



2016 杭州·云栖大会
THE COMPUTING CONFERENCE

云栖社区
yq.aliyun.com

余额宝超大规模基金文件处理

——分布式调度组件：余额宝大白组件

叶恺
蚂蚁金服财富线技术专家

主办单位： 杭州

 Alibaba Group
阿里巴巴集团

战略合作伙伴：



扫码观看大会视频

目录

content

- 一、组件背景
- 二、组件功能
- 三、技术实现
- 四、未来展望





一、组件背景



基于单元化部署上的分布式环境无法充分利用机器资源。

- 余额宝现有几十种任务类型，每日生成上百个文件，多种批量任务，几十种核对任务。
- 现有分布式调度中心不区分资源类别，无法合理利用资源

大量类似的文件导入、导出任务，但文件格式多样，基础流程复用性低。

- 多种文件格式
- 任务类别可简单分为导入、导出、导入且导出类别

文件存储异构，且有异构存储容灾需求。

- 多种存储类型，且需支持异构存储容灾切换

任务之间业务依赖关系，无统一通用解决方案。

- 文件生成顺序存在依赖。



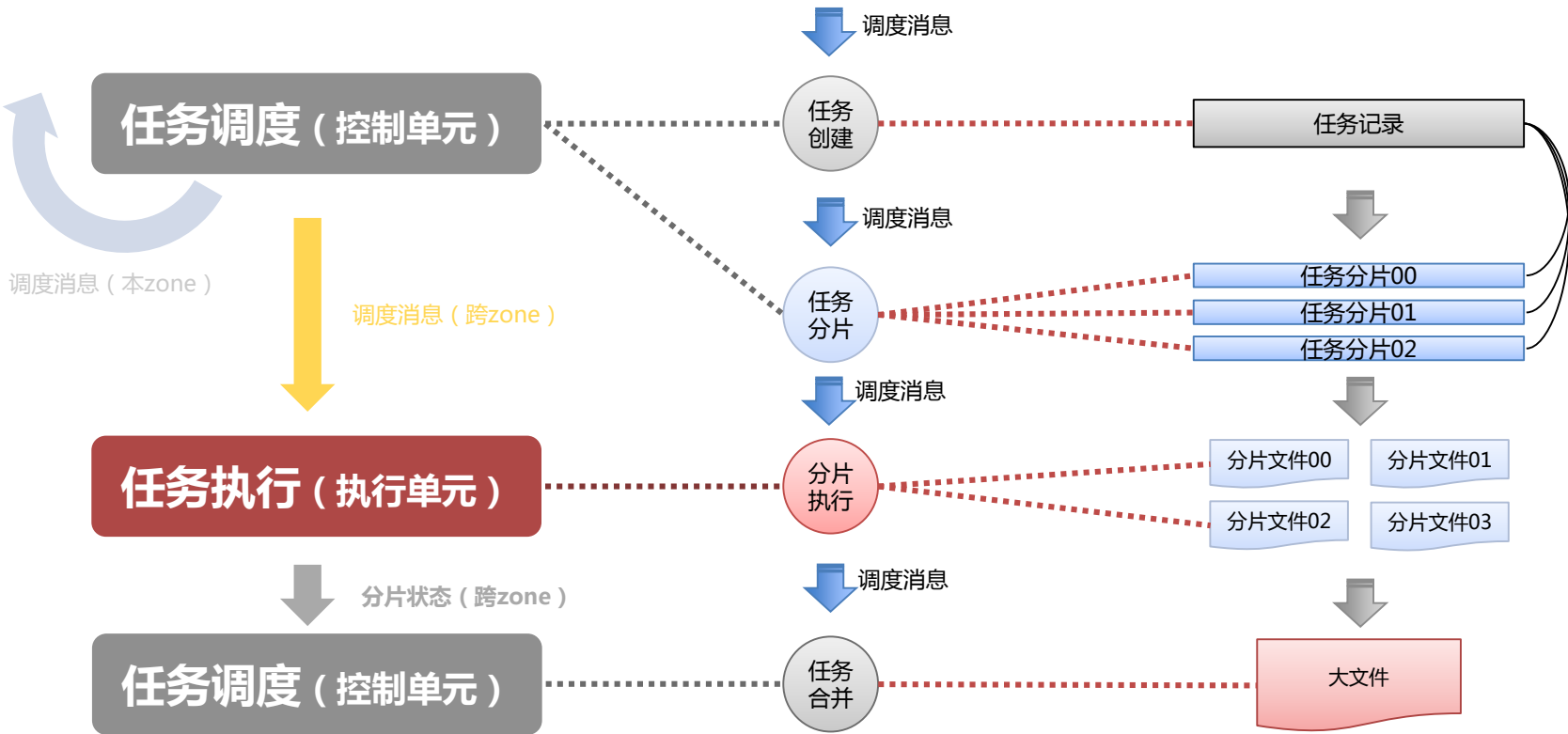


二、组件功能



组件功能：大文件导出

思路：并行生成小文件后，合并成大文件；任务调度，任务执行独立



组件功能：自适应调度

思路：高时效任务单独线程池，忙时集群内单机自发调度，自发调度开关自适应

自扫描开关：

自触发根据系统繁忙程度自行开关

待处理任务

任务记录

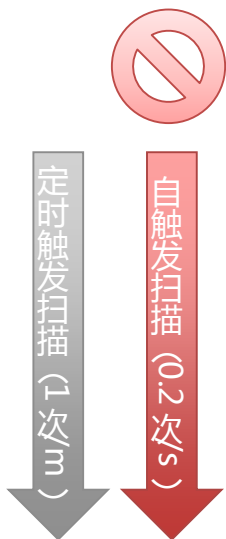
任务记录

任务记录

任务记录

任务记录

任务记录



任务扫描：

- ✓ 定时触发（慢）
- ✓ 自触发（快）
- ✓ 任务扫描数自适应

一般任务执行线程

一般任务执行线程

一般任务执行线程

公有线程池

- ✓ 一般任务执行
- ✓ 线程数可调
- ✓ 队列数可调

紧急任务执行线程

紧急任务执行线程

紧急任务执行线程

专用线程池

- ✓ 紧急任务执行
- ✓ 线程、队列数可调
- ✓ 动态创建线程池



组件功能：任务配置

思路：建立任务模板，配置业务参数，建立任务依赖



使用场景：文件操作及存储

思路：构造文件模板语法、封装文件操作接口，屏蔽底层存储差异

FileConfig

FileTemplate

```
{  
  "head":["caption|lable|type"],  
  "body":["caption|label|type"],  
  "tail":["caption|label|type"],  
  "summaryColumnPairs":["caption1|caption2"]  
}
```

FileProtocol

DE (天弘)
FUND (标准基金)

FileStorage

NAS (网络存储)
OSS (分布式对象存储)

OSS

NAS

文件
工具

FUND

FileWriter	FUND:OSS
FileReader	FUND:OSS

FileWriter	FUND:NAS
FileReader	FUND:NAS

FileMerger	FUND:NAS/OSS
FileSplitor	FUND:NAS/OSS

DE

FileWriter	DE:OSS
FileReader	DE:OSS

FileWriter	DE:NAS
FileReader	DE:NAS

FileMerger	DE:NAS/OSS
FileSplitor	DE:NAS/OSS



扫码观看大会视频



三、技术实现原理



实时任务管理

掉单任务

流水捞取

流水执行

文件任务管理

导入文件任务

简单导入

分块导入

分库导入

文件夹导入

导出文件任务

分库导出

混合文件任务

文件夹导入多文件导出

文件夹导入单文件导出

单文件导入多文件导出

实时调度

周期管理

RG调度

任务管理

容量管理

定时调度

调度对象

任务

任务分片

任务锁

调度控制

本地调度

调度消息

调度分发

异常控制

任务重试

R状态保持

锁恢复

容量控制

线程池

任务调度量

分片调度量

文件读写

文件存储

NAS

OSS

文件协议

天弘格式

Ib格式

基金格式

文件读写

文件头读

文件分割

文件合并

运维

DRM

调度运维

文件参数

应急运维

日志

业务日志

监控日志

接入适配



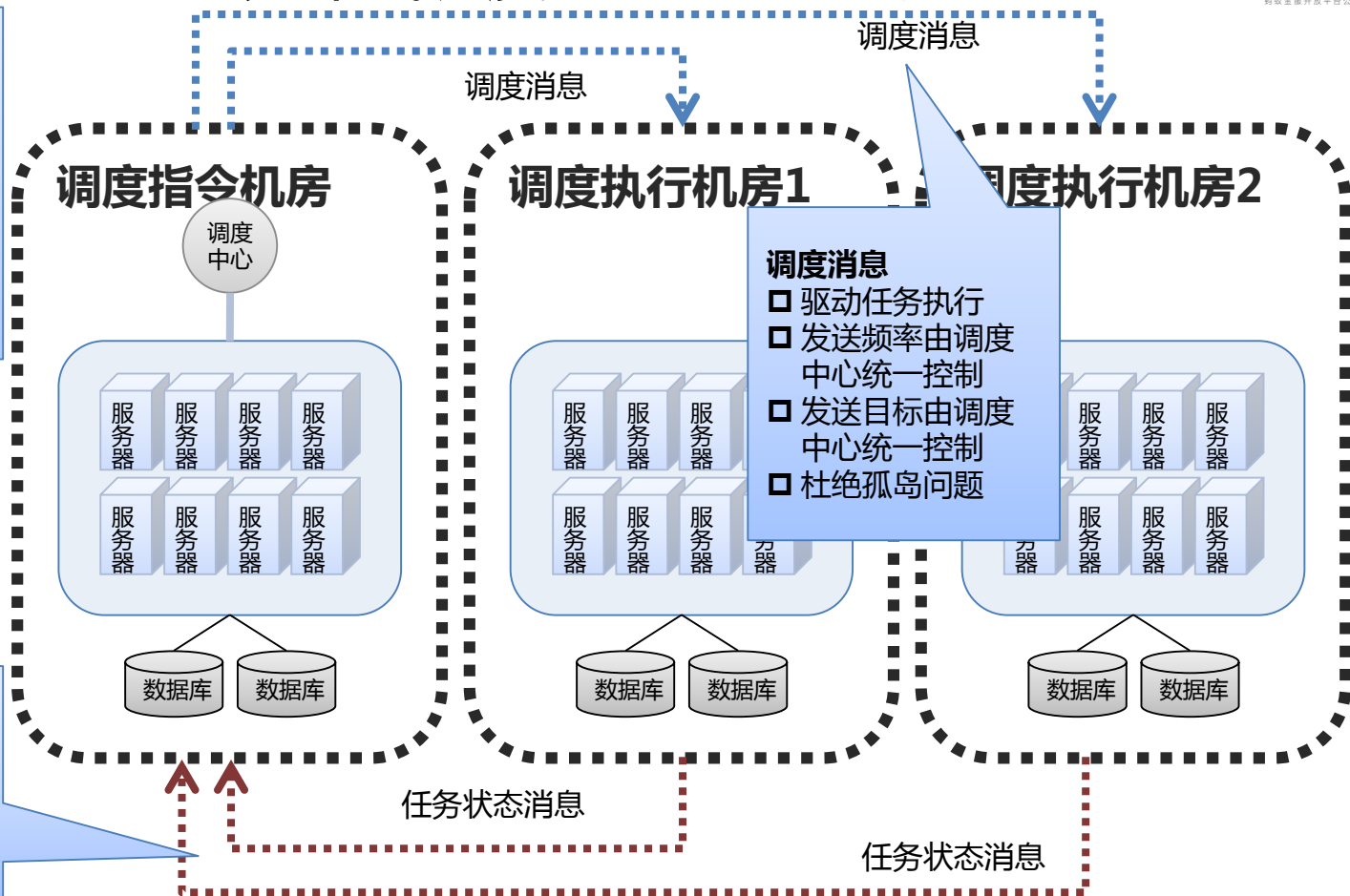
扫码观看大会视频

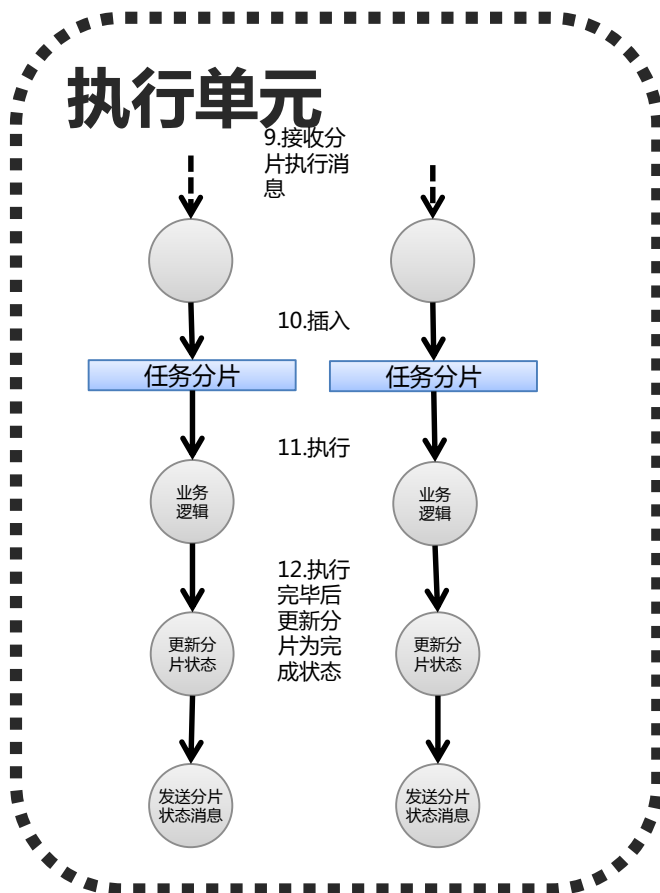
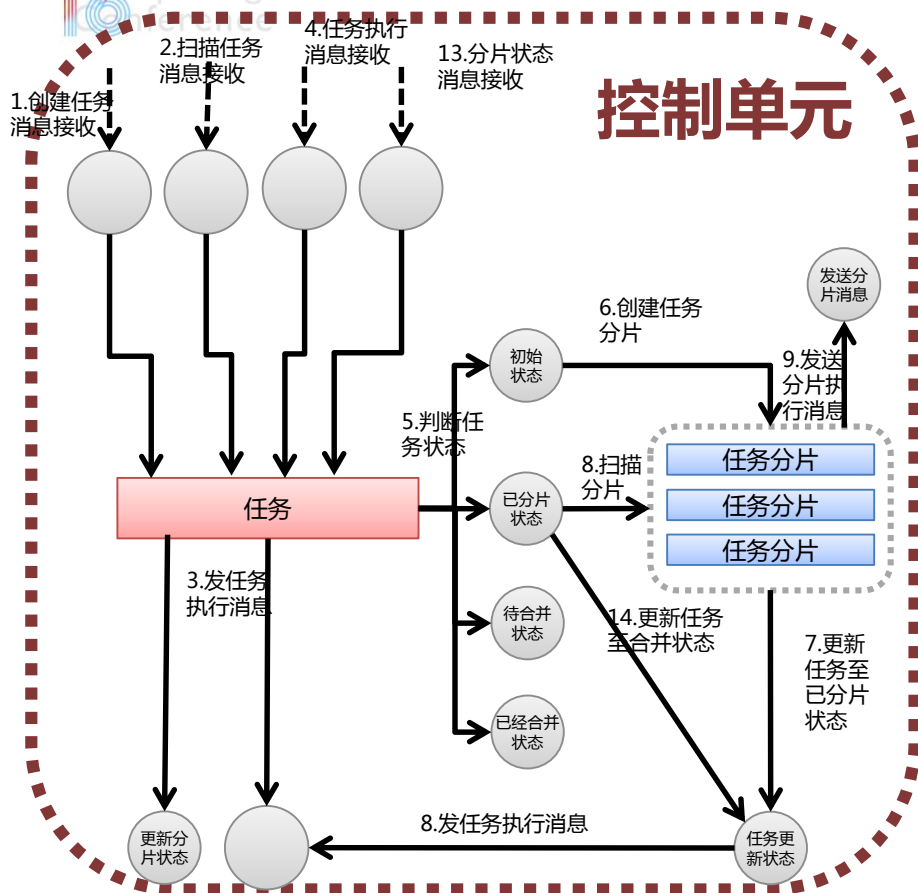
总控调度中心

- 全局控制资源利用
- 调度频率
- 容灾指令
- 任务拆分
- 全局管控
- 自适应调度

任务状态消息

- 任务执行情况反馈
- 调度中心根据反馈调整调度策略

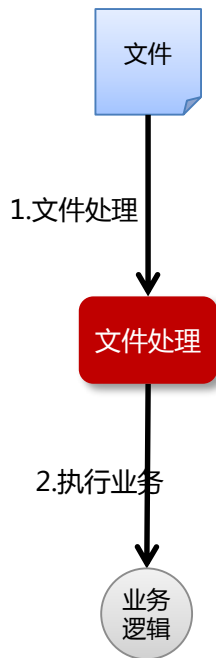




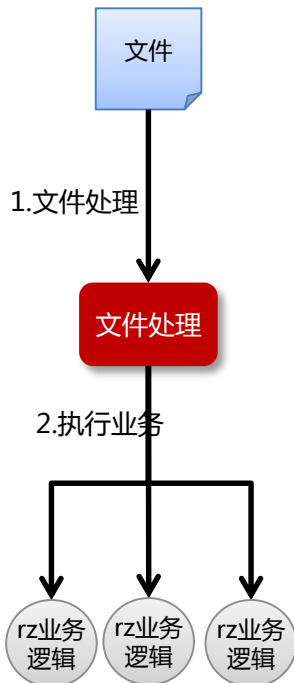
文件任务管理



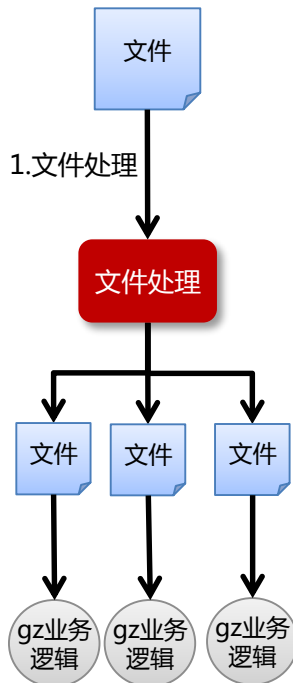
简单导入任务



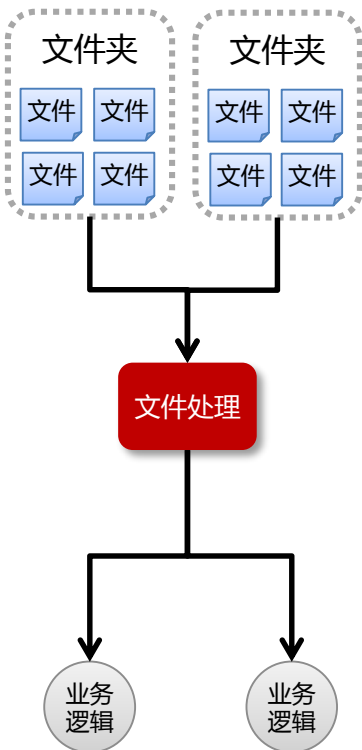
分库导入任务



分块导入任务



文件夹导入任务



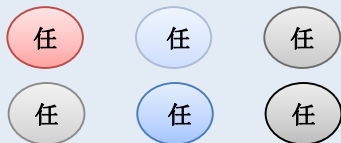


四、未来展望



两级调度：任务调度+资源调度，更加合理使用集群资源

任务调度



工作流配置

定义任务的
资源需求，
调度策略可
通过扩展由
业务实现

资源调度

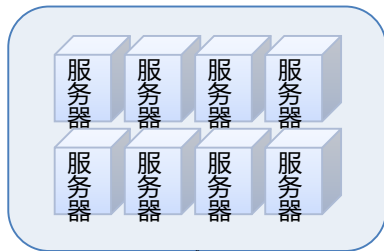


配置中心

由配置中心
获取所有服
务器资源使
用情况，分
配空闲资源

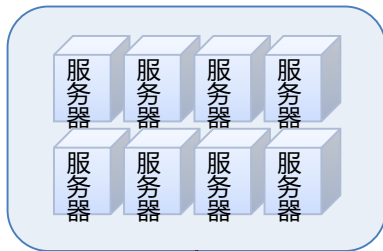
oneway

调度执行（系统1）



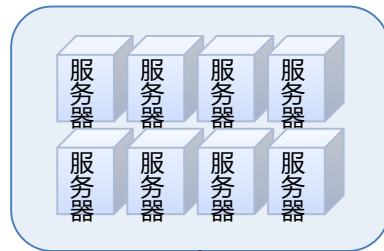
oneway

调度执行（系统2）



oneway

调度执行（系统3）



扫码观看大会视频



2016 杭州·云栖大会
THE COMPUTING CONFERENCE



2016 杭州·云栖大会
THE COMPUTING CONFERENCE



云栖社区
yq.aliyun.com



欢迎扫码关注今蚂讲支付宝服务窗

主办单位:



战略合作伙伴:



扫码观看大会视频

2016 The
Computing
Conference
THANKS

