PERCONA

MongoDB in the Cloud: How to Reduce Infrastructure Costs Using Open-Source Components and Kubernetes - Percona

Michał Nosek

Senior Enterprise Architect @ Percona

Agenda

I. Why costs are getting out of control?

II. K8s and Operators 5 min crash course

III. Practical approaches to cost reduction

IV. Summary

PERCONA

I. Why costs are getting out of control?

Cloud DBaaS convenience fee

No backup, no data transfer fees, yearly

Cloud and region	MongoDB Atlas	IaaS pay-as-you-go	IaaS reserved (3y)
M30 (3-node RS, 2vCPU, 8GB RAM, 40GB disk)			
AWS us-east1			
Azure centralus			
GCP us-east1			
M200 (4x3-node RS, 64vCPU, 256GB RAM, 2TB storage)			
AWS us-east1			
Azure centralus			
GCP us-east1			

Cloud DBaaS convenience fee

No backup, no data transfer fees, yearly

Cloud and region	MongoDB Atlas	IaaS pay-as-you-go	IaaS reserved (3y)
M30 (3-node RS, 2vCPU, 8GB RAM, 40GB disk)			
AWS us-east1	$4,730	$2,666	$1,204
Azure centralus	$5,256	$2,450	$1,049
GCP us-east1	$3,854	$2,796	$1,393
M200 (4x3-node RS, 64vCPU, 256GB RAM, 2TB storage)			
AWS us-east1	$555,822	$355,086	$159,674
Azure centralus	$626,515	$366,672	$162,575
GCP us-east1	$525,755	$379,376	$198,303

Cloud DBaaS convenience fee

No backup, no data transfer fees, yearly

Cloud and region	MongoDB Atlas	IaaS pay-as-you-go	IaaS reserved (3y)
M30 (3-node RS, 2vCPU, 8GB RAM, 40GB disk)			
AWS us-east1	**177%**	**$2,666**	$1,204
Azure centralus	**214%**	**$2,450**	$1,049
GCP us-east1	**137%**	**$2,796**	$1,393
M200 (4x3-node RS, 64vCPU, 256GB RAM, 2TB storage)			
AWS us-east1	**156%**	**$355,086**	$159,674
Azure centralus	**171%**	**$366,672**	$162,575
GCP us-east1	**138%**	**$379,376**	$198,303

PERCONA

Cloud DBaaS convenience fee

No backup, no data transfer fees, yearly

Cloud and region	MongoDB Atlas	IaaS pay-as-you-go	IaaS reserved (3y)
M30 (3-node RS, 2vCPU, 8GB RAM, 40GB disk)			
AWS us-east1	**392%**	$2,666	**$1,204**
Azure centralus	**501%**	$2,450	**$1,049**
GCP us-east1	**276%**	$2,796	**$1,393**
M200 (4x3-node RS, 64vCPU, 256GB RAM, 2TB storage)			
AWS us-east1	**349%**	$355,086	**$159,674**
Azure centralus	**386%**	$366,672	**$162,575**
GCP us-east1	**265%**	$379,376	**$198,303**

PERCONA

Ease of use (and scale) trap

- Cloud resources are virtually unlimited
- Pay-as-you go
- Less careful planning
- Building 20-shard 5-node each cluster?

  Easy!

What's the best tool to resolve MongoDB performance issues quickly?

PERCONA

Resolving problems with a credit card

- Over Indexing
- Suboptimal shard keys
  - Unbalanced shards
  - Queries spanning all shards
- No data archival

(Couple of) sizes fit all

- Any upgrade costs a lot (double!)
- Only 10 production-grade tiers
  - Some with low-cpu options
  - Some with NVMe options

M50	8vCPUs	32	$2.16/h
⇩+97%			
M60	16vCPUs	64	$4.27/h
⇩+85%			
M80	32vCPUs	128	$7.90/h

MongoDB Atlas, 3-node RS, AWS eu-north-1

Cloud DBaaS topologies flexibility

- Lack of arbiter nodes
- No way to differentiate nodes
- Can’t run single node

Data transfer costs

- Hidden
- Hard to estimate
- Hard to verify
- Inter AZ -> Cross AZ -> Cross Region -> Internet

- AWS: $0.01 - $0.09
- Azure: $0.01 - $0.18
- GCP: $0.01 -$0.23

Backups

AWS

- Snapshots: $0.05GB/month
- S3: $0.023GB/month

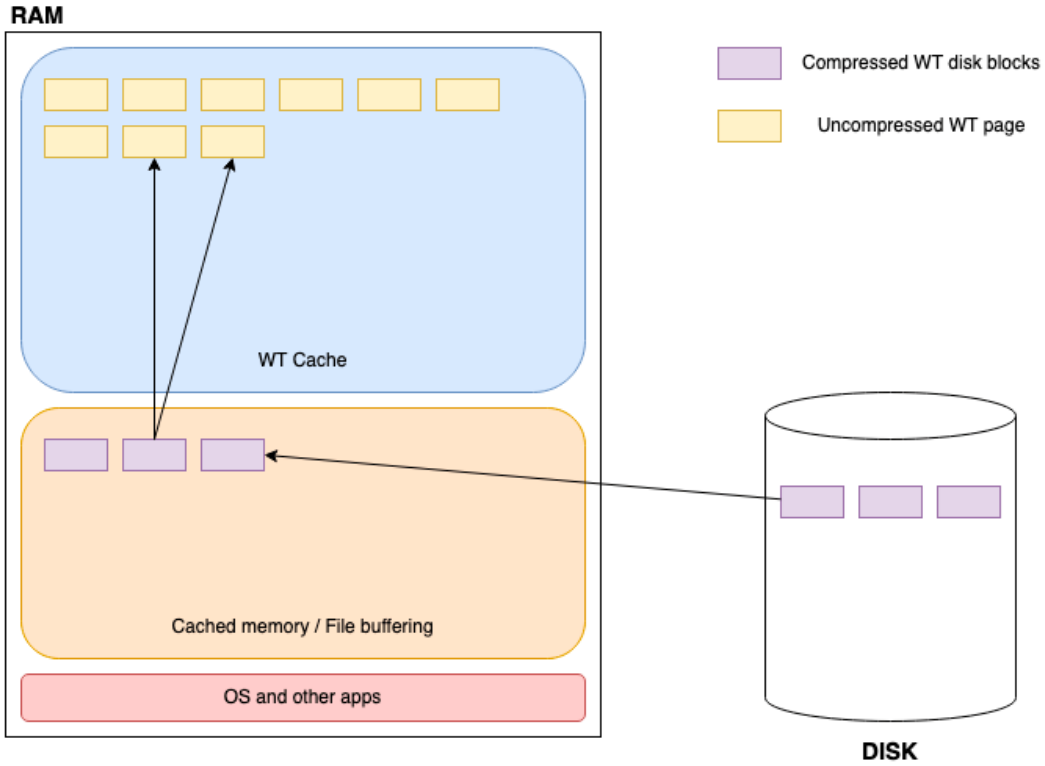
MongoDB Atlas on AWS

- Snapshots: $0.19GB/month
- PITR: $1.55 - $0.40 (!)

Suboptimal configuration

- Overprovisioning config replica set nodes
- Too many shards for the dataset
- Lack of tuning and sticking to defaults
- Multiple mongod services on a single node with a default cache options

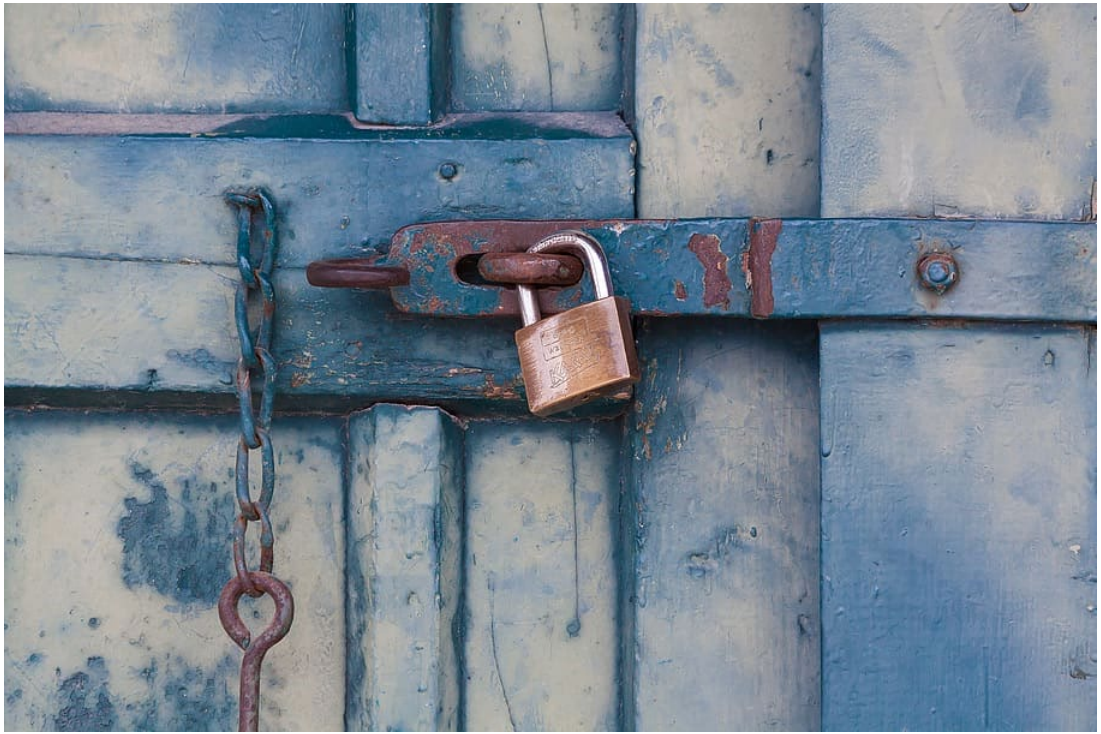
Suboptimal configuration



“Enterprise-grade” features needs

- LDAP Authentication/Authorization
- Data at rest encryption
- Own encryption key management
- Audit log
- Advanced backups
- Support?
- Automation for scale

Cloud DBaaS lock-in

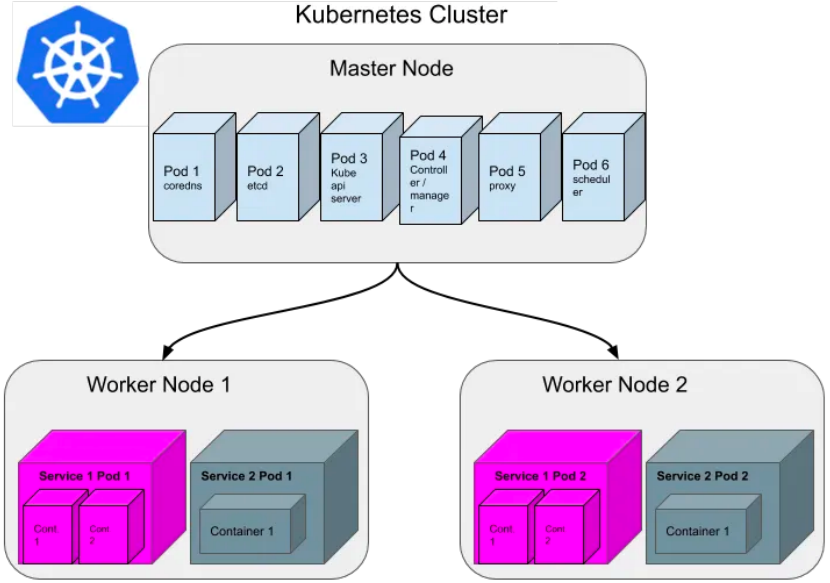


PERCONA

PERCONA

II. K8s and Operators 5 min crash course

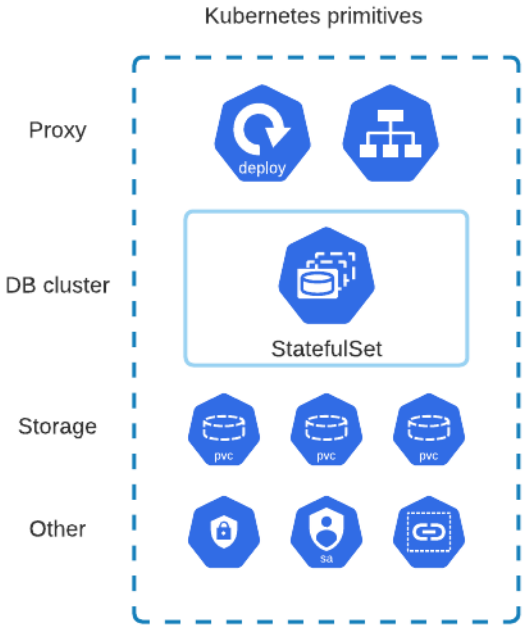
What is Kubernetes?



- Worker nodes
- Pods
- StatefulSets
- ReplicaSets
- PersistentVolumes
- PersitentVolumeClaims
- Secrets

PERCONA

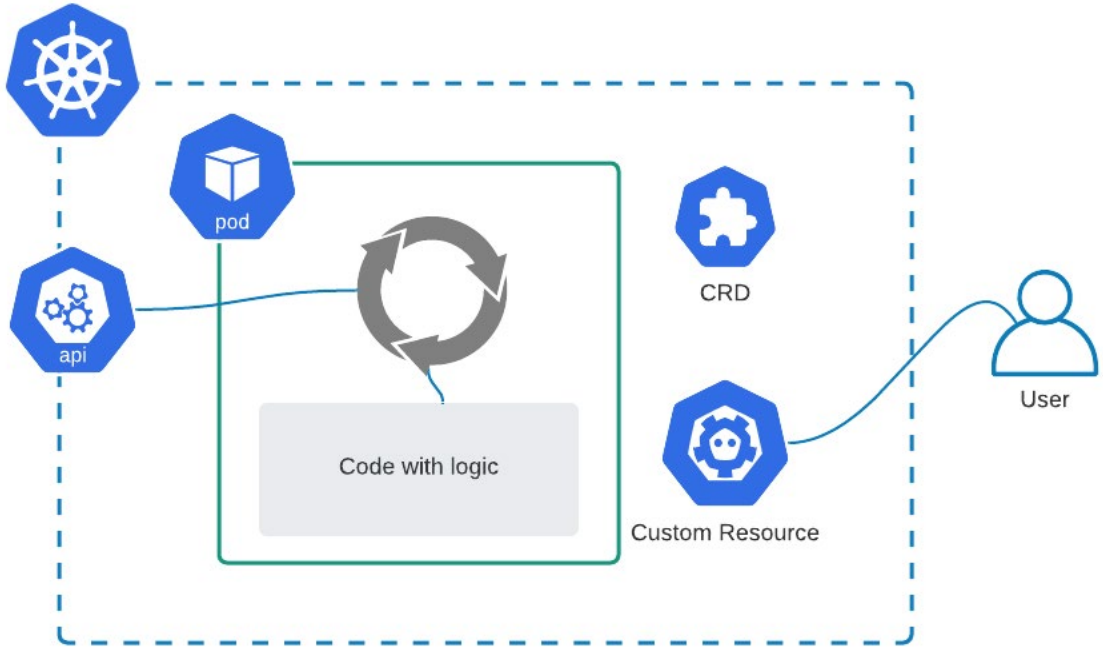
MongoDB on K8s



1. Underlying Kubernetes
2. StatefulSet for Primaries/Replicas/Mongo Config
3. High Availability and failover
4. DR
5. Storage (PVC/Hostpath)
6. Endpoints
7. Configuration
8. Monitoring
9. Backups
10. Upgrades
11. Encryption
12. …..

PERCONA

What is Kubernetes Operator?

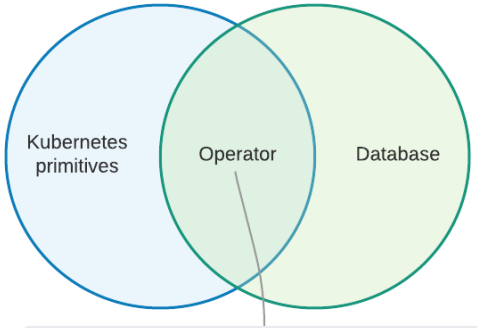


PERCONA

What is Kubernetes Operator?

- Operator controls database and k8s primitives

- Day-1 simplified to one step
- Day-2 operations automated

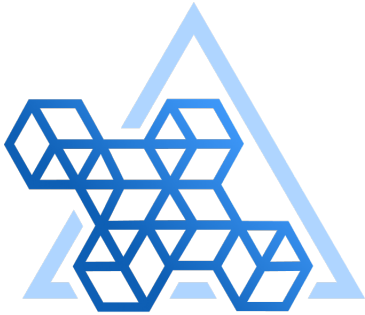



```
spec:
  image: percona/percona-server-mongodb:4.4.6-8
  replsets:
   - name: rs0
     size: 3
    sharding:
       mongos:
          size: 3
        configsvrReplSet:
           size: 3
      backups:
         ...
```


PERCONA

What is Kubernetes Operator?

1. Deploy easily: replica sets, shards, (mongo/d/s/c)
2. Topology management (arbiters, node affinity, scaling)
3. Monitoring integration
4. Network exposure and load balancing
5. Backups management with Percona Backup for MongoDB
6. Self-healing
7. Upgrade automation (minor, manual major)
8. Configuration adjustments



PERCONA

Kubernetes
Operators

kubernetes OPENSHIFT Google Cloud Platform aws Microsoft Azure VMware Tanzu

How to use it?

```
apiVersion: psmdb.percona.com/v1
kind: PerconaServerMongoDB
metadata:
  name: percona-live-cluster
spec:
  crVersion: 1.15.0
  image: percona/percona-server-mongodb:6.0.4-3
  secrets:
    users: minimal-cluster
  replsets:
  - name: shard1
    size: 3
    resources:
      limits:
        cpu: "4"
        memory: "8G"
      requests:
        cpu: "4"
        memory: "8G"
    volumeSpec:
      persistentVolumeClaim:
        resources:
          requests:
            storage: 30Gi
```

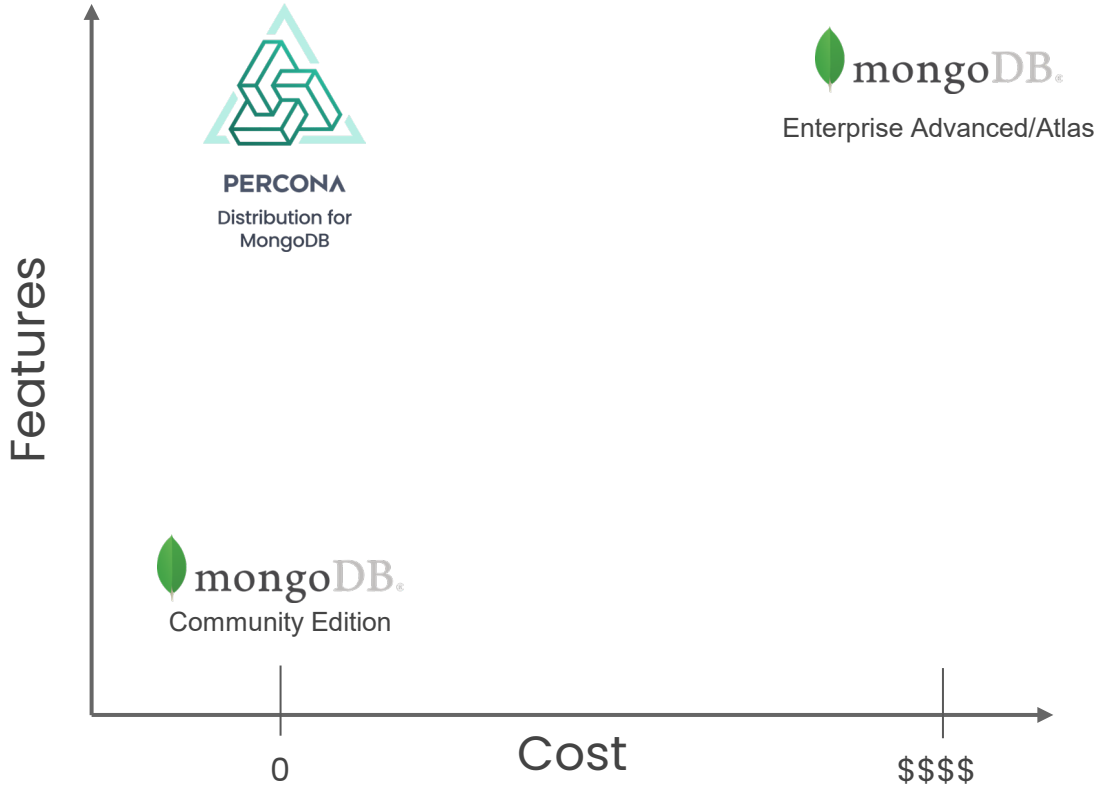
```
sharding:
  enabled: true
  configsvrReplSet:
    size: 3
    resources:
      limits:
        cpu: "2"
        memory: "4G"
      requests:
        cpu: "2"
        memory: "4G"
    volumeSpec:
      persistentVolumeClaim:
        resources:
          requests:
            storage: 3Gi
  mongos:
    size: 1
```

PERCONA

III. Practical approaches to cost reduction

Use open source

1. Advanced backups
2. LDAP
3. Data-at-rest encryption
4. KMIP integration
5. Auditing
6. Monitoring

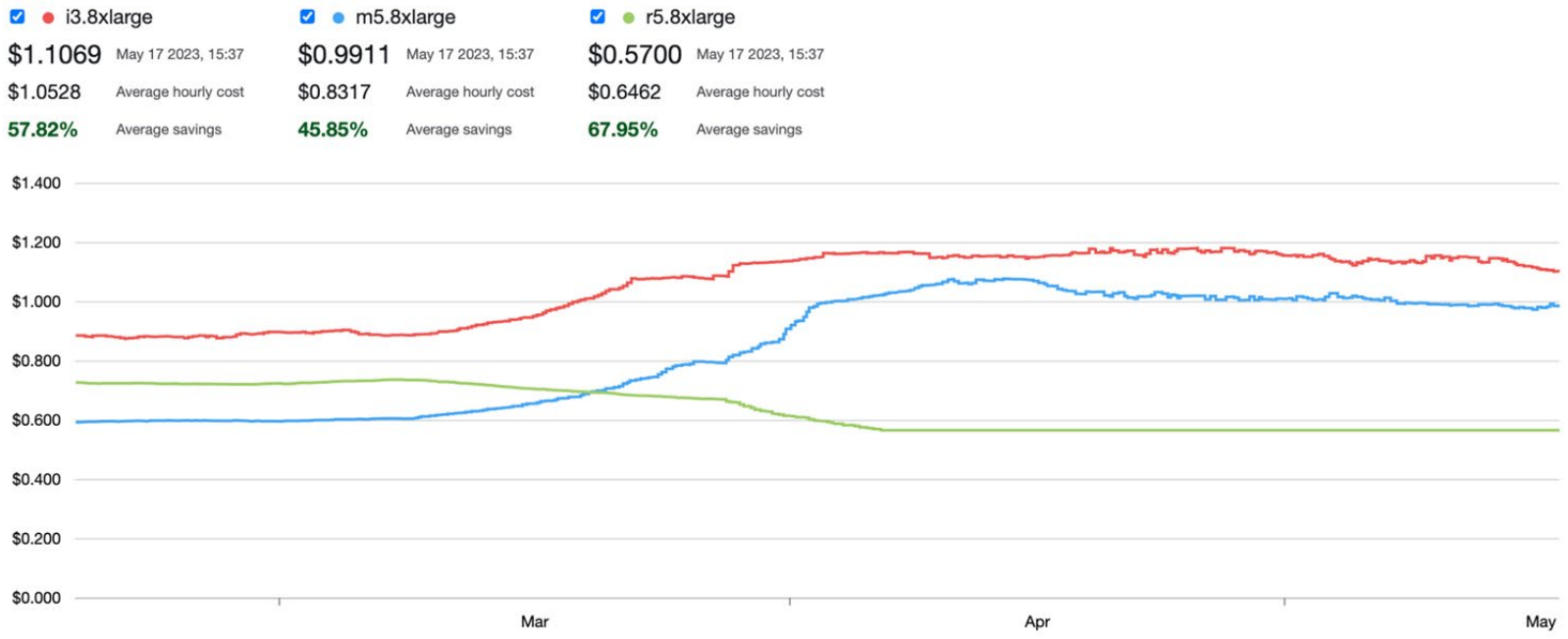


K8s resources are cost-efficient

- Managed K8s is ~$70/month at any major cloud provider
- Utilize “raw” resources at “raw” resource prices
- Embrace discounts:
  - Reserved prices (41% 1 year, 62% 3-year)
  - Spot instances

K8s resources are cost-efficient

Spot instances



Source: https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#SpotInstances:

PERCONA

Embrace the variety of available cloud compute resources

a. Create node groups with instance types that fit your workload
b. Utilize multiple node groups
c. Choose CPU/memory ratio wisely

Embrace the variety of available cloud compute resources

```
eks.amazonaws.com/capacityType=ON_DEMAND
eks.amazonaws.com/nodegroup=critical
node.kubernetes.io/instance-type=r5.2xlarge
topology.kubernetes.io/region=eu-west-1
topology.kubernetes.io/zone=eu-west-1b
```

Assign resources granularly

4 CPU -> 8 CPU -> 16 CPU -> 32 CPU
16GB -> 32GB -> 64GB -> 128 GB

Assign resources granularly

~~4 CPU -> 8 CPU -> 16 CPU -> 32 CPU~~
~~16GB -> 32GB -> 64GB -> 128 GB~~

1m (0.001 CPU) and 1 (CPU)
k, M, G, T...

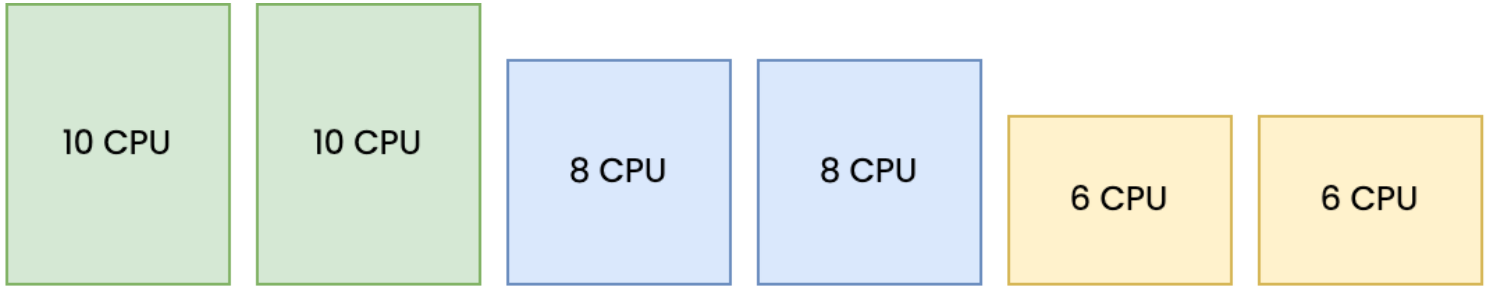
12,5 CPU and 19GB RAM

Assign resources granularly

```
replsets:
- name: shard1
  size: 3
  resources:
    limits:
      cpu: "12.5"
      memory: "19G"
    requests:
      cpu: "12.5"
      memory: "19G"
```

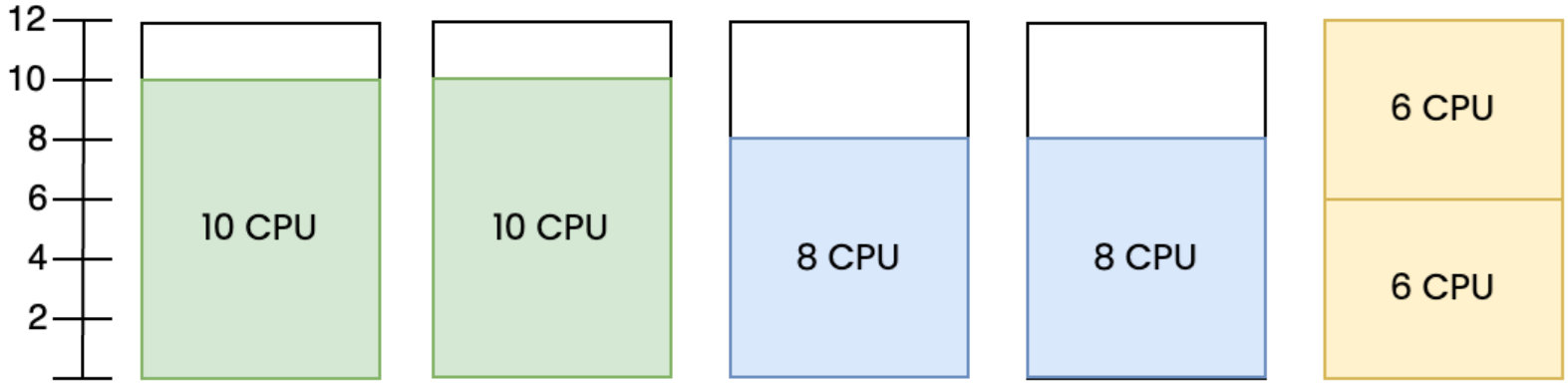
1. Each shard
    a. Data nodes
    b. Hidden nodes
    c. Arbiters
2. Config RS
3. Mongos
4. Backup pod

Play container tetris



48 cores to schedule

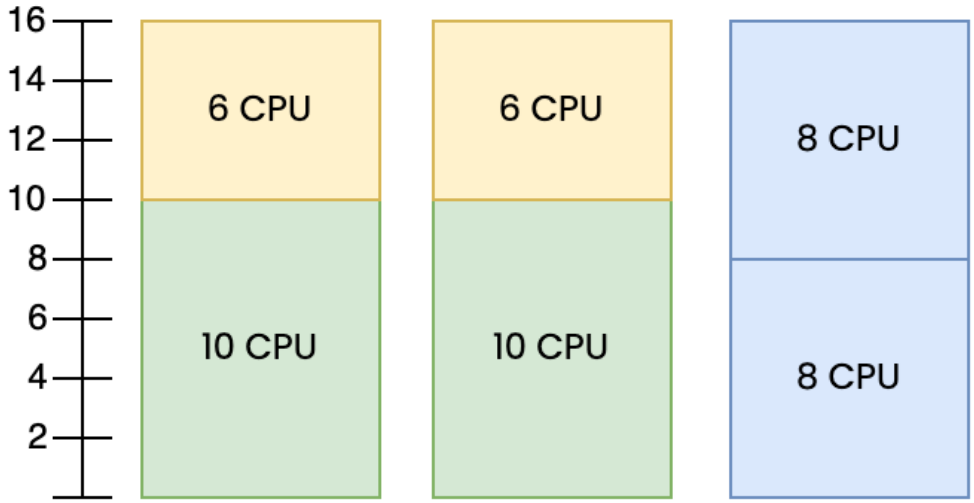
Play container tetris



48 cores to schedule, 60 cores to pay for

PERCONA

Play container tetris

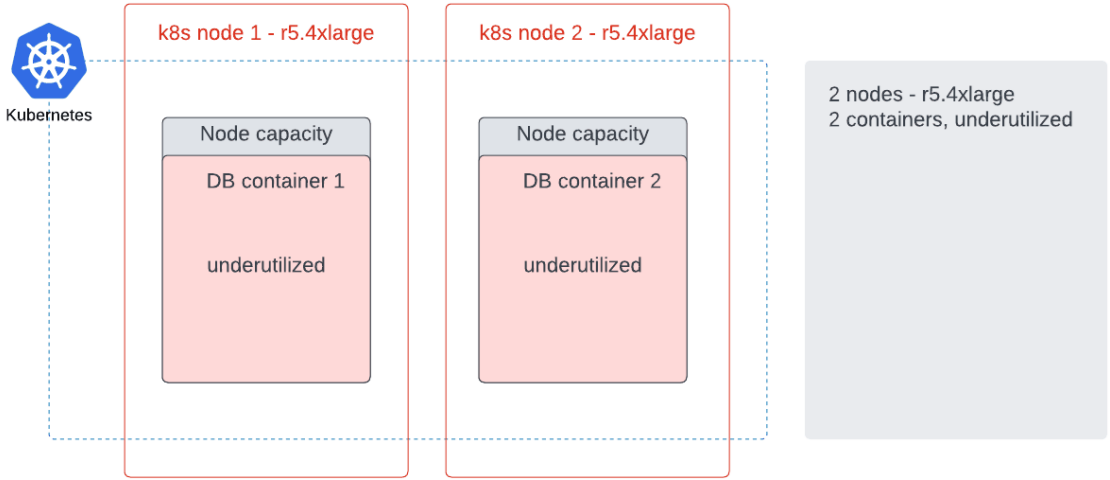


48 cores to schedule, 48 cores to pay for

PERCONA

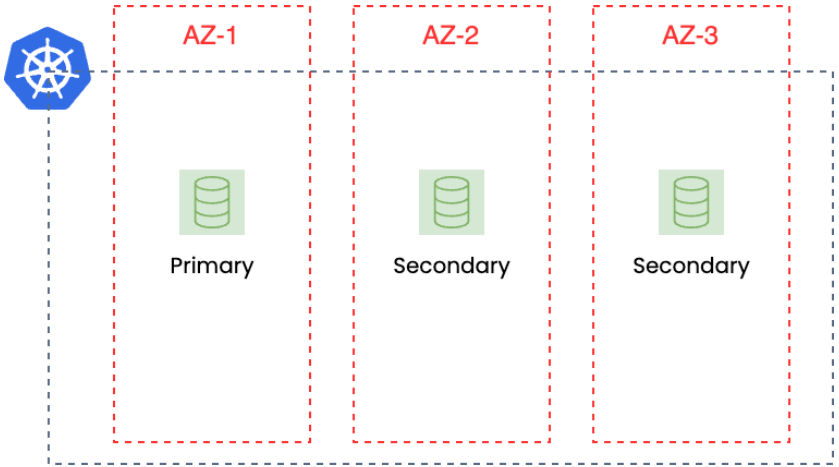
Autoscaling

1. VPA
2. HPA/KEDA
3. Cluster Autoscaler



PERCONA

Choose the right HA-level and topology



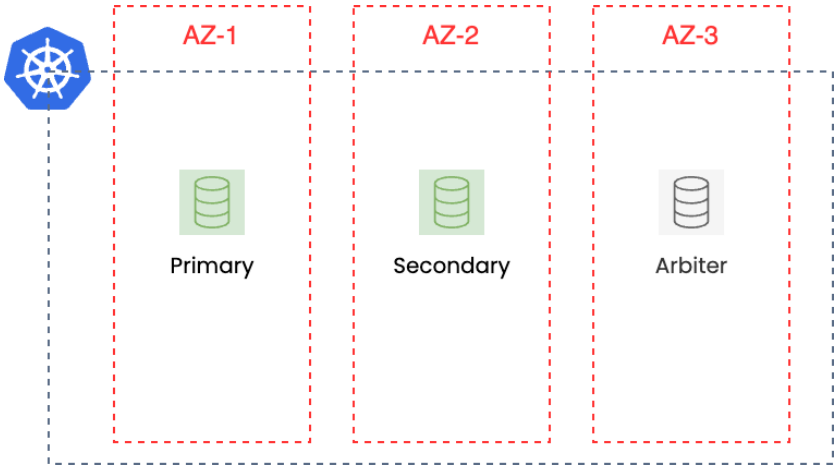

```
spec:
  allowUnsafeConfigurations: false

  replsets:
    - name: shard1
      size: 3
      affinity:
        antiAffinityTopologyKey:
          "kubernetes.io/zone"

      arbiter:
        enabled: false
        size: 1
```


PERCONA

Choose the right HA-level and topology



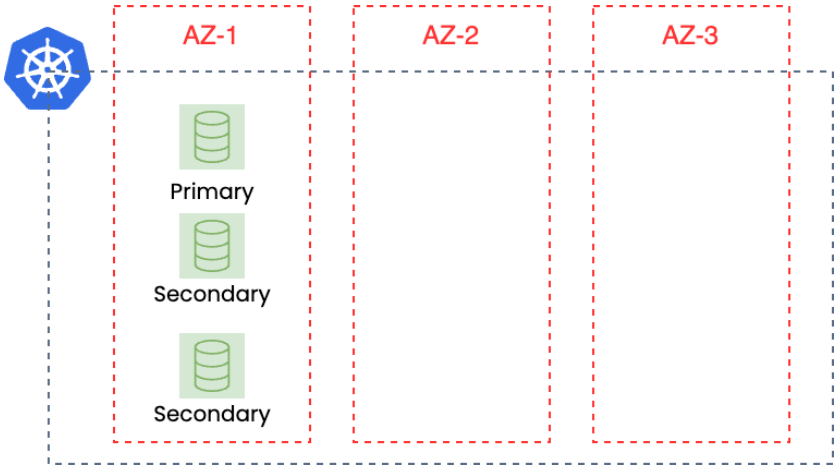

```
spec:
  allowUnsafeConfigurations: false

  replsets:
    - name: shard1
      size: 3
      affinity:
        antiAffinityTopologyKey:
          "kubernetes.io/zone"

      arbiter:
        enabled: true
        size: 1
```


PERCONA

Choose the right HA-level and topology



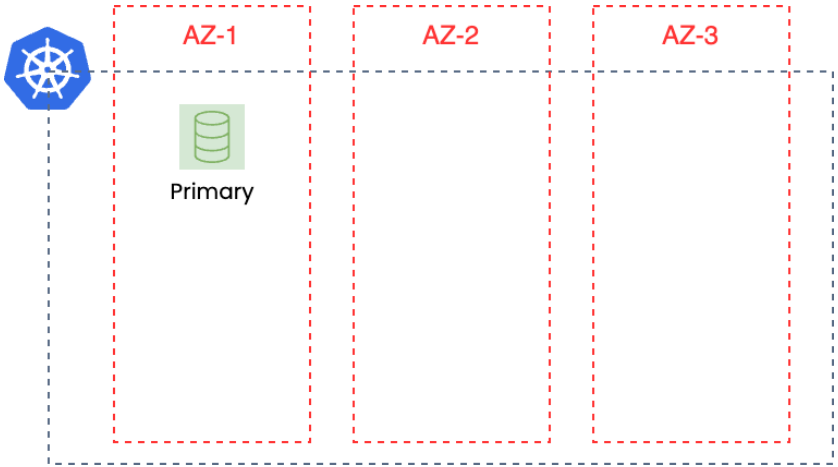

```
spec:
  allowUnsafeConfigurations: false

  replsets:
    - name: shard1
      size: 3
      affinity:
        antiAffinityTopologyKey:
          "kubernetes.io/host"

      arbiter:
        enabled: false
        size: 1
```


PERCONA

Choose the right HA-level and topology




```
spec:
  allowUnsafeConfigurations: true

  replsets:
    - name: shard1
      size: 1
      affinity:
        antiAffinityTopologyKey:
          "kubernetes.io/host"

      arbiter:
        enabled: false
        size: 1
```


PERCONA

Use multiple storage classes

```
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
  name: bronze
parameters:
  type: gp2
  fsType: xfs
```

```
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
  name: silver
parameters:
  type: gp3
  iops: 6000
  fsType: xfs
```

```
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
  name: gold
parameters:
  type: io2
  iopsPerGB: 40
  allowAutoIOPSPerGBIncrease: “true”
  fsType: xfs
```


PERCONA

Use multiple storage classes

```
replsets:
  - name: shard1
    size: 3
    volumeSpec:
      persistentVolumeClaim:
        storageClassName: gold
        resources:
          requests:
            storage: 300G
```


PERCONA

Local NVMe storage is cheap!

Local storage

- i3.8xlarge
- 32 vCPU
- 244GB RAM

- 4x1900GB NVMe SSD included
- 1,650k Read + 720k Write IOPS

$1,822/month

EBS storage

- r5.8xlarge
- 32 vCPU
- 256GB RAM

- EBS 4x600GB
- 64k IOPS each

$15,915/month

Local NVMe storage is cheap!

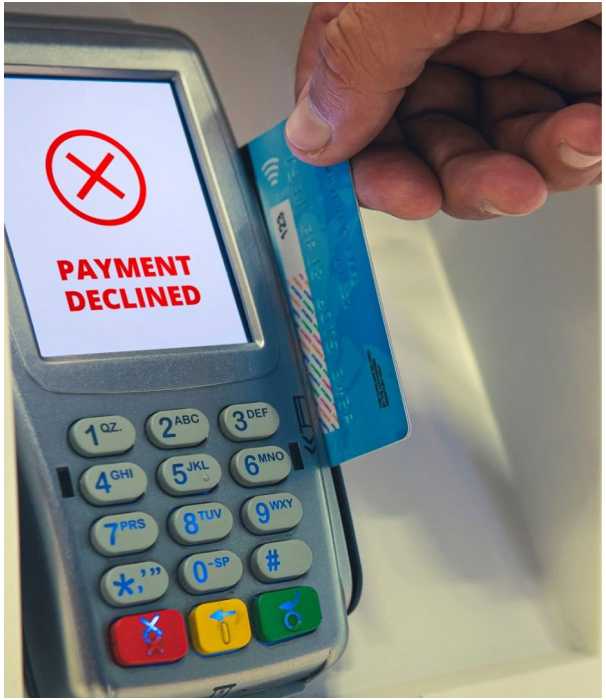
1. It's ephemeral. You lose the data when the node dies
2. You can't extend its size
3. It depends strictly on the instance type
4. Pods won't be automatically rescheduled

Recommendations:

1. Always use HA mode (at least 3 nodes)
2. OpenEBS for simplified management
3. Set Write Concern properly

Stop solving issues with credit card

- Indexes:
  - Check usage stats
  - Consider partial indexes
- Sharding key:
  - Plan it
  - Adjust it
- Data archival
  - Capped collection
  - TTL indexes



PERCONA

Understanding is a key

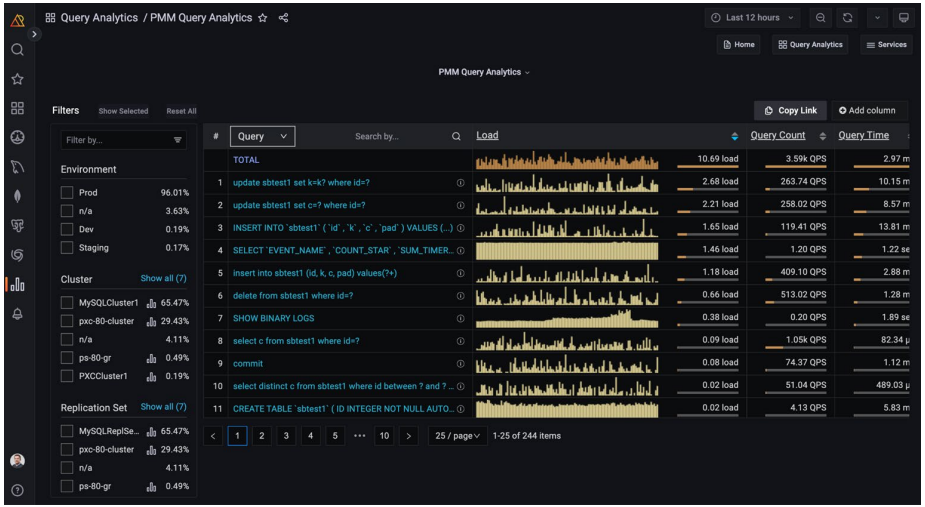
- Understand the workload
- Review outstanding queries
- Look for underutilization, too

```
helm repo add percona
https://percona.github.io/percona-helm-charts/
helm install pmm \
--set secret.create=false \
--set secret.name=pmm-secret \
percona/pmm
```

```
pmm:
    enabled: true
    image: percona/pmm-client:2.37.0
    serverHost: monitoring-service
```

Understanding is a key

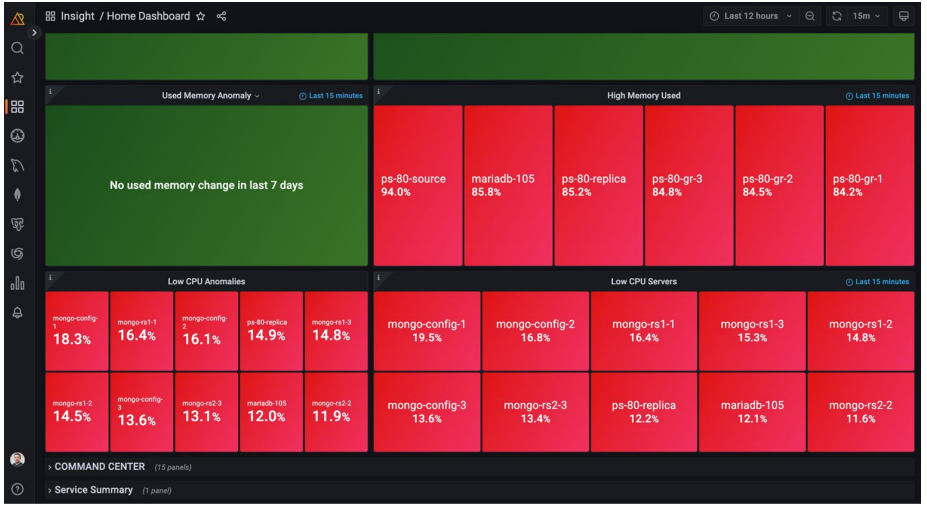
- Understand the workload
- Review outstanding queries
- Look for underutilization, too



PERCONA

Understanding is a key

- Understand the workload
- Review outstanding queries
- Look for underutilization, too



PERCONA

PERCONA

IV. Summary

Why costs are getting out of control?

1. Convenience fee
2. Licenses/extra features
3. Suboptimal configuration/usage

What can you do to reduce costs?

1. Use open source K8s Operator + PSMDB
2. Utilise cheap resources efficiently
3. Eliminate common mistakes in configuration and usage

What’s the potential?

What's the potential?

It depends.

Key takeaways

1. Cloud DBaaS such as MongoDB Atlas comes with a huge premium fee

   In some cases 5x cloud resources cost.

1. Hidden costs and suboptimal MongoDB configuration make the situation worse

   Cloud and DBaaS fees are difficult to predict and understand.

1. Kubernetes and Percona MongoDB Operator can handle any MongoDB workload

   The solution has been proven in production for many years.

1. Running MongoDB on K8s has a significant cost saving potential

   It makes it possible to utilize cheap “raw” cloud resources without giving up convenience and automation

PERCONA

Want to know more about

Percona Kubernetes Operators?

PERCONA

Databases run better with Percona