



DTCC

2016中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2016

数据定义未来

SequeMedia
盛拓传媒

IT168.com

ChinaUnix

ITPUB

MongoDB服务化及在 网易游戏应用

网易游戏：黄文宇



about me

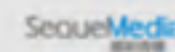
- 网易游戏2005-2016
- 天下、天谕、CC、UU
- 乱斗西游、功夫熊猫、率土之滨
- 游戏运维自动化平台负责人
- MongoDB Team Leader



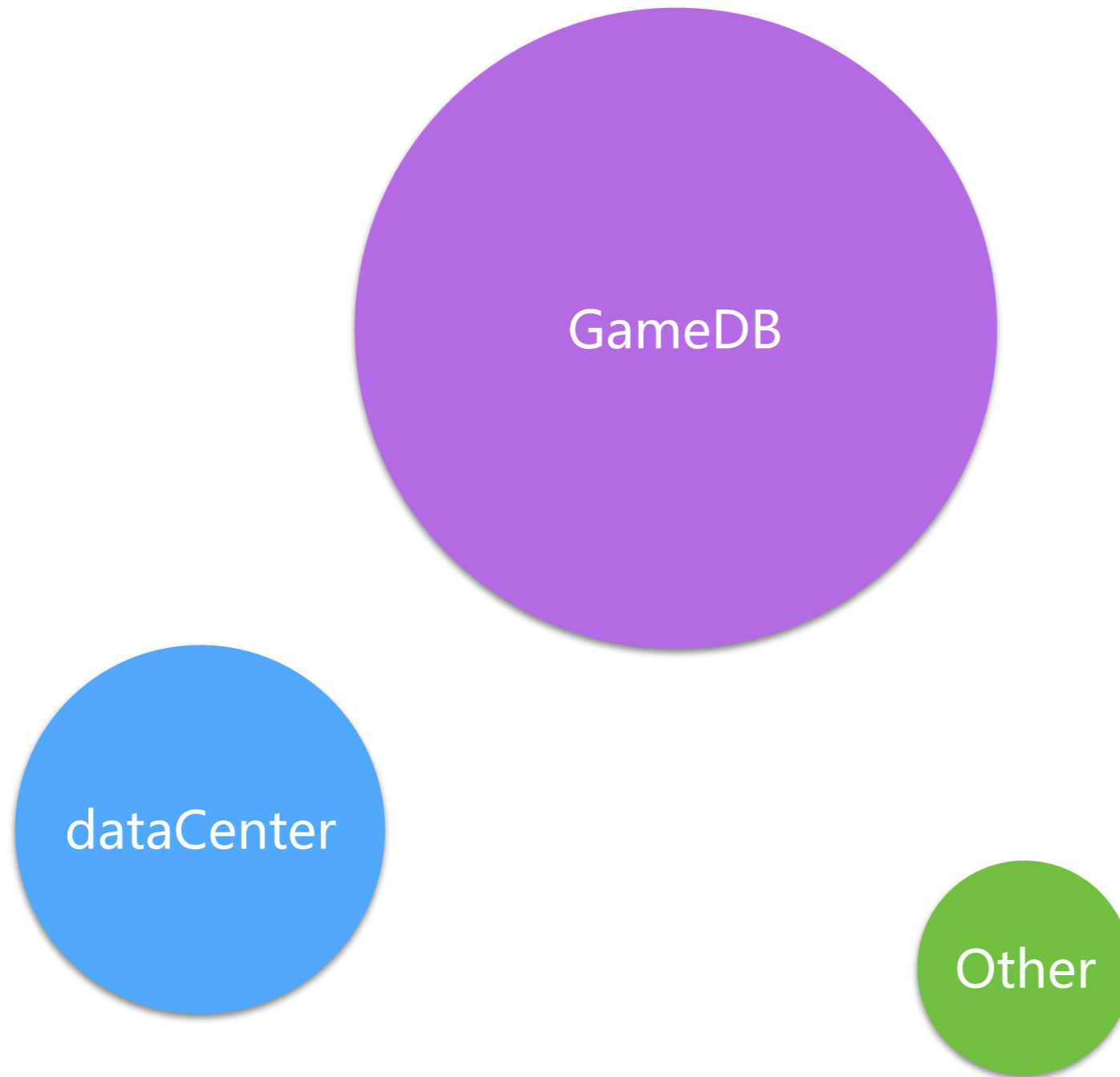
MongoDB 在网易游戏



2016年中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2015



MongoDB at NeteaseGame



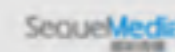
- 由项目SA/DBA搭建与管理MongoDB
 - 学习&运维成本高
 - 资源浪费严重
 - 交付效率低
 - 可用性低



MongoDB SaaS -Ocean



2016年中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2015



服务化目标

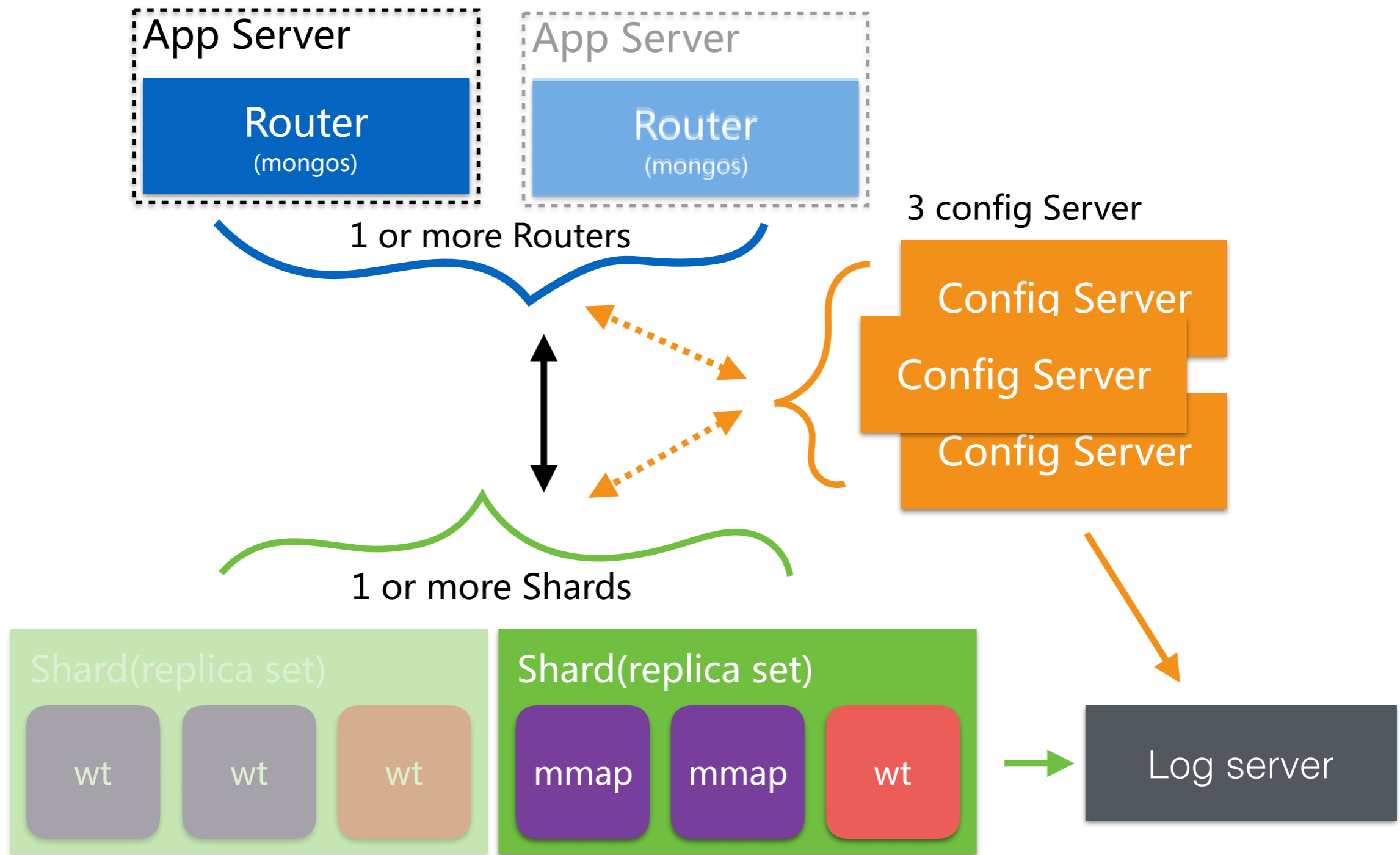
- 减轻项目DBA/SA的学习及配置成本
- 为每个使用MongoDB的项目提供标准化的服务
- 统一调配及资源整合降低人力及硬件成本



服务提供形式



3 config + N Shards + N mongos + 1 logserver



- 2 node + arbiter or 3 node ?
- replica set or shard ?



Why not RS

- 扩展性不强
- 不同程序驱动下表现不一致
- 不同的资源概念容易引起用户混淆



资源划分问题

- 是否提供灵活的CPU、内存定制服务？
- 磁盘以什么形式挂载？



For Game

- 用户对各类性能指标无概念，定制内容过多反而影响判断与选择
- 业务操作频繁update，对IO要求高



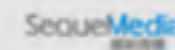
- 提供标准化node，不提供纵向伸缩，只提供shard方式的横向伸缩
- 高性能（SSD）
- 大存储（flash cache SSD+sata）
- 硬盘挂载宿主机



架构



2016年中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2015



内网主交换

交换

config

mongod

mongod

mongod

交换

config

mongod

mongod

mongod

交换

config

mongod

mongod

mongod

每个机柜独立交换机, 独立电源

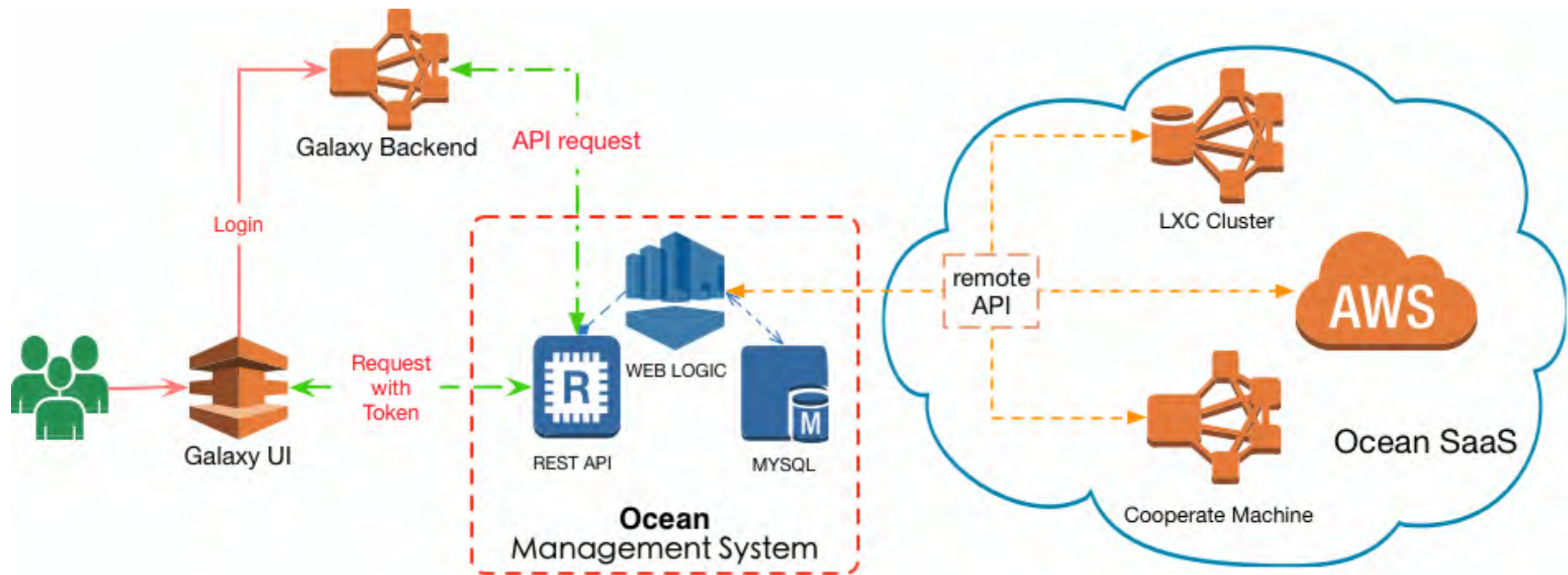
交换机需要保证带宽

- Ocean资源采用LXC半虚拟化解决方案
- 自己开发的Paas调度平台，确保每个shard的3个node来自不同机柜
- 资源通过cgroups划分，虚拟化资源消耗小
- 灵活扩展，理论上可以极端情况下可在线动态增加虚拟机的CPU、内存、硬盘、带宽资源



异构资源问题





Ocean管理系统架构图



来源

合作方

资源类型



区域

亚太-Seoul

项目

h2kr-天下HD韩国

用途

开发测试

Mongo版本

2.6.10

分片数量

2



常规备份时间

03:00



临时备份



是



否

取消

提交



状态: **运行中**

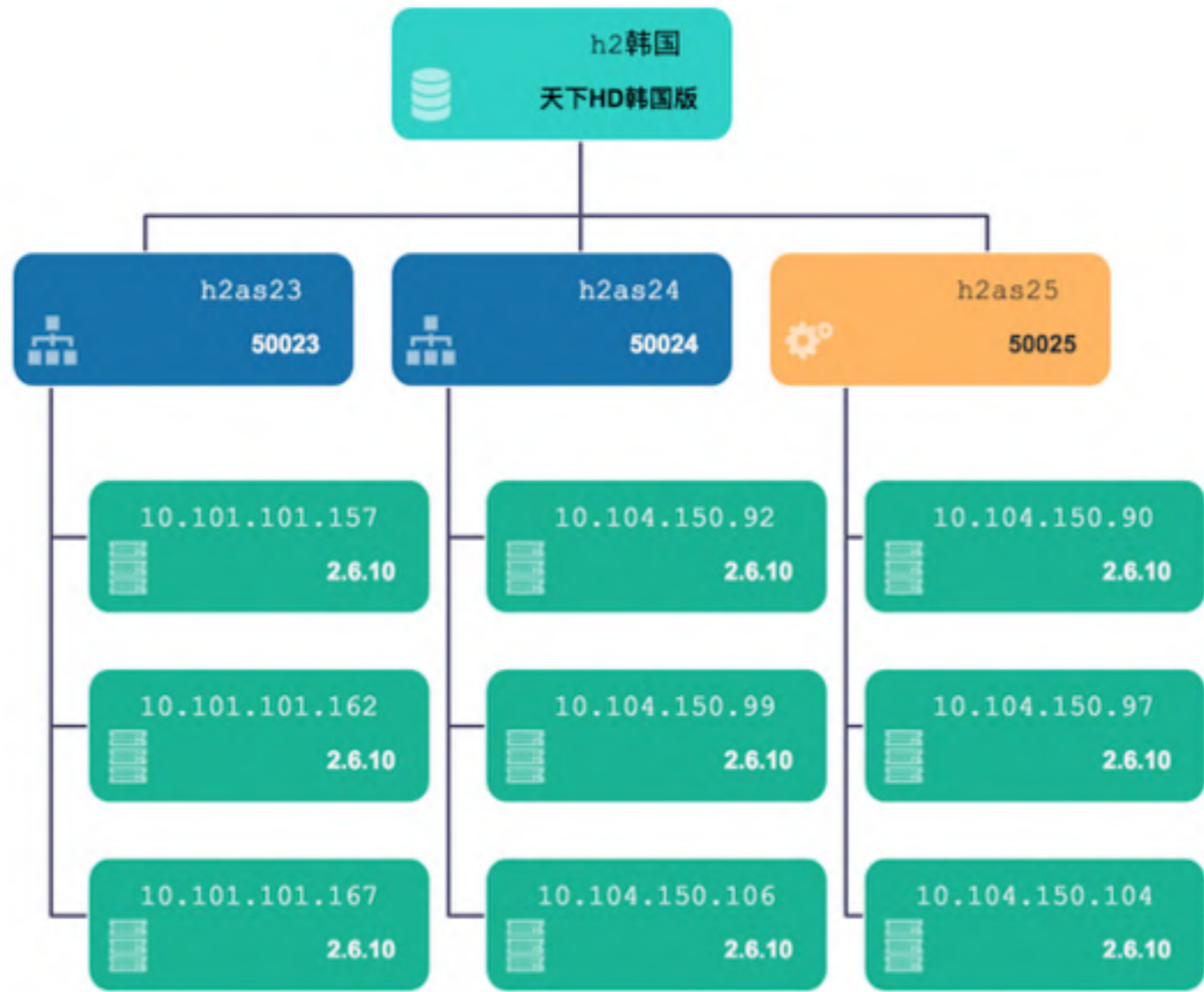
集群类型: 海外 [redacted] 管理机IP: [redacted]

Mongo 主版本: 2.6.10 引擎配比: 1 WT 2 MMAP

进度: 

集群创建流程已完成 60%。这个过程一般不会超过30分钟。如果长时间没有反应，请联系Ocean管理员。

架构图 详情



基本信息

集群ID: 60007

所属项目: h2kr-天下HD韩国

管理机IP: [redacted]

日志服: [redacted]

集群类型: 海外/AWS mongo

引擎配比: 1 WT 2 MMAP

用途: 高性能集群

Mongo版本: 2.6.10

备份计划

是否备份: 是

备份周期: 一天一次

备份时间: 03:00:00

恢复备份 添加分片

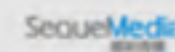
- 项目运维不需要关注数据库的搭建与维护
- 硬件成本下降73%
- 资源交付秒达
- 2014年可用性99.99%
- 2015年可用性99.999%
- 2016年可用性99.99%



使用建议



2016年中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2015



一些使用建议

- 集群内shard分片尽量不要超过五个
 - Balance 速度缓慢
 - 数据恢复困难
 - 宿主硬件故障影响范围加大



一些使用建议

- 慎重对待小于总量64M的表
 - 无法通过chunk平衡出去，容易造成单shard压力过大
 - 可考虑通过服务端预加载解决



两次全面故障

- rsyslog 远程写入日志服失败，导致全服mongod卡死
- LXC无法限制连接数导致宿主机连接数爆满，然后网络无法访问



- 云服务需要充分考虑所有小细节
 - 模拟每个业务环节故障时的影响
 - 确认每个业务环节是否存在参数或硬件限制





TIANKO

SequeMedia
融拓传媒

IT168.com

ChinaUnix

ITPUB

天科