

Running Java and Grails applications on Amazon EC2

Chris Richardson

Head of Cloud Development

SpringSource, a division of VMware

[@crichardson](#)

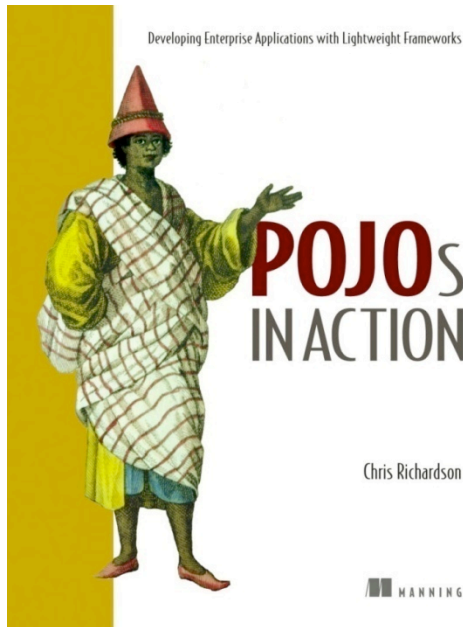
Chris.Richardson@SpringSource.com

<http://www.cloudfoundry.com>

Overall presentation goal

How to deploy multi-tier
Java and Grails
applications on clouds
such as Amazon EC2

About Chris



- Grew up in England and live in Oakland, CA
- Over 25+ years of software development experience including 14 years of Java
- Speaker at JavaOne, SpringOne, NFJS, JavaPolis, Spring Experience, etc.
- Organize the Oakland JUG and the Groovy Grails meetup



http://www.theregister.co.uk/2009/08/19/springsource_cloud_foundry/

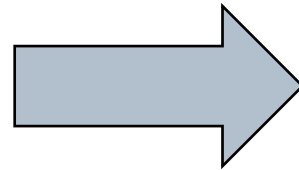
Agenda

- The future is cloudy**
- Using Amazon EC2
- Deploying on Amazon EC2
- Running the web tier
- Deploying app servers
- Deploying a database
- Handling security
- Building highly available systems
- Moving to Platform as a Service

Development is much easier and faster...

Better:

- Languages
- Frameworks
- Tools
- Processes
- ...



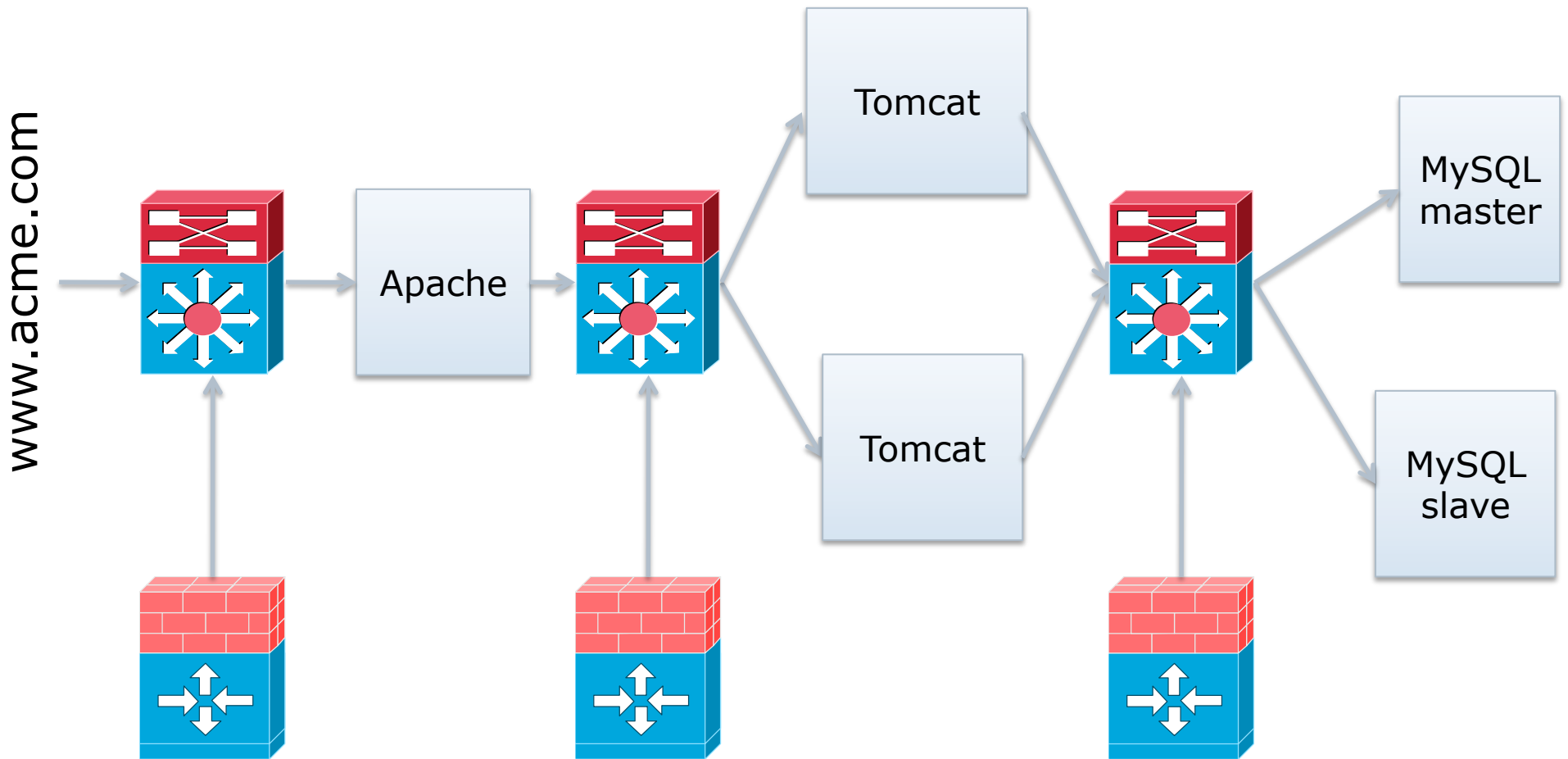
Simpler,
faster
development

Agile

POJOs



But deployment is a challenge



Deployment challenges

- ❑ **Uncertainty:** how much hardware do we need?
- ❑ **Risk: under-provisioning => success catastrophe**
- ❑ **Upfront Cost:** Can we afford it?
- ❑ **Skills deficit:** Who is going to set up and maintain it?
- ❑ **Friction:** How much time and effort does it take to approve, buy and install?

The future is cloudy

A pool of highly scalable, abstracted infrastructure that hosts your application, and is billed by consumption

By James
Staten
of Forrester
Research

**AND
is managed via a self-service API**

me

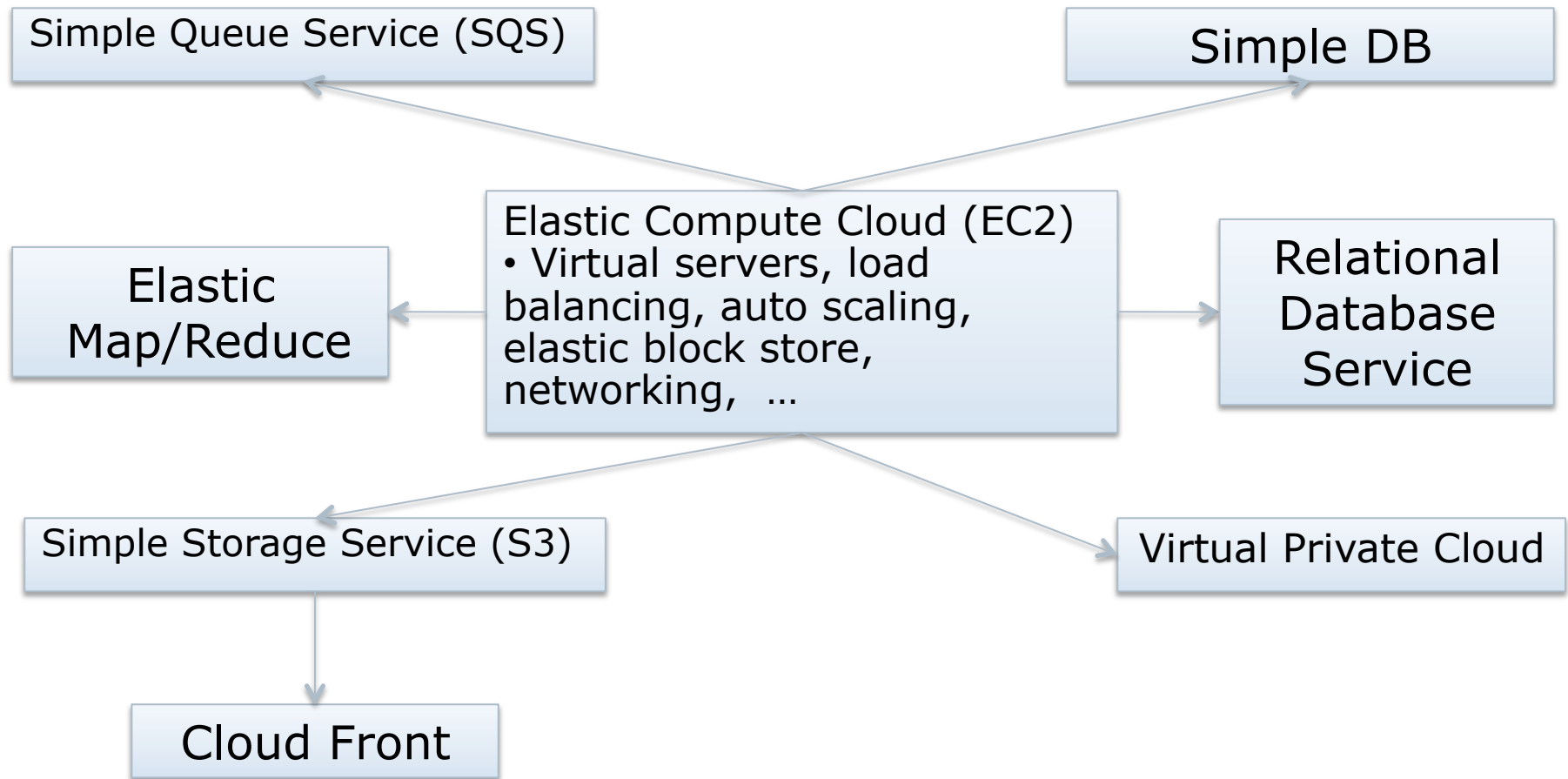
Public Clouds

- Amazon EC2
- VMware vCloud Express partners, e.g. Terremark
- Attractive to smaller businesses and enterprise IT application developers

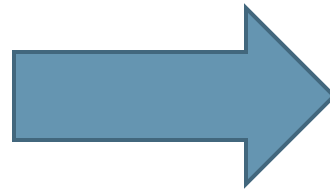
Private Clouds

- In-house data center managed as a cloud
- Vendors:
 - Eucalyptus
 - VMware vCloud
- Attractive to enterprise IT operations

Pay per use web services managed by Amazon



Sign up



- ❑ Login using your existing Amazon account
- ❑ Select the web services you want to use
- ❑ Only takes a few minutes

Make web service call to create instances...

End point

```
https://us-east-1.amazonaws.com?  
  Action=RunInstances  
&AWSAccessKeyId=...  
&Signature=...  
&ImageId=ami-3795705e  
&InstanceType=c1.medium  
&MinCount=1  
&MinCount=1  
...
```

OS image
Linux/
Windows
Preinstalled
applications

CPU/
Memory
/Storage

... Get a response...

```
<RunInstancesResponse>
...
  <item>
    <instanceId>i-4ef21327</instanceId>
    <instanceState>...
      <name>pending</name>
    </instanceState>
    <placement>
      <availabilityZone>us-east-1b
      </availabilityZone>
    </placement>
    <dnsName/>
  </item>
</RunInstancesResponse>
```

The diagram illustrates an XML response from the AWS RunInstances API. A red box highlights the XML content. Three callout boxes on the right point to specific parts of the XML: 'Your instance' points to the `<instanceId>` element, 'It's state' points to the `<name>` element within `<instanceState>`, and 'It's DNS names' points to the `<dnsName/>` element. The XML content is as follows:

... a few minutes later

```
cer@arrakis ~
$ ssh ... root@ec2-67-202-41-150.compute-1.amazonaws.com
Last login: Sun Dec 30 18:54:43 2007 from 71.131.29.181
[root@domU-12-31-36-00-38-23:~] yum install mysql-server
[root@domU-12-31-36-00-38-23:~] yum install tomcat5
[root@domU-12-31-36-00-38-23:~] yum install httpd
[root@domU-12-31-36-00-38-23:~]
```

WS call to terminate instances

```
https://us-east-1.amazonaws.com?  
  Action=TerminateInstances  
&InstanceId.1=i-4ef21327
```

...

Pay monthly bill

This Month's Activity as of February 10, 2010

The billing cycle for this report is February 1 - February 28, 2010. The AWS service usage charges on this page currently show activity through approximately 02/10/2010 17:59 GMT.

Expand All Services Collapse All Services		Printer Friendly Version
		Totals
+ Amazon Elastic Compute Cloud	View/Edit Service	
Download Usage Report >>		55.80
+ Amazon SimpleDB	View/Edit Service	
Download Usage Report >>		0.00
+ Amazon Simple Storage Service	View/Edit Service	
Download Usage Report >>		0.16
Amazon CloudFront	View/Edit Service	
Download Usage Report >>		0.00
Amazon Simple Queue Service	View/Edit Service	
Download Usage Report >>		0.00
Amazon Virtual Private Cloud	View/Edit Service	
Download Usage Report >>		0.00
Taxes		
Estimated Taxes		0.00
<small>(Due March 1, 2010)</small>		
Total Charges due on March 1, 2010†		\$55.96

Instance types

Type	Virtual Cores	Compute Units /core*	32/64 Bit	Memory	Storage
Small	1	1	32 bit	1.7G	160G
High-CPU Medium	2	2.5	32 bit	1.7G	350G
Large	2	2	64 bit	7.5G	850G
Extra Large	4	2	64 bit	15G	1690G
High-CPU XL	8	2.5	64 bit	7G	1690G
HiMem/XL	2	3.25	64 bit	17.1	420G
HiMem/Double XL	4	3.25	64 bit	34.2G	850G
HiMem/QuadXL	8	3.25	64 bit	68.4G	1690G

* EC2 Compute Unit = 1.0-1.2 GHz 2007 Opteron/
Xeon processor

Pricing models

- On-demand instances
 - Pay by the hour
 - \$0.085/hour (small instances) – \$2.40/hour (Hi Mem. Quad XL)
- Reserved instances
 - Upfront payment
 - The right to run instances at reduced rate for 1-3 year term
 - Small instances: \$227.50/\$350 + \$0.03/hr
 - Hi Mem. Quad XL: \$6370/\$9800 + \$0.84/hr
 - Up to 30%-50% cheaper
- Spot instances
 - Each instance type has a spot price – reflects unused capacity
 - Launch request: N instances, max price, valid time period
 - Spot price \leq max price \rightarrow instances launched
 - Spot price $>$ max price \rightarrow instances terminated

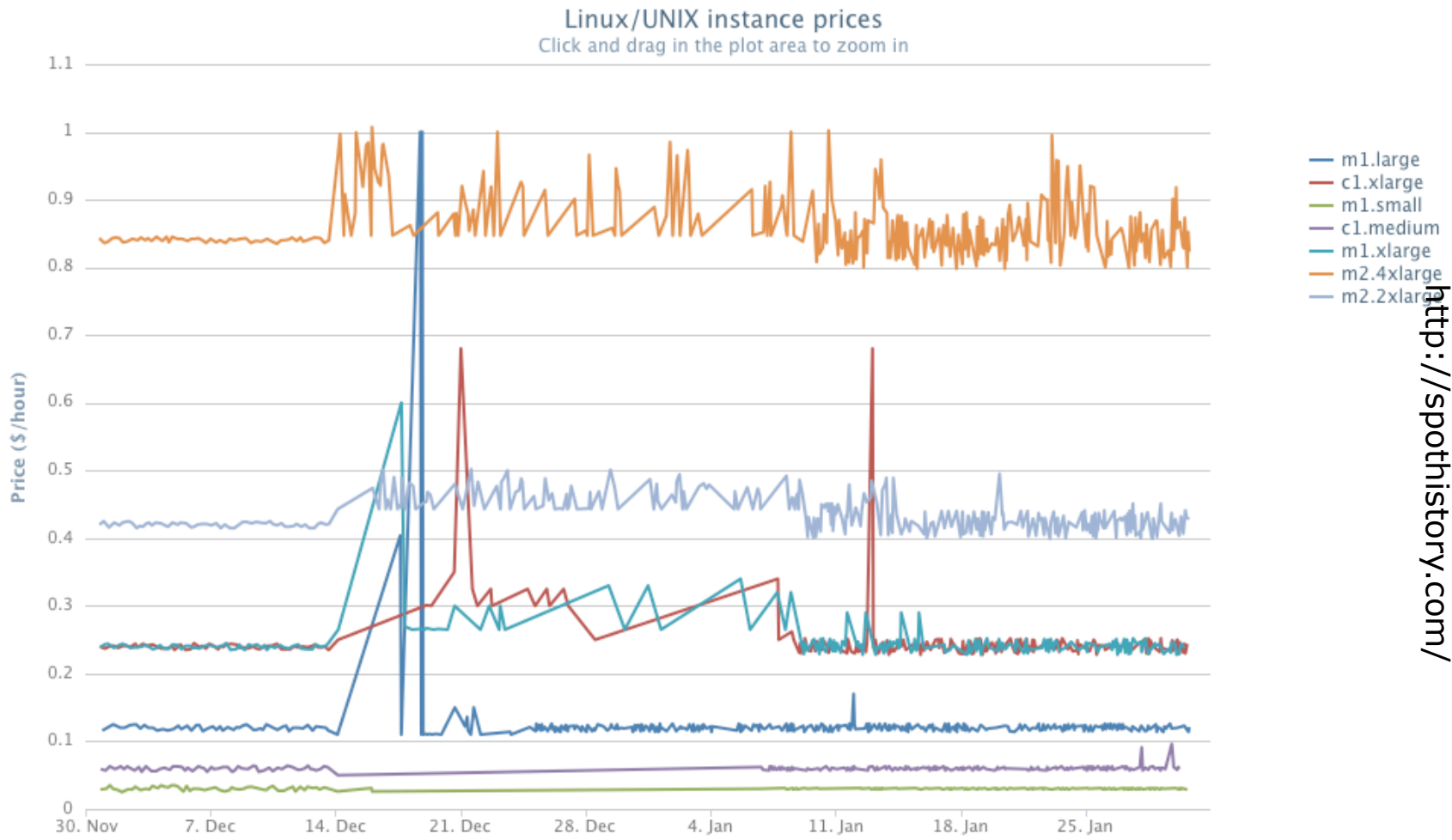
Windows instances are more expensive. Regional price variations

On-demand and reserved instance pricing

Type	Size	On demand \$/hr (/ month)	Reserved \$/hr (/w 3 year)
Small	1 / 1 / 32 bit / 1.7G 160G	0.085 (61)	0.04
High-CPU Medium	2 / 2.5 / 32 bit / 1.7G/ 350G	0.17 (122)	0.09
Large	2 / 2 / 64 bit / 7.5G / 850G	0.34 (245)	0.17
Extra Large	4 / 2 / 64 bit / 15G / 1690G	0.68 (490)	0.35
High-CPU XL	8 / 2.5 / 64 bit / 7G / 1690G	0.68 (490)	0.35
HiMem/ML	2 / 3.25 / 64bit / 17.1G / 420G	0.50 (360)	0.34
HiMem/ Double XL	4 / 3.25 / 64 bit / 34.2G / 850G	1.20 (864)	0.61
HiMem/ QuadXL	8 / 3.25 / 64 bit / 68.4G / 1690G	2.40 (1728)	1.21

** Windows and Europe are more expensive. bandwidth: \$0.08-0.15/ Gbyte

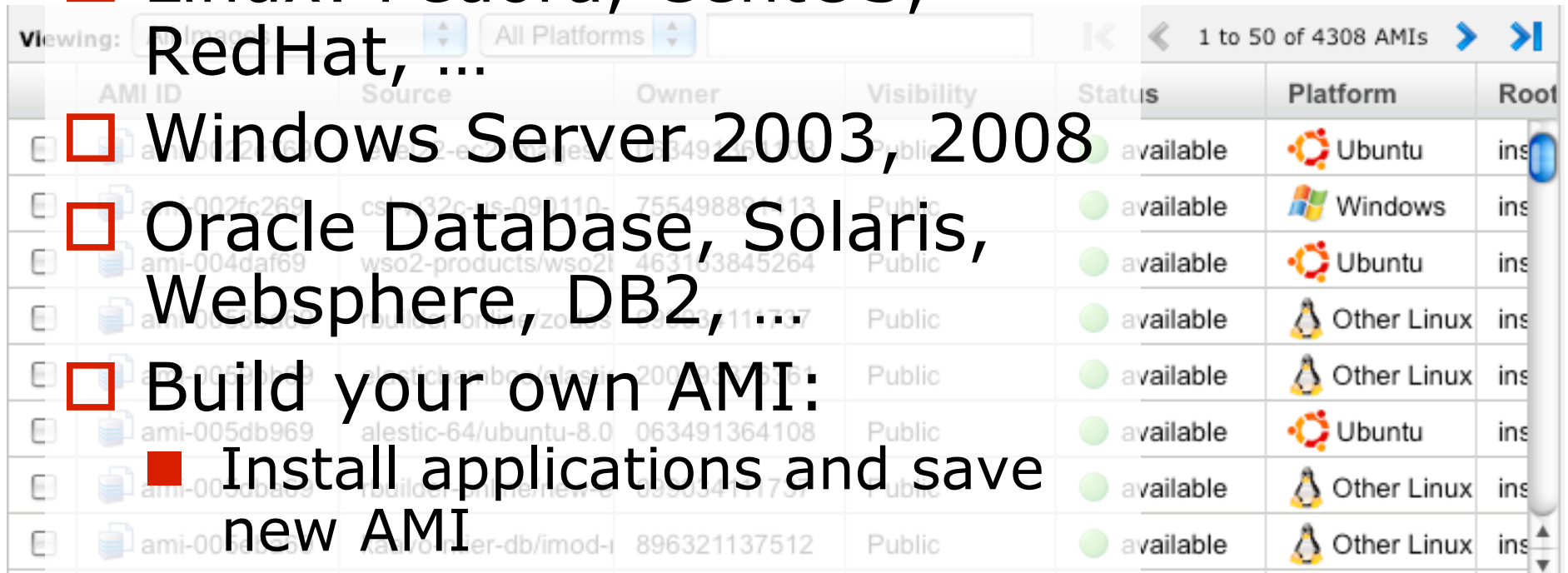
Spot instance pricing



<http://spothistory.com/>

Amazon Machine Image (AMI)

- ❑ Contains OS and applications
- ❑ Linux: Fedora, CentOS, RedHat, ...



The screenshot shows the Amazon Machine Image (AMI) console interface. At the top, it says "Viewing: All Platforms" and "1 to 50 of 4308 AMIs". Below this is a table with columns: AMI ID, Source, Owner, Visibility, Status, Platform, and Root. The table lists several AMIs, including Windows Server 2003, 2008, Oracle Database, Solaris, Websphere, DB2, and various Linux distributions like Ubuntu and Other Linux.

AMI ID	Source	Owner	Visibility	Status	Platform	Root
ami-002fc269	cs-32c-us-090110-	75549889113	Public	available	Ubuntu	ins
ami-004daf69	wso2-products/wso2i	463103845264	Public	available	Ubuntu	ins
ami-005db969	alestic-64/ubuntu-8.0	063491364108	Public	available	Ubuntu	ins
ami-005db969	builder-im/new-e	05503411757	Public	available	Other Linux	ins
ami-005db969	builder-im/new-e	05503411757	Public	available	Other Linux	ins
ami-005db969	builder-im/new-e	05503411757	Public	available	Other Linux	ins

- ❑ Windows Server 2003, 2008
- ❑ Oracle Database, Solaris, Websphere, DB2, ...
- ❑ Build your own AMI:
 - Install applications and save new AMI
 - Create an AMI from scratch

Benefits of cloud deployment

- ❑ Frictionless, agile deployment
- ❑ No upfront cost
- ❑ Leverage the expertise of the cloud provider
- ❑ Easily scale up/down based on load
- ❑ Reduces risk of a success catastrophe
- ❑ No long-term commitment
- ❑ Minimal downtime from hardware failure

Issues with public clouds

- Security:
 - AWS is SAS70 Type II certified
 - Runs HIPAA compliant apps

BUT

 - Lack of PCI compliance
 - Discomfort with sending customer data to a 3rd party
- Instance types:
 - Lack of small machines
 - Lack of very large machines, e.g. 128G memory
- Sophisticated networking

Cloud Computing Survey: IT Leaders See Big Promise, Have Big Security Questions

Greatest Concerns Surrounding Cloud Adoption at Your Company

Security	45%
Integration with existing systems	26%
Loss of control over data	26%
Availability concerns	25%
Performance issues	24%
IT governance issues	19%
Regulatory/compliance concerns	19%
Dissatisfaction with vendor offerings/pricing	12%
Ability to bring systems back in-house	11%
Lack of customization opportunities	11%
Measuring ROI	11%
Not sure	7%
Other	6%

*Respondents selected up to three criteria.

SOURCE: CIO Research

www.cio.com/article/455832/Cloud_Computing_Survey_IT_Leaders_See_Big_Promise_Have_Big_Security_Questions

Using a public cloud seems expensive

- ❑ Running larger servers 24 x 7 (e.g. \$490/month)
 - ❑ Storing data (\$150/TB/month)
 - ❑ Bandwidth (\$0.08-0.15/GB)
- BUT** using your own hardware
- ❑ Is often just as expensive
 - ❑ Lacks elasticity/agility

Example – beer on the cloud

- ❑ Grails application
- ❑ Short-term marketing campaign site
- ❑ Fluctuating load
 - Sat/Sun 4 servers
 - Mon-Fri 1 server



iTelliSeek.com – wine on the cloud

The screenshot shows the iTelliSeek.com website with a green background. At the top is a navigation bar with buttons for HOME, BE A TELLER, PARTNERS, ABOUT US, CONTACT, LOGIN, MY ACCOUNT, and MY REVIEWS, along with an FAQ icon. The main content area features a central panel with 'iTELL' and 'iSEEK' buttons, a 'REVIEW-O-METER' gauge, and a 'NUMBER OF REVIEWS' display showing '000000000070'. Text on the left describes the site as a place for wine connoisseurs and amateurs to submit and search reviews, and mentions 'Challenging the Status Quo'. Text on the right explains how to use the site: 'To review a wine, click on iTELL. To search for a wine review, click on iSEEK.' A vertical 'iTelliSeek!' logo is on the right side. At the bottom, there are social media links for Facebook, Twitter, Ning, and Feed, and a footer with copyright information and links to Terms of Use and Privacy Policy.

HOME BE A TELLER PARTNERS ABOUT US CONTACT LOGIN MY ACCOUNT MY REVIEWS FAQ ?

A place for wine connoisseurs and amateurs to submit and search unique and consistent wine reviews.

*Wine opinion has been democratized: We aim to open wine criticism to all people who have taste buds and an opinion. **Challenging the Status Quo***

To review a wine, click on iTELL.
To search for a wine review, click on iSEEK.

iTELL **iSEEK**

REVIEW-O-METER

000000000070
NUMBER OF REVIEWS

Our goal is to reach critical mass before we open up the search to our seekers. We award our telling community with ownership in our company. (see [TELLER's reward program](#))

Facebook Twitter Ning Feed

© 2008-2010 iTELL iSEEK V0.70 [Terms of Use](#) [Privacy Policy](#)

iTELL
iSEEK!

Agenda

- The future is cloudy
- Using Amazon EC2**
- Deploying on Amazon EC2
- Running the web tier
- Deploying app servers
- Deploying a database
- Handling security
- Building highly available systems
- Moving to Platform as a Service

AWS Tools

- Amazon provided command line tools
 - CLI equivalents of APIs
 - AMI creation tools
- AWS CLI tools from Tim Kay
 - CLI for S3 and EC2
 - Alternatives to Amazon CLI tools
- AWS Console - very slick
- ElasticFox
 - Awesome Firefox plugin
 - Launch and manage instances
- S3 Organizer
 - Firefox plugin
 - Manipulate S3 buckets and objects
- AWS Eclipse plugin
- ...

Some AWS features are unavailable in the GUI tools – must use CLI

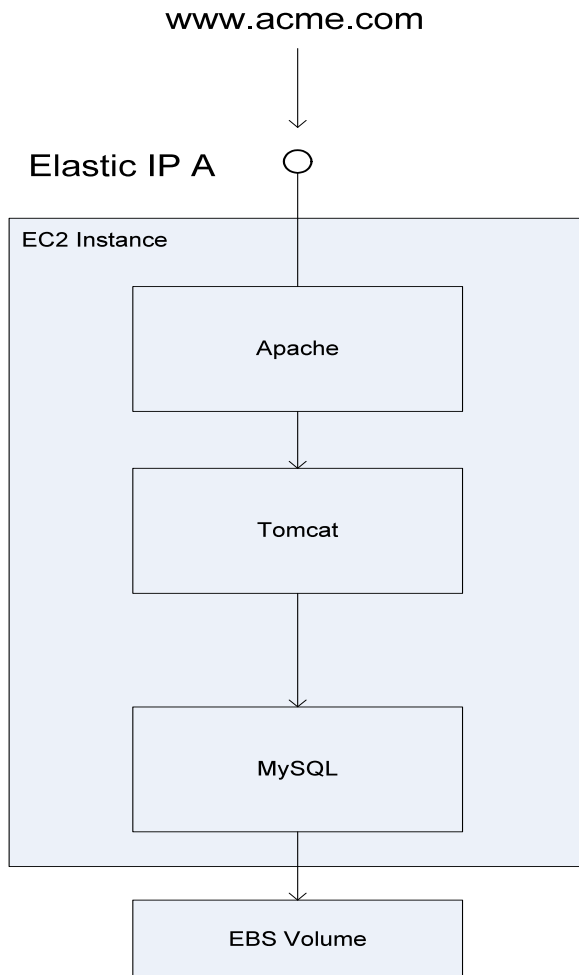
Launch an instance

DEMO

Agenda

- The future is cloudy
- Using Amazon EC2
- Deploying on Amazon EC2**
- Running the web tier
- Deploying app servers
- Deploying a database
- Handling security
- Building highly available systems
- Moving to Platform as a Service

Starter website - \$

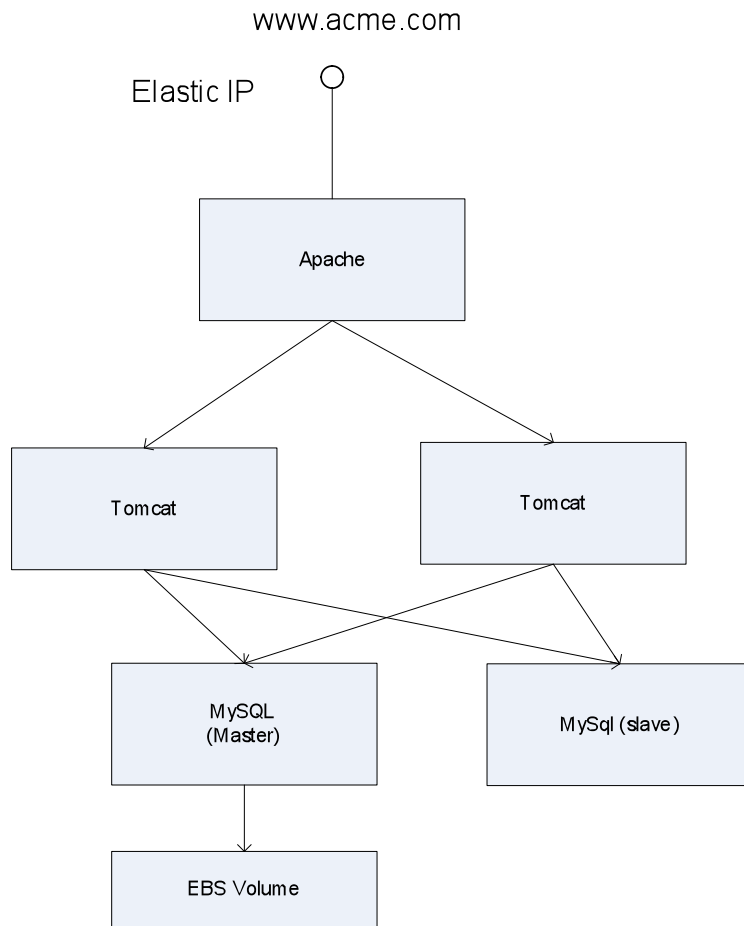


Low cost - \$61/month

Elastic - load changes \Rightarrow change instance type in a few minutes

Available - instance crashes \Rightarrow replace in a few minutes

Higher capacity website - \$\$

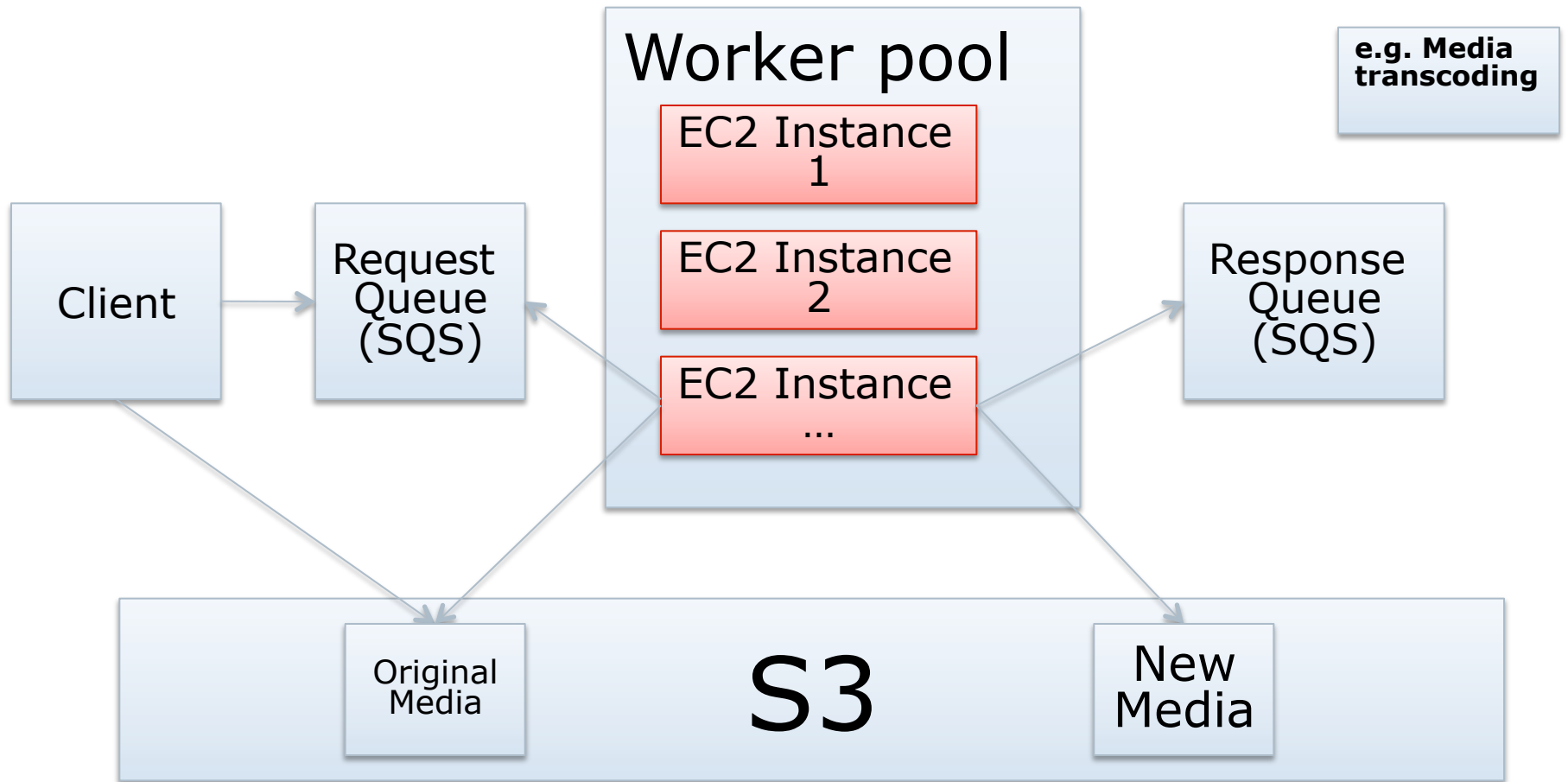


Low cost - > ~\$180/month
(1 or more Tomcats, 0 or more Slaves)

Elastic - load changes ⇒ quickly expand/subtract Tomcats with no downtime

Available - instance crashes ⇒ replace in a few minutes

Batch processing architecture



Easy upgrades

- ❑ Clone production environment
 - Make read-only or turn off
 - Snapshot EBS volumes and create new volumes
- ❑ Apply upgrades to clone
- ❑ Test clone
- ❑ Move elastic IP addresses to clone
- ❑ Terminate old instances once you are sure that everything works

Agenda

- The future is cloudy
- Using Amazon EC2
- Deploying on Amazon EC2
- Running the web tier**
- Deploying app servers
- Deploying a database
- Handling security
- Building highly available systems
- Moving to Platform as a Service

Overview of the web tier

- ❑ Load balancing and request routing to app servers
- ❑ Serving static content
- ❑ Content caching
- ❑ SSL termination
- ❑ Needs static IP address
- ❑ Options:
 - Single (Apache) Web Server
 - Elastic Load Balancer
 - Elastic Load Balancer + one or more web servers

Giving Apache a static IP address

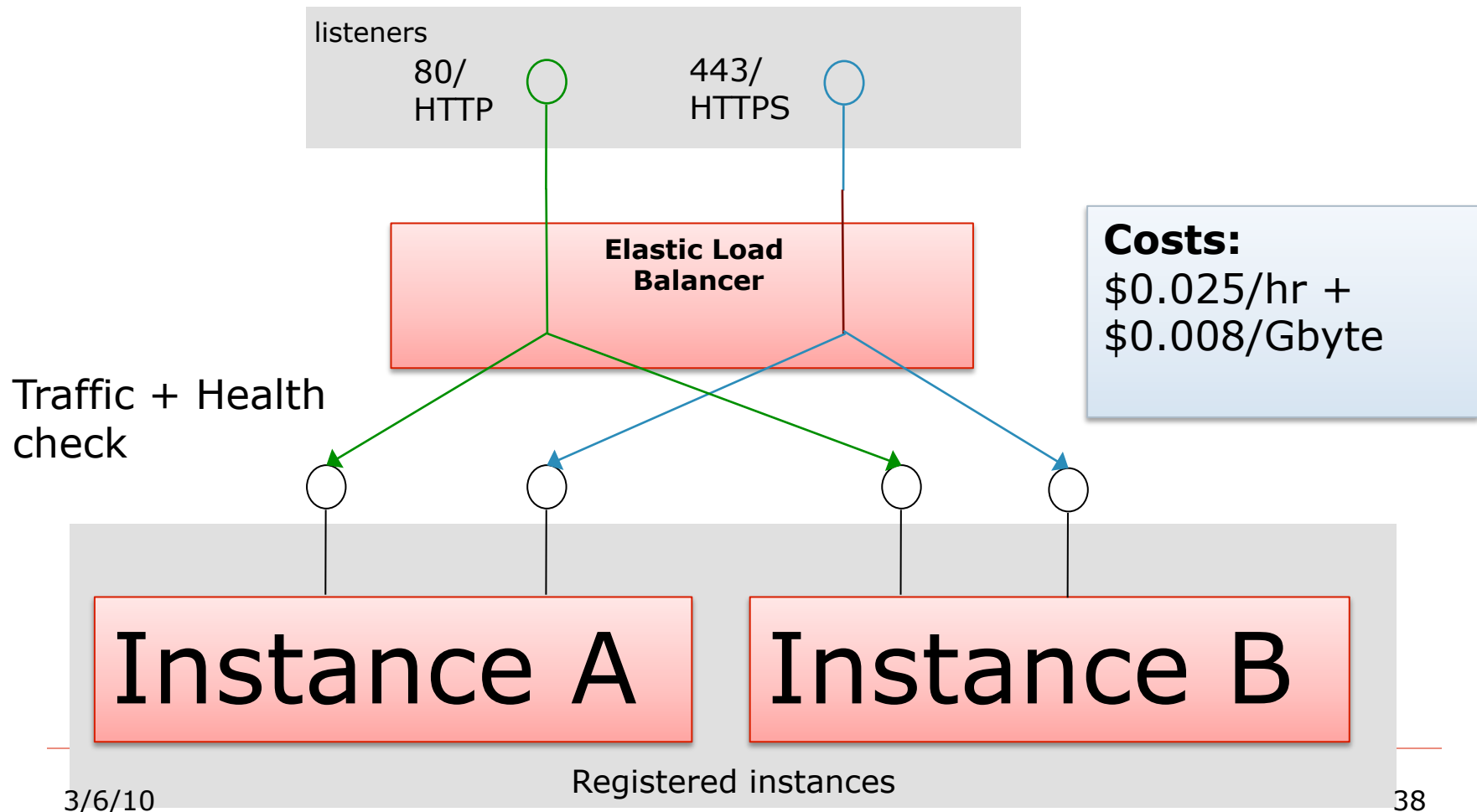
- Instance IP addresses are dynamically allocated
- Elastic IP addresses
 - Static public IP addresses that belong to your account
 - Make WS request to allocate
 - Associate with instance (e.g. web server) = it's public IP address
 - You configure DNS to resolve to the elastic IP address
- You pay for unused EIPs

Beware of lag with : EIP association, DNS, Connectivity

Elastic load balancer

www.acme.com CNAME

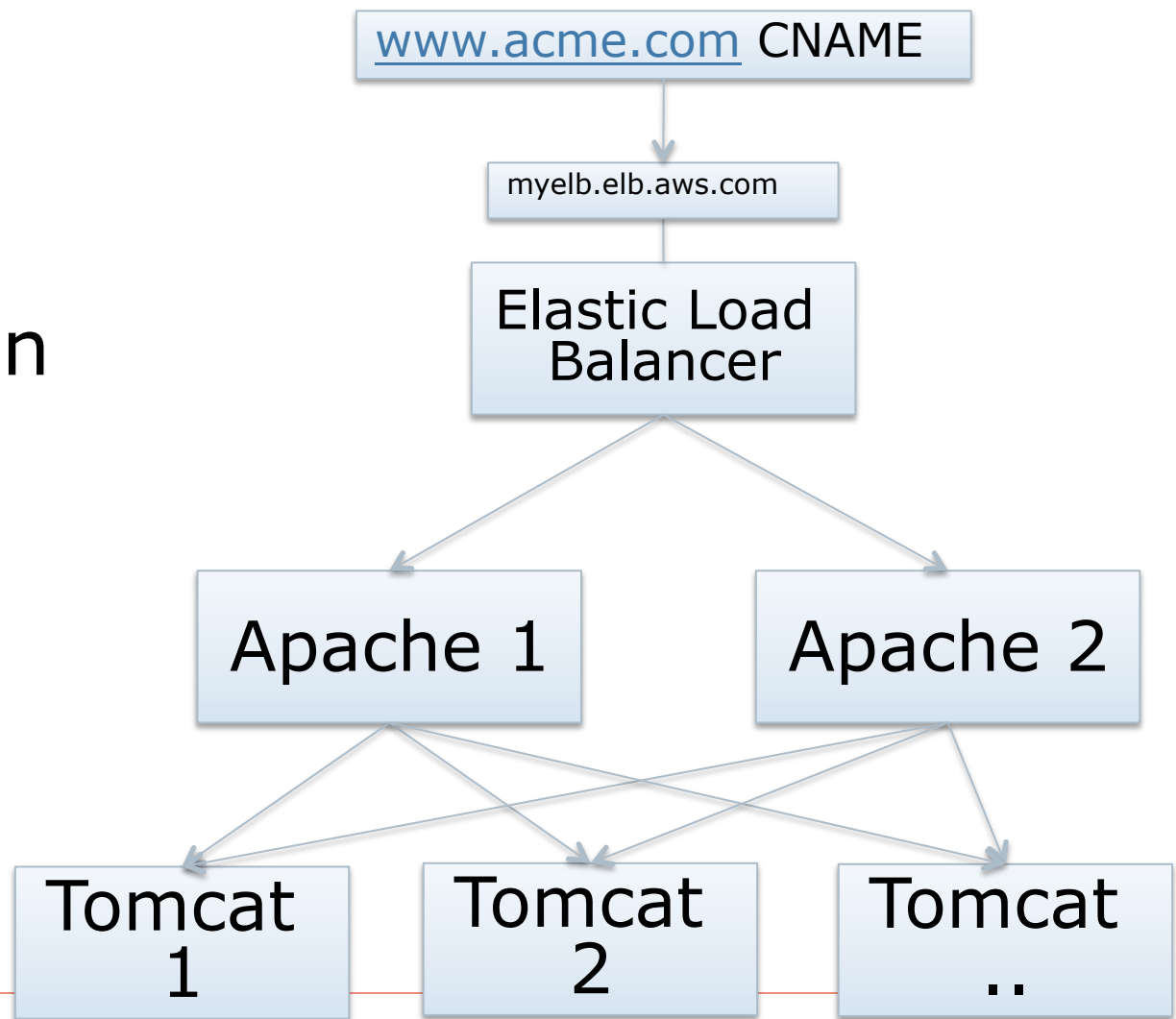
MyLB-148691016.us-east-1.elb.amazonaws.com



Elastic Load Balancers are insufficient

- ❑ No sticky sessions
- ❑ No SSL termination

➔ Use with Apache



Agenda

- The future is cloudy
- Using Amazon EC2
- Deploying on Amazon EC2
- Running the web tier
- Deploying app servers**
- Deploying a database
- Handling security
- Building highly available systems
- Moving to Platform as a Service

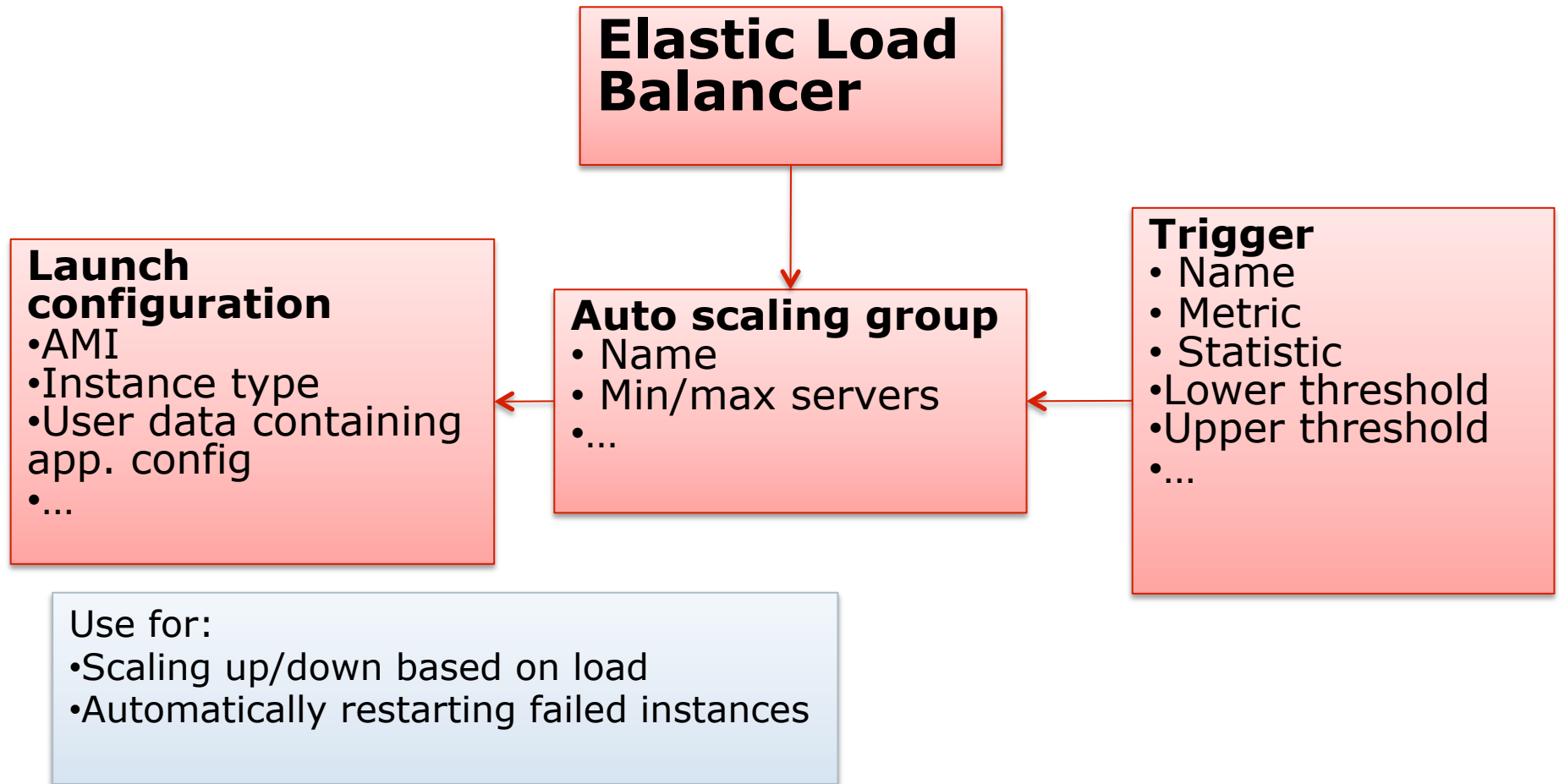
The app server tier

- Multiple load balanced application servers:
 - e.g. Tomcat or SpringSource tc Server
- Sometimes clustered:
 - Session-state replication
 - Distributed/replicated caches
 - ...
- Ideally, auto-scaled

No multicast for resource discovery

- ❑ Prevents the use of standard clustered resource discovery: e.g. JGroups multicast etc
- ❑ Use a registry, e.g.:
 - The database, S3
 - security groups, auto scaling group, ...
 - Extend JGroups to read registry
- ❑ JGroups with TCP
- ❑ Use Terracotta to cluster Tomcat

Amazon Auto Scaling



Issues with Amazon Auto Scaling

- Instances must be self-configuring via user data
 - App server - wars to deploy, database connection information, ...
 - Apache – static content, SSL certs, ...
- Decisions driven by only what the hypervisor can see:
 - CPU, I/O, Response time
 - Not from application-level metrics, e.g. JMX
- Need app server registration/discovery mechanism
 - ➔ Less useful for Java PaaS

Agenda

- The future is cloudy
- Using Amazon EC2
- Deploying on Amazon EC2
- Running the web tier
- Deploying app servers
- Deploying a database**
- Handling security
- High availability
- Moving to Platform as a Service

The database tier

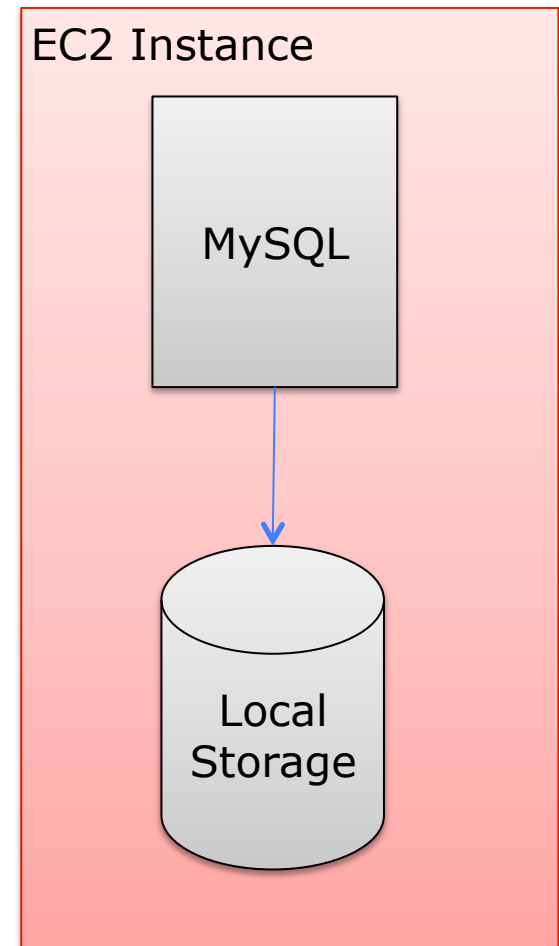
- ❑ Run database server such as MySQL or Oracle
- ❑ Need reliable storage
- ❑ Need a reliable backup mechanism
- ❑ Two choices – convenience vs. flexibility:
 - Run your own database
 - Using Amazon's relational database service

MySql + Local storage

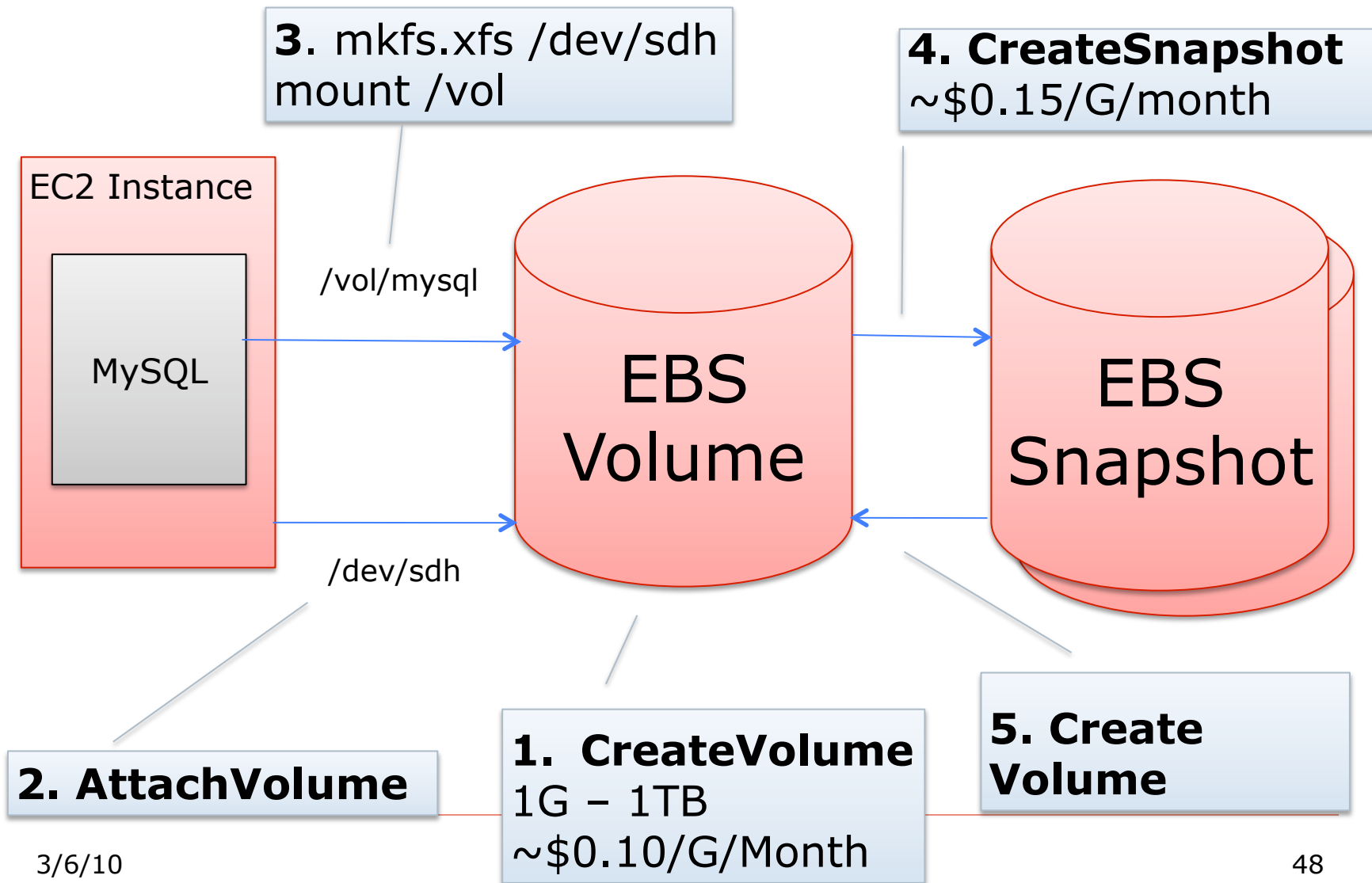
- ❑ It's plentiful (160G to 1690G)

BUT

- ❑ Local storage is ephemeral
- ❑ First write performance penalty
- ❑ Need to backup (to S3)



MySql + Elastic Block Store Volume



Issues with EBS

- ❑ Attachment lag
- ❑ Accessed by a single instance
- ❑ Performance
 - EBS volume = single disk drive
 - Strip for performance
 - Over Gigabit Ethernet -> potentially I/O bound
- ❑ Number of snapshots
 - Limited to 1000

What you can buy:

12x140G drives @
15KRPM drives, RAID 10,
Battery backed 2G cache

Terremark vCloud
Express: fiber attached
storage

Amazon Relational Database Service

- ❑ MySQL 5.1 as a web service
- ❑ Database Instance = EC2 instance + EBS volume
- ❑ Preferred maintenance window:
 - 4 hour weekly window
 - For patches etc
- ❑ Backups
 - Preferred backup window (2 hour daily window)
 - Continuous log file backups -> point in time recovery
- ❑ Pricing:
 - Compute = \$0.11 - \$3.10/hour
 - Storage = EBS pricing

Amazon RDS is very convenient

BUT

- ❑ Only MySQL
- ❑ No replication
- ❑ Four hour weekly maintenance window:
 - Amazon claim the outage will be brief
 - But you can't control if and when it happens

Agenda

- The future is cloudy
- Using Amazon EC2
- Deploying on Amazon EC2
- Running the web tier
- Deploying app servers
- Deploying a database
- Handling security**
- Building highly available systems
- Moving to Platform as a Service

Use the usual security best practices

- Turn off unused services
- File ownership and permissions
- Disabling password based ssh login
- Standard Linux, Apache, Tomcat and MySQL best practices

EC2 Network security

Amazon:

- Monitor and manage the network
- Prevent An instance cannot sniff traffic for other instances

You:

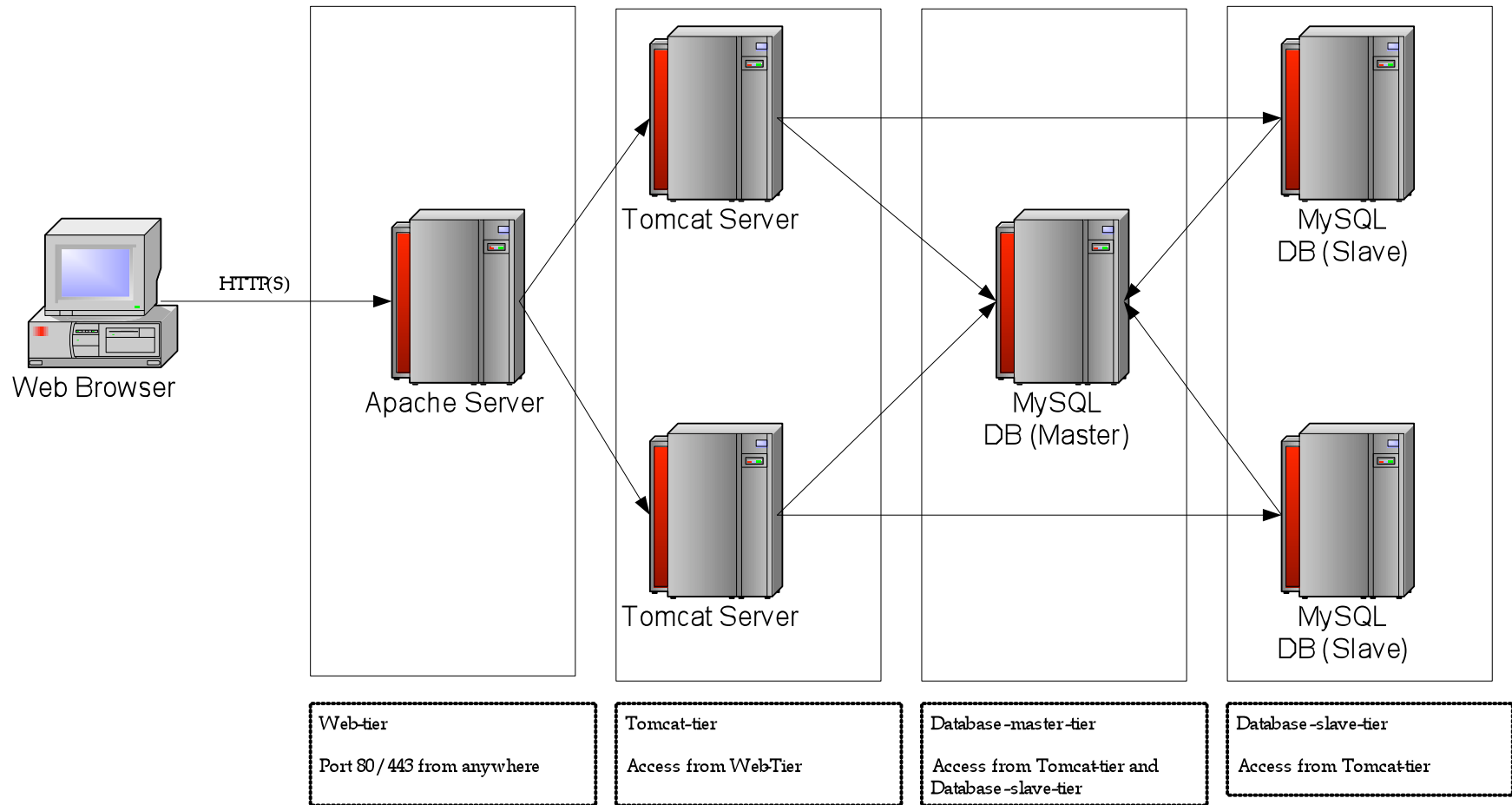
- Consider encrypting network traffic
- Use EC2 firewall – aka. security groups
- Limit SSH access to only your location

Security Groups

- ❑ Named set of firewall rules associated with your account
- ❑ An instance
 - Belongs to one or more security groups
 - Defaults to “default” security group
- ❑ Permits inbound traffic
 - Protocol: tcp, udp
 - Range of ports
- ❑ From:
 - Anywhere – specific port range
 - An IP address (range) – specific port range
 - Another group - **all ports**
- ❑ Common usage
 - Port 80 (http)/443 (https) – anywhere
 - Port 22 (ssh) – just from your location

```
?Action=RunInstances  
&SecurityGroup.1=g1  
&SecurityGroup.2=g2
```

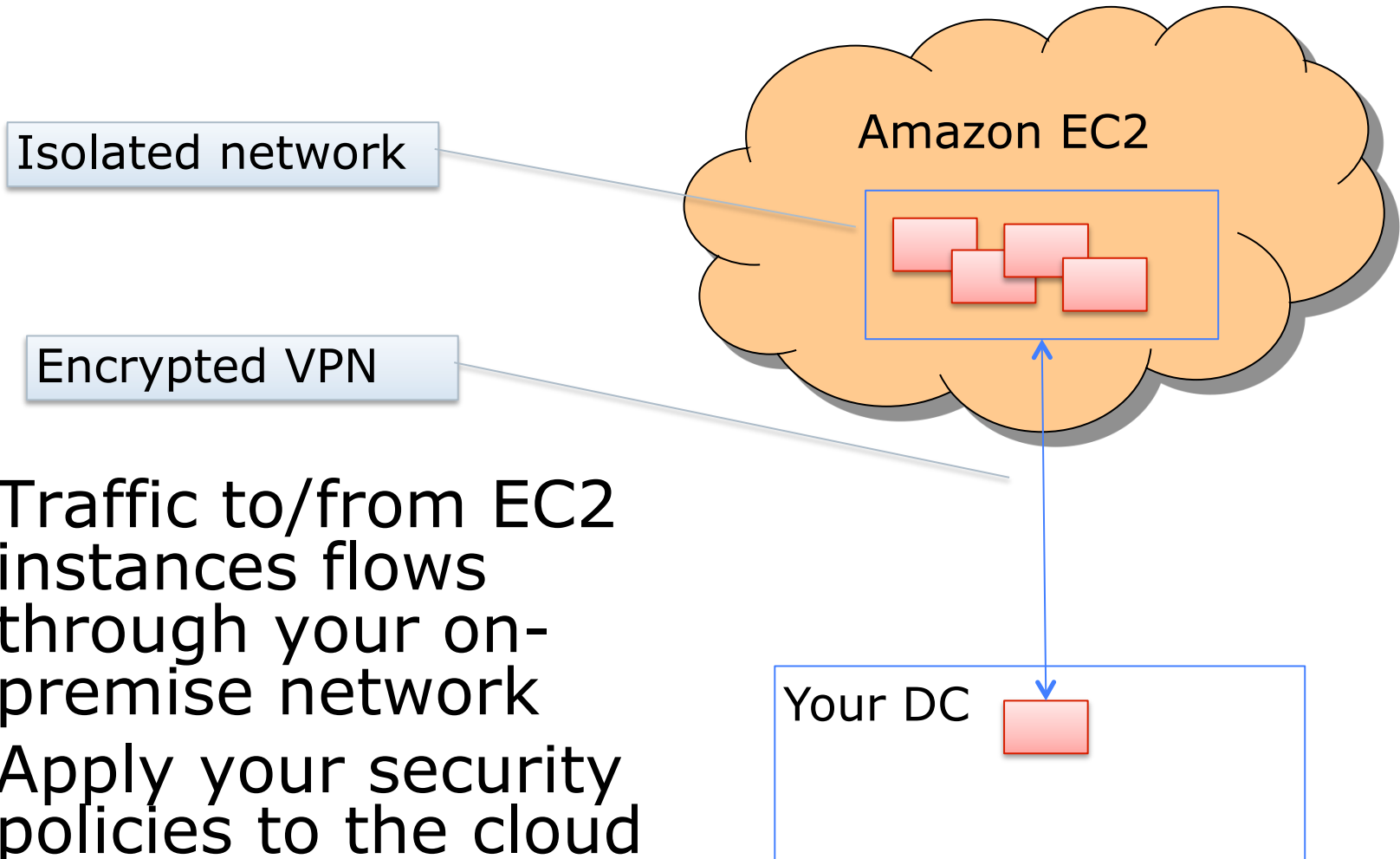
Using security groups



Use a software firewall

- With a security group:
 - Tomcat Servers are only accessible from Apache Server
 - But this means all ports!
- Use iptables:
 - Tomcat servers only allow port 22 and port 8009 (AJP)

Use Virtual Private Cloud



- ❑ Traffic to/from EC2 instances flows through your on-premise network
- ❑ Apply your security policies to the cloud

EC2 Storage security

- Amazon wipes virtual disks so one customer cannot see another's data
- **But**
 - You don't know where your data is
 - Amazon could be subpoena'd
- Consider encrypting data
 - Encrypted file systems
 - Encrypting sensitive data in DB
 - Encrypting backups in S3

Agenda

- The future is cloudy
- Using Amazon EC2
- Deploying on Amazon EC2
- Running the web tier
- Deploying app servers
- Deploying a database
- Handling security
- Building highly available systems**
- Moving to Platform as a Service

Deploying highly available applications

- ❑ AWS has had very well publicized outages

BUT...

- ❑ Is internal IT really any better?
- ❑ In reality: AWS is (more) reliable
- ❑ Don't forget:
 - You are not responsible for the hardware
 - Instance fails ⇒ Launch a new one in a few minutes

But once in a blue moon 1

Hello,

We have noticed that one or more of your instances are running on a **host degraded due to hardware failure.**

i-14d00b7d

The host needs to undergo maintenance and will be taken down at 12:00 GMT on 2009-04-03. Your instances will be terminated at this point.

The risk of your instances failing is increased at this point. We cannot determine the health of any applications running on the instances. We recommend that you launch replacement instances and start migrating to them.

Feel free to terminate the instances with the ec2-terminate-instance API when you are done with them.

Let us know if you have any questions.

Sincerely,

The Amazon EC2 Team

But once in a blue moon 2

Hello,

One of your instances in the us-east-1 region is on hardware that requires network related maintenance. Your other instances that are not listed here will not be affected.

i-83d31feb

For the above instance, we recommend migrating to a replacement instance to avoid any downtime. Your replacement instance would not be subject to this maintenance.

If you leave your instance running, you will lose network connectivity for up to two hours. The maintenance will occur during a 12-hour window starting at 12:00am PST on Monday, February 15, 2010. After the maintenance is complete, network connectivity will be restored to your instance.

As always, we recommend keeping current backups of data stored on your instance.

Sincerely,

The Amazon EC2 Team

Using virtual IP addresses for failover

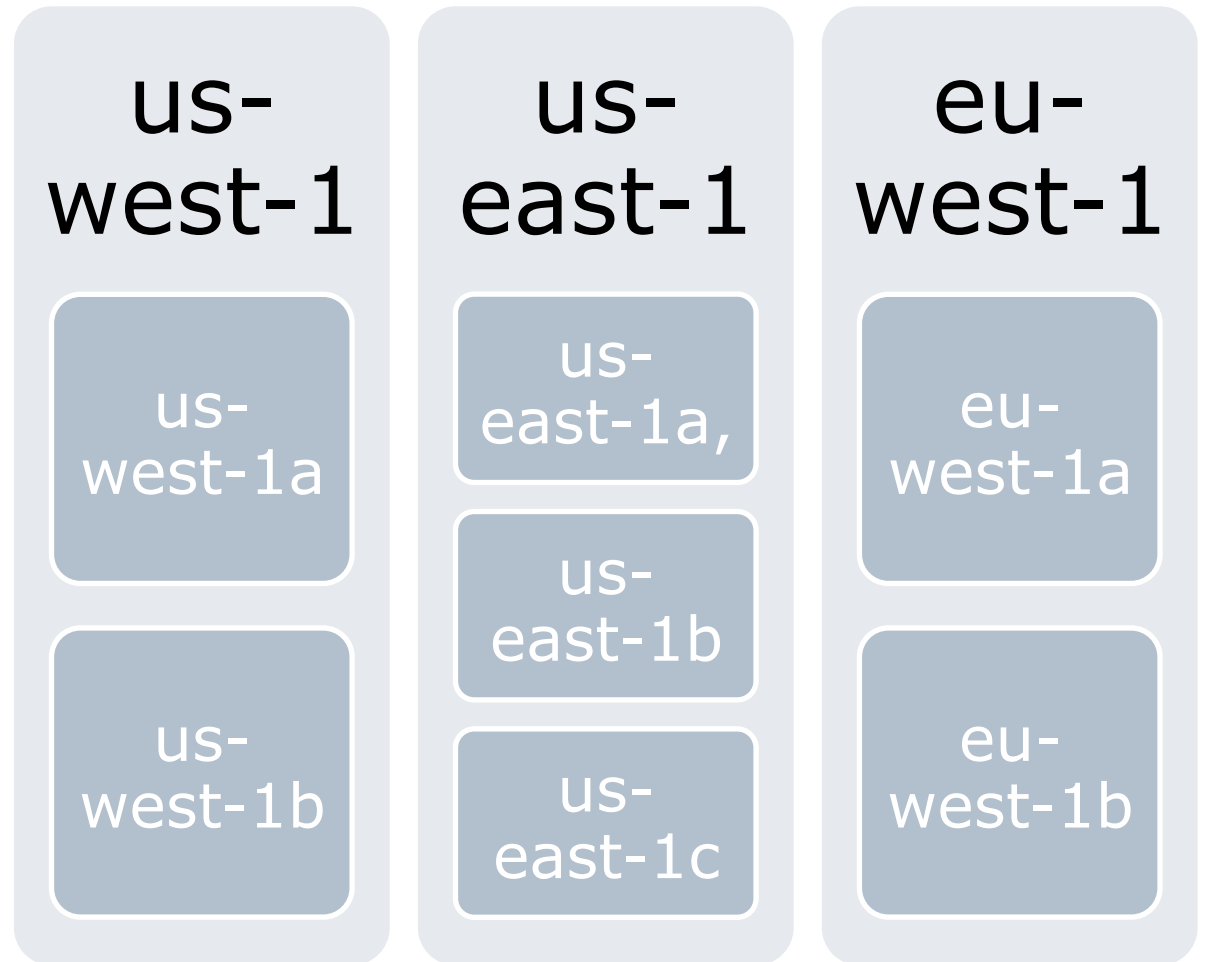
- ❑ EC2 does not have private VIP addresses
- ❑ Elastic IP addresses behave like VIP addresses
 - Assign EIP to your active server
 - Reference active server via public DNS name (not EIP) to avoid charges
 - Fail-over by moving EIP to standby server

BUT

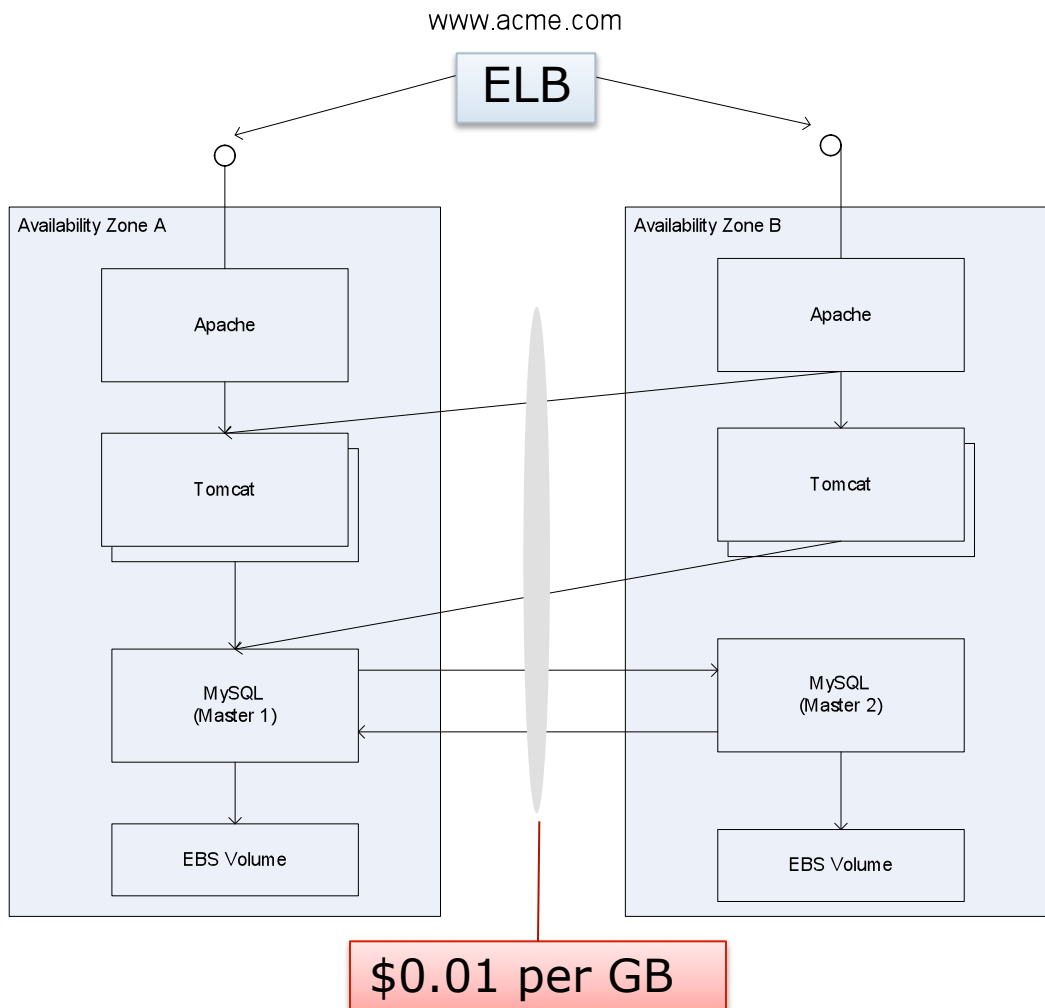
- Amazon might not give you more than 5 EIPs
- The EIP is the server's sole public IP and there is often a significant lag when assigning a new dynamic IP
- ❑ Use a DNS based approach instead, e.g.
 - Update /etc/hosts
 - Run DNS server

Regions and availability zones

- ❑ Regions - geographically dispersed locations
- ❑ Availability zone - engineered to be insulated from failure in other zones
- ❑ Specify availability zone when launching instances



Highly available - \$\$\$



EC2 SLA with
99.95% availability

Higher cost - > ~\$ 300/
month (2 Apaches, 2 MySQLs,
1 or more Tomcats, 0 or more
Slaves)

Elastic - load changes =>
quickly expand/subtract
Tomcats with no downtime

Available - No SPOF, instance
crashes => replace in a few
minutes

Agenda

- The future is cloudy
- Using Amazon EC2
- Deploying on Amazon EC2
- Running the web tier
- Deploying app servers
- Deploying a database
- Handling security
- Building highly available systems
- Moving to Platform as a Service**

Beyond infrastructure as a service

- ❑ Virtual servers and other IaaS resources are great building blocks

BUT

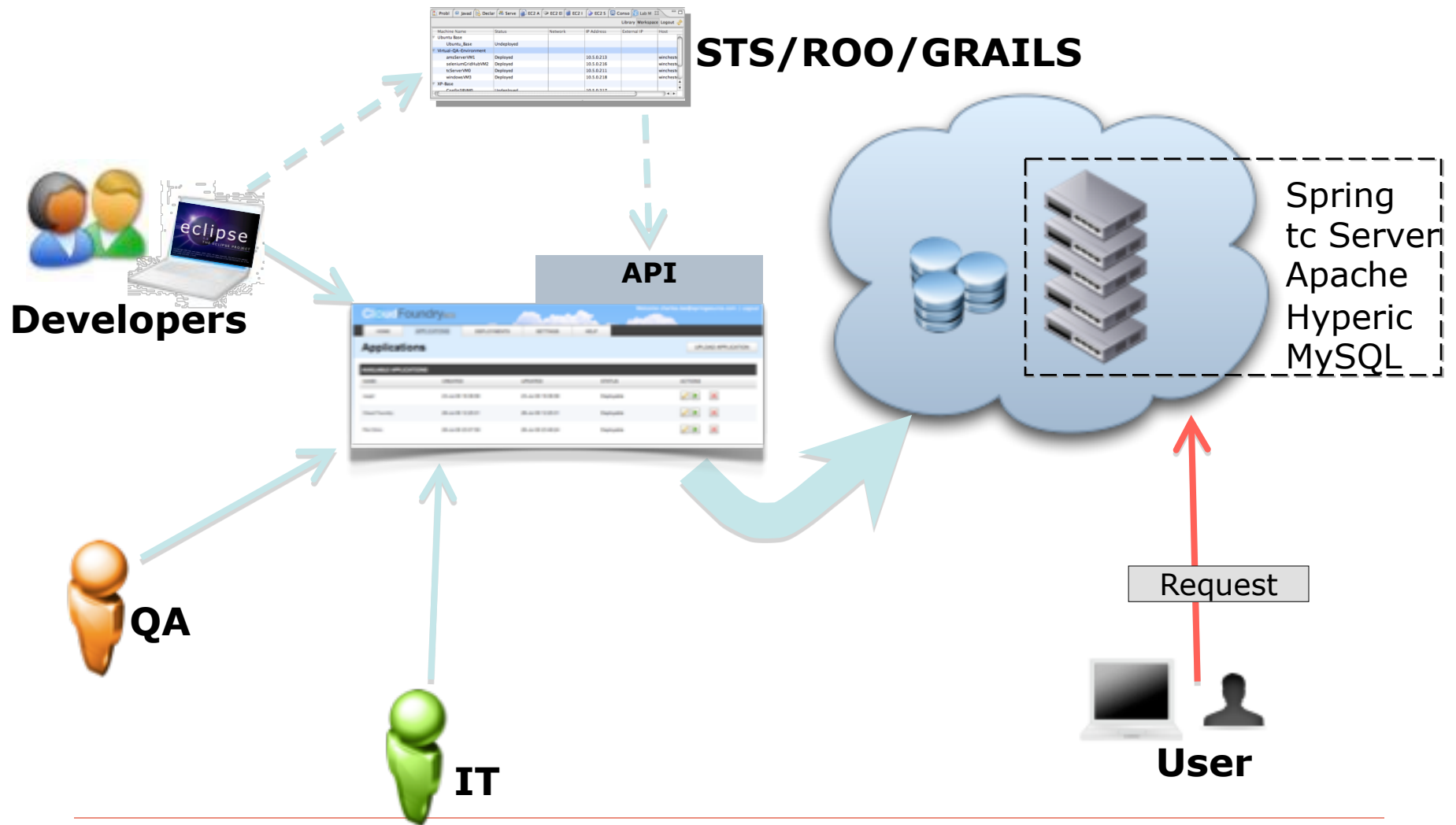
- ❑ Who wants to mess around at that level?
- ❑ Who has the skills, the time, etc.?

➔ Platform-as-a-Service (PaaS)

Platform as a Service offerings

- Microsoft Windows Azure
 - Microsoft hosted datacenters
 - .NET services
 - SQL Azure
 - ...
- Google App Engine
 - Java subset
 - Non-relational database
 - Restrictive transactions

SpringSource Cloud Foundry





Cloud Foundry Demo

DEMO

Summary

- Infrastructure as a service:
 - Enables frictionless, agile deployment
 - Pay as you go – no upfront investment/commitment required
 - Enables scale up/down
 - Hardware is someone else's problem
- Platform as a Service
 - Builds on infrastructure as a service
 - Provides a developer-focused experience
 - Available in public clouds today

Final thoughts



Download or contribute to Cloud Tools today :

www.cloudtools.org

Request a free trial of Cloud Foundry:

<http://bit.ly/cftrial0310>

Buy my book ☺

Send email:

chris.richardson@springsource.com

[@crichardson](https://twitter.com/crichardson)

