

whamcloud

The logo for Whamcloud features the word "whamcloud" in a bold, lowercase, sans-serif font. A thick blue horizontal line underlines the text. To the right of the text, a blue graphic element consists of two curved segments: the top one is a simple arc, and the bottom one is a larger, more complex shape that loops around the end of the underline.

Lustre 管理工具 Chroma

- 杨升

Whamcloud, Inc
ys@whamcloud.com

目录

- Chroma 简介
- 系统结构
- 用户界面
- 插件技术简介

Chroma 简介

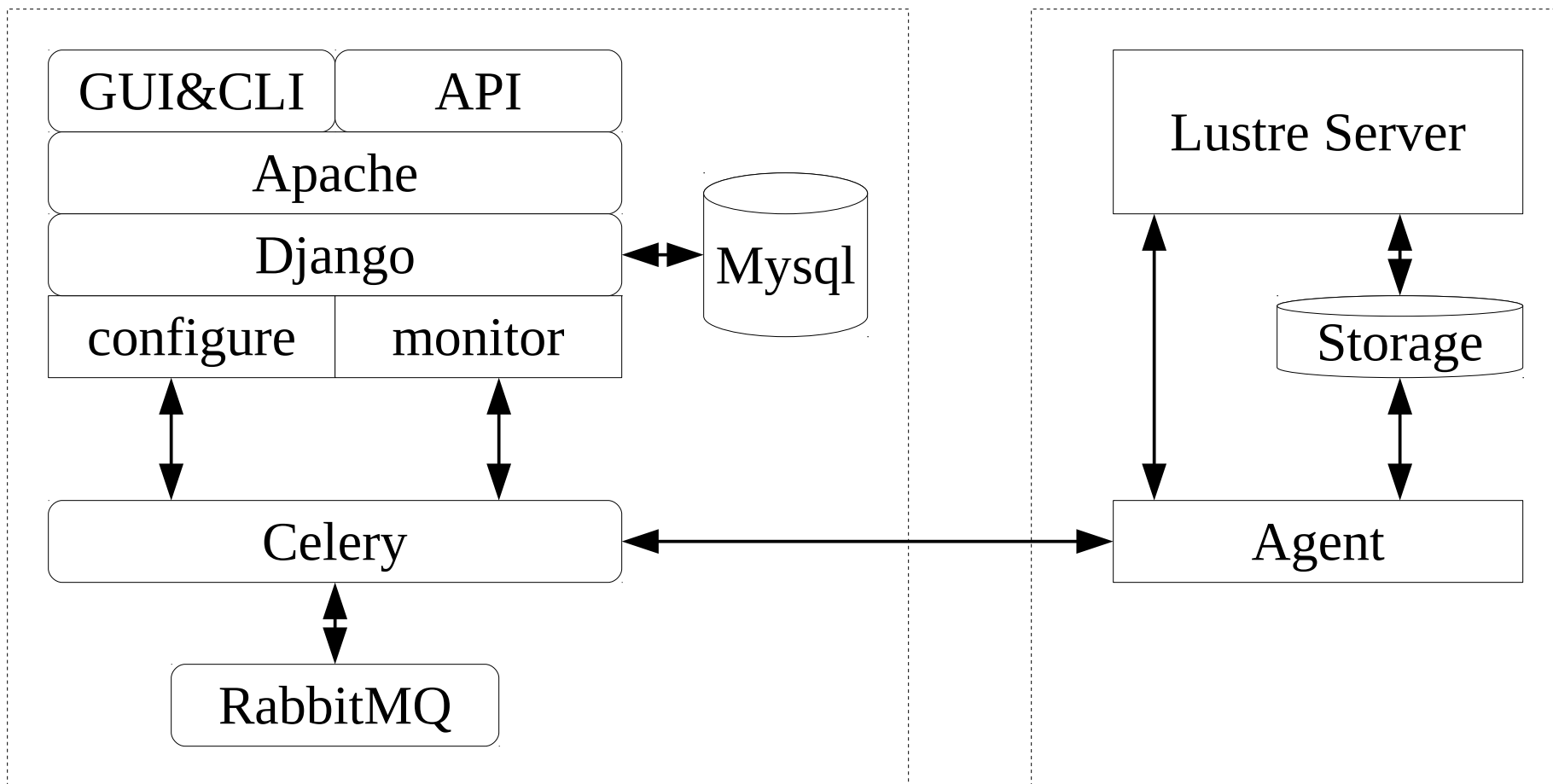
- Lustre 社区需要一个管理工具
- 主要用于监控和配置
- 数据管理和复制
- 同时管理多个 Lustre 系统
- 直接访问存储硬件
- 可集成第三方软件

开发工具

- Python
- Django
- Celery
- RabbitMQ
- Paramiko
 - SSH 协议客户端

倾向于使用通用技术，尽可能保持核心代码的精悍。

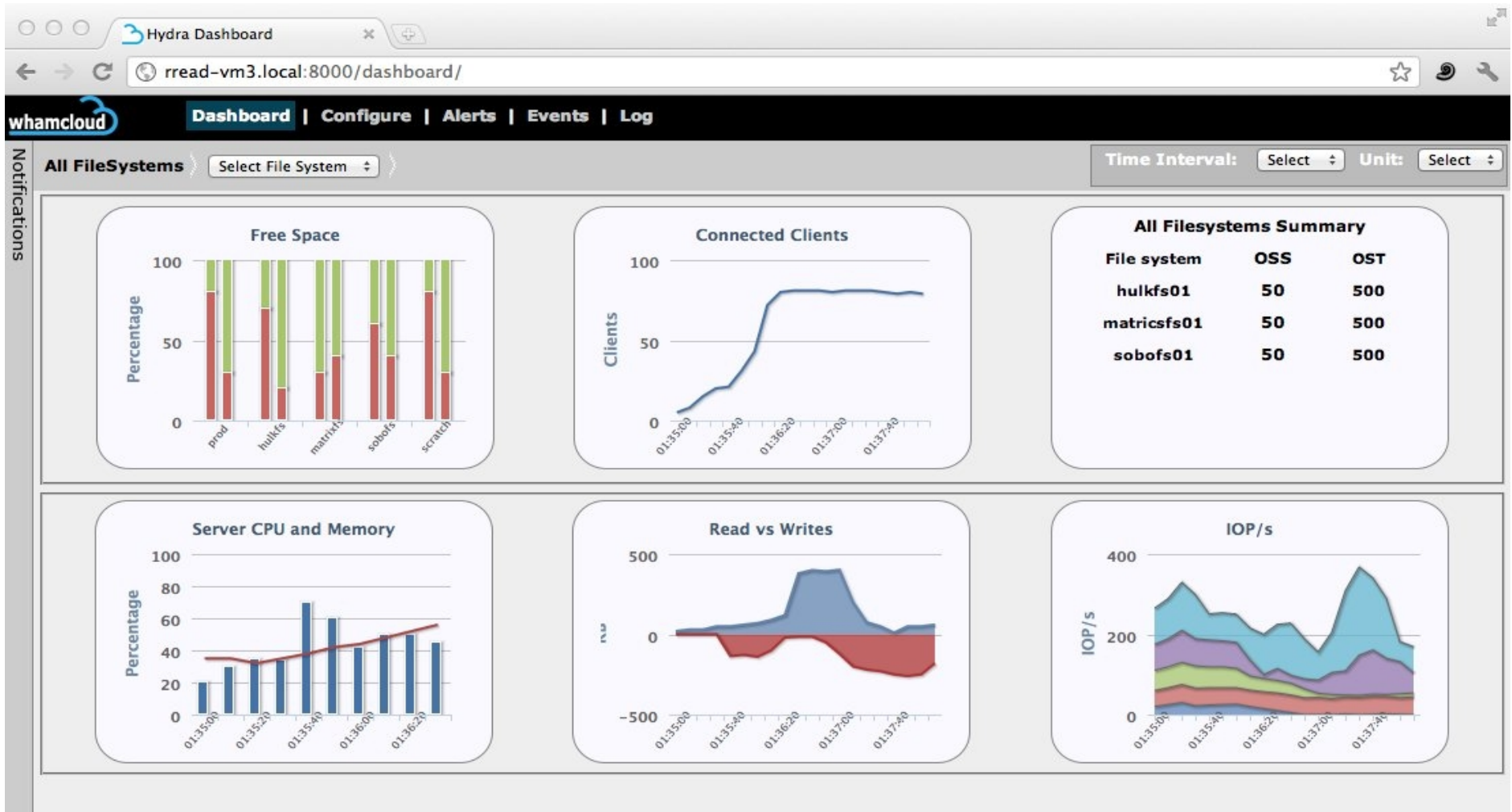
系统结构



用户界面

- Django 框架下基于 HTML/jQuery 的 GUI 界面
 - 调用 Chroma API
 - 中心化管理，一个界面实现监控和配置
 - 提供多种类型的图表（Charts，Graphs，Heat-Maps）
 - 实时反馈的监控信息
 - 可灵活自定义的界面组件

Dashboard



Configure

Hydra Configuration | 115.113.226.203/hydracm/

whamcloud Dashboard | **Configure** | Alerts | Events | Log

Notifications

Lustre FS Configuration | MGT Configuration | Volume Configuration | Server Configuration

FS name :

File Used: 0	Total Capacity: 0
Inodes: 0	Total no OSS: 1
	Total no OST: 4

All File System Space Usage

Free Used

Files Vs Free Nodes

Free Used

Select MGT Configuration

Target Path	Target Name	Server Name	Failover Partner	Action
clo-pune-lin01:/dev/sdb6	MGS	clo-pune-lin01		<input type="button" value="⊖"/>

Showing 1 to 1 of 1 entries

Select MDT Configuration

Target Path	Target Name	Server Name	Failover Partner	Action
clo-pune-lin04:/dev/sdb6	neofs01-MDT0000	clo-pune-lin04		<input type="button" value="⊖"/>

Showing 1 to 1 of 1 entries

Select OST Configuration

Show entries Search:

Target Path	Target Name	Server Name	Failover Partner	Action
clo-pune-lin04:/dev/sdb1	neofs01-OST0003	clo-pune-lin04		<input type="button" value="⊖"/>
clo-pune-lin04:/dev/sdb2	neofs01-OST0002	clo-pune-lin04		<input type="button" value="⊖"/>
clo-pune-lin04:/dev/sdb3	neofs01-OST0001	clo-pune-lin04		<input type="button" value="⊖"/>
clo-pune-lin04:/dev/sdb5	neofs01-OST0000	clo-pune-lin04		<input type="button" value="⊖"/>

Showing 1 to 4 of 4 entries

Chroma API

- Server 通过 Web 服务的方式提供 API

- 使用和 CLI 接口同样的模块和功能
- 可以轻松实现远程控制工具
- Server 的资源和功能可以通过 Web API 来向用户提供
- 易于实现不同版本的兼容

- 使用 JSON 格式交换数据

- HTTP GET 方法返回数据
- HTTP POST 方法发送数据

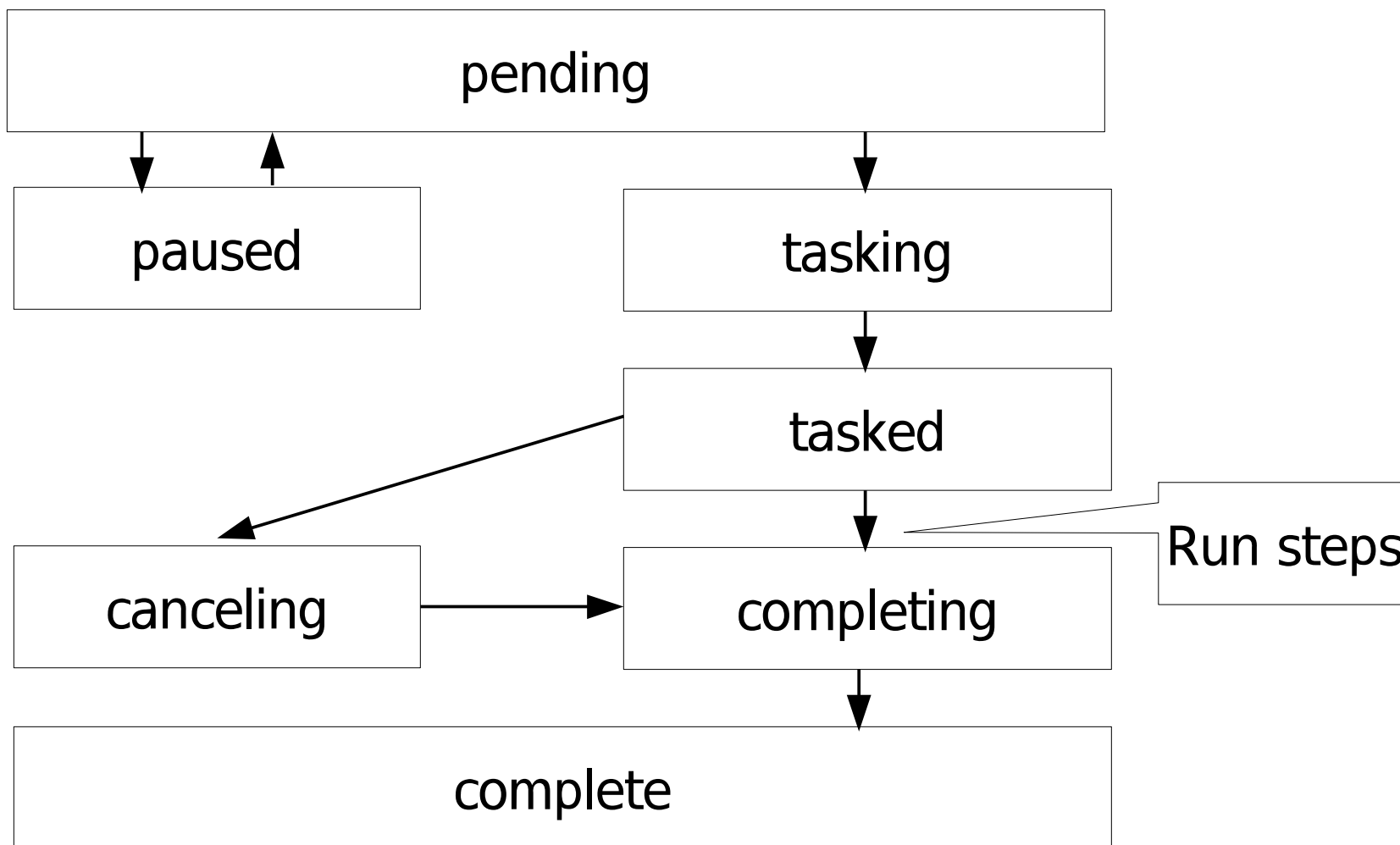
```
{  
  "errors": [],  
  "response": [{"status": "OK", "id": 1, "name": "puneofs"},  
               {"status": "OK", "id": 2, "name": "neofs"},  
  "success": true  
}
```

发生错误时，错误信息会返回并且 success 字段会指示错误状态

在 **Lustre server** 上执行命令

- 一个单独的任务被定义为 **JOB**
 - 用户看到的是单独的任务
 - 每个任务会被定义成一系列 **STEP**
 - **STEP** 必须顺序执行，后续操作依赖前置操作的状态
 - 不冲突的 **JOB** 可以并行执行
- **JOB** 管理是关键部分
 - 针对 **Crash** 状况下的处理
 - 利用 **Celery** 的机制可以在 **Crash** 以后自动恢复运行
 - 可以被暂停或者取消
 - **JOB** 之间可以定义依赖关系和是否冲突

Chroma JOB 管理



调整 **Lustre** 的参数

- 目前只管理永久性的参数
 - 保存在 MGS 上，使用 `lctl conf_param` 调整
- 参数管理的方法
 - 在 MGS 和 Chroma 数据库之间保持一致
 - 目前没有检查参数值的合法性
 - 建立 `conf_param` 和 `get_param` 之间的对应关系
 - 可以修改一个值而不是重新输入
 - 并不是所有的 `conf_param` 都有规则对应的 `get_param`
 - Agent 会扫描 MGS 的 `llog` 并且返回给 Chroma

failover 管理

- Pacemaker 用来监测所有 failover 的节点
 - 决定是否进行 failover
 - 通知 agent 启动和停止的时间
 - 配置管理由 agent 负责
- Chroma server 观测是否发生 failover
 - 具体的配置信息由 Chroma server 通过 agent 提供。
 - 目前的模式是由 Chroma server 轮询 agent，今后会由 agent 直接发送信息到 Chroma server 端
 - 没有任何加密和安全措施，默认网络是安全的

Plugins

- 提供 storage 的具体信息
 - 名称，类型，属性。
 - 实体间的关系
- 可以提供底层硬件的运行信息
 - 性能统计数据
 - 错误信息
- 特性
 - 在存储硬件和 Lustre 之间还有 OS 层
 - 需要完整的描述实体关系
 - 允许不同级别的实现
 - 基础型：检测硬件资源
 - 监控型：报告运行状态和数据
 - 完全型：能够直接配置存储设备



谢 谢