

# The Cloud and Google

International Supercomputing Conference '09

June 24th, 2009

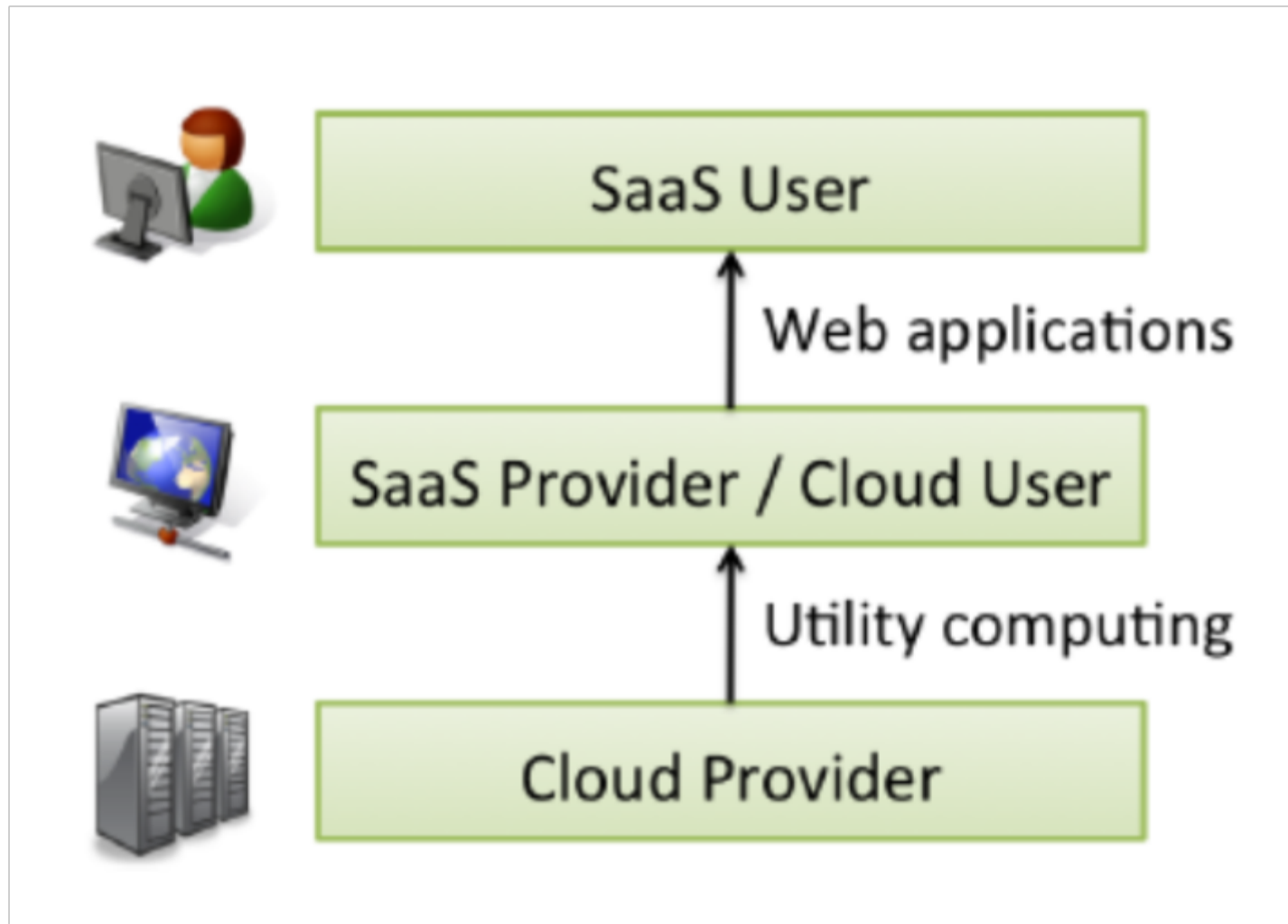
Dr. Robin Williamson  
Google Inc.



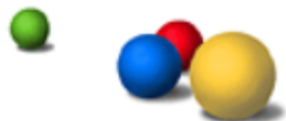
- The Cloud to Google
- Technology Stack
  - Hardware computing platform
  - Distributed systems software infrastructure
  - Products
- Development Platform and Tools
  - Google App Engine



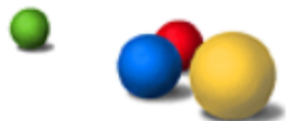
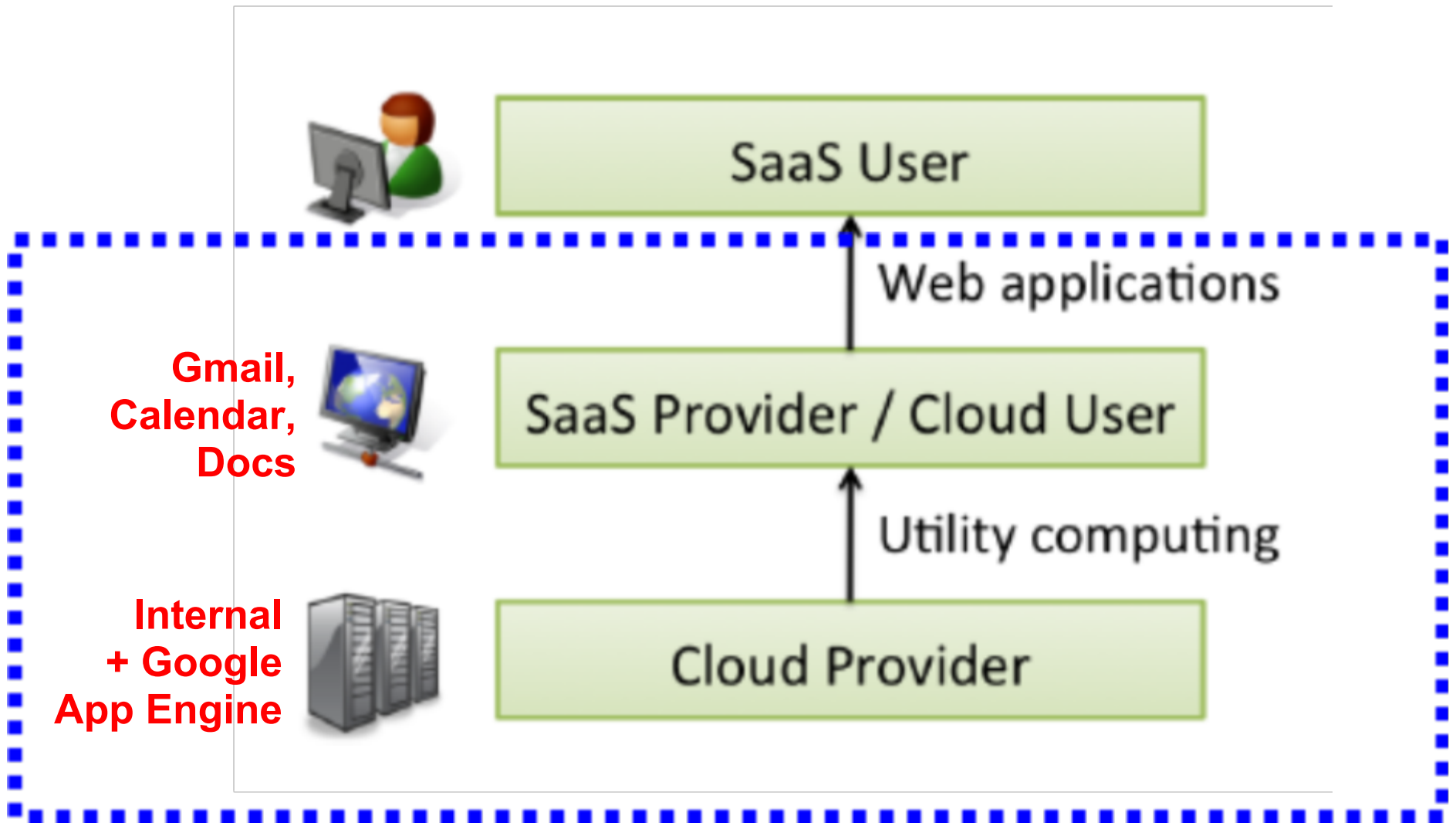
# Berkeley View of Cloud Computing



[Above the Clouds: A Berkeley View of Cloud Computing, Armbrust et al. 2009]

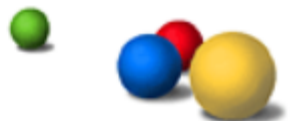


# Where We Play



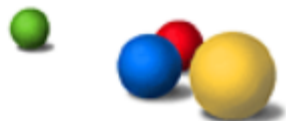
# Google's Mission Statement

To organize the world's information and  
make it universally accessible and  
useful.

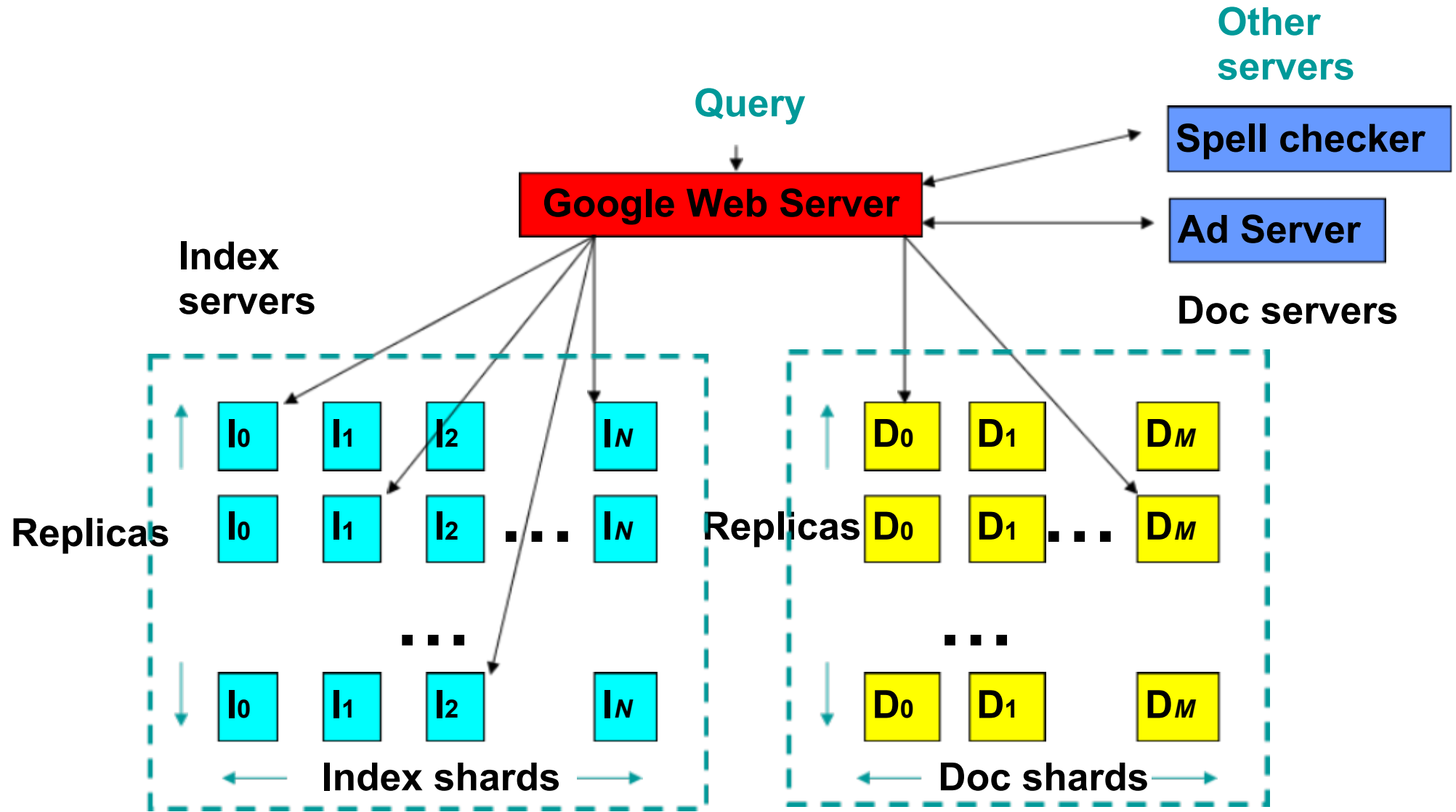


# The *World's* Information

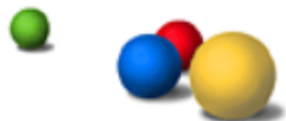
- There's *lots* of data—100s of TBs just for the Web
- Much of the data is common across large numbers of users
- Bandwidth is slow and expensive in the last mile (and improving more slowly than other dimensions)
- **Analyzing, transforming, querying all best done centrally (adjacent to the data)**



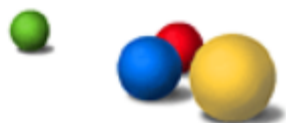
# Query Serving Infrastructure



Elapsed time: 250ms, machines involved: 1000+



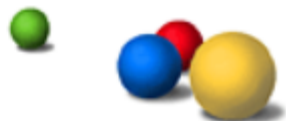
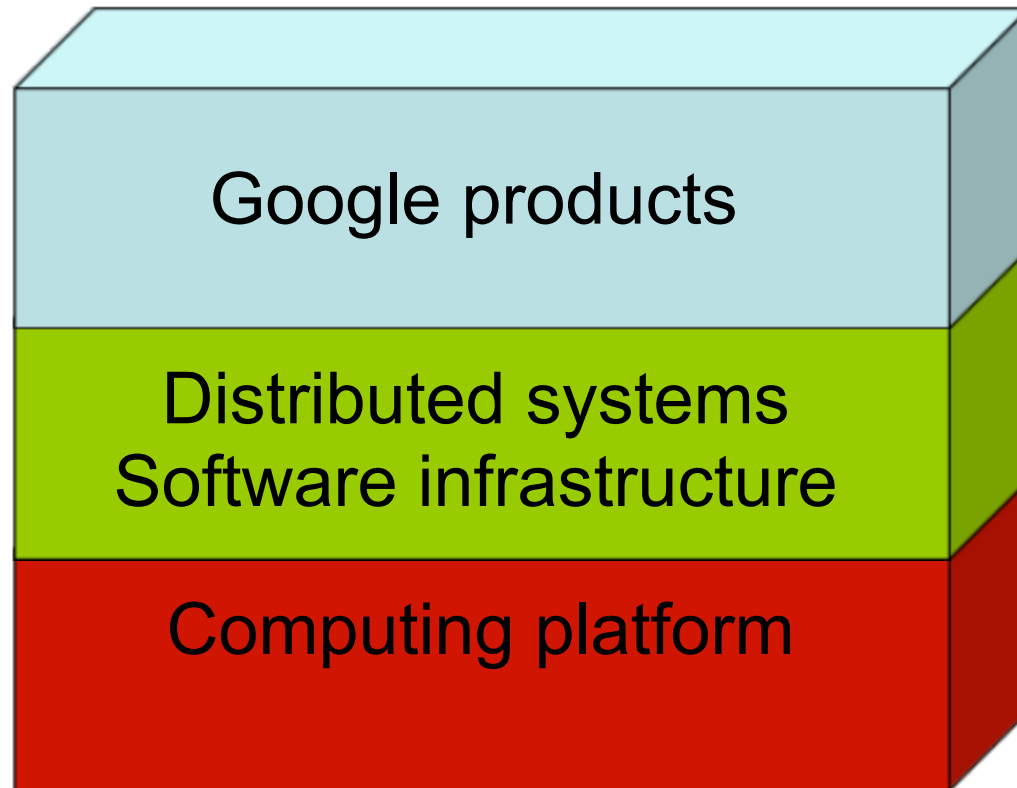
- The Cloud to Google
- **Technology Stack**
  - Hardware computing platform
  - Distributed systems software infrastructure
  - Products
- Development Platform and Tools
  - Google App Engine





# Technology Stack

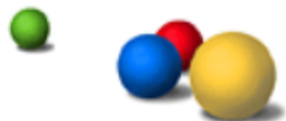
- Innovation at all layers...



# Computing Platform

- Single-threaded performance matters less
- Moore's law manifests as more cores
- The computer is the datacenter

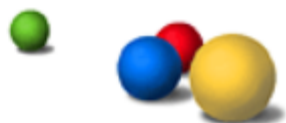
- 
- Cost-efficiency
  - Server design
  - Networking
  - Datacenter technology



# Higher-level Programming Abstractions

## **Systems Infrastructure**

- Google file system (GFS)
- MapReduce
- BigTable



# System Infrastructure

- **Google File System (GFS):**

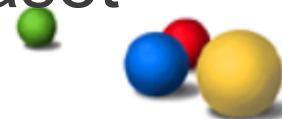
- Fault tolerant distributed disk storage
- Optimized for high-bandwidth sequential read/writes

- **BigTable:**

- A large-scale storage system for semi-structured data
- Database-like model, but data stored on thousands of machines..

- **MapReduce:**

- A programming model and library to simplify large-scale computations on large clusters
- Distribute data among many machines, execute same computation at each machine on its dataset



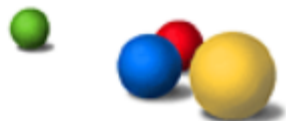
# Top of the Stack

## Google products

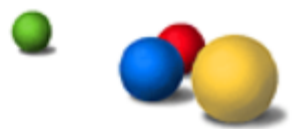
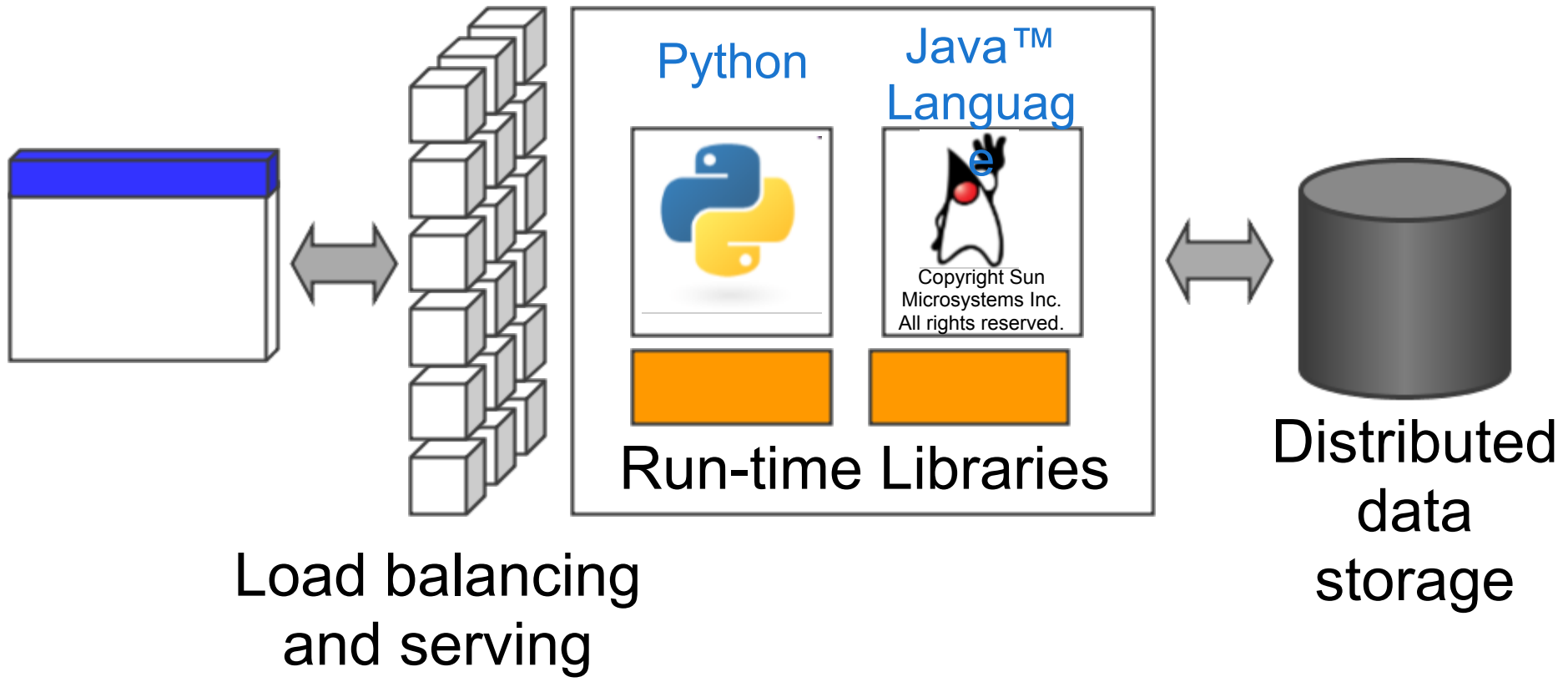
- Search (Web, Books, Products, ...)
- Ads
- Gmail, Calendar
- Docs, Sheets
- Machine Translation



- The Cloud to Google
- Technology Stack
  - Hardware computing platform
  - Distributed systems software infrastructure
  - Products
- **Development Platform and Tools**
  - Google App Engine

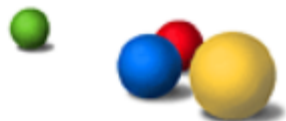


## Your code running on Google infrastructure



# Key Features

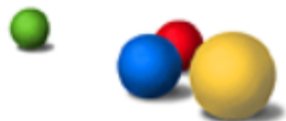
- Python or Java™ Language source code
- Develop locally, deploy to Cloud seamlessly
- Write once, scale automatically
- Local SDK & Eclipse Plugin
- Free quota of 1 GB storage and ~5M pageviews / month



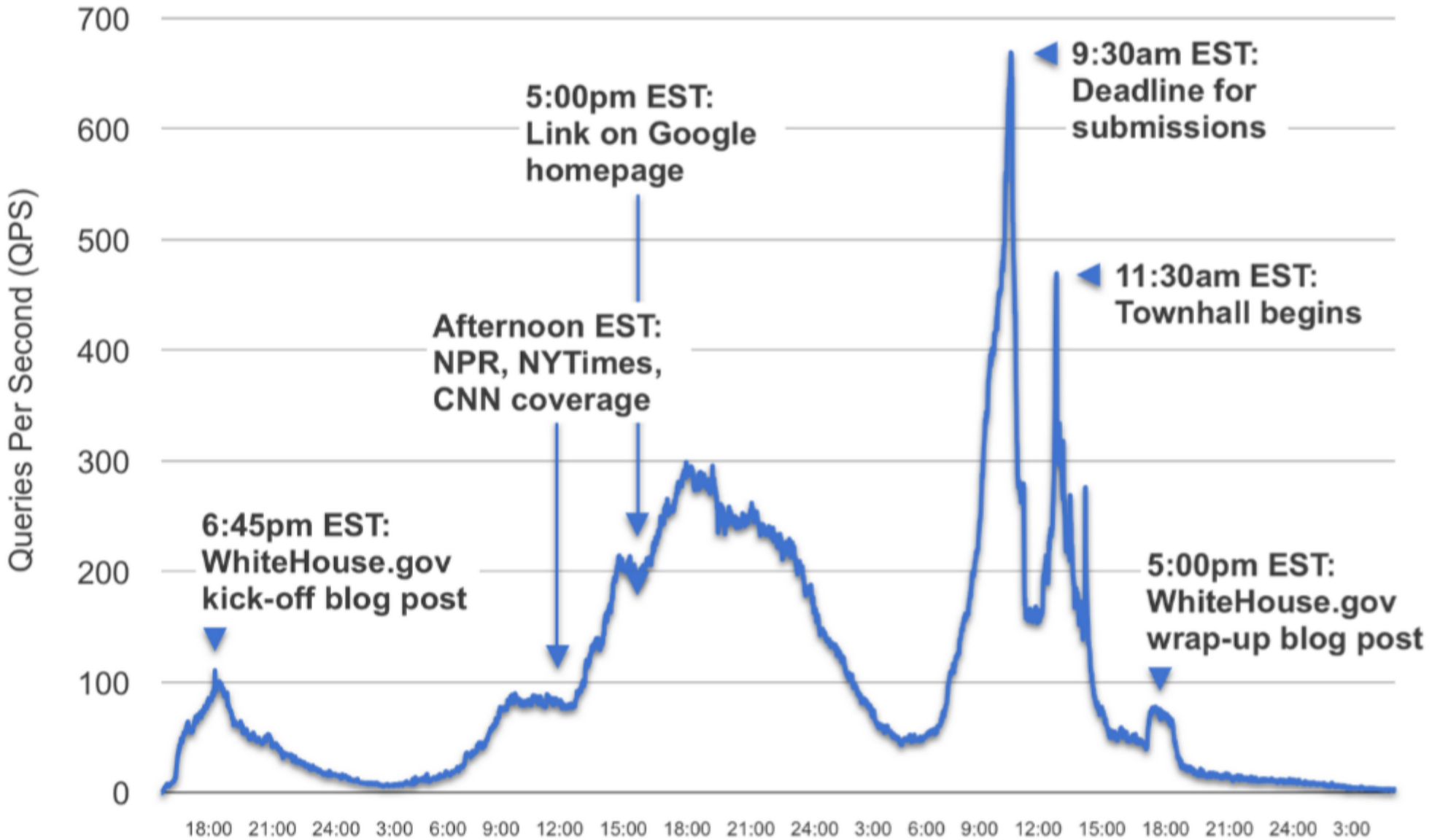


# Programming & Run-time Model

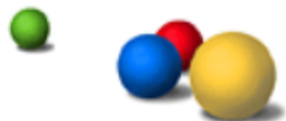
- Responds to HTTP requests
- A programming platform, not “raw iron”
- API support for
  - User login and identity
  - Persistent state (on top of Bigtable, not RDBMS)
  - memcache
  - Mail, Images, URL Fetch
  - Django Templates / JSP



# Does it Scale?



[WhiteHouse.gov/openforquestions](http://WhiteHouse.gov/openforquestions)



# 5M Free PVs + Pay as you Go



# Admin Interface

onthafly ▾ Version: 1.54

[« Show All Applications](#)

## Main

[Dashboard](#)

[Quota Details](#)

[Logs](#)

## Datastore

[Indexes](#)

[Data Viewer](#)

## Administration

[Application Settings](#)

[Developers](#)

[Versions](#)

[Admin Logs](#)

## Billing

[Billing Settings](#)

[Billing History](#)

## Resources

[Documentation](#)

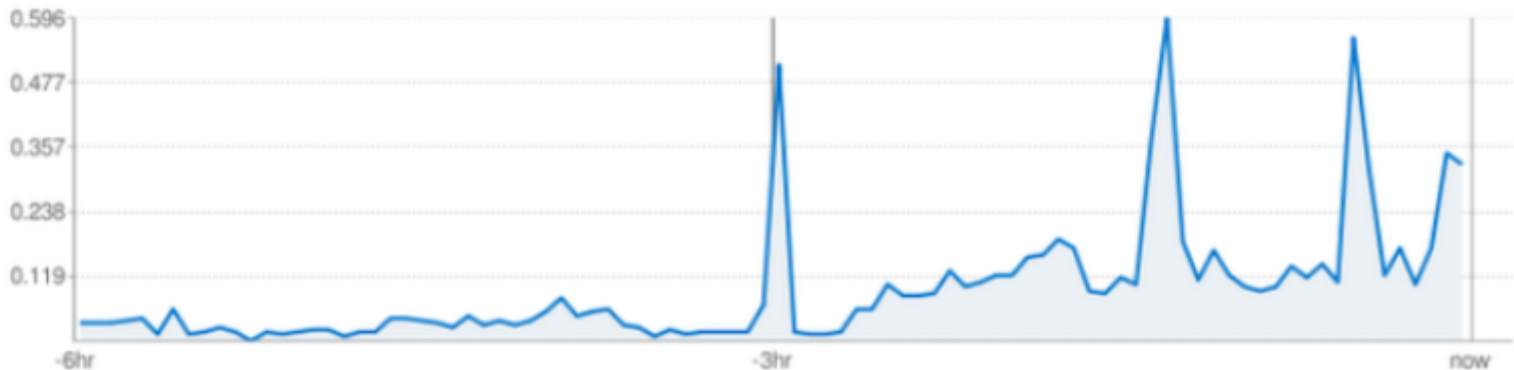
[Developer Forum](#)

[Downloads](#)

[System Status](#)

## Charts ?

Requests/Second ▾ all 24 hr 12 hr 6 hr



## Billing Status: Enabled - [Settings](#)

Resource (reset every 24 hours. Next reset: 10 hrs) <span>?</span>	Usage	Cost / Budget
Processor \$0.10/CPU hour	<div style="width: 94%; background-color: red;">94%</div> 48.30 of 51.30 hours	\$0.20 / \$0.40
Bandwidth In \$0.10/Gbyte	<div style="width: 80%; background-color: green;">80%</div> 12.00 of 15.00 Gbytes	\$0.20 / \$0.40
Bandwidth Out \$0.12/Gbyte	<div style="width: 99%; background-color: red;">99%</div> 14.10 of 14.17 Gbytes	\$0.17 / \$0.40
Storage \$0.005/Gbyte	<div style="width: 25%; background-color: green;">25%</div> 25.12 of 100.50 Gbytes	\$0.12 / \$0.40
Email \$0.0001/Message	<div style="width: 20%; background-color: green;">20%</div> 500 of 2500 Messages	\$0.00 / \$0.40

↑ = Free quota

Cost for the last 14 hours: **\$0.69**

## Current Load ?

URI	Req/Sec current	Requests last 12 hrs	Avg CPU last hr	% CPU last 12 hrs
/	450.0	450	2	0%

## Errors ?

URI	Count	% Errors last 12 hrs
/	39	9%





# Thank you

