302 Google Compute Engine

Running large scale computing workloads on Google's infrastructure

Craig McLuckie
Lead Product Manager

Google I/O 12

Google Compute Engine is Infrastructure-as-a-Service

A little about our infrastructure...

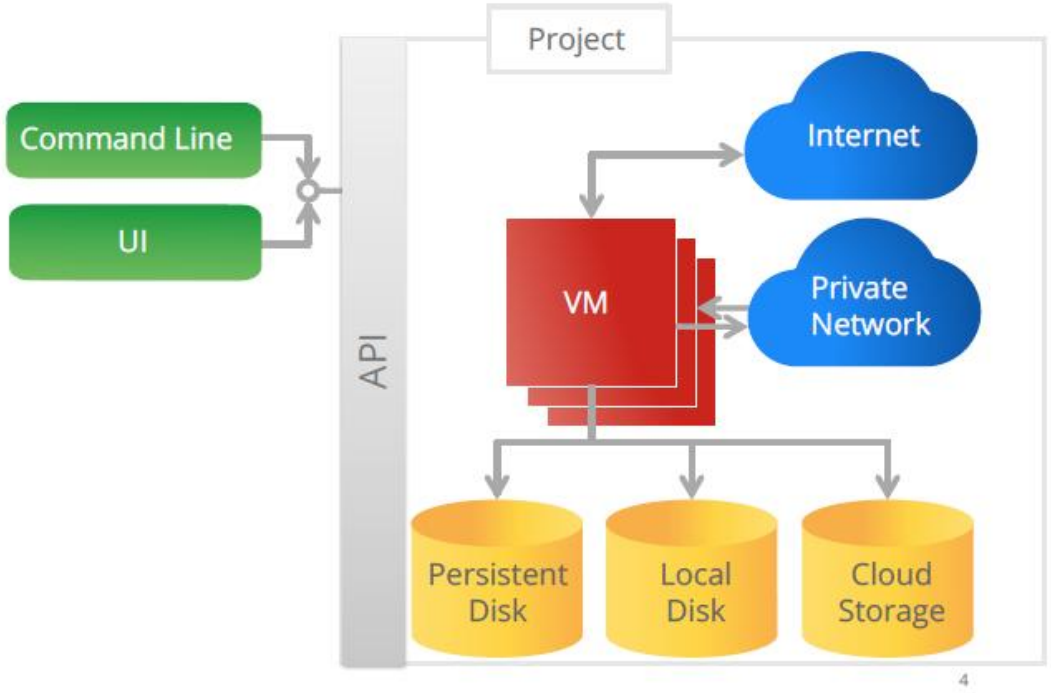
- The internet requires big infrastructure
  - Indexing billions of web pages
  - Serving 3 billion hours of video per month
  - Offering 350 million Gmail users 10GB of storage
- Our infrastructure is amongst the most scalable and efficient
- Google Compute Engine makes our infrastructure available to you



Introducing Google Compute Engine

Infrastructure-as-a-Service: Linux virtual machines

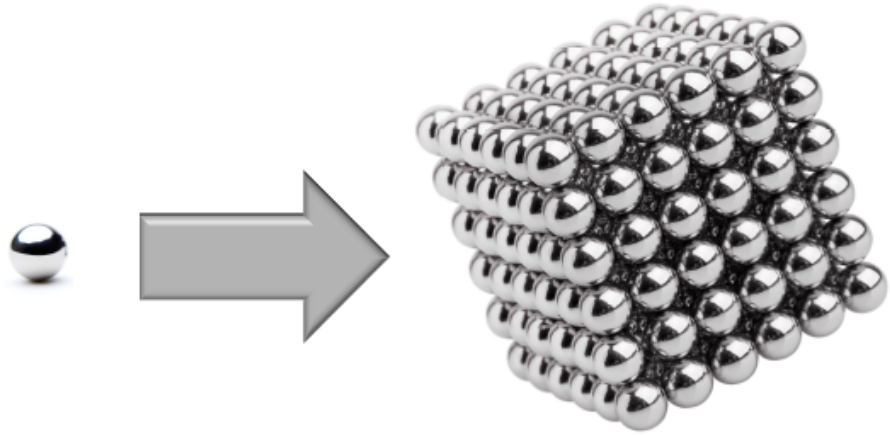
- Compute
- Storage
- Network
- Tools



Introducing Google Compute Engine

Lots of virtual machines...

- Highly scalable
- High performance
- Affordable



Guiding principles

How we thought about building this system

Strong Security

- Network encapsulation
- Disk encryption of data at rest

Consistency

- All about performance
  - Processor allocations
  - Storage latencies and throughput
  - Network architecture

Open and Flexible

- Open API
- Open tooling
- Focus on the ecosystem

Proven

- One VM infrastructure
  - For Google production workloads
  - For your workloads
- Currently running Google production services

Powered by the Ecosystem

- Focus on building a vibrant ecosystem
- Strong core infrastructure
- Our ecosystem partners
  - Support mobility to the cloud
  - Richer experiences
- Think about services not servers

RightScale.launch(Clouds)
Google Compute Engine

Michael Crandell – RightScale, CEO and Founder

RIGHTSCALE®

= Cloud Management

= 4M servers, 55k users

= Largest scaling events: 3.5k servers

= Cloud migrations: 20k+ servers

Google Compute Engine

Global Private Networks: It Matters

- Google infrastructure “secret sauce”
- Deploy globally
- Replicate for DR

Google Compute Engine

Fast Boot Times: It Matters

- Scale up quickly
- Shorten test cycles

2 minutes

Google Compute Engine

Encrypted Data at Rest: It Matters

- Top security requirement
- Simpler and faster

RIGHTSCALE®

3 key principles for the cloud:

- Usable Stuff
- Automation
- Workload Liberation

.launch

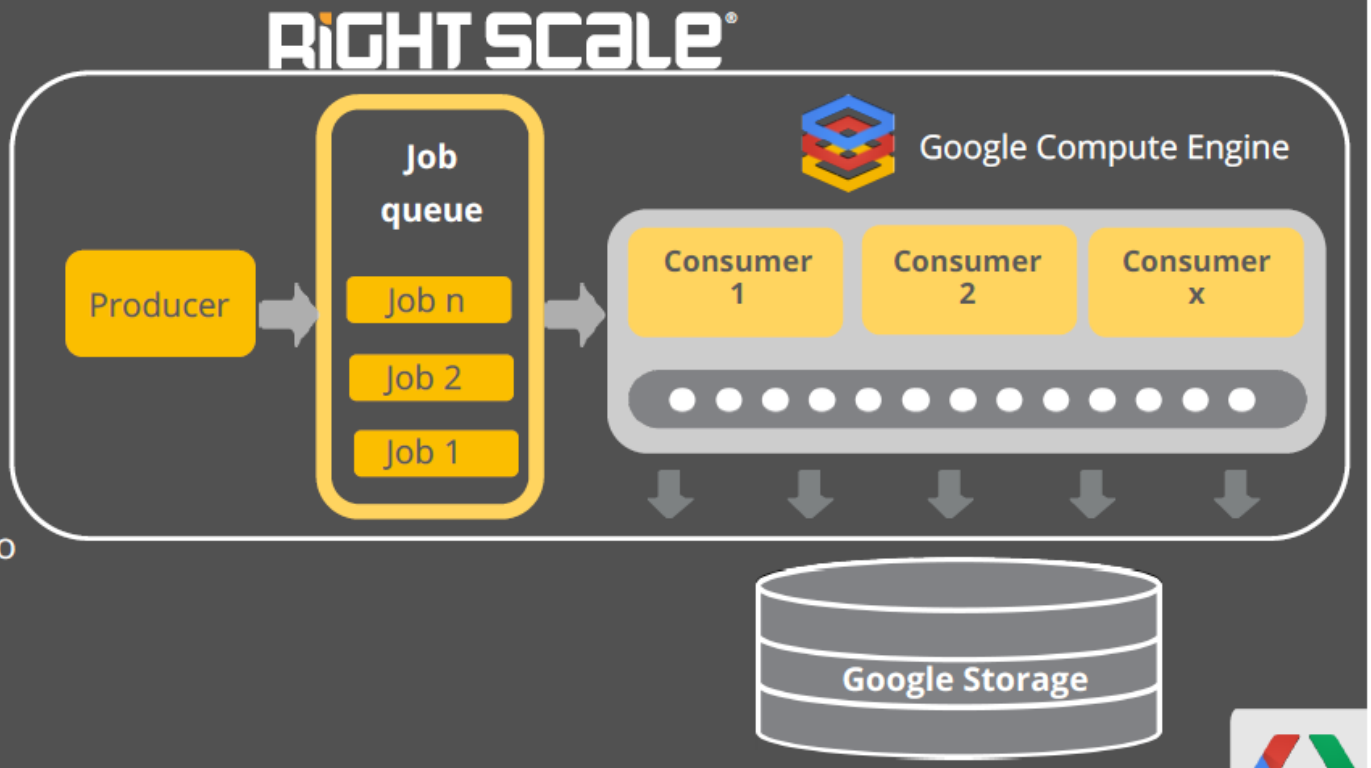
Shivan Bindal – RightScale, Product Manager

Demo: Video Transcoding

Producer pushes transcoding jobs to queue

Consumer pops instructions off the queue, downloads video and runs the rendering engine

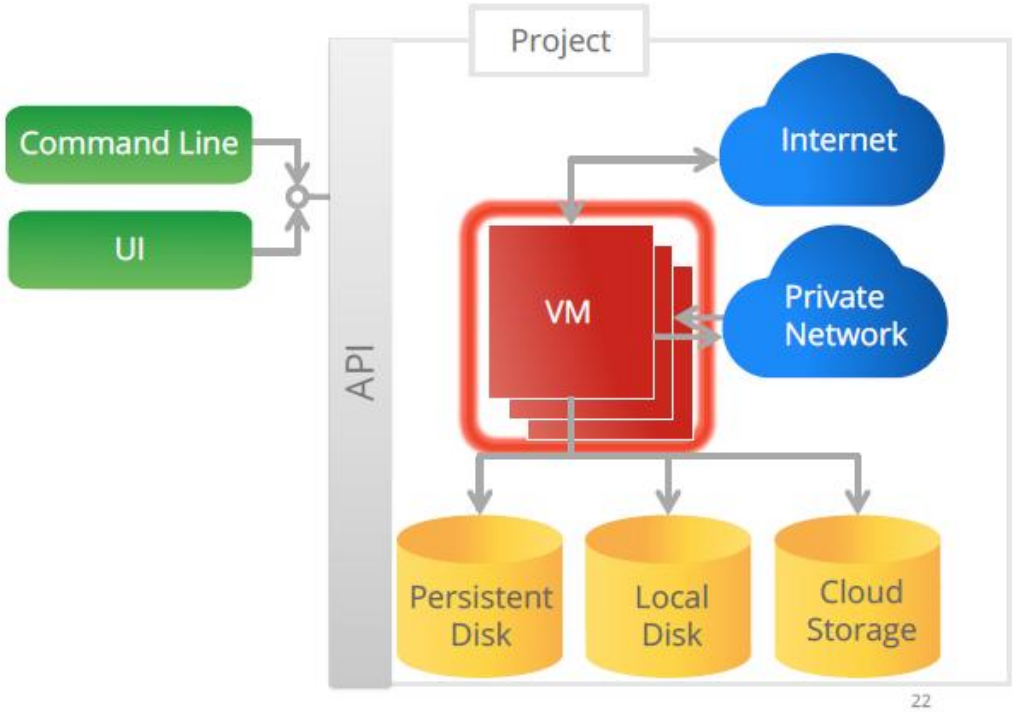
After rendering, video is sent to Google Storage



Technical Walk Through

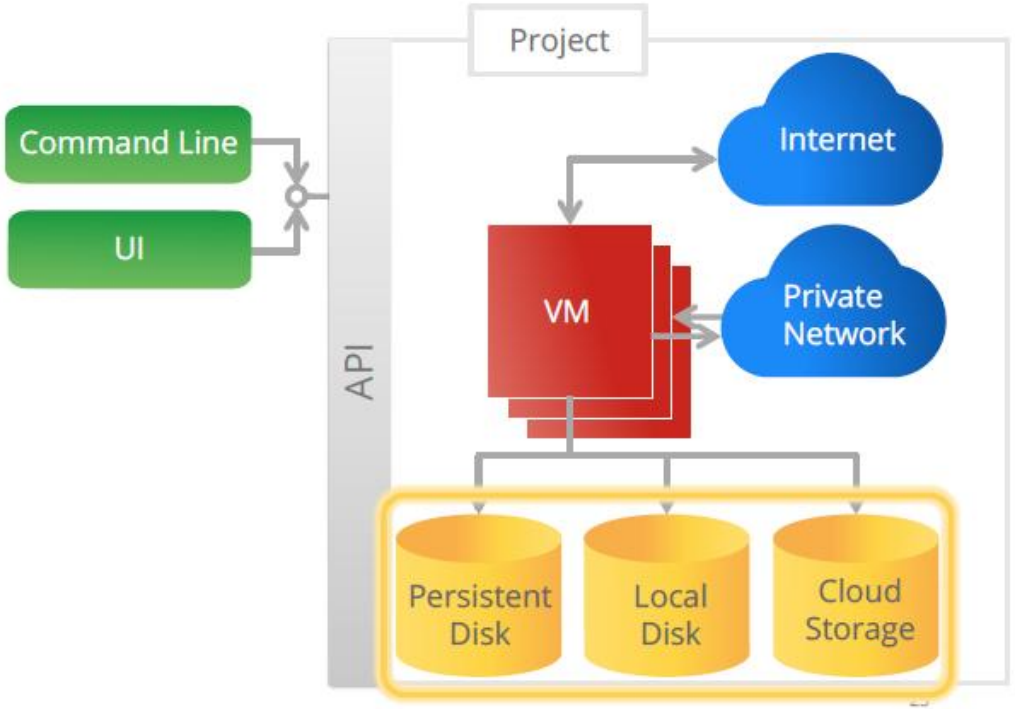
Compute

- KVM based Linux VMs
- Multiple sizes



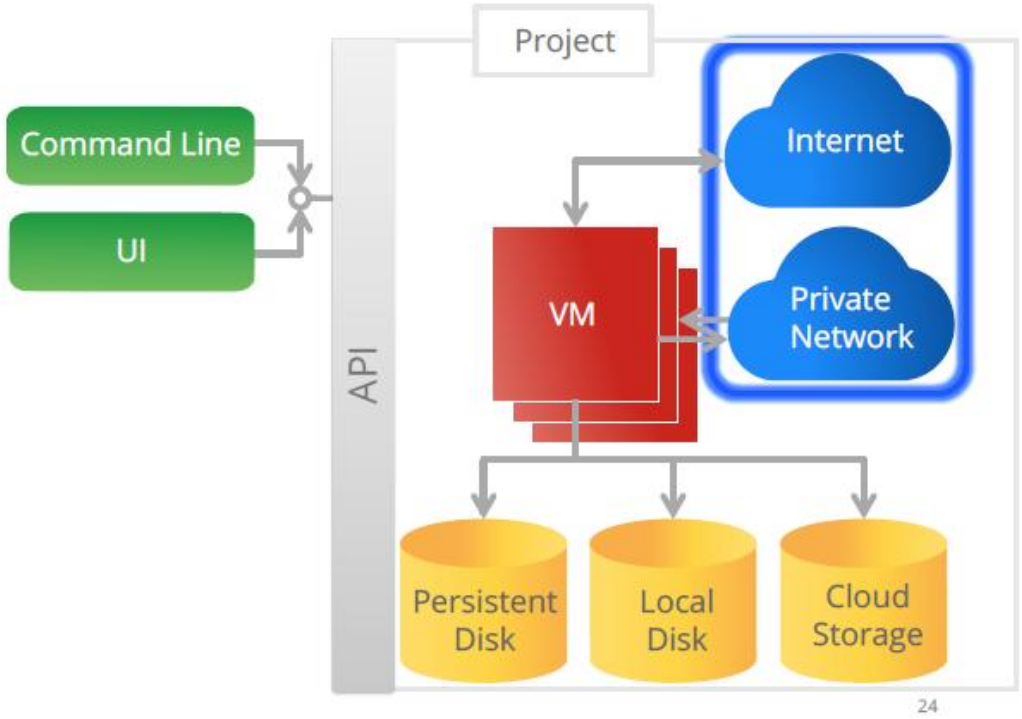
Storage

- Network block
- Local ephemeral
- Object Store



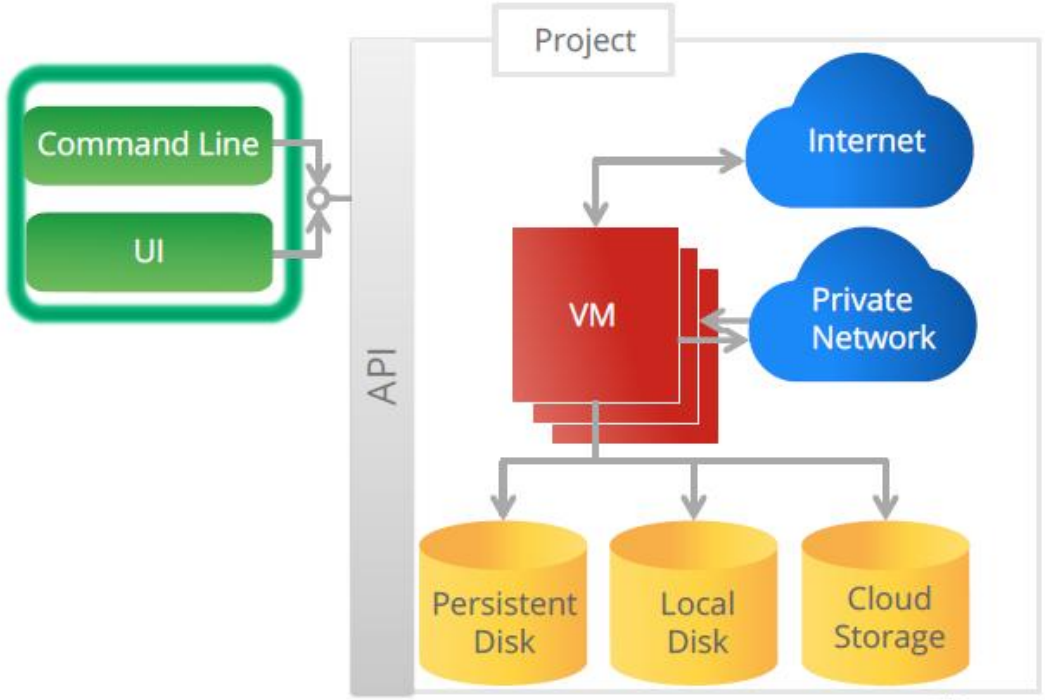
Network

- Connect to internet
- Connect internally
  - Global private network



Tools

- REST API
- Open Tools



MapR

John Schroeder – MapR CEO and Founder
MC Srivas – MapR CTO and Founder

MapR's Experience with Google Compute Engine

- **Fast**
  - Virtualized public cloud rivals on premise physical
- **Easy**
  - Provision 1,000s of servers in minutes
- **Cost effective**
  - Pay only for what you use

MapR Technologies

- **Open, enterprise-grade distribution for Hadoop**
  - Easy, dependable and fast
  - Open source with standards-based extensions
- **Hadoop**
  - Big data analytics
  - Hadoop inspired by MapReduce paper published by Google scientists Jeffrey Dean and Sanjay Ghemawat in 2004
- **MapR is deployed at 1000's of companies**
- **MapR Hadoop Cloud Service now available on Google Compute Engine**

Demo

Let's run a large sort

Run TeraSort on a **1250**-node MapR Hadoop cluster on Google Compute Engine

How does this Compare to Terasort Records?

	Record on physical hardware	MapR on Google Compute Engine
Hardware	Physical	Virtual/Cloud
Servers	1460	1256
Disks	5840	1256
Cores	11680	5024
Time	1:02 min	1:20 min

Deployment Comparison

Current Record

1460 physical servers
Prepare datacenter
Rack and stack servers
Maintain hardware

Months

Google MAPR TECHNOLOGIES

1256 instances
Invoke `gcutil` command

Minutes

Cost Comparison

Current Record

1460 1U servers x
$4K/server =

$5,840,000

Google MAPR TECHNOLOGIES

1256 *n1-standard-4-d* x
$.58/instance hour x
80 seconds =

$16

Try MapR on Google Compute Engine

www.mapr.com/google

Compute Pricing

Up to 50% more compute for your money than other leading cloud providers

	Virtual CPU/ GCEU	RAM	Disk	Price/Hour
n1-standard-1-d	1 / 2.75	3.75 GB	420 GB	$0.145
n1-standard-2-d	2 / 5.5	7.5 GB	870 GB	$0.29
n1-standard-4-d	4 / 11	15 GB	1770 GB	$0.58
n1-standard-8-d	8 / 22	30 GB	2 x 1770 GB	$1.16

Limited preview program

Apply for access today!

- Focused on large scale workloads
- Complimentary access for a limited period
- SLA and support available to commercial customers
- Apply for program at cloud.google.com

Thank You!

Please visit cloud.google.com
See other Google Compute Engine sessions: 308 and 313
Look for me (Craig McLuckie) on Google +