The Costs of Cloud Migration

Omer Rana, Cardiff University

The cost of outsourcing computing infrastructure requires consideration not only of potential savings in operational and capital expenditures, but also of human and management costs.

Cloud computing is one of the most potent examples of how we can use computing as a utility. The ability to outsource computing infrastructure to one or more providers, with varying levels of trust, has allowed various companies to successfully adopt the cloud computing utility model. Smaller companies unable to afford an in-house computing infrastructure (and operational support for maintaining and managing such an infrastructure) are often cited as potential benefactors of this model. By migrating to a cloud infrastructure, such companies can save operational expenditures and focus on their core business rather than their computing systems.

However, the actual costs, especially for long-term use of an outsourced computing infrastructure, can be unfavorable for smaller companies. Of course, human issues can also influence cloud migration decisions, such as concern about job security for existing in-house administrative staff and lack of understanding of how the provider operates. By outsourcing, many companies surrender important systems management skills that they could have developed in house.

Understanding the true cost of outsourcing infrastructure therefore requires more detailed consideration than many organizations undertake when performing cost-saving analyses to decide whether cloud migration would benefit them.

A Multicriteria Economic Perspective

Taking a wider, multicriteria economic perspective is essential to ensuring that sustained, long-term use of cloud systems makes sense. Such analyses also offer important insights into which part of the local system should be moved to an external provider and which should remain in house. In this context, it's also useful to understand the organizational changes that cloud computing would generate. For instance, how would departments within a company deal with "pay-as-you-go" pricing?

Economic models therefore play an essential role in determining whether such outsourcing is likely to benefit the company. They strongly influence the decision of many companies as to whether they'll migrate (either partially or fully)

 2325-6095/14/$31.00

their infrastructure or services to a cloud provider. Such decisions factor in issues of pricing/cost, reputation/trust, performance/availability, energy savings, and security and privacy. Each of these decision factors impacts both shorter-term revenue and cost savings, and longer-term reputation and strategic operation.

Cloud providers also need to estimate the cost of provisioning infrastructure and services to clients, accounting for their own operational and capital expenditures, as well as potential reputation concerns (such as how potential clients perceive them in terms of reliability and their ability to deliver what they advertise) that impact their long-term survivability in the marketplace. Energy costs are increasingly important in this equation for many cloud providers and have influenced where they build their datacenters as well as potential alliances with energy providers offering special pricing.

There is also often a cloud supply chain, in which a single company uses services that are provisioned by others (in various service mashups), as well as associated dependencies within the supply chain. For instance, a company might run its own website but outsource storage to an infrastructure-as-a-service (IaaS) provider, establishing mutually beneficial service-level agreements (SLAs) that provide financial security for the company running the website (allowing them to establish penalty clauses that could lead to crediting customers in case of unavailability, for instance). Are users who access such websites fully aware of the different providers in the supply chain? Do service providers fully disclose their dependencies within their supply chains to their users/customers? Brokers play an important role in establishing these supply chains, "matchmaking" service requests to providers based on factors such as cost, operational history (for example, uptime and availability), and user feedback.

Pricing and Usage Models

Understanding how external infrastructure and service platforms should be compared also remains a challenge for companies. Difficulty arises when cloud providers use different names/terms for computing and storage resources, or bundle provider-specific services, making an overall comparison of providers a nontrivial process. Today, a limited number of providers dominate the market, offering users a range of configurable options.

Understanding how resource requirements can map to products from such providers, often available in a range of pricing bands (current versus older instances) and market models (spot market versus reserved instances, and so on), often requires input from economic and technical experts working in collaboration. Standards (see the "StandardsNow" column in this magazine) play an important role in creating suitable terminology that can be shared across providers. In the research community, there is also significant interest in providing auction-based models for improving utilization of spare capacity in the cloud market. Simulation-based approaches are generally used to demonstrate the benefit of these auction-based models and how they can improve utility for both consumers (cheaper resources) and providers (increased utilization of an otherwise rarely used resource). However, in practice, cloud providers have been reluctant to adopt auction-based models. Why are most cloud providers unwilling to offer underutilized capacity in an auctions market?

The multitenancy nature of cloud computing is often a leading reason for revenue generation by cloud providers.

The multitenancy nature of cloud computing is often a leading reason for revenue generation by cloud providers. By sharing common aspects of an infrastructure (using virtualization technologies) across multiple users, providers can benefit from economies of scale (and management efficiency). However, resource sharing across users also provides a key limitation if used inefficiently or incorrectly. Interference between virtual machine (VM) instances, potential data leakage (due to "dirty disks"), and VM/hypervisor "escape"—that is, when a VM goes beyond its defined boundaries and interacts directly with the operating system (and other VMs)—could negatively affect the potential cost benefits for customers. It's important to understand how such data hosting risks can be quantified and presented to users (and providers), so users can factor in these risks when choosing a cloud provider.

Existing cloud providers generally require users to register for their services using a credit card, and the number of instance hours used are charged to that card. With the emergence of digital currencies such as Bitcoin, cloud providers might begin offering exchangeable credit schemes. With this approach,

users could purchase several instance tokens from a cloud marketplace and redeem them at a number of different providers. The mapping between token value and number of instance hours received could vary depending on the popularity (and demand) for offerings from a particular provider. More popular providers could charge more tokens for their services than less popular providers, thereby creating a marketplace driven by supply-demand principles. A token exchange, in which unused tokens are auctioned to potential bidders, could create a dynamic marketplace for services/resources.

Another example is the use of provision-point (assurance) cloud provider contracts, in which members of a group pledge to contribute to an action if some prespecified threshold condition is met. Once the threshold point is passed, the action is taken based on the advertised capability (service or product is delivered); otherwise, no party is required to perform the action and any fees paid are refunded. Examples of this include offers from "daily deal" websites (making special-price offers of products/services that must be accepted within a limited time frame). Such contracts allow cloud providers to reduce operational expenditure through economies of scale. It's essentially a form of advance best-effort lease (that is, the provider can't guarantee in advance that a resource will be available). The consumer requests a resource to use in the future, but it is not guaranteed. The provider only delivers the resource if it can benefit from having a certain minimal number of consumers.[1]

> Distributed clouds extend the marketplace beyond the datacenter to devices owned by individuals or consortia.

Emerging Issues

Recently, interest in the creation of cloud "havens" (that is, cloud computing service providers operating in countries with lax or nonexistent security and privacy regulations) creates the potential for the coexistence of multiple marketplaces (with varying degrees of guarantees provided on the privacy of the stored data). Understanding how these cloud havens impact a cloud marketplace, how they influence the operation of cloud providers, and how customers perceive them remain interesting research questions.

Interest in "distributed clouds," which attempt to integrate datacenter and edge device capabilities, along with the potential of using software-defined network models for accessing network components such as routers and switches (using GENI OpenFlow, for instance), opens up new possibilities for cloud economics. The marketplace can now extend beyond the datacenter to devices owned by individuals or consortia, along with access to services made available by backbone network providers. Understanding how resource sharing and revenue models can factor into such multilayer capability remain important research challenges for the future.

The highly configurable nature of cloud computing—especially software-as-a-service (SaaS) offerings—suggests the potential for new business models that have yet to be fully realized. According to a Gartner report, global spending on SaaS will likely reach US$250 billion by 2017.[2] How much current SaaS offerings will need to change to reach such spending targets is unclear. Currently, SaaS business models are often similar to traditional models for selling software products (or licensing), and customers have limited ability to negotiate or alter these models. Existing SaaS vendors also limit their products' configurability, requiring clients to use standard capabilities offered by the vendor. This limits the benefit to the client of using the SaaS capability. Delivery of software capability through the Internet (hence, outsourcing deployment and hosting) could offer a range of potential capabilities, such as

- combining software hosting and development as a key part of the business offering;
- establishing long-term strategic relationships with customers through negotiated provisioning agreements; and
- providing subscription-based pricing that attempts to understand how consumers use the system, and how providers can use data from such usage to enhance their offerings.

How companies integrate SaaS capability into their existing in-house systems also remains an important question. For instance, are SaaS offerings primarily obtained for services that are seen as noncritical for a company's operation? Of course, it is useful to remember that cloud service consumers aren't just companies, but also individuals. Should business models for consumers differ from those for providers? How different should they be?

Call for Submissions

Cloud Computing magazine seeks submissions in a variety of areas, covering both micro- and macro-economic issues. Authors are encouraged to contribute articles demonstrating novel thinking on how economic models adapted from other domains could improve the use of multilayered cloud systems, and how novel resource provisioning strategies could lead to the development of new economic models. Authors are also encouraged to report on the practical use of economic models for provisioning cloud services/infrastructure and their experiences using pricing/economic models from commercial providers.

Cloud Computing magazine also seeks articles covering aspects of risk, especially risk from an economic perspective, given new cloud providers (some of which use services from more established companies, such as Amazon and Google). Risk-assessment strategies are often influenced by criteria that have a particular bearing on a company's operation (because a risk-versus-opportunity assessment is often needed). Different users are therefore likely to place varying degrees of emphasis on particular factors that impact their operation.

Brokerage-based risk assessment also remains an important challenge. Here, intermediate brokers can consider concerns relevant to particular users or companies and subsequently compare these with offerings (at particular price bands) from various cloud providers. It's also important to note that an organization's ability to understand such risk factors can change with maturity—that is, how long they've been using cloud services.

These are still early days for the cloud computing research and development community—in particular, how this community perceives and uses utility computing. The IEEE/ACM "Utility and Cloud Computing" conference, for instance, is only in its seventh year, compared with conferences in areas such as parallel computing that have been occurring for decades—for example, the International Conference on Parallel Processing is now in its 43rd edition. There is still plenty of room for innovation. With improved understanding of how cloud computing systems and services are used in practice, intermediate (brokerage) organizations can find numerous opportunities for interacting with users and cloud providers. Increasing interest and adoption of cloud standards can also be an important catalyst for generating a more sustainable cloud market.

References

1. O. Rogers, "Improved Public Cloud Capacity Planning through the Sale of Options, Forwards and Provision Point Contracts," PhD thesis, Dept. of Computer Science, Bristol Univ., UK, 2013.
2. Gartner, "Forecast: Public Cloud Services, Worldwide, 2011-2017, 4Q13 Update," 26 Dec. 2013; www.gartner.com/doc/2642020/forecast-public-cloud-services-worldwide.

OMER RANA *is a professor of performance engineering in the School of Computer Science and Informatics at Cardiff University. He also currently acts as an advisor to CBNine, a company specializing in cloud computing for the architecture, engineering, and construction sector. His research interests include high-performance distributed computing, data analysis/mining, and multiagent systems. Rana has a PhD in neural and parallel computing from Imperial College (London University). He is a member of the* IEEE Cloud Computing *editorial board and a member of IEEE. Contact him at o.f.rana@cs.cardiff.ac.uk.*