say, "The red wire that goes into that Cisco device should be in slot number 4." We have only one operational group—so we essentially have one virtual data center.

Originally we thought we could manage infrastructure diversity in software. But we decided that we couldn't do it. If we had tried, we would still be sitting here writing business cases for it, I'm convinced. The best, most direct way forward was to start solving the complexity problem from the infrastructure layer up, and when we have one of those applications that doesn't port to our new environment, we envelope that one and we find a new solution for that.

PwC: Did you need to make any changes in how you select and do IT projects?

GR: One thing that was fundamental in this change was to avoid designing solutions for the lowest common denominator. Rather than focusing designs on a functionality that has to work for all, we moved to designs that covered 95 percent of our business needs. We decided to design a base-level architecture for the masses and find a solution for the uniqueness.

Of course, we heard, "What about that project that sits in the Saudi Arabian desert? This will not work for them." And we basically said, "OK, it will not, but let us plan for the other 95 percent of our business and get special solutions for the unique needs." That was a huge fundamental change in how we developed our design criteria.

For example, to support unique circumstances, we created what we call a SNAP, which is a server network access point. It's approximately a half rack, and it's an absolutely consistent extension of our highly standardized data centers. That SNAP has all the networking gear that the project needs—wire, telephone, all that stuff. It can have 30 to 40 processors and terabytes of storage, and, most important, it has exactly the same kind of hardware that's in the three data centers. It runs the same software. And the operations staff is one group that sits in two locations (US and India) to provide 24x7 service.

We just ship it to a project location and all it needs is input of power and some form of connectivity. We have a new project in a very remote location where the only connectivity we have is satellite. The operations can't run on one of our three data centers, so we put a SNAP in there instead.

PwC: One of the problems with legacy solutions is that you have to keep running them even when usage is minimal. What's your approach to solving that issue?

GR: It's true. Today we run hundreds of servers in support of projects that are no longer active because we don't know what to do with them. For example, we have projects that might be within the warranty period. Utilization is extremely low. We run it, we back it up, and we maintain it as if it is a real production system,

"The best, most direct way forward was to start solving the complexity problem from the infrastructure layer up."

because we haven't had a reasonable alternative. Now we're figuring out how we can put a "time capsule" around it.

PwC: What do you mean by time capsule?

GR: It's an interesting concept that something like the Amazon EC2 service may solve. Let's say you need a server and you configure it, and, after you're finished running what you need, you shut it down. In the process of shutting it down, you have the option of creating a current image of the environment, which stores the entire state of your application. Then you can come back and say, "That server we ran a few months ago that had this stack, we called it ABC." And you say, "Fire that up again," and in a matter of a few minutes, that server is up and running exactly as it was when you shut it down.

What you've done is started untangling the huge, complex array of legacy systems. You can take this down to the lowest level of where the software is and extract it, and plug it in again. It might not go to the same infrastructure next time at all, but it will be up again exactly the same way as it was before.

PwC: And do you rely on vendors for that, or have you needed to build that yourself?

GR: Here we lucked out a little bit. We started working with Simon Crosby at XenSource to build the virtual server environment more than three years ago. We became one of their first enterprise customers. When Citrix bought XenSource, we took a proactive approach and said, "We have had this relationship with XenSource, and we're interested in creating a much more strategic partnership with you as a company, because we have a vested interest."

We told them where we really wanted to go—that we wanted to virtualize the server and the desktop—and we assumed that Citrix bought XenSource because they wanted to get to that state. So we built up a very strong relationship with Citrix that is just great at this point.

I believe we have come to a stage in enterprise IT where all the great young companies that were once legacy killers have become legacy companies themselves. In essence, they've all won that previous battle, but now they're invested in holding enterprise IT to their own approaches. The result is a new legacy that is destroying all the flexibility these companies offered when they were the emergent providers. All that flexibility is gone.

PwC: Is there an approach that doesn't recycle the old legacy for a new legacy? Something that might be called Evergreen IT, where you can have a future where your IT is legacy free. Do you think what you have done allows you to avoid recycling into a new legacy base?

GR: I think Evergreen IT is a good way of describing our ultimate goal.

PwC: Cloud providers such as Amazon.com appear to be much closer to this vision of Evergreen IT by investing in standardization and automation of IT operations, particularly the provisioning and management of infrastructure. How much scope is there to automate in enterprise IT?

GR: There's a lot, because we still "tinker" too much. There is still too much knowledge in the heads of operations staff that needs to be formalized and standardized in software. I wish we could run to our internal servers the way you can with Amazon.com—where you can fire up machines in their data center. I don't know where the servers actually sit, and, at one point, I will not care, but suddenly I have an environment running on my IP [Internet Protocol] space.

I'd like to take our peak loading requirements and offload that to somebody, like some of our time capsule stuff. You start to get into cloud bursting here. In the Evergreen IT sense, in the foreseeable future, you might have a little core IT capacity, and then you tap into external providers for whatever else you need beyond that.

In our business, it could be for a specific project, or it might be for some specific load—like one payroll process that bursts far above normal for a few hours twice a month. In an enterprise IT environment, you buy everything for peak loading. In 2002 we did an analysis of our servers parked around the world, and our

average utilization was 2.3 percent. With our second-generation virtualization, we're starting to get that up into 60 and 70 percent.

PwC: There are many concerns about risks of security in the cloud. How did you mitigate risk in your case?

GR: The point is perceived risk. I believe that the information in our systems is more secure today than it was in the typical legacy environment, because we have tightened up that back end tremendously. Think about it. How many times do you hear about people hacking into the back end of Amazon.com? It's very secure. You don't hack into the bottom layer of where Google is operating. It's very secure.

Basically, you change just the demarcation of where you have flexibility and where you don't have the flexibility. Behind the demarcation, you tighten up much more than you ever did before. Like I said, we had security policies before. It was just too difficult for anybody to find intruders when they came in. The sheer complexity of the legacy environment defeats all reasonable efforts to control intrusions.

In the old way, you had to make all these exceptions. You had to open up this for this purpose, that for that purpose, and then you would forget to close it again. We now operate in a much more standardized way, and in the process have become more secure than in the past.

PwC: How have you financed your transition?

GR: We decided that we were going to do the transition without asking the company for money. We realized that we needed a financial architecture just as we needed a technical architecture. So we planned our new activities as an integrated part of the existing operations and budgets, and, as a result, we had no overall cost increase. Over time, we strongly believe that this will give us more financial flexibility. It should increase our variable cost and decrease the fixed cost component that IT has tried to tackle for years. I firmly believe that one of my responsibilities is to build an agile operations model that includes financial flexibility, and such operations will have greater success in responding to business needs.

PwC: What were some of the changes that you had to make to your IT organization?

GR: We began with an IT organization that was extremely distributed, both from a technical and operational standpoint. Then we started finding ways to leverage off each other to deliver services. We called these global leveraged services. So you have a global leveraged service such as data center operations, and one for networking. There are real advantages to this model beyond cost reduction.

We formerly had 33 help desks in the company and none of them had 24-hour service. We couldn't afford it. But now, with one virtual help desk you can have 24-hour service. You get the back-end systems so that they all can work on a global fabric. They have only one help ticket system, and we have only one phone number, even though it moves with the clock.

PwC: Have you been successful in rationalizing your legacy portfolio of applications?

GR: During our first phase of rationalization, we took our application count down from about 1,600 to 230. We still maintain 800 applications, because for the 230, on average, there are three-and-a-half versions that still need support.

A key decision has been to develop multitenant versions of our core software. One key application is InfoWorks, and we formerly maintained 16 different versions of it. Very early on we created a new version of InfoWorks that is a multitenant environment. This allows us to run InfoWorks in a software-as-a-service model, where we are the service provider for any of the projects. Multitenancy was a requirement due to the nature of our business.

With most of these services, we're finding that you can buy many of them and they serve our subcontractors very well. But commercial services can't support us internally when it comes to our complexity. So the procurement system will be a modern multitenant system, and we are working on that today. We're not rewriting it, but we are modifying it to the new requirements. A full rewrite would be costly and unnecessary, because this large, well-designed suite of applications has 15 years of development in it.

PwC: But you're still going to run it on the standardized infrastructure, right?

GR: Yes, and we will run it on one set of code instead of our current situation that requires multiple versions. We will reduce operating costs, and we will be able to deliver faster.

PwC: What would you say at the highest level is your vision for IT?

GR: Our vision is that we want anybody to be able to have access to the right resources at any place at any time with any device. And we want to provide this all in a secure and cost-effective manner—cost-effective meaning that we have the flexibility we need, not necessarily that we want to be lowest cost, but we want that flexibility.

I'm after the tens of thousands of end-user devices. Look at this scenario. You can say that your recycle rate for end-user devices is about three years. The cost over that period is roughly $10,000, of which your hardware upfront is about $1,500 to $2,000. The $3,000 a year is all the tinkering we have to do, often driven by changes from the vendor community.

When you decouple that, when you move away from enterprise-controlled and -supported end-user devices, we are not worried about where you come from, because we might not even give you the data to your machine any longer. That data might all sit in our environment, and, as I said earlier, our data centers today are more secure than they ever were before.

And then we can say, "Well, you know, this person will never be able to download anything." It's just a totally different end-user environment strategy and philosophy. You can come to the BYOC, bring your own computer, model where you just say, "Buy anything. I don't care what you buy. As long as it can connect to the network, we can offer our services to it."

PwC: In other words, from an Evergreen IT standpoint, we're talking about complete decoupling so that you're legacy free. Would that be a requirement?

GR: Oh, absolutely. You can't get there without it. You must have that to be evergreen. If not, you can't deliver on any information, any device, anyplace.

PwC: You mentioned four phases. You've talked about three: disruption, alignment, and transformation. What's the fourth phase?

GR: We call it the expansion phase. It's when IT adds profound value to the customers who buy our products and services. That is when somebody chooses to buy, say, an oil refinery from us over somebody else because we could deliver differentiated, value-add IT services. At that time, it's mostly information, actually. For example, it might be when a customer starts operating faster, because the system is already in place by the time they take it over from us.

We did this for a project in China, and our key customer said they would never do it any other way after that. They shortened their startup time by several months and they reached higher capacity than they had planned for. In this stage, you have to take a life cycle view of everything, including infrastructure, applications, and information that is generated or used both inside and outside your organization. In addition, what you really need to do is virtualize your workspace, so that, as long as you're connected to the Internet, you can get your work done. In addition to your basic computer environment, this will include all the other resources you are accustomed to having, including telephony, print services, and file sharing. ■

"Our vision is that we want anybody to be able to have access to the right resources at any place at any time with any device."

Enabling Evergreen IT

Technologies to create separation in the IT vertical stack and to automate IT operations form the basis for pursuing Evergreen IT.

While most IT environments choke on costly and inflexible legacy systems, the on-demand, self-service offerings from external cloud computing providers promise agility, flexibility, and cost-effectiveness. In the vision of the external cloud, IT organizations, corporate departments, and individuals gain push-button access to nearly unlimited computing, storage, and network capacity, all of it managed efficiently. Business-class applications are similarly on tap, configured-to-order, easy to access and use, and priced by the drink.

It's enough to give even skeptical IT executives a case of cloud computing envy.

The vision that PricewaterhouseCoopers (PwC) calls Evergreen IT will give IT managers the best of both worlds: the security, control, and focus that legacy systems promised (if not always delivered) and the flexibility, ease of management, and cost-effectiveness of cloud computing.

The Evergreen IT vision is a pattern of IT provisioning, architecture, and operational management designed to deliver loose coupling between logically distinct layers of the IT stack. It is indeed inspired by, and will eventually use, external cloud providers. However, the goal of Evergreen IT is not cloud computing per se. The goal is to transition IT operations from predominantly manual to predominantly automated processes driven by intelligent software. The result is an approach where incremental IT investment no longer creates legacy systems. Instead, each layer of the IT stack can be continually refreshed without worrying about interdependencies between layers.

Two sets of technological advances are vital to this vision of Evergreen IT: the use of virtualization to decouple legacy IT's largely customized technology stack, and the formalization and modeling of data center operations to facilitate automation. (See Figure 1.)

This article takes a close look at each set of technologies and how they contribute to Evergreen IT.

Virtualization's role in achieving Evergreen IT

The most "real" technology enabler for Evergreen IT today is virtualization, a group of technologies that are fundamental to separating the morass of technology deployments stacked and integrated in the typical IT environment.

The goal is to transition IT operations from predominantly manual to predominantly automated processes driven by intelligent software. The result is an approach where incremental IT investment no longer creates legacy systems.

There are several types of virtualization, each focused on a different type of computing:

- **Server virtualization**—Server virtualization puts an entire server environment (operating system [OS], hardware, and connectivity) onto a virtual machine so it can be run on any compatible server, typically in parallel with other virtual machines on the same server to increase hardware utilization.
- **Storage virtualization**—Storage virtualization pools together multiple physical data stores into one logical pool.
- **Network virtualization**—Network virtualization creates logically separate communication channels within and across local area networks (LANs) to allow multiple types of network channels on the same infrastructure.

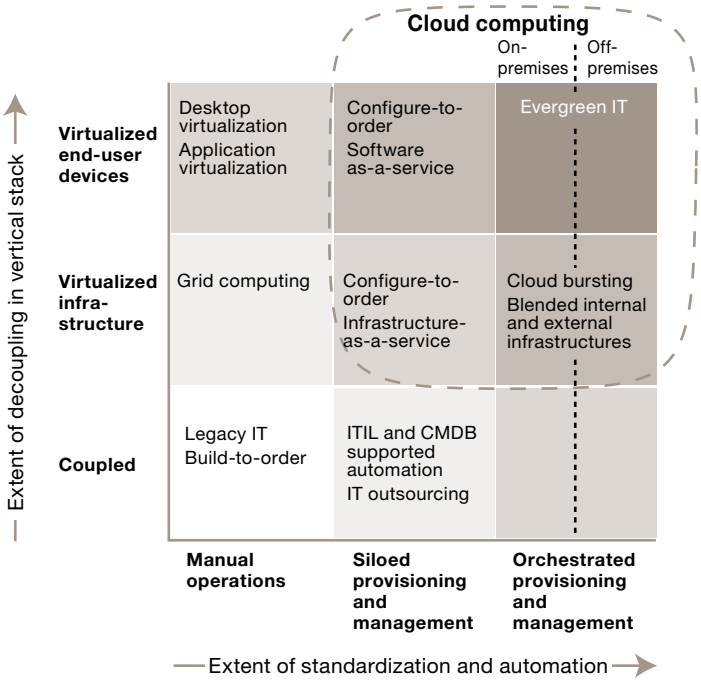


Figure 1: **Evergreen IT results from the decoupling of the vertical stack and the automation of IT operations. Cloud computing technologies are enablers of the vision of Evergreen IT.**

- **Desktop virtualization**—Desktop virtualization puts an entire PC environment on a virtual machine so it can be provisioned to users from a common, centrally stored "image."
- **Application virtualization**—Application virtualization takes applications and their supporting OS capabilities and creates a logical package for each application that runs as an independent layer on an OS so applications don't affect each other or the underlying OS.
- **I/O and memory virtualization**—Emerging types of virtualization include I/O and memory virtualization, both of which break down physically separate sets of computing resources into more flexible logical groupings.

In the Evergreen IT vision, the infrastructure virtualization group—server, storage, network, I/O, and memory—is critical to the ability to both standardize the IT infrastructure and make it more flexible to handle the legitimately nonstandard technology platforms. The desktop and application virtualization technologies extend this flexible standardization to the client side while increasing IT's ability to manage and control the far-flung client systems. Most CIOs today are focused on infrastructure virtualization for valid, but ultimately overly narrow, efficiency reasons.

Eventually, virtualization can make IT more strategic by bringing focus to layers that an enterprise should own and manage for differentiation. "If you apply virtualization rigorously up the stack, you get an opportunity to cross those vertical silos at the appropriate level for your enterprise," says Simon Crosby, CTO of the Virtualization and Management Division of Citrix Systems. "You can then examine whether it is your core competence or mission to manage the level below that particular virtualization abstraction."

The value of infrastructure virtualization

Each application workload needs a particular environment—processor, memory, operating system (OS), storage, networking, and power—to run. Virtualization technology is a software layer that runs

directly on the hardware or in combination with the OS, and this software layer creates and manages virtual machines. A virtual machine is a software simulation of a specified complete server, storage, or other computing environment. It tricks OSes and applications into "seeing" the same hardware interfaces that a real computer would present—hardware instruction set, data storage, network access cards, and so on—thereby removing the dependency between the application and specific hardware resources. (See Figure 2.)

The concept of virtualization has been around for four decades. In the 1970s, IBM offered hypervisor software to allow one mainframe to host scores or hundreds of virtual machines at once, each appearing to its user as a self-contained computer running its own OS, disk drives, and peripherals. In the 1990s, IBM extended the hypervisor to run UNIX and related OSes on its mainframes, not just IBM's OSes. By the end of the decade, other manufacturers such as Sun Microsystems offered similar virtualization capabilities on their hardware.

As client-server computing took hold, interest in using virtualization for UNIX and Linux server environments surged. The issue was server sprawl—hundreds or thousands of standalone systems, each often running at low utilization levels. The standard operations pattern of the time was to dedicate one server (or cluster of servers) to one application, optimizing it for one application's maximum workload requirement. Because average utilization levels are always lower than peak levels, most servers ran at extremely low utilization levels on average (less than 20 percent was typical, and some companies have reported average utilization rates of less than 3 percent).

Clustering servers produced limited benefits because clusters typically needed to run the same OS on all servers. Clustering across administrative domains and security contexts was not possible without a new layer of abstraction. The development of hypervisors for the x86 instruction-set machines and other hardware platforms provided this new abstraction. Today, hypervisors—such as Citrix's XenServer, Microsoft's Hyper-V and VMware's ESX—allow one physical server or group of servers to support multiple OSes, including Windows, UNIX, and Linux, in what is now called server virtualization.

The concept of virtualization can apply to the entire IT stack, including storage and networking. Instead of associating specific disk and tape drives with specific servers, storage virtualization creates virtual drives. To programmers, applications, servers, end users, and administrators, these software-defined drives look and act like the real thing. They offer push-button provisioning, easier administration, more-efficient utilization of storage hardware, and improved automation of tasks such as backups. Major storage system providers—EMC, Fujitsu, Hewlett-Packard (HP), Hitachi, and IBM—have gone well down the path of storage virtualization.

Network connections within the data center also can be virtualized. For instance, in systems running virtualization software from VMware or Cisco Systems, physical Ethernet adapters can be configured to run many virtual adapters and switches, respectively. Each device can be dedicated to a specific virtual machine, offering programmers familiar interfaces and operating characteristics. The big advantage is being able to virtualize Ethernet switches normally used to create LANs of servers. With virtual switches, creating and reconfiguring LANs is a matter of changing software settings—not pulling wires, moving and reconnecting physical devices, and keeping track of physical addresses. A virtual network is open to automated management in ways that a traditional network is not.

Going beyond infrastructure virtualization

Virtualizing the connection between applications and user devices is also becoming more common. (See Figure 2.) The Wintel (Windows/Intel) PC environment has offered enterprise IT a relatively standard client or device environment. However, the increasing use of smart phones, netbooks, tablets, and other devices creates a growing imperative that applications be accessible securely and reliably on any device by any user anywhere. And even the fairly stable Wintel environment is rife with variation, from multiple versions of Windows in use to multiple brands of hardware that

have different components in use. Even when companies have fully standardized their PCs, their contractors, customers, employees, and others use their own PCs, bringing variability back into the mix.

One approach to addressing this variation is referred to as desktop virtualization, which creates a virtual machine that contains the OS, drivers, and an abstraction layer. To the OS and its drivers and applications, the abstraction layer "looks" like a PC. Developers often use such virtual machines to run multiple OSes, or multiple instances of the same OS, for testing across platforms and configurations. Mac OS X users often use this technology to run Windows on their computers.

But the promise for IT is greater: Because a virtual machine is software, it can be provisioned into a computer as needed from a standard "image" library. This means that IT can ensure standard deployments across all users. IT can even remotely load and unload the virtual machine on user desktops from the data center, ensuring security, predictability, and compliance. Ceedo, Citrix Systems, EMC's VMware unit, InstallFree, and Xenocode are all active in this technology area.

Application virtualization takes the desktop virtualization concept to the next level, encapsulating applications and their supporting drivers into mini virtual machines that can be loaded and unloaded independently, without affecting the OS registry or other applications that might use different versions of the same drivers. IT could centrally provision applications as needed and isolate rogue or personal applications from the supported applications, again bolstering security and compliance in addition to easing management.

With a fully virtualized, software-defined client, users are free to move seamlessly among office and home PCs, laptops, and handheld devices, and among different brands of hardware and OSes. IT managers like the approach because security, backing up data, and maintaining client software can all be addressed centrally, uniformly, and at a lower cost.

The decoupling benefits of virtualization

The various virtualization approaches all decouple technology layers from each other, reducing or eliminating dependencies among them—a critical enabler of the Evergreen IT vision. This decoupling has the paradoxical effect of allowing greater standardization and greater flexibility. For example, a standard Windows image could run on any number of PCs, with no need for individual image files for each hardware configuration. And multiple OSes can run on the same hardware, simplifying hardware acquisition and reducing IT management efforts.

Decoupling the application workloads from the underlying infrastructure or from the client devices creates many benefits not previously available. These benefits include the following:

- **Flexibility**—Virtualization maintains simple but useful abstractions of enormously complex and constantly changing computing environments. To an application and its programmers, the virtual machine offers exactly the same interfaces as the real computer. Thus, the app's coding and the programmer's skills and knowledge don't need to change when the underlying hardware undergoes substantial change. At the same time, hardware can be refreshed continually to benefit from advances in processing and management without being constrained by the application or its configuration. The infrastructure remains evergreen rather than contributing to the creation of a legacy stack.
- **Mobility**—The abstraction level of a virtual machine captures the entire execution environment, so any given virtual machine can be moved to another physical server in the same data center or, possibly, elsewhere. Typically, IT staff would perform such a move to provide the virtual machine with more computing capacity and boost the performance of the applications. Such mobility is also an important element of virtualization for business continuity and disaster recovery.
- **Standardization**—Although virtualization solutions can orchestrate and mediate a multivendor hardware environment, such an environment is not the desired end state for Evergreen IT. Achieving the price/

performance benchmarks of external cloud providers over the long term requires a commitment to hardware standardization as well. Standardization is also a principal way to address the complexity that is a reason for lack of IT responsiveness.

How virtualization facilitates legacy migration to Evergreen IT

Beyond enabling the Evergreen IT infrastructure, the use of virtualization also helps enterprises migrate their legacy applications to the Evergreen IT environment and shed their dependency on a complex and costly infrastructure. (As discussed in the next section, the formalization and modeling of data center operations also play a key role in this migration.) Absent virtualization, enterprises must convert legacy applications to run in the new environment, often a costly and uneconomical chore.

Most migrations entail multiple steps. First is physical consolidation (as enabled by virtualization products) to pull together the servers into a single compute fabric. Second is tailoring resources to the needs of the applications (also called hypervisor automation). Third is enabling on-demand capability so applications are executed only when they are requested (such as by using technology from Cassatt, a company CA recently acquired). Fourth is dynamically imaging and moving applications to provide business continuity (such as by using VMware vSphere). The fifth and final step is managing applications so enough instances are available to meet the demand (such as by using DataSynapse technology).

The fifth step is a more difficult effort than the other steps because it requires the parallelization of the application—rewriting at least some of the application to run in multiple, parallel instances to take advantage of greater resources for scaling. Many enterprises blur

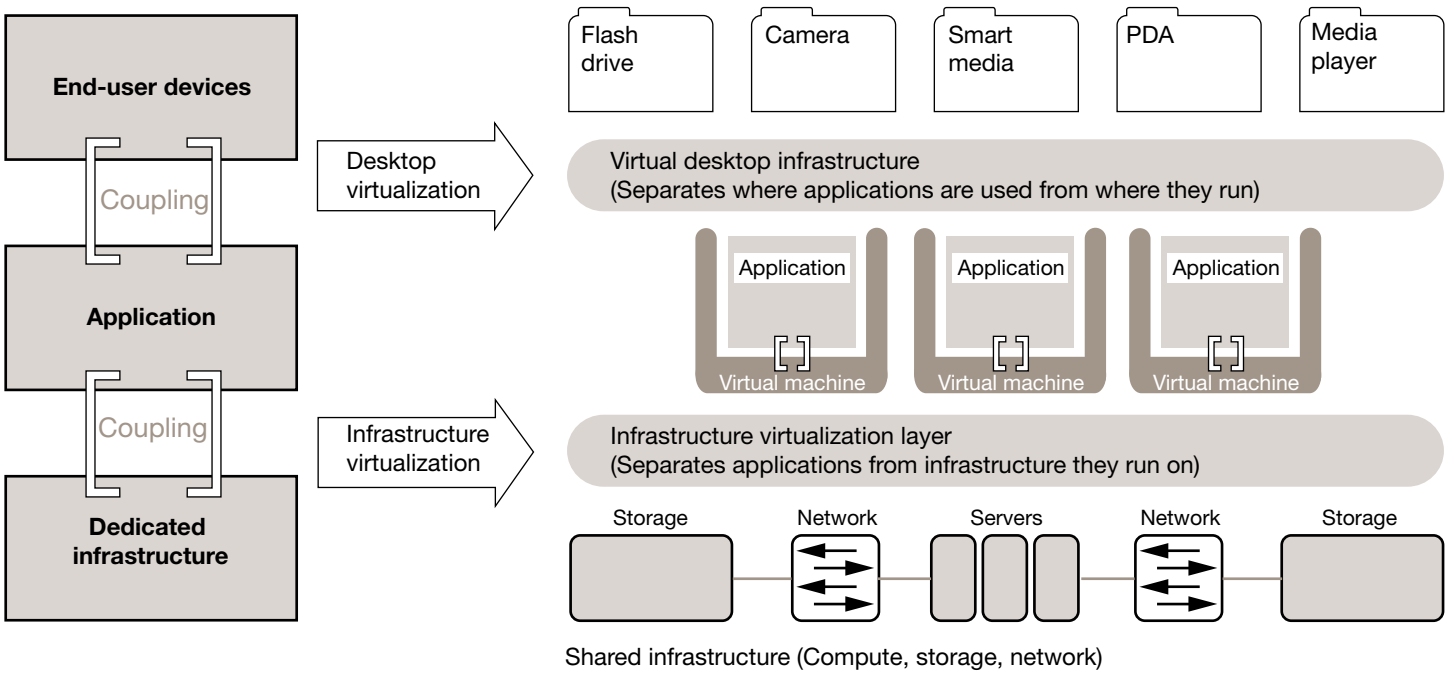


Figure 2: **Virtualization at two key layers is necessary for the vision of Evergreen IT. One is to create the separation between the application workloads and the infrastructure they need to run. The other is the separation between applications and end-user devices.**

the line between moving an application into a virtualized infrastructure and parallelizing the application to improve its scalability, attempting to do both in the early migration stages. In fact, most applications can be moved into a virtualized infrastructure as is without parallelization; enterprises can then decide later which applications need the parallelization effort.

Over time, technological advancements from vendors should enable the migration of a greater number of application workloads to virtualized environments. (See Figure 3.) But many impediments remain. Non-technical impediments occur primarily in two areas: One is a matter of internal IT interest in and readiness to embrace change; the other is the lack of vendor support for applications running in virtualized environments, both in terms of licensing flexibility and applications being certified to operate in a virtualized environment—and ultimately remaining under contracted support by the vendor.

The popular IT press has focused on broad performance and security issues as technical impediments to virtualization. Typically ignored are the issues related to the unexamined assumptions from the world of fixed and tightly coupled resources that break in a world of

A virtualization vendor sampler

Many established and emerging vendors offer virtualization solutions that enable decoupling at the various levels of the IT stack.

Ceedo: This company's virtualization solution allows IT administrators to create and deploy a desktop application workspace to end-user devices such as a PC without requiring installation on it. Workspaces can be deployed directly to a PC, virtual PC, a network drive, or a portable device; decoupling the application and data from the end-user devices.

Citrix Systems: XenServer is an infrastructure virtualization product. The base version is available for free and supports features such as live relocation and multi-server management. A more feature-rich offering is available for a price and adds features, such as high availability and dynamic resource balancing. Many external cloud providers use and support virtualization based on XenServer.

InstallFree: Focused on the Windows environment, InstallFree Virtual (IFV) is the company's core product for application virtualization. Their product will decouple, encapsulate, and distribute stateless applications, add-ons, and updates in self-contained modules that dynamically bind to each other at run time on the host machine and integrate with user-specific customizations.

Microsoft: Hyper-V is a role in Windows Server that provides tools and services to create a virtualized server computing environment. It is a hypervisor that supports Microsoft operating systems and applications and Linux as a guest operating system.

VMOps: VMOps Cloud Stack is designed for enterprises and cloud service providers. It includes a hypervisor for virtualizing processors, storage, and networking, and a cloud server for administering the provisioning of virtual machines while managing physical resources and balancing workloads.

VMware: ESX Server was an early entrant in the x86 virtualization market and is the most widely deployed hypervisor that supports Windows, Linux, Netware, Solaris, and other guest operating systems. ESX Server is the foundation for the vSphere product that adds features such as centralized management, load balancing, business continuity, and live migration.

Xenocode: Xenocode Virtual Desktop allows application components and settings to be deployed in pluggable "layers" that can be dynamically enabled or disabled. Their solution automates the steps a skilled Windows administrator takes to package and deploy an application.

virtual and variable resources. For example, applications and operating systems are designed with a series of assumptions about the resources, such as clock time, network addresses, and physical connectivity as well as operating characteristics of latency and function-call response times. When a virtual machine is stopped and then started later, the operating environment created by the virtual machine must manage any application requirements for the appearance of continuous time without unexplained time gaps.

Citrix CTO Crosby warns that virtualization presents a challenge for storage and networking, too: "It used to be that the workload was statically configurable to the server, but now that workload can move anywhere, and that breaks the networking abstraction. Because the state [of the application] is required to protect and enforce policies, it [must] be [adjustable], too. [This mobility also] breaks storage, because the unit of storage must be visible to the server that it's about to run on."

Solutions from virtualization vendors have proprietary interfaces, raising CIO concerns for a new round of vendor lock-ins, this time at the virtualization level. IT also has unmet needs for unified management interfaces to reduce the complexity around a growing pool of virtual resources, in particular how the management tools for virtual resources integrate with existing management tools in use at most enterprises.

Vendors are working to address these technical impediments. The sidebar "A virtualization vendor sampler" presents a sampling of vendors that are active in providing virtualization solutions.

Modeling's role in achieving Evergreen IT

In addition to virtualization, Evergreen IT relies on a more automated IT operation. The goal is to replace labor-intensive, nonstandard processes with formally defined, automated processes that are invoked and monitored automatically by software. The main benefits are speed and flexibility achieved through improved orchestration—the ability to provision the right mix of IT resources to handle changing application workloads.

As Russ Daniels, CTO of EDS, an HP company, puts it, the data center is undergoing a fundamental shift from build-to-order to configure-to-order, and automation is playing a crucial role in enabling that shift. But the automation is not possible without a model of IT operations that forms the basis for the automation software's actions and that the software uses to monitor and trigger its actions.

To be configurable, an IT environment should have a standardized infrastructure and set of processes; otherwise, enterprises build a series of one-off customizations that quickly become too complex to manage, much less model or automate. The use of virtualization can lead toward the necessary standardization, as well as make the technology components more malleable and thus appropriate for being managed through automation. But without a clear set of operational processes, defined parameters, and expected behaviors—the IT operational model—enterprises cannot take advantage of that malleability in any significant way via automation.

The beginnings of a model of the data center

Fully automating data center operations requires new thinking, and several approaches are in the air. Perhaps the most promising is to model the IT infrastructure and all related processes in exacting detail: all physical and virtual components; all significant inter-relationships among those elements; and all operational processes and managerial policies that affect, or "touch," these elements and govern their usage. An example of such an IT policy is the requirement that all enterprise resource planning (ERP) data be backed up every hour on the hour. This approach seeks to capture the details of every significant planning and operational activity in data and to use software to intelligently automate the management of IT.

Modeling like this is commonplace in factories, and it is a foundation for the design of manufacturing processes and their continual improvement. In an automobile manufacturing plant, for instance, chassis and components such as bumpers and engines move along carefully engineered paths, each timed to arrive exactly when and where it's needed. Once the basic process is

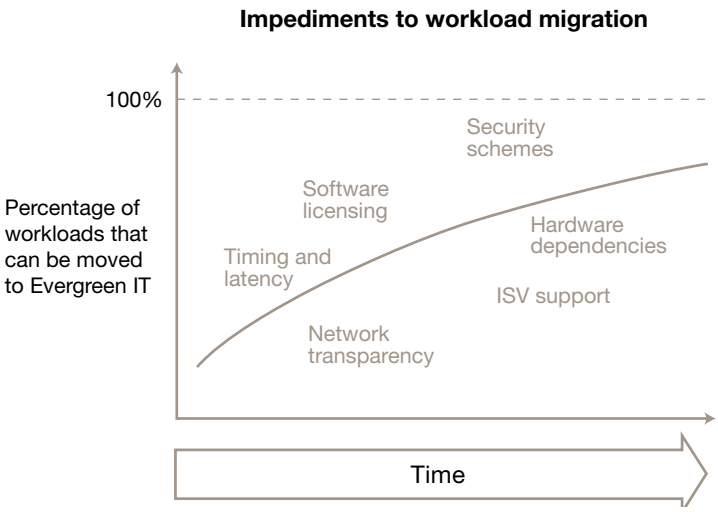


Figure 3: **The proportions of workloads that can migrate to an Evergreen IT environment will grow over time as vendors address many of the technical and non-technical impediments.**

in place, its many paths and steps can be repeatedly optimized and further automated as conditions demand.

Large data centers are missing this kind of thinking, primarily because they've accumulated an array of disparate systems and components. Even without so much diversity, the sheer number of elements, interdependencies, policies, and procedures is staggering. The time has come, and the technology is maturing, to comprehensively model the data center and prepare it for the industrialization it lacks.

"The traditional approaches to automation are significantly focused on imperative [how to do it] logic. So [for provisioning], either it's scripts, and scripts are software with a particular notation; or it's workflows, which are software with a particular notation. It's all imperative logic," says HP's Daniels. "To successfully automate more and more complex kinds of configure-to-order deliveries, you really need to focus on forms that are more declarative [what to do] rather than imperative in their expression," Daniels continues. "We have a strong focus on how to capture the relevant information in information models and then allow systems software to perform introspections on those models and to understand what must happen in a sequential workflow-centric approach."

Data centers already have something like this in the typical network management system. It is as useful as it is because it provides a complete description, or map, of the network: every router and switch, and every circuit—real or virtual—that connects those boxes. With this map in machine-readable form, it is possible to automate processes—such as provisioning new services—so the processes become a simple matter of reconfiguring the appropriate network elements.

Taking this idea a step further leads to the creation of a complete model of the data center that describes all hardware and software elements, their logical interdependencies, and the policies that govern each element (for example, the data-backup policy for a key transaction-processing application). With such a model, enterprises can take the automation of data center operations to a new level.

Some IT organizations are more ready than others for this automation. One test is how far they are in adopting the processes and best practices known as the IT Infrastructure Library (ITIL). ITIL describes processes for delivering IT in the form of services along with strategies for continually improving those processes and the delivery of services over time.

ITIL documents describe processes in some detail, but in an abstract way that leaves each enterprise to work out its own implementation. For any existing data center infrastructure of size and scope, an as-is state must be captured and defined first before moving to an ITIL-inspired future state. Individual enterprise requirements will include unique IT deliverables that no generic framework can be expected to cover. What has been lacking are data center management applications that incorporate management frameworks, such as ITIL best practices, while also facilitating an understanding of the as-is state and supporting unique management challenges. Moving workloads between self-hosted and service provider infrastructures quickly and reliably will require a level of automation not present in most data centers today.

Elastra, a San Francisco startup funded in part by Amazon.com, is one company working to realize this vision. It envisions an executable or machine-readable model of the data center that enables IT to also have the standard, automated processes found in other corporate functions. Kirill Sheynkman, Elastra's founder, president, and CEO, says: "HR has an automation system, the sales force has an automation system, and finance has always had one. Supply chain management has had an automation system. IT has not. There are bits and pieces, but there's nothing that holistically looks at IT as a process that creates a product."

When an IT organization has fully modeled its data center and operations as Elastra envisions, automated procedures could arrange for applications to be moved between servers, new IT services to be provisioned, and workloads to be moved to other data centers that have radically different hardware resources. For example, an enterprise might wish to run the Web-facing portion of a certain transaction-processing e-commerce application in the external cloud while keeping the back-end database, containing sensitive customer information, in-house.

According to Elastra, the modeling of infrastructure gives IT managers tremendous flexibility in where to run existing workloads. The first step is to describe the existing IT environment, which will consist mainly of dedicated vertical stacks. (The company has developed its own markup languages, called ECML and EDM, to build such models.) The result is a mapping of the infrastructure's capabilities to meet the needs of a certain workload. For example, a model may capture the performance thresholds used to determine when a specific application requires a change in hardware capacity.

The model may be used to configure other infrastructure—even if built from different components—so it, too, can run the same workload and adhere to specified policies. Most often today, setting up new infrastructure to handle an existing workload is accomplished manually. In a decoupled, highly virtualized infrastructure, the provisioning can be fully automated.

As Elastra's Sheynkman points out, new, virtualized IT setups could be created on demand as the business needs them. Once all the life cycles for deploying a certain system are described, he explains, "we can write an optimizer that determines the optimal way of deploying that system, and we can version it and modify it on the fly. We can use software to define an allocation policy—for example, how resources are deployed—and then map a design to an allocation policy according to the situation for which it's needed."

The cloud, Sheynkman says, is forcing IT managers to think of infrastructure "as a pool of processing, a pool of storage, a pool of networking resources that I can use to do things. Abstraction of the machine is what [the] cloud gave us, so [Elastra] said, 'Well, the next interesting thing for enterprise IT is the abstraction of the application.' "

Elastra is not alone in developing these tools. Several other companies—mostly startups—are developing solutions to address the automation of IT operations. See the sidebar "An IT automation vendor sampler."

The path to IT automation

The automation of IT processes will unfold in two stages. (See Figure 1.) The first will address the provisioning of infrastructure necessary to run a specified workload.

According to Citrix's Crosby, the basic objective of automation is to have an administrative interface that IT can use to do fault, configuration, accounting, performance, and security (FCAPS) management, and a user interface that allows business users, on the basis of their roles, to drive their own applications without requiring IT involvement.

External cloud service providers today claim to have achieved this level of automation, but where they in fact deliver on their claims it is because they offer self-service, on-demand provisioning for a limited set of configurations and/or workloads. (Enterprises today have much more complex sets of workloads to handle, so replicating a cloud provider's capabilities across that complexity is no more than a fantasy today.) Over time, PwC expects the external cloud providers to offer a

wider range of configurations for more hardware platforms, OSes, middleware platforms, and so on. And this will expand the providers' ability to support the range of applications enterprises use. In other words, as IT simplifies its operational complexity in the pursuit of Evergreen IT, external cloud providers will be able to handle increasingly complex environments—and at some point the two will have comparable environments that allow the intermingled use of internal and external cloud technologies.

In the meantime, IT can create its own cloud using technologies such as virtualization and emerging management tools. These, combined with a consistent set of IT operational processes, will allow IT to take advantage of model-based automation tools as they emerge.

Automation also will expand to include the orchestration of internal and external resources to optimize their use according to policies and business objectives. Although PwC expects most enterprises to embrace Evergreen IT by building the capability on-premises, we also expect that, over time, they will also use resources from external cloud service providers for business continuity, disaster recovery, cloud bursting, and other purposes. As choices expand, automated orchestration driven by policies becomes necessary for the sophisticated use of resources balanced with business needs and objectives.

Note the change from "provisioning" to "orchestration." Provisioning focuses on configuring the infrastructure to run a workload, whereas orchestration focuses on deployment, run-time monitoring, and policy-driven changes during the workload life cycle. In other words, the kind of agile, "smart" technology management long sought by enterprises.

With the ability to jointly orchestrate internal and external resources, techniques such as cloud bursting will become easier to employ. In cloud bursting, an enterprise keeps cloud-based computing in reserve, ready to call it into action to handle a sudden spike in demand or to keep the business running after a disaster. Orchestrating also will make possible new forms of distributed computing, in which applications are managed as single entities running across internal and external clouds.

An IT automation vendor sampler

Many vendors offer IT process automation products. Most listed here enable automation in the data center as enterprises adopt or migrate to private clouds.

3tera: The company's AppLogic solution can be used as a hosted service or to build on-premise private clouds or as a hybrid cloud solution. AppLogic provides a catalog of preconfigured virtual appliances and applications templates, run-time cloud application manager and monitoring tools, to allow service providers and enterprises to create and operate scalable applications without reengineering specifically for cloud..

Arjuna Technologies: The company provides tools, called Arjuna Agility, for moving legacy IT systems to a range of cloud infrastructures and for managing their operations there. It includes an XML-based scheme for describing applications and automatically configuring their resources.

Elastra: Elastra Enterprise Cloud Server 2.0 enables enterprises to design, deploy, and manage applications and virtualized infrastructure in a holistic way in internal and external clouds. The core technology is a method for semantically modeling applications, infrastructure, and all policies—including minimum performance levels, backup schedules, and such—that relate the two.

Enomaly: The company's Elastic Computing Platform, derived from an open-source project, automates the configuration, management, and deployment of virtual machines in private, public, and hybrid computing clouds. This Platform, has a virtual machine management system written in Python, and the back-end storage is on MySQL.

GridGain Technologies: GridGain, written in Java, is an open-source development platform that helps construct new cloud-based applications and grid-enable existing ones. It uses a software framework that supports large data set distributed computing by including data grids, affinity load balancing, zero deployment, peer-to-peer class loading, and other concepts.

Majitek: GridSystem software helps IT rapidly provision and scale high-performance business applications and services on virtualized, heterogeneous infrastructure. The company also provides tools and services to securely provision, license, and bill for digital content or services.

RightScale: The company offers a suite of tools for managing cloud-based applications. Included are predefined server templates to help get new applications running quickly, a real-time performance monitor, and customizable scripts that can scale infrastructure automatically according to demand for applications.

Skytap: Skytap supplies cloud-based IT lab services to help enterprises develop, test, and demonstrate large-scale applications. Its virtual, on-demand infrastructure pools can be used in conjunction with or as a replacement for on-premises IT resources.

Surgient: Software from Surgient automates the life cycle of creating, testing, and delivering IT services. This includes managing and optimizing heterogeneous virtualized infrastructure, managing libraries of IT service configurations, and enabling self-service, policy-driven provisioning of services.

VMware: vCloud is a data center operating system to enable enterprises to create on-premises private cloud infrastructure that can federate computing between on-premises and off-premises resources. VMware has partnered with technology and cloud service providers to enable migration of application services to a cloud environment regardless of operating system or application.

Zimory: Zimory has developed software for combining virtual servers, even in disparate data centers, into a homogeneous computing cloud. The technology is aimed at enterprises and is the basis for an Internet marketplace that trades in on-demand computing capacity aggregated from enterprises around the world.

Conclusion

Legacy IT is the fundamental reason IT is not flexible, responsive, and efficient, as we explain in the article, "The real promise of cloud computing," on page 4. IT departments need to rewrite their playbooks for infrastructure design and data center automation, using the Evergreen IT principles outlined in this issue of the *Technology Forecast*. Only by doing so can enterprises break the cycle of creating more legacy systems to manage and maintain. The enabling technologies—notably virtualization and model-based automation—are now beginning to emerge to permit enterprises to begin the Evergreen IT migration.

Although the full vision of Evergreen IT will take many years to achieve, the journey starts with the infrastructure. By starting there and moving up the stack as experience and technology mature, enterprise IT will reap tremendous short-term benefits and will be ready for ever greater advantages in the future.

For more information on the topics discussed in this article, contact Larry Best at +1 646 471 4889 or Bob Zukis at +1 213 217 3222.

Using virtualization to break vertical integration in IT silos

Simon Crosby of Citrix explains how virtualization creates separation in the IT vertical stack.

Interview conducted by Vinod Baya and Bo Parker

Simon Crosby is the CTO of the Virtualization and Management Division of Citrix Systems, and he was the founder and CTO of XenSource prior to the acquisition of XenSource by Citrix Systems. Prior to XenSource, Crosby was a principal engineer at Intel where he led strategic research in distributed autonomic computing, platform security, and trust. Previously, he was the founder of CPlane Inc., a network optimization software vendor. Before joining the private sector, Crosby was a tenured faculty member at the University of Cambridge, UK, where he led research on network performance and control, and multimedia operating systems. He is the author of more than 35 research papers and patents on data center and networking topics, including security, network and server virtualization, resource optimization, and performance.

In this interview, Crosby shares how the vertically integrated stacks in IT are a key barrier to agility and how virtualization technology is the software layer that will allow enterprises to bridge IT silos at the appropriate level based on their core competence.

PwC: Cloud computing is getting a lot of attention these days. How much disruption does it represent? Is it of the same scale as the Internet?

SC: It's a profound change of the same order as the Internet—maybe somewhat smaller in the sense that today, cloud is pretty much articulated in the context of enterprise uses. But it's of the same order in the sense that the Internet built the pipes to everybody—consumer and enterprise—and the cloud can deliver the apps to anybody—consumer or enterprise. It's a different way to deliver all the applications that we consume today. I think it's going to take a long time to get there, but it is a built-in process of maturing IT along the lines of maturing any utility function, so it's about IT maturing into a planetary-wide utility.

PwC: We're interested in the role IT infrastructure plays in enabling business agility. What are some of the key barriers to IT being responsive to businesses today?

SC: Today's IT infrastructures are built from the equipment of today or 10 years ago and the successes of a scaled-out environment, which was entirely dependent on a growing economy. Given the dynamics and the fast pace of business, throwing in a new server and all the gunk that ran on it, and racking and stacking, was easier than addressing the fundamental management challenges of a scaled-out environment.

Now enterprises find themselves with these messily scaled infrastructures and an incapacity to deal with the ensuing complexity. Wherever the technology doesn't provide the right degree of scale or visibility or anything

else, you build an organization of humans around it to make sure that it runs. You end up with an organization that has a practiced excellence at handling servers, storage, the network, the various functions in the network, security, compliance, and all of these different things in the enterprise.

Additionally, vertically integrated groups understand how to make firewalls work, say, or how to make servers rack and stack and show up in data centers. But unfortunately, that vertical stack of knowledge about how to make a particular function work in the IT infrastructure is a direct enemy of deploying a new thing, because new things change every one of those silos, and to get a new thing deployed is going to require the IT team to change how a firewall works or the configuration of a particular set of devices and so on. Every one of the silos gets hit by a new deployment.

The enemy of agility is the existing vertical integration of expertise and function in the organization structures of IT. Even if you could get somebody to put a new VM [virtual machine] on the server, a network person will need to configure the VLANs [virtual local area networks] so all the traffic will show up, or configure the firewall purposely to let the traffic through. The organizational structures that have evolved are incompatible with the desire to be agile, and that is nobody's fault. It's property of the fact that the technologies were not mature and not self-scaling. What cloud gives us is a very interesting opportunity to change that.

Moreover, the IT infrastructure of today's enterprise is massively diverse. The multivendor nature of it also conspires to cause complexity. Previously, multivendor sourcing was the enterprise's friend in terms of retaining cost or pricing advantage from multiple vendors. Multivendor sourcing is now rearing its head in terms of complexity, and this is the primary reason why IT is so expensive.

PwC: You have led the development of virtualization solutions, and it is one of the key technologies in cloud computing developments. What role does virtualization have in addressing the concerns you have raised?

SC: The concept of virtualization is much richer than virtualization between an OS [operating system] and a hardware component, which is the most common use of it today. Virtualization is a layering function horizontally up the enterprise IT stack, between, say, servers and OSes, between OSes and applications, and between applications and end-user binding.

If you apply virtualization rigorously up the stack, you get an opportunity to cross those vertical silos at the appropriate level for your enterprise and examine whether it is your core competence or mission to manage the level below that particular virtualization abstraction. If it isn't, then you can outsource it.

So, for example, if it is not your core mission to manage large data centers full of servers, then you can outsource that, focus on the things above it, and simply rely on the hypervisor as a way to separate and allow the IT stack to be dynamically bound with any set of servers run by somebody else.

PwC: It is common to see the cloud universe segmented into infrastructure-as-a-service [IaaS], platform-as-a-service [PaaS], and software-as-a-service [SaaS] layers. Which level will be more important going forward?

SC: Yes, there can be many layers. At the end of the day, I'm passionate about one layer, which is the bare-bones IT infrastructure-as-a-service layer. It's extremely powerful, because it addresses some key needs for scale and cost efficiency, while not locking in the

"The enemy of agility is the existing vertical integration of expertise and function in the organization structures of IT."

customer. That is, the notion of compatibility and portability is much easier to achieve than if you write an app against somebody's platform as a service. It's very important to preserve the notion of compatibility, and compatibility at the virtual hardware level is much easier than compatibility at the platform level.

The other reason I like IaaS is to deliver multitenancy. In general, the application needs to be rewritten to get multitenancy to it in the cloud. If an application is offered as a SaaS solution and multitenancy is a feature, then you need to preserve that security and other abstractions all the way down through the application stack. It's very, very difficult. On the other hand, it's easy to provide multitenancy down at the hypervisor, network infrastructure, and virtual switching infrastructure level. Then you can have absolutely guaranteed separation from a security perspective, as well as auditability and compliance—without actually doing anything to your apps. You'd get there in an evolutionary way, and quicker, so I like that part.

PwC: Are we still very much in the early days of managing APIs [application programming interfaces] for a cloud virtualization infrastructure?

SC: In the [public] cloud, absolutely we are. In terms of the enterprise, VMware has the enterprise lead in virtualization, and the company has done an extremely good job of turning its administrator into the new hotshot of the data center and delivering a direct benefit to the business. In general, before virtualization, you had a bunch of people managing storage, a bunch of people managing servers, and so on.

With virtualization, the VM becomes a file. The storage person is just a provider of blocks who doesn't do backup or DR [disaster recovery] anymore. It all happens out of virtualization. And the VM administrator is also the person who's delivering the hard savings in terms of consolidation, power savings, et cetera. So this person is the new power elite in the data center.

PwC: As virtualization becomes pervasive, do enterprises create value and differentiation through the management of the virtualized infrastructure?

SC: Yes, though in general, I'm not wild about the term "management." In the context of cloud and agility, I think there's another abstraction that needs to be part of the IT infrastructure system. Once you have virtualization deployed, you need an orchestration layer that allows you to automate the deployment and running of applications. That is a key component of how you build this internal cloud function. I don't think it's management in the traditional sense of FCAPS [fault, configuration, accounting, performance, and security] style or getting alerts. That's still drag and drop with the GUI [graphical user interface]. At the largest end of the scale, it's all about automation, because that's the only way that you get rid of the vertical sets of skills you must traverse every time you want to deploy something.

PwC: What is the current state of automation in the IT infrastructure in the data center? What capabilities lie ahead?

SC: The basic objective of automation is to get to the point where you have an administrative interface that IT can use to do its usual FCAPS-style management. You then expose to lines of business or users, according to their roles, a completely lights-out automated IT infrastructure where they can drive their own applications through their life cycle without requiring anybody in the IT infrastructure to flip switches, install software, or configure anything. That's entirely automated beneath that line.

Where it gets challenging is in two areas: storage and network. The workload used to be statically configurable to the server, but now, with virtualization, the workload can move anywhere and that breaks the networking abstraction. Because the state [of the workload] is required to protect and enforce policies,

the state must be agile too. Moving the workload also breaks the storage abstraction, because the unit of storage, which is the bootable entity of the VM, must be visible to the server that the workload is about to run on. Plus, you need to be able to perform backup, DR, and other activities there. Those layers of orchestration are required.

PwC: What will drive the adoption of cloud computing—greenfield solutions or the migration of existing solutions in the enterprise?

SC: Right now, four categories are relevant. One is surely greenfield solutions. Second is the stuff that you can throw up in the cloud very easily and not care if you lose it. Web servers and HPC [high-performance computing] scale up very nicely. There's no issue around latency. You just want to consume some computing—you want to have easily scalable computing resources, and if somebody compromises a blob of binary data, they wouldn't be able to use it anyway.

There are two other interesting categories. One is essentially DR as a category, so you have a site that can be used to recover if a disaster occurs. The other category is when an application is running locally, but if things go pear shaped [fails], it will be reconstituted in the cloud. Then there's bursting to clouds. Now, you see bursting with HPC, but it's happening already with our portfolio—XenApp and primarily cases that have some predictable workload and that have backup or additional entities up in the cloud, not running. When a flash of demand occurs, a portion of the traffic automatically redirects up into the cloud and fires up additional service for the load.

We see that one as a fairly important category. Arguably, it's an extension of the same core technologies as the DR scenario, which has one instance of a typical workload, but not double the hardware costs, because you can have mission capacity on demand in the cloud. It's just not running, and you don't pay for it when you don't use it.

PwC: All waves of technology adoption have been constrained by the legacy base of applications. What opportunity exists for enterprises to use cloud computing for their legacy workloads? Is there an opportunity to modernize the legacy base?

SC: The rate at which we lag—the tail is so long. When you go to the airport and check into any major airline, every piece of software that you interact with—from the time you walk in the airport until you get on the plane—was written in Windows 3.1, and that was written when I was in high school. That stuff is not moving. It can't move; it's too difficult. So there's going to be this preservation notion of technology, which will adapt to those infrastructures. I think that's a very, very interesting area to play in.

As per modernizing the base, how high up the stack do you have to go? Arguably, that's Citrix's core business. You have things that are legacy apps—strategic, where you cannot deliver them to end users—so instead of figuring out how to rewrite the app or anything else, you basically send a picture of the app from a server to an end user. That's XenApp. You could take advantage of scale and various other things by doing it in the data center—reduced costs, reduced complexity, reduced management, and so on. Centralization reduces the number of things that you have to manage.

PwC: Given all that is unfolding with cloud computing, what should CIOs and their organizations be doing right now?

SC: The biggest barrier right now is lack of knowledge, and so I think people should just get going and do it. There are all sorts of issues. There are issues around licensing, scale, management, and everything else. But what doesn't help is a lack of understanding of those issues and a lack of familiarity. I tell everybody that I meet who has had some budget canceled for buying new servers or not much demand from business: "Just find some stuff to throw out there. Go and use it."

I think there's a real knowledge gap to be filled. And because every vendor wants to be relevant, every vendor stretches the definition of cloud to include themselves. Again, that creates an opportunity for the CIO to analyze and articulate very concrete, very clear, concise benefits in the context of existing infrastructure that value-added providers can bring to customers today. ■

Modeling to automate data center operations

Kirill Sheynkman of Elastra discusses how modeling IT environments in the data center will enable enterprises to migrate legacy workloads to the cloud.

Interview conducted by Vinod Baya and Bo Parker

Kirill Sheynkman is the founder, president, and CEO of Elastra. A database software executive, he spent the early part of his career in the database sales and technology organizations of Oracle and Microsoft. Sheynkman was co-founder, president, and CEO of two successful startups: Stanford Technology Group, which created the world's first relational online analytical processing (OLAP) engine (acquired by Informix/IBM in 1995), and Plumtree Software, which created the world's first corporate portal for heterogeneous information (acquired by BEA Systems in 2005). Between startups, Sheynkman was entrepreneur in residence at Sequoia Capital. He holds a degree in electrical engineering and computer science from Stanford University and an MBA from the Haas School of Business at the University of California, Berkeley.

In this interview, Sheynkman shares how semantic modeling techniques are emerging to bring automation to the data center and agility to IT operations.

PwC: Kirill, can you please tell us about your company, Elastra?

KS: Elastra is about two years old. In a sentence, we are focused on the next generation of data center automation and IT service management.

The way things typically work in IT departments is that application architects draw something on a whiteboard or in Visio or, even more sophisticated, in Rational Rose. They create a design, talk and comment about it, and then hand it off to another group that deploys that design and makes it real. That process tends to be lengthy and error prone. Iterations of that process take a long time, and also there is a disconnect in the thinking or the perspective on how applications work.

For example, application architects think about scalability: Do we want horizontal or vertical scalability for this application? Because we're going to get X demand. Hardware people or data center architects think about what kind of servers they need, the horsepower and flexibility of those servers, and what happens when one goes down. It's a very different perspective. So, at Elastra, we said we could take something that used to be just design and drawings and actually make them into real applications.

To approach the problem, we started with the idea of capturing the data center architecture and infrastructure in a model, a data model. We are using semantic technology to model the data center—not just settings, not just packages, not just computers and networks,

but the life cycles, the processes, the dependencies, the procedures for backing up and restoring and so on. The model also includes policies that are being used geographically, information about where things are located, and so on.

It's an interesting approach, because there have been CMDBs [configuration management databases] and things like that for managing configurations. There are package management systems for managing packages and patches. But when it actually came down to making the thing run, it boiled down to scripts or run books or something like that. It was still this process of code that needed to be versioned, that needed to be run.

PwC: How does having a model-based approach help?

KS: If we could capture [model] everything, then one, we could automate the entire data center operation via a software server, and two, we could do interesting things. For example, if we know the life cycle of deploying a system, we can write an optimizer that says this is the optimal way of deploying this thing, without ever touching the script. We can version it, and we can modify these things on the fly. We can use software to define an allocation policy—how resources are deployed. And we can use software to map a design to an allocation policy depending on the situation.

Here's another way to look at our opportunity: HR has an automation system, the sales force has an automation system, and finance has always had one. Supply chain management has had an automation system. IT has not. There are bits and pieces, but there's nothing that holistically looks at IT as a process that creates a product. You have contracts and licenses with suppliers—they bring these components together that need to be composed. It needs to be put through the assembly line in the right order to come out with the result at the end.

PwC: How will the IT organization benefit from your products?

KS: Here's an example. I can give you a design that's a document on a USB [universal serial bus] key chain, and you can run it. It can be very, very complex, and you can run it in many different ways, in many different setups. So, say you're a consulting company and you have a repository of designs and practices that you have implemented, such as how to deploy a scalable data warehouse that involves a particular kind of data. Launch five of them, and look at them in different sizes. Look at how they perform. Everything is created just-in-time.

PwC: How does cloud computing fit with what you are trying to do?

KS: We are focused on automated deployments of systems, using resource pools that are called clouds. There are a couple of characteristics of cloud that creates the fit. One is the notion of disassociating the hardware resource pool from the [application] software requirements. That's the first thing that cloud forces you to do. You no longer think in terms of, "I have these five machines, and we're going to put application A on this machine and B on this machine." I think of them as a pool of processing, a pool of storage, a pool of

"HR has an automation system, the sales force has an automation system, and finance has always had one. Supply chain management has had an automation system. IT has not."

networking resources that I can use to do things. The other characteristic is the fact that these resources are networked by design. You can't have it without a network, so they are connected and so can be orchestrated in a variety of combinations.

The big IT organizations, they look at companies like Google with admiration, not just because of Google's market cap, but because Google has been able to build a very scalable data center, where a huge engineering organization doesn't really care about hardware or specific machines. What do you mean requisition a machine to do a proof of concept? I have this pool of compute, and I have a pool of talent, and I can actually create things on my own schedule.

So cloud is conducive to that kind of thinking. It's abstraction of the machine. Abstraction of the machine is what cloud gave us, so we said, well, the next interesting thing for enterprise IT is the abstraction of the application. Abstraction of application deployment and design—from the enterprise architect who is concerned with policy and governance, down to the application architect who is concerned with what pieces are needed to meet the requirements that were given.

PwC: We are interested in the question of how IT enables business agility. How can IT be more responsive to business changes, and what opportunity do you see in this regard for yourselves?

KS: I'll break it up into two pieces. The first piece is about being more responsive to changes. Anytime you can do things faster and go live faster, there's value to it. We have an ROI [return on investment] calculator on our Web site that looks at costs in terms of time saved. You can tweak the parameters and say, "If I can save 15 percent in my development time versus my testing time on a project, what does this really translate to based on the cash flows that I anticipate from this project? What does this get me in terms of realizing that project sooner?"

For the second part, our opportunity is that IT is seen as a bottleneck, but IT is a bottleneck sometimes for the right reason. The governance role that IT plays and the policies that IT puts in place are very significant. So the question was: How can we make the business agile and let it innovate, while at the same time making sure that IT keeps track of the house?

And that's what governance is all about. It's about delegating responsibility while maintaining control. With early clouds—like Amazon.com's, for example, and others—you see engineers and maybe even business units whip out their credit cards and they start running applications. The cool thing about it is, "Wow, we got it up in two days, and it's running, and it's our wiki, and we're doing all this nice stuff." That's wonderful. You want that kind of reaction, because they're doing it for a reason.

However, IT cringes. Why? Because if the information goes out where it shouldn't go, who's going to support and maintain it? If they start relying on it, how is this going to fit into the overall architecture? So everyone from the CIO down doesn't like that. The alternative, we think, is not to ban that notion that you can get things done on demand. The idea is to merge policy and control with the flexibility of an on-demand offering, public or private, and a cloud format for your data center—and an application design process that's meshed with the two.

PwC: Can you provide an example of how you are working with customers now and solving these problems?

KS: I'll give you an example of a large software company in the San Francisco Bay Area. They're moving an entire data center from an old one to a new one that they built using the latest and greatest technologies. They need to move a significant chunk of the operations of one of their major business units to the new data center. We need to put it in so that it can evolve from the design of the old data center to the new one and we can speed up the iterations.

How do you do that? How do you move thousands of machines and systems coherently? They're building a model of how things interrelate to each other, what the process of this movement is going to be, and they're automatically deploying it in the new data center. What do we bring to the table for them? It's that they can actually get it done in time using our solutions. The

model is being used to drive the migration. Think of the model of the old as a design, as a drawing. Now run this drawing on the new data center and take advantage of the capabilities that data center has.

"We have a matchmaking algorithm that says, 'These are the requirements for these software components, and this is what this cloud or this set of hardware resources provides.'"

PwC: Can you tell us how the model works? What are some of the key characteristics?

KS: We have a matchmaking algorithm that says, "These are the requirements for these software components, and this is what this cloud or this set of hardware resources provides." So it's requirements and capabilities.

For example, say I want to have a highly scaled or redundant database—one that really scales. Let's say Oracle RAC [Real Application Clusters] fits the bill. Now, where do I run it? Getting into technical detail, Oracle RAC needs three NICs [network interface cards] to function. That's what you have to have on the machine. Well, Amazon.com, for example, doesn't let you do that yet. You cannot run it there, so you find a different solution to the same problem. Your search path from the model lets you run it on a VMware virtualized data center, which lets you have three NICs, and then you need to provision them and hook them up together. Say Amazon.com gets that capability tomorrow—well, you can use either one then.

PwC: What innovation did you have to do to develop the models needed? What are some of the key characteristics of the models?

KS: We started with a notion of building an ontology for the whole system. That was at the very beginning. It's a very difficult process, because there are so many things to track and capture and think about. We went with an ontology-based approach because it enables different viewpoints and different sets of knowledge to be readily incorporated. It would be very difficult if we said, "We're going to have an XML [Extensible Markup Language] schema that describes the data center." That would be extremely difficult, and it would break down at the first customer site.

The model needs to be dynamic. It needs to be distributed. It needs to have rule bases and to acknowledge that the way I see things is not the way you see things, and the way this customer deals with information is not the way another customer deals with information. So we started with a notion of accommodating such differences. As we have been building the product and talking to prospects and to customers, we augment the model, and these changes are easy to do because of the technologies that we've chosen.

We also went with a notion of loose schema, and that's what you need to do the reasoning and the processing. I have spent my entire life in the relational database world, so it was not an easy thing. But our chief architect and I, we said, "OK, this really does make sense for this particular problem." And perhaps that's one of the reasons why people haven't been able to tackle it before.

I think a lot of innovation will take place in terms of how these network systems talk to each other and how they communicate, how they exchange their capabilities, and how they share a workload between themselves. We're scratching the tip of the iceberg. Automating the processes—right now we basically have three allocation mechanisms, and we have a couple of rudimentary planning engines that do that. But to actually use real inference rules and so on to automate the IT process further, that's very interesting. As in automating manufacturing, your tolerances become better, your error rates become smaller. That's where the innovation takes place. That's what we would expect with IT processes. ■

A CIO's road map to adopting Evergreen IT

There's no shortcut to make your data center and IT infrastructure legacy free and agile, but you can map a sensible route to that goal.

Your IT organization is treading water just to keep everything running. Resources have been cut, but the business is eager for new capabilities that might provide a critical edge in tough times. So the last thing a CIO wants to hear right now is that you should be planning a transformation of your IT infrastructure.

But now is the time to begin moving to what PricewaterhouseCoopers (PwC) calls Evergreen IT. Evergreen IT is a pattern of IT provisioning, architecture, and operational management designed to deliver loose coupling between distinct layers of the IT stack. Evergreen IT also is a systematic transition of IT operations from predominantly manual to predominantly automated using intelligent software. In this approach, incremental investment does not create an ongoing burden of legacy systems.

Adopting the Evergreen IT approach prevents the accumulation of a mass of difficult-to-maintain legacy systems. Evergreen IT becomes a simpler, flexible platform on which to deliver IT's core value to the business using fewer resources.

Of course, you've heard this promise before. Indeed, you've probably already embarked on a few transformation or rationalization efforts in the past. That effort simply reset the baseline onto which new layers of legacy barnacles grew, right?

However, emerging technologies under the umbrella of cloud computing are new factors that promise Evergreen IT won't become the platform for new legacy systems. They provide the basis for Evergreen IT that breaks from the legacy traits of the past and positions IT for the future. (See the article, "Enabling Evergreen IT," on page 24.)

None of these emerging technologies will lead you to Evergreen IT by themselves. The CIO must harness them to achieve the vision. Several CIOs have begun this journey and can demonstrate it's a path worth taking. In the first article in this issue of the *Technology Forecast*, we profile Bechtel, but this company is not alone: Companies as diverse as Amerisure, an insurance firm, and Revlon, a cosmetics maker, are following similar approaches.

Preparing for a move to Evergreen IT

No matter how beneficial the Evergreen IT concept is, it's a rare CIO who would rebuild the IT organization in that mold without business buy-in from the CEO, the CFO, and the leaders of the key business units. It's mostly an education process. Top executives and business unit leadership need to understand that Evergreen IT means that they no longer need to be concerned about the specific hardware platforms their applications run on. In addition, Evergreen IT may have upfront costs you can't hide in routine project and refresh budgets.

To gain C-suite buy-in to a fundamental change in technology strategy, the CIO must focus on the promise of increased business agility and lower cost. IT is pervasive in all business functions, and as a result, business agility depends on IT flexibility. IT cannot be

"Governance is going to be key in this transition. If the governance is right, the benefits will be realized." —Erich Clementi, IBM

flexible if it continually ties itself in legacy knots to lash the pieces together and keep them running.

Business units must understand that the status quo is not sustainable, and they will lose in the long run by not having this flexibility. Their largest concern is likely to be sharing common infrastructure instead of having separate systems presumably insulated from whatever else might occur. But a shared infrastructure increases IT reliability, as it is better designed for failover, flexibility, and scale. Plus, virtualization allows the continued use of any specialty applications. Finally, the move to a configure-to-order approach should meet business needs faster than the old-style build-to-order approach. CIOs will be able to provide a convincing argument through actual results.

It is important to develop a good, detailed design of the goal while planning the transformation. At Bechtel, CIO Geir Ramleth worked backward from the goal. "We designed the desired end stage and built it, and we did no modifications. That meant we built new data centers, we built new networks, we built everything new, and—very important—nothing was allowed to come from the old state to the new one unless it passed a certification process."

The C-suite, the business units, and the IT organization will all see risks in Evergreen IT. Isn't this just another promise of reliability and responsiveness from IT that CIOs have struggled to deliver on before? Yes, if CIOs frame this entirely as a problem created by IT alone, and therefore one that needs to be solved by IT alone. CIOs need business leadership and IT to jointly own the problem of inexorable increases in complexity and cost. Everyone needs to understand how their past decisions are the root causes of poorly performing legacy IT. And through that understanding will come the necessary engagement for pursuing the Evergreen IT vision.

Of course, the CIO may want to consider first revamping systems under his or her direct control in the Evergreen IT style. Then the CIO can start building proofs of concept that can help convince early business adopters to take the risk.

Changes to the role of the CIO

Evergreen IT changes the roles of the CIO and the IT organization as well. The CIO becomes more strategic and visionary as mundane tasks are automated. He or she becomes much more of an integrator of internal and external services, making sure integration happens in the right way. "This is a very interesting and powerful role, but it needs to be actively managed," says Erich Clementi, general manager of Enterprise Initiatives at IBM. "Governance, in other words, is going to be key in this transition. If the governance is right, the benefits will be realized." In other words, Evergreen IT is emerging in the context of business units seeing more options for sourcing IT support from external service providers. Lack of coordination between the CIO and the business will re-create a complexity crisis caused by nonstandard point solutions, inconsistent processes, and inconstent data.

Successful implementation of the Evergreen IT approach requires a change in thinking and behavior for the entire organization. "For the IT function, the opportunities and the significant impacts are about shifting your culture from a build-to-order world, where everything is done in a bespoke model, to a configure-to-order world," says Russ Daniels, CTO of EDS, an HP company.

IT staff members' roles will need to change also to support new areas of expertise—virtualization, orchestration, modeling, design for configurability, and so on—and to work within new sourcing, deployment, and planning styles. Skill sets must change to favor the areas of expertise over the break/fix and manual maintenance skills that most IT organizations must disproportionately favor today.

IT staff must truly understand the business and work with it at more than a fulfill-the-requirements level. The CIO must redistribute the IT skill set away from the lower-level operations focus that most IT organizations have and instead to the upper layers of the technology stack that are closer to the business and aligned with direct value. “The IT department will be less about running the infrastructure and more about increasing the value of the platform for the business,” says Doug Hauger, Microsoft’s general manager of Business Strategy for Cloud Infrastructure Services.

The path to Evergreen IT

PwC has developed a road map to implement Evergreen IT. (See Table 1.) Each CIO should adapt the five stages of Evergreen IT adoption to his or her environment. Note that the stages do not have to be sequential, but in total they should apply to most organizations. The road map begins with efforts that you likely have already undertaken, and shows how they can become the basis for Evergreen IT. Where the road map ends is uncertain. What is clear is that the road to Evergreen IT is enabled by many of the emerging technologies under the umbrella of cloud computing. (See Figure 1.)

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Title	Rationalization	Modernizing infrastructure	Formalizing and modeling IT operations	Migration of application workloads	Reaching out to external resources
Key tasks	Consolidation Elimination of redundancies Elimination of heterogeneity	Virtualization of servers, storage, networking Standardization of infrastructure components	Making dependencies between applications and infrastructure explicit Formalization and modeling of key operations	Identifying workloads that should be migrated Certification to run in modernized infrastructure New security paradigm	Integrating internal and external resources Governance
Key benefits	Reduction in complexity Lower operations costs	Configurable provisioning Scalable resources on demand Infrastructure independent of applications	Automation and industrialization of key operations	Independence from infrastructure	Optimized use of resources based on business policies
Impact on agility	Limited impact	Quick provisioning of IT resources and services Self-service, on-demand response	Model-driven changes, quicker deployment of applications in new infrastructure environments	Changes deployed quickly Quicker scalability	Greater flexibility in resource usage Financial flexibility

Table 1: **The road map to Evergreen IT**

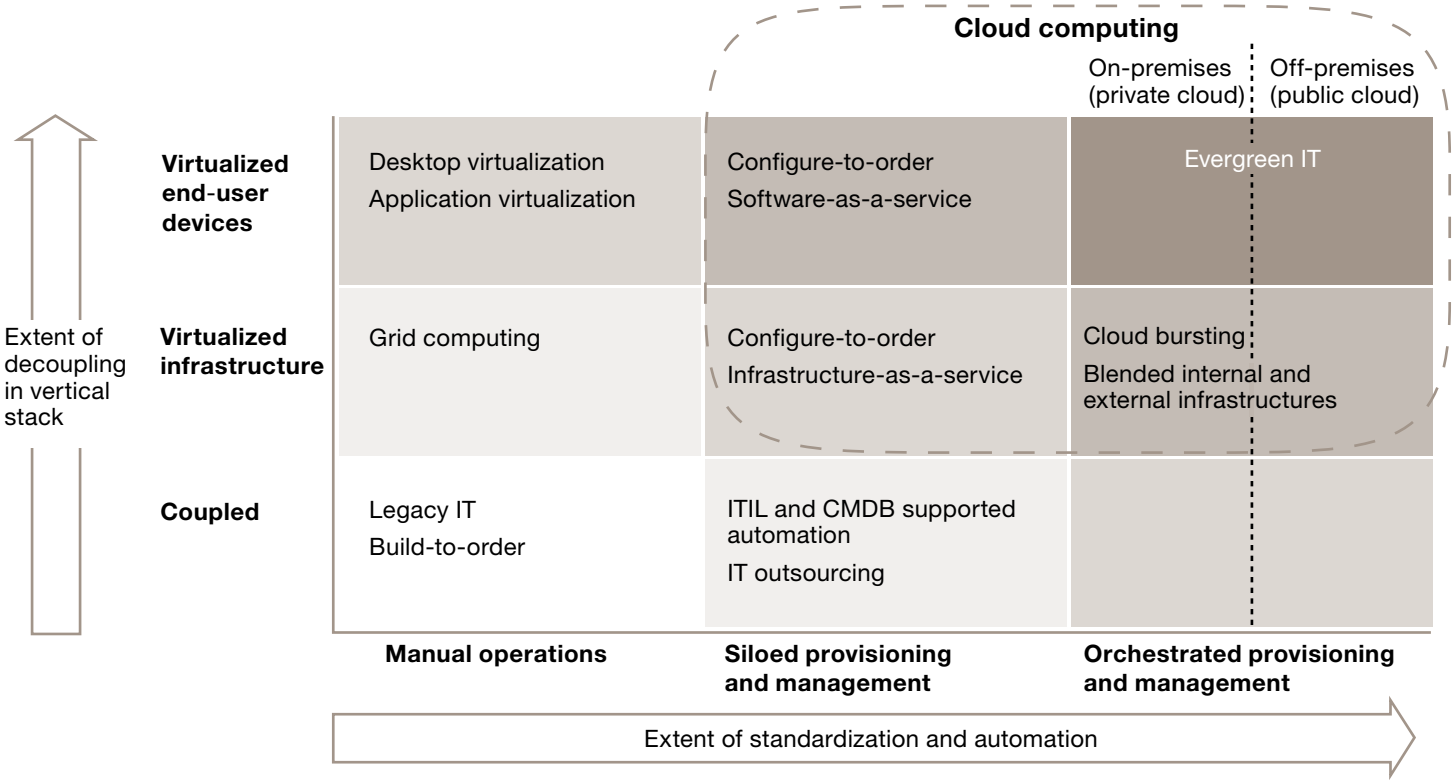


Figure 1: **Many of the emerging technologies that enable Evergreen IT come from cloud computing trends**

Stage 1: Clean up the mess

Most organizations have an unwieldy collection of technologies accumulated from mergers, departmental purchases, project-specific efforts, and so on. Even organizations that took a disciplined approach to mergers and acquisitions by consolidating on the same technology have a mishmash of systems—the result of bringing in technology for specific needs. Over time, the effort to keep systems running consumes increasingly more resources,[1] making the organization less flexible and less rich. You need to clean up the mess by removing the value-destroying complexity.

The cleanup should cover all IT systems, from infrastructure through applications. Given the enormous number of legacy systems, there's no way to do this all at once. But you can make the cleanup part of the requirements of your normal refresh cycles for hardware and software. Use these refresh cycles as opportunities to standardize servers, operating systems, and applications (and multiple versions of the same apps). You can also use new IT projects as a catalyst for rationalization, such as introducing a standard database system or standard Web applications platform to which you migrate the rest of operations over a defined period. You can use efficiency efforts

1. "Strategies for Dealing with IT Complexity," *CIO*, Nov. 27, 2007, http://www.cio.com/article/158356/Strategies_for_Dealing_With_IT_Complexity.

such as virtualization to increase utilization and to lower various data center costs ranging from energy consumption to management overhead.

Most IT organizations know how to do this, whether they call it rationalization, alignment, portfolio management, complexity reduction, or something else. The difference between rationalization in the typical context and rationalization in the Evergreen IT context is that the goal is a platform that doesn't become another basis for legacy-making. It needs to have several different characteristics to become evergreen, which is the focus in Stage 2.

Stage 2: Start with the infrastructure

The fundamental benefit of Evergreen IT is that you don't create new legacy to manage. Legacy means out-of-control complexity caused by custom and nonstandard IT that results in the many exceptions that drain IT organizations. The more variations, the more exceptions that will occur. The incentive to standardize is strong, and standardization should start with the infrastructure.

Decoupling infrastructure from applications provides the flexibility to make changes to one independently of the others. Standardized infrastructure will lend itself to greater automation in provisioning, management, and orchestration.

There is a tradeoff for simplification. The CIO must learn to rely on just a handful of providers, limiting pricing power and losing some control over technology direction. By standardizing on specific models of products, however, you gain the ability to hot-swap failed components and add new ones without tackling legacy-style integration issues.

When choosing the hardware, you should take into account the trends that Evergreen IT exploits. For example, storage and server processing are becoming decoupled: Servers are no longer an integrated environment for application processing capability, storage capacity, and an operating system. The trend is storage-less servers that run a single virtualization technology that lets applications run in virtual machines with whatever operating systems they need, accessing whatever storage they need from a virtualized pool.

By moving storage outside the server box, storage loses its ultra high-speed connection on the internal server bus and travels over a shared network. This shift could create a huge I/O bottleneck, which is why you need to consider 10 Gigabit Ethernet and perhaps even Terabit Ethernet for your networks, so you can run application workloads across disconnected but aggregated-on-the-fly storage and server pools. To ease I/O issues, vendors now package separate servers and storage as "pods" to create essentially a subnet within a pod. Nevertheless, as your data center grows or gets distributed across global operations and providers, network issues will become a gating factor in data center design.

Standardization in infrastructure is not only about servers, storage, or networking components and technologies. It extends to the whole system. Consider Revlon's recent change-out of five data centers with the same standard "data center in a box" shipped to each in a container.[2] Each data center has a consistent platform for computers, storage, and networking; uses virtualization to run applications and to let local IT launch additional virtual servers as needed; and links them all via a secure wide area network (WAN) connection. In this way, Revlon ensures that any data center could take over for any other almost instantly, and that any application that ran in one would run in any other. Revlon needed to cross-train formerly siloed IT staff, but the new data centers required a smaller staff. So Revlon redeployed some server administrators to support roles, where they could help the business staff better use the company's technology.

2. "Revlon creates a global IT network using 'Mini Me' datacenters," *InfoWorld*, June 1, 2009, http://www.infoworld.com/t/it-management/revlon-creates-global-it-network-using-mini-me-datacenters-145.

"For the IT function, the opportunities and the significant impacts are about shifting your culture from a build-to-order world, where everything is done in a bespoke model, to a configure-to-order world." —Russ Daniels, EDS, an HP Company

Bechtel took a similar approach, building data centers in three locations with the same designs and equipment in each. Remote offices that couldn't get appropriate connections to the three data centers received small subset facilities based on the same standards. "We have virtually one data center, and only one operational group to manage it," says Ramleth.

Stage 3: Create a model for automated management

Evergreen IT relies on the automation of key operations such as the provisioning, management, and orchestration of IT resources. Automation cannot proceed without formalization of the many processes and making the interdependencies among resources explicit. As explained in the article, "Enabling Evergreen IT," on page 24, emerging methods and technologies for modeling data center operations show great promise for the industrialization of data center operations.

CIOs should work to rationalize and then simplify their IT management processes. The adoption of IT management processes such as those in the IT Infrastructure Library (ITIL) can provide the framework for intelligent rationalization, so long as such standards are viewed as true process frameworks and not as workflow checklists that concretize a convoluted process.

The process ideas behind ITIL and other frameworks suggest that CIOs can have a model of their IT operations. And having a model suggests that the operations can be managed and orchestrated like any other functions controlled by applications. Although plenty of IT automation tools are available today, they typically are designed to manage a closed environment—you have lots of them for many purposes. They suffer from the same legacy integration issues as everything else. The challenge that many organizations have when implementing configuration management database (CMDB) systems shows that a spaghetti-like environment will defy even sophisticated tools' ability to manage it.

"There have been CMDBs and things like that for managing configurations. There are also package management systems for managing packages and patches. But when it actually came down to making [an IT process] run, it boiled down to scripts or run books or something like that. It was still this process of code that needed to be versioned, that needed to be run," explains Kirill Sheynkman, CEO of Elastra.

Modeling tools, such as those from the startups DataSynapse, Elastra, and Enigmatic, show how IT operations management may evolve. If the IT infrastructure is rationalized, simplified, and standardized, it should be easier to model and thus manage through software. Although these technologies are in early stages, gaining familiarity with them will let your organization move quickly to gain their benefits when they're mature.

Standardization in infrastructure is not only about servers, storage, or networking components and technologies. It extends to the whole system.

Stage 4: Begin to evergreen your applications

Large enterprises have mature collections of legacy applications developed in the build-to-order model. Early experiences (such as those at Bechtel) suggest that some application workloads are more amenable to an Evergreen IT infrastructure than others. The reasons are not just technical, but economical. "When you analyze the cost of delivery, you need to do it for a specific service or workload. You will discover that the knee in the cost curve occurs at different points for different services," says IBM's Clementi.

Cloud technology providers are developing solutions to allow the migration of legacy workloads to cloud environments (internal or external) with no or few changes. Table 2 offers guidelines for positioning representative workloads today and a time frame for migrating them to an Evergreen IT environment.

Application workload types	Technology characteristics	Examples	Potential time frame
High-throughput enterprisewide transaction capture	Interactive, monolithic software, strong referential integrity, sensitive to response time, complex security/authority models	Enterprise resource planning (ERP), operational vertical applications	Long term (5+ years)
Data intensive, compute intensive	Batch or interactive, large memory footprint, large data throughput requirements	Data warehousing, business intelligence, enterprise search, business dashboards	Medium term (3–5 years)
Communications centric, groupware, departmental apps	Distributed, WAN dependent, independent software vendor (ISV) sourced, locally coded, interactive	E-mail, instant messaging, social networking, departmental applications	Short term (1–3 years)
Internal Web applications, Web components of enterprise apps	Modular, bursty, multimedia, frequent code changes, decentralized, homegrown, complex security and authority models	Internal portals, internal Web sites	Short term (1–3 years)
External Web	Modular, bursty, multimedia, frequent code changes, centralized, homegrown	External Web	Short term (1–3 years)

Table 2: **Guidelines on when certain workload types will be amenable to deployment in cloud environments**

A conversation with Eran Feigenbaum of Google on security issues in cloud services

Eran Feigenbaum is director of security for Google Apps at Google. In this role, he defines and implements security strategy for Google's enterprise products. He formerly was the chief security officer for PricewaterhouseCoopers.

PwC: What are some of the key security issues with cloud services?

EF: When moving to the cloud, enterprises need to understand where their data is. It's a little bit of a misnomer that the data is in the cloud. Data is not typically floating randomly through the Internet. Rather, it's in the cloud service provider's data center and under that provider's watch. It's critical that a company understands the controls its cloud provider has in place. It's still the enterprise's data, and if something bad happens to that data, the enterprise is still going to have to answer to its customers.

PwC: What benefits do cloud services provide from a security perspective?

EF: One of the things I spent a lot of time and effort on as a chief security officer was patch management, making sure that all applications and operating systems are patched and up-to-date. Software vendors release security patches on a regular basis. It was my responsibility to understand what those patches were and whether they were applicable to our environment, make sure they didn't break anything, and then deploy them on all relevant systems more quickly than the bad guys trying to reverse-engineer the patches. Once there's a patch, everybody knows there's a vulnerability in those systems. It becomes an arms race. What's faster? Can I as a company deploy the appropriate patches, or do the bad guys figure out what the vulnerability is and break into the system?

Now, with Google Apps specifically, we have a very homogeneous or standardized environment. All of our servers look alike, and we control the entire stack. The operating system, the applications, and the user repository are all written by Google for Google. It's very homogeneous. So when it is time to patch, I can do so in a uniform manner across all of the servers—unlike in the traditional environment where I'm trying to understand whether the patch is relevant to me and whether it works on this or that server. That's a tremendous advantage.

PwC: Cloud architecture aggregates users and data, and therefore any security mishaps have the potential to affect a large number of users and their data. How are such risks being mitigated?

EF: We maintain multiple copies of your data—multiple copies within a single data center and multiple copies within a secondary data center—to help prevent against a problem like that. We fully expect a drive to fail, or to not be recoverable, but we account for that possibility from the software level on up by having the multiple copies and no single point of failure.

The other part is the unique way in which we've used cloud technology. Other people may use it the same way as well. In the traditional environment, you would have a mail server that would be dedicated to you. That would always be the server that you spoke to, and that server housed all your mail. We've taken a different approach. We've fragmented all of your data, and we've spread it across our infrastructure in a system designed such that your data cannot become compromised by gaining access to a single location. The data is spread across our many servers. It's like a needle that's been chopped up into small pieces and put in a haystack, impossible to find.

PwC: What can enterprises and government organizations that have the scale to build a private cloud learn from the experience of Google in offering and securing cloud services?

EF: First, there are many lessons that we all learned from the years of computing that existed before us. Lessons from all the security drawbacks that exist in the client-server model—the patching process being just one.

The next lesson is around the use of standardized or homogeneous infrastructure, and the ease of maintenance and support that follows. A lesson can also be drawn from some high-profile losses of data on portable devices, USB sticks, laptops, CDs, and DVDs. Users are working from home on weekends and so on. To do that, what are they doing? They're putting data on USB sticks. They're mailing it to their personal Hotmail, Yahoo, or Gmail accounts. As soon as they do that, they've completely broken their company's security model and taken the data out of that company's control. Putting it in the cloud and making it available anytime, anywhere, while it's still maintained in the security of the cloud, presents a tremendous advantage.

PwC: What can the information security officers at an enterprise get from the cloud today that they are not getting in their own environments?

EF: Let me give you an example. Depending on which statistic you believe, 70 to 95 percent of e-mail is spam, right? With spam comes viruses. A cloud provider like Google processes billions of SMTP [Simple Mail Transfer Protocol] transactions every day, bringing a tremendous amount of knowledge—knowledge that a single CISO [chief information security officer] of a single company can't get. So with this knowledge, the cloud has visibility. We can block spam, botnet attacks, and viruses in the cloud. Because of all the traffic that we're processing, we can block our customers from viruses a couple of hours before the antivirus vendors have even seen those viruses.

PwC: Surveys have shown that security remains a primary concern with cloud services. What message would you like to send to the CISOs who might be exploring private or public cloud solutions?

EF: The message to them is don't dismiss the cloud because of the hype. The fact that the FUD [fear, uncertainty, and doubt] factor is security, is mainly FUD. Do your research and understand the security model of your cloud provider. Not all clouds are created equal. From a security perspective, look and weigh the advantages and disadvantages of a cloud provider against those of in-house solutions today.

I speak every day to CIOs and CISOs of Fortune 1000 organizations, and I think many of them come to the conclusion that the cloud often can be more secure than their existing operations. It's very much the paradigm. Imagine if we were in the banking industry 100 years ago, where you had to decide to take the money from underneath your mattress and put it in the bank. If you chose the mattress, you could go home and look at it every day and know that it was still there and still safe. But the bank really had the economies of scale and could afford the armed guards and the big industrial safes that each individual user couldn't. That's the same kind of world we're in now.

PwC: Looking forward five to ten years, as the cloud unfolds, public or private, how different do you think the conversation around security will be?

EF: I hope that a lot more CISOs and CIOs use these new concepts of the cloud that take care of a lot of the security issues that we spend so much time and money on and still do wrong today—such as passwords and portable media. There are a lot of basic applications that we don't do very differently. The way you do e-mail and the way I do e-mail are pretty similar, so let's enjoy those economies in scale.

In the next five years, I'd also like to see CIOs and CISOs figuring out what services they can move to the cloud or to other service providers, and focus their efforts on things that are unique to their businesses that can really drive a competitive advantage. ■

The road to Evergreen IT will take years, but its benefits begin immediately and grow over time. Most important is the shift in how Evergreen IT requires CIOs and other IT pros to think about their work.

One thing you can do today is to favor configurable applications in what you buy and develop. Such applications do not need to be customized for specific projects or users; instead, they have capabilities for a broader range of uses and can be configured through setting preferences in a functionality and user interface (UI) database system. Thus, there is only one code base now and in the future. This is the model of software-as-a-service (SaaS) applications like those from salesforce.com, but it is also the model of many packaged applications.

Irving Wladawsky-Berger, chairman emeritus of the IBM Academy of Technology, uses the term "mass customization" where PwC says "configuration." He explains: "Mass customization is sort of a new concept in the IT world. It's not a new concept to the telcos or cable. In fact, I would even say retail companies such as Amazon.com and Netflix are all about mass customization. Notice everything Amazon.com does is mass customized, and it's self-provisioned, but it's self-provisioned within the boundaries that they give you." That's the goal CIOs should have for the applications they run.

Migrating applications to Evergreen IT will likely affect the security architecture and processes. In various surveys, CIOs list security issues as their top concern about adopting external clouds. The concern is legitimate, but Evergreen IT can improve security. With a standard environment at the core, fewer exceptions to security policies occur, which significantly reduces the risk of intentional or accidental lapses and breaches.

"We found that you just have to accept a whole new security paradigm. That's really the biggest challenge on the technical side," Bechtel's Ramleth says. "The point is perceived risk. I can prove that the information in our core systems is more secure today than it was in a typical legacy environment because we have tightened up that back end tremendously."

For more on the state of security in clouds and what to expect, see our conversation with Eran Feigenbaum, director of security for Google Apps at Google, on page 52.

Not only does Evergreen IT offer the potential to decouple applications from infrastructure, but it also presents the potential to decouple applications from client devices (PCs, smart phones, and so on). You can use technologies such as Web services and presentation-level front ends such as asynchronous JavaScript and XML (AJAX) and terminal emulation (à la Citrix) to confine the applications to your data center. The client-side decoupling can reduce the effort of PC management, increase security (because data is no longer local), and allow more flexible access (by contractors and employees at home or at a customer site, without regard to their specific client hardware).

Amerisure, for example, uses application and desktop virtualization to reduce complexity and cost, and to support telecommuting from home. The company no longer issues PCs or laptops. Each worker gets a $300 thin-client device. There's no PC for IT to configure and maintain. Desktop virtualization connects the thin clients—as well as employees' home PCs and customers' PCs when they use them—to the back end. This left the IT group with two-thirds fewer physical servers to run, just two server platforms to manage, and all applications under central control. When the IT group moved the remote office servers into the central data center and had no desktop PCs to support, Amerisure was able to end an expensive third-party IT support contract.[3]

3. "Amerisure uses virtualization as a strategic technology, not mere cost-cutter," *InfoWorld,* June 1, 2009, http://www.infoworld.com/t/virtualization/amerisure-uses-virtualization-strategic-technology-not-mere-cost-cutter-399.

Stage 5: Reach out to external clouds

For most enterprises, the journey to Evergreen IT will likely start with the expansion of virtualization software to mediate application workloads so they no longer are tied to specific hardware environments. This virtualization is already happening at many companies. But the IT infrastructure remains a management challenge. The big step toward Evergreen IT is the move to a radically simplified and standardized server/storage infrastructure mediated by virtualization software. The last step is to fully automate the provisioning and management of the infrastructure—a task far easier to accomplish when hardware is highly standardized.

Does this vision include a role for external cloud services? Today, many CIOs are rightfully skeptical of external cloud computing. In the next decade or so, however, a compelling set of reliable, scalable, secure, service-level-guaranteed cloud services may be available. They will likely range from transaction-oriented computing capacity to enterprise-class application services for financial management and reporting, from on-demand storage to composable logic services (such as widgets) that plug into your application architecture without integration work.

Even if not, broad adoption of the principles of Evergreen IT will mean that your business partners, suppliers, and customers will also have agile IT environments that you will want to interact with as part of a private, group, or ad hoc cloud.

Lack of industry standards in external cloud services creates the additional concern of vendor lock-in. Until external cloud service providers offer guarantees of application and data portability, CIOs will rightly be hesitant to port mission-critical applications to them. Today, each external cloud provider has its own infrastructure, so using its on-demand services requires the legacy-style, tightly coupled vertical integration you want to avoid. However, as test beds for gaining experience with such systems, they are worthwhile.

Looking to the future, combining internal and external compute resources will be beneficial, economical, and a source of agility in many scenarios. A common example being explored today is cloud bursting, a way to access external resources to accommodate overflow load during peak times. In time, moving virtual machines from one provider to another will become more possible, in which case dissimilar infrastructures will not matter or compromise the benefits from your own Evergreen IT infrastructure. In the meantime, external clouds remain perfectly acceptable for activities you don't intend to run on your infrastructure anyway.

The future is orchestration

The road to Evergreen IT will take years, but its benefits begin immediately and grow over time. Most important is the shift in how Evergreen IT requires CIOs and other IT pros to think about their work. It's no longer about tinkering and seeing how long you can keep your finger in the dike while colleagues try to bail out the water. It's not about having the forensics chops to figure out how the current system was put together, what caused it to stop working, and how to fix it without breaking something else.

Evergreen IT is about building technology systems with as few dependencies as possible, so maintenance and integration costs decrease and the ability to flex with business needs increases. Building such systems requires more discipline in design and greater modeling capabilities, but the result is an IT infrastructure that can be more easily managed, automated, and adapted.

Although much of the technology that will make Evergreen IT is still in early stages, the starting points—especially virtualization—are in place. You just need to figure out what your evergreen destination is to begin moving toward it.

For more information on the topics discussed in this article, contact Bob Zukis at +1 213 217 3222 or Gerard Verweij at +1 617 530 7015.

Standardization and mass customization of IT

Erich Clementi and Irving Wladawsky-Berger of IBM discuss how cloud computing is a fundamental change in IT that represents a new business model as well as a new computing model.

Interview conducted by Vinod Baya

Erich Clementi was named general manager of Enterprise Initiatives at IBM in October 2008. Prior to that, he was general manager of the Business Systems division in IBM's Systems and Technology Group, where he was responsible for providing small and medium-sized businesses with systems and storage solutions—solutions that later became some of the first pieces of IBM's cloud computing portfolio, called Smart Business. Since the 1990s, Clementi has worked in several other roles at IBM, including general manager of IBM's Managed Business Process Services division, general manager of IBM's System z division, and leadership roles in corporate strategy, marketing, and sales.

Dr. Irving Wladawsky-Berger is chairman emeritus of the IBM Academy of Technology and visiting professor of Engineering Systems at the Massachusetts Institute of Technology (MIT). Prior to retiring from IBM in June 2007, Wladawsky-Berger was responsible for identifying emerging technologies and marketplace developments critical to the future of the IT industry, and organizing appropriate activities in and outside IBM to capitalize on them. In his emeritus role with the IBM Academy, he participates in a number of technical strategy and innovation initiatives. At MIT, he is involved in multidisciplinary research and teaching activities focused on how information technologies are helping transform business organizations and the institutions of society.

In this interview, Clementi and Wladawsky-Berger share their insights on how cloud computing borrows from the principles of the mainframe era, but in doing so, it brings greater flexibility to the IT infrastructure and applications. They also forecast how, with standardization and mass customization, clould computing is changing the interface between the service providers and service consumers.

"Cloud computing is doing to the IT supply chain what Henry Ford did with the conveyor belt. Otherwise you cannot explain the economics." —Erich Clementi

PwC: How much disruption does cloud computing represent when compared with the Internet?

IWB: I think cloud is the name we are giving to a new computing model in IT. That's important, because if you look at the last 50 to 60 years of the IT industry, we have had only two previous computing models. We've had the central model, where the mainframe was the biggest product. We've had the client-server model, where the Windows- and Intel-based computer was the biggest product. And now cloud is emerging as a new computing model.

The central model was based on mainframes and supercomputing built for sharing the machines with multiple applications and users. Acquisition was very much controlled by a central group, since they were spending millions of dollars and were supporting lots of users. The applications were back-office, mission-critical applications. The central group didn't let many applications onto the machine, since one application could bring down everything. They had very strict guidelines about what you're allowed to have in the machine. That was the mainframe model.

In the client-server model, instead of the systems being optimized for sharing, now they were optimized for low cost, because they were built out of PC and workstation technologies. Instead of a central organization, departments did their own acquisitions, which they found appealing because they didn't have to go to a central bureaucracy. Since these systems were built out of microprocessors, you tended to run a dedicated application on them. As a result of that, there was a lot more flexibility in what applications you could run. So there was a lot of innovation in new applications, because it was so much easier to get the applications in.

Now, a different model is emerging. The cloud model is an Internet-based model. It has taken a while to work itself out, but we realized that something started to change even in the late '90s. With the Internet, you had so many more users that all of a sudden the servers got far bigger, people started putting them into central computing data centers, and they had to pay a lot more attention to scalability, systems management, and things like that. It started to change with the Internet.

EC: IBM believes this is a fundamental shift in the way services are consumed and delivered. We have seen computing model shifts before, but they didn't change the IT industry's own business model the way cloud computing is. We sold a centralized solution in the "mainframe era," and then we all sold a decentralized solution when the client-server model became mainstream. This move here is different, because it changes the business model on how IT value is consumed. And that is a very significant shift.

It has similarities in impact to the Internet in terms of deployment. Everybody loved the Internet for content, but there were certainly concerns over security, privacy, control, and processes when it first came on the scene. The CIO was nervous. Similar sentiments are true today with cloud.

PwC: It is often said that many of the characteristics of cloud computing have been with us from the mainframe days. What is new now that cloud computing is gaining traction and interest?

EC: Coupled with the delivery, I'd say the first new element of cloud computing is the self-service nature of it. Think about what ATMs [automated teller machines] did to the finance industry. Initially, ATMs were nothing other than cash dispensing machines, and you could say that they just substituted the human interface with a machine interface. But that is a superficial way of looking at it because the human interface did much more. The teller was a built-in security system because the teller identified you and handed you the money because he or she knew you. The teller also managed the whole process of restocking, of accounting. By substituting that human for a machine, it forced the standardization of the entire process within the organization. Then when customers wanted to get their money from any bank ATM and not just their own bank, messaging and the transactions had to be standardized across the entire industry. Otherwise it couldn't work. So this change standardized the process around cash.

Cloud computing applies the same discipline to IT and changes the way you consume. So far, IT has been something like an art. Cloud computing is doing to the IT supply chain what Henry Ford did with the conveyor belt. Otherwise you cannot explain the economics.

The second new aspect of cloud computing is a different business model. I hesitate to go to variable pricing directly because there are different levels of variability. Many enterprise customers will probably want to have a subscription or some fixed, guaranteed fee and then some variability—but a higher variability for flexibility than before. The third element that is coming into play is the hybrid nature of the sourcing. Until recently, you basically had to have all the technology in your department to consume IT. This is what created the IT industry. Now there will be an increasing number of services that you can consume without actually being in charge of delivering them.

IWB: The rise of mobile computing and sensor-based computing makes this new. By mobile computing I mean BlackBerry and iPhone devices and e-book readers with Amazon.com and others, and now netbooks.

At some level, one of the most tangible differences between central computing and client-server computing was going from text-based terminals used by a relatively small number of people to PCs with GUIs [graphical user interfaces] used by many, many more people. Now these new mobile devices, let alone all the sensor stuff, is orders of magnitude more people, and more things accessing the services.

Also, in the previous two eras, the services were relatively custom. With cloud, the need to support billions of people has put us into the mass customization era, which, by the way, the telcos [telephone companies] know how to do very well. Mass customization is sort of a new concept in the IT world. It's not a new concept to the telcos or cable. In fact, I would even say retail companies such as Amazon.com and Netflix are all about mass customization. Notice everything Amazon.com does is mass customized, and it's self-provisioned, but it's self-provisioned within the boundaries that they give you.

I also think that you need cloud now because the client things that you're supporting are now exploding, if you include sensors, into the trillions. The central mainframe model will be way too expensive for this, and in the client-server model, the system management costs would be too high. So cloud had to be invented to handle this explosion of new things that come into it.

PwC: What existing problem is cloud computing solving that enterprises should pay attention to?

EC: What is ultimately driving cloud computing is a crisis of complexity, driven by the distributed computing model. So why did the distributed computing models supplant to some extent or complement to a big extent the centralized one? The centralized computing model was too rigid. But before enterprises realized it, they

"Reduction in complexity is a huge part of cloud computing. This is a very important point, because what cloud is using is identical to what the telcos have done for their services, which is standardization and mass customization." —Dr. Irving Wladawsky-Berger

had server sprawl, inefficiency, and low utilization, and soon they lost the flexibility of the distributed model. Remember, there wasn't a service problem when it was centralized. There was no inefficiency, because it was highly utilized. But the perceived cheapness of the distributed environment came back to haunt us.

So cloud computing now offers an opportunity to deal with this complexity, and guess how cloud computing is doing it? It is doing it with the same virtues, if you want, that the centralized computing had, but on top, giving flexibility. The virtues are the virtualization, the standardization, the automation, the discipline. So in that sense, from an infrastructure point of view, for sure, cloud computing modernizes the legacy.

IWB: Reduction in complexity is a huge part of cloud computing. This is a very important point, because what cloud is using is identical to what the telcos have done for their services, which is standardization and mass customization. In other words, if you say, "There is a general-purpose computer underneath. I can do whatever I want to do. Why do you tell me I can have only an 18-minute plan or a 120-minute plan? I want a 93-minute plan." It is because the cost of administering a plan that's different from the available options is so high that it's not worth giving it to you. It would be overwhelming and add to the complexity of their operations. They do their segmentation, they give you these plans, and if you don't like it, you're into custom territory but you pay a lot more.

PwC: What will be the impact of the current economic environment on cloud computing, if any?

EC: One indication on how to read this is that the enabling technologies for cloud computing have been in the making for a very long time—virtualization, service orientation, automation, the networking, the bandwidth characteristics, the whole infrastructure. This has been in development for 40 years. People have been talking about putting together something that virtually works like a mainframe for a long time. So this is something that was inevitable in its economic impact. Now you come to a moment in which cost focus is heightened due to the economy, so this will accelerate the demand for these efficiencies and maybe it will force some second look at what level of customization you really need.

IWB: One of the things that makes new models happen is cultural resistance. People will say, "I've always done it this way. Why are you asking me to change?" Now, if I say, "Well, because it's much cheaper," a customer might say, "Well, but— it's not as good as what I did before" or "It won't work for us." But when you have a crisis, all of a sudden those excuses don't count anymore. If you can really get something done that's good enough and cheaper, then people will pay a lot more attention to it. I think the tough economic situation will push companies to be much more disciplined about making these changes.

PwC: How will the adoption of or migration to cloud-based services take place?

EC: First, computing models impose themselves for economic reasons. If the economics are right, it's going to be successful. If it's not, it's not. Second, things take time. When you analyze the cost of delivery, you need

to do it for a specific service or workload. You will discover that the knee in the cost curve occurs at different points for different services, so that is why in the industry everybody says the SMB—the small and medium-sized business that can never get to that knee in the curve—is the absolute first that will go for services delivered through cloud computing. They also happen to be the users that, while needing a full range of services, usually have fewer requirements for customization or configuration.

There is a whole set of services where you can get very, very significant economic benefit from private clouds. Workloads that might be regulated, have security requirements beyond what is available, or that need to undergo significant migration costs won't work well for a public cloud model. But you might be able to reap significant economics in a private cloud. For other workloads that have massive economies of scale, you might go directly outside. Take high-performance computing, for example. High-performance computing has a characteristic that when you really need it, you never have enough power, and most of the time you have an issue with the economics of it, because you simply cannot load the machinery sufficiently. So what better model than going to a public or hybrid cloud? Where the real cost benefits are, where the cost elements come to the asymptotic range, is very different for different workloads. That is our experience.

IWB: I think client-server may have been an interim solution for the more front-end kinds of services. It may be interim because in many ways, cloud is a kind of scaled-out mainframe. If you look at classic mainframes as being very good for scale-up, SMP [symmetric multiprocessing] kinds of applications, what clouds have done is let you consolidate all these highly distributed applications in a more centrally managed, very disciplined environment, but using commodity technology. Since it's scale-out, you don't need to pay the kind of sharing costs that you have to pay with SMPs, and you get much higher utilization of your hardware, much lower system management costs, and far better scalability. So you get things that people associate with mainframes, with central computing, but now applied to this commodity-based scale-up model.

PwC: All enterprises have a significant legacy base of applications. How should enterprises look at the cloud computing opportunity with reference to the base of solutions in which they've already invested?

EC: Take your workloads and put them into three buckets. First are the workloads that for reasons of security, migration costs, maybe regulation, or risk are not the primary target for cloud computing. Second, there are workloads that could benefit both from the technology model and from the potential outsourced delivery model. Take those first. For example, 30 to 50 percent of servers are used for testing development, and these systems notoriously have single-digit utilization rates. You can move that whole workload, which usually has fewer security and viability requirements—they're not so stringent—and reap a host of benefits. Similarly, collaboration workloads can benefit from this kind of scale.

Then there is a third set of workloads that you are not doing today, simply because you never thought you could afford it. But with cloud, these now become possible. For example, take the risk analysis done in financial services. Today it is largely done on a

"There is a whole set of services where you can get very, very significant economic benefit from private clouds. Workloads that might be regulated, have security requirements beyond what is available, or that need to undergo significant migration costs won't work well for a public cloud model." —Erich Clementi

stochastic basis. Why? For two reasons. Either the customer would have to buy such an amount of computing that it wouldn't be economical, or it would take so long that they would realize the risk after the event happened.

But if the customer could get sufficient computing power to do the analysis in time and at a decent price, they would most probably do it. Think about inline analytics. Think about risk management. Think about extreme high-performance computing. None of these usually have stringent requirements in terms of security, compliance, or migration costs. Remember, when I say stringent, regulation, and security, all of this is a statement in time.

PwC: Customer lock-in—whether it's an application or data or other elements of IT—has surfaced as one of the big concerns with clouds. Customer lock-in has also plagued users for a long time, with traditional delivery models. Does cloud represent another round of lock-in of a different nature or a freedom from lock-in?

EC: We cannot allow lock-in for cloud computing. I think we need two things. We need interoperability and we need standards. Every maturing industry evolves around standards. The lock-in in the past was in part driven by the fact that there was nothing else. In a nascent industry, that's one thing. Lock-in could be an inhibitor to its growth.

Customers would want to know how they can get the data out of the cloud and how they would interoperate. The industry needs to know what they find on the other clouds—for example, they can communicate with a cloud, exchange data, collaborate with a cloud, and there are mechanisms to do these. For sure we are going to push for openness and interoperability.

IWB: We will prevent lock-in if we agree on standards. With the Web, there is no lock-in, because everybody agreed to a common set of standards, such as HTTP [HyperText Transport Protocol] and others. With cloud, this is the perfect time to come together around standards. We must have standards, and we may not even know all the places we need standards. That's why we need the industries to work together and identify and start developing them.

PwC: If standards are essential, what is the opportunity for vendors to differentiate?

IWB: Customer satisfaction, customer service, add-on services that they provide. Just because you have standards doesn't mean your total offering needs to be standard. It's perfectly fine that the basic offering is standard, and then if you want to lock people in, it should be because you are adding extra benefits that they wouldn't be able to live without—value-added benefits. I think that will be the battle—the value-added features. But I honestly think the biggest value of all is how well you provide customer service. If you do a really good job, you'll retain your customers. And if you do a poor job, I don't think you will.

PwC: How far along are we in automating IT and data center operations? What challenges lie ahead?

EC: Automation can happen by different means. We can build hardware that is self-healing and autonomic. When that reaches a limit, we can build software systems that are so resilient that they can jump in when the hardware underneath fails. There are a lot of ways to go, but then consolidate it all for high availability. When you have 99.999 availability and you are the user, you don't care how that is reached. Today, the best are at 99.999. We can go to 99.9999. When you measure it in minutes of unavailability per year for the whole system, that should be what ultimately matters.

IWB: I think it's an Alice in Wonderland situation that we are running like crazy just to keep up, because in the meantime the workloads keep exploding. If you look at Smart Planet kinds of workloads, congestion management, infrastructure management, energy efficiency, and things like that, those are enormous workloads. To keep the system management needs

"The discussion would not be that everything becomes a cloud. The discussion would be about picking those services you offer that are growing the fastest that can benefit the most from being standardized and mass customized, and moving them to be more cloudlike."
—Dr. Irving Wladawsky-Berger

from growing with the workloads requires massive innovation, so we are constantly running. And then new workloads will come up.

I think system management, autonomic management—I don't see an end to them. Those are things that will constantly be innovated. We'll keep making the hardware smarter, the software smarter, the management console smarter.

Also, the probability of security problems, hacking problems, reliability, self-configuring—there's a whole set of things that we need to keep adding to make them easier to use, more reliable, and lower cost.

PwC: How will cloud computing change the mission and role of CIOs and their organizations?

IWB: I think cloud computing brings much more discipline to the CIO role and the CIO's organization. The discussion would not be that everything becomes a cloud. The discussion would be about picking those services you offer that are growing the fastest that can benefit the most from being standardized and mass customized, and moving them to be more cloudlike. You can buy cloud systems and software, or you can have a company like IBM run it for you and have a plan to phase that over time. It's possible that a lot of what they do will never move, either because they are more like back-end transaction services that are not cloudlike, or because they are legacy client-server things that are stabilized and the effort to migrate them isn't worth it. So cloud computing imposes a discipline that's very healthy.

To manage this discipline, you will need a very good CIO to decide what to do on-premises and off-premises, to coordinate the vendors, and to make sure the vendors are good. It's like if somebody is manufacturing something for you and they use lead paint, it's your problem. So you need somebody to make sure they don't use lead paint and they do what they promised to do. And if you have IT suppliers, you need people to manage that very carefully and to interface between them and the rest of the business. So the IT function changes.

EC: When CIOs managed the centralized computing, distributed computing came in through the departments. Why? Because centralized computing was not responsive enough for everyone and departments shifted to a more flexible model. Cloud presents the same challenge and there is a possibility that departments will take matters into their own hands if flexibile, secure, and "sanctioned" options are not provided. CIOs can do two things. CIOs can ignore the growing interest in cloud computing, and that is not a good idea because the economics tend to be overwhelming and if you have overwhelming economics, usually sooner or later it's going to happen. Or they can evolve into much more of an integrator of internal and external services, making sure that the integration happens in the right way. This is a very interesting and powerful role, but it needs to be actively managed.

Governance, in other words, is going to be key in this transition. If the governance is right, the benefits will be realized. If the governance is wrong, they will incur hidden costs of integration. Customers love the idea

that they can swipe their credit card to provision some systems on Amazon.com, but they are telling us that the acquisition of these virtual machines needs to go through the right internal processes, to be auditable, and to be able to be monitored. If CIOs want the benefit, they need to integrate cloud activities with IT governance. Otherwise it will be a problem.

We see the biggest demand in helping customers adopt the model while still fulfilling enterprise demands. Individual developers can swipe their card and develop a little application for the iPhone, but if an IT project inside of a corporation deals with its IP [intellectual property], has a timeline, and has a group process, then you need a progress report, you need to have ITIL [IT Infrastructure Library]-certified processes, and you need to have a SAS 70–certified provider. Enterprise computing needs enterprise computing standards.

PwC: As your vision of cloud computing unfolds, how will your customers be different? How do businesses change as a result of cloud computing trends?

EC: They will look like services customers, so in many cases they will contract the outcome versus the how to do it. Second, instead of talking to me about lock-ins, they will talk to me about customer satisfaction and service level agreements. That is at least as big a transition for the company that delivers the service as it is for the customer. And we should not be surprised, because this is a maturing industry. In a maturing industry, the differentiation moves to the interface with the customer and the interaction is on results—not on how to do it.

IWB: I think businesses are spending way too much money and time on things that don't bring them value, but there are things that they need to do, such as managing technology environments. As a result, they don't have enough time to focus on the things that do bring them value. I think that's why cloud computing will grow. Eventually everybody must move in this direction, because if more and more businesses are able to make the shift, then in principle they should do better than their competitors who are still wasting their time managing technologies. Eventually, everybody has to do it this way. ■

"They [CIO] can evolve into much more of an integrator of internal and external services . . . This is a very interesting and powerful role, but it needs to be actively managed." —Erich Clementi

From build-to-order IT to configure-to-order IT

Russ Daniels of EDS, an HP company, discusses how cloud computing is creating opportunity in industry ecosystem interactions.

Interview conducted by Vinod Baya, Bo Parker, and Victoria Huff

Russ Daniels is CTO of EDS, an HP company. In this role, announced in June 2009, he continues the momentum he built as vice president and Cloud Services CTO at HP, driving the cloud technology strategy and architecture for HP's outsourcing services portfolio. Daniels has more than 25 years of experience in the technology industry, specializing in software architecture, enterprise management, and software development methodologies. He has filled a wide range of staff and line management roles and run his own Internet services business. From 2002 to 2007, Daniels was CTO of Software for the Technology Solutions Group at HP. During his tenure, the business tripled in revenue and emerged as a significant player in the software industry. In 2006, *InfoWorld* named Daniels one of the industry's top 25 CTOs.

In this interview, Daniels shares insights on how IT will create agility by moving from a predominantly build-to-order model to a configure-to-order model. He also discusses how cloud computing represents an opportunity to create new value at the level of industry ecosystems.

PwC: What level of disruption does cloud computing represent? Does it represent the same level as the Internet, something smaller, or even something bigger?

RD: We think two things are going on. The first primarily concerns what a year ago we would have called utility computing rather than cloud computing: the fundamental questions of how do you improve the efficiencies and the agility of the IT function as it sources and delivers the services the business needs. It's a very important agenda, but it isn't new. It provides some new opportunities, some new choices that businesses can make about how they source and deliver those services. We think calling that cloud, though, is confusing, because it obscures or takes away the ability to talk about the second thing that's going on that we think has more long-term and transformational implications.

A 2x2 diagram is useful to show how we see things. [See Figure 1.] The left column describes on-premises, which is to say the compute capacity, the data center, the servers, the storage, the network capacity, the software licenses, and the operational responsibilities live within the enterprise, within the consumer—the consumer and producer are the same. Off-premises means some third party takes over more of those production responsibilities.

The bottom row is labeled "Dedicated," and the one above it is labeled "Shared." The idea is that in the "dedicated" row, the resources that are being applied are bound to specific application workloads, and making changes to those bindings requires manual intervention. In the "shared" row, there's some mechanism that allows you to bind those resources more flexibly and to adjust those bindings dynamically.

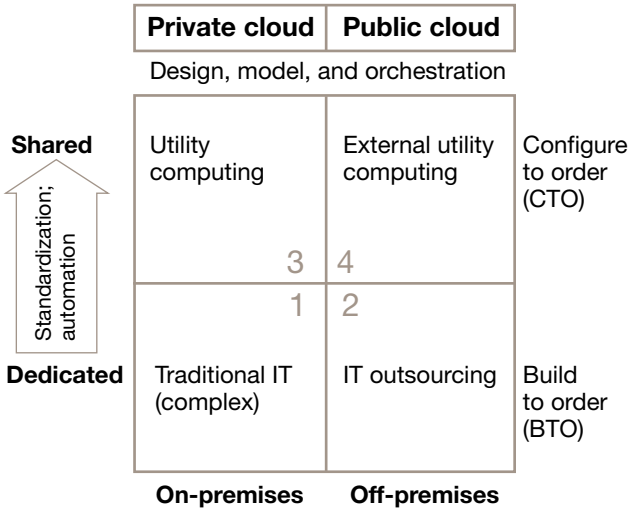


Figure 1: **HP's view of IT transformation and cloud computing**

PwC: What do each of the quadrants represent in this framing?

RD: If I walk you through the quadrants, the quadrant labeled 1 is the traditional way we built things in the past, which is the customer buys all the assets, constructs a complex execution container for a particular workload, and then effectively pours that container into cement—it's very difficult to change. That approach is really good for some things. It's great if you have concerns around performance, for example, because you can optimize everything in that environment for performance, availability, and throughput. So, for things that are mission critical—which is to say something such that when it goes down, your business can't book revenue, can't produce products, services, et cetera—you might really like this model, because it provides you with the ultimate control.

If you need some of those characteristics, but you don't have the competencies or don't want to maintain the competencies to do that, you might move to the right to quadrant 2, where somebody else does some of that stuff for you.

For most of our customers, when they think of the cloud, they're thinking of moving from quadrant 1 to 3, where they typically take advantage of some combination of virtualization and automation technologies. They can pool the resources and then dynamically bind them to workloads and adjust those bindings as business need dictates.

The upper right quadrant, quadrant 4, is where you take advantage of shared capabilities off-premises, so some third party offers the services that the business is consuming. Here you get the advantages of scaled multitenancy and potentially specialization to be able to do particular things at a cost efficiency that makes it attractive.

PwC: How does the framework in the diagram relate to the responsiveness of IT to changes?

RD: The best way to answer that is to change these labels from dedicated and shared, and say what's really going on in the context of this frame. If you use the manufacturing metaphor, the bottom world is the world of build to order [BTO]. Build to order is distinct in that design is an integral aspect of delivery. It's custom. It's bespoke tailoring. You want a suit? No problem. Here are some bolts of cloth. Which ones do you like? What felt would you like under the collar? What buttons would you like? We will create exactly the suit you want. It's going to be a little bit more expensive than if you go and buy it off the shelf. The delivery might be six weeks from now rather than later today—up to you.

The world in the top row is the world of configure to order [CTO], and configure to order is a world that—again, in the tailoring world, it means you buy off the rack. There's a set of designs. You might be able to get a little bit of alteration done, but if you're a size 44 jacket and a size 42 pants, then you're going to have to find the right shop to buy from. So there are all these different things that you give up, but you get lower cost and faster availability.

So, for the IT function, the opportunities and the significant impacts are about shifting your culture from a

build-to-order world, where everything is done in a bespoke model, to a configure-to-order world. This is just as relevant within an IT organization as it is for a service provider. For example, EDS has a lot of culture and competencies around build to order, and a lot of the opportunities that we can address from a business perspective concern how we capture that best-practice knowledge in a form that allows us then to apply it in a configure-to-order delivery model.

PwC: How does your framework apply to cloud computing?

RD: We don't think any of this has anything to do with the cloud other than one important dependency. We think quadrant 1 is the world of classic IT. Quadrant 3 is the world of utility computing. It is what HP talks about as adaptive infrastructure, what Gartner calls real-time infrastructure, what Forrester calls organic IT. There's a huge amount of pressure in the industry to grab onto a new term and to recharacterize everything we already do as embodying the values of that new term. So, people all over the industry have taken their existing slide decks and slapped "cloud" around liberally, primarily for the things that are here in quadrants 3 and 4.

It's hard not to call it cloud. I'm just saying that it doesn't really clarify anything. But there is something else going on, and let's call that cloud. If you look at the figure again, you'll see I've added above the four existing quadrants a private cloud and a public cloud.

And so the key question to ask is what was added to our four existing quadrants to justify a shift from just talking about utility to talking about cloud? Our answer is that it's about how you design. It's primarily around software, but you can think about software and systems. You get fundamentally different results and expressiveness, allowing you to solve a different class of problems—problems that when you try to solve them in traditional delivery models, the projects tend to be either so expensive they never get started or so complex that they underdeliver.

PwC: To move from BTO to CTO, what is the role of automation of IT processes? What are the unsolved challenges in automation to achieve a configure-to-order future?

RD: We think automation ultimately ends up being more important, because you can automate both physical as well as virtual environments and support configure to order of more and more complex application architectures. Additionally, if you have the ability to rethink the way you architect workloads, it then allows you to attack a whole set of problems that you can never address in the context of traditional application designs.

Another thing about automation is that the traditional approaches to automation are significantly focused on imperative logic. So [for provisioning], either it's scripts, and scripts are software with a particular notation; or it's workflows, which are software with a particular notation—it's all imperative logic.

And our view is that to successfully automate more and more complex kinds of configure-to-order deliveries, you really need to focus on forms that are more declarative rather than imperative in their expression. We have a strong focus on how to capture the relevant information in information models and then allow systems software to perform introspections on those models and to understand what must happen in a sequential workflow-centric approach.

The advantage of that information model is that you can leverage it in multiple ways. It also ends up being just as simple. It's much easier to generate data than it is to generate code.

PwC: HP has been supporting the management of IT environments via its OpenView platform. Where are we with managing the infrastructure in cloud environments?

RD: In many cases, the key to understanding the transformation that's going on right now is to realize that most of the management technologies in the market today have focused on addressing the needs that come from this build-to-order world—complex heterogeneous environments with a huge degree of customization. Another key component is that in the BTO world, there's a very strong distinction between the creation of the service versus the ongoing delivery and operation of the service. The classic model is that the software developer builds something and then throws it over the wall to the operations team who then has to figure out how to make it work well enough to meet the business needs. Because of the convergence that we see between those development and deployment activities [in CTO], you then have the ability to shift more of the responsibility back in the pre-production concerns.

So if you look at our management portfolio, it isn't just about operations management. When we brought in Mercury, we brought in a lot of capabilities related to quality management, performance testing, and validation. "How can I do more work earlier in the life cycle to characterize a workload to be able to take out risk in operations?" And the key thing that happens then is you turn down some of the pressure on improved process maturity of IT organizations and, instead, push it into, "Let's try to get more of these things addressed architecturally."

PwC: What will be the impact of cloud computing on enterprises? What is the new opportunity for them?

RD: During the last 20 years, there's been a lot of increased growth and interest around enterprise architecture, where enterprise architecture really focuses on understanding the behavior of the enterprise itself and where the organizing abstraction is the business process. So the highest-level organizing structure within enterprise architecture comprises those fundamental business processes, order to cash, and so on.

We think the cloud allows you to step up to yet another level of architectural design concern, where, rather than thinking about the processes the business uses to accomplish its end, you think about the business in the context of the ecosystems in which it creates value. So it's the business in context. How do you interact with your customers? How do you interact with your channels, your partners, your suppliers? It's those abilities to understand, to share information, to encourage the useful collaborations that you need within those ecosystems—that's why we think the cloud is so important.

PwC: Does that need to be in a shared cloud, ideally, and therefore something that is a result of public cloud services?

RD: We think the cloud ultimately is all about using information and focusing on sharing information with the intended purpose of improving collaboration across an ecosystem. And, again, there's this subtle distinction that within an ecosystem you can't require conformity around a business process.

It's those ecosystem interactions, and this comes down to our definition of the cloud, which is that it's the next phase of evolution of the Internet. The characteristic of the next phase of the Internet is the ability to capture huge amounts of data—adding and extending, and creating more and more context. That characteristic enables you to create much richer kinds of interactions in the context of ecosystems that you care about.

PwC: Can you provide an example of how cloud architecture will enable new interactions and value in an ecosystem?

RD: Let me give you one practical example. Think about the issues of product traceability in the produce supply chain. Was the salmonella outbreak because of the tomatoes or the jalapeños, or maybe it was the onions or the cilantro—because it was found in salsa, right? We figured that out. If you try to think about how you would answer that kind of question from a technical implementation perspective in our traditional architectures, you would need point-to-point integrations between the supply chain management systems of all the participants in that extended supply chain. That's a very expensive and complex thing to do, and the value of being able to answer these questions isn't great enough that you'll go through that cost and complexity. So when we try to answer them today, practically speaking, people are pouring through printed reports to try to figure out what happened.

Instead, the cloud approach says each of the participants in that extended supply chain just publishes a data stream to a cloud service, where that data is captured. Now, when we need to answer the questions, rather than trying to bring all this data into an enterprise data warehouse for analysis, instead you apply searchlike algorithms. You write code using JSON [JavaScript Object Notation—a lightweight data-interchange format] data structures, and you use MapReduce as your parallel distribution algorithm. You go through this two-phase process, and you come up with answers. The answers aren't transactional, they're not right to the last penny, but they are probably far better than where you would start—where you would get to in any bounded time.

You've minimized the complexity for all the participants. There was no point-to-point integration. All you had to do was generate a report, which is an easy thing to do, and if you add a new player in the supply chain, it's just one more report that gets generated. Somebody leaves, and that report's no longer generated. You don't need to do anything more at the nodes than that. All the cost and the complexity of the analysis goes to the cloud service itself. ■

Abstracting IT closer to business value

Doug Hauger of Microsoft explains how cloud computing is a combination of a new abstraction of resources to developers and new business models.

Interview conducted by Vinod Baya and Bo Parker

Doug Hauger is the general manager of Business Strategy in Cloud Infrastructure Services at Microsoft. Before joining Microsoft, Hauger was the director of technology at Cambridge Technology Partners, where he focused on implementing Microsoft technology in large enterprise environments. Hauger is also the author of several books on Microsoft technologies.

In this interview, Hauger shares his insights on how cloud computing disrupts from both a technological and business model perspective and will transform IT organizations toward bringing more value to their enterprises.

PwC: Do you see cloud computing as a disruption of the same magnitude as the Internet, or is it different?

DH: I think there certainly will be a lot of disruption from what people are terming cloud computing. There is a lot of confusion in the marketplace, because everyone is now calling everything cloud computing, and, quite frankly, there should be two ways to look at cloud—as cloud computing and cloud services. The disruption is twofold: One is from a technology perspective, and one is from a business model perspective.

On the technology side, there's some disruption because cloud computing is simply the ongoing abstraction of the developer and the end user from the resources of computing. We've seen this over time: Operating systems, virtualization, the Internet, and so on move people up the stack and abstract them from the underlying infrastructure, platform, or software.

That disruption, I believe, is not going to be incredibly dramatic, although somewhat disruptive. I say that because IT departments, end users, developers, and ISPs [Internet service providers] have grown somewhat accustomed to the fact that every five or seven or ten years there is another cycle of innovation in the IT industry. We're sort of on the fifth generation of computing as we think about it at Microsoft, and in this fifth generation people now just know, "Yeah, there's a fifth generation. There will be a sixth one; there will be a seventh."

I think the greater disruption is around the business model. How do you think about transferring cap-ex to op-ex? How do you think about shared risk with your vendors? How do you think about the pay-as-you-go or subscription model to IT services, which really moves these services to more of a utility type of engagement rather than a capital investment in infrastructure type of engagement?

PwC: Some argue that many of the characteristics that cloud computing promises have been with us since the mainframe days. What do you think is new here?

DH: There's something new on the business model side that enterprise customers in particular are becoming more and more comfortable with: the pay-as-you-go concept. When I say pay-as-you-go, it doesn't necessarily mean that you can't prepay or you can't have a subscription service, but it's really paying for the unit of resource that you consume. So what's new is that you're seeing acceptance of that concept, even though that concept has been around for 30 years.

PwC: It's well accepted that the pace of business change is accelerating. How will Windows Azure enable agility for your customers compared to dedicated in-house Windows Server deployments?

DH: What it provides versus on-premises is an abstracted platform as a service for your internal developers. With Windows Azure, we do that for you, and we do it in a globally scalable, distributed way. The main differences are who is doing the heavy lifting to build the platform as a service and who is taking the risk to build out the data center capacity to meet the demands. That's us in the case of Windows Azure, and it's the IT department in the case of, say, Windows Server System Center.

The goal we have is to radically accelerate or increase the effectiveness of developers—decreasing their time to market or time to service realization and, most importantly, increasing their ability to respond to business needs.

PwC: What will drive the adoption of Azure? Is it new applications that enterprises want to create or the migration of applications they have right now?

DH: The early adoption curve is new applications to be built on the platform, because there's obvious simplicity there. If you have a legacy or existing application on-premises that you want to move into a cloud computing platform, there's a lot of re-architecture that will have to take place. So it will be easier to build a new application on the platform. That said, I have been surprised by the level of enthusiasm of enterprise customers to move existing applications onto the platform and do the work required to actually get the benefits of platform as a service.

PwC: What will be the impact of cloud computing on the role of CIOs and their organizations?

DH: The impact will be different depending on whether people move toward using the off-premises public cloud or building an on-premises private cloud. If there's a wholesale migration toward public cloud, the IT

"There is a lot of confusion in the marketplace, because everyone is now calling everything cloud computing, and, quite frankly, there should be two ways to look at cloud—as cloud computing and cloud services."

"The role of IT departments will be less about running the infrastructure and more about getting value out of the platform for the business."

departments first must become smaller as some of the functionality is not needed and second must learn how to become a value-add line of business around application architecture—public cloud computing. The role of IT departments will be less about running the infrastructure and more about getting value out of the platform for the business.

If businesses are more focused on building private clouds, clearly there is a huge amount of work required to retrain the system administrators and the networking teams to understand the level of expertise needed to build that private cloud and deliver it to services. In the short term, it will require significant reskilling or retooling of the IT teams.

PwC: What are some of the differences between Windows Azure as a cloud operating system and the Windows Server operating system?

DH: The major difference is the level of abstraction from the infrastructure to turn it into a cloud operating system. You're moving developers up the stack, so they're more removed from the infrastructure. Another difference is the global scale and reach of Azure—the ability to scale out across tens of thousands, hundreds of thousands of servers around the world.

PwC: How does this abstraction simplify what developers do?

DH: They don't have to manage the operating system. If you're a developer today, what do you do? You install Windows Server, you load a bunch of libraries, you install SQL Server, you configure a bunch of server-side stuff, you register things in the registry, then you write some code, and then you test it and run it. Whereas, when you have platform as a service, you don't install the operating system, you don't register a bunch of libraries, you don't do things in the registry, you don't install SQL Server. You don't open up the box and put in 16 gigs of RAM when you had 8 gigs of RAM. You don't decide: "Oh, I should have had four 12-terabyte drives instead of two 12-terabyte drives." You just write your application.

PwC: What will be the impact of cloud computing on application architectures? Is there a possibility of running applications distributed across off and on-premises cloud environments?

DH: Absolutely. I think some applications will move wholesale [to cloud], and some won't move at all. There certainly will be a disaggregation of functionality, where some will be in the cloud, some will be on-premises, and something that I've heard from customers a lot is that some will be in one cloud and some will be in another cloud. That's probably not as likely a scenario, simply because the only reason to do that is cost arbitrage. I think the marketplace will be such that there's not going to be a lot of advantage from a cost perspective, going from one cloud to another.

PwC: How do you anticipate you might be talking about cloud computing five years from now?

DH: Well, in some ways, I believe we won't be talking about it, because it's like many other technology phases. It will be implemented and passé. It will be part of the fabric of computing. We don't necessarily try to define every day what client-server is, because we use it. Cloud computing will be just one more way for you to get access to computing resources. ■

Acknowledgments

Advisory

Sponsor & Technology Leader
Paul Horowitz

US Thought Leadership

Partner-in-Charge
Tom Craren

Center for Technology and Innovation

Managing Editor
Bo Parker

Editors
Vinod Baya, Alan Morrison

Contributors
Brian Butte, Galen Gruman, Larry Marion, Bill Roberts, John Verity

Editorial Advisers
Larry Best, Gary Delooze, Glen Hobbs, Jim Kinsman, Bud Mathaisel, Justin McPherson, Jonathan Reichental, Terry Retter, Deepak Sahi, Joe Tagliaferro, Gerard Verweij

Copyedit

Lea Anne Bantsari

Transcription

Paula Burns

Graphic Design

Art Director
Van Collins

Designer
Suzanne Lau

Illustrator
Donald R. Bernhardt

Photographers
Diana Lira
Marina Waltz

Online

Director, Online Marketing
Jack Teuber

Designer and Producer
Joe Breen

Review

Paula Adler, Richard Beaumont,
Sean Middleton, Bob Zukis

Marketing

Bob Kramer

Special thanks to

Gregg Agens, Robert Moline, Victoria Huff,
Mark A. Lobel, James O'Shaughnessy

Industry perspectives

During the preparation of this report, we benefited greatly from interviews and conversations with the following executives:

Erich Clementi, general manager of
Enterprise Initiatives, IBM

Simon Crosby, chief technology officer,
Virtualization and Management Division,
Citrix Systems

Russ Daniels, chief technology officer,
EDS, an HP company

Eran Feigenbaum, director of security,
Google Apps, Google

Doug Hauger, general manager of Business Strategy,
Cloud Infrastructure Services, Microsoft

Geir Ramleth, chief information officer,
Bechtel Corporation

Kirill Sheynkman, founder, president,
and chief executive officer, Elastra

Irving Wladawsky-Berger, chairman emeritus,
IBM Academy of Technology, IBM

pwc.com/us

To have a deeper conversation about how this subject may affect your business, please contact:

Paul Horowitz
Principal, Technology Leader
PricewaterhouseCoopers
+1 646 471 2401
paul.j.horowitz@us.pwc.com

This publication is printed on Coronado Stipple Cover made from 30% recycled fiber; and Endeavor Velvet Book made from 50% recycled fiber, a Forest Stewardship Council (FSC) certified stock using 25% post-consumer waste.

Recycled paper

Subtext

Cloud computing	An umbrella term that encompasses the delivery of information technology resources as a service distinguished by self-service, on-demand provisioning and usage-based pricing schemes.
Evergreen IT	A future-state vision of enterprise IT that anticipates dramatic reductions in IT complexity, delivering disruptively higher levels of flexibility in capacity without long-term financial commitments. Evergreen IT will leverage radical standardization of the infrastructure, virtualization software to support varied workloads, and data center operations managed by powerful software automation.
Virtualization software	A software environment that harnesses collections of servers and storage devices to present the distinct operating environments needed to execute a variety of compute workloads, thereby separating the traditional dependencies between applications and hardware.
Data center automation	Management of the IT infrastructure using sophisticated software that fully models and controls the complex interdependencies between and among applications and IT infrastructure.

Comments or requests? Please visit www.pwc.com/techforecast OR send e-mail to: techforecasteditors@us.pwc.com

PricewaterhouseCoopers (www.pwc.com) provides industry-focused assurance, tax and advisory services to build public trust and enhance value for its clients and their stakeholders. More than 155,000 people in 153 countries across our network share their thinking, experience and solutions to develop fresh perspectives and practical advice.