



Toward an Open Cloud Marketplace

Vision and First Steps

Azer Bestavros and Orran Krieger, *Boston University*

The Open Cloud Exchange (OCX) is envisioned as a public cloud marketplace in which many stakeholders, rather than just a single provider, participate in implementing and operating the cloud. This ecosystem would enable innovation from the broader academic and industry communities, resulting in a much healthier and more efficient cloud marketplace.

The notion of computing as a commodity that can be bought and sold isn't new. As early as the 1960s, luminaries such as John McCarthy¹ and Douglas Parkhill foresaw the potential for computation to "someday be organized as a public utility"² that paralleled the electricity industry and its complex set of marketplace players. In recent years, the advent of virtualization and the emergence of many public cloud offerings at various levels of the "cloud stack" – that is, infrastructure, platform, and software as a service (IaaS, PaaS, and SaaS) – have fueled significant interest and new research in mechanisms for supporting the cloud marketplace. Yet, much of this research remains "academic," given provider-imposed constraints that make it difficult, if not impossible, for customers to leverage such innovation and benefit from the efficiencies of a truly open marketplace. This article overviews this status quo and outlines a path forward to create a truly open marketplace that fosters competition and innovation.

The Challenges of Public Clouds

Today's public cloud is dominated by the relatively few IaaS providers willing and able to make the massive capital and automation investments needed to create an at-scale offering. Only a small number of these providers have access to operational data about their clouds, have rich insights into the applications they

run, can perform at-scale experiments on large-scale computing infrastructure, and have a sufficiently large enough share of the cloud market to implement sustainable economic operation models. This natural oligopoly stifles innovation and limits research, mainly owing to three main aspects. First, these clouds are largely homogeneous and are operated by a single provider. Second, the provider alone has access to the operational data. Third, rather than competing solely at the infrastructure level, IaaS providers operate higher-level services as well.

The first two aspects mean that research and innovation is limited largely to cloud providers who tend to design their own hardware and develop their own software, thus sacrificing compatibility with private clouds (those internal to large companies). In the long run, if only a handful of major providers continue to dominate the public cloud marketplace, then any innovation can only be realized through one of them. Thus, even if (say) a new hardware platform has huge value for a user community, no big provider is likely to deploy it unless it has broad utility.

This single-provider model results in strong homogeneity in cloud providers' offerings. This not only limits research but also results in a monoculture that's susceptible to security threats. Security concerns also arise because public clouds are designed under the assumption that the cloud operator is fundamentally trusted. No documented technologies or policies keep a

The OCX as a Platform for Value-Added Data Services

Data collection is becoming “standard operating procedure” for businesses and organizations, alike. While a given dataset (or stream) is typically collected for a particular purpose, such data could be valuable for other purposes as well — especially if combined creatively with data from other sources. This is precisely the value proposition from big data analytics. For such creative uses of (big) data to flourish, however — and for the associated economic development to materialize — we must develop open platforms that provide

the incentives and means (including secure/private access/mining platforms that are currently only of academic value) to “monetize” data assets. In an Open Cloud Exchange (OCX) setting, data as well as computational resources become assets that value-added providers (VAPs) can combine in creative ways. More importantly, lowering the barrier to entry for new value-added data services, and consequently increasing competitiveness among such services’ providers, will ensure that data assets’ true market value is realized.

cloud provider, or even a disgruntled employee, from introspecting on a customer’s network traffic, computers, or even private datasets. We argue that an open cloud would deliver some of the same long-recognized security benefits as has open source software development.³

The lock that cloud providers have on operational data not only limits third parties’ ability to innovate, but also prevents or at best significantly encumbers important services, such as those offered in other marketplaces by value-added resellers, aggregators, or auditors. Many recognize the lack of such services as a main hurdle to an even broader embrace of cloud computing in many economic sectors.

With exclusive access to design and operational data, cloud providers gain another significant advantage: they can offer highly competitive higher-level services such as big data analytics (see the sidebar, “The OCX as a Platform for Value-Added Data Services”) and other PaaS platforms and applications. A competitive advantage for one party (resulting in an effective monopoly) tends to drive researchers and innovators away from working on related problems. In addition, because each provider develops different higher-level services, such services will increasingly become incompatible. A customer dependent on a service from one cloud will find it difficult to move to another.

Given that each provider operates its own isolated, vertical silo, with no

internal competition, the economic models that providers have developed are designed to pull customers into a single cloud. For example, Amazon’s pricing of bandwidth into and out of its cloud incentivizes customers to put more and more of their computing resources in the cloud to avoid external bandwidth charges. Richer pricing models in more open marketplaces (such as those used to trade in electricity) have yielded enormous efficiencies that remain unexplored in today’s public clouds (see the sidebar, “The OCX as a Prerequisite for Rational IaaS Pricing”).

Although many predict that the public cloud model will eventually dominate all computing, its limitations have resulted in big customers (government, large enterprises, universities, and so on) developing their own private clouds. In fact, by most estimates, the private cloud market today is far more important than the public one. If we’re to realize the vision of public clouds, we must somehow find a way to engage a broad research community and allow the resulting innovation to be relevant.

The Open Cloud Exchange

The cloud could have as big an impact on our society as has the Internet. Some of the Internet’s foremost architects have argued that a key factor in its success has been that it accommodates the “tussle” between conflicting interests,⁴ which plays out in an arena that’s neither

unregulated — thus favoring only the largest actors — nor overconstrained, with a predetermined winner. This tussle is limited in today’s public cloud: few companies have the resources to deploy a competitive public cloud, and vendor lock-in imposes high costs on customers wishing to switch providers. Can we reproduce a dynamic similar to what fueled Internet innovation to support cloud computing? Can a broad community participate and innovate, as was critical to the Internet’s success?

We believe that the answer to these questions is yes. The public cloud can evolve into a marketplace in which many actors can compete and cooperate with each other, and innovation can flourish. We refer to the marketplace model we propose for the public cloud as the Open Cloud Exchange (OCX). In an OCX, many stakeholders, rather than just a single provider, participate in implementing and operating the cloud. This creates a multisided marketplace in which participants freely cooperate and compete with each other, and customers can choose among numerous competing services and solutions that are differentiated not only by the features they offer, but also by their applicability to particular applications, their offerings’ reliability, price, or security, and other characteristics.

Our vision and ongoing work setting up an OCX at scale will expose researchers to problems that are invisible (because they require access

The OCX as a Prerequisite for Rational IaaS Pricing

Today's vertically integrated public clouds offer simple fixed prices for a (virtual) computer. Although researchers have explored alternatives that can offer greater efficiency or encourage performance or power optimization in applications, none of this research has been relevant to large public clouds. In addition, the kind of economic models possible in other marketplaces — such as the electric grid — aren't possible in today's public clouds.

As an example, consider current cloud resource pricing. Although the operational costs providers incur for operating such resources depend on variables such as energy cost, cooling strategies, and aggregate demand, the pricing models extended to customers don't reflect these variables — at least not directly. This disconnect between supply and demand results in an economically inefficient marketplace. In particular, it doesn't provide incentives for customers to express workload

scheduling flexibilities that could benefit them as well as providers. In recent work,¹ we showed how a dynamic pricing model that leverages Shapley valuation can address this inefficiency by incentivizing customers to exploit any flexibility they might have as regards provisioning their workloads. In the absence of an Open Cloud Exchange (OCX), implementing such an approach will remain mostly an academic exercise.

In an OCX setting, every research and industry partner can specify how its service is priced. For example, some partners might charge a fixed rate for the resources used, whereas others could charge based on demand, power used, or specific workload constraints (such as colocation) or flexibilities.

Reference

1. V. Ishakian et al., "CloudPack: Exploiting Workload Flexibility through Rational Pricing," *Proc. 13th Int'l Middleware Conf.*, Springer, 2012, pp. 374–393.

to operational data available only to providers in today's closed cloud model) as well as to opportunities for innovations that are currently of only academic value (because researchers can't implement and experiment with them at scale). Note that the Internet stack and the Linux operating system are two examples of the innovation unlocked when the research community was able to contribute to open systems operating at scale.

The OCX Ecosystem

The OCX is a marketplace open to software, hardware, and services vendors who collectively participate in operating open-cloud solutions; we call these participants *exchange service providers* (XSPs). For example, the OCX might have several hardware companies as XSPs that would install their equipment in the OCX and make it available to other partners or customers on a pay-as-you-go basis. The OCX might also have numerous software companies as XSPs to provide and manage IaaS stacks or PaaS solutions that use equipment provided by other (not necessarily a fixed set of) XSPs. For example, an XSP could expose a simple IaaS abstraction, just like today's single-provider IaaS clouds,

that exploits whatever capacity is cheapest in other underlying IaaS offerings, or even within a single IaaS offering (such as Amazon Web Services Elastic Compute Cloud versus Spot instances). In the OCX, multiple such XSPs could compete, thus providing customers (or other XSPs) with a range of options. Finally, the OCX could have a number of solution providers as XSPs that offer customers tailored SaaS solutions.

From a customer's perspective, navigating the OCX marketplace with all its potential XSPs could be daunting. (Indeed, you could argue that one value of today's public clouds is that they offer a simple economic model and a few offerings that a cloud consumer can easily understand.) One way that other marketplaces have addressed this challenge is through third-party intermediaries, which we call *value-added providers*. Such VAPs offer new services, capabilities, or different pricing models by

- integrating resources and services from XSPs and other VAPs;
- acting as agents for multiple customers; or
- acting as brokers connecting customers with specific needs to

XSPs with resources that meet (or could be made to meet) those needs.

Thus, we can view VAPs as providing a simple consumption model for a particular class of application or customer, while spanning and integrating many lower-level services, thus hiding the complexities of the OCX marketplace from customers.

As we envision them, VAPs are third-party intermediaries leveraging and taking to a different level emerging XSP offerings. In today's cloud marketplace, we can view offerings such as Heroku, Cloud Foundry, Rightscale, Engine Yard, OpenShift, packaged Hadoop, and StarCluster as XSPs that provide customers with new or specialized abstractions. We believe that the role of intermediaries and the opportunity for innovation in their creation is even more important in an open cloud, which has a rich diversity of underlying services, and thus presents an opportunity for intermediaries to go well beyond the typical IaaS, PaaS, and SaaS landscape. In this sense, VAPs are distinct from XSPs because they offer services unavailable in today's highly stratified public cloud marketplace. Among other examples

of potential capabilities, we single out third-party auditing, brokerage, aggregation, and auctioneering.

A salient feature of our envisioned OCX ecosystem is that it provides an equal-opportunity platform not only for production, but also for experimental solutions. In particular, while well-established industry players might provide production-level services based on private cloud software that they've developed, researchers and aspiring software entrepreneurs can offer alternative services on an experimental or trial basis. Customers are thus able to compare and contrast, and ultimately choose the solutions that fit their needs best. Within the OCX marketplace, we can view a researcher or innovator as yet another XSP or VAP, thus offering a path for evaluating and adopting new solutions.

The OCX's operation will be subject to rules and protocols that ensure a level playing field. As an open platform, the OCX will make all operational data available to stakeholders (unless otherwise spelled out in specific provisions of a service-level agreement between a provider and a consumer). This includes the operational workload data as well as performance and problem reports. This data is critical to enable research into both how clouds are used and the limitations of existing cloud hardware and software infrastructures. Moreover, it lets customers assess the value proposition of competing platform or solution providers. Finally, the availability of such data would let third-party auditors engage in this cloud ecosystem.

The OCX OpenStack Prototype

Although much of an OCX's value depends on its research and vendor partners, it must provide a core infrastructure that all marketplace participants share. This common set of services is necessary for multiple

partners to participate in the cloud, and for customers to move between one partner and another, if needed. We've developed a set of core functionalities that we've begun implementing in the OpenStack framework.

Automation

This is a cross-cutting concern for the OCX, as for any public or private cloud, because the economics depend on minimizing the need for human intervention in all but the most extreme situations. We single out two aspects of automation.

Orchestration. This includes configuring, monitoring, and upgrading the systems. OpenStack contains mature support for these functions, but will need integration and development for the use cases and additional services required for the OCX.

This is an ongoing task, given that OpenStack's components are constantly evolving, and the entry of new industrial participants in the OCX testbed will bring new services requiring integration.

Automated request generation, handling, and tracking. The OCX will need to handle a large class of requests automatically, dispatching most to the appropriate entity controlling the infrastructure and notifying appropriate users and providers of any problems that arise and resolutions thereof.

Network Operations Center (NOC) Software

The OCX needs its own NOC for its operators to control core infrastructure and services. This operations center will in turn provide information to providers controlling their own services in the shared cloud.

Scheduling

Infrastructure that different partners introduce into the OCX could be subject to specific constraints that must

be observed and satisfied by scheduling and provisioning services. For example, an academic institution introducing federally funded hardware into the OCX might impose constraints on its use – for instance, restrict it for use by federally funded research projects, or in support of specific subjects. Vendors could impose similar constraints to limit liability or access to their technologies.

Metering and Billing Infrastructure

The OCX will leverage the rich metering functionality recently added to OpenStack as well as the billing solutions available from third parties. However, these solutions must be integrated with core OCX services, and will likely require modifications, for instance, to handle scheduling constraints.

Shared Provisioning

With multiple administrative entities deploying OCX services, the OCX must have a layer of software that lets an entity provision part of the physical infrastructure, installing whatever OS and higher-level services it wants on that hardware without interference with or risk to other entities. This includes both production and experimental services. We're developing a prototype of such a service to let trusted administrators partition the infrastructure and distribute credentials to partners and aggregators who would then be able to provision appropriate subsets of the physical infrastructure.

Compatibility Testing

A key attribute of the OCX as a platform for innovation is that implementation of new services will need to be constantly incorporated into its offerings. Thus, at-scale testing of these new services will be necessary before they become broadly available to customers. This area is undergoing significant effort in both industry and academia, but additional

OCX-Enabled Research

Addressing cloud computing challenges will engage areas of computer science, engineering, and many other disciplines. It isn't our role to define the solutions, but rather pose some important questions. A production-level service comparable to today's public clouds is the platform needed to address these challenges; we don't need to wait to solve these problems before developing the OCX testbed.

Data-Driven Analytics

One core value proposition of the cloud is that it will enable the sharing of datasets and make public datasets broadly available. Can big data platforms perform well on top of existing virtualization technology? What networking technologies and assumptions (such as nonoversubscribed networks) are necessary? What storage services are necessary? What kind of provenance information should be built into storage services? At what levels should availability and fault tolerance be available?

Security

Can we protect users not only from each other but also from individual malicious providers? Can we distribute data or compute power across multiple providers without trusting any of them? How do we expose operational and performance data in a privacy-preserving fashion? Can we expose performance data while protecting vendor proprietary information? How can we federate across untrusting data centers managed by different institutions?

High-Performance Computing

Can clouds provide the high-performance computing capacity that traditional HPC clusters have been used for? What networking requirements are needed? Can we exploit a cloud's massive computational capacity to move toward a more interactive rather than batch HPC model?

Software-Defined Networking

A key enabling cloud technology is software-defined networking (SDN). How do we integrate across multiple heterogeneous

SDN solutions? Can we define an SDN approach so that third parties develop services? Can we use SDN techniques across heterogeneous federated clouds? Does the customer have to trust the SDN solution?

System Software

How do we deal with accelerators, such as GP-GPUs or Intel Phi processors, in virtualized platforms? How do we efficiently support high-performance remote direct memory access in a virtualized platform? Can we support highly elastic applications? Can we develop a lingua franca for the services available at different federated sites? Can we provide end-to-end lossless service-level agreements (SLAs) for distributed applications over circuit-switched wireless area networks? What are the key federation services we need to design? Examples include replicated storage across availability zones, appliance marketplace, content distribution, content ingestion, and so on. What control does each data center need to cede to the larger cloud?

Economics

How do we expose different services' characteristics to aggregators? Can we develop a lingua franca for this purpose? How do we represent other constraints, for example, capacity provided for specific customer segments? What regulation can or should we impose at the marketplace level to ensure free competition? How do we keep users from gaming the marketplace?

Domain-Specific Platforms

Most of cloud's value is in the innovation that will occur on top of it. We've already seen platforms such as CloudMan (<http://wiki.galaxyproject.org/CloudMan>) have a dramatic effect on particular domains. What turnkey services can we develop for other environments, such as fisheries, brain mapping, chemistry, and so on? What community services can we develop? What lower-level platforms and interfaces will enable these platforms?

work is required to support the OCX multiprovider environment.

OpenStack Services Directory

The services directory in OpenStack enables the identification of instances of compute, storage, and other services to fulfill a customer request. The current OpenStack services directory assumes a homogeneous cloud, with a single instance of each service type. We'll need to extend this directory to support multiple instances,

and will require a way to describe the different instances available. On top of this work, we must explore how characteristics such as pricing, SLA, performance, and security should be exposed, and create a service definition language for this purpose.

User Interfaces

The changes to the OpenStack services directory will make the platform's heterogeneity visible to programs. We must also make this

heterogeneity visible to users, along with higher-level information such as reputation, ranking, price information, and so on. A significant body of open source software packages is available for this purpose; we are evaluating candidates for integration and deployment in the OCX.

The OCX Testbed at Boston University

One problem that has crippled many areas of systems research in the

past is that researchers could only focus on applications to which they had direct access. For the cloud, the user community must include not only scientific researchers, but also companies and even individuals. These entities must be able to trust their computation and datasets to the platform. Hence, while research can occur at all levels – and while experimental services should be exposed – at-scale production-level services must also be available. To tackle this challenge, we’re designing and implementing an OCX testbed that would involve researchers with both vendors and customers in a production setting (see the “OCX-Enabled Research” sidebar for more on the many issues an OCX must address.).

Although the concept of an OCX marketplace might have merit, bootstrapping it and setting it up as an operational testbed is a major undertaking that requires a public/private partnership to succeed. In particular, an OCX testbed must provide value not only to researchers, but also to a broad community of technology vendors and users.

The initial vision for this partnership is being driven by a multi-university collaboration anchored in the Massachusetts Green High-Performance Computing Center (MGHPCC; www.mghpcc.org) and incubated at Boston University. The MGHPCC is a nationally unique model of collaboration that brings together industry, state government, and public and private universities around a single project – building and operating a research computing data center supporting the growing needs of five of the most research-intensive universities in Massachusetts: Boston University, Harvard University,

Massachusetts Institute of Technology, Northeastern University, and the University of Massachusetts.

To implement the OCX testbed, we’re inviting multiple technology vendors to set up their (primarily OpenStack) IaaS offerings on top of their own equipment racks, and to make such services available (either granted or on a pay-as-you-go basis) to the participating universities and regional users. Technology vendors are willing, and in fact eager, to participate in setting up the OCX testbed. Much of the computer industry’s interest is, fortunately, aligned with that of the research community in the OCX concept. Both parties are excluded from today’s public cloud marketplace. Major vendors’ efforts to stand up their own public clouds have been largely unsuccessful, partly because few customers are interested in locking themselves into another vendor-specific offering. At the same time, major strides were made in developing IaaS platforms for the private cloud that offer services similar to those needed to stand a major public cloud.

The OCX concept offers a path forward for deploying such services. The OCX testbed is also leveraging some shared MGHPCC research computing capacity, which would be managed as a cloud. The Commonwealth of Massachusetts is supporting the effort, with both funds for developing and operating shared services, and some capital funds.

As a regional initiative, the OCX testbed is giving us insights into both the model and underlying technologies of our envisioned OCX marketplace. Moreover, we have a


unique opportunity to house the OCX testbed in a modern, at-scale data center. □

References

1. S. Garfinkel, “The Cloud Imperative,” *MIT Technology Rev.*, MIT, 3 Oct. 2011.
2. D.F. Parkhill, *The Challenge of the Computer Utility*, Addison-Wesley, 1966.
3. B. Witten, C. Landwehr, and M. Caloyannides, “Does Open Source Improve System Security?” *IEEE Software*, vol. 18, no. 5, 2001, pp. 57–61; <http://dx.doi.org/10.1109/52.951496>.
4. D.D. Clark et al., “Tussle in Cyberspace: Defining Tomorrow’s Internet,” *IEEE/ACM Trans. Networking*, vol. 13, no. 3, 2005, pp. 462–475.

Azer Bestavros is a professor of computer science and director of the Hariri Institute for Computing at Boston University. His research interests include networking, distributed systems, and cloud computing. Bestavros has a PhD in computer science from Harvard University. He’s a member of ACM and IEEE, and former chair of the IEEE Technical Committee on the Internet. Contact him at best@bu.edu.

Orran Krieger is director of the Cloud Computing Initiative at Boston University. His research interests include cloud computing, operating systems, and distributed systems. Krieger has a PhD in electrical and computer engineering from the University of Toronto. He’s a member of ACM, IEEE, and Usenix. Contact him at okrieg@bu.edu.

 Selected CS articles and columns are also available for free at <http://ComputingNow.computer.org>.