

# 七牛云大数据平台的架构与演进

陈超

七牛云 技术总监



- Pandora简介
- Workflow设计与优化
- 总结

- **Pandora简介**
- Workflow设计与优化
- 总结





# 实时工作流引擎

产品列表

数据统计 文档中心 工单 站内信 个人面板

## 实时计算 workflows

plugin列表 更新

数据源

计算任务

消息队列

- HTTP
- 日志检索
- 时序数据库
- 对象存储

导出任务

计算任务

查看数据

复制

粘贴

### 数据源

名称 \*

lit\_test\_1

字段 \*

csv导出

字段名称 *	类型 *
ip	string
date	date
name	string

添加新字段

# 离线工作流引擎

88 产品列表

数据统计 工单 站内信 | 个人面板

离线计算 workflow

计算任务列表 查看日志 更新

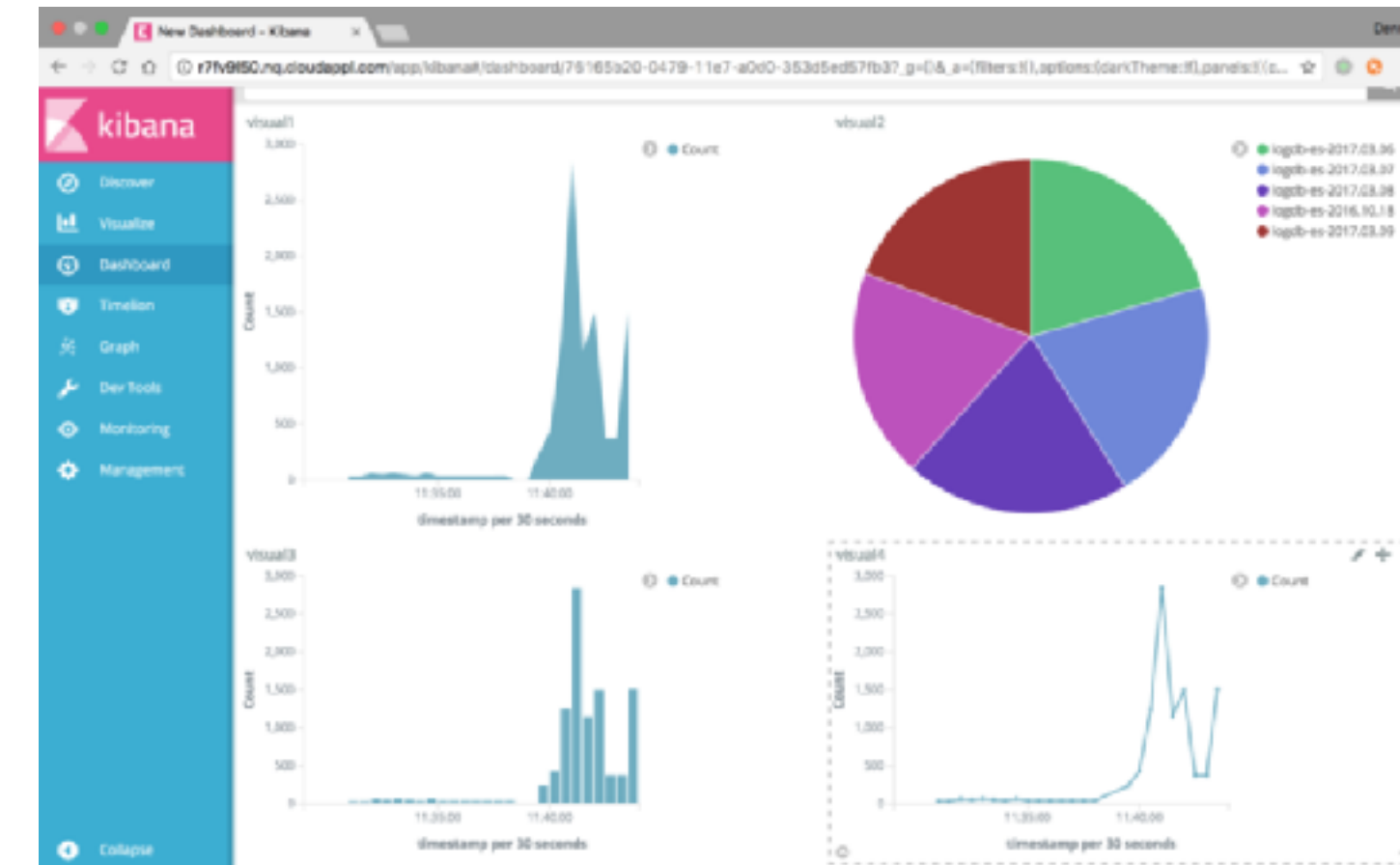
数据源

计算任务

数据加载中

# 日志检索服务(LogDB)

- 高扩展及高稳定的日志检索SaaS服务
- 兼容ES查询语法
- 与开源生态紧密结合
- 已经经过超大规模数据量检验





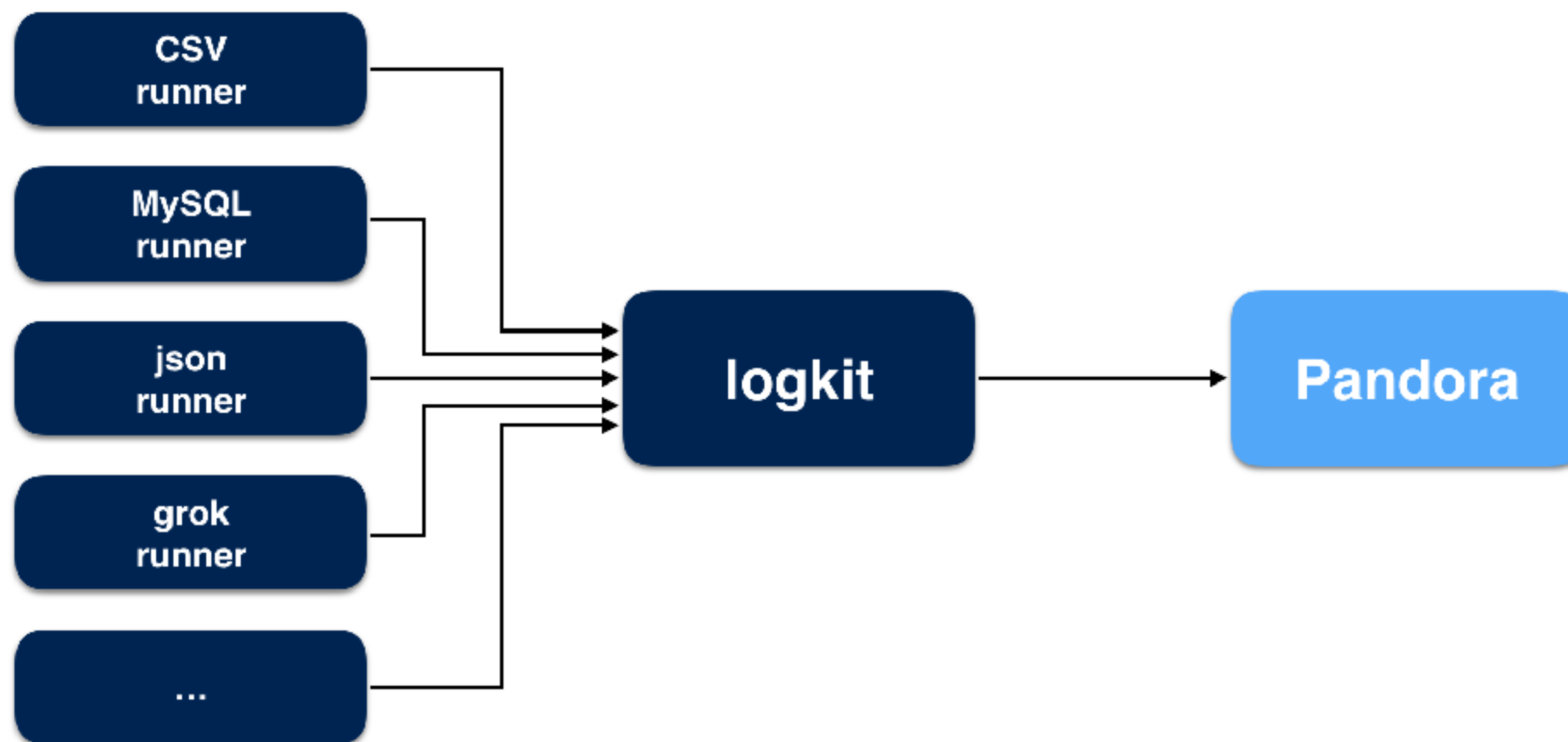
## 其它服务

- **TSDB**: 高性能分布式时序数据库
- **BI Studio**: 报表工作室
- **XSpark**: Spark服务



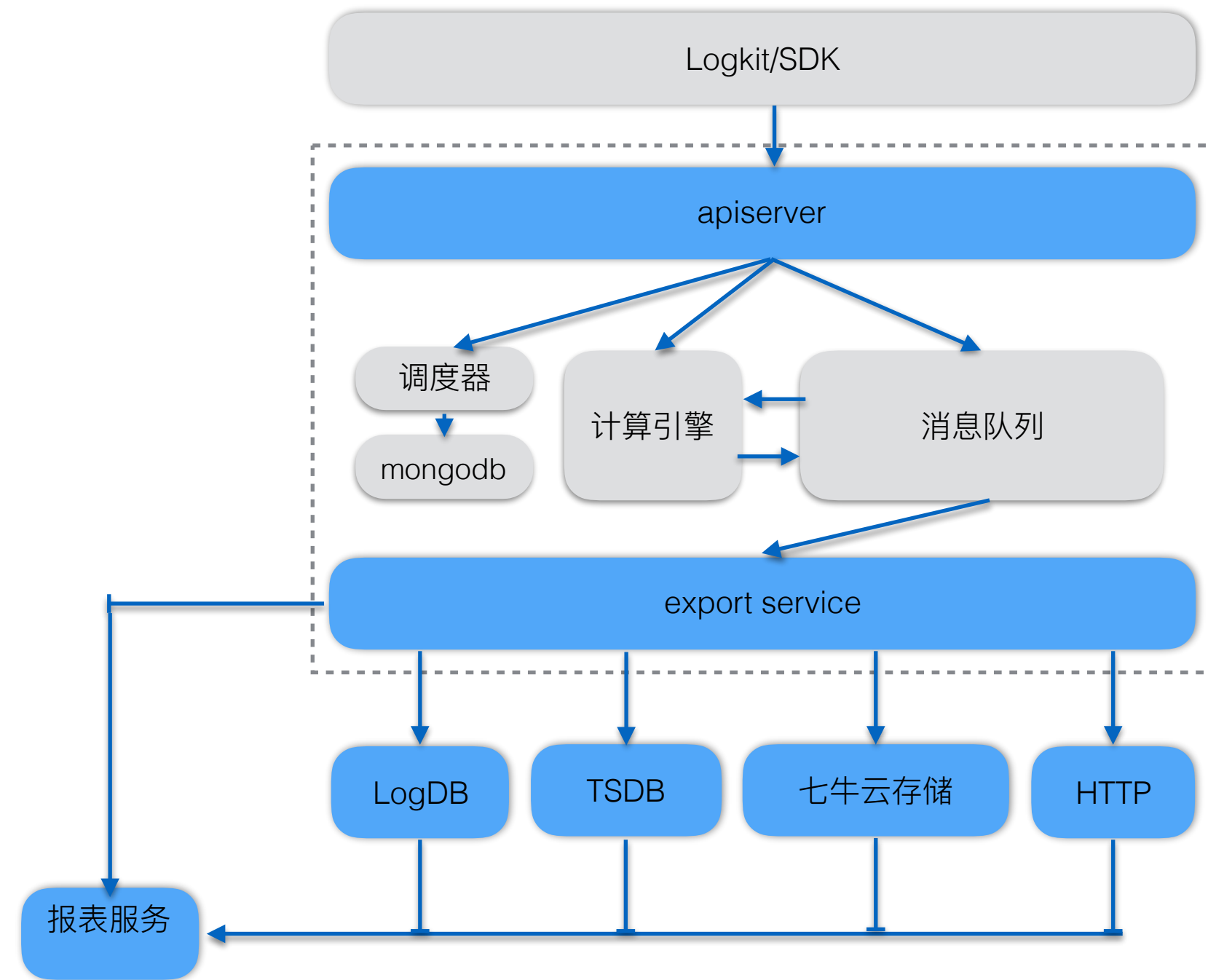
## 特别介绍(Logkit)

- Logkit是七牛Pandora开发的一个通用的数据采集工具，可以将不同数据源的数据方便的发送到 Pandora进行数据分析，除了基本的数据发送功能，Logkit还有容错、并发、监控、删除等功能。
- Logkit支持文件、MySQL、MSSQL、ES、MongoDB、Kafka 及Redis等一系列数据源。
- [github.com/qiniu/logkit](https://github.com/qiniu/logkit)



- Pandora简介
- **Workflow设计与优化**
- 总结

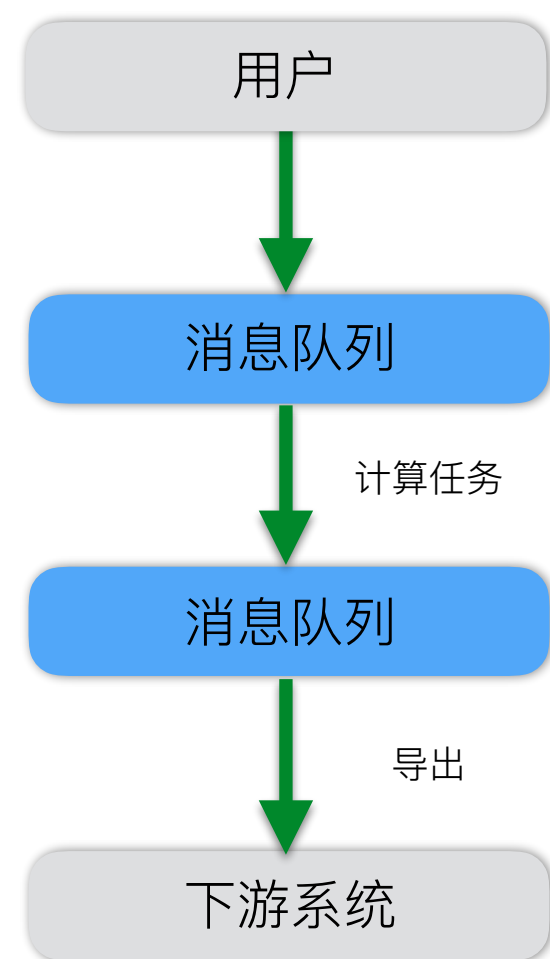
# 架构简图



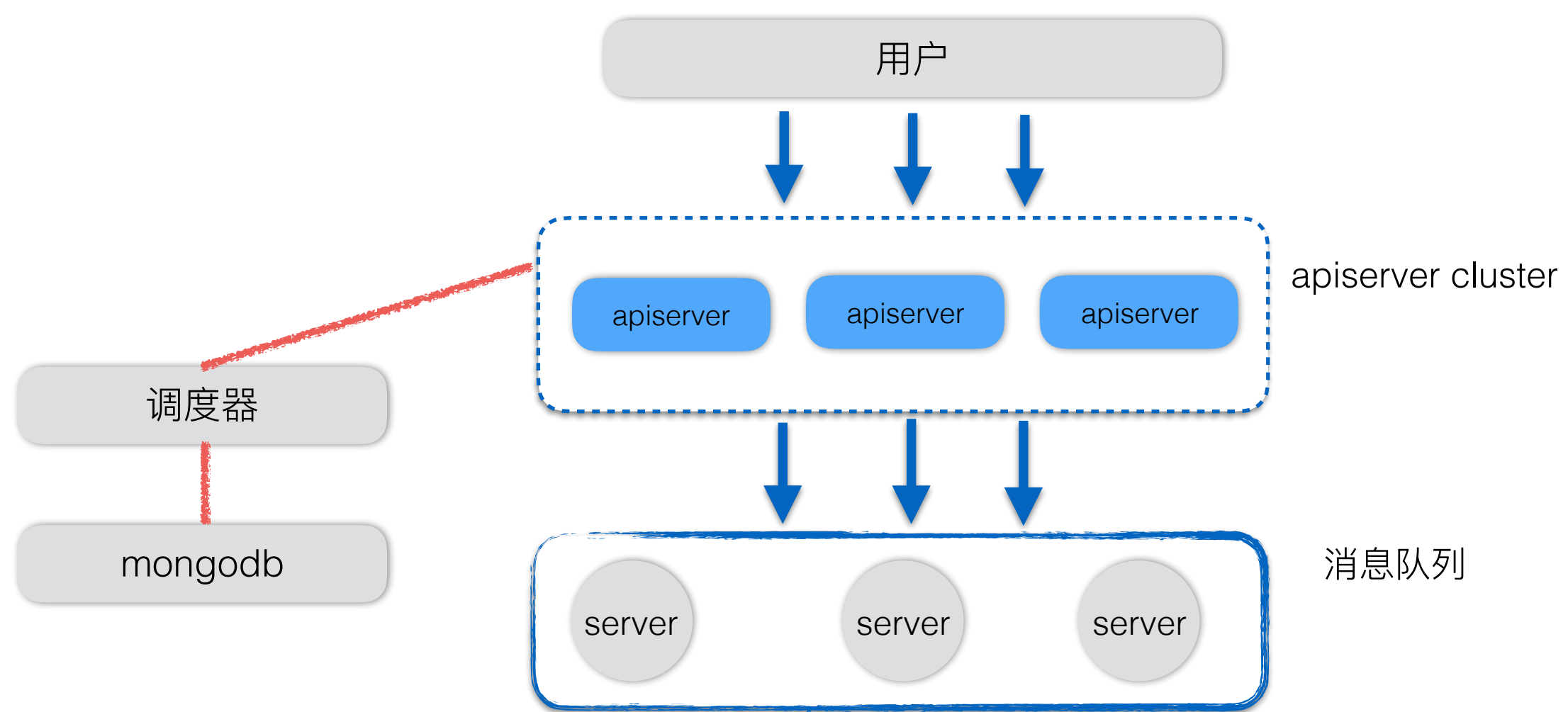
# 数据流剖析

## 一般影响因素

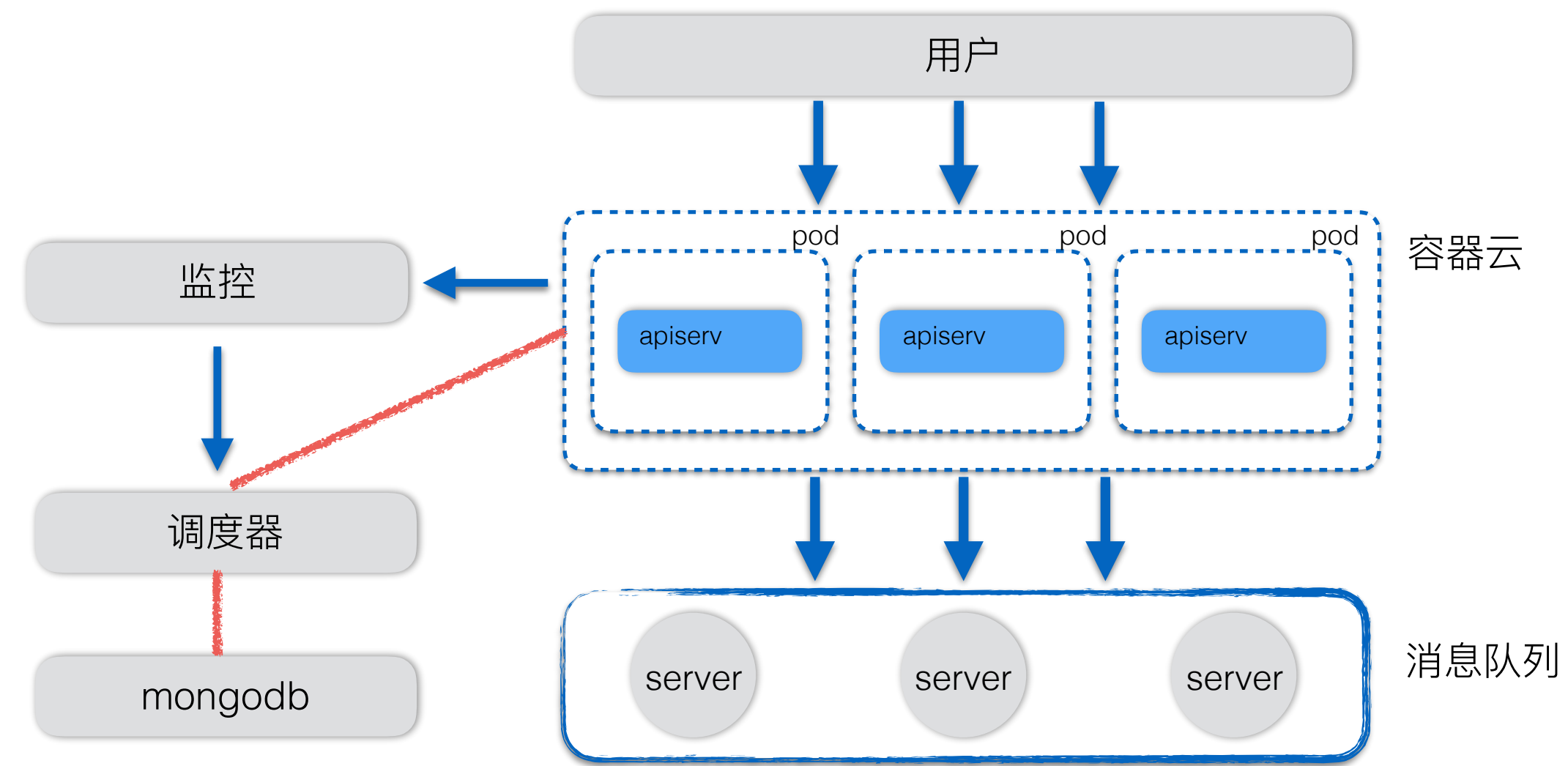
- 资源利用率
- 处理效率
- 木桶效应
- 链路损耗
- 其他



# 数据接入层

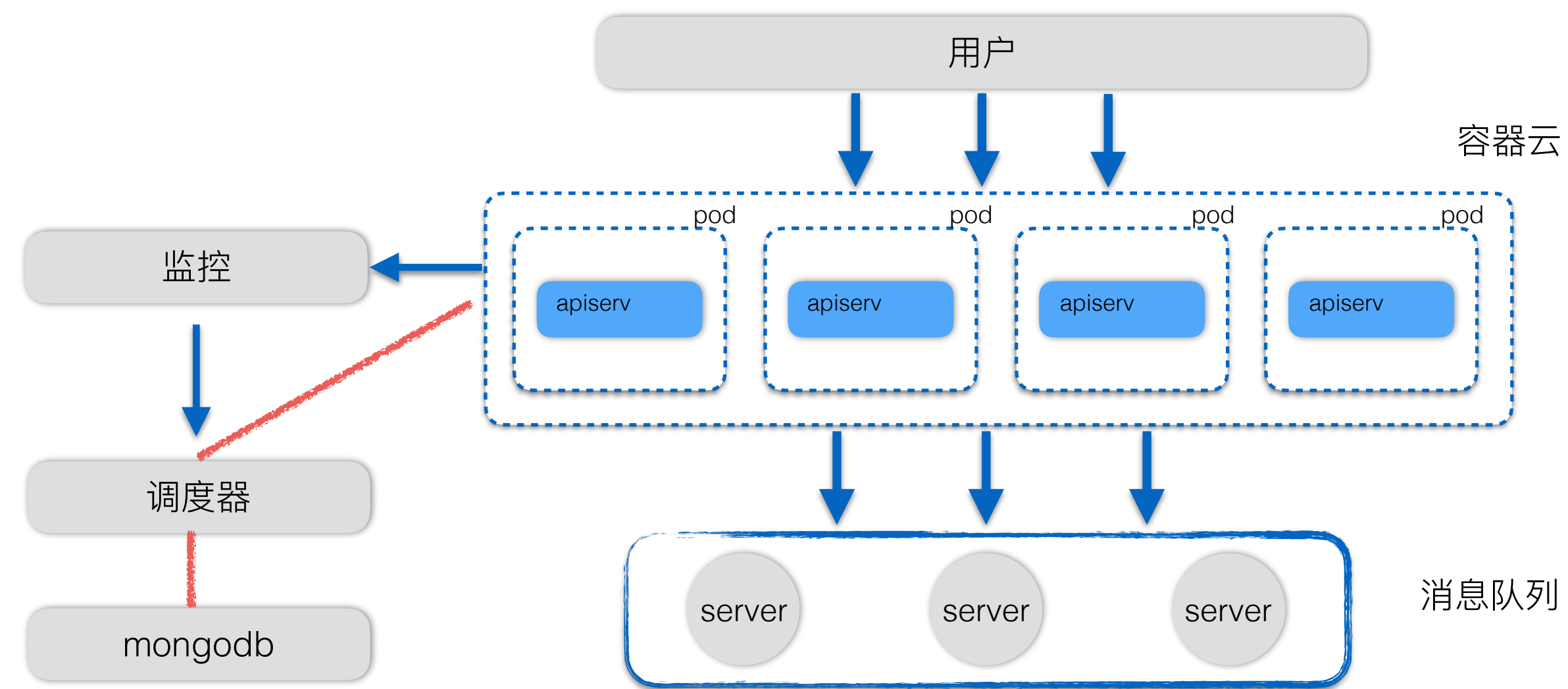


# 容器化



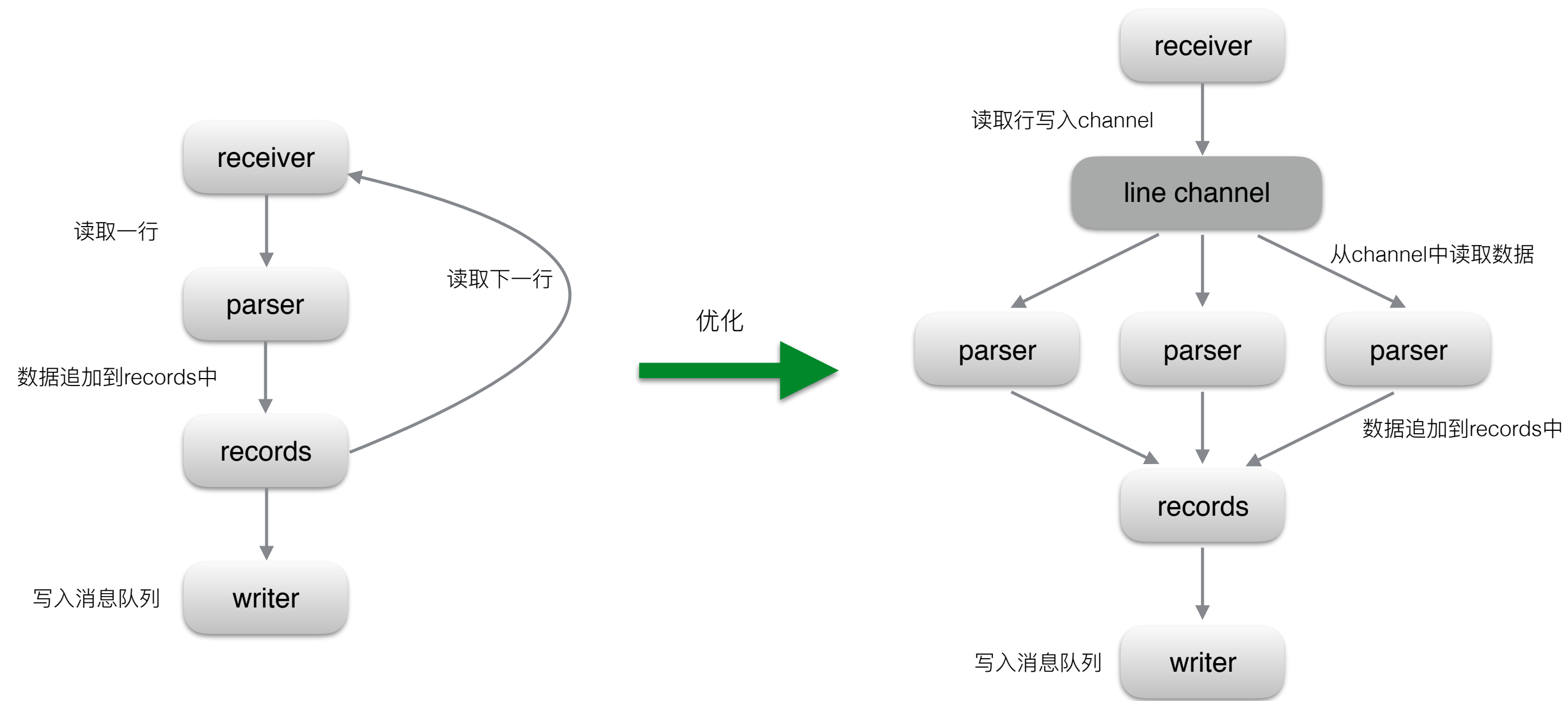
# 动态扩容

- 基于时序数据的监控
- 基于监控数据自动扩容缩容

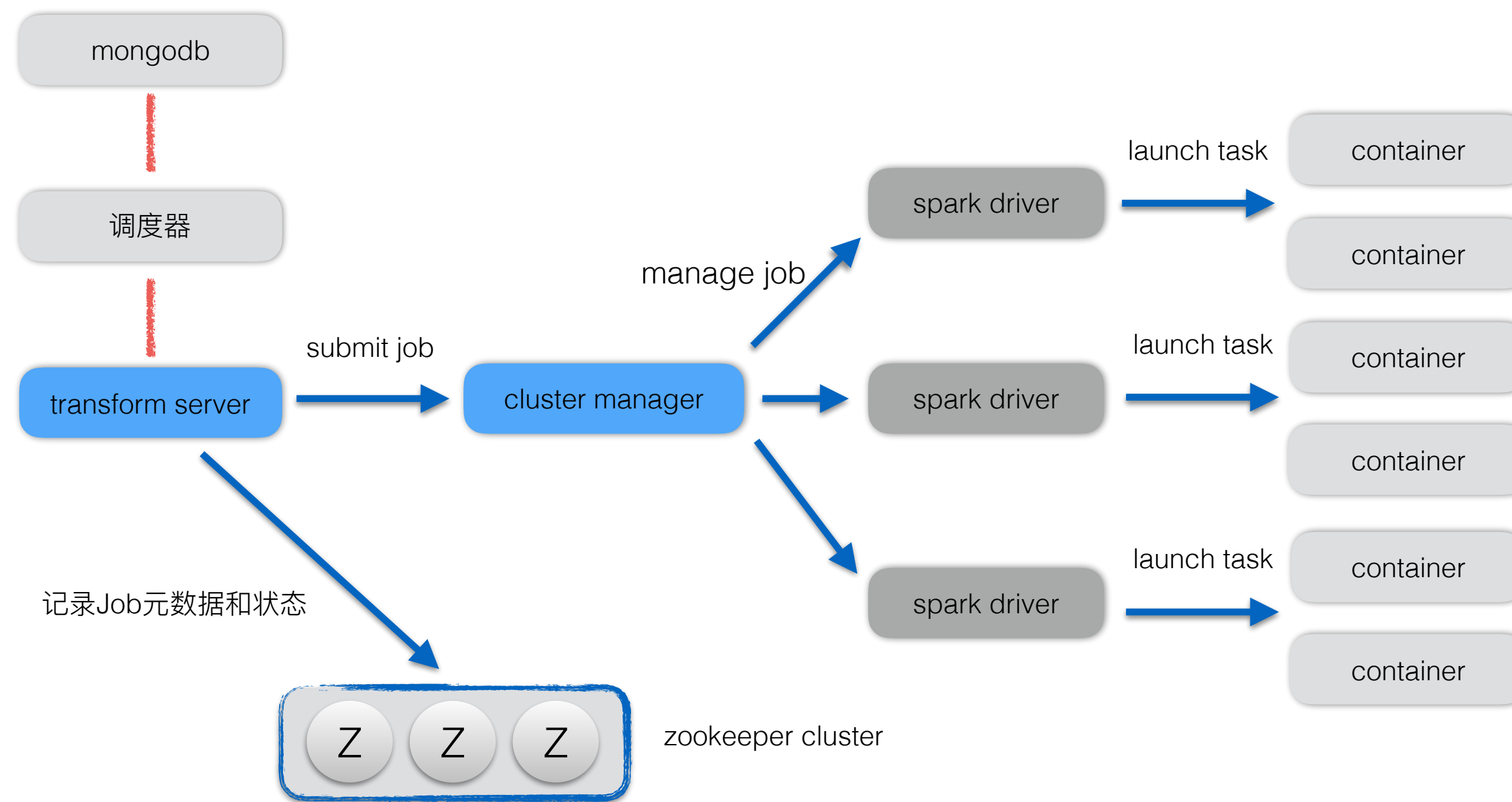




# 数据写入优化



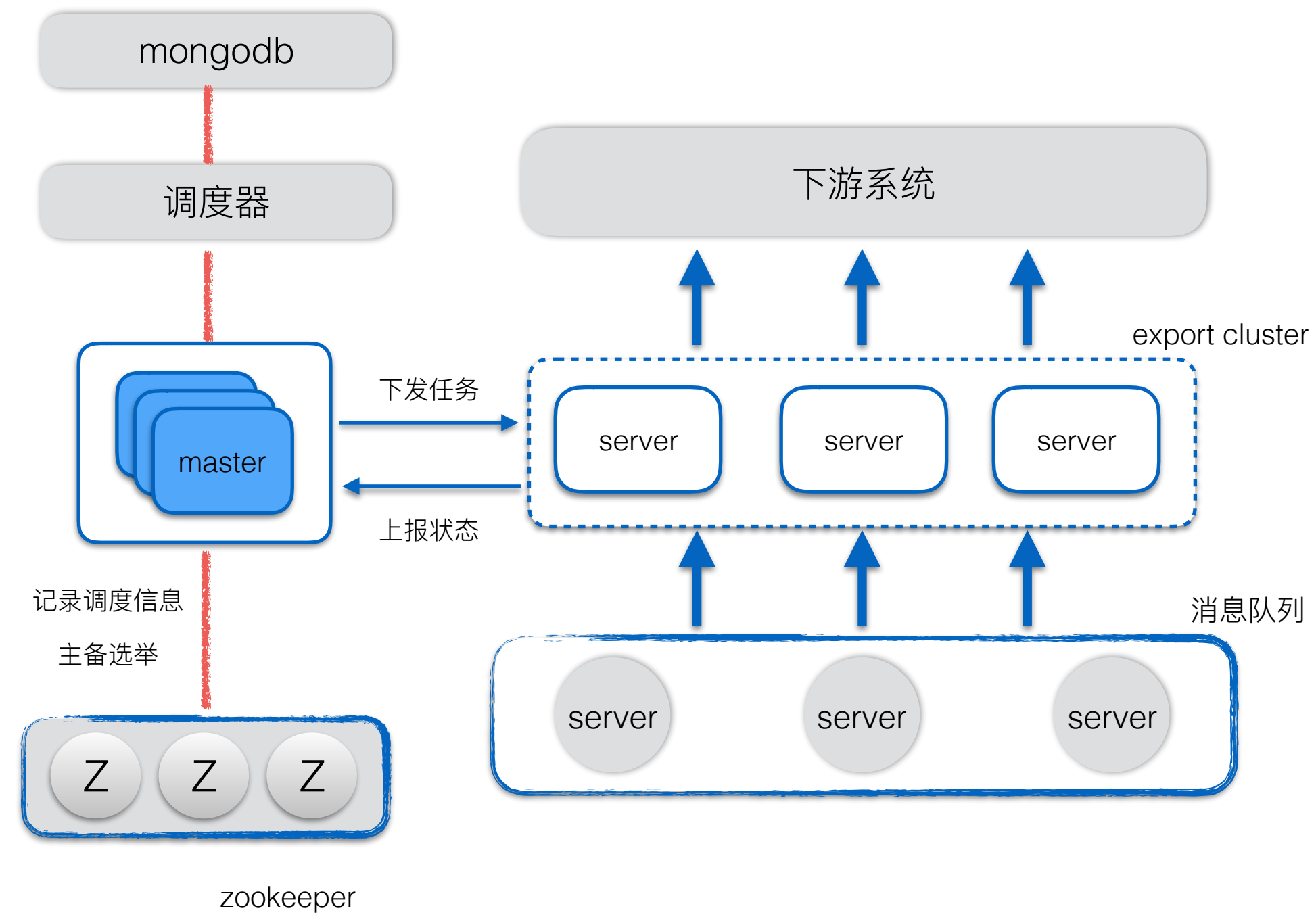
# 计算



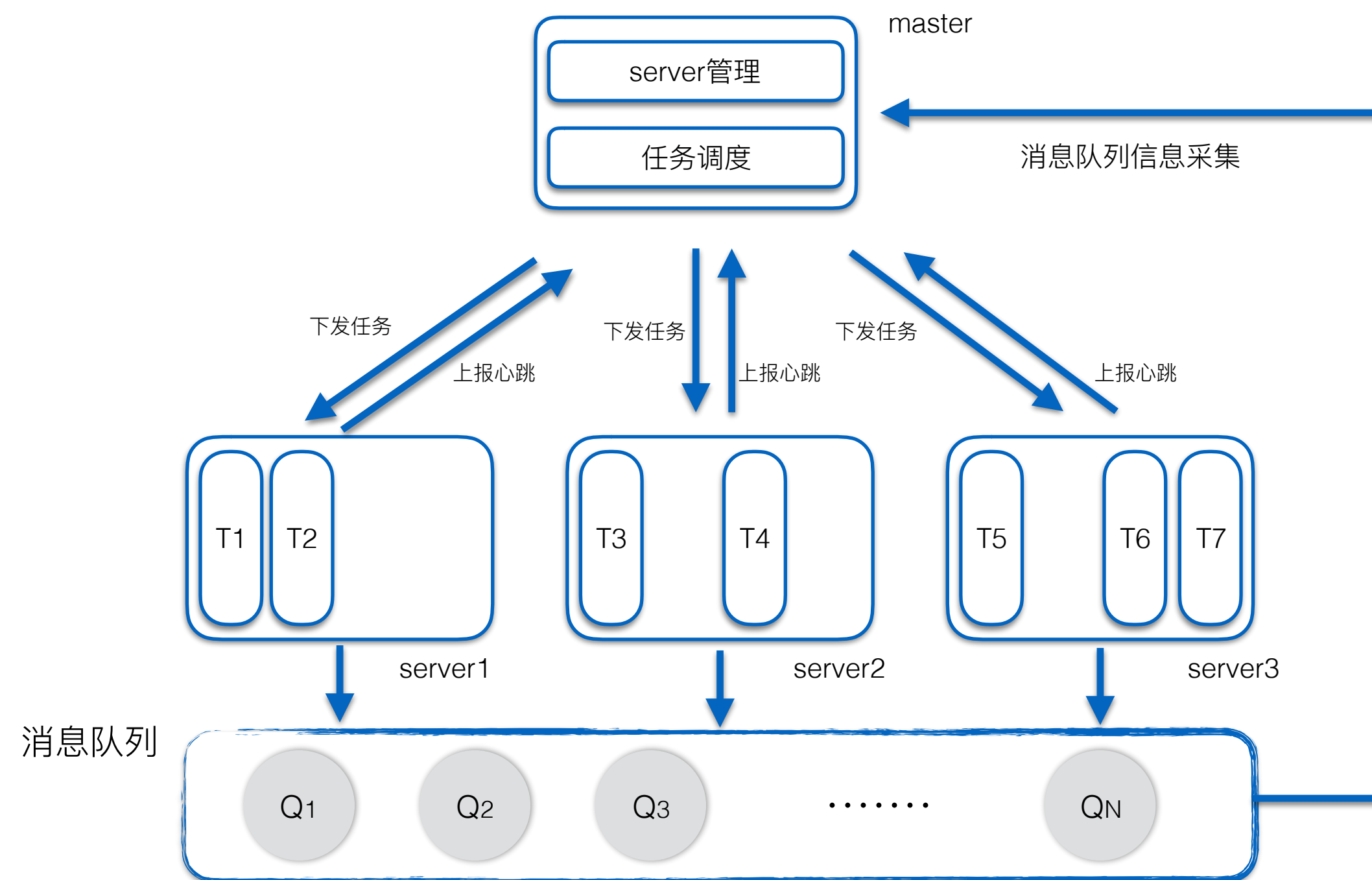
# 导出

连接上下游

- 任务切分
- 调度
- 任务自动均衡
- 水平扩展
- 资源隔离
- 高可用



# 任务切分与管理

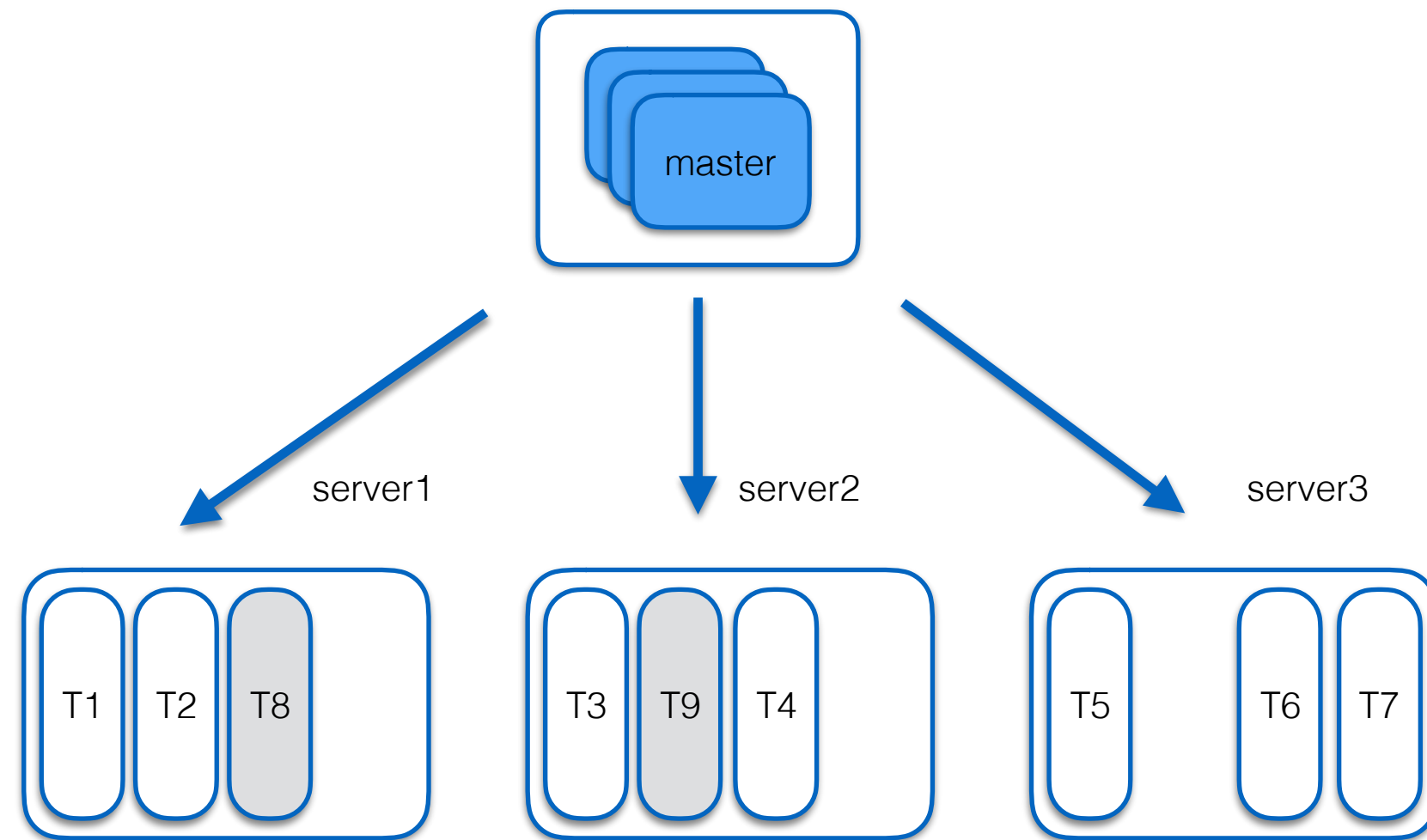


## 调度方法

- 面向资源
- 充分利用异构机器
- 自动调整

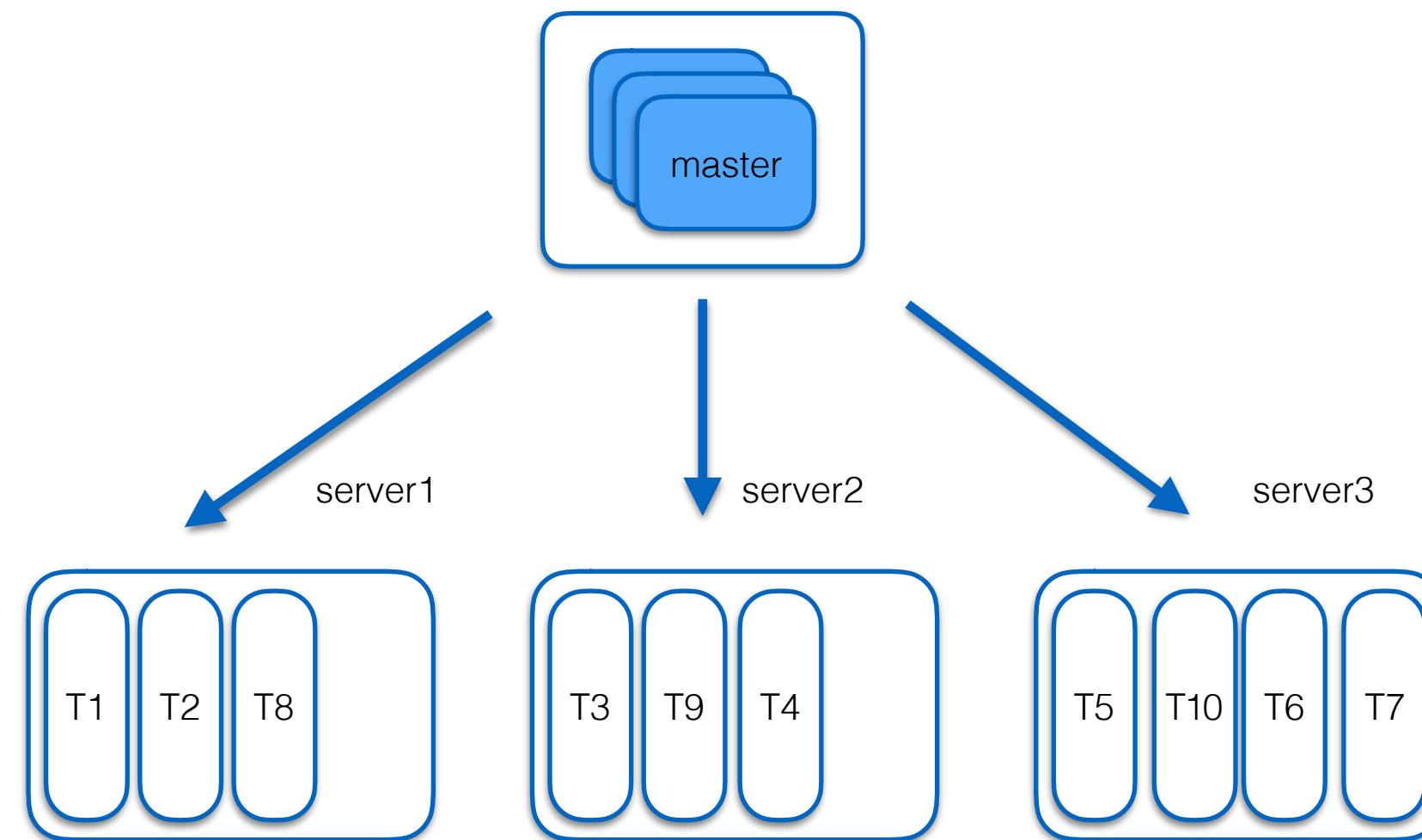
# 任务分配

- 任务均匀分配在server上
- T8和T9加入



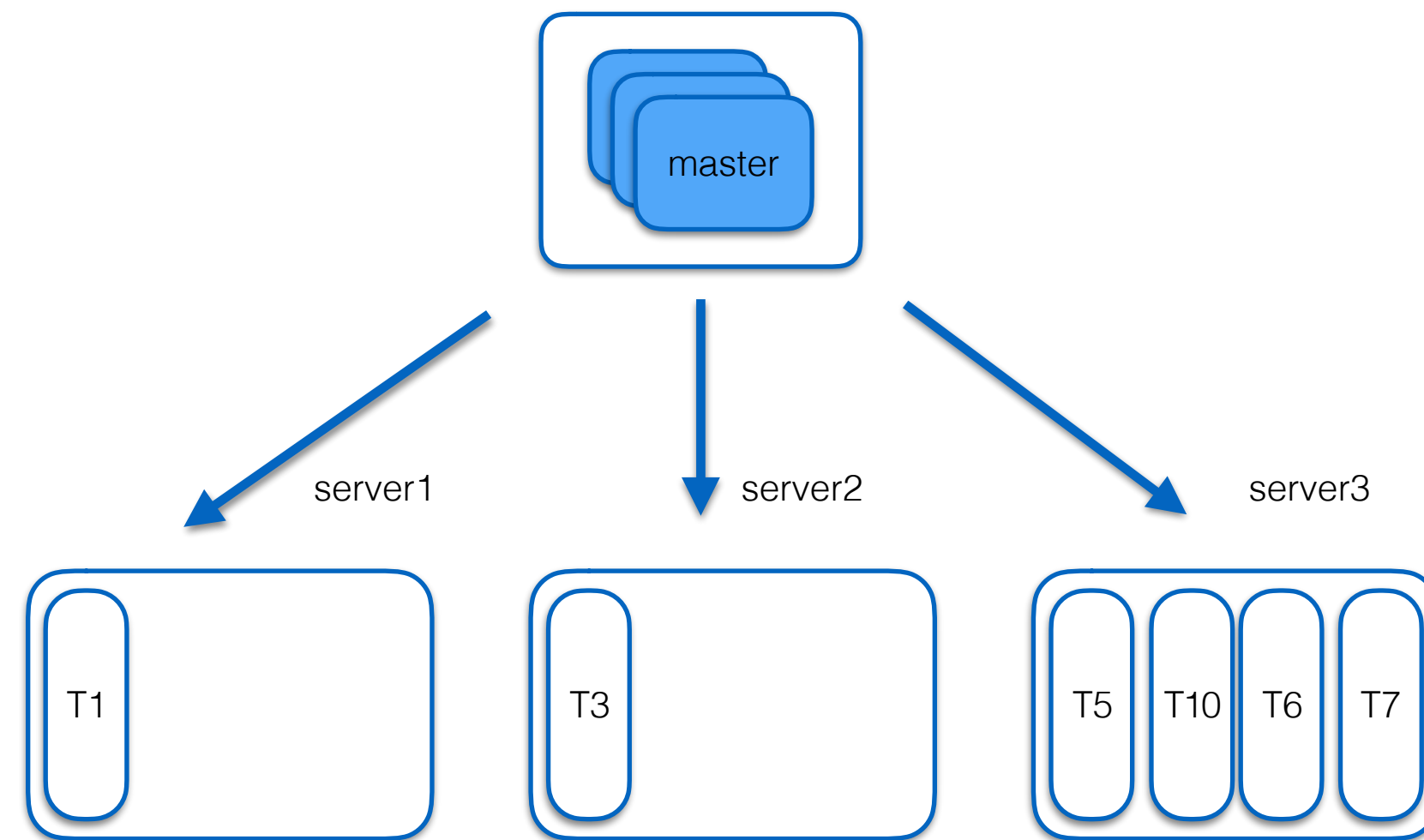
# 自动调整

- 任务均匀分配在3台server上



## 自动调整

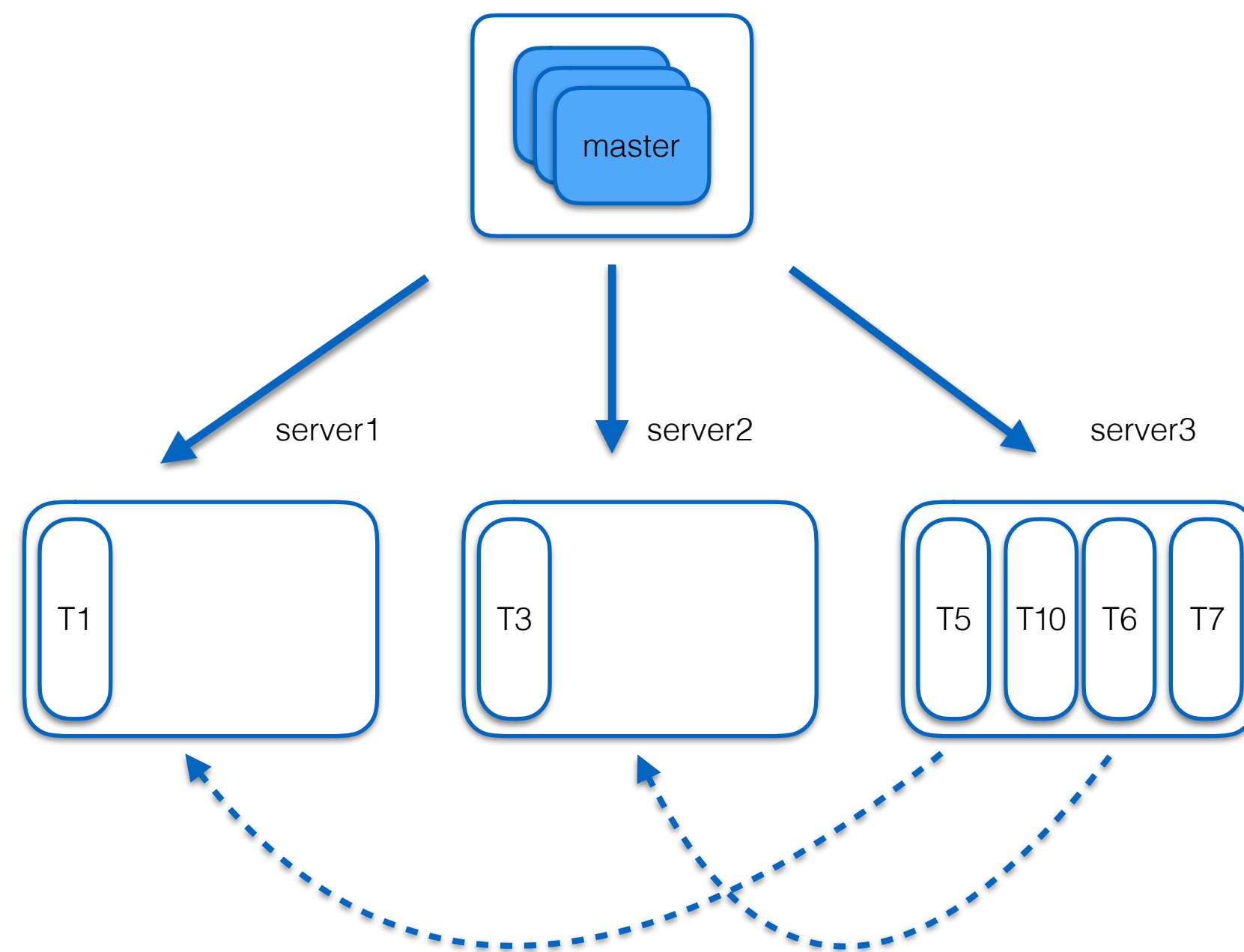
- 任务均匀分配在3台server上
- T2、T8、T4、T9被删除





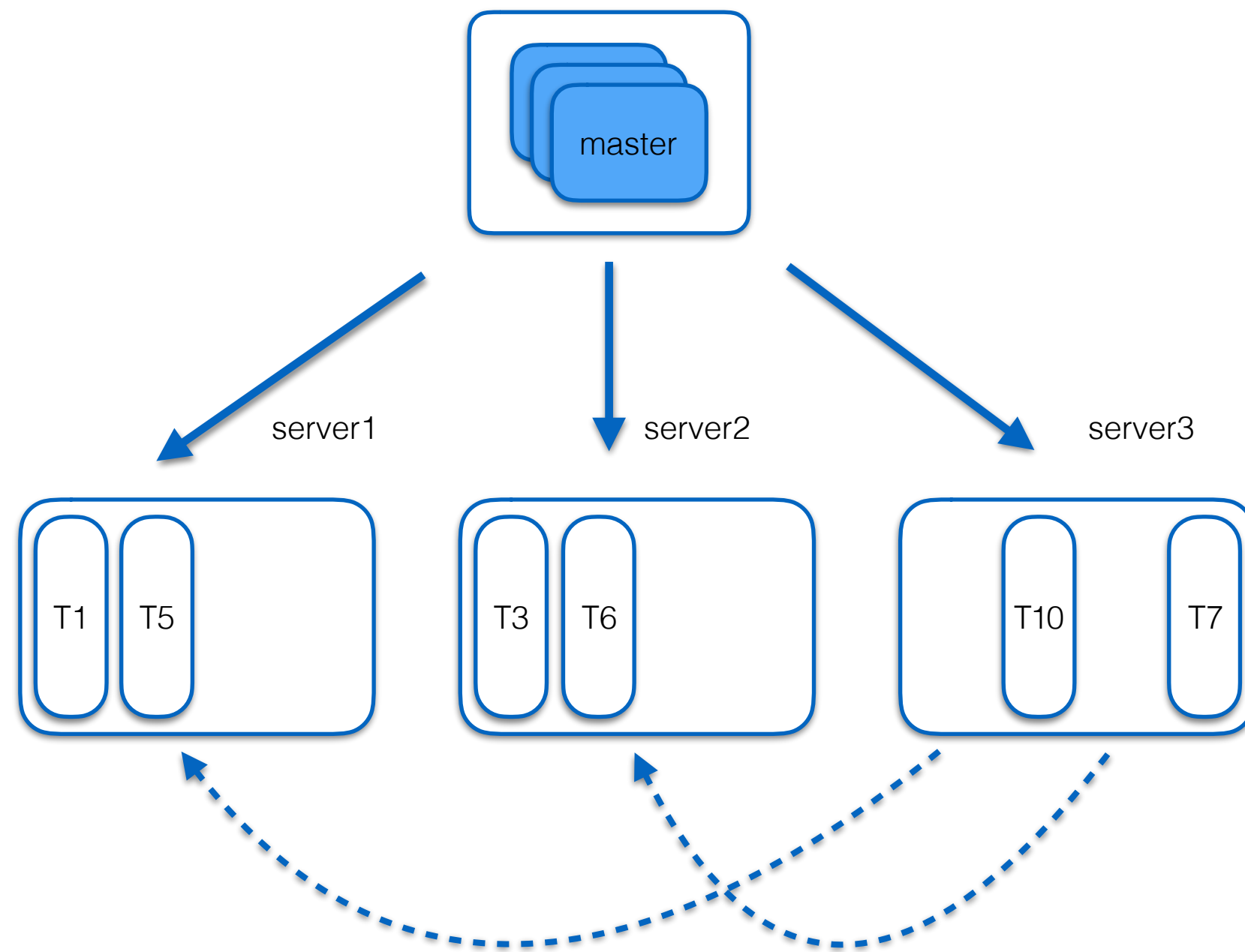
# 自动调整

- 任务均匀分配在3台server上
- T2、T8、T4、T9被删除
- 资源出现不均衡的情况，触发任务自动均衡



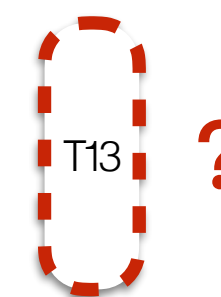
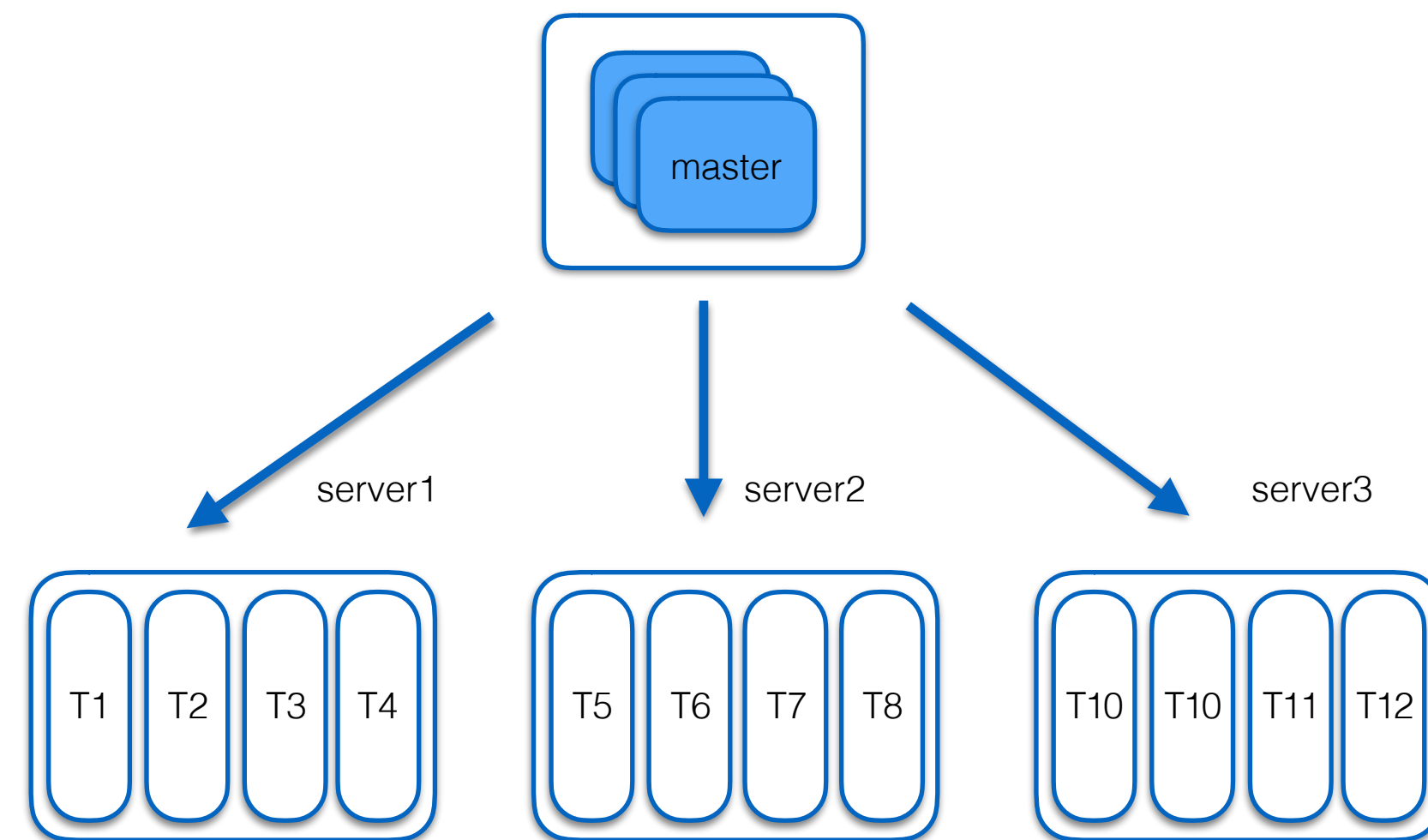
# 自动调整

- 任务均匀分配在3台server上
- T2、T8、T4、T9被删除
- 资源出现不均衡的情况，触发任务自动均衡
- 调度任务至空闲机器

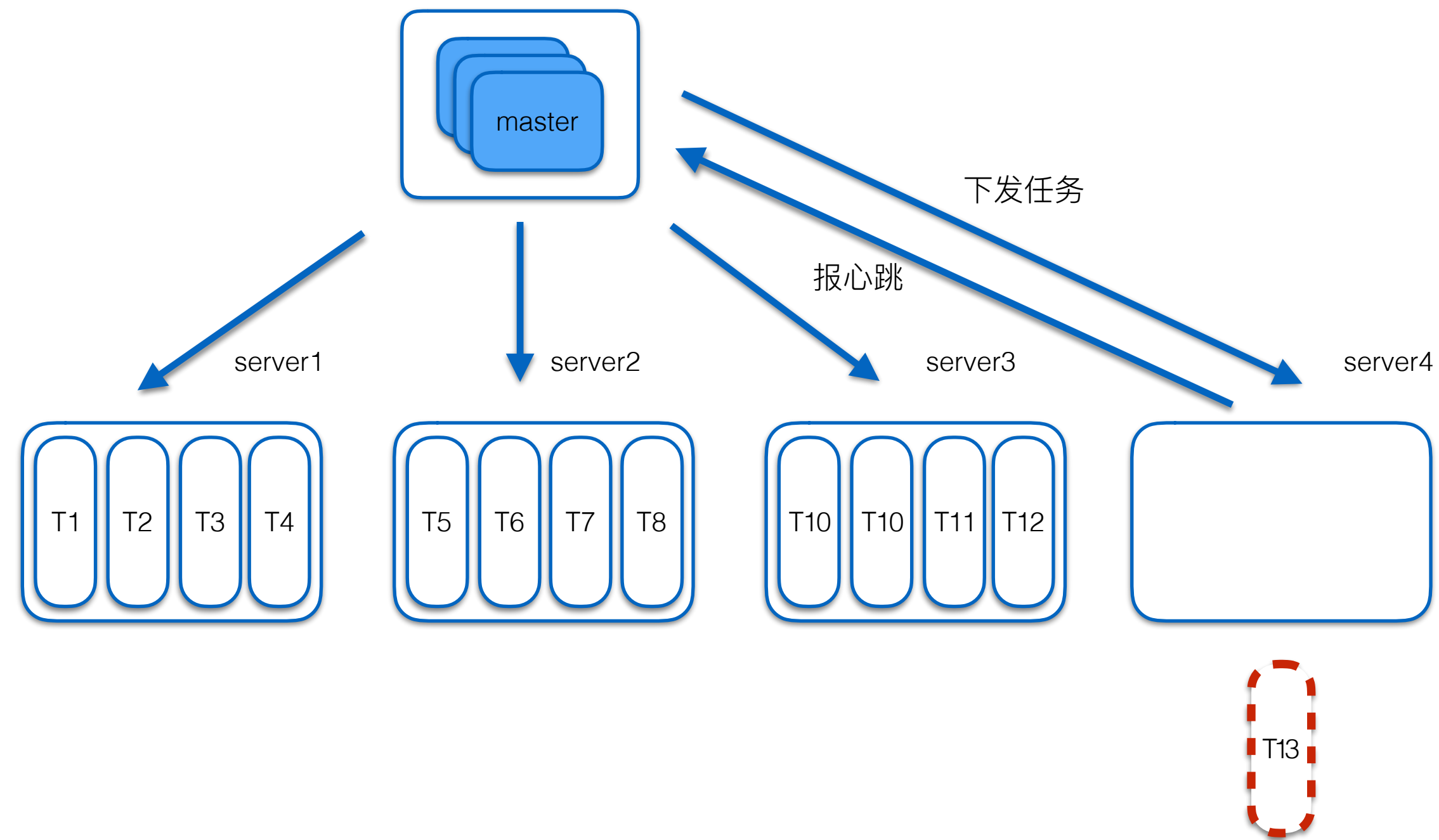


# 水平扩展

- 3台server已经全部处于满负载情况
- 新加入的任务T13无法被有效处理

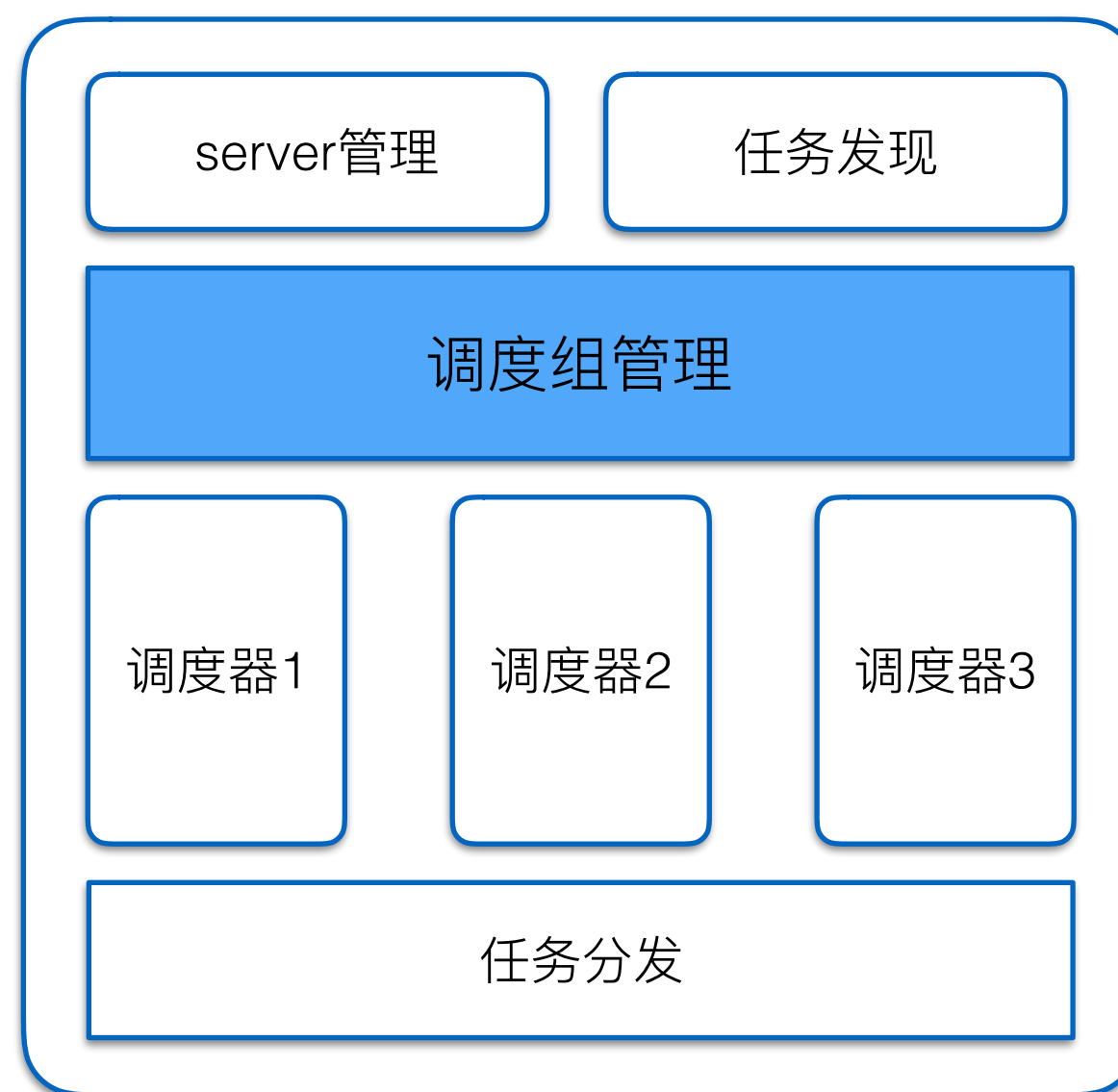


# 水平扩展



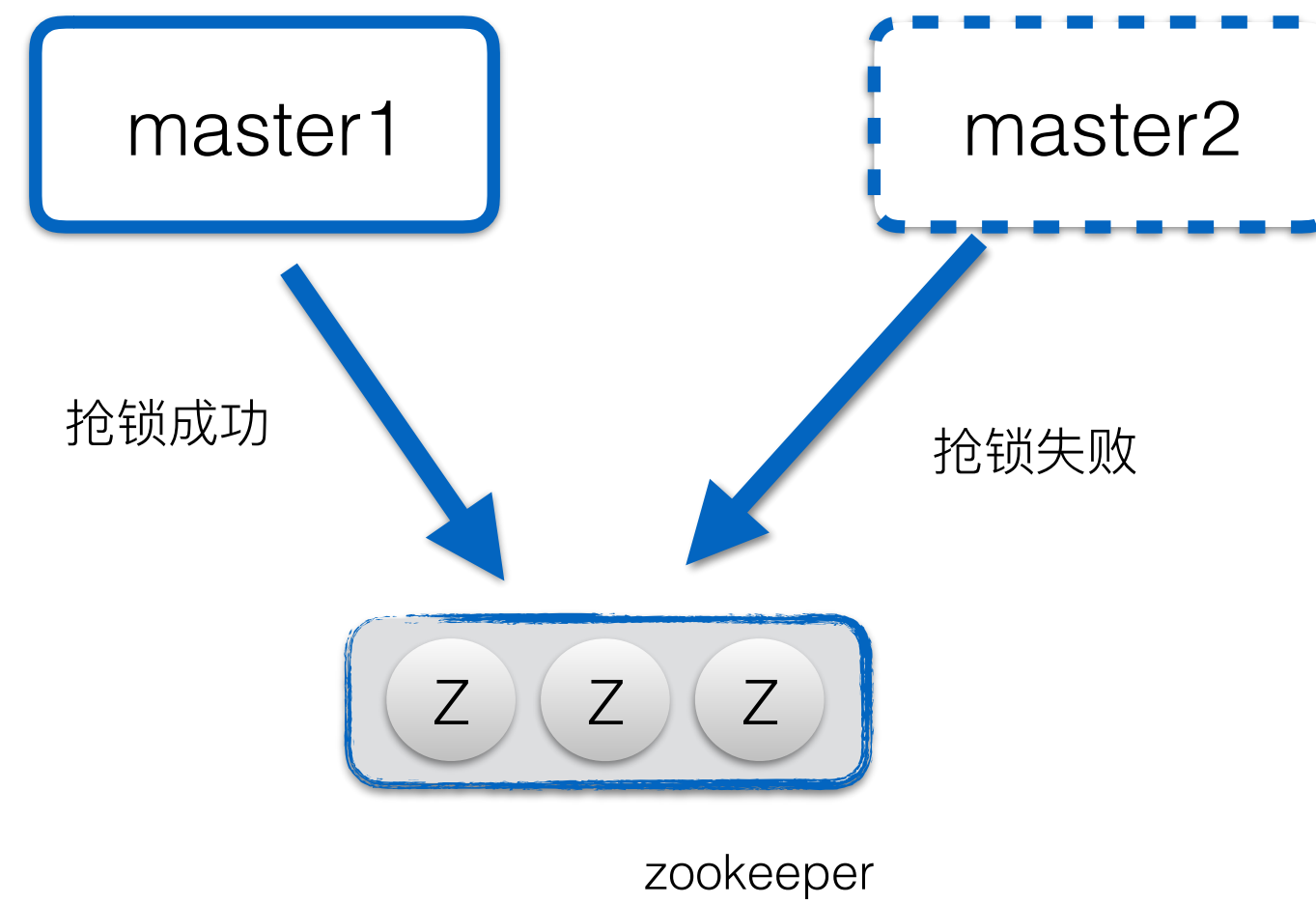
## 资源隔离

- 隔离特殊类型任务
- 利用特殊硬件资源



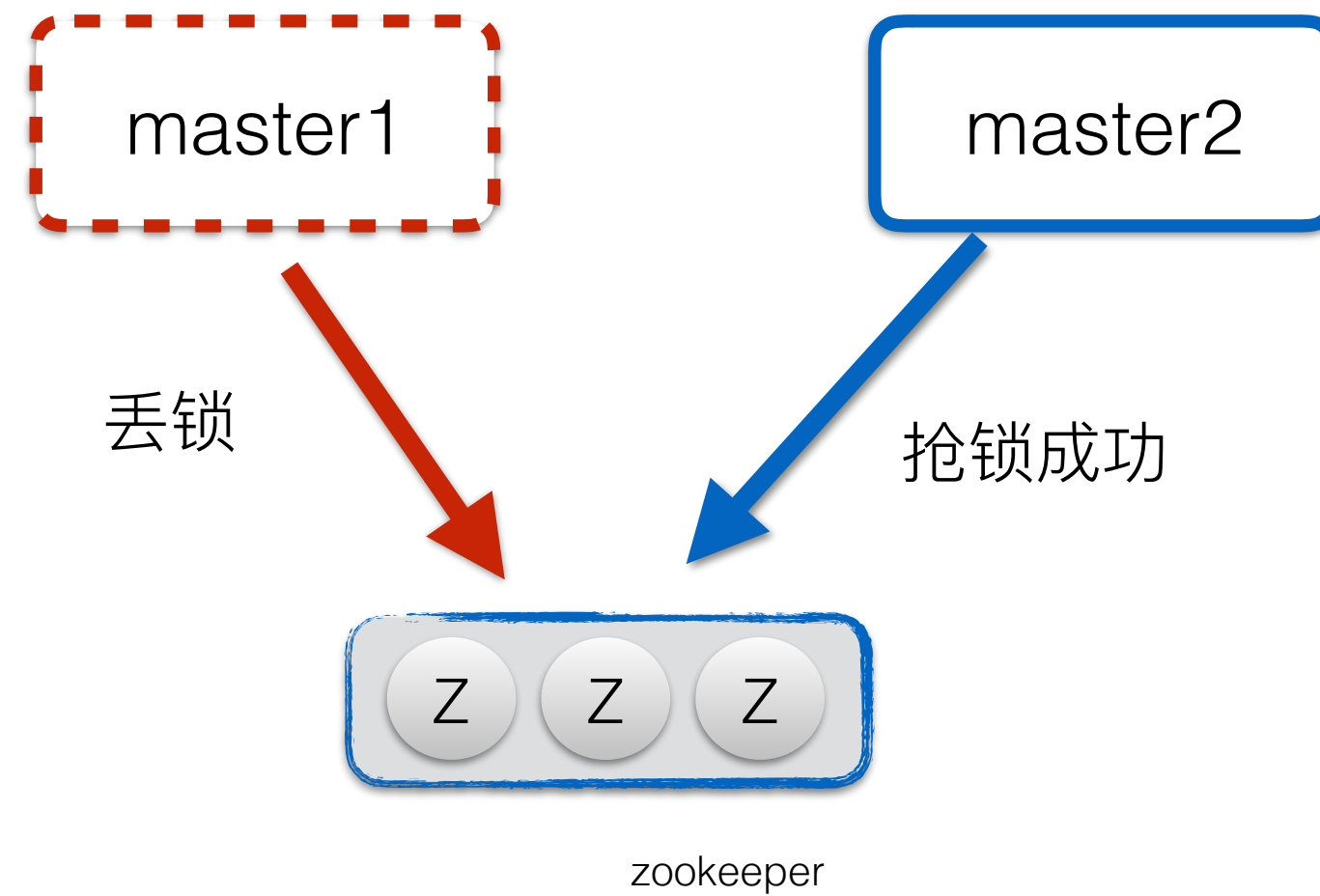
## master高可用

- master通过抢锁来决定主和备
- 主master注册自己的身份到zk

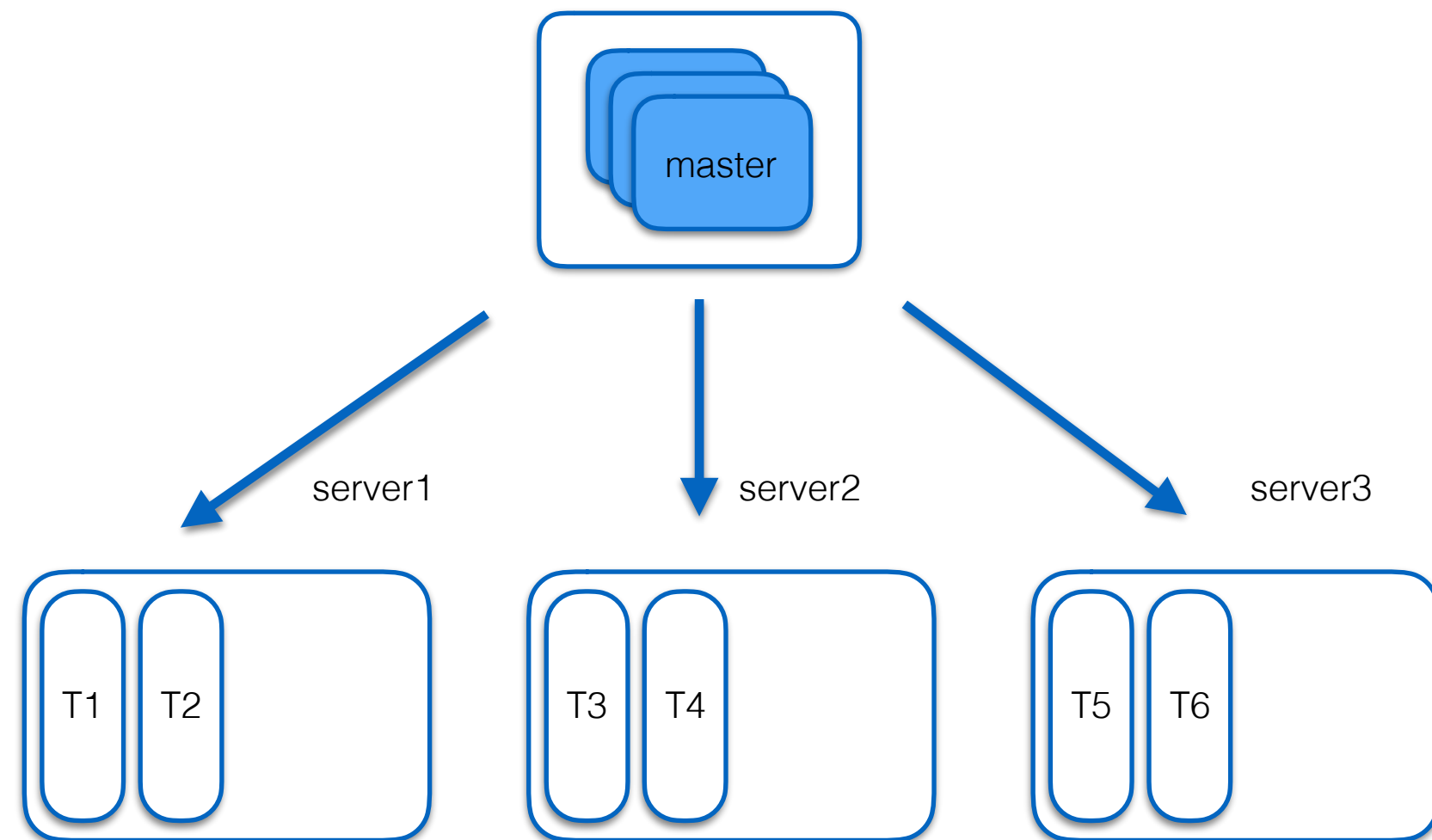


## master高可用

- master通过抢锁来决定主和备
- 主master注册自己的身份到zk

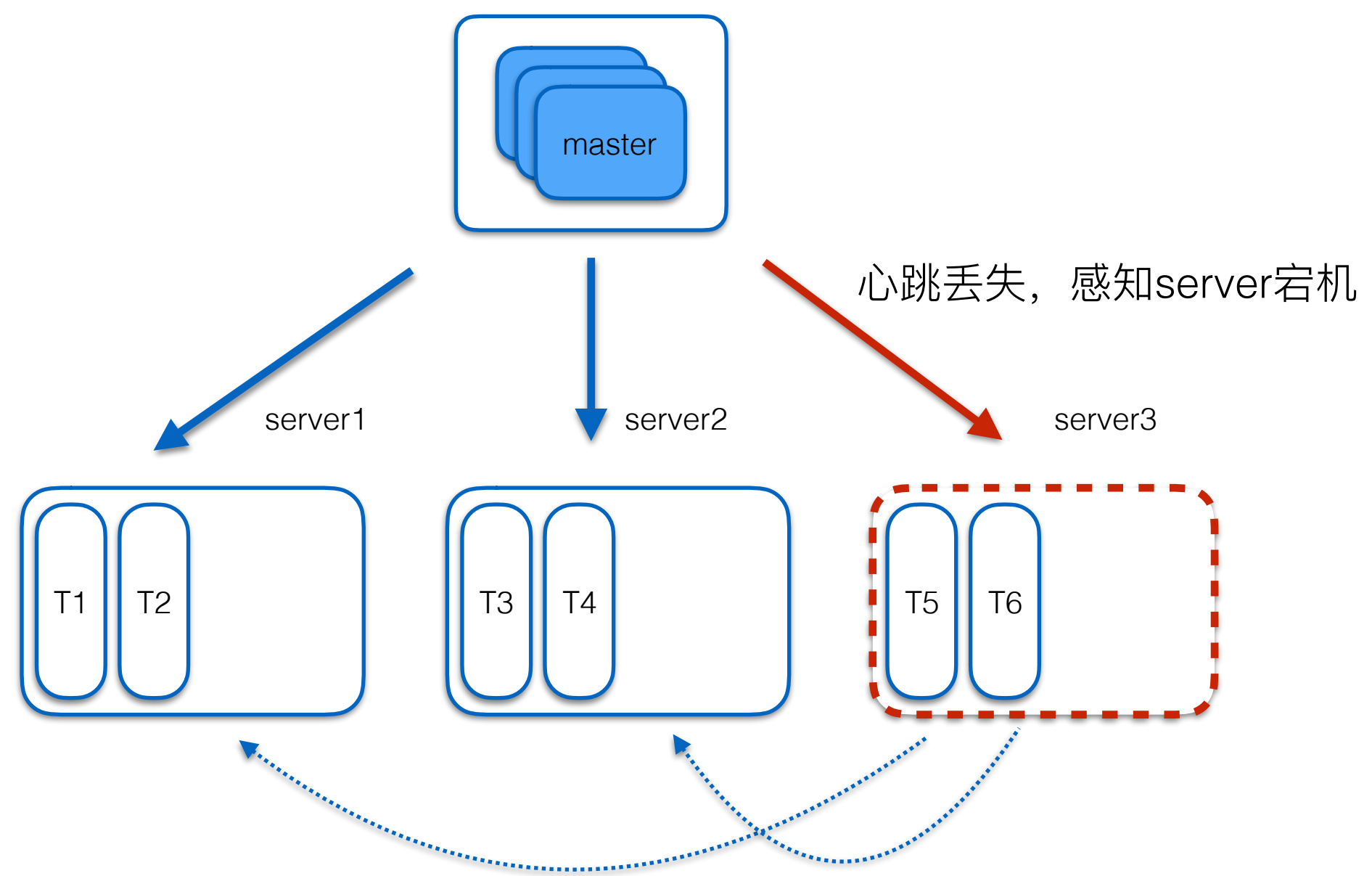


# server高可用



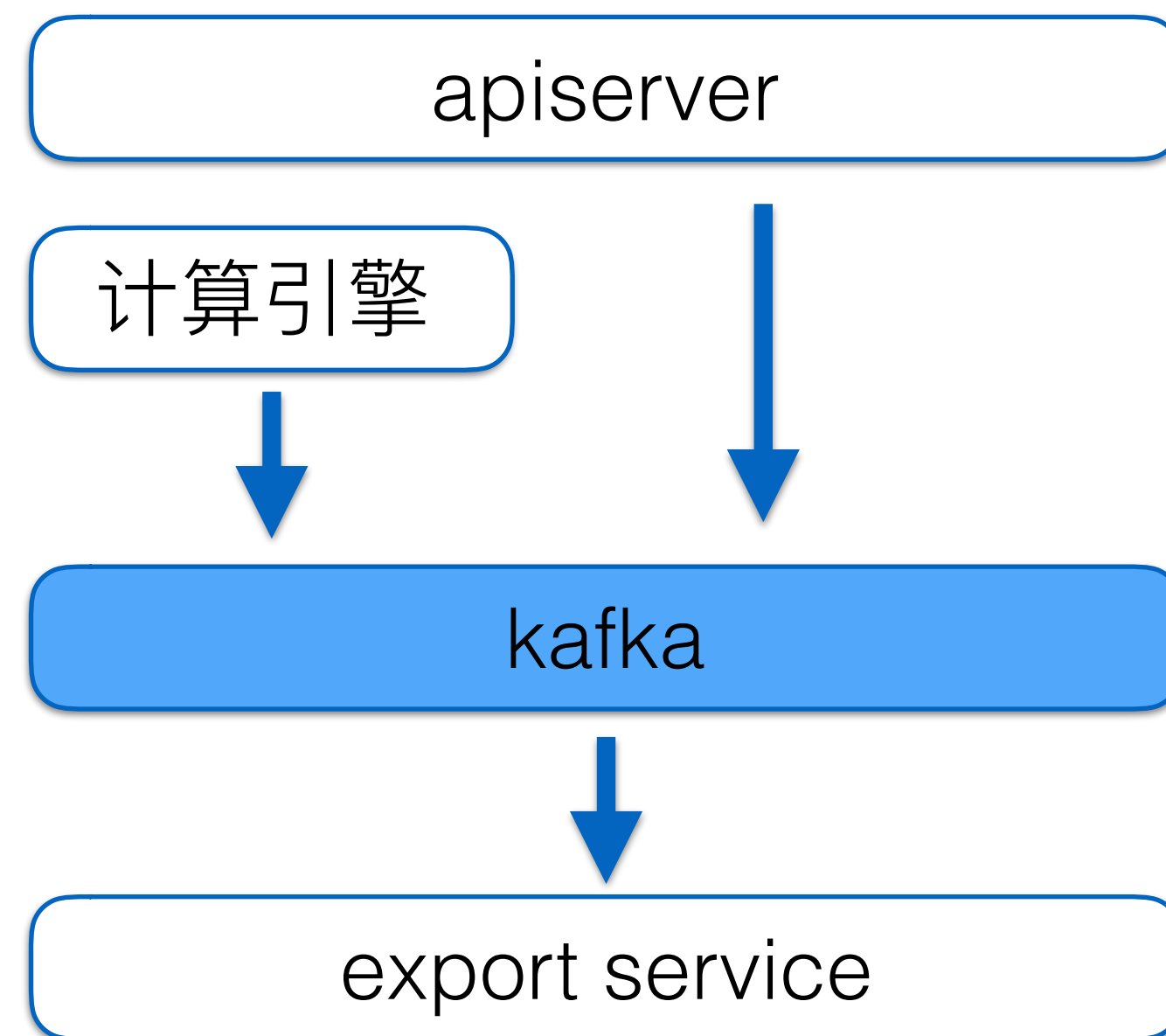


# server高可用



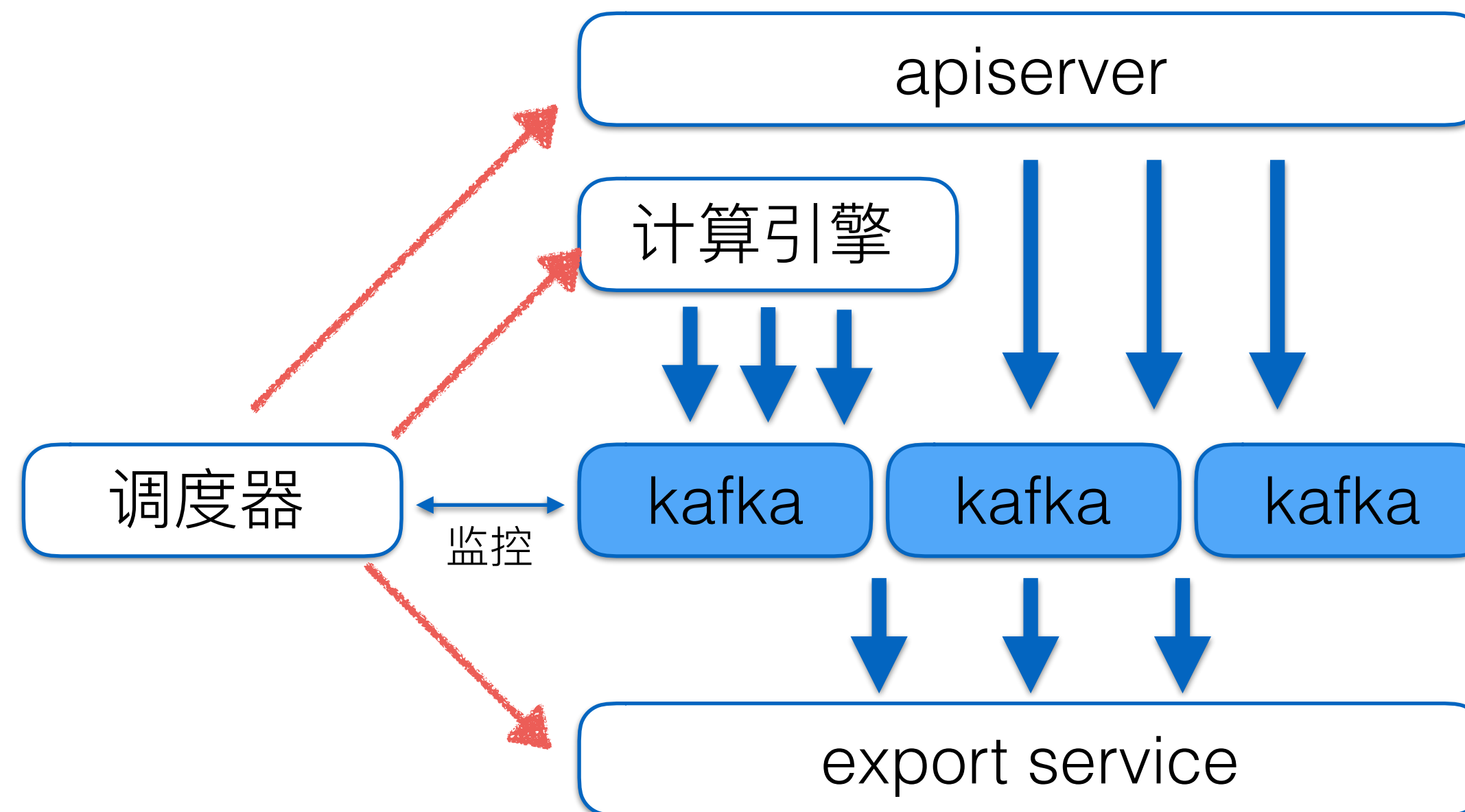
## 系统级水平扩展

- kafka单集群天花板
- 集群级scale
- 灵活路由



## 系统级水平扩展

- kafka单集群天花板
- 集群级scale
- 灵活路由



# 上下游协议优化

## • Json vs Protobuf

```

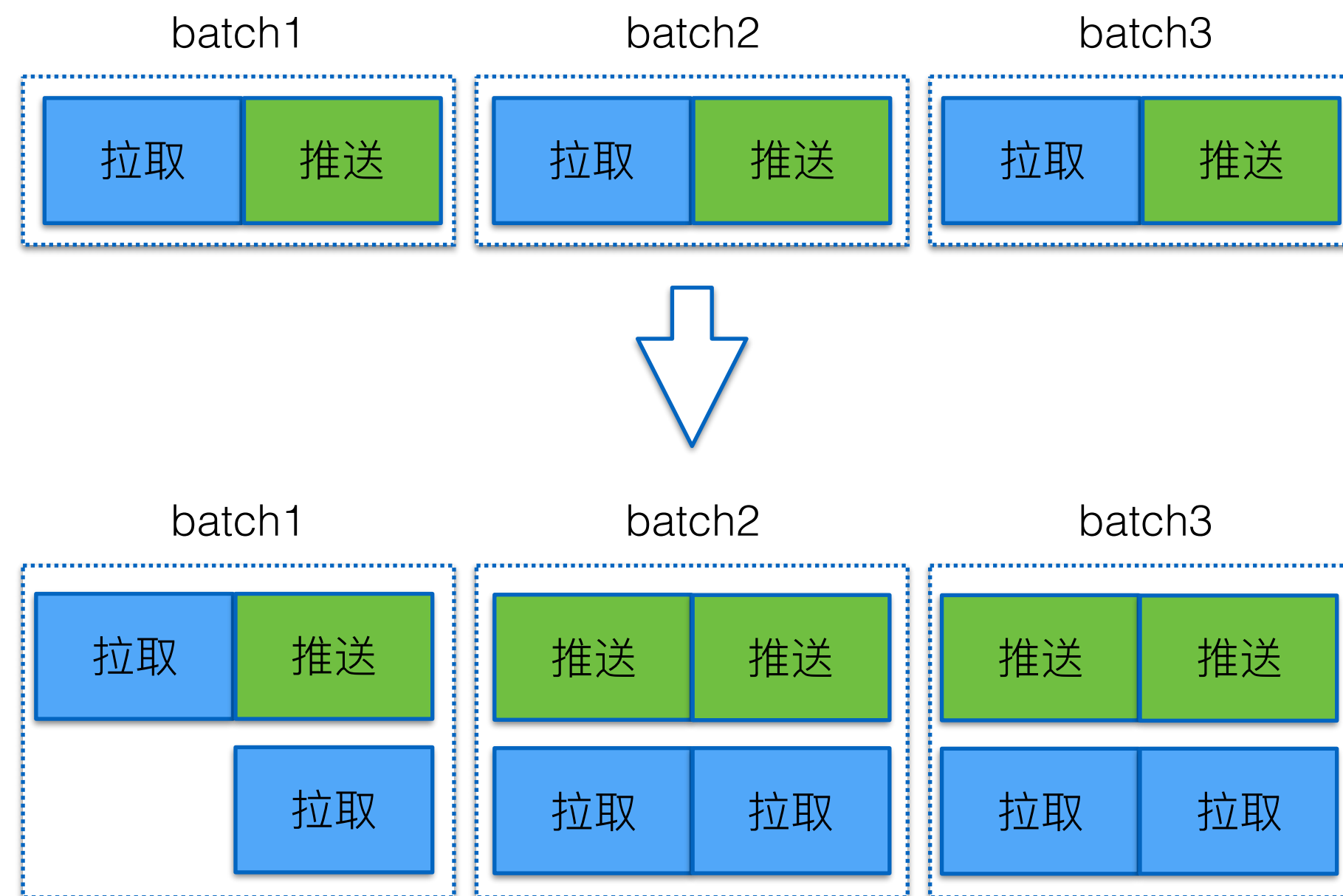
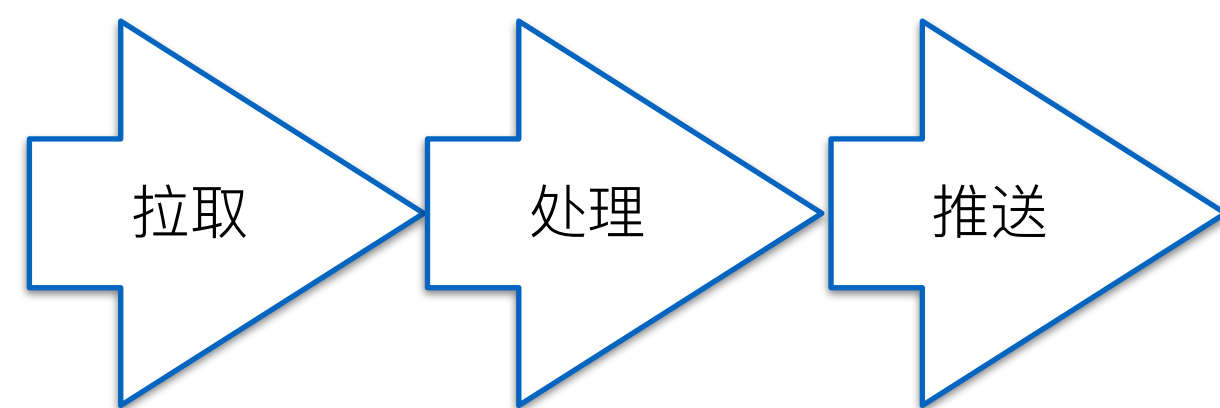
type Test struct {
    Uid          string `json:"uid"`
    BatchSize    int64  `json:"batchSize"`
    Hostname     string `json:"hostname"`
    Method       string `json:"method"`
    Operation    string `json:"operation"`
    Instance     string `json:"instance"`
    ReqBodyLength int64  `json:"reqBodyLength"`
    ReqId        string `json:"reqId"`
    ResBodyLength int64  `json:"resBodyLength"`
    ResCode      int64  `json:"resCode"`
    ResTime      int64  `json:"resTime"`
    Timestamp    int64  `json:"timestamp"`
}
    
```

项目	Json	Protobuf
序列化 (ns/op)	82161	67833
反序列化 (ns/op)	36380	7705
序列化长度 (byte)	259	100



# 流水线处理

- 导出处理模型
- 流水线并行处理



## Golang GC

- stop the world
- sync.Pool
- 重用对象
- Golang版本升级



## 有限资源假设

- 单位资源服务能力
- 资源使用评估
- 资源规划

- Pandora简介
- Workflow设计与优化
- 总结



## 成果

- 每天支撑万亿级数据点、PB级数据量
- 支持海量用户
- 极低的系统延迟
- 自动化运维
- 可用性达到99.9%



## 成果

- 七牛云所有业务的日志全部迁移至Pandora
- 大量外部客户接入

