



大规模OpenResty SaaS服务的搭建与运维技巧

2018 OpenResty Con @ 杭州 Hangzhou



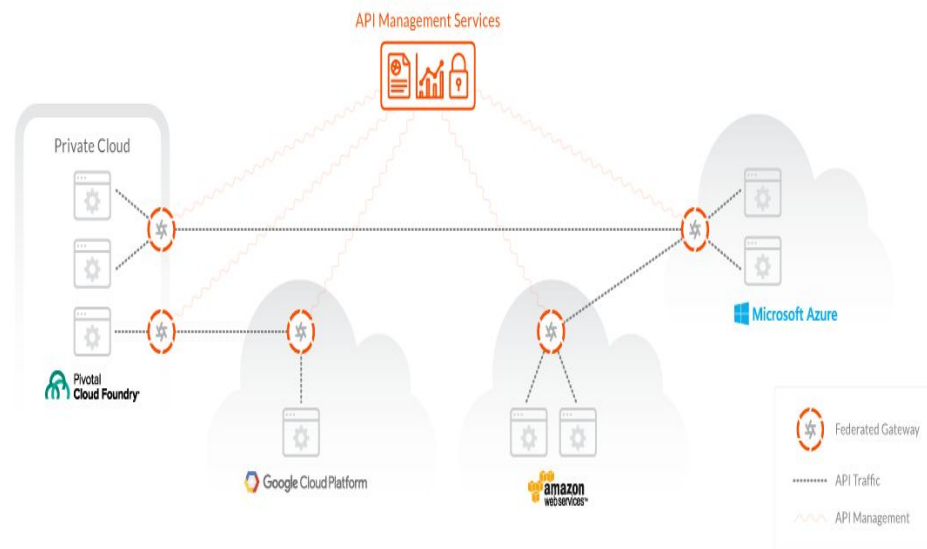
Who Am I?

戴冠兰

- OpenResty 贡献者
- Kong担任工程经理及技术负责人, 带领 Cloud、人工智能、机器学习团队
- 曾在Cloudflare 担任Tech Lead, Edge/WAF/CDN, 万亿级别每月请求规模
- 铁人三项, 户外越野, 自由搏击

如何打造一个好的SaaS产品？

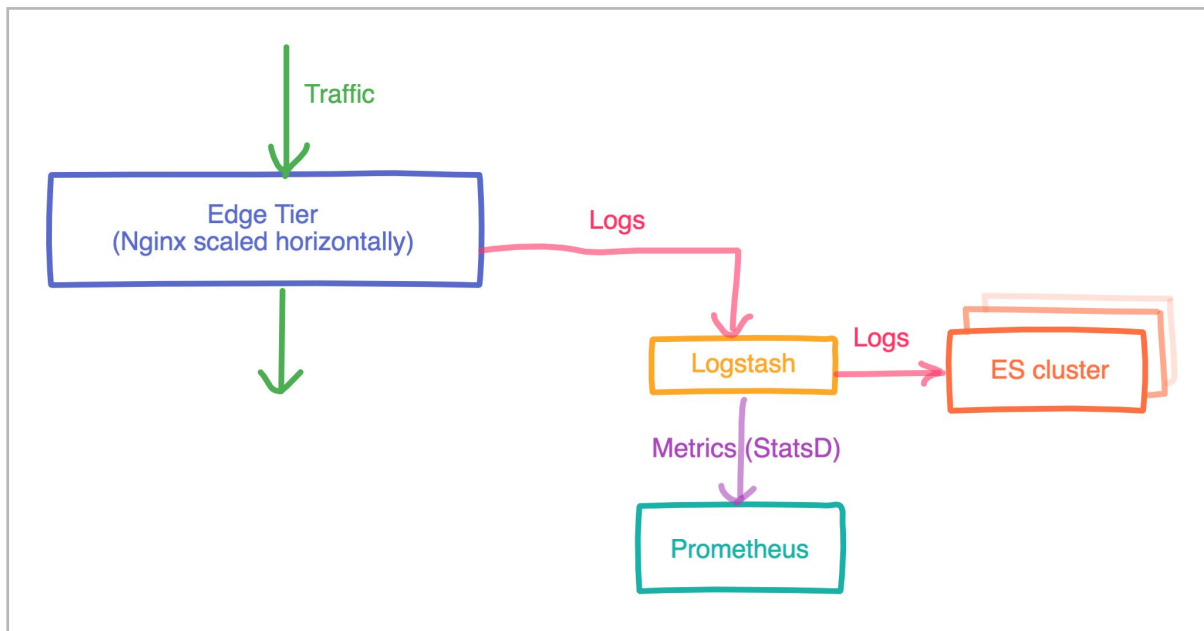
- 可观测性 (Observability)
- 快速部署迭代 (The latest and greatest)
- 飞一般的速度 (Blazingly Fast)
- 多云平台支持 (Cloud Agnostic)



 **THE LATEST AND GREATEST**

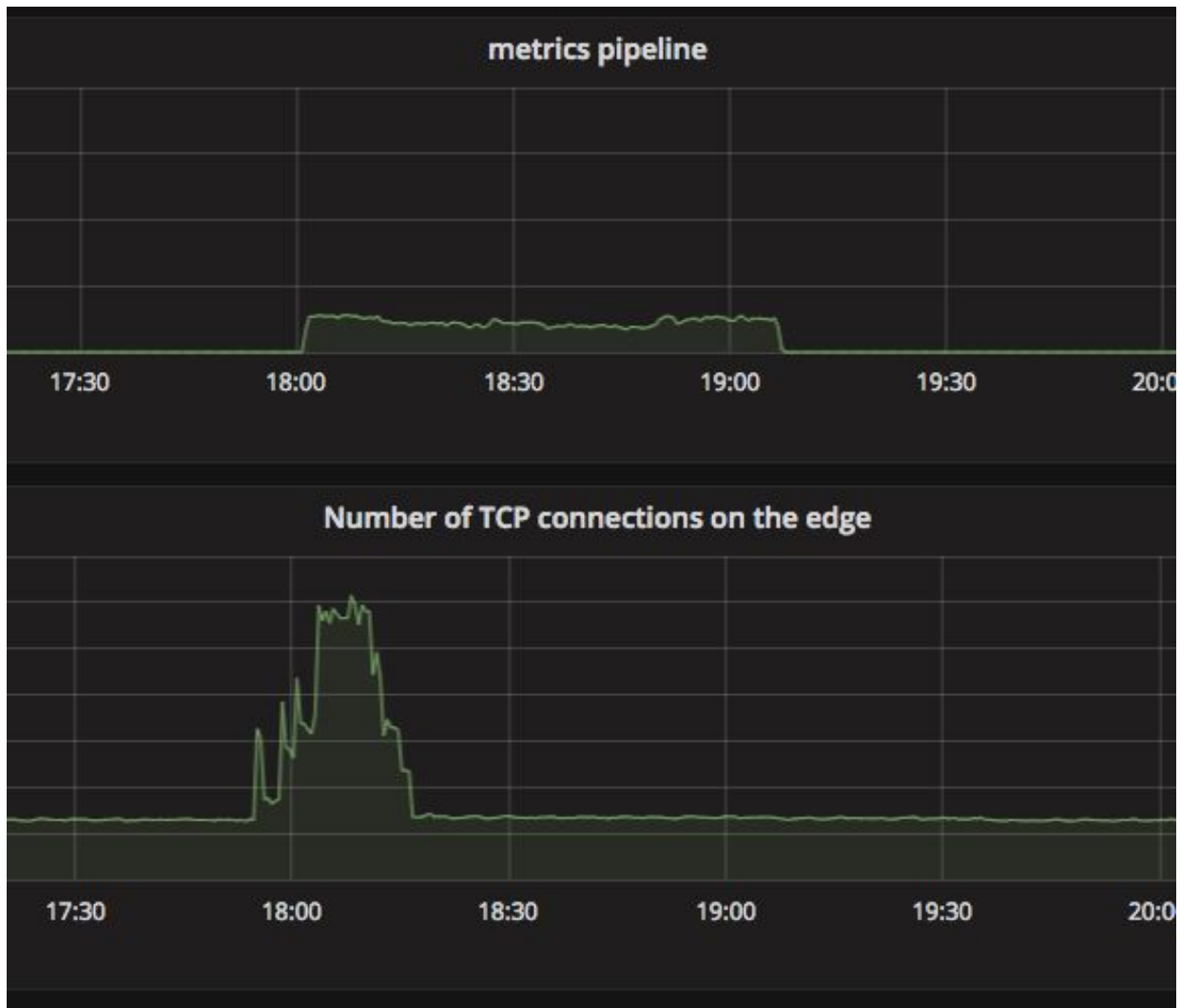
可观测性 (Observability)

Metrics系统搭建实战 - 初版架构设计



局限

- logs != metrics
- 复用了logging pipeline
- Log parse 很慢很昂贵



优化尝试

- Filebeat批处理
- 提升Elasticsearch索引速率

用定制状态机代替Exporter正则引擎

```
mappings:  
- match: client.*.*.request.edge_status.*  
  name: "edge_status"  
  labels:  
    client: "$1"  
    dc: "$2"  
    status: "$3"  
    job: "client"
```

StatsD Event:

```
client.client1.kong-aws-us-west-1.request.edge_status.200
```

Prometheus metric:

```
edge_status{client="client1",dc="kong-aws-us-west-1",status="200"}
```

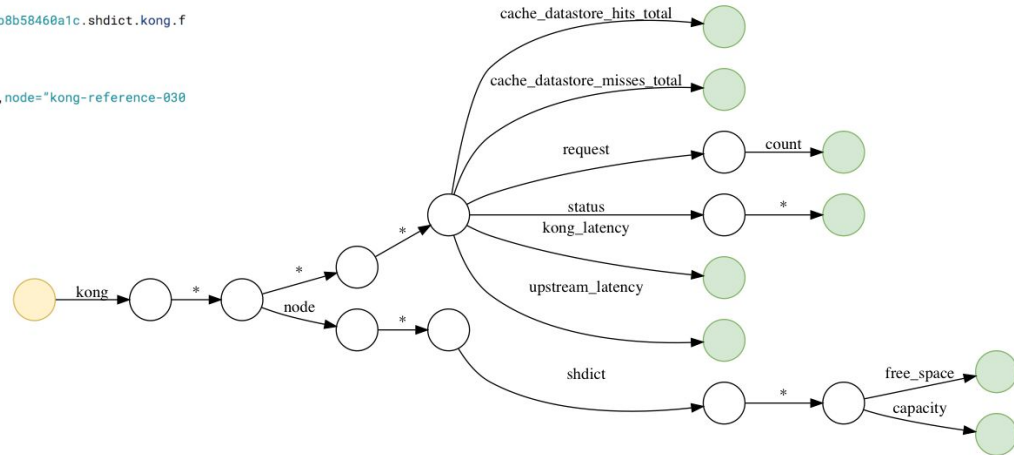
```
- match: kong.*.node.*.shdict.*.free_space  
  name: "kong_shdict_free_space"  
  labels:  
    client: "$1"  
    node: "$2"  
    shdict: "$3"  
    job: "kong_metrics"
```

StatsD Event:

```
kong.reference.node.kong-reference-030f95b8b58460a1c.shdict.kong.free_space
```

Prometheus metric:

```
kong_shdict_free_space{client="reference",node="kong-reference-030f95b8b58460a1c",shdict="kong"}
```




Benchmark结果


StatsD Exporter version	Syscall CPU percentage (prorated)	Time taken to finish 100000 mapping iterations
Stock Binary	20.36%	N/A
Go 1.10.3	23.39%	1.655s
Our Version	42.23%	1.003s
	+19%	-39%




优化已经合并回开源社区

- https://github.com/prometheus/statsd_exporter

Tag: v0.8.0 ▾ [statsd_exporter](#) / CHANGELOG.md Find file Copy path

 **matthiasr** Release 0.8.0 555cd98 6 days ago

4 contributors 

124 lines (95 sloc) | 5.85 KB Raw Blame History   

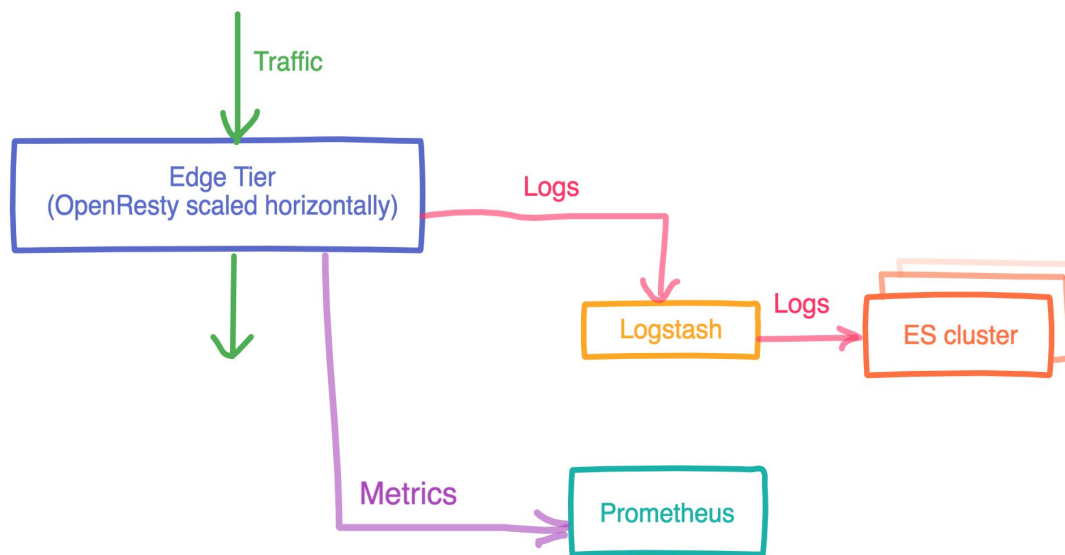
0.8.0 / 2018-10-12

- [ENHANCEMENT] Speed up glob matching (#157)

This release replaces the implementation of the glob matching mechanism, speeding it up significantly. In certain sub-optimal configurations, a warning is logged.

This major enhancement was contributed by [Wangchong Zhou](#).

Metrics系统搭建实战 - 改进架构设计



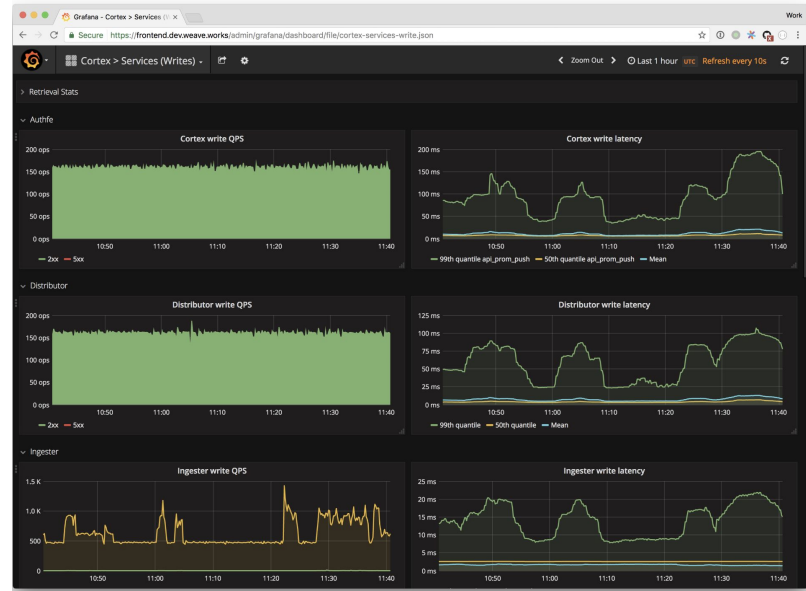
实现

- Log phase 记录每个请求的metrics
- 转接到Prometheus后端



Metrics指标设计之RED大法

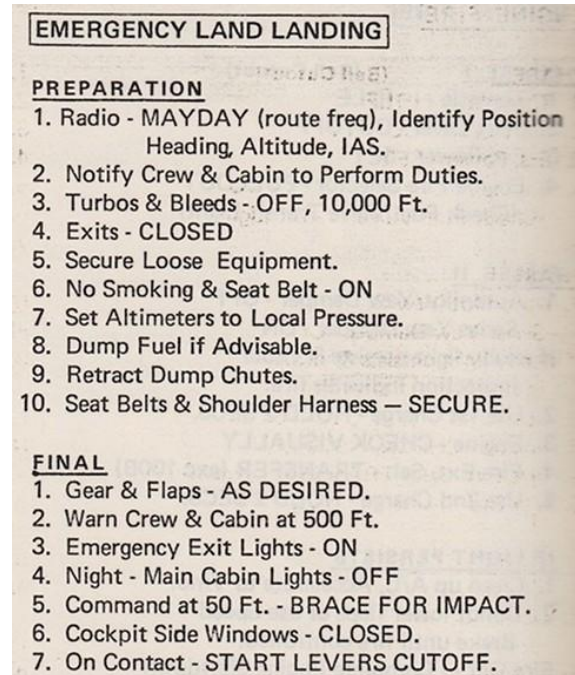
- **R**ate 速率
- **E**rrors 错误
- **D**uration 持续时间



<https://www.weave.works/blog/the-red-method-key-metrics-for-microservices-architecture/>

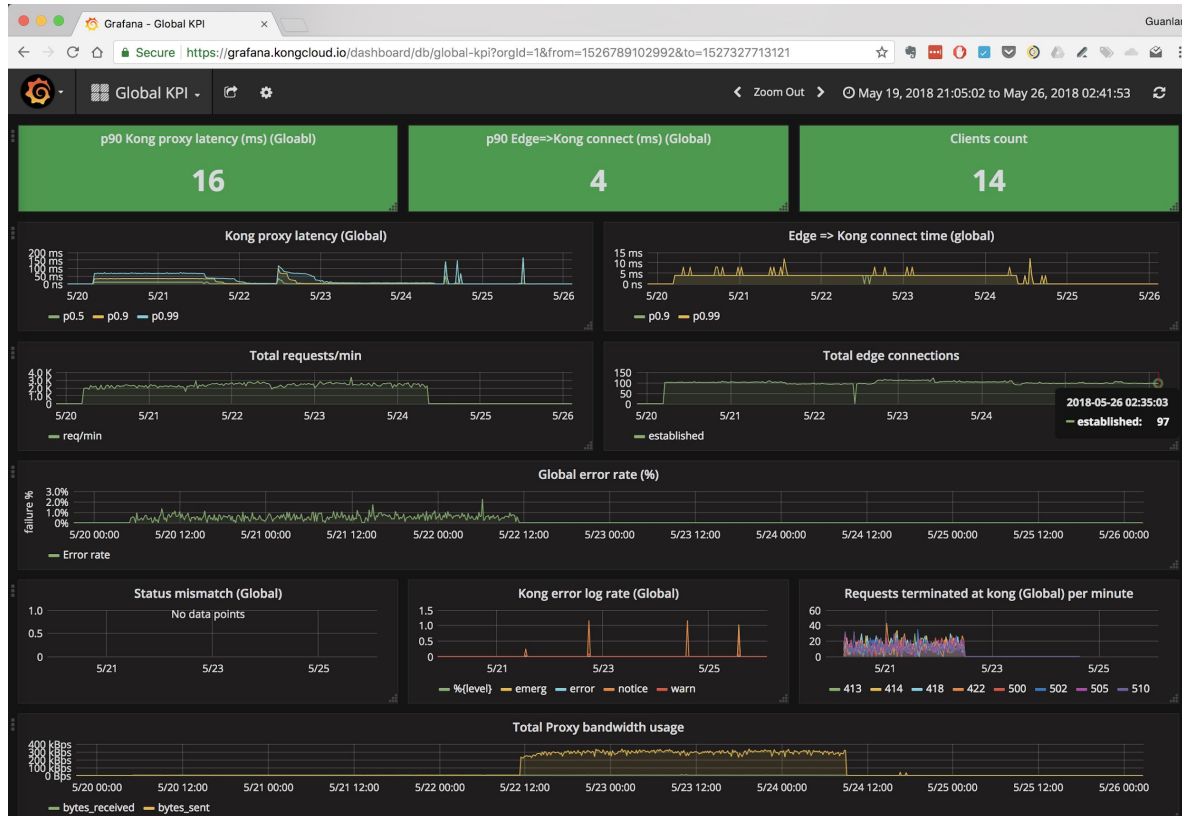
Metrics指标设计之USE大法

- **U**tilization 利用率
- **S**aturation 饱和状态
- **E**rrors 事件错误

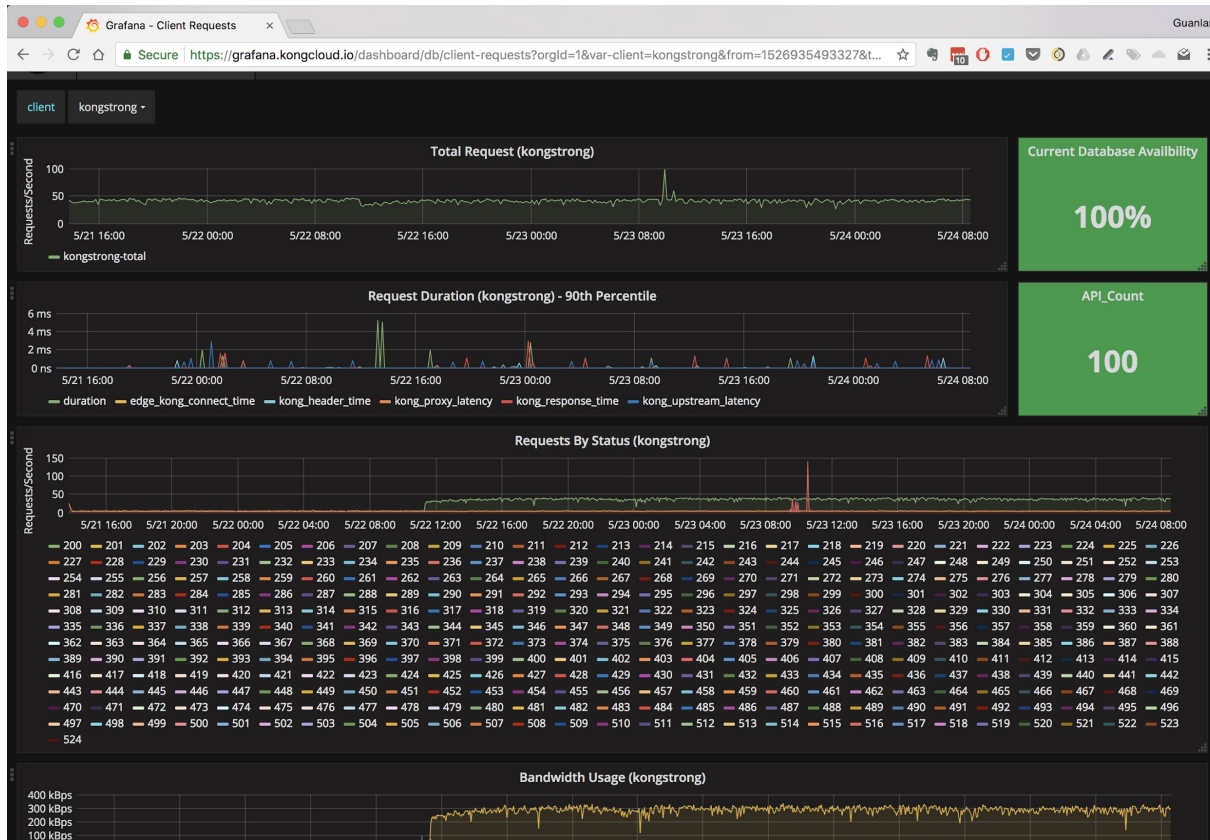


<http://www.brendangregg.com/usemethod.html>

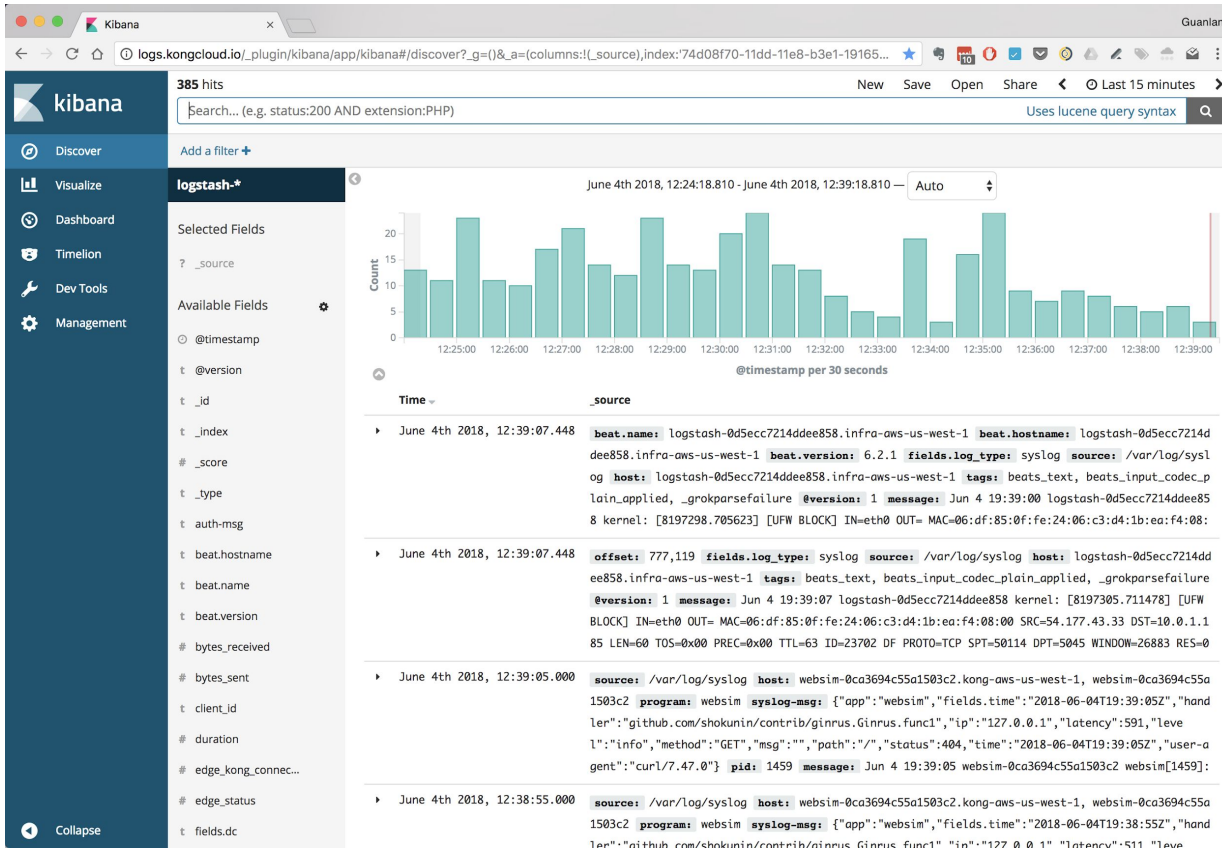
Observability - Grafana (Global)



Observability - Grafana (Per Client)



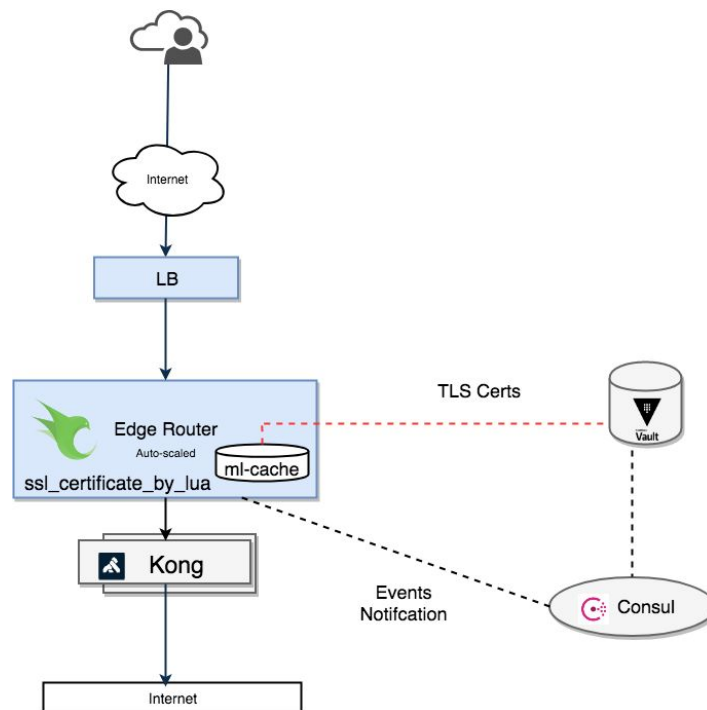
Observability - Logstash



快速部署迭代 (The latest and greatest)

动态客户端证书部署

- 利用ssl_certificate_by_lua
- Vault 存储证书
- Consul 证书事件更新通知



优点

- 避免硬盘存储明文证书
- 证书热更新
- 简化证书部署逻辑

Consul OpenResty支持

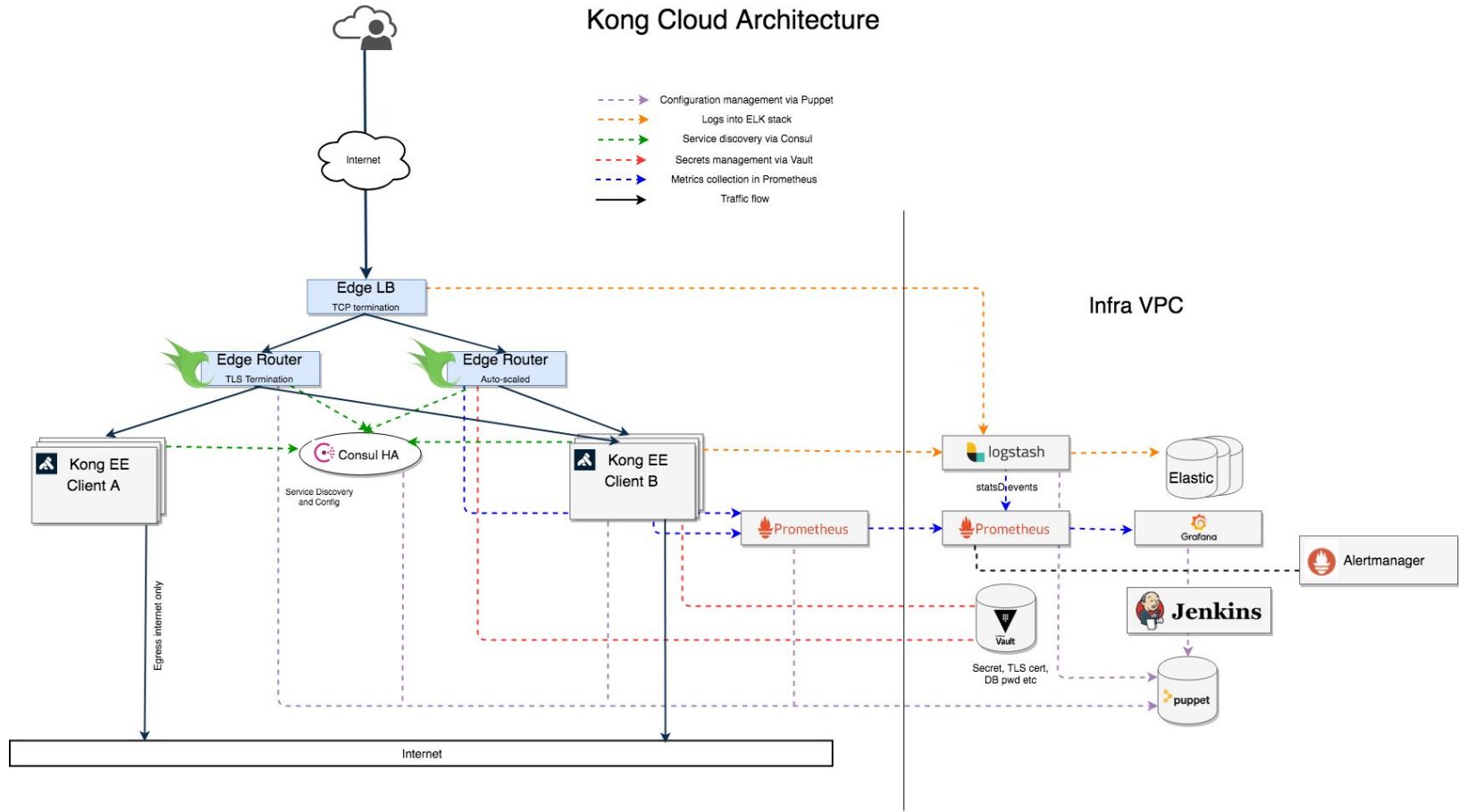
<https://github.com/Kong/lua-resty-consul-events/>
<https://github.com/Kong/lua-resty-vault/> (WIP)

Automation - CI/CD

The screenshot shows the Jenkins web interface. The top navigation bar includes the Jenkins logo, a search bar, and the user name 'Guanlan'. The left sidebar contains navigation options like 'New Item', 'People', 'Build History', and 'Build Queue'. The main content area is titled 'Build History of Jenkins' and features a Gantt chart showing build timelines from June 5th to June 7th. Below the chart is a table listing the build details.

Build	Time Since ↑	Status
UpgradeDemoAMIs #3	17 min	back to normal
packer-base-kong-ee #77	40 min	stable
tf-kong-cloud #87	44 min	stable
packer-base-kong-ee #76	1 hr 7 min	stable
tf-kong-cloud #86	1 hr 9 min	stable
puppet-kong_ee #121	1 hr 19 min	stable
Deprovision Kong #55	5 hr 37 min	stable

Kong Cloud Architecture



飞一般的速度 (Blazingly Fast)

OpenResty系统调优

- CPU affinity 亲和性配置
- NGINX 配置优化
- Linux sysctl 设置
- Intel Quickassist

NGINX 指令配置

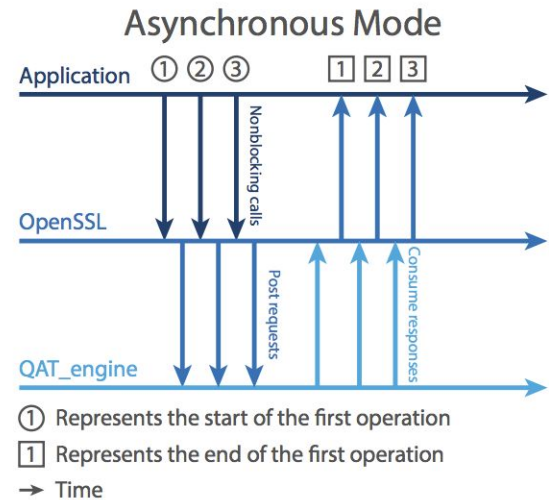
- keepalive_timeout, keepalive_requests
- sendfile on - copy fd at kernel level
- tcp_nopush, tcp_nodelay
- listen backlog reuserport

Linux sysctl 设置

- Memory
- Size of processor queue
- Maximum TCP buffer size
- Disable TCP timestamps

Intel Quickassist

- https://github.com/intel/asynch_mode_nginx
- 异步SSL/TLS 处理加速
- 需要硬件支持
 - Intel® C62X Series Chipset
 - Intel® Communications Chipset 8925 to 8955 Series



<https://01.org/sites/default/files/downloads/intelr-quickassist-technology/intelquickassisttechnologyopensslperformance.pdf>

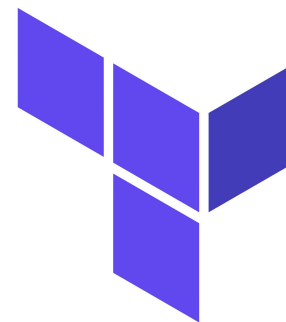
多云平台支持 (Cloud Agnostic)

多云平台设计技巧

- Infrastructure as code
- 避免依赖云平台独占功能
- Kubernetes

Infrastructure as code - Terraform实例

```
provider "aws" {  
  region = "us-west-2"  
  profile = "dev"  
}  
module "kong" {  
  source = "github.com/zilllowgroup/kong-terraform"  
  vpc_name      = "my-vpc"  
  environment   = "dev"  
  ec2_instance_type = "t2.small"  
  ec2_ebs_optimized = false  
  ec2_key_name   = "my-key"  
  ssl_cert_external = "*.domain.name"  
  ssl_cert_internal = "*.domain.name"  
  ssl_cert_internal_gui = "*.domain.name"  
  enable_internal_lb = true  
  db_instance_count = 3  
  tags = {  
    Owner = "devops@domain.name"  
    Team = "DevOps"  
  }  
}
```



HashiCorp
Terraform

Terraform Kong支持

<https://github.com/zillowgroup/kong-terraform>

Q&A