Clouds at the Crossroads: Research Perspectives and Challenges

Ymir Vigfusson and Gregory Chockler
IBM Research, Haifa Labs

Despite its promise, most cloud computing innovations have been almost exclusively driven by a few industry leaders, such as Google, Amazon, Yahoo!, Microsoft and IBM. The involvement of a wider research community (both in academia and industrial labs) has so far been patchy without a clear agenda. In our opinion, the limited participation stems from the prevalent view that clouds are mostly an engineering and business-oriented phenomenon based on stitching together existing technologies and tools.

Here, we take a different stance and claim that clouds are now mature enough to become first-class research subjects, posing a range of unique and exciting challenges deserving collective attention from the research community. For example, the realization of privacy in clouds is a cross-cutting interdisciplinary challenge, permeating the entire stack of any imaginable cloud architecture.

The goal of this article is to present some of the research directions that are fundamental for cloud computing. We pose various challenges that span multiple domains and disciplines. We hope these questions will provoke interest from a larger group of researchers and academics who wish to help shape the course of the new technology.

Cloud Computing: an Architectural View

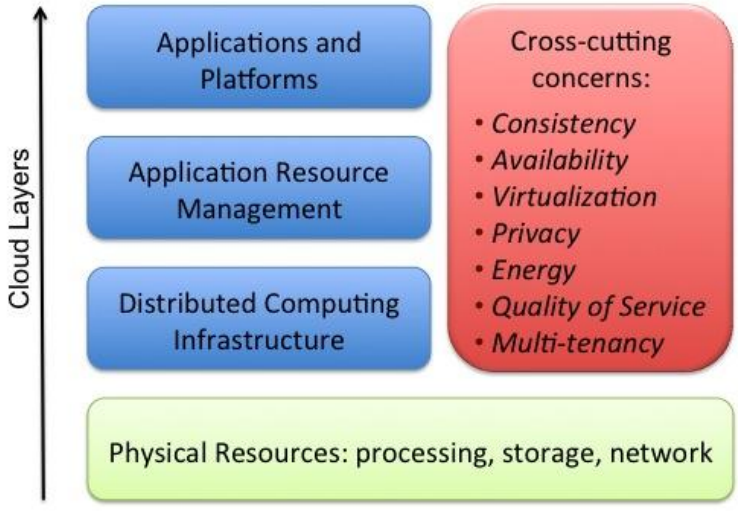

Figure 1: Internal Organization of a Typical Cloud System

Figure 1 provides a basic reference point for the internal organization of a typical cloud. The **physical resources** is simply a collection of machines, storage, and networking resources collectively representing the physical infrastructure of a data center(s) hosting the cloud computing system. Large clouds may contain some hundreds of thousands of computers.

The **distributed computing infrastructure** offers a collection of core services that simplify the development of robust and scalable services on top of a widely distributed failure-prone physical platform. The services supported by this layer typically include communication (such as multicast and publish-subscribe), failure detection, resource usage monitoring, group membership, data storage (such as distributed file systems and key-value lookup services), distributed agreement (such as consensus) and locking.

The **application resource management** layer manages allocation of physical resources to the actual **applications and platforms**, including higher-level service abstractions (such as virtual machines) offered to end-users. The management layer deals with problems related to the application placement, load balancing, task scheduling, service-level agreements and others.

Finally, we enumerate some **cross-cutting concerns** that dissect the entire cloud infrastructure. We will focus on three of these issues: energy, privacy and consistency, as well as the lack of standards, benchmarks and testbeds for conducting cloud related research.

Energy

Large cloud providers are natural power hogs. To reduce the carbon footprint, data centers are frequently deployed in proximity to hydroelectric plants and other clean energy sources. Such deployment has become easier: Microsoft, Sun and Dell have advocated putting data centers in shipping containers consisting of several thousand nodes at a time. Although multi-tenancy and the use of virtualization improve on resource utilization over traditional data centers, the growth of cloud provider services has been rapid and power consumption is a major operating expense for the large industry leaders.

A fundamental question is how and where we can reduce power consumption in the cloud, and at what cost. Here are three examples to illustrate potential directions.

Solid-state disks (SSDs) have substantially faster access times and draw less power than regular mechanical disks. The downside is that SSDs are more expensive and lack durability because blocks can become corrupted after 100,000-1,000,000 write-erase cycles. SSDs have made their way into the laptop market, the next question is whether cloud data centers will follow [1]. Can we engineer mechanisms to store read-intensive data on SSDs instead of disks?

Google has taken steps to revamp energy use in hardware, for instance by producing custom power supplies for computers which have more than double the efficiency of regular ones [2]. They even patented a "water-based" data center on a boat which

harnesses energy from ocean tides to power the nodes and uses the sea for cooling. How can we better design future hardware and infrastructure for improved energy efficiency, and how can we minimize energy loss in the commodity machines currently deployed in data centers?

In the same fashion as laptop processors adapt the CPU frequency to the workload being performed, data center nodes can be powered up and down to adapt to variable access patterns, for instance due to diurnal cycles or flash crowds. Some CPUs and disk arrays have more flexible power management controls than simple on/off switches, thus permitting intermediate levels of power consumption [21]. File systems spanning multiple disks could for instance bundle infrequently accessed objects together on "sleeper" disks [4]. More generally, how should data and computation be organized on nodes to permit software to decrease energy use without reducing performance?

Privacy concerns

Storing personal information in the cloud clearly raises privacy and security concerns. Sensitive data are no longer barred by physical obscurity or obstructions, instead exact copies can be made in an instant. Technological advances have reduced the ability of an individual to exercise personal control over his or her personal information, making it elusive to define privacy within clouds [6]. The companies which gather information to deliver targeted advertisements are working towards their ultimate product: *you*. The amount of information known by large cloud providers about individuals is staggering and the lack of transparent knowledge about how this information is used has provoked concerns. Are there reasonable notions of privacy that would still allow businesses to collect and store personal information about their customers in a trustworthy fashion? How much are users willing to pay for additional privacy?

We could trust the cloud partially, but implement auditing and accountability mechanisms. If privacy leaks have serious legal repercussions, then cloud providers would have incentives to deploy secure information flow techniques (even if they are heavy-handed) to limit access to sensitive data and to devise tools to locate the responsible culprits if a breach is detected [3]. How can such mechanisms be made practical? Is the threat of penalty to those individuals who are caught compromising privacy satisfactory, or should the cloud be considered an untrusted entity altogether?

If we choose not to trust the cloud, then one avenue of research is to abstract it as a storage and computing device for encrypted information. We could use a recent invention in cryptography called fully homomorphic encryption [19], a scheme allowing the sum and multiplication (and hence arbitrary Boolean circuits) to be performed on encrypted data without needing to decrypt it first. Unfortunately, the first implementations are entirely impractical, but beg the question whether homomorphic encryption can be made practical. Another approach is to sacrifice the generality of homomorphic encryption. We can identify the most important functions that need to be computed on the private data and devise a practical encryption scheme to supports these functions -- think MapReduce [20] on encrypted data. As a high-level example, if all e-mails in Gmail (Google Mail) were encrypted by the user's public key and decrypted by the user's web browser, then Gmail could not produce a

search index for the mailbox. However, if each individual word in the e-mail were encrypted, Gmail could produce an index (the encrypted words would just look like a foreign language) but would not understand the message contents. The latter case implies that Gmail could not serve targeted ads to the user. What are the practical points on the privacy vs. functionality spectrum with respect to computational complexity and with respect to a feasible cloud business model? Secure multiparty computation (SMC) allows mutually distrusting agents to compute a function on their collective inputs without revealing their inputs to other agents [18]. Could we partition sensitive information across clouds, perhaps including a trusted third-party service, and perform SMC on the sensitive data? Is SMC the right model?

Consistency issues

In a broad sense, consistency governs the semantics of accessing the cloud-based services as perceived by both the developers and end-users. The consistency related issues are particularly relevant to the distributed computing infrastructure services (see Figure 1), such as data storage. The most stringent consistency semantics, known as *serializability* or *strong* consistency [8], globally orders the service requests and presents them as occurring in an imaginary global sequence. For example, suppose Alice deposits $5 to a bank account with the initial balance of $0 concurrently with Bob's deposit of $10 to the same account. If Carol checks the account balance twice and discovers it first to be $10 and then $15, then no user would ever see $5 as the valid balance of that account (since in this case, Bob's deposit gets sequenced before that of Alice). In the database community, this type of semantics is typically implied by **ACID** (**A**tomicity, **C**onsistency, **I**solation, and **D**urability).

Intuitively, supporting serializability requires the participants to maintain global agreement about the command ordering. Since cloud services are typically massively distributed and replicated (for scalability and availability), reaching global agreement may be infeasible. Brewer's celebrated **CAP** theorem [9] asserts that it is impossible in a large distributed system to simultaneously maintain (strong) **C**onsistency, **A**vailability, and to tolerate **P**artitions (that is, network connectivity losses).

Researchers have looked for practical ways of circumventing the CAP theorem. Most work has so far focused on relaxing the consistency semantics, basically substituting serializability or (some of) the ACID properties with weaker guarantees. For instance, it does not matter if Carol and Darel in the example above would see either $5 or $10 as the intermediate balances, as long as both of them will eventually see $15 as the final balance. This observation underlies the notion of *eventual* consistency [10], which allows states of the concurrently updated objects to diverge provided that eventually the differences are reconciled, for example when the network connectivity is restored. Apart from eventual consistency, other ways of weakening consistency semantics have looked into replacing single global ordering with multiple orderings. For instance, *causal* consistency [11] allows different clients to observe different request sequences as long as each observed sequence is consistent with the partial cause-effect order.

Weaker consistency semantics work well only for specific types of applications, such as cooperative editing, but do not easily generalize to arbitrary services. (Just imagine what would happen if withdrawals were allowed in the bank account

example above.) Moreover, semantics that are weaker than serializability (or ACID) tend to be difficult to explain to users and developers lacking the necessary technical background. Yet another problem is that for certain types of data, such as the meta-data of a distributed file system, it might be inherently impossible to compromise on strong consistency without risking catastrophic data losses at a massive scale. The possible research questions here would address the following points.

Can we produce a comprehensive and rigorous framework to define and reason about the diverse consistency guarantees? The framework should unify both weaker and stronger models and could serve as a basis for rigorous study of various consistency semantics of cloud services and their relative power. It should be expressive enough to allow new properties to be both easily introduced, for example by composing the existing basic properties, and understood by both developers and consumers of the cloud services. It should also help to bridge diverse perspectives on consistency that exist today within different research communities, like the database and distributed systems communities. Although it is well understood that a cloud architecture should accommodate both strongly and weakly consistent services, it is unclear how the two can be meaningfully combined within a single system, how they should interact or what implications such a model has on performance and scalability.

The current approaches to supporting strong consistency, including Chubby and Zookeeper, primarily focus on isolating the problem into "islands" of server replicas. While beneficial for scalability, such an approach creates an extra dependency on a set of servers that have to be carefully configured and maintained. Can we make strong consistency services that are more dynamic and easier to reconfigure, providing a simpler and more robust solution?

Standards, benchmarks and testbeds

Technical innovations are often followed by standards wars, and cloud computing is no exception. There is a plethora of cloud interoperability alliances and consortia (Open Cloud Manifesto, DTMF Open Cloud Standards Incubator, Open Group's Cloud Work Group, e.g.); the largest incumbents in the market are nevertheless reluctant to follow suit and have chosen to define their own standards. Whereas the strategy is understandable, the lack of interoperability may have adverse effect on consumers who become locked-in on a single vendor. The worry is that clouds become natural monopolies.

The Internet was built on open standards, the question is whether clouds will be as well. Making cloud services open and interoperable may stimulate competition and allow new entrants to enter the cloud market. Customers would be free to migrate their data from a stagnant provider to a new or promising one without difficulty when they so choose. Can the smaller players leverage their collective power to lobby for an open and flexible cloud computing standard that fosters competition while still allowing businesses to profit? Or can this be accomplished by the larger companies or governments? What business models are suitable for an open cloud? On the technical side, could users switch between providers without needing their support, for instance by using a third-party service?

Different cloud providers often adopt similar application programming interfaces (APIs) for physical resources and the distributed computing infrastructure. For

instance, MapReduce and Hadoop expose a similar API, as do the various key-value lookup services (Amazon's Dynamo [16], Yahoo!'s PNUTS [15], memcached [17]). Other components have more diverse APIs, for instance locking services like Google's Chubby [12] and Yahoo!'s Zookeeper [13], and real-time event dissemination services. The broad question asks what components and interfaces are the "right" way to provide the cloud properties mentioned in the introduction. A more specific question is how we can compare and contrast different implementations of similar components. For instance, how can we evaluate the properties of key-value stores like PNUTS and Facebook's Cassandra [7]? The most appealing approach is to compare well-defined metrics on benchmark traces, such as the TPC benchmark for databases[1]. How can we obtain such traces, or perhaps synthetically generated until real ones are produced? Also, consensus benchmarks enable researchers outside the major incumbent companies to advance the core cloud technologies.

Developing distributed computing infrastructure layers or data storage systems is a hard task, but evaluating them for the massive scale imposed by clouds without access to real nodes is next to impossible. Academics who work on peer-to-peer systems (P2P), for example, rely heavily on the PlanetLab[2] testbed for deployment. PlanetLab constitutes over 1000 nodes distributed across nearly 500 sites, making it ideal an ideal resource for experimental validation of geographically networked systems which sustain heavy churn (peer arrivals and departures). The nodes in the data centers underlying the cloud tend to be numerous, hierarchically structured with respect to networking equipment, and face limited random churn but occasionally suffer from large-scale correlated failures. PlanetLab's focus on wide-area networks is suboptimal for cloud platform research, unfortunately, and the same holds true for other similar resources. A handful of testbeds appropriate for cloud research have made their debut recently, including Open Cirrus from HP, Intel and Yahoo! and the Open Cloud Testbed. We encourage other players to participate and contribute resources to cloud research, with the goal of providing a standard testbed with open-access, at least for academia, including researchers from underrepresented universities. Who will create the future "CloudLab"?

Conclusion

Clouds may be young but the role of cloud computing in society is becoming increasingly important. Their scale, both technically and socially, presents both challenges and enormous opportunities for contributions by the research community. Those interested should consider participating in the LADIS[3] or HotCloud[4] workshops, or the upcoming Symposium on Cloud Computing[5] (SoCC). Large industry players are currently driving the research bandwagon for cloud computing, but the journey is only beginning. A concerted multi-disciplinary effort is needed to turn the cloud computing promise into a success.

---

[1] http://www.tpc.org
[2] http://www.planet-lab.org
[3] http://www.cs.cornell.edu/projects/ladis2010
[4] http://www.usenix.org/events/hotcloud10
[5] http://research.microsoft.com/en-us/um/redmond/events/socc2010

References

[1] D. Narayanan, A. Donnelly, E. Thereska, S. Elnikety, and A. Rowstron. Migrating Server Storage to SSDs: Analysis of Tradeoffs. *Proceedings of EuroSys 2009*. Nuremberg, Germany, April 2009

[2] Urs Hoelzle and Bill Weihl. High-Efficiency Power Supplies for Home Computers and Servers. Urs Hoelzle and Bill Weihl. Google Inc. September 2006. Available at http://services.google.com/blog_resources/PSU_white_paper.pdf

[3] Geoffrey Smith. Principles of Secure Information Flow Analysis. Chapter 13 (pp. 291-307) of Malware Detection, edited by Mihai Christodorescu, Somesh Jha, Douglas Maughan, Dawn Song, and Cliff Wang, Springer-Verlag, 2007.

[4] L. Ganesh, H. Weatherspoon, M. Balakrishnan, K. Birman. Optimizing Power Consumption in Large-Scale Storage Systems. *In Proc. of HotOS*. 2007.

[5] Roy Campbell, et al. Open Cirrus Cloud Computing Testbed: Federated Data Centers for Open Source Systems and Services Research, *USENIX HotCloud '09*, San Diego, June, 2009.

[6] Ann Cavoukian. Privacy in the Clouds. A White Paper on Privacy and Digital Identity: Implications for the Internet, 2008. Available at http://www.ipc.on.ca/images/Resources/privacyintheclouds.pdf

[7] Raghu Ramakrishnan. Data Management Challenges in the Cloud (Invited Talk). *ACM SIGOPS LADIS 2009*. Available at http://www.cs.cornell.edu/projects/ladis2009/talks/ramakrishnan-keynote-ladis2009.pdf

[8] Jim Gray, Andreas Reuter. Isolation Concepts. Chapter 7 of Transaction Processing: Concepts and Techniques. Morgan Kaufman Publishers, 1993.

[9] Eric Brewer. Towards Robust Distributed Systems (Invited Talk). *Principles of Distributed Computing (PODC)*. Portland, Oregon, July 2000.

[10] W. Vogels. Eventually Consistent. ACM Queue vol. 6, no. 6, December 2008

[11] M. Ahamad, P. W. Hutto, G. Neiger, J. E. Burns, and P. Kohli. Causal Memory: Definitions, Implementations and Programming. Distributed Computing, 9:37-49, 1995.

[12] M. Burrows. The Chubby lock service for loosely-coupled distributed systems. *OSDI '06: Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation* (pp. 335-350). Seattle, WA: USENIX Association.

[13] B. Reed, F. P. Junqueira. A simple totally ordered broadcast protocol. *Second Workshop on Large-Scale Distributed Systems and Middleware (LADIS 2008).* Yorktown Heights, NY: ACM. ISBN: 978-1-60558-296-2.

[14] Nicholas Carr. The Big Switch: Rewiring the World, from Edison to Google. W.W.Norton. January 2008

[15] B. Cooper, R. Ramakrishnan, et al. PNUTS: Yahoo!'s hosted data serving platform. *Proceedings of VLDB Endowment*. Vol. 1, No. 2. (2008), pp. 1277-1288.

[16] DeCandia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A., et al. Dynamo: Amazon's highly available key-value store. *SOSP '07: Proceedings of the twenty-first ACM SIGOPS Symposium on Operating Systems Principles* (pp. 205-220). Stevenson, Washington: ACM. 2007.

[17] *memcached: a Distributed Memory Object Caching System*. http://www.danga.com/memcached/

[18] Wenliang Du, Mikhail J. Atallah. Secure multi-party computation problems and their applications: a review and open problems. *Proceedings of the 2001 Workshop on New Security Paradigms.* Cloudcroft, New Mexico. September, 2001.

[19] Craig Gentry. Fully homomorphic encryption using ideal lattices. *In Proceedings of the ACM Symposium on Theory of Computing (STOC '09)*. Bethesda, MD, USA. June 2009.

[20] J. Dean, S. Ghemawat. MapReduce: simplified data processing on large clusters. *Commun. ACM, 51* (1), 107-113. 2008.
[21] S. Khuller, J. Li, B. Saha. Energy Efficient Scheduling via Partial Shutdown. *In Proc. of ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2010.