

# MongoDB in Docker

袁晓沛

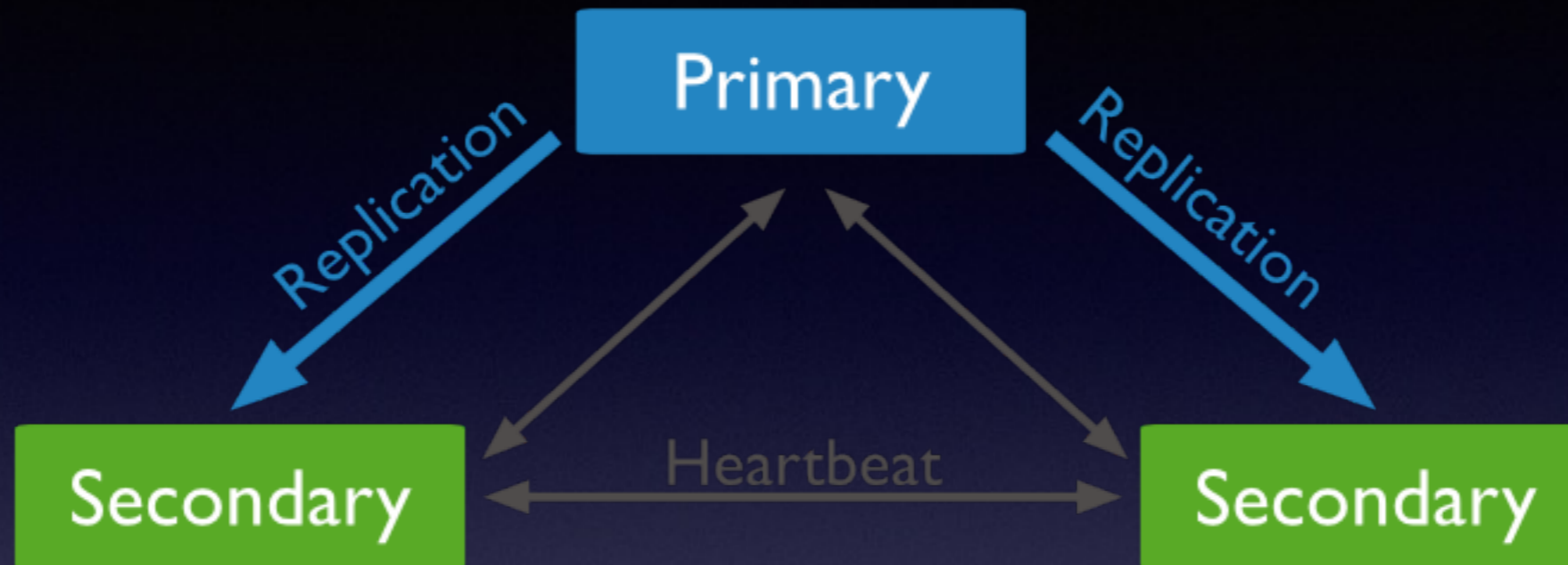
七牛 通用计算 / 应用部

# Agenda

- MongoDB Arch
- Pain Points
- Dockerize MongoDB
- Benefits
- Traps

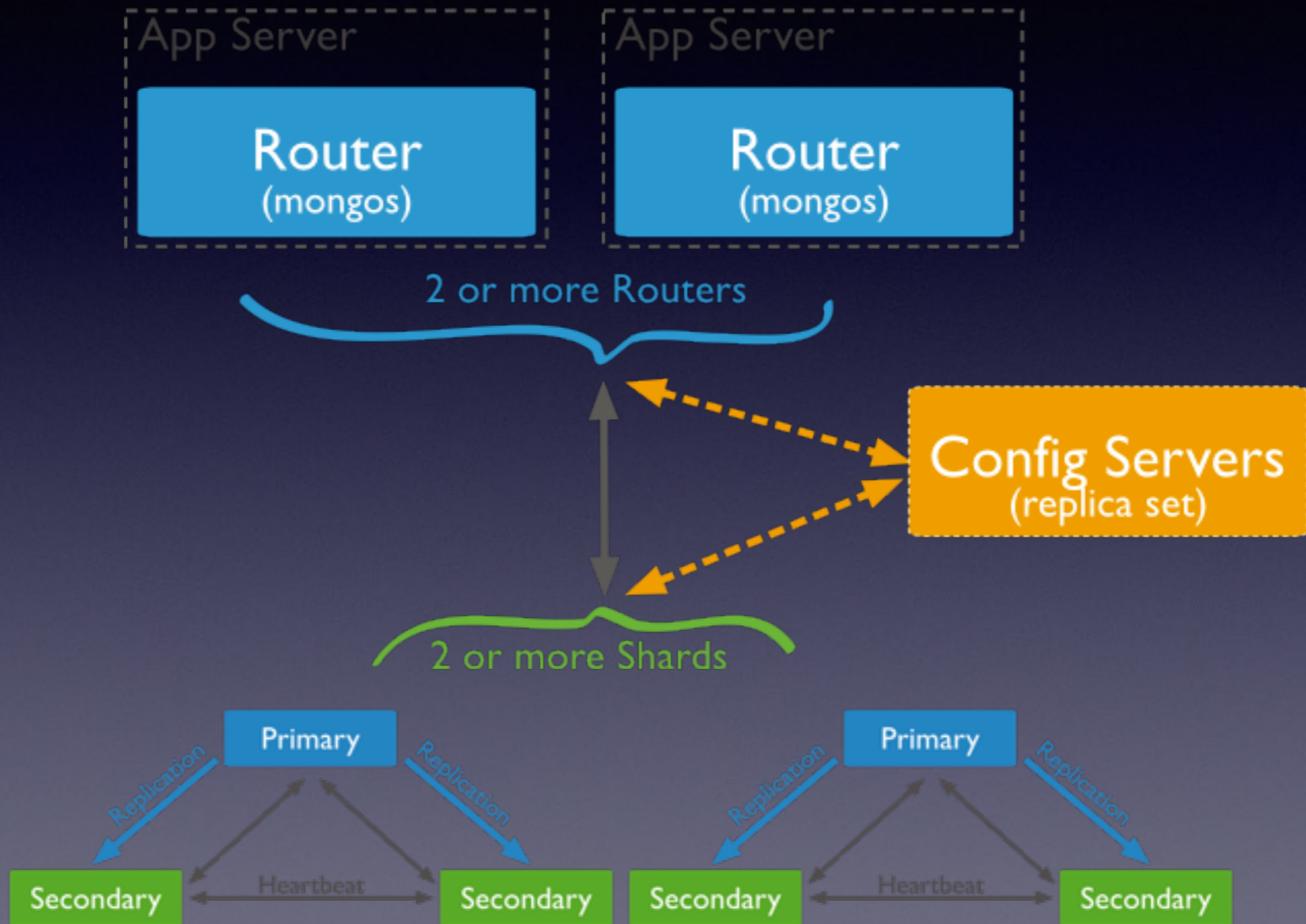
# MongoDB Arch

# Replica Set



- 一主两从
- 主故障，从自动切主
- 写入主节点
- 读主或者读从取决于一致性模式

# Sharding Cluster



# Pain Points

# 痛点一：部署繁琐

- 准备物理机
- 配置软件
- 难以自动化

## 痛点二：运维麻烦

- 配置监控
- 备份策略
- 手动运维，难以自动化



# 痛点三：资源浪费

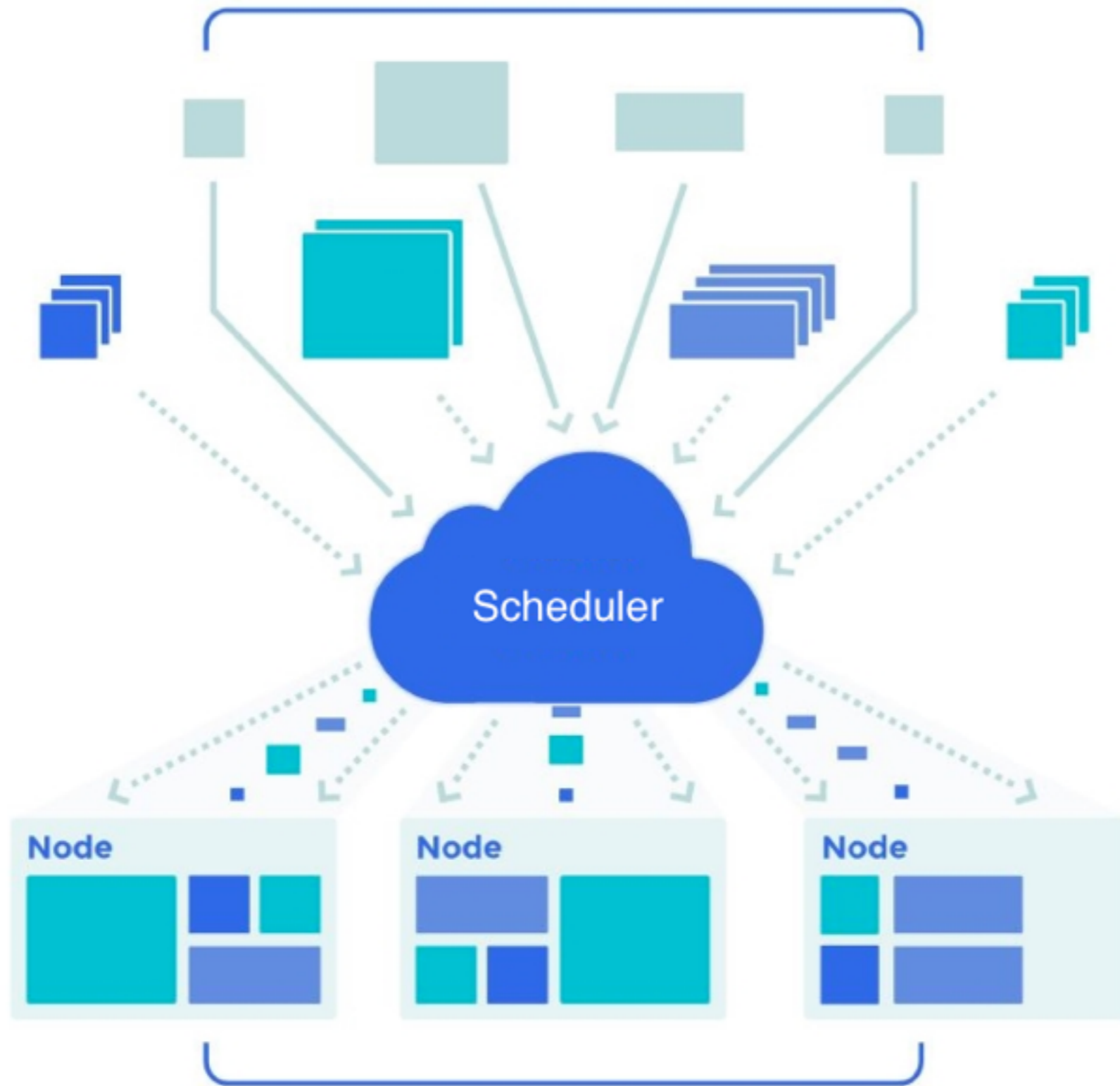
- 不同节点资源消耗不同
- 混合部署增大运维复杂性
- 资源利用率低

# 痛点四：很难支持多租户

- 共用db集群?
  - 业务相互影响性
- 自己实现DB中间件?
  - 实现复杂，开发周期长
  - 投入大、产出小

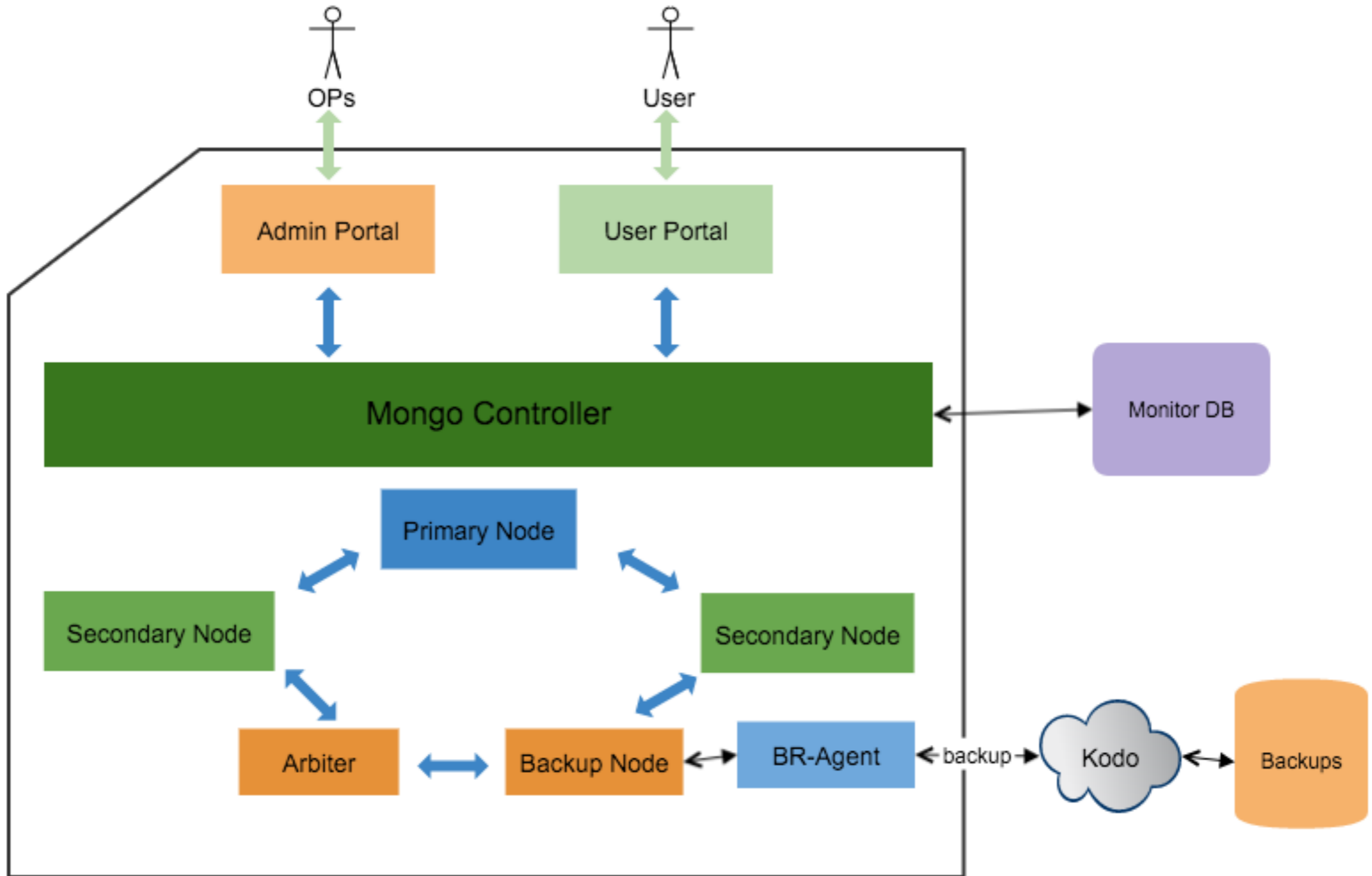
# Dockerize MongoDB

# An ocean of user containers



Scheduled and packed dynamically onto nodes

# Dockerize Everything



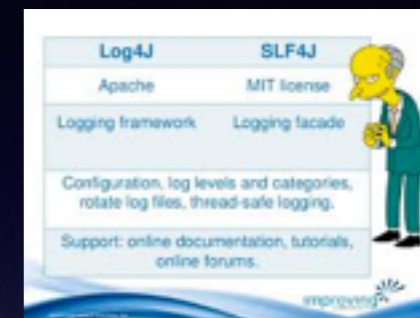
# 编排系统

- 自定义编排系统
  - 资源纬度：定义各种资源
  - 应用纬度：定义应用节点，以及和资源间的关系
  - 描述式语法



# 日志聚合

- 容器输出日志
  - fluentd, 归档
- 应用多个日志目录
  - 编排扩展描述中, 配置额外的log挂载目录
  - 归档相应目录中的日志



# 存储管理

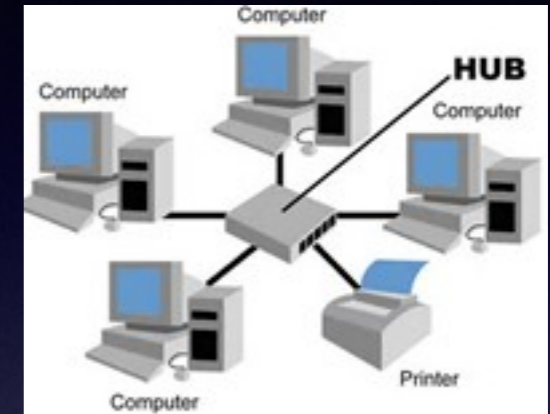
- 实现存储管理子系统
  - 目前支持本地LVM
  - 未来可以拓展到网络存储





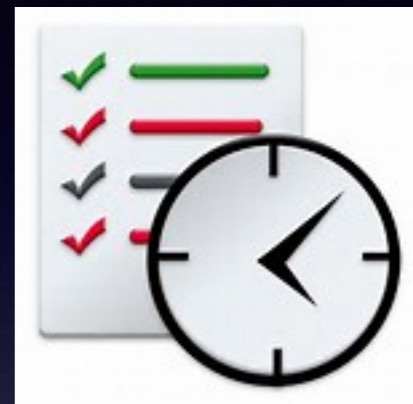
# 网络管理

- SDN
  - OpenVSwitch, VXLAN
  - 多租户支持
  - 负载均衡



# 容器调度

- 根据资源需求调度
  - CPU, Mem, Storage, Network
- 非亲缘性
  - 消减硬件故障的影响面



Benefits

# Benefits

- 资源隔离

# Benefits

- 资源隔离
- 自动化部署

# Benefits

- 资源隔离
- 自动化部署
- 自动化监控

# Benefits

- 资源隔离
- 自动化部署
- 自动化监控
- 半自动化运维

# Benefits

- 资源隔离
- 自动化部署
- 自动化监控
- 半自动化运维
- 横向扩展 (Scale Out)



# Benefits

- 资源隔离
- 自动化部署
- 自动化监控
- 半自动化运维
- 横向扩展 (Scale Out)
- 纵向扩展 (Scale Up)

# Traps

# 避免 docker 镜像缓存

## Wrong

```
-RUN curl -o jdk.tar.gz https://dn-qcos.qbox.me/jdk-7u15-linux-x64.tar.gz
```

```
-RUN mkdir -p opt && tar -xf jdk.tar.gz -C /opt && rm -rf jdk.tar.gz
```

## Correct

```
+RUN curl -o jdk.tar.gz https://dn-qcos.qbox.me/jdk-7u15-linux-x64.tar.gz \
```

```
+    && mkdir -p opt \
```

```
+    && tar -xf jdk.tar.gz -C /opt \
```

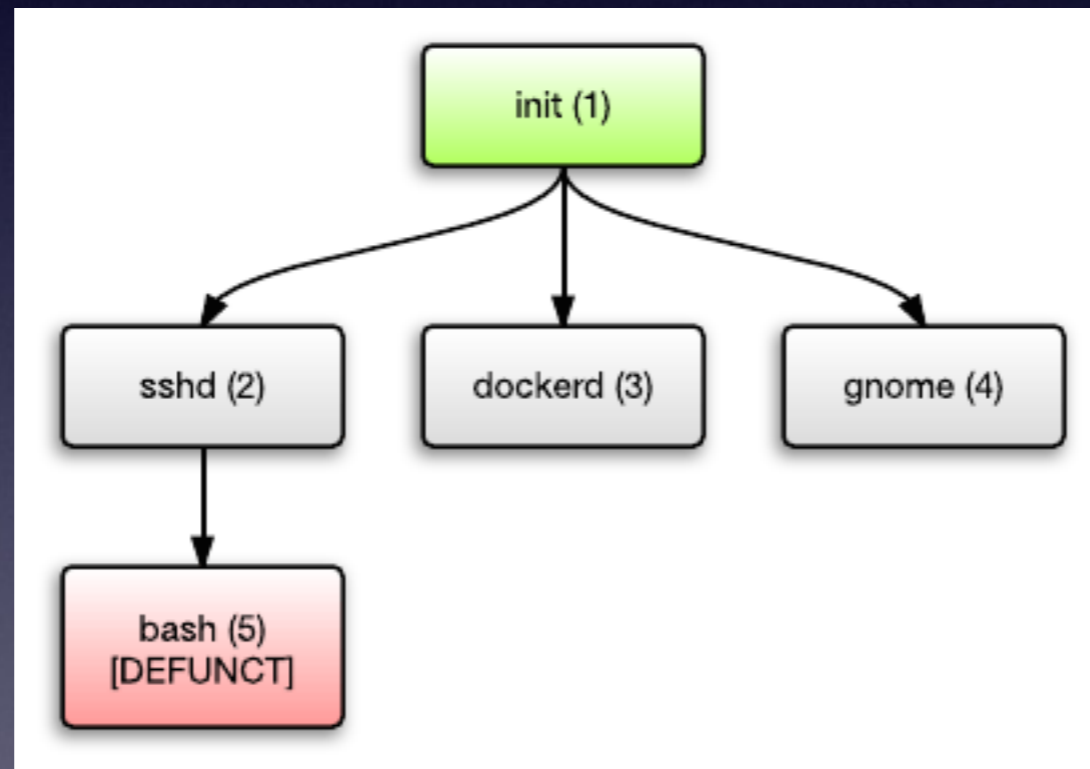
```
+    && rm -rf jdk.tar.gz
```

# 容器重启策略

- 设置恰当的重启策略
  - no
  - on-failure[:max-retries] (DB类)
  - always (无状态类、缓存类)

# 容器初始进程

- 初始进程退出前，应waitpid



# stopWait

- 设置恰当的stopWait时间
  - `docker stop -t <stopWait>`
    - Seconds to wait for stop before killing it



Q & A

[yuanxiaopei@qiniu.com](mailto:yuanxiaopei@qiniu.com)