

**CONTENTS INCLUDE:**

- About Cloud Computing
- Usage Scenarios
- Underlying Concepts
- Cost
- Data Tier Technologies
- Platform Management and more...

# Getting Started with Cloud Computing

By Daniel Rubio

## ABOUT CLOUD COMPUTING

Web applications have always been deployed on servers connected to what is now deemed the 'cloud'.

However, the demands and technology used on such servers has changed substantially in recent years, especially with the entrance of service providers like Amazon, Google and Microsoft.

These companies have long deployed web applications that adapt and scale to large user bases, making them knowledgeable in many aspects related to cloud computing.

This Refcard will introduce you to cloud computing, with an emphasis on these providers, so you can better understand what it is a cloud computing platform can offer your web applications.

## USAGE SCENARIOS

### Pay only what you consume

Web application deployment until a few years ago was similar to most phone services: plans with allotted resources, with an incurred cost whether such resources were consumed or not.

Cloud computing as it's known today has changed this. The various resources consumed by web applications (e.g. bandwidth, memory, CPU) are tallied on a per-unit basis (starting from zero) by all major cloud computing platforms.

This can be beneficial for web applications that have disproportionate resource requirements (e.g. bandwidth intensive vs. memory intensive), since only consumed resources incur in cost.

### One time event provisioning

Web applications are often subject to traffic spikes due to one time events (e.g. National broadcast exposure, SuperBowl commercial). Not only can this type of provisioning be expensive, but often times difficult to achieve.

By using a cloud computing platform, provisioning of this sort can be greatly simplified.

Cloud computing platforms allow web applications "on tap" access to resources without an application owner (i.e. you) footing the bill for stand-by equipment.

Additionally, since the underlying architecture of a web application is built around a cloud computing platform, this

also minimizes the need to make design changes to support one time events.

### Automated growth & scalable technologies

Having the capability to support one time events, cloud computing platforms also facilitate the gradual growth curves faced by web applications.

Large scale growth scenarios involving specialized equipment (e.g. load balancers and clusters) are all but abstracted away by relying on a cloud computing platform's technology.

In addition, several cloud computing platforms support data tier technologies that exceed the precedent set by Relational Database Systems (RDBMS): Map Reduce, web service APIs, etc. Some platforms support large scale RDBMS deployments.

## CLOUD COMPUTING PLATFORMS AND UNDERLYING CONCEPTS

### Amazon EC2: Industry standard software and virtualization

Amazon's cloud computing platform is heavily based on industry standard software and virtualization technology.

Virtualization allows a physical piece of hardware to be utilized by multiple operating systems. This allows resources (e.g. bandwidth, memory, CPU) to be allocated exclusively to individual operating system instances.

As a user of Amazon's EC2 cloud computing platform, you are assigned an operating system in the same way as on all hosting providers that preceded cloud computing platforms.



**Don't Miss An Issue!**  
Get over 70 DZone Refcardz  
**FREE** from Refcardz.com!

**Visit Refcardz.com to get them all Free!**

The primary difference is that such an instance is highly customizable, in addition to having its resources tallied on a per unit basis, as well as being equipped to scale to larger loads on a case by case basis.

#### Key characteristics of Amazon EC2

- Choice of industry standard server operating system (e.g. Windows, Linux, Solaris)
- Deployment building block consists of an Amazon Machine Image(AMI). An AMI is a standard server operating system image with pre-selected applications. AMI's can be found at: <http://developer.amazonwebservices.com/connect/kbcategory.jspa?categoryID=171>
- Application development open to any server-side development tool, compatible with industry standard server operating system.

#### Google App Engine: Google infrastructure & SDK

Google's cloud computing platform is heavily based on Google's own server infrastructure.

As a user of Google's App Engine, your web applications are built on the same principles as Google applications.

#### Key Characteristics of Google App Engine

- Built on Google infrastructure (i.e. No commercially available server operating system).
- Choice of either Python or Java run-time for running web applications. Other pre-selected applications are available via services (e.g. Mail, Memcache).
- Application development tightly pegged to Google's Software Development Kit (SDK). ([http://code.google.com/appengine/downloads.html#Download\\_the\\_Google\\_App\\_Engine\\_SDK](http://code.google.com/appengine/downloads.html#Download_the_Google_App_Engine_SDK))
- Tightly integrated with Google's web services APIs (e.g. For authenticating users and sending email).
- Free quotas for applications limited to: 500MB of persistent storage and CPU & bandwidth for approximately 5 million page views a month.

#### Microsoft Azure: Azure & Visual Studio

Microsoft's cloud computing platform is tightly integrated with Microsoft's product line.

As a user of Microsoft Azure's cloud computing platform, you can expect your web applications to have streamlined integration with Microsoft's product line.

#### Key Characteristics of Microsoft Azure

- Operates on Microsoft's virtualized 64-bit Windows Server 2008 operating system.
- Support for .NET applications, as well as other third party applications available for the same OS running on a standard server (i.e. unmanaged code apps).
- Support for .NET services: .NET Access Control Service & .NET Service Bus. Originally known as BizTalk services, focused on enterprise application scenarios.

- Application development tightly integrated with Microsoft's Visual Studio, in addition to having its own Software Development Kit (SDK)  
<http://go.microsoft.com/fwlink/?LinkID=128752>
- Free usage under CTP (Community Technology Preview), but limited to 2000 hours, 50 GB of persistent storage and 20GB/day bandwidth.

#### Selection Grid by Web Application Language

Web application language	Amazon EC2	Google App Engine	Microsoft Azure
PHP	✓		
.NET	✓		✓
Java	✓	✓	
Python	✓	✓	
Ruby	✓		

#### Resources (Bandwidth, CPU, I/O)

Cloud computing providers keep track of consumed resources on a more granular basis than traditional service providers. The following list illustrates a series of consumption units:

- Server – Per Hour
- Bandwidth – Per Gigabyte
- Storage – Per Gigabyte
- CPU/Memory – Per unit
- Emails – Per recipient

This approach gives an application owner (i.e. you) greater leverage and cost effectiveness. The next section on 'Costs' illustrates case scenarios with side by side comparisons for the various cloud computing platforms.

#### Other cloud computing providers

In addition to Amazon's EC2 , Google's App Engine and Microsoft's Azure cloud computing platforms, other providers in this space have also emerged.

Some of these providers include:

- **Slice Host** - <http://www.slicehost.com/>
- **Linode** - <http://www.linode.com/>
- **Prgmr** - <http://prgmr.com/>
- **Heroku** - <http://heroku.com/>
- **Rackspace** - <http://www.rackspacecloud.com/>
- **GoGrid** - <http://www.gogrid.com/>

Many of these providers rely on industry standard virtualization and operating system technology, making them close competitors to Amazon's EC2 cloud computing platform.

Comparing these other providers to Google's App Engine or Microsoft's Azure cloud computing platforms can be more difficult. This in light of the greater proprietary nature of both Google's and Microsoft's platforms.

Still, with the brand recognition and breadth of companies like Amazon, Google and Microsoft, these other cloud computing providers can often fall short of being deemed 'platforms'.

This can be due to a lack of end-to-end integration (e.g. application development, tools and application deployment),

lack of scalable data tier technology options, to service level agreements (e.g. uptime and indemnity) that can only be offered by large corporations the size of Amazon, Google and Microsoft.

Nevertheless, some of these other cloud computing providers have carved out niche markets in the cloud computing market. Some do so by adopting more aggressive pricing structures, catering to the specific needs of certain communities (e.g. Ruby/Rails, or Linux), or providing better customer service than their larger rivals.

## COSTS

Cloud computing platform costs are fairly competitive. However, some metrics used by providers are sufficiently different from others to make holistic cost comparisons difficult.

For example, stored data can have added costs related to the number of Input/Output operations or transactions. Other aspects, like CPU consumption, can also vary in the form they are tallied by provider. The following table illustrates comparable resources and their associated costs in each cloud computing platform.

Resources	Amazon EC2 (Small instance)	Google App Engine	Microsoft Azure
Outgoing bandwidth (Gigabyte)	\$0.10 (Over 150 TB) ~\$0.17(First 10 TB)	\$0.12	\$0.15
Incoming bandwidth (Gigabyte)	\$0.10	\$0.10	\$0.10
CPU time (hours)	\$0.085 (Unix/Linux)~ 0.12 (Windows)	\$0.10	\$0.12
Stored data (Gigabytes per month)	\$0.10 (+ \$0.10 per 1 million I/O requests)	\$0.15	\$0.15 (+\$0.01 for 10K transactions)
Recipients emailed (Recipients)	N/A	\$0.0001	N/A

### Cost Calculators

For an accurate cost estimate pertaining to each cloud computing platform, I recommend you use the following calculators offered by each provider:

- Amazon EC2 - <http://calculator.s3.amazonaws.com/calc5.html>
- Google App Engine - <http://code.google.com/appengine/docs/billing.html> (ONLY budgeting resources – No calculator)
- Microsoft Azure - <http://www.microsoft.com/windowsazure/tco/>

### Cost case scenarios: Mailing list or report processing

To give added cost context to the use of cloud computing platforms in web applications, let's take the case of common one-time events in web applications.

Mailing list or end of month report processing can consume substantial resources from a web application's main environment, in addition to being short-lived tasks.

Instead of leasing a stand-alone server for such tasks or hampering the performance level of a web application's main environment, a cloud computing platform can be a cost effective solution.

Assuming the data for a mailing list or report batch is already stored on a cloud computing platform: A conservative estimate of 1 day (24 hours) for processing and 5GB of outgoing bandwidth, would equal approximately \$3.00 in cost from each of the previous cloud computing providers.

As you can surely attest, at this price point it's only such cloud computing providers that are able to offer dedicated resources at such competitive rates, especially compared to leasing your own hardware or using one of the many commercial hosting providers.

### Spot pricing on Amazon EC2

Providing what can potentially be the most competitive rates among cloud computing platforms, Amazon EC2 offers what it calls 'spot instances'.

A spot instance allows you to make a bid for unused Amazon EC2 capacity and run applications for as long as your bid exceeds the current spot price.

For web application tasks that are not time sensitive (e.g. long-running scientific calculations or historical reports) this approach can substantially reduce a web application's running costs.

Since spot prices change based on supply and demand, this allows you to obtain the most competitive rates at any given time, without exceeding your maximum bid.

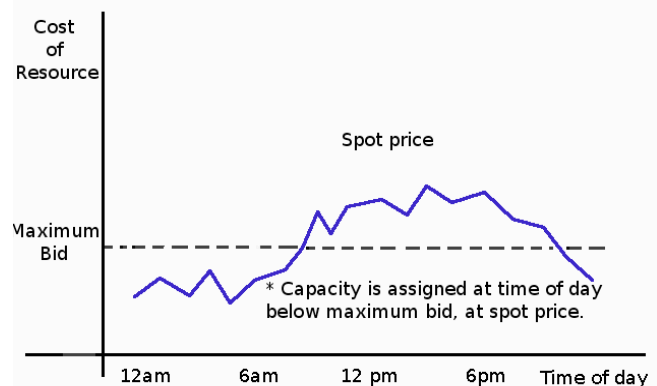


Figure - Amazon EC2 spot pricing behavior

More information on Amazon EC2 Spot instances can be found at: <http://aws.amazon.com/ec2/spot-instances/>

## CLOUD COMPUTING PLATFORMS & DATA TIER TECHNOLOGIES

Scaling a web application's data tier entails a different approach than scaling its business logic and web tier. This is due to limitations and features pertaining to specific data tier technologies.

Most web applications are underpinned by Relational Database Management Systems (RDBMS) that use Structured Query Language (SQL) as their access mechanism.

Though a series of cloud computing platforms now offer RDBMS/SQL data tier support, many cloud computing platforms grew to address data tier demands for which RDBMS/SQL technology had limiting factors. Namely those pertaining to data mining and the complexities involved in providing fault-tolerant & high-availability RDBMS/SQL solutions.

**NoSQL movement**  
 The industry has blossomed healthy debates over the suitability of RDBMS/SQL vs. alternate data tier technologies for developing large scale web applications. Now often cataloged as the NoSQL movement <http://en.wikipedia.org/wiki/nosql>

**Hot Tip**

**Amazon EC2 Data Tier**

Amazon’s cloud computing platform offers the largest array of data tier technologies.

**Amazon SimpleDB**

SimpleDB technology has the following characteristics:

- Storage and retrieval based on Amazon API; available via web service.
- Low administrative overhead compared to RDBMS (e.g. No index maintenance and performance tuning required)
- Schema-less; requiring no up-front data modeling tasks.
- Provides the building block for querying Amazon S3 data.

**Amazon Simple Storage Service (S3)**

Whereas Amazon SimpleDB provides the foundations for querying data in Amazon’s EC2 cloud computing platform, Amazon’s Simple Storage Service (S3) is used for the actual storage of data.

Simple Storage Service (S3) has the following characteristics:

- Storage of objects between 1 byte and 5 gigabytes.
- REST and SOAP interfaces, as well as authentication mechanisms.
- Objects are assigned a unique ID, with meta-data assignment done in Amazon SimpleDB for querying purposes.
- Built on Amazon infrastructure.

**Amazon Simple Queue Service**

Provides data tier capabilities similar to those of message orientated middleware ( [http://en.wikipedia.org/wiki/Message-oriented\\_middleware](http://en.wikipedia.org/wiki/Message-oriented_middleware) ) for web applications.

Amazon Simple Queue Service has the following characteristics:

- Messages can contain up to 8 KB of text in any format.
- Messages can be sent and read simultaneously.
- Access is supported through standard SOAP web services.

**Amazon Elastic MapReduce**

Provides data tier capabilities based on Google’s MapReduce framework ( <http://en.wikipedia.org/wiki/MapReduce> ) built on Amazon’s EC2 cloud computing platform.

Amazon Elastic MapReduce has the following characteristics:

- Out-of-the-box MapReduce capabilities built on Apache’s MapReduce implementation Hadoop.
- Depends on Amazon Simple Storage Service (S3).
- Support for third party MapReduce tools (e.g. Karmasphere)

**Amazon Relational Database Service**

Provides data tier capabilities for deploying RDBMS/SQL web applications.

Amazon Relational Database Service has the following characteristics:

- Out-of-the-box RDBMS/SQL capabilities built on MySQL.
- Scale and compute capacity managed through Amazon APIs.
- Automated backup and patch management.

**Google App Engine Data Tier**

Google’s cloud computing platform is built entirely on Google’s data tier technology stack.

Google’s App Engine data tier has the following characteristics:

- Storage and retrieval based on either Java – available via Java Data Objects (JDO), Java Persistence API (JPA) or low-level datastore API – as well as Python – available via a data modeling API and a SQL-like query language called GQL.
- Schema-less; requiring no up-front data modeling tasks.
- Built on Google infrastructure (i.e. BigTable, Google File System).

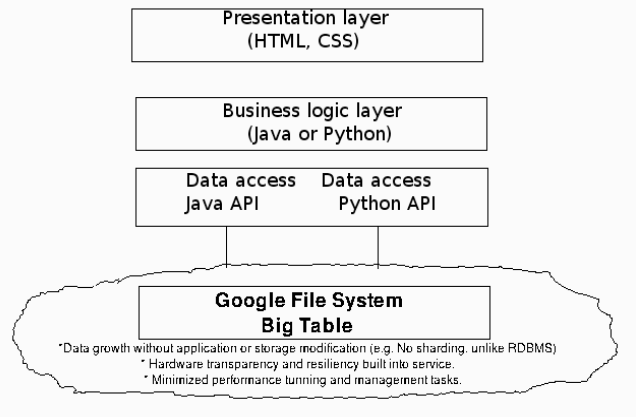


Figure - Google App Engine Data Tier Advantages

**Microsoft Azure Data Tier**

Microsoft’s cloud computing platform offers similar data tier solutions to the previous cloud computing platforms, based on Microsoft technology.

**Windows Azure Storage Service**

- Storage and retrieval based on .NET API: ADO.NET or LINQ, as well as web services (e.g. REST).
- Schema-less; requiring no up-front data modeling tasks.
- Built on Microsoft infrastructure, including storage replication.

**Windows SQL Azure**

- Out-of-the-box RDBMS/SQL capabilities built on Microsoft SQL Server.
- Minimal operational management (e.g. Disk usage, log files)
- Synchronization availability between various RDBMS instances (a.k.a ‘Huron Data Sync’)

## CLoud COMPUTING PLATFORM MANAGEMENT

For all the benefits of cloud computing platforms, the term 'cloud' often comes with the connotation of losing control over one's web applications and being at the mercy of a service provider.

While it's true that some cloud computing platforms have certain proprietary elements that can lock-in your applications to their service offerings, cloud computing management and security concerns are often unfounded.

### Cloud computing platform management

Management of cloud computing platforms – which is to say provisioning or modifying (e.g. starting, stopping or deleting) an underlying environment – is achieved by either a provider's administrative web console, through APIs or other third party tools.

Administrative web consoles provide practical access to standard cloud computing tasks. APIs on the other hand allow the execution of more sophisticated cloud management chores, such as the integration of tasks into custom applications or automation of tasks altogether. Third party tools can range from browser plug-ins to open source libraries.

#### Amazon EC2 management

Amazon's cloud computing platform can be managed through the following means:

- Amazon EC2 Administrative console: Basic web console for managing EC2 instances, Elastic Block Store volumes and modifying configuration settings (e.g. I.P addresses). <http://aws.amazon.com/console/>
- Amazon CloudWatch: Advanced web console – billed separately – for determining resource utilization, operational performance, and demand metrics (e.g. CPU utilization, disk reads and writes, and network traffic). <http://aws.amazon.com/cloudwatch/>
- Amazon EC2 API: Web services API for inspecting and modifying EC2 instances from remote/custom applications. <http://docs.amazonwebservices.com/AWSEC2/latest/APIReference/>
- Libcloud API: Python API for inspecting and modifying EC2 instances from remote/custom applications. <http://incubator.apache.org/libcloud/>
- Elasticfox & S3Fox browser plug-ins: Firefox plug-ins for managing EC2 instances & EC3 data.  
Elasticfox - <http://developer.amazonwebservices.com/connect/entry.jspa?externalID=609>  
S3Fox - <http://developer.amazonwebservices.com/connect/entry.jspa?externalID=771>
- Lifeguard: provides an automatic, Spring based monitoring solution to dynamically scale EC2 resources based on load. <http://code.google.com/p/lifeguard/issues/list>

#### Google App Engine

Google's App Engine computing platform can be managed through the following means:

- Google App Engine Administrative console: Basic web console for managing Google App Engine. <https://appengine.google.com/>
- Google App Engine API: Google's App Engine development kit (SDK) includes an API to communicate remotely with Google App Engine servers.  
Python - <http://code.google.com/appengine/docs/python/tools/>  
Java - <http://code.google.com/appengine/docs/java/tools/>

#### Microsoft Azure

Microsoft's Azure computing platform can be managed through the following means:

- Microsoft Azure Administrative console: Basic web console for managing Windows Azure instances. <https://windows.azure.com/Cloud/Provisioning/Default.aspx>
- Windows Azure API: Windows Azure development kit (SDK) includes an API to communicate remotely with Windows Azure servers. <http://msdn.microsoft.com/en-us/library/dd179367.aspx>
- Windows Azure Management Tool: Provides a desktop (i.e. fat-client) to communicate remotely with Windows Azure servers. <http://code.msdn.microsoft.com/windowsazuremmc>

## CLoud COMPUTING PLATFORM SECURITY

Generally speaking, security for web applications running on cloud computing platforms is no different than security pertaining to any web application accessible to the public at large.

Issues such as code injection ( [http://en.wikipedia.org/wiki/Code\\_injection](http://en.wikipedia.org/wiki/Code_injection) ) or cross-site scripting ( [http://en.wikipedia.org/wiki/Cross\\_site\\_scripting](http://en.wikipedia.org/wiki/Cross_site_scripting) ) can just as easily present themselves in web applications running on cloud computing platforms, given they are issues entirely under the control of an application's designer.

As a user of a cloud computing platform, your security concerns should span to contemplate the security vulnerabilities and security limitations inherent to a provider's services, in addition to those of web applications in general.

The following sections enumerate key security characteristics to take into account when choosing a cloud computing platform.

#### Amazon EC2 security characteristics

- Full access to host operating system instance. Vulnerability and 'hardening' policies are the responsibility of a user, as with any other public operating system.
- Amazon Security groups to facilitate and limit access to instances by port, protocol and or incoming IP.
- Optional multi-factor authentication, to limit access through a six-digit, single-use code from an authentication device in your physical possession ( <http://aws.amazon.com/mfa/> )

#### Google App Engine security characteristics

- Access to underlying host provided entirely through a Google account. Limiting a user's security accountability (e.g. no operating system to 'harden')



- No custom domain SSL certificate support (i.e. <https:// access>). SSL is supported, but only routed via a domain in the form <https://your-app-id.appspot.com>
- Google Secure Data Connector (SDC) support. Allows data encryption between applications running on Google App
- Engine and a corporate network.

**Microsoft Azure security characteristics**

- Access to underlying host provided entirely through Windows Live ID account, limiting a user's security accountability.
- Windows host operating system instance with limited security accountability. Updates are performed automatically.

- Role based access mechanisms. Supported are Web roles as defined by ASP.NET – and Worker roles for general purpose tasks.

**CLOUD COMPUTING TEAM BLOGS**

In order to keep abreast on the latest offerings made by cloud computing providers, I recommend you consult each platform's team blog.

**Google App Engine team blog:**

<http://googleappengine.blogspot.com/>

**Amazon EC2 team blog:**

<http://aws.typepad.com/>

**Microsoft Azure team blog:**

<http://blogsmsdn.com/windowazure/>

**ABOUT THE AUTHOR**



**Daniel Rubio** is an independent technology consultant with over 10 years of experience in enterprise and web based systems.

He is the author of several books focused on enterprise Java, in addition to participating as technical writer and editor for several online technical publishers. He maintains a blog covering software platforms and emerging technologies at <http://www.webforefront.com/>

**RECOMMENDED BOOK**

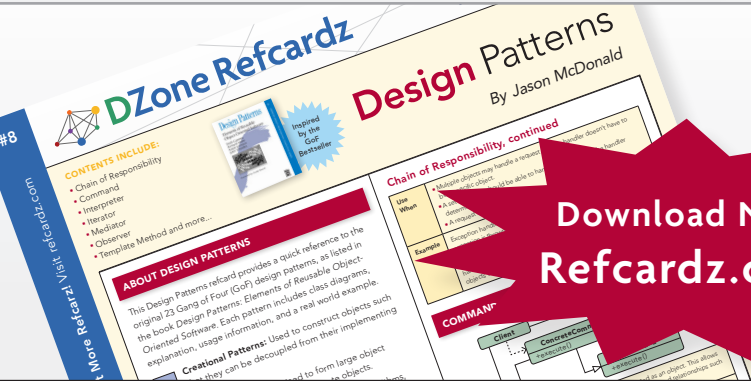


This book is an industry-leading primer on cloud computing: its background, the purpose it serves, how the cloud can be best utilized, which platforms offer which features, and how to get started.

**BUY NOW**

[books.dzone.com/books/cloud-computing](http://books.dzone.com/books/cloud-computing)

**Professional Cheat Sheets You Can Trust**



**Download Now Refcardz.com**

*"Exactly what busy developers need: simple, short, and to the point."*

James Ward, Adobe Systems

**Upcoming Titles**

- Blaze DS
- Domain Driven Design
- Virtualization
- Java Performance Tuning
- Expression Web
- Spring Web Flow
- BPEL

**Most Popular**

- Spring Configuration
- jQuery Selectors
- Windows Powershell
- Dependency Injection with EJB 3
- Netbeans IDE JavaEditor
- Getting Started with Eclipse
- Very First Steps in Flex



DZone communities deliver over 6 million pages each month to more than 3.3 million software developers, architects and decision makers. DZone offers something for everyone, including news, tutorials, cheatsheets, blogs, feature articles, source code and more.

**"DZone is a developer's dream,"** says PC Magazine.

DZone, Inc.  
140 Preston Executive Dr.  
Suite 100  
Cary, NC 27513  
888.678.0399  
919.678.0300

**Refcardz Feedback Welcome**  
[refcardz@dzone.com](mailto:refcardz@dzone.com)

**Sponsorship Opportunities**  
[sales@dzone.com](mailto:sales@dzone.com)



\$7.95