

# Introduction to Compute Cloud

Tao Zou

CS 5220  
Applications of Parallel Computers

# About me

- 3<sup>Rd</sup> year PhD student in Computer Science
  - Databases/Cloud Computing/Distributed Systems
- Have been burning thousands of dollars in
  - Amazon Web Services (*Amazon Pays*)
  - Microsoft Azure (*Microsoft Pays*)
  - Rackspace Cloud (*My advisor Pays*)
- Who is paying for the cost?

# Where is the Cloud?



# Where is the Cloud?



- Datacenter



# Where is the Cloud?



# What is a Cloud?

- Service, rather than a product
- Cloud Models:

## **Infrastructure** as a Service

- Provides raw computing resources.
- Amazon EC2
- Rackspace

**Do** anything

## **Platform** as a Service

- Provides programming languages and tools.
- Microsoft Azure
- Google AppEngine

**Program** anything

## **Software** as a Service

- Provides applications.
- Salesforce.com
- Google Docs

**Use** anything

hardware focus, higher flexibility

application focus, less flexibility,

# Amazon Web Service: Elastic Compute Cloud(EC2)

## Compute

Amazon Elastic Compute Cloud (EC2)

Amazon Elastic MapReduce

Auto Scaling

## Content Delivery

Amazon CloudFront

## Database

Amazon SimpleDB

Amazon Relational Database Service (RDS)

Amazon ElastiCache

## Deployment & Management

AWS Elastic Beanstalk

AWS CloudFormation

## E-Commerce

Amazon Fulfillment Web Service (FWS)

## Industry-specific Clouds

AWS GovCloud (US)

## Messaging

Amazon Simple Queue Service (SQS)

Amazon Simple Notification Service (SNS)

Amazon Simple Email Service (SES)

## Monitoring

Amazon CloudWatch

## Networking

Amazon Route 53

Amazon Virtual Private Cloud (VPC)

Elastic Load Balancing

AWS Direct Connect

## Payments & Billing

Amazon Flexible Payments Service (FPS)

Amazon DevPay

## Storage

Amazon Simple Storage Service (S3)

Amazon Elastic Block Store (EBS)

AWS Import/Export

## Support

AWS Premium Support

## Web Traffic

Alexa Web Information Service

Alexa Top Sites

## Workforce

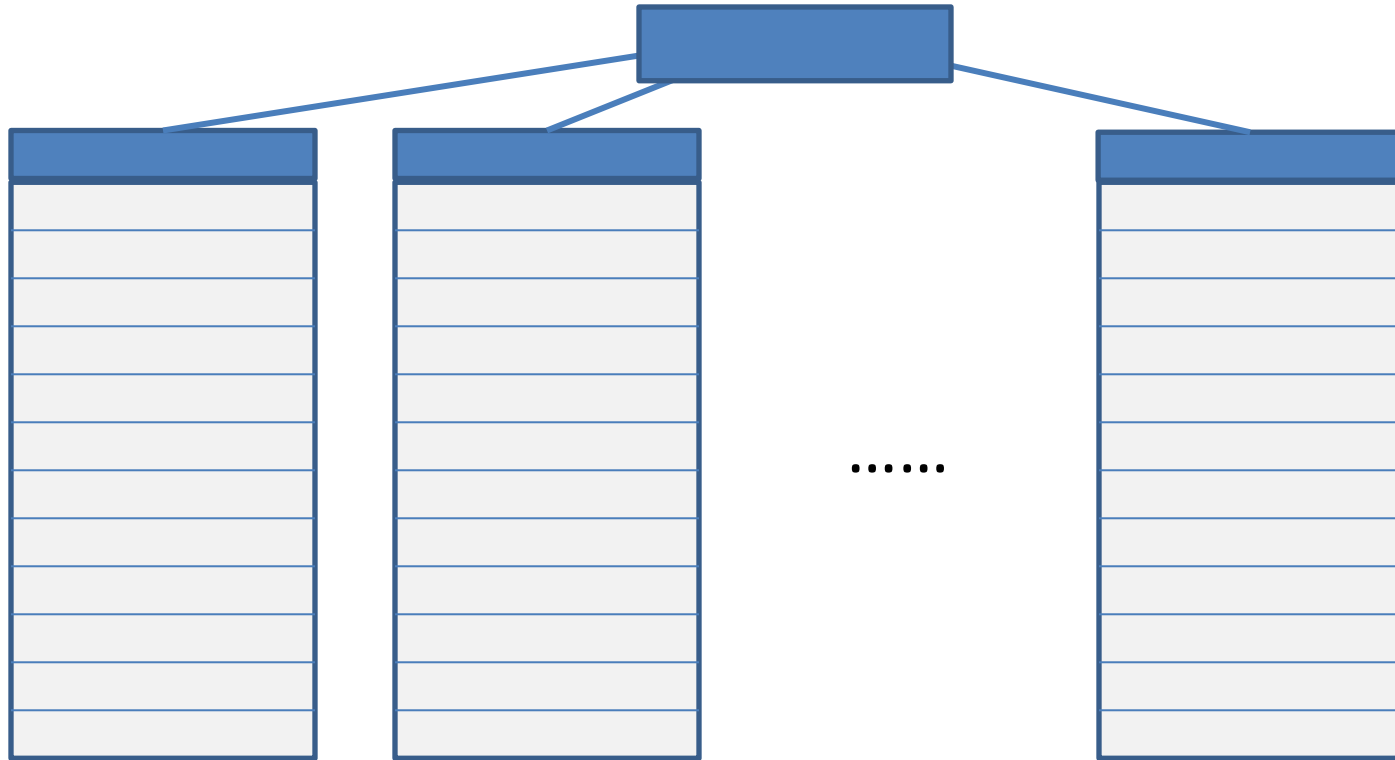
Amazon Mechanical Turk

# Amazon Elastic Compute Cloud(EC2)

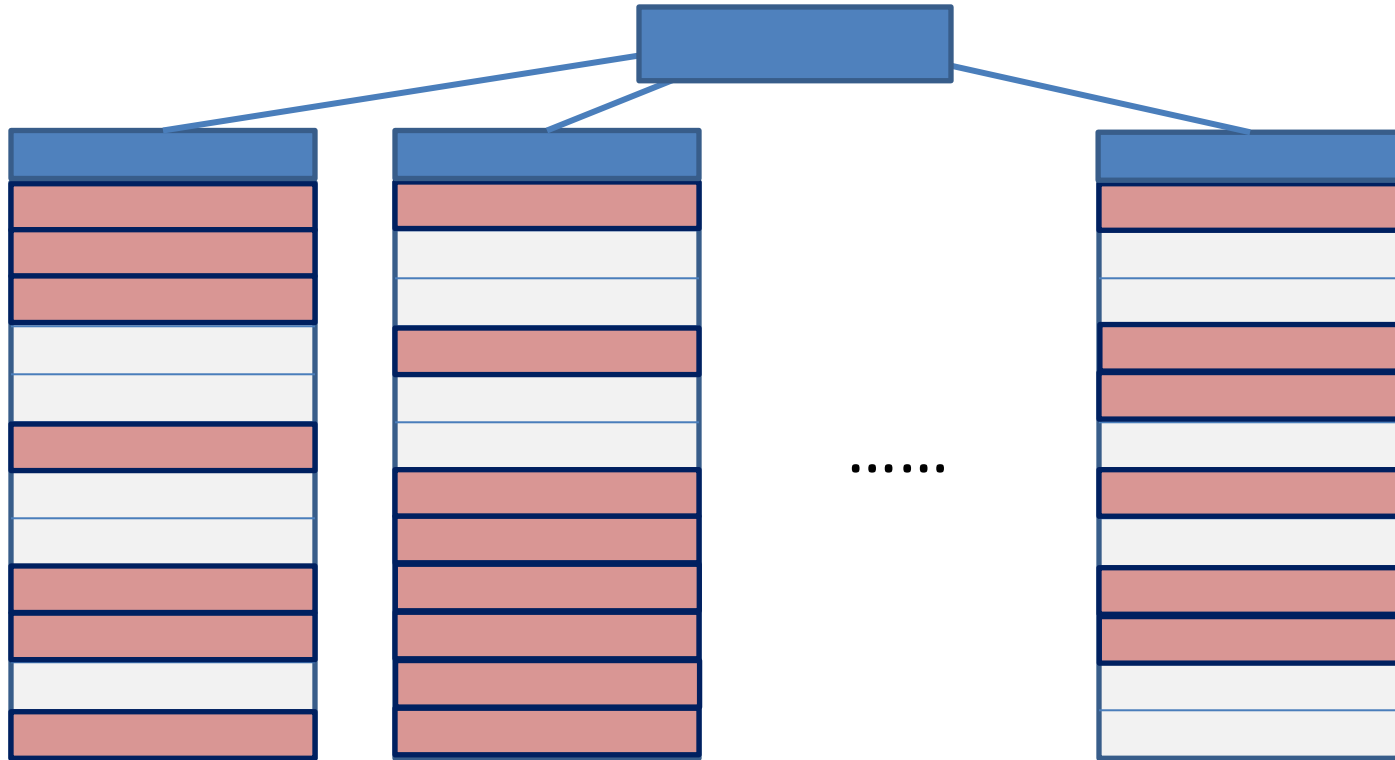




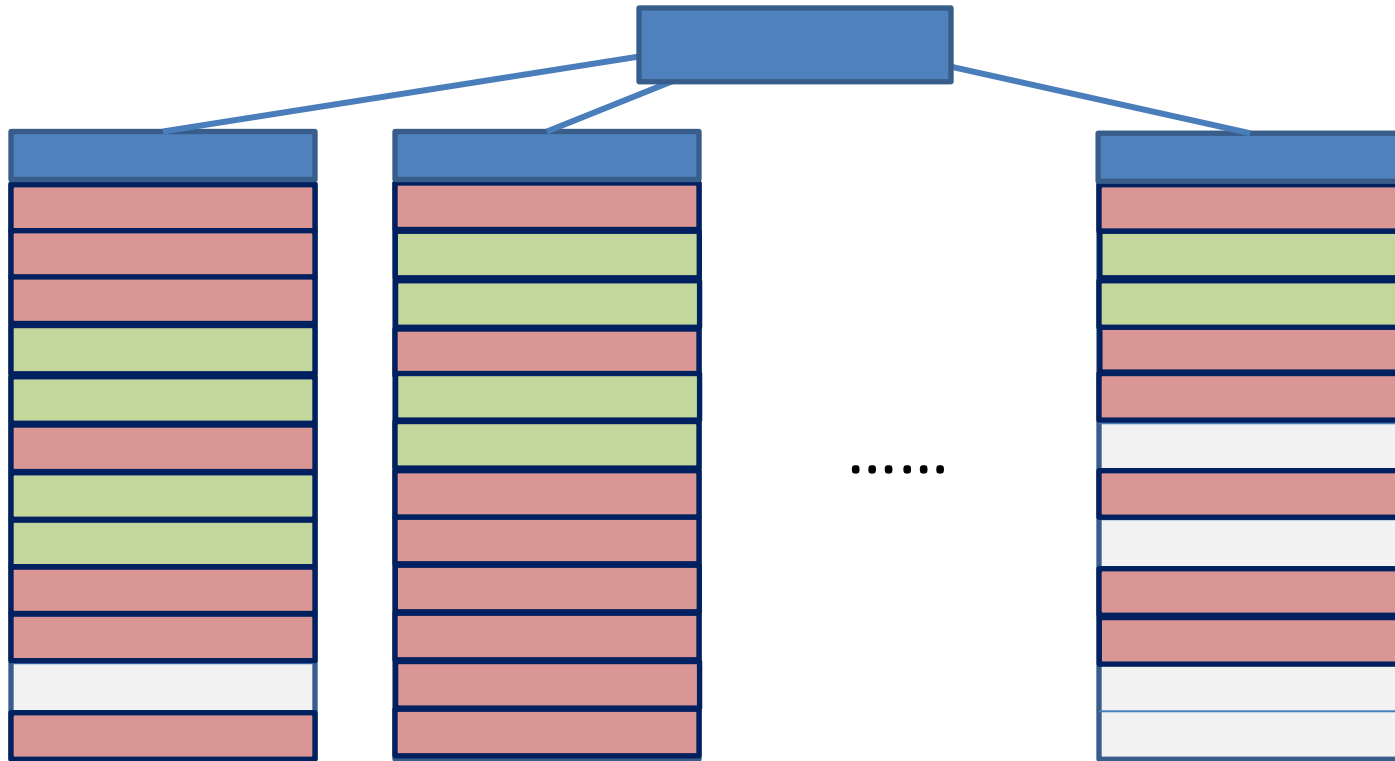
# Amazon Elastic Compute Cloud(EC2)



# Amazon Elastic Compute Cloud(EC2)

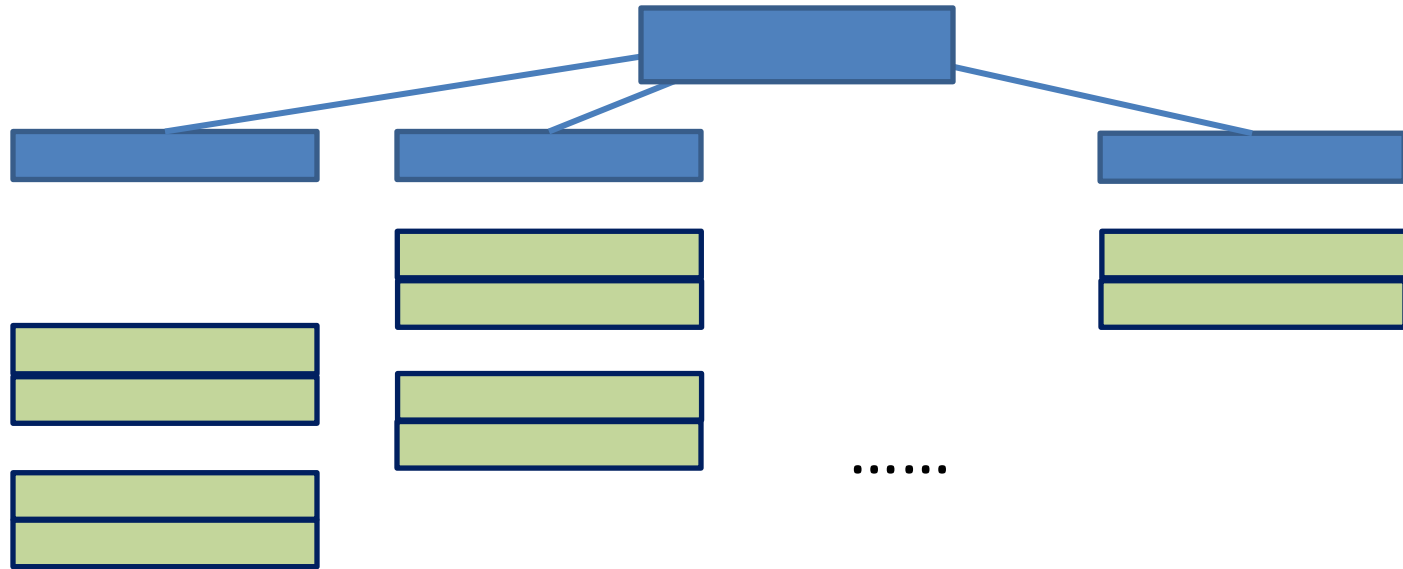


# Amazon Elastic Compute Cloud(EC2)



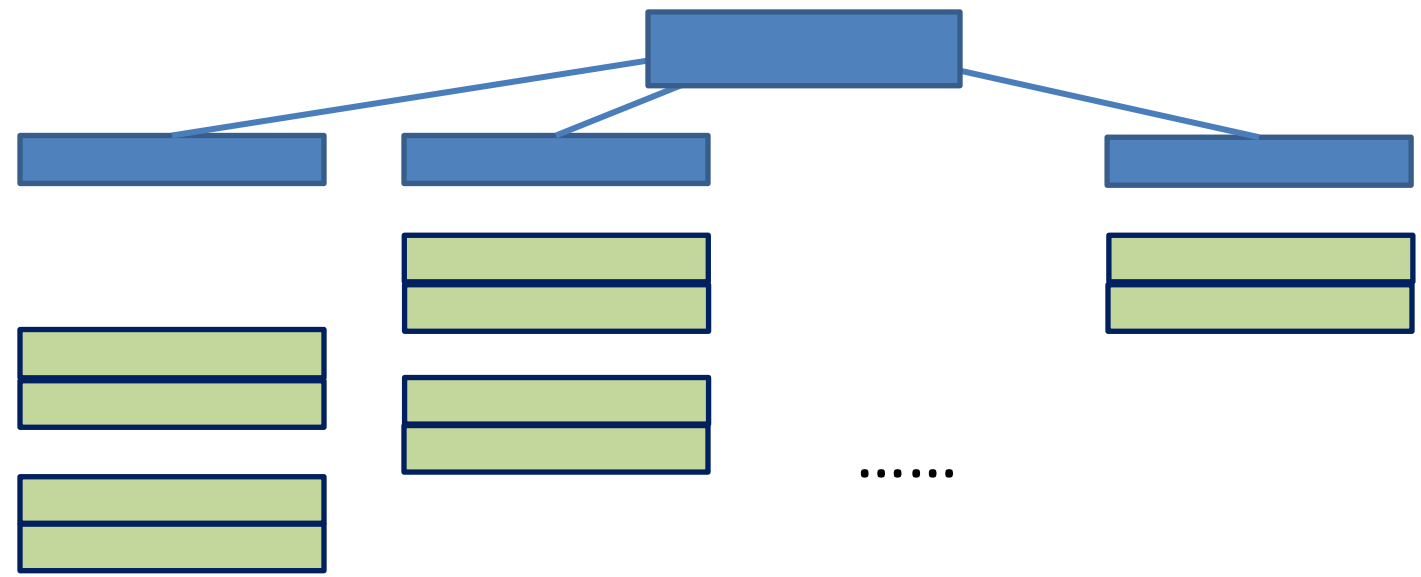
Give me 10 instances

# Amazon Elastic Compute Cloud(EC2)



Give me 10 instances

# Amazon Elastic Compute Cloud(EC2)

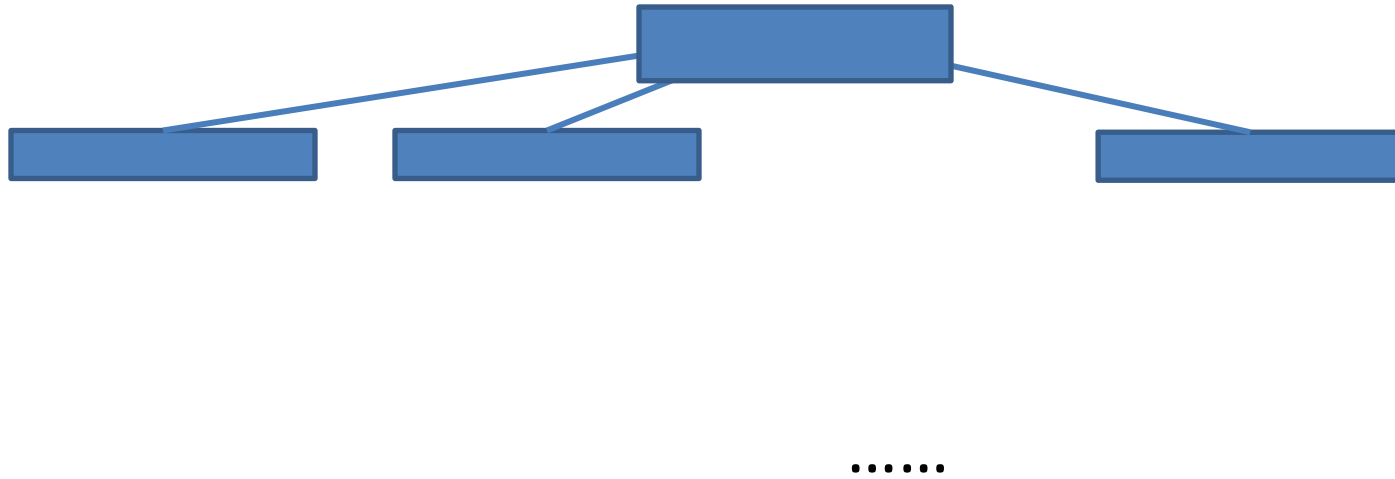


Give me 10 instances

Run Computation → takes X hours

Shutdown all my instances

# Amazon Elastic Compute Cloud(EC2)



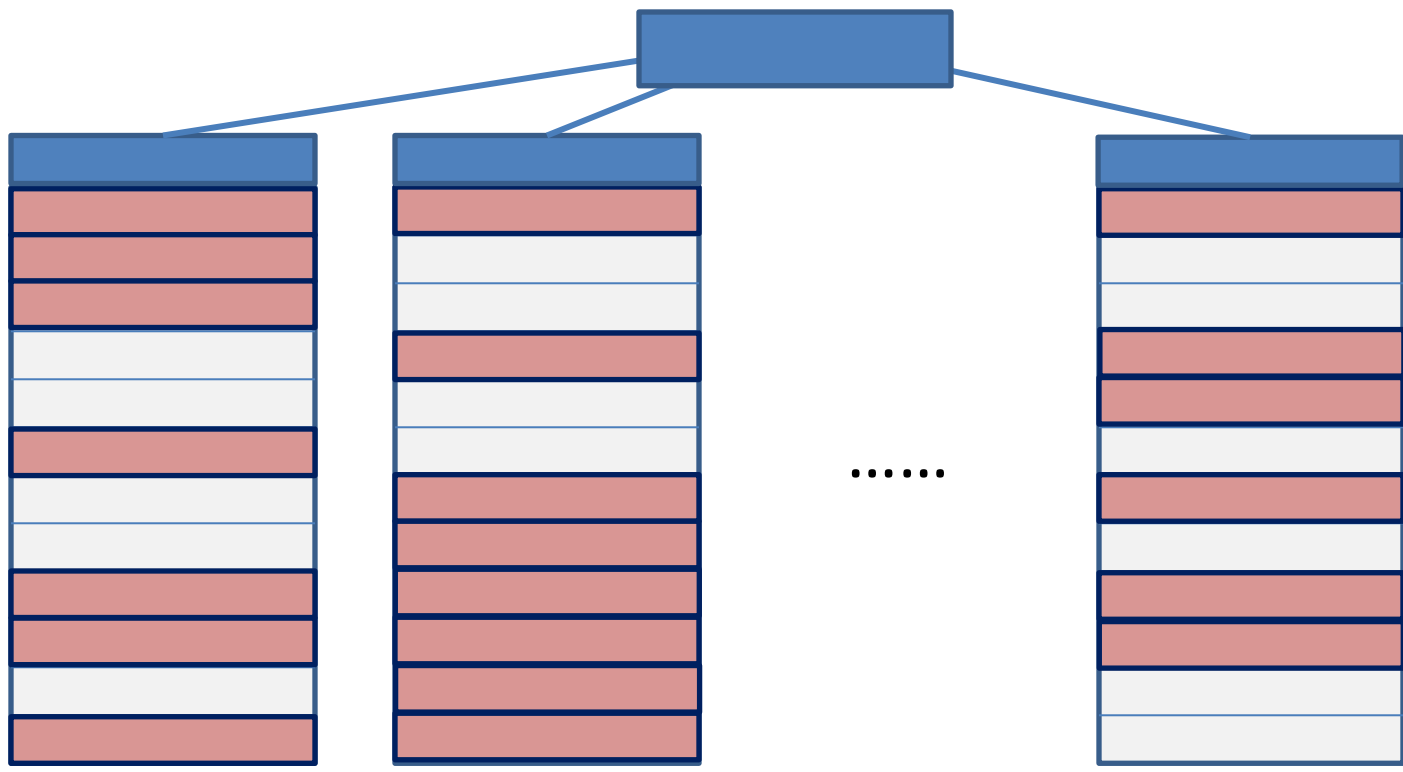
Give me 10 instances

Run Computation → takes X hours

Shutdown all my instances



# Amazon Elastic Compute Cloud(EC2)



Give me 10 instances

Run Computation → takes X hours

Shutdown all my instances

# EC2 Pricing Model: Pay As You Go

- Pay only for what you use
  - Machine hours (  $10 \cdot [X]$  )
  - Type of instances

Region: <input type="text" value="US East (Virginia)"/>	Linux/UNIX Usage	Windows Usage
<b>Standard On-Demand Instances</b>		
Small (Default)	\$0.085 per hour	\$0.12 per hour
Large	\$0.34 per hour	\$0.48 per hour
Extra Large	\$0.68 per hour	\$0.96 per hour
<b>Micro On-Demand Instances</b>		
Micro	\$0.02 per hour	\$0.03 per hour
<b>Hi-Memory On-Demand Instances</b>		
Extra Large	\$0.50 per hour	\$0.62 per hour
Double Extra Large	\$1.00 per hour	\$1.24 per hour
Quadruple Extra Large	\$2.00 per hour	\$2.48 per hour
<b>Hi-CPU On-Demand Instances</b>		
Medium	\$0.17 per hour	\$0.29 per hour
Extra Large	\$0.68 per hour	\$1.16 per hour
<b>Cluster Compute Instances</b>		
Quadruple Extra Large	\$1.60 per hour	\$1.98 per hour
<b>Cluster GPU Instances</b>		
Quadruple Extra Large	\$2.10 per hour	\$2.60 per hour

- Large instances  $\rightarrow$  total cost (  $0.34 \cdot 10 \cdot [X]$  )

# EC2 Pricing Model: Instance Types

- **Large Instance**
  - 7.5 GB memory
  - 4 EC2 Compute Units (2 virtual cores with 2 EC2 Compute Units each)
  - 850 GB instance storage
  - I/O Performance: High
- **Cluster Compute Quadruple Extra Large Instance**
  - 23 GB of memory
  - 33.5 EC2 Compute Units (2 x Intel Xeon X5570, quad-core “Nehalem” architecture)
  - 1690 GB of instance storage
  - 64-bit platform
  - I/O Performance: Very High (10 Gigabit Ethernet)
- EC2 Compute Units = CPU capacity of a 1.0-1.2 GHz 2007 Xeon processor

# Why Cloud?

- Time-to-solution
  - $1000 \text{ cores} * 1 \text{ hour} = 1 \text{ core} * 1000 \text{ hours}$
- Scalability: Can easily get 1000 cores
  - For \$170 per hour
- Instance Availability
  - No queuing time

# How to Use EC2

# AWS Management Console

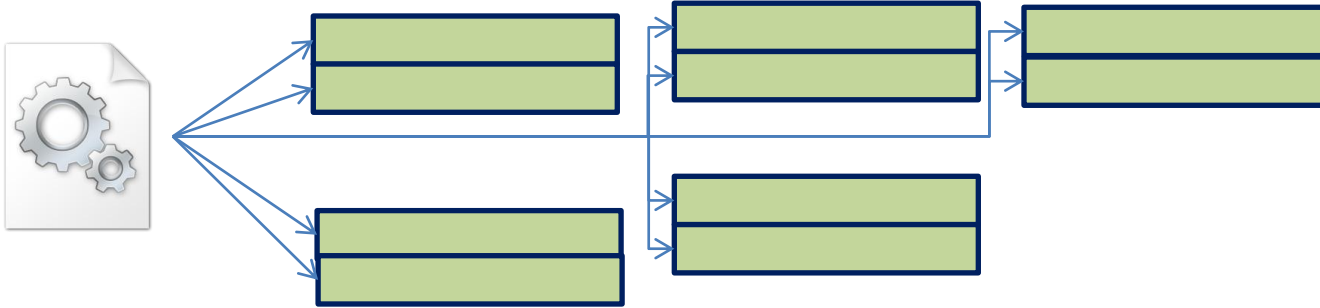
The screenshot shows the AWS Management Console interface. At the top, there is a navigation bar with various AWS services. Below this, the 'My Instances' page is displayed. The left sidebar contains a 'Navigation' menu with a 'Region' dropdown set to 'US East (Virginia)'. The main content area features a 'Launch Instance' button and a table of instances. The table has the following columns: Name, Instance, AMI ID, Root Device, Type, Status, Security Groups, Key Pair Name, Monitoring, and Virtualization. One instance is listed with the following details:

Name	Instance	AMI ID	Root Device	Type	Status	Security Groups	Key Pair Name	Monitoring	Virtualization
empty	i-7f05851c	ami-1b814f72	ebs	t1.micro	running	quick-start-1	tao	basic	paravirtual

- OR using command line



# EC2: Start Instance



- Load OS Image
  - Similar to VMware, VirtualBox
  - But to all instances in parallel
- Takes < 1 minutes

# EC2: Start Instance

**Create a new instance** Cancel

Select an option below:

- Launch Classic Wizard**  
Continue to the classic wizard which provides you with the full list of AMIs as well as fine-grained control over how you would like your instance to be launched.
- Quicklaunch**  
Select from a list of popular configurations to launch your instance into the cloud as quickly as possible.




[Submit feedback](#)

**Name your instance:**  Pick a meaningful name, e.g. Web Server


**Choose a Key Pair:**  
Public/private key pairs allow you to securely connect to your instance after it launches.

- Select an existing Key Pair**
- Create new Key Pair**

**Choose a Launch Configuration:**

 <b>Amazon Linux</b> Includes the EC2 AMI Tools.	64 bit <input checked="" type="radio"/> 32 bit <input type="radio"/>
 <b>Microsoft Windows Server 2008</b> Microsoft Windows 2008 Datacenter edition	64 bit <input checked="" type="radio"/> 32 bit <input type="radio"/>
 <b>Microsoft Windows Server 2008 with SQL Server Express</b> Includes Microsoft SQLServer 2008 Express, Internet Information Services 7, and ASP.NET 3.5.	64 bit <input checked="" type="radio"/> 32 bit <input type="radio"/>

Note: You can customize your settings in the next step.

**Continue** 

- **ec2-run-instances** AMI [-n INSTANCE\_COUNT] .....

# EC2: Login using SSH

- `ssh -i KeyPair root@ec2-107-20-54-150.compute-1.amazonaws.com`

```
  _ |  _ |  ) Fedora 8
  _ | (  _ /   64-bit
  _ | \ _ | _ |
Welcome to an EC2 Public Image
      :-)
Base
--[ see /etc/ec2/release-notes ]--
[root@ip-10-203-26-198 ~]# █
```

- `mpirun -hostfile ~/hosts -n 100 fishSim.exec`

# EC2: Save Image

- ec2-bundle-instance
- Takes ~10 minutes (except Cluster/GPU Compute Instance)
- Typical Procedure
  - Start one instance from a default image
  - Install MPI, OpenMP,...
  - Save Image
  - Start 100 instances from the saved image

# EC2: Terminate Instances

The screenshot shows the AWS Management Console for Amazon EC2. The navigation pane on the left includes sections for INSTANCES (Instances, Spot Requests, Reserved Instances), IMAGES (AMIs, Bundle Tasks), ELASTIC BLOCK STORE (Volumes, Snapshots), and NETWORK & SECURITY (Security Groups, Elastic IPs, Placement Groups, Load Balancers, Key Pairs). The main content area displays a table of instances. The first instance is named 'empty' with Instance ID 'i-7f05851c', AMI ID 'ami-1b814f72', Root Device 'ebs', Type 't1.micro', and Status 'running'. A context menu is open over this instance, listing various actions under 'Instance Management' (Connect, Get System Log, Create Image (EBS AMI), Add/Edit Tags, Change Security Groups, Change Source / Dest Check, Launch More Like This, Disassociate IP Address, Change Termination Protection, View/Change User Data, Change Instance Type, Change Shutdown Behavior), 'Instance Lifecycle' (Terminate, Reboot, Stop, Start), and 'CloudWatch Monitoring' (Enable Detailed Monitoring, Disable Detailed Monitoring). The 'Terminate' option is highlighted.

Name	Instance	AMI ID	Root Device	Type	Status	Security
empty	i-7f05851c	ami-1b814f72	ebs	t1.micro	running	quick-sta

- **ec2-terminate-instances** INSTANCEID [INSTANCEID ...]

# How to get it for free

- AWS in Education
  - <http://aws.amazon.com/education/>
- For teaching:
  - CS 5300 - The Architecture of Large- Scale Information Systems
- For research:
  - Quarterly research grant applications
    - Next deadline: February 10, 2012
    - Renewable: maximum of two grant awards per academic year