Eucalyptus

Eucalyptus: An Open-source Infrastructure for Cloud Computing

Rich Wolski
Eucalyptus Systems Inc.
www.eucalyptus.com

Exciting Weather Forecasts

Commercial Cloud Formation

What is a cloud?

SLAs

Web Services

FOR RENT

FOR RENT
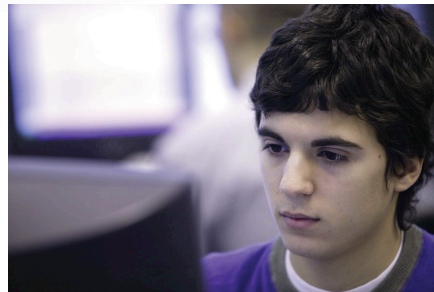
Virtualization

Cloudy issues

- **Public clouds are opaque**
  - What applications will work well in a cloud?
- **Many of the advantages offered by Public Clouds appear useful for "on premise" IT**
  - Self-service provisioning
  - Legacy support
  - Flexible resource allocation
- **What extensions or modifications are required to support a wider variety of services and applications?**
  - Data assimilation
  - Multiplayer gaming
  - Mobile devices

Open-source Cloud Infrastructure

Eucalyptus

- *Idea:* **Develop an open-source, freely available cloud platform for commodity hardware and software environments**
  - Stimulate interest and build community knowledge
  - Quickly identify useful innovations
  - Act to dampen the "hype"
- **First-principles cloud implementation**
  - Not a refactorization of previously developed technology
- **Linux or Anti-Linux?**
  - Linux: open-source platform supporting all cloud applications changes the software stack in the data center
  - Anti-Linux: transparency of the platform makes it clear that clouds do not belong in the data center

On a clear day…
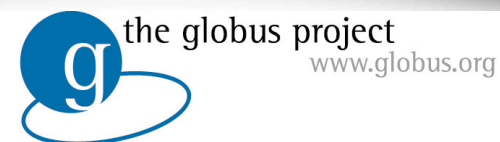
- **Globus/Nimbus**
  - Client-side cloud-computing interface to Globus-enabled TeraPort cluster at U of C
  - Based on GT4 and the Globus Virtual Workspace Service
  - Shares upsides and downsides of Globus-based grid technologies
- **Enomalism (now called ECP)**
  - Start-up company distributing open source
  - REST APIs
- **Reservoir**
  - European open cloud project
  - Many layers of cloud services and tools
  - Ambitious and wide-reaching but not yet accessible as an implementation

What's in a name?

Eucalyptus

- **Elastic Utility Computing Architecture Linking Your Programs To Useful Systems**
- **Web services based implementation of elastic/utility/cloud computing infrastructure**
  - Linux image hosting ala Amazon
- ***How do we know if it is a cloud?***
  - Try and emulate an existing cloud: Amazon AWS
- **Functions as a software overlay**
  - Existing installation should not be violated (too much)
- **Focus on installation and maintenance**
  - *"System Administrators are people too."*

Goals for Eucalyptus

- **Foster greater understanding and uptake of cloud computing**
  - Provide a vehicle for extending what is known about the utility model of computing
- **Experimentation vehicle prior to buying commercial services**
  - Provide development, debugging, and "tech preview" platform for Public Clouds
- **Homogenize local IT environment with Public Clouds**
  - AWS functionality locally makes moving using Amazon AWS easier, cheaper, and more sustainable
- **Provide a basic software development platform for the open source community**
  - E.g. the "Linux Experience"
- **Not designed as a replacement technology for AWS or any other Public Cloud service**

Requirements

- **Implement cloud abstractions and semantics**
  - Must be a cloud (inarguably)

- **Simple**
  - Must be transparent and easy to understand

- **Scalable**
  - Interesting effects are observed at scale (e.g. not an SDK)

- **Extensible**
  - Must promote experimentation

- **Non-invasive**
  - Must not violate local control policies

- **System Portable**
  - Must not mandate a system software stack change

- **Configurable**
  - Must be able to run in the maximal number of settings

- **Easy**
  - To distribute, install, secure, and maintain

Open-source Cloud Anatomy

Eucalyptus

- **Extensibility**
  - Simple architecture and open internal APIs
- **Client-side interface**
  - Amazon's AWS interface and functionality (familiar and testable)
- **Networking**
  - Virtual private network per cloud
  - Must function as an overlay => cannot supplant local networking
- **Security**
  - Must be compatible with local security policies
- **Packaging, installation, maintenance**
  - system administration staff is an important constituency for uptake

Architecture

Eucalyptus

Client-side API Translator

Cloud Controller

Database

Walrus (S3)

Cluster Controller

Node Controller

Storage Controller
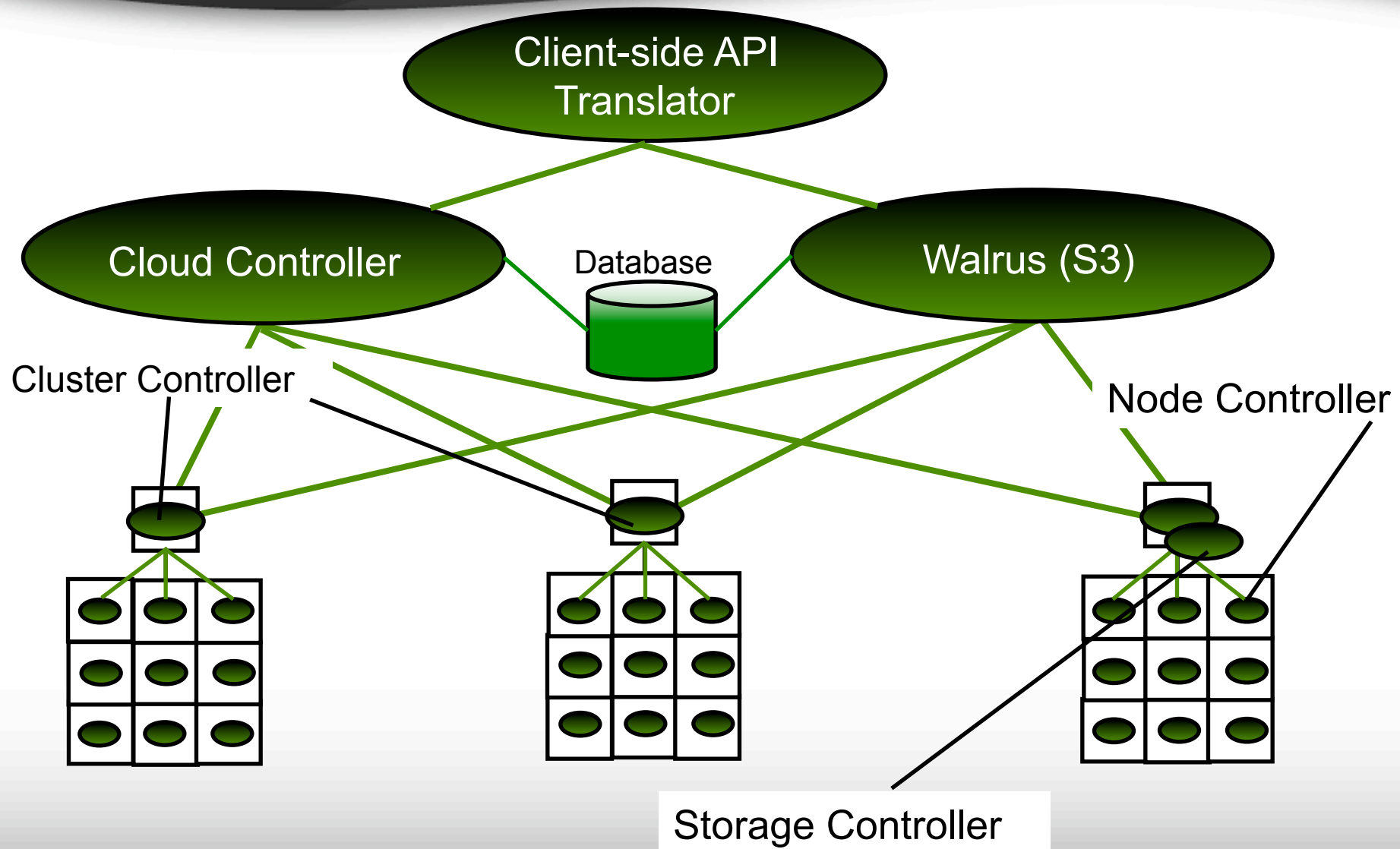
Notes from the Open-source Cloud

Eucalyptus

- **Private clouds are really hybrid clouds**
  - Users want private clouds to export the same APIs as the public clouds

- **In the Enterprise, the storage model is key**
  - Scalable "blob" storage doesn't quite fit the notion of "data file."

- **Cloud Federation is a policy mediation problem**
  - No good way to translate SLAs in a cloud allocation chain
  - "Cloud Bursting" will only work if SLAs are congruent

- **Customer SLAs allow applications to consider cost as first-class principle**
  - Buy the computational, network, and storage capabilities that are required

Cloud Myths

- **Cloud computing infrastructure is just a web service interface to operating system virtualization.**

  – "I'm running Xen in my data center – I'm running a private cloud."

- **Clouds and Grids are equivalent**

  – "In the mid 1990s, the term grid was coined to describe technologies that would allow consumers to obtain computing power on demand."

- **Cloud computing imposes a significant performance penalty over "bare metal" provisioning.**

 – "I won't be able to run a private cloud because my users will not tolerate the performance hit."

Clouds and Virtualization

Eucalyptus

- **Operating System virtualization (Xen, KVM, VMWare, HyperV) is only apparent for IaaS**
  - AppEngine = BigTable
- **Hypervisors virtualize CPU, Memory, and local device access as a <u>single</u> virtual machine (VM)**
- **IaaS Cloud allocation is**
  - Set of VMs
  - Set of storage resources
  - Private network

Requires more than
A set of Hypervisors

- **Allocation is atomic**
- **Each user gets an SLA**
  - Server consolidation must respect SLAs given to users

Datacenter Virtualization

Virtualization Control

Eucalyptus

Cloud Services
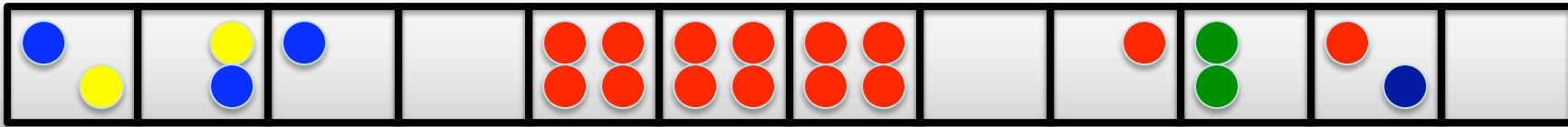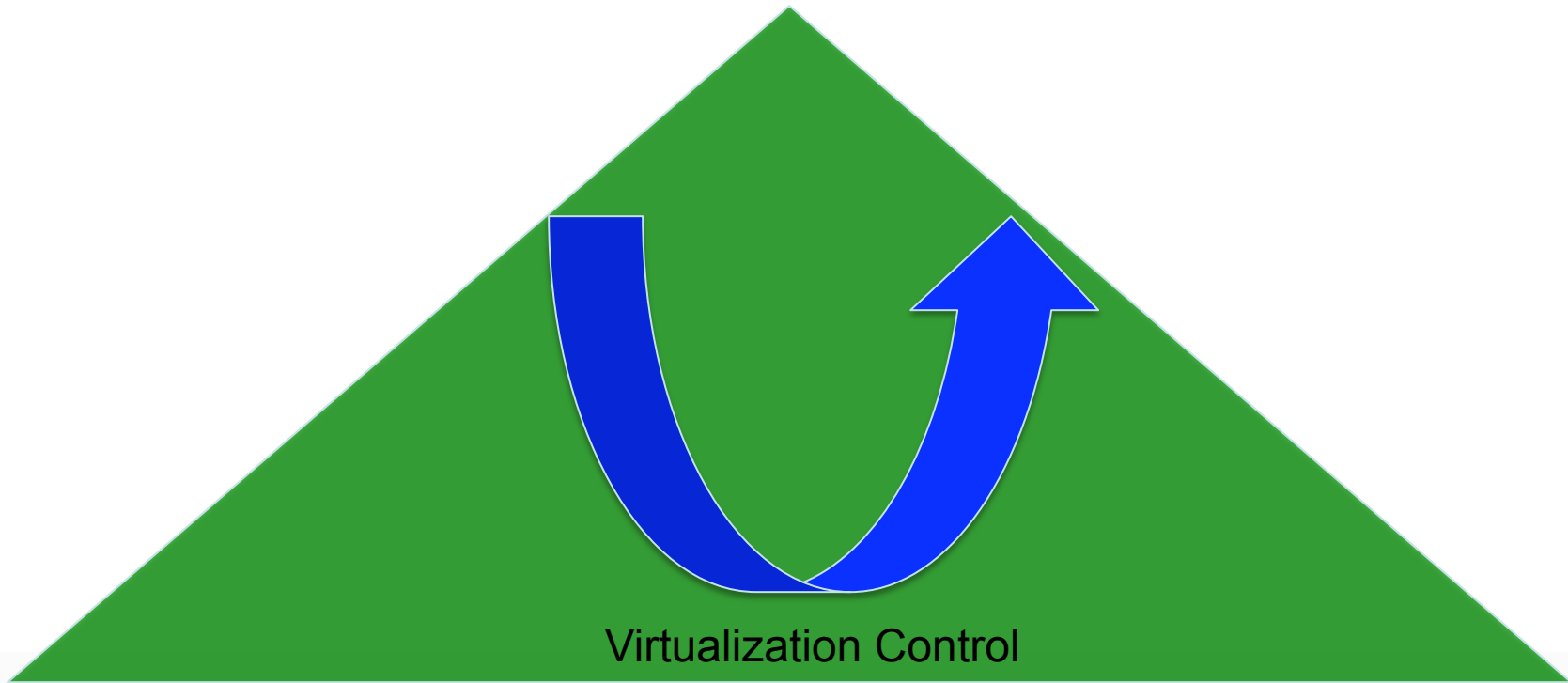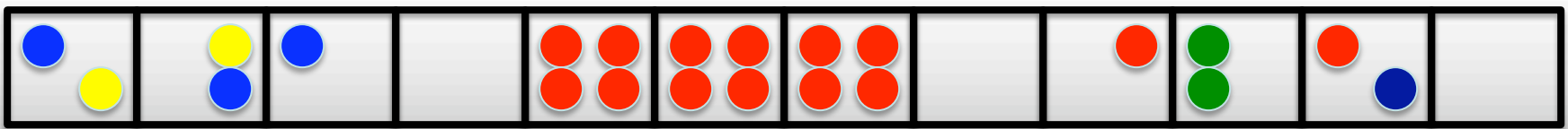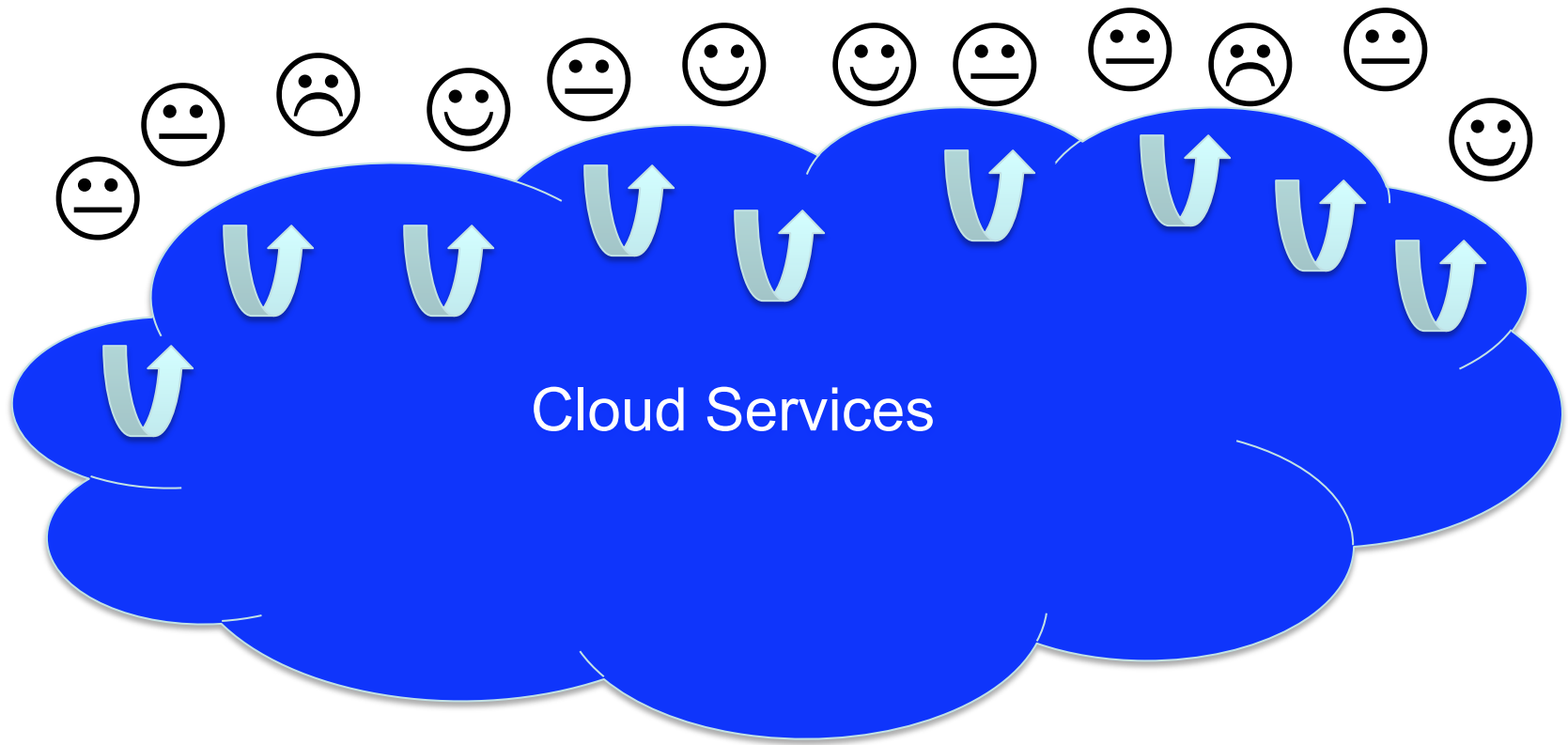
Clouds vs. Grids

Eucalyptus

- **Rich's assertion: Clouds and Grids are distinct**
- **Cloud**
  - Individual user can only get a tiny fraction of the total resource pool
  - No support for cloud federation except through the client interface
  - Opaque with respect to resources
- **Grid**
  - Built so that individual users can get most, if not all of the resources in a single request
  - Middleware approach takes federation as a first principle
  - Resources are exposed, often as bare metal
- **These differences mandate different architectures for each**

Cloud Speed

- **Extensive performance study using HPC applications and benchmarks**

- **Two questions:**
  - *What is the performance impact of virtualization?*
  - *What is the performance impact of cloud infrastructure?*

- **Tested Xen, Eucalyptus, and AWS (small SLA)**

- **Many answers:**
  - Random access disk is slower with Xen
  - CPU bound can be *faster* with Xen -> depends on configuration
  - Kernel version is far more important
  - Eucalyptus imposes no statistically detectable overhead
  - AWS small appears to throttle network bandwidth and (maybe) disk bandwidth -> *$0.10 / CPU hour*

Comparing TCP Performance between EC2 and EPC
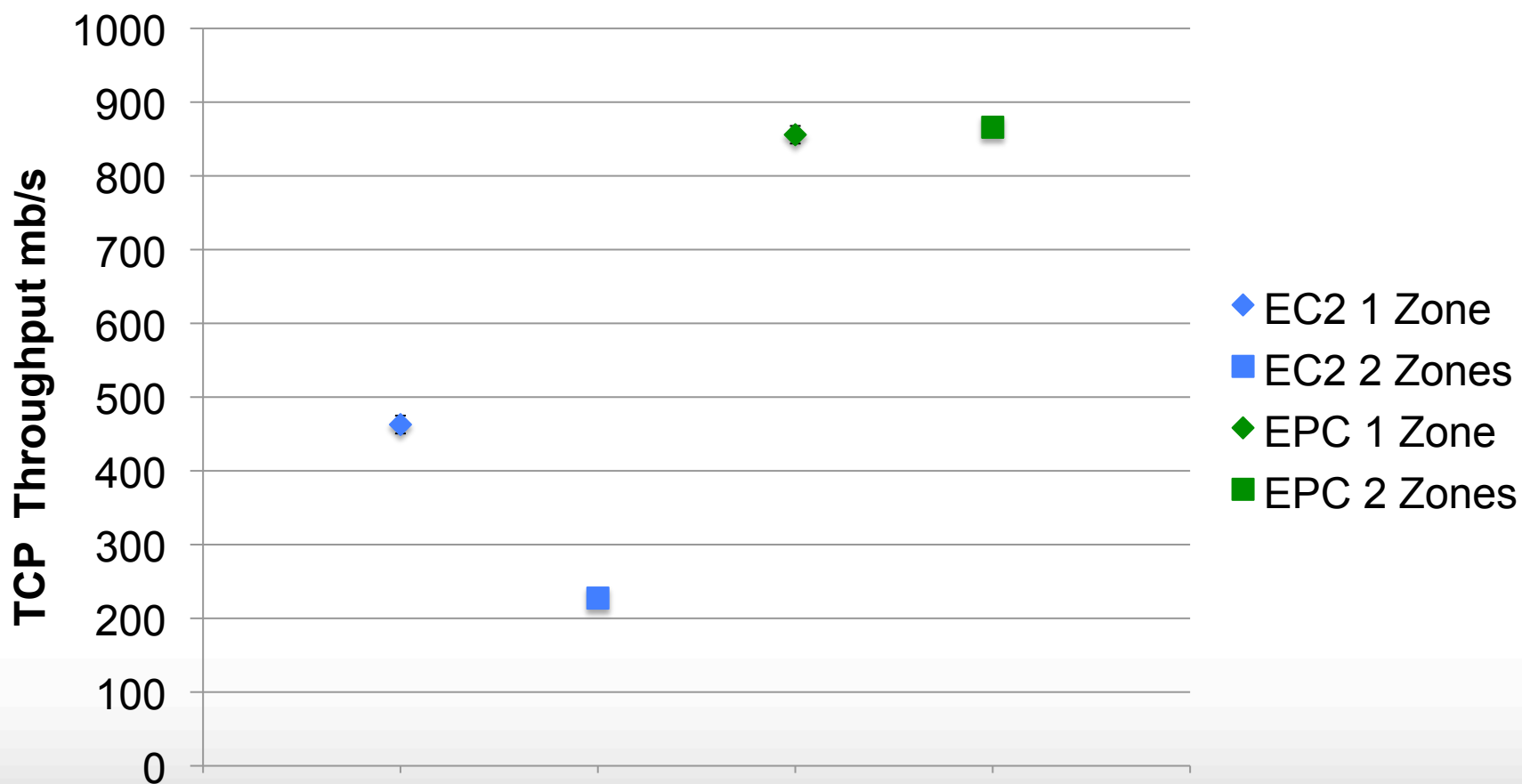
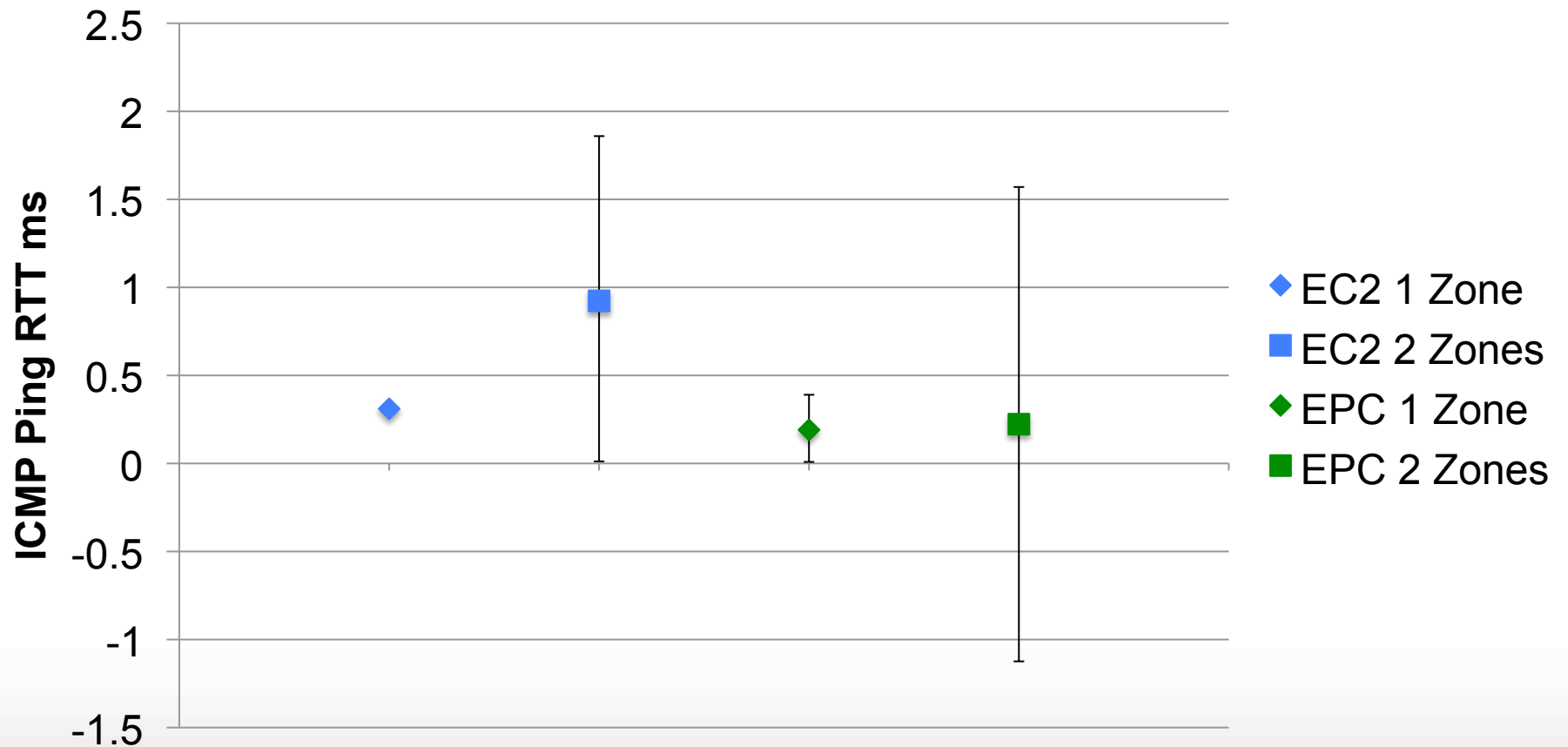# TCP RTT

Comparing ICMP Ping Performance between EC2 and EPC

Image Start-up
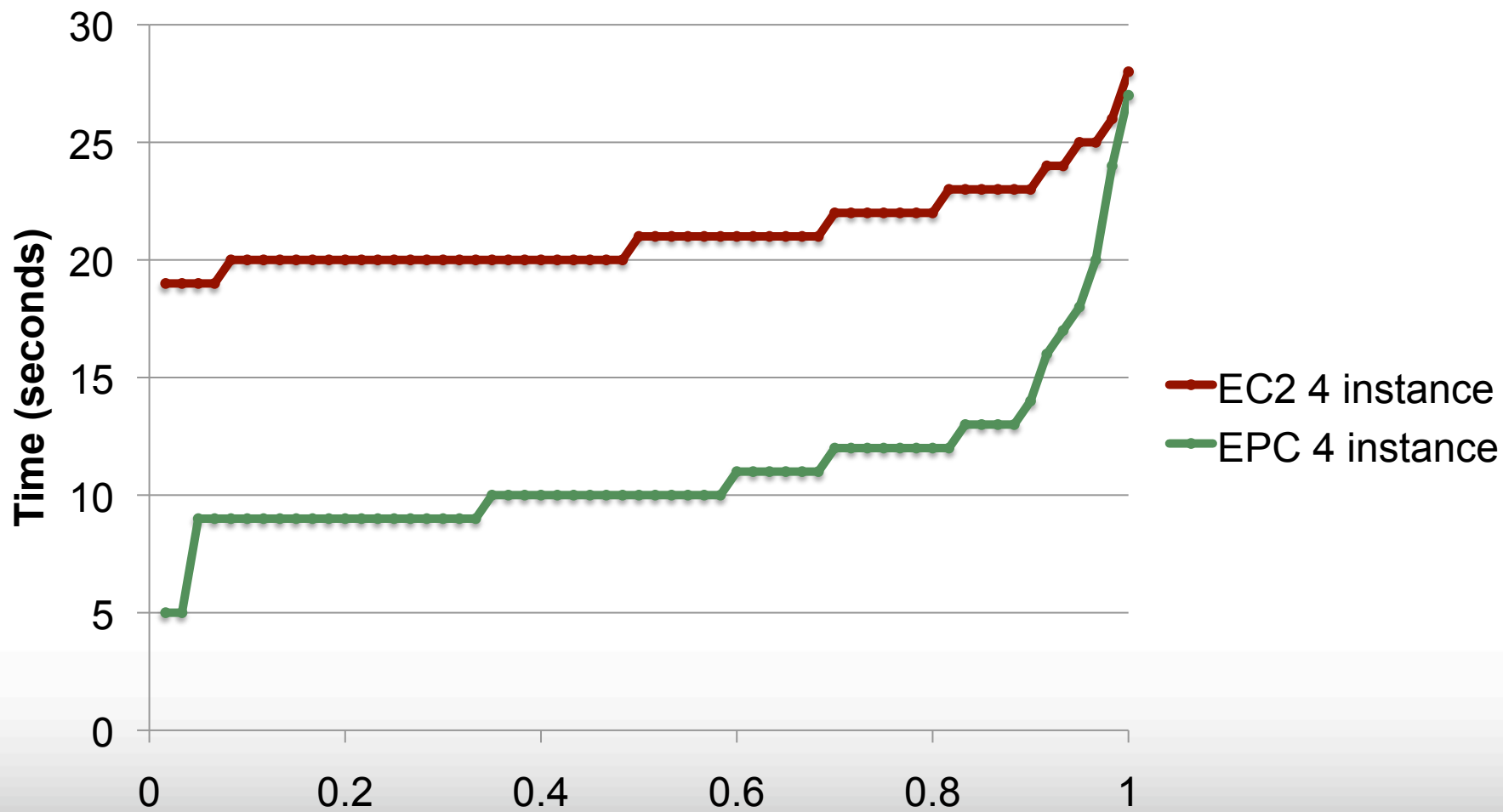
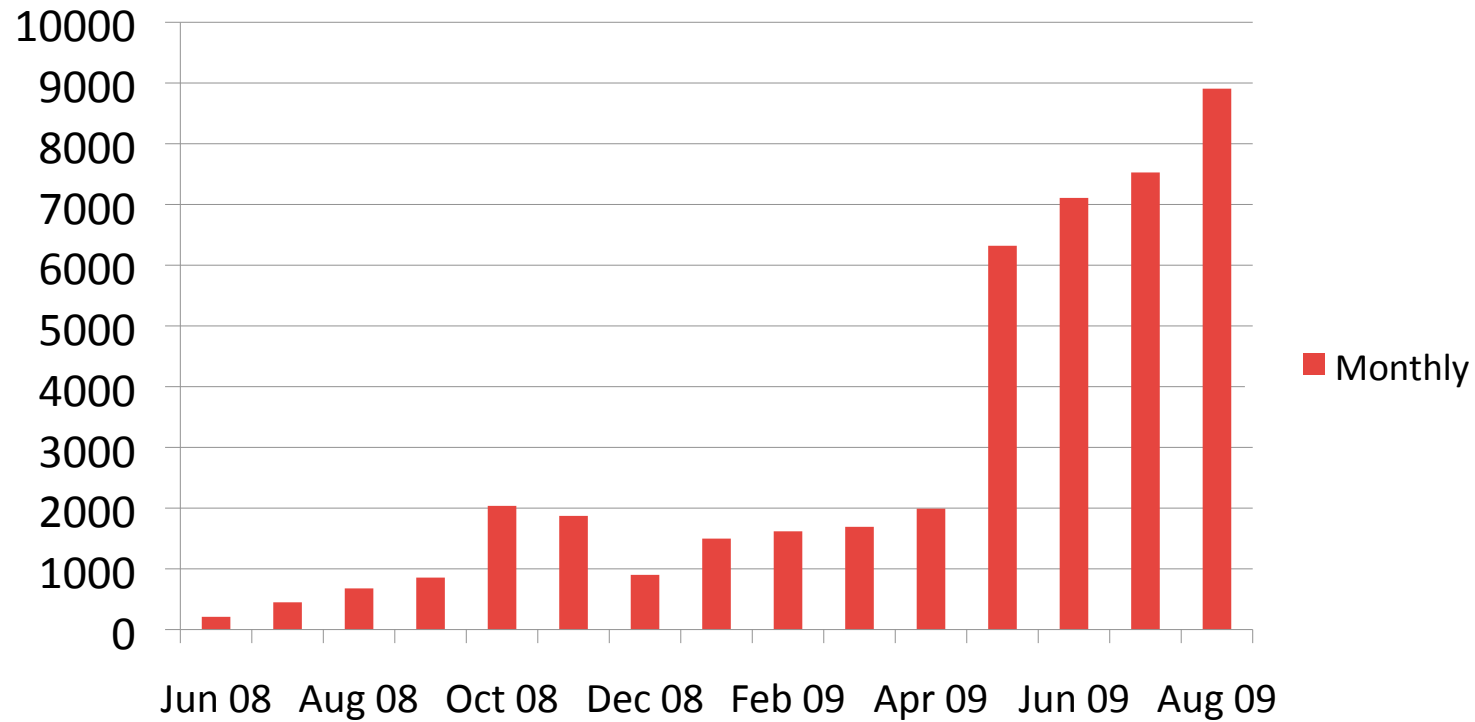Eucalyptus

Four Instance Start-up Time (CDF^{-1})

- EC2 4 instance
- EPC 4 instance

Uptake – 55K so far

Downloads (excluding Ubuntu 9.04)

No Eucalyptus in Antarctica (yet)

Open-source Distribution

Linux Distribution: *Ubuntu* and Eucalyptus

- Jaunty Jackalope "Powered by Eucalyptus"
  - April 23, 2009
  - Complete build-from-source
- Karmic Koala
  - October 23, 2009
  - Full-featured Eucalyptus
- Fundamental technology
  - "Ubuntu Enterprise Cloud" ecosystem surrounding Eucalyptus
- 10,000,000 potential downloads
- *Debian* "squeeze"
  - Source release packaging under way
- Packaged for *CentOS*, *OpenSUSE*, *Debian*, and *Ubuntu* as "binary" release as well

> " Make Eucalyptus the open source reference implementation for cloud computing. "
>
> Simon Wardley (head of cloud strategy), Canonical

Eucalyptus Ecosystem
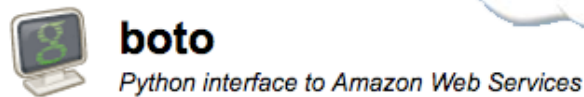
Open-source Roadmap

- **5/28/08 – Release 1.0 shipped**
- **8/28/08 – EC2 API and initial installation model in V1.3**
  - Completes overlay version
- **12/16/08 – Security groups, Elastic IPs, AMI, S3 in V1.4**
- **4/19/09 – EBS, Metadata service in V1.5.1**
- **4/23/09 - Ubuntu release**
- **4/27/09 – www.eucalyptus.com**
- **7/17/09 – Bug fix release in V1.5.2**
- **10/23/09 – Karmic Koala release**
  - 10^7 downloads from "main" archive
- **11/5/09 – Final feature release as V1.6.1**
  - Completes AWS specification as of 1/1/2009
- **1/1/09 – Final bug-fix release V1.6.2**
  - "core" opens for community contribution

You Are Here

Eucalyptus Systems Inc.

PHASE 1 – Services and Consulting

Bootstrap with high profile product and service agreements

- Enterprise credibility
- Market validation for product development
- Develop sales leads
- Current customers

PHASE 2 – Enterprise Products

Enterprise cloud products leveraging open source base

- Customized resource allocation and SLA's
- Enterprise cloud management tools
- Go to Market: Direct sales and download conversions

Thanks!

- **Thanks to our original research sponsors…**

UCSB   **VGrADS**   **NSF**

- **…and to our new commercial friends**

BENCHMARK CAPITAL   **BV CAPITAL** SILICON VALLEY EUROPE

www.eucalyptus.com
805-845-8000
rich@eucalyptus.com