

www.qconferences.com
www.qconbeijing.com



伦敦 | 北京 | 东京 | 纽约 | 圣保罗 | 上海 | 旧金山
London · Beijing · Tokyo · New York · Sao Paulo · Shanghai · San Francisco

QCon全球软件开发大会

International Software Development Conference

Erlang开发实践

淘宝核心系统资深专家 余锋

<http://yufeng.info>



@淘宝褚霸

2013-05-25



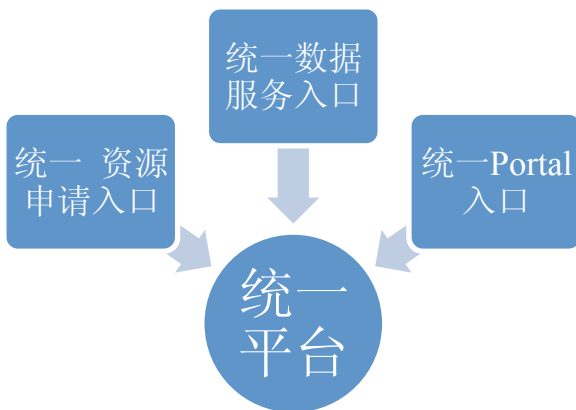
平台架构 工程实践 讨论

MySQL平台(UMP)要解决什么？

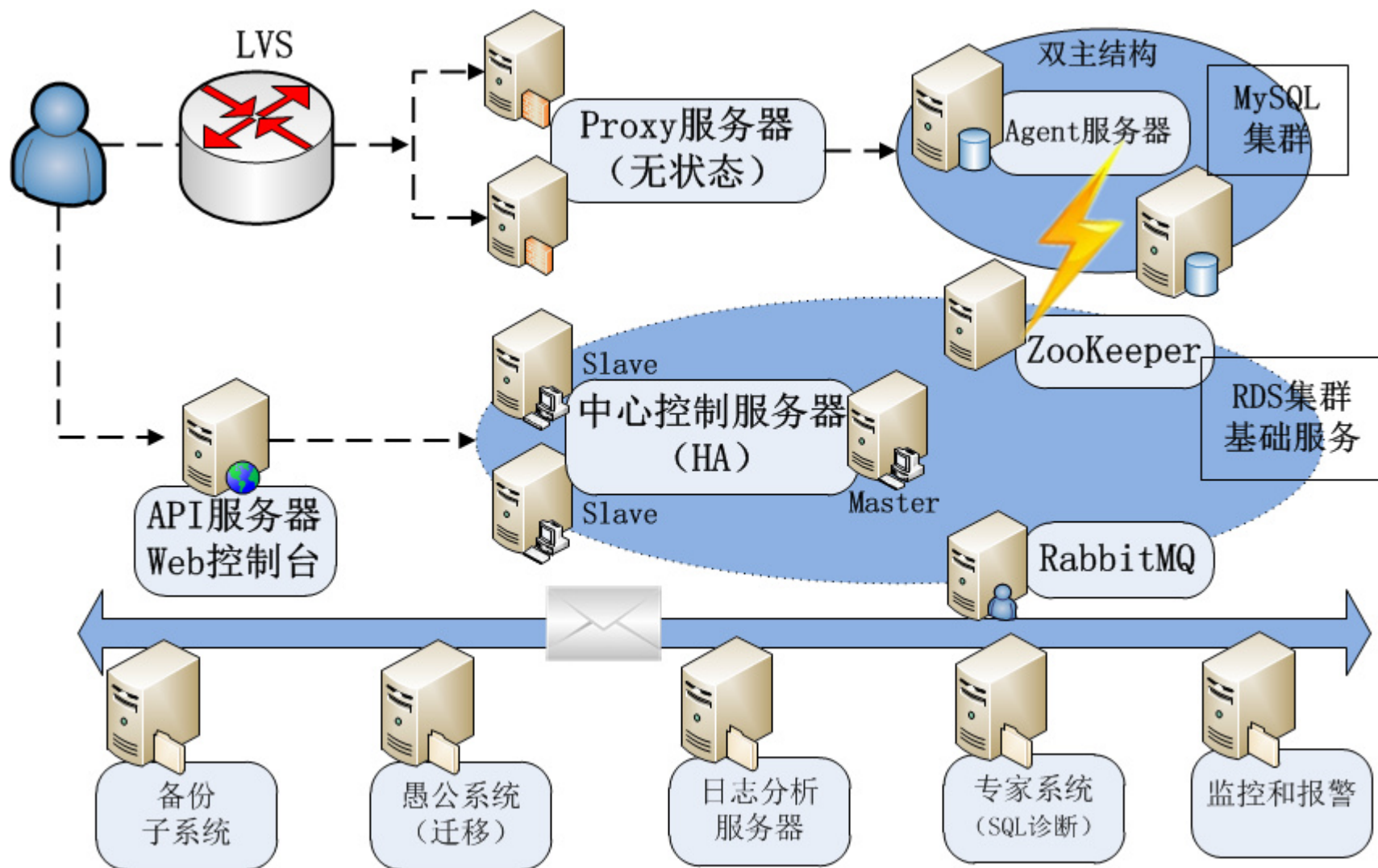
- 高效MySQL数据库服务支持，自动化运维
- 提高资源利用率，降低平台成本
- 7*24长期运行，屏蔽软硬件的变化



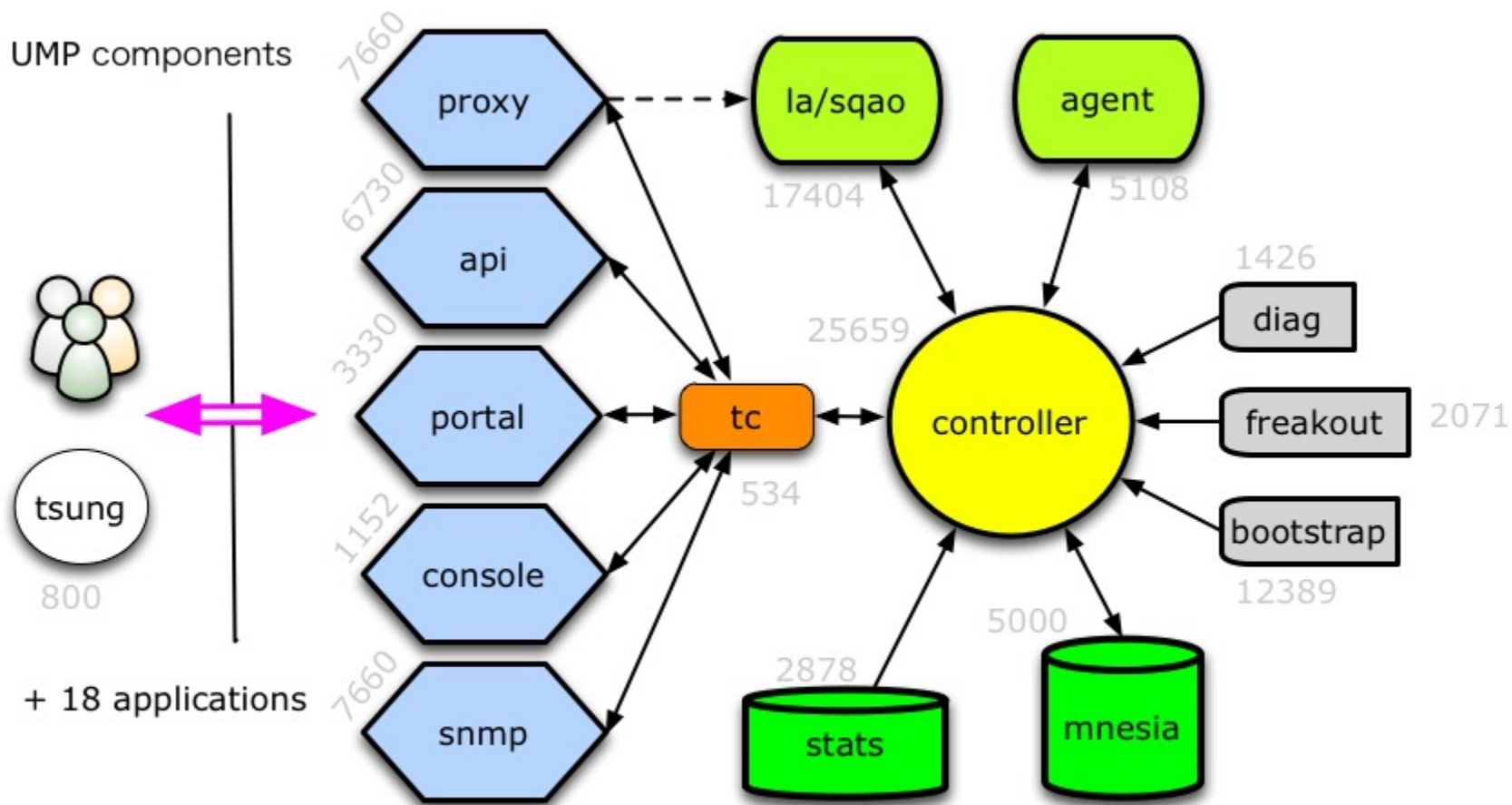
- 平台对外保持单一入口，对内维护单一的资源池。
- 保证服务的高可用性，消除单点故障
- 保证系统是弹性可伸缩的，可以动态的增加、删减计算与存储节点。
- 保证分配给用户的资源也是弹性可伸缩的，资源之间相互隔离。



UMP整体架构



UMP内部组件图

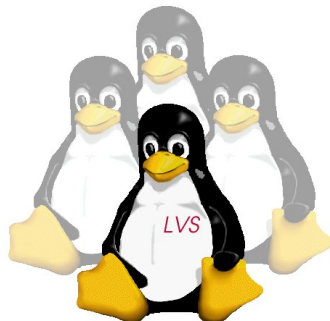


为什么要用Erlang实现

- 高并发、高性能、集群易扩展
- 时间检验的高可靠
- 强大的管理功能，方便的问题定位支持
- 强大的交互性，与其他系统整合能力



- 经验
 - 采用开放成熟的第三方部件的好处
 - 开放的平台，方便用户扩展
 - 热部署和升级对不停机维护的意义
 - 容错系统设计的重要性

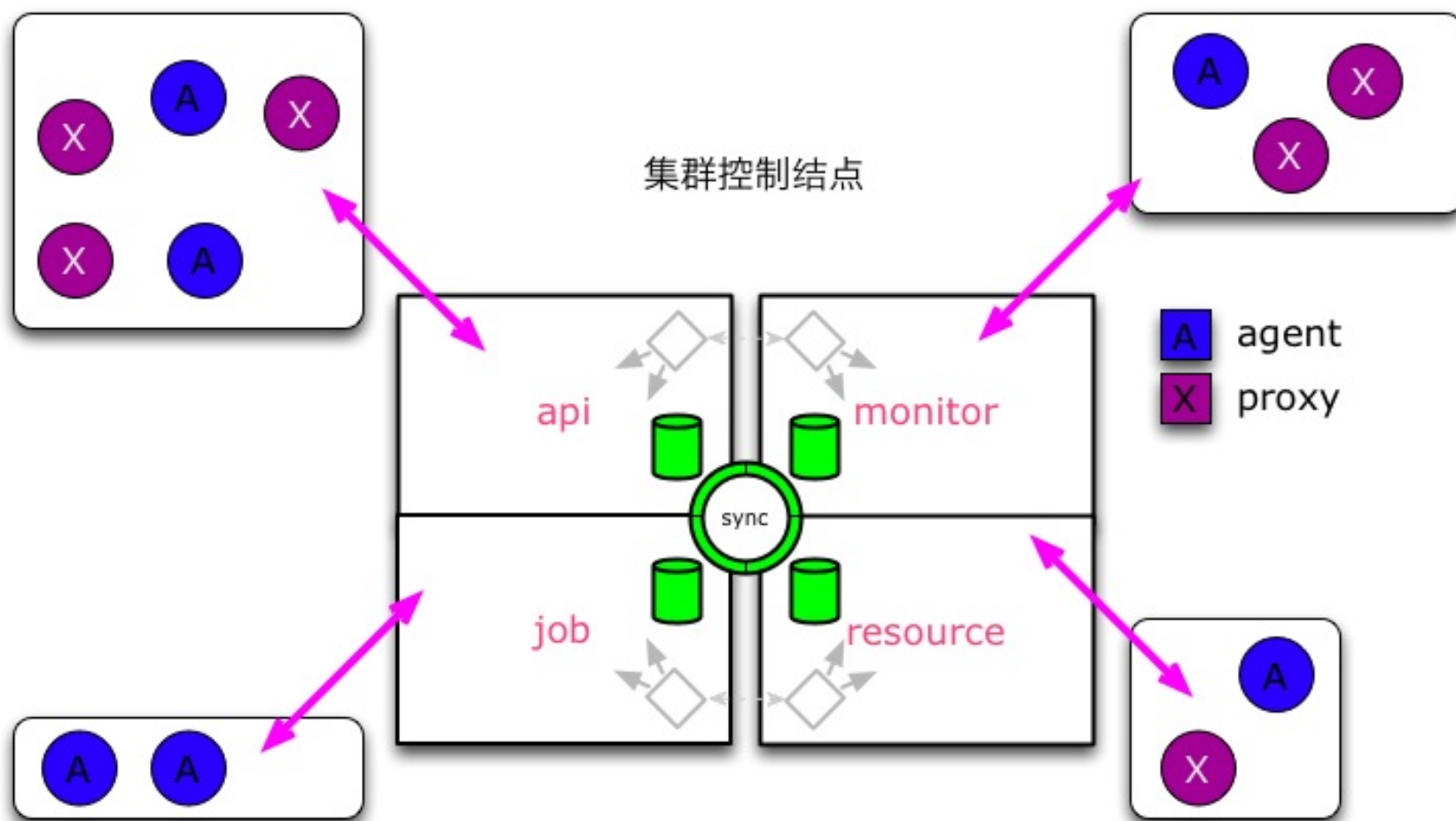




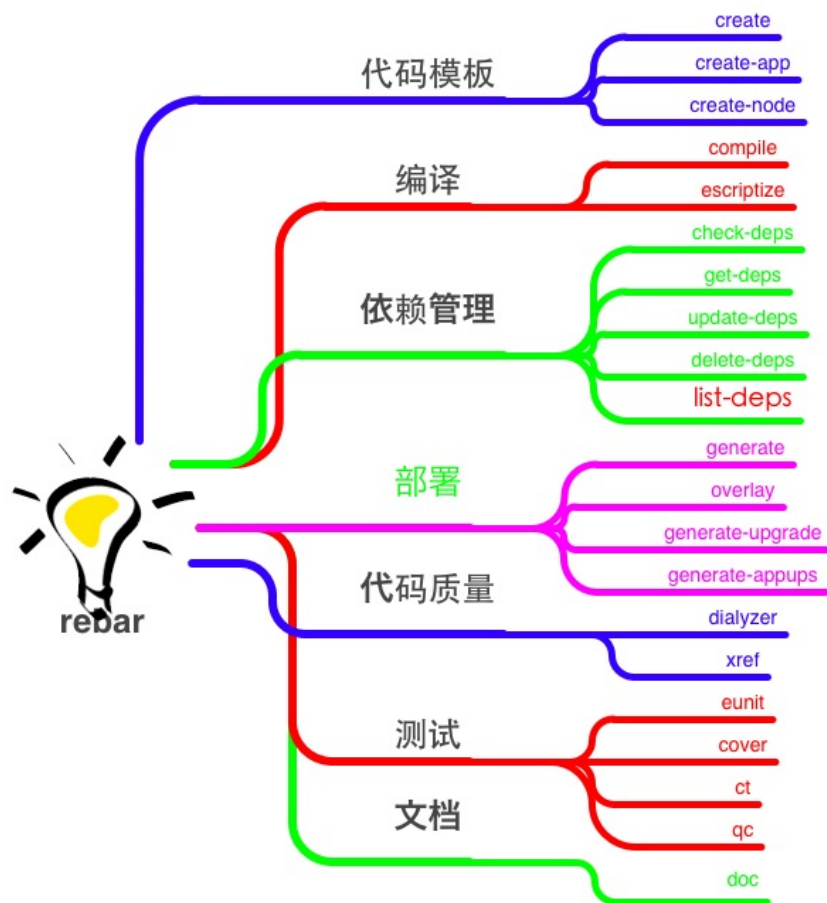
平台架构 工程实践 讨论

- 哲学
 - 小而美
 - 细节是魔鬼
 - fail fast
- 三高
 - HIGH cost performance
 - HIGH Reliability
 - HIGH Scalability

设计—典型三层



- 三化
 - rebar化
 - application化
 - 插件化
- 代码组织
 - api化
 - rpc松散化
 - 接口平民化
- 代码质量
 - 警告零容忍
 - dialyzer
 - tidier



```
$ ls -l deps/ump_controller/  
include  
priv  
README.md  
rebar  
rebar.config  
src  
THANKS  
...
```

实现—以app为单位扁平组织

```
$ tree -L 1 deps
```

```
deps
```

```
|—— ...
```

```
|—— ump
```

```
|—— ump_agent
```

```
|—— ump_alarm
```

```
|—— ump_bs
```

```
|—— ump_freakout
```

```
|—— ump_la
```

```
|—— ump_tc
```

```
|—— ...
```



```
$ ls -al deps/ump_controller/src
```

```
...
```

```
ump_instance_job.erl
```

```
ump_job_clone.erl
```

```
ump_job_delete_share_user.erl
```

```
ump_metastore.erl
```

```
ump_metastore_instance.erl
```

```
ump_mnesia_monitor.erl
```

```
ump_monitor_node_api.erl
```

```
ump_monitor_resource_api.erl
```

```
$ cat ump_diag_nodes.erl
```

```
...
```

```
get_nodes() ->
```

```
    case ump_diag_rpc:call(ump_monitor_node_api,  
registered_node, []) of
```

```
        {badrpc, _Reason} ->
```

```
            [];
```

```
        Nodes ->
```

```
            Nodes
```

```
end.
```

```
$ cat deps/ump_diag/Makefile
all: compile escriptize
compile:
    ./rebar get-deps compile
escriptize:
    ./rebar escriptize
```

```
$ dialyzer ebin
```

```
Checking whether the PLT /home/  
chuba/.dialyzer_plt is up-to-date... yes
```

```
Proceeding with analysis... done in 0m5.67s  
done (passed successfully)
```

- reltool 和 release_handler <=> 打包和热部署
- ssh/sftp <=> 安装和升级包传递
- ump_bootstrap <=> 大规模自动部署

- 白盒 $\Leftrightarrow$ eunit
- 黑盒 $\Leftrightarrow$ common test
- 覆盖率 $\Leftrightarrow$ cover
- 压力测试 $\Leftrightarrow$ tsung
 - api
 - mysql
 - console

- 系统基本信息<=>observer
 - crashdump
 - etop
- 业务信息<=>console/portal
- 主动诊断 <=> ump_diagnostic
- 主动破坏 <=> ump_freakout

- 热更新 <=> 不停机
- snmp监控 <=> 方便对接现有监控系统
- 节点管理 <=> 随时添加、移除、启停结点
- 集群安全 <=> 审计

- 消息流动<=>dbg
- 函数级别<=>eprof
- 驱动层面<=>dtrace
- 锁竞争<=>lcnt
- 角色协作<=>percept



平台架构 工程实践 讨论

- mnesia brain split
- 节点间rpc通道压力
- 小众语言

谢谢大家！

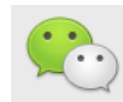
 @淘宝褚霸

www.infoq.com/cn

InfoQ^{ueue}



@InfoQ



infoqchina

软件
正在改变世界!