

CS15-319 / 15-619

Cloud Computing

Recitation 2

January 20 & 22, 2015

Overview

- **Setup your instance for the demo**
 - Information in the handout
- **Administrative Issues**
 - TA hours, guidelines on Piazza posts
- **Last Week's Reflection**
 - Project Primer, OLI Unit 1, Module 1
- **This Week's Schedule**
 - Deadlines for OLI Unit 1, Module 2, Quiz 1 & Project 1.1
- **Demo**
- **Questions**

Administrative

- TA office hours are posted on Piazza and [Google calendar](#).
- Suggestions for using Piazza
 - Discussion forum, contribute questions and answers
 - Read the Piazza Post Guidelines ([@20](#)) before asking
 - Read Piazza questions & answers carefully to avoid duplicate ones
 - Don't ask a public question about a quiz question
 - Try to ask a public question if possible

Platforms

- Open Learning Initiative (OLI)
 - Access through Blackboard
 - Contains Units and Quizzes
- Amazon Web Services (AWS) Account
 - Create AWS account ([@8](#))
 - Complete [Account Linking Form](#)
 - Receive email request and **click link to confirm!**
- <http://theproject.zone>
 - Project write up, submissions and scoreboard
 - Registration Link in Email
- Piazza
 - Discussion forum
- If you do not have access to all of these platforms, please contact us immediately!

FAQ this week, 1

- Q: What support option should I use when creating my AWS Account?
 - A: Choose “basic” support level when signing up for your AWS account.
 - If you didn’t, change it ASAP.
- Q: How do I Link the AWS Account to the Course?
 - Read the Primer and [@8](#)
- Q: How do I Know if I’m linked yet?
 - Check the “My Account” section in AWS console
 - More information in [@8](#)

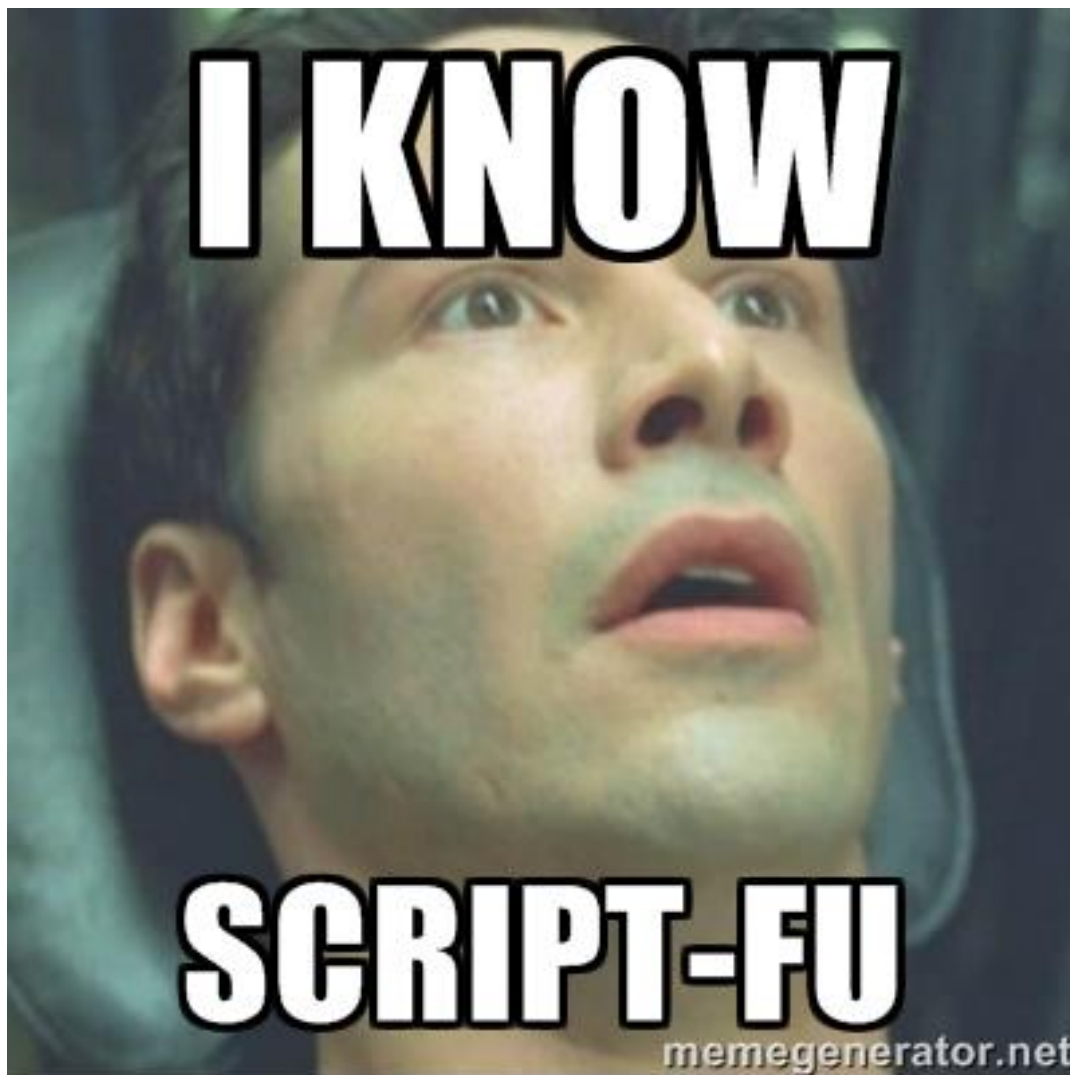
FAQ this week, 2

- Working with EC2 Instances
 - Q: Permission denied for SSH to EC2 instances. Why?
 - A: Change access permission of .pem file to 600. Connect using the appropriate username (ubuntu or ec2-user)
 - Q: How do I connect to EC2 instances in Windows?
 - A: Follow instructions in the Primer. Recommended clients are [XShell4](#) and [Putty](#) ([instructions](#))

FAQ this week, 3

- Q: Do I need to learn bash, python, perl, ruby, java, OCaml, haskell, erlang, FORTRAN77, ADA, ALGOL, COBOL for this course?
 - Use any language that works with AWS
 - Staff can support bash, python and Java
 - It's extremely useful develop your scripting skills
 - Read the Primer for details
- **Note:** We have a MapReduce project which requires knowledge of Java, at the end of the course.

So, when you finish the course, you can proudly say, among other skills...



Amazon Web Services (AWS) Account –1

- Students who are just joining us or who have not completed the AWS Account Setup:

=== ONLY IF YOU HAVEN'T DONE SO ALREADY ===

- Setup an AWS Account
- Complete AWS information using the link in [@8](#)
- Wait to receive Consolidated Billing Request email from Amazon
 - Manual process, waiting time varies
- Click the link to verify the linked billing
 - Many students have not clicked on the link yet!

Amazon Web Services (AWS) Account - 2

- **ALL STUDENTS:**

- If you have created an AWS account and have not received notification that it was linked to our account
 - Your credit card on file may get charged!
 - We CANNOT reimburse you!!!
 - It is your responsibility to verify that your account is linked to us correctly.
 - Contact us if there are any issues related to your account.

TheProject.Zone

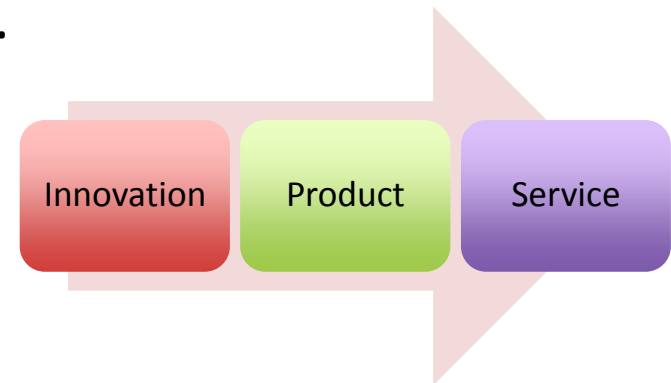
- Contains Project
 - Writeups
 - Submissions
 - Scoreboard
- You should be registered and have a login at this point.
- Let us know if you have not gotten access yet

Last Week Reflection

- Reading: Unit 1: Introduction to Cloud Computing
 - Module 1 : Introduction to Cloud Computing
- Project: Project Primer
 - Expected Linux Skills
 - AWS: Checking expenditure on AWS
 - EC2: Launching and remote logging into an EC2 micro instances
 - Tagging all instances and AWS resources
 - Tags provided in the project write-up on OLI
 - S3: Creating and modifying S3 Buckets

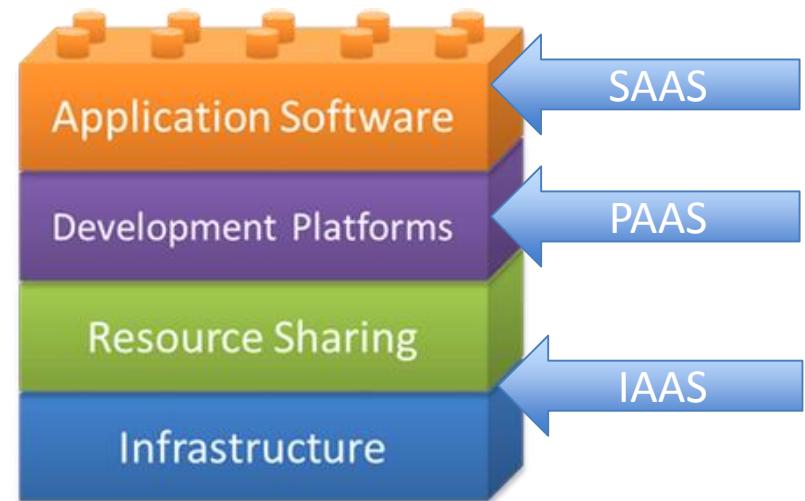
Module 1: Introduction to Cloud Computing

- Definition
 - delivery of computing as a service over a network whereby distributed resources are provided to the end user as a utility.
- Enabling Technologies
 - Networking, virtualization, etc.
- Deployment Models
 - Private, Public, Hybrid
- Benefits and Risks
- Economies of Scale



Module 2: Building Blocks and Service Models

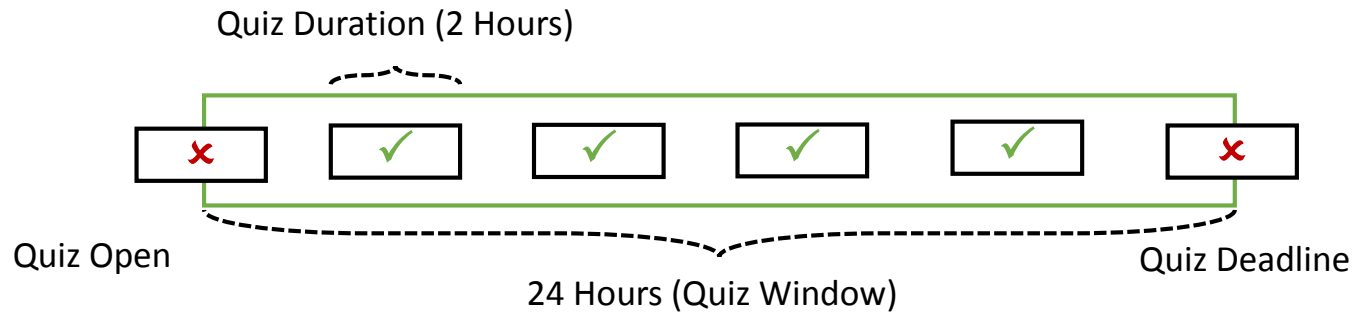
- Building Blocks
- Service Models
 - IAAS, PAAS, SAAS
- SLAs, SLOs and Auditing
- Cloud Security
- Popular Cloud Stacks
- Use Cases



This Week's Schedule

- Complete Unit 1 (Modules 1 & 2) by Thursday
 - Complete activities on each page
 - In-module activities are not graded
 - If you encounter a bug in the OLI write-up
 - Provide feedback at the end of each OLI page
- **Take Checkpoint Quiz 1**
 - Submission Deadline, Friday, Jan 23, 11:59pm EST
- **Complete Project 1.1 (Sequential Analysis)**
 - Submission Deadline, Sunday, Jan 25, 11:59pm EST

Quiz 1



- Quiz 1 will be open for 24 hours, Friday, Jan 23
 - Quiz 1 becomes available on **Jan 23, 00:01 AM EST.**
 - Deadline for submission is **Jan 23, 11:59 PM EST.**
 - Once open, you have **120 min** to complete the quiz.
 - Late submissions are NOT accepted.
 - You may not start the quiz after the deadline has passed.
 - **Maintain your own timer from when you start the quiz.**
 - **Click submit before deadline passes. No Exceptions!**

Location	Silicon Valley	Pittsburgh	Rwanda	Adeliade
Open	Jan 22, 09:01 PM	Jan 23, 00:01 AM	Jan 23 07:01 AM	Jan 23 03:31 PM
Deadline	Jan 23, 08:59 PM	Jan 23, 11:59 PM	Jan 24 06:59 AM	Jan 24 03:29 PM

Quiz 1

- 5% of your Overall Grade
- You only have 1 attempt
- You can save your Quiz answers
 - Highly recommended
 - Save prompt every 15 minutes
- What can I expect from the Quiz?
 - Questions similar to the activities in the Units
 - multiple choice, fill-in-the-blanks, numeric questions, ...
- Feedback for Quiz 1 is released after the deadline passes

Completing Projects in this Course

- Provision EC2 instances
 - Use the AMIs we provide for the project
 - Launch on-demand or spot
 - Tag the instances!
- Monitor their cost
 - Manage your instances
- Complete tasks for each project module
- Submit your work
 - Pledge of integrity
 - Results in scoreboard
- Terminate all instances

Amazon Machine Images (AMIs)

- What are AMIs?
 - An image of the OS with other software bundled up.
 - We have created custom AMIs for this course.
 - Everything that you do is tracked, so **DO NOT CHEAT!**
 - AMIs are specified when creating virtual machines within EC2
- Finding AMIs
 - Course AMIs are Community AMIs
- Potential issues
 - Regions: use US-East-1 (N. Virginia) for this course
 - Security Group : manage inbound ports (e.g. 22, 80, 8080)

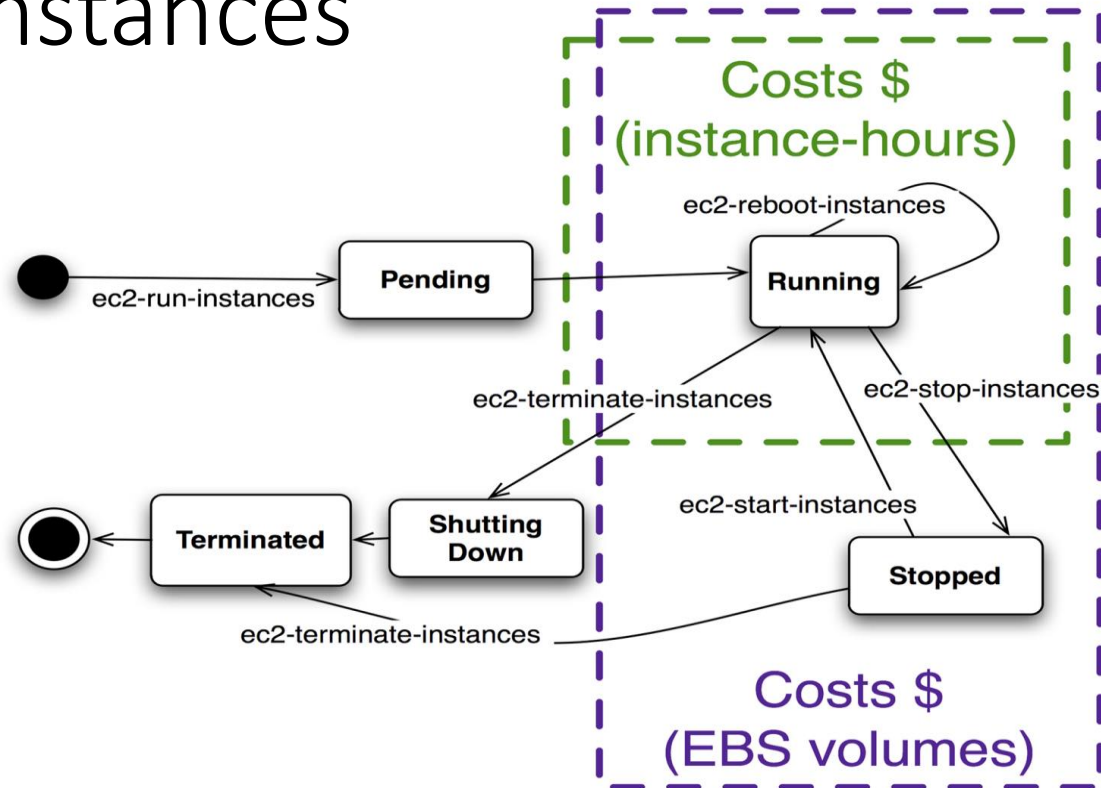
AMI Authentication

- Our AMIs authenticate during launch
 - Transparent process
 - You need to launch the instance from an account which is associated with your profile in <http://theproject.zone>
- Instance will shutdown if:
 - You launch an instance from an AWS account that is not registered in <http://theproject.zone>
 - You try to clone the AMI and launch it from your account

Using Spot Instances

- How to use spot instances
 - Spot instances allow you to bid on unused capacity
 - Specify a maximum bid for an instance
 - Your bid should be larger than the current price
 - Click [here](#) to find current spot prices (vary for each zone)
- Charging model for spot instances
 - Amazon EC2 instances are billed on a non-prorated basis
 - Each partial hour is billed as a full hour.
 - If Amazon terminates your instance, you do not get charged for that hour
 - If you terminate, you are charged for that hour

Manage Your Instances



- When you are done using an instance (e.g. finish the project), terminate
 - All data is deleted
- When you stop an instance, you incur EBS costs.
 - Data is persisted in EBS

Tagging AWS Resources

- How to tag instances & other resources
 - Format: Key: **Project** Value: **x.x**
 - Click [here](#) for more info
- Why tagging
 - Manage resources & costs
- Issues faced with tagging
 - Lower case/upper case
 - Misspellings
- This week: **Key: Project, Value: 1.1**

TPZ Scoreboard

THEPROJECTZONE

[Home](#)

[Primer](#)

[Project 1](#)

[Project 2](#)

[Project 3](#)

[Project 4](#)

[15619 Project](#)

[Profile](#)

[Log Out](#)

[suhail_the_man](#)

1.1 Sequential Analysis

[Writeup](#)

[Submissions](#)

[ScoreBoard](#)

Column headers: **click** to **sort**.

Search:

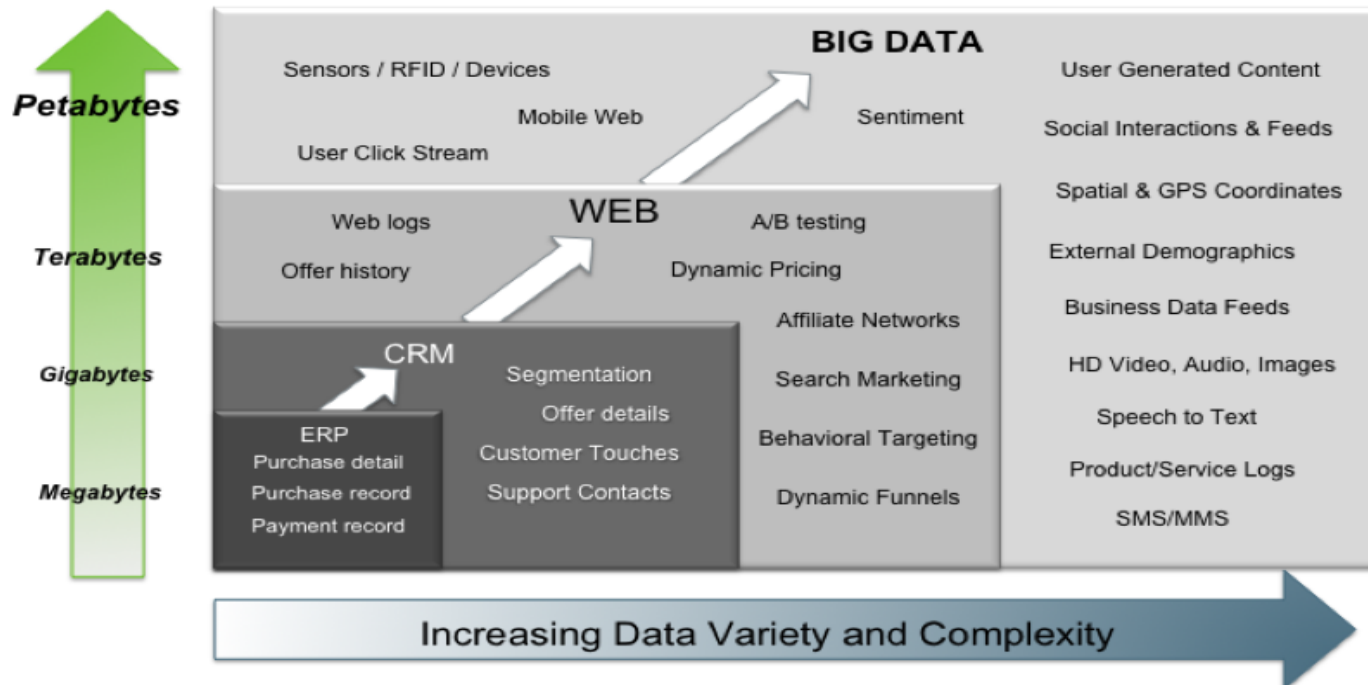
Nick Name	Q1(25)	Q2(5)	Q3(10)	Q4(20)	Q5(10)	Q6(10)	Q7(10)	Q8(10)	Total	Attempts
a_creative_nickname	25	5	10	20	10	10	10	10	100	1



Project 1 Motivation: Big Data

- What is Big Data?
 - It is high volume, high velocity, and/or high variety information assets.

Big Data = Transactions + Interactions + Observations

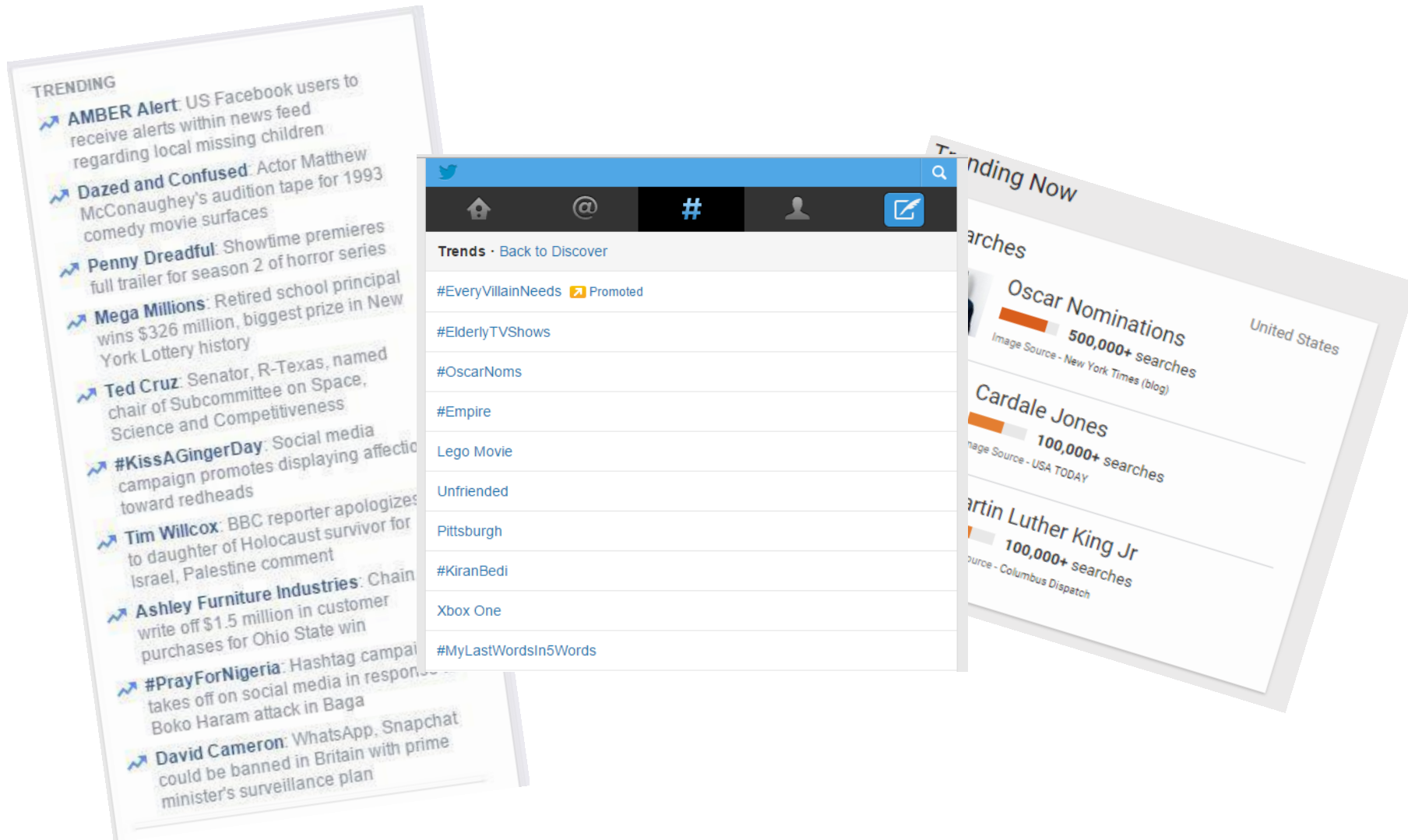


Source: Contents of above graphic created in partnership with Teradata, Inc.

Use Cases: Big Data Analysis

- Online retailers are analyzing consumer spending habits to learn trends and offer personalized marketing campaigns and offers to individual customers.
- Companies such as Time Warner, Comcast etc. are using big data to track media consumption habits of their subscribers and trends to provide value-added information to advertisers and customers.

Trending Topics are Everywhere!



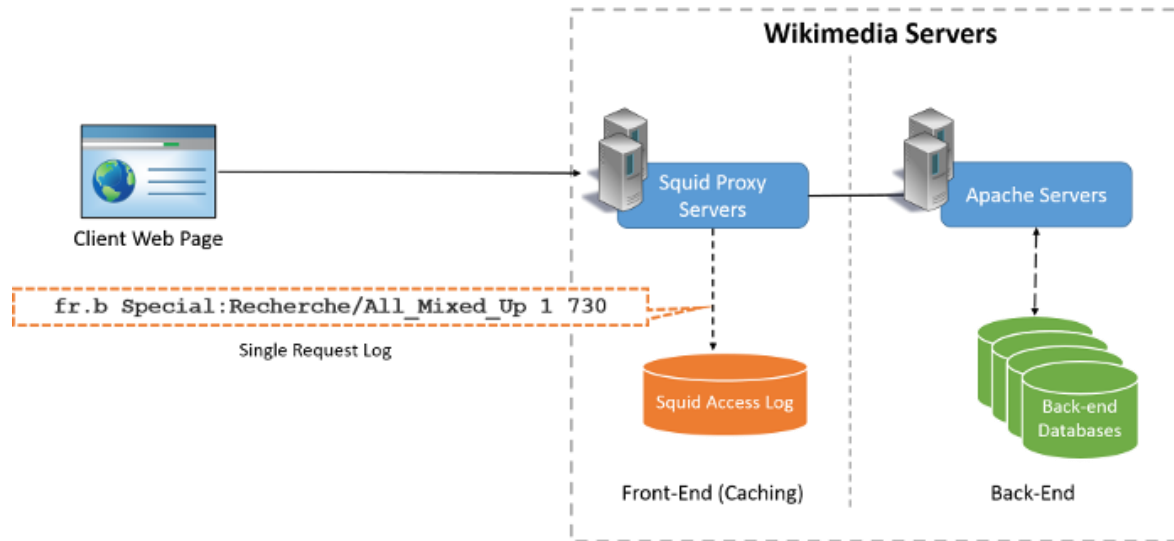


But Why
Trending
Topics?

Why Trending Topics?

- Identify trends and viral content
- Maximize ad placement opportunities
- Search Engine Optimization (SEO)
- And more....

Project 1



- Identify Trending Topics on Wikipedia
- Project 1.1: (This Week)
 - Find trends from a single hour of data.
- Project 1.2: (Next Week)
 - Find trends for an entire month.

Project 1.1

- Data set
 - [Wikimedia raw page views data](#)
- One file generated every hour.
- Explore the data set to learn the format

```
<project name> <page title> <number of accesses> <total  
data returned in bytes>
```
- You need to parse and filter the data
 - As described in P1.1 writeup on <http://theproject.zone>
- Launch the AMI and get started
- Sort the data and save the output to a file
- Answer 8 Questions and Submit for Autograding

P1.1 Autograder

- P1.1 code submissions are auto-graded
- Scores will be made available on <http://theproject.zone> after submission.
- We will grade all the code (both auto and manually)
 - Be sure to make your code readable
 - Preface each function with a header that describes what it does
 - Use whitespace well
 - Indent when using loops or conditional statements
 - Use descriptive variable names
 - For more detail, please refer to www.cs.cmu.edu/~213/codeStyle.html
 - If your code is not well documented and is not readable, we will deduct points
 - Documentation shows us that you know what your code does!
 - The idea is also NOT to comment every line of code

Project 1 – Budgets and Penalties

- Configure **t1.micro** EC2 instances in **US-East-1** (N. Virginia)
- Tag all resources with **Key: Project** and **Value: 1.1**
 - No tags → **10%** grade penalty
- Budget
 - For P1.1, each student's budget is \$5
 - Exceeding \$5 → **10%** project penalty
 - Exceeding \$10 → **100%** project penalty
- Plagiarism → the lowest penalty is **200%** & potential dismissal
 - Other students, previous students, Internet (e.g. Stackoverflow)
 - Do not work on code together
 - This is about learning
 - Penalty for cheating is SEVERE – don't do it!

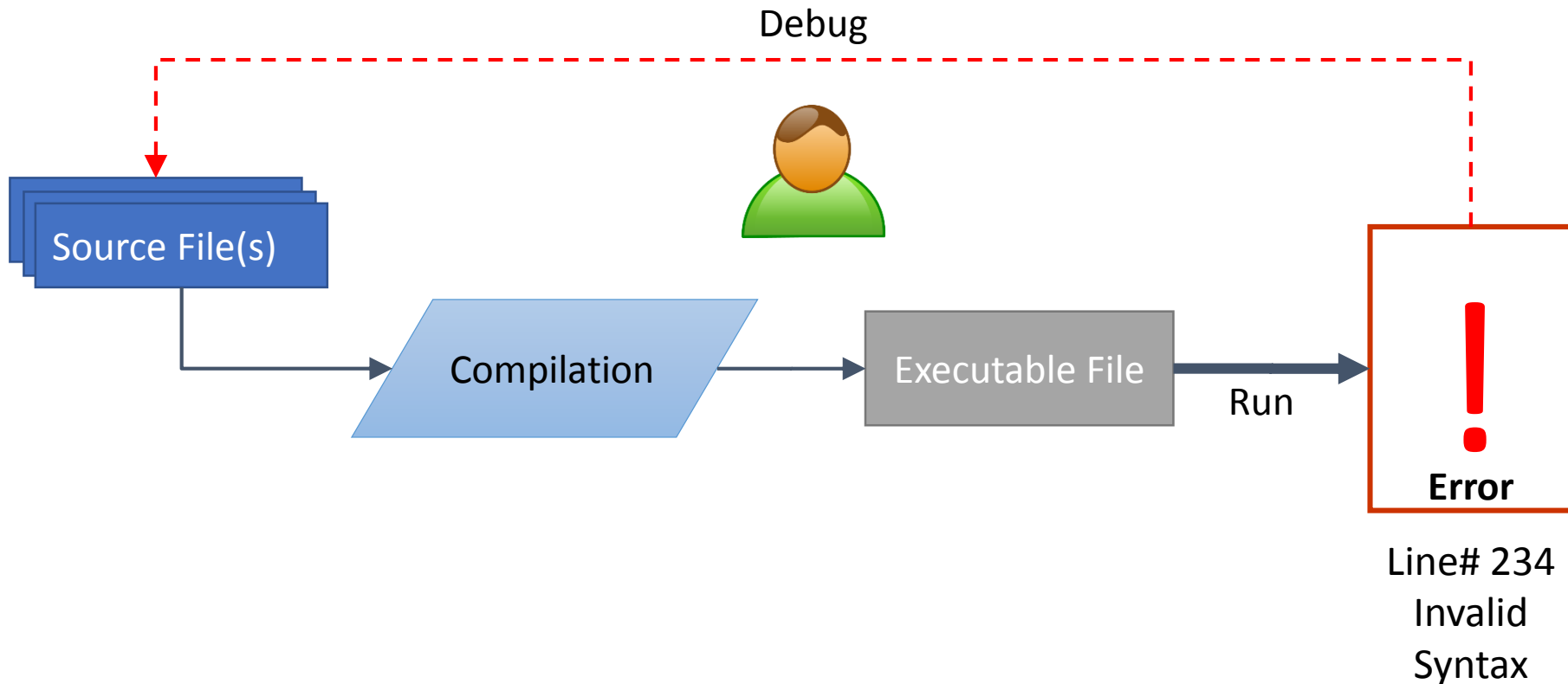
How to Work on a Budget

- P1.1 Budget → \$5
- You are only allowed to use t1.micro
 - \$0.02 per hour (on demand)
 - Total time you have: 250 hours of t1.micro
- Other Costs to consider:
 - EBS is \$0.1 per GB/month
 - t1.micro has a default of 8 GB EBS attached.
 - Data transfer costs (minimal)
- **Note:** Free Tier does not apply to any of the linked account.

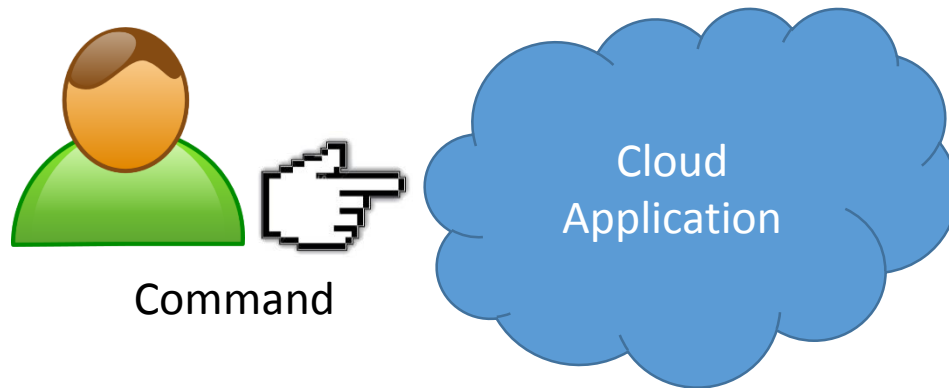
Debugging Tips

Typical Programming Workflow:

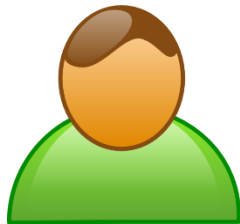
- For most courses:



In the Cloud



In the Cloud



In the Cloud



Y U NO WORK?!?

How do I even
begin to fix this?



In the Cloud



Y U NO WORK?!?

How do I even
begin to fix this?

AWS
Instances

Load
Balancers

!
Error

Databases

Front-End
Services

Suggested Error Debugging Workflow

What information can I get about the error?

- Read Error messages, Look through Logs, other information

How can I isolate the source of the problem?

- What component seems to have the problem?

What remedial action can I take?

- The error messages and other information should have clues.
- Configuration changes, command parameters

Am I Still Stuck?

- Google, Piazza, TA Office Hours (In that order!)

Demo

- Quick Tour of AWS
 - EC2
 - On-Demand and Spot Instances
 - Billing and Monitoring Costs
- Launch Course AMI on EC2
 - Configuration and Tagging
- P1.1 - Sequential Analysis
 - One Simple Task
- Auto-grader for P1.1
 - How to make a submission

Questions?

- Reminder: Office hours are posted on Piazza.

Upcoming Deadlines

- Quiz 1: Introduction to Cloud Computing
 - Checkpoint Available Now
 - **Due 01/23/15 11:59 PM EST**
- Project 1: Introduction to Big Data Analysis
 - Sequential Analysis
 - **Due 01/25/15 11:59 PM EST**