УДК 4И (Англ): 62

CLOUD COMPUTING WITH AMAZON WEB SERVICES: A VIABLE WAY FORWARD?

Kurbatova E.A., Tolstoy A.B.
Associate professor - Kurbatova E.A.
Siberian Federal University, Krasnoyarsk

Over the last years, there has been a lot of media hype around such disruptive technology as 'cloud computing'. This hype has led people to question its real application; however such web giant as Amazon, one of the leaders in Infrastructure-as-a-Service market, with its *Amazon Web Services (AWS)* is unlikely to be jumping into the cloud just due to the buzz. But the question is: how can one benefit from using Amazon's infrastructure web services platform in the cloud?

Amazon Web Services provides a collection of highly reliable and scalable infrastructure services with more flexibility than a traditional data center. It offers a rich set of Application Programming Interface (API) to take advantage of Amazon's global computing infrastructure.

Amazon emphasizes several benefits an IT organization or a developer can get:

- *Cost-effective*. No up-front commitments are required, you pay only for what you use and as you use.
- *Dependable*. One can rely on the expertise of Amazon in building, operating, and maintaining distributed, secure and resilient cloud.
- *Flexible*. Amazon Web Services provides you with ready to use services for managing databases, queues, etc. Despite this, it does not force you to use any specific operating system, development platform or programming model. AWS is not only a rich platform to build web-scale products but also a platform to build platforms on top of it.

Amazon Elastic Compute Cloud (Amazon EC2)

Amazon EC2 service allows users to rent virtual servers (instances) and launch a variety of operating systems through a web service interface. A user can select a pre-configured, templated Amazon Machine Image (AMI) to get up and running immediately, or create it manually with all the applications, libraries, data and configuration required. Using AMIs an application can scale out any time by launching another instance with the same AMI. Amazon EC2 uses capabilities of Amazon S3 to store AMIs both privately and publicly.

Amazon EC2 uses Xen virtualization meaning that every 'instance' functions as a virtual private server. Several types of instances are on offer at Amazon EC2 at three different regions (US - North Virginia, US - North California, EU - Ireland) with several availability zones.

- **Standard instances**, which suit most applications.
  - *Small* Instance (Default) 1.7 GB of memory, 1 EC2 Compute Unit (1 virtual core with 1 EC2 Compute Unit), 160 GB of local instance storage, 32-bit platform.
  - *Large* Instance 7.5 GB of memory, 4 EC2 Compute Units (2 virtual cores with 2 EC2 Compute Units each), 850 GB of local instance storage, 64-bit platform.
  - *Extra Large* Instance 15 GB of memory, 8 EC2 Compute Units (4 virtual cores with 2 EC2 Compute Units each), 1690 GB of local instance storage, 64-bit platform.
- **High-Memory instances** offer large memory sizes for high throughput applications.

- High-Memory *Extra Large* Instance 17.1 GB memory, 6.5 ECU (2 virtual cores with 3.25 EC2 Compute Units each), 420 GB of local instance storage, 64-bit platform.
- High-Memory *Double Extra Large* Instance 34.2 GB of memory, 13 EC2 Compute Units (4 virtual cores with 3.25 EC2 Compute Units each), 850 GB of local instance storage, 64-bit platform.
- High-Memory *Quadruple Extra Large* Instance 68.4 GB of memory, 26 EC2 Compute Units (8 virtual cores with 3.25 EC2 Compute Units each), 1690 GB of local instance storage, 64-bit platform.

- **High-CPU instances** offer more CPU resource and suit compute-intensive applications.
  - High-CPU *Medium* Instance 1.7 GB of memory, 5 EC2 Compute Units (2 virtual cores with 2.5 EC2 Compute Units each), 350 GB of local instance storage, 32-bit platform.
  - High-CPU *Extra Large* Instance 7 GB of memory, 20 EC2 Compute Units (8 virtual cores with 2.5 EC2 Compute Units each), 1690 GB of local instance storage, 64-bit platform.

One EC2 Compute Unit (ECU) provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor.

Amazon EC2 provides customers with three different instance purchasing models:

- **On-Demand Instances** allow users to pay a fixed rate for an instance-hour consumed (each partial instance-hour will be billed as full hour). *($0.085 - $2.88 per hour)*
- **Reserved Instances** allow users to get a discount on the instance-hour usage price by paying a low one-time fee with an upfront commitment for a one- or three-year term. *($0.03 - $1.10 per hour)*
- **Spot instances** provide customers with an ability to purchase compute capacity with no upfront commitment and at hourly rates usually lower than the On-Demand rate by specifying the maximum hourly price that you are willing to pay to run a particular instance type. *($0.031 - $1.081 per hour)*

There are also additional charges for incoming/outgoing Internet traffic and several optional services:

- Amazon Elastic Block Storage (Amazon EBS), service providing block level persistent storage volumes for use with Amazon EC2 instances.
- Elastic IP address
- Amazon CloudWatch, service for monitoring instances.
- Elastic Load Balancing

Amazon EC2 management can be performed in several ways including AWS Management Console (web interface), Amazon EC2 API Tools (command-line tools serving as a client interface) and ElasticFox (Firefox extension for Amazon EC2). One is free to choose any developer tool to control its EC2 instances, etc.

As for the projects and companies that use Amazon EC2, benefits and applications of this service are innumerable. For instance, Twitter, a popular micro-blogging service, solved its scalability issues on Amazon cloud. On the other hand, Harvard's Laboratory for Personalized Medicine uses customized Oracle AMIs on EC2 to run genetic testing models and simulations.

Amazon Simple Storage Service (Amazon S3)

Amazon S3 is one of the main AWS services that most people tend to call 'cloud storage'. It provides users with unlimited storage through a simple web services interface. Amazon claims that it gives any developer access to the same highly scalable, reliable, fast, inex-

pensive data storage infrastructure that Amazon uses to run its own global network of web sites.

Objects are redundantly stored on multiple devices, with the only limitation of 5Gb data per object. Each object is stored in a bucket and identified by a unique, user-assigned key. A bucket can be created in one of several regions (covered in the Amazon EC2 section). Buckets and objects can be created, listed and retrieved using either a REST-style HTTP or a SOAP interfaces. Additionally, objects can be downloaded using the BitTorrent protocol and the HTTP GET interface. Apart from that, Amazon has recently allowed versioning for objects stored.

Since objects are accessible by HTTP, Amazon S3 may be used as static web hosting service. Moreover, using ACLs (Access Control Lists) users can limit access to the objects.

Storage cost is calculated per gigabyte used, starting from $0.055/Gb with additional charges for data transfer (in/out) and for the requests (e.g. PUT, COPY, LIST, GET).

Ubuntu One and Dropbox, online backup and synchronization utilities use S3 as their storage and transfer facility. Twitter and many other websites use Amazon S3 to host static content such as images. Moreover, S3 is very widely used as data storage for MapReduce tasks.

Amazon Elastic MapReduce

Amazon Elastic MapReduce is based on Apache Hadoop, opensource Java software framework for data-intensive distributed applications. Hadoop implements Google's computational model called 'MapReduce' in which job is divided into many chunks, each of which can be executed on any node in the cluster.

A developer can write jobs in SQL-like languages such as Hive and Pig, or use the power of almost any programming language including Java, Ruby, Perl, Python, PHP, R, C, C++, etc.

Amazon Elastic MapReduce uses the advantages of Amazon EC2 and Amazon S3 meaning that the charges correspond with those applied for these services.

Test case on Amazon Web Services

A task of generating 3-grams corpus (subsequences of 3 words with the number of times they appeared) from all the articles in English Wikipedia is not so trivial for a regular PC. Nearly 6Gb of plain-text articles parsed from the Freebase Wikipedia Extraction (WEX) were uploaded to Amazon S3. Simple (but optimized for usage with Hadoop) Python code was written to map and reduce 3-grams.

The task was executed on a 10 node (+1 master node) cluster of m1.small instances with 2/1 (map/reduce) tasks per node. The execution finished within 2 hours. Map step produced 16Gb of 3-grams, while reduce step resulted in 3.5Gb of data.

As for the cost issues, Amazon Web Services console does not provide users with a bill for a fixed period (only a total for a billing period of a month), however approximate charges can be calculated: 11 m1.small instances at $0.085/hour (in the US East region) plus 10Gb S3 storage at $0.150/Gb. Clearly, the total cost is around $3.50.

All things considered, this experiment highlights the ease with which data intensive computation can be handled. One can gain access to massive computing (Amazon EC2) and storage (Amazon S3) resources and to Hadoop framework with just a few command line invocations (or just a few mouse clicks alternatively).

Conclusion

Ultimately, cloud computing is something that goes well beyond a hype. This shift in IT might take some time to become mature. Reliable and flexible services available on-demand at a fixed price are attractive and Amazon Web Services provides a customer with innumerable advantages regardless of whether one uses Amazon's services for backups and storage, application hosting, content delivery, high performance computing, media hosting,

on-demand workforce or any other application in the cloud. On the condition that a customer needs a cost-effective, dependable, flexible platform, Amazon is among the leaders.