Kubernetes in Production

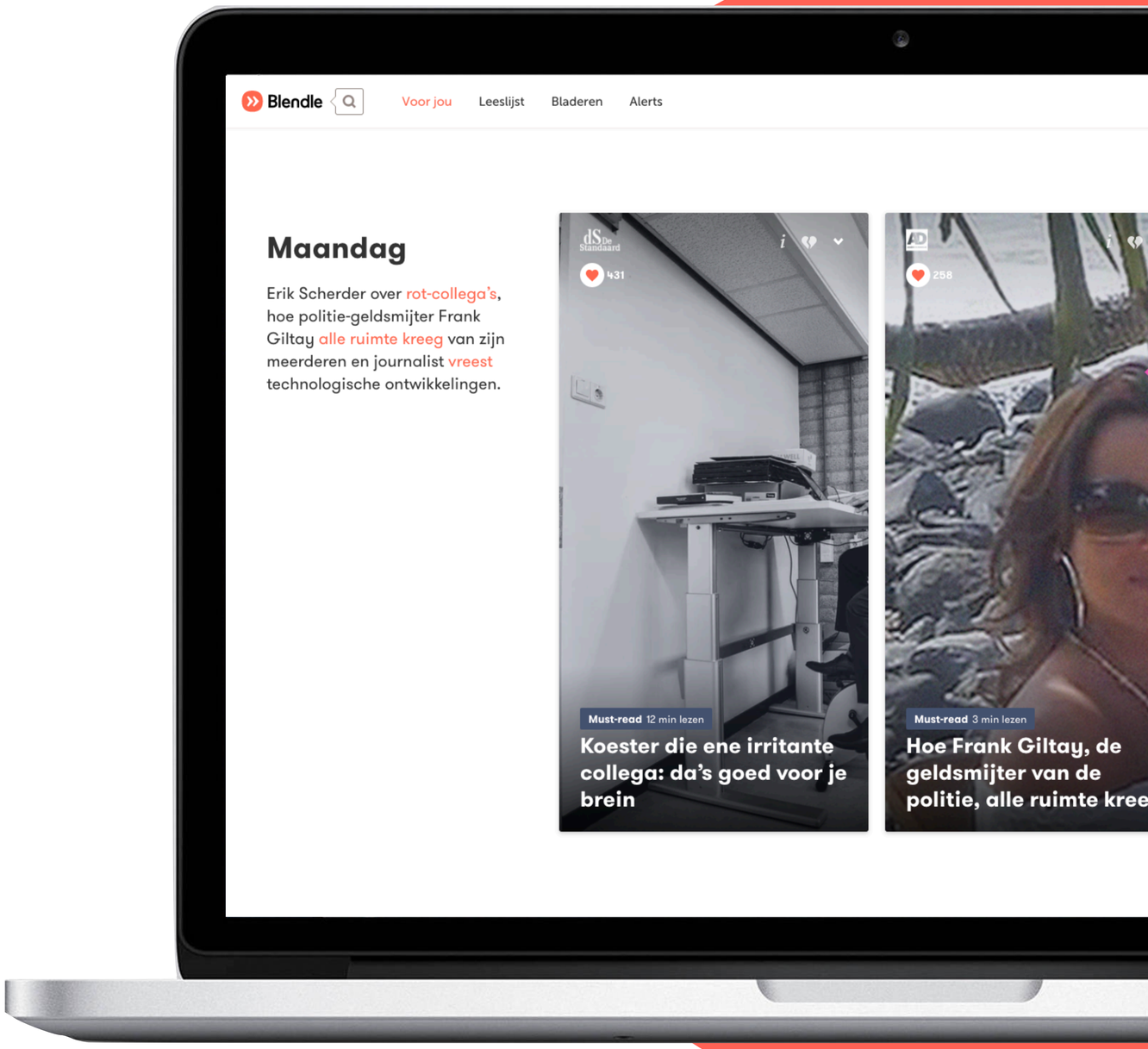
Niels Stevens

Blendle

Hello!

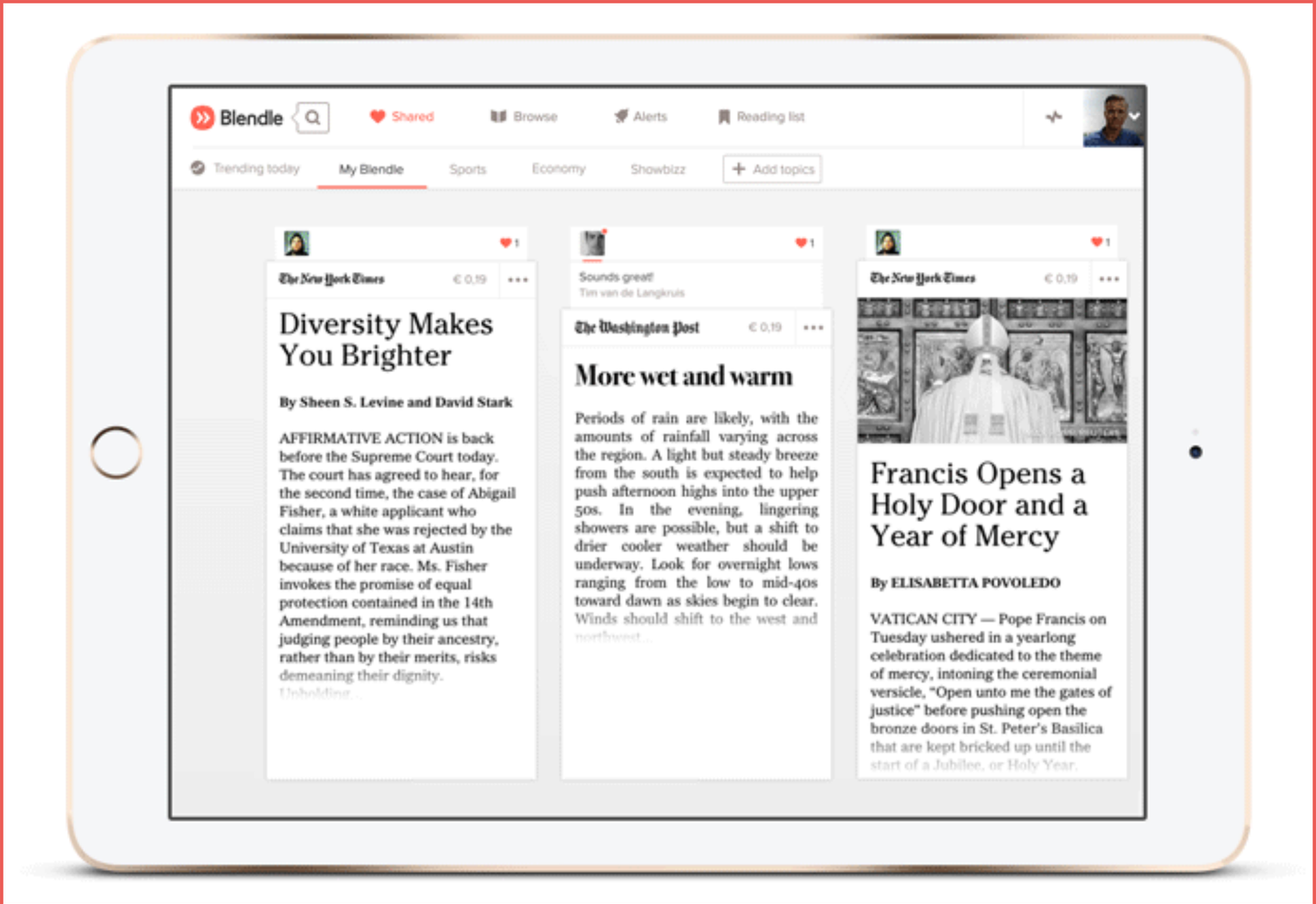
Niels Stevens
Software Engineer at Blendle

- What is Blendle?
- Our journey towards Kubernetes
- Lessons learned
- What's next?



Blendle @stvnsn niels-s

What is Blendle?

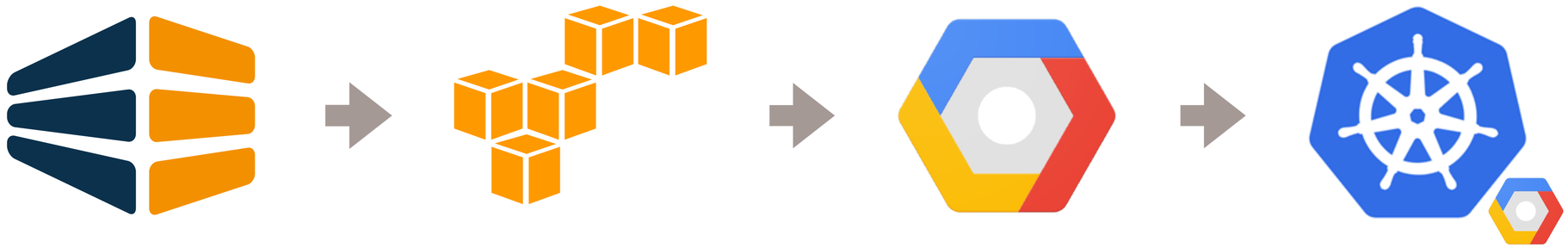


Mission: "Help you discover and support the worlds best journalism"

Blendle @stvnsn niels-s

Our infrastructure history

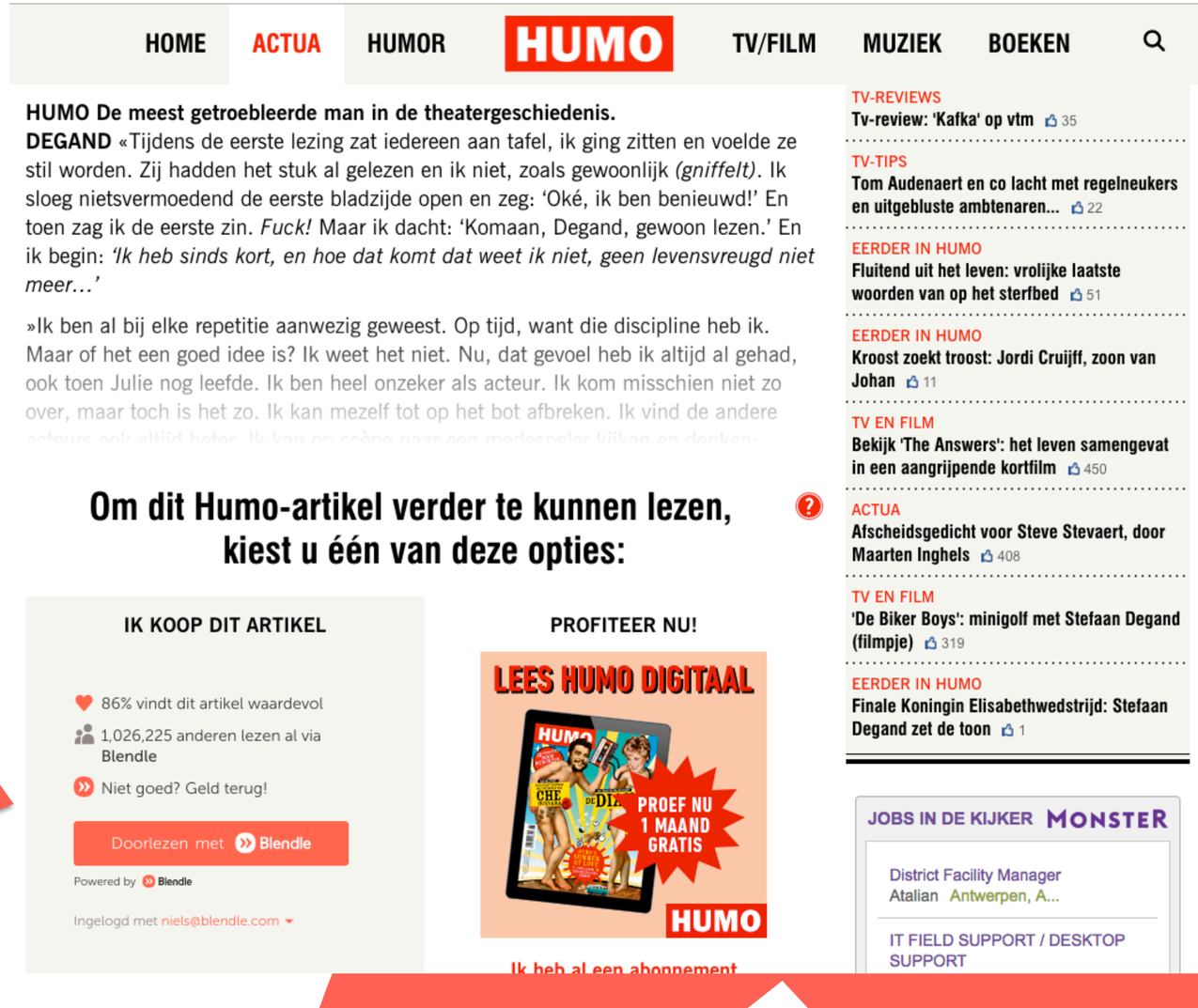
- Started out from bare metal
- Moved over to AWS
- Lift/shift to GCP
- Move “everything” to GKE (k8s)



When we started with k8s

- We launched a separate product
- New API & Client
- Kubernetes version 0.9

This button!



Blendle @stvnsn niels-s

Our Current Setup

- 3 clusters, hosted by GKE
  - “old” production cluster, 1.6.4, 16 nodes
  - “new” production cluster, 1.7.5, 5 nodes
  - ci cluster, 1.6.9, 5 nodes

Some numbers

```
$ kubectl get namespaces | wc -l
    109

$ kubectl get ingress --all-namespaces | wc -l
    69

$ kubectl get services --all-namespaces | wc -l
    174

$ kubectl get deployments --all-namespaces | wc -l
    316

$ kubectl get pods --all-namespaces | wc -l
    843
```

* combined numbers of our 2 production clusters

Preemptible Nodes

They…

- ... are cheap » saves around 75% in $$$
- ... help us (force us) to make robust and versatile software

But needed to solve some issues

- improve resiliency of kube-dns
- increase replicas of pods >= 3
- use podAntiAffinity
- and nodeAffinity
- more smaller nodes

```
spec:
  affinity:
    nodeAffinity:
      requiredDuringSchedulingIgnoredDuringExecution:
        nodeSelectorTerms:
        - matchExpressions:
          - key: cloud.google.com/gke-preemptible
            operator: DoesNotExist
    podAntiAffinity:
      preferredDuringSchedulingIgnoredDuringExecution:
        - podAffinityTerm:
          labelSelector:
            matchExpressions:
            - key: processor
              operator: In
              values:
              - select-users
          topologyKey: kubernetes.io/hostname
        weight: 100
```

Liveness / Readiness Probes

- Don’t use “empty” health checks
- If you project has dependencies, check them!
- Don’t limit the probes to api /web facing services

```
readinessProbe:
  exec:
    command:
    - /bin/health-check
    - --quiet
    - --readiness
  initialDelaySeconds: 5
  periodSeconds: 15
livenessProbe:
  exec:
    command:
    - /bin/health-check
    - --quiet
  initialDelaySeconds: 15
  periodSeconds: 15
```

Deployments

- CI?
  - » Our own Jenkins setup
  - » “previously” on Travis / Wercker
- github/scripts-to-rule-them-all
  - » script/test script/deploy
- Secrets fetched from password store on every deploy
- K8s Resources stored on Github » every change via a pull-request

Namespaces

```
$ kubectl get namespaces | wc -l
      109
```

- Split between production, staging, approval, ….

  ex: core-api-production, experiment-api-staging
- A namespace per project or service
- experimenting with RBAC

github.com/blendle/kns

#protip: Get a tool to switch between namespaces!

Templating language

- First used simple Environment Variable substitution
- https://github.com/blendle/epp

```
{% if RACK_ENV == "production" %}
---
apiVersion: autoscaling/v1
kind: HorizontalPodAutoscaler
metadata:
  name: api
  namespace: {{ KUBERNETES_NAMESPACE }}
spec:
  scaleTargetRef:
    kind: Deployment
    name: api
  minReplicas: {{ API_REPLICA_COUNT }}
  maxReplicas: {{ API_REPLICA_COUNT * 2 }}
  targetCPUUtilizationPercentage: 50
{% endif %}
```

```
---
kind: Secret
apiVersion: v1
metadata:
  name: example
data:
  example: "{{ example | b64enc }}"
```

Some small stuff

- Configure you resource request and limits!
    - LimitRange
    - Resource Quotas
- Configure a RevisionHistoryLimit
- Structured logging with StackDriver
- Please don’t ever use latest tag for your images!

```
resources:
  requests:
    cpu: 200m
    memory: 1Gi
  limits:
    cpu: 2
    memory: 5Gi
```

Cattle Shepherd

```yaml
kind: Deployment
metadata:
  annotations:
    blendle-metadata: |
      {
        "repository": "https://github.com/blendle/example",
        "contacts": ["[redacted]@blendle.com","[redacted]@blendle.com"]
      }
```

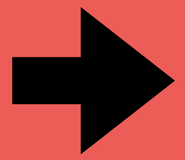
Name

	A (Niels Stevens)	B	C	D
1	**Name**	**Namespace**	**Repository**	**Contacts**
2	sidekiq-web	core-vodafone-prod	No blendle-metadata found	
3	cron	vodafone-resetter	No blendle-metadata found	
4	worker	blendle-push-development	https://github.com/blendle/blendle-push	[redacted]@blendle.com
5	cron	convert-workflow-production	https://github.com/blendle/convert-workflow	[redacted]@blendle.com,[redacted]@blendle.com
6	api	core-approval	https://github.com/blendle/core-api	[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com
7	sidekiq	subscription-production	https://github.com/blendle/blendle-subscription	[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com
8	stream-processor-alerts	stream-processors-staging	https://github.com/blendle/blendle-streaming	[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com
9	resque	core-approval	https://github.com/blendle/core-api	[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com
10	stream-processor-item-features	stream-processors-production	https://github.com/blendle/blendle-streaming	[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com
11	api	core-production	https://github.com/blendle/core-api	[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com
12	cron	core-development	https://github.com/blendle/core-api	[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com,[redacted]@blendle.com

Blendle @stvnsn niels-s

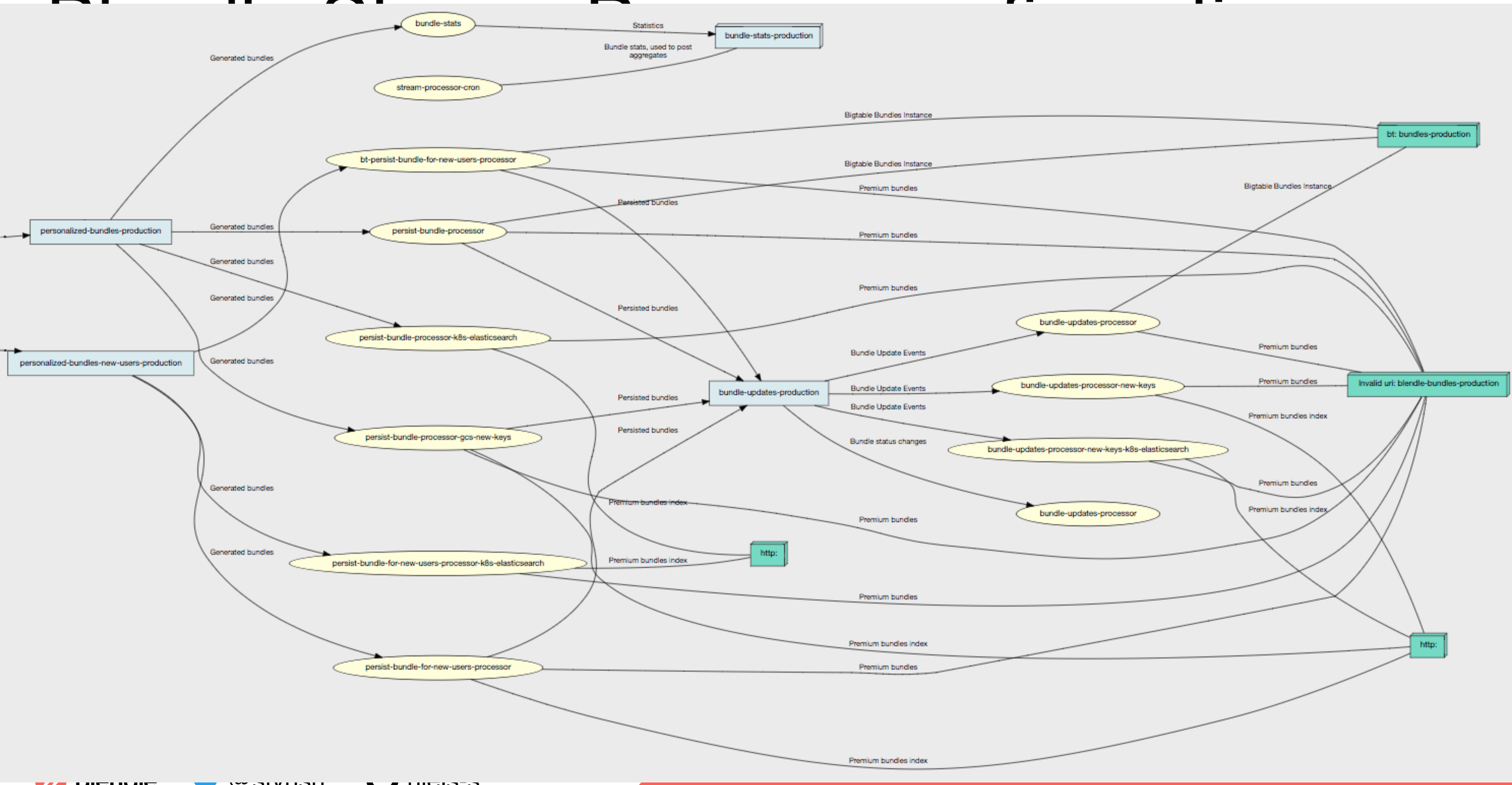
Kubecrt

```yaml
apiVersion: v1
charts:
- blendle/web:
    version: ~> 0.8.0
    values:
      public: false
      internalLoadBalancer: true
      contacts:
      - [redacted]@blendle.com
      - [redacted]@blendle.com
      imageVersion: "{{ env "GIT_COMMIT" }}"
      metrics: true
      envs:
        ACCESS_CONTROL_ALLOWED_ORIGINS: "{{ env "ACCESS_CONTROL_ALLOWED_ORIGINS" }}"
        APP_ENV: "{{ env "APP_ENV" }}"
        BUNDLE_SERVICE_URL: "{{ env "BUNDLE_SERVICE_URL" }}"
        GIT_COMMIT: "{{ env "GIT_COMMIT" }}"
      secrets:
        NEW_RELIC_LICENSE_KEY: "{{ env "NEW_RELIC_LICENSE_KEY" }}"
      resources:
        limits:
          cpu: 1
          memory: 512Mi
        requests:
          cpu: 500m
          memory: 256Mi
```



```yaml
kind: Deployment
metadata:
  annotations:
    blendle-metadata: |
      {
        "repository": "https://github.com/blendle/page-api-v2",
        "contacts": ["[redacted]@blendle.com","[redacted]@blendle.com"]
      }
  labels:
    app: page-api-v2
    component: web
  name: web
spec:
  replicas: 3
  revisionHistoryLimit: 3
  selector:
    matchLabels:
      app: page-api-v2
      component: web
  template:
    metadata:
      annotations:
        prometheus.io/scrape: "true"
      creationTimestamp: null
      labels:
        app: page-api-v2
        component: web
    spec:
      containers:
      - env:
        - name: APP_NAME
          value: page-api-v2
        - name: APP_COMPONENT
          value: web
        envFrom:
        - configMapRef:
            name: web
        - secretRef:
            name: web
        image: eu.gcr.io/bnl-blendle/page-api-v2:892d971
        imagePullPolicy: IfNotPresent
        name: page-api-v2-web
        ports:
        - containerPort: 80
          protocol: TCP
        readinessProbe:
          failureThreshold: 3
          httpGet:
            path: /health
            port: 80
            scheme: HTTP
          periodSeconds: 10
          successThreshold: 1
          timeoutSeconds: 1
        resources:
          limits:
            cpu: "1"
            memory: 512Mi
          requests:
            cpu: 500m
            memory: 256Mi
```

Blendle @stvnsn niels-s


bundle-stats
Statistics
bundle-stats-production
Generated bundles
Bundle stats, used to post aggregates
stream-processor-cron
Bigtable Bundles Instance
bt: bundles-production
bt-persist-bundle-for-new-users-processor
Bigtable Bundles Instance
Premium bundles
Bigtable Bundles Instance
Persisted bundles
personalized-bundles-production
Generated bundles
persist-bundle-processor
Premium bundles
Generated bundles
Premium bundles
Generated bundles
Persisted bundles
persist-bundle-processor-k8s-elasticsearch
bundle-updates-processor
personalized-bundles-new-users-production
Generated bundles
Bundle Update Events
Premium bundles
bundle-updates-production
Bundle Update Events
bundle-updates-processor-new-keys
Premium bundles
Invalid uri: blendle-bundles-production
Persisted bundles
Bundle Update Events
Premium bundles index
Persisted bundles
persist-bundle-processor-gcs-new-keys
Bundle status changes
bundle-updates-processor-new-keys-k8s-elasticsearch
Premium bundles
Generated bundles
Premium bundles index
Premium bundles
Premium bundles index
bundle-updates-processor
Generated bundles
http:
persist-bundle-for-new-users-processor-k8s-elasticsearch
Premium bundles index
Premium bundles
Premium bundles index
http:
persist-bundle-for-new-users-processor
Premium bundles
Premium bundles index

What’s next?

- External Admission Webhooks / CustomResourceDefinitions
- Istio
- Prometheus
- Cron Jobs (beta 1.8 🙏)
- Network policies
- HorizontalPodAutoscaling with new Metrics API
- Chatops/ GH Bot to setup PR to test namespace

Thanks!

niels@blendle.com

namespace switcher: github.com/blendle/kns

kubecrt: github.com/blendle/kubecrt

epp: github.com/blendle/epp

Blendle @stvnsn niels-s