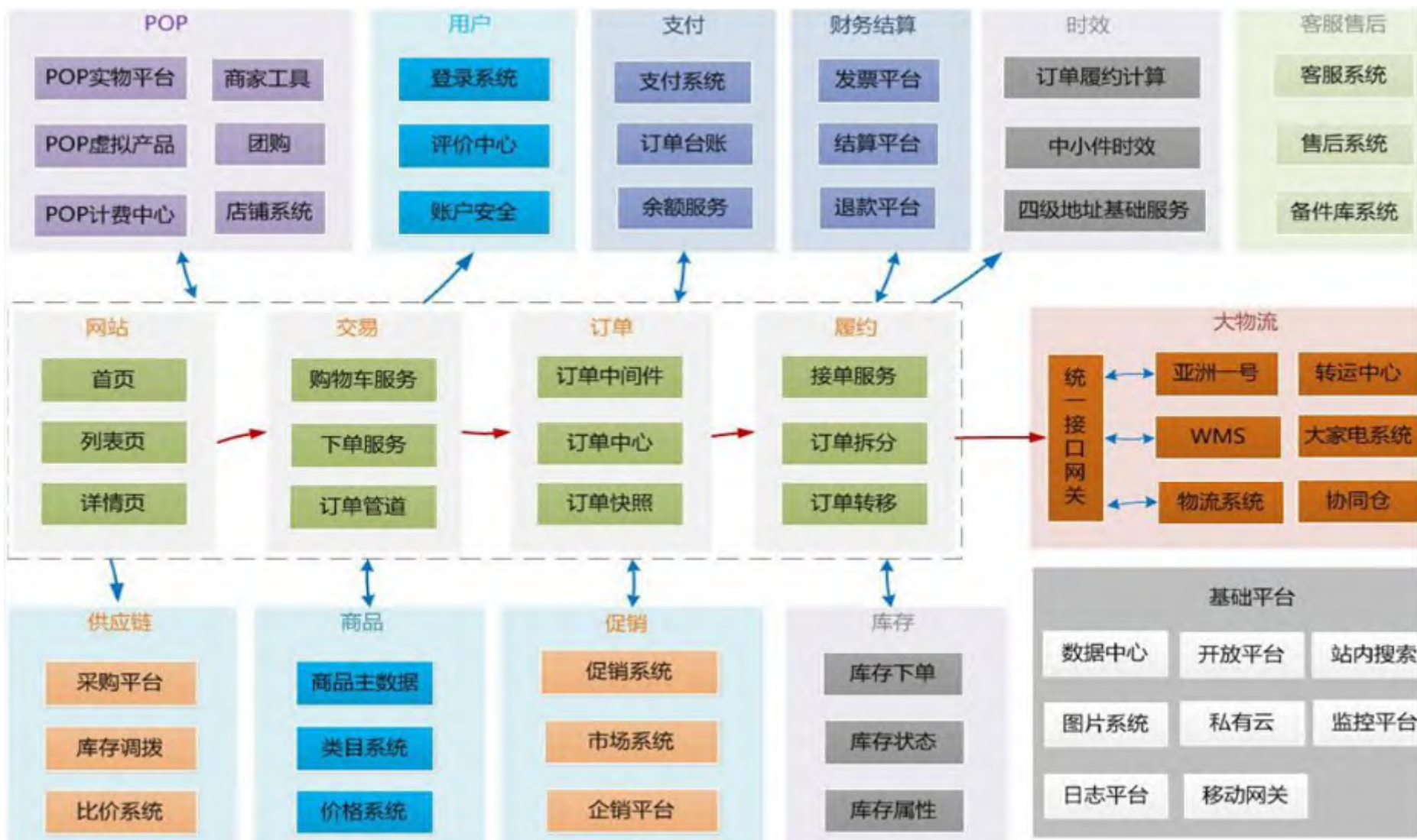


# 京东大规模高访问量下的技术及架构支撑

田琪





## 架构原则

### 1 稳定性原则

- 一切以稳定为中心
- 架构尽可能简单、清晰
- 不过度设计

### 2 解耦/拆分

- 稳定部分与易变部分分离
- 核心业务与非核心业务分离
- 电商主流程与辅流程分离
- 应用与数据分离
- 服务与实现细节分离

### 3 抽象化

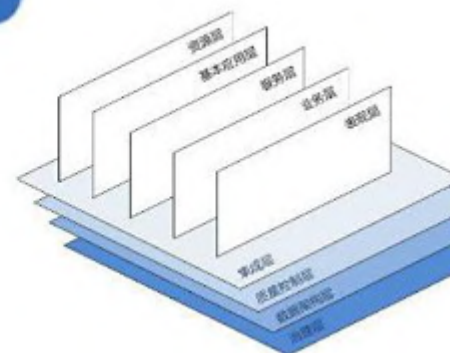
- 应用抽象化：应用只依赖服务抽象，不依赖服务实现细节、位置
- 数据库抽象化：应用只依赖逻辑数据库，不需要关心物理库的位置和分片
- 服务器抽象化：应用虚拟化部署，不需要关心实体机配置，动态调配资源

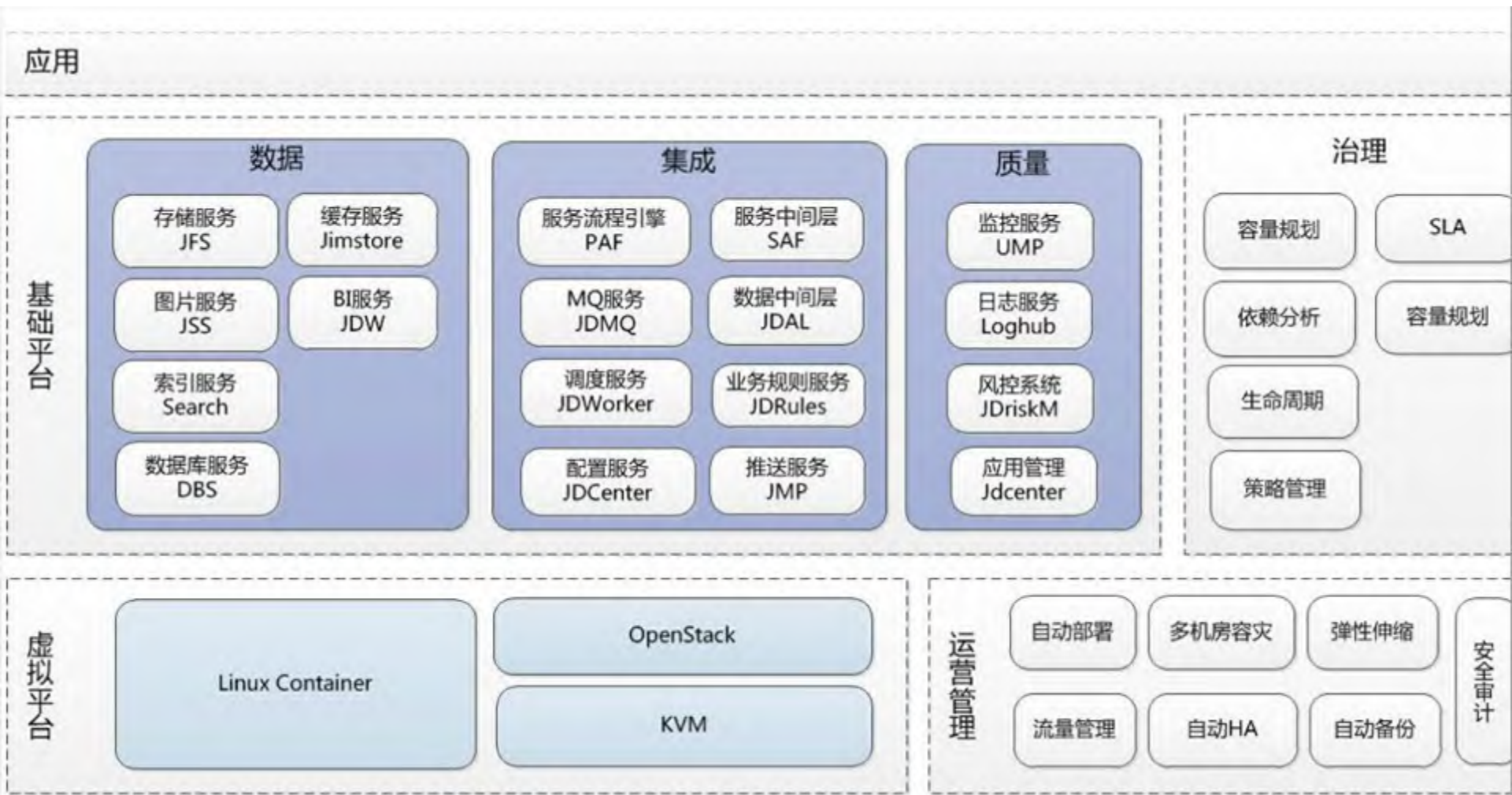
### 4 松耦合

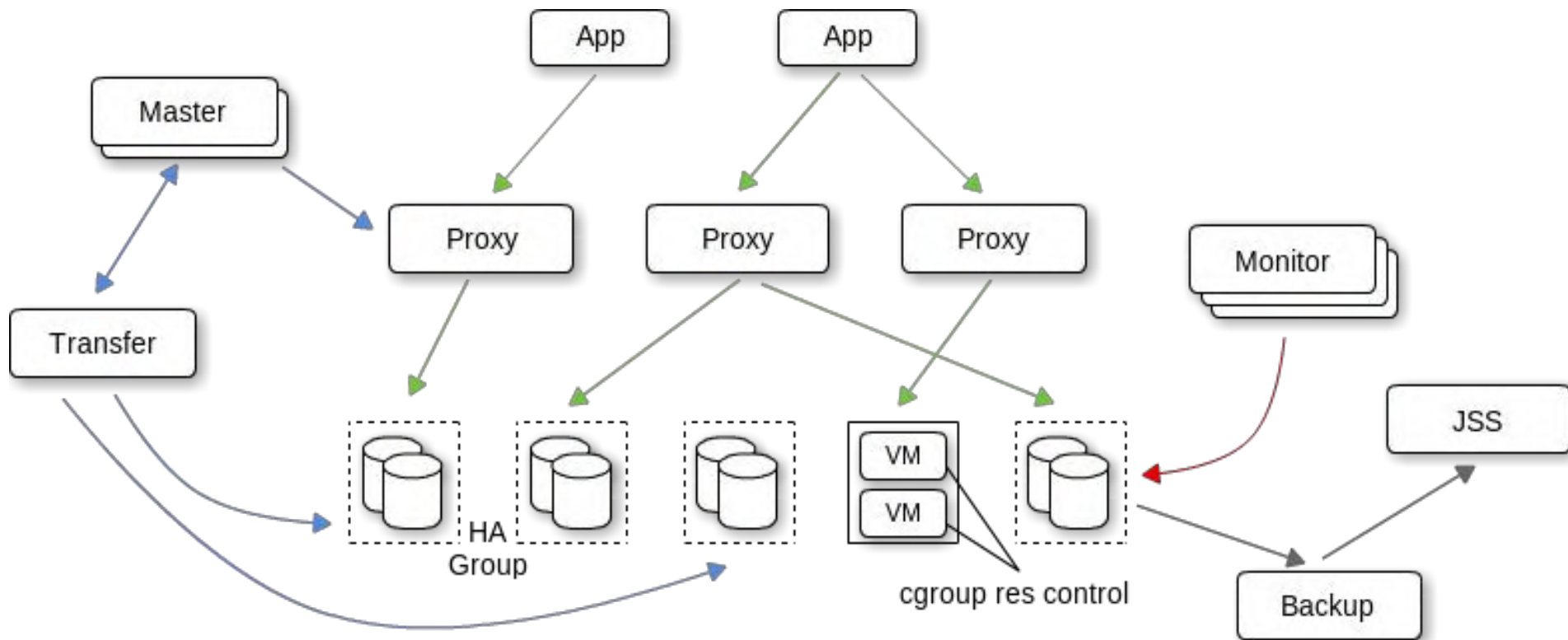
- 跨域调用异步化：不同业务域之间尽量异步解耦。
- 非核心业务尽量异步化：核心、非核心业务之间，尽量异步解耦
- 必须同步调用时，需要设置超时时间和任务队列长度

### 5 容错设计

- 服务自治：服务能彼此独立修改、部署、发布和管理。避免引发连锁反应
- 集群容错：应用系统集群，避免单点
- 多机房容灾：多机房部署，多活







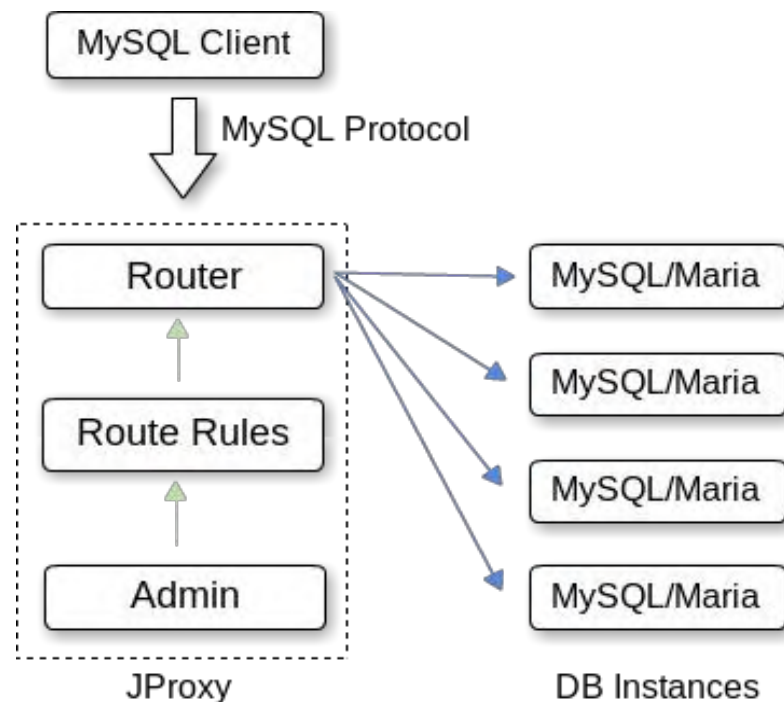
- 支持MySQL, MariaDB, MongoDB等数据库
- 服务高可用，主库故障，从库自动切换
- 数据高可靠，定期快照备份，增量备份
- 数据自动拆分，一键无缝迁移扩容
- 针对特殊业务需求，定制优化特殊的数据库版本

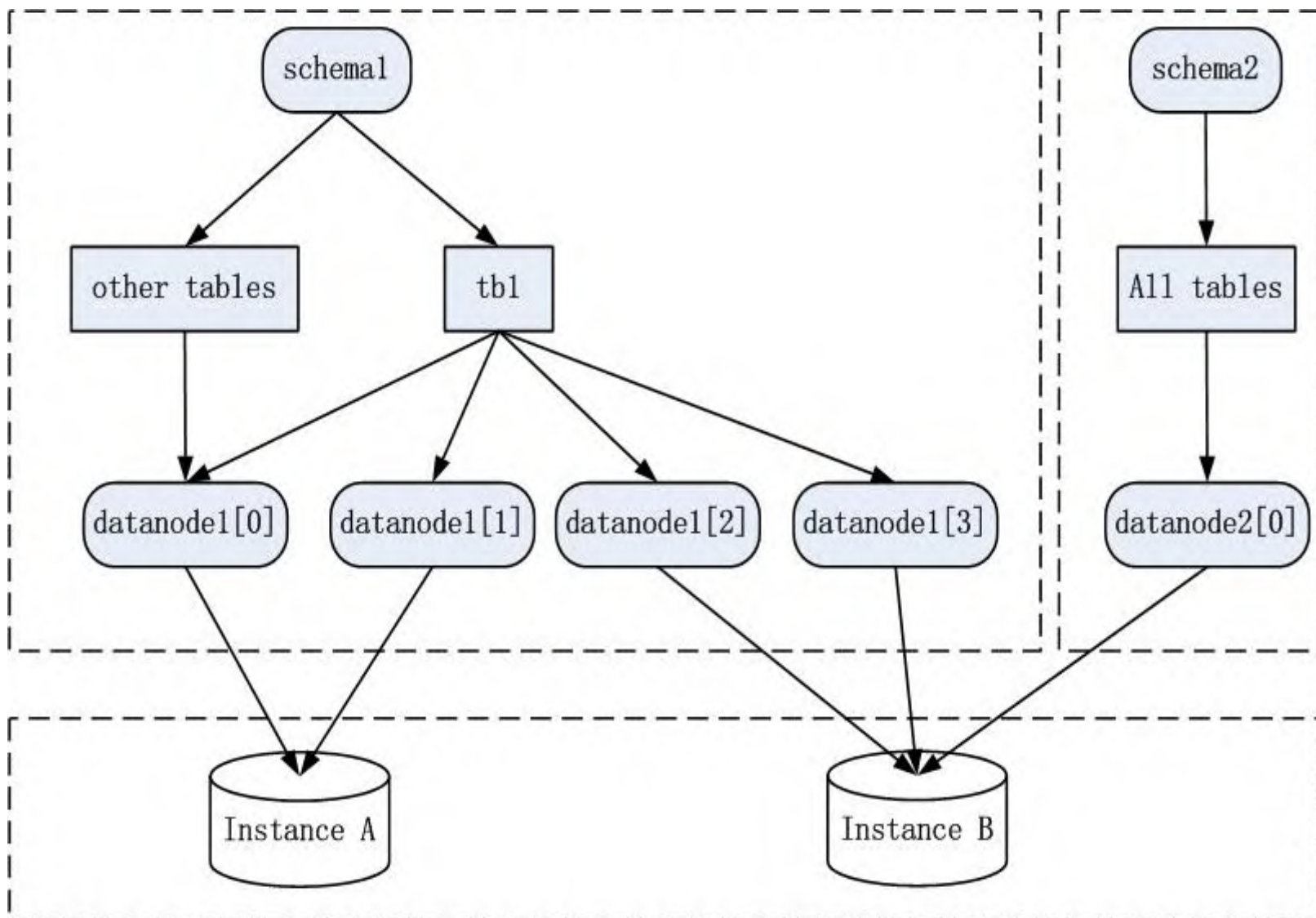


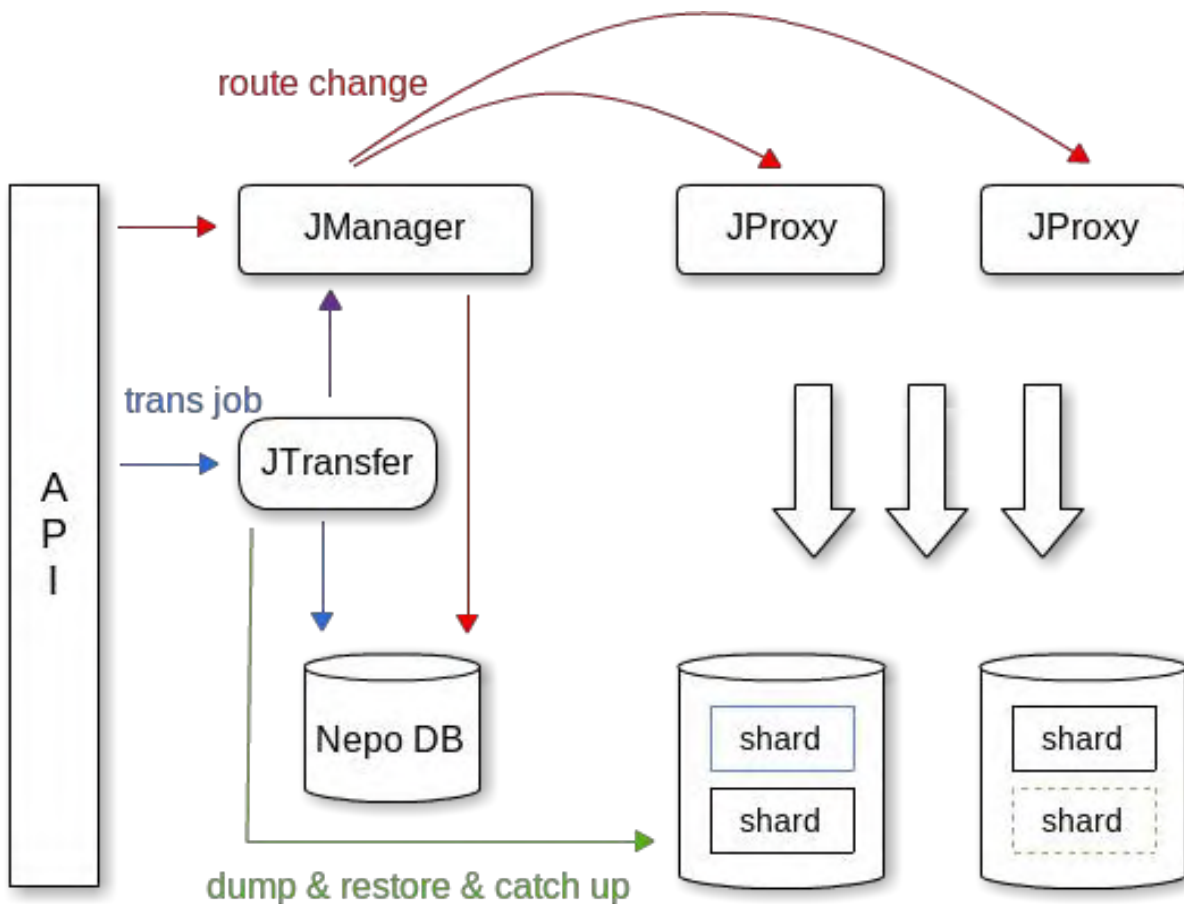
原生MySQL协议，接入使用标准MySQL客户端

数据根据路由规则分库分表，对业务访问透明

单库容量满，可以快速在线无缝迁移，不影响业务

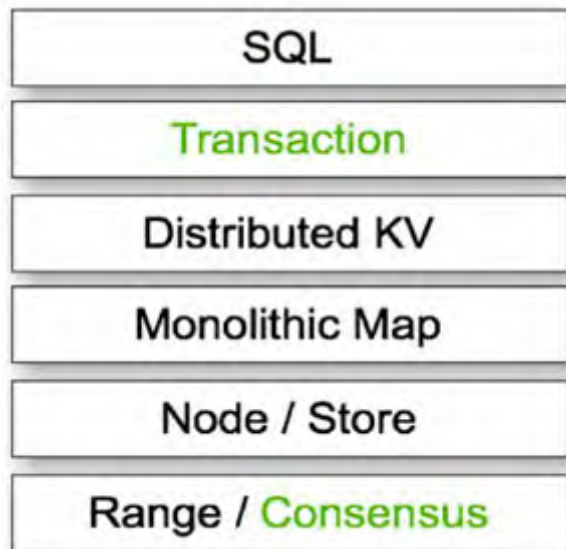




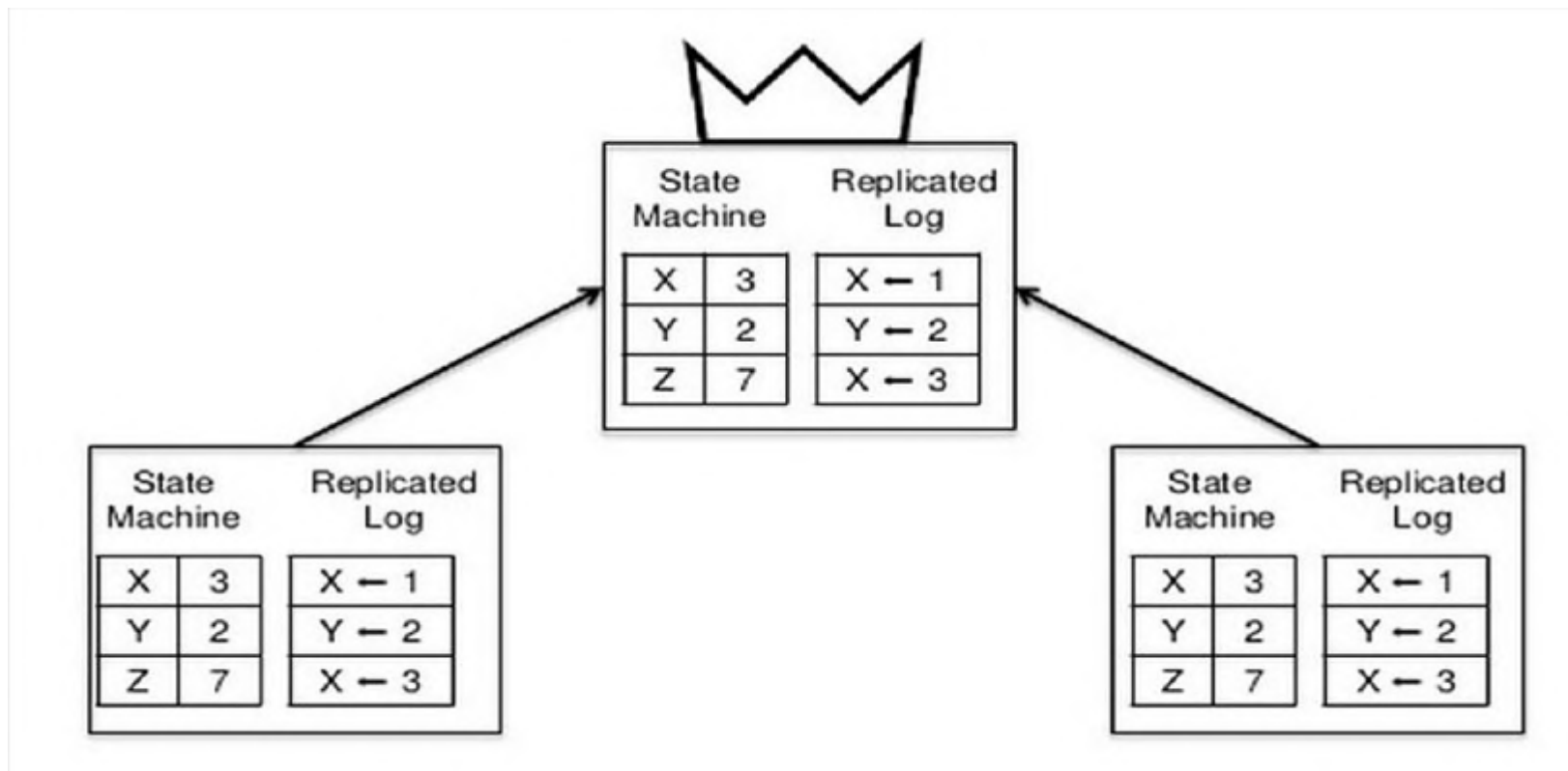


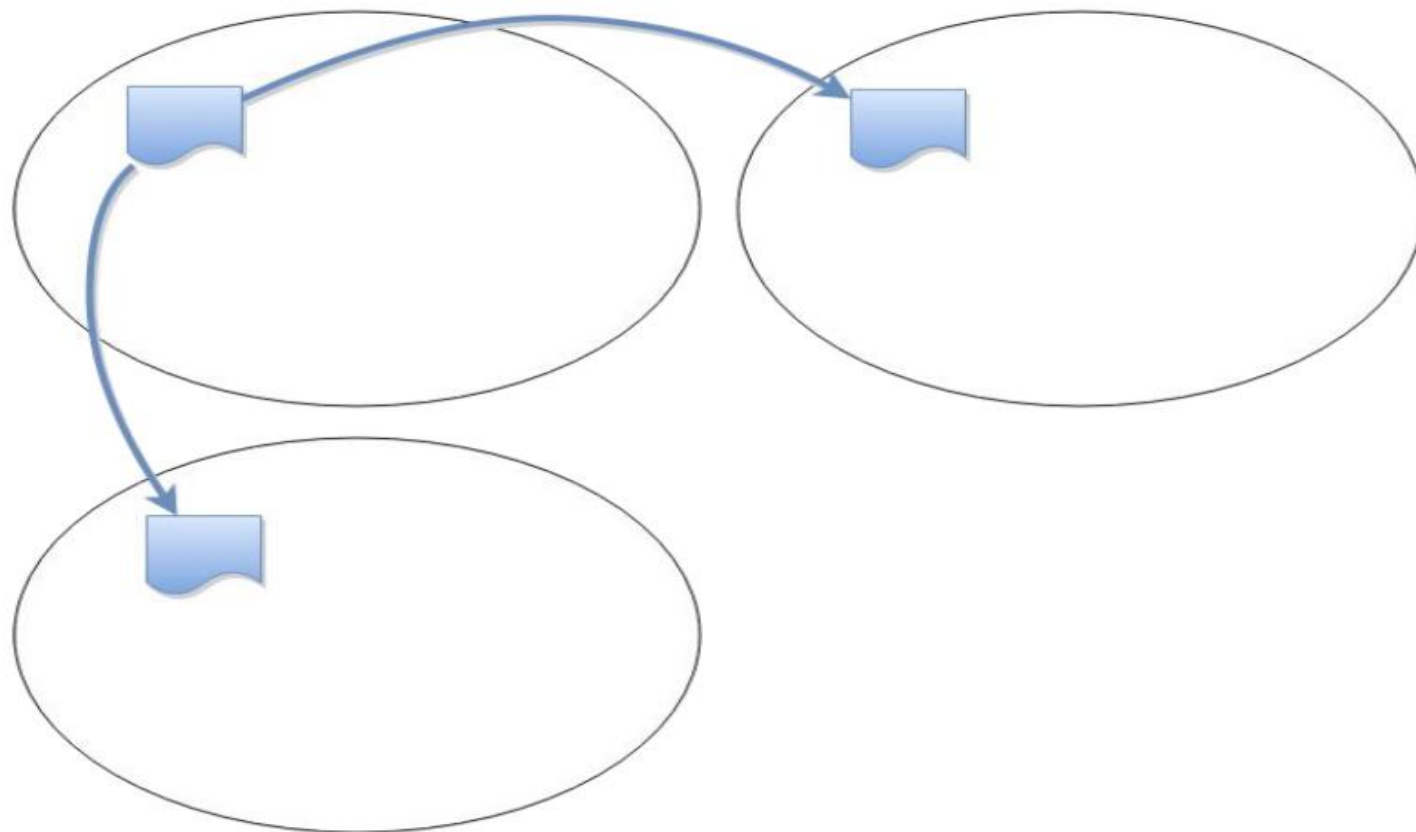
- 应用不关心
  - sharding
  - failover
  - eventual consistency
- 应用需要事务
- 应用需要SQL
- 应用需要ACID

- 空间分为若干个range
- 使用RSM保障一致性
- 每个node由range组成
- 对外暴露一个完整的map
- 在map基础上实现分布式事务

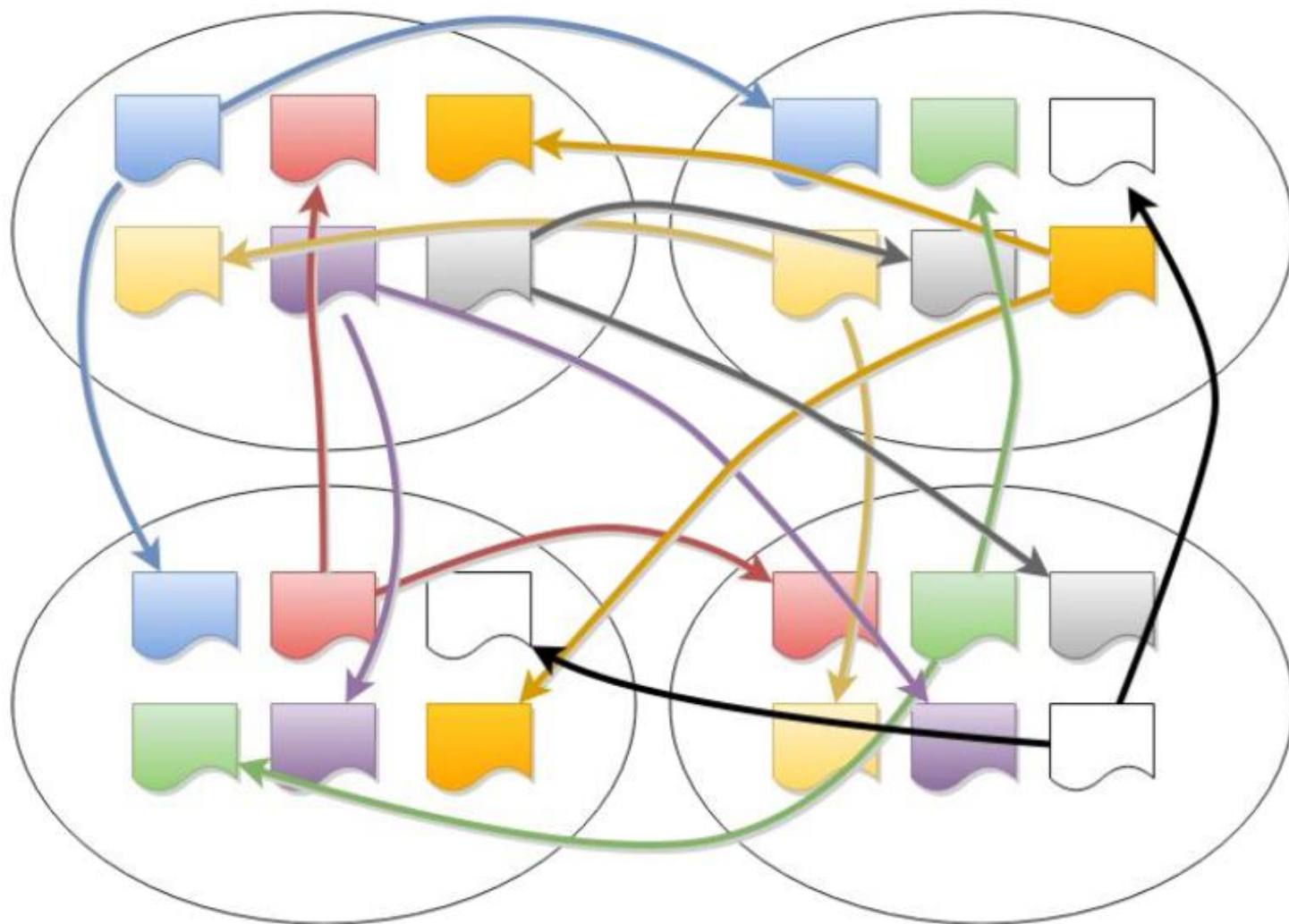


- 至少支持snapshot isolation级别
- 使用2阶段提交协议+2阶段锁协议
- 数据多版本并存
- 每个key由时间戳+真实key组成
- 提交时间有可能会被推迟







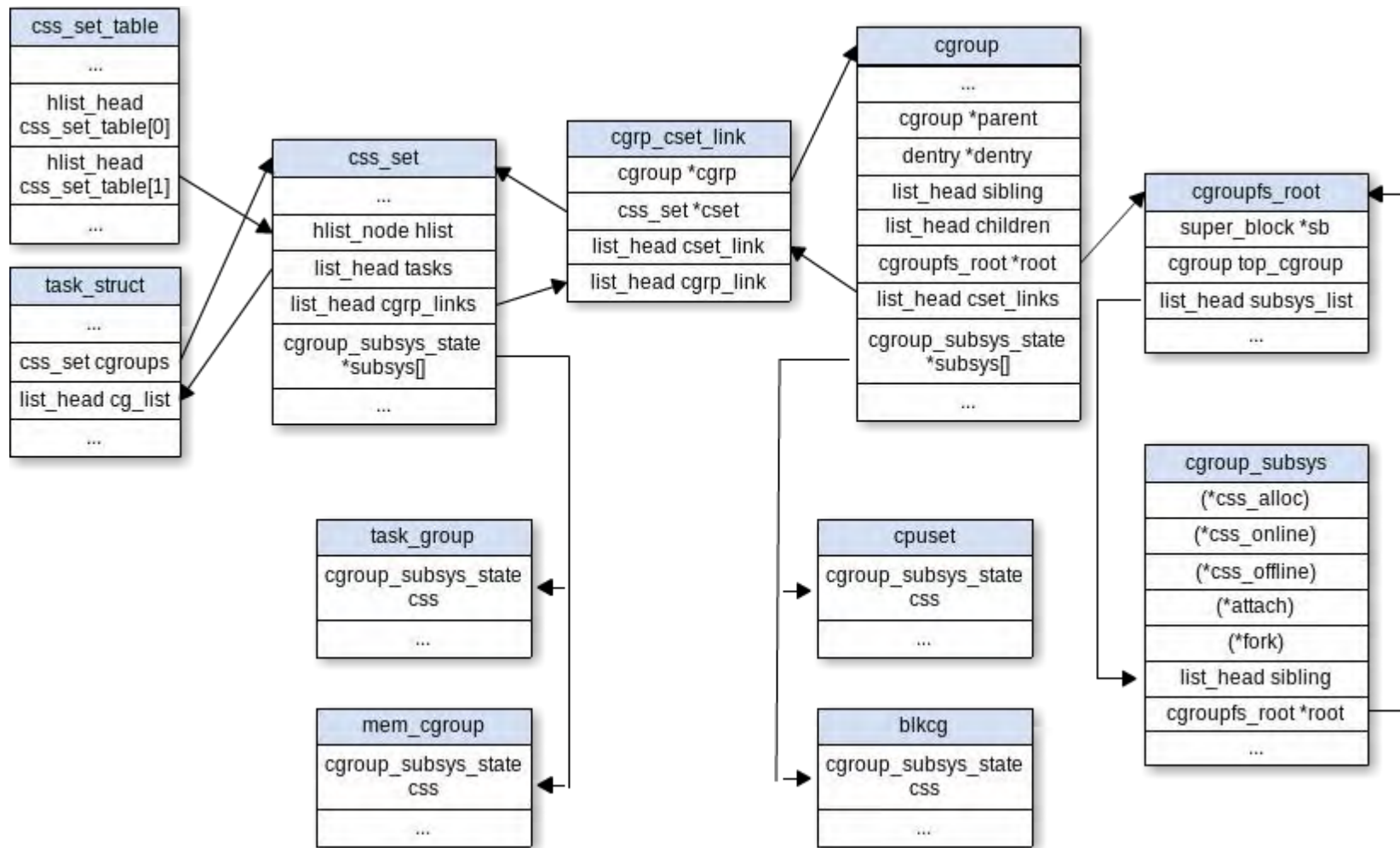


- clone/setns/unshare三个系统调用是容器资源隔离技术的核心实现
- 内核新增nsproxy数据结构使不同进程可以拥有各自不同的命名空间

```
struct nsproxy {
    atomic_t count;
    struct uts_namespace *uts_ns;
    struct ipc_namespace *ipc_ns;
    struct mnt_namespace *mnt_ns;
    struct pid_namespace *pid_ns_for_children;
    struct net *net_ns;
};
```

```
struct mnt_namespace {
    atomic_t count;
    unsigned int proc_inum;
    struct mount *root;
    struct list_head list;
    struct user_namespace *user_ns;
    u64 seq; /* Sequence number */
    wait_queue_head_t poll;
    int event;
};
```

CLONE_NEWNS	2.4.19	CAP_SYS_ADMIN
CLONE_NEWUTS	2.6.19	CAP_SYS_ADMIN
CLONE_NEWIPC	2.6.19	CAP_SYS_ADMIN
CLONE_NEWPID	2.6.24	CAP_SYS_ADMIN
CLONE_NEWNET	2.6.29	CAP_SYS_ADMIN
CLONE_NEWUSER	3.8	No capability is required



- 依赖较新内核版本（lxc需要至少3.8 以上）
- IO控制方面问题比较多
  - Buffer IO如何处理
  - IO带宽控制只能用CFQ IO调度器，不适合高速硬件设备
  - deadline IO调度器只支持限流,缺乏弹性
- 混合存储之后应该在哪一层进行限制

谢谢

JD.COM 京东

Q & A

[www.jd.com](http://www.jd.com)

