DRUPALCON

AMSTERDAM

2014

Turning Drupal Into a Machine for Automated Deployment

Nick Veenhof: Nick_vh
Peter Wolanin: pwolanin

Track: DevOps
https://amsterdam2014.drupal.org/session/turning-drupal-machine-automated-deployment

THE SITUATION:

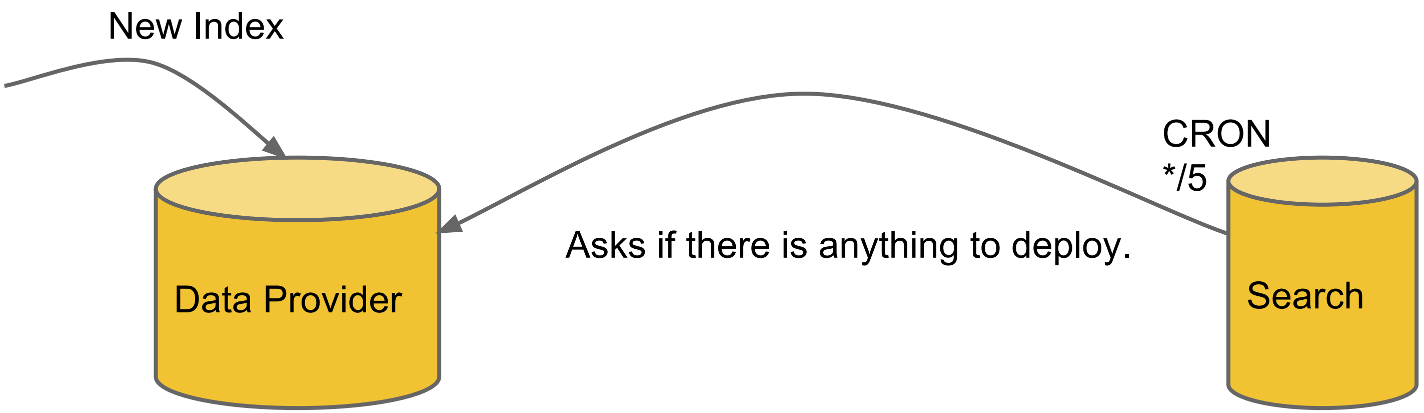
- We need to manage a growing number of servers and 4k+ individual Apache Solr Search Indexes.
  - Thousands of indexes
  - Hundreds of servers, incl. load balancers.
  - Multiple AWS Regions where customers are deployed and want low latency.

THE SITUATION (con'td):

- The system built in 2009 predated our own AWS management tools.
- It had not been built with high scale or evolving needs in mind, leading to operational stress and development friction.

Mechanics at Charles Street shops, 1958 https://www.flickr.com/photos/seattlemunicipalarchives/4089729743/ CC

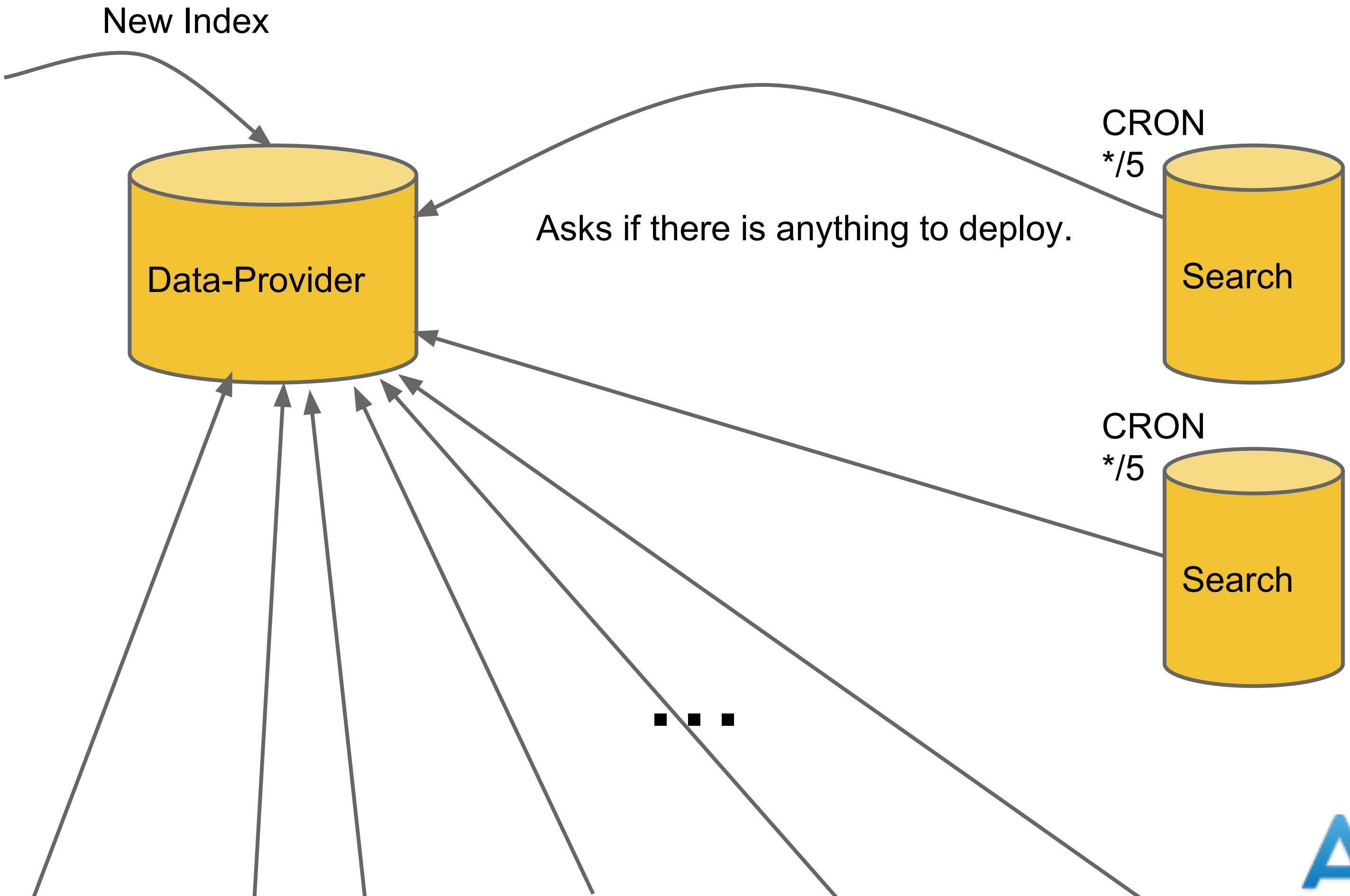
Dev+Ops Problems (Deploying)



- Poll model for index deploy/update
  - Had a 5 to 10 minute turn-around time before any action was taken
  - Hard to see errors, needed a lot of (manual) audit checks

Dev+Ops Problems (Scaling):

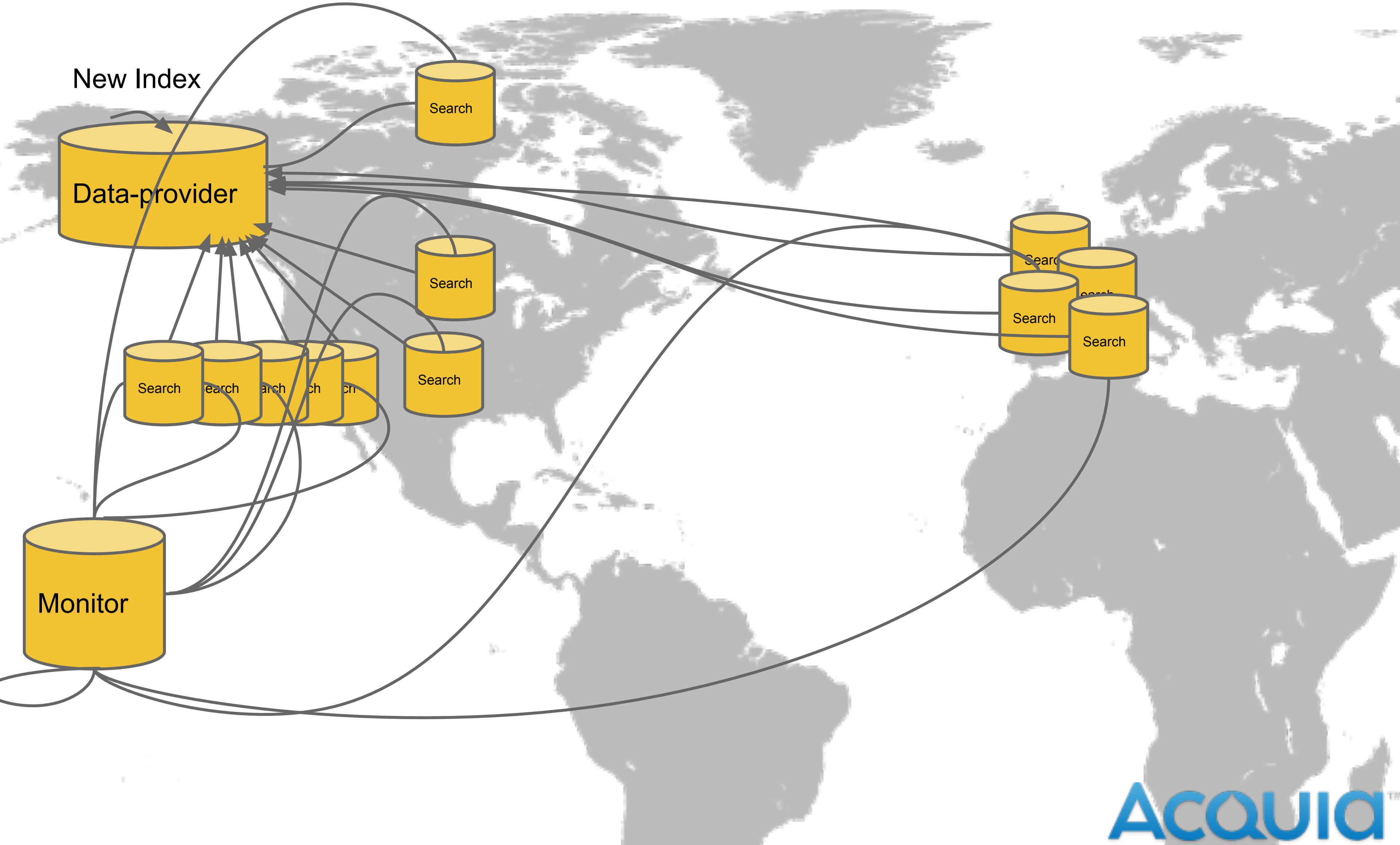
- Poll model doesn’t scale



Dev+Ops Problems (Scaling):

- Manual configuration for new or re-launched servers
  - Scaling balancing of server load takes hours of manual operations time.
  - Hard to recover when you lose a server in the "Amazon".
  - 24/7 SLA hard to meet for search.

Dev+Ops Problems (Monitoring - Distribution):



Dev+Ops Problems (Monitoring - Distribution):

- No single, canonical record of the expected state of a given index.
- Monitoring was hard.
- Instances confined to one region
- Due to hard-coded single index per customer we were not able to flexibly provision additional indexes.

https://www.flickr.com/photos/sarabbit/4549185468/

Context

Drupal Gardens

- We manage 10’s of thousands of Drupal sites for Drupal Gardens.
- acquia.com/blog/minimizing-maintenance-time-while-updating-thousands-drupal-gardens-sites
- slideshare.net/AcquiaInc/enterprise-drupal-gardens-opensaas-site-factory

Aegir

- www.aegirproject.org
- drupal.org/project/devshop

Goals for our Deployment Machine:

- Make an independent machine for managing Acquia Search based on API’s
  - Testable, "Black box", API endpoint for creating search indexes
- Build using mostly pre-existing hosting platform and tools
  - Acquia Cloud is designed for Drupal (php)
- Build using community code or contributing back as much new code as possible

Why Drupal as Deployment Machine?

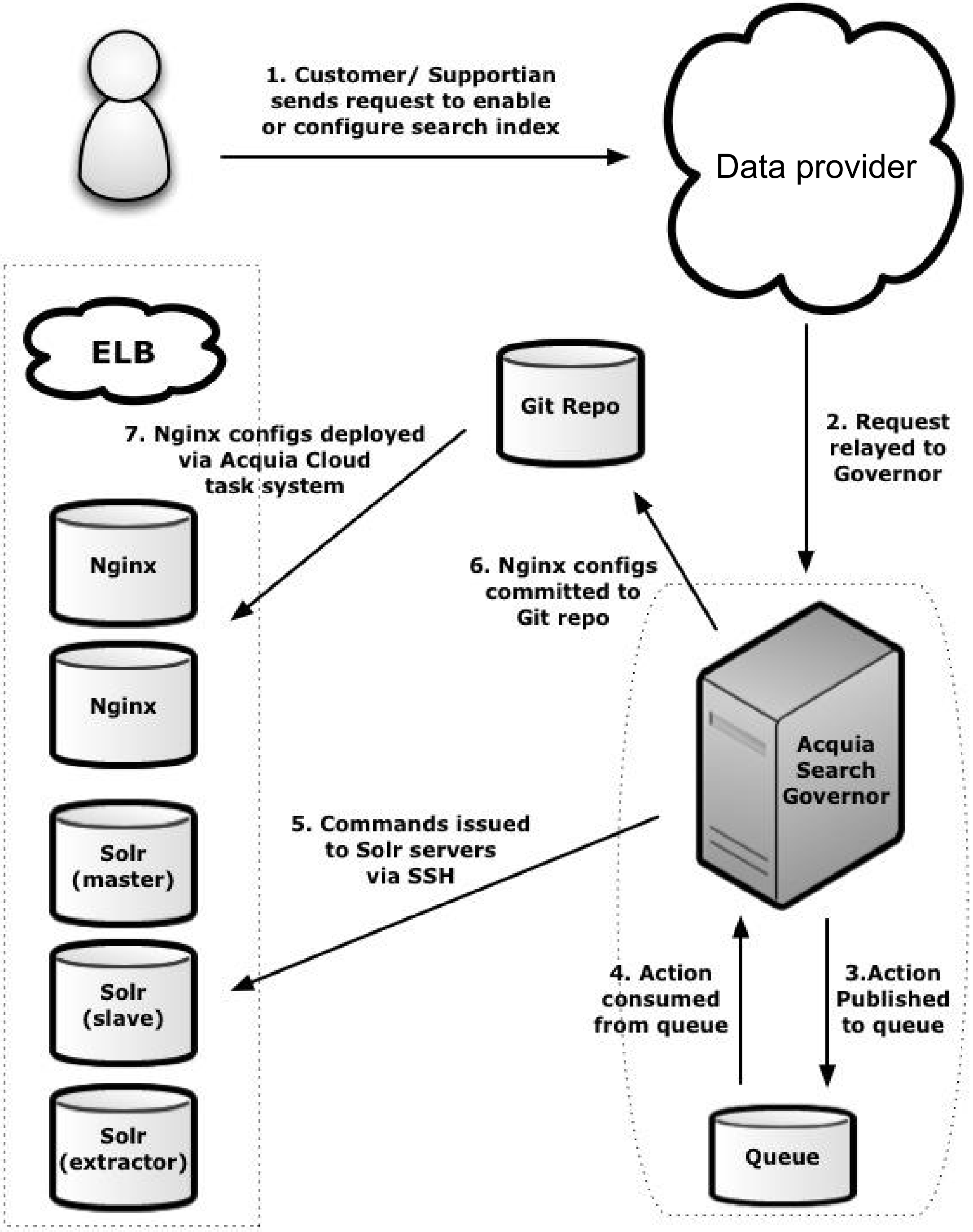
- Familiarity
- UI for manual intervention and debugging
- Data-driven model
- Synchronize the "data" with other Drupal sites
- Easy to integrate modules in it for customizing.
- Views for consuming customer statistics, and easily find the data you are looking for.

DRUPALCON
AMSTERDAM
2014

“Zaphod attacked everything in life with a mixture of extraordinary genius and naive incompetence and it was often difficult to tell which was which.”

– Douglass Adams

The Big Picture



Data provider

ELB

Git Repo

Nginx

Nginx

Solr (master)

Solr (slave)

Solr (extractor)

Acquia Search Governor

Queue

1. Customer/ Supportian sends request to enable or configure search index

2. Request relayed to Governor

3.Action Published to queue

4. Action consumed from queue

5. Commands issued to Solr servers via SSH

6. Nginx configs committed to Git repo

7. Nginx configs deployed via Acquia Cloud task system

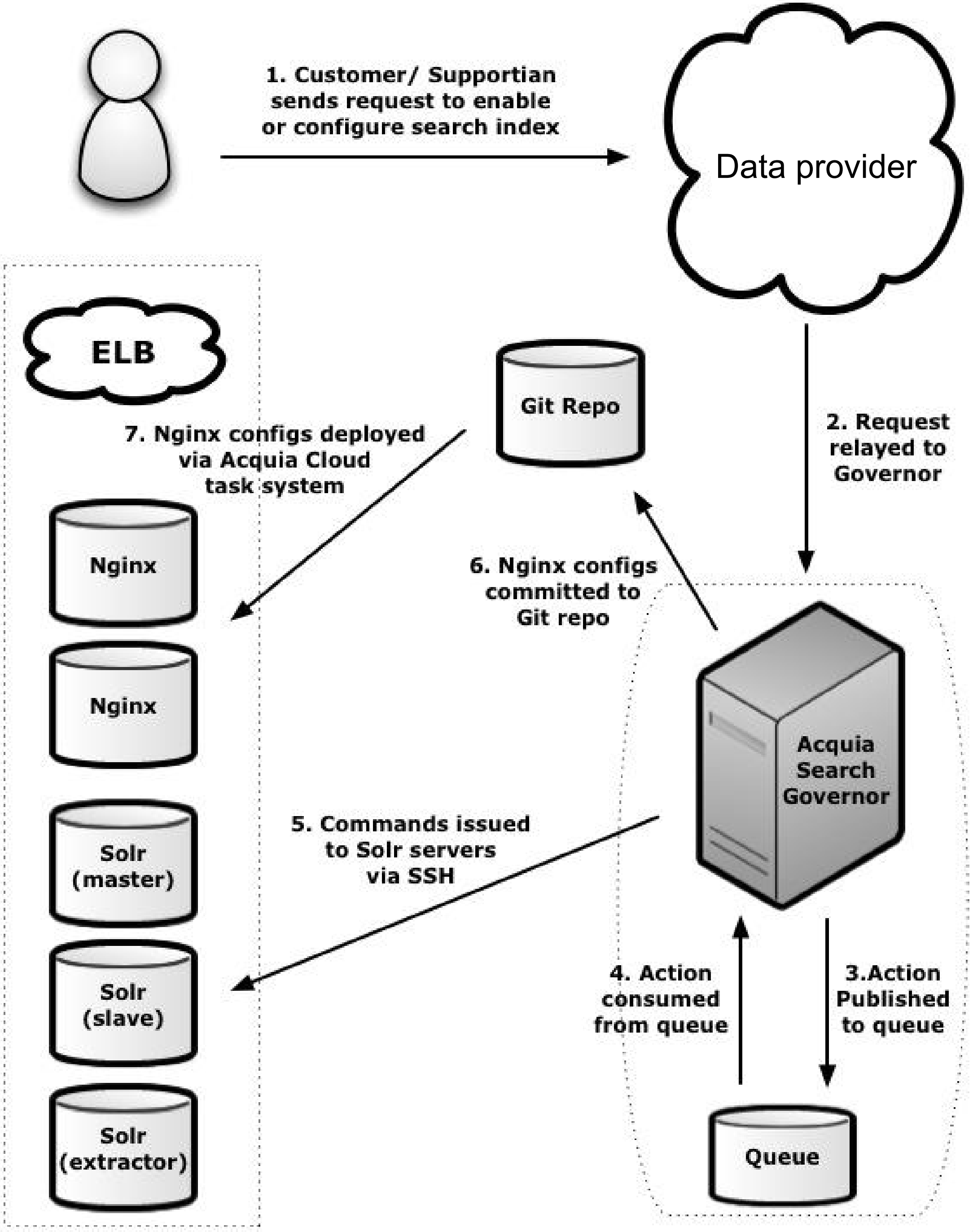
Drupal Content Types

- **Search Index**: name and authentication details
- **Server**: Any type of master, slave of extractor
- **Cluster**: used to define a subset of servers, with 1 entry point being a domain name. Could be different regions or dedicated to one customer.
- **Search Configuration Set**: schema.xml, etc

Drupal 7: The Coordinator

- Content drives server and index configuration
- Use Drupal queues and ssh to push data or operations.
- Drupal receives API calls and returns data based on content in Drupal.
- Load balancer "site" configuration from git - pushed by Drupal based on content.
- Drupal code deployed from git.
- Solr "site" code deployed from git.

The Big Picture



1. Customer/ Supportian sends request to enable or configure search index
Data provider
2. Request relayed to Governor
ELB
Git Repo
7. Nginx configs deployed via Acquia Cloud task system
Nginx
Nginx
6. Nginx configs committed to Git repo
Acquia Search Governor
5. Commands issued to Solr servers via SSH
Solr (master)
Solr (slave)
Solr (extractor)
4. Action consumed from queue
3.Action Published to queue
Queue

Existing and Enhanced Modules

- Queue Runner
- SSH Helper
- Git Wrapper
- Services
- Composer Manager
- EFQ Extra Fields

Queue Runner

Finite state machines and more

The main goal of the Queue Runner module is that it runs for x time and executes as many tasks as possible within the given timeframe, with a specific interval between tasks. This is effectively a fake sort of daemon that continuously executes tasks in the queue.

Copyright violation by Andrew Yang

No task could ever escape the finite state machine….

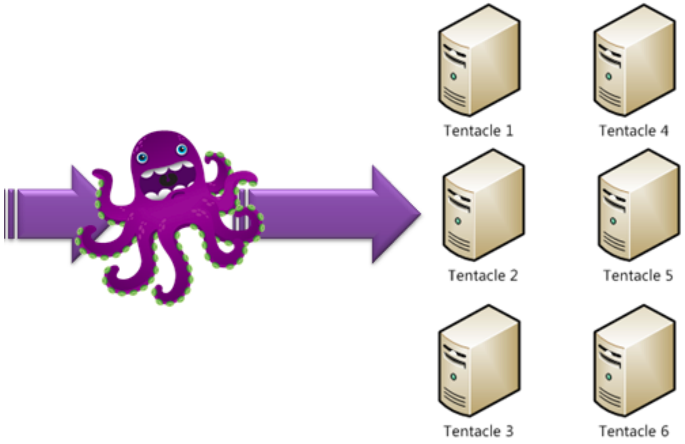
https://www.drupal.org/project/queue_runner

SSH Helper

First encounter to Travis

Allows you to generate a SSH key pair (public/private) for a site using ssh-keygen and uses the phpseclib library.

Pull to Push solution



http://paulstovell.com/blog/octopus/intro

Control all the things!!

https://www.drupal.org/project/ssh_helper
https://github.com/phpseclib/phpseclib

Git Wrapper

First encounter to Composer

This project is an API module that integrates with the Git Wrapper library. This module doesn't do anything by itself and is usually extended by other projects that do something useful with Git.

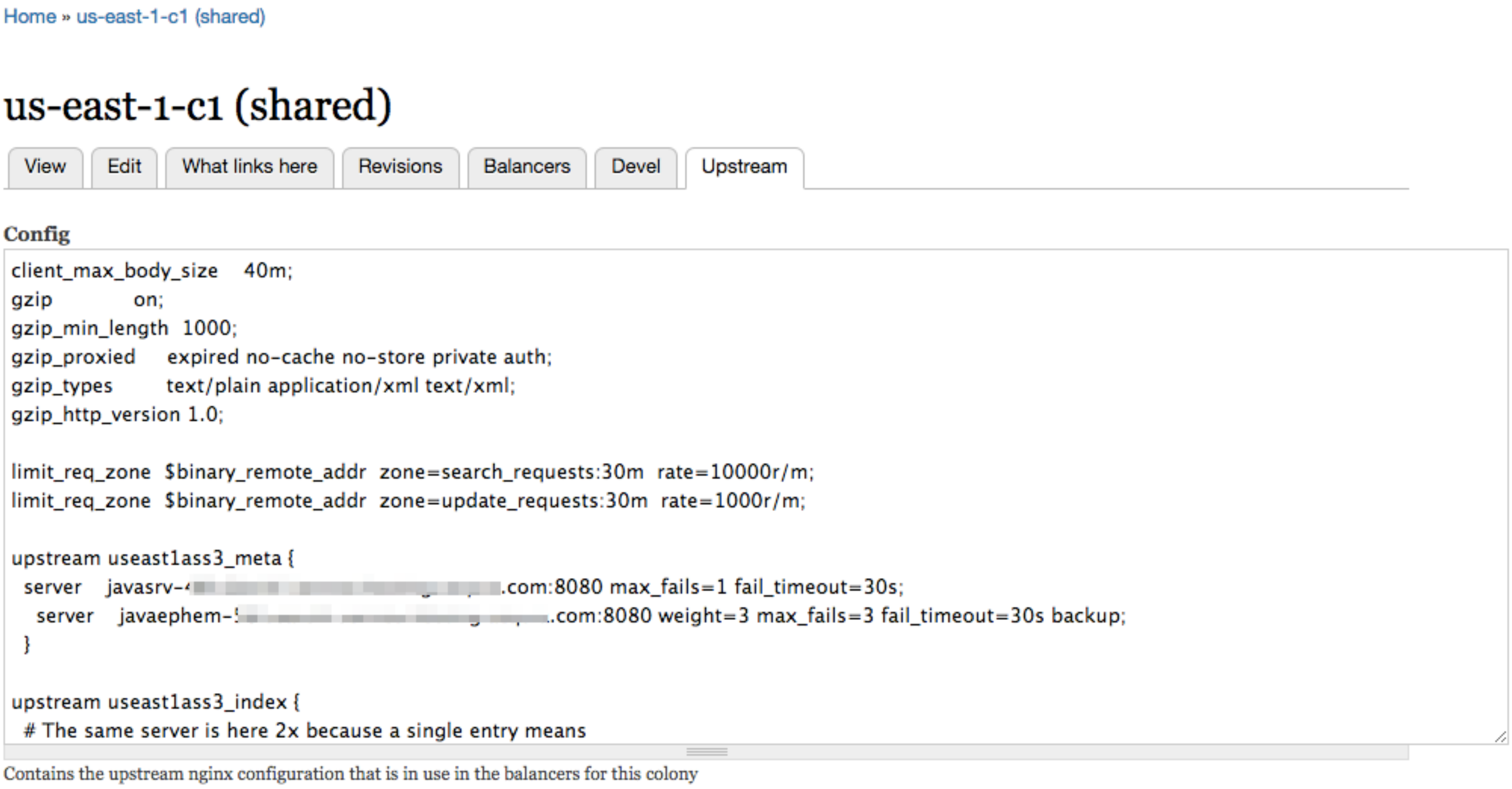
Solved manual balancer configuration issues. Pushes out configs to many servers at once. Scalable.

http://blog.ncsudining.com/wordpress/2010/08/its-a-wrap/

Hmmmmmm… Git

https://www.drupal.org/project/git_wrapper
https://github.com/cpliakas/git-wrapper

Git Wrapper



us-east-1-c1 (shared)

Config

```
client_max_body_size    40m;
gzip            on;
gzip_min_length  1000;
gzip_proxied     expired no-cache no-store private auth;
gzip_types       text/plain application/xml text/xml;
gzip_http_version 1.0;

limit_req_zone  $binary_remote_addr  zone=search_requests:30m  rate=10000r/m;
limit_req_zone  $binary_remote_addr  zone=update_requests:30m  rate=1000r/m;

upstream useast1ass3_meta {
  server    javasrv-[redacted].com:8080 max_fails=1 fail_timeout=30s;
    server    javaephem-[redacted].com:8080 weight=3 max_fails=3 fail_timeout=30s backup;
  }

upstream useast1ass3_index {
  # The same server is here 2x because a single entry means
```

Contains the upstream nginx configuration that is in use in the balancers for this colony

Composer Manager

cpliakas's orchestra

Composer Manager provides a gateway to the larger PHP community by enabling Drupal modules to more easily use best-in-breed libraries that are managed by Composer.

Solves “Not-invented-here” syndrome. Use what is available.

Chris Pliakas's addition: "Someone asked me, if I were stranded on a desert island what book would I bring... 'How to Build a Boat.'"
~Steven Wright

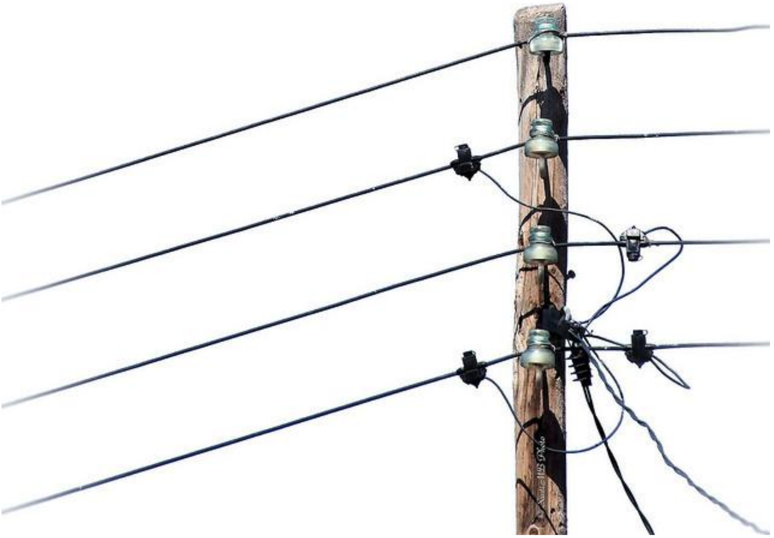
Elk veugeltje zinckt zoo het gebekt is.

http://drupal.org/project/composer_manager

Services & Services Client

The services client module answers that age old question: how do I arbitrarily interface with the services module without having to build a new module from the ground up?

Solves data synchronization between data-storages.



https://www.flickr.com/photos/smb_flickr/2373172675/in/

Wire it all up

http://drupal.org/project/services_client
http://drupal.org/project/services

EntityFieldQuery Extra Fields

Not how EFQ was intended.

```
$query = new EntityFieldQueryExtraFields();
$query->entityCondition('entity_type', 'node')
  ->propertyCondition('type', 'my_bundle_type')
  ->addExtraField('field_myfield', value', value')
  ->execute();
```

Solves the need for quick and painless MySQL queries using Drupalisms.

https://www.flickr.com/photos/73853155@N00/419812216

https://www.drupal.org/project/efq_extra_field

Servers

- Puppet
- Tomcat 6 or 7, Java 6 or 7, Solr 3 or 4.
- Configs set up by Operations and are then launched in EC2 cloud.
- Servers reach out themselves to ask what role they have and ask to be registered as worker

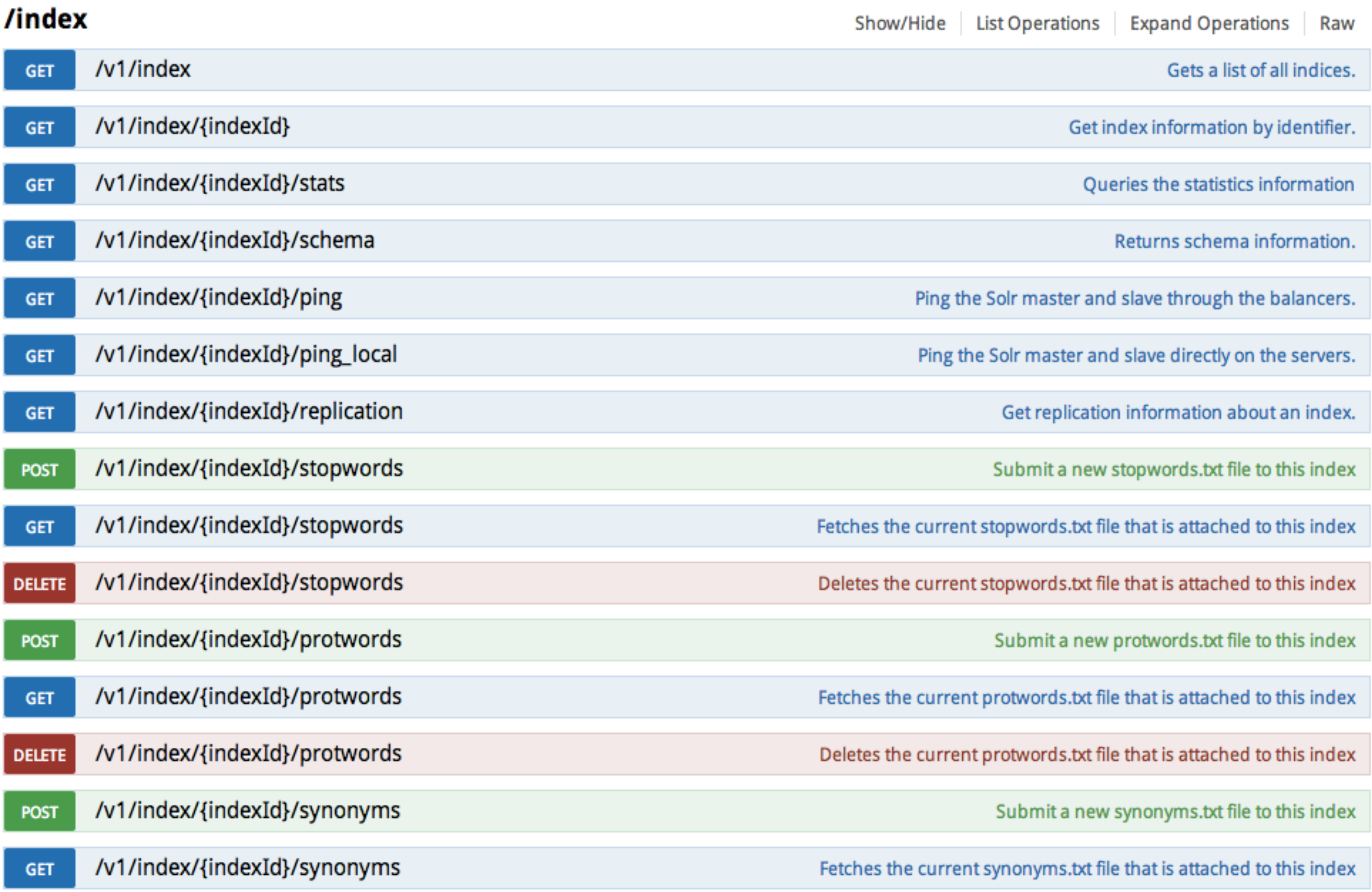
Reference

- Drupalcon Munich
  - http://munich2012.drupal.org/program/sessions/bootstrapping-solr-search-clusters-and-maintaining-them.html

API’s

- '**delivery callback**' => 'drupal_json_output'
- Serves core list/config to each server, if it needs to be rebuilt.
- List of cores/ping URLs to Operations for nagios
- Registration of new Servers, posted by servers asking for a their role.
- Add configuration files to Search Indexes by Support

API’s

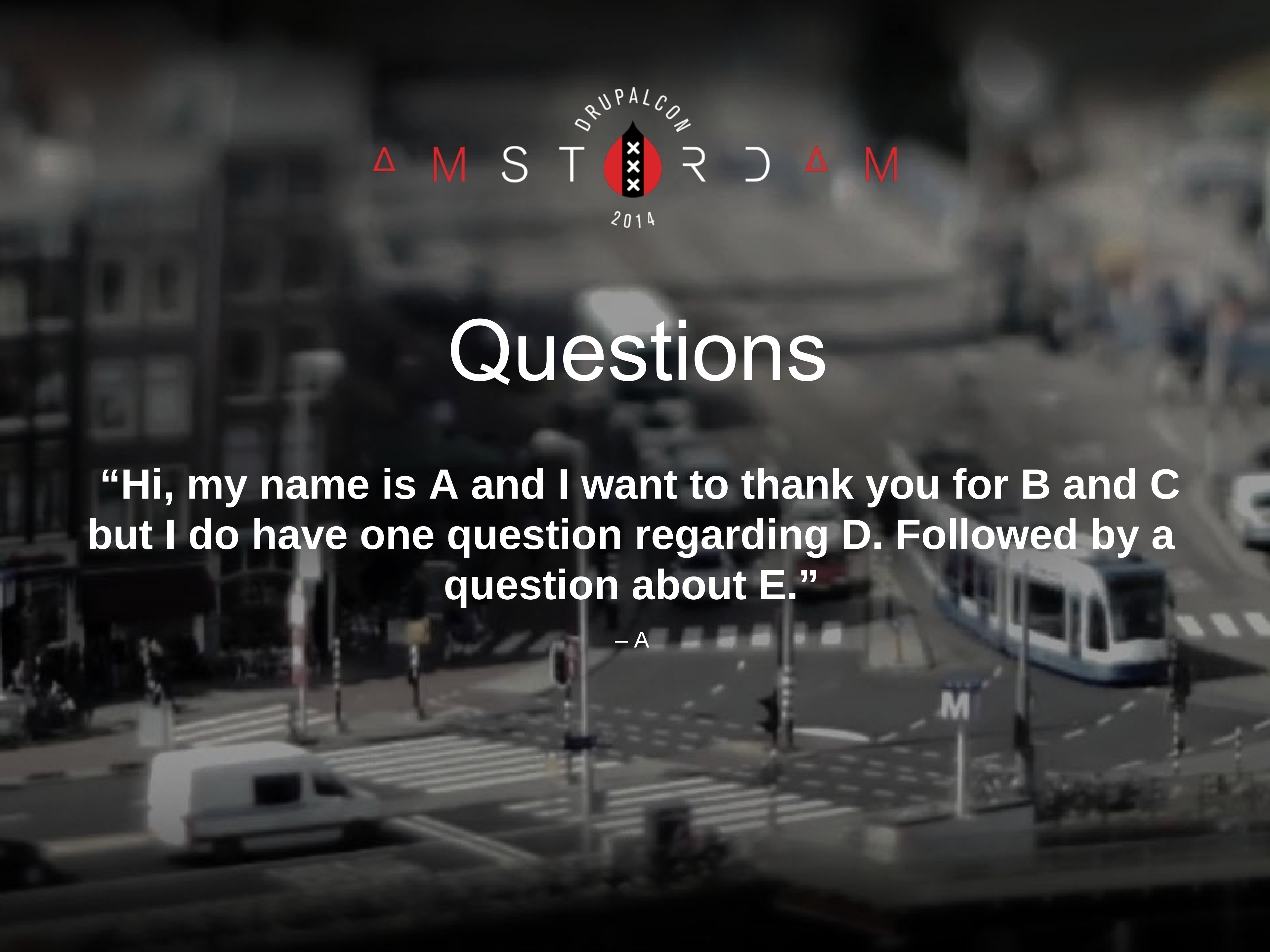


/index

Show/Hide | List Operations | Expand Operations | Raw

GET	/v1/index	Gets a list of all indices.
GET	/v1/index/{indexId}	Get index information by identifier.
GET	/v1/index/{indexId}/stats	Queries the statistics information
GET	/v1/index/{indexId}/schema	Returns schema information.
GET	/v1/index/{indexId}/ping	Ping the Solr master and slave through the balancers.
GET	/v1/index/{indexId}/ping_local	Ping the Solr master and slave directly on the servers.
GET	/v1/index/{indexId}/replication	Get replication information about an index.
POST	/v1/index/{indexId}/stopwords	Submit a new stopwords.txt file to this index
GET	/v1/index/{indexId}/stopwords	Fetches the current stopwords.txt file that is attached to this index
DELETE	/v1/index/{indexId}/stopwords	Deletes the current stopwords.txt file that is attached to this index
POST	/v1/index/{indexId}/protwords	Submit a new protwords.txt file to this index
GET	/v1/index/{indexId}/protwords	Fetches the current protwords.txt file that is attached to this index
DELETE	/v1/index/{indexId}/protwords	Deletes the current protwords.txt file that is attached to this index
POST	/v1/index/{indexId}/synonyms	Submit a new synonyms.txt file to this index
GET	/v1/index/{indexId}/synonyms	Fetches the current synonyms.txt file that is attached to this index

DRUPALCON AMSTERDAM 2014



Questions

> “Hi, my name is A and I want to thank you for B and C but I do have one question regarding D. Followed by a question about E.”
>
> – A

WE ARE HIRING

Acquia™

DRUPALCON

AMSTERDAM

2014

WHAT DID YOU THINK?

EVALUATE THIS SESSION amsterdam2014.
drupal.org/node/2153

THANK YOU!