

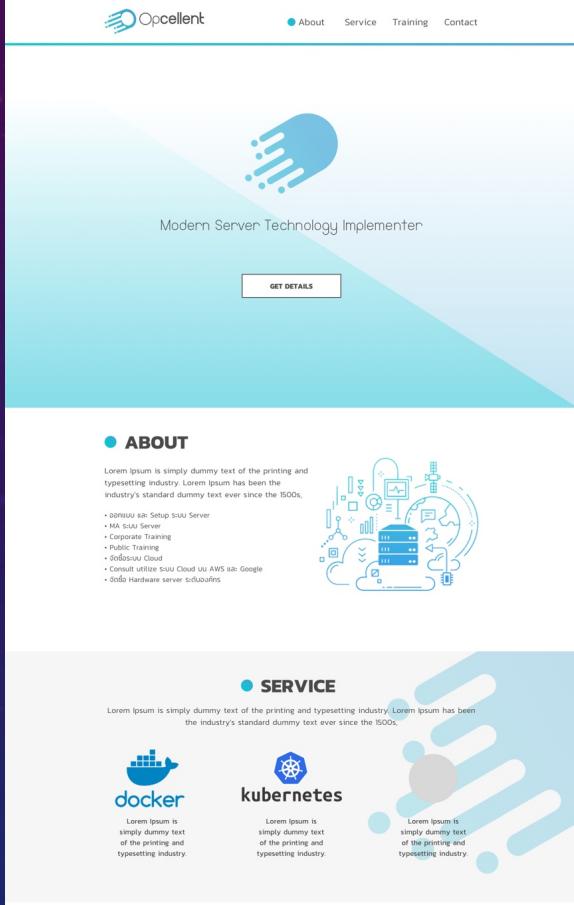
K8S Storage Built-In Manage Storage Cluster for real-world

By Praparn Luengphoonlap
Email: praparn@opcellent.com

AGENDA

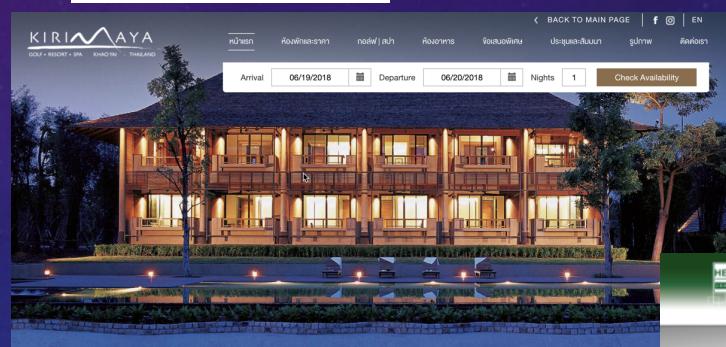
- Fantastic disk and where to find them
- Kubernetes for all (All-in-one solution)
- Case study: Push message processing
 - Solution design
 - Enhancement
 - Future improve
- Demo Case: Wordpress on scale
- Q&A

WHO ARE WE ? (OPCELLENT)



The website for Opcellent features a teal header with the company logo and navigation links for About, Service, Training, and Contact. Below the header is a large blue graphic element resembling a stylized flame or data stream. The main content area includes a section titled "Modern Server Technology Implementer" with a "GET DETAILS" button, an "ABOUT" section with a brief description and a cloud icon, and a "SERVICE" section featuring icons for Docker and Kubernetes.

Present by: Praparn L. (eva10409@gmail.com)



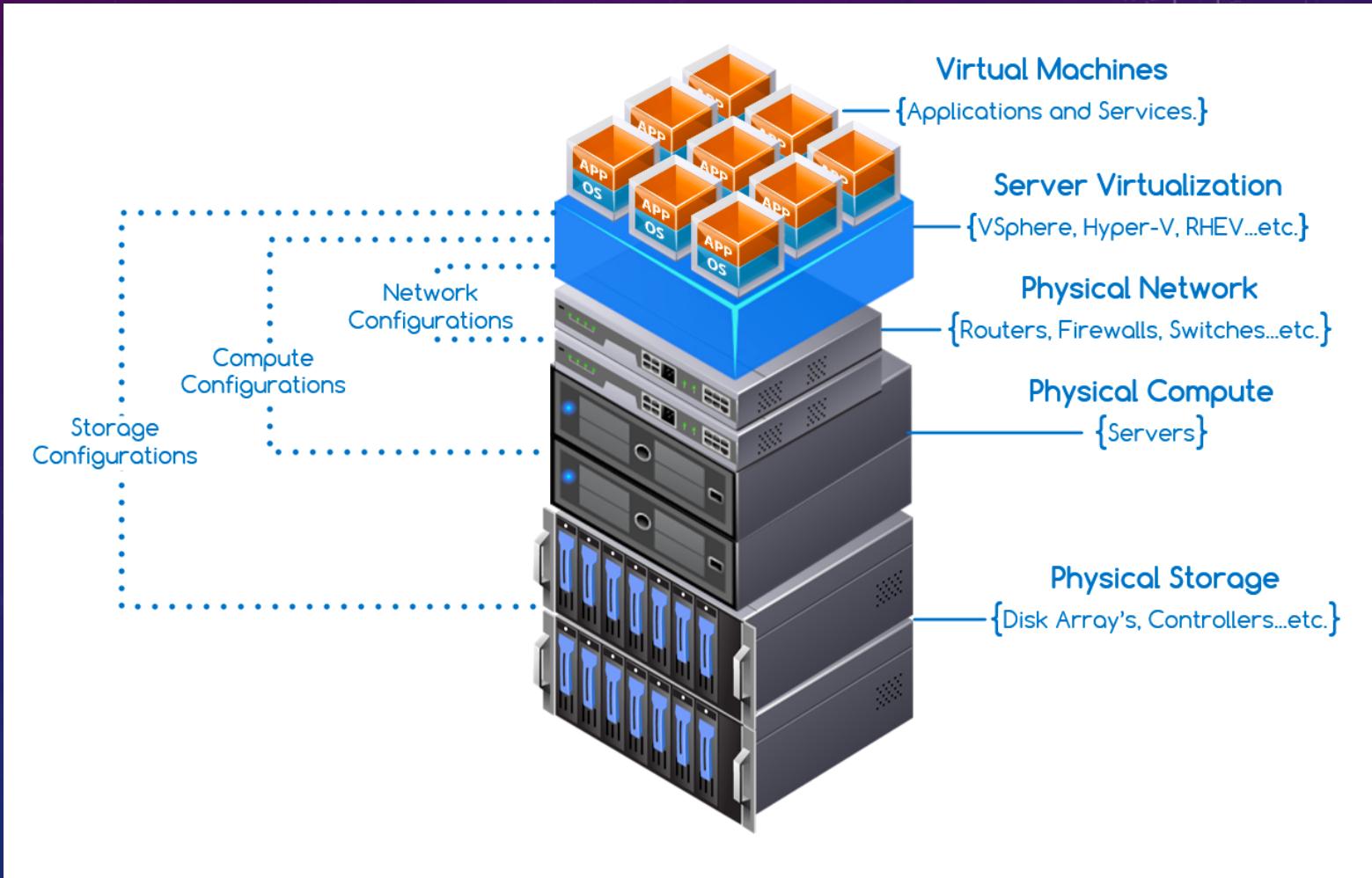
FANTASTIC DISK AND WHERE TO FIND THEM

- Find disk (Storage) for enterprise is like find monster
- Storage consist many in product
 - IBM V9000, V7000, V5000
 - HP 3PAR Storeserve
 - Dell EMC
 - NFS Storage etc
- Each product need special skill
- Discussion ?
 - How many disk space ?
 - Is it need redundancy disk ?
 - What the characteristic of storage
 - How about disk zoning ?
 - Ahhhh!!!



FANTASTIC DISK AND WHERE TO FIND THEM

- How about other component ?



FANTASTIC DISK AND WHERE TO FIND THEM

Imaginary that ...

- How many staff need to join?
- How many skill be need ?
- How about complexity be here
- ...
- ...
- Is it necessary to like there ?



Kubernetes for All



KUBERNETES FOR ALL

- Kubernetes is orchestrator for control container running in multiple server like cluster system
- Kubernetes will short name with “K8S”
- K8S will provide all component in single platform

VMWare / Physical Server	Kubernetes
Application/Database install	Pods (Set of container)
Server / Guest OS	No need
Networking	SDN (Calico, Flennel etc)
Load Balance	Service (native,Istio,traefik etc)
Firewall	Service (Calico, etc)
Storage	All you need 😊

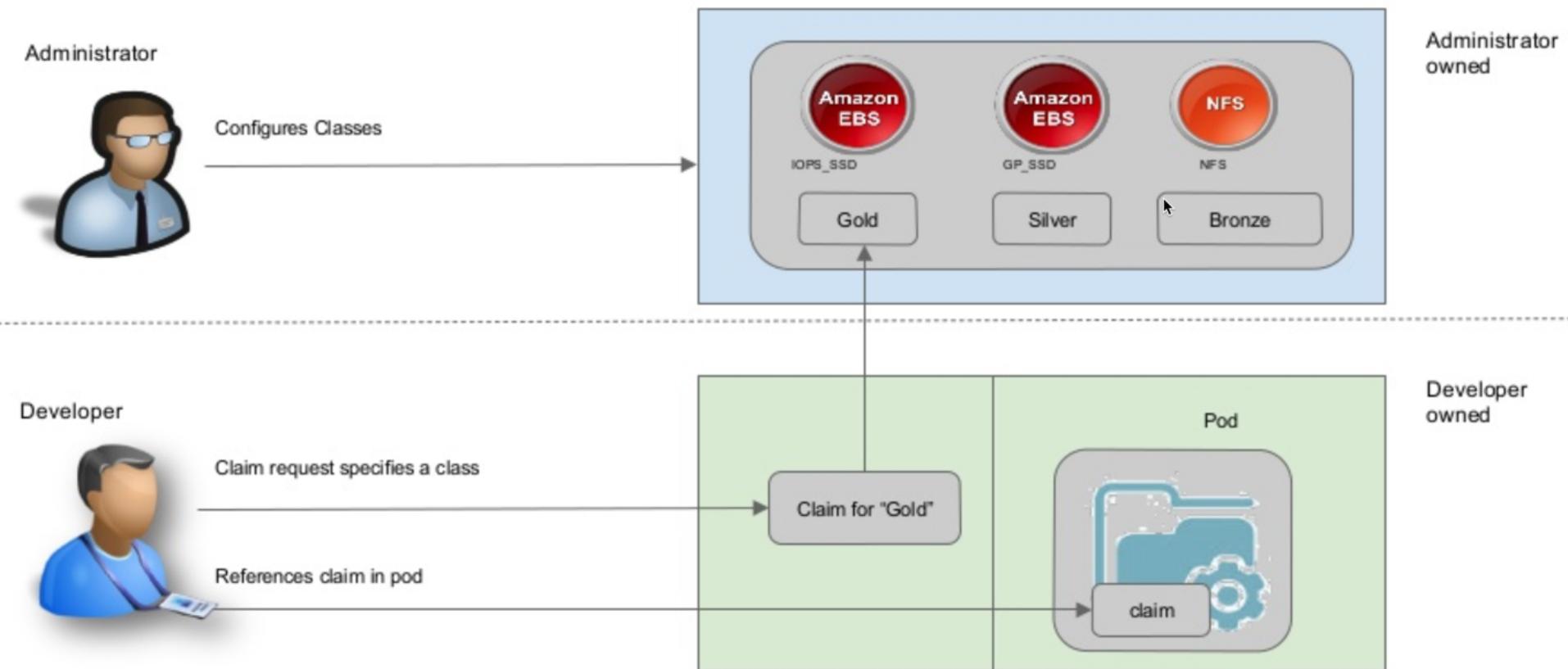
K8S Storage Built-In

KUBERNETES FOR ALL

- Kubernetes support storage in many solution upon our request
 - Dynamic Provision by “StorageClass”
 - Create storage real-time when application request

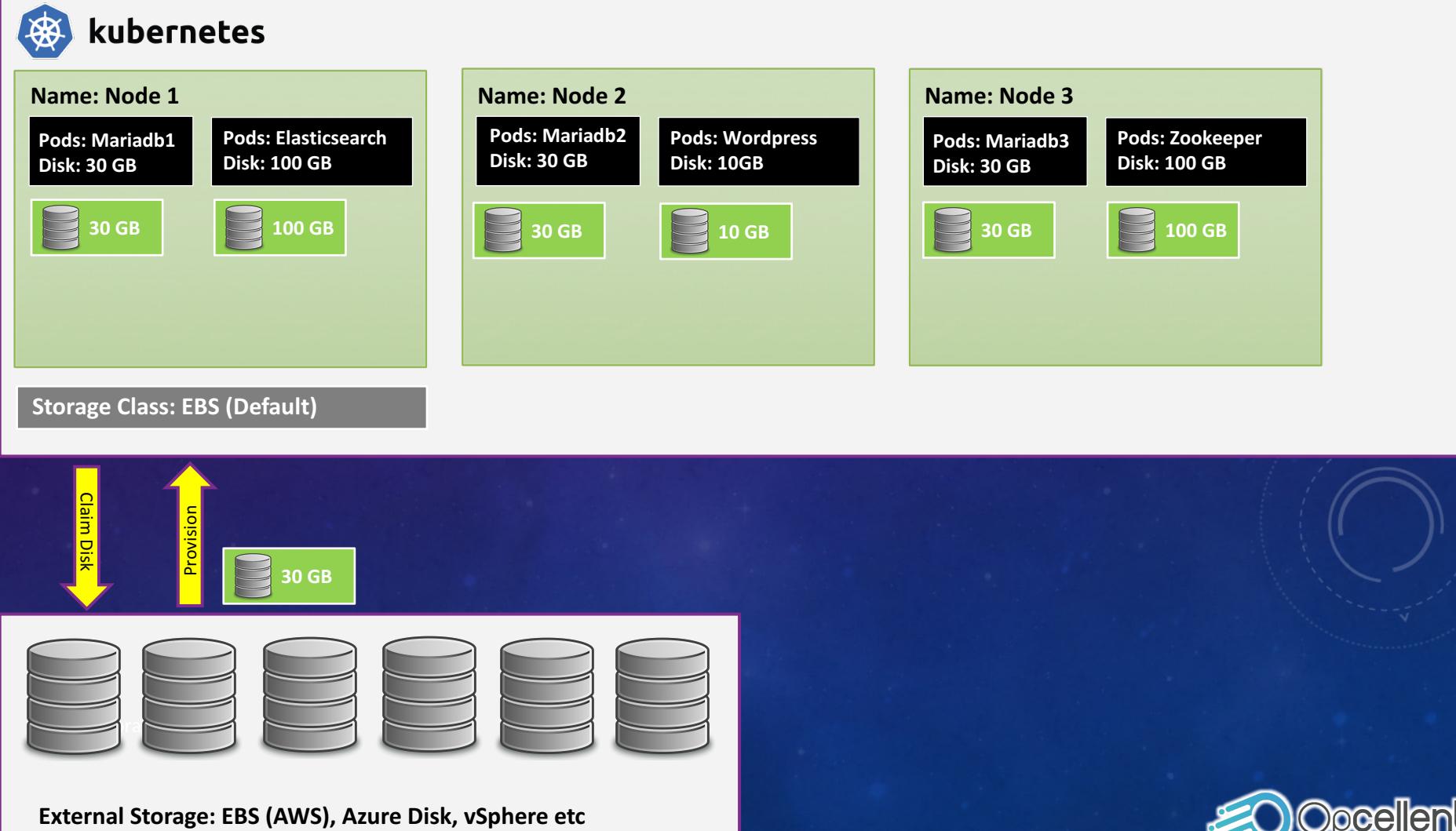
Volume Plugin	Internal Provisioner	Config Example
AWSElasticBlockStore	✓	AWS
AzureFile	✓	Azure File
AzureDisk	✓	Azure Disk
CephFS	-	-
Cinder	✓	OpenStack Cinder
FC	-	-
FlexVolume	-	-
Flocker	✓	-
GCEPersistentDisk	✓	GCE
Glusterfs	✓	Glusterfs
iSCSI	-	-
Quobyte	✓	Quobyte
NFS	-	-
RBD	✓	Ceph RBD
VsphereVolume	✓	vSphere
PortworxVolume	✓	Portworx Volume
ScaleIO	✓	ScaleIO
StorageOS	✓	StorageOS
Local	-	Local

KUBERNETES FOR ALL



KUBERNETES FOR ALL

- External Storage: Kubernetes will request storage from outside

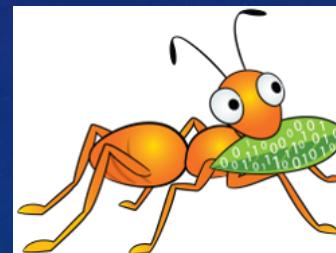


KUBERNETES FOR ALL

- Storage Built-In: Kubernetes will manage disk with SDS (software define storage)
- Manage end-to-end with single platform
 - Manage raw disk to be storage system
 - Provision storage on application demand
 - Manage storage network to application (End-Point)
 - No overhead network outside kubernetes cluster
 - Total open-source solution



K8S Storage Built-In



KUBERNETES FOR ALL



kubernetes

Name: Node 1

Pods: Mariadb1
Disk: 30 GB



Pods: Elasticsearch
Disk: 100 GB



Name: Node 2

Pods: Mariadb2
Disk: 30 GB



Pods: Wordpress1
Disk: 10GB (Share)

Pods: Wordpress2
Disk: 10GB (Share)

Pods: Wordpress3
Disk: 10GB (Share)



Deployment

Name: Node 3

Pods: Mariadb3
Disk: 30 GB



Pods: Zookeeper
Disk: 100 GB



Storage Class: SDS

Claim Storage

Provision

Software Define Storage (SDS) on Kubernetes

Pods: Manage Storage

Pods: Manage Storage

Pods: Manage Storage



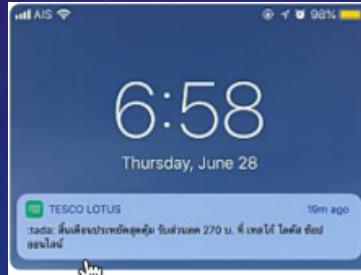
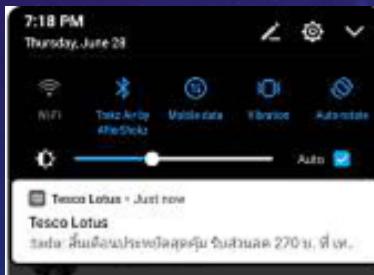
K8S Storage Options

Case Study: Push Message Processing



PUSH MESSAGING PROCESSING

- Business Purpose
 - Operate push message notification to all mobile's customer
 - Message is about 50,000 – 1M message each round
 - Push notification need to keep logging for cross-check message
 - System need to handling all message

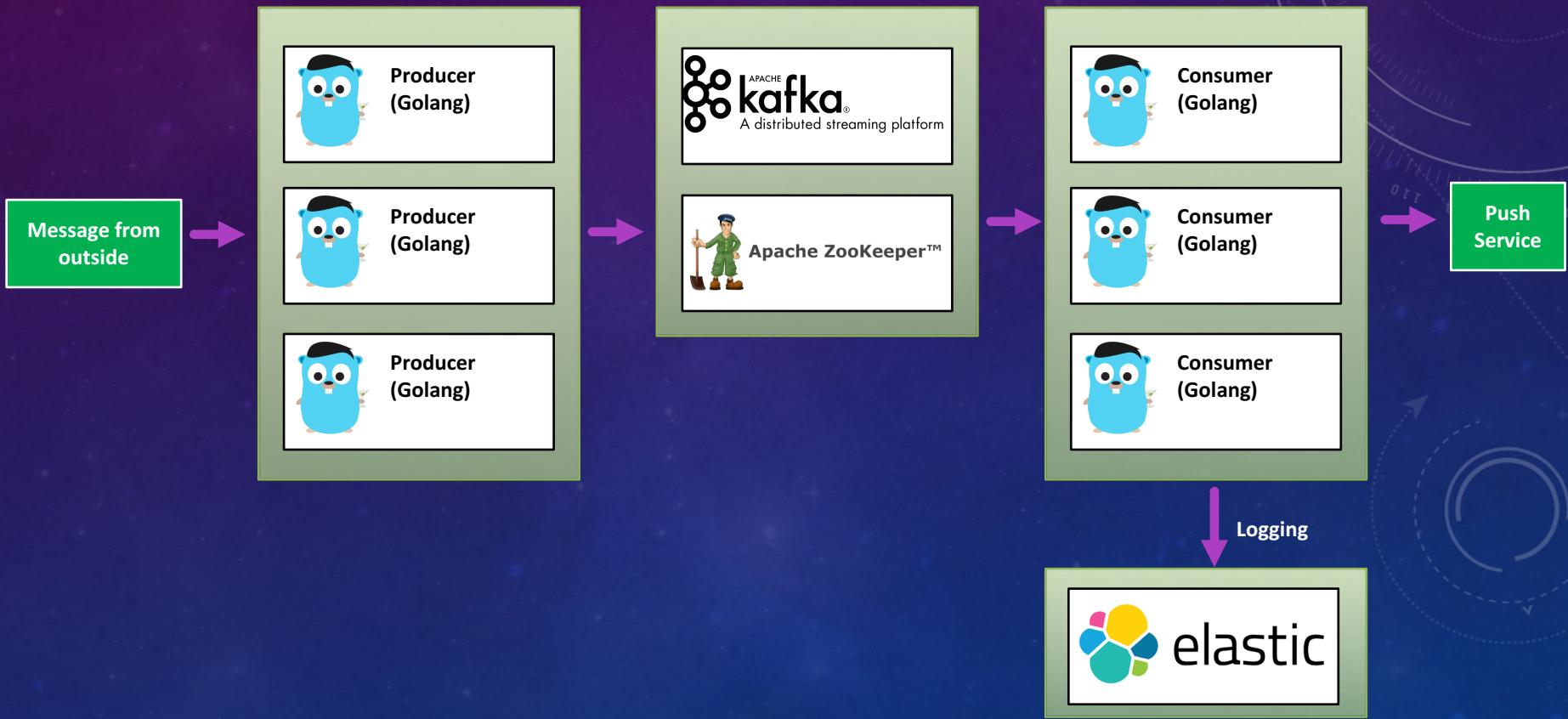


PUSH MESSAGING PROCESSING

- Solution Design
 - We're select to handling message queue via "Apache Kafka" and keep data on zookeeper
 - Producer/Consumer was developed base on "golang"
 - Output will send to external push notification service
 - Logging will keep to elasticsearch
 - Solution was design base on "CNCF Landscape" for operate this system anywhere (On-prem/On-Cloud)

PUSH MESSAGING PROCESSING

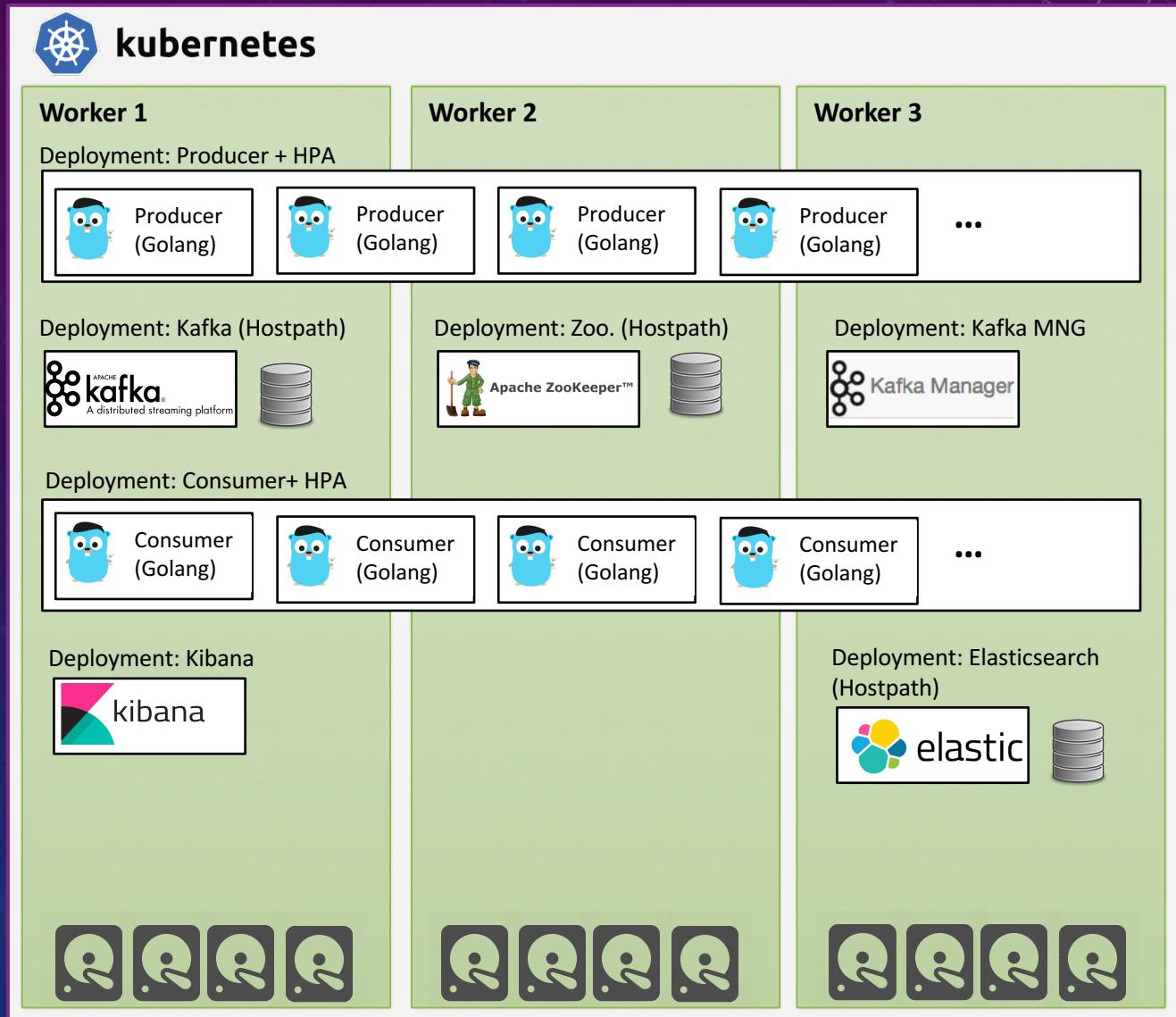
- Solution Design



K8S Storage Built-In

PUSH MESSAGING PROCESSING

Version 1.0



PUSH MESSAGING PROCESSING

- Enhancement
 - Change “Hostpath” to “StorageClass” for dynamic provision storage.
 - For avoid dependency component. We will provide storage by kubernetes itself.
 - We choose “GlusterFS” to provide storage system from 3 worker node with standard
 - All GlusterFS need “heketi” for provide RESTful volume management interface

PUSH MESSAGING PROCESSING

- Enhancement
 - <https://github.com/gluster/gluster-kubernetes>
 - GlusterFS official project “gluster-kubernetes”
 - Component
 - Kubernetes Farm (With RAW disk)
 - GlusterFS (Pods)
 - Heketi (Pods)
 - Control disk layout by define “topology.json”

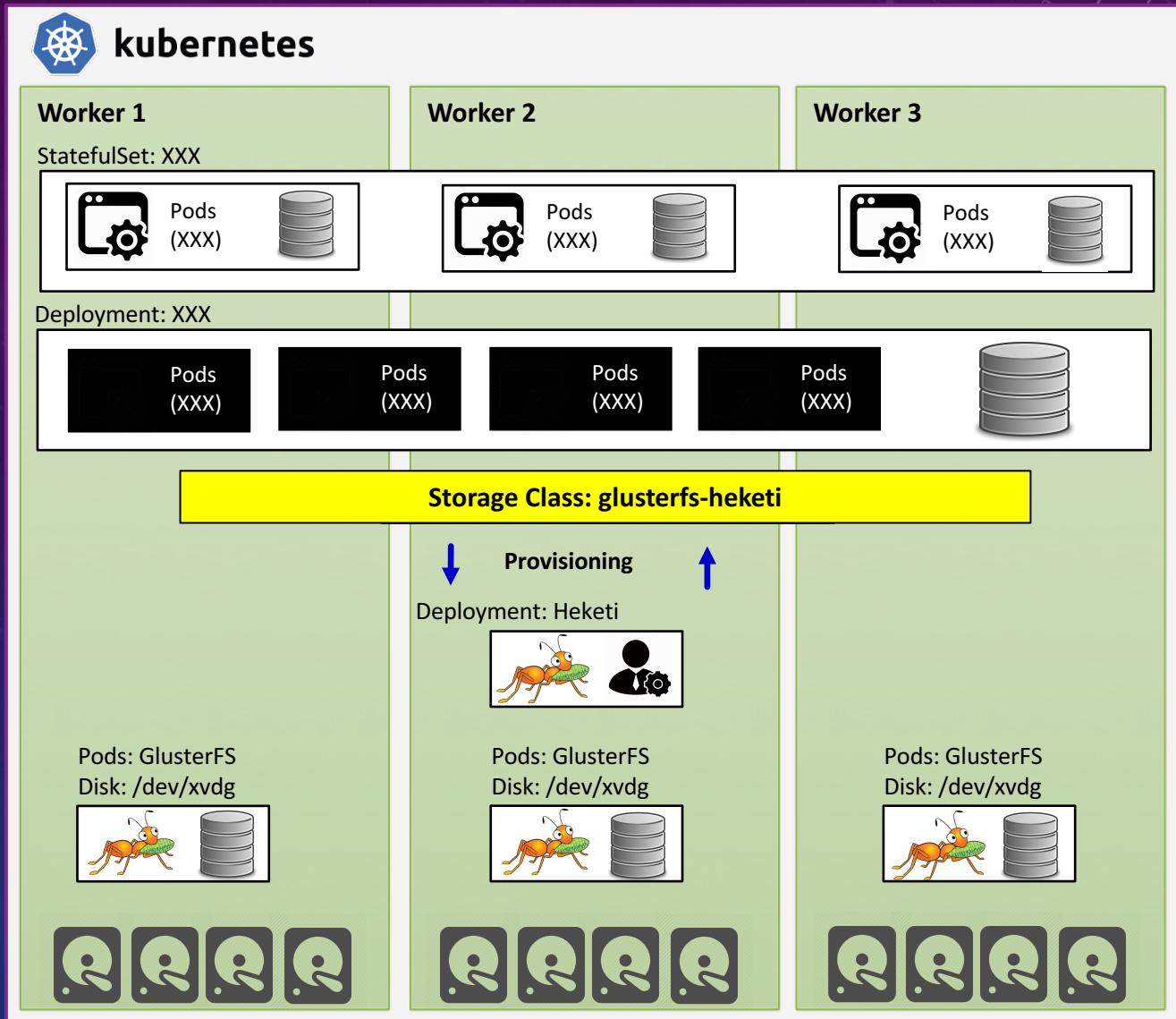
PUSH MESSAGING PROCESSING

```
Shortnote    {} topology.json x  ⌂ instructions  
[  
  "clusters": [  
    {  
      "nodes": [  
        {  
          "node": {  
            "hostnames": {  
              "manage": [  
                "ip-10-21-1-164"  
              ],  
              "storage": [  
                "10.21.1.164"  
              ]  
            },  
            "zone": 1  
          },  
          "devices": [  
            "/dev/xvddg"  
          ]  
        },  
        {  
          "node": {  
            "hostnames": {  
              "manage": [  
                "ip-10-21-1-213"  
              ],  
              "storage": [  
                "10.21.1.213"  
              ]  
            },  
            "zone": 1  
          },  
          "devices": [  
            "/dev/xvddg"  
          ]  
        }  
      ]  
    ]  
  ]  
]
```

```
37   {  
38     "node": [  
39       "hostnames": {  
40         "manage": [  
41           "ip-10-21-1-97"  
42         ],  
43         "storage": [  
44           "10.21.1.97"  
45         ]  
46       },  
47       "zone": 1  
48     ],  
49     "devices": [  
50       "/dev/xvddg"  
51     ]  
52   }  
53 ]  
54 }  
55 ]  
56 }
```

K8S

PUSH MESSAGING PROCESSING



PUSH MESSAGING PROCESSING

```
ubuntu@ip-10-21-1-32:~/gluster-kubernetes/deploy$ kubectl get pods
```

NAME	READY	STATUS	RESTARTS	AGE
glusterfs-78nls	1/1	Running	0	4m
glusterfs-phn98	1/1	Running	0	4m
glusterfs-rwqtt	1/1	Running	0	4m
heketi-86f98754c-xmf2	1/1	Running	0	1m

```
ubuntu@ip-10-21-1-32:~/gluster-kubernetes/deploy$ kubectl get svc
```

NAME	TYPE	CLUSTER-IP	EXTERNAL-IP	PORT(S)	AGE
heketi	ClusterIP	10.96.128.6	<none>	8080/TCP	1m
heketi-storage-endpoints	ClusterIP	10.98.108.0	<none>	1/TCP	2m
kubernetes	ClusterIP	10.96.0.1	<none>	443/TCP	2h

```
ubuntu@ip-10-21-1-32:~/gluster-kubernetes/deploy$ █
```

█

```
ubuntu@ip-10-21-1-32:~/DevOpsThailand2018_Storage_K8S$ kubectl create -f gluster-pvc.yaml  
persistentvolumeclaim/gluster-pvc created
```

```
ubuntu@ip-10-21-1-32:~/DevOpsThailand2018_Storage_K8S$ kubectl get pvc
```

NAME	STATUS	VOLUME	CAPACITY	ACCESS MODES	STORAGECLASS
LAST	AGE				
gluster-pvc	Bound	pvc-332c80fc-ac50-11e8-a961-028fc43bfd2c	5Gi	RWO	glusterfs-heketi
s-heketi	5s				

```
ubuntu@ip-10-21-1-32:~/DevOpsThailand2018_Storage_K8S$ kubectl get pv
```

NAME	STORAGECLASS	REASON	CAPACITY	ACCESS MODES	RECLAIM POLICY	STATUS	CLAIM
			AGE				
pvc-332c80fc-ac50-11e8-a961-028fc43bfd2c	glusterfs-heketi		5Gi	RWO	Delete	Bound	default
lt/gluster-pvc			8s				

```
ubuntu@ip-10-21-1-32:~/DevOpsThailand2018_Storage_K8S$ █
```

PUSH MESSAGING PROCESSING

```
ubuntu@ip-10-21-1-32:~/DevOpsThailand2018_Storage_K8S$ kubectl get statefulsets
NAME      DESIRED   CURRENT   AGE
galera-mariadb   3          3         11m

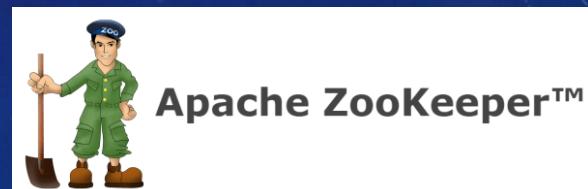
ubuntu@ip-10-21-1-32:~/DevOpsThailand2018_Storage_K8S$ kubectl get pvc
NAME           STATUS    VOLUME                                     CAPACITY   ACCE
SS MODES      STORAGECLASS   AGE
mysql-datadir-galera-mariadb-0   Bound    pvc-89d5a8dd-ac52-11e8-a961-028fc43bfd2c   10Gi      RWO
          glusterfs-heketi   11m
mysql-datadir-galera-mariadb-1   Bound    pvc-086b0341-ac53-11e8-a961-028fc43bfd2c   10Gi      RWO
          glusterfs-heketi   7m
mysql-datadir-galera-mariadb-2   Bound    pvc-89a42e0e-ac53-11e8-a961-028fc43bfd2c   10Gi      RWO
          glusterfs-heketi   3m

ubuntu@ip-10-21-1-32:~/DevOpsThailand2018_Storage_K8S$ kubectl get pv
NAME           CAPACITY   ACCESS MODES   RECLAIM POLICY   STATUS   CLAIM
STORAGECLASS   REASON     AGE
pvc-086b0341-ac53-11e8-a961-028fc43bfd2c   10Gi      RWO          Delete   Bound   defau
lt/mysql-datadir-galera-mariadb-1   glusterfs-heketi   7m
pvc-89a42e0e-ac53-11e8-a961-028fc43bfd2c   10Gi      RWO          Delete   Bound   defau
lt/mysql-datadir-galera-mariadb-2   glusterfs-heketi   3m
pvc-89d5a8dd-ac52-11e8-a961-028fc43bfd2c   10Gi      RWO          Delete   Bound   defau
lt/mysql-datadir-galera-mariadb-0   glusterfs-heketi   11m

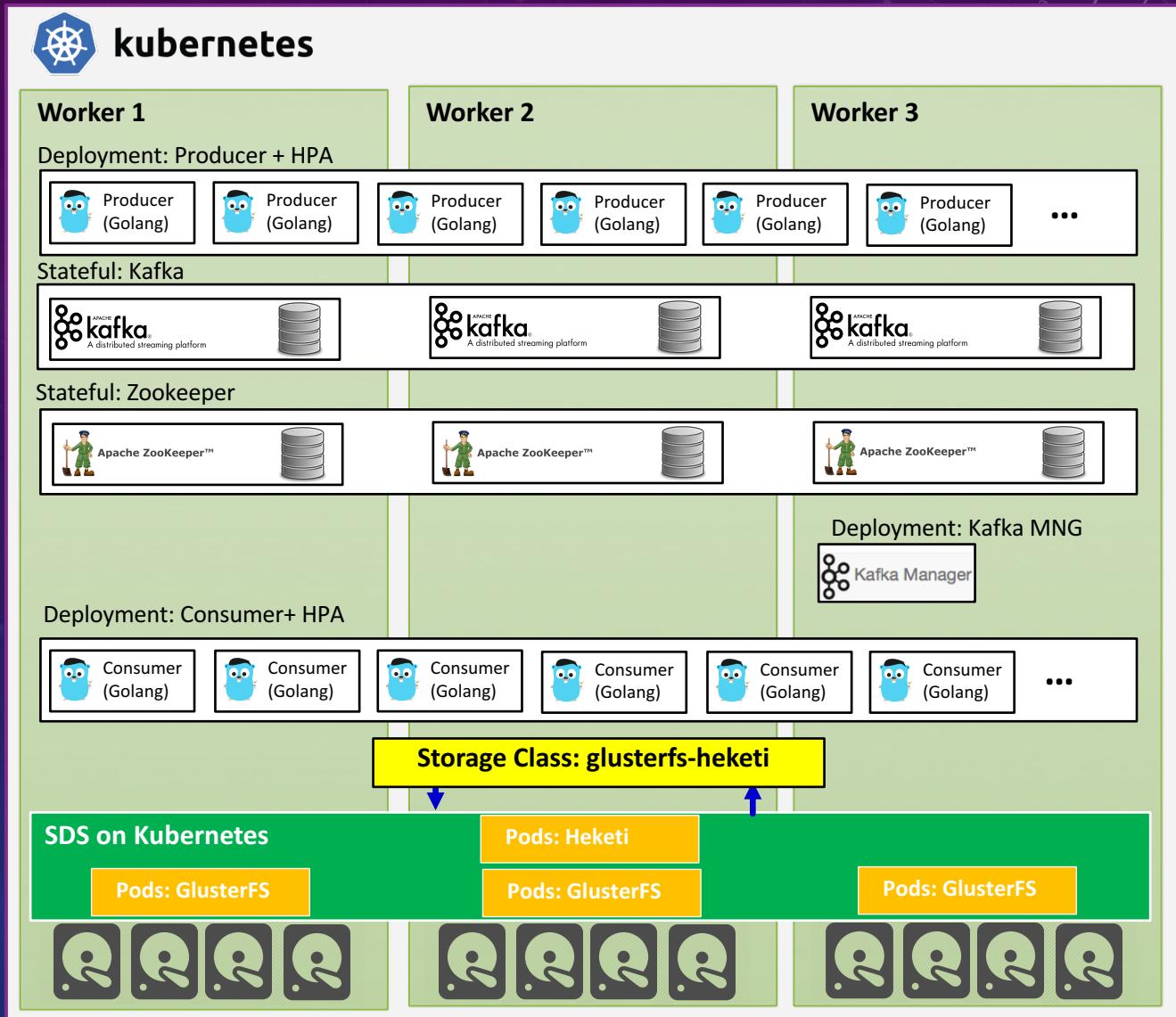
ubuntu@ip-10-21-1-32:~/DevOpsThailand2018_Storage_K8S$ █
```

PUSH MESSAGING PROCESSING

- Enhancement
 - Component:
 - Kafka create cluster component:
 - Kafka (Stateful) with 3 Replicas
 - Zookeeper (Stateful) with 3 Replicas
 - Kafka MNG (Deployment) with 1 Replica
 - Producer (Deployment) with 10 Replicas + HPA
 - Consumer (Deployment) with 6 Replicas + HPA



PUSH MESSAGING PROCESSING

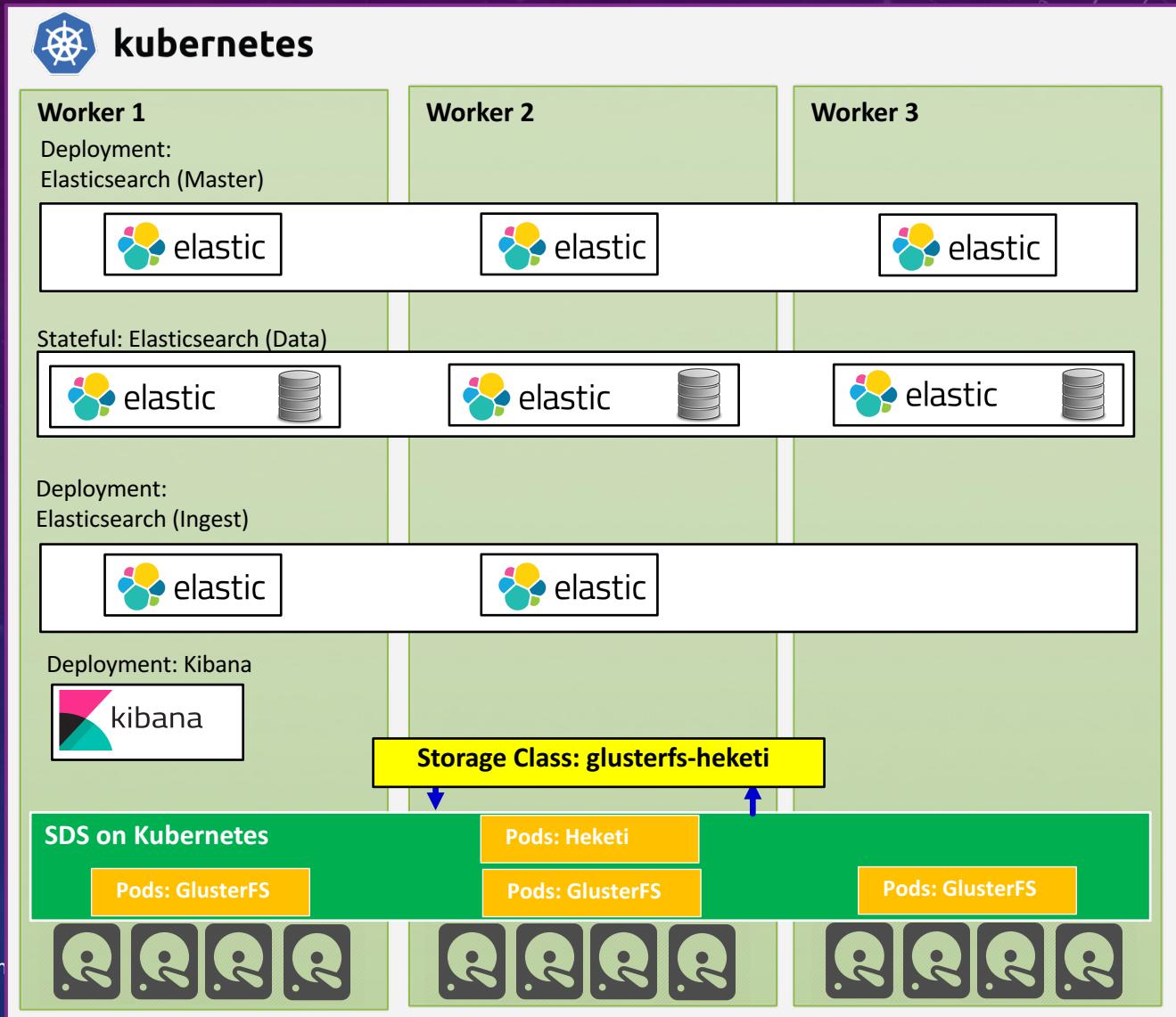


K8S Storage Built-In

PUSH MESSAGING PROCESSING

- Enhancement
 - Component:
 - Elasticsearch create cluster and separate role as best practice:
 - Master Node (Deployment) with 3 Replicas
 - Data Node (Stateful) with 3 Replicas
 - Ingest Node (Deployment) with 2 Replicas
 - Kibana (Deployment) with 1 Repica
 - Ref:<https://www.elastic.co/guide/en/elasticsearch/reference/6.2/modules-node.html>

PUSH MESSAGING PROCESSING



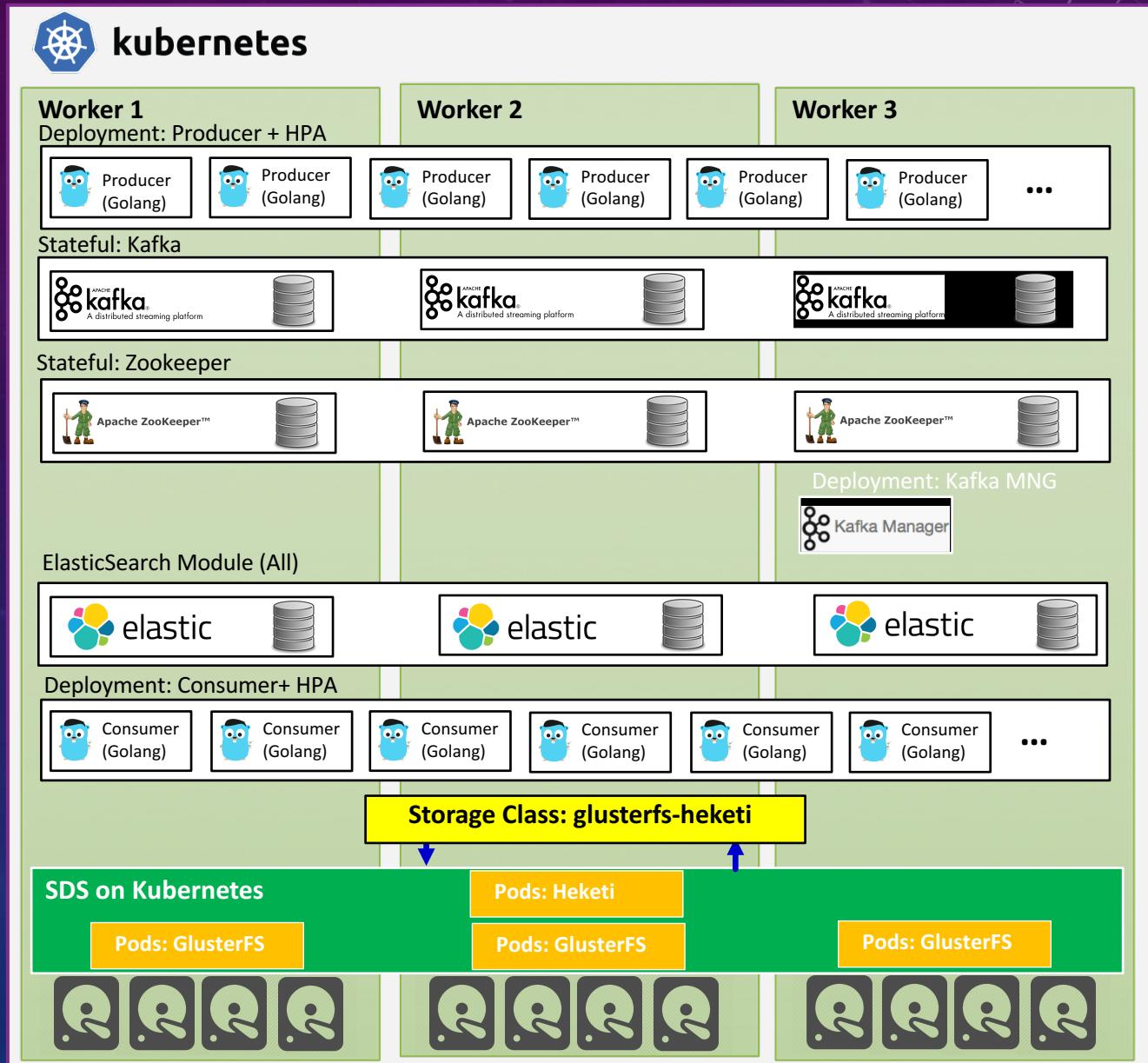
K8S Storage Built-In

PUSH MESSAGING PROCESSING

- Benefit for enhancement
 - Dedicate storage for each application by Storage Class
 - Remove all single point failure component
 - Compute
 - Network
 - Storage
 - Complete cluster for all component
 - Extend capability for support message queue

PUSH MESSAGING PROCESSING

Version 2.0



PUSH MESSAGING PROCESSING

- Future Improvement
 - Change storage provision to “Ceph RBD” (Block Storage)
 - Optional integrate with external storage on AzureDisk, AWS(EBS) etc
 - Enhance feature for online file system extension via kubernetes (“ExpandInUsePersistentVolumes”,Alpha feature on Version 1.11)

PUSH MESSAGING PROCESSING

```
~ — ubuntu@ip-10-21-1-193:~/gluster-kubernetes/deploy — zsh ... scluster/group_vars — ssh -i scblife-access.pem ubuntu@10.21.219.9 ... ubectl --kubeconfig ./adminconfig_dev.conf proxy --accept-hosts *  
## That can be useful for example to setup a keepalived virtual IP  
# supplementary_addresses_in_ssl_keys: [10.0.0.1, 10.0.0.2, 10.0.0.3]  
  
## Running on top of openstack vms with cinder enabled may lead to unschedulable pods due to NoVolumeZoneConflict restriction in kube-scheduler.  
## See https://github.com/kubernetes-incubator/kubespray/issues/2141  
## Set this variable to true to get rid of this issue  
volume_cross_zone_attachment: false  
[ubuntu@ip-10-21-219-9:~/kubespray/inventory/k8scluster/group_vars$ ceph osd status  
+-----+-----+-----+-----+-----+-----+-----+-----+  
| id | host | used | avail | wr ops | wr data | rd ops | rd data | state |  
+-----+-----+-----+-----+-----+-----+-----+-----+  
| 0 | scblf-dev-kubernetes-worker-1 | 1113 | 598 | | | | exists,up |  
| 1 | scblf-dev-kubernetes-worker-2 | 1096 | 598 | | | | exists,up |  
| 2 | scblf-dev-kubernetes-worker-3 | 1129 | 598 | | | | exists,up |  
+-----+-----+-----+-----+-----+-----+-----+-----+  
[ubuntu@ip-10-21-219-9:~/kubespray/inventory/k8scluster/group_vars$ ceph osd status  
+-----+-----+-----+-----+-----+-----+-----+-----+  
| id | host | used | avail | wr ops | wr data | rd ops | rd data | state |  
+-----+-----+-----+-----+-----+-----+-----+-----+  
| 0 | scblf-dev-kubernetes-worker-1 | 1113 | 598 | | | | exists,up |  
| 1 | scblf-dev-kubernetes-worker-2 | 1096 | 598 | | | | exists,up |  
| 2 | scblf-dev-kubernetes-worker-3 | 1129 | 598 | | | | exists,up |  
+-----+-----+-----+-----+-----+-----+-----+-----+  
ubuntu@ip-10-21-219-9:~/kubespray/inventory/k8scluster/group_vars$ []  
● ● ● ↑ praparnlueangphoonlap — ubuntu@scblf-dev-kubernetes-master-1: ~ — ssh -i scblife-access.pem ubuntu@10.21.219.59 — 193x24  
ubuntu@scblf-dev-kubernetes-master-1:~$ kubectl get nodes  
NAME STATUS ROLES AGE VERSION  
ip-10-21-219-10.ap-southeast-1.compute.internal Ready node 6h v1.11.2  
ip-10-21-219-18.ap-southeast-1.compute.internal Ready node 6h v1.11.2  
ip-10-21-219-24.ap-southeast-1.compute.internal Ready node 6h v1.11.2  
ip-10-21-219-35.ap-southeast-1.compute.internal Ready node 6h v1.11.2  
ip-10-21-219-41.ap-southeast-1.compute.internal Ready node 6h v1.11.2  
ip-10-21-219-43.ap-southeast-1.compute.internal Ready master 6h v1.11.2  
ip-10-21-219-48.ap-southeast-1.compute.internal Ready master 6h v1.11.2  
ip-10-21-219-57.ap-southeast-1.compute.internal Ready node 6h v1.11.2  
ip-10-21-219-59.ap-southeast-1.compute.internal Ready master 6h v1.11.2  
ubuntu@scblf-dev-kubernetes-master-1:~$ kubectl get pvc  
NAME STATUS VOLUME CAPACITY ACCESS MODES STORAGECLASS AGE  
aws-ebs-pvc Bound pvc-64664273-ab4f-11e8-a77a-06cd925c4206 5Gi RWO aws-ebs 5h  
rbd-pvc Bound pvc-eba0537b-ab72-11e8-a77a-06cd925c4206 20Gi RWO rbd-ceph 1h  
ubuntu@scblf-dev-kubernetes-master-1:~$ kubectl get pv  
NAME CAPACITY ACCESS MODES RECLAIM POLICY STATUS CLAIM STORAGECLASS REASON AGE  
pvc-64664273-ab4f-11e8-a77a-06cd925c4206 5Gi RWO Delete Bound default/aws-ebs-pvc aws-ebs 5h  
pvc-eba0537b-ab72-11e8-a77a-06cd925c4206 20Gi RWO Delete Bound default/rbd-pvc rbd-ceph 1h  
ubuntu@scblf-dev-kubernetes-master-1:~$ kubectl get pv  
NAME CAPACITY ACCESS MODES RECLAIM POLICY STATUS CLAIM STORAGECLASS REASON AGE  
pvc-64664273-ab4f-11e8-a77a-06cd925c4206 5Gi RWO Delete Bound default/aws-ebs-pvc aws-ebs 6h  
pvc-eba0537b-ab72-11e8-a77a-06cd925c4206 20Gi RWO Delete Bound default/rbd-pvc rbd-ceph 1h  
ubuntu@scblf-dev-kubernetes-master-1:~$ []  
K8S Storage
```

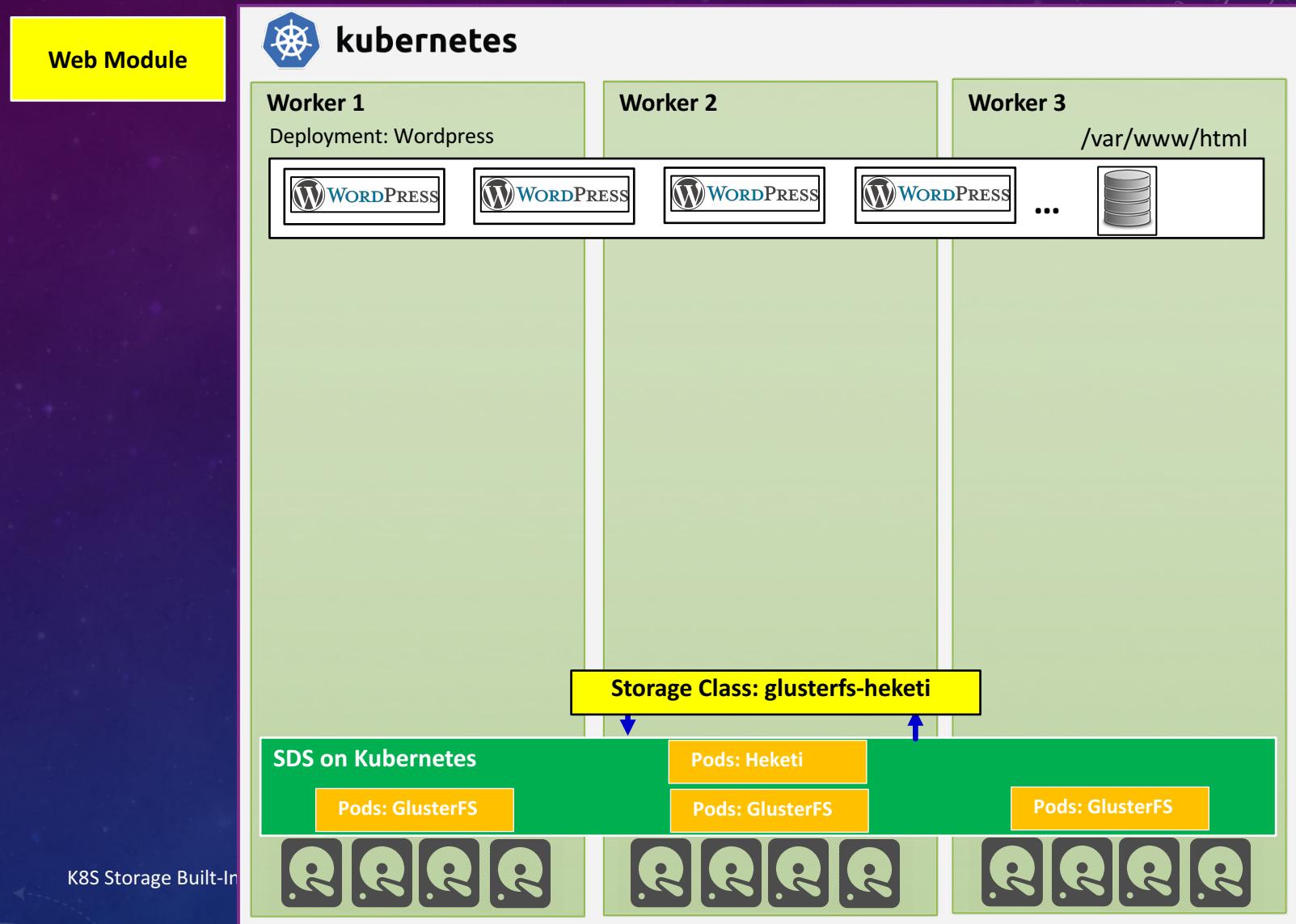
Demo: Wordpress on Scale



WORDPRESS ONSCALE

- Solution Design
 - Wordpress major container 2 components
 - Web Module: Apache + PHP (Last update also support NGINX)
 - Database Module: MariaDB / Mysql
 - How we scale it ?
 - Web Module: Will use deployment and share storage on path “/var/www/html” for all wordpress.
 - This will make all wordpress share session on same place and resolve problem of stateful at all

PUSH MESSAGING PROCESSING



WORDPRESS ONSCALE

- How we scale it ?
 - Database Module:
 - We select to use MariaDB with Galera Cluster
 - Provide Active/Active multi master technology
 - Provide 3 active database node (Read/Write)
 - Application read/write database normally
 - Keep state on “etcd” for all member of mariadb
 - wsrep_local_state_comment = “Synced”
 - wsrep_last_committed=<value>

WORDPRESS ONSCALE



Products Services Resources About Us Contact [Get Started](#) [Download](#)

Home » Resources » Knowledge Base » Library » MariaDB Documentation » High Availability & Performance Tuning » MariaDB Galera Cluster » What is MariaDB Galera Cluster?

What is MariaDB Galera Cluster?

The most recent release of MariaDB 10.3 is:
MariaDB 10.3.8 Stable (GA) [Download Now](#)

The most recent release of MariaDB 10.2 is:
MariaDB 10.2.16 Stable (GA) [Download Now](#)

The most recent release of MariaDB 10.1 is:
MariaDB 10.1.35 Stable (GA) [Download Now](#)

The most recent MariaDB Galera Cluster 10.0 release is:
MariaDB Galera Cluster 10.0.36 [Download Now](#)

The most recent MariaDB Galera Cluster 5.5 release is:
MariaDB Galera Cluster 5.5.61 [Download Now](#)

About

MariaDB Galera Cluster is a synchronous multi-master cluster for MariaDB. It is available on Linux only, and only supports the XtraDB/InnoDB storage engines (although there is experimental support for MyISAM - see the wsrep_replicate_myisam system variable).

Starting with MariaDB 10.1, the wsrep API for Galera Cluster is included by default. This is available as a separate download for MariaDB 10.0 and MariaDB 5.5.

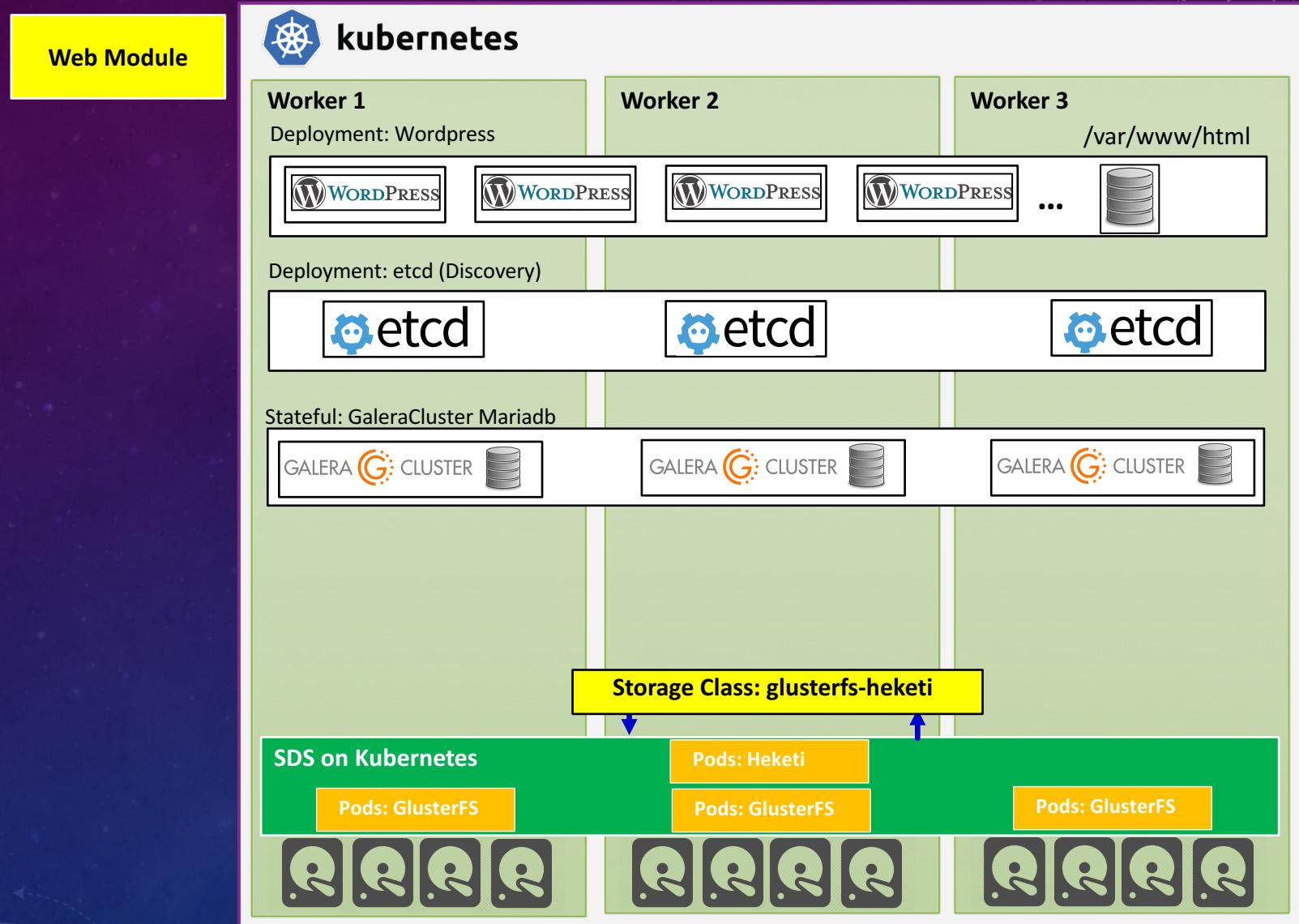
Features

Contents

1. About
2. Features
3. Benefits
4. Technology
 - 1. Galera wsrep provider Versions
5. See Also

↑ [MariaDB Galera Cluster](#)
↑ [What is MariaDB Galera Cluster?](#)
[About Galera Replication](#)
[Galera Use Cases](#)
[Getting Started with MariaDB Galera Cluster](#)
[MariaDB Galera Cluster - Known Limitations](#)
[Galera Cluster Status Variables](#)
[Galera Cluster System Variables](#)
[Building the Galera wsrep Package on Ubuntu and Debian](#)
[Building the Galera wsrep Package on Fedora](#)
[Installing Galera from Source](#)
[Galera Test Repositories](#)
[wsrep_provider_options](#)
[Galera Cluster Address](#)
[Galera Load Balancer](#)
[MariaDB Galera 10.0 Release Notes](#)
[MariaDB Galera 10.0 Changelogs](#)
[MariaDB Galera 5.5 Release Notes](#)
[MariaDB Galera 5.5 Changelogs](#)
[Tips on Converting to Galera](#)
[Upgrading from MariaDB Galera Cluster 10.0 to MariaDB 10.1](#)
[Upgrading from MariaDB Galera Cluster 5.5 to MariaDB Galera Cluster 10.0](#)

PUSH MESSAGING PROCESSING

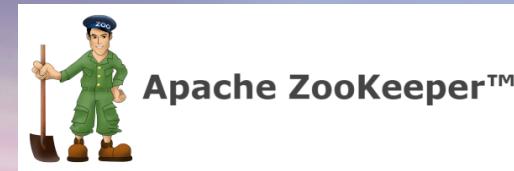


WORDPRESS ONSCALE

The image shows a desktop interface with several windows open:

- Terminal Window:** Shows command-line output for Kubernetes operations. It includes:
 - `kubectl get pods -o wide` listing various pods like `deploy-heketi-7c4898d9cd-99s7t`, `etcd0`, `etcd1`, `etcd2`, `galera-mariadb-0`, etc.
 - `kubectl delete pods/galera-mariadb-0` deleting the `galera-mariadb-0` pod.
 - `kubectl -n DevOpsThailand2018_Storage_K8S get svc` listing services.
- Kubernetes Dashboard:** A browser-based dashboard showing the status of StatefulSets, including `galera-mariadb`. It displays CPU and Memory usage for the pods.
- Code Editor:** A terminal window showing a script or log file with numbered lines. The script details the steps for deploying a Wordpress site, including:
 - Create StorageClass for Dynamic Provision.
 - Create Etcd Cluster for Discovery Service.
 - Create Galera Cluster MariaDB by command: (5 - 10 min)
 - Create Wordpress for Access and Get by command: (5 - 10 min)
 - Check Service by Command: `kubectl get svc` and tried to access via browser==> `http://<public ip address>:31000`
- WordPress Site Preview:** A browser window showing a test WordPress site with the title "TEST" and the subtext "Just another WordPress site".

Docker: The Next-Gen of Virtualization



Q&A (We are Hiring!)



By Praparn Luengphoonlap
Email: paparn@opcellent.com