GOPS 2021
Shenzhen

GOPS

GOPS 全球运维大会

2021
-XOPS 风向标

深圳站

中国·深圳

指导单位：

主办单位：

时间：2021年5月21日-22日

GOPS 2021
shenzhen

如何在K8S中用好Nginx?

陶辉　　智链达CTO

陶辉

智链达联合创始人 & CTO

《深入理解Nginx》作者，极客时间《系统性能优化必知必会》专栏作者，视频课《Web协议详解与抓包实战》

《Nginx核心知识100讲》讲师，腾讯云TVP

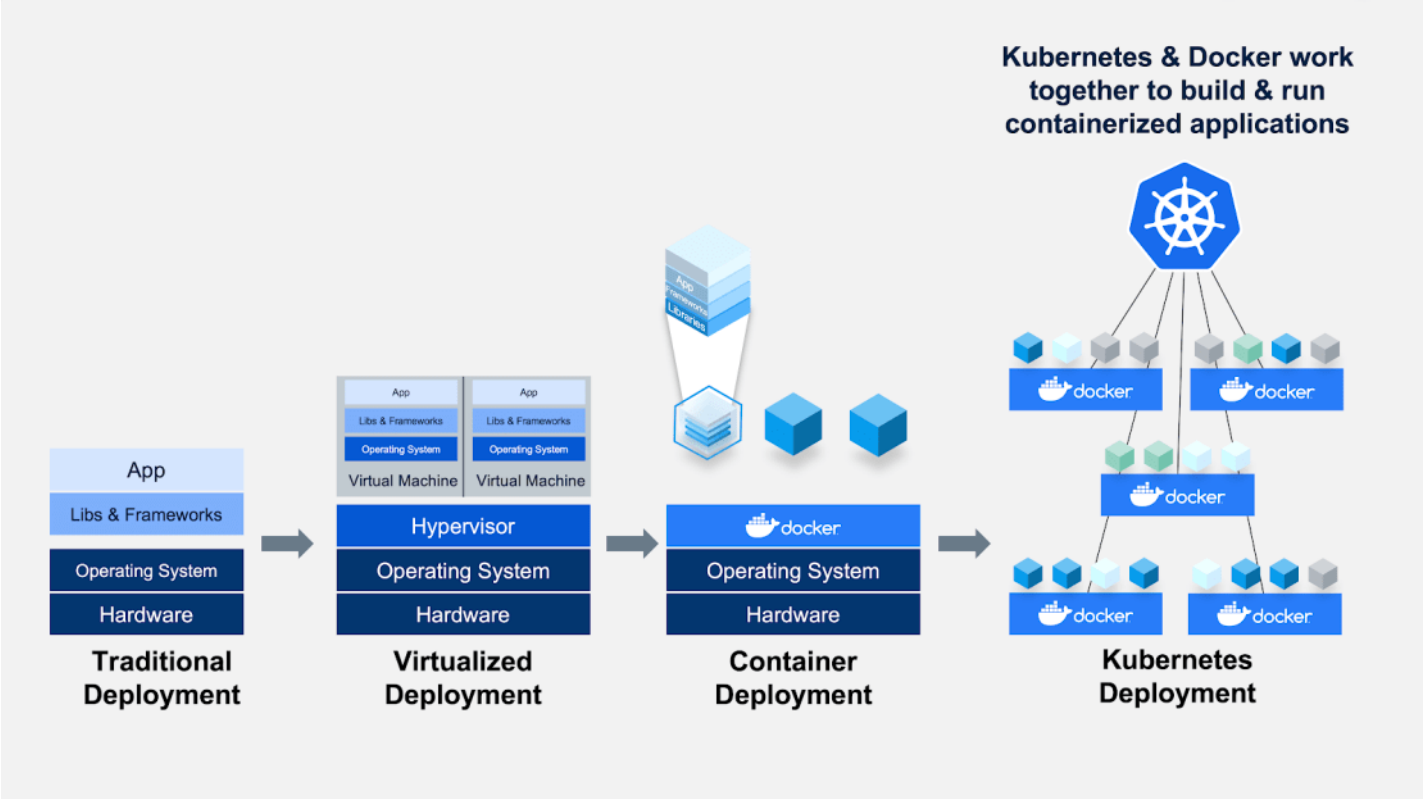
目录
CONTENTS

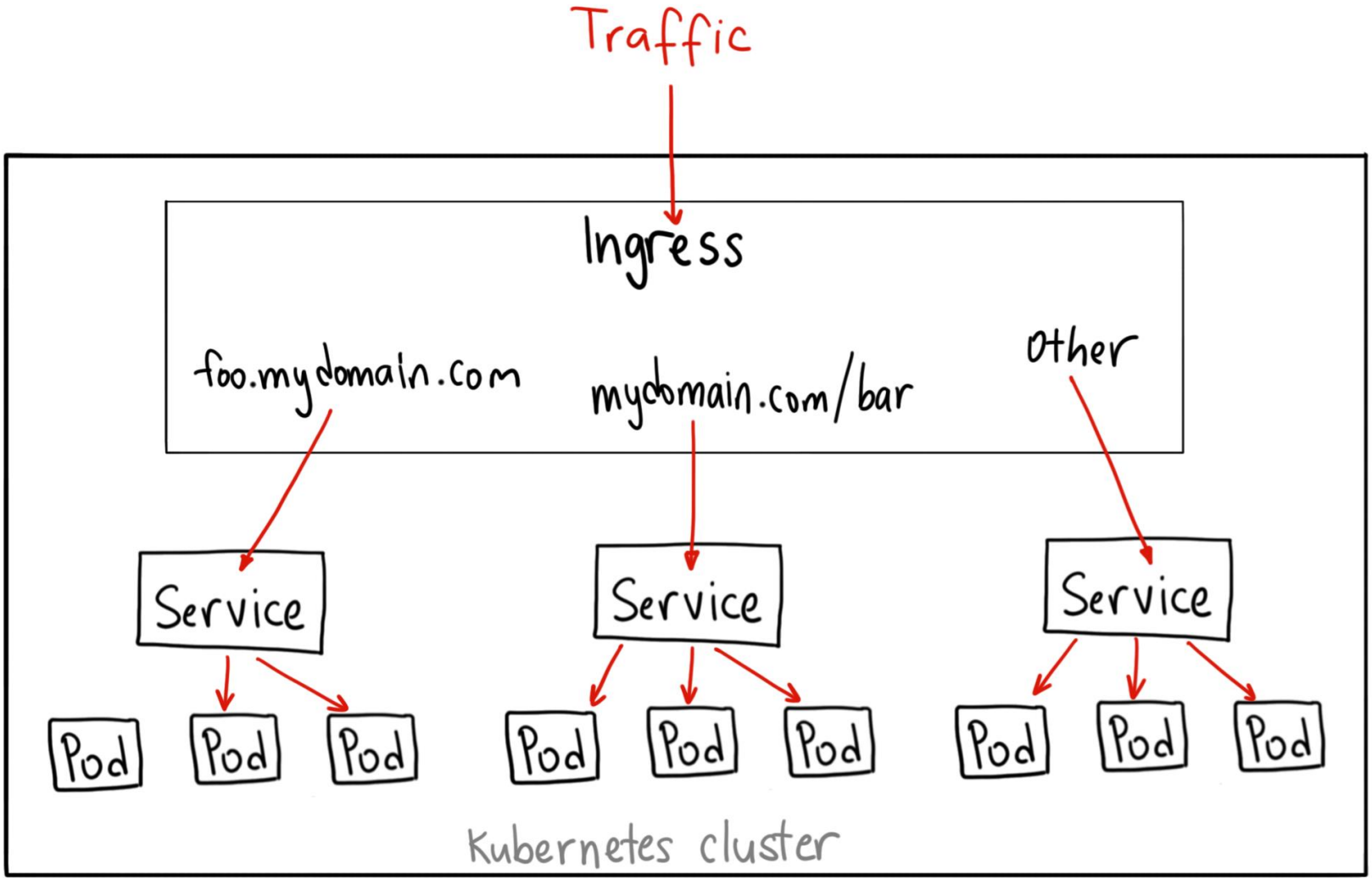
K8S如何同步Nginx配置?

nginx.conf的修改流程

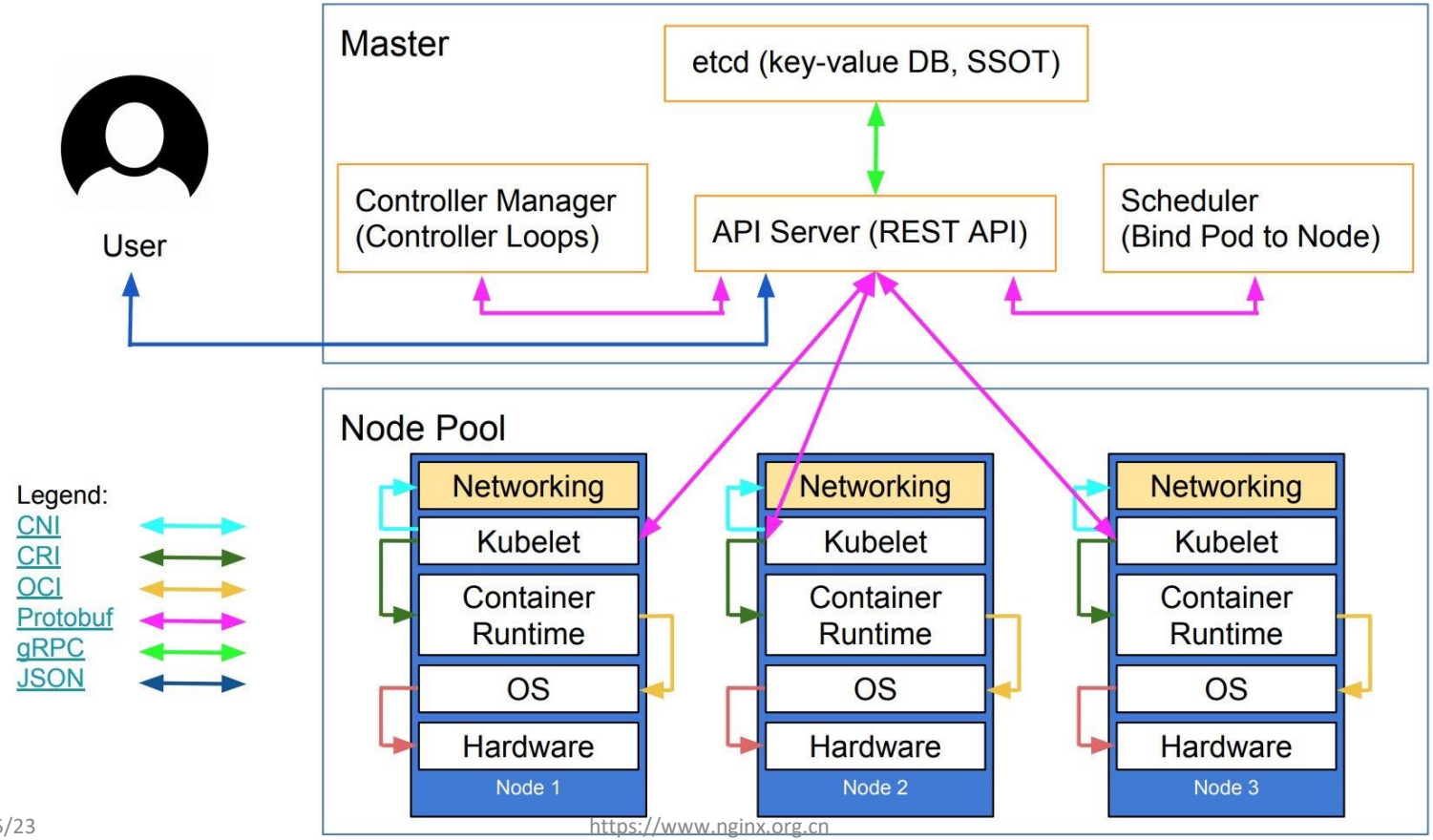
01

容器集群的特点


Kubernetes & Docker work together to build & run containerized applications
App
Libs & Frameworks
Operating System
Hardware
Traditional Deployment
App
Libs & Frameworks
Operating System
Virtual Machine
App
Libs & Frameworks
Operating System
Virtual Machine
Hypervisor
Operating System
Hardware
Virtualized Deployment
docker
Operating System
Hardware
Container Deployment
docker
docker
docker
docker
docker
Kubernetes Deployment

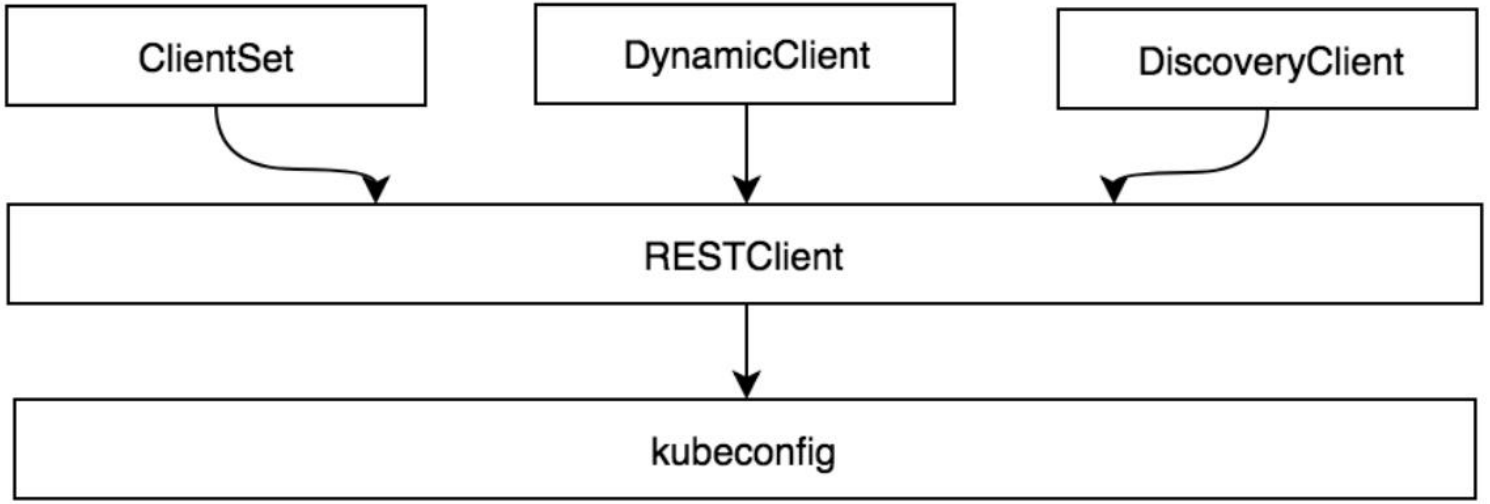

Traffic
Ingress
foo.mydomain.com
mydomain.com/bar
Other
Service
Service
Service
Pod
Pod
Pod
Pod
Pod
Pod
Pod
Pod
Pod
Kubernetes cluster

Kubernetes' high-level component architecture

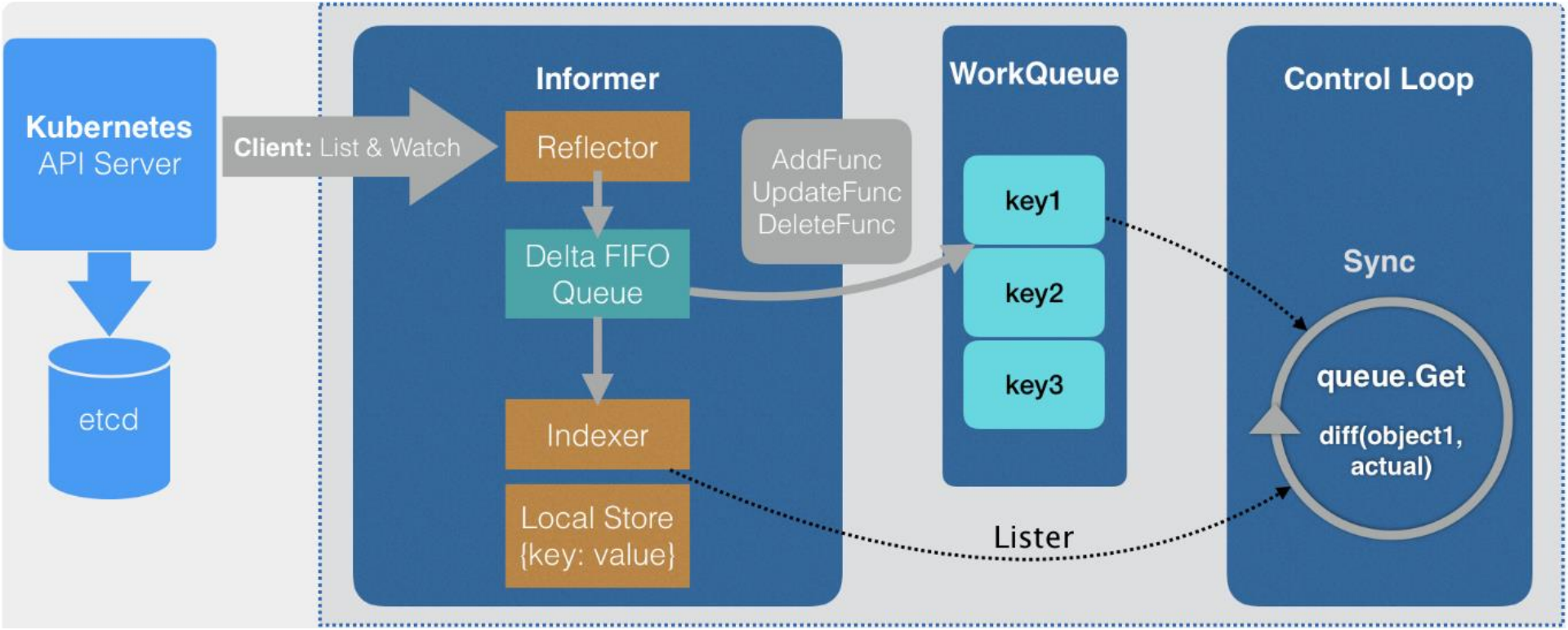


client-go

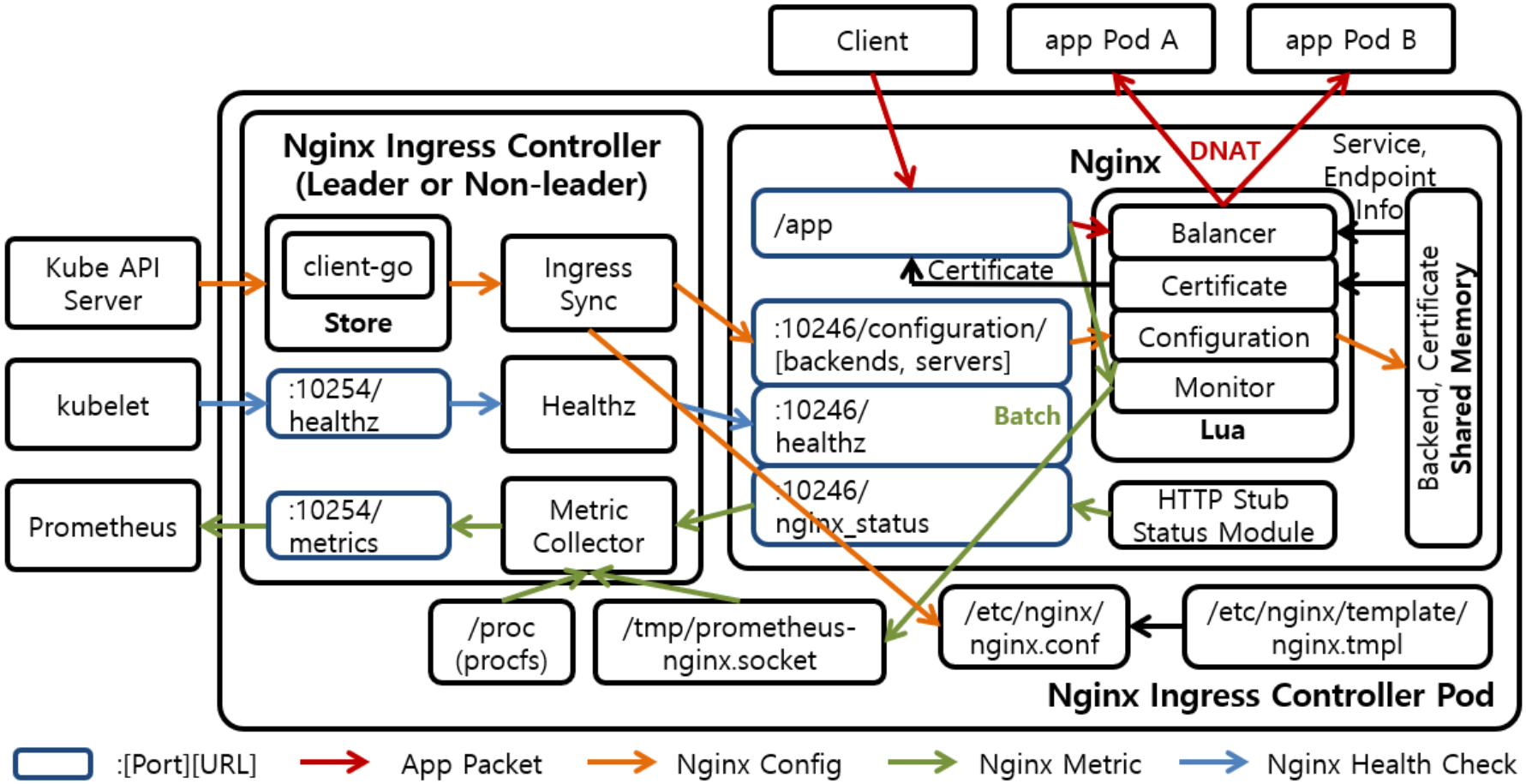
- https://github.com/kubernetes/client-go



client-go informer原理



K8S社区Nginx Ingress Controller架构



怎样在yaml中加入第三方模块指令?

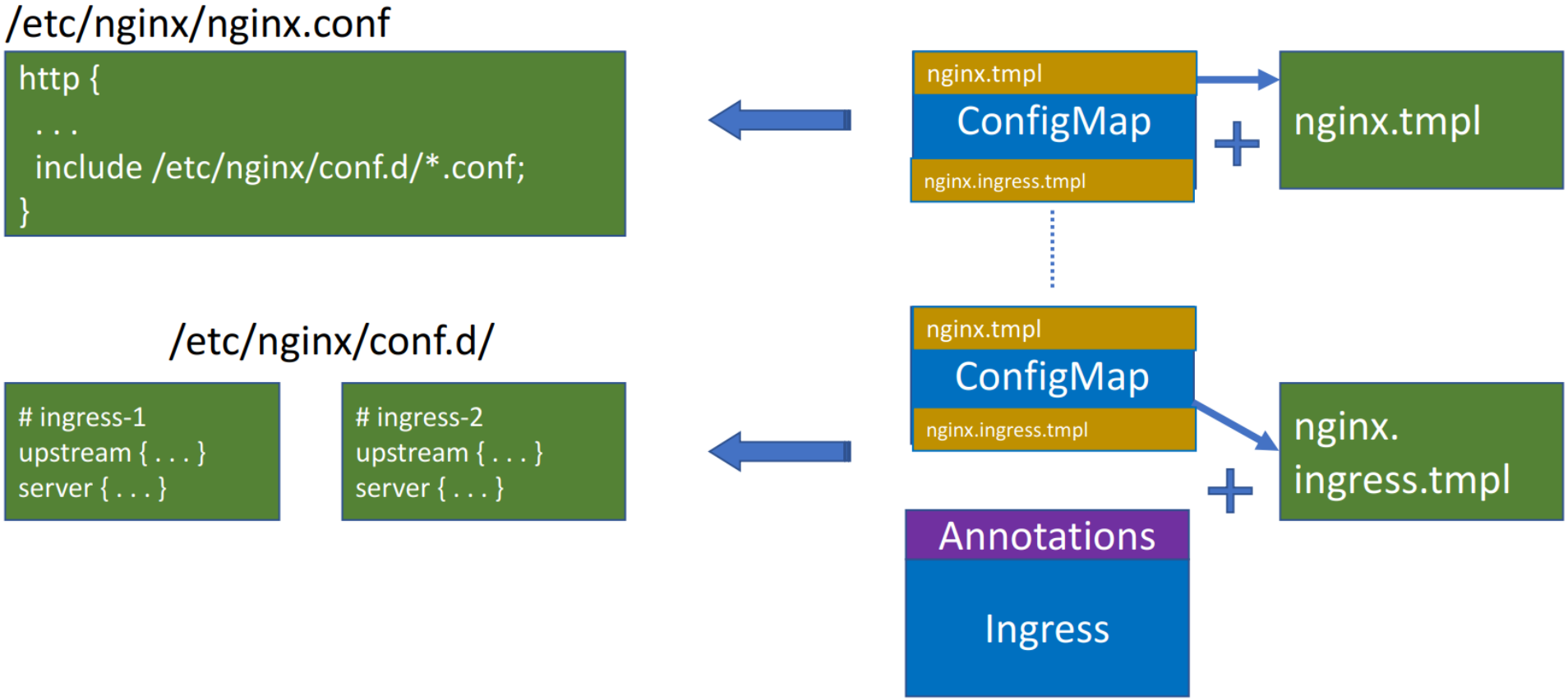
随心所欲修改nginx.conf

02

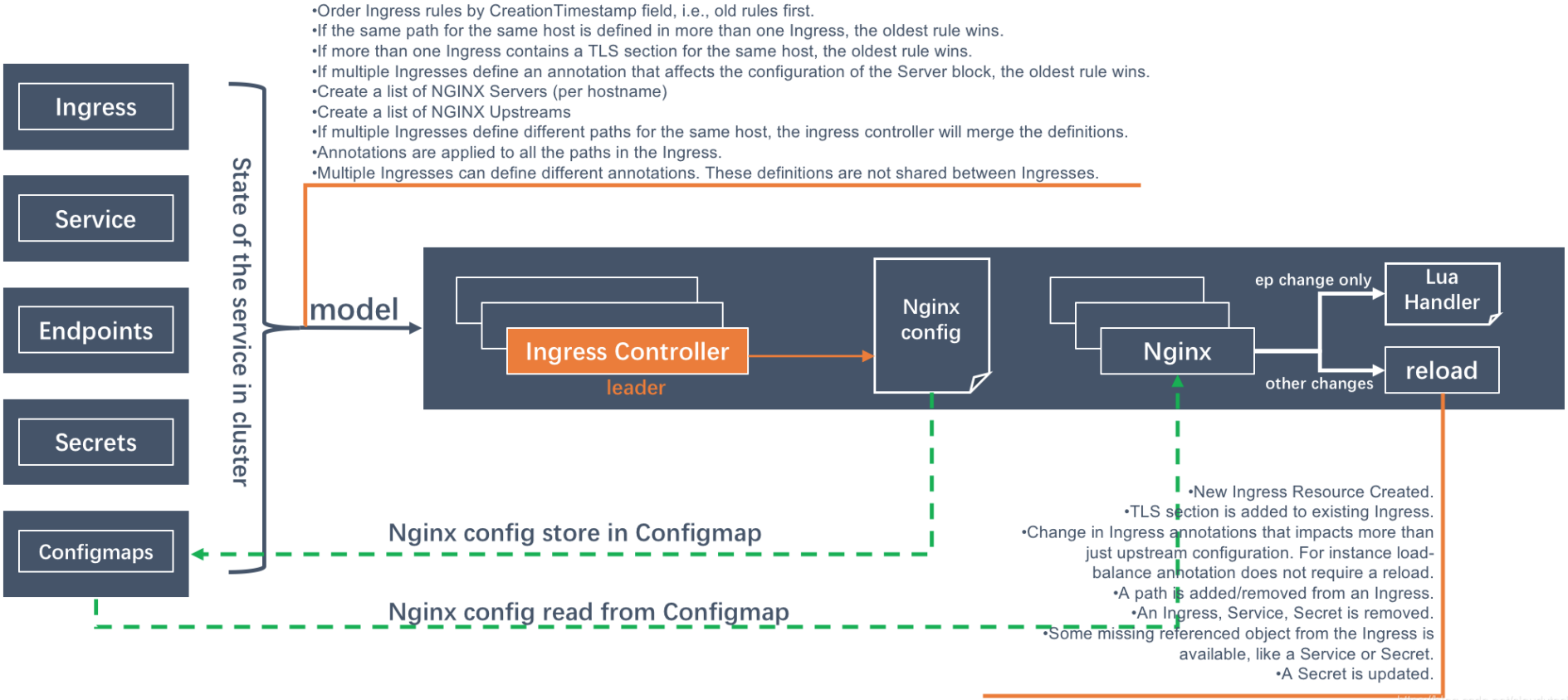
nginx.conf怎么搬到yaml声明配置中?

```
worker_processes    4;
worker_cpu_affinity 0001 0010 0100 1000;
events {  worker_connections 2048; }
http {
        upstream backend {    10.244.1.3 weight=1;     10.244.0.9 weight=3;    least_conn;    }
        server {
                listen 80;
                server_name www.taohui.pub;
                subs_filter_types text/html text/css text/xml;
                subs_filter st(\d*).taohui.tech $1.taohui.org.cn ir;
                location /private/ { auth_request /auth; ... }
                location = /auth {     proxy_pass http://auth.taohui.tech;     }
                location /ajp {   ajp_pass backend;  }
        }
}
stream {
}
```

Ingress/ConfigMap/Annotations/Template



还有Lua...


•Order Ingress rules by CreationTimestamp field, i.e., old rules first.
•If the same path for the same host is defined in more than one Ingress, the oldest rule wins.
•If more than one Ingress contains a TLS section for the same host, the oldest rule wins.
•If multiple Ingresses define an annotation that affects the configuration of the Server block, the oldest rule wins.
•Create a list of NGINX Servers (per hostname)
•Create a list of NGINX Upstreams
•If multiple Ingresses define different paths for the same host, the ingress controller will merge the definitions.
•Annotations are applied to all the paths in the Ingress.
•Multiple Ingresses can define different annotations. These definitions are not shared between Ingresses.
Ingress
Service
Endpoints
Secrets
Configmaps
State of the service in cluster
model
Ingress Controller
leader
Nginx config
Nginx
ep change only
Lua Handler
other changes
reload
Nginx config store in Configmap
Nginx config read from Configmap
•New Ingress Resource Created.
•TLS section is added to existing Ingress.
•Change in Ingress annotations that impacts more than just upstream configuration. For instance load-balance annotation does not require a reload.
•A path is added/removed from an Ingress.
•An Ingress, Service, Secret is removed.
•Some missing referenced object from the Ingress is available, like a Service or Secret.
•A Secret is updated.

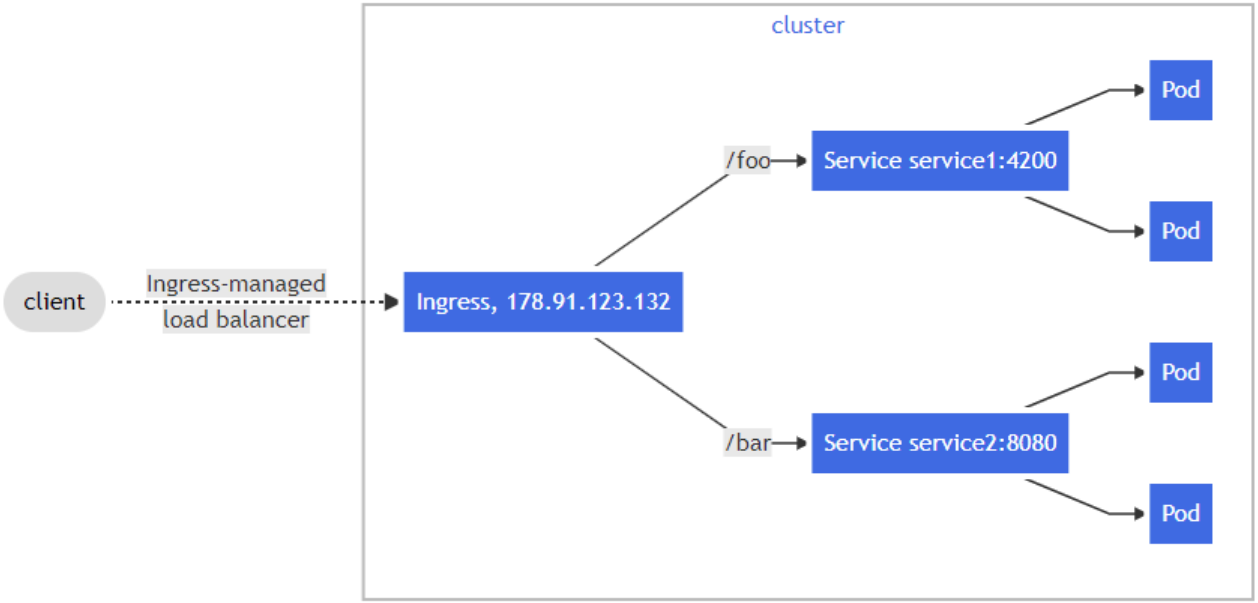
社区Controller提供了哪些模块?

默认Controller的赋能方式

03

Ingress Controller要完成哪些工作?

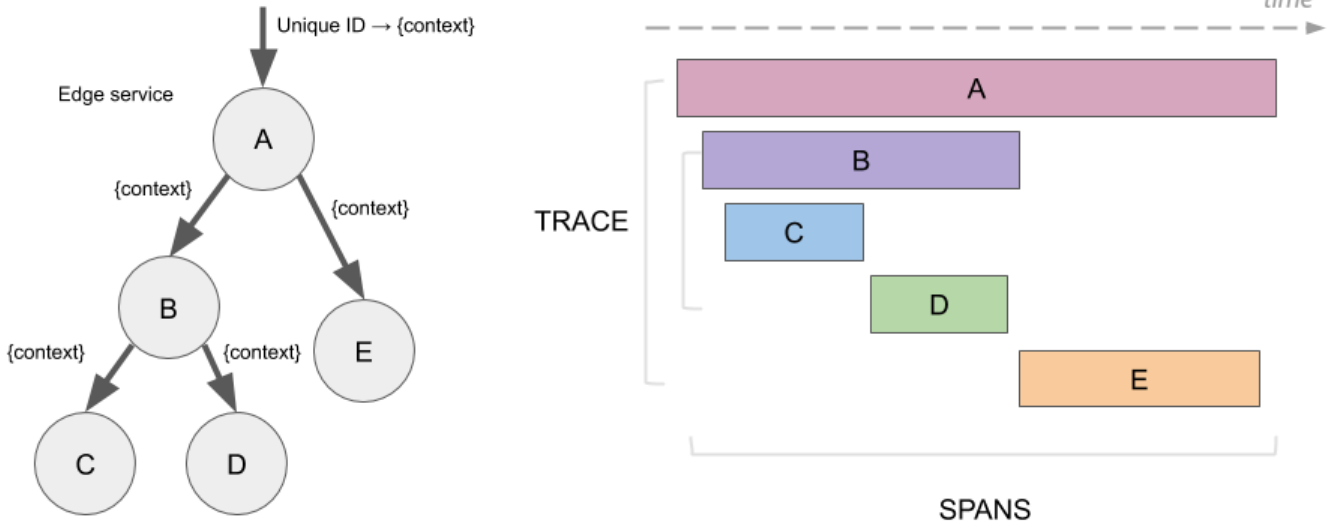
- 负载均衡/会话保持
- 协议转换
- TLS卸载与认证
- 文本压缩
- 请求认证
- 限流限速
- WAF
- 全链路跟踪
- 日志上报



Nginx模块钩子

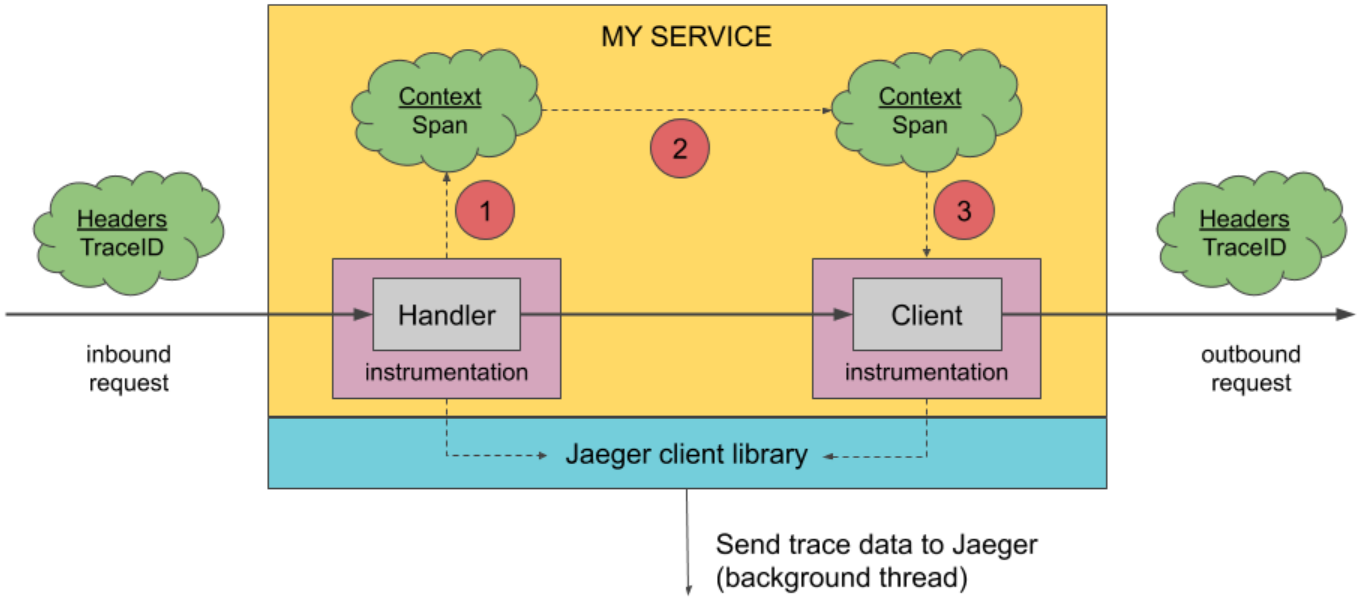
处理阶段	子阶段	模块举例	指令
openssl		lua_nginx_module	ssl_certificated_by_lua
rewrite		ngx_http_opentracing_module	enable-opentracing: "true"
preaccess		ngx_http_limit_req_module ngx_http_limit_conn_module	limit-rps “100” limit-connections “5”
access		ngx_http_auth_digest_module ngx_http_auth_basic_module ngx_http_modsecurity_module	auth-type: “digest” auth-url: “url” modsecurity
content		nginx_ajp_module	ajp_pass tomcats;
	upstream	lua-upstream-nginx-module	balance_by_lua
	filter	ngx_http_brotli_filter_module	enable-brotli “true”
log		nginx-influxdb-module	enable-influxdb “true”

全链路跟踪的实现原理

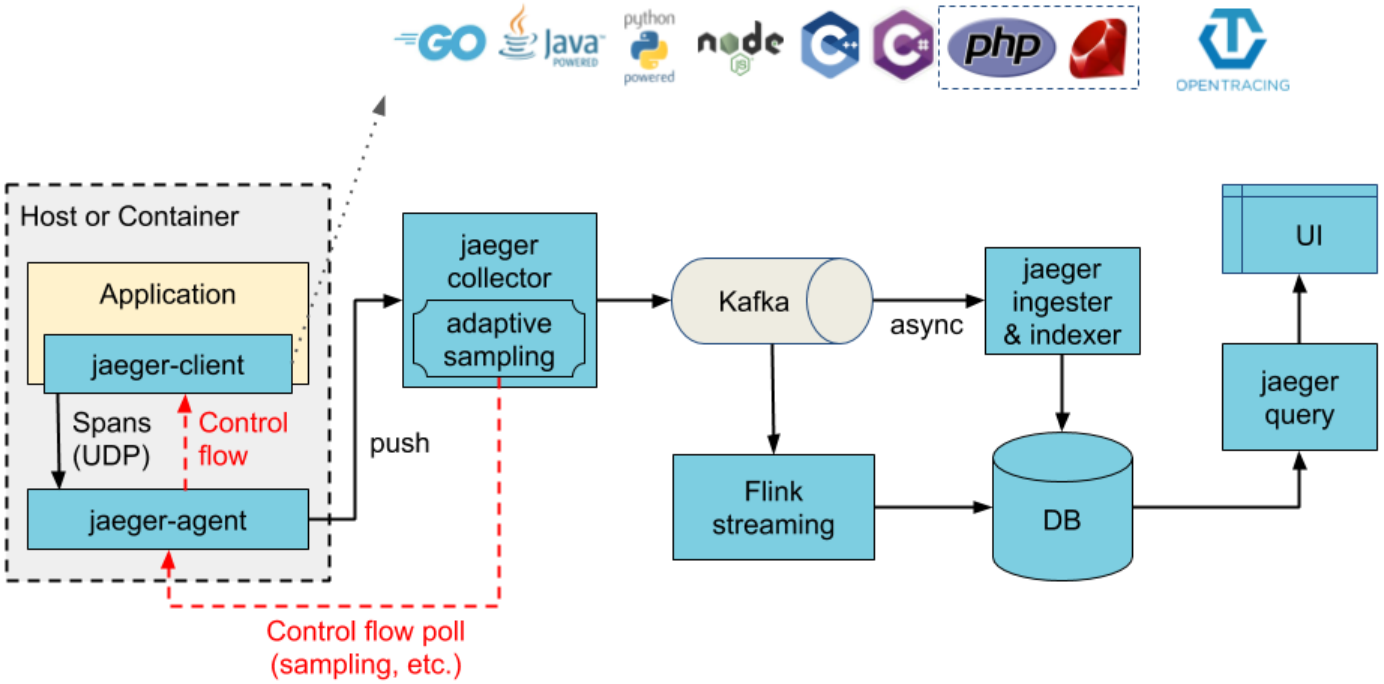

Unique ID → {context}
Edge service
A
{context}
{context}
B
{context}
{context}
E
C
D
time
TRACE
A
B
C
D
E
SPANS

opentracing规范

https://opentracing.io/



全链路跟踪的实现者：jaeger

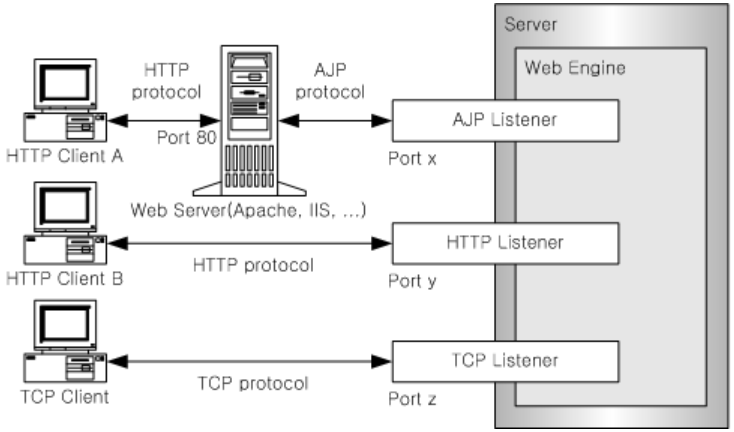


基于opentracing接口实现全链路跟踪

- nginx-opentracing模块
  - https://github.com/opentracing-contrib/nginx-opentracing
    - https://github.com/opentracing/opentracing-cpp
    - https://github.com/jaegertracing/jaeger-client-cpp
    - https://github.com/rnburn/zipkin-cpp-opentracing
    - https://github.com/DataDog/dd-opentracing-cpp
    - ~~https://github.com/lightstep/lightstep-tracer-cpp~~
  - C++库，支持Jaeger, Zipkin, ~~LightStep~~, Datadog.
  - 指令
    - opentracing_load_tracer libjaegertracing_plugin.so jaeger-nginx-config.json
    - opentracing on;
    - opentracing_tag http_user_agent $http_user_agent;
    - opentracing_operation_name $uri;
    - opentracing_(_fastcgi/_grpc)_propagate_context/;
    - opentracing_trust_incoming_span
    - opentracing_location_operation_name
    - opentracing_trace_locations

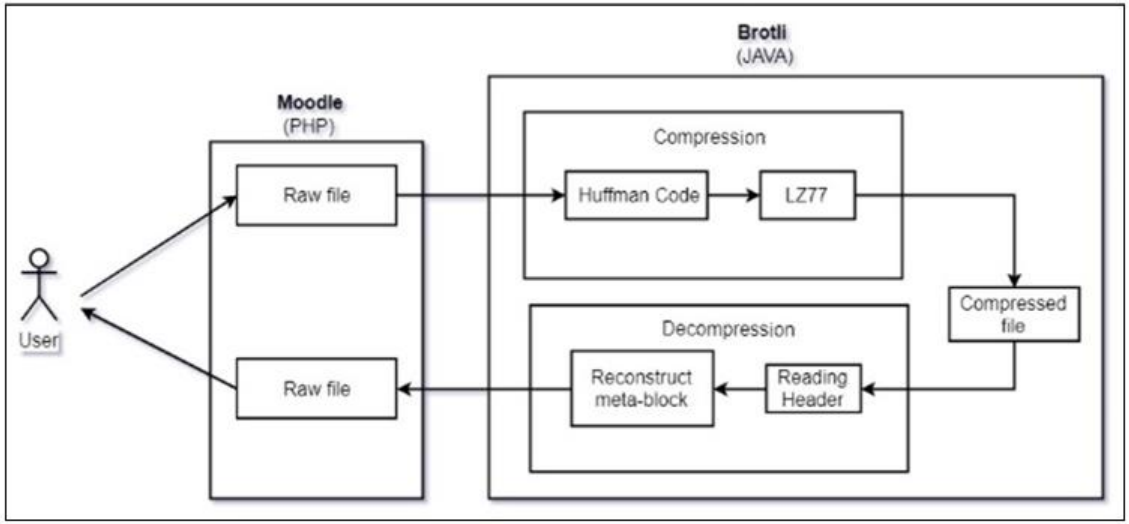
AJP: Apache Jserv Protocol

- backend-protocol：AJP
  - HTTP/HTTPS
  - GRPC/GRPCS
  - FASTCGI
- nginx_ajp_module模块
  - https://github.com/yaoweibin/nginx_ajp_module



brotli压缩算法

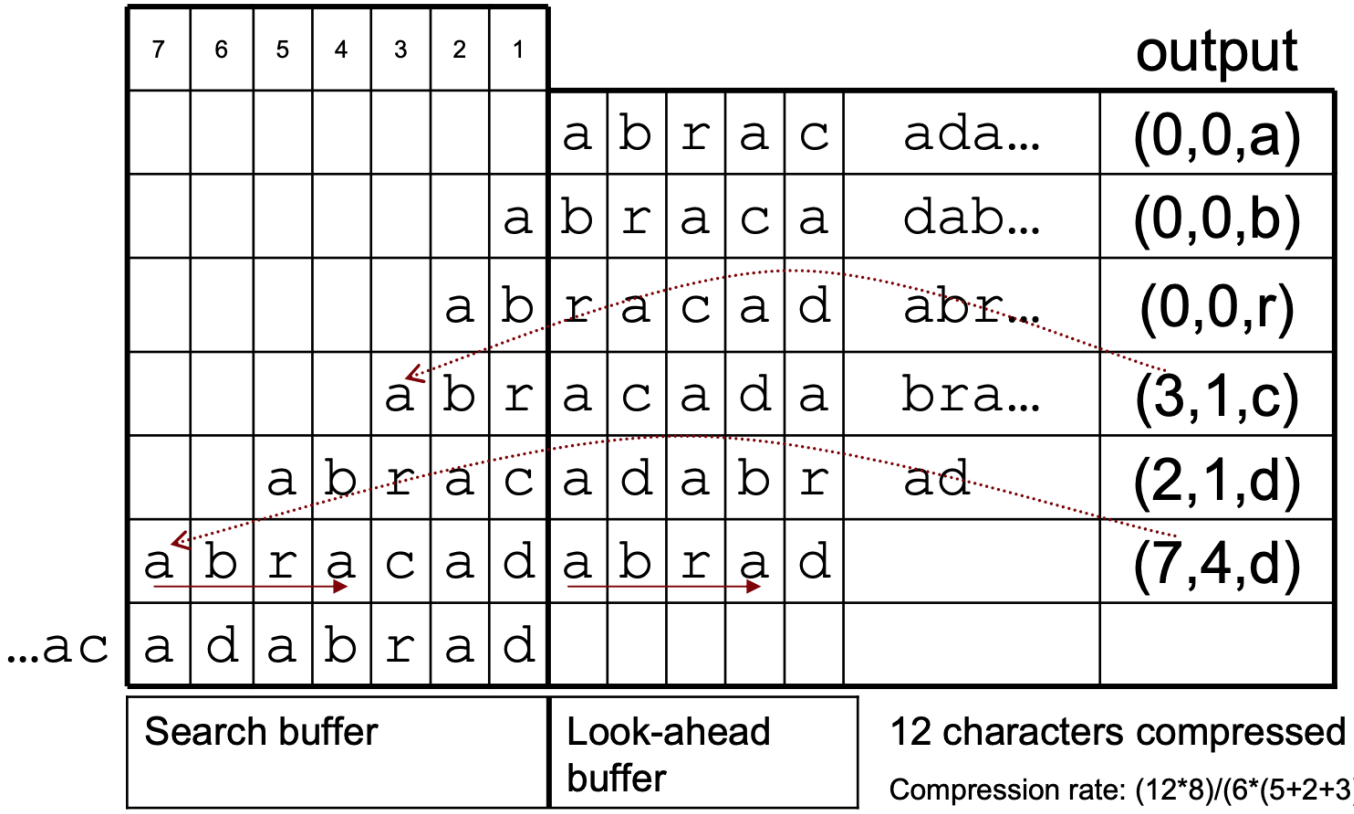
- http://www.icicelb.org/ellb/contents/2019/11/elb-10-11-02.pdf
- https://github.com/google/brotli



LZ77动态字典

Encoding of the string:
abracadabrad

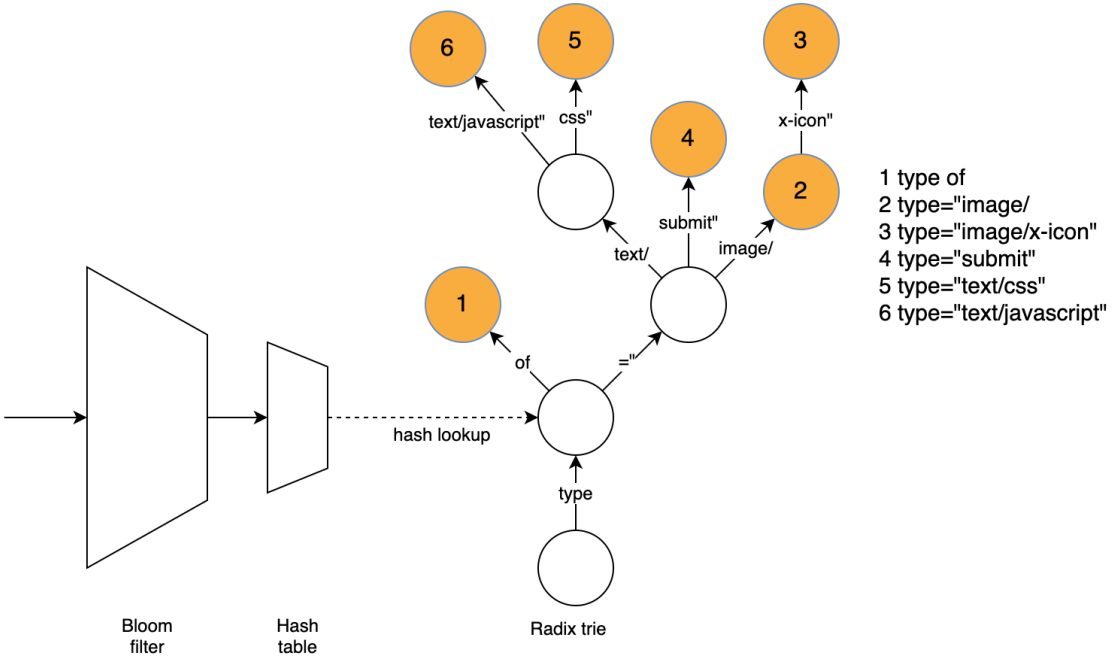
output tuple: (offset, length, symbol)



12 characters compressed into 6 tuples

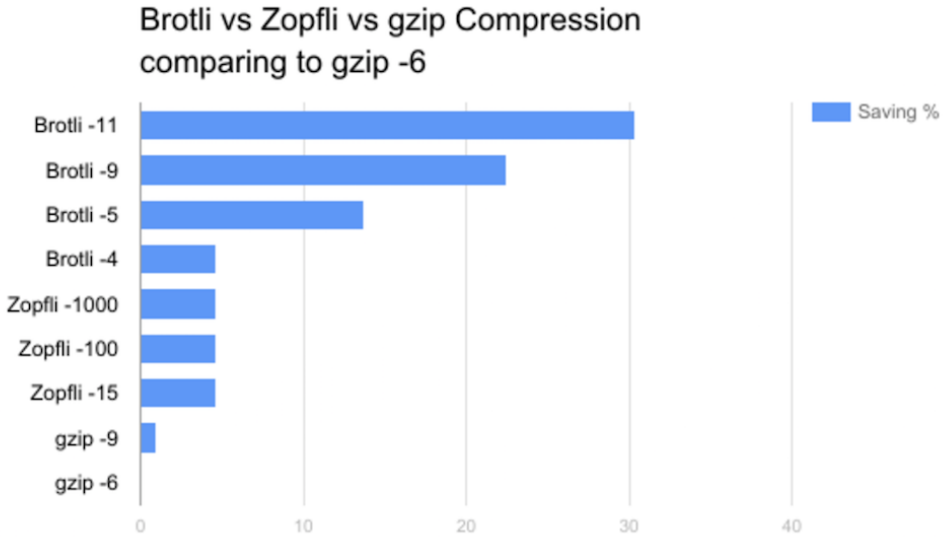
Compression rate: (12*8)/(6*(5+2+3))=96/60=1,6=60%.

brotli静态字典


6
5
3
text/javascript"
css"
4
x-icon"
2
submit"
text/
image/
1
of
="
hash lookup
type
Bloom filter
Hash table
Radix trie
1 type of
2 type="image/
3 type="image/x-icon"
4 type="submit"
5 type="text/css"
6 type="text/javascript"

ngx_brotli模块

- https://github.com/google/ngx_brotli



K8S官方Nginx默认开启了哪些模块?

- Nginx官方默认未编译模块
  - --with-http_ssl_module \
  - --with-http_stub_status_module \
  - --with-http_realip_module \
  - --with-http_auth_request_module \
  - --with-http_addition_module \
  - --with-http_geoip_module \
  - --with-http_gzip_static_module \
  - --with-http_sub_module \
  - --with-http_v2_module \
  - --with-stream \
  - --with-stream_ssl_module \
  - --with-stream_realip_module \
  - --with-stream_ssl_preread_module \
  - --with-threads \
  - --with-http_secure_link_module \
  - --with-http_gunzip_module"

- **第三方C模块**
  - nginx-influxdb-module
  - ngx_http_geoip2_module
  - nginx_ajp_module
  - ModSecurity-nginx
  - msgpack-c
  - ngx_devel_kit
  - set-misc-nginx-module
  - headers-more-nginx-module
  - nginx-http-auth-digest
  - ngx_http_substitutions_filter_module
  - nginx-opentracing
  - lua-nginx-module
  - stream-lua-nginx-module
  - lua-upstream-nginx-module
  - ngx_brotli

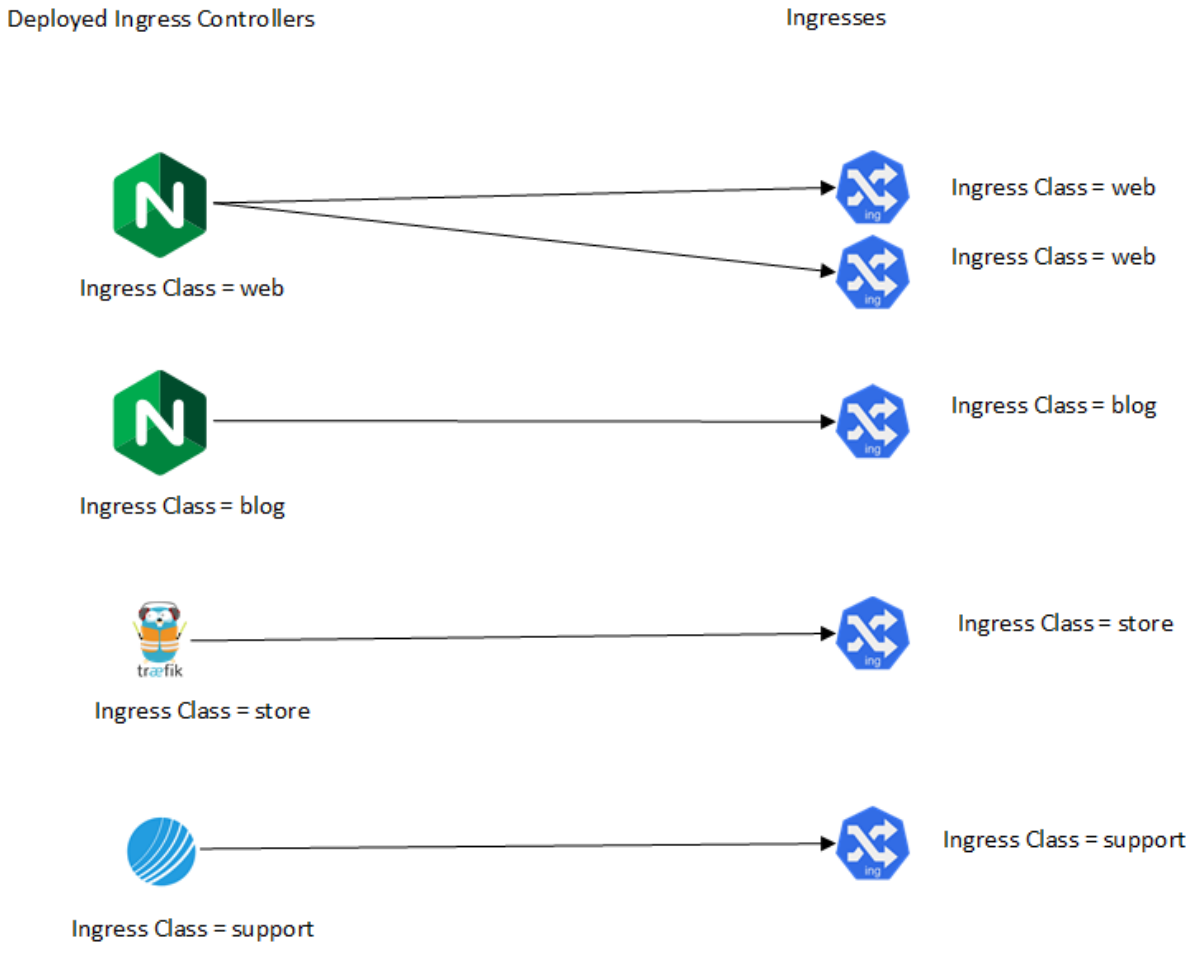
K8S官方Nginx提供了哪些Lua模块?

- lua-resty-upload
- lua-resty-string
- lua-resty-balancer
- lua-resty-core
- lua-cjson
- lua-resty-cookie
- lua-resty-lrucache
- lua-resty-lock
- lua-resty-dns
- lua-resty-http
- lua-resty-http
- lua-resty-redis
- lua-resty-ipmatcher
- lua-resty-global-throttle

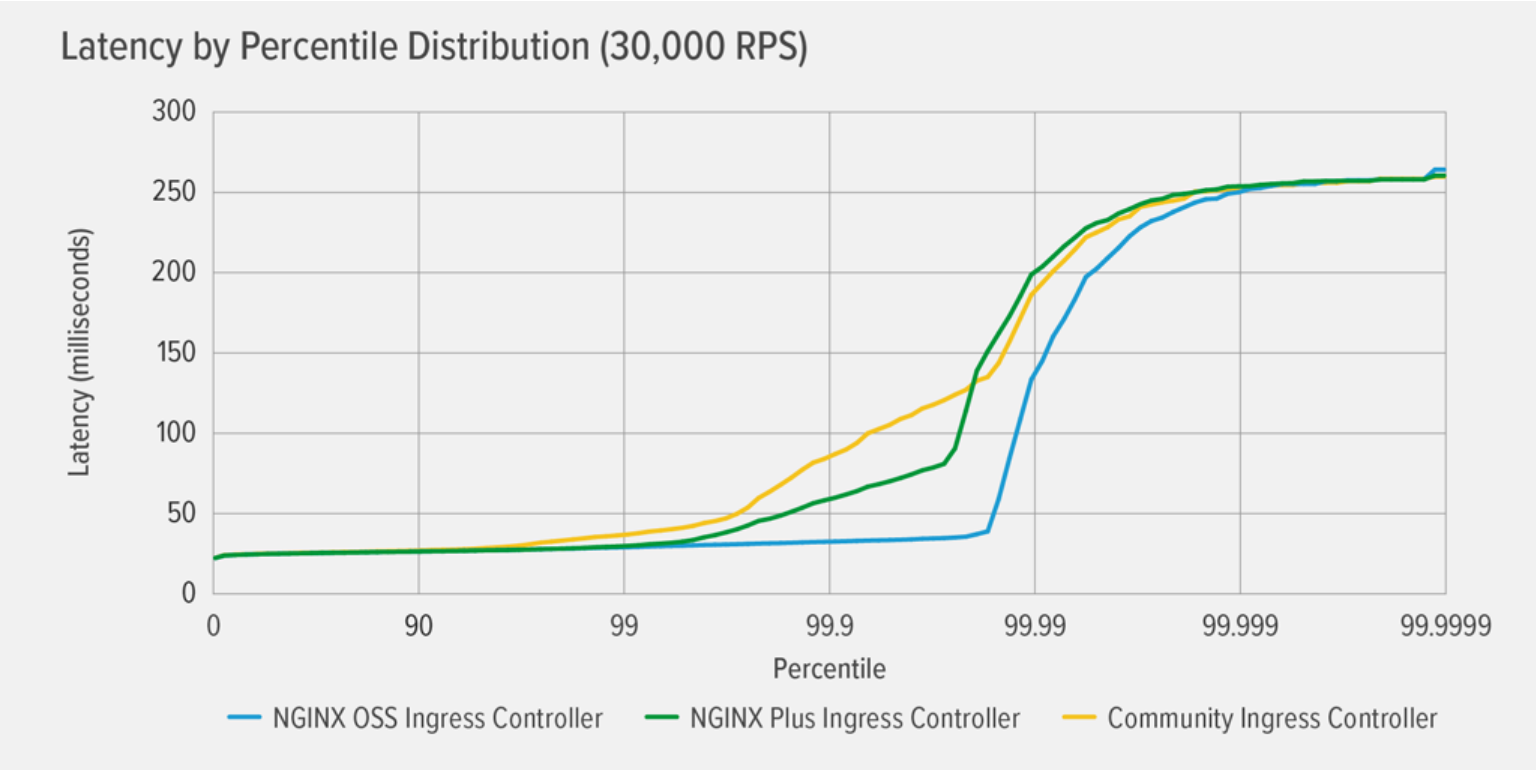
Nginx官方Controller有哪些新功能?

Ingress Controller横向对比

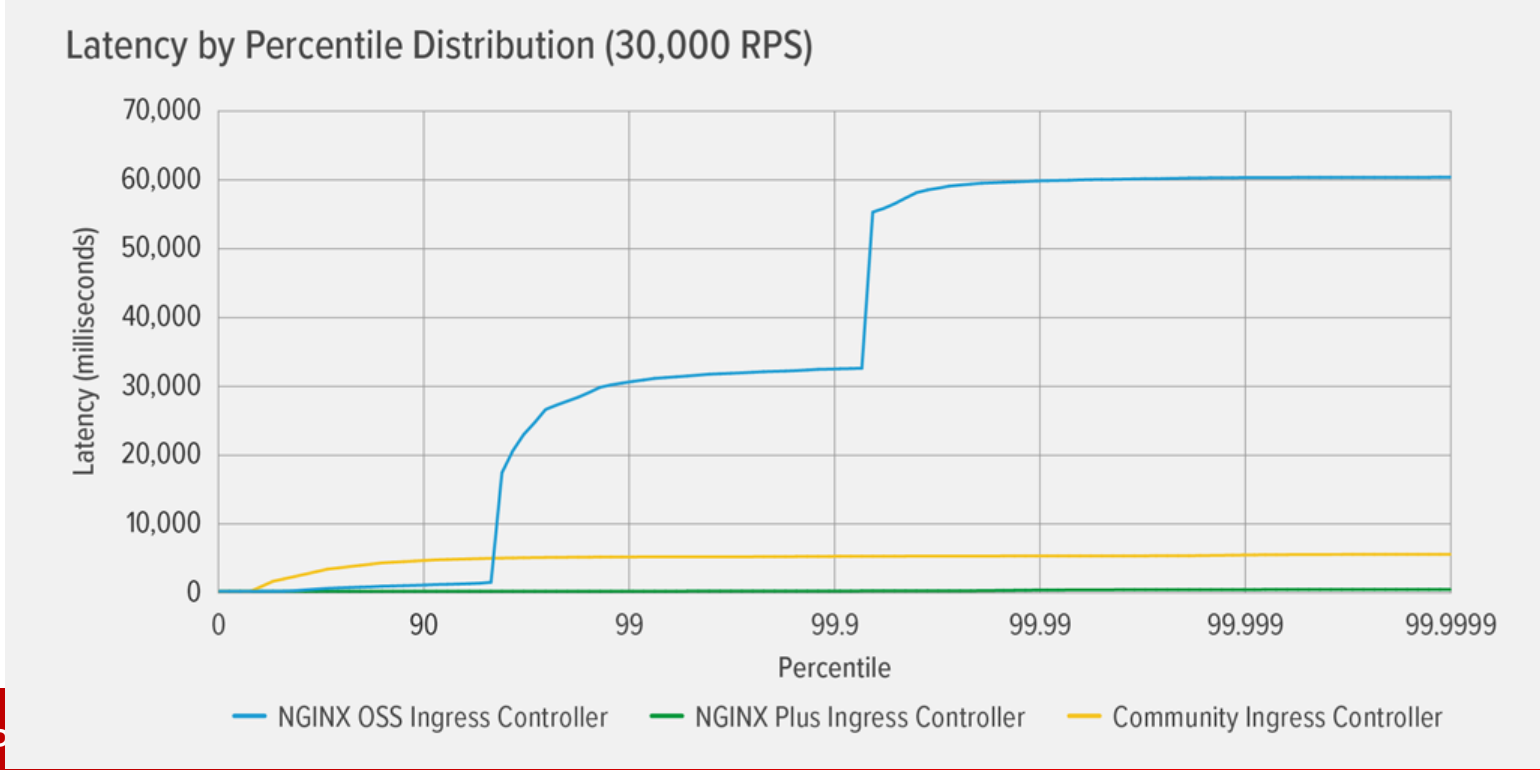
04


Deployed Ingress Controllers
Ingresses
Ingress Class = web
Ingress Class = web
Ingress Class = web
Ingress Class = blog
Ingress Class = blog
Ingress Class = store
træfik
Ingress Class = store
Ingress Class = support
Ingress Class = support

静态部署测试：时延对比


Latency by Percentile Distribution (30,000 RPS)
Latency (milliseconds)
300
250
200
150
100
50
0
0
90
99
99.9
99.99
99.999
99.9999
Percentile
NGINX OSS Ingress Controller
NGINX Plus Ingress Controller
Community Ingress Controller

动态部署测试：时延对比


Latency by Percentile Distribution (30,000 RPS)
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
Latency (milliseconds)
0
90
99
99.9
99.99
99.999
99.9999
Percentile
NGINX OSS Ingress Controller
NGINX Plus Ingress Controller
Community Ingress Controller

GOPS 2021
shenzhen

https://www.taohui.pub

Thanks

高效运维社区

开放运维联盟

荣誉出品