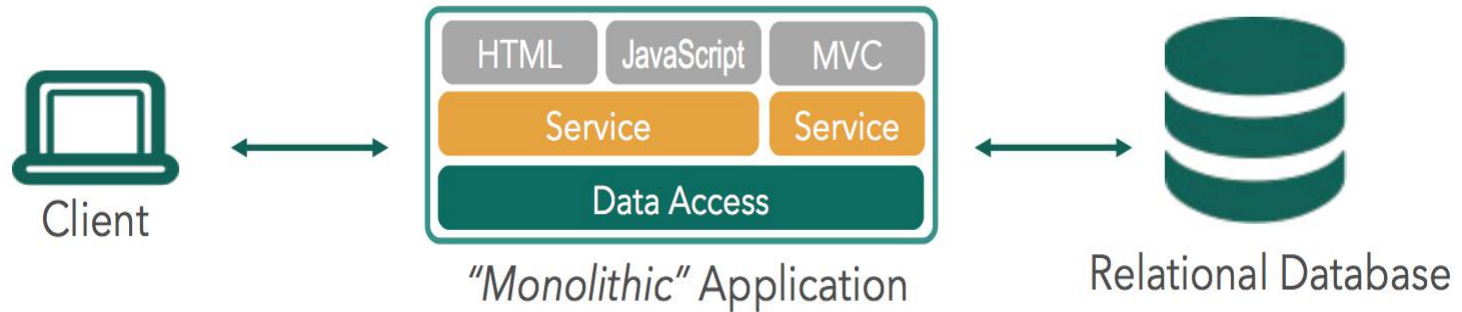


# Spring Cloud Services

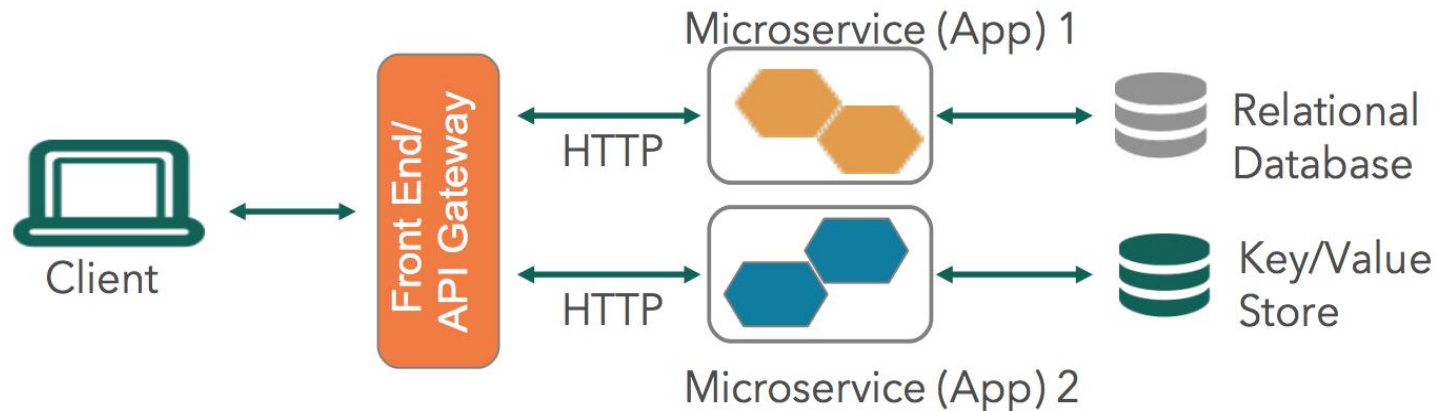
for Pivotal Cloud Foundry

# Monolithic vs. Microservice Architecture

- Classic 3-tier application

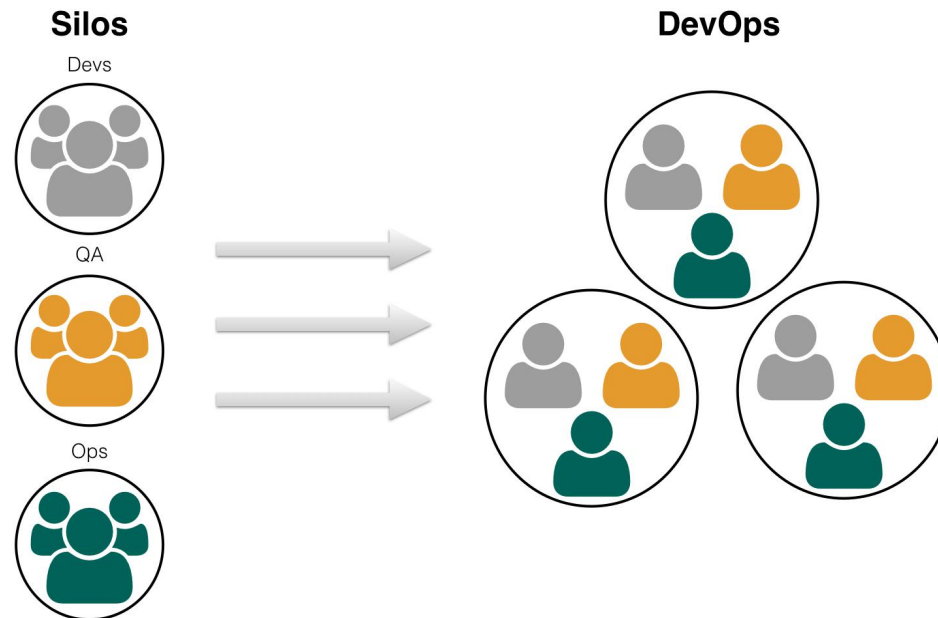


- Microservice architecture



# Silos to DevOps

- Goal: Deliver value rapidly and safely
- Shared vocabulary, tools, and incentive structures
- Bureaucratic processes replaced with trust & accountability
- Common leadership



# Spring Cloud



- **Spring Cloud Netflix:**
  - **Hystrix:** circuit breaker
  - **Eureka:** service discovery
  - **Ribbon:** client-side load balancer
  - **Feign:** declarative REST Client
  - **Zuul:** API proxy server
- **Spring Cloud Config Server:** configuration as a service
- **Spring Cloud Sleuth:** distributed tracing
- **Spring Cloud Contract:** facilitates contract testing

# Spring Cloud Services (SCS)



A PCF **managed service** for deploying  
Spring Cloud infrastructure services  
**on-demand** in the cloud

*Currently supported:*

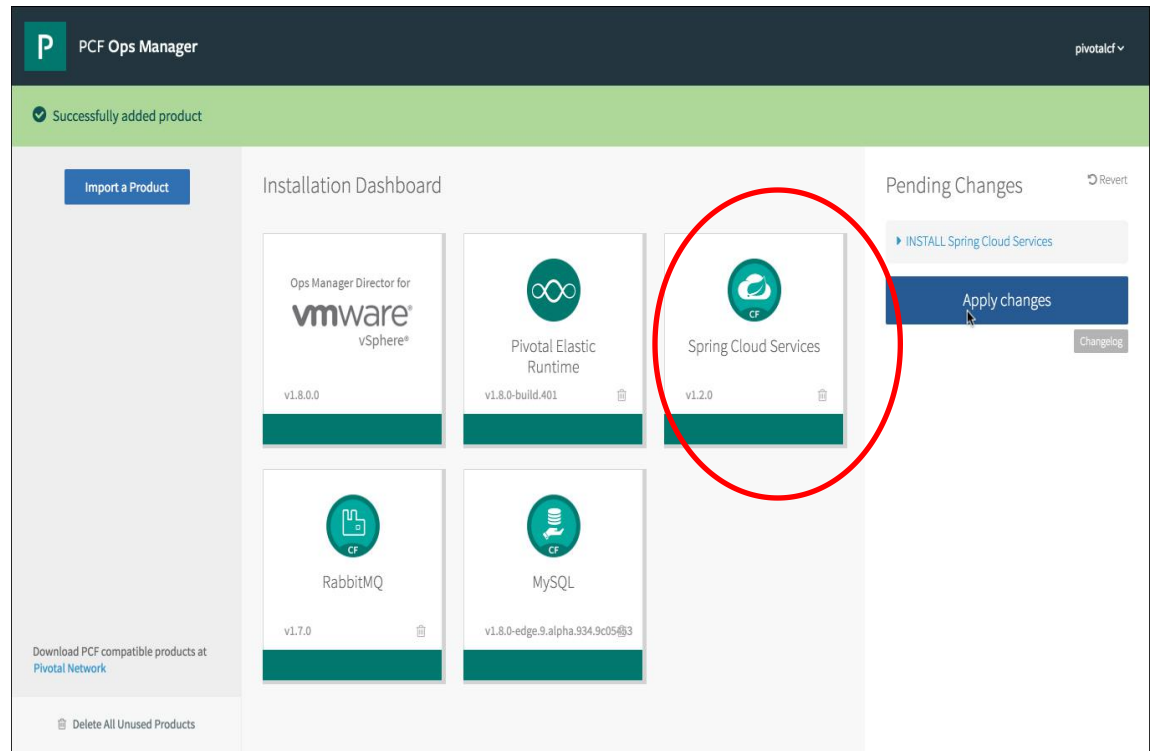
- *Config Server*
- *Service Registry*
- *Circuit Breaker*

<https://docs.pivotal.io/spring-cloud-services>

# Installation

Spring Cloud Services is packaged as a PCF “tile”, installed by a PCF administrator as a CF extension

The installation involves the deployment of the service broker and registration of services into the PCF marketplace




*PCF Operations Manager*

Can verify if a PCF instance has the Spring Cloud Services installed by looking for the presence of these services in the Cloud Foundry Marketplace


## Marketplace

Get started with our free marketplace services. Upgrade select plans to gain access to premium service plans.


### Services ^



**Circuit Breaker**  
Circuit Breaker Dashboard for Spring Cloud Applications



**Config Server**  
Config Server for Spring Cloud Applications



**Service Registry**  
Service Registry for Spring Cloud Applications

# cf plugin for SCS

A plugin is available for the cf CLI, that provides the following commands:

<code>config-server-encrypt-value, csev</code>	Encrypt a string using a Spring Cloud Services configuration server
<code>service-registry-deregister, srdr</code>	Deregister an application registered with a Spring Cloud Services service registry
<code>service-registry-disable, srda</code>	Disable an application registered with a Spring Cloud Services service registry so that it is unavailable for traffic
<code>service-registry-enable, sren</code>	Enable an application registered with a Spring Cloud Services service registry so that it is available for traffic
<code>service-registry-info, sri</code>	Display Spring Cloud Services service registry instance information
<code>service-registry-list, srl</code>	Display all applications registered with a Spring Cloud Services service registry
<code>spring-cloud-service-restage, scs-restage</code>	Restage a Spring Cloud Services service instance
<code>spring-cloud-service-restart, scs-restart</code>	Restart a Spring Cloud Services service instance
<code>spring-cloud-service-start, scs-start</code>	Start a Spring Cloud Services service instance
<code>spring-cloud-service-stop, scs-stop</code>	Stop a Spring Cloud Services service instance
<code>spring-cloud-service-view, scs-view</code>	Display health and status for a Spring Cloud Services service instance

These commands can be useful for analysis and troubleshooting of provisioned services.

<https://plugins.cloudfoundry.org/>

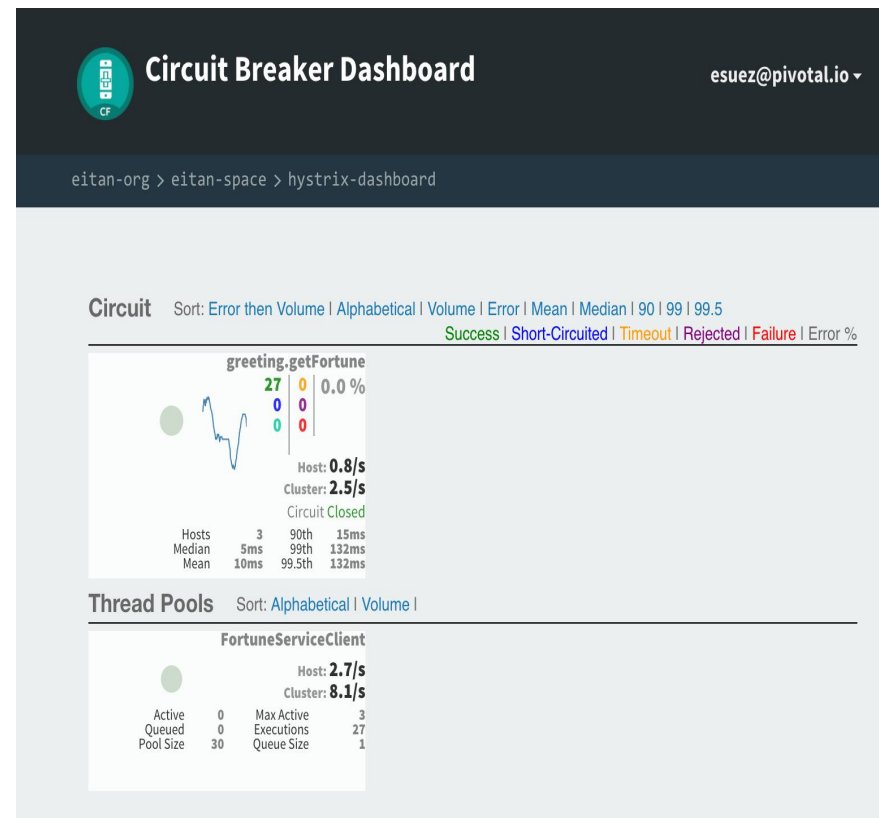


# Provisioning a Circuit Breaker

Example:

```
cf create-service p-circuit-breaker—dashboard standard cb-dashboard
```

After service has been provisioned,  
the circuit breaker dashboard is accessible  
directly from PCF *Apps Manager*

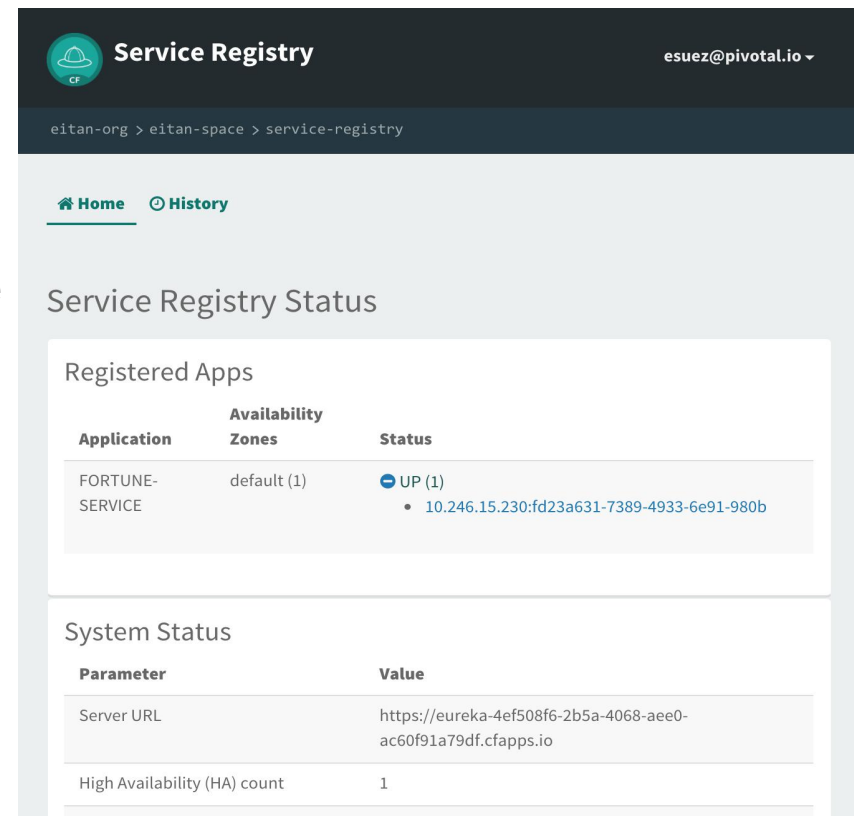


# Provisioning a Service Registry

Example:

```
cf create-service p-service-registry standard service-registry
```

After service has been provisioned,  
the service registry dashboard is accessible  
directly from PCF *Apps Manager*



The screenshot shows the Service Registry dashboard. At the top, there's a header with the Service Registry logo and the user 'esuez@pivotal.io'. Below the header, the breadcrumb 'eitan-org > eitan-space > service-registry' is visible. The main content area has two tabs: 'Home' (selected) and 'History'. The 'Service Registry Status' section contains a 'Registered Apps' table and a 'System Status' table.

Application	Availability Zones	Status
FORTUNE-SERVICE	default (1)	UP (1) <ul style="list-style-type: none"><li>10.246.15.230:fd23a631-7389-4933-6e91-980b</li></ul>

Parameter	Value
Server URL	https://eureka-4ef508f6-2b5a-4068-ae0-ac60f91a79df.cfapps.io
High Availability (HA) count	1

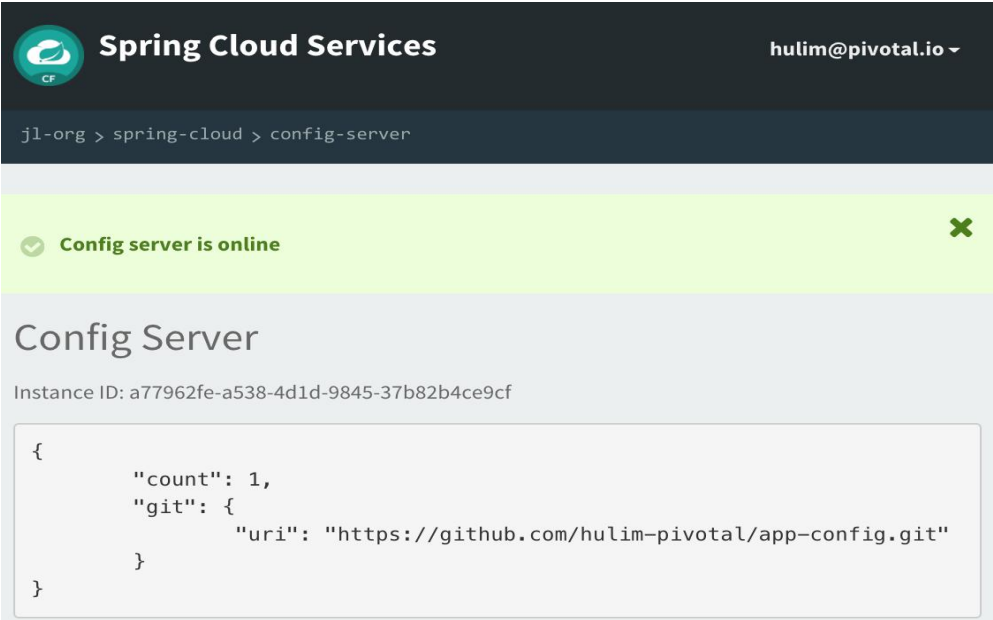
# Provisioning a Config Server

- Example:

```
cf create-service p-config-server standard config-server -c config.json
```

- minimal *config.json* file contents:

```
{"git": {"uri": "https://github.com/<username>/config-repo.git" }}
```



The screenshot shows the Spring Cloud Services console interface. At the top, the header includes the Spring Cloud Services logo, the text "Spring Cloud Services", and the user email "hulim@pivotal.io". Below the header, a breadcrumb trail reads "j1-org > spring-cloud > config-server". A green status bar with a checkmark icon and the text "Config server is online" is displayed. Below this, the section "Config Server" is shown with the "Instance ID: a77962fe-a538-4d1d-9845-37b82b4ce9cf". A code block displays the JSON configuration for the config server:

```
{
  "count": 1,
  "git": {
    "uri": "https://github.com/hulim-pivotal/app-config.git"
  }
}
```

# Circuit Breakers

Netflix Hystrix

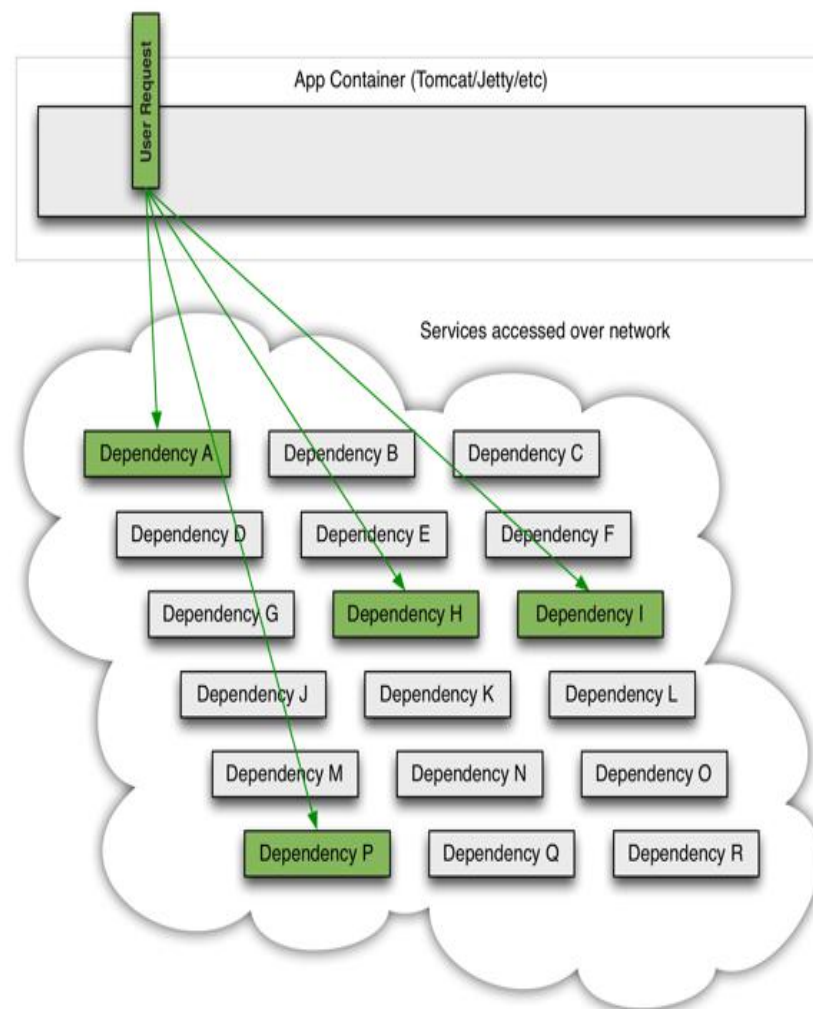
# Motivation

“Applications in complex distributed architectures have dozens of dependencies, each of which will inevitably fail at some point.

If the host application is not isolated from these external failures, it risks being taken down with them.”

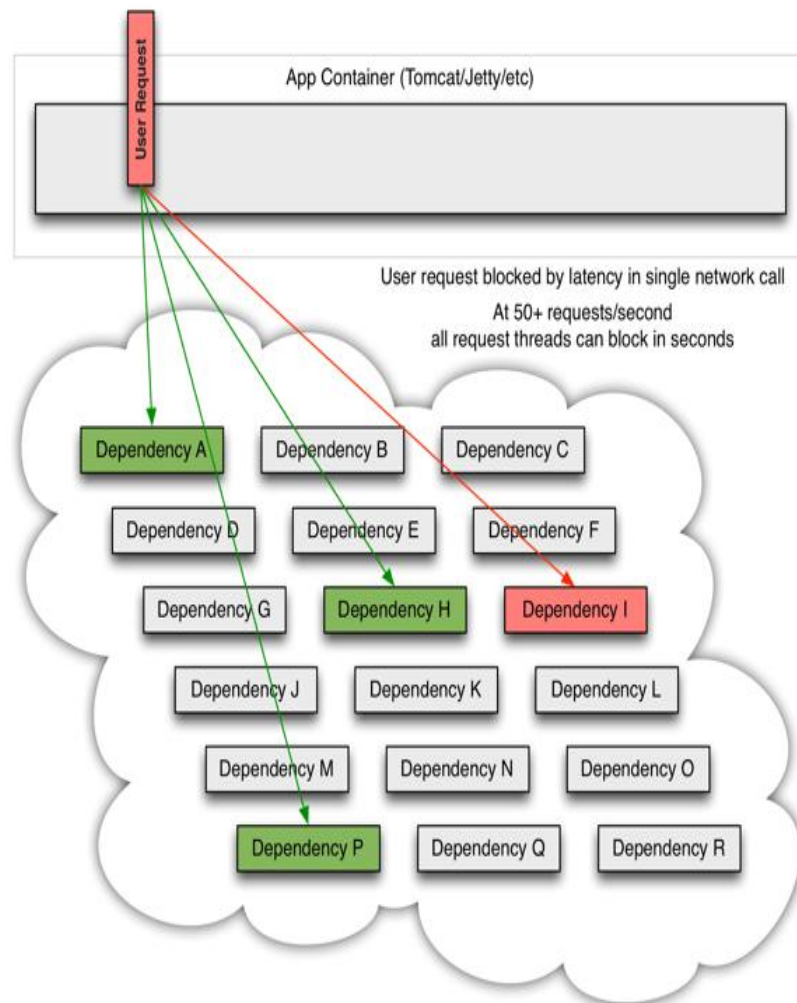
# Services Dependency Scenario

- A typical application depending on a number of backing services
- All services are up and behaving normally
- Circuit is *Closed*



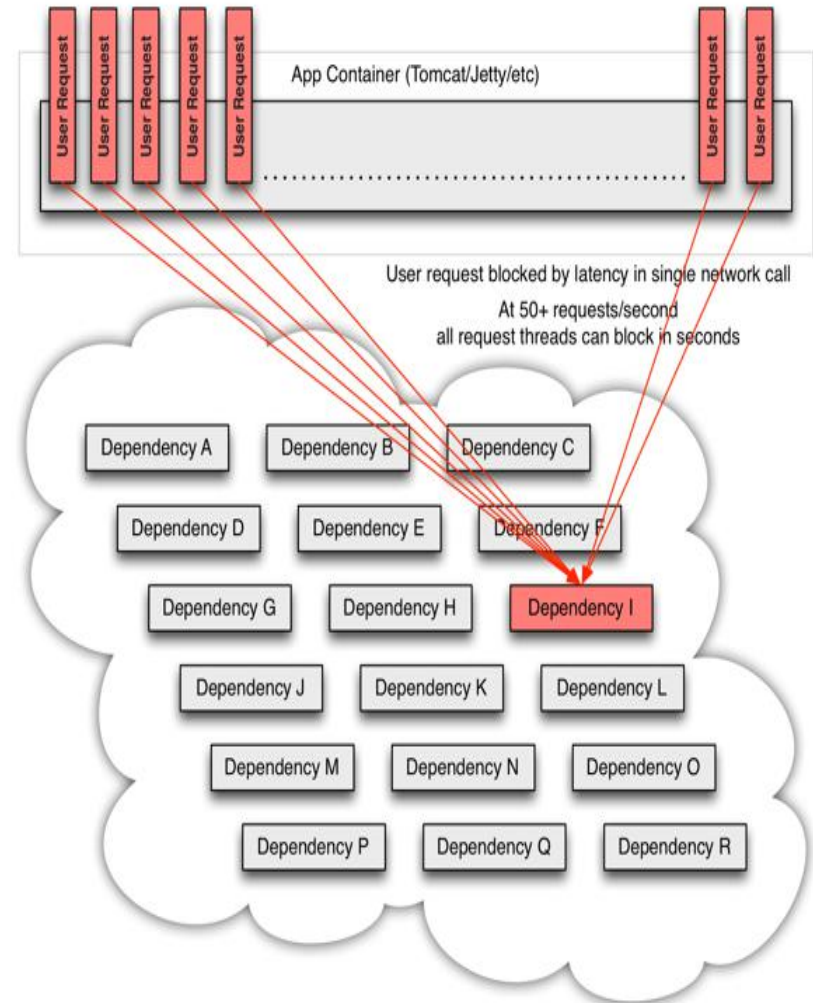
# Failing Dependency

- A dependency begins misbehaves
- Response latency increases, tying up thread in calling application



# Failure Cascades to Caller

- Calling application's thread pool is exhausted waiting on misbehaving dependency
- Failure cascades to caller





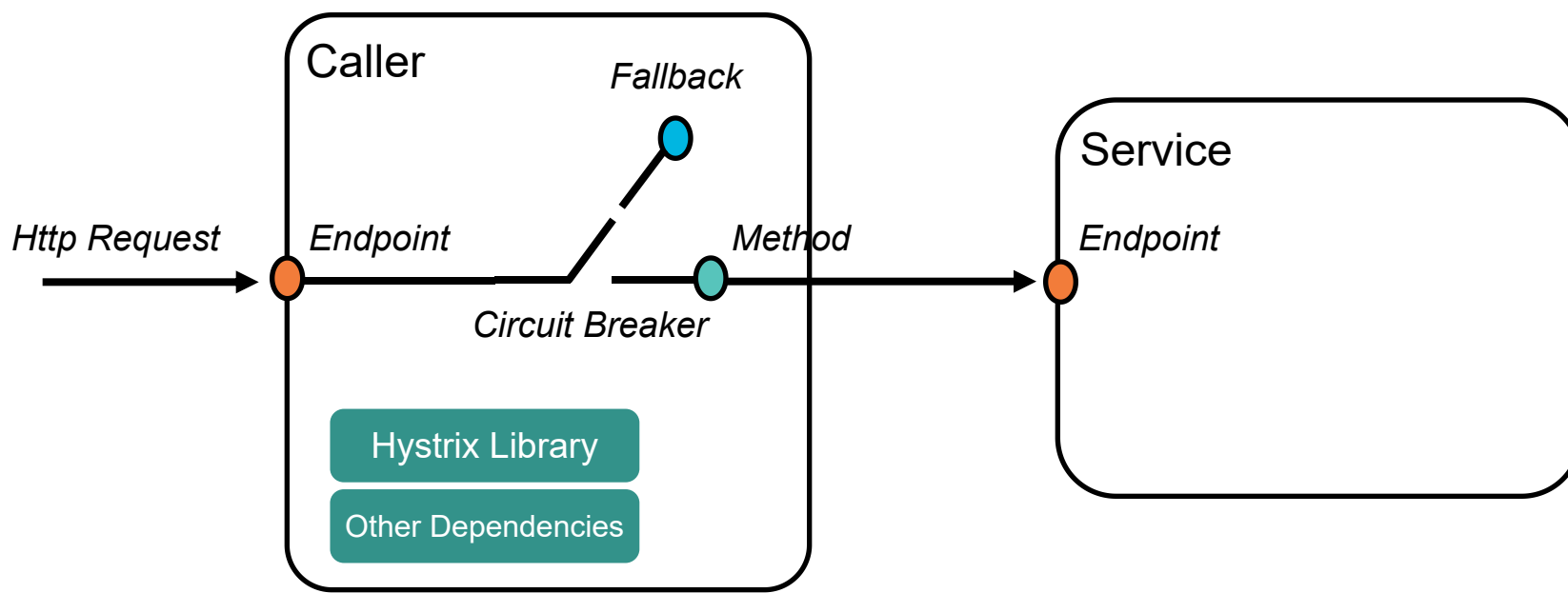
# What is it for?

- Give protection & control over latency & failure from dependencies accessed via 3rd-party client libraries
- Stop cascading failures in a complex distributed system
- Fail fast and rapidly recover
- Fallback and gracefully degrade when possible
- Enable near real-time monitoring, alerting, operational control

# Circuit Breaker isolates calls to other services

Application is isolated from a misbehaving backing service

When backing service health is restored, calling application will automatically reconfigure itself to call it once more



# Annotating a service call

Add build dependency: `spring-cloud-starter-hystrix`

```
25     @HystrixCommand(fallbackMethod = "defaultFortune")
26     String getFortune() {
27         Map map = restTemplate.getForObject(fortuneUrl, Map.class);
28         return (String) map.get("fortune");
29     }
30
31     String defaultFortune() {
32         log.info("Default fortune used");
33         return "Your future is uncertain";
34     }
```

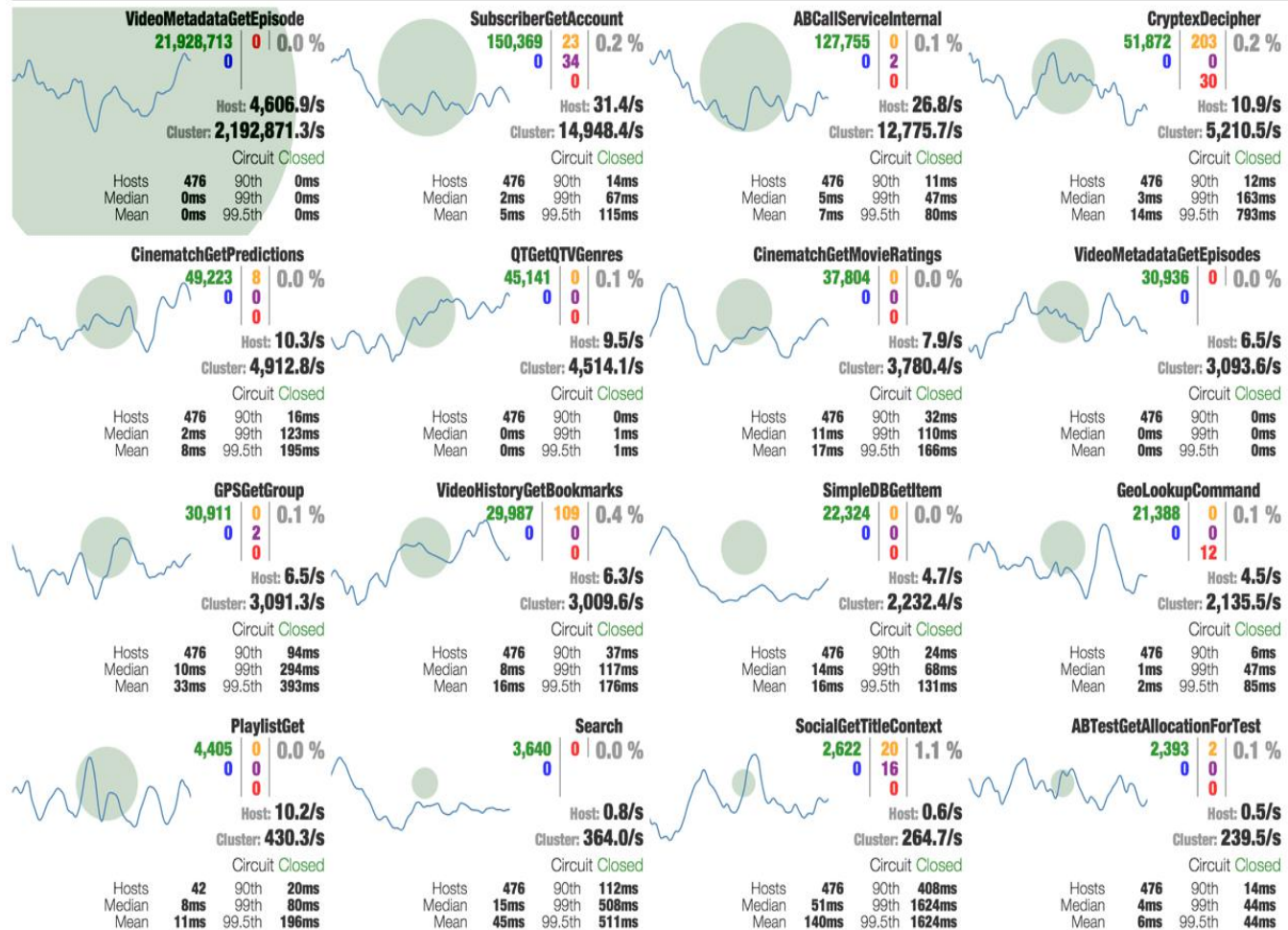


# Configuration

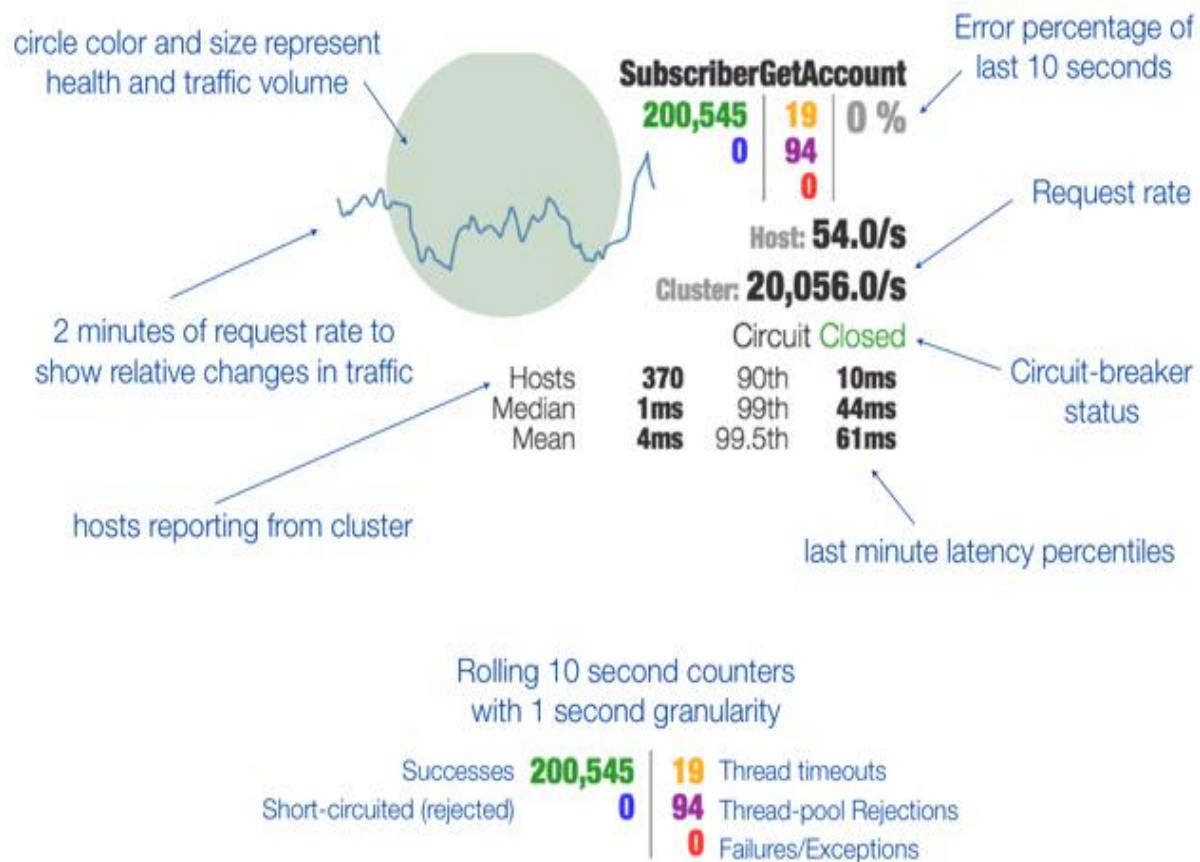
	Default
<b><code>execution.isolation.thread.timeoutInMilliseconds</code></b> The time in milliseconds after which the caller will observe a timeout and walk away from the command execution	1000 ms
<b><code>circuitBreaker.requestVolumeThreshold</code></b> The minimum number of requests in a rolling window that will trip the circuit	20 requests min in rolling window
<b><code>circuitBreaker.sleepWindowInMilliseconds</code></b> The amount of time, after tripping the circuit, to reject requests before allowing attempts again to determine if the circuit should again be closed	5000 ms
<b><code>circuitBreaker.errorThresholdPercentage</code></b> The error percentage at or above which the circuit should trip open and start short-circuiting requests to fallback logic	50%
<b><code>hystrix.threadpool.HystrixThreadPoolKey.maximumSize</code></b> The maximum thread-pool size. This is the maximum amount of concurrency that can be supported without starting to reject <code>HystrixCommands</code>	10

# Hystrix Dashboard

**Circuit Breakers** Sort: [Error then Volume](#) | [Alphabetical](#) | [Volume](#) | [Error](#) | [Mean](#) | [Median](#) | [90](#) | [99](#) | [99.5](#) [Success](#) | [Latent](#) | [Short-Circuited](#) | [Timeout](#) | [Rejected](#) | [Failure](#) | [Error %](#)

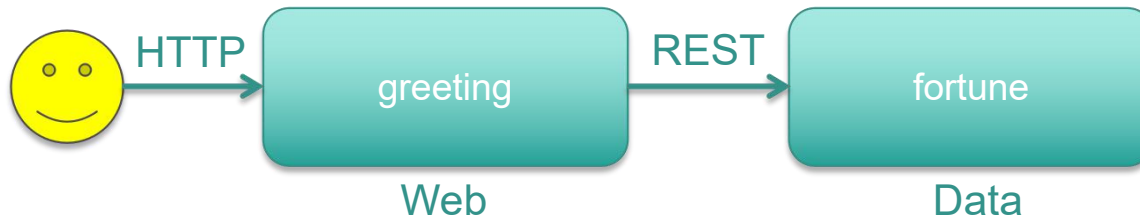


# Circuit Breaker Monitoring



# Lab / Demo

## Hystrix Circuit Breakers



# Service Discovery

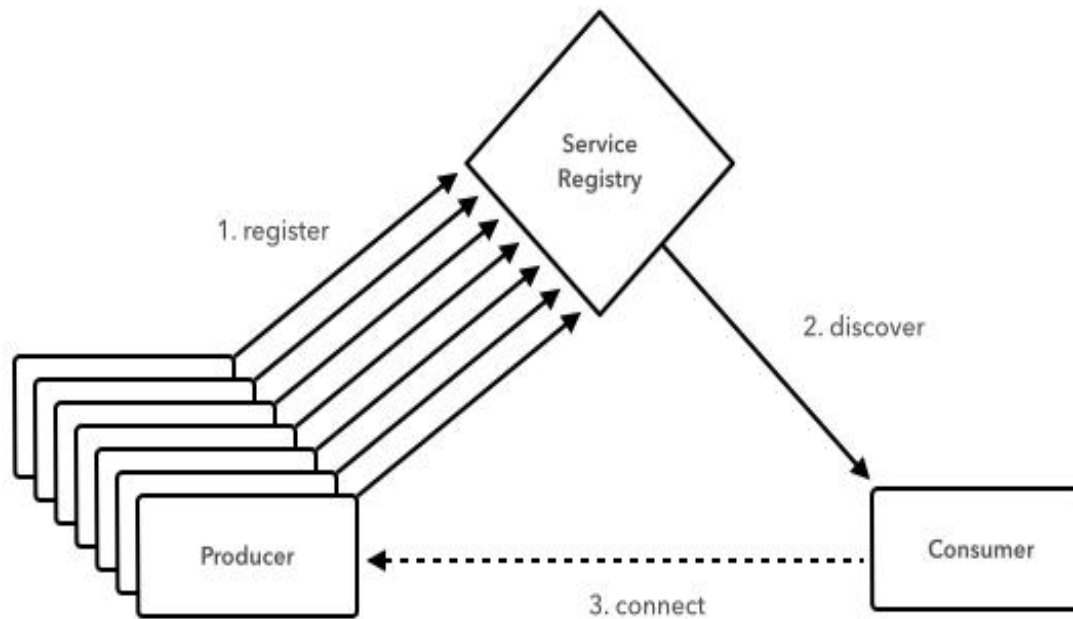
Netflix Eureka



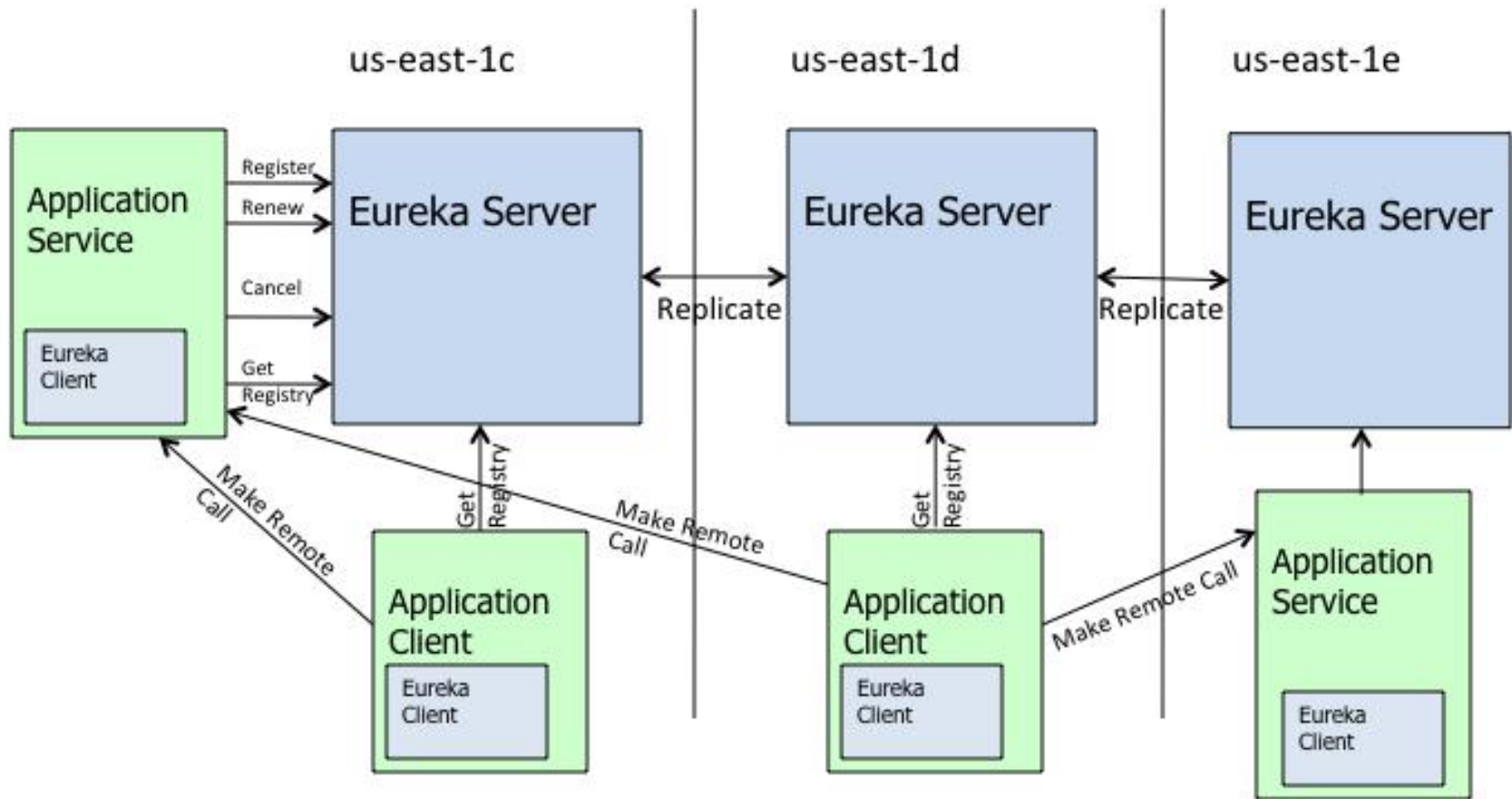
# Service Registries

- A microservice architecture consists of many collaborating service instances that must know each others' address
- A cloud environment implies application instances that come and go, that are dynamically scaled
- Service registries provide dynamic application instance lookup capabilities
- Pattern prevalent in distributed systems: Service Locators, Membership Coordinators
- Examples: HashiCorp Consul, Apache ZooKeeper, Netflix Eureka

# General Concept



# Eureka Architecture




# Renew Registration

- Services must periodically renew their registration, which would otherwise expire
- aka “Heartbeats”
- The configuration property `eureka.instance.leaseRenewalIntervalInSeconds` governs how often a service renews their registration

# Fetch Registry

- Clients fetch a copy of the registry periodically
- An optimization, allows lookups to be performed directly against a cached copy
- `eureka.client.fetchRegistry` can be used to control whether to fetch the registry
- `eureka.client.registryFetchIntervalSeconds` controls how frequently to fetch a new copy

# Eureka Dashboard



HOME    LAST 1000 SINCE STARTUP

## System Status

Environment	test	Current time	2017-11-21T13:13:25 -0600
Data center	default	Uptime	00:00
		Lease expiration enabled	false
		Renews threshold	5
		Renews (last min)	0

## DS Replicas

### Instances currently registered with Eureka

Application	AMIs	Availability Zones	Status
FORTUNE	n/a (1)	(1)	UP (1) - eitans-mbp:fortune:8081
GREETING	n/a (1)	(1)	UP (1) - eitans-mbp:greeting

# Configuring a eureka instance or client

- Add build dependency: `spring-cloud-starter-eureka`
- Configure service with `spring.application.name` property
- Annotate Spring Boot Application class with `@EnableDiscoveryClient`
- Clients auto-wire a `EurekaClient` instance`

# Eureka Lookup Example

```
29 String getFortune() {  
30     String fortuneUrl = lookupUrlFor( appName: "FORTUNE");  
31     Map map = restTemplate.getForObject(fortuneUrl, Map.class);  
32     return (String) map.get("fortune");  
33 }  
34  
35 private String lookupUrlFor(String appName) {  
36     InstanceInfo instance = eurekaClient.getNextServerFromEureka(appName, secure: false);  
37     return instance.getHomePageUrl();  
38 }  
39
```



# Lab / Demo

## Eureka Service Discovery



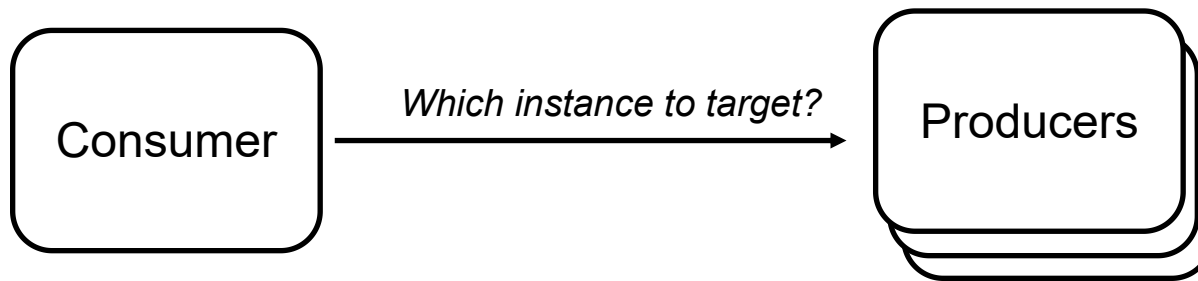
# Client-side Load Balancer

Netflix Ribbon

# Traditional vs Microservice Load Balancing

- Traditional LB:
  - LB performed by dedicated appliance e.g. F5 or HAProxy
  - Configured manually
  - Entry point for HTTP requests from end users (public-facing)
  - Fronts monolithic server instances
- Microservice LB:
  - Embed LB logic in consumer (caller)
  - Configuration is dynamic and automatic
  - Not public-facing
  - Load balancing is between services (inter-service)

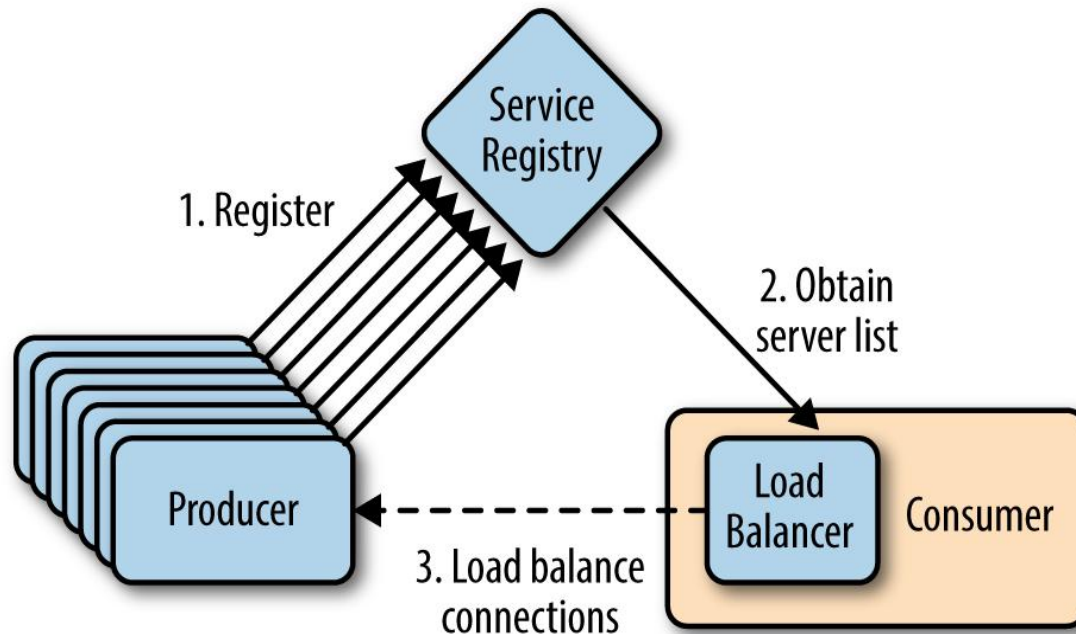
# Service Instances are scaled out



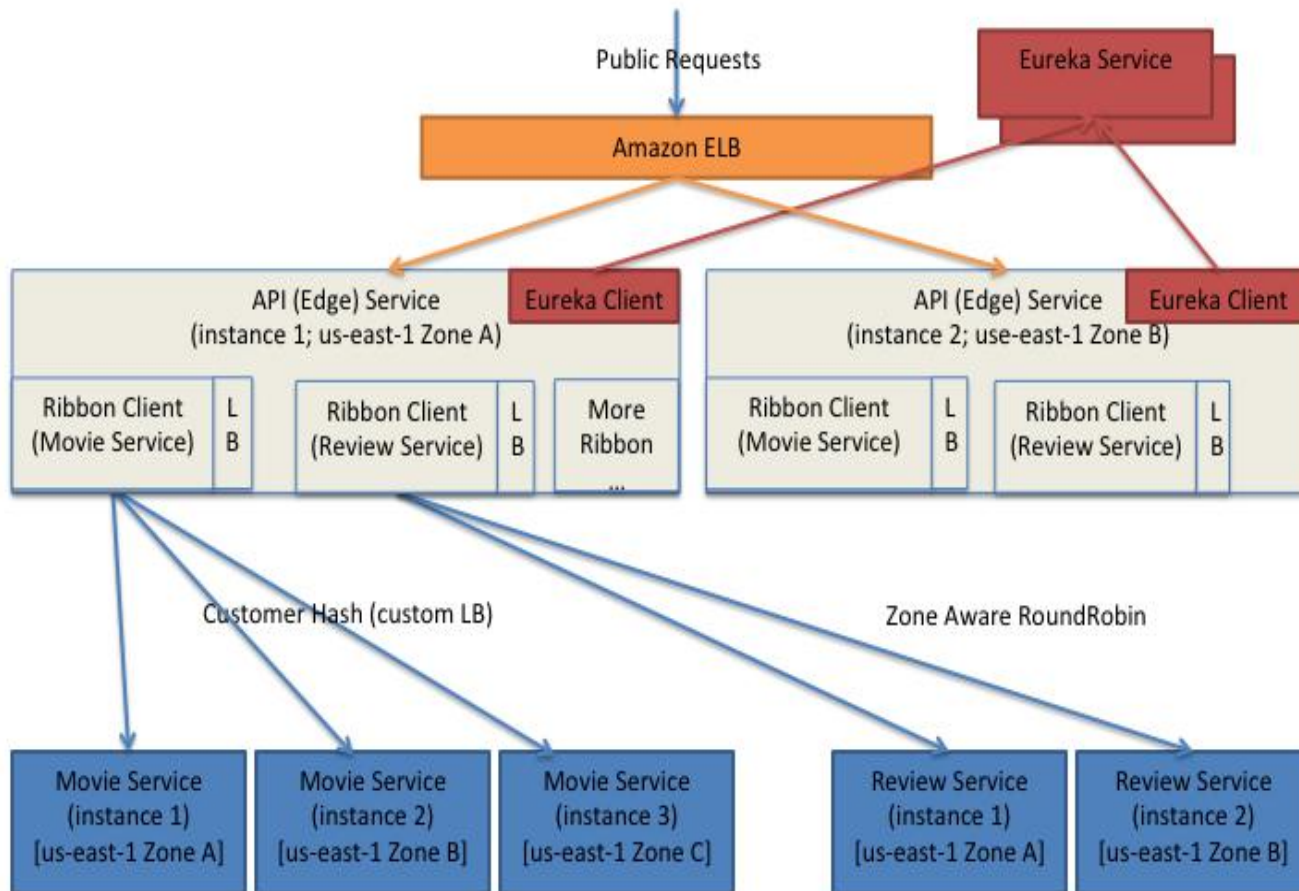
A eureka lookup yields multiple service instances for a given service name

# Netflix Ribbon with Eureka

Ribbon provides several LB algos, runs in-process in the consumer, gets its list of producers from Eureka, and so does not require manual configuration of the server list



# Inter-service Load Balancing



# Load Balancing Rule Options

- RoundRobinRule
- WeightedResponseTimeRule
- RandomRule
- BestAvailableRule
- AvailabilityFilteringRule

*<https://github.com/netflix/ribbon/wiki/working-with-load-balancers>*

# Configuration

	Default
<code>myclient.ribbon.ServerListRefreshInterval</code> The time in milliseconds after which the caller will observe a timeout and walk away from the command execution	30 seconds
<code>myclient.ribbon.NFLoadBalancerRuleClassName</code> The implementation of the load balancing Rule (strategy)	AvailabilityFilteringRule
<code>myclient.ribbon.NFLoadBalancerPingClassName</code> Strategy for pinging servers	NoOpPing
<code>myclient.ribbon.MaxAutoRetriesNextServer</code> Max number of next servers to retry (excluding the first server)	1

See: <https://github.com/Netflix/ribbon/wiki/Getting-Started>



# Ribbon Load Balancing Example

```
20 private final LoadBalancerClient loadBalancerClient;
21
22 public FortuneServiceClient(RestTemplate restTemplate, LoadBalancerClient loadBalancerClient) {
23     this.restTemplate = restTemplate;
24     this.loadBalancerClient = loadBalancerClient;
25 }
26
27 @HystrixCommand(fallbackMethod = "defaultFortune")
28 String getFortune() {
29     String fortuneUrl = lookupUrlFor( appName: "FORTUNE");
30     Map map = restTemplate.getForObject(fortuneUrl, Map.class);
31     return (String) map.get("fortune");
32 }
33
34 private String lookupUrlFor(String appName) {
35     ServiceInstance instance = loadBalancerClient.choose(appName);
36     return String.format("http://%s:%s/", instance.getHost(), instance.getPort());
37 }
38
```

Basically, swap EurekaClient with LoadBalancerClient

API changes slightly: use the choose() method, which returns a ServiceInstance type

# Alternative: @LoadBalanced RestTemplate

```
11  @SpringBootApplication
12  @EnableCircuitBreaker
13  @EnableDiscoveryClient
14  public class GreetingApplication {
15
16  public static void main(String[] args) {
17      SpringApplication.run(GreetingApplication.class, args);
18  }
19
20  @Bean
21  @LoadBalanced
22  public RestTemplate restTemplate() {
23      return new RestTemplate();
24  }
25
26  }
27
```

# RestTemplate Usage

```
16     private final RestTemplate restTemplate;
17
18     public FortuneServiceClient(RestTemplate restTemplate) {
19         this.restTemplate = restTemplate;
20     }
21
22     @HystrixCommand(fallbackMethod = "defaultFortune")
23     String getFortune() {
24         Map map = restTemplate.getForObject(url: "http://fortune/", Map.class);
25         return (String) map.get("fortune");
26     }
27
```

- URL encodes service name (as registered in Eureka)
- Replacement of key with actual service instance returned by load balancing strategy is performed automatically internally to the restTemplate API call (delegates to LoadBalancerClient)

# Lab / Demo

## Ribbon client-side LB



# Declarative REST Client

Netflix Feign

# Spring Cloud Feign

- Encapsulates details of REST API calls behind an interface
- Integrated with Eureka and Ribbon

## Steps:

- Add dependency: `spring-cloud-starter-feign`
- Annotate Spring Boot app class with `@EnableFeignClients`
- Define the interface

# Feign Interface: Simple Example

- Annotate class with `@FeignClient` and indicate the service id of the backing service this interface represents
- Map REST API calls to interface methods
- Annotate each method with the familiar `@RequestMapping` annotation

```
8      @FeignClient("fortune")
9      public interface FortuneAPI {
10
11          @RequestMapping("/")
12          Map<String, String> getFortune();
13      }
14
```

# Usage

- Auto-wire the interface into any client class
- To make REST API call, invoke corresponding interface method instead
- Eureka URL lookup and Ribbon load balancing still take place
- Encapsulate a set of related REST API calls behind an interface

```
15     private final FortuneAPI fortuneAPI;  
16  
17     public FortuneServiceClient(FortuneAPI fortuneAPI) {  
18         this.fortuneAPI = fortuneAPI;  
19     }  
20  
21     @HystrixCommand(fallbackMethod = "defaultFortune")  
22     String getFortune() {  
23         Map map = fortuneAPI.getFortune();  
24         return (String) map.get("fortune");  
25     }  
26
```



# Contrast with RestTemplate

```
16 private final RestTemplate restTemplate;  
17  
18 public FortuneServiceClient(RestTemplate restTemplate) {  
19     this.restTemplate = restTemplate;  
20 }  
21  
22 @HystrixCommand(fallbackMethod = "defaultFortune")  
23 String getFortune() {  
24     Map map = restTemplate.getForObject("http://fortune/", Map.class);  
25     return (String) map.get("fortune");  
26 }  
27
```

```
15 private final FortuneAPI fortuneAPI;  
16  
17 public FortuneServiceClient(FortuneAPI fortuneAPI) {  
18     this.fortuneAPI = fortuneAPI;  
19 }  
20  
21 @HystrixCommand(fallbackMethod = "defaultFortune")  
22 String getFortune() {  
23     Map map = fortuneAPI.getFortune();  
24     return (String) map.get("fortune");  
25 }  
26
```

# Configuration as a Service

## Spring Config Server

# Spring Application Configuration

Traditionally:

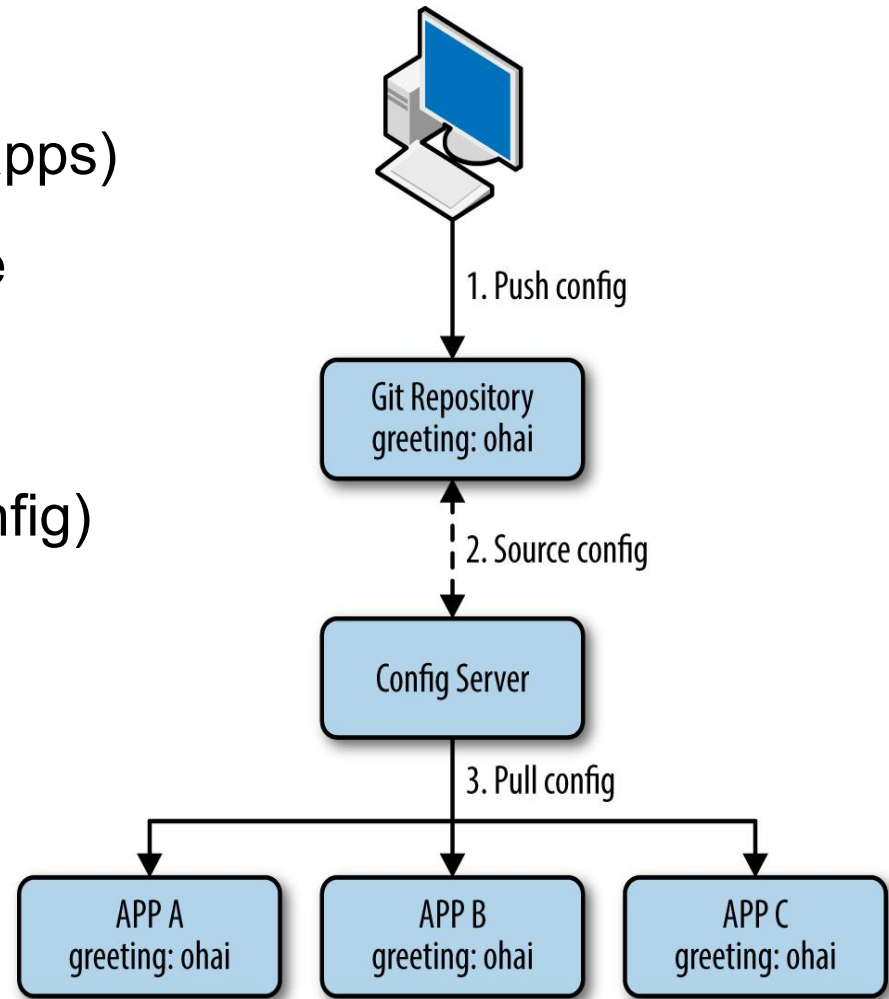
- configuration is stored with the application, and fetched from the classpath
- configuration in a .properties file under src/main/resources

Evolution:

- YAML files (.yml)
- Java system properties
- Environment variables

# Spring Config Server: Concepts

- Externalized config (outside apps)
- Centralized config for multiple services & environments
- Config as a service (REST endpoints for reading app config)



# Backends

- Supports multiple types of backends:
  - Git
  - Subversion
  - HashiCorp Vault
  - File System
  - JDBC
- Also supports a composite of backends

# Backend File Naming

- Default Pattern:

`{application}-{profile}.[properties|yml]`

- Example:

spring.application.name: *greeting*  
spring.profiles.active: *qa*

Configuration stored in a file: ***greeting-qa.yml***  
(or ***greeting-qa.properties***)

*NOTE: The pattern can be customized via a configuration property named [searchPaths](#)*

# HTTP Service Endpoints

`/ {application} / {profile} [ / {label} ]`

`/ {application} - {profile} . yml`

`/ {label} / {application} - {profile} . yml`

`/ {application} - {profile} . properties`

`/ {label} / {application} - {profile} . properties`

- *With the Git backend, {label} maps to a branch name.*
- *{label} is optional, and if not specified, defaults to master*

# Example

`spring.application.name: greeting`

`spring.profiles.active: qa`

Endpoint → `http://{host}:{port}/greeting/qa`

If any of these files exist, their config are returned as one JSON response:

*application.yml (or .properties)*

*application-qa.yml*

*greeting.yml*

*greeting-qa.yml*

Properties in more specifically-named files override those in more general file



```
{
  "name": "greeting",
  "profiles": [
    "qa"
  ],
  "label": null,
  "version": "30cc374c619628d33ac7aada95961fcaca30f568",
  "state": null,
  "propertySources": [
    {
      "name": "file:///Users/esuez/work/config-repo/greeting-qa.yml",
      "source": {
        "greeting.displayFortune": true,
        "fortune.ribbon.NFLoadBalancerRuleClassName": "com.netflix.loadbalancer.RoundRobinRule"
      }
    },
    {
      "name": "file:///Users/esuez/work/config-repo/greeting.yml",
      "source": {
        "greeting.displayFortune": false,
        "fortune.ribbon.NFLoadBalancerRuleClassName": "com.netflix.loadbalancer.WeightedResponseTimeRule"
      }
    },
    {
      "name": "file:///Users/esuez/work/config-repo/application.yml",
      "source": {
        "management.security.enabled": false,
        "security.basic.enabled": false,
        "spring.cloud.services.registrationMethod": "direct",
        "logging.level.io.pivotal.training": "INFO"
      }
    }
  ]
}
```



# Setting up the Server

- Add dependency: `spring-cloud-config-server`
- Annotate Spring Boot app class with `@EnableConfigServer`
- Example configuration of git backend to a public repository:  
*application.properties*:  
`spring.cloud.config.server.git.uri=https://github.com/{username}/config-repo.git`

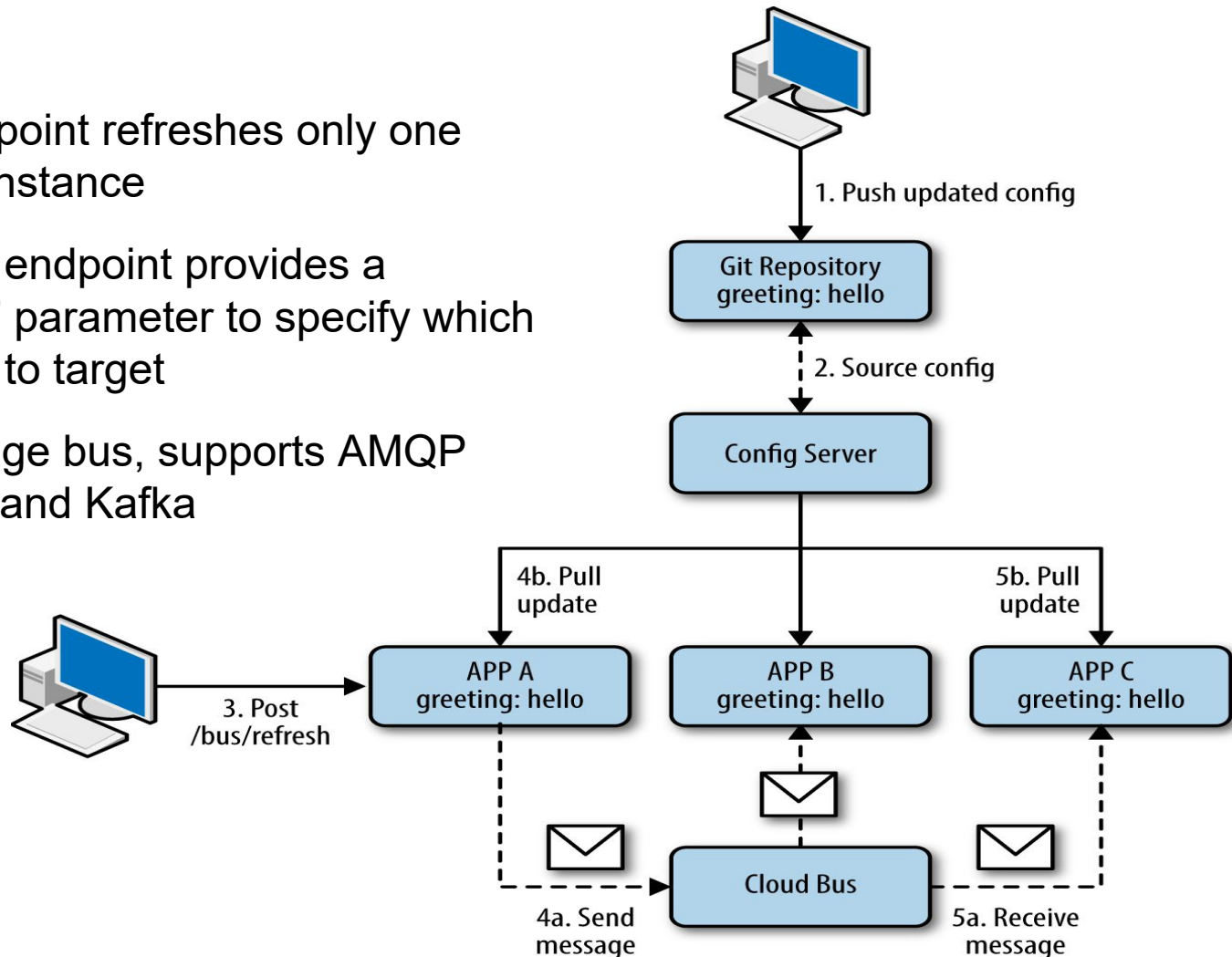
*Many more configuration options exist for the server (consult project reference manual).*

# Configuring Clients

- Add dependency: `spring-cloud-starter-config`
- Set `spring.cloud.config.uri` for location of config server
- Set `spring.application.name` in *bootstrap.yml*,  
not *application.yml*
- Alternatively, set `cloud.config.discovery.enabled`  
to lookup Config Server via Eureka  
(if config server is registered with Eureka)

# Config Server + Spring Cloud Bus

- `/refresh` endpoint refreshes only one application instance
- `/bus/refresh` endpoint provides a “destination” parameter to specify which applications to target
- Uses message bus, supports AMQP (RabbitMQ) and Kafka



# Benefits of Config Server

- All configuration is available in one place
- Separation of app dev lifecycle from configuration lifecycle
- Re-configure aspects of running apps without downtime (e.g. log level, feature toggles)
- Supports encryption of sensitive properties using various mechanisms (symmetric encryption, asymmetric key pair)
- Choice of Git backend provides complete configuration auditability

# Lab / Demo

## Spring Config Server



A dark, atmospheric photograph of the Golden Gate Bridge in San Francisco, with the bridge's towers and suspension cables visible against a misty background. The bridge is the central visual element, with its red-orange color contrasting with the dark, moody tones of the sky and water.

# Pivotal®

Transforming How The World Builds Software

Feedback @ <http://tinyurl.com/apj-eval>

(course: Customized Session Not Shown in the Above Listing)