Kubernetes

Arun

May 27, 2019

First bit of basics

kubernetes

First bit of basics

- pods
- deployment
- service
- horizontal pot autoscaler (HPA)

Before k8s

Before k8s

- 600+ servers
- External load balancers see 5K+ requests per *second*
- Internal Amplification of 8x to 12x

Kubernetes

kubernetes

- Partial Kubernetes deployment since Oct, 2016
- Full Production deployment since Nov, 2016
- Using Google Container Engine

kubernetes

Kubernetes

kubernetes

- Partial Kubernetes deployment since Oct, 2016
- Full Production deployment since Nov, 2016
- Using Google Container Engine
- 5+ Clusters
- 75+ deployments
- 500+ containers (At Peak)

kubernetes

Kubernetes

kubernetes

- Partial Kubernetes deployment since Oct, 2016
- Full Production deployment since Nov, 2016
- Using Google Container Engine
- 5+ Clusters
- 75+ deployments
- 500+ containers (At Peak)
- All new services directly onboard to kubernetes
- Autoscaled on CPU targets

kubernetes

Kubernetes

```yaml
apiVersion: extensions/v1beta1
kind: Deployment
metadata:
  annotations:
    deployment.kubernetes.io/revision: "159"
    kubernetes.io/change-cause: kubectl set image deployment/main-app main-app=xxxxxxxxxxx
      --record=true
  creationTimestamp: 2017-05-11T08:17:58Z
  generation: 3757
  labels:
    app: main-app
    track: stable
  name: main-app
  namespace: default
  resourceVersion: "48500686"
  selfLink: /apis/extensions/v1beta1/namespaces/default/deployments/main-app
  uid: 57d8c2f8-3622-11e7-87c9-420096
spec:
  replicas: 212
  revisionHistoryLimit: 15
  selector:
    matchLabels:
      app: main-app
      track: stable
  strategy:
    rollingUpdate:
      maxSurge: 20%
      maxUnavailable: 5%
    type: RollingUpdate
  template:
    metadata:
      creationTimestamp: null
      labels:
        app: main-app
        track: stable
    spec:
      containers:
      - args:
        - start_prod
        env:
        - name: DISABLE_CONSUL
          value: "true"
        - name: SERVICE_NAME
          value: app
        - name: GUNICORN_THREADS
          value: "14"
        - name: DISABLE_STATSD_EXPORTER
          value: "true"
        image: xxxxxxxxxxx
        imagePullPolicy: Always
        livenessProbe:
          failureThreshold: 3
          httpGet:
            path: /healthcheck/
            port: 8000
            scheme: HTTP
```

- Rolling updates across containers

Kubernetes

```yaml
apiVersion: extensions/v1beta1
kind: Deployment
metadata:
  annotations:
    deployment.kubernetes.io/revision: "159"
    kubernetes.io/change-cause: kubectl set image deployment/main-app main-app=xxxxxxxxxxx
      --record=true
  creationTimestamp: 2017-05-11T08:17:58Z
  generation: 3757
  labels:
    app: main-app
    track: stable
  name: main-app
  namespace: default
  resourceVersion: "48500686"
  selfLink: /apis/extensions/v1beta1/namespaces/default/deployments/main-app
  uid: 57d8c2f8-3622-11e7-87c9-420096
spec:
  replicas: 212
  revisionHistoryLimit: 15
  selector:
    matchLabels:
      app: main-app
      track: stable
  strategy:
    rollingUpdate:
      maxSurge: 20%
      maxUnavailable: 5%
    type: RollingUpdate
  template:
    metadata:
      creationTimestamp: null
      labels:
        app: main-app
        track: stable
    spec:
      containers:
      - args:
        - start_prod
        env:
        - name: DISABLE_CONSUL
          value: "true"
        - name: SERVICE_NAME
          value: app
        - name: GUNICORN_THREADS
          value: "14"
        - name: DISABLE_STATSD_EXPORTER
          value: "true"
        image: xxxxxxxxxxx
        imagePullPolicy: Always
        livenessProbe:
          failureThreshold: 3
          httpGet:
            path: /healthcheck/
            port: 8000
            scheme: HTTP
```

- Rolling updates across containers
- Custom commands to run same container across dev/stage/prod environments

Kubernetes

```yaml
apiVersion: extensions/v1beta1
kind: Deployment
metadata:
  annotations:
    deployment.kubernetes.io/revision: "159"
    kubernetes.io/change-cause: kubectl set image deployment/main-app main-app=xxxxxxxxxxx
      --record=true
  creationTimestamp: 2017-05-11T08:17:58Z
  generation: 3757
  labels:
    app: main-app
    track: stable
  name: main-app
  namespace: default
  resourceVersion: "48500686"
  selfLink: /apis/extensions/v1beta1/namespaces/default/deployments/main-app
  uid: 57d8c2f8-3622-11e7-87c9-420096
spec:
  replicas: 212
  revisionHistoryLimit: 15
  selector:
    matchLabels:
      app: main-app
      track: stable
  strategy:
    rollingUpdate:
      maxSurge: 20%
      maxUnavailable: 5%
    type: RollingUpdate
  template:
    metadata:
      creationTimestamp: null
      labels:
        app: main-app
        track: stable
    spec:
      containers:
      - args:
        - start_prod
        env:
        - name: DISABLE_CONSUL
          value: "true"
        - name: SERVICE_NAME
          value: app
        - name: GUNICORN_THREADS
          value: "14"
        - name: DISABLE_STATSD_EXPORTER
          value: "true"
        image: xxxxxxxxxxx
        imagePullPolicy: Always
        livenessProbe:
          failureThreshold: 3
          httpGet:
            path: /healthcheck/
            port: 8000
            scheme: HTTP
```

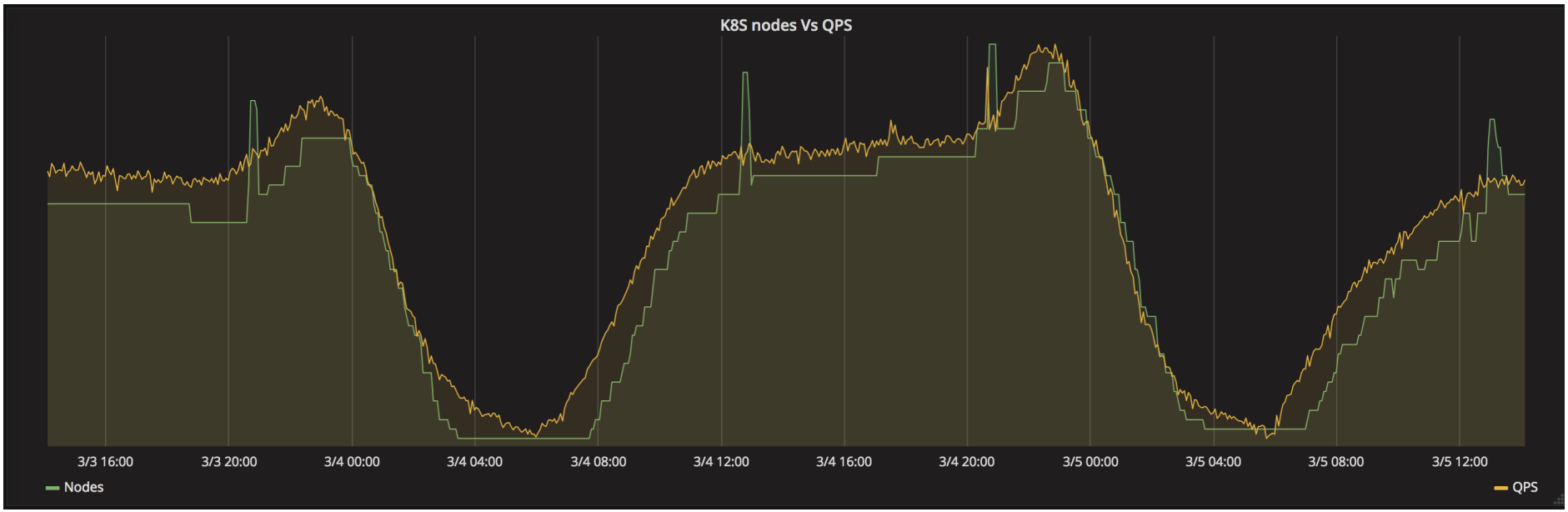
- Rolling updates across containers
- Custom commands to run same container across dev/stage/prod environments
- Environment variables to override 'config-service' values

Kubernetes

```yaml
apiVersion: extensions/v1beta1
kind: Deployment
metadata:
  annotations:
    deployment.kubernetes.io/revision: "159"
    kubernetes.io/change-cause: kubectl set image deployment/main-app main-app=xxxxxxxxxxx
      --record=true
  creationTimestamp: 2017-05-11T08:17:58Z
  generation: 3757
  labels:
    app: main-app
    track: stable
  name: main-app
  namespace: default
  resourceVersion: "48500686"
  selfLink: /apis/extensions/v1beta1/namespaces/default/deployments/main-app
  uid: 57d8c2f8-3622-11e7-87c9-420096
spec:
  replicas: 212
  revisionHistoryLimit: 15
  selector:
    matchLabels:
      app: main-app
      track: stable
  strategy:
    rollingUpdate:
      maxSurge: 20%
      maxUnavailable: 5%
    type: RollingUpdate
  template:
    metadata:
      creationTimestamp: null
      labels:
        app: main-app
        track: stable
    spec:
      containers:
      - args:
        - start_prod
        env:
        - name: DISABLE_CONSUL
          value: "true"
        - name: SERVICE_NAME
          value: app
        - name: GUNICORN_THREADS
          value: "14"
        - name: DISABLE_STATSD_EXPORTER
          value: "true"
        image: xxxxxxxxxxx
        imagePullPolicy: Always
        livenessProbe:
          failureThreshold: 3
          httpGet:
            path: /healthcheck/
            port: 8000
            scheme: HTTP
```

- Rolling updates across containers
- Custom commands to run same container across dev/stage/prod environments
- Environment variables to override 'config-service' values
- Liveness and Readiness checks are must!

Kubernetes Autoscaling



kubernetes

Kubernetes

NAME	DESIRED	CURRENT	UP-TO-DATE	AVAILABLE	AGE
chat-app	30	30	30	30	75d
main-app	204	204	204	204	75d
search-app	36	36	36	36	75d
workers-accounts	9	9	9	9	22d
workers-accounts-auth	1	1	1	1	22d
workers-activities	1	1	1	1	22d
workers-apns	1	1	1	1	22d
workers-apns-feedback	1	1	1	1	22d
workers-beat	1	1	1	1	22d
workers-bulk-index-products	4	4	4	4	22d
workers-dump	4	4	4	4	22d
workers-chat-emails	10	10	10	10	22d
workers-message	3	3	3	3	22d
workers-comments	3	3	3	3	22d
workers-delay-images	5	5	5	5	22d
workers-duplicate-products	6	6	6	6	22d
workers-emailalias	1	1	1	1	22d
workers-fblistings	5	5	5	5	22d
workers-fbshare	11	11	11	11	22d
workers-flaggings	2	2	2	2	22d
workers-grouplistings	2	2	2	2	22d
workers-listings	5	5	5	5	22d
workers-market	1	1	1	1	22d
workers-purchases	1	1	1	1	22d
workers-transactions	1	1	1	1	22d
workers-notifications	9	9	9	9	10d
workers-notifications-1	3	3	3	3	22d
workers-password	1	1	1	1	22d
workers-post-category-check	2	2	2	2	22d
workers-post-price-check	2	2	2	2	22d
workers-price-consumer	1	1	1	1	22d
workers-product-history	2	2	2	2	22d
workers-product-updates-stream	4	4	4	4	22d
workers-productlk	7	7	7	7	22d
workers-push-apns	3	3	3	3	22d
workers-push-notification	2	2	2	2	22d
workers-push-gcm	8	8	8	8	22d
workers-register-android	3	3	3	3	22d
workers-register-ios	5	5	5	5	22d
workers-runconsumer	1	1	1	1	22d
workers-sls	4	4	4	4	22d
workers-scheduler	1	1	1	1	22d
workers-seo	1	1	1	1	22d
workers-services-fulfill	1	1	1	1	22d
workers-services-visibility	2	2	2	2	22d
workers-signals-handler	4	4	4	4	22d
workers-listings	4	4	4	4	22d
workers-social	1	1	1	1	22d
workers-support	6	6	6	6	22d
workers-verification	1	1	1	1	22d
workers-wback	2	2	2	2	22d
workers-worker	10	10	10	10	22d

- Not using **K8S Ingress/Service**

Kubernetes

NAME	DESIRED	CURRENT	UP-TO-DATE	AVAILABLE	AGE
chat-app	30	30	30	30	75d
main-app	204	204	204	204	75d
search-app	36	36	36	36	75d
workers-accounts	9	9	9	9	22d
workers-accounts-auth	1	1	1	1	22d
workers-activities	1	1	1	1	22d
workers-apns	1	1	1	1	22d
workers-apns-feedback	1	1	1	1	22d
workers-beat	1	1	1	1	22d
workers-bulk-index-products	4	4	4	4	22d
workers-dump	4	4	4	4	22d
workers-chat-emails	10	10	10	10	22d
workers-message	3	3	3	3	22d
workers-comments	3	3	3	3	22d
workers-delay-images	5	5	5	5	22d
workers-duplicate-products	6	6	6	6	22d
workers-emailalias	1	1	1	1	22d
workers-fblistings	5	5	5	5	22d
workers-fbshare	11	11	11	11	22d
workers-flaggings	2	2	2	2	22d
workers-grouplistings	2	2	2	2	22d
workers-listings	5	5	5	5	22d
workers-market	1	1	1	1	22d
workers-purchases	1	1	1	1	22d
workers-transactions	1	1	1	1	22d
workers-notifications	9	9	9	9	10d
workers-notifications-1	3	3	3	3	22d
workers-password	1	1	1	1	22d
workers-post-category-check	2	2	2	2	22d
workers-post-price-check	2	2	2	2	22d
workers-price-consumer	1	1	1	1	22d
workers-product-history	2	2	2	2	22d
workers-product-updates-stream	4	4	4	4	22d
workers-productlk	7	7	7	7	22d
workers-push-apns	3	3	3	3	22d
workers-push-notification	2	2	2	2	22d
workers-push-gcm	8	8	8	8	22d
workers-register-android	3	3	3	3	22d
workers-register-ios	5	5	5	5	22d
workers-runconsumer	1	1	1	1	22d
workers-sls	4	4	4	4	22d
workers-scheduler	1	1	1	1	22d
workers-seo	1	1	1	1	22d
workers-services-fulfill	1	1	1	1	22d
workers-services-visibility	2	2	2	2	22d
workers-signals-handler	4	4	4	4	22d
workers-listings	4	4	4	4	22d
workers-social	1	1	1	1	22d
workers-support	6	6	6	6	22d
workers-verification	1	1	1	1	22d
workers-wback	2	2	2	2	22d
workers-worker	10	10	10	10	22d

- Not using **K8S Ingress/Service**
- Config-Service as DaemonSet
- Containers get registered on Config-Service (*NodePort*) from health check

Kubernetes

NAME	DESIRED	CURRENT	UP-TO-DATE	AVAILABLE	AGE
chat-app	30	30	30	30	75d
main-app	204	204	204	204	75d
search-app	36	36	36	36	75d
workers-accounts	9	9	9	9	22d
workers-accounts-auth	1	1	1	1	22d
workers-activities	1	1	1	1	22d
workers-apns	1	1	1	1	22d
workers-apns-feedback	1	1	1	1	22d
workers-beat	1	1	1	1	22d
workers-bulk-index-products	4	4	4	4	22d
workers-dump	4	4	4	4	22d
workers-chat-emails	10	10	10	10	22d
workers-message	3	3	3	3	22d
workers-comments	3	3	3	3	22d
workers-delay-images	5	5	5	5	22d
workers-duplicate-products	6	6	6	6	22d
workers-emailalias	1	1	1	1	22d
workers-fblistings	5	5	5	5	22d
workers-fbshare	11	11	11	11	22d
workers-flaggings	2	2	2	2	22d
workers-grouplistings	2	2	2	2	22d
workers-listings	5	5	5	5	22d
workers-market	1	1	1	1	22d
workers-purchases	1	1	1	1	22d
workers-transactions	1	1	1	1	22d
workers-notifications	9	9	9	9	10d
workers-notifications-1	3	3	3	3	22d
workers-password	1	1	1	1	22d
workers-post-category-check	2	2	2	2	22d
workers-post-price-check	2	2	2	2	22d
workers-price-consumer	1	1	1	1	22d
workers-product-history	2	2	2	2	22d
workers-product-updates-stream	4	4	4	4	22d
workers-productlk	7	7	7	7	22d
workers-push-apns	3	3	3	3	22d
workers-push-notification	2	2	2	2	22d
workers-push-gcm	8	8	8	8	22d
workers-register-android	3	3	3	3	22d
workers-register-ios	5	5	5	5	22d
workers-runconsumer	1	1	1	1	22d
workers-sls	4	4	4	4	22d
workers-scheduler	1	1	1	1	22d
workers-seo	1	1	1	1	22d
workers-services-fulfill	1	1	1	1	22d
workers-services-visibility	2	2	2	2	22d
workers-signals-handler	4	4	4	4	22d
workers-listings	4	4	4	4	22d
workers-social	1	1	1	1	22d
workers-support	6	6	6	6	22d
workers-verification	1	1	1	1	22d
workers-wback	2	2	2	2	22d
workers-worker	10	10	10	10	22d

- Not using **K8S Ingress/Service**
- Config-Service as DaemonSet
- Containers get registered on Config-Service (*NodePort*) from health check
- No change in existing architecture needed
- Service discovery from Internal/External load balancers

Kubernetes

```
NAME          REFERENCE                  TARGETS     MINPODS   MAXPODS   REPLICAS   AGE
chat-app      Deployment/chat-app        48% / 48%   6         90        30         75d
main-app      Deployment/main-app        50% / 48%   18        350       204        75d
search-app    Deployment/search-app      41% / 43%   6         90        36         74d
NAME        DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
main-app    204       204       204          204         75d
NAME        DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
chat-app    30        30        30           30          75d
NAME          DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
search-app    36        36        36           36          75d
```

- 'Config-Service' allows us to have hybrid model
- Services can be deployed in multiple clusters
- Instance groups can coexist with Kubernetes

Kubernetes

```
NAME          REFERENCE                  TARGETS     MINPODS   MAXPODS   REPLICAS   AGE
chat-app      Deployment/chat-app        48% / 48%   6         90        30         75d
main-app      Deployment/main-app        50% / 48%   18        350       204        75d
search-app    Deployment/search-app      41% / 43%   6         90        36         74d
NAME        DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
main-app    204       204       204          204         75d
NAME        DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
chat-app    30        30        30           30          75d
NAME          DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
search-app    36        36        36           36          75d
```

- 'Config-Service' allows us to have hybrid model
- Services can be deployed in multiple clusters
- Instance groups can coexist with Kubernetes
- Instances run the same container image

Kubernetes

```
NAME          REFERENCE                  TARGETS     MINPODS   MAXPODS   REPLICAS   AGE
chat-app      Deployment/chat-app        48% / 48%   6         90        30         75d
main-app      Deployment/main-app        50% / 48%   18        350       204        75d
search-app    Deployment/search-app      41% / 43%   6         90        36         74d
NAME        DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
main-app    204       204       204          204         75d
NAME        DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
chat-app    30        30        30           30          75d
NAME          DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
search-app    36        36        36           36          75d
```

- 'Config-Service' allows us to have hybrid model
- Services can be deployed in multiple clusters
- Instance groups can coexist with Kubernetes
- Instances run the same container image
- Recovery mechanism / Transitioning

Kubernetes

```
NAME          REFERENCE                  TARGETS     MINPODS   MAXPODS   REPLICAS   AGE
chat-app      Deployment/chat-app        48% / 48%   6         90        30         75d
main-app      Deployment/main-app        50% / 48%   18        350       204        75d
search-app    Deployment/search-app      41% / 43%   6         90        36         74d
NAME       DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
main-app   204       204       204          204         75d
NAME       DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
chat-app   30        30        30           30          75d
NAME          DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE
search-app    36        36        36           36          75d
```

- 'Config-Service' allows us to have hybrid model
- Services can be deployed in multiple clusters
- Instance groups can coexist with Kubernetes
- Instances run the same container image
- Recovery mechanism / Transitioning
- Easy Migration - Create new K8S cluster and terminate old cluster

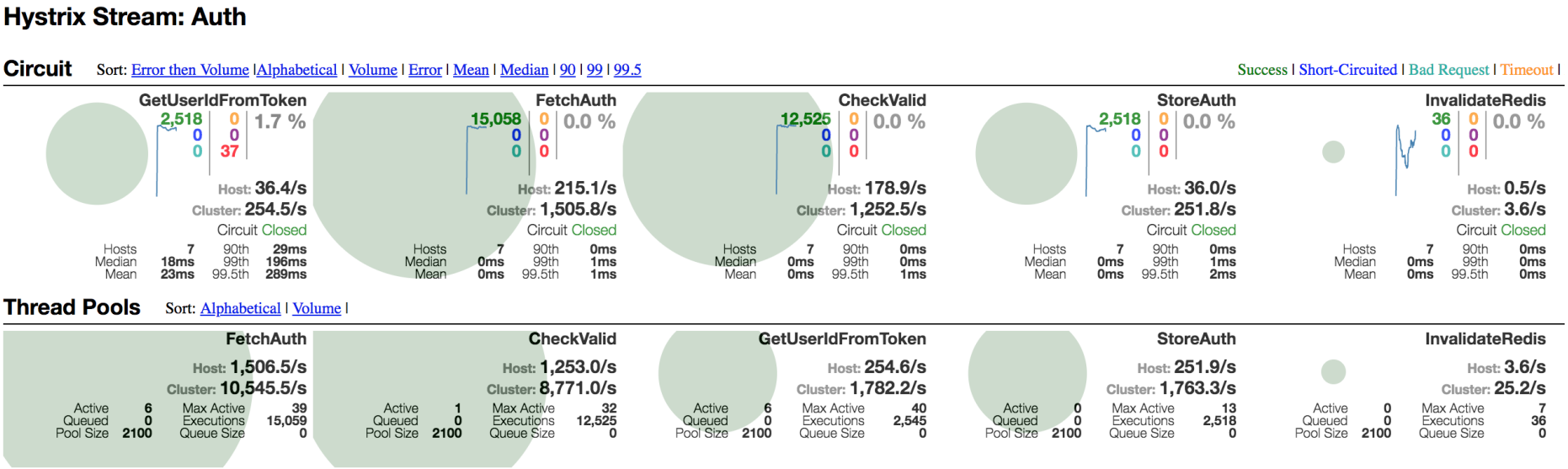
Kubernetes

- Envoy for communication between micro-services (gRPC)

Kubernetes

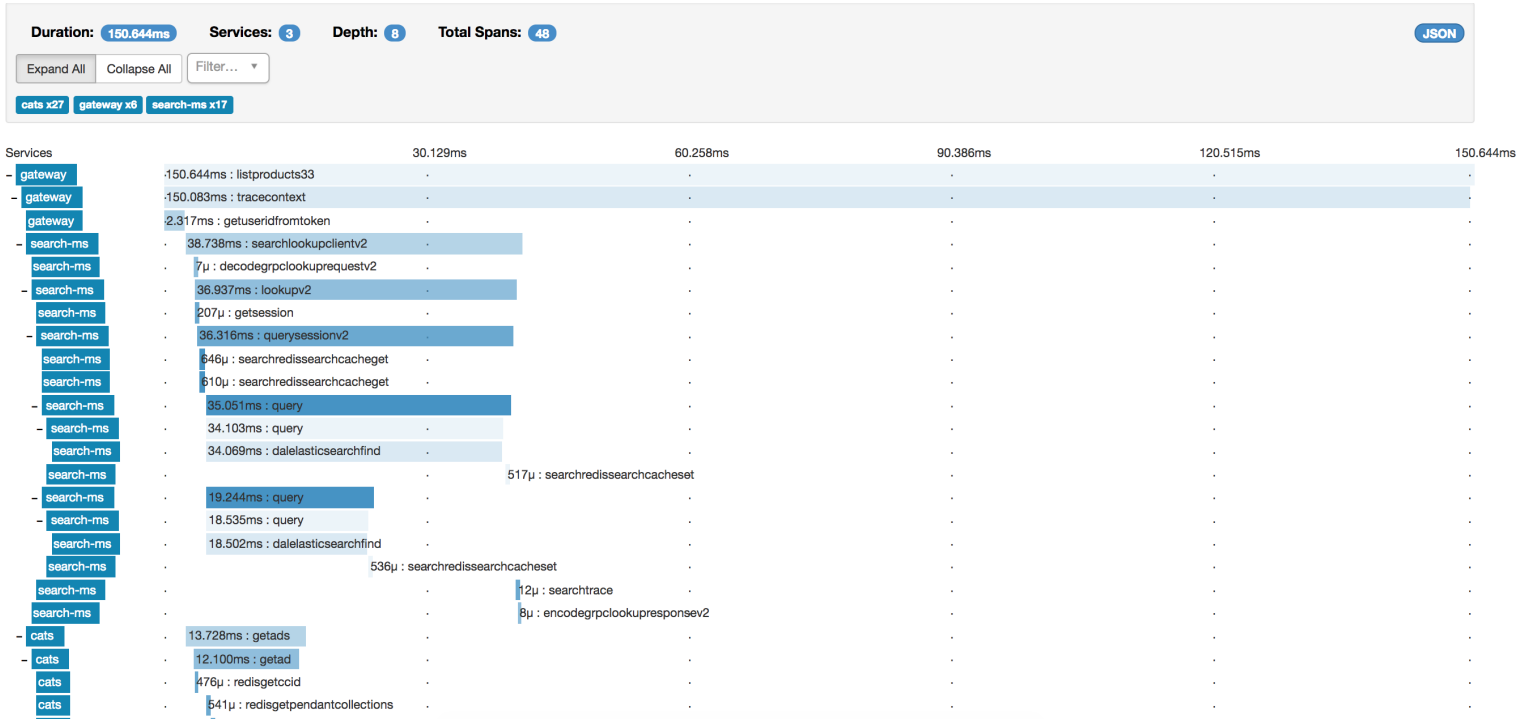
- Envoy for communication between micro-services (gRPC)
- Prometheus for monitoring

Kubernetes



- Envoy for communication between micro-services (gRPC)
- Prometheus for monitoring
- Turbine for hystrix stream aggregation

Kubernetes



- Envoy for communication between micro-services (gRPC)
- Prometheus for monitoring
- Turbine for hystrix stream aggregation
- Zipkin Spans are written to external kafka

Deployment Pipeline

	Build Docker Image	Deploy Image To Canary	Promote Image To Latest	Deploy to Prod	Deploy to Kubernetes	Deploy to Workers
Average stage times: (Average full run time: ~30min 23s)	52s	53s	13s	55s	8min 28s	44s
#1027 Mar 05 15:22 — No Changes	57s	47s (paused for 34s)	14s (paused for 7min 0s)	56s	6min 38s	45s
#1026 Mar 05 14:59 — No Changes	1min 1s	49s (paused for 2min 47s)	12s (paused for 8min 51s)	1min 1s	6min 30s	41s
#1025 Mar 05 14:05 — No Changes	50s	52s (paused for 9min 9s)	14s (paused for 30min 32s)	58s	7min 48s	45s
#1024 Mar 05 12:32 — No Changes	45s	49s (paused for 1min 49s)	15s (paused for 11min 48s)	53s	10min 1s	37s
#1023 Mar 02 17:12 — No Changes	54s	51s (paused for 22s)	10s (paused for 7min 11s)	52s	9min 8s	44s
#1022 Mar 02 15:40 — No Changes	51s	50s (paused for 8s)	15s (paused for 35min 44s)	53s	8min 25s	41s

kubernetes

Thank You

Q&A