Lowering IT Costs While Expanding Services:
IBM’s Own Evolution to a Smarter Enterprise

Executive Summary
IBM Corporation is in the midst of an amazing information technology (IT) transformation project that should enable the company to recover approximately $2.3 billion in IT capital/operational costs by 2015 while continuing to expand and improve IT services. IBM is reducing its computing costs by:

- Using a workload optimization model (known as the enterprise computing model or ECM) to analyze application workloads — and then placing those workloads on the systems architectures best suited to most efficiently execute those workloads. To date, IBM has migrated approximately 8,000 servers to more efficient scale-up consolidated and virtualized servers. The company plans to expand its migration scope to include a total of 15,000 servers over the next few years;
- Reducing the number of software licenses the company had been using by 93%. Scale-up virtualized servers perform more work than smaller, underutilized servers — and this means that more work is being done per core on scale-up servers. Software licensing costs are based on core-cost. More efficient use of cores means that fewer licenses are required — and this results in huge software cost savings;
- Reducing IBM’s application portfolio from approximately 15,000 in 1998 to only 4,500 today (this means IBM systems managers have fewer applications to manage — and this reduces human-labor management costs);
- Reducing energy costs at the systems level (by doing more work on fewer servers), and by reducing energy use within the data center (by using more efficient power supplies and cooling equipment). Energy usage has now been driven down by thirty thousand megawatt hours (MWh) per year;
- Reducing the number of physical network connections needed to link smaller distributed systems by 55% (significant cost savings are achieved by lowering the cost of physical plant as well as the cost of labor to deploy and maintain networks);
- Reducing floor space by 74,000 square feet (this averts having to build new multi-million dollar data centers to accommodate increasing computing demands — and the additional space is often allocated to IBM’s services organization to accommodate increased demand for IBM outsourcing services);
- Consolidating data centers (IBM has consolidated 155 data centers down to 7);
- Greatly reducing maintenance costs by using far fewer servers;
- Improving the utilization rate of its storage (through storage virtualization and provisioning) — and by placing data in appropriate tiers (tiering helps reduce
Lowering IT Costs While Expanding Services: IBM’s Own Evolution to a Smarter Enterprise

storage costs by placing lesser used data on less expensive storage devices). These efforts have resulting in an increase in storage utilization from 50% to 90%;

- Greatly reducing management costs (for instance, migrations to System z reduce management labor requirements by approximately 50%. Human-related labor often represents more than half a given organization’s IT budget. So taking human labor costs out of the equation by using sophisticated management tools serves to greatly reduce operational costs);

- Automating systems/storage management wherever possible. IBM uses IT service management practices to automate costly, human labor-intensive management tasks — thus driving down management costs; and by

- Focusing on cloud service delivery models to expand service offerings while keeping costs under control.

Together, all of these activities have enabled IBM to save almost $300,000,000 in IT capital/operational costs since 2007. AND REMEMBER, THESE SAVINGS ARE CUMULATIVE! So if IBM takes $300,000,000 out of its cost structure this year, this also reduces the IT budget burden by $300,000,000 next year and the year after and the year after. Due to this compounding cumulative effect, it is easy to see how IBM can achieve billions of dollars in savings...

But taking capital/operational costs down is only half the story. IBM’s IT organization is also supporting a 25% compound growth in data — AT FLAT COST — and improving IT service at the same time (for instance, IBM has been able to reduce the provisioning time for a new computing environment from 5 days back in 2007 to 1 hour today).

In this Case Study, Clabby Analytics shows how IBM’s approach to transformation evolved from an early systems migration effort to a company-wide IT transformation. We describe how three IBM projects and initiatives (IBM’s ECM project, Project Big Green, and IBM’s new enterprise data center [NEDC]) have enabled IBM to tremendously reduce computing costs — and we show how these projects/initiatives have contributed to IBM’s overall go-to-market strategy (known as “smarter computing”). Finally, we discuss how IBM’s deep migrational expertise, and the company’s comprehensive set of methods and services can be used to help IBM customers systematically reduce IT costs while improving overall IT service delivery (every CIO’s dream!).

IBM’s Transformational Journey to Date
Clabby Analytics has been following IBM’s IT optimization efforts fairly closely over the past five years — and the way we see it, IBM’s approach to IT transformation was not based on a formal, up-front master plan, but instead was formed using an iterative approach. First IBM experimented with large scale application migration from aging pSeries and System x servers to mainframes — finding ways to reduce expedite migrations while reducing costs along the way. Then IBM attacked data center costs — initially taking aim at reducing power and cooling costs — followed by reducing the size of data centers. And finally, IBM took aim at reducing information systems management costs through centralized management and by automating previously manual management tasks. We associate each of these three steps with specific IBM projects: 1) the enterprise computing model; 2) Project Big Green; and, 3) the new enterprise data center. Each of these projects is described in greater detail in the following subsections.
2007 to the Present: The Enterprise Computing Model

In early 2007, IBM set out to reduce capital and operational costs through consolidation and virtualization of server resources. Its goal was to migrate approximately 4,000 aging pSeries and System x servers to Linux on several System z mainframes (the initial target was to migrate these workloads to about 30 mainframes). And what IBM was about to find surprised even IBM…

IBM intuitively knew that its pSeries and System x servers were being underutilized — which is why it embarked on its large migration program in the first place. But imagine IBM’s shock to find that some utilization rates for these servers were as low as 5%. Back in 2007, distributed server utilization rates of 10-20% were quite common — but for IBM, the market leader in scale-up computer systems to have deployed servers with such low utilization rates was almost unconscionable. IBM knew that it had to significantly increase its system utilization rates in order to contain/reduce costs.

As IBM proceeded with its migration efforts, the company soon learned that not all applications should be moved to mainframes. Windows applications, for instance, could not transition to Linux. And, some AIX (Unix) workloads were easier to “flip” than to migrate. (Flipping refers to moving an application that may run on a heavily underutilized pSeries to a more modern, consolidated and virtualized Power Systems server environment). Further, IBM learned that some applications just plain perform better on Power Systems than on a mainframe (floating point calculation applications, for instance).

It was in this timeframe that IBM started to talk about how matching application workloads to the appropriate server architecture help maximize efficiency. Initially IBM referred to this concept as “fit-for-purpose”, but the newer parlance for this activity is called “tuned to the task”. Tuned to the task is now one three pillars of IBM’s overarching “smarter computing” go-to-market strategy.

To streamline the migration process, IBM developed a model for assessing applications and for moving applications to other server environments. This model became known as IBM’s enterprise computing model. Back in 2008, it took IBM about 200 days to migrate applications off of one server to a new environment — and it cost approximately $43K per server to do so. By the end of 2011, ECM methods and best practices had developed to the point that it now takes less than 90 days to migrate a workload from one server to another — at a cost of less than $12K per server.

It should be noted that IBM's current application distribution shows that 48% of applications are moving to System z, 40% to Power Systems, and 12% to System x. Also note that in 2008, it took approximately 200 days to effect a System p AIX (Unix) to System z Linux migration at an average cost of $43K per server. In 2011, due to efficiencies in the migration process, a typical server takes less than 90 days to be migrated, at a cost of only $12K per server. Finally, it is important to note that IBM has seen an average reduction in total-cost-of-ownership of a whopping 70% by migrating older servers to more efficient, better utilized, consolidated, virtualized servers.

Project Big Green and the New Enterprise Data Center

Fueled by rising energy costs, in mid-2007 IBM launched another initiative known as Project Big Green — a collection of products and services aimed at lowering data center energy costs. This program complemented IBM’s ECM efforts as it, too, focused on lowering IT operating costs. The goal of Project Big Green was to focus IT executive
Lowering IT Costs While Expanding Services:  
IBM’s Own Evolution to a Smarter Enterprise

attention on how their data centers operated (illustrating power inefficiencies, fragmentation, lax distributed system design practices, …) — and then show how IBM could help data centers operate more efficiently through more efficient information systems and data center designs.

Project Big Green was followed by a third initiative known as the new enterprise data center (NEDC) — an initiative that combined elements of ECM and Project Big Green, while introducing the concept of integrated service management. IT service management is an approach to information systems management that leads to streamlined process flows; enables IT managers and administrators to better manage and secure applications, resources and services; and helps optimize systems and improve Quality-of-Service (QoS includes reliability, availability, security, resiliency, and so on).

Based upon the lessons learned during its own internal migration and optimization efforts, as well as lessons learned in building energy efficient data centers, and lessons learned from reducing management costs using IT service management, IBM is now positioned to help its customers transform their IT operations. Using knowledge gained from its own migration/transformation efforts, IBM can help its customers to learn from IBM’s experience, enabling these customers to build “smarter enterprises”.

Go-to-Market: IBM’s Smarter Computing Initiative

IBM’s own IT transformation experiences, combined with lessons learned helping its customers streamline their own computing environments, led IBM to build a marketing campaign around three principles that it had observed during the course of thousands of migrations. This campaign, known as “smarter computing”, focuses on:

1. Deploying applications on systems best suited to serve them (IBM calls this “tuned to the task”);
2. Building data/storage management designs up-front before a migration takes place (this practice is known as “designed for data”); and,
3. Automate systems, storage, and process management using an IT service management approach (this is known as “manage as a cloud”).

Tuned to the Task

From our perspective there is a clear linkage between IBM’s ECM program and smarter computing’s tuned to the task message. IBM learned through its ECM experiences that workloads should be placed on systems best suited to execute those workloads. IBM now offers its customers tools and utilities, as well as migration services, to help customers analyze their workloads and place them on the best server to execute those workloads.

Designed for Data

Designing for data implies that IT executives examine their organizational structure; analyze the variety of data that needs to be managed (structured and/or unstructured); examine the sources of data (to come up with one version of the truth); examine data volume; and analyze where data should be placed (the closer data is to the CPU, the faster it can be processed). IBM encourages its customers to come up with a data deployment
and management plan before they undertake a system migration. We see this as a sage approach because insufficient planning can result in much higher migration costs if customers have to go back and rebuild information systems and related infrastructure to meet performance service levels or to handle more data volume than initially planned. 

IBM practices what it preaches when it comes to designing for data as manifest in IBM’s Blue Insight cloud environment. IBM’s Blue Insight cloud was designed for data up front — allowing a large user population (over 180,000 users) to perform business analytics on a very large database. IBM looked at its organizational structure, rationalized its database, built a cloud that can handle the volume of data that needs to be analyzed, and tiered its data such that the most used data is placed in closer proximity to the processors for faster processing. IBM estimates that will save $20 million in system and management costs using this environment.

*IBM Takes the “Designed for Data” Concept a Step Further*

When we interviewed a member of IBM’s Office of the CIO we found that IBM is also heavily focused on data management and storage efficiency. In order to achieve a compound annual growth rate of 25% — at a flat cost — the Office of the CIO is focused on:

1. Reducing the volume of data to be stored;
2. Storing more with what is on the floor already; and,
3. Moving data to the right place (tiering).

To reduce the volume of data to be stored, IBM uses advanced techniques to compress data — and to get rid of duplicate data.

To store more on the floor, IBM makes heavy use of a storage virtualization technique known as “thin provisioning” to maximize the utilization on its storage devices. Thin provisioning is about allowing systems to claim more storage than is physically available — based on the assumption that not all workloads will be demanding all of their allotted storage at the same time. Using thin provisioning combined with compression, IBM has been able to drive storage utilization rates from 50% to almost 90%.

IBM’s storage tiering effort is also quite interesting. In the past, storage tiering was a largely manual effort that involved a lot of human analysis and labor in order to move storage to the right storage devices (oft used data would be placed on fast disk or solid state storage in order to speed performance; lesser utilized data could be offloaded onto comparatively inexpensive tape devices). IT storage administrators had to figure out what data was being used and how often that data was being used in order to make decisions on where that data should be located. This process was cumbersome and costly — and, accordingly, enterprises chose to buy more storage rather than deal with the management complexity of manually tiering storage.

Over the past few years, however, sophisticated storage tiering products have arrived to market, performing storage analysis and tiering tasks automatically. In fact, IBM offers several commercial products that enable customers to locate storage on the right tier to ensure the right balance between efficient data delivery and storage costs. IBM’s Office of the CIO uses these products to tier IBM storage worldwide.
Lowering IT Costs While Expanding Services:  
IBM’s Own Evolution to a Smarter Enterprise

We also found that IBM’s IT organization is taking advantage of even more advanced storage automation, migration and tiering products (known as “Spark” and “Sage”) that have been developed by IBM Research and are being used on an experimental basis. As an example of how effective IBM’s storage tiering is, in a case study IBM found that the time it took to tier a 57 Terabyte database could be reduced from 235 hours of labor to 6 hours. Storage tiering, therefore, plays a huge role in helping IBM reduce its management costs, which in turn helps IBM lower operational costs as it works to grow its storage at 25% compound annual growth at a flat cost.

Manage as a Cloud
IBM’s Office of the CIO is a big believer in the benefits of cloud architecture. Clouds can help lower computing costs through virtualization and resource pooling — and they can help lower management costs. Accordingly, IBM’s IT organization deploys clouds wherever cloud architecture makes sense. As examples of IBM clouds, consider:

- **IBM’s R&D Cloud** — Cloud architecture makes it easy for users to configure their own resources (a concept known as “self-service”). IBM uses this self-service approach in its research and development organization, and has set up a development and test cloud environment. Using this cloud, developers can reduce the time it takes to set up a server environment from five days to an hour or less. This significantly lowers labor and management costs — while at the same time delivering better service to users.

- **IBM’s Storage Cloud** — IBM’s storage cloud enables IBM employees to automatically provision their own file storage — offering a storage self-service to approximately 130,000 IBMers globally.

- **IBM’s Compute Cloud** — A final example of another cloud environment that has been deployed at IBM is its production compute cloud environment. This hybrid cloud (it uses several different computing platforms) helps users find computing resources to run initial implementations and/or to support many non-critical applications.

- **IBM’s Collaboration Cloud** — 85% of IBM’s web meeting minutes are conducted through IBM’s collaboration cloud using IBM’s LotusLive environment. Using Lotus Live Engage, 6,000 registered users can share files internally and with clients. This cloud is considered a Software-as-a-Service cloud implementation.

The Importance of IT Service Management
As stated earlier, one of the largest costs in data center operations is related to management (management costs can be over 50% of an enterprise’s IT budget). To drive down management costs, IBM has adopted an “IT service management” approach to applications/systems management that dramatically shifts the way that its computing environment is managed.

Key to this IT service management approach is the concept of setting up standardized services. Quality-of-Service (QoS) requirements such as high availability, security, and performance vary by workload. For instance, mission critical run-the-business applications must be highly available and secure, whereas a spreadsheet application may not need to always be available, nor may it have high security requirements. To accommodate
Lowering IT Costs While Expanding Services:
IBM’s Own Evolution to a Smarter Enterprise

different QoS levels, IBM has created standardized tiered “service classifications” into which applications and workloads can be placed (see Figure 1).

**Figure 1 — IT Service Management and Tiered Service Classifications**

<table>
<thead>
<tr>
<th>Service Classification</th>
<th>Tier 1</th>
<th>Tier 2</th>
<th>Tier 3</th>
<th>Tier 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Applications</strong></td>
<td>Most Critical Gold and Silver</td>
<td>Gold, Silver multi country</td>
<td>Bronze, silver single country</td>
<td>Blue</td>
</tr>
<tr>
<td><strong>Portfolio Coverage</strong></td>
<td>3 %</td>
<td>16 %</td>
<td>16 %</td>
<td>65 %</td>
</tr>
<tr>
<td><strong>Availability</strong></td>
<td>Near Continuous</td>
<td></td>
<td></td>
<td>Best Effort</td>
</tr>
<tr>
<td><strong>Support</strong></td>
<td>24/7 Sev 1</td>
<td></td>
<td></td>
<td>Best Effort</td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td>Multi-site</td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td><strong>Monitoring</strong></td>
<td>Fully Monitored</td>
<td></td>
<td></td>
<td>Minimal / none</td>
</tr>
<tr>
<td><strong>Cloud</strong></td>
<td>Fully managed</td>
<td></td>
<td></td>
<td>Self-managed allowed</td>
</tr>
</tbody>
</table>

*Source: IBM Corporation — July, 2012*

There are a number of important concepts in Figure 1 to be understood. First, notice the application tiers (gold and silver through blue). And notice that, at IBM, only 3% of the company’s applications must have the highest degree of reliability, availability, security, performance, etc. Meanwhile, 65% of the company’s applications fit into the blue category (with lower QoS requirements and lower support implications). Enterprises that do not tier their applications may be paying for service levels that don’t match their application requirements — or for levels of support that exceed their application requirements.

Using this approach, IBM is able to standardize the infrastructure services needed to support each tier (so instead of building infrastructure to support every new server deployment, a server is linked with the infrastructure services needed to meet its workload QoS requirements). Using this model, IBM gets out of having to operate a bunch of stand-alone applications, and instead can focus on end-to-end integration of its entire portfolio. And, instead of managing a bunch of application silos (the ERP silo, the CRM silo, the SCM silo), IBM focuses on managing the underlying infrastructure, ensuring that the infrastructure can deliver the QoS needed by each workload type.

Using this approach, IBM claims that it can support 80% of its application portfolio with standard services from a single service catalog! Using this IT service management approach and associated service catalog IBM can reduce application development costs while accelerating development by using common infrastructure services found in its catalog. This catalog also leads to a 36% improvement in the time it takes to deploy simple applications. Maintenance costs are reduced because IBM needs to maintain a service catalog rather than thousands of individual applications. And, when it comes time to retire an application, an application can be easily shut down and the resources that application used can be easily returned to the systems/storage/networking pool.
Lowering IT Costs While Expanding Services: IBM’s Own Evolution to a Smarter Enterprise

Summary Observations
The way we see it, IBM’s IT transformation is focused on systems efficiency (this attacks capital costs), and on reducing information systems/workload management costs (this reduces the cost of human labor).

From a systems efficiency point of view, IBM has focused strongly on consolidation and virtualization — placing workloads on servers that can most efficiently execute those workloads. And by doing this, IBM is able to run more workloads more efficiently than by deploying a variety of workloads on smaller, underutilized systems. By deploying workloads on scalable, virtualized systems, IBM is getting more work done per system — and this is driving down software licensing, power/cooling, and maintenance costs — while also forestalling the need to build more data centers. And, in IBM’s case, the data center real estate it saves can be used by IBM to provide outsourcing services for its customers.

IBM’s efforts in storage efficiency should not be overlooked. Yes, the company “designs for data” by doing the necessary work up front to determine organizational requirements; address data variety; cleanse data; address data volume requirements, and analyze where data should be placed. But the company is also aggressive in maximizing storage efficiency (through compression, virtualization, and tiering) — and this approach is making it possible to deliver a 25% compound annual growth in data services at FLAT cost!

From a management efficiency point of view, IBM heavily uses its Tivoli product offerings to automate management tasks wherever possible, to manage changes and configurations, to manage incidents and problems — and to manage service requests and provisioning. But we are equally impressed with IBM’s IT service level approach that allows applications and workloads to be classified within 4 tiers — with each tier receiving different levels of QoS and support. IBM has created a catalog of standard services that support each application tier — and by standardizing services, the company takes itself out of creating custom infrastructure and support for each workload it deploys. Using IT service management, the company focuses on standardizing services, end-to-end integration of its application portfolio, supporting service level objectives, and maximizing its total cost of ownership in its infrastructure investment. Without IT service level management, the company focused on one-off builds, stand-alone applications, adhering to service level agreement requirements as opposed to focusing on meeting service level objectives, and on managing application/workload silos.

Given that management is the largest cost when it comes to operating a computing environment, enterprises looking to drive down operational costs should be looking very closely at IBM’s IT service management model.

There are many lessons to be learned from IBM’s IT transformation. But probably the most important is that it is indeed possible to tremendously drive down IT capital and operational costs while improving service delivery. CIOs everywhere should take notice!