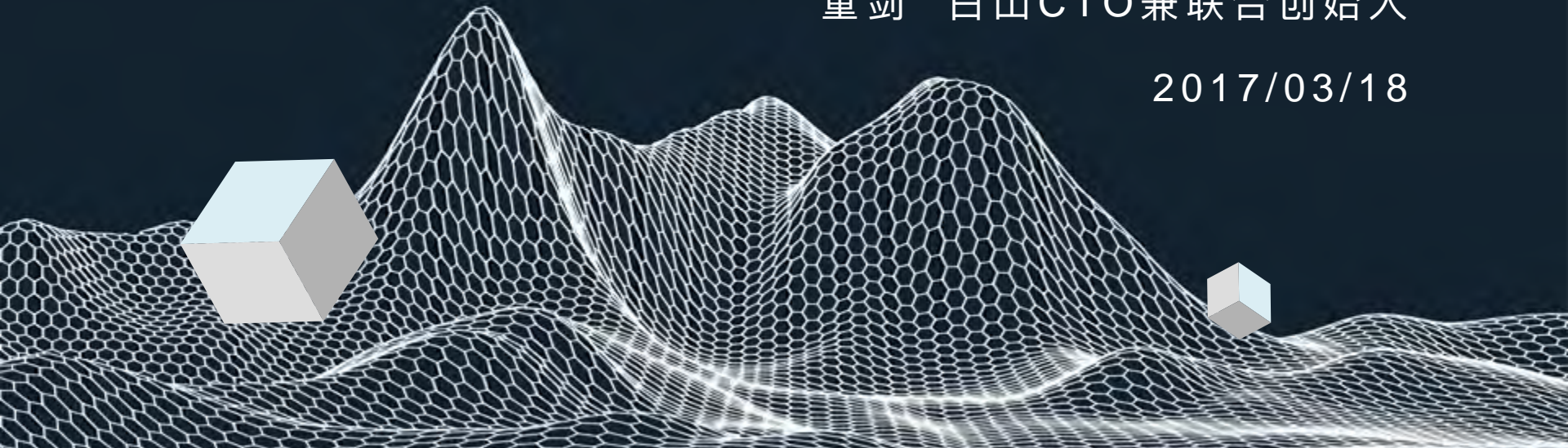




API网关在大数据开放中的应用

童剑 白山CTO兼联合创始人

2017/03/18



关于大数据的一些说法

- 数据是新的石油
- 数据是一种能源
- 马云：人类正从IT时代走向DT 时代

数据越来越多之后。。。

- 各国政府也在推动政务数据的开放，包括中国政府
- 企业和公共服务机构的数据如何二次利用？
- 数据开放共享有多种形式，该如何选择？

本次分享关键词：数据开放共享、API网关、API全生命周期管理

内容大纲

1. 目的、形式和趋势
2. 开发的难点所在
3. API网关和新的开发工具



大数据开放共享的目的

- 挖掘数据的商业价值
- 提升数据资源应用的能力

数据开放共享的形式

应用

文件

数据库

API

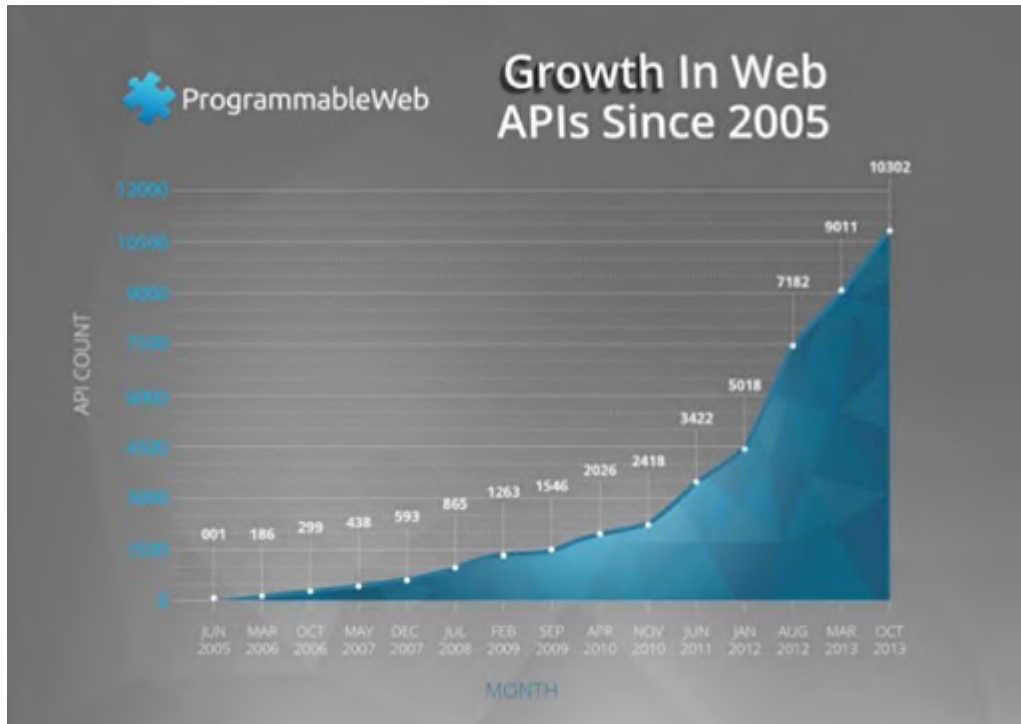
API正在变的越来越重要



API:
Application Programming Interface
应用程序编程接口

从最早的本地库函数API，到RPC和SQL模式，再到Web API模式，新的API应用形式越来越多，可以说API是未来数字生活和经济的基石

趋势一：数量的快速发展



数据统计来源：

ProgrammableWeb
该公司收录了超过
15000个开放的API

私有API数量更加巨大

趋势二：移动应用和云服务大量用API



- 1** 以API形式交付的云服务
- 2** 企业用API开放数据和功能
- 3** 应用的多平台化
- 4** 云服务之间相互集成

趋势三：越来越被物联网所依赖



NO API
NO IOT

物联网设备

依赖 云端 API

实现数据与通讯服务



更重要的趋势： API经济（互联网+）

API经济 (Economy)



国务院印发《促进大数据发展行动纲要》

企业正通过API经济重塑商业模式：保险、金融、交通

智能家居、智能汽车、物联网、工业4.0等，将会更多使用API

2016年美国50%的B2B贸易通过Web API完成 (Gartner)



API 开放数据也有点难

数据开放平台的架构



数据开放的各种顾虑

如何保证
安全性

开发需
要投入

代码和文档
维护

认证授权

数据
保密

运营支持

API的开发体验也很糟糕



- 性能和稳定性差，一个接口问题引起全局雪崩
- 数据源格式多样化，对接和管理复杂
- 没有API生命周期的概念，旧API下线很麻烦
- API无具体规范，沟通、调试、接入成本巨大
- 文档不全，或和代码不同步
- 升级过程中无法灵活控制新老版本的更替
- 测试成本较大，自己开发测试脚本，或人肉测试



怎么做好数据开放平台

API要这样开发才够愉快



- 统一的规范
- 自动生成文档、SDK、后端
- 统一的认证方式，不用重复实现
- 自动生成测试服务
- 前后端开发分离、无缝衔接
- 新老版本无缝上下线，或者多版本同时在线，逐步下线老版本
- 方便接入各种云服务，比如可以直接用 RunScope 进行API监控，而无需复杂的设置
- 配合网关、微服务以及新兴结构化存储服务简化后端开发

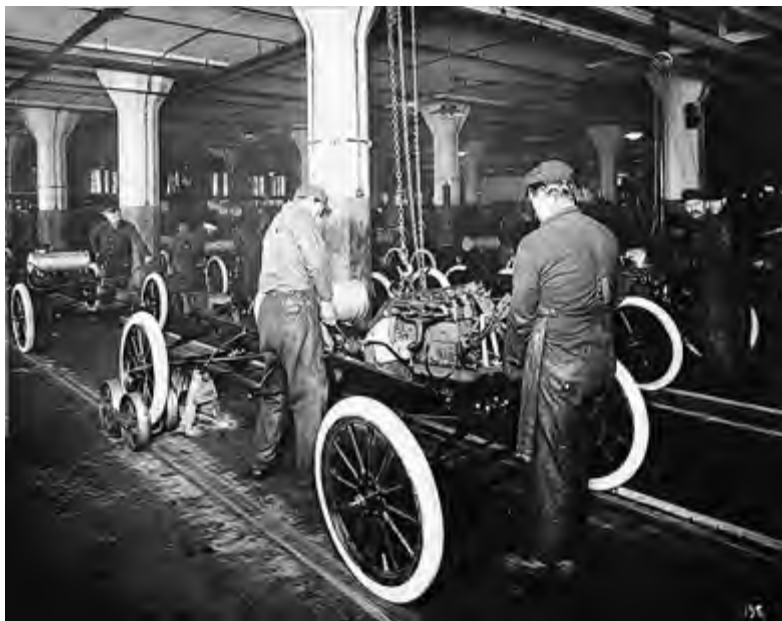
API开发和管理方法在变革



人工生产线

VS

标准化、自动化



VS





API 全生命周期管理

Full Lifecycle Management

- 生命周期十个阶段
- 管理体系架构
- API网关 - Gateway
- 设计和开发：Swagger/OAI
- 相关工具

API完整生命周期



重视生命周期的
每个阶段

API平台的管理体系

开发者服务

- Developer Portal
- 新手引导
- API目录
- API文档
- 服务条款
- 讨论区

API 生命周期管理

- 设计
- 开发
- 发布
- 管理和退休

API 管理

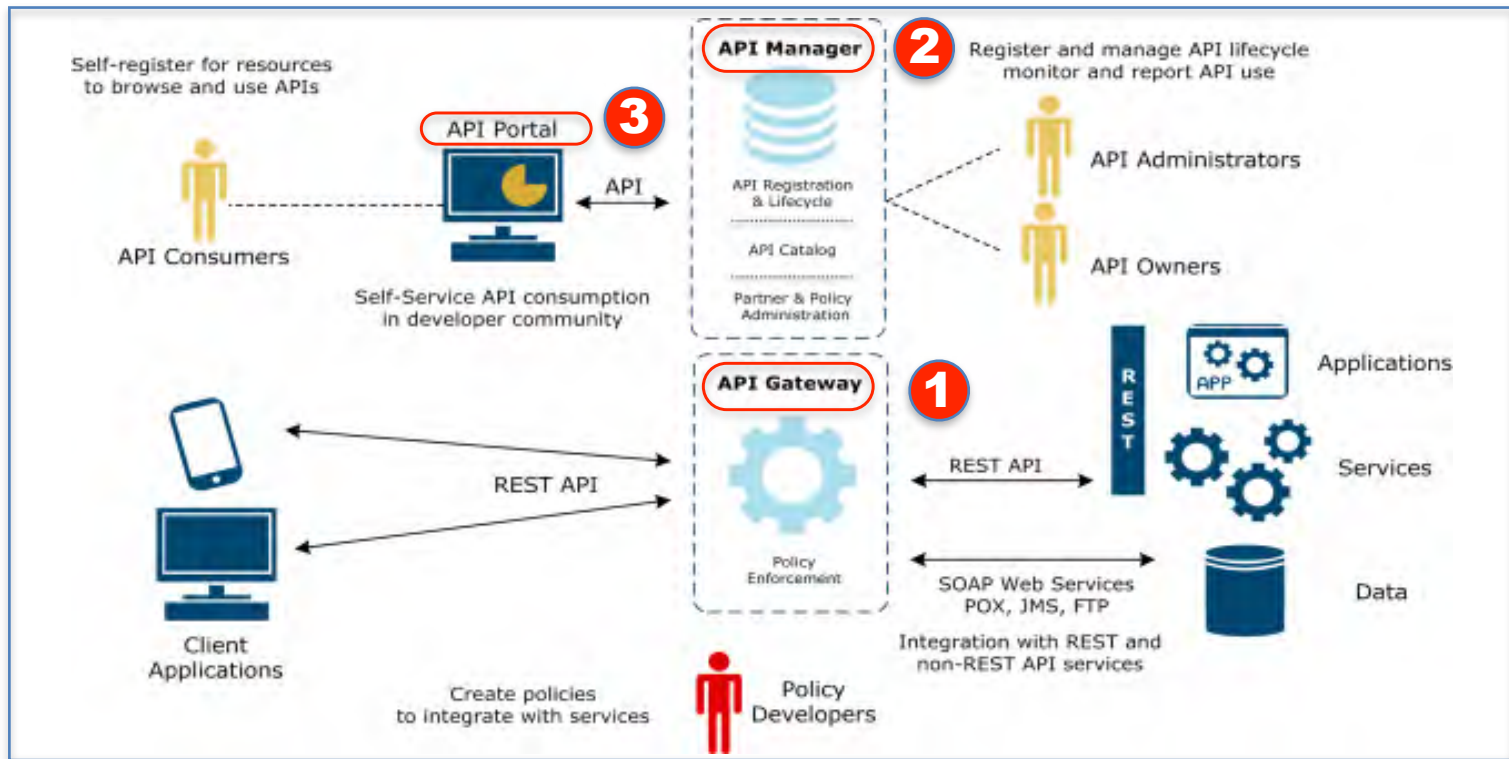
API 网关

- 认证和授权
- 加密
- 攻击防护
- 配额、流量控制

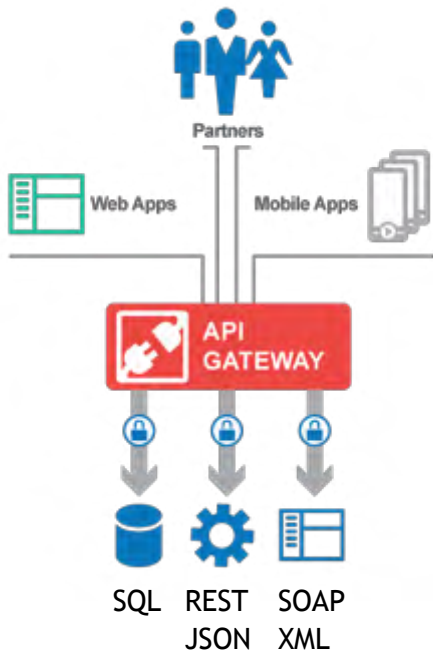
API 运维

- 监控和异常处理
- 版本管理
- 关键指标、SLA、分析
- 访问量、计费

API平台的技术架构



API网关(Gateway)



将所有API共性的功能统一放在Gateway上实现

1. 认证授权: OAuth2、OpenID、API Key、IP限制
2. 传输安全: SSL、TLS
3. 流量控制: 配额、过载保护、频率限制
4. 威胁防护: 防抓站、防撞库、DDoS/CC防护
5. 编排转换: 服务分组、7层路由、协议格式转换
6. 监控分析: 访问量实时统计、性能监控、访问分析
7. 统计计费: 针对不同用户、不同API分组
8. 缓存加速: 毫秒级缓存、请求队列、异步化

API设计和开发

前后端开发分离提高效率



目标：后端只负责数据的计算和输出，前端负责数据组织和展现

但一些问题是：

- 最后集成时商量好的数据结构发生了变化，需要反复
- API没有统一标准，前端人员花时间读文档和适配，重复性工作
- 前端需要等待后端完成API开发才能集成，说好的并行开发呢？

持续集成、持续交付、快速迭代如何实现？

前后端并行开发又如何实现？

随着移动App开发的兴起，Swagger、RAML等描述语言和工具出现了。

OAI和Swagger

Open API 计划 (OAI) 由前瞻性行业专家组成的联盟创建，他们认识到标准化 REST API 的标准的巨大价值，致力于推广其统一的 API 描述格式 (Swagger)。SmartBear 将 Swagger 规范直接捐赠给 OAI 作为其规范的基础。

Swagger 是一个规范和完整的框架，用于生成、描述、调用和可视化 RESTful 风格的 Web 服务，它由文档生成器、编辑器、代码生成器以及很多第三方工具组成。

The Swagger specification is a powerful definition format to describe RESTful APIs. The Swagger specification creates a RESTful interface for easily developing and consuming an API by effectively mapping all the resources and operations associated with it. It's easy-to-learn, language agnostic, and both human and machine readable.

用Swagger进行API设计和开发

Swagger带来的好处：

- 自动生成一个优美的API文档，实时预览
- 让API文档总是与API定义同步更新，是一件非常有价值的事
- 自动生成服务器所需要的代码结构，并可以快速实现Mock API
- 自动生成不同语言的客户端SDK

Swagger相关的开源工具

- Swagger Editor：所见即所得编辑器，用于API设计，基于Swagger描述语言
- Swagger Codegen：自动生成前后端代码、SDK的工具，支持主流编程语言
- Swagger UI：通过Swagger描述自动生成文档
- 更多的开源项目实现的功能(<http://swagger.io/open-source-integrations/>)
 - 参数验证
 - 通过代码注释自动生成文档
 - 远程调用框架(grpc)
 - 开发测试等等

有了Swagger“契约”还能做什么？

API设计

(RESTful API design, without coding)

- <https://restlet.com/products/restlet-studio/>
- <http://stoptlight.io/>
- <https://gelato.io/>

API自动生成，运行环境托管，微服务等 (Easy deploy, secure and run APIs)

- <https://restlet.com/products/apispark/>
- AWS Gateway + Lambda
- DreamFactory

Swagger

API 自动测试

- [runscope](#)
- [postman](#)

API 市场

- <https://market.mashape.com/>
- <https://www.rapidapi.com/>

目标：

- Without coding
- 自动化
- 安全
- Serverless

Swagger和其它描述语言的对比



| | API-Blueprint | RAML | Swagger |
|---------------------|---------------|--------------|------------|
| Format | Markdown | YAML | JSON |
| Spec License | MIT | ASL 2.0 / TM | ASL 2.0 |
| Available at | GitHub | GitHub | GitHub |
| Sponsored by | Apiary | Mulesoft | Reverb |
| Current Version | 1A3 | 0.8 | 1.2 |
| Initial commit | April, 2013 | Sep, 2013 | July, 2011 |
| Commercial Offering | Yes | Yes | No |
| API Design Approach | Top-down | Top-down | Bottom-up |

技术社区的对比



| | API-Blueprint | RAML | Swagger |
|---|---------------|------|---------|
| Stackoverflow questions | 37 | 18 | 596 |
| Most Github Stars (on single project) | 613 | 478 | 1859 |
| Total Github projects (search on name) | 72 | 52 | 340 |
| Google search (name + "rest") | 807K | 64K | 5M |
| | | | |
| | | | |

参考资料: http://www.slideshare.net/SmartBear_Software/api-strat-2014metadataformatsshort

API管理的资源

用作API网关的开源软件



提供API管理产品的公司



其他可供参考的资源

参考链接:

<https://www.openapis.org/>

<http://swagger.io/>

<http://www.infoq.com/cn/articles/document-description-formats-web-apis?winzoom=1>

支持Swagger的云服务:

- Baishancloud: <http://juhe.baishancloud.com/gateway.html>
- SwaggerHub: <https://swaggerhub.com/>
- Akana: <https://www.akana.com/>
- Runscope: <https://www.runscope.com/>
- Restlet: <http://restlet.com/>
- Soapui: <https://www.soapui.org/>
- Axway: <http://axway.com/>
- Stoptlight: <http://stoptlight.io/>
- Postman: <https://www.getpostman.com/>
- Apigee: <https://apigee.com>

白山云聚合解决方案

CLN-X Cloud Linkage Network

云聚合为企业提供应用API加速、API防护、API网关和管理等技术服务，有效保障网站安全，提升访问速度，帮助企业快速构建数据开放平台，帮助数据发挥更大的价值。

云聚合产品是基于高性能可扩展的技术架构开发，也拥有大数据实时分析的安全技术，代表了当今最为前沿的技术趋势。

网站链接：<http://juhe.baishancloud.com/gateway.html>



THANKS!

