Economics of the Cloud and Open Challenges

Dr. Prasenjit Chanak
Assistant Professor

Department of Computer Science and Engineering
Indian Institute of Technology (BHU), Varanasi-221005

Economics of the Cloud

- The main drivers of cloud computing are economy of scale and simplicity of software delivery and its operation
- In fact, the biggest benefit of this phenomenon is financial: the pay-as-you-go model offered by cloud providers
- In particular, cloud computing allows:
  - Reducing the capital costs associated to the IT infrastructure
  - Eliminating the depreciation or lifetime costs associated with IT capital assets
  - Replacing software licensing with subscriptions
  - Cutting the maintenance and administrative costs of IT resources

Open Challenges

- Still in its infancy, cloud computing presents many challenges for industry and academia
- There is a significant amount of work in academia focused on defining the challenges brought by this phenomenon
- This lecture highlights the most important ones: the definition and the formalization of cloud computing, the interoperation between different clouds, the creation of standards, security, scalability, fault tolerance, and organizational aspects

Cloud Definition

- As discussed earlier, there have been several attempts made to define cloud computing and to pro- vide a classification of all the services and technologies identified as such. One of the most comprehensive formalizations is noted in the NIST working definition of cloud computing
- It characterizes cloud computing as on-demand self-service, broad network access, resource-pooling, rapid elasticity, and measured service; classifies services as SaaS, PaaS, and IaaS; and categorizes deployment models as public, private, community, and hybrid clouds.
- The view is in line with our discussion and shared by many IT practitioners and academics
- Despite the general agreement on the NIST definition, there are alternative taxonomies for cloud services

Cloud Definition

- David Linthicum, founder of BlueMountains Labs, provides a more detailed classification, which comprehends 10 different classes and better suits the vision of cloud computing within the enterprise
- A different approach has been taken at the University of California, Santa Barbara (UCSB), which departs from the XaaS concept and tries to define an ontology for cloud computing
- In their work the concept of a cloud is dissected into five main layers: applications, soft- ware environments, software infrastructure, software kernel, and hardware
- Each layer addresses the needs of a different class of users within the cloud computing community and most likely builds on the underlying layers. According to the authors, this work constitutes the first effort to provide a more robust interaction model between the different cloud entities on both the functional level and the semantic level

Cloud Definition

- These characterizations and taxonomies reflect what is meant by cloud computing at the present time, but being in its infancy the phenomenon is constantly evolving, and the same will happen to the attempts to capture the real nature of cloud computing

Cloud Interoperability and Standards

- The current state of standards and interoperability in cloud computing resembles the early Internet era, when there was no common agreement on the protocols and technologies used and each organization had its own network
- Yet the first steps toward a standardization process have been made, and a few organizations, such as the Cloud Computing Interoperability Forum (CCIF), the Open Cloud Consortium, and the DMTF Cloud Standards Incubator, are leading the path
- Another interesting initiative is the Open Cloud Manifesto, which embodies the point of view of various stakeholders on the benefits of open standards in the field
- The standardization efforts are mostly concerned with the lower level of the cloud computing architecture, which is the most popular and developed

Cloud Interoperability and Standards

- In particular, in the IaaS market, the use of a proprietary virtual machine format constitutes the major reasons for the vendor lock-in, and efforts to provide virtual machine image compatibility between IaaS vendors can possibly improve the level of interoperability among them
- The Open Virtualization Format (OVF) is an attempt to provide a common format for storing the information and metadata describing a virtual machine image
- Even though the OVF provides a full specification for packaging and distributing virtual machine images in completely platform-independent fashion, it is supported by few vendors that use it to import static virtual machine images

Cloud Interoperability and Standards

- The challenge is providing standards for supporting the migration of running instances, thus allowing the real ability of switching from one infrastructure vendor to another in a completely transparent manner
- Another direction in which standards try to move is devising a general reference architecture for cloud computing systems and providing a standard interface through which one can interact with them
- At the moment the compatibility between different solutions is quite restricted, and the lack of a common set of APIs make the interaction with cloud-based solutions vendor specific
- In the IaaS market, Amazon Web Services plays a leading role, and other IaaS solutions, mostly open source, provide AWS-compatible APIs, thus constituting themselves as valid alternatives
- Even in this case, there is no consistent trend in devising some common APIs for interfacing with IaaS (and, in general, XaaS), and this constitutes one of the areas in which a considerable improvement can be made in the future

Scalability and Fault Tolerance

- The ability to scale on demand constitutes one of the most attractive features of cloud computing. Clouds allow scaling beyond the limits of the existing in-house IT resources, whether they are infrastructure (compute and storage) or applications services
- To implement such a capability, the cloud middleware has to be designed with the principle of scalability along different dimensions in mind—for example, performance, size, and load
- The cloud middleware manages a huge number of resource and users, which rely on the cloud to obtain the horsepower that they cannot obtain within the premises without bearing considerable administrative and maintenance costs
- These costs are a reality for whomever develops, manages, and maintains the cloud middleware and offers the service to customers
- In this scenario, the ability to tolerate failure becomes fundamental, sometimes even more important than providing an extremely efficient and optimized system. Hence, the challenge in this case is designing highly scalable and fault-tolerant systems that are easy to manage and at the same time provide competitive performance

Security, Trust, and Privacy

- On one side we need to decide whether to trust the provider itself; on the other side, specific regulations can simply prevail over the agreement the provider is willing to establish with us concerning the privacy of the information managed on our behalf
- Moreover, cloud services delivered to the end user can be the result of a complex stack of services that are obtained by third parties via the primary cloud service provider
- In this case there is a chain of responsibilities in terms of service delivery that can introduce more vulnerability for the secure management of data, the enforcement of privacy rules, and the trust given to the service provider
- In particular, when a violation of privacy or illegal access to sensitive information is detected, it could become difficult to identify who is liable for such violations. The challenges in this area are, then, mostly concerned with devising secure and trustable systems from different perspectives: technical, social, and legal

Organizational Aspects

- From an organizational point of view, the lack of control over the management of data and processes poses not only security threats but also new problems that previously did not exist
- Traditionally, when there was a problem with computer systems, organizations developed strategies and solutions to cope with them, often by relying on local expertise and knowledge
- One of the major advantages of moving IT infrastructure and services to the cloud is to reduce or completely remove the costs related to maintenance and support. As a result, users of such infrastructure and services lose a reference to deal with for IT troubleshooting
- At the same time, the existing IT staff is required to have a different kind of competency and, in general, fewer skills, thus reducing their value. These are the challenges from an organizational point of view that must be faced and that will significantly change the relationships within the enterprise itself among the various groups of people working together