



QCon

QCon全球软件开发大会（杭州站）2012

讲师资料下载



Alibaba

technology
Association

淘宝分布式服务框架

玄宵



- 分布式服务框架基础数据

| 参数 | 值 |
|---------|--------|
| 每天调用量 | 300+亿 |
| 提供的服务数量 | 3k+ |
| 机器数量 | 8k+ |
| 机房分布 | 6,7个机房 |
| 应用 | 1000+ |
| 使用者 | 整个阿里系 |



- 淘宝分布式服务框架（**HSF**）演进过程
- 软负载体系
- 服务治理
- 分布式跟踪系统（**Eagleeye**）



- 初始版本
 - 服务发布，订阅以xml文件形式配置
 - Xml文件与应用分离
 - 通讯层基于JbossRemoting
 - 负载通过硬件设备负载



产生的问题

- 使用起来非常复杂，部署维护成本高
- Jboss Remoting量大，不稳定，而且不可控
- 硬件负载设备成本高，易出问题。



- 配置使用方式的改进
 - 使用示例

```
<bean id="helloWorld" class="com.taobao.hsf.test.HelloWorldImpl" />
```



- 发布服务

```
<bean class="com.taobao.hsf.app.spring.util.HSFSpringProviderBean">
  <property name="serviceInterface">
    <value> com.taobao.hsf.test.HelloWorld </value>
  </property>
  <property name="target">
    <ref bean="helloWorld"/>
  </property>
</bean>
```

- 消费服务

```
<bean id="hello"
class="com.taobao.hsf.app.spring.util.HSFSpringConsumerBean">
  <property name="interfaceName">
    <value>com.taobao.hsf.test.HelloWorld </value>
  </property>
</bean>
```




- 部署及隔离方式改进
 - 与应用分开部署，运行期依赖
 - 外部采用与应用独立的classloader隔离，内部采用OSGI隔离
- 优点vs缺点？



- 网络通讯改进
 - 基于mina封装TB-Remoting
 - 分阶段序列化（java，hessian）
 - 连接采用长连接



- 负载均衡改进
 - 采用基于配置中心（configserver）订阅推送
 - 客户端软负载
 - 容灾，失效恢复
 - 路由等规则支持



- **Configserver**
 - 面向动态数据推送
- **Diamond**
 - 面向静态数据推送



- 跨语言改进
 - Webservice
 - Protocol buffer
 - Hessian



- 演进过程中的一些小功能
 - 客户端线程池控制（稳定性开关）
 - 日志放置的目录
 - 日志刷屏
 - 服务本机优先调用
 - 服务调用及执行统计（**logstat**）
 - 服务端线程池隔离（防止雪崩）
 - 线程池满，自动执行**jstack**，**jmap**
 - 服务端及客户端配置交互（超时，序列化类型）
 - **Core+plugin**模式



- 演进过程中的一些小功能
 - 服务动态归组
 - 服务限流
 - 服务延迟注册
 - 服务调用上下文支持
 - **Rpc**框架与业务交互（常见如：**remotehost**）
 - 服务**NDI**方式调用
 - 运行期动态发布数据
 - 服务降级
 - **Jar**包升级



- 演进过程中走过的弯路
 - 动态热部署
 - 跨语言支持（protocol buffer）
 - 版本及其他信息存放
 - 基于应用粒度的订阅
 - 权重规则
 - 分布式事务（补偿，二阶段提交）
 - 订阅从configserver迁移diamond

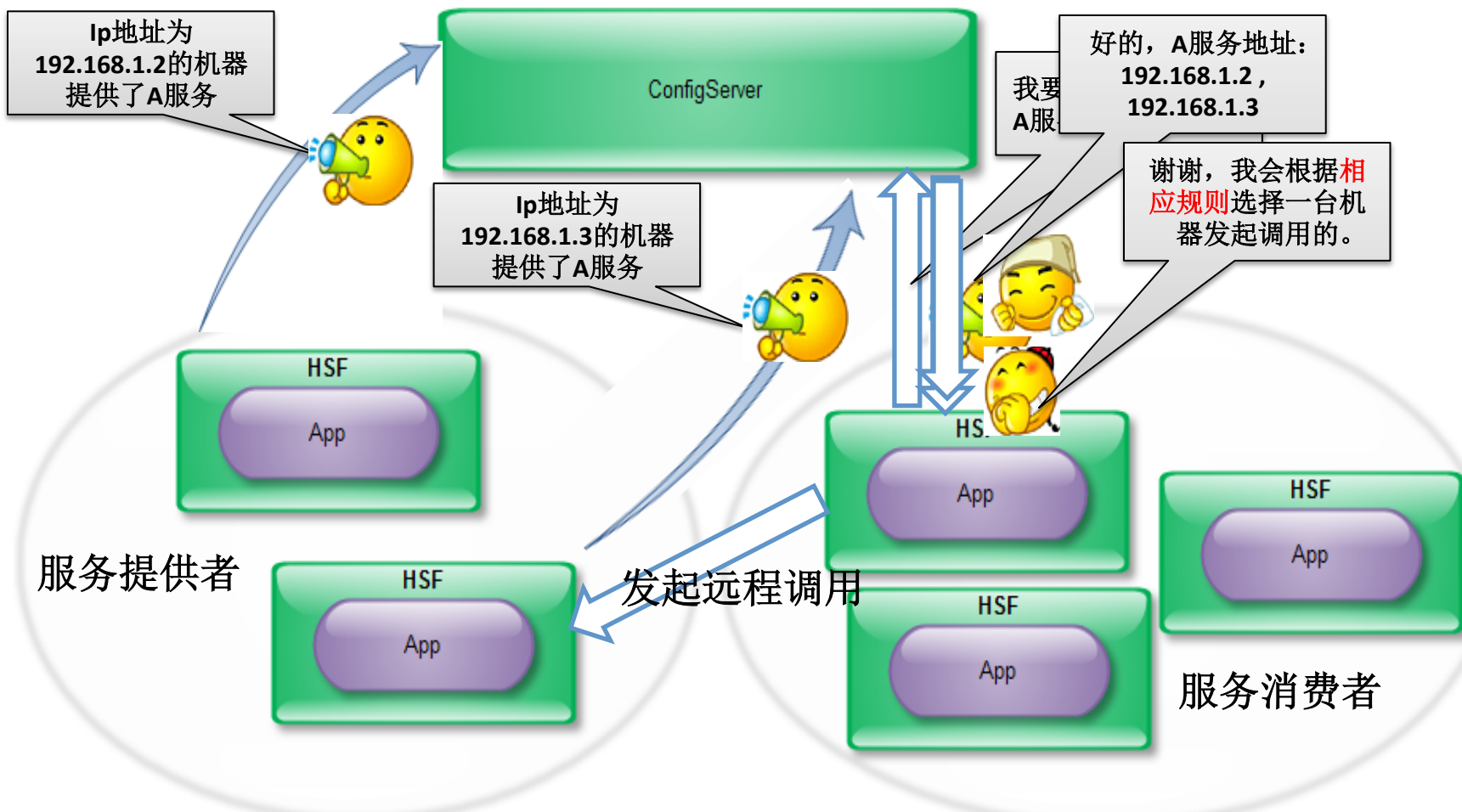


- 严重事故
 - Configserver地址归组错乱
 - Configserver地址推空
 - Configserver网卡瓶颈





HSF运行原理





- 路由规则
 - 接口路由
 - 方法路由
 - 参数路由
- 选址算法
 - 随机
 - 权重

RoutingRule

```
Groovy_v200907@package hqm.test.groovy
public class RoutingRule{

    Map<String, List<String>> routingRuleMap(){
        return [
            "read_method_address_filter_Key":["192.168.1.3:*"],
            "write_method_address_filter_Key":["192.168.1.2:*"]
        ]
    }

    String methodRoutingRule(String methodName, String[] paramTypeStrs){
        if(methodName.matches("get.*"))
            return "read_method_address_filter_Key";
        else if (methodName.matches("put.*"))
            return "write_method_address_filter_Key";
        else
            return null;
    }
}
```



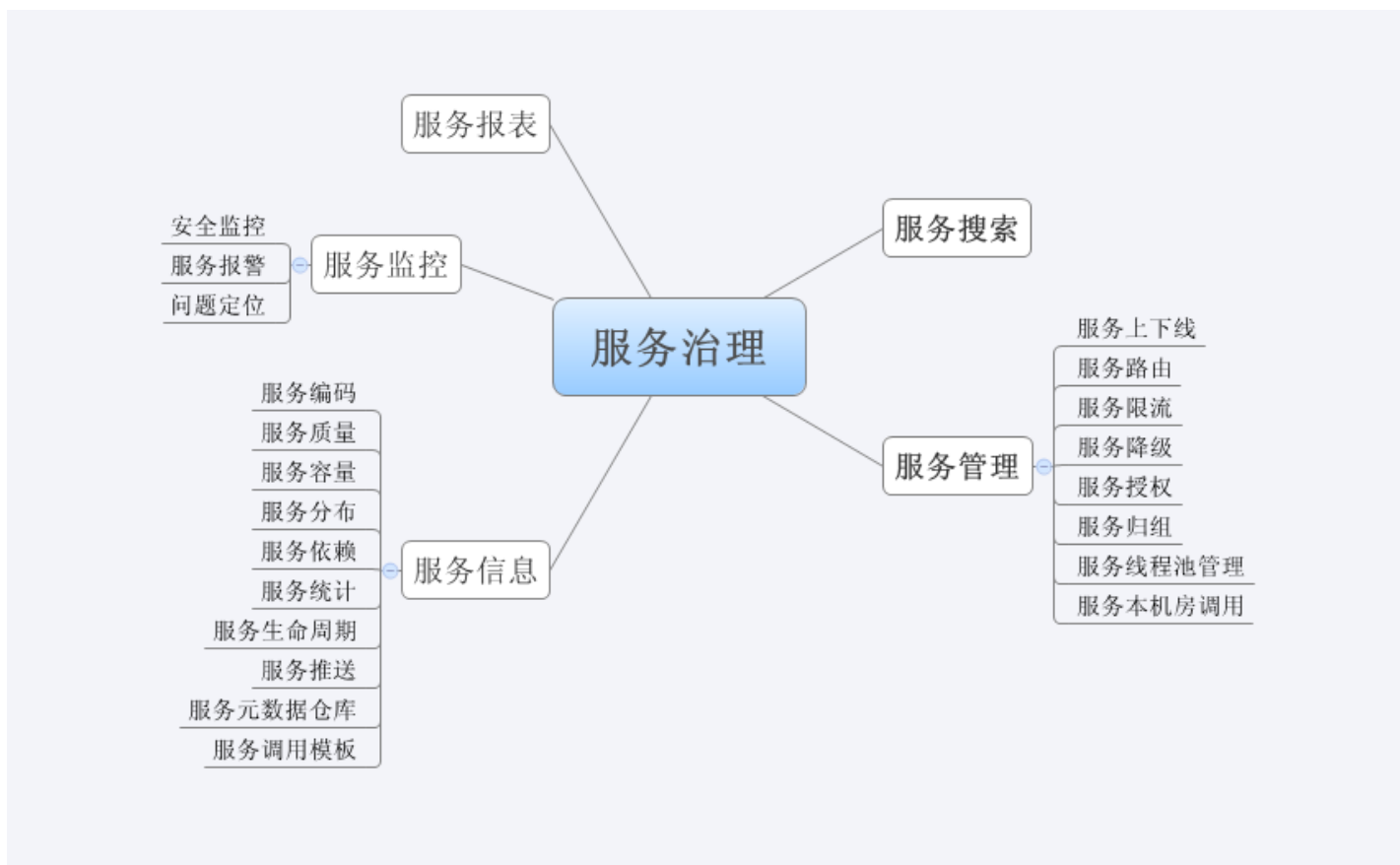
- 服务本机房调用
 - 基于服务名或者应用名订阅机房调用规则
 - 默认非本机房调用
 - 业务场景： 机房容灾演习
- 虚机房规则
 - 将几个机房看成一个机房做调用
 - 业务场景： 双11，机房容量不足



- 服务限流
 - 应用白名单
 - 阈值规则
 - 消费者级别限流



• 整体结构图





- 服务搜索
 - 以服务为粒度搜索服务相关信息

HSF服务治理

服务搜索

服务管理

报表

容灾演习

服务信息

小工具

日常

服务

应用

IP

com.taobao.item.service.ItemService:1.0.0.daily-L1

搜索

服务全称: **[com.taobao.item.service.ItemService:1.0.0.daily-L1]**

应用名: **itemcenter**

应用负责人: **玄宵**

测试负责人: **玄宵**

SCM: **玄宵**

PE: **玄宵**

版本分布: **1.4.9[1台], 1.4.9.2[3台],**

发布者

消费者

机器

超时

序列化方式

组别

10.232.12.113:12200

10000

java

[HSF]

10.232.12.114:12200

10000

java

[HSF]



- 服务管理
 - 服务上下线
 - 服务路由
 - 服务降级
 - 服务归组
 - 服务线程池管理
 - 虚机房规则
 - 服务授权



- 服务信息
 - 服务编码
 - 服务质量
 - 服务容量
 - 服务机房分布
 - 服务统计
 - 服务生命周期
 - 服务推送
 - 服务依赖
 - 服务调用模板
 - 服务元数据仓库



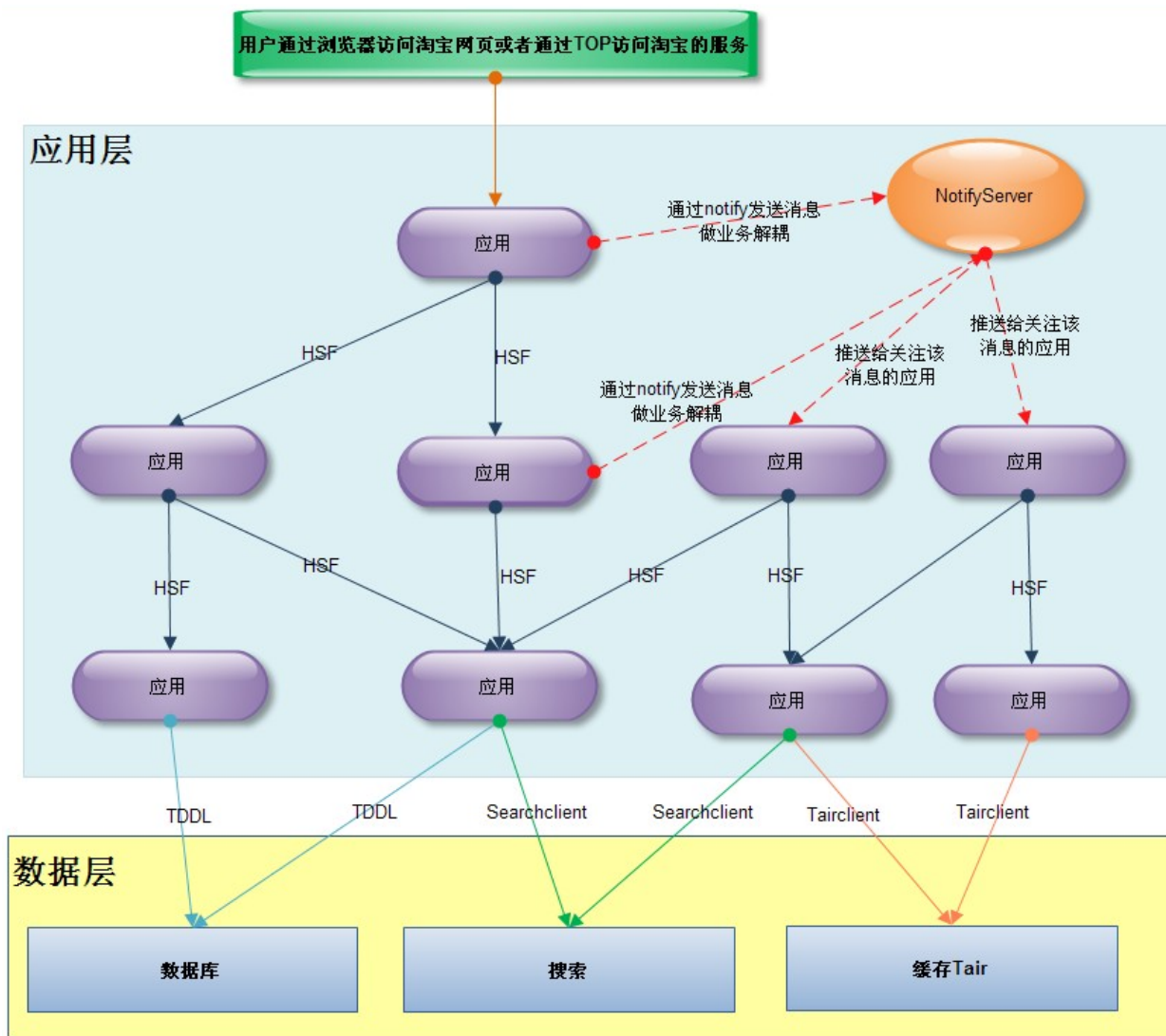
- 服务监控
 - 安全监控
 - 报警
 - 问题定位



- 类似google的[dapper](#)， Twitter [Zipkin](#)
- 基于tcp方式， http方式支持但是未全局推广

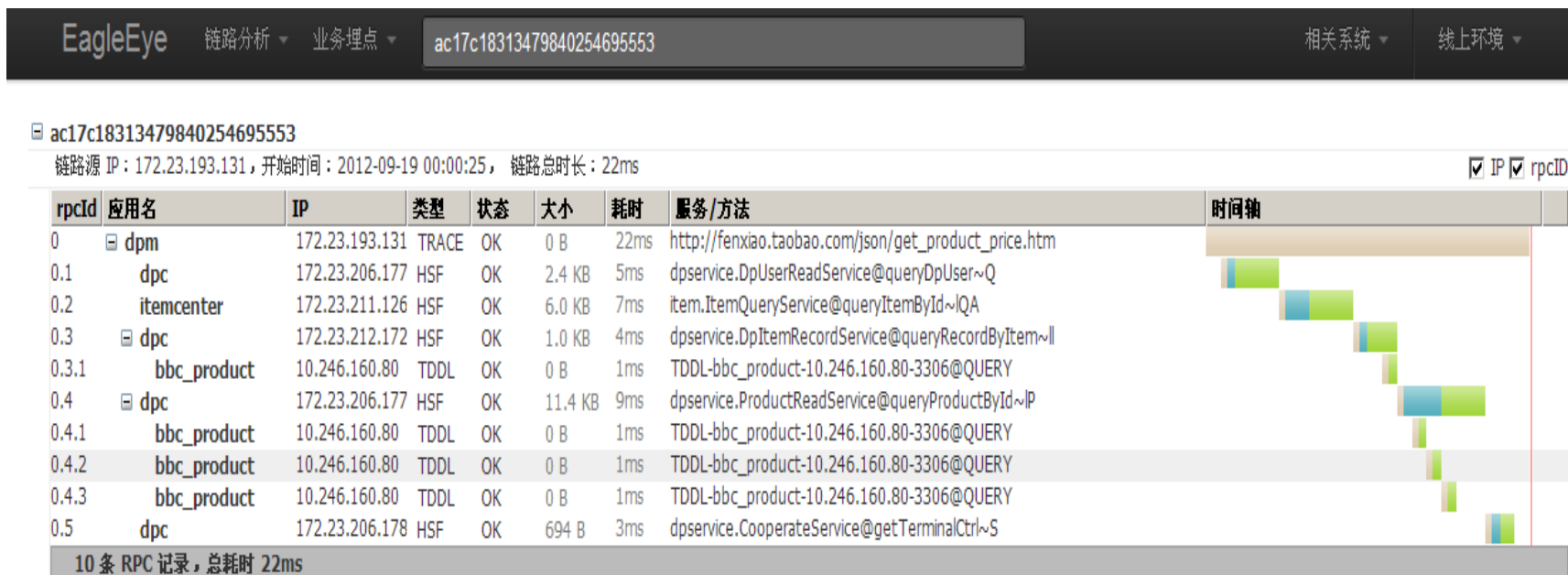


分布式跟踪系统





• 分布式跟踪系统链路图





- 分布式跟踪系统基本元素
 - 全局唯一ID
 - 链路顺序的rpcID
 - 响应时间
 - 请求，响应大小
 -



- 分布式跟踪系统带来的价值
 - RPC层
 - 应用调用链路分析
 - 服务依赖检测
 - 性能优化
 - 用户行为分析
 -
 - 业务层
 - 子账户系统
 - 账号追踪
 - 风险控制
 -



Alibaba
technology Association

问题

QA?



QCon全球软件开发大会（北京站）2013

即将启动敬请关注

www.qconbeijing.com

全球企业开发大会

INTERNATIONAL
SOFTWARE DEVELOPMENT
CONFERENCE