The Fifth Utility

An Opinion

V V Preetham

Executive Summary

Utility computing -- often called the fifth utility -- seems to be positioned as the current cool new model for enterprise computing. As always, whenever a new business model is proposed, the industry will try to commoditize the concept so as to align vendor specific products to cash in on the wave.

In this paper I will explain the reasons towards a general shift towards a utility model. I will derive a generic utility computing model definition that is not biased by vendor specific definitions. I will also debate on the key ingredients involved in making utility computing a reality.

The idea of service provisioning for the clients is based on the same grounds as using electricity service or water, wherein clients can consume as much resources as needed but are only required to pay for what is being used. This is 'Utility Computing'.

Whenever an operation or a service becomes an absolute necessity for an enterprise to function, and the abundance and standardization of such service drives the ubiquity of such service, then it will make the most sense for an organization to obtain such services from a external source rather than harnessing such services in-house. For example, we acknowledge that electricity is very much necessary throughout the industry and the general public alike. But, do we ever try to generate our own proprietary electrical systems to power our on-going operations? The same holds true for utilities.

On the same premise, I have based this paper on an axiom suggesting that utility computing model is very much an outsourced model rather than an internally hosted service. Considering IT as a utility is a very highly debatable topic. I do want to suggest that IT as a whole does not fit into the utility model. IT services can be broadly categorized into two systems.

1) Systems that provide strategic benefits for an enterprise.
2) Systems, which are ubiquitous across enterprises.

The utility computing model best fits the second category of services. In my view, the utility computing model implies that majority of the services that are ubiquitous across enterprises can be outsourced. Such outsourced services can be used based on a pay-by-the-drink usage model. The idea of paying for using applications or compute resources or storage, based on the pay-by-the-drink model sounds great in theory.

But does it really work in practice?

To arrive at finding an answer to the above stated question, we need to first understand the concept of a generic utility computing model (without vendor bias). Most importantly, we must try to find the candidate attributes of the applications that best fit into the utility computing model. Eventually, we can contemplate on the practicality of all the variables in the equation that can make the utility computing model usable.

In the Past

Surprisingly, utility computing is not really a new business model!

In the 90s, computer servers and server farming was a proposition that was positioned in the IT industry as a compelling business model for outsourcing some of the resource needs of business organizations. The business model of the 90s proposed that most of the components of a business entity could be hosted outside the organization. Software vendors and ASPs geared themselves to offer such services based on differing usage models. Some proposed a subscription service while others offered metering and lease. But in reality where are all these service providers now? What happened to the proposed model?

Well, only few such hosted services made it to the present day. The industry finally found out that email, and Web presence were the only few services that really made any sense to be hosted outside the company. There are many reasons for the industry not embracing the business model of the 90s. For one, the industry was still not ready to accept such a rental model. Clients were worried about issues on security, efficiency and even the very reliability of such a model. Clients were also apprehensive about the pricing for such a service. Secondly, the 90s ASP model lacked a value-for-the-customer approach. Gartner analyst Peter Dueck, stated: "*ASPs failed for a number of reasons. They did not think about customers and their value proposition and were too much driven by technology.*"

Later in 2002-2003 we saw a downslide in the equity markets. Many companies adopted the growth-by-acquisition model. Major mergers and acquisitions such as Ameritrade-Datek or Bank of America-FleetBoston gave rise to mammoth systems that were complex and hard to integrate and maintain. There was a general trend towards debt reduction through disinvestments. All these factors during the year 02-03 forced the enterprises to bear-proof their investment against volatility in the market. In effect, enterprises froze most of their spending on IT Systems.

The Present

Currently, the concept of service provisioning as an outsourced model is back again. Though this time it is not as vanilla as what was offered earlier, the business model is still looked upon with the same amount of skepticism as in the 90s. The key factors for skepticism still remain the same. Regulatory compliance, security, reliability and pricing are some of the major FUD factors which make clients reluctant to embrace this model. But late 2003 seemed to change some of the variables of the equation to favor the service-provisioning model.

Late 2003 sees the change in the macro economic indicator wherein the IT sector seems to be slowly and steadily recovering from the economic slump. Lot of corrections in the equity market and also the drive down of interest rates are making enterprises to rethink their business models. Outsourcing has become a clear alternative model to leverage cost disparity and lack of expertise. Offshoring (outsourcing to offsite locations like India) has also caught on to benefit from labor arbitrage costs.

Meanwhile, we can also notice some obvious problem-patterns within the enterprises that are emerging.

Pattern 1 - Most of the enterprises are realizing that the IT investments they made during the dot.com bubble are being wasted without proper accountability. Also the OPEX for the existing resources seems to be comparatively higher than what was projected. A break down of the OPEX on a 5-year overlay suggests that the cost paid towards FTEs to maintain the existing systems are seemingly higher than necessary.

Pattern 2 - Mainly, those enterprises that are subjected to the seasonal variance of demand are finding it difficult to load IT system resources only for the sake of a transactional peak that occurs seasonally. For one, the frontal load on the CAPEX and IMPEX is really big time for large systems. Also as suggested, the OPEX for such systems will apparently chew the bottom-line anyway. Such costs are very deterring for large enterprises to venture into. On the flip side, if these enterprises do not cater to the variance in demand, they will lose the competitive edge and more importantly their customer base . This is the current digital paradox for most of the enterprises.

Pattern 3 - Also, due to all the M&As, enterprise managers are finding out that integrating and managing the disparate data centers is becoming increasingly complex. At the same time, mangers have their CFOs breathing down their necks asking them to reduce the cost of integration and maintenance. Modifying the existing IT infrastructure to meet the dynamic (read volatile) business environment seems to be intangible. Focusing on managing such a huge amount of complex IT architecture seems to be impractical for companies that conduct non-IT business. When in business, they want to focus on their core business strengths rather than worry about the IT systems and Infrastructure.

In the following section of the article, I will present a couple of business cases, which portray the problems specified in the patterns. I will also present the solution proposing a general shift towards the utility model. Also, within the business case, I will emphasize the sections that will aid in defining a generic utility computing model.

Case 1:

"Enterprises lack proper accountability and mapping of applications to leverage existing IT assets."

It is a know fact that large enterprise will have their systems spread across different geographies. Server and system proliferation has raised the underutilization level even higher. Gartner analyst says 'many companies don't know where all their servers are located, who controls and owns them and the main functions and applications running on them'. Due to this, there is no proper accountability for resources. Without proper accountability, the resource usage in such organizations will be under planned.

Let us take the case of processor usage for large enterprises. Average processor usage by the applications deployed on widely scattered systems will be anywhere in the range of 40% to 60%. This is a software problem. The applications are not built to take optimal advantage of the IT infrastructure. This problem also mostly stems from improper planning. Due to this, an average of 50% of the resources are not used. But the enterprises still need to bear the full time cost of maintaining such resources independent of its improper utilization. This thoughtlessly adds to the OPEX that bloats the TCO. The ROI on such systems does also decline.

The solution for this problem is as follows:

- The enterprise must consolidate and standardize all the resources across the organization and make them accountable, then the enterprise can map the applications to fully utilize the capacity of the IT infrastructure.

- In order to map the applications to IT assets transparently, the enterprise must **'virtualize'** the existing IT assets so as to provide a clean separation of concern between the applications that support business initiatives and the hardware, storage and networking assets that support the applications.

This poses a twofold challenge.

Challenge 1 - Consolidating the existing systems and standardizing them is by itself a complex and daunting task. To achieve this, the enterprise first of all, needs to audit all the existing hardware and software across geographies. They need to use IT asset management tools to perform such audits. Once they track down the inventory on hand, they must start rationalizing the applications in their portfolio. They must then start identifying the relationships between applications and the IT assets. Then the enterprise needs to create a consolidated logical topology to map the services and flow of information. This way, on the schematic level, enterprises can design their enterprise wide IT architecture more efficiently. Analytic cubes can be planned and positioned optimally for mining and decision support.

Challenge 2 - There is a requirement to **'virtualize'** the resources so that the application can run transparently without worrying about the underlying resource details such as hardware architecture or resource location. Virtualizing the IT systems requires that the enterprise captures all the data about distributed systems at the instrumentation level. They need to obtain information about routers, hubs, switches, servers, IPLBs, firewalls and the likes that exist on a network. They must also capture the bandwidth and the bottlenecks that exist across geographies. Eventually they must also collect information about failure modes, backup and disaster recovery plans for the same. Also, there is a requirement to trace the process boundaries of the currently running applications over the network, middleware and back-office servers.

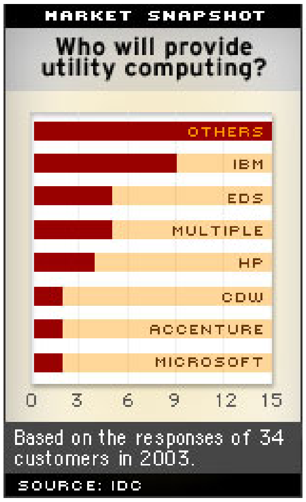
Proper capacity planning and re-architecture can overcome the first challenge. This may or may not require business process re-engineering and rationalization of enterprise wide IT portfolio. Mostly, an enterprise wide IT architecture and a governance framework needs to be in place. A big bang approach to do the same will prove futile. Instead a gradual approach of migrating and re-architecting small portions of the systems must be executed. The enterprise can choose to execute the re-architecture effort on their own, or outsource the efforts to vendors who are best suited to execute such challenges.

Outsourcing such challenges proves worthwhile considering the time and effort spent and also considering the expertise required to execute such transformation. An enterprise wide IT architecture definition is best handled by integrated tech and ops companies, who are strong in technology as well as business consulting services. As an example, we at iGATE Global Solutions offer Enterprise Architecture solutions for our clients to rationalize the application services. We also conduct portfolio analysis to identify the stack of applications that are core to the organization and segregate them from other mission critical processes that are running. iGATE Global Solutions also uses a discriminant analysis tool called pH-Matrix (Process Health Matrix), to identify the process level of the organization. Such detailed analysis will help the clients to identify the processes and services that need to be outsourced. In effect, such analysis will improve the existing processes for the clients that will enhance their revenue, while at the same time provide great cost savings due to outsourcing.

As for the second challenge, to achieve virtualization, one requires a combination of hardware, software, storage and network devices to work in harmony.

Aid to definition: The ‘utility computing model’ is in effect a drive towards ‘virtualization’.

Virtualization tools and services are what the software vendors in the industry are focusing on right now. Vendors are providing tools to trace the business processes that exist within an enterprise. Such tools will help identify the spread of a business process across network boundaries and servers. Such tools also propose to help the enterprises in identifying the relationship between applications and help build an optimal enterprise wide IT architecture.



Major vendors like SUN, HP, IBM, Oracle and others are coming out with a wave of products and platforms to provide virtualization services. Also many service-provisioning vendors are positioning themselves to provide utility computing services for enterprises. The figure provides a market snapshot on the utility computing offering from different vendors as perceived by the market (source: news.com)

Also many arcane concepts like ‘Grid Computing’ and ‘Autonomic Computing’ are positioning themselves as the saviors for the enterprise problem-patterns. In fact ‘Utility Computing’ and ‘Grid Computing’ are similar concepts with different approaches. I will go a step ahead and claim that these are complementary to each other. While utility computing concentrates on service provisioning on a metered or leased basis, Grid computing proposes the means to achieve the same. Grid computing is based on the concept of clustering where one would set up a Grid by clustering distributed computers across the network over the Internet.

Most of the vendor platforms provide service-provisioning systems that directly reduce the number of FTEs required to maintain such a ‘**virtual enterprise**’. As Bill Gates states “*There's something common between the IBM message, the Sun message and the Microsoft message: Some of the things that you do with personnel to operate these systems today should be done automatically with software. We all agree it's a software breakthrough that will let people free up part of their IT budget that now goes toward operations and apply it toward new things. What's interesting is that everyone admits it's a software problem, not a hardware problem.*” This means that software replaces the jobs of system administrators. All this adds to greater reduction in the OPEX.

In effect, when the enterprises succeed in collaborating and consolidating IT architecture using the concept of virtualization, the managers can improve the utilization of their resources, which relatively reduces the OPEX.

Table –1 provides a brief description of the platforms offered by major vendors.

Table – 1

Company	Products	Brief Description
IBM	E-business On Demand	IBM is focusing to position most of its software product line to be based on the 'Autonomic Computing' model. The autonomic computing model is defined to be a systems model to provision services to businesses at real time, based on the dynamic demands of the enterprise. Also systems built on the Autonomic computing model is supposed to be self healing and resilient to change. IBM is also offering a range of hardware options to its clients based on the concept of 'Server Blades'. In effect, IBM is planning to offer utility solutions for its clients by aligning its software and hardware towards the utility model.
Sun Microsystems	N1	Sun's N1 is supposed to be the next-gen architecture for simplifying data center operations. N1 simplifies the operations of a data center by automating most part of the data center management. N1 focuses to provide an integrated management console for the entire IT infrastructure instead of providing individual consoles for relevant storages and servers. Such consolidated view benefits the administrator to look at the IT infrastructure as a whole instead as the individual parts. This in effect virtualizes the infrastructure so that application deployment and mapping can be automated. In effect, the effort is to provide hardware service-provisioning system that simplifies the management of a data center.
Hewlett Packard	Adaptive Enterprise	The major technology theme behind HP's adaptive enterprise is to provide utility computing options to the enterprises to reduce the IT OPEX. HP was first in market to provide 'Server Blades' so as to provide hardware service provisioning options to its clients. HP was also quick to capitalize on the concepts of virtualization of data centers based on the utility model. The HP Utility Data Center (UDC) enables large computing environments to be flexible and cost effective while delivering consistent QoS.
Oracle	10g	Oracle is offering Grid-computing solutions for its clients under the 10g platform. Grid computing is supposed to be a coordinated use of a large number of servers and storage acting as one computer. The 10g platform is based on a modular software component platform that helps enterprises to build capacity by starting small, and adding components as and when the business demands increases. Oracle focuses to provide a data centric approach towards the utility model based on Grid computing paradigm. In effect, the solution will enable enterprise to obtain computing power on demand, obtain automatic load-balancing and easy management of the entire IT infrastructure through Grid control.

Case 2:

"Enterprises are unable to justify the IT system costs to sustain peak seasonal demands."

During the NFL super bowl season of Jan 2003, I remember frantically logging on to the super bowl site to buy some Falcons merchandise. I wanted to get myself a Michael Vick jersey (QB for Falcons). I was based out of Atlanta and rooted for Falcons. The pure thrill of wearing an oversized jersey with Falcons No 7 on the back was indeed exhilarating. Especially when Falcons beat Greenbay Packers on Packers home ground in Lambeau stadium. This was a history in the making. I am sure most of the Falcons' fan will agree with me. Mostly, 10s of thousands of Falcon's fans who would have tried to get some merchandise for themselves.

The point I am making here is… if you are hosting the nflshop.com, an online merchandising shop, how will you be prepared for such an onslaught of transitions. Well, you might say that you would rent a super-sized hardware for the game seasons and return it to the vendors later. The question is… how do you really know how much to super size?

The same question holds true for many businesses. Especially when the businesses are influenced by seasonal variance. Even on an average, any normal enterprise will have a critical window of peak loads, which will be 4 to 6 times the average load of the system. These critical windows are pattern based such as weekdays 4:00 PM to 6:00 PM or weekend 8:00 AM to 4:00 PM.

Other times, the critical sections (same as critical windows) will be unpredictable. When the critical sections hit you, you do not want to bog down to the hits and say you are sorry. You will leave a lot of irate customers who will go elsewhere to buy their stuff. For one, you will lose out on cashing in the opportunity, and for other you may lose your clients for good.

Enterprises usually buy IT systems to accommodate the peak loads of their transactions. In fact the peak loads are pattern based or infrequent wherein it may only account to 20% of the operations. The other 80% of the time, enterprises may be running on a average count of transactions which may be 6x times lesser than the peak. This is usually an 80-20 rule.

Isn't there something wrong with this equation? Duh!! yes of course. We all agree that it isn't rocket science to see where the problem lies. If so, how do you justify the overall costs for such large systems only to accommodate your critical windows?

Well, most of the enterprises had no other go apart from souping up their hardware for the critical sections. Some of the vendors did provide options like blade servers to enable servers in real time for increased transactional loads. These blade servers can be racked up (hot deployed) in real time without bringing down the systems.

This brings down the equation for the enterprises to maintain the 'blade inventory' so that they are available at critical sections. Some vendors allowed renting the blades on demand. For pattern based critical windows, the rental model seems to work fine. In other cases, how do enterprises be prepared for transactional loads that hit them like a lightning from the blue?

There are two things that need to be justified in this equation.

1. How to justify the OPEX of inventory management of the blades?
2. How to be prepared for unforeseen events of real time transactional peaks without super sizing the hardware?

The solution for the both the problems is again the utility model.

> **Aid to definition:** The 'utility computing model' is a drive towards accommodating as much resources as required, on-demand, real-time.

The utility model works well for this business case. Hardware vendors like HP, IBM, SUN and others are providing hardware provisioning options with varying usage models such as Metered usage, Subscription usage and Managed usage.

The generic idea behind hardware service provisioning is that, the vendor agrees to provide hardware systems with active and inactive processors and cells (cells are boards with embedded processors and memory). Active processors are provided to accommodate the average transactions that occurs 80% of the time. Since these transactions are 6x times lesser than the predicted peak, the number of processors required to support the average will be considerable less. The inactive processors come at no cost to the enterprises. These inactive processors are provided to accommodate the peak load that occurs at critical sections.

Given a scenario like the game seasons, whenever a peak load hits the enterprise box, and the tolerance limit set on the active processors are exceeded, then, the inactive processors gets activated. When this happens, a meter built into the system starts clocking the processor usage of the inactive processors. At the end of the month, the enterprise will receive a bill from the hardware-provisioning vendor for the amount of inactive processors used for the period. How cool.

The financial of the prescribed model is a risk-sharing model. In other words, the hardware-provisioning vendor will provide you the extra hardware at no upfront cost for the enterprise. The enterprise only pays for the amount of additional processors that were activated. The OPEX for the additional processors is accounted into the billing period. The justification for OPEX is based on usage only. The enterprise in effect saves the CAPEX, IMPEX and OPEX for additional IT assets that they only use seasonally.

The prescribed model need not necessarily work on the grounds of an active versus inactive processors embedded into a monolithic box. Other technologies such as 'Grid computing' may also be used to cluster disparate hardware devices including printers, scanners, storage, and network devices that can account into the risk-sharing model. Eventually the enterprise will get a single bill categorized on the basis of different devices used and the amount of usage.

In effect

Based on the 'aid to definitions' highlighted during the business case and also the explanations provided throughout the case, we can arrive at a brief definition of the utility model. The definition is as follows:

Utility computing is part of a drive towards creating 'virtual enterprises' whose dynamic demands can be met at real time by provisioning services through a Metered or Leased usage model, mostly hosted outside the enterprise by a service provisioning vendor.

The utility computing model is not just for hardware provisioning. The same logic holds true for software usage as well. In late 2003, you already see a number of hosted services such as Payroll or CRM being offered by the software provisioning companies with a strong revenue model. Salesforce.com is a successful and apt example for a CRM service hosted online.

Imagine the amount of investments that goes into implementing a CRM or an ERP in-house. If a software-provisioning vendor, as a risk-sharing model can host the same as an outsourced service, then the two fold benefits of avoiding the frontal load as well as the OPEX for such systems are immediately realized.

Using the utility model, software service provisioning vendors will be able to provide metered access for CRM systems, ERP systems or even something as basic as spread sheet processing. Considering the flexibility offered by discrete systems, it is in effect possible to build outsourced services for any software system an enterprise may desire.

It is not the possibility of building outsourced services for internally hosted software systems, which needs to be answered. Instead the question that really needs to be answered is, "what type of services fit the utility model and will it make financial and legal sense to outsource such services".

Financial Justification

As per Gartner report, enterprises spend 70% of their IT costs towards OPEX. Only 22% is spent on the CAPEX and remaining 8% on hidden IS related expenses. Table – 2 provides the cost breakups for IS spending. This is based on Gartner's survey on annual IT staffing and spending.

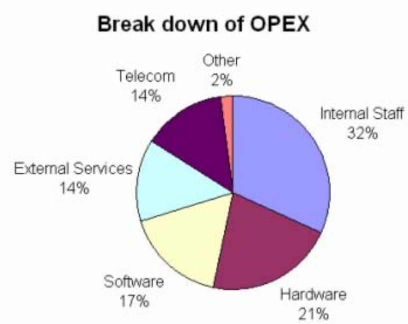
Table – 2

70 % of IS Spending

Break down of OPEX		Break Down of CAPEX	
Internal Staff	32%	Hardware	35%
Hardware	21%	Software	25%
Software	17%	External Services	14%
External Services	14%	Network/Telecom	20%
Telecom	14%	Other	6%
Other	2%	Hidden IS Related Expense	

22 % of IS Spending

Remaining 8 %



OPEX being 70% of the overall IT spending, 32% of the OPEX is based on internal staff costs. This is 22.4% of the whole IT spend. Enterprises expend 22.4% of their total IT costs towards FTEs!!! Also a 36.4% is accounted towards OPEX for Hardware, Software and Network (**HSN**). Together a whopping 58.8% accounted from the OPEX towards the overall IS costs, goes only towards FTEs and HSN. For huge enterprises this is a very large number in dollars.

Also, notice the fact that the total IS spending towards external services is only 12.88% of the overall costs (9.8% from OPEX and 3.08% from CAPEX).

All these numbers suggest that enterprises need to cut down on the OPEX of their IT. To cut down the OPEX, enterprises needs to bring down the cost spent towards FTEs to support the systems. Also, it seems to make financial sense for enterprises to stop adding IT infrastructure in-house. Meanwhile, the OPEX on HSN amounting to 36.4% of the overall spend needs to be cut down.

An obvious choice to reduce costs is to outsource. More so, it does provide greater benefits to offshore the systems for higher cost savings. Even more so, it will make the most sense if you are able to offshore the services to a service-provisioning vendor who is ready to offer a Risk-Sharing Model. How can a vendor justify the Risk-sharing model? Well, the answer is utility computing.

Using the utility computing model, a Business Service-Provisioning (**BSP**) vendor, based in an offshore site can setup hosting services to provide expert solutions, while reducing OPEX for the enterprises. To do this, the BSP vendor must implement Hardware, Software, Storage and Network systems that comply with the utility computing requirements. The vendor must also centralize resources to optimize the logistics of operation.

Once a service-provisioning shop is setup, the BSP vendor will be able to service many enterprises by sharing the virtual unit across many enterprises. In fact a global BSP vendor will be able to take advantage of the time zone disparity across countries and plan the capacity of his systems to be optimal with very less resources. A BSP vendor can save greater costs by not only sharing the capacity across clients, but also take advantage of non-overlapping critical windows. Such cost savings will be passed on to the enterprise clients.

Using the utility model, the vendor will also be able to provide a crisply defined security context for executing an application process. This is possible due to the 'design-by-contract' approach required to virtualize the hardware from the running processes to provide transparency. Since the process boundaries are well defined, and also due to a well-defined mapping of the process to a schematic topology, the administrators of the utility model will have managed control of the processes supported by automated tools. This way, no two processes from the same enterprise or across different enterprises will be able to run over each other. As a whole, the technologies that eventually come into play to setup a utility model will provide better security, scalability, reliability and performance for the overall system. The utility model also enables the enterprise client to better handle any dynamic business demands, which may create critical sections in their transaction patterns.

Consequently, the enterprise clients serviced by the BSP vendors greatly benefits from not only reducing their OPEX (by outsourcing the deal), but also obtain a bigger bang for the buck by getting a scalable, reliable service that performs better than the internally housed systems. At most, out of 58.8% of the IT costs accounted from the OPEX of FTE and HSN (leaving out 'external services' and 'others'), the IS spending can be reduced by 50% of 58.8% which is approx 30%. The 30% reduction already takes into account the service charges borne towards the BSP vendor. Since I did not account for the reduction in the CAPEX, I believe that the projected estimate is conservative, and is a practical proposition. A total cost saving close to 30% is a very healthy indicator by any norms. In fact, the industry average suggests a 28% cost savings for the utility model. The best part of the deal is that, an enterprise client not only has better control over their cost, due to the pay-by-the-drink model, but also would have reduced his risk of investment, courtesy; the risk-sharing model.

Service Rationalization for the Fifth Utility

We have already understood that the utility model works using the façade of virtualization. It is possible to map the existing applications, seamlessly over the IT assets provided that the assets are virtualized. The application services need to execute transparently over a network of servers, middle wares, and storage independent of their locations.

So how does it matter if IT assets are in-house or outsourced? Well, you may say that printers and scanners need to be at the place of use. Off course it will not make sense to print something over the network in India and call them to courier it to you in few minutes. But these are miniature exceptions in the grand scheme of things. The utility model best works for applications, which are core to the business operations but are not strategic. We are talking about huge and complex data centers, large analytic cubes and heterogeneous maze of devices. In fact any system, which uses large storage and local network bandwidth and heavy processing, is a good candidate for utility computing.

Transitioning such services to a BSP will require careful planning and analysis of the portfolio. Once a portfolio analysis is conducted, it becomes easier for the enterprises to identify the stack of services, which defines the essential applications in the enterprise that are core to operations but are not strategic. Again, based on the 80-20 rule, it is true to notice that 80% of the enterprise resources are dedicated to provide the core operations while 20% of the resources provide strategic advantage. So, while transitioning the services to a BSP vendor, it will make sense to disinvest in 50-60% of the hardware, software, storage and networking costs and still retain other systems to support strategic operations. Such disinvestments are in tune with the projected cost savings in the financial justification section. As an example, we at iGate manage the entire hardware, software, networking and operations of the ING bank in India (ING-Vysya). iGate provides a great revenue enhancement and cost savings for the bank by managing their operations.

One another factor that may need to be accounted into the equation while transitioning the services is 'Regulatory Compliance'. Regulatory compliance can be one of the hurdles to outsource utility computing. US regulatory laws such as the Gramm-Leach-Bliley Act (GLBA) require U.S. financial institutions to ensure the security and confidentiality of customer records and information. Enterprises will demand strict security policies and compliance to GLBA by the BSP vendors. Also if for some reason, enterprises are not able to let out information for a third party outsourcer, then, a captive unit can be setup. A captive unit is a JV (Joint Venture) between the BSP and the enterprise so as to provide an in-sourcing deal for a utility setup. An in-sourcing deal can be hosted in places such as onsite, offshore or also near-shore.

BSP vendors can use different business models to provide compliance for regulatory laws such as Patriots Act, Basel II Capital Accord, Sarbanes-Oxley, Section 508 and host of others. In fact, compliance to regulations is a part of the hosting deals for most of the services.

Another hurdle may be the cost of telecom. According to a research paper by Jim-Gray of Microsoft (source:

http://www.clustercomputing.org/content/tfcc-5-1-gray.html), he points out that the economic issues in moving a computer task have four characteristic demands. They are, Networking, Computation, Database Access, and Database storage. He goes ahead to point that the ratios between these quantities and their relative costs are pivotal. He says, "*It is fine to send a GB of data over the network if it saves years of computation – but it is not economical to send a kilobyte question if the answer can be computed locally in seconds.*"

What Jim Gray suggests, makes sense. In fact most of the complex data center operations and analytic queries do include huge processing efforts. The queries by themselves will be in order of few kilobytes. The cost of processing such queries versus the bandwidth required to transport such queries will be a defining factor for utility computing.

Jim Gray concludes that if the cost of telecom drops at the same rate as the price of hardware, then outsourcing utility services makes sense. Currently, this seems to be happening in a smaller proportion. Off late, the broadband industry has come out with various options and packages in the industry. The BSP vendor (both hardware and software) is now having a range of options to choose from so as to enable a high-bandwidth connectivity for the enterprises to use such externally hosted services. The cost of telecom is reducing drastically by the day.

Due to lowered telecom costs and cheap labor costs, the Indian sub continent seems to be a breeding ground for many of the call centers and BPO centers (Business Process Outsourcing). A BSP vendor hosting a utility center from India can take advantage of the lowered telecom costs and labor costs to provide great benefits for the enterprise clients.

Conclusion

Utility computing, in effect, will be one of the best-fit execution models to provide the pay-by-the-drink service provisioning offerings. In the coming days, there will be a number of BSP vendors positioning their service around the utility computing model. There will be a rise in channel partnering between BSPs and System software vendors to provide a federated utility service for the end clients. Multiple tools, software and hardware will be branded as utility computing compliant. A host of protocols will be built specifically for the utility model. System and service standardization will be a prime focus that will be driven by heterogeneous group of vendors. Service stack and communication standardization efforts will be consolidate by major players in the software industry. Other technologies such as Web Services, Peer-To-Peer and the likes will be positioned within the utility computing model to define higher benefits for the proposition.

In fact, all these things are happening right now. Already vendors such as SUN and ACS are collaborating together to provide a risk sharing, pay-by-the-drink model for the clients. Other vendors such as IBM, Cisco, VM ware, Oracle, HP, Veritas, Computer Associates and the likes are coming out with a plethora of options for the clients. The main drive for the utility computing model will be Revenue enhancement and major savings on OPEX. All in all, the fifth utility will emerge as a good revenue model for both the BSPs and the Enterprise clients alike.

A truly win-win proposition, it is.

Acknowledgements

First and foremost, I would like to thank Jessie Paul, Head of Marketing, iGATE Global Solutions, who provided the idea and guidance to write on utility computing. Thanks for all the 'sanity check' efforts by Jessie. In fact, thanks for Jessie's patience to edit the article for all the iterations while she had busy schedules. Thanks to Venkat Ramanan (Financial Markets Group, iGATE) for providing timely and balanced inputs on the article. Also thanks to Lalitha Narayanan, (Marketing, iGATE) for her efforts in collecting research data from different sources to help put the article together.

About the Author

V V Preetham is a Technical Architect and can be reached for further information and comments at
preetham.venkatesh@igate.com