

# Introduction to Google's Cloud Platform Technologies



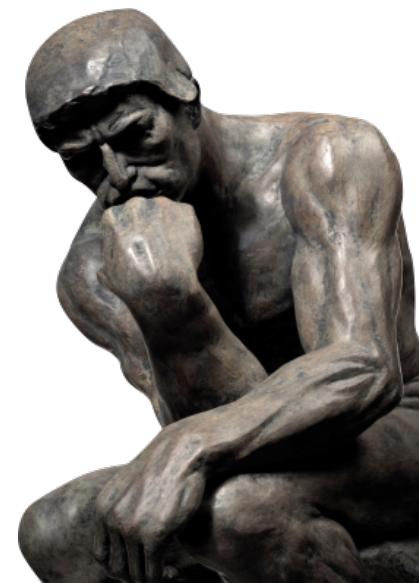
Chris Schalk  
Google Developer Advocate

Cloudstock  
Monday Dec 6<sup>th</sup>, 2010

# What is cloud computing?



**Just Kidding ;-)**



# Google Cloud Platform Technologies at Glance

Existing



Google App Engine

Google App Engine for Business **(new)**

---

New!

Google BigQuery

Google  
Prediction API

Google Storage

# Agenda

- Part I - Intro to App Engine
  - App Engine Details
  - Development Tools
  - App Engine for Business
- Part II – Google's new cloud technologies
  - Google Storage
  - Prediction API
  - BigQuery

# Part I – Intro to App Engine

## Topics covered

- App Engine a PaaS
- App Engine usage/customers
- App Engine Technical Details

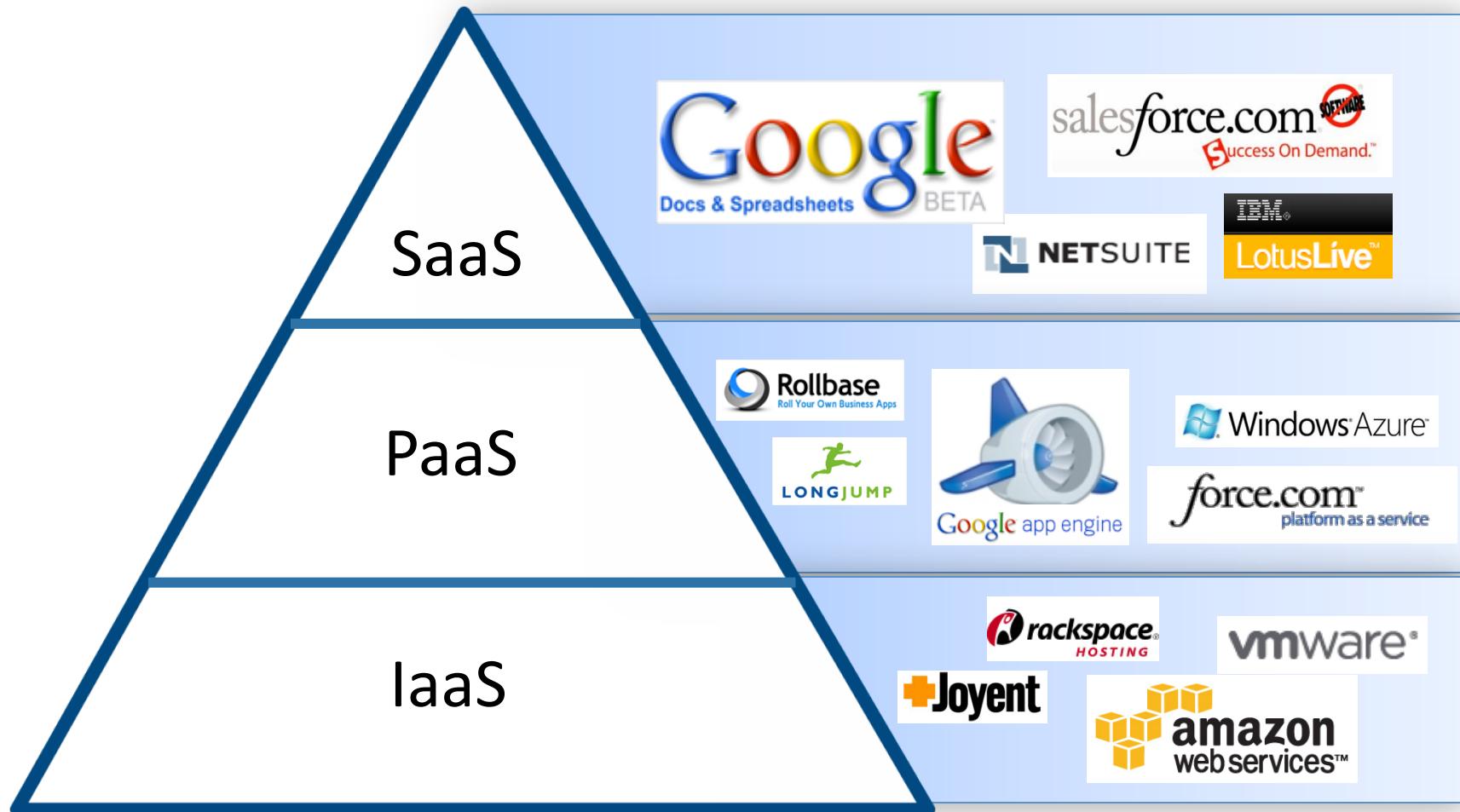


# Google App Engine

Build your own applications in Google's cloud

Google™

# Cloud Computing as Gartner Sees It



Source: Gartner AADI Summit Dec 2009

# Why Google App Engine?

- Easy to build
- Easy to maintain
- Easy to scale



# By the Numbers

588M<sup>+</sup>  
daily  
Appelopers  
Pageviews

It's never been easier to print what you want. Print your own photos, documents, and more with just one click. It's fast, easy, and reliable.

Drive money. Drive one environment.  
Tired of printing web pages only to find your printer is full of junk, empty space and other junk you don't want?  
PrintMeNow is a free online editor that lets you format any web page for printing in seconds!

Send faxes via PDF or Connect faxes to your location at your preferred time with our Fax Center... Read More

Add your blog to the fastest growing directory on the web!  
Get Found



What You'll  
The News  
The Last Word  
John Oliver  
The Special  
Trevor Noah  
The Daily Show  
The Real政治  
Piers Morgan  
Lorraine  
Piers Morgan  
Piers Morgan  
Piers Morgan  
The World

Chat live with your friends across the Web in Pix Chat!

New! [Get Started](#)



Pixverse - Now part of FOX's Content Unit  
© 2007-2008 Pixverse Inc. All Rights Reserved.



Watch Full Episodes

8: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

10: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

11: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

12: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

13: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

14: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

15: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

16: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

17: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

18: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

19: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

20: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

21: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

22: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

23: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

24: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

25: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

26: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

27: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

28: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

29: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

30: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

31: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

32: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

33: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

34: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

35: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

36: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

37: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

38: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

39: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

40: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

41: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

42: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

43: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

44: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

45: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

46: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

47: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

48: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

49: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

50: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

51: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

52: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

53: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

54: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

55: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

56: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

57: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

58: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

59: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

60: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

61: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

62: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

63: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

64: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

65: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

66: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

67: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

68: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

69: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

70: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

71: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

72: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

73: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

74: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

75: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

76: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

77: Hell's Kitchen - Day 16, 1:45 Eastern

2: Chopped, Season 11, Ep 15, 4:45 ET

Rating: 0.00/0.00

# Some App Engine Partners



Help Widget for Your Website



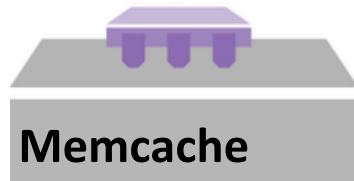
# App Engine Details

# Cloud Development in a Box

- Downloadable SDK
- Application runtimes
  - Java, Python
- Local development tools
  - Eclipse plugin, AppEngine Launcher
- Specialized application services
- Cloud based dashboard
- Ready to scale
- Built in fault tolerance, load balancing



# Specialized Services



**Memcache**



**Datastore**



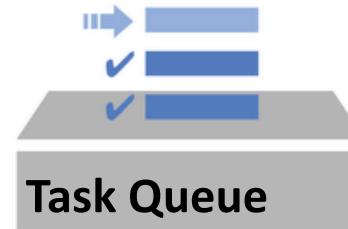
**URL Fetch**



**Mail**



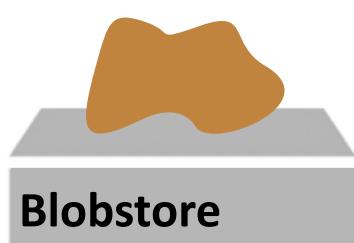
**XMPP**



**Task Queue**



**Images**

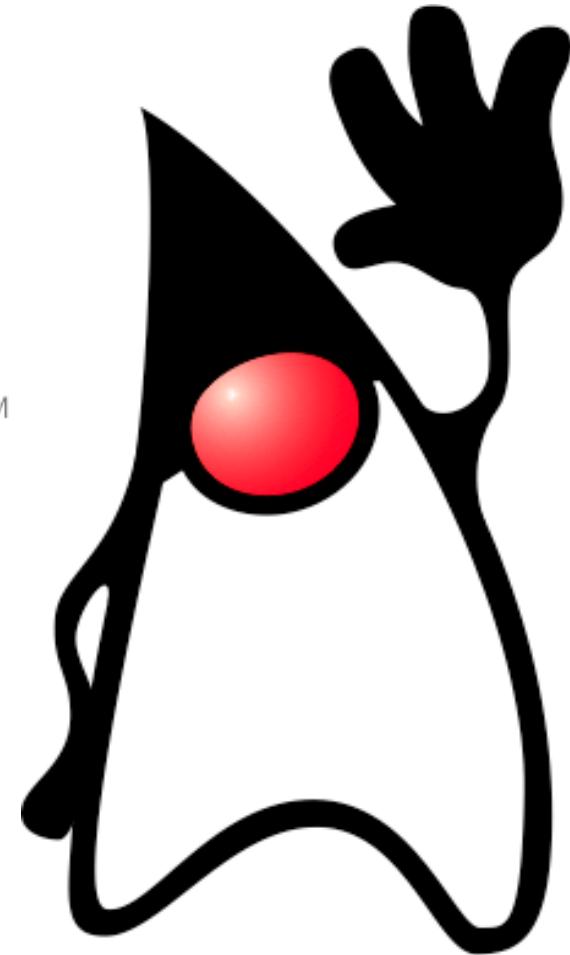


**Blobstore**



**User Service**

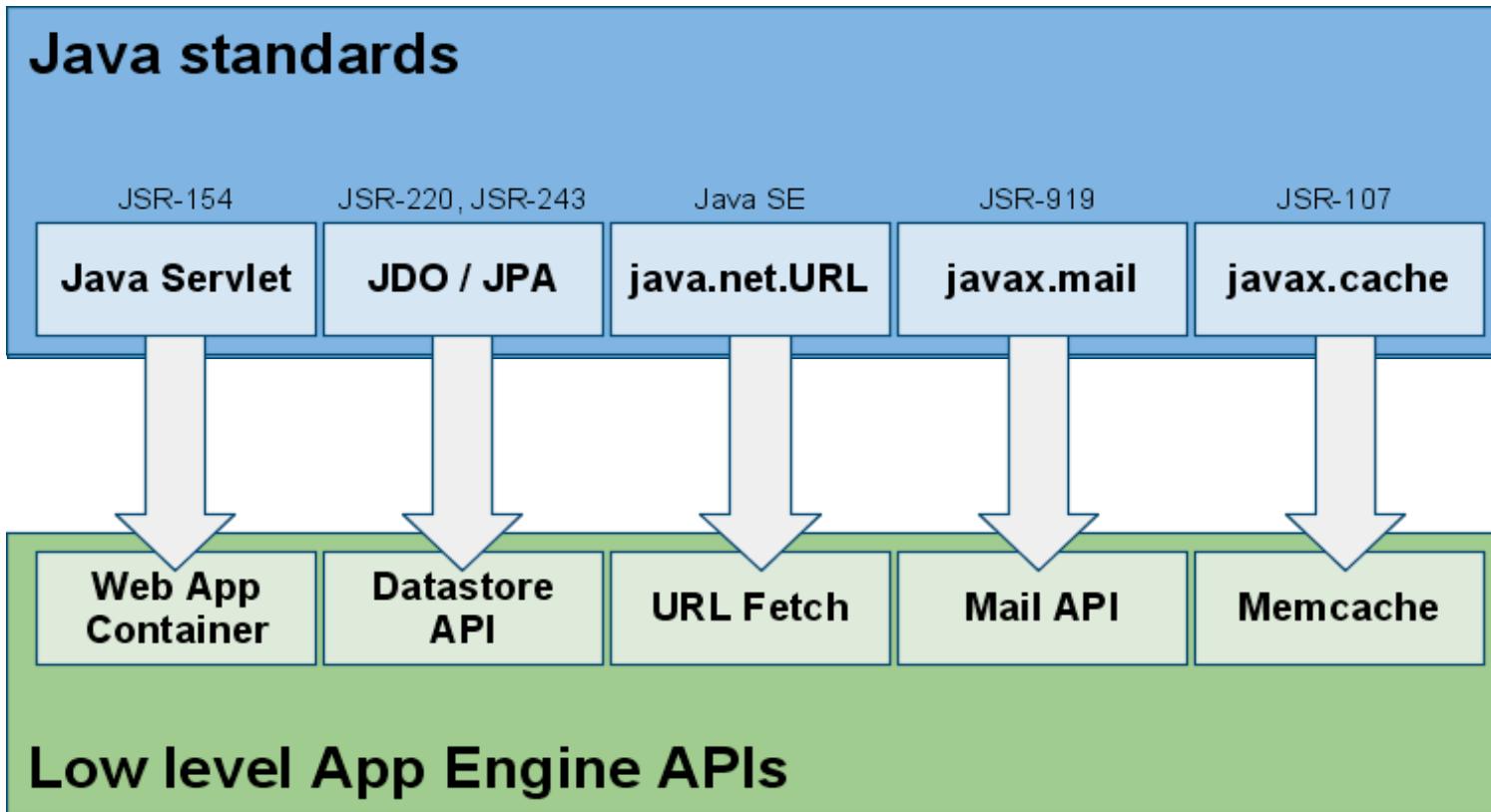
# Language Runtimes



Duke, the Java mascot  
Copyright © Sun Microsystems Inc., all rights reserved.

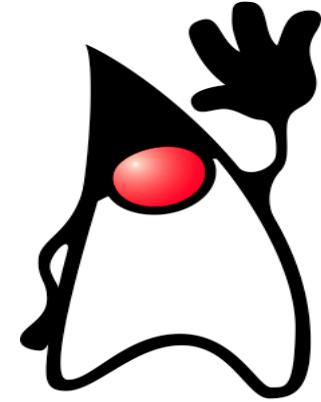


# Ensuring Portability



# Extended Language support through JVM

- Java
- Scala
- JRuby (Ruby)
- Groovy
- Quercus (PHP)
- Rhino (JavaScript)
- Jython (Python)



Duke, the Java mascot  
Copyright © Sun Microsystems Inc., all rights reserved.

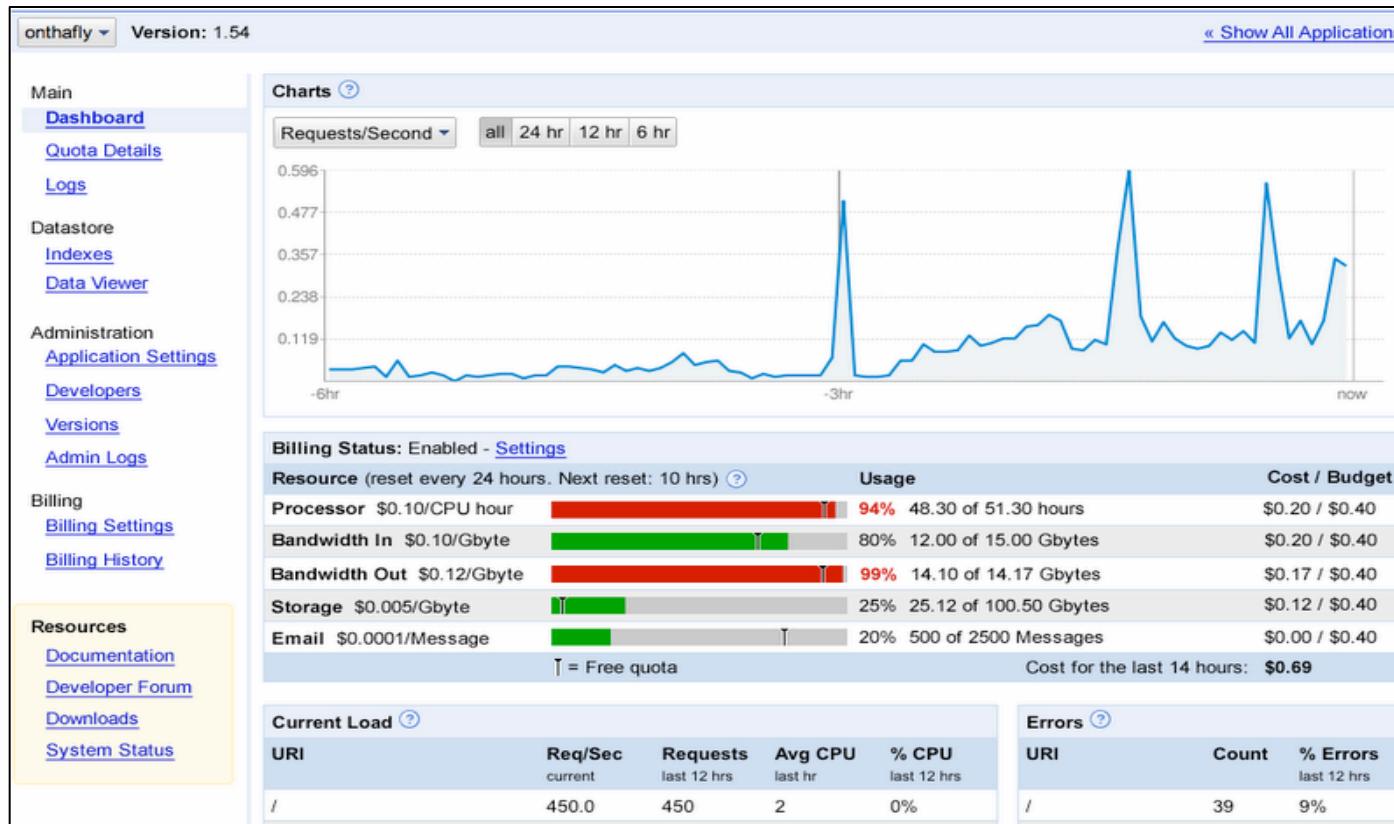
# Always free to get started

- ~5M pageviews/month
- 6.5 CPU hrs/day
- 1 GB storage
- 650K URL Fetch calls/day
- 2,000 recipients emailed
- 1 GB/day bandwidth
- 100,000 tasks enqueued
- 650K XMPP messages/day

# Application Platform Management



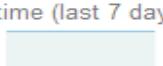
# App Engine Dashboard



# App Engine Health History

Google App Engine [Help | Admin Console](#)

## System Status

Current Availability **100%** Uptime (last 7 days)  Read latency (today)  Write latency (today) 

	03/20/09	03/21/09	03/22/09	03/23/09	03/24/09	03/25/09	03/26/09	03/27/09	Now
<b>Datastore</b>	✓	✓	✓	✓	✓	✓	✓	✓	Normal
<b>Images</b>	✓	✓	✓	✓	✓	✓	✓	✓	Normal
<b>Mail</b>	✓	✓	✓	✓	✓	✓	✓	✓	Normal
<b>Memcache</b>	✓	✓	✓	✓	✓	✓	✓	✓	Normal
<b>Serving</b>	✓	✓	✓	✓	✓	✓	✓	✓	Normal
<b>Urlfetch</b>	✓	✓	✓	✓	✓	✓	✓	✓	Normal
<b>Users</b>	✓	✓	✓	✓	✓	✓	✓	✓	Normal

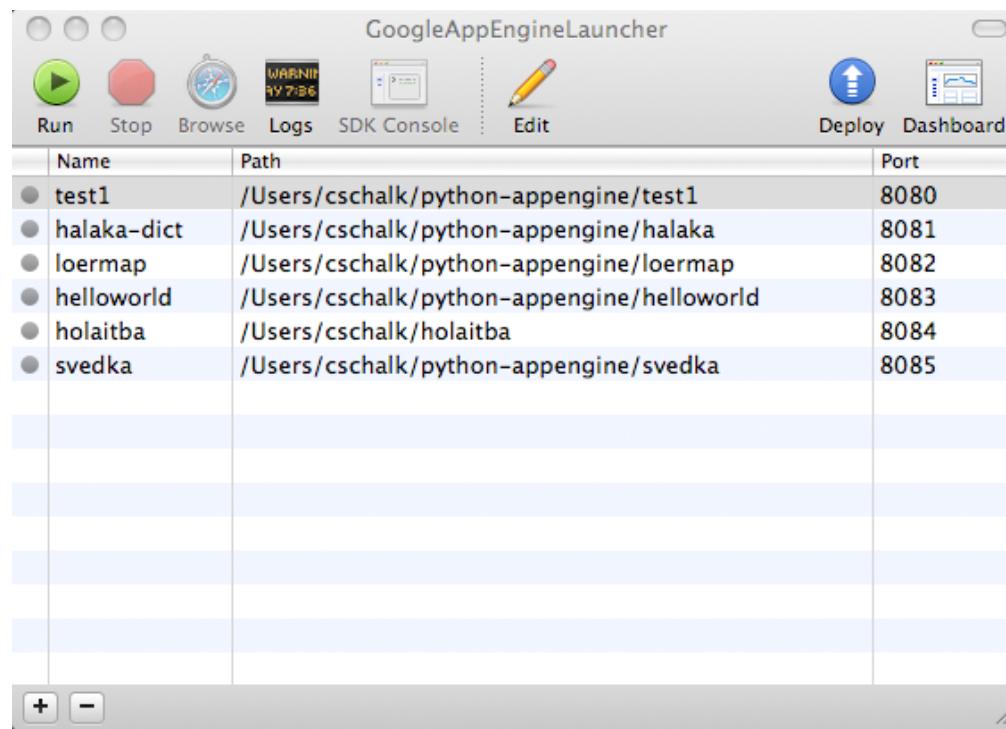
The following symbols signify the most severe issue (if any) encountered during that day. Click a symbol in the table above to view a day's performance graphs.

✓ No issues or minor performance issues   ? Investigating   ! Service disruption   🤔 Unknown

# Development Tools for App Engine



# Google App Engine Launcher



# SDK Console

The screenshot shows a web browser window titled "hellofbtest1 Development Console". The URL in the address bar is "http://localhost:8086/\_ah/admin/queues". The main content area is titled "Task Queues" and displays the following information:

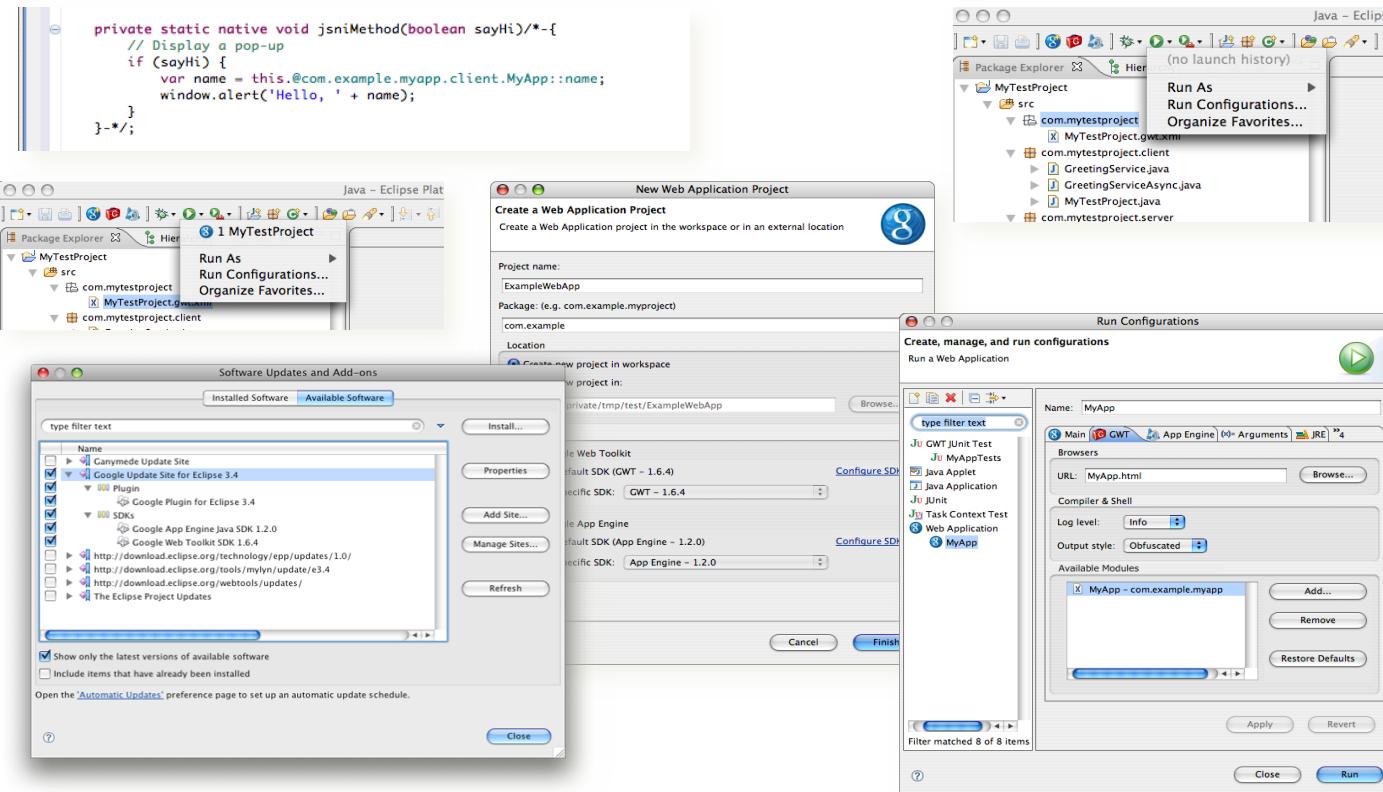
Tasks will not run automatically. Select a queue to run tasks manually.

Queue Name	Maximum Rate	Bucket Size	Oldest Task (UTC)	Tasks in Queue	Action
default	5.00/s	5	None	0	Flush Queue

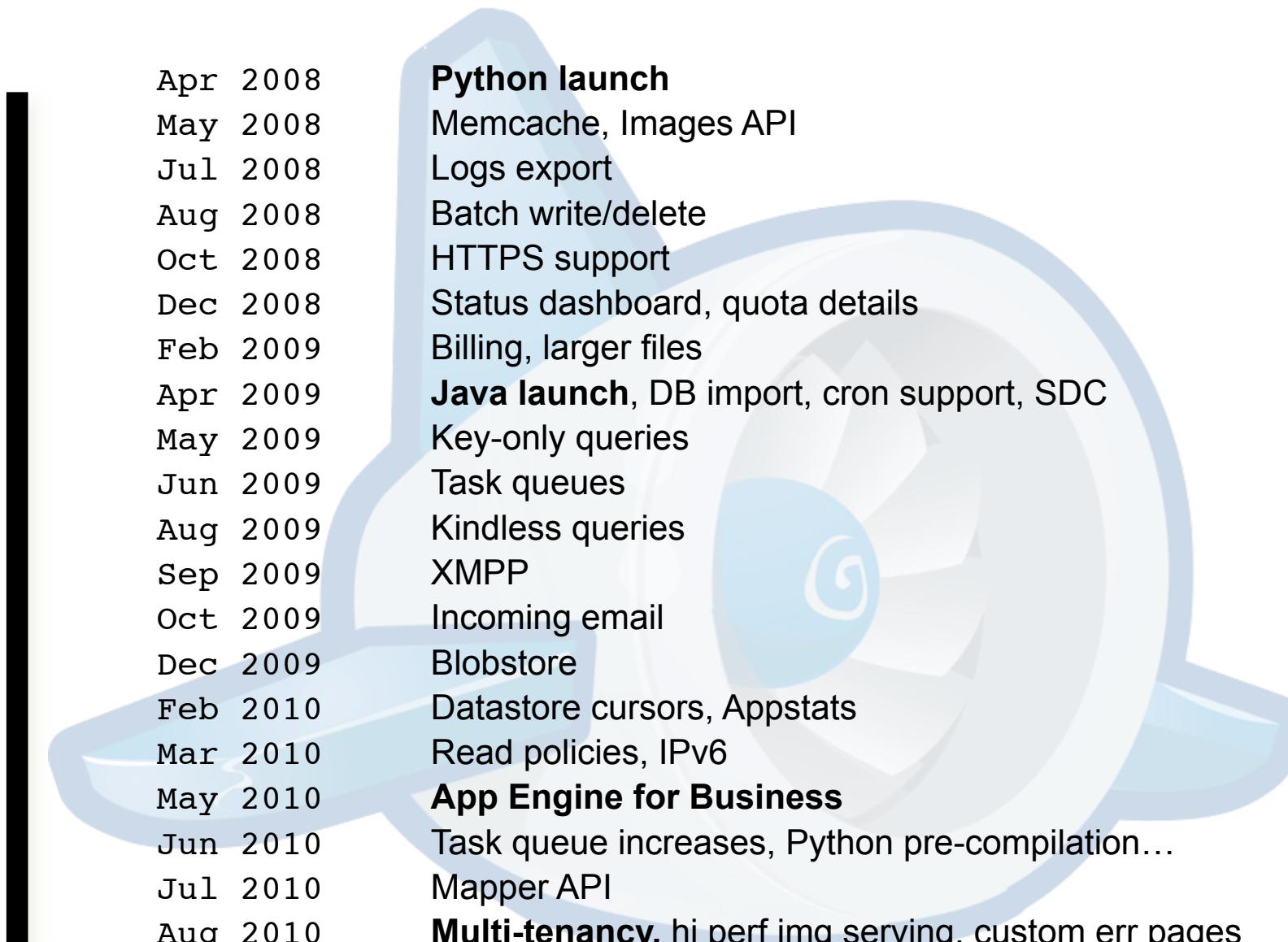
The sidebar on the left contains links: Datastore Viewer, Interactive Console, Memcache Viewer, Task Queues (which is selected), Cron Jobs, XMPP, and Inbound Mail.

At the bottom of the page, there is a copyright notice: ©2009 Google

# Google Plugin for Eclipse



# Two+ years in review



Apr 2008	<b>Python launch</b>
May 2008	Memcache, Images API
Jul 2008	Logs export
Aug 2008	Batch write/delete
Oct 2008	HTTPS support
Dec 2008	Status dashboard, quota details
Feb 2009	Billing, larger files
Apr 2009	<b>Java launch</b> , DB import, cron support, SDC
May 2009	Key-only queries
Jun 2009	Task queues
Aug 2009	Kindless queries
Sep 2009	XMPP
Oct 2009	Incoming email
Dec 2009	Blobstore
Feb 2010	Datastore cursors, Appstats
Mar 2010	Read policies, IPv6
May 2010	<b>App Engine for Business</b>
Jun 2010	Task queue increases, Python pre-compilation...
Jul 2010	Mapper API
Aug 2010	<b>Multi-tenancy</b> , hi perf img serving, custom err pages
Oct 2010	Instances Console, Delete Kind/App Data

# Introducing App Engine for Business



App Engine for Business

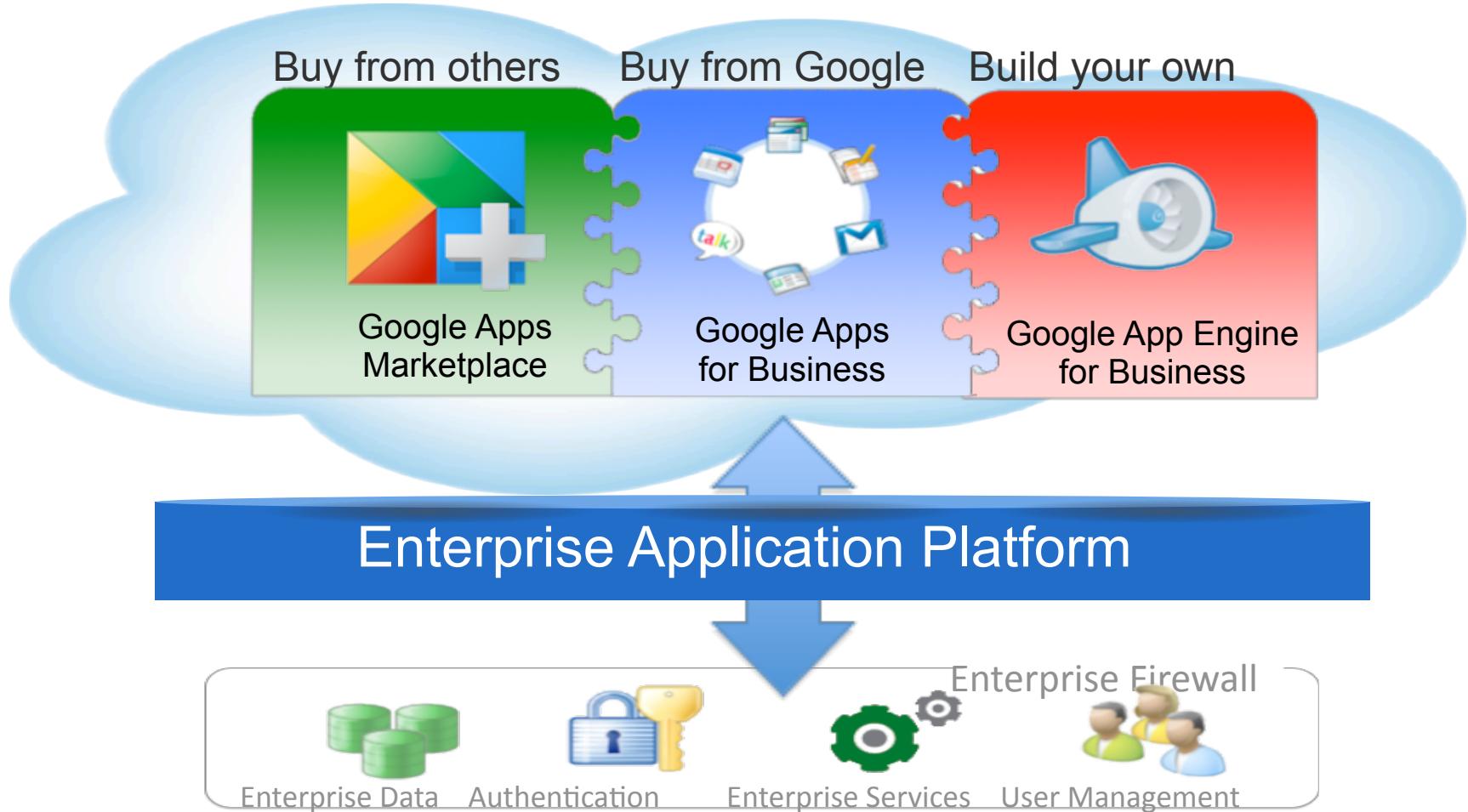
*Same scalable cloud platform, but designed for the Enterprise*

# Google App Engine for Business Details

- Enterprise application management
  - Centralized domain console (**preview available**)
- Enterprise reliability and support
  - 99.9% Service Level Agreement
  - Direct support
- Hosted SQL
  - Relational SQL database in the cloud (**preview available**)
- SSL on your domain
- Extremely Secure by default
  - Integrated Single Sign On (SSO)
- Pricing that makes sense
  - Apps cost \$8 per user, up to \$1000 max per month



# Enterprise App Development with Google



# App Engine for Business Roadmap



Enterprise Administration Console	Preview (signups available)
Direct Support	Preview (signups available)
Hosted SQL	Preview (signups available)
Service Level Agreement	Available Q4 2010 (Draft published)
Enterprise billing	Available Q4 2010
Custom Domain SSL	Limited Release EOY 2010

# App Engine Resources

Get started with App Engine

- <http://code.google.com/appengine>

Read up on App Engine for Business and become a trusted tester

- <http://code.google.com/appengine/business>

- [bit.ly/gae4btt](http://bit.ly/gae4btt) <- sign up!

# Part II - Google's new Cloud Technologies

## Topics covered

- Google Storage for Developers
- Prediction API (machine learning)
- BigQuery

# Google Storage for Developers

Store your data in Google's cloud



# What Is Google Storage?

- Store your data in Google's cloud
  - any format, any amount, any time
- You control access to your data
  - private, shared, or public
- Access via Google APIs or 3rd party tools/libraries

# Sample Use Cases

**Static content hosting**

e.g. static html, images, music, video

**Backup and recovery**

e.g. personal data, business records

**Sharing**

e.g. share data with your customers

**Data storage for applications**

e.g. used as storage backend for Android, AppEngine, Cloud based apps

**Storage for Computation**

e.g. BigQuery, Prediction API

# Google Storage Benefits



**High Performance and Scalability**  
Backed by Google infrastructure

**Strong Security and Privacy**  
Control access to your data



**Easy to Use**  
Get started fast with Google & 3rd party tools

# Google Storage Technical Details

- RESTful API
  - Verbs: GET, PUT, POST, HEAD, DELETE
  - Resources: identified by URI
  - Compatible with S3
- Buckets
  - Flat containers
- Objects
  - Any type
  - Size: 100 GB / object
- Access Control for Google Accounts
  - For individuals and groups
- Two Ways to Authenticate Requests
  - Sign request using access keys
  - Web browser login

# Performance and Scalability

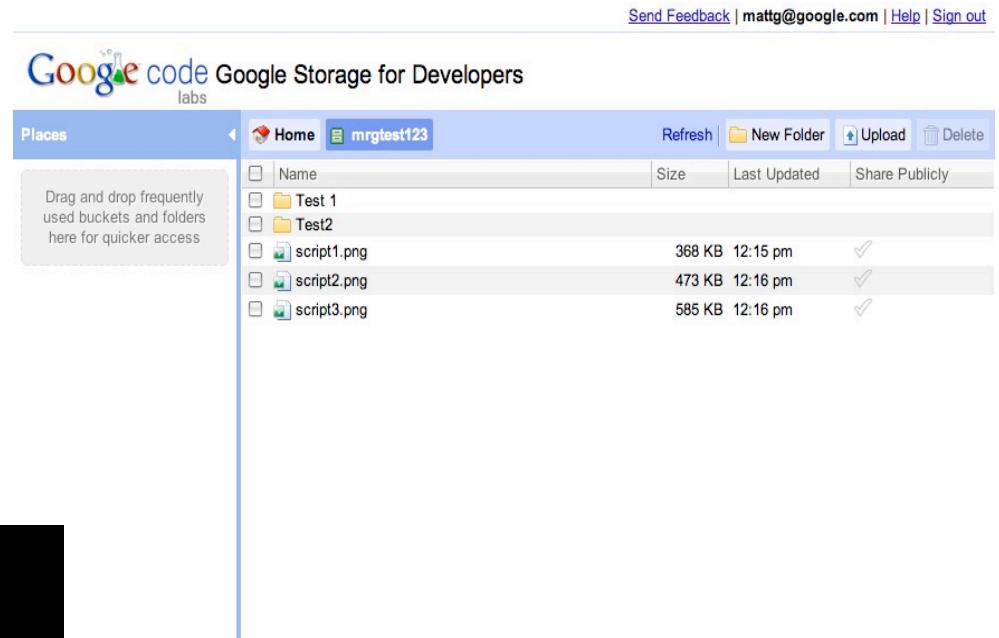
- Objects of any type and 100 GB / Object
- Unlimited numbers of objects, 1000s of buckets
- All data replicated to multiple US data centers
- Utilizes Google's worldwide network for data delivery
- Only you can use bucket names with your domain names
- Read-your-writes data consistency
- Range Get

# Demo

- Tools:
  - GS Manager
  - GSUtil
- Upload / Download

```
dhcp-172-19-3-109:~ wferrell$ gsutil
SYNOPSIS
  gsutil [-d] [-h header]... command args
  -d option shows HTTP protocol detail.
  -h option allows you to specify additional HTTP headers, for example:
    gsutil -h "Cache-Control:public,max-age=3600" -h "Content-Type:gzip" cp * g
  s://bucket

  Commands:
    Concatenate object content to stdout:
      cat [-h] uri...
      -h Prints short header for each object.
  Copy objects:
    cp [-a canned_acl] [-t] [-z ext1,ext2,...] src_uri dst_uri
    - or -
    cp [-a canned_acl] [-t] [-z extensions] uri... dst_uri
      -a Sets named canned_acl when uploaded objects created (list below).
      -t Sets MIME type based on file extension.
      -z 'txt,html' Compresses file uploads with the given extensions.
  Get ACL XML for a bucket or object (save and edit for "setacl" command):
```



# Google Storage usage within Google

Google BigQuery

Google  
Prediction API



double  
click

Partner Reporting



Partner Reporting

# Some Early Google Storage Adopters



# Google Storage - Pricing

- Storage
  - \$0.17/GB/Month
- Network
  - Upload - \$0.10/GB
  - Download
    - \$0.15/GB Americas / EMEA
    - \$0.30/GB APAC
- Requests
  - PUT, POST, LIST - \$0.01 / 1000 Requests
  - GET, HEAD - \$0.01 / 10000 Requests

# Google Storage - Availability

- Limited preview in US currently
  - 100GB free storage and network from Google per account
  - Sign up for waitlist at <http://code.google.com/apis/storage/>
- Note: Non US preview available on case-by-case basis

# Google Prediction API

Google's prediction engine in the cloud



# Introducing the Google Prediction API

- Google's sophisticated machine learning technology
- Available as an on-demand RESTful HTTP web service



# How does it work?

The Prediction API finds relevant ***features*** in the sample data during training.

"english"	The quick brown fox jumped over <b>the</b> lazy dog.
"english"	To err <b>is</b> human, but <b>to</b> really foul things up you need a computer.
"spanish"	No hay mal <b>que por</b> bien <b>no</b> venga.
"spanish"	La tercera <b>es la</b> vencida.

The Prediction API later searches for those ***features*** during prediction.

?	To be or not <b>to</b> be, that <b>is the</b> question.
?	<b>La</b> fe mueve montañas.

# A virtually endless number of applications...

				
Customer Sentiment	Transaction Risk	Species Identification	Message Routing	Diagnostics
				
Churn Prediction	Legal Docket Classification	Suspicious Activity	Work Roster Assignment	Inappropriate Content
				
Recommend Products	Political Bias	Uplift Marketing	Email Filtering	Career Counselling

... and many more ...

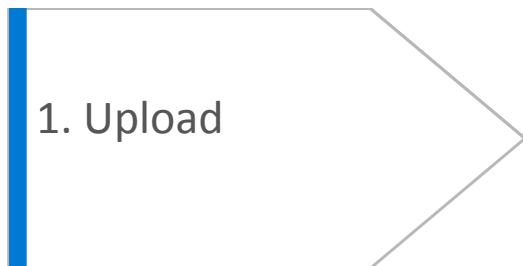
# A Prediction API Example

Automatically categorize and respond to emails by language

- **Customer:** ACME Corp, a multinational organization
- **Goal:** Respond to customer emails in their language
- **Data:** Many emails, tagged with their languages
- **Outcome:** Predict language and respond accordingly

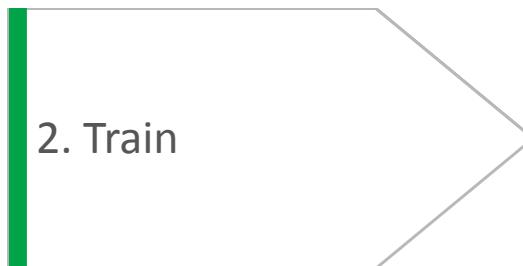
# Using the Prediction API

A simple three step process...



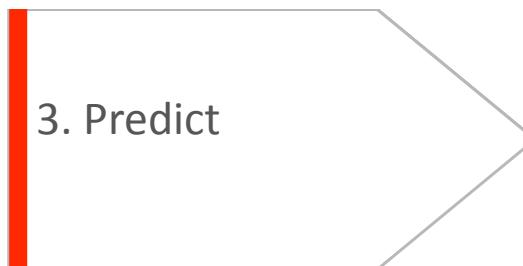
1. Upload

Upload your training data to  
**Google Storage**



2. Train

Build a model from your data



3. Predict

Make new predictions

## Step 1: Upload

Upload your training data to Google Storage

- Training data: **outputs** and **input features**
- Data format: comma separated value format (CSV)

"english","To err is human, but to really ..."

"spanish","No hay mal que por bien no venga."

...

Upload to Google Storage

`gsutil cp ${data} gs://yourbucket/${data}`

## **Step 2: Train**

Create a new model by training on data

To train a model:

**POST prediction/v1.1/training?data=mybucket%2Fmydata**

Training runs asynchronously. To see if it has finished:

**GET prediction/v1.1/training/mybucket%2Fmydata**

```
{"data":{  
  "data":"mybucket/mydata",  
  "modelinfo":"estimated accuracy: 0.xx"}}}
```

## Step 3: Predict

Apply the trained model to make predictions on new data

**POST prediction/v1.1/query/mybucket%2Fmydata/predict**

```
{ "data":{  
  "input": { "text" : [  
    "J'aime X! C'est le meilleur" ]}}}
```

## Step 3: Predict

Apply the trained model to make predictions on new data

**POST prediction/v1.1/query/mybucket%2Fmydata/predict**

```
{ "data":{  
    "input": { "text" : [  
        "J'aime X! C'est le meilleur" ]}}}
```

```
{ data : {  
    "kind" : "prediction#output",  
    "outputLabel":"French",  
    "outputMulti": [  
        {"label":"French", "score": x.xx}  
        {"label":"English", "score": x.xx}  
        {"label":"Spanish", "score": x.xx}]]}
```

## Step 3: Predict

Apply the trained model to make predictions on new data

An example using Python

```
import httplib

header = {"Content-Type" : "application/json"}

#...put new data in JSON format in params variable
conn = httplib.HTTPConnection("www.googleapis.com")conn.request("POST",
"/prediction/v1.1/query/mybucket%2Fmydata/predict", params, header)

print conn.getresponse()
```

# Prediction API Capabilities

## Data

- Input Features: numeric or unstructured text
- Output: up to hundreds of discrete categories

## Training

- Many machine learning techniques
- Automatically selected
- Performed asynchronously

## Access from many platforms:

- Web app from Google App Engine
- Apps Script (e.g. from Google Spreadsheet)
- Desktop app

# Prediction API v1.1 - new features

- Updated Syntax
- Multi-category prediction
  - Tag entry with multiple labels
- Continuous Output
  - Finer grained prediction rankings based on multiple labels
- Mixed Inputs
  - Both numeric and text inputs are now supported

*Can combine continuous output with mixed inputs*

# Google BigQuery

Interactive analysis of large datasets in Google's cloud

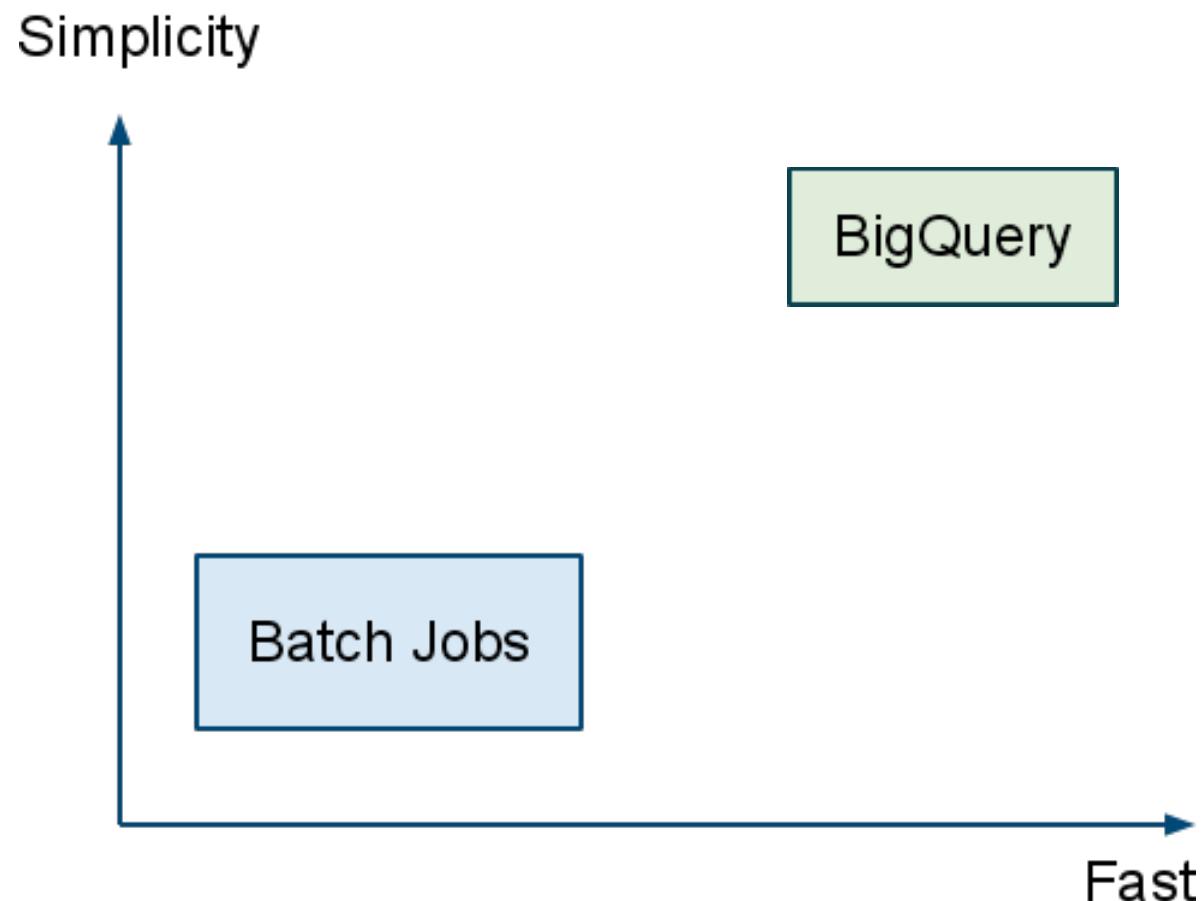


# Introducing Google BigQuery

- Google's large data adhoc analysis technology
  - Analyze massive amounts of data in seconds
- Simple SQL-like query language
- Flexible access
  - REST APIs, JSON-RPC, Google Apps Script

# Why BigQuery?

Working with large data is a challenge



## Many Use Cases ...



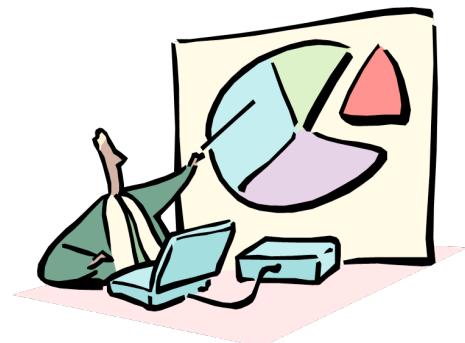
Interactive Tools



Spam



Trends  
Detection



Web Dashboards



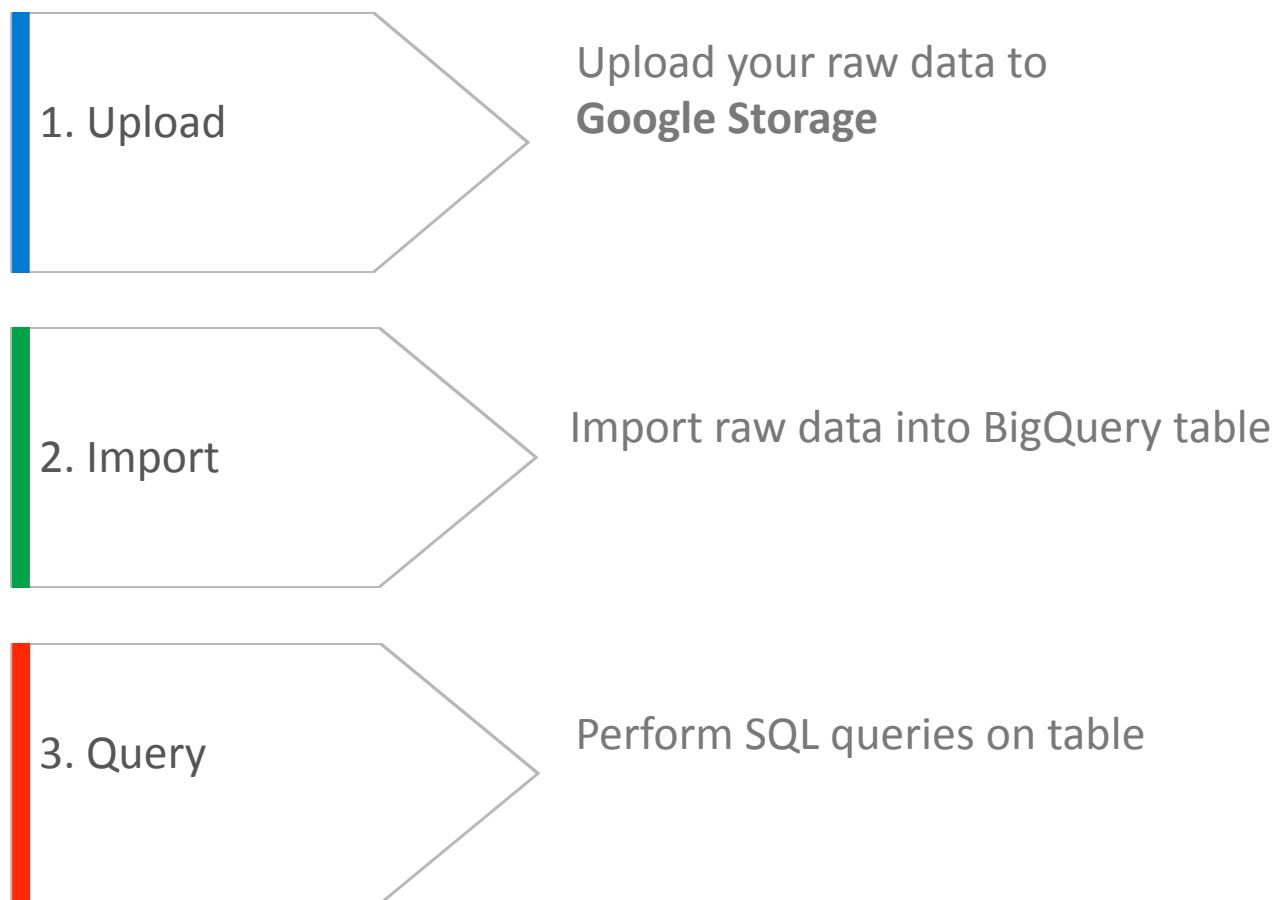
Network  
Optimization

# Key Capabilities of BigQuery

- Scalable: Billions of rows
- Fast: Response in seconds
- Simple: Queries in SQL
- Web Service
  - REST
  - JSON-RPC
  - Google App Scripts

# Using BigQuery

Another simple three step process...



# Writing Queries

Compact subset of SQL

- **SELECT ... FROM ...**
- **WHERE ...**
- **GROUP BY ... ORDER BY ...**
- **LIMIT ...;**

Common functions

- **Math, String, Time, ...**

Statistical approximations

- **TOP**
- **COUNT DISTINCT**

# BigQuery via REST

GET /bigquery/v1/tables/{table name}

GET /bigquery/v1/query?q={query}

Sample JSON Reply:

```
{  
  "results": {  
    "fields": { [  
      {"id":"COUNT(*)","type":"uint64"}, ... ]  
    },  
    "rows": [  
      {"f": [{"v": "2949"}, ...]},  
      {"f": [{"v": "5387"}, ...]}, ... ]  
  }  
}
```

Also supports JSON-RPC

# Security and Privacy

## Standard Google Authentication

- Client Login
- OAuth
- AuthSub

## HTTPS support

- protects your credentials
- protects your data

Relies on Google Storage to manage access

# Large Data Analysis Example

## Wikimedia Revision History

BigQuery

Sign out

Enter SQL query

```
SELECT TOP(title, 5), COUNT(*)
FROM [wikipedia]
WHERE wp_namespace = 0;
```

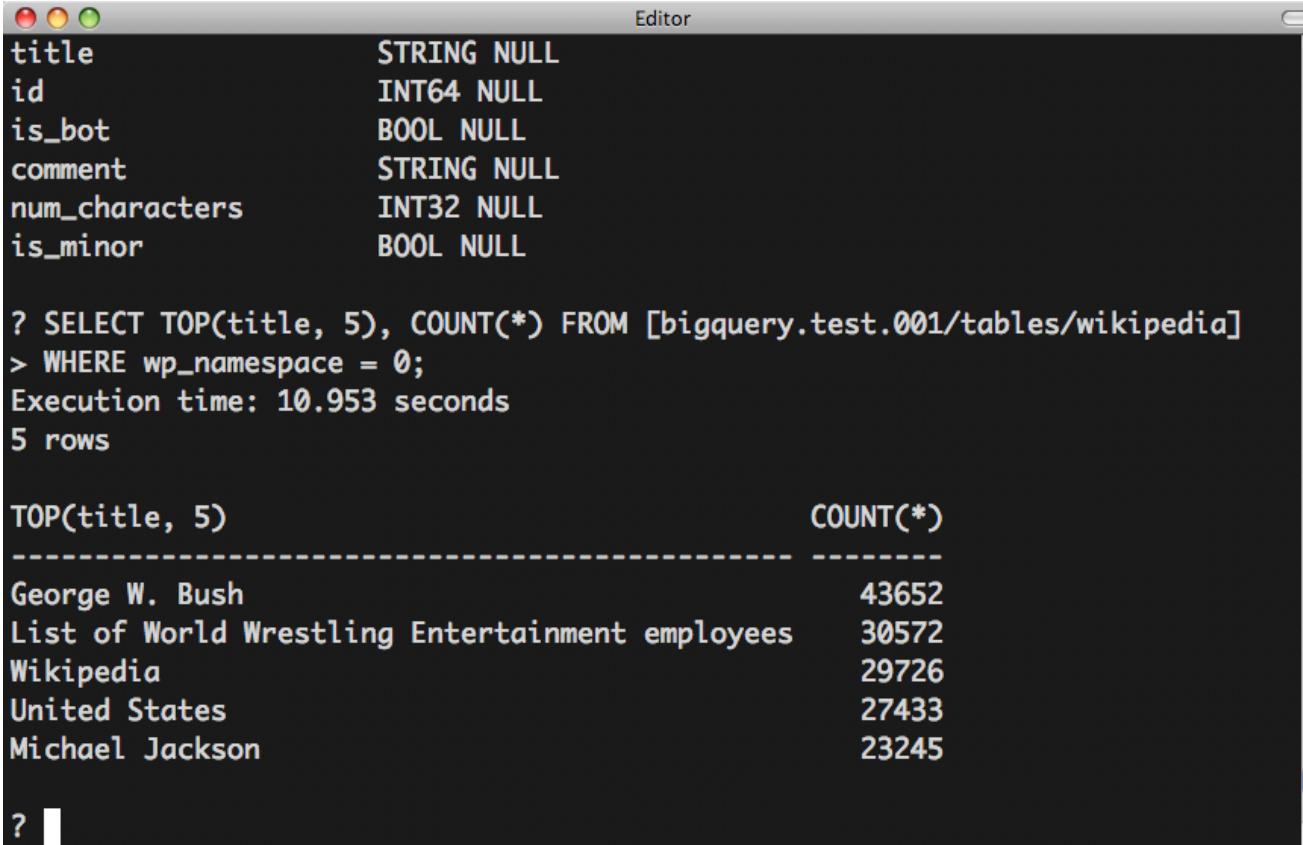
Run

Wikimedia Revision history data from: <http://download.wikimedia.org/enwiki/latest/enwiki-latest-pages-meta-history.xml.7z>

# Using BigQuery Shell

Python DB API 2.0 + B. Clapper's **sqlcmd**

<http://www.clapper.org/software/python/sqlcmd/>



```
Editor

title      STRING NULL
id        INT64 NULL
is_bot    BOOL NULL
comment   STRING NULL
num_characters  INT32 NULL
is_minor  BOOL NULL

? SELECT TOP(title, 5), COUNT(*) FROM [bigquery.test.001/tables/wikipedia]
> WHERE wp_namespace = 0;
Execution time: 10.953 seconds
5 rows

TOP(title, 5)          COUNT(*)
-----
George W. Bush          43652
List of World Wrestling Entertainment employees 30572
Wikipedia                29726
United States              27433
Michael Jackson            23245

?
```

# BigQuery from a Spreadsheet

# BigQuery from a Spreadsheet

AppScript BigQuery Demo

BigQuery

```
2 importData(concatenate('http://www.googleapis.com/bigquery/v1/query?q=',$B4))
```

Search Term	Num_Revs
Title	
Google	8755
Google search	4261
Google Earth	3874
Google Chrome	2687
Google Maps	2617
--	
--	
--	
--	
--	

Num\_Revs

Product	Num_Revs
Google	8755
Google search	4261
Google Earth	3874
Google Chrome	2687
Google Maps	2617

Wikipedia

=CONTINUE(A5, 1, 2)

Further info available at:

- Google Storage for Developers
  - <http://code.google.com/apis/storage>
- Prediction API
  - <http://code.google.com/apis/predict>
- BigQuery
  - <http://code.google.com/apis/bigquery>

# Recap

- Google App Engine
  - Google's PaaS cloud development platform
- Google App Engine for Business
  - *New* enterprise version of App Engine
- Google Storage
  - *New* high speed data storage on Google Cloud
- Prediction API
  - *New* machine learning technology able to predict outcomes based on sample data
- BigQuery
  - *New* service for Interactive analysis of very large data sets using SQL

# Q&A

# Thank You!

**Chris Schalk**  
Google Developer  
Advocate

<http://twitter.com/cschalk>