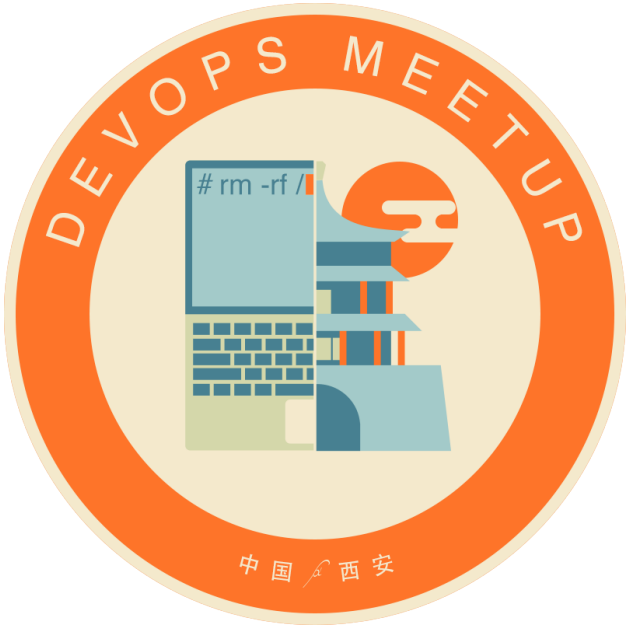
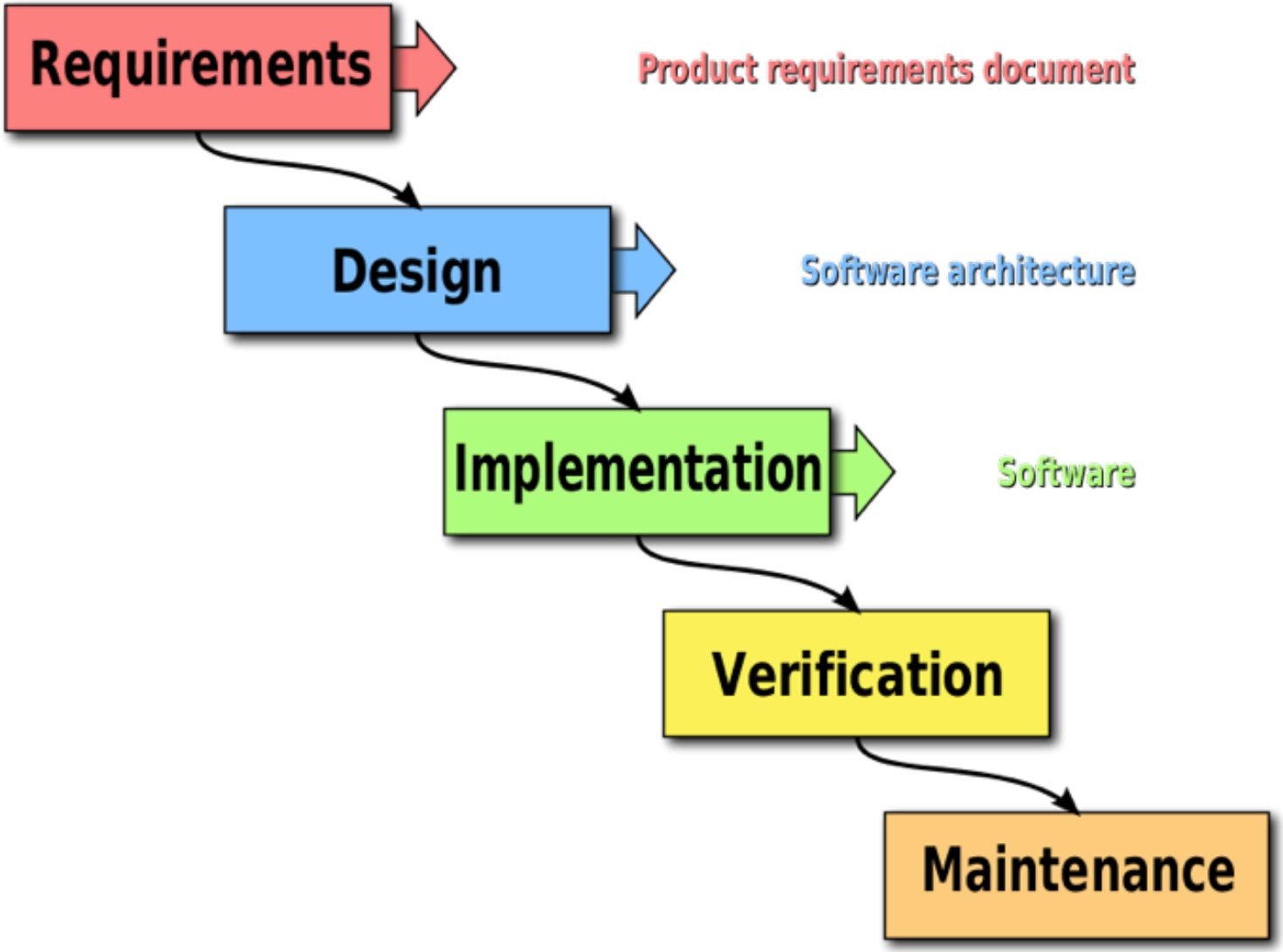
西安DevOps Meetup #5

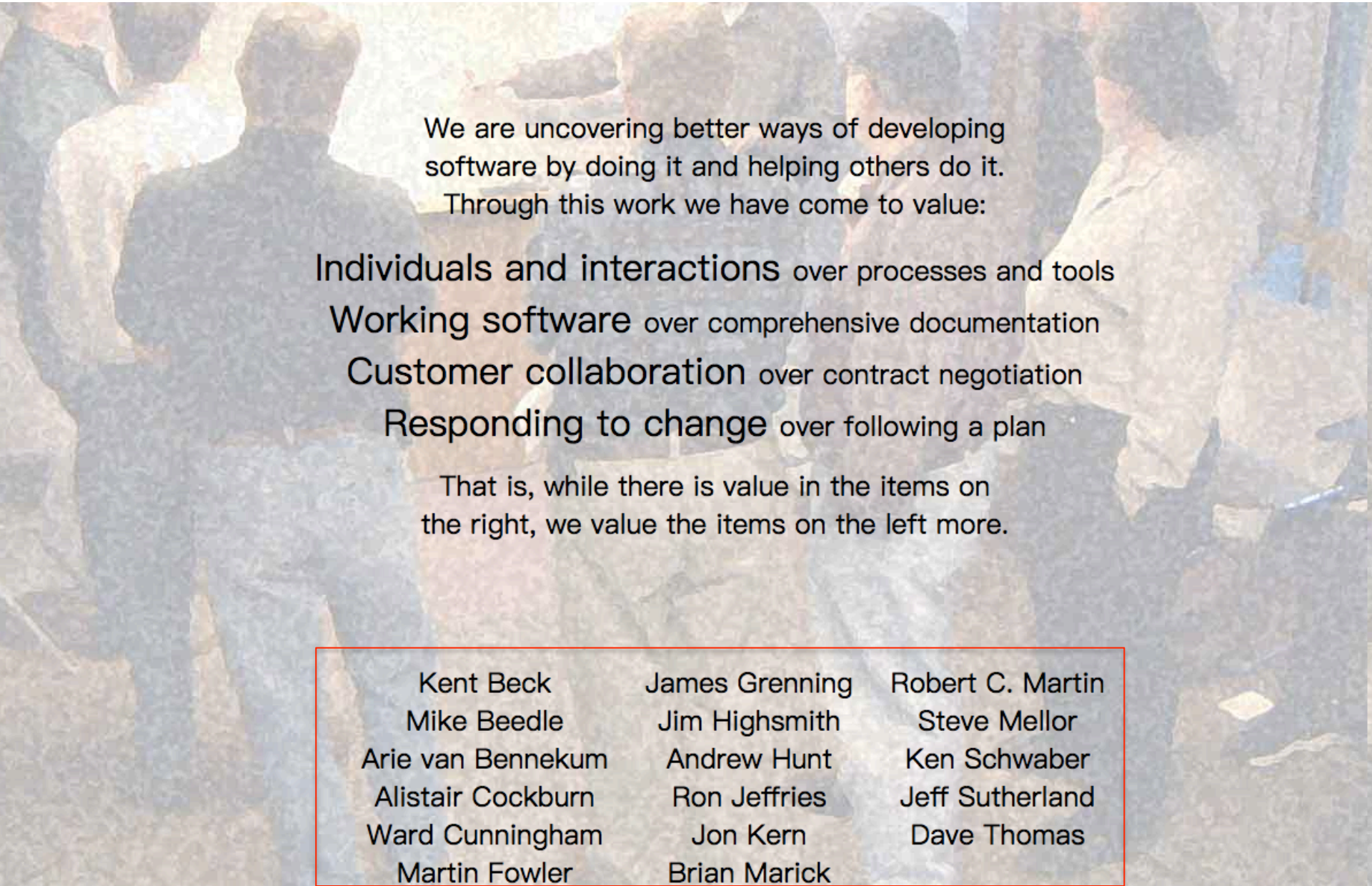


IT大咖说
知识分享平台

What’s DevOps



Agile Manifesto



We are uncovering better ways of developing
software by doing it and helping others do it.
Through this work we have come to value:

Individuals and interactions over processes and tools

Working software over comprehensive documentation

Customer collaboration over contract negotiation

Responding to change over following a plan

That is, while there is value in the items on
the right, we value the items on the left more.

Kent Beck	James Grenning	Robert C. Martin
Mike Beedle	Jim Highsmith	Steve Mellor
Arie van Bennekum	Andrew Hunt	Ken Schwaber
Alistair Cockburn	Ron Jeffries	Jeff Sutherland
Ward Cunningham	Jon Kern	Dave Thomas
Martin Fowler	Brian Marick	

快速交付 VS 稳定

DEV

扎心了，老铁呀

OPS

DevOps起源

2009: Patrick Debois <Agile infrastructure and operations: how infra-gile are you?>

2009: Velocity <10+ Deploys Per Day: Dev and Ops Cooperation at Flickr>

2009: First DevOpsDay

2009: Agile Admin: What's DevOps

2010: Jez Humble <Continuous Delivery>(持续交付)

https://theagileadmin.com/what-is-devops/

什么是DevOps

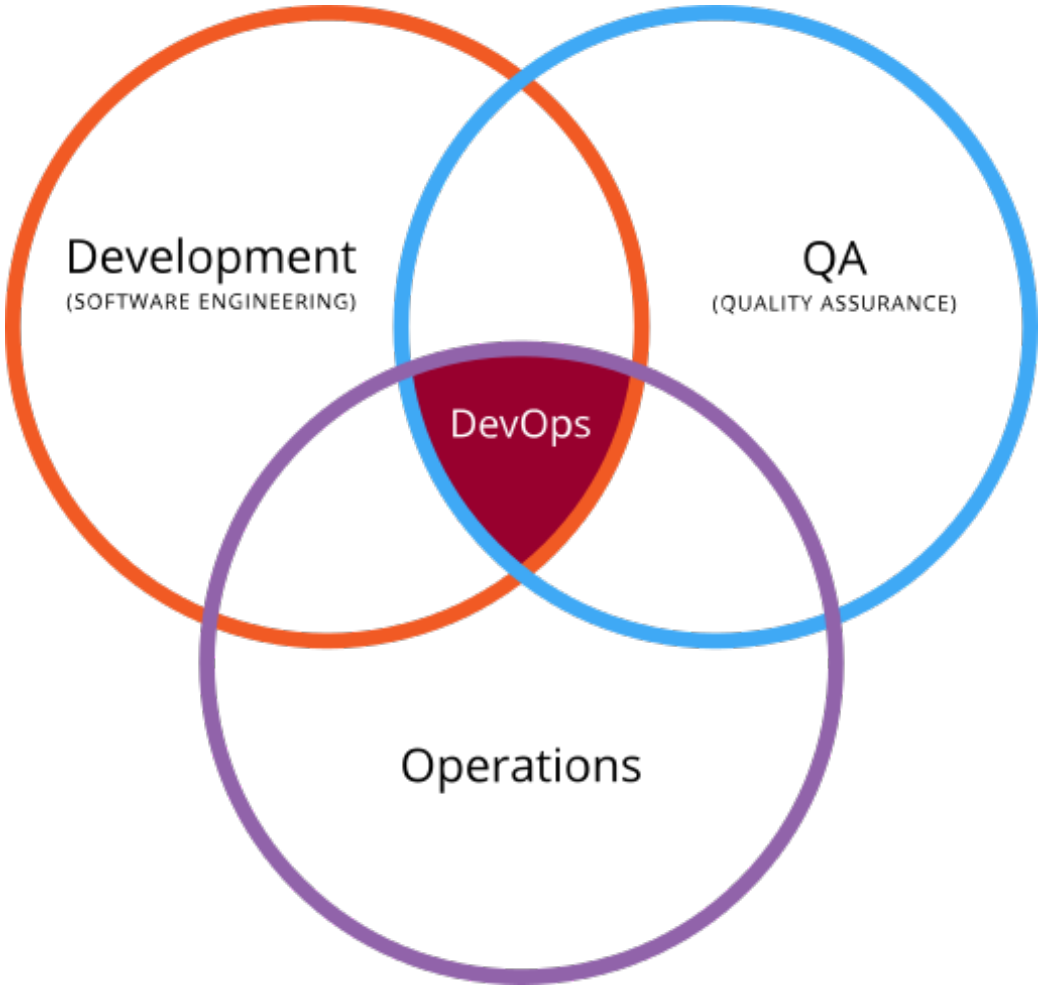
DevOps is the practice of operations and development engineers participating together in the entire service lifecycle, from design through the development process to production support.

https://theagileadmin.com/what-is-devops/

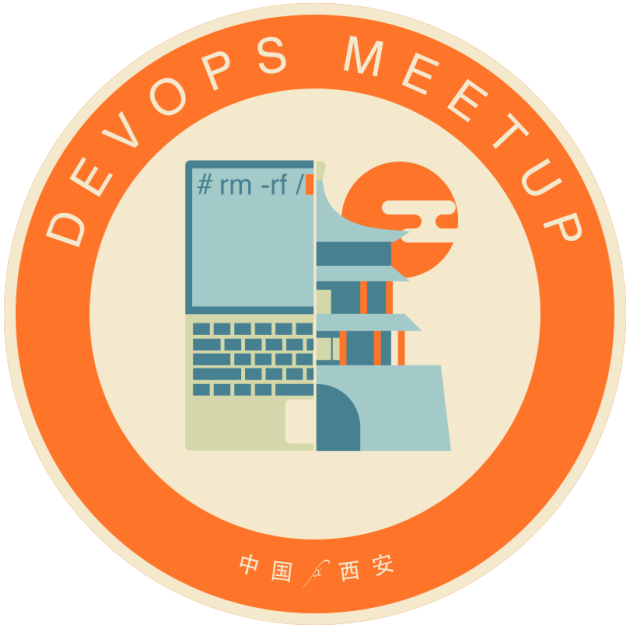
DevOps核心实践

- Culture
- Automation
- Lean
- Measurement
- Sharing

为什么叫DevOps


Development
(SOFTWARE ENGINEERING)
QA
(QUALITY ASSURANCE)
DevOps
Operations

西安DevOps Meetup



宗旨

- 促进西安DevOps领域的技术分享和学习
- 促进西安DevOps领域的技术分享和学习
- 促进西安DevOps领域的技术分享和学习
- 将西安技术社区活动通过直播分享到全国

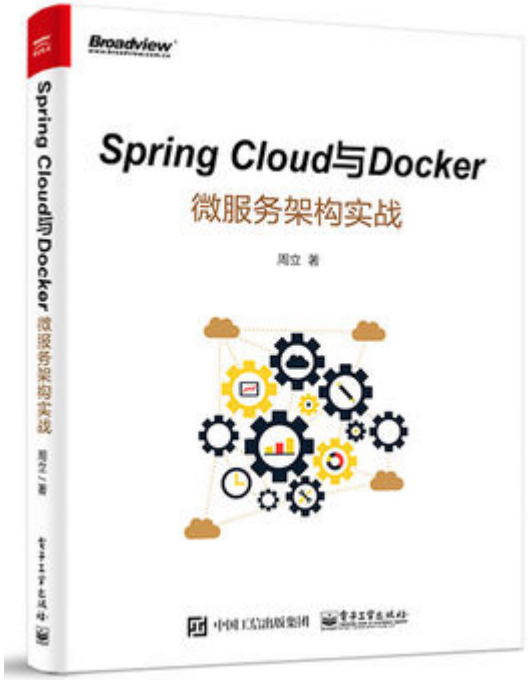
感谢金牌赞助商们

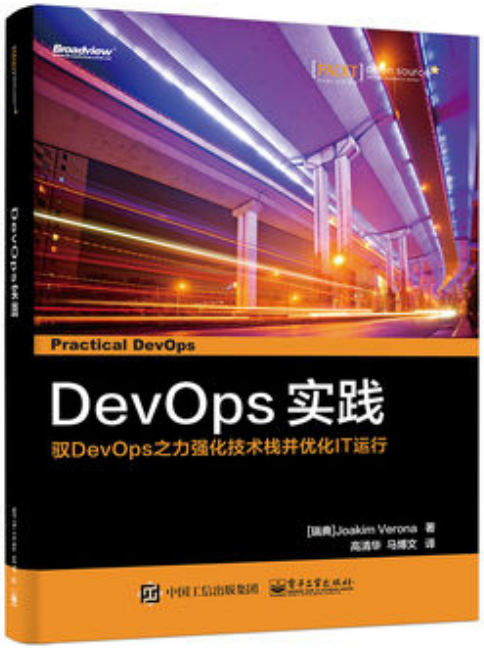
ThoughtWorks®

感谢博文视点

Broadview®

www.broadview.com.cn



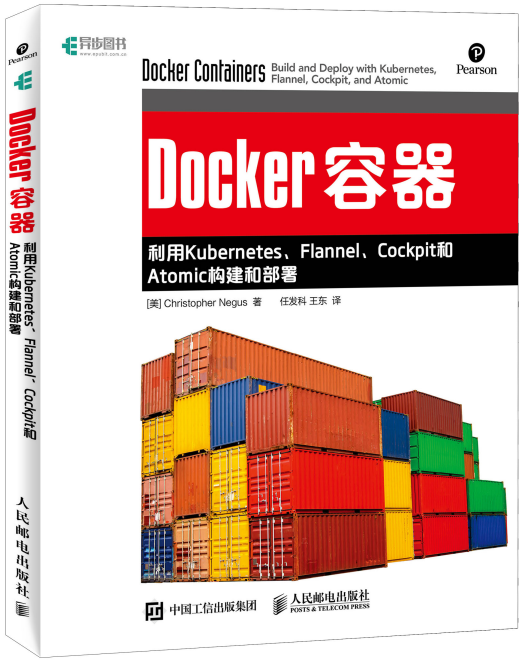


感谢异步社区

异步社区

人民邮电出版社

www.epubit.com.cn



感谢IT大咖说提供直播支持

IT大咖说
知识分享平台

欢迎赞助 DevOps Meetup

活动收集的信息在会后会删除

《12factor APP理解与实践》

By ThoughtWorks咨询师 张羽辰

《使用CodeDeploy实现自动化软件部署》

By AWS解决方案架构师 蒙维

IT大咖说
知识分享平台

《SRECon Asia 见闻》

By 马博文

西安DevOps Meetup #5

SRECon Asia 见闻

Agenda

- 什么是SRE
- SRECon介绍
- 专题分享
- SRECon 总结

Agenda

- 什么是SRE
- SRECon介绍
- 专题分享
- SRECon 总结

什么是SRE

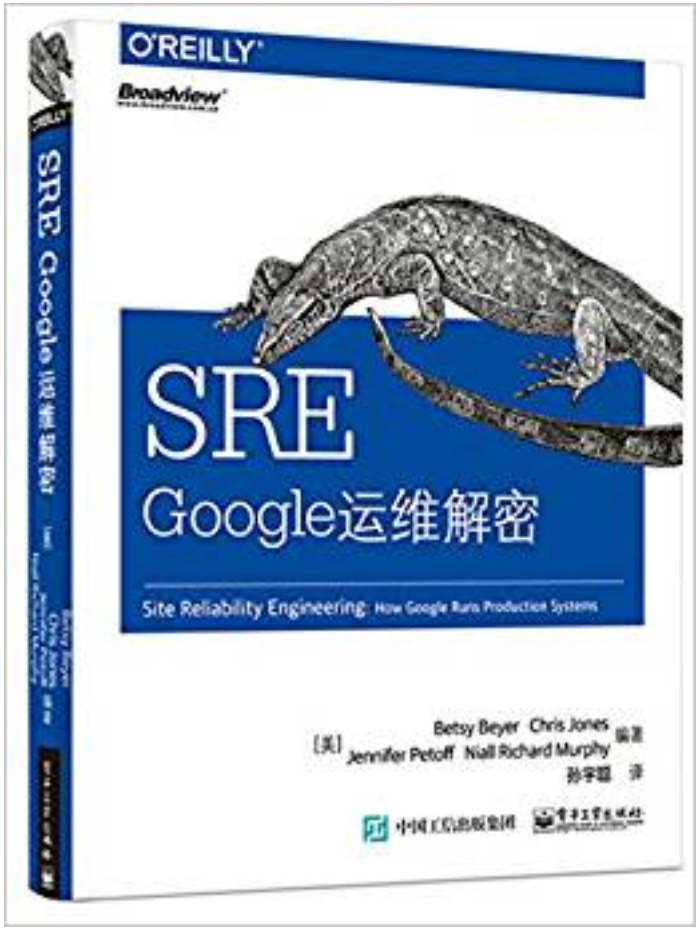
- 网站可靠性工程师(Site Reliability Engineer)
- 广泛的技能(e.g Google)
  - 50%-60%标准软件工程师
  - 满足80%-90%软件工程师要求，unix细节以及网络
- 用软件工程的思维去解决运维领域问题
- 负责可用性、性能、效率、监控、事务处理等

SRE方法论

- 关注研发工作(e.g Borg, BigTable)
- 保障服务SLO前提下最大化迭代速度
- 监控系统
- 应急事件处理
- 变更管理
- 需求预测和容量规划
- 资源部署
- 效率和性能

From： SRE Google运维解密

SRE Google 运维解密



Agenda

- 什么是SRE
- SRECon介绍
- 专题分享
- SRECon 总结

SRECon Asia

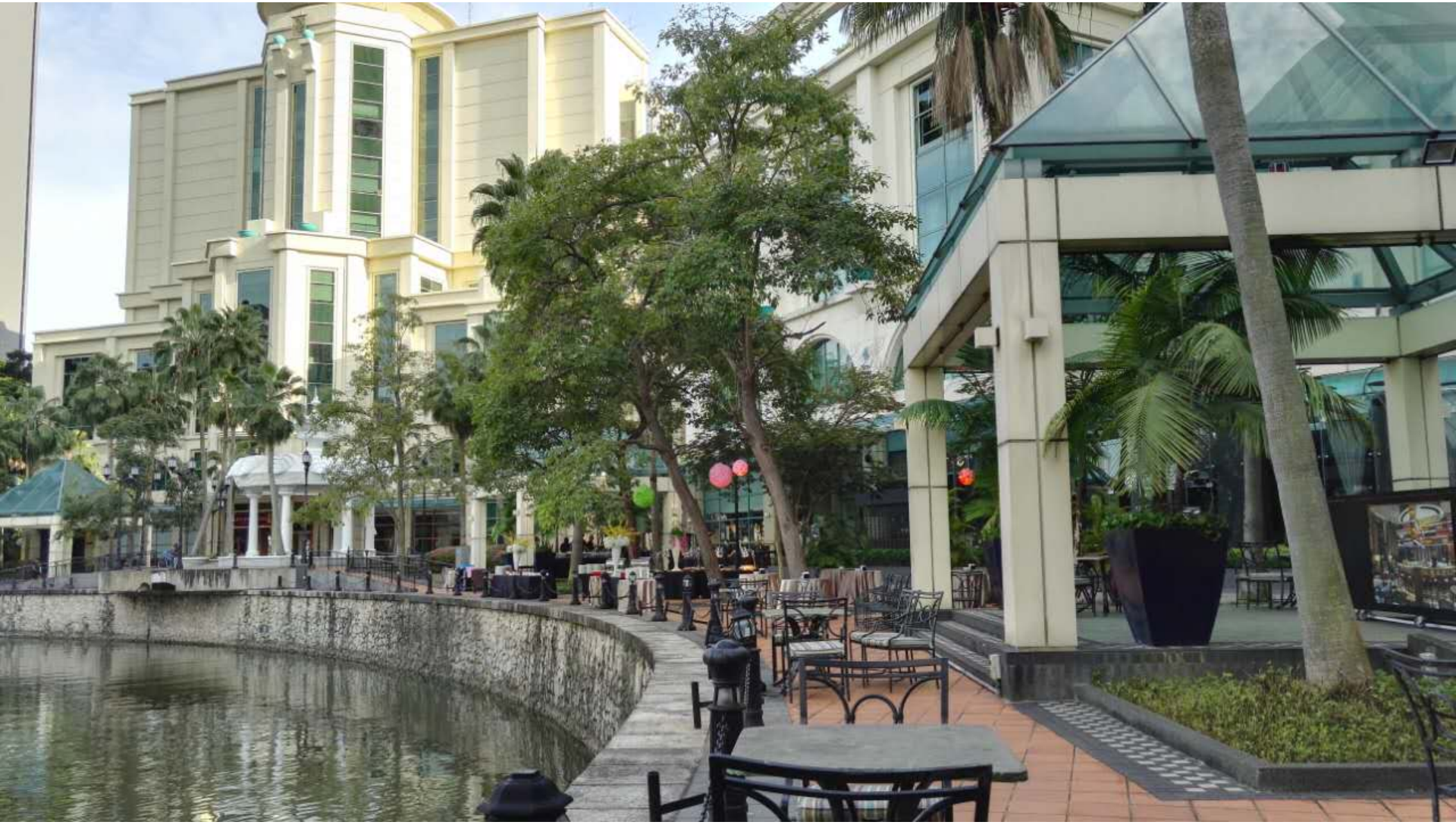
- 主办方: USENIX

- 主赞助商: Baidu/Facebook/Linkedin
- 到会人数: 250人左右
- 贡献话题的公司:
  - Google/FB/Linkedin/PayPal/CloudFlare/ Dropbox/Yahoo/Atlassian/REA Group等
  - Baidu/Alibaba/Didi/QiNiu/Tingyun/Tsinghua

SRECon Asia

- 时间: 5.22-5.24
- 地点: Grand Copthorne Waterfront Singapore(国敦河畔)
- 门票: 600$
  - 3天会议
  - 停不下来的茶歇
  - 两天自助午餐加一天的晚上酒会

感受下环境



感受下伙食

Agenda

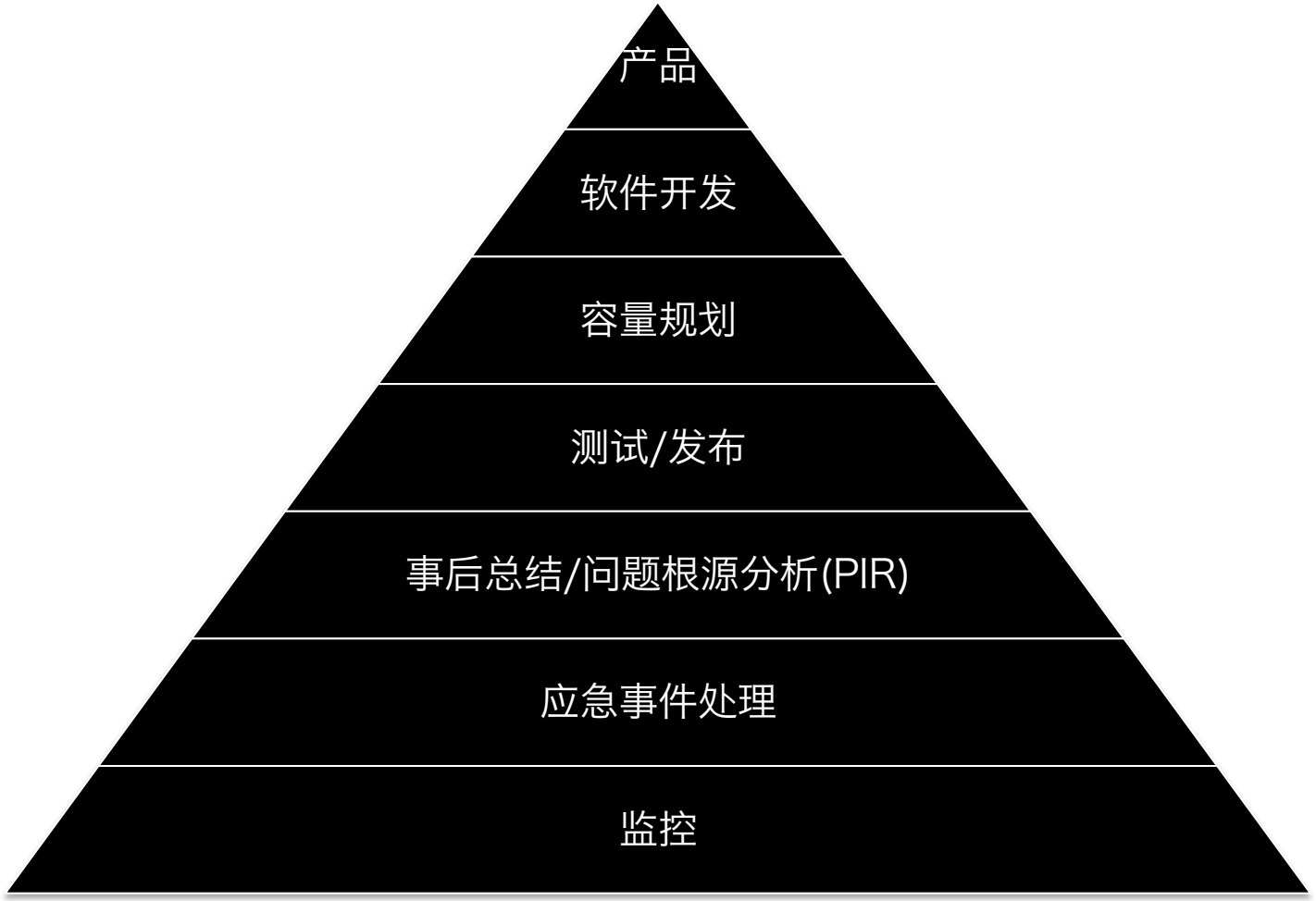
- 什么是SRE
- SRECon介绍
- 专题分享
- SRECon 总结

SRECon专题

- 监控和告警(Monitoring and Alerting)
- 服务生命周期(Service Lifecycle)
- 事故管理(Incident Management)
- 服务扩展(Service Expansion)
- 容量规划/性能调优
- SRE组织构建、文化等

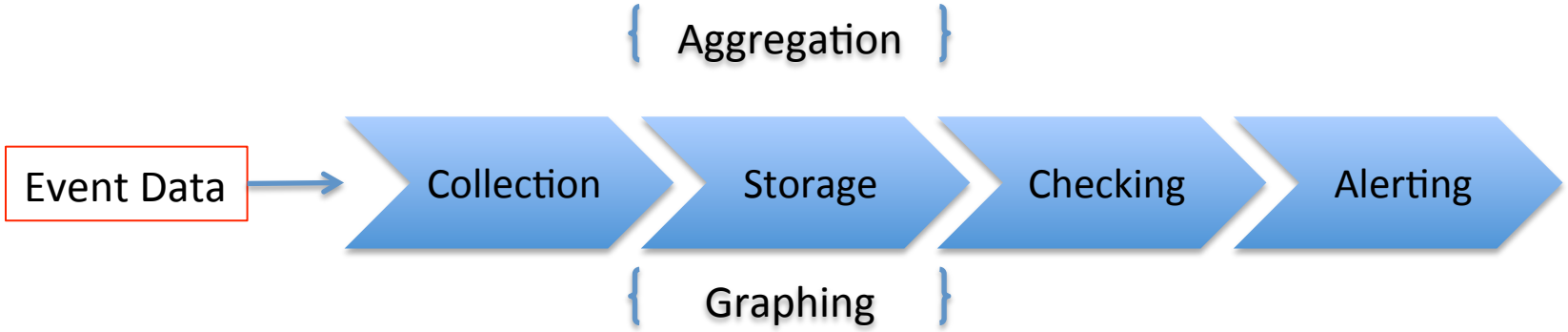
监控与告警

软件可靠性层级



《Google SRE》 p90

监控流水线



http://www.ustream.tv/recorded/54694069

Open-Falcon: A Distributed and High-Performance Monitoring System

https://www.usenix.org/conference/srecon17asia/program/presentation/ou

Open-Falcon: Motivation

- Zabbix
  - 水平扩展困难(>2000)
- OpenTSDB
  - 写性能，水平扩展好，query慢
- InfluxDB
  - Query性能，aggregator强大，水平扩展难

https://www.usenix.org/conference/srecon17asia/program/presentation/ou

Open-Falcon: Motivation

- Zabbix
  - 水平扩展困难(>2000)
- OpenTSDB
  - 写性能，水平扩展好，query慢
- InfluxDB
  - Query性能，aggregator强大，水平扩展难

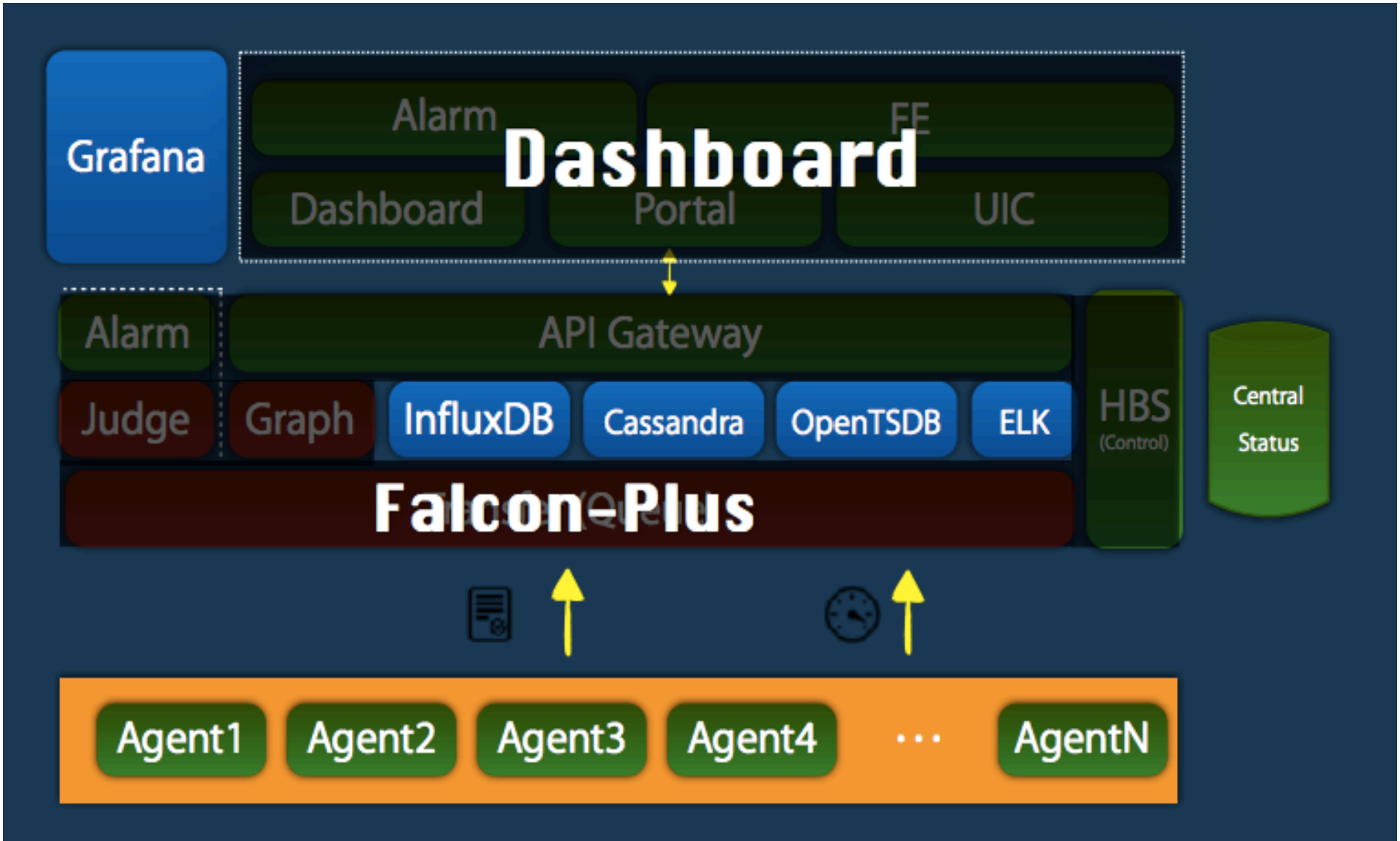
https://www.usenix.org/conference/srecon17asia/program/presentation/ou

Open-Falcon: Performance

- 容易水平扩展
- 每分钟百万transaction (query/ judge/ store/search)
-
- 轻松支持超过100,000主机
- RRA机制，查询1年历史数据，100+ metric，秒级响应时间
- 可以存储10+年的metric历史数据

https://www.usenix.org/conference/srecon17asia/program/presentation/ou

Open-Falcon: Architecture(别做动画)



Open-Falcon vs Prometheus

OPEN-FALCON	PROMETHEUS
Abundant APIs	Metrics API
Push Model: Auto Discovery	Pull Model: Manual configuration
Easy to scale out	Harder to scale out
simple alert management of own dashboard	Alertmanager offers **grouping, deduplication and silencing** functionality
Faster query performance of RRA	Slower, Recording rules
Simple shellscript as plugin	A bit learning curve to write exporter and collector
Limited expression	**PromQL**

https://www.usenix.org/conference/srecon17asia/program/presentation/ou

《Talking to an OpenStack Cluster in Plain English》

by Xu Wei From Tsinghua

https://www.usenix.org/conference/srecon17asia/program/presentation/xu

Lab Infrastructures

- ~ 300 servers
  200* 2U servers
  100* Facebook OCP

- OpenStack (125 nodes)
  Ceph (60 nodes)
  Hadoop
  Spark

https://www.usenix.org/conference/srecon17asia/program/presentation/xu

问题

- 运维OpenStack，修复问题所需要的知识复杂，操作过多

- 这些知识很难Transfer

https://www.usenix.org/conference/srecon17asia/program/presentation/xu

解决思路

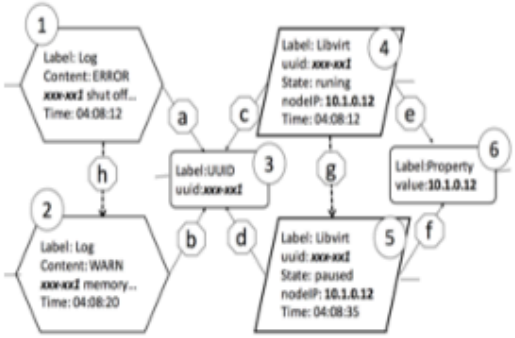
- 使用自然语言查询系统状态
  - 好于CLI和Regex
- 使用最基本的规则自动发现系统知识
  - SOSG(System Operation State Graph)
  - 将特定系统的查询转化为图遍历
  - 异常检测发现隐藏的问题。

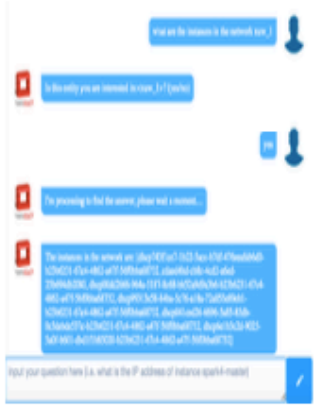
https://www.usenix.org/conference/srecon17asia/program/presentation/xu

构建知识图谱

Raw States

Logs



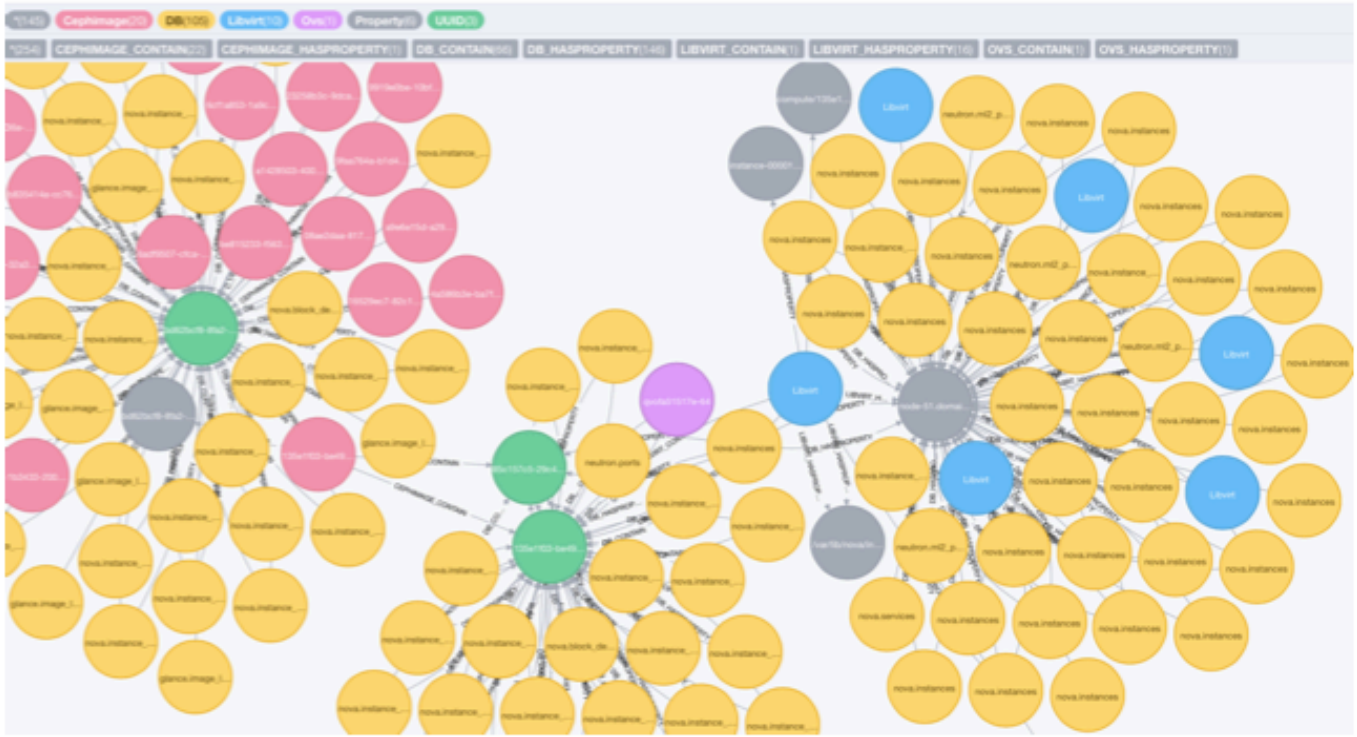


Automatically generated
“Knowledge graph” for OpenStack

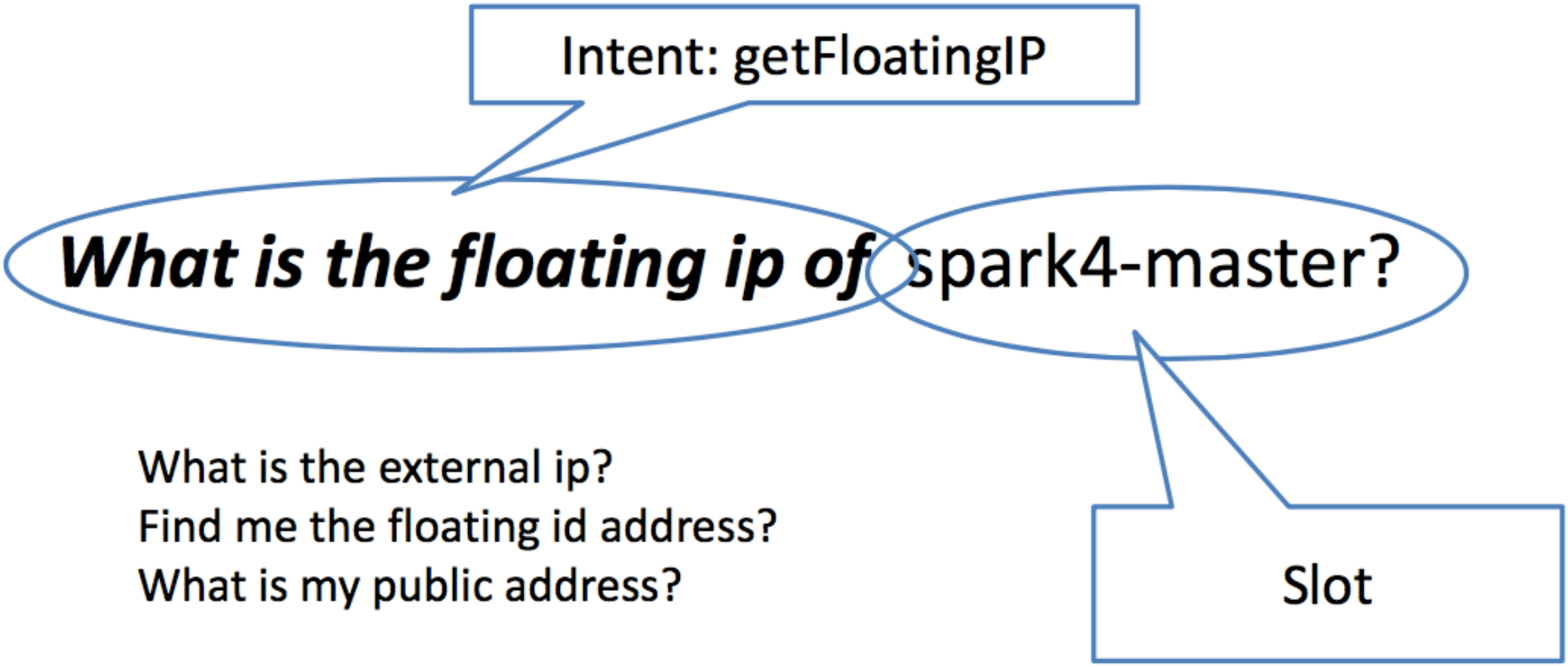
https://www.usenix.org/conference/srecon17asia/program/presentation/xu

OpenStack Data

- 3-day operation data, about 40 GB
- Graph size: 43.3 million vertices, 56.6 million edges

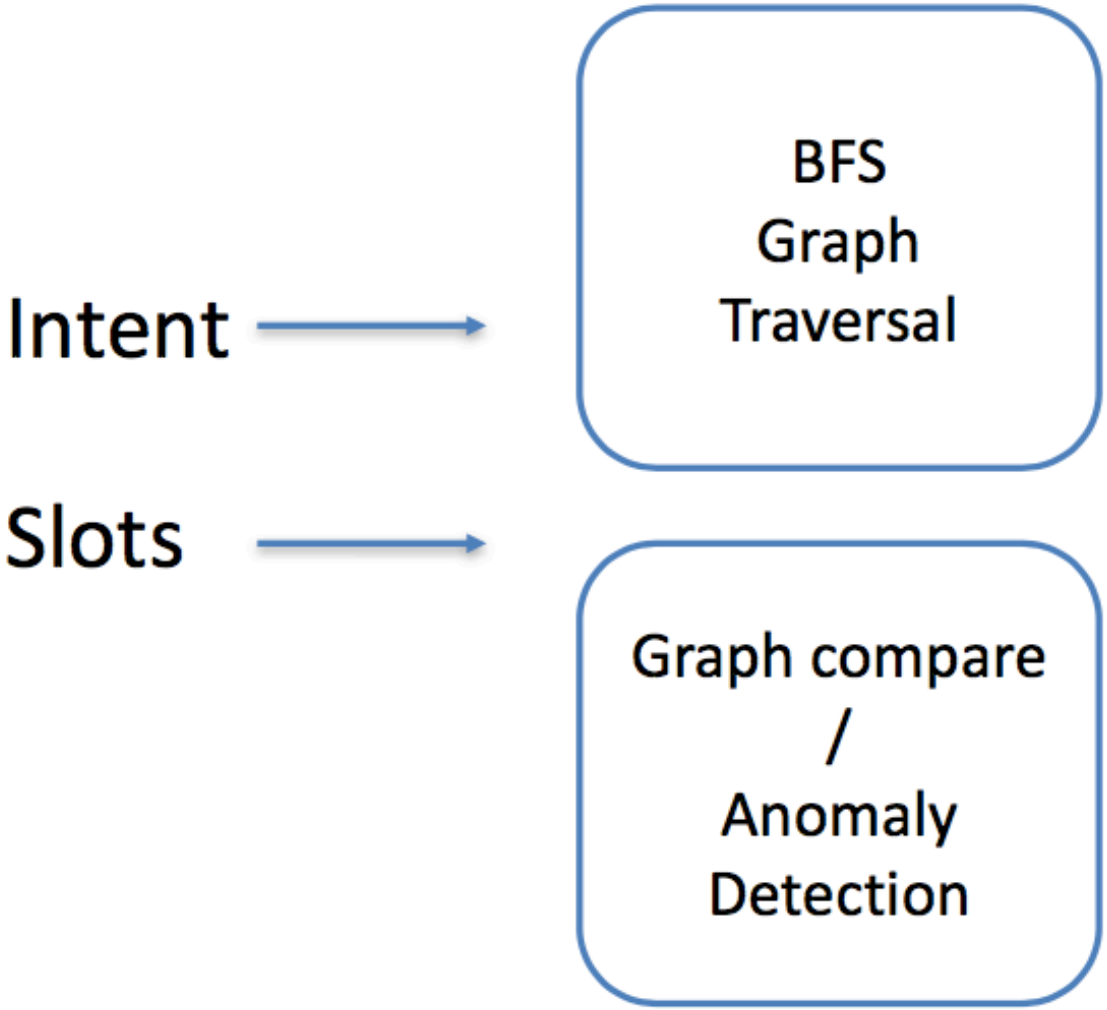


支持自然语言



https://www.usenix.org/conference/srecon17asia/program/presentation/xu

在状态图上执行自然语言



https://www.usenix.org/conference/srecon17asia/program/presentation/xu

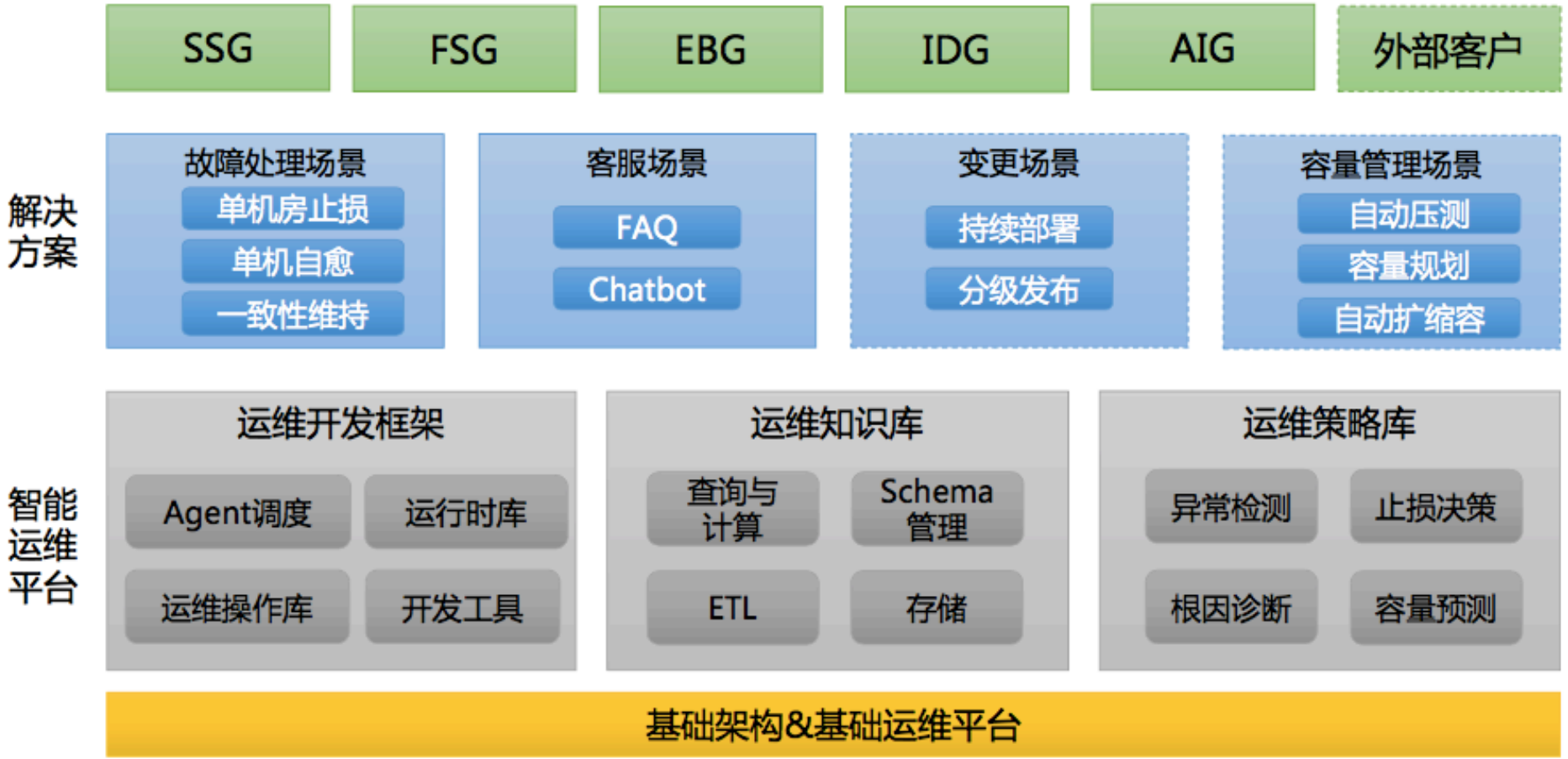
结论

- Automation -> Intelligence
- Regular expressions -> natural language understanding
- Scripting -> training models
- Manually examine logs -> systematic query and traversal
- Time to adopt AIs in dev-ops!

https://www.usenix.org/conference/srecon17asia/program/presentation/xu

AIOps

百度AIOps框架



https://www.usenix.org/conference/srecon17asia/program/presentation/qu

服务生命周期

服务生命周期

- Diistributed Consensus Algorithms
  - Paxos算法
- Reliable Launches at Scale
  - Launch Check List
- Managing Changes Seamlessly on Yahoo's Hadoop Infrastructure Servers
  - 45,000 nodes managed by Chef

https://www.usenix.org/conference/srecon17asia/program/presentation/kirsch
https://www.usenix.org/conference/srecon17asia/program/presentation/vadrevu
https://www.usenix.org/conference/srecon17asia/program/presentation/nolan

Reliable Launches at Scale

- 1. 架构
- 2. 容量
- 3. 可靠性
- 4. 监控
- 5. 自动化程度
- 6. 增长趋势
- 7. 第三方（google内部）服务是否准备好
- 8. 上线

https://www.usenix.org/conference/srecon17asia/program/presentation/kirsch

Managing Server Secrets at Scale with a Vaultless Password Manager

https://www.usenix.org/conference/srecon17asia/program/presentation/korgachin

密码/Key管理工具

- Password/Key Management Tool(some vault)
  - OnePassword/LastPassword
  - KeepassX
  - RatticDB
  - Hashicop Vault
  - AWS KMS

https://www.usenix.org/conference/srecon17asia/program/presentation/kirsch

Not In Cloud

- Key/Credentials随着服务器增多而增多
- 在配置管理工具中保存Secrets
  - 启动配置管理工具需要key/pair etc
  - 无法scale key(每个服务器密码不能相同)
  - Key Rotation
- 保存在服务器上(服务器启动时生成)
  - root password
  - 磁盘加密比较困难
  - 无状态/磁盘的服务器无法存储(e.g image)

https://www.usenix.org/conference/srecon17asia/program/presentation/kirsch

UEFI

- BIOS 2.0
- Standard pre-OS environment
- Extensible (you can write your own apps)
- Supported by most major Oses
- Provides many advanced features
  - UEFI variables

https://www.usenix.org/conference/srecon17asia/program/presentation/kirsch

UEFI Variables

- Backed by flash memory on platform firmware chip
- Can store standard and custom (OEM/user) data
- Can be accessed after OS kernel booted
- Have built-in support in Linux

https://www.usenix.org/conference/srecon17asia/program/presentation/kirsch

KDF

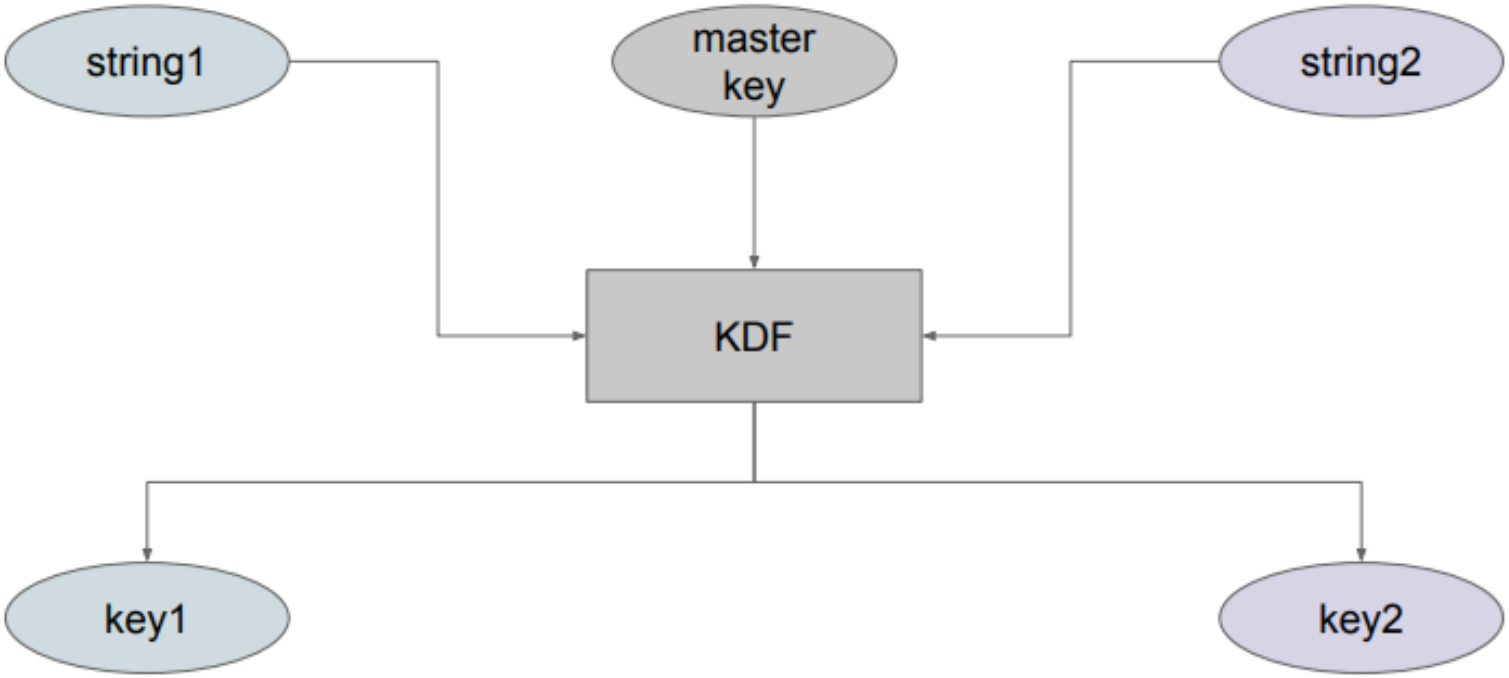
- In cryptography, a key derivation function (KDF) derives one or more secret keys from a secret value such as a master key, a password, or a passphrase using a pseudo-random function.

- DK=KDF(Key, Salt, Iterations)

https://en.wikipedia.org/wiki/Key_derivation_function

For key/password

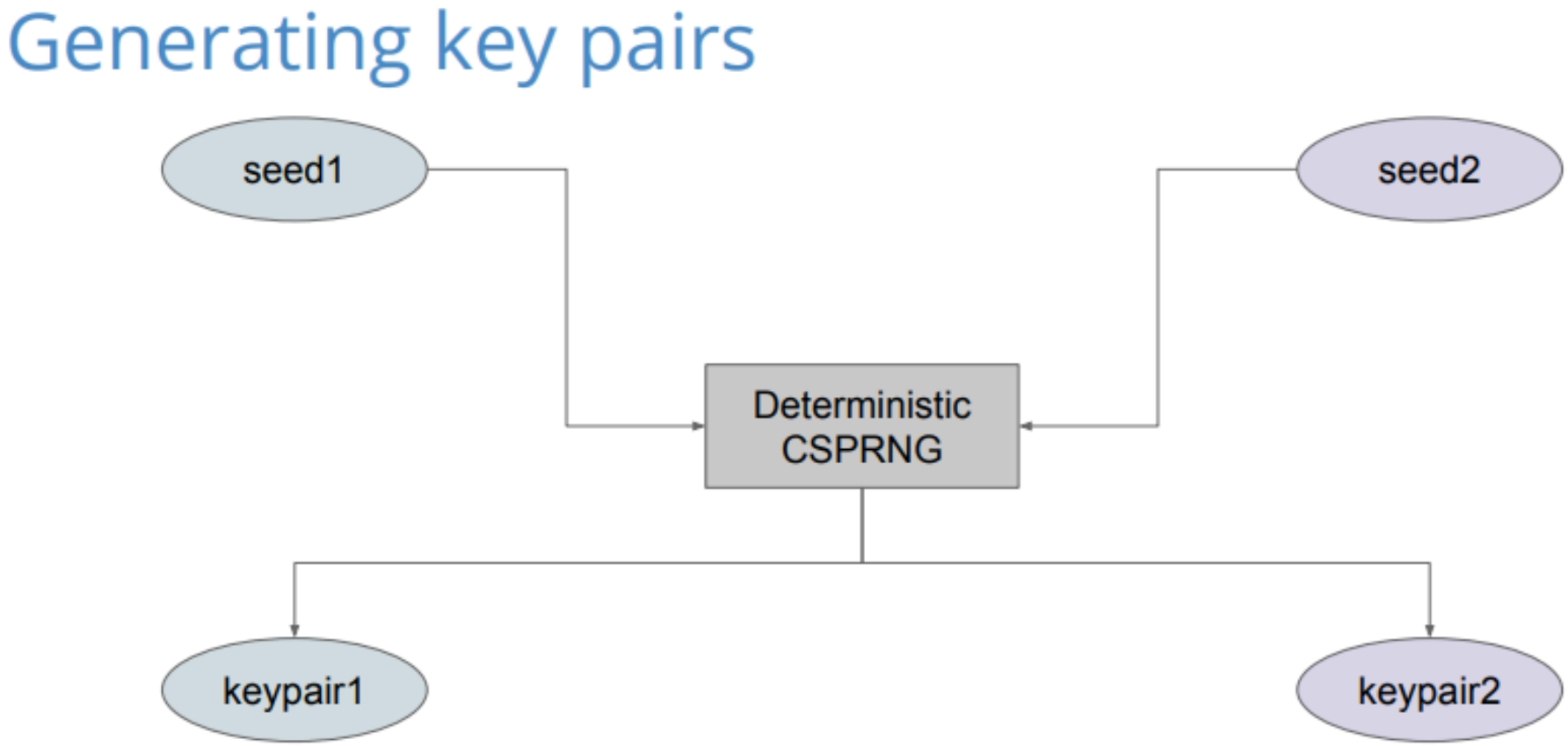
Key derivation functions



CLOUDFLARE

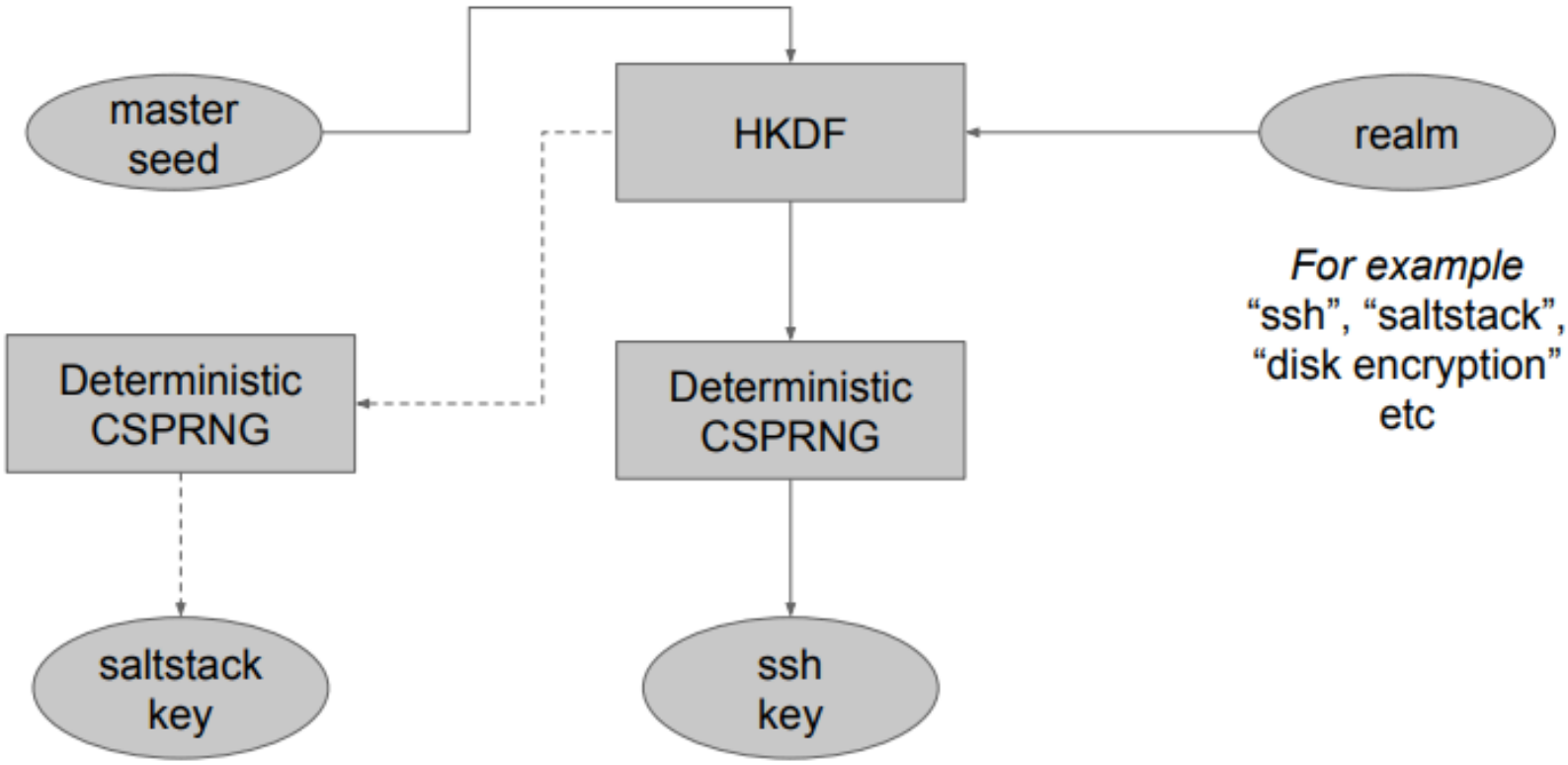
https://en.wikipedia.org/wiki/Key_derivation_function

Deterministic CSPRNG(伪随机数生成器) for key pairs



https://en.wikipedia.org/wiki/Key_derivation_function

Introducing gokey tool



https://www.usenix.org/conference/srecon17asia/program/presentation/korgachin

Key management

- Provisioning process ensures a master seed is generated and stored in UEFI on first boot
- Startup scripts “recover” (derive from master seed) configuration management credential (key)
- Configuration management “recovers” all other keys

https://github.com/cloudflare/gokey

事故管理

事故管理的一些挑战

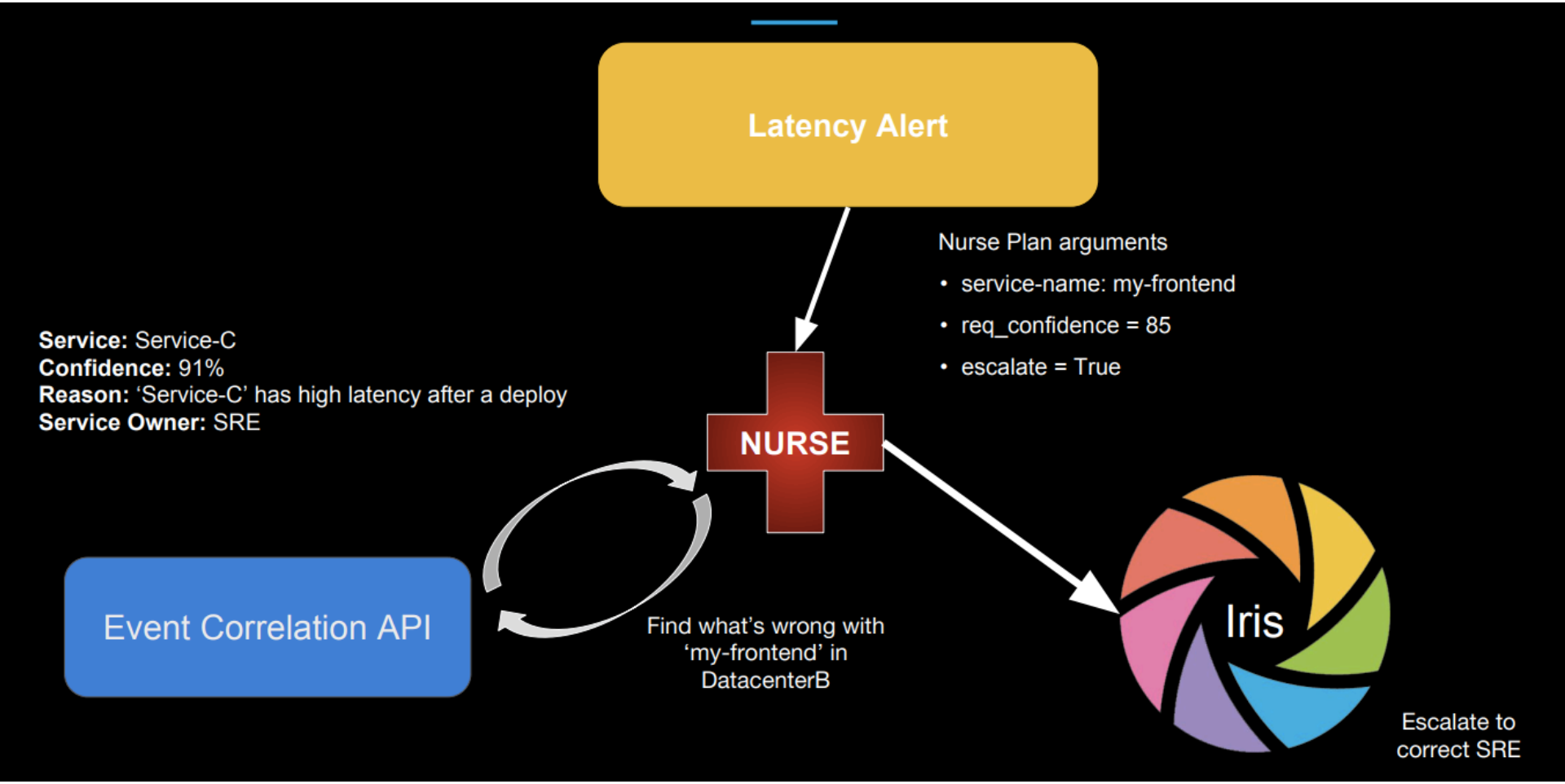
- 如何达成更短的MTTR
- 很多事故的处理比较简单，如重启等，如何自动处理这些事故
- false alarms如何减少
- 报警如何给出正确信息，快速定位问题

https://github.com/cloudflare/gokey

事故管理

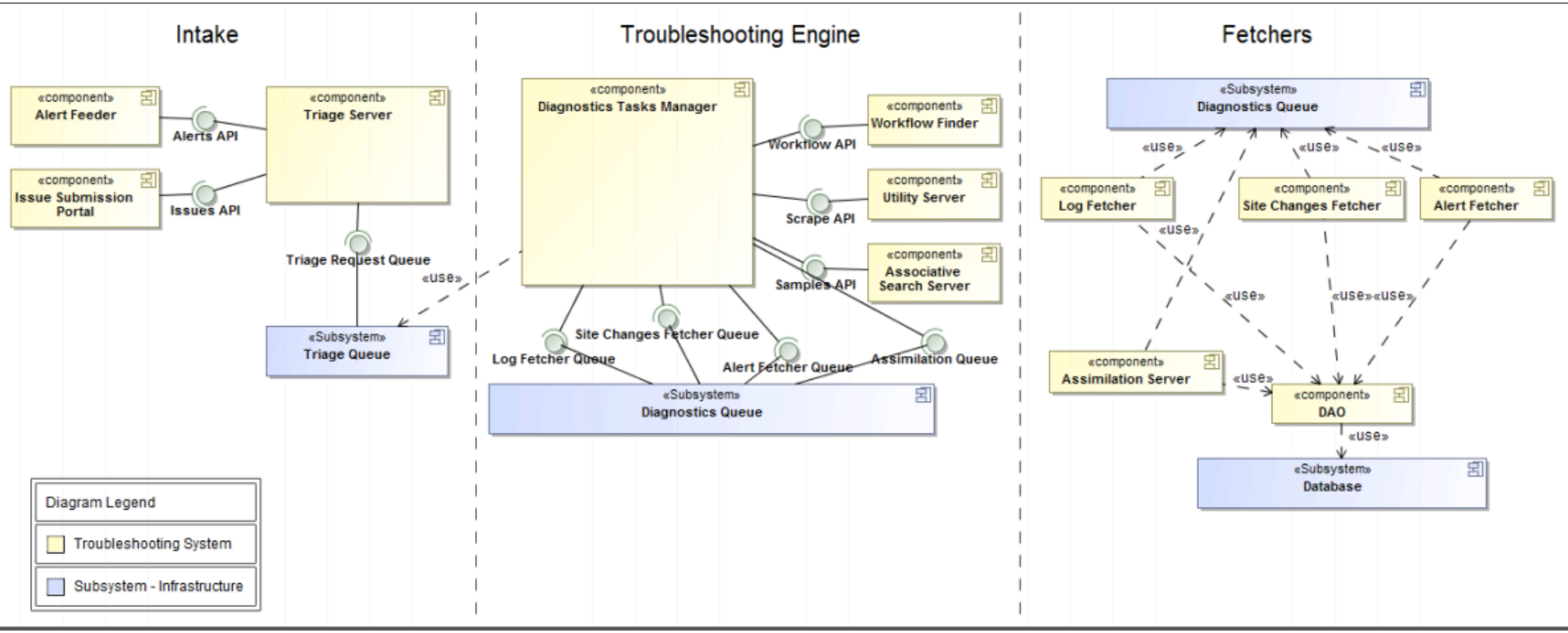
- Event Correlation: A Fresh Approach towards Reducing MTTR
- Automated Troubleshooting of Live Site Issues
- Accept Partial Failures, Minimize Service Loss
- Azure SREBot: More than a Chatbot—an Intelligent Bot to Crush Mitigation Time

Event Correlation: A Fresh Appro towards Reducing MTTR



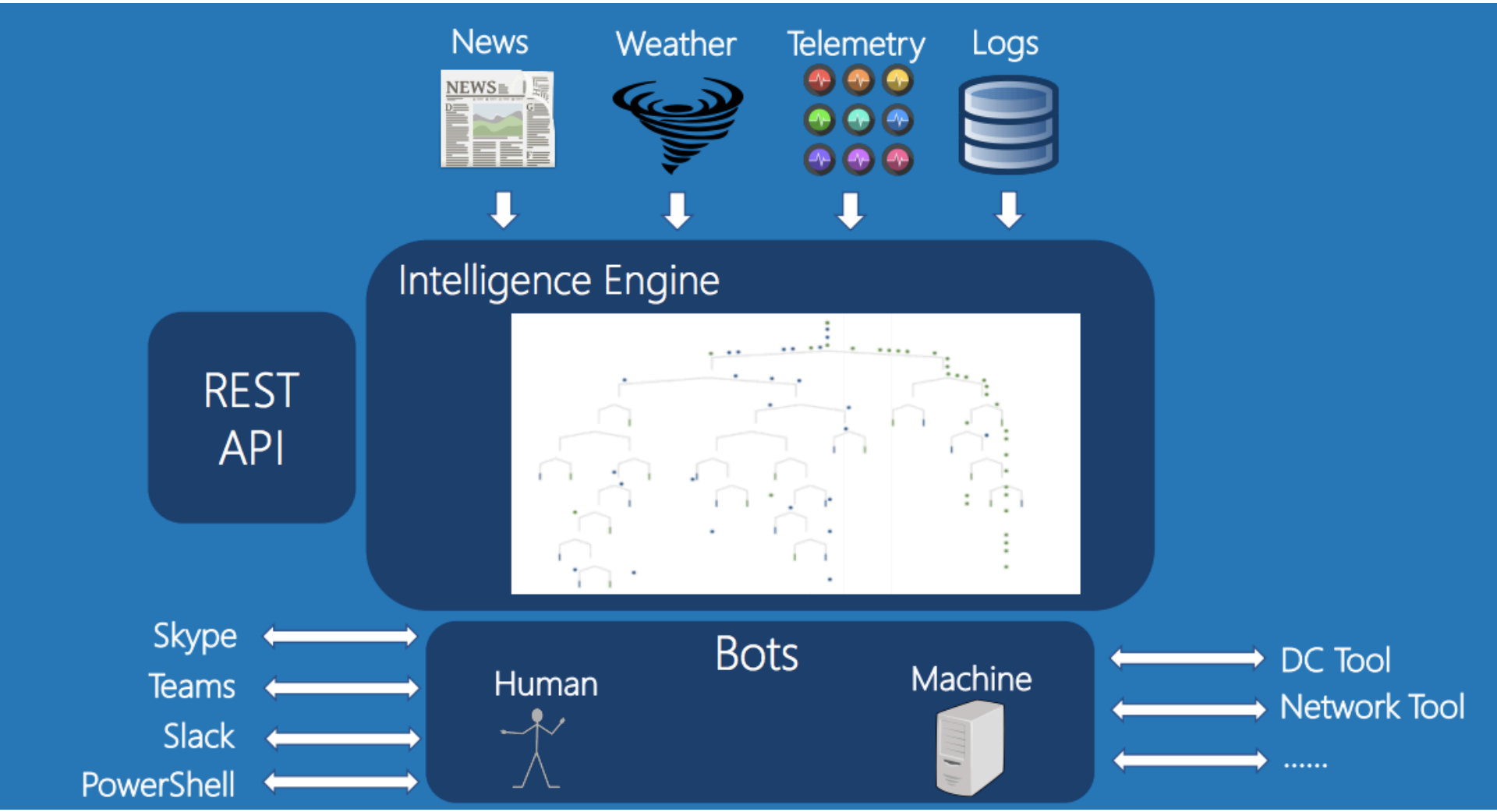
https://www.usenix.org/conference/srecon17asia/program/presentation/rajan

Automated Troubleshooting of Live Site Issues



https://www.usenix.org/conference/srecon17asia/program/presentation/xu

Azure SREBot: More than a Chatbot– Intelligent Bot to Crush Mitigation Time



https://www.usenix.org/conference/srecon17asia/program/presentation/guimaraes

服务扩展

服务扩展

- Small, Cheap, and Effective Testing for Production Engineers
- Merou: A Decentralized, Audited Authorization Service
- Shame on facebook and dropbox

容量规划/性能调优

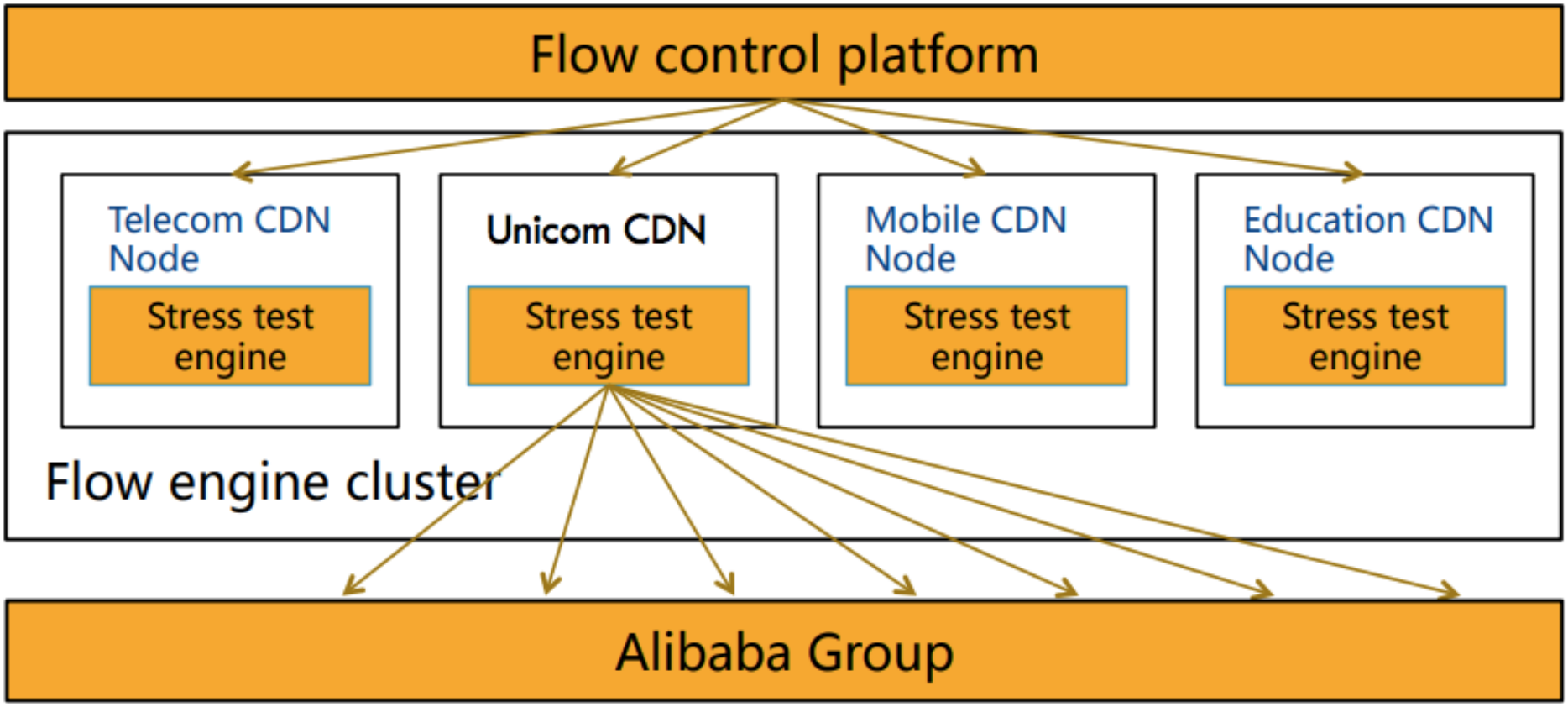
Capacity Planning and Flow Control

- 容量估算：单机压测
  - 模拟: ab/jmeter/gatling
  - 复制：复制生产环境流量
  - 重定向
  - 负载均衡: weight

https://www.usenix.org/sites/default/files/conference/protected-files/srecon17asia_slides_zhang.pdf

Capacity Planning and Flow Control

- 全链路压测: (10m/s)



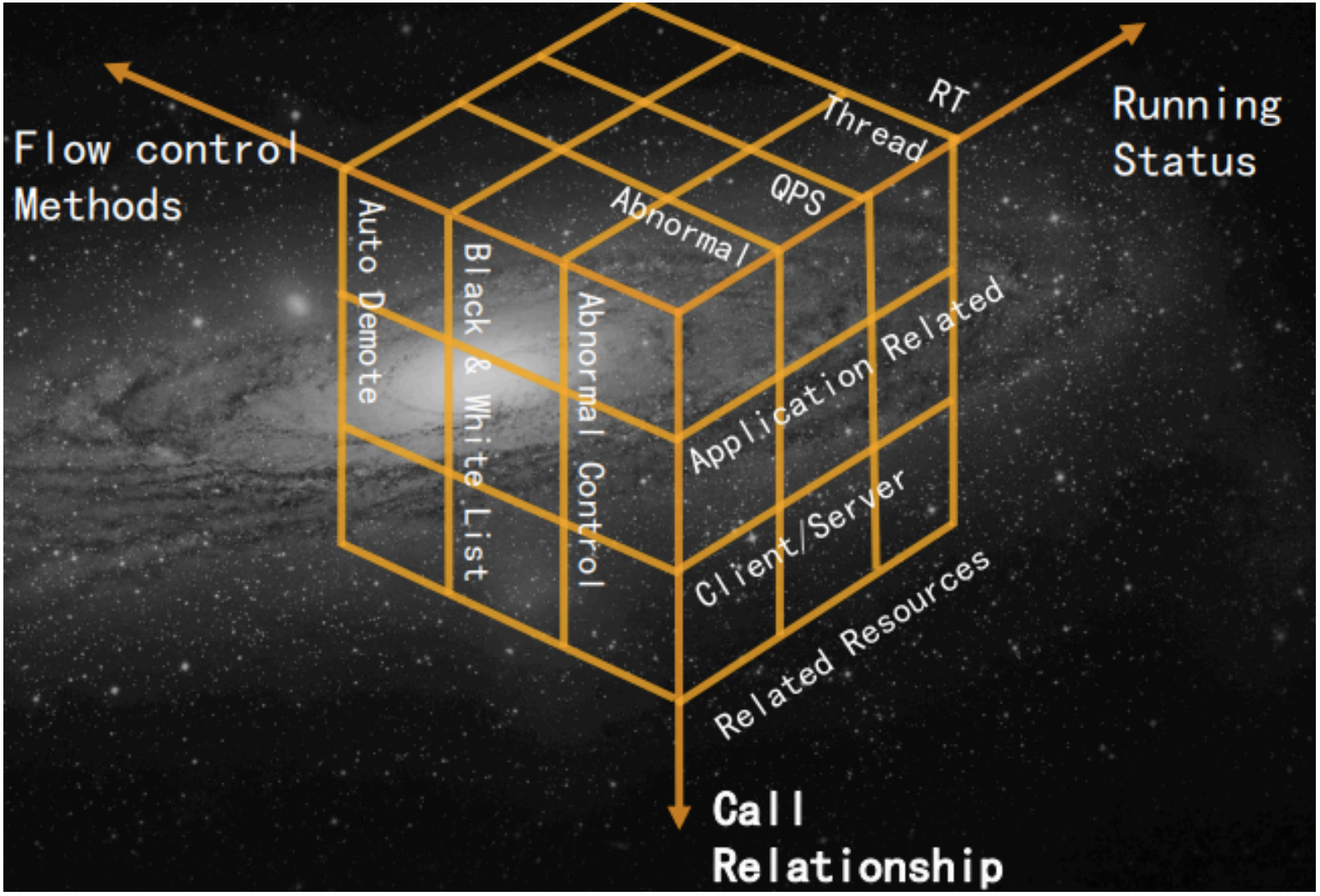
https://www.usenix.org/sites/default/files/conference/protected-files/srecon17asia_slides_zhang.pdf

Why Flow Control

- 队列堆积
  - 服务器性能降低
  - 响应时间增加
  - 影响应用以及用户体验
- 雪崩效应
- 需要限制过载的流量

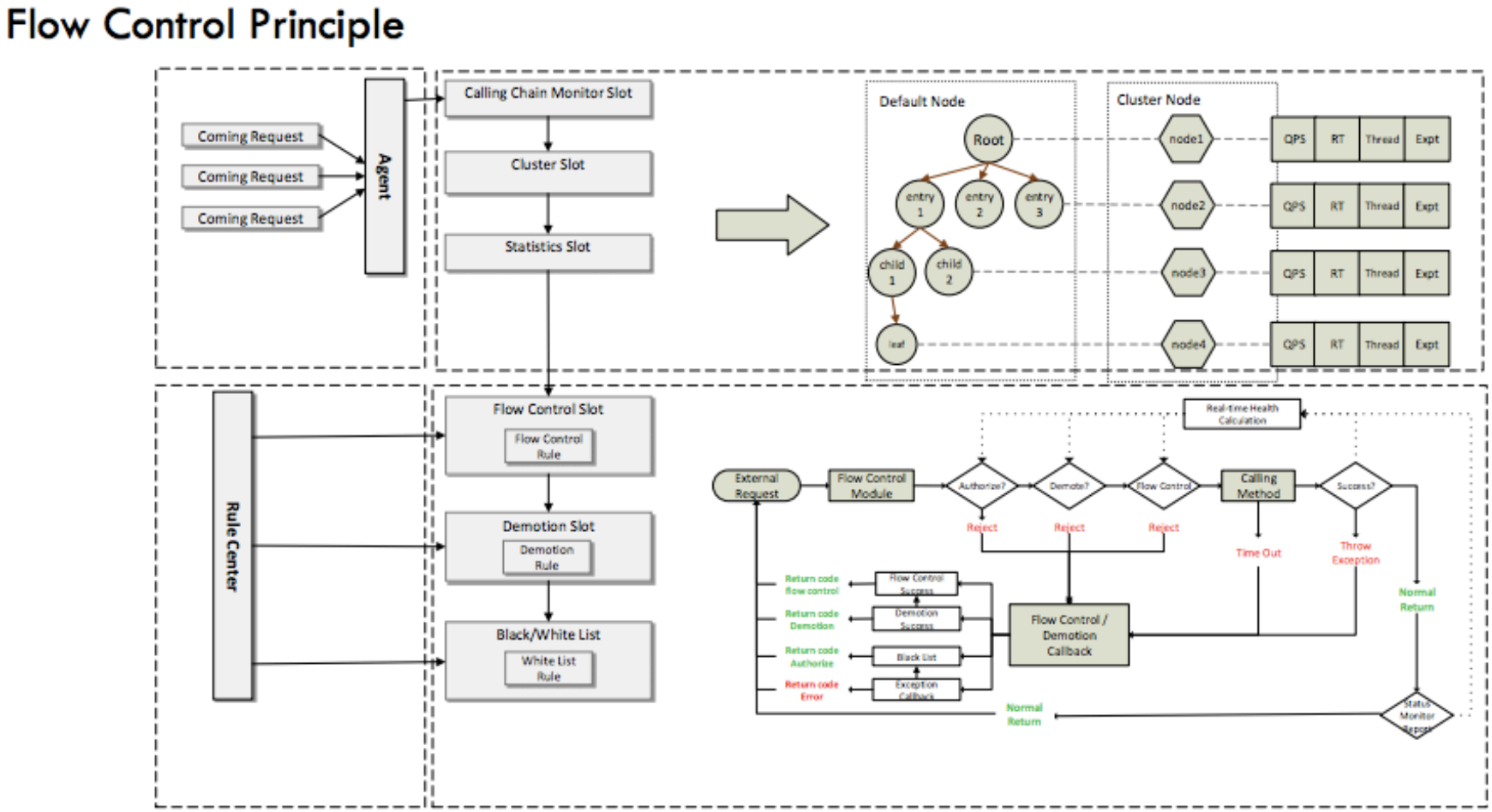
https://www.usenix.org/sites/default/files/conference/protected-files/srecon17asia_slides_zhang.pdf

Flow Control Method



https://www.usenix.org/sites/default/files/conference/protected-files/srecon17asia_slides_zhang.pdf

Flow Control Principle



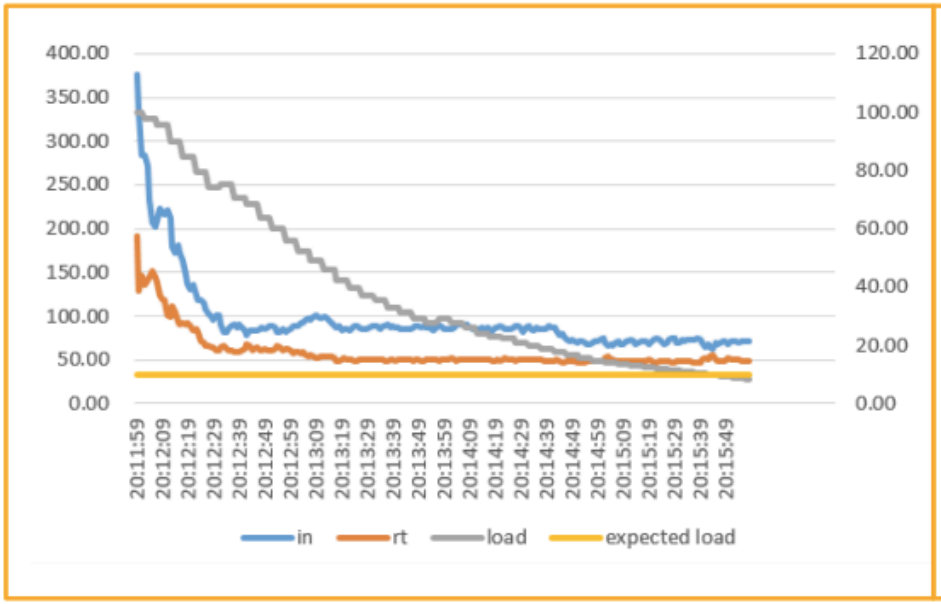
https://www.usenix.org/sites/default/files/conference/protected-files/srecon17asia_slides_zhang.pdf

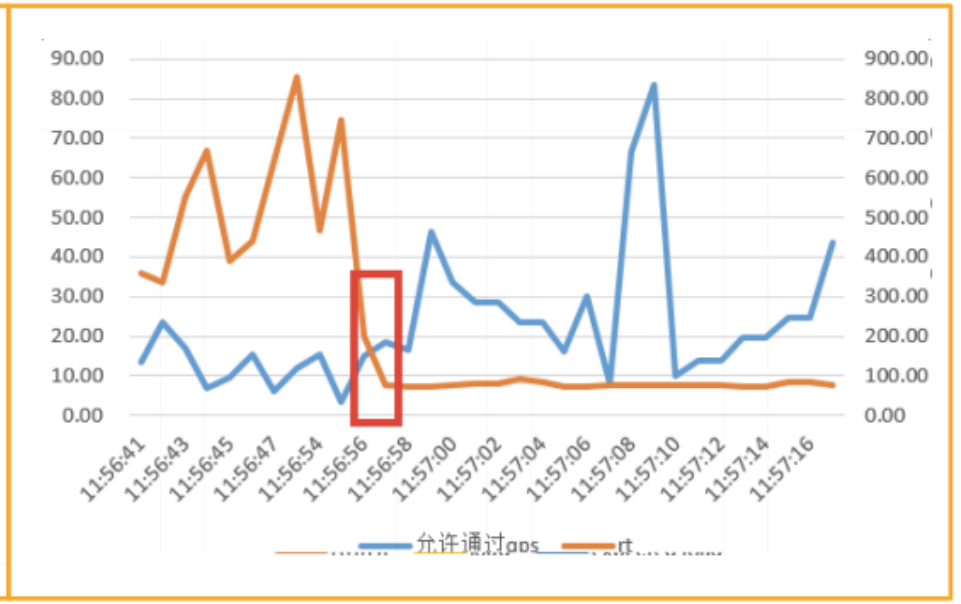
And a Formula!

- 计算原则:
  - Entrance Size= volume ＊ RT(response time)
  - Requests = constants ＊ LOAD ＊ RT
- 流量控制原则:
  - 系统超载则限制volume
  - 负载正常则去掉限制
- 使用动态阈值控制

https://www.usenix.org/sites/default/files/conference/protected-files/srecon17asia_slides_zhang.pdf

结果





https://www.usenix.org/sites/default/files/conference/protected-files/srecon17asia_slides_zhang.pdf

服务扩展

- Small, Cheap, and Effective Testing for Production Engineers
- Merou: A Decentralized, Audited Authorization Service
- Shame on facebook and dropbox

SRE组织构建、文化等

- LinkedIn SRE: From Inception to Global Scale
  - linkedin SRE的发展
  - 如何将这种能力transfer到印度
- Operationalizing DevOps Teaching
  - REA如何培养新人成为DevOps
- Scaling Reliability at Dropbox
  - PPT不错

https://www.usenix.org/conference/srecon17asia/program/presentation/connelly
https://www.usenix.org/conference/srecon17asia/program/presentation/contad
https://www.usenix.org/conference/srecon17asia/program/presentation/khalsa

其他: gRPC

- 高性能、开源通用RPC框架(http2/protobuf)
- 客户端负载均衡、流量控制
- Streaming等

https://www.usenix.org/conference/srecon17asia/program/presentation/khalsa

Agenda

- 什么是SRE
- SRECon介绍
- 专题分享
- SRECon 总结

总结

- 参会人数少，交流效果比较好
- 了解到不同的公司，e.g，a9，彭博社
- 大部分的话题都有可学习的地方
- 数据流水线+大数据+机器学习+AI+Bot

http://www.jianshu.com/u/206a6b0af1c5

Fin & QA

线上反馈