DAE

PyCon China 2013

About Me

- 洪强宁 / hongqn
- 2002年开始接触Python
- 2004年开始完全使用Python工作
- 2006年加入豆瓣，现任首席架构师
- www.douban.com/people/hongqn/
- hongqn@douban.com

议程

- DAE是什么
- 为什么要开发DAE
- DAE的架构
- DAE的特点
- DAE的未来

DAE是什么

Douban

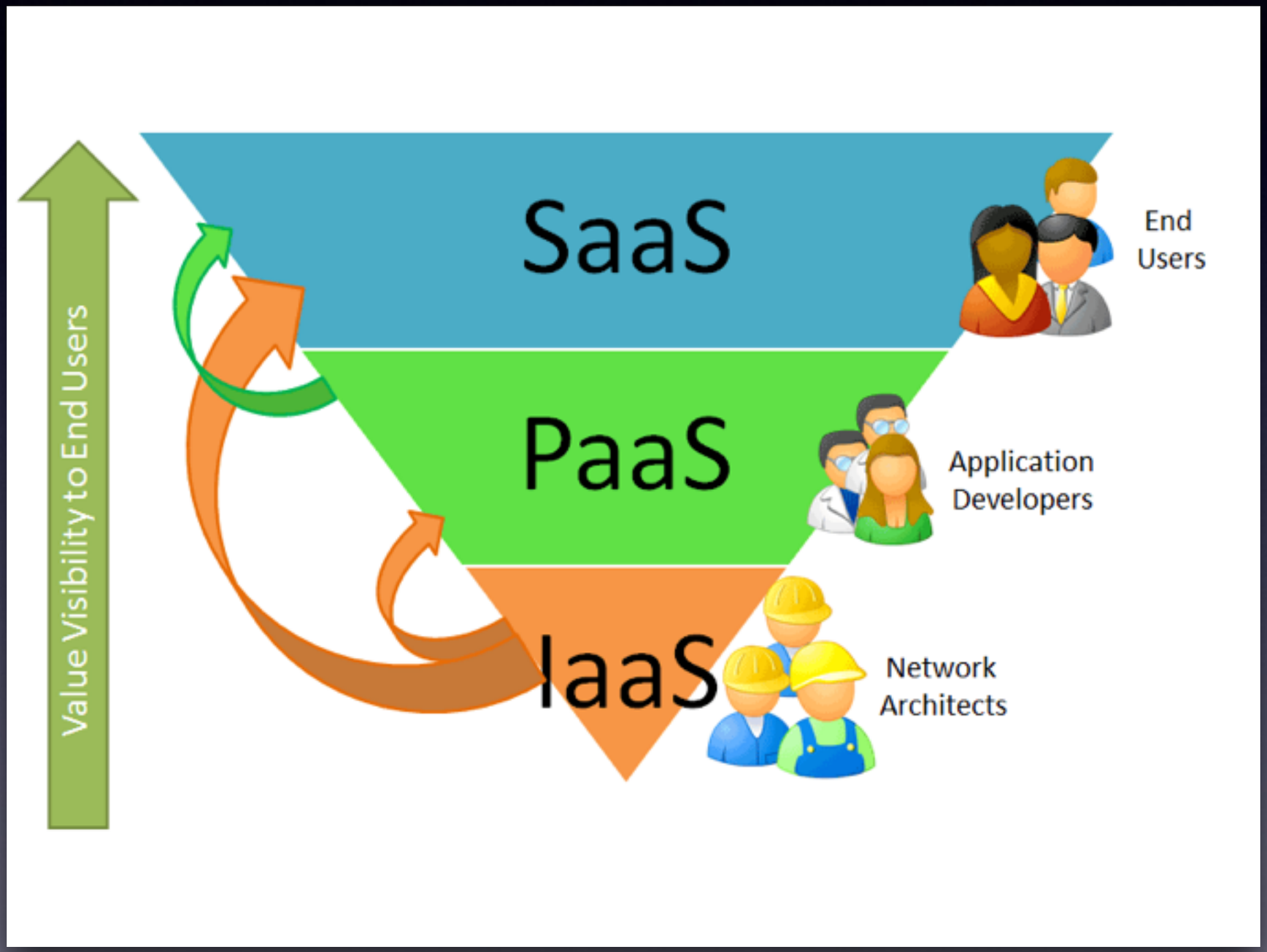
Adorable

App

Engine

PaaS

Platform as a Service (PaaS)



Private Python PaaS

Current State

427 apps

126 对外应用

bubbler.labs.douban.com

豆瓣 douban

首页 | 我的豆瓣 | 豆邮(0) | 我的小组>>

hongqn,说点什么吧

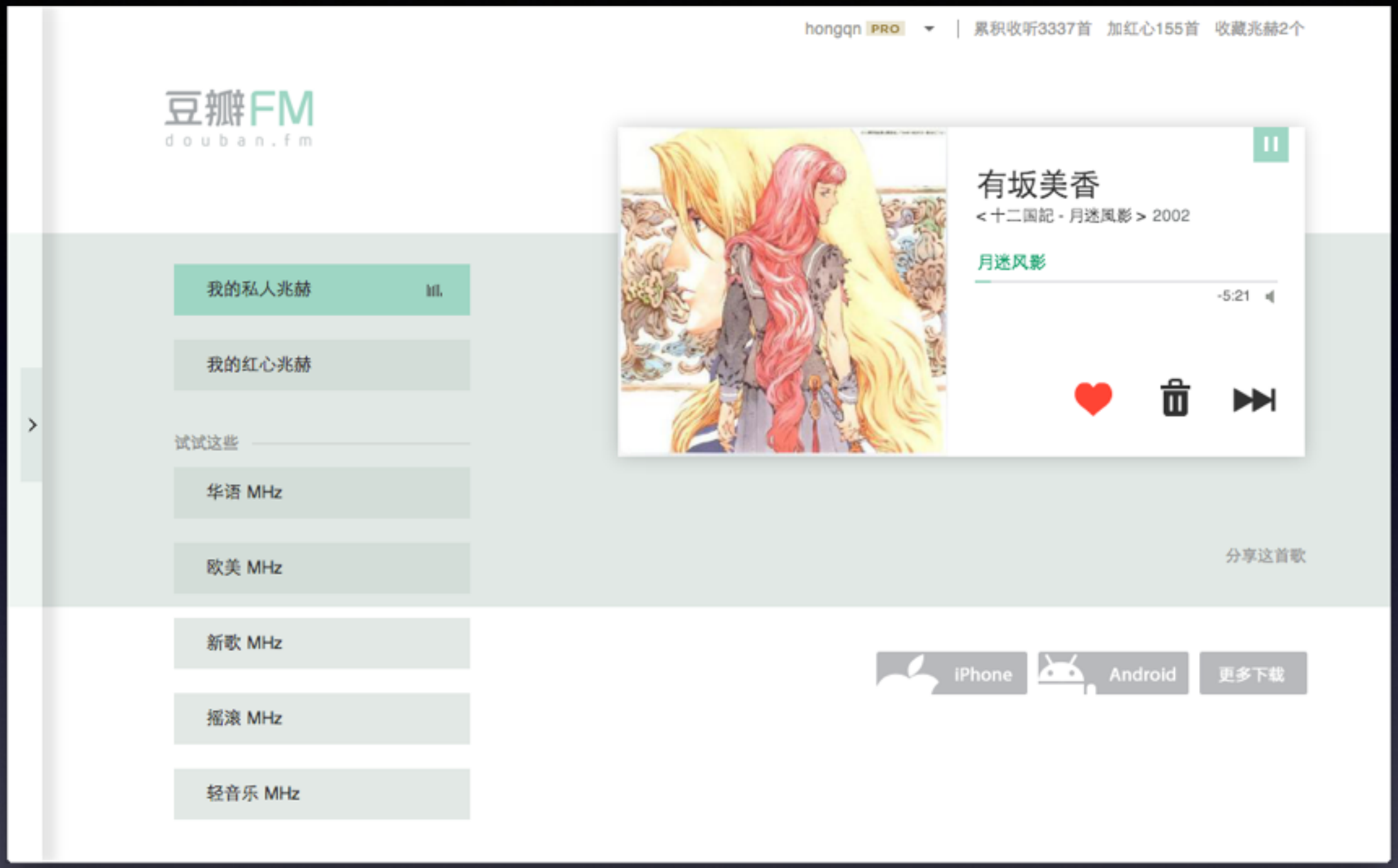
我说

天川六壁坂：为了维系新选组，早年的土方岁三一直强迫自己戴着“魔鬼”的面具【石假面】，成为了冷酷而残忍的“魔鬼副长”，让人闻风丧胆，使新选组成为了史上最强大的武士团体。维新多年后的他终于得以释放自己，露出了温和善良的本性，改名为【荒木飞吕彦】开始了漫画的创作。

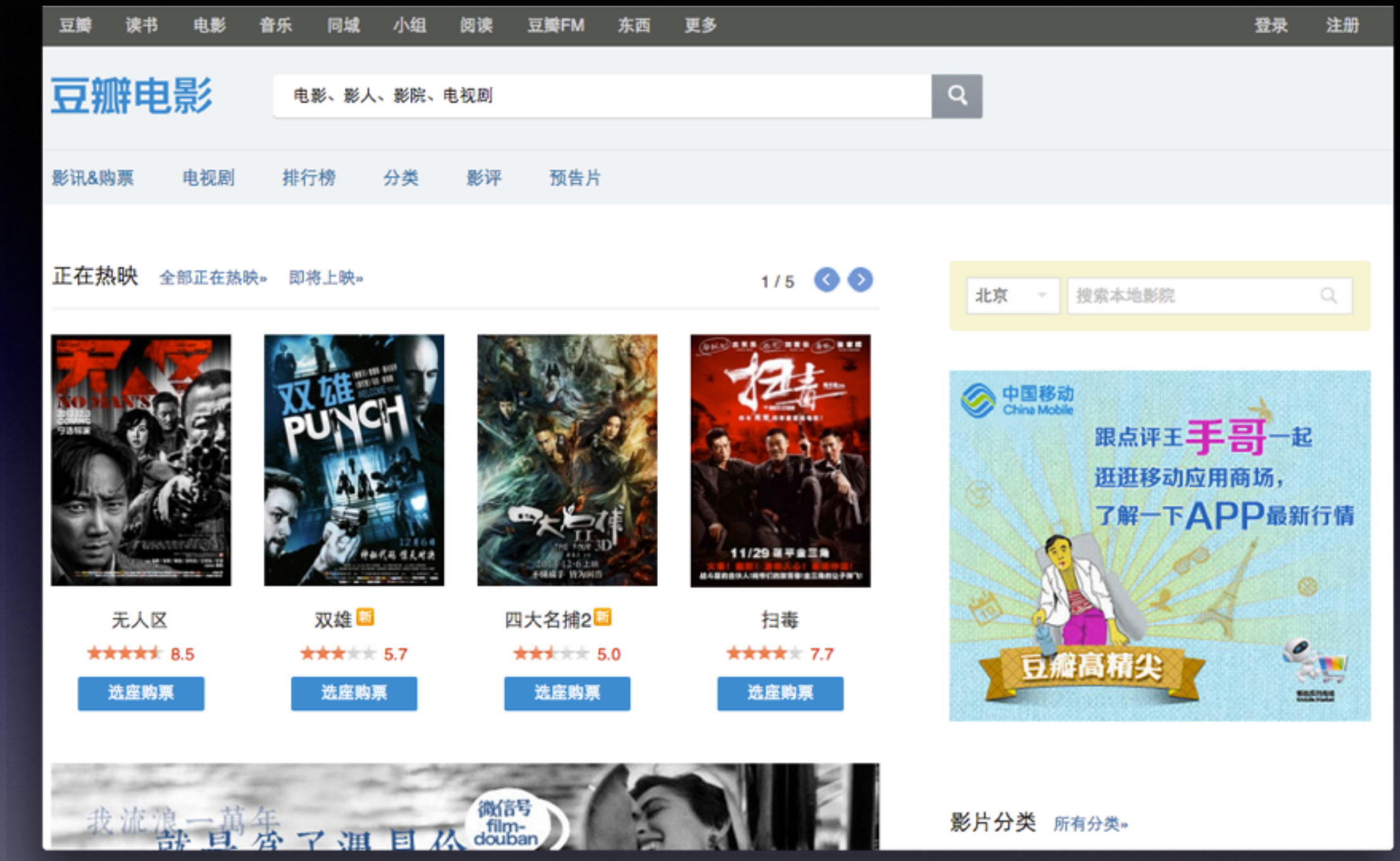
12-06 16:42:59 · 1回复

由iano转播

m.douban.com


hongqn PRO
累积收听3337首 加红心155首 收藏兆赫2个
豆瓣FM
douban.fm
有坂美香
<十二国記 - 月迷風影> 2002
月迷风影
-5:21
我的私人兆赫
我的红心兆赫
试试这些
华语 MHz
欧美 MHz
新歌 MHz
摇滚 MHz
轻音乐 MHz
分享这首歌
iPhone
Android
更多下载

douban.fm



movie.douban.com

group.douban.com

Search... Timeline Teams Fair Gist Extra 542 ms 101

dae

Not Released 2 New Issue Pull Request Unwatch 38 Fork 34

Code | User Manual | Dev Doc | Blog | Pull Requests 12 | Issues 137 | Graphs

Open You're asking to pull 2 commits into `dae : master` from `liutianwei/dae_t… : add-webtest-flag-in…` 3小时前

Discussion | Commits 2 | Files Changed 3 | Notification ON OFF

liutianwei opened 2013-12-06 14:15:51

add webtest flag in dae test command

- 进行webtest和apitest时，需要启动dae server，此时使用的mc，beansdb应该与dae server 保持一致，所以需要对其进行monkey patch，但是对于unittest， 依赖于shire环境的，会使用corelib.tests.framework中的mc，beansdb等设置，此时不需要进行monkey patch, 所以加一个标志进行区分
- 补充test的相关文档

@dae pls, review, thanks

2 participants

liutianwei added some commits 2013-12-06 14:15:55

liutianwei	add test flag for apitest or webtest	2013-12-06 14:12	3c985a2
liutianwei	update doc	2013-12-06 14:12	fdfe5ea

code

豆瓣上线系统 shire

hongqn,退出

正在上线

169719, 2013-12-06 17:43:20

紧急终止

详情

最近的上线

资料更新于 12-06 17:49:27

版本	时间	Schedule id	Action
169719	2013-12-06 17:43:19 [5]		
169705	2013-12-06 16:54:47 [4]		
169688	2013-12-06 14:12:19 [3]		
169658	2013-12-06 10:19:07 [2]		
169656	2013-12-06 09:11:37 [1]		

最近的灰度上线

stage 0 rollback to last release

版本	时间	level	状态	promotion
169719	2013-12-06 17:41:14	1	成功	stage 0 stage 1 promote
169705	2013-12-06 16:51:47	1	成功	stage 0 stage 1 promote
169705	2013-12-06 16:50:08	0	成功	stage 0 stage 1 promote

avgerage: 1, 5, 15(mins)

0.00 0.00 0.00

top errors(1 minute):

recent errors:

up

搜索姓名，豆瓣id，邮箱
0
广播动态
最新图片
事件
群组
我的卡片
全部成员
时间线
豆瓣花名册
2027 ms
豆瓣事件 (创建新事件)
Griffin Riders Camp
对儿勾
18张照片 · 2013-11-25
Q4扫雷+用户服务+内网三部门联合TB
某S
199张照片 · 2013-11-23

花名册

电影-基于长评 | 关于Silmarils | movie_rl:1292720

Silmarils

欢迎来到 Silmarils(发音：seel'mareels)，即《指环王》中的精灵宝钻，该词来自昆雅语，其意为"纯光之辉"。据传，精灵宝钻共有三颗，所以用其复数形式 Silmarils，也代表本框架可用于多种产品。

采用Silmarils来命名有几个原因：条目之间的多重相似性就像钻石的多个刻面；silmarils像是similars的变位词，刚好用以表达"多重相似性"；很喜欢《Silmarillion》这本书。

> *在那段时期，精灵制造了最著名的一件作品传世。正处於盛年的费诺，心中兴起了新的念头，或者说，一些事先已经预知的命运阴影，逐渐接近笼罩住他。 他不断思考，双圣树的光辉，这蒙福之地的荣耀，该如何永远保存不灭。於是，他展开了一场漫长又辛苦的工作，殚精竭虑，运用他所有的力量、知识、与精微的技巧，终於创作出了「精灵宝钻」。*
>
> *它们的模样看起来像是三颗巨大的钻石。但是，除非直到未了，直到在日头被造之前消亡的费诺折返，静候在亡灵的殿堂，不再见於亲族之间； 直到太阳成为过去，月亮永不升起之时，制成这三颗宝石的物质才会揭晓。..... 即便是在最深最黑暗的宝库中，精灵宝钻所放射出来的光芒，仍如瓦尔妲的星辰一般闪亮， 何况，它*

使用帮助——条目多重相似性

- 可以搜索特定条目的多重相似性，试试在搜索框输入movie_rl:1292720。
- 在条目页面点击某个特定的条目，将显示该条目的多重相似性。
- 在条目页面点击条目旁边的图标小箭头，将显示该条目的豆瓣页面。

使用帮助——整体聚类

- 可以查看每个产品的聚类结果，试试点击导航栏中的"电影-基于长评"。
- 可以在聚类结果中查看一级分类下包含的二级分类。
- 可以查看各个类别对应的关键词，例如沉醉（可点击）将对应这些词：想入非非，心动不已，心醉，沉醉，浮想，神往，迷醉，心驰神往.....
- 可以查看每个类别的代表条目，列在类别关键词之下。

后续工作

1. 由专业编辑进行类目标定。
2. 目前只使用了影评信息，接下来会把影评和标签结合起来。
3. 应用于图书条目，获得多重相似性。
4. 应用于豆瓣阅读，根据文本内容自动生成多重标签。
5. 应用于电商，基于东西的文本计算商品之间的多重相似性。
6. 基于文本内容的跨领域推荐。
7. 探索应用于其他产品的可能性。

精灵宝钻

…等等

每天处理 2.4 亿动态请求

5K 峰值 qps

on 32 nodes

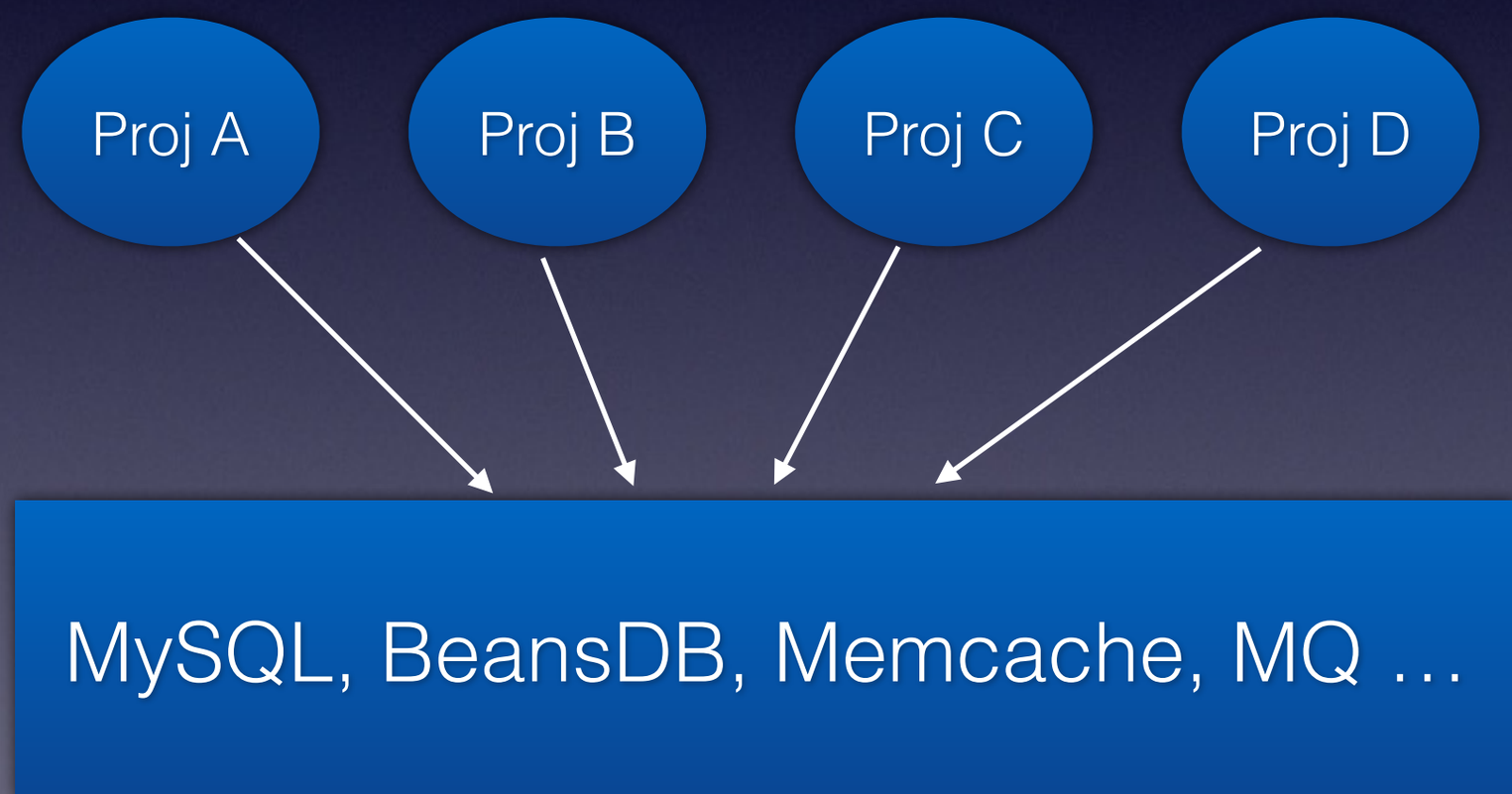
为什么要开发DAE

代码拆分

1 big svn repo (17万提交)

4602 git repo (2881 fork)

重用基础设施



简化新项目启动

```
$ dae create
$ dae serve
$ dae deploy
```

最佳实践对
每个app开箱即用

持续集成

分级上线

故障报告

状态收集

Scale SA

4

DAE的架构

app.yaml

```yaml
application: movie
version: 1
runtime: python
api_version: 1
enable_shire: True
worker: sync
sentry_dsn: udp://[redacted]@onimaru-udp.intra.douban.com:4008/77

irc_report_channel: movie
report_observer_emails: fuyuquan@douban.com

release_managers:
- torpedoallen
- fuyuquan
- guohaochuan
- py

handlers:
- url: /swf/(.*)
  static_files: static/swf/\1
  upload: static/swf/.*

- url: /api/.*
  wsgi_app: app:api_app

- url: /.*
  wsgi_app: app:app

instances:
- url: /api/.*
  instance: api

- url: /service/api/.*
  instance: api

mqs:
- name: _async
  handler: foobar
```

PyPI 镜像
pypi.douban.com

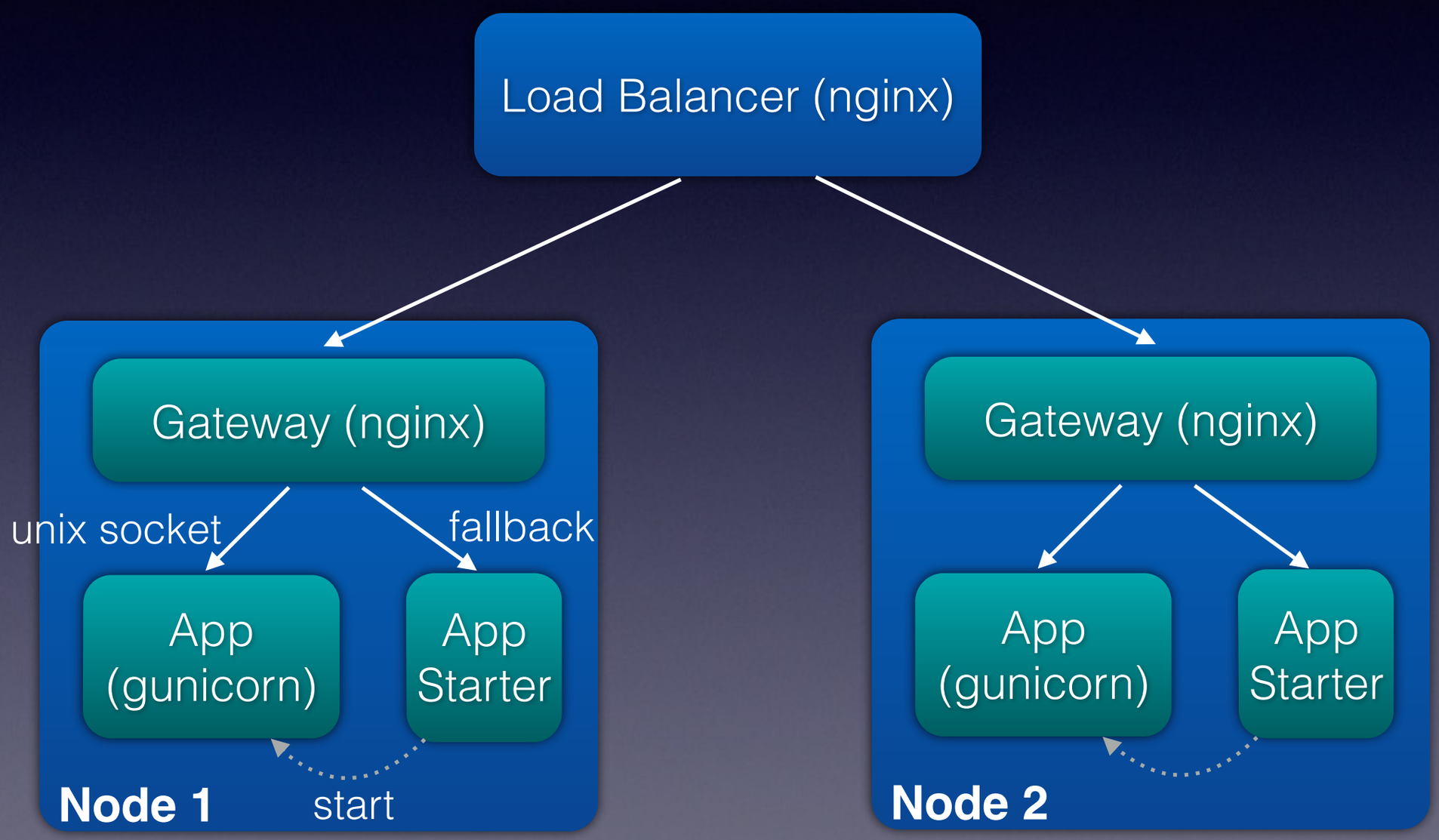
pip-req.txt

```
$ dae install flask
```

Instance

- web
  - gunicorn
- service
  - gunicorn + customized worker
- daemon
  - zookeeper

两级结构


Load Balancer (nginx)
Gateway (nginx)
unix socket
fallback
App (gunicorn)
App Starter
start
Node 1
Gateway (nginx)
App (gunicorn)
App Starter
Node 2

资源池API

- mysql
- memcache
- doubandb, doubanfs
- moosefs
- beanstalkd
- cdn
- statsd + graphite
- mail
- elastic search
- irc

waylife
监控API调用

DAE的特点

Low Overhead

- 资源分配 — 进程
- 资源隔离 — UNIX帐号
- 资源限制 — 外部监控

Customizable

- 每个应用从模板生成nginx配置
- 使用 hook point 定制
- 或者只使用生成的 nginx 配置片段

自依赖

- 通过一个应用部署应用
- 通过一个应用管理应用
- 一个应用 scale 所有应用
- 一个应用 serve 跨应用的静态文件
- DAE 的代码托管在一个 DAE 应用上

greenify 已开源
github.com/douban/greenify

gevent

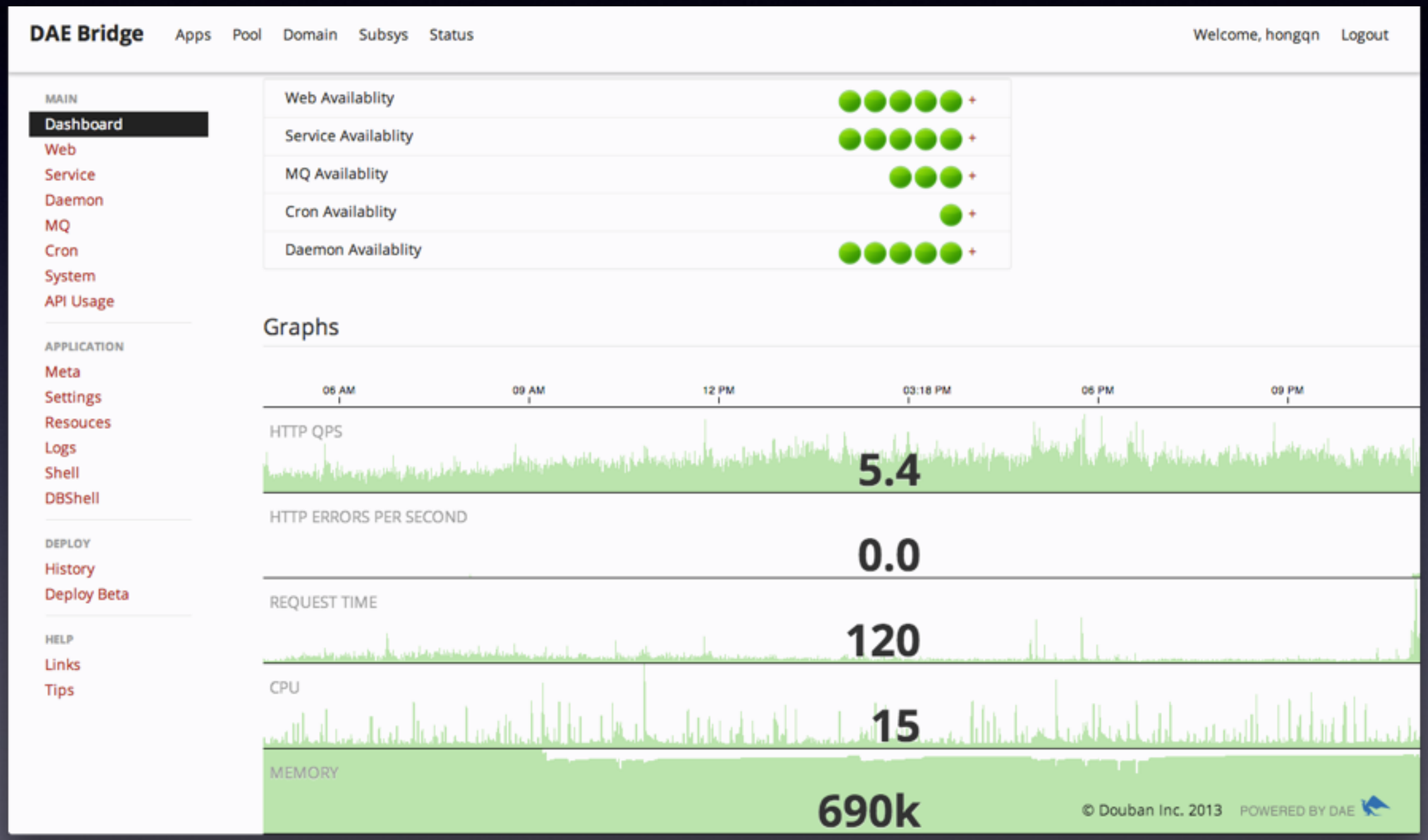
- 默认启用
- 支持长连接，如 websocket
- 节省内存
- 很多坑

thriftclient 即将开源

ice 是一等公民

- Thrift
- 在线客户端生成
- 依赖注册
- 升级通知

收集大量状态数据


DAE Bridge
Apps
Pool
Domain
Subsys
Status
Welcome, hongqn
Logout
MAIN
Dashboard
Web
Service
Daemon
MQ
Cron
System
API Usage
APPLICATION
Meta
Settings
Resouces
Logs
Shell
DBShell
DEPLOY
History
Deploy Beta
HELP
Links
Tips
Web Availablity
Service Availablity
MQ Availablity
Cron Availablity
Daemon Availablity
Graphs
06 AM
09 AM
12 PM
03:18 PM
06 PM
09 PM
HTTP QPS
5.4
HTTP ERRORS PER SECOND
0.0
REQUEST TIME
120
CPU
15
MEMORY
690k
© Douban Inc. 2013
POWERED BY DAE

cfgpusher 即将开源

最佳实践

- 错误收集 - sentry
- 在线 profile 工具
- 预发布环境
- 持续集成

- Code集成
- 打包发布
- 服务地址实时推送
- 分级上线

曹娥已开源
github.com/douban/CaoE

的Python环境

- 可用C扩展
- 可用 fork, subprocess, multiprocessing…

DAE的未来

豆瓣所有应用的平台

cgroups

离线大数据计算

更好的 service 体系

跨 IDC

QoS

公有云服务?

开源?

谢谢！