

CS 188/219

Scalable Internet Services

Andrew Mutz
October 15, 2015



Today's Agenda

Motivation

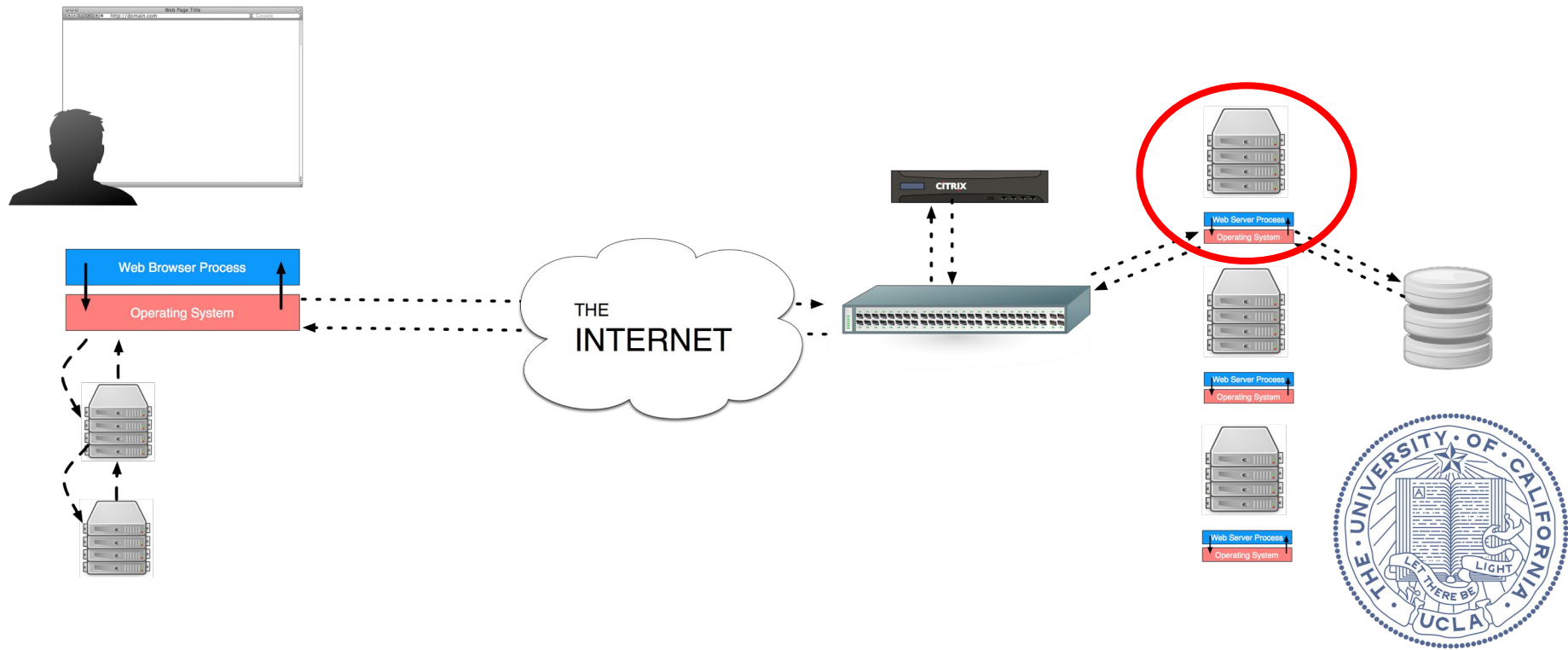
Server-side Caching

Deploying on AWS

For Next Time



Server-side Caching



Motivation

After today you should understand

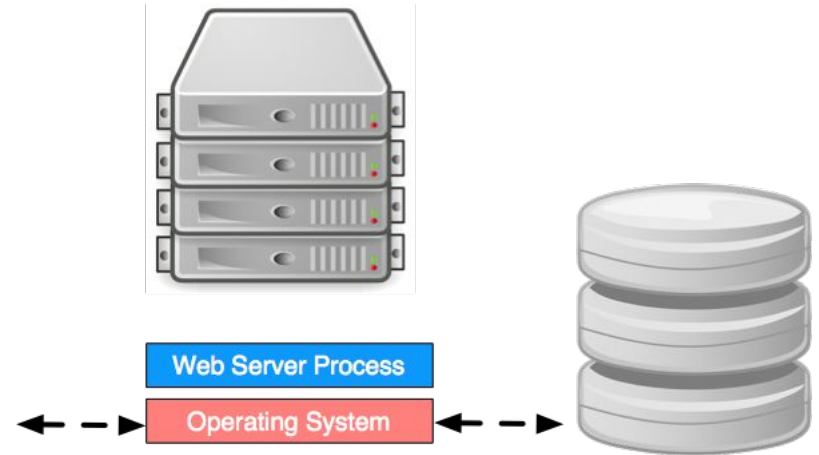
- Why server side caching exists
- What options you have when using server side caching
- How to use this in your projects
- How to deploy on AWS using CloudFormation



Server-side Caching

We have a web server process that is repeatedly responding to HTTP requests from a variety of clients.

Responding to each request requires computation and I/O to be performed, and this can be expensive.

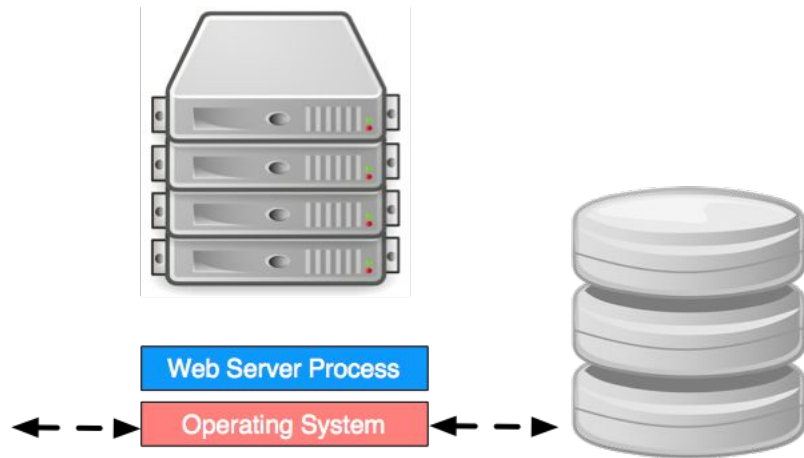


Server-side Caching

In practice, there is a great deal of similarity between responses.

In the last lecture (HTTP Caching) we looked at optimizing scenarios where repeated responses are identical.

In this lecture we will look at optimizing scenarios where repeated responses are not identical, but are similar.

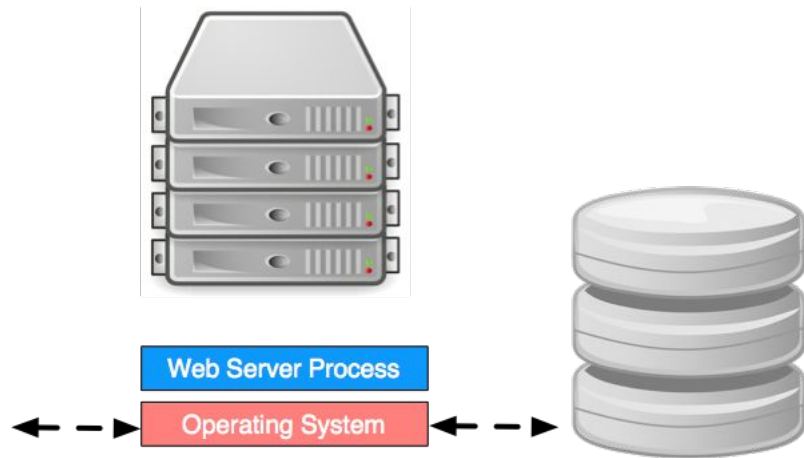


Server-side Caching

There are many parts of a response that are similar.

There are many steps to creating a response that are repeated.

What can you think of?



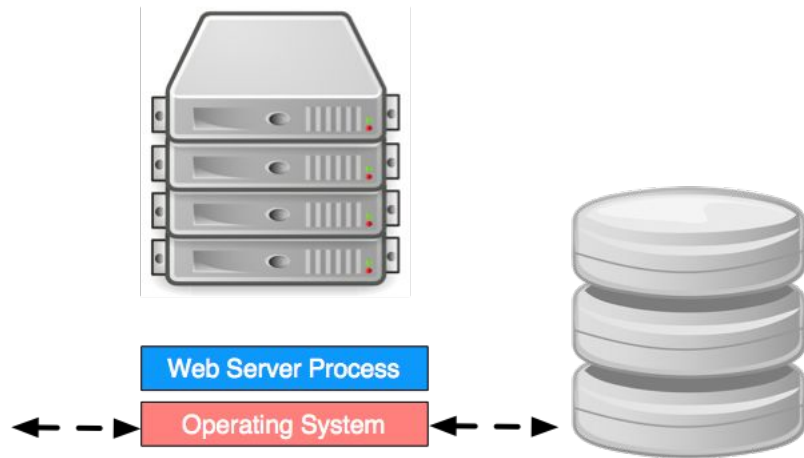
Server-side Caching

There are many parts of a response that are similar.

There are many steps to creating a response that are repeated.

What can you think of?

- View Fragments
- Rarely modified ORM objects
- Any summarized data that is expensive to compute



Server-side Caching

View Fragments: Similarity between pages

```
<html><head><meta name="referrer" content="origin"><link rel="stylesheet" type="text/css" href="news.css?Dc4WHhDIHIntIL20h45C">

<link rel="shortcut icon" href="favicon.ico">

<link rel="alternate" type="application/rss+xml" title="RSS" href="rss">

<script type="text/javascript">

    function hide(id) { var el = document.getElementById(id); if (el) { el.style.visibility = 'hidden'; } }

    function vote(node) { var v = node.id.split(/_/); var item = v[1]; hide('up_' + item); hide('down_' + item); var ping = new Image(); ping.src = node.href;
return false;}

</script><title>Hacker News</title></head><body><center><table id="hnmain" op="news" border="0" cellpadding="0" cellspacing="0" width="85%" bgcolor="#f6f6ef">

<tr><td bgcolor="#fff600"><table border="0" cellpadding="0" cellspacing="0" width="100%" style="padding:2px"><tr><td style="width:18px;padding-right:4px"><a href="http://www.ycombinator.com"></a></td>

<td style="line-height:12pt; height:10px;"><span class="pagetop">

        <b><a href="news">Hacker News</a></b><a href="newest">new</a> | <a href="newcomments">comments</a> | <a href="show">show</a> | <a href="ask">ask</a> | <a href="jobs">jobs</a> | <a href="submit">submit</a></span></td><td style="text-align:right;padding-right:4px;"><span class="pagetop">

        <a href="login?goto=news">login</a></span></td></tr></table></td></tr><tr style="height:10px"></tr><tr><td><table border="0" cellpadding="0" cellspacing="0">

<tr><td align="right" valign="top" class="title"><span class="rank">1.</span></td>

<td><center><a id="up_9403571" href="vote?for=9403571&dir=up&goto=news"><div class="votearrow" title="upvote"></div></a></center></td><td class="title">
<span class="deadmark"></span><a href="http://fossdroid.com/">Fossdroid.com: Free and open source Android applications</a><span class="sitebit comhead"> (fossdroid.com)
</span></td></tr><tr><td colspan="2"></td><td class="subtext">

    <span class="score" id="score_9403571">164 points</span> by <a href="user?id=SnaKeZ">SnaKeZ</a> <a href="item?id=9403571">4 hours ago</a>

    | <a href="item?id=9403571">38 comments</a></td></tr>

<tr class="spacer" style="height:5px"></tr><tr><tr>.....
```



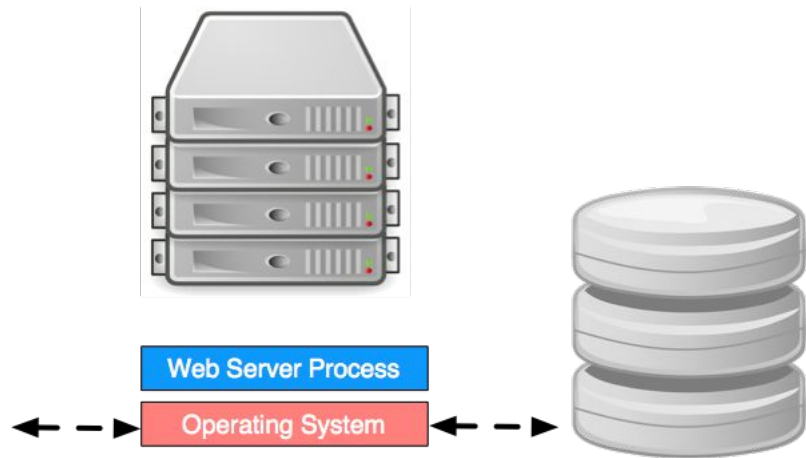
Server-side Caching

Rarely modified ORM objects?

- User permissions
- Configuration options
- Any database-backed data that changes rarely

Summarized data that is difficult to compute

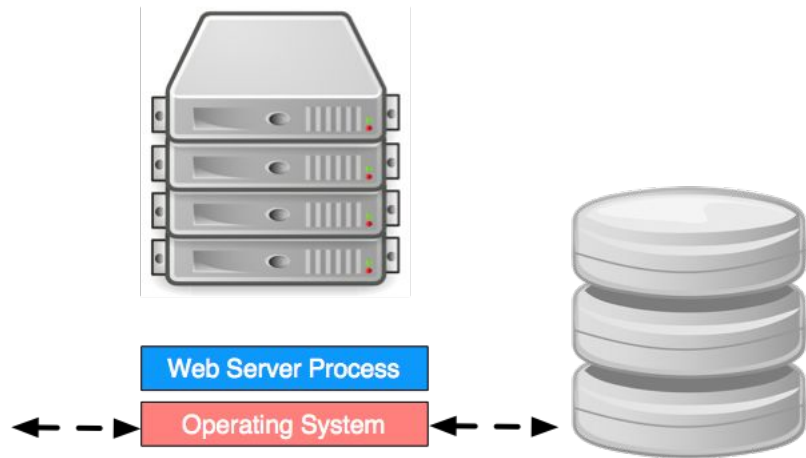
- Any particularly heavyweight SQL query
- Example: Total account balance on Mint.com



Server-side Caching

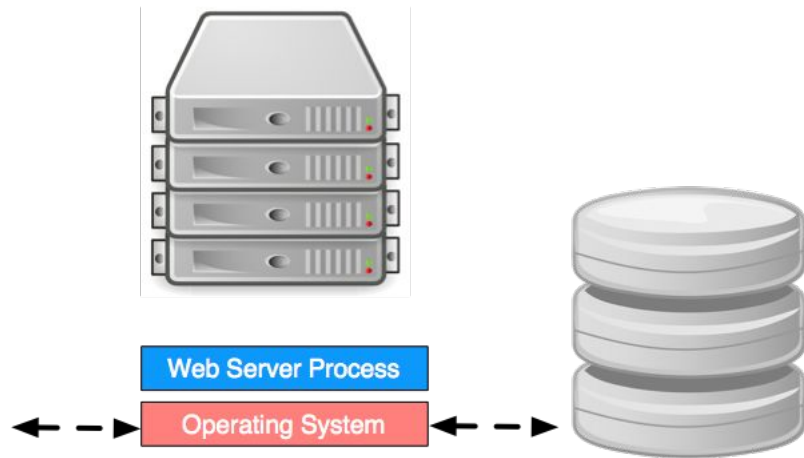
Some of these things are expensive to materialize

- View fragments are produced by extensive string manipulation.
 - Ruby optimizes humans over CPU
- The database can be a bottleneck in our current architecture
- Some SQL queries are necessarily heavyweight



Server-side Caching

So if we want to keep previously computed results around between requests, how should we do it? Where should we put it?

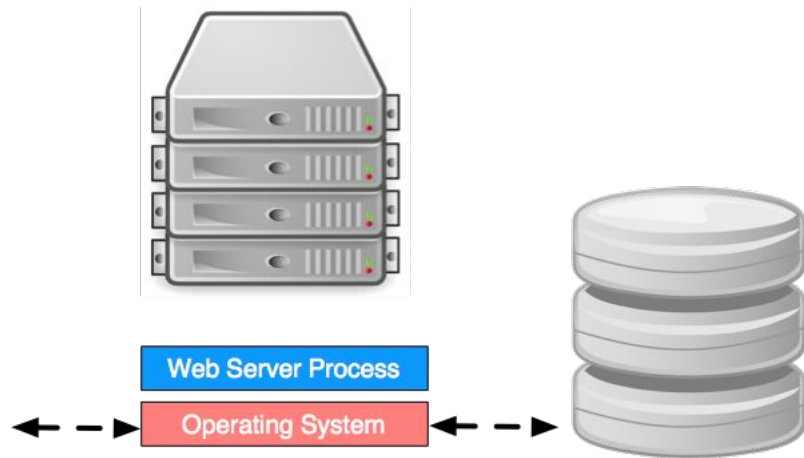


Server-side Caching

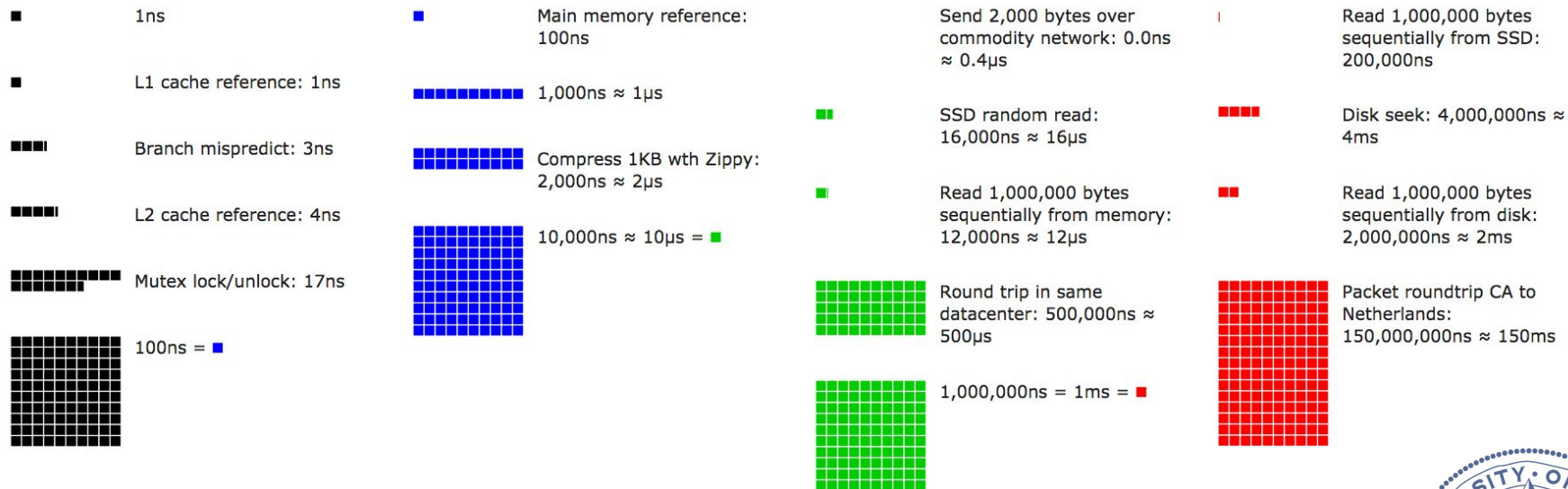
So if we want to keep previously computed results around between requests, how should we do it? Where should we put it?

- Just keep it in memory between requests?
- Store it on the filesystem?
- Store it in memory on another machine?

All of these are reasonable options. Lets look into each in more depth.



Server-side Caching



Server-side Caching

What can we conclude from these numbers?

- Storing in memory and reading later is fast:
 - Random reads from memory will be $0.1\mu\text{s}$, reading 1MB will be $12\mu\text{s}$
- Storing on disk is slow *without SSD*:
 - Disk seek is $4000\mu\text{s}$, subsequent sequential read of 1MB is $2000\mu\text{s}$
- Storing on disk *with SSD* is much more reasonable:
 - Random read is $16\mu\text{s}$, sequential read of 1MB will be $200\mu\text{s}$
- Storing on another machine is reasonable:
 - Round trip within datacenter is $500\mu\text{s}$.



Server-side Caching

Summary

- In memory: tens of μs
- On SSD: hundreds of μs
- On Disk: thousands of μs
- And if it's on a remote machine, add hundreds of μs .

Conclusion

- Always use SSD
- Memory > local SSD > Remote?



Server-side Caching

It's not that simple. Why?



Server-side Caching

It's not that simple.

What effect on the cache hit rate does each of these designs have?

- In memory: Cache per process
- On local SSD: Cache per machine
- On (single) remote machine: Cache per cluster



Server-side Caching

Conclusion:

- **In memory:** highest performance, lowest hit rate
- **On SSD:** lower performance, higher hit rate
- **On remote cache server:** lowest performance, highest hit rate

There is no silver bullet. How will each of these affect system performance:

- Number of processes per machine?
- Concurrency model of Application Server: threads vs. processes?
- Number of machines per cluster?



Memcached

Memcached is a commonly used implementation of a remote cache server

- Keeps a cache in memory
- Accepts TCP connections and returns lookup requests
- Distributed key-value store
 - Keys can be up to 250 bytes, values can be up to 1MB
 - Can scale horizontally
- When it runs out of space, it uses a simple LRU mechanism to make more space
- Lightweight features, everything is constant time.
- Originally developed at LiveJournal

Another commonly used tool is Redis



Rails Caching

The good news for your projects is that Rails has great support for server-side caching.

Rails emphasizes three types of caching:

- HTTP caching
- Fragment caching
- Low level caching

We covered HTTP caching in the last lecture, so today we will talk about fragment caching and low-level caching



Rails Caching

By default, caching is disabled in development and test, and enabled in production

- If you want to use it in development mode, add this to your environment:

```
config.action_controller.perform_caching = true
```

Rails can be configured to store cached data in a few different places:

- In memory
- Local file system
- Remote in-memory store



Rails Caching

ActiveSupport::Cache::MemoryStore

- Cached data is stored in memory, in the same address space as the ruby process and is retained between requests.
- Defaults to 32 megs, but is configurable.

ActiveSupport::Cache::FileStore

- Cached data is stored on the local file system.
- Can configure the location of the storage in Rails environment:
 - `config.cache_store = :file_store, "/path/to/cache/"`



Rails Caching

ActiveSupport::Cache::MemcacheStore

- Cached data is stored in memory on another machine.
- Can configure the location of the server in Rails environment:
 - `config.cache_store = :mem_cache_store, "cache-1.example.com"`



Rails Caching - Fragment Caching

Fragment caching caches a portion of a rendered view for reuse on future requests.

Let's take a look at the demo app...



Rails Caching - Fragment Caching

Demo App		
Submissions		
Title	Url	Community
A id ea officia.	http://williamson.net/rebekah	Est distinctio qui aut quis doloribus dignissimos sit sed. 20 comments
Earum sequi veniam libero quibusdam est corporis omnis voluptatem.	http://dare.biz/adolph_pouros	Est distinctio qui aut quis doloribus dignissimos sit sed. 20 comments
Deleniti vel expedita sint voluptate repellat.	http://schamberger.org/raphaelle	Est distinctio qui aut quis doloribus dignissimos sit sed. 20 comments
Neque eum sint cum magni.	http://vonruedenborer.name/edythe	Est distinctio qui aut quis doloribus dignissimos sit sed. 20 comments

We can cache each line of this markup.

Regardless of anything else that changes on the page, we can rerender this if it stays fresh



Rails Caching - Fragment Caching

```
<tbody>
  <% @submissions.each do |submission| %>
    <tr>
      <td><%= link_to(submission.title, submission.url) %></td>
      <td><%= submission.url %></td>
      <td><%= submission.community.name %></td>
      <td><%= link_to "#{submission.comments.size} comments", submission, class: 'btn btn-
primary btn-xs' %></td>
    </tr>
  <% end %>
</tbody>
</table>
```



Rails Caching - Fragment Caching

```
<tbody>
  <% @submissions.each do |submission| %>
    <% cache(cache_key_for_submission_row(submission)) do %>
      <tr>
        <td><%= link_to(submission.title, submission.url) %></td>
        <td><%= submission.url %></td>
        <td><%= submission.community.name %></td>
        <td><%= link_to "#{submission.comments.size} comments", submission, class: 'btn btn-
primary btn-xs' %></td>
      </tr>
    <% end %>
  <% end %>
</tbody>
</table>
```



Rails Caching - Fragment Caching

How should we choose a cache key?

```
module SubmissionsHelper
  def cache_key_for_submission_row(submission)
    "submission-#{submission.id}"
  end
end
```

What are the weaknesses with the above approach?



Rails Caching - Fragment Caching

How should we choose a cache key?

```
module SubmissionsHelper
  def cache_key_for_submission_row(submission)
    "submission-#{submission.id}"
  end
end
```

What are the weaknesses with the above approach?

- Invalidation will be annoying: clear out the cache on possible action causing staleness?



Rails Caching - Fragment Caching

Instead, let's make the key change whenever the data gets stale.

```
module SubmissionsHelper
  def cache_key_for_submission_row(submission)
    "submission-#{submission.id}-#{submission.updated_at}-#{submission.comments.count}"
  end
end
```

There is no action needed to invalidate the cache: the cache key changes.



Rails Caching - Fragment Caching

Demo App		
Submissions		
Title	Url	Community
A id ea officia.	http://williamson.net/rebekah	Est distinctio qui aut quis doloribus dignissimos sit sed. 20 comments
Earum sequi veniam libero quibusdam est corporis omnis voluptatem.	http://dare.biz/adolph_pouros	Est distinctio qui aut quis doloribus dignissimos sit sed. 20 comments
Deleniti vel expedita sint voluptate repellat.	http://schamberger.org/raphaele	Est distinctio qui aut quis doloribus dignissimos sit sed. 20 comments
Neque eum sint cum magni.	http://vonruedenborer.name/edythe	Est distinctio qui aut quis doloribus dignissimos sit sed. 20 comments

If we step back and look at this page, we can observe that the whole table is expensive to compute and stays fresh for awhile.



Rails Caching - Fragment Caching

```
<h3>Submissions</h3>
<table class="table">
  <thead>
    <tr>
      <th>Title</th>
      <th>Url</th>
      <th>Community</th>
      <th colspan="3"></th>
    </tr>
  </thead>

  <tbody>
    <% @submissions.each do |submission| %>
      <% cache(cache_key_for_submission_row(submission)) do %>
        <tr>
          <td><%= link_to(submission.title, submission.url) %></td>
          <td><%= submission.url %></td>
          <td><%= submission.community.name %></td>
          <td><%= link_to "#{submission.comments.size} comments", submission, class: 'btn btn-primary btn-xs' %></td>
        </tr>
      <% end %>
    <% end %>
  </tbody>
</table>

<br>
<%= link_to 'New Submission', new_submission_path, class: 'btn btn-primary' %>
<%= link_to 'New Community', new_community_path, class: 'btn btn-primary' %>
```



Rails Caching - Fragment Caching

```
<% cache(cache_key_for_submission_table) do %>
  <h3>Submissions</h3>
  <table class="table">
    <thead>
      <tr>
        <th>Title</th>
        <th>Url</th>
        <th>Community</th>
        <th colspan="3"></th>
      </tr>
    </thead>

    <tbody>
      <% @submissions.each do |submission| %>
        <% cache(cache_key_for_submission_row(submission)) do %>
          <tr>
            <td><%= link_to(submission.title, submission.url) %></td>
            <td><%= submission.url %></td>
            <td><%= submission.community.name %></td>
            <td><%= link_to "#{submission.comments.size} comments", submission, class: 'btn btn-primary btn-xs' %></td>
          </tr>
        <% end %>
      <% end %>
    </tbody>
  </table>

  <br>
  <%= link_to 'New Submission', new_submission_path, class: 'btn btn-primary' %>
  <%= link_to 'New Community', new_community_path, class: 'btn btn-primary' %>
<% end %>
```



Rails Caching - Fragment Caching

```
module SubmissionsHelper
  def cache_key_for_submission_row(submission)
    "submission-#{submission.id}-#{submission.updated_at}-#{submission.comments.count}"
  end
  def cache_key_for_submission_table
    "submission-table-#{Submission.maximum(:updated_at)}-#{Comment.maximum(:updated_at)}"
  end
end
```

This technique of nesting cache fragments is known as “Russian Doll” caching.



Rails Caching - Low-level Caching

You can use the same mechanisms to cache anything:

```
class Product < ActiveRecord::Base
  def competing_price
    Rails.cache.fetch("#{cache_key}/competing_price", expires_in: 12.hours) do
      Competitor::API.find_price(id)
    end
  end
end
```



Rails Caching - Low-level Caching

Let's compare the demo app's performance!

For these tests I will compare the performance of the branch master, with the branch “server_side_caching”, which implements the caching shown in the previous slides.

Master intentionally includes no optimizations to the way we interact with the database.

We will use the default (memory) caching.



Rails Caching

We'll use an M3-medium instance with the usual workload.

When we test we will have 12 phases, of 60 seconds each:

Phase	1	2	3	4	5	6	7	8	9	10	11	12
Users/sec	1	1.5	2	4	6	10	16	20	25	35	45	55

Deployed using nginx & passenger.



Rails Caching

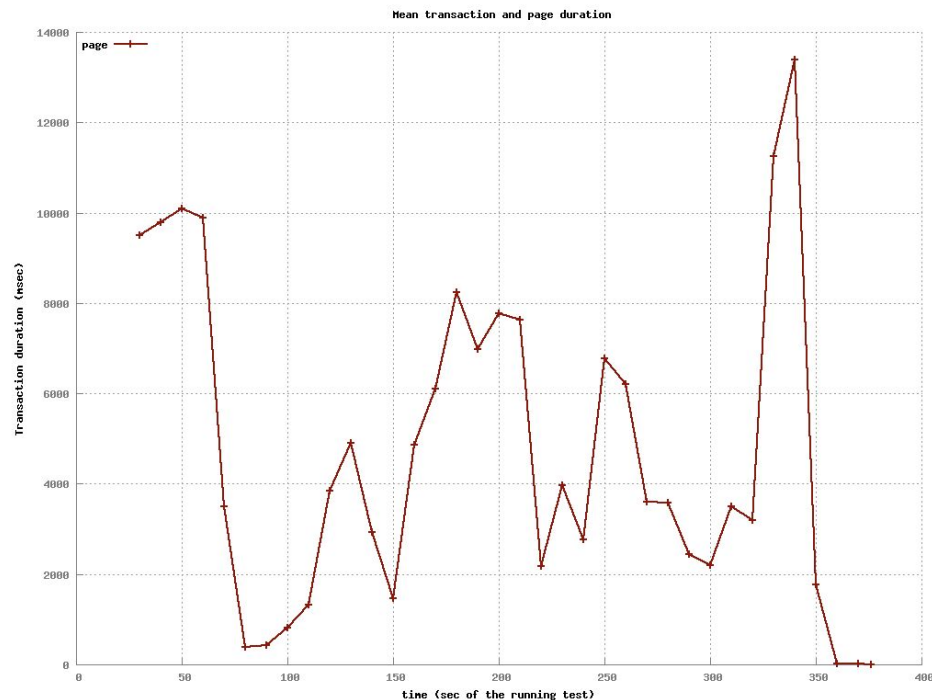
We will use the same testing script from before...

1. Going to the homepage
2. Waiting for up to 2 seconds
3. Requesting a form to create a new community
4. Waiting for up to 2 seconds
5. Submitting the new community
6. Requesting a form to create a new link submission
7. Waiting for up to 2 seconds
8. Submitting the new link
9. Waiting for up to 2 seconds
10. Delete the link
11. Waiting for up to 2 seconds
12. Delete the community



Rails Caching

Performance on master:

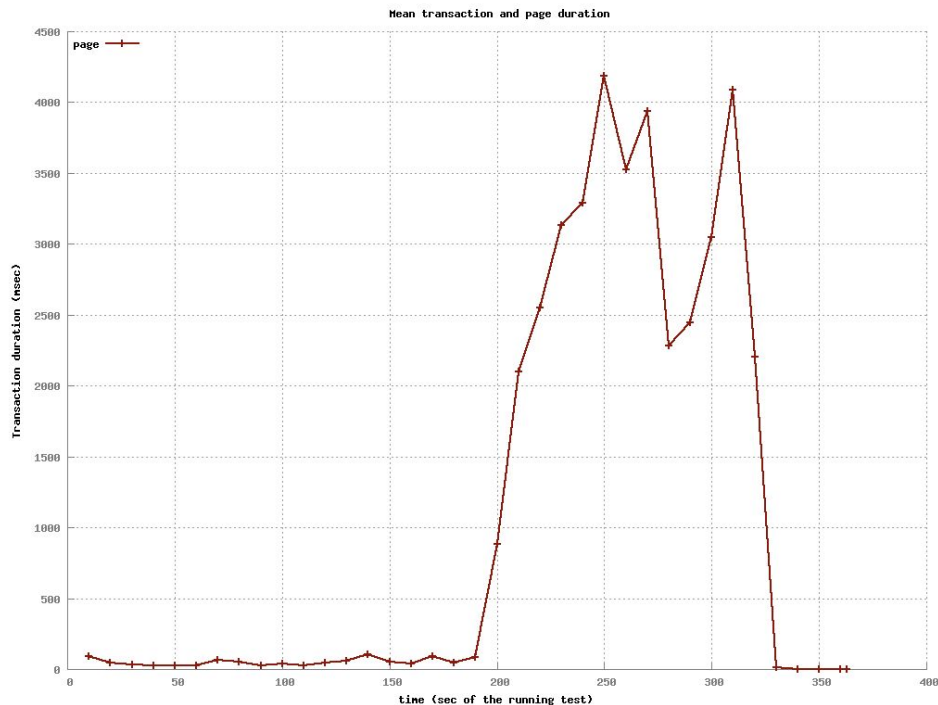


From the start, the application can't handle a single user arriving each second



Rails Caching

Performance on server_side_caching:



With the server-side caching implemented, the server can handle up to two new users a second easily.



AWS Instructions

Now that you've all got a blank Rails app (or more) pushed to Github, it's time to learn how to deploy to AWS.

We will be using CloudFormation, and here are step-by-step instructions for deploying.

But first, some warnings and rules.



AWS Instructions

These are scarce resources

- This is free time on Amazon's infrastructure, and it's not unlimited.
- Unless you have a specific reason to do otherwise, always use micro instances.
 - Example of a good reason: testing vertical scaling.



AWS Instructions

These are scarce resources

- Our AWS budget has a fixed limit.
- Whenever you are done with an instance, shut it down.
- Never keep important data on the instance, because it can go down at any time.
 - SCP important data back to your laptop.
- I will periodically run a script that terminates all instances that have been up longer than 8 hours.



AWS Instructions

Treat these credentials as secrets.

- Do not check into any publicly accessible repository
- There are automated scripts that actively seek out AWS credentials
 - Why?



AWS Instructions

Today you will receive an email with AWS credentials. These include:

- A.txt file
 - Username: what you use to log in to the AWS web interface
 - Password: the password you use to log in to the AWS web interface
 - AccessKey: not needed yet, but very sensitive
 - SecretKey: not needed yet, but very sensitive
- A.pem file
 - Contains a private key that you will use to ssh into the instances that you launch.



AWS Instructions

Go to <https://scalable-internet-services.signin.aws.amazon.com/console>



Account:

User Name:

Password:

☐ I have an MFA Token [\(more info\)](#)

Sign In

[Sign-in using root account credentials](#)

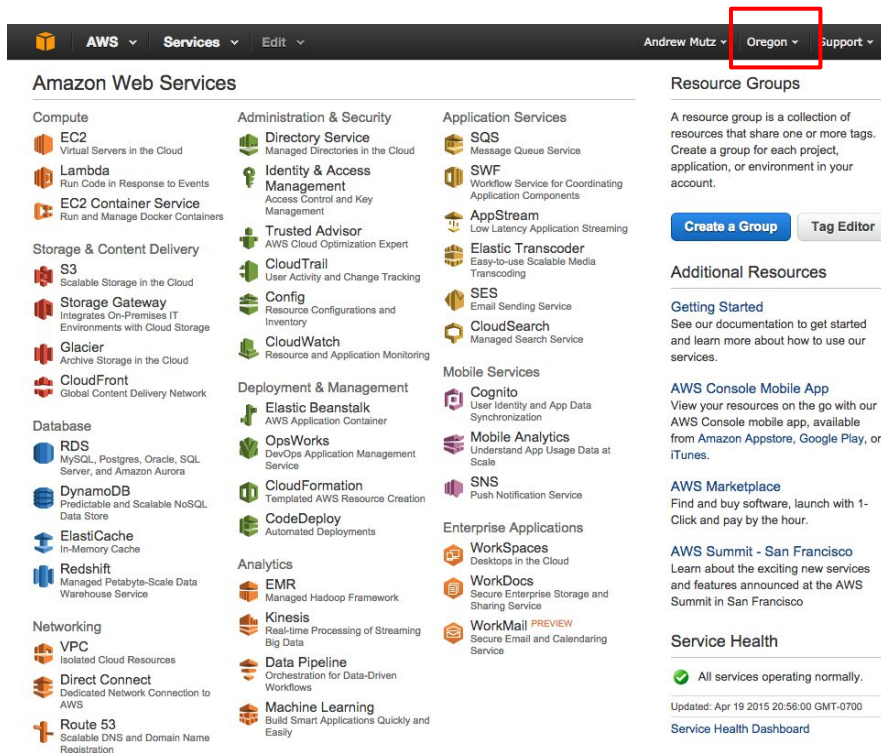
English

[Terms of Use](#) [Privacy Policy](#) © 1996-2015, Amazon Web Services, Inc. or its affiliates.

Login with Username and Password provided in txt file in email.



AWS Instructions



The screenshot shows the AWS Management Console interface. At the top, the navigation bar includes the AWS logo, 'AWS' dropdown, 'Services' dropdown, 'Edit' dropdown, the user name 'Andrew Mutz', the region 'Oregon' (highlighted with a red box), and 'Support' dropdown. Below the navigation bar, the 'Amazon Web Services' section lists various services categorized into Compute, Storage & Content Delivery, Database, Networking, Administration & Security, Application Services, Mobile Services, Enterprise Applications, Analytics, and Machine Learning. The 'Resource Groups' section on the right provides a description of resource groups and includes a 'Create a Group' button and a 'Tag Editor' link. The 'Additional Resources' section includes links for 'Getting Started', 'AWS Console Mobile App', 'AWS Marketplace', 'AWS Summit - San Francisco', and 'Service Health'.

Amazon Web Services

- Compute**
 - EC2**
Virtual Servers in the Cloud
 - Lambda**
Run Code in Response to Events
 - EC2 Container Service**
Run and Manage Docker Containers
- Storage & Content Delivery**
 - S3**
Scalable Storage in the Cloud
 - Storage Gateway**
Integrates On-Premises IT Environments with Cloud Storage
 - Glacier**
Archive Storage in the Cloud
 - CloudFront**
Global Content Delivery Network
- Database**
 - RDS**
MySQL, PostgreSQL, Oracle, SQL Server, and Amazon Aurora
 - DynamoDB**
Predictable and Scalable NoSQL Data Store
 - ElastiCache**
In-Memory Cache
 - Redshift**
Managed Petabyte-Scale Data Warehouse Service
- Networking**
 - VPC**
Isolated Cloud Resources
 - Direct Connect**
Dedicated Network Connection to AWS
 - Route 53**
Scalable DNS and Domain Name Registration
- Administration & Security**
 - Directory Service**
Managed Directories in the Cloud
 - Identity & Access Management**
Access Control and Key Management
 - Trusted Advisor**
AWS Cloud Optimization Expert
 - CloudTrail**
User Activity and Change Tracking
 - Config**
Resource Configurations and Inventory
 - CloudWatch**
Resource and Application Monitoring
- Application Services**
 - SQS**
Message Queue Service
 - SWF**
Workflow Service for Coordinating Application Components
 - AppStream**
Low Latency Application Streaming
 - Elastic Transcoder**
Easy-to-use Scalable Media Transcoding
 - SES**
Email Sending Service
 - CloudSearch**
Managed Search Service
- Mobile Services**
 - Cognito**
User Identity and App Data Synchronization
 - Mobile Analytics**
Understand App Usage Data at Scale
 - SNS**
Push Notification Service
- Enterprise Applications**
 - WorkSpaces**
Desktops in the Cloud
 - WorkDocs**
Secure Enterprise Storage and Sharing Service
 - WorkMail** PREVIEW
Secure Email and Calendaring Service
- Analytics**
 - EMR**
Managed Hadoop Framework
 - Kinesis**
Real-time Processing of Streaming Big Data
 - Data Pipeline**
Orchestration for Data-Driven Workflows
 - Machine Learning**
Build Smart Applications Quickly and Easily

Resource Groups

A resource group is a collection of resources that share one or more tags. Create a group for each project, application, or environment in your account.

[Create a Group](#) [Tag Editor](#)

Additional Resources

Getting Started
See our documentation to get started and learn more about how to use our services.

AWS Console Mobile App
View your resources on the go with our AWS Console mobile app, available from Amazon Appstore, Google Play, or iTunes.

AWS Marketplace
Find and buy software, launch with 1-Click and pay by the hour.

AWS Summit - San Francisco
Learn about the exciting new services and features announced at the AWS Summit in San Francisco

Service Health

✓ All services operating normally.

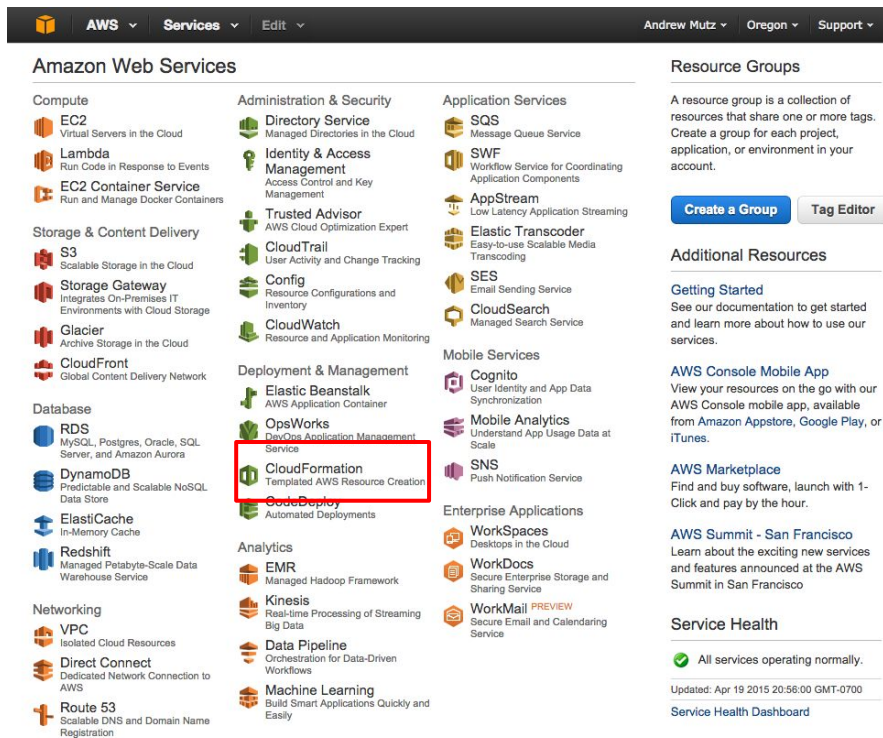
Updated: Apr 19 2015 20:56:00 GMT-0700

[Service Health Dashboard](#)

Make sure your region is set to
“US-West Oregon”



AWS Instructions



The screenshot shows the AWS Management Console interface. At the top, there's a navigation bar with the AWS logo, 'Services' dropdown, 'Edit' dropdown, and user information 'Andrew Mutz', 'Oregon', and 'Support'. Below the navigation bar, the 'Amazon Web Services' section is displayed, organized into columns. The services listed include:

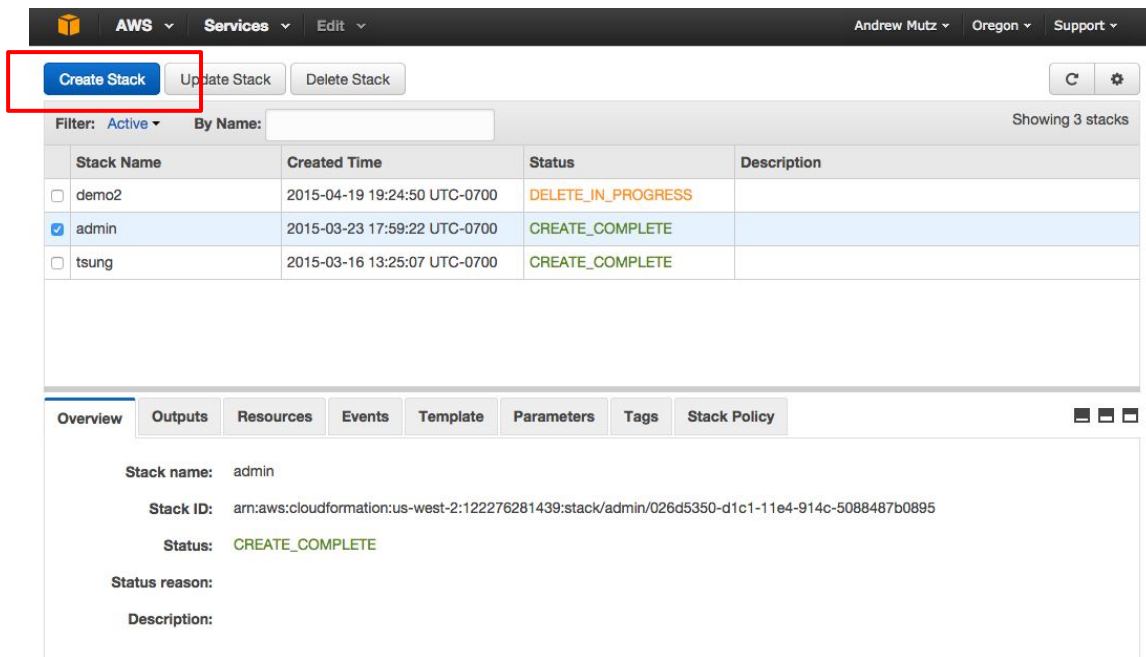
- Compute:** EC2 (Virtual Servers in the Cloud), Lambda (Run Code in Response to Events), EC2 Container Service (Run and Manage Docker Containers).
- Storage & Content Delivery:** S3 (Scalable Storage in the Cloud), Storage Gateway (Integrates On-Premises IT Environments with Cloud Storage), Glacier (Archive Storage in the Cloud), CloudFront (Global Content Delivery Network).
- Database:** RDS (MySQL, PostgreSQL, Oracle, SQL Server, and Amazon Aurora), DynamoDB (Predictable and Scalable NoSQL Data Store), ElastiCache (In-Memory Cache), Redshift (Managed Petabyte-Scale Data Warehouse Service).
- Networking:** VPC (Isolated Cloud Resources), Direct Connect (Dedicated Network Connection to AWS), Route 53 (Scalable DNS and Domain Name Registration).
- Administration & Security:** Directory Service (Managed Directories in the Cloud), Identity & Access Management (Access Control and Key Management), Trusted Advisor (AWS Cloud Optimization Expert), CloudTrail (User Activity and Change Tracking), Config (Resource Configurations and Inventory), CloudWatch (Resource and Application Monitoring).
- Deployment & Management:** Elastic Beanstalk (AWS Application Container), OpsWorks (DevOps Application Management Service), **CloudFormation** (Templated AWS Resource Creation) - highlighted with a red box, CodeDeploy (Automated Deployments).
- Analytics:** EMR (Managed Hadoop Framework), Kinesis (Real-time Processing of Streaming Big Data), Data Pipeline (Orchestration for Data-Driven Workflows), Machine Learning (Build Smart Applications Quickly and Easily).
- Application Services:** SQS (Message Queue Service), SWF (Workflow Service for Coordinating Application Components), AppStream (Low Latency Application Streaming), Elastic Transcoder (Easy-to-use Scalable Media Transcoding), SES (Email Sending Service), CloudSearch (Managed Search Service).
- Mobile Services:** Cognito (User Identity and App Data Synchronization), Mobile Analytics (Understand App Usage Data at Scale), SNS (Push Notification Service).
- Enterprise Applications:** WorkSpaces (Desktops in the Cloud), WorkDocs (Secure Enterprise Storage and Sharing Service), WorkMail (Secure Email and Calendaring Service).

On the right side of the console, there's a 'Resource Groups' section with a description: 'A resource group is a collection of resources that share one or more tags. Create a group for each project, application, or environment in your account.' Below this are buttons for 'Create a Group' and 'Tag Editor'. Further down is the 'Additional Resources' section, which includes links for 'Getting Started', 'AWS Console Mobile App', 'AWS Marketplace', 'AWS Summit - San Francisco', and 'Service Health'. The 'Service Health' section shows a green checkmark and the text 'All services operating normally.' with a timestamp 'Updated: Apr 19 2015 20:56:00 GMT-0700' and a link to the 'Service Health Dashboard'.

We will be deploying with
cloudformation, so click on that



AWS Instructions



The screenshot shows the AWS CloudFormation console interface. At the top, there's a navigation bar with 'AWS', 'Services', and 'Edit' dropdowns, along with user information 'Andrew Mutz', region 'Oregon', and 'Support'. Below this, there are three buttons: 'Create Stack' (highlighted with a red box), 'Update Stack', and 'Delete Stack'. To the right of these buttons are icons for refresh and settings. Below the buttons, there's a filter section with 'Filter: Active' and a search box 'By Name:'. A table lists three stacks:

Stack Name	Created Time	Status	Description
<input type="checkbox"/> demo2	2015-04-19 19:24:50 UTC-0700	DELETE_IN_PROGRESS	
<input checked="" type="checkbox"/> admin	2015-03-23 17:59:22 UTC-0700	CREATE_COMPLETE	
<input type="checkbox"/> tsung	2015-03-16 13:25:07 UTC-0700	CREATE_COMPLETE	


Below the table, there's a section for the selected stack 'admin'. It includes tabs for 'Overview', 'Outputs', 'Resources', 'Events', 'Template', 'Parameters', 'Tags', and 'Stack Policy'. The 'Overview' tab is active, showing details for the stack 'admin':

- Stack name: admin
- Stack ID: arn:aws:cloudformation:us-west-2:122276281439:stack/admin/026d5350-d1c1-11e4-914c-5088487b0895
- Status: CREATE_COMPLETE
- Status reason:
- Description:

Click on “Create Stack”



AWS Instructions

 AWS

Services

Edit

demo @ scalable-Internet-servi... Oregon Support

Select Template

Specify Details

Options

Review

Select Template

Select the template that describes the stack that you want to create. A stack is a group of related resources that you manage as a single unit.

Design a template

Use AWS CloudFormation Designer to create or modify an existing template. [Learn more.](#)

Design template

Choose a template

A template is a JSON-formatted text file that describes your stack's resources and their properties. [Learn more.](#)

☒ Select a sample template

☐ Upload a template to Amazon S3

☐ Specify an Amazon S3 template URL

Choose Fileno file selected

Cancel

Next

We want to specify an Amazon S3 template URL.



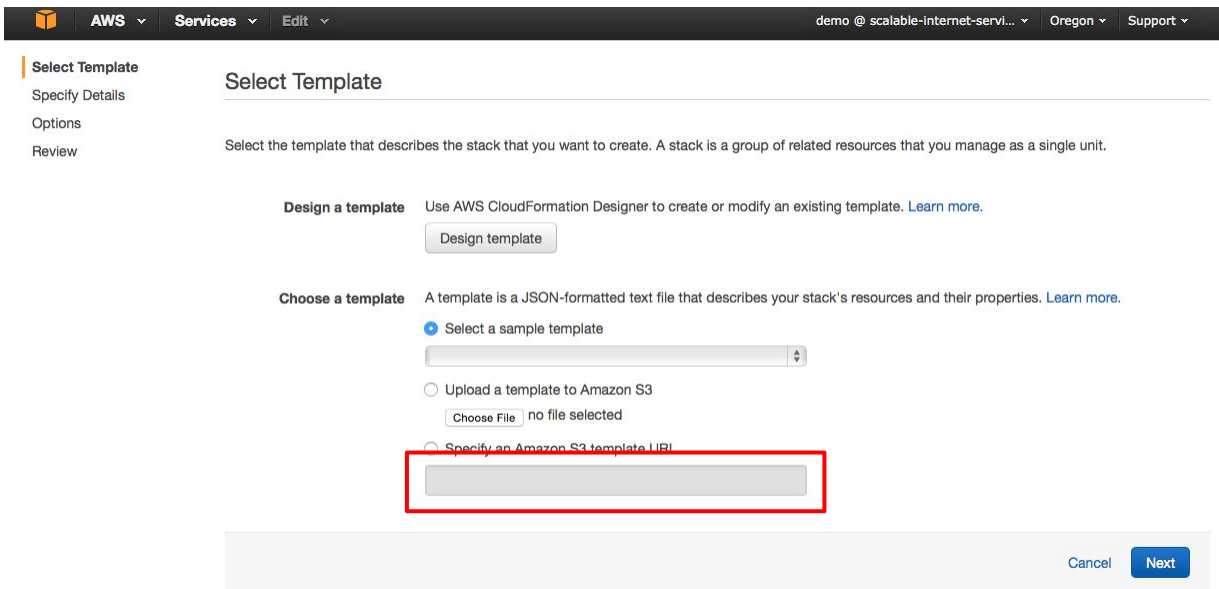
Go to
<https://github.com/scalableinternetservices/utls>
and choose your template.

Today just choose the Webrick template.

Copy the link.



AWS Instructions



AWS ▾ **Services** ▾ **Edit** ▾ demo @ scalable-Internet-servi... ▾ Oregon ▾ Support ▾

Select Template

Specify Details
Options
Review

Select the template that describes the stack that you want to create. A stack is a group of related resources that you manage as a single unit.

Design a template Use AWS CloudFormation Designer to create or modify an existing template. [Learn more.](#)

Design template

Choose a template A template is a JSON-formatted text file that describes your stack's resources and their properties. [Learn more.](#)

☒ Select a sample template

☐ Upload a template to Amazon S3

Choose File no file selected

☐ Specify an Amazon S3 template URL

Cancel Next

Paste the link to the template in the field labeled “Specify an Amazon S3 Template URL”



AWS Instructions

AWS Services Edit demo @ scalable-internet-servi... Oregon Support

Select Template
Specify Details
Options
Review

Specify Details

Specify a stack name and parameter values. You can use or change the default parameter values, which are defined in the AWS CloudFormation template. [Learn more.](#)

Stack name

Parameters

AppInstanceType The AppServer instance type.

Branch The git branch to deploy.

TeamName Your team name.

▲ Parameters with AllowedValues must not be empty

▲ TeamName has an invalid value.

Cancel Previous Next

The stack name must begin with your team name.

For example: “demo-test”



AWS Instructions

Select Template

Specify Details

Options

Review

demo @ scalable-internet-servi... Oregon Support

Specify a stack name and parameter values. You can use or change the default parameter values, which are defined in the AWS CloudFormation template. [Learn more.](#)

Stack name

Parameters

AppInstanceType The AppServer instance type.

Branch The git branch to deploy.

TeamName Your team name.

▲ Parameters with AllowedValues must not be empty

▲ TeamName has an invalid value.

Cancel Previous Next

Choose your fields.

Select your TeamName from the dropdown.

The teamname is how the template knows where to get your code.



AWS Instructions

Options

Tags

You can specify tags (key-value pairs) for resources in your stack. You can add up to 10 unique key-value pairs for each stack. [Learn more.](#)

	Key (127 characters maximum)	Value (255 characters maximum)	
1	<input type="text"/>	<input type="text"/>	<input data-bbox="807 442 830 455" type="button" value="+"/>

Advanced

You can set additional options for your stack, like notification options and a stack policy. [Learn more.](#)

Notification options

☒ No notification

☐ New Amazon SNS topic

Topic

Email

☐ Existing Amazon SNS topic

Timeout ⓘ

Minutes

Rollback on failure ⓘ

☒ Yes

☐ No

Stack policy ⓘ

☐ Enter policy

☒ Upload policy file

Choose File No file chosen

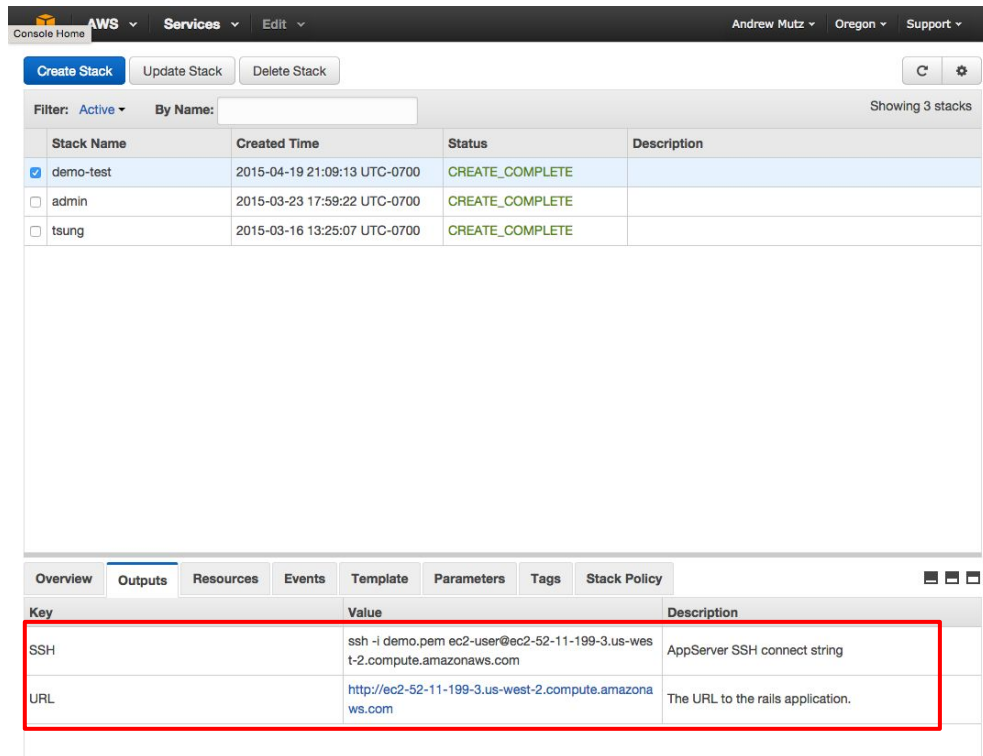
[Learn more](#)

On the next screen just choose “Next”

If you are having problems deploying, you can disable rollback.



AWS Instructions



The screenshot shows the AWS Management Console interface. At the top, there's a navigation bar with 'AWS', 'Services', and 'Edit' dropdowns, along with the user 'Andrew Mutz' and region 'Oregon'. Below this, there are buttons for 'Create Stack', 'Update Stack', and 'Delete Stack'. A filter section shows 'Filter: Active' and 'By Name:'. A table lists three stacks: 'demo-test' (created 2015-04-19, status 'CREATE_COMPLETE'), 'admin' (created 2015-03-23, status 'CREATE_COMPLETE'), and 'tsung' (created 2015-03-16, status 'CREATE_COMPLETE'). The 'demo-test' stack is selected. Below the table, there are tabs for 'Overview', 'Outputs', 'Resources', 'Events', 'Template', 'Parameters', 'Tags', and 'Stack Policy'. The 'Outputs' tab is active, showing a table with two outputs: 'SSH' and 'URL'. The 'SSH' output has a value 'ssh -i demo.pem ec2-user@ec2-52-11-199-3.us-west-2.compute.amazonaws.com' and a description 'AppServer SSH connect string'. The 'URL' output has a value 'http://ec2-52-11-199-3.us-west-2.compute.amazonaws.com' and a description 'The URL to the rails application.'.

Stack Name	Created Time	Status	Description
<input checked="" type="checkbox"/> demo-test	2015-04-19 21:09:13 UTC-0700	CREATE_COMPLETE	
<input type="checkbox"/> admin	2015-03-23 17:59:22 UTC-0700	CREATE_COMPLETE	
<input type="checkbox"/> tsung	2015-03-16 13:25:07 UTC-0700	CREATE_COMPLETE	

Key	Value	Description
SSH	ssh -i demo.pem ec2-user@ec2-52-11-199-3.us-west-2.compute.amazonaws.com	AppServer SSH connect string
URL	http://ec2-52-11-199-3.us-west-2.compute.amazonaws.com	The URL to the rails application.

After creation, the outputs tab in the bottom pane will tell you how to reach your server via HTTP and SSH.

The PEM file mentioned in the SSH command was emailed to you.

SCP accepts the same -i FILE.pem argument



AWS Instructions

How do I SSH into my instance?

- `ssh -i [your pemfile here] ec2-user@ec2-something.us-west-2.compute.amazonaws.com`

How do I copy files to/from my instance?

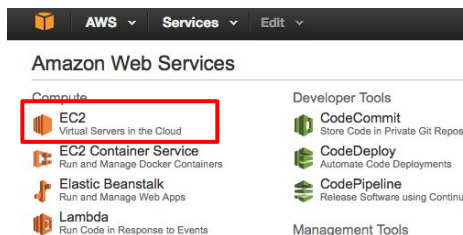
- `scp -i [your pemfile here] ec2-user@ec2-something.us-west-2.compute.amazonaws.com:fromfile tofile`



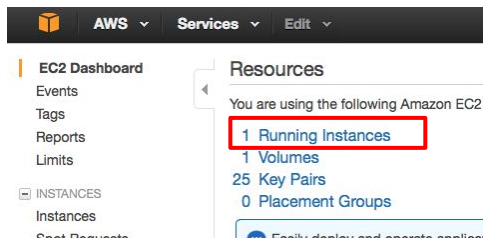
AWS Instructions

My cloudformation stack failed to deploy. How do I debug this?

1. Go to the AWS dashboard and select EC2



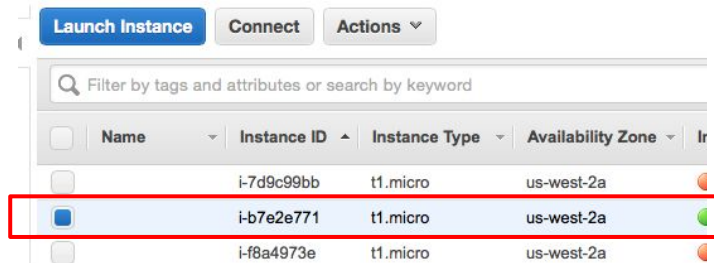
2. Click on “X Running Instances”



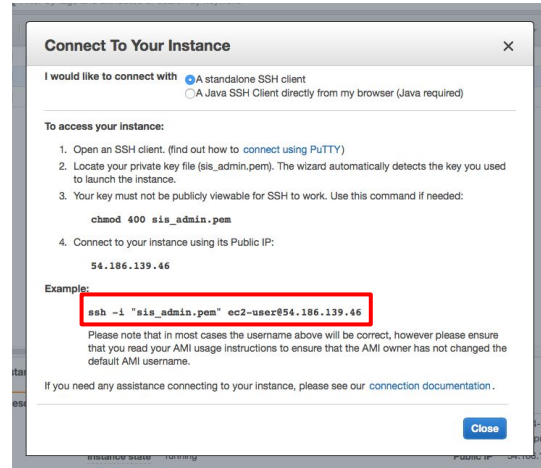
AWS Instructions

My cloudformation stack failed to deploy. How do I debug this?

3. Select the instance corresponding to your team



4. Click the connect button and copy and paste the ssh command.



AWS Instructions

My cloudformation stack failed to deploy. How do I debug this?

Once you have SSHed into your instance, you can find details of problems at:

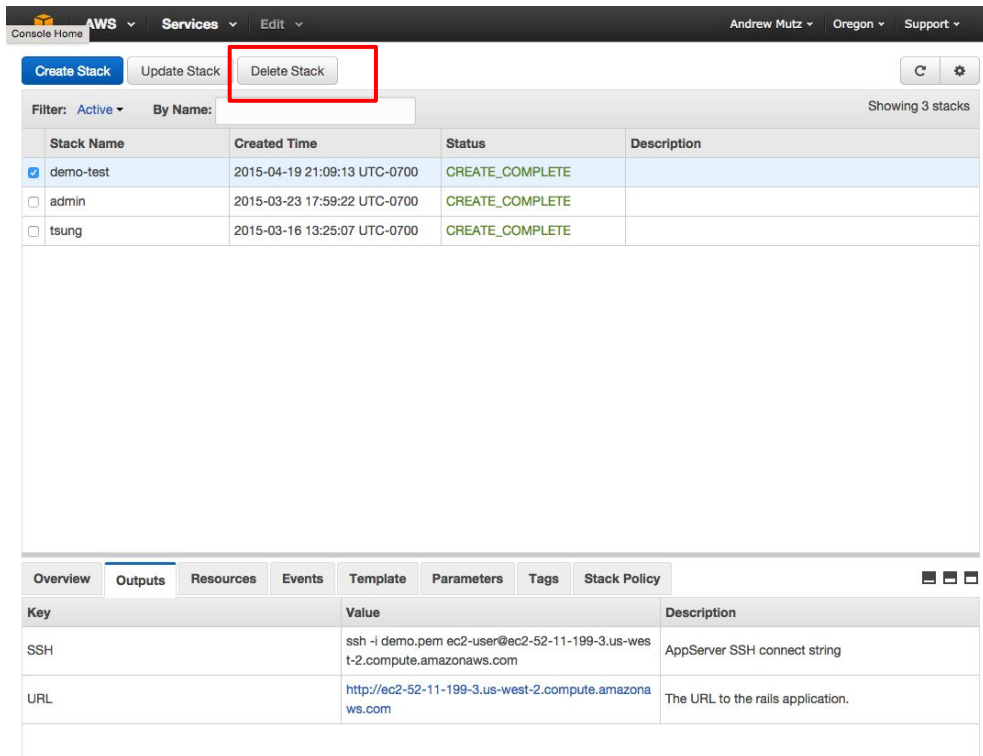
```
/var/log/cloud-init-output.log
```

You can also read the script that was executed on CF startup here:

```
/var/lib/cloud/instance/scripts/part-001
```



AWS Instructions



The screenshot shows the AWS Management Console interface. At the top, there's a navigation bar with 'AWS', 'Services', and 'Edit' dropdowns, along with user information 'Andrew Mutz', region 'Oregon', and 'Support'. Below this, there are three buttons: 'Create Stack', 'Update Stack', and 'Delete Stack'. The 'Delete Stack' button is highlighted with a red rectangular box. Below the buttons, there's a filter section with 'Filter: Active' and 'By Name:' followed by a search input field. A table lists three stacks: 'demo-test', 'admin', and 'tsung'. The 'demo-test' stack is selected with a checkbox. The table columns are 'Stack Name', 'Created Time', 'Status', and 'Description'. Below the table, there's a section for 'Outputs' with tabs for 'Overview', 'Outputs', 'Resources', 'Events', 'Template', 'Parameters', 'Tags', and 'Stack Policy'. The 'Outputs' tab is active, showing a table with columns 'Key', 'Value', and 'Description'. The table lists two outputs: 'SSH' and 'URL'.

Stack Name	Created Time	Status	Description
<input checked="" type="checkbox"/> demo-test	2015-04-19 21:09:13 UTC-0700	CREATE_COMPLETE	
<input type="checkbox"/> admin	2015-03-23 17:59:22 UTC-0700	CREATE_COMPLETE	
<input type="checkbox"/> tsung	2015-03-16 13:25:07 UTC-0700	CREATE_COMPLETE	

Key	Value	Description
SSH	ssh -i demo.pem ec2-user@ec2-52-11-199-3.us-west-2.compute.amazonaws.com	AppServer SSH connect string
URL	http://ec2-52-11-199-3.us-west-2.compute.amazonaws.com	The URL to the rails application.

When you are done with your stack, remember to delete it!

If you have any questions, please post to Piazza!



Motivation

After today you should understand

- Why server side caching exists
- What options you have when using server side caching
- How to use this in your projects
- How to deploy on AWS using CloudFormation



For Next Time...

Continue to work on sprint 1 stories. We will demo your progress at tomorrow's lab.

Try and get your app deployed on AWS. If you run into problems:

- Post on Piazza
- Come to Franz 1260 tomorrow at noon

