The utility business model and the future of computing services

by M. A. Rappa

The utility business model is shaped by a number of characteristics that are typical in public services: users consider the service a necessity, high reliability of service is critical, the ability to fully utilize capacity is limited, and services are scalable and benefit from economies of scale. This paper examines the utility business model and its future role in the provision of computing services.

The idea of utility computing has received attention recently and for good reason. The use of computers continues to be a rapidly expanding feature of modern society, and industry has come to rely on computers to perform a multitude of tasks beyond simple data processing and storage. Computer networks have extended the reach of computing to connect businesses across the supply chain and, in many instances, directly to the consumer. With the growth of the Internet, the computer has come to play an even greater role in commerce.

Computing has also become a larger and more intimate part of daily life for many people. Individuals now use computers to accomplish a wide array of tasks, from the complex to the mundane.[1–3] Whether it is used for communicating by e-mail and instant messaging, paying bills and managing personal finances, or the pursuit of hobbies and entertainment, the computer has become an essential tool. Indeed, the variety of tasks performed with computers today would have been difficult to foresee as little as two decades ago.

With all this progress has come a greater degree of reliance on computers and their connectivity to networks, and this reliance has bred high expectations for the availability and performance of computing and networking services. This expectation is not unlike that seen in other areas of technology to which modern society has grown accustomed; for example, the dependence on a ready availability of affordably priced electricity. Long ago a curiosity and a luxury, over the last century we have seen electricity grow beyond a modern everyday convenience to become a necessity in the lives of most people.

The prominence of computers in society and our growing reliance on them raises an interesting question: Is computing the next utility? The answer to this question has broad implications for the future of computing. Already, the idea of utility computing has begun to influence the development of computer technology in such areas as the auto-provisioning of computing resources and resource sharing across a computing grid.[4–6] Its potential role in the evolution of business models for computing services is of equal importance, and that role is addressed in this paper.

Common characteristics of utilities

In many parts of the world, although by no means everywhere, services such as water, power, heat, light, common carrier transportation (airlines, buses, and railroads), and telephone access are typically provided by a public utility. What makes any particular

0018-8670/04/$5.00 © 2004 IBM

Table 1 Requirements common to utility services

	Necessity	Reliability	Usability	Utilization	Scalability	Exclusivity
Water	H	H	H	M	H	H
Electricity	H	H	H	H	H	H
Common Carrier Transportation	M	H	H	H	M	H
Telephone:						
POTS	H	H	H	H	H	M
Cellular	M	M	H	H	H	L
Radio and Television Broadcasting:						
Terrestrial	M	H	H	L	H	M
Satellite	M	M	M	L	H	L
Cable	M	H	H	L	H	H
Internet Access:						
DSL	H	H	H	H	H	H
Cable	H	H	H	H	H	H
Dial-up	M	M	H	H	H	L

L = Low relevance
M = Medium relevance
H = High relevance

service a utility is shaped by a combination of requirements (see Table 1), most notably: users consider it a necessity; high reliability of service is critical; ease of use is a significant factor; the full utilization of capacity is limited; services are scalable (leading to economies of scale); and exclusive rights are granted for providing service in a given area.

Necessity. Users depend on utility services to fulfill their day-to-day needs. Doing without service is an unwelcome option for them. Of course, seldom do utility services start out as essential. Its takes time for distribution networks to spread and costs to decline. It also may take time for users to adapt to the service. Once a service does take hold, it may grow in importance as users discover new ways to use it to their benefit. How crucial a service becomes may ultimately depend on the circumstances of the individual user. But once users do come to depend on a service, it can become a transparent part of their everyday reality.

Reliability. The service provided by a utility must be readily available when and where the user needs it. A temporary or intermittent loss of service may cause more than a trivial inconvenience to the user; a prolonged loss of service may cause severe hardship. Because a failure in service has undesirable consequences, utilities must operate with an exceptionally high degree of reliability.

Providing continuous service in the face of various contingencies is a huge technological challenge that utilities face. Because some kinds of services may not be easily or cheaply inventoried, if at all, redundancy must be built into production capacity to make up for the inevitable equipment failure. Furthermore, because utilities provide on demand services, they must deploy transparent failover mechanisms and standby services to ensure continuous availability to the user. If one area of a service grid fails, the system must be able to compensate and respond instantaneously to the shortfall, thereby preventing the disruption of the service.

Necessity drives user expectations of utility services beyond what may be typical in other industries. Whether or not these expectations are realistic, utilities must do their best to buffer users from the predictable problems that could cause a discontinuity in service.

Usability. No matter how technologically complex they may be on the production end, utility services are characteristically simple at the point of use. Users have what could be called a "plug-and-play" mentality. This is not to say that devices connected to a service are unsophisticated, but the utility service itself tends to exist only in the background. Users may become mindful of a utility only in those rare instances when the service fails to meet their expectations. This may explain why the public perception of a utility is not always positive.

One ingredient in making a service simple at the user interface is a high level of technical standardization.

Devices that add user functionality to the service must conform to the specifications of the network. Plug compatibility, independent of the vendor, is a common feature of utility services. Even so, technical standardization can be extremely difficult to achieve. In marketplaces where proprietary innovation is strong, the incentive for competitors to agree on standards is weak. Although a lack of standardization is costly and inconvenient, premature consensus on a standard may forestall significant innovation that can be of benefit to users.

To the extent that incompatible standards take hold, in some cases the consequences can endure for long periods, as fixed investments in infrastructure grow. Just how long this condition can last is illustrated by the case of the difference in voltage standards around the world. In such situations, technologies that enable the conversion between standards become a regular and cumbersome aspect of the user experience.

Utilization rates. Utilities are driven by a need to carefully manage utilization rates. User demand for utility services can fluctuate widely over time and across the service region. Because sufficient production capacity must be installed to handle periods of peak demand, overall utilization rates are typically well below full capacity.

In addition to fluctuations in usage, there may be discrete incidents of an exceptional nature when demand spikes sharply upward. Such spikes can occur when large numbers of users suddenly want to use the service simultaneously. Other spikes may occur when users fear a shortage in supply and begin hoarding, to the extent that it is possible.

Underutilization in off-peak periods provides a strong economic rationale for service providers to shift user demand from peak to off-peak periods. By pricing services according to actual metered usage and by providing off-peak price discounts, fluctuations in user demand can be smoothed out over the cycle. How a service is billed may also create incentives for users to limit their usage.

Scalability. Utilities are commodity businesses. Therefore, utility services can exhibit significant economies of scale that favor larger producers over smaller ones. As production capacity rises, the unit cost of production falls. There may be other size-related benefits as well. It might be expected that as the demand for a service increases beyond some threshold, the quality of service may decline as users begin to compete with each other. However, with some types of utilities, service can become more and more useful as the number of users of increases.

Service exclusivity. The economies of scale in a utility can benefit from a monopolistic provision of services. When this is the case, the government may step in to grant an exclusive franchise in a geographic region. Government regulation of the service and how it is priced typically accompanies such a sanction. Cost-based pricing is a common formula. With the benefits of an exclusive franchise comes the obligation to serve any and all users regardless of how profitable it may be for the utility.

Some of the common characteristics of a utility derive from its relationship with its customers. Other characteristics are derived from technological and business aspects of how the service is produced and distributed. The preceding list of characteristics, while important, is not meant to exclude other possible factors that may be relevant to particular types of utility.

Each of the characteristics described here may or may not play an equal role in shaping any particular type of utility service. Table 1 provides an evaluation in the most general terms of the potential relevance of each factor for public utility services, including water, electricity, and common carrier (or public) transportation. In addition, the comparison is extended to examine a few businesses that have some characteristics in common with public utilities, namely radio and television broadcasting and Internet access services.

The utility business model

The factors of user necessity, reliability, usability, utilization, scalability, and exclusivity, when taken together, shape the business model for utility services. To understand the nature of the utility model, it is useful to place it in the context of business models in general. A business model is a method of doing business. All business models specify what a company does to create value, how it is situated among upstream and downstream partners in the value chain, and the type of arrangement it has with its customers to generate revenue. In any given industry, the methods of doing business may vary, but there are limits imposed by technological factors, by the competitive dynamic among companies and between companies and their channel partners, and by customer expectations and preferences, among other things.

There have been a number of attempts to create schema for classifying the various types of business models seen in practice, particularly in relation to the Internet.[7–11] The commercialization of the Internet during the 1990s drew a great deal of attention to business models. The Internet opened the door to new business opportunities, but many Internet-based enterprises failed because they had not clearly thought through their model — particularly, how money would be made. Nonetheless, given the rapid adoption of the Internet, it may no longer be possible to discuss business models without taking it fully into account.

To understand the nature of the utility model, it is useful to place it in the context of business models in general.

One approach to the classification of e-business models is a comprehensive taxonomy using the customer relationship as the primary dimension for defining categories.[7] Although by no means the only approach, this has proven to be a useful framework because it builds upon a common parlance already used in many industries to describe methods of business. Although other approaches may be more suitable for other purposes, it is unreasonable to expect that any single taxonomy can account for the vast diversity of business models found in practice without becoming unwieldy.

Nine major categories are used to classify a number of different types of business models that have been identified in practice among Web-based enterprises (see Table 2):

Brokerage model. Brokers are market makers: they bring buyers and sellers together and facilitate transactions. Brokers play a frequent role in business-to-business (B2B), business-to-consumer (B2C), or consumer-to-consumer (C2C) markets. Usually, a broker charges a fee or commission for each transaction it enables. The formula for fees can vary. Brokerage models include exchanges, demand collection systems, and auction brokerages.

Advertising model. The advertising model on the Web is an extension of the traditional media broadcast model. The broadcaster, in this case a Web site, provides content (usually, but not necessarily, for free) and services (like e-mail, chat, forums) mixed with advertising messages in the form of banner ads. The banner ads may be the major or sole source of revenue for the broadcaster. The broadcaster may be a content creator or a distributor of content created elsewhere. The advertising model works best when the volume of traffic is large or highly specialized. Advertising models include portals, query-based paid placement, contextual advertising, and content-targeted advertising.

Information-intermediary model. Data about consumers and their consumption habits are valuable, especially when that information is carefully analyzed and used to target marketing campaigns. Independently collected data about producers and their products are useful to consumers who are considering a purchase. Some firms function as "infomediaries" (information intermediaries) assisting buyers and/or sellers to understand a given market.

Merchant model. Merchants are wholesalers and retailers of goods and services. Sales may be made based on list prices or through auctioning. Merchant models include virtual merchants or "e-tailers", mail-order businesses with a Web-based catalog, and traditional brick-and-mortar retail establishments with Web storefronts.

Manufacturer Direct model. The maker of a product or service may sell (by purchase, lease, or license) directly to the consumer. The manufacturer or direct model is based on the power of the Web to allow a manufacturer to reach buyers directly and thereby compress the distribution channel. The manufacturer model may be chosen for its efficiency, improved customer service, or due to a better understanding of customer preferences.

Affiliate model. The affiliate model provides purchase opportunities wherever people may be surfing the Web. Financial incentives (in the form of a percentage of revenue) are offered to affiliated partner sites. The affiliates provide purchase-point click-through (i.e. direct linking) from their Web sites to the merchant's Web site. It is a pay-for-performance model — if an affiliate does not generate sales, no cost to the merchant is incurred. The affiliate model is inherently well suited to the Web, which explains its popularity. Variations of this model include banner

Table 2 Taxonomy of e-business models

Type	Model
Brokerage	*Marketplace Exchange*—Offers a full range of services covering the transaction process, from market assessment to negotiation and fulfillment. Exchanges operate independently or are backed by an industry consortium. *Buy/Sell Fulfillment*—Takes customer orders to buy or sell a product or service, including terms like price and delivery. *Demand Collection System*—The patented "name-your-price" model pioneered by Priceline.com**. Prospective buyer makes a final (binding) bid for a specified good or service, and the broker arranges fulfillment. *Auction Broker*—Conducts auctions for sellers (individuals or merchants). Broker charges the seller a listing fee and commission based on the value of the transaction. Auctions vary widely in terms of the offering and bidding rules. *Transaction Broker*—Provides a third-party payment mechanism for buyers and sellers to settle a transaction. *Distributor*—A catalog operation that connects a large number of product manufacturers with volume and retail buyers. Broker facilitates business transactions between franchised distributors and their trading partners. *Search Agent*—A software agent used to search for the price and availability of goods or a service specified by the buyer or to locate hard-to-find information. *Virtual Mall*—A hosting service for on-line merchants that charges setup, monthly listing, and/or transaction fees. May also provide automated transaction and relationship marketing services.
Advertising	*Portal*—Usually a search engine that may include varied content or services. A high volume of user traffic makes advertising profitable and permits further diversification of site services. A personalized portal allows customization of the interface and content to the user. A niche portal cultivates a well-defined user demographic. *Classifieds*—List of items for sale or wanted for purchase. Listing fees are common, but there also may be a membership fee. *Registered User*—Content-based sites that are free to access but require users to register and provide demographic data. Registration allows inter-session tracking of user surfing habits and thereby generates data of potential value in targeted advertising campaigns. *Query-based Paid Placement*—Sells favorable link positioning (i.e., sponsored links) or advertising keyed to particular search terms in a user query, such as the Overture** trademark "pay-for-performance" model. *Contextual Advertising*—Freeware developers who bundle ads with their product. For example, a browser extension that automates authentication and form fill-ins may also deliver advertising links or pop-ups as the user surfs the Web. Contextual advertisers can sell targeted advertising based on an individual's surfing behavior. *Content-Targeted Advertising*—Pioneered by Google**, the precision of search advertising is extended to the rest of the Web. Google identifies the content of a Web page and then automatically delivers relevant ads when a user visits that page. *Ultramercials***—Interactive online ads that require user interaction to reach the intended content.
Information Intermediary	*Advertising Networks*—A service feeding banner ads to a network of member sites, thereby enabling advertisers to deploy large marketing campaigns. Ad networks collect data about Web users that can be used to analyze marketing effectiveness. *Audience Measurement Service*—On-line audience market research. *Incentive Marketing*—Customer loyalty programs providing incentives to customers such as redeemable points or coupons for making purchases from associated retailers. Data collected about users are sold for targeted advertising.
Merchant	*Virtual Merchant*—A retail merchant that operates solely over the Web (also known as an "e-tailer"). *Catalog Merchant*—Mail-order business with a Web-based catalog which combines mail, telephone, and on-line ordering. *Click and Mortar*—Traditional brick-and-mortar retail establishment with a Web storefront. *Bit Vendor*—A merchant who deals strictly in digital products and services and, in its purest form, conducts both sales and distribution over the Web.
Manufacturer Direct	*Purchase Model*—A manufacturer that sells its products or services directly to the consumer. *Lease Model*—A manufacturer that finances the sale or rental of its products directly to the consumer. *Licensing Model*—A manufacturer, such as a software maker, that licenses its product directly to the consumer. *Brand-Integrated Content*—In contrast to the sponsored-content approach (i.e., the advertising model), brand-integrated content is created by the manufacturer itself for the sole purpose of product placement.

Table 2 Taxonomy of e-business models (continued)

Type	Model
Affiliate	*Banner Exchange*—Trades banner placement among a network of affiliated sites. *Pay-per-Click*—Site that pays affiliates for a user click-through. *Revenue Sharing*—Offers a percent-of-sale commission based on a user click-through in which the user subsequently purchases a product.
Community	*Open Source*—Software developed voluntarily by a global community of programmers who share code openly. Instead of licensing code for a fee, open source relies on revenue generated from related services like systems integration, product support, tutorials, and user documentation. *Public Broadcasting*—User contributor model used by not-for-profit radio and television broadcasting extended to the Web. The model is based on the creation of a community of users who support the site through voluntary donations. *Knowledge Networks*—Discussion sites that provide a source of information based on the sharing of expertise among professionals.
Subscription	*Content Service*—Provides text, audio, or video content to users who subscribe for a fee to gain access to the service. *Person-to-Person Networking Service*—Conduit for the distribution of user-submitted information, for example, individuals searching for former schoolmates. *Trust Service*—Membership association that abides by an explicit code of conduct and to which members pay a subscription fee. *Internet Service Provider*—Provides network connectivity and related services.
Utility	*Metered Usage*—Measures and bills users based on actual usage of a service. *Metered Subscription*—Allows subscribers to purchase access to content in metered amounts (e.g., numbers of pages viewed).

Source: Rappa, M. "Business Models on the Web," *Managing the Digital Enterprise* (http://digitalenterprise.org), May 2003.

exchange, pay-per-click, and revenue sharing programs.

Community model. The community model is based on user loyalty. Loyal users invest both their time and emotions in a business. Revenue can be generated based on the sale of ancillary products and services or voluntary contributions. The best known example of a community model is that of "open source" computing. The businesses that have emerged around open source products rely on revenue generated from related services such as systems integration, product support, tutorials, and user documentation. Another example is the traditional public broadcasting model, the listener or viewer-contributor method used in not-for-profit radio and television broadcasting. The model is based on the creation of a community of interested users who support the site through voluntary donations.

Subscription model. Users are charged a periodic daily, monthly, or annual fee to subscribe to a service. It is not uncommon for sites to combine free content with "premium" (i.e., subscriber only or member only) content. Subscription fees are incurred regardless of actual usage rates. Subscription and advertising models are frequently combined. Examples include content services, person-to-person networking services, trust services, and Internet service providers (ISPs).

Utility and hybrid models. The utility model is based on metering usage and constitutes a "pay as you go" approach. Unlike subscription services, metered services are based on actual usage rates. For example, an ISP may use a utility model, charging customers for connection minutes, though the subscription model is more common among ISPs operating in the United States. An interesting hybrid model on the Web, the metered subscription, allows subscribers to purchase access to content in metered portions, such as the number of pages viewed.

Metering customer usage is one characteristic that figures prominently in the utility business model and sets it apart from other models. But utilities in the off-line world are not limited to the approach of metering usage (see Table 3 and Figure 1). One example is residential telephone services. For some time, the so-called "plain old telephone system" (or POTS) adopted a combination of metered usage for long distance services, a subscription model for local calling services, and a lease model for the usage of telephone equipment (though nowadays equipment is typically purchased outright). Under a subscription

Table 3 Business models of utility services

Type of Service	Business Models
Water	Metered usage of service
Electricity	Metered usage of service
Common Carrier Transportation	Basic pay-as-you-go fare for one-way or roundtrip service; subscription for commuter service
Telephone:	
POTS	Subscription for local service; metered usage of long distance service; equipment is leased or purchased
Cellular	Subscription with usage limits; metered usage in excess of the subscription limit; equipment purchased or bundled with subscription
Radio and Television Broadcasting:	
Terrestrial	Advertiser-sponsored, community-sponsored
Satellite	Subscription with basic package and premium services Lease or purchase equipment
Cable	Subscription with basic package and premium services Pay-per-view for special event programming and movie selections Leased equipment is bundled with service
Internet Access:	
DSL	Subscription for unlimited ("always on") service Leased equipment is bundled with service
Cable	Subscription for unlimited ("always on") service Leased equipment is bundled with service
Dial-up	Subscription for limited service or metered usage based upon connection time Equipment is purchased

model, users pay a flat rate for monthly service regardless of actual usage levels.

Cellular phone services have adopted yet another combination of the subscription and utility models. A monthly subscriber fee for both local and long distance service is tied to a maximum level of usage (i.e., connection minutes), beyond which usage is metered and billed accordingly. The subscription may come with a minimum-length service contract, and may also include equipment as part of the agreement. The popularity of the cellular business model has recently led telephone service providers to consider the adoption of a similar approach with the introduction of a flat-rate subscription for both local and long distance calling services bundled together.

Radio and television broadcasting offers another example of how different business models can be combined in a service that exhibits characteristics similar to a utility. Terrestrial broadcasting services have typically depended on sponsorship in the form of commercial advertisements that are interspersed with programming. There is also a form of terrestrial broadcasting, which is publicly sponsored, that could be classified with community business models. Satellite and cable broadcasters use a subscription model in which the user is charged for a basic package of bundled services and can also choose from a menu of premium content.

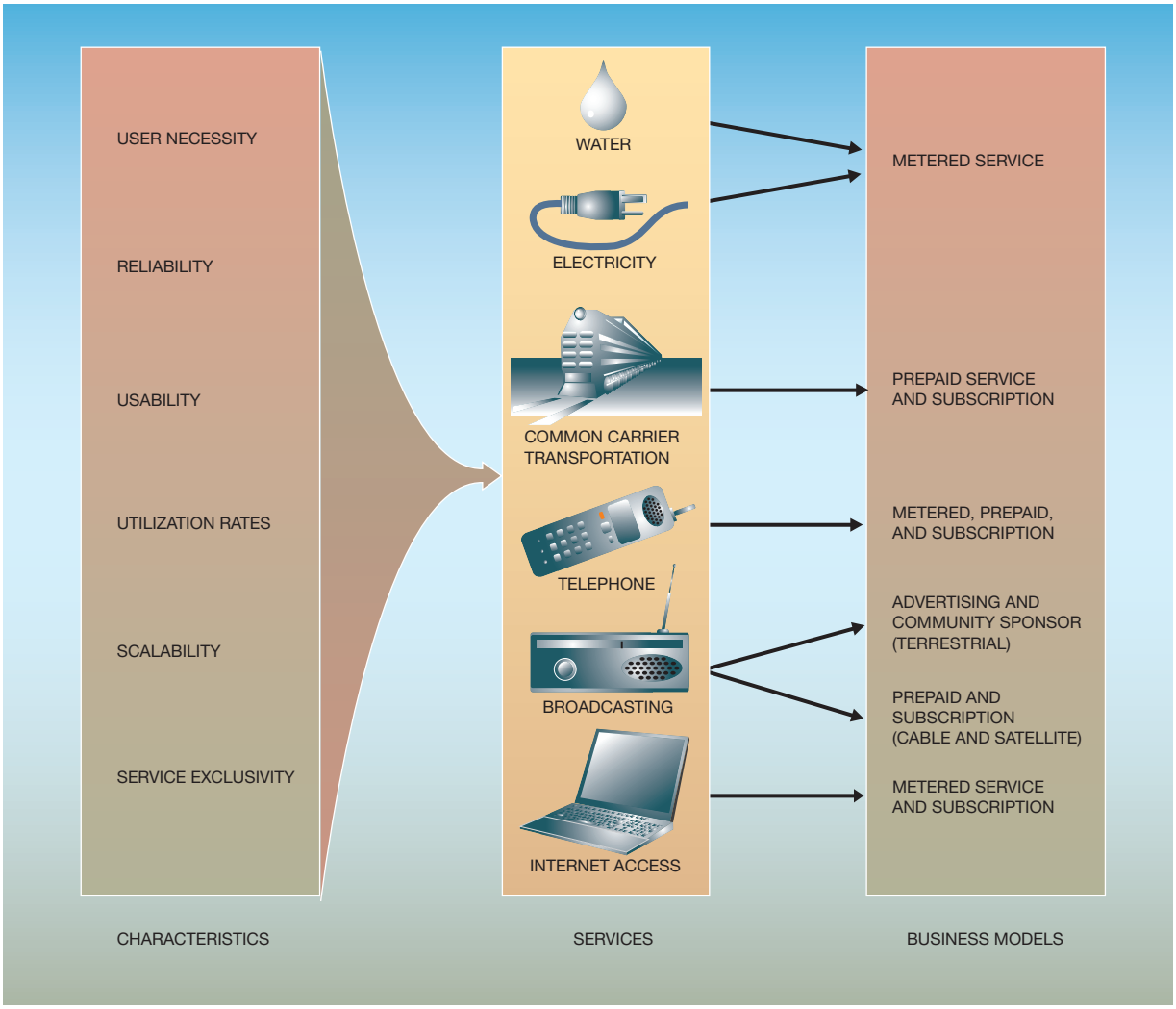
Internet access services provide an example of how business models adapt to technological changes in how a service is provided. Early in the commercial rollout of the Internet, services offered dial-up access by using a combination of business models not unlike the telephone service on which that access depended. A user could pay according to actual usage or pay a subscriber fee for limited (and later unlimited) service. With the advent of residential broadband services, Internet access providers offer "always on" service using a subscription model.

Utility computing

Recent projections from IBM have envisioned utility computing as an integral part of the future of information technology. IBM Global Services provides the following definition:

> Utility computing is the on demand delivery of infrastructure, applications, and business processes in a security-rich, shared, scaleable, and standards-based computer environment over the Internet for

Figure 1 Utility-service characteristics and business models



> a fee. Customers will tap into IT resources–and pay for them–as easily as they now get their electricity or water.[12]

In a recent interview, Irving Wladawsky-Berger, General Manager for e-business on demand, said:

> The idea of accessing computing and data without having to own the computers probably is even older than the 1980s because at some level that's what time-sharing computing was about, and that probably was back in the 1970s. It's a very nice thought, which comes from watching other successful utilities. You know, people are used to the fact that they use electricity, they turn on water faucets, they use the telephone, they get access to television, and in none of these cases do they have to own the equipment that generates electricity. They don't have to own the water supply, or the telephone switches, or the broadcasting companies. Everybody says, "Gee wouldn't it be nice if I can similarly plug in to get access to applications and information?"[13]

Recent discussions of computing have begun to point to its similarity to conventional public service util-

ities.[14,15] To be sure, the analogy can be taken too far, but it is instructive to examine the general service requirements of utilities in the context of computing to better gauge where business models may be heading in the future. It may very well be that current business models based on purchasing, leasing, and licensing of products will ultimately give way to a utility model of computing based on subscriber fees and metered services.

The importance of computers for business has become an indisputable fact. Computer and network services are an end-to-end component of many business processes. To be without service is not merely an inconvenience; it is a potential financial disaster. It should therefore come as no surprise that businesses both large and small have come to view computers as a necessity in the same manner as they might view utility services. Furthermore, it is likely that the reliance of individual users on computing services will soon rival if not surpass their reliance on public utilities like the residential telephone as an essential service.

As reliance on computing grows, so will the expectation among users for reliable service. Improvements in computer reliability have been made, but there is still much work to be done. Software remains a source of instability in the heterogeneous computer environments that exist today. As software has become more powerful, it has also become more complex in terms of the underlying code. This complexity creates a major challenge in engineering reliable software even under benign conditions, and this is exacerbated by the unrelenting security threat to computer networks. Current approaches to security, which can rely heavily on patching software after it is deployed, may be severely flawed as a method for achieving high reliability.

The growth of personal computers and the Internet have made computing a mainstream activity. Today the computer user population cuts across a wide spectrum in terms of age, education, and other demographic dimensions. Fortunately enormous progress has been made to improve computer usability. The user interface for personal computing is good evidence of success in making computers easier to use. Both hardware and software makers are moving quickly in the direction of creating products with "plug and play" convenience. But even so, computers have not yet achieved the goal of becoming as simple to use as a common household appliance. Advances in usability have been offset by the rapid technological innovation that has kept designs and standards in flux.

The substantial investment made in computing infrastructure has spurred an interest in increasing the overall rate of its utilization. Current progress in technology is making this possible. An example is the virtualization of server and storage capacity and the advent of grid computing supported by open standards. Grid computing will ultimately present utility-computing service providers with capacity planning issues similar to those faced by public utilities. Managing peak demand and the economical utilization of capacity will require incentives to modify usage patterns. This will favor the adoption of metered usage as a core element in the business model for utility computing.

It is likely that the reliance of users on computing services will soon rival, if not surpass, their reliance on public utilities.

To the extent that technology enables computing services to be scalable, the economies of scale typical of public utility services should also apply to utility computing. The benefits may come on several fronts, because the fixed cost of services of the utility can be amortized across a larger population of users, thereby reducing the unit cost per user. Utility computing may also benefit from the ability to retain the necessary skilled workforce to manage and maintain computing services in a way that is difficult for small and medium-sized enterprises. This may be of particular importance in dealing with network security, a field where there is a scarcity of talent.

One manner in which utility computing may differ from a public utility is with respect to service exclusivity. The trend in recent decades has been to deregulate and encourage competition in the public utility sector, for example, telephone and electricity service. It may be too soon to judge the overall effect deregulation has had on the provision of public services. At least some of the challenge of deregulation is the result of the tumultuous transition from a regulated environment. Utility computing may ben-

Table 4 Factors favoring subscription versus metered utility model for computing services

Metered Model	Subscription Model
Usage measures are easy to define, monitor, and verify	Usage measures are difficult to define, monitor, or verify
Strong managerial controls on usage patterns	Weak managerial controls on usage patterns
Commoditized, low value-added services	Proprietary, high value-added services
Favored by cost-conscious users with an ethic to conserve resources	Favored by users who are less conscious of resource costs and the need to conserve
Easy to forecast resource usage patterns	Hard to forecast resource usage patterns

efit from an opportunity to grow in a highly competitive marketplace. To the extent that technology evolves in a direction that permits competition in the provision of services, such as with open network protocols, the need for service exclusivity may be lessened. However, it should be expected that de facto technical standards and competitive advantages among enterprises might eventually lead to the emergence of dominant (if not exclusive) utility computing service providers.

It is interesting to speculate on what shape the utility-computing business model might take in the foreseeable future. The provision of computing services presents a matrix of opportunities that goes well beyond any comparison to traditional public utilities like electricity. Although it may be technically feasible to meter some kinds of computing services, there remains the question of which services to meter and how this can be done. At the level of computing infrastructure, it is possible to envision the metered usage of CPU resources, for example. At the application layer, there is already a move away from a pure license model toward subscription-based services. It is also conceivable that some kinds of applications could be adapted to a metered usage model, or a combination of subscription and metering. Lastly, it may be advantageous to meter computing services based on the completion rate of discrete business processes, such as the number of customer transactions.

From the customer's point of view, the business logic of metering usage can be compelling: one pays only for what one uses. This is something that IT managers who are faced with an escalating cost-of-ownership can appreciate. But the initial move away from the ownership model to the utility computing model will be hampered unless there exist clear operational measures of the underlying demand function and, therefore, a way to determine the cost to an organization when the meter is turned on. Depending on the kind of metering employed, end users may have to adjust their computing habits to cultivate an ethic of resource conservation and be mindful of the usage costs they incur. It is likely that commoditized services will be more easily adapted to a metered usage model.

In comparison, the subscription model provides users with more flexibility, and provides managers with a more accountable, if intermediate, approach on the path toward utility computing. A subscription-based service is more amenable to high value-added proprietary services that require more elaborate service level agreements. In a situation where there remains uncertainty about the upside demand for services or where there are weak controls on usage levels, we are more likely to see adoption of a subscription approach. Table 4 provides a summary comparing the factors that may favor deployment of a subscription model versus those that favor metered usage of computing services.

Conclusion

A vision of the future of computing services based on the utility-computing business model has already begun to take shape. Application service providers, managed services, and hosting are an increasingly common part of the computing landscape. Users have come to depend on computers and have high expectations of their reliability; they look toward a day when the use of computers matches the ease of other everyday appliances. Furthermore, the provision of computing services is increasingly driven by economies of scale and the effective utilization of resources.

The kind of utility-computing business model that will find favor with the customer remains to be seen. Already large enterprise customers are taking the first steps toward a model based on multiyear subscription contracts. However the metered use of com-

puting services is a significant leap from the current model of purchasing or leasing computer hardware, accompanied by software licensing.

Clearly, there are trade-offs involved in migrating to new business models. Customers must be convinced of the tangible benefits in making a change. What is ceded in terms of ownership and control must be more than modestly offset by advantages in procuring computing services, such as flexibility, speed of deployment, and cost savings.

Acknowledgments

The author is grateful to Meeta Yadav for her research assistance, John Killela for his encouragement, and the helpful comments of three anonymous reviewers.

**Trademark or registered trademark of priceline.com Incorporated, Google Inc., Ultramercial, Inc., or Overture Services Inc.

Cited references

1. *A Nation Online,* U.S. Department of Commerce (February 2002).
2. *Surveying the Digital Future,* The UCLA Internet Report, UCLA Center for Communication Policy (January 2003).
3. J. B. Horrigan and L. Raine, *Counting on the Internet,* The Pew Internet and American Life Project (December 2002).
4. V. Berstis, *Fundamentals of Grid Computing*, IBM Redbook (October 2002).
5. M. Milenkovic, S. H. Robinson, R. C. Knauerhase, D. Barkai, S. Garg, V. Tewari, T. A. Anderson, and M. Bowman, "Toward Internet Distributed Computing," *Computer*, **36,** No. 5, 38–46 (May 2003).
6. I. Foster, C. Kesselman, J. M. Nick, and S. Tueckel., "The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration," http://www.globus.org/research/papers/ogsa.pdf (June 22, 2002).
7. M. Rappa, "Business Models on the Web," *Managing the Digital Enterprise,* http://digitalenterprise.org (May 2003).
8. M. Vitale, P. Weill, and J. Ross, "From Place to Space: Migrating to Profitable Electronic Commerce Business Models," Center for Information Systems Research Working Paper No. 324, Massachusetts Institute of Technology (June 2002).
9. A. Osterwalder and Y. Pigneur, "An e-Business Model Ontology for Modeling e-Business," *Proceedings of the 15th Bled Electronic Commerce Conference,* 1–12, Bled, Slovenia (June 2002).
10. P. Timmers, "Business Models for Electronic Markets," *Electronic Markets*, **8,** No. 2, 3–8 (1998).
11. P. Bambury, "A Taxonomy of Internet Commerce," *First Monday*, Issue 3.10 (1998).
12. *FAQs: IBM e-Business On Demand,* IBM Global Services (March 2002), http://www-1.ibm.com/services/strategy/files2/FAQ.ondemand.pdf
13. R. Jernigan, "Interview: Irving Wladawsky-Berger Explains Grid Computing," *IBM Linux Portal,* http://www-1.ibm.com/linux/news31.shtml (March 2002).
14. *e-Business On Demand; The Next Wave of IT Services,* IBM Global Services, http://www-1.ibm.com/services/ondemand/files/nextwave.pdf (2002).
15. L. Siegele, "At Your Service," *The Economist,* May 10, 2003.
16. T. Eilam, K. Appleby, J. Breh, G. Breiter, H. Daur, S. A. Fakhouri, G. D. H. Hunt, T. Lu, S. D. Miller, L. B. Mummert, J. A. Pershing, and H. Wagner, "Using a Utility Computing Framework to Develop Utility Systems," *IBM Systems Journal* **43,** No. 1, 97–120 (2003, this issue)

Accepted for publication September 5, 2003.

Michael A. Rappa *North Carolina State University, Box 7229, Raleigh, NC 27695 (michael_rappa@ncsu.edu).* Dr. Rappa is the Alan T. Dickson Distinguished University Professor of Technology Management and Director of the Open Courseware Laboratory at North Carolina State University. Prior to joining NCSU, for nine years he was a professor at the Massachusetts Institute of Technology. In addition to overseeing the Open Courseware Lab, Dr. Rappa is co-director of North Carolina State's E-Commerce initiative. He is the creator of the award-winning Web site, Managing the Digital Enterprise, digitalenterprise.org, one of the Web's most extensive and widely used sites devoted to understanding the digital world and a definitive source for research on Internet business models.