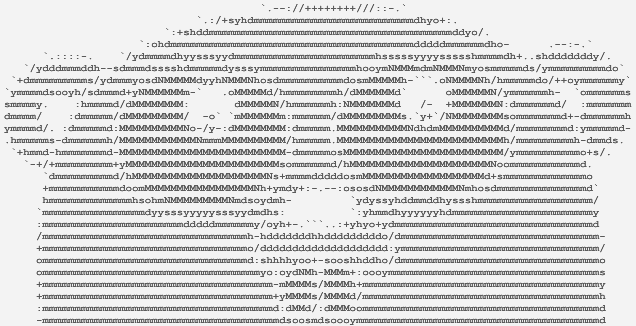
A bit faster logging

Hacking K8S logging on GKE

Golang Warsaw #32
15 Jan 2020



Tl;dr

Me, me, me

12 factor ML API in Go

“Tuning” GKE for Go apps to get x2 more bang for a buck

Go-ing upstream

A bit about myself

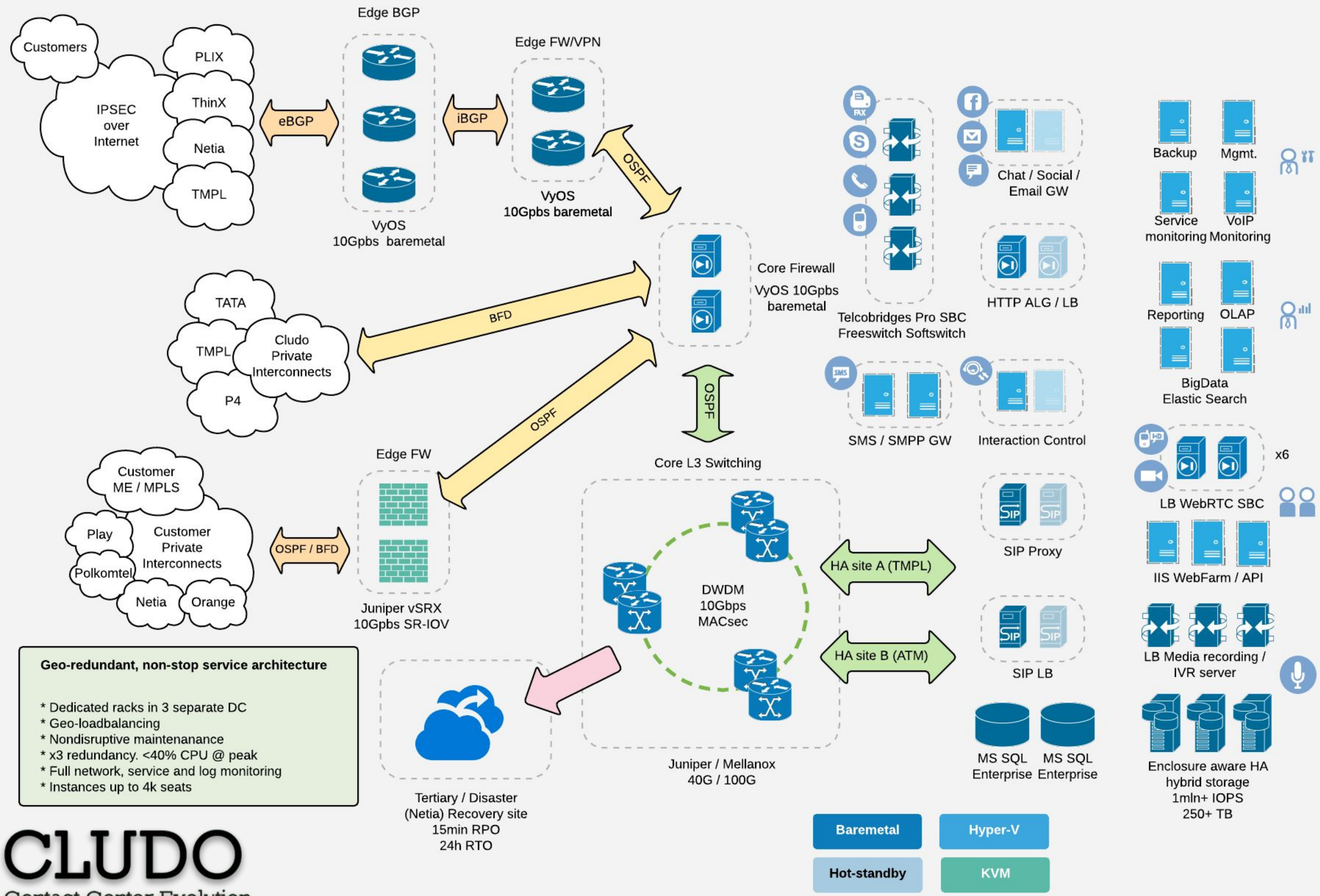
Paweł
Pierścionek
CTO @ Cludo

/in/urtho

/urtho

invalid@data.pl

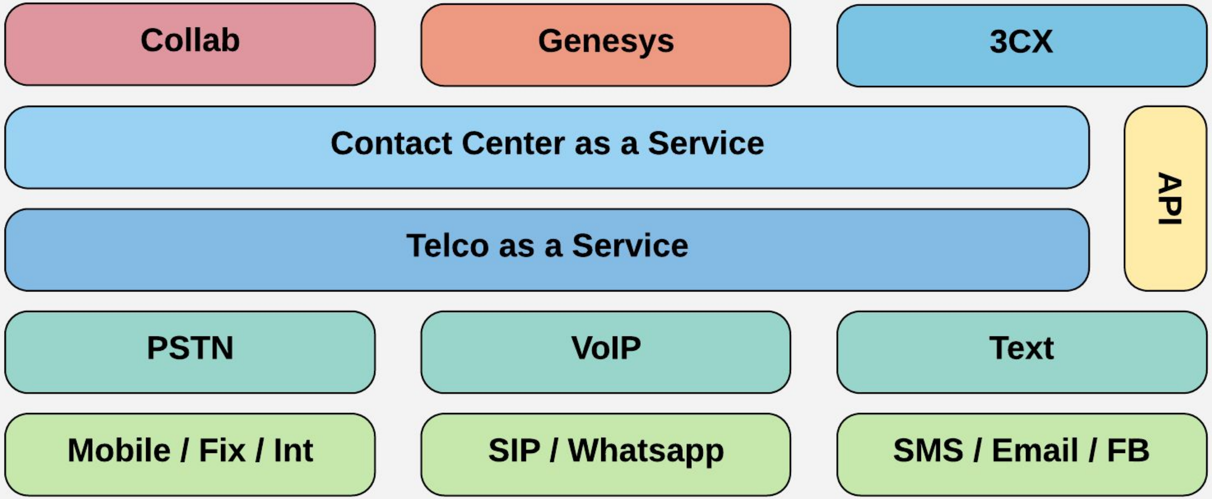
Cludo



Geo-redundant, non-stop service architecture

* Dedicated racks in 3 separate DC
* Geo-loadbalancing
* Nondisruptive maintenanance
* x3 redundancy. <40% CPU @ peak
* Full network, service and log monitoring
* Instances up to 4k seats

CLUDO
Contact Center Evolution



Collab
Genesys
3CX

Contact Center as a Service

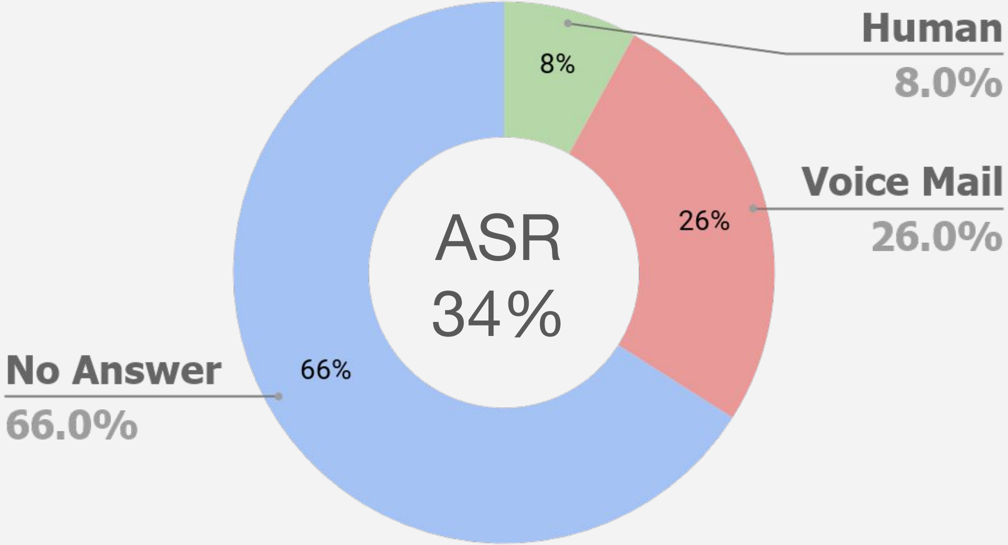
Telco as a Service

API

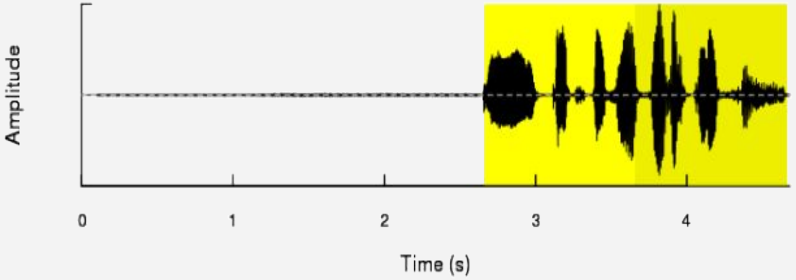
PSTN
VoIP
Text

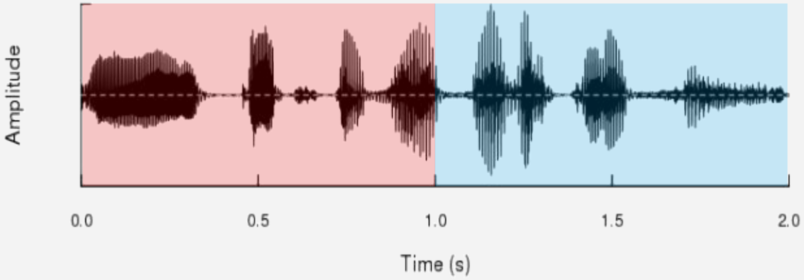
Mobile / Fix / Int
SIP / Whatsapp
SMS / Email / FB

VoiceMail issue



aiAMD - private cloud @ Cludo



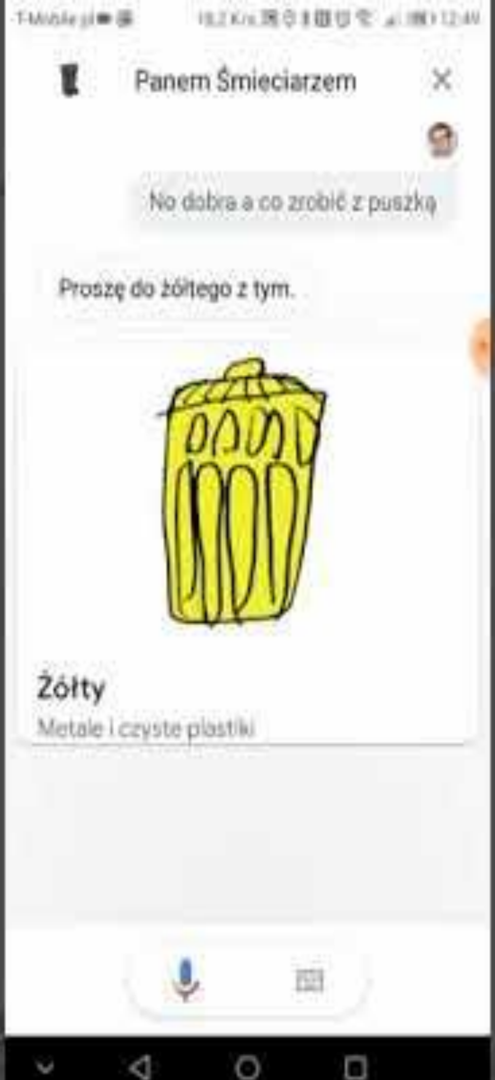


FreeSwitch module (~10k of C) - 4 phases - only models for Poland:

- Custom DSP algorithms (VAD/Tone)
- Human matching
- VoiceMail matching
- Optional (DNN-1D + RNN)

AMD - solutions in the wild (Poland)

Solution	False Positive	FN/(TP+TN+FN)
Open Source (2.5s)	< 15%	< 85%
Commercial (2.5s)	< 15%	< 95%
aiAMD B2B (1s,2s)	< 0.0002%	> 95%
aiAMD B2C (1s,2s)	< 2%	> 98%
feo.ai B2C (1s,2s)	< 1%	> 99.2%



Panem Śmieciarzem

No dobra a co zrobić z puszką

Proszę do żółtego z tym.

Żółty

Metale i czyste plastiki

Żółty

Metale i czys

Go REST word2vec category server

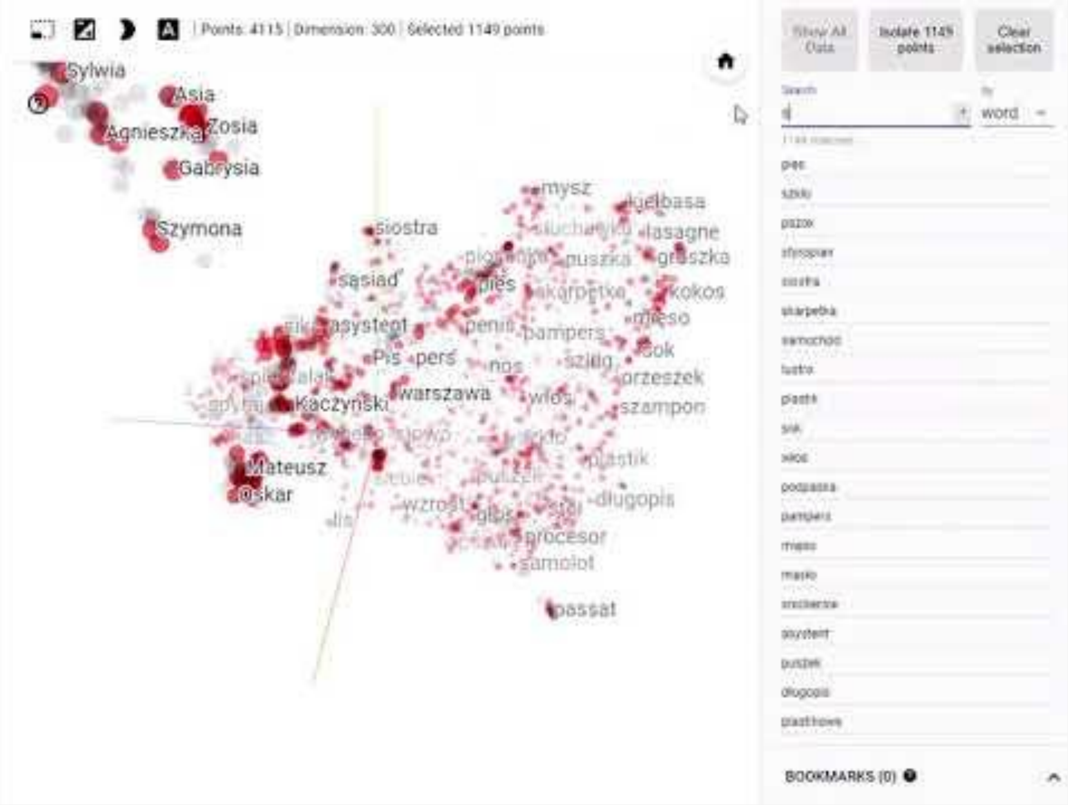
```
{ 'brown-vege': ['warzywo','warzywa','kartofel','ziemniak']},
{ 'brown-fruit': ['jabłko','banan','gruszka','pomarańcz']},
{ 'brown-fruit': ['jabłkami','bananami','gruszkami','pomarańczami']},
{ 'brown-1': ['chleb','pieczywo','bułka','bagietka']},
{ 'brown-2': ['bio','obierki','resztki','ryż']},
{ 'brown-3': ['kawa','herbata','fusy']},
{ 'meat-1': ['mięso','kości','kość','mięsne','krew']},
{ 'meat-2': ['ser','krowa','koza','owca','serek','masło']},
{ 'meat-6': ['frytki','McDonald','hamburger']},
{ 'meat-7': ['ryba','ości','tuńczyk']},
{ 'radioactive-1': ['atomowa','radioaktywny','promieniotwórczy','pluton','uran']},
{ 'pszok-danger': ['pluton','uran','ołów','rtęć','azbest','eternit']},
{ 'pszok-1': ['termometr','odczynniki','budowlane','ropę']},
{ 'battery': ['baterie','bateria','akumulator','akumulatorkami']},
{ 'electro-1': ['żarówka','jarzeniówka','świetlówka','led']},
{ 'electro-9': ['telefon','komórka','iphone','android','smartfon']},
{ 'electro-10': ['kuchenka','mikrofalówka','mikrofala','piekarnik','piec']},
{ 'black-1': ['styropian','ceramika','ceramikę','doniczka']},
{ 'black-2': ['zabawka','pluszak','miś','lalka']},
{ 'yellow-metal': ['metal','metale','aluminium','miedź','cynk','żelazo','chrom','drut']},
{ 'yellow-metal': ['sztućce','gwóźdź','śruba','klucze']},
{ 'yellow-metal': ['łyżka','widelec','nóż','sztućce']},
{ 'yellow-plastic': ['worek','torebki','torba','reklamówka','folia']},
{ 'green-glass': ['szkło','szklaną','słoiki','słoik','wódce']},
{ 'wear-1': ['ubrania','spodnie','koszula','bluzka','buty','plecak','kapeć','pasek']},
{ 'wear-2': ['tkanina','ręcznik','bawełna','frote']},
{ 'wear-3': ['koc','poduszka','prześcieradło','kołdra','jasiek','pidżama']},
```

```
{ 'drugs-1': ['lekarstwo','leki','lekarstwie','lekami','apap','ibuprom']},
{ 'drugs-2': ['narkotyki','amfetamina','mefedron','dragi']},
{ 'drugs-3': ['strzykawka','igła','narkotyki']},
{ 'drugs-4': ['aspiryna','apap','ibuprom','bepanthen']},
{ 'police-1': ['bomba','karabin','pistolet','broń','naboje','petarda']},
{ 'police-2': ['petarda','rakieta','pocisk','bomba','pistolet','wybuch']},
{ 'policeman': ['policjant','strażak','oficer','żandarm','ochroniarz','gliniarz','komendant']},
{ 'human-1': ['dziecko','człowieka','kobietę','brata','siostrę','teściową','noworodka']},
{ 'human-2': ['osoba','dziecko','człowiek','dorosły','kobieta','brat','siostra','teściowa']},
{ 'human-3': ['Paweł','Krzysztof','Agata','Bartek','Dorota','Mikołaja','Ola']},
{ 'large-1': ['biurko','szafa','stół','krzesło','kanapa','sofa','drzwi','obraz','lustro','materac']},
{ 'large-2': ['deska','grabie','łopata','miotła','dywan']},
{ 'large-3': ['duży','długi','ogromny','kolosalny','wysoki']},
{ 'large-4': ['altana','dom','budynek','balkon','szopa']},
{ 'scrap-1': ['auto','złom','samochód','puchar','radiowóz']},
{ 'scrap-2': ['felgi','alufelgi','koła','samochód','ciężarówkę']},
{ 'money-1': ['pieniądze','złotówki','kasa','mamona','dolary','złoto']},
{ 'sell-1': ['szlachetne','ametyst','diament','szafir','rubin','bursztyn']},
{ 'hygiene-1': ['pieluszka','podpaska','pampers']},
{ 'hygiene-3': ['prezerwatywa','kondom','kondon']},
{ 'religion': ['dekalog','przykazanie','religia','papież','ksiądz','kościół']},
{ 'toilet-1': ['wino','piwo','alkohol','nalewka']},
{ 'underwear-1':
['majtki','gacie','pończochy','podkoszulek','slipy','kalesony','stanik','biustonosz']},
{ 'magic-1': ['różdżka','magia','czary','czarownica','czarodziej']},
{ 'noseg-1': ['murzyn','chińczyk','indianin','rosjanin','ukraińca','menel']},
{ 'noseg-2': ['menel','bezrobotny','murzyn']},
{ 'head': ['ułamki','nauka','matematyka','historia','liczby','logarytmy','wzory','naukę']},
```

Go REST category server

```
{ 'app-exp': ['aplikacj[aeę]','aplikacjami','program.*komputer','(?:^|\\s)apkę']},
{ 'askflat-exp': ['ziem.*płaska']},
{ 'askicecream-exp': ['zrobisz loda','robić loda','zrób loda']},
{ 'askname-exp': ['(?:^|\\s)imię$','jak.*nazywasz','jak masz na imię','ja ci na imię','jak na imię.*(ci|tobie|śmiecia)']},
{ 'askwhen-exp': ['kiedy']},
{ 'askwhy-exp': ['(?:^|\\s)czemu','(?:^|\\s)dlaczego','(?:^|\\s)czy są']},
{ 'askwife-exp': ['czy.* żonę','masz żonę','czy.*pedał','kochasz mnie','czy.*męża','masz męża','czy.*żonaty']},
{ 'card-exp': ['kart[aę].*płatni','kart.*VISA','kart.*MASTERCARD','kart.*debet','kart.*kredytow','kartę SIM']},
{ 'china-exp': ['chińskie','chińską']},
{ 'cmon-exp': ['.{70,}']},
{ 'deadanimal-exp': ['zdechł','zwierzę','zwierzątko','mart.* psa','mart.* kota','mart.* zwie.*']},
{ 'explain-exp', fun: 'explain': ['(?:^|\\s)czemu$','(?:^|\\s)dlaczego$','wytłumacz$','wyjaśnij$']},
{ 'focus-exp': ['(?:^|\\s)czy ','(?:^|\\s)dlaczego ','(?:^|\\s)ile ','^która.*godzina']},
{ 'fuck-exp': ['kurwa','pojebie']},
{ 'magic-exp': ['czarodziejsk']},
{ 'noseg-exp': ['żyda?(?=$|\\s)','żydów']},
{ 'noway-exp': ['(?:^|\\s)pedała','(?:^|\\s)płód(?=$|\\s)','(?:^|\\s)płody(?=$|\\s)']},
{ 'police-exp': ['człowieka','(?:^|\\s)ciało','trupa','granat(?=$|\\s)','(?:^|\\s)szkielet']},
{ 'politics-exp': ['prezydenta','premiera','polityka','Kaczyńskiego','króla','królową','(?:^|\\s)sejm','(?:^|\\s)senat','(?:^|\\s)posła','(?:^|\\s)rząd(?=$|\\s)']},
{ 'radioactive-exp': ['promieniotwórcze','radioaktywne']},
{ 'religion-exp': ['dekalog','przykazania','religię','papieża','księdza','kościół']},
{ 'schrodinger-exp': ['Schrodinger']},
{ 'sell-exp': ['biżuteri[aę]','bankomat','sejf','lotto','LEGO']},
{ 'should-exp': ['czy (trzeba|należy|powinienem|powinnam|muszę|musimy|powinno się) u?myć']},
{ 'sorry-exp': ['wkurzasz','okłamałeś','okłaman[ya]$','popierdliło','chyba.*spadłeś']},
{ 'terminate-exp': ['króla ciszy','nie podsłuchuj'], fun:'terminate'},
{ 'trashman-exp': ['było pierwsze','kto.*rządzi','kto.*szy(?=$|\\s)']},
{ 'washer-exp': ['bru.* skarp','bru.* ubra','bru.* koszu','bru.* ubran','bru.* majt','bru.* bieli']},
```

Waste-space projector ;)



| Points: 4115 | Dimension: 300 | Selected 1149 points

pies
szkło
pszok
styropian
siostra
skarpetka
samochód
lustro
plastik
sok
włos
podpaska
pampers
mięso
masło
śmietana
asystent
puszek
długopis
plastikowe

https://wastemaster-gcp.cludo.pl/projector/

https://github.com/urtho/pansmieciarz-projector/

Vacation time :)

Learn GCP by passing the Architect exam

Burn my GCP credits by migrating a service to GCP

Join the Cloud Native cult

Learn K8S ?

Switching to Cloud Native / GCP

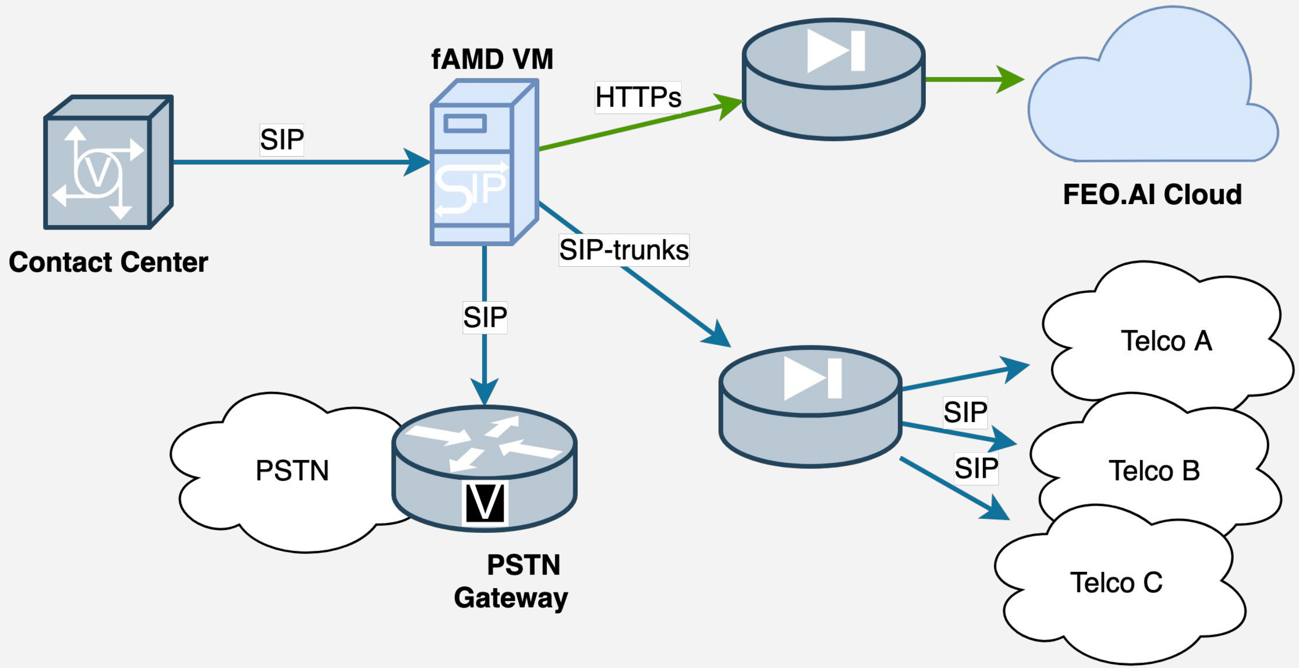
Velocity

Time to market - 2mo

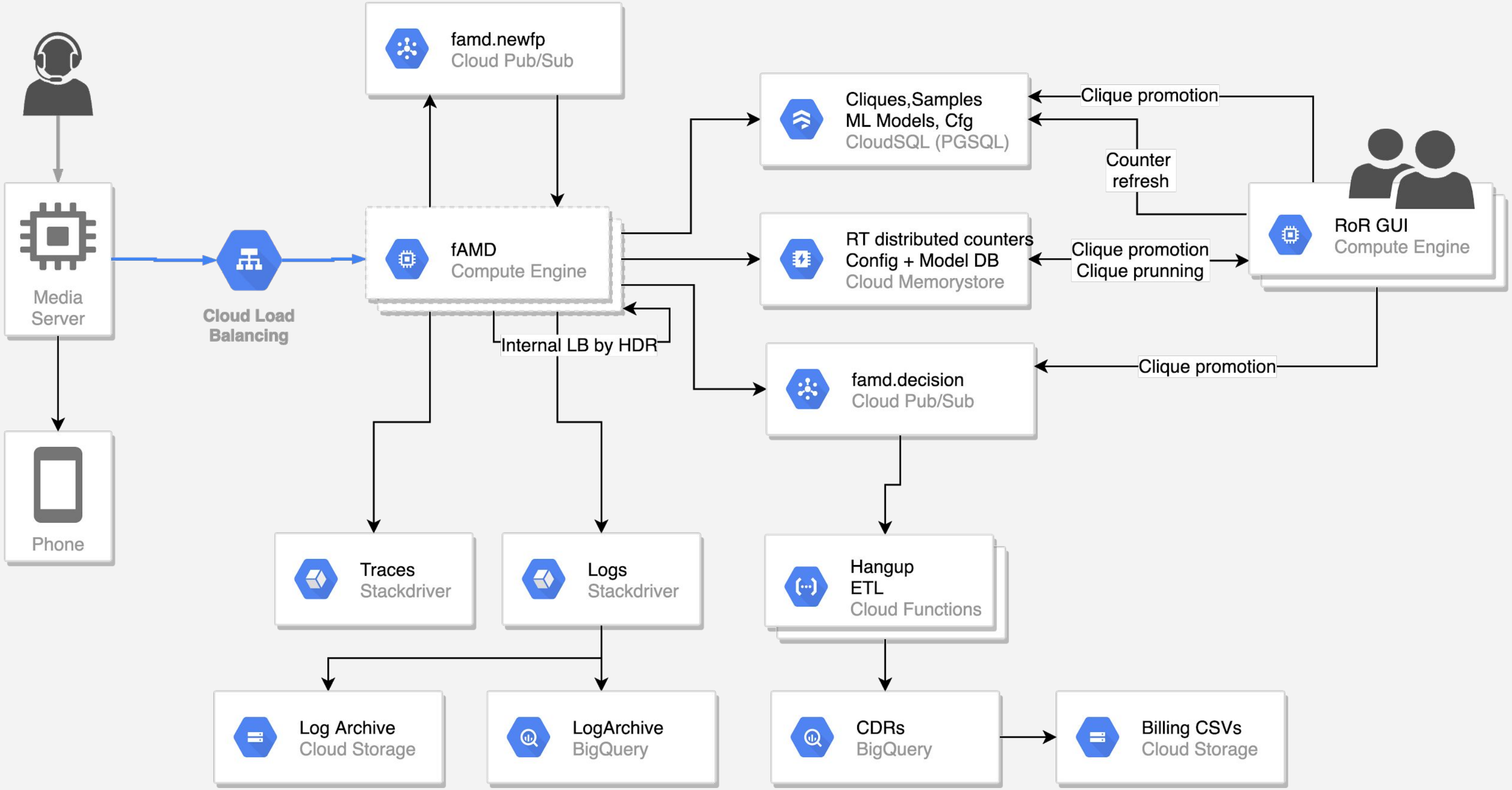
Security (Privacy/Continuity/ACLs/DiD)

Technical (E2E Latency/Debugging/LoadBalancing)

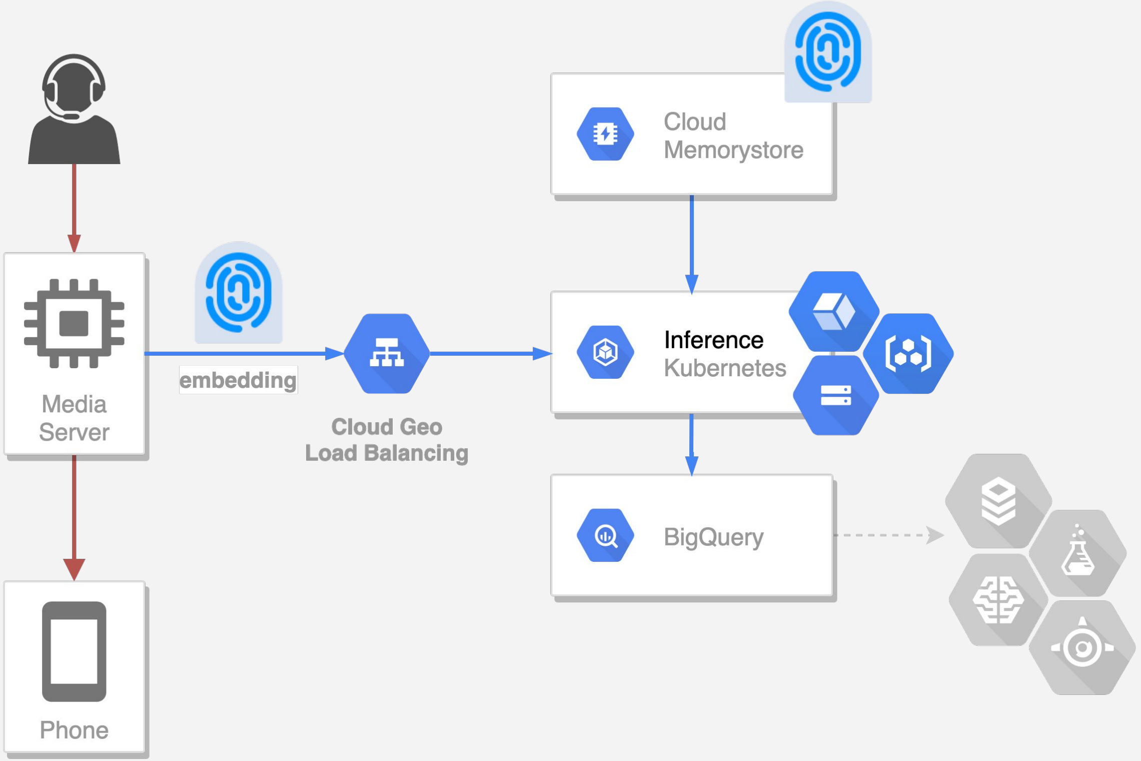
A hybrid service



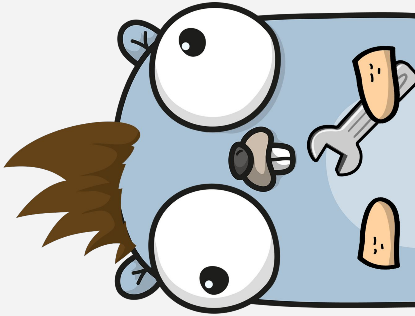
GCP the wrong way



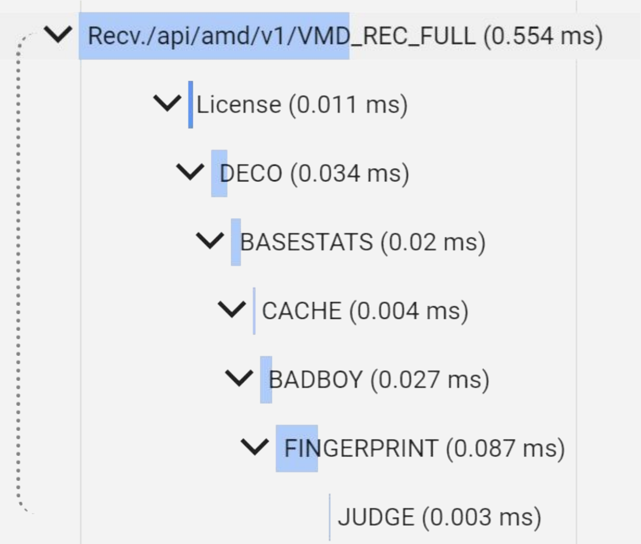
GCP the agile way



… 8 crazy weeks later



“Cloud refactoring”



```
Recv./api/amd/v1/VMD_REC_FULL (0.554 ms)
  License (0.011 ms)
   DECO (0.034 ms)
    BASESTATS (0.02 ms)
     CACHE (0.004 ms)
      BADBOY (0.027 ms)
       FINGERPRINT (0.087 ms)
        JUDGE (0.003 ms)
```

C: 10k -> 2k lines

12f Go @K8S: 0k -> 4k lines

SVCTime: 0.6ms avg , 1.2ms 95th, 3ms 99th

Audio embedding instead of PCM

SVM instead of DNN-1D + RRN

New algorithms in the decision pipeline (anti fraud)

Country specific models !!! (crowd-supervised learning)

Chaos ~~engineering~~

Security from the start

1 vCPU preemptible nodes, random restarts

Two DR runs, 1 week of mirrored run, 2 days of canary

Migrate everyone , don’t tell anyone (90k calls / day)

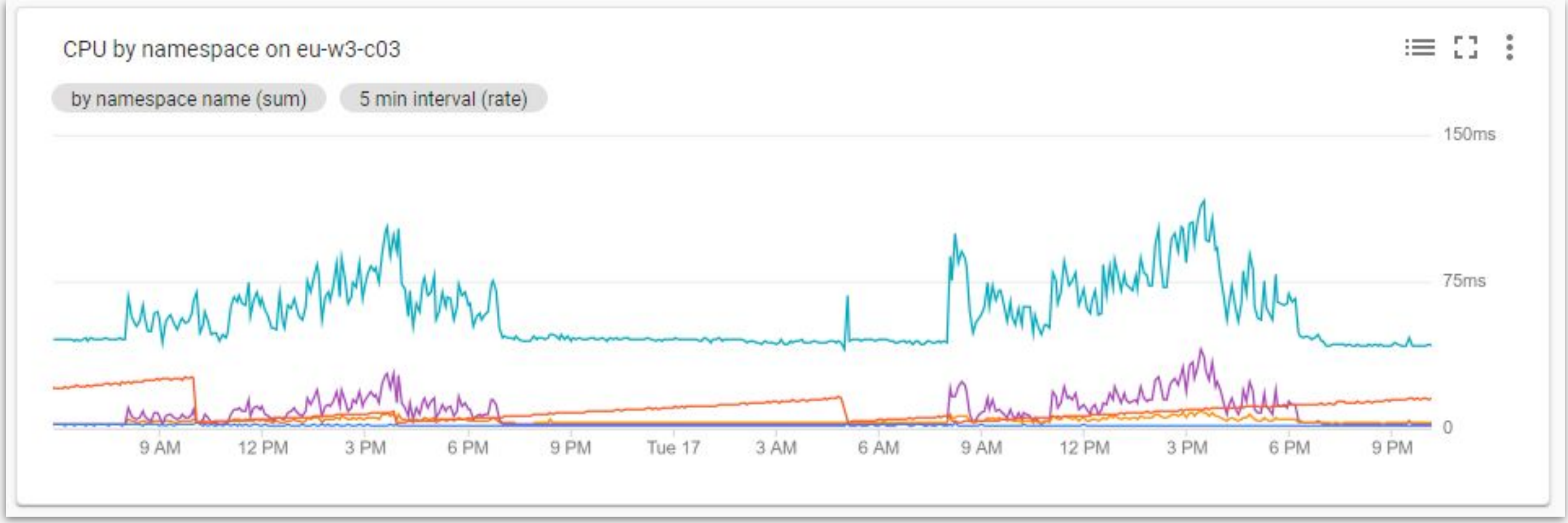
No load tests ….

= 4 minutes total downtime

Ingress

Fuentd

Scalability issues

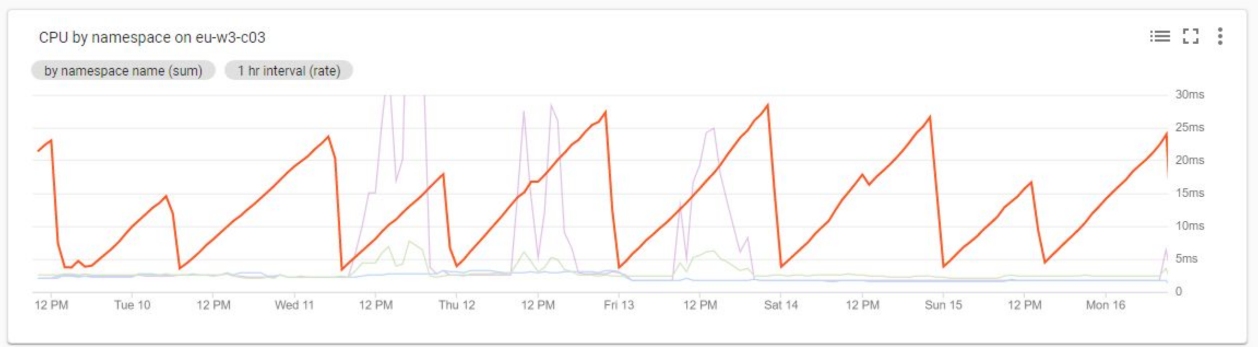


CPU by namespace on eu-w3-c03

by namespace name (sum)

5 min interval (rate)

Cert-manager 0.12

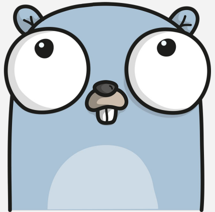


```
func wait10() {
    ticker := time.Tick(time.Second)
    for count := 1; count <= 10; count++ {
        fmt.Println(<- ticker, count, "seconds elapsed")
    }
}
```

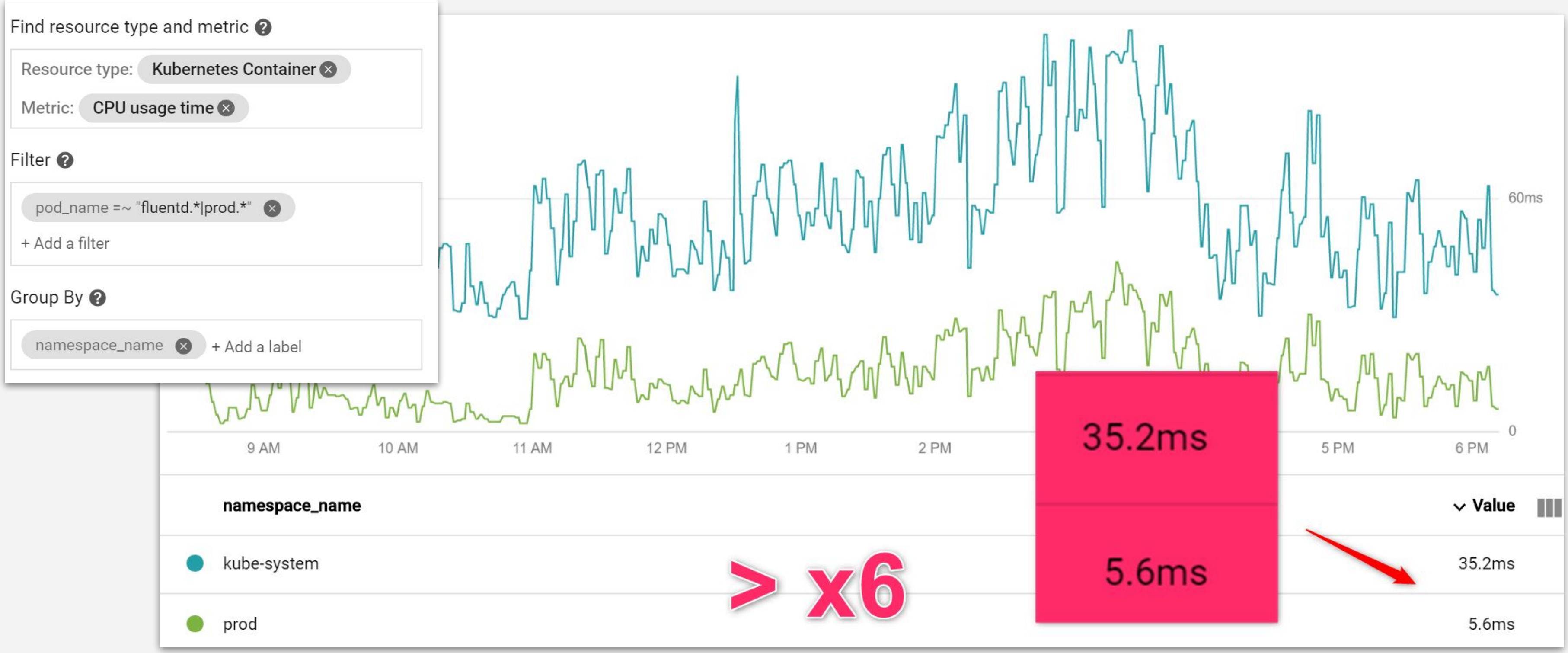
defer .Stop()

```
func wait10() {
    ticker := time.NewTicker(time.Second)
    defer ticker.Stop()

    for count := 1; count <= 10; count++ {
        fmt.Println(<- ticker.C, count, "seconds elapsed")
    }
}
```

```go
func init() {
    log.SetFormatter(lgrsdf.NewFormatter(
        lgrsdf.WithService("amdws"),
        lgrsdf.WithVersion(g.GitCommit),
    ))

    gcpid := os.Getenv("GOOGLE_CLOUD_PROJECT")
    ctx := context.Background()
    if credentials, err := google.FindDefaultCredentials(ctx); err == nil {
        gcpid = credentials.ProjectID
    }

    log.SetLevel(log.DebugLevel)
    log.WithFields(log.Fields{"action":"boot","ver": g.GitCommit}).Warnf("Booting in debug mode")

    exporter, err := stackdriver.NewExporter(stackdriver.Options{
        ProjectID: gcpid,
    })

    if err != nil {
        log.Fatal(err)
    }

    trace.RegisterExporter(exporter)
    trace.ApplyConfig(trace.Config{DefaultSampler: trace.AlwaysSample()})
}
```



Logging CPU = 6x APP CPU



Find resource type and metric ?

Resource type: Kubernetes Container ⓧ

Metric: CPU usage time ⓧ

Filter ?

pod_name =~ "fluentd.*|prod.*" ⓧ

+ Add a filter

Group By ?

namespace_name ⓧ + Add a label

	namespace_name	Value
●	kube-system	35.2ms
●	prod	5.6ms

35.2ms

5.6ms

> x6

Fluentd vs Fluent-bit

	Fluentd	Fluent-bit
Language	C & Ruby	C
Memory	40MB	0.5MB

K8S res	`requests:`
	`cpu: 100m`
	`memory: 200Mi`
		`requests:`
		`cpu: 5m`
		`memory: 10Mi`

Plugins	>700	> 45
Stackdriver	Yes / supported	Yes / not-usable

JSON Transmogrification

```
{
 "jsonPayload": {
  "context": {
   "data": {
    "action": "statsupdskip",
    "inactive": 0,
    "model": "100f20ms:HU:default",
   }
  },
  "kubernetes": {
   "container_hash": "eu.gcr.io/feoai-253119/feo/amdws@sha256:...",
   "container_image": "eu.gcr.io/feoai-253119/feo/amdws:latest",
   "container_name": "demo-amdws",
   "labels": { "app": "demo-amdws", "pod-template-hash": "7787c9577b" },
   "namespace_name": "demo"     ,
   "pod_id": "51646286-29cd-11ea-8143-4201ac100104"     ,
   "pod_name": "demo-amdws-7787c9577b-bz7t9"     ,
  },
  "log": "{\"timestamp\":\"2020-01-02T17:10:01Z\",\"message\":\"Model stats update
skipped\", \"severity\":\"DEBUG\", \"context\": {\"data\":
{\"action\":\"statsupdskip\",\"inactive\":0,\"model\":\"100f20ms:IT:default\"}}}" ,
  "message": "Model stats update skipped"    ,
  "severity": "DEBUG"    ,
  "stream": "stderr"    ,
  "time": "2020-01-02T17:10:01.673407221Z"    ,
  "timestamp": "2020-01-02T17:10:01Z"    ,
 },
 "logName": "projects/feoai-253119/logs/demo-amdws-7787c9577b-bz7t9.log",
 "timestamp": "2020-01-02T17:10:01.673407221Z"
}
```

```
{
 jsonPayload: {
  context: {
   data: {
    action: "statsupdskip"
    inactive: 0
    model: "100f20ms:HU:default"
   }
  }
  message: "Model stats update skipped"
  timestamp: "2019-12-29T20:25:14Z"
 }
 labels: {
  k8s-pod/app: "demo-amdws"
  k8s-pod/pod-template-hash: "7b9cd645c4"
 }
 logName: "projects/feoai-253119/logs/stderr"
 resource: {
  labels: {
   cluster_name: "feo-eu-w3-c03"
   container_name: "demo-amdws"
   location: "europe-west3"
   namespace_name: "demo"
   pod_name: "demo-amdws-7b9cd645c4-bgx5l"
   project_id: "feoai-253119"
  }
  type: "k8s_container"
 }
 severity: "DEBUG"
 timestamp: "2019-12-29T20:25:14.990941189Z"
}
```

Fluentd configmap

```
<filter reform.**>
  @type record_modifier
  enable_ruby true
  <record>
    _dummy_ ${if record.is_a?(Hash) && record.has_key?('kubernetes') &&
record['kubernetes'].has_key?('labels') && record['kubernetes']['labels'].is_a?(Hash); then;
record["logging.googleapis.com/labels"] = record['kubernetes']['labels'].map{ |k, v| ["k8s-pod/#{k}",
v]}.to_h; end; nil}
  </record>
  remove_keys _dummy_,kubernetes,docker
</filter>
```

Fun fact

```
~/$ curl -H "Metadata-Flavor: Google" -H "Content-Type: application/json" \
http://metadata.google.internal/computeMetadata/v1/instance/service-accounts/default/token

{"access_token":"...","expires_in":2156,"token_type":"Bearer"}
```

```
~/$ curl -0 -H "Host:" -H "Metadata-Flavor: Google" -H "Content-Type: application/json" \
http://metadata.google.internal/computeMetadata/v1/instance/service-accounts/default/token

access_token …
expires_in 2188
token_type Bearer
```

Setting up stuff

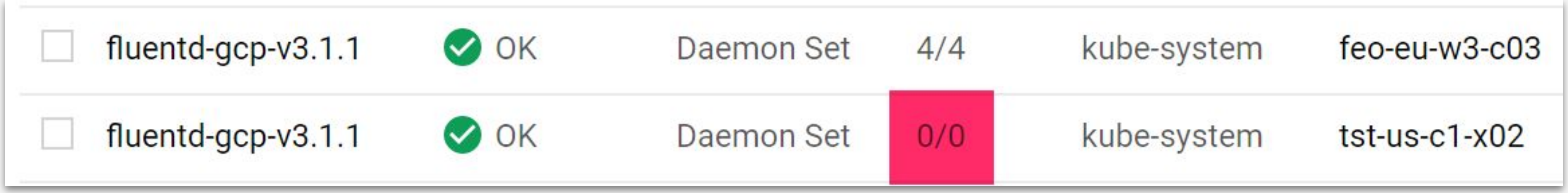
```
git clone https://github.com/urtho/fluent-bit.git
git checkout k8s_container
docker build .

git clone https://github.com/pingcap/k8s-fluent-bit-stackdriver.git
```

```
[FILTER]
    Name                kubernetes
    Match               kube.*
    Kube_URL            https://kubernetes.default.svc.cluster.local:443
    Merge_Log_Trim      On
    Merge_Log           On
    Keep_Log            Off

[OUTPUT]
    Name            stackdriver
    Match           *
    K8s_container   kubernetes
```

Blowing managed K8S up

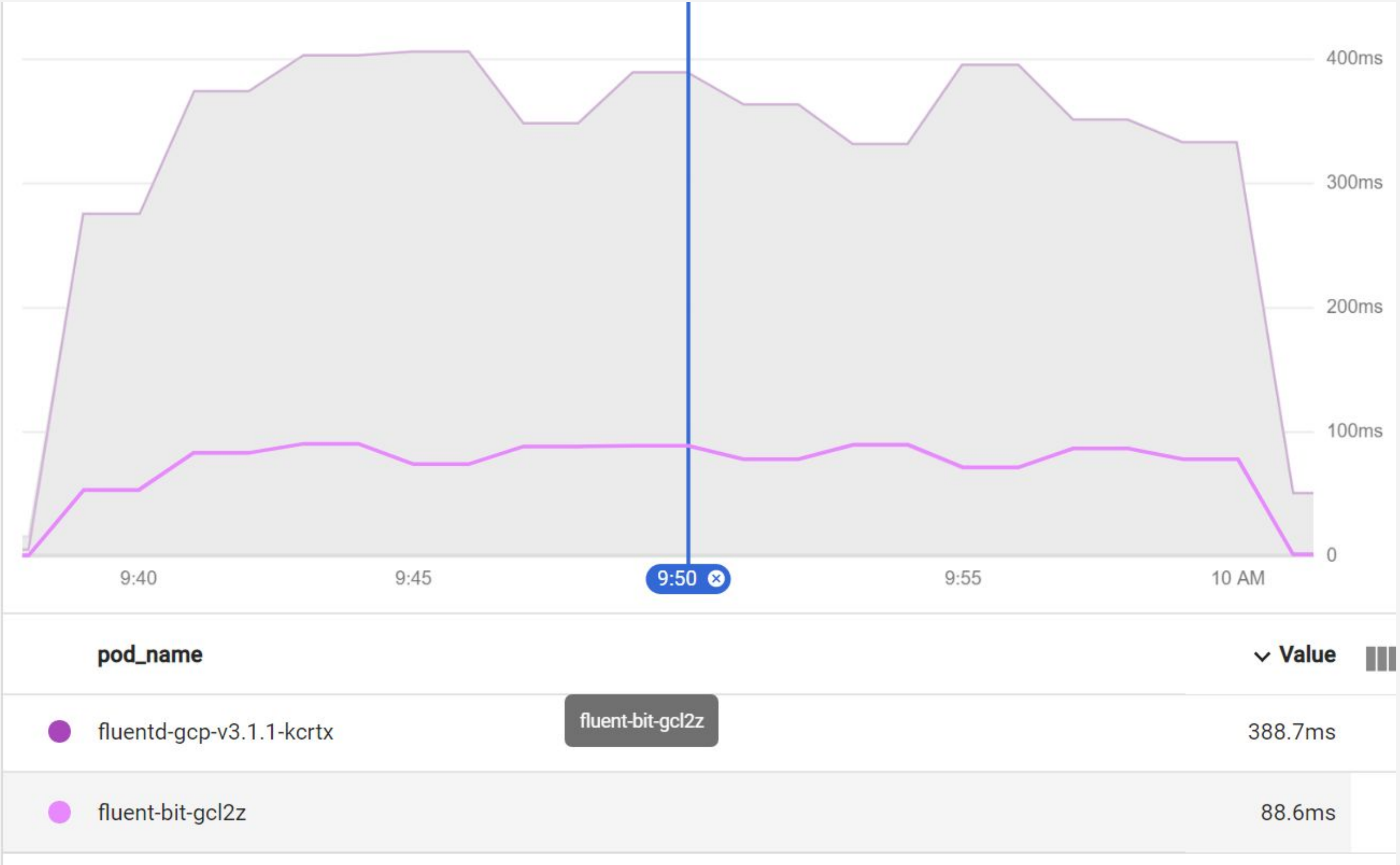
GKE >= 1.15 - all or nothing

```
kubectl -n kube-system patch daemonset fluentd-gcp-v3.1.1 \
-p '{"spec": {"template": {"spec": {"nodeSelector": {"non-existing": "true"}}}}}
```



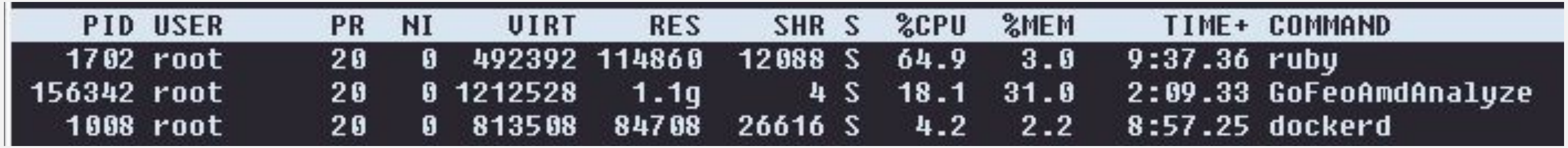
□	fluentd-gcp-v3.1.1	✓ OK	Daemon Set	4/4	kube-system	feo-eu-w3-c03
□	fluentd-gcp-v3.1.1	✓ OK	Daemon Set	0/0	kube-system	tst-us-c1-x02

Let’s test it

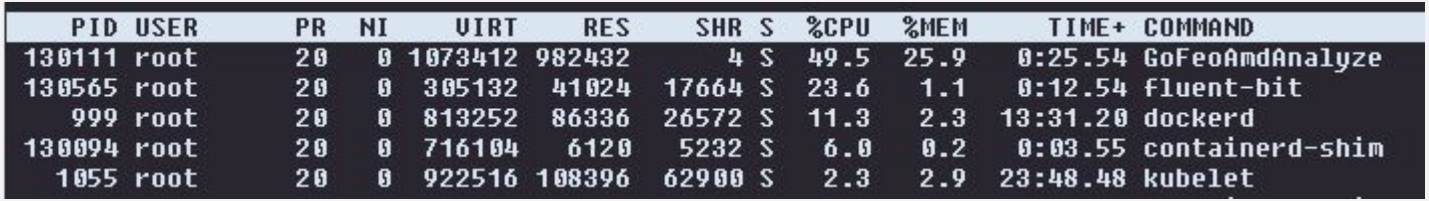


	pod_name	Value
●	fluentd-gcp-v3.1.1-kcrtx	388.7ms
●	fluent-bit-gcl2z	88.6ms

Let’s test it



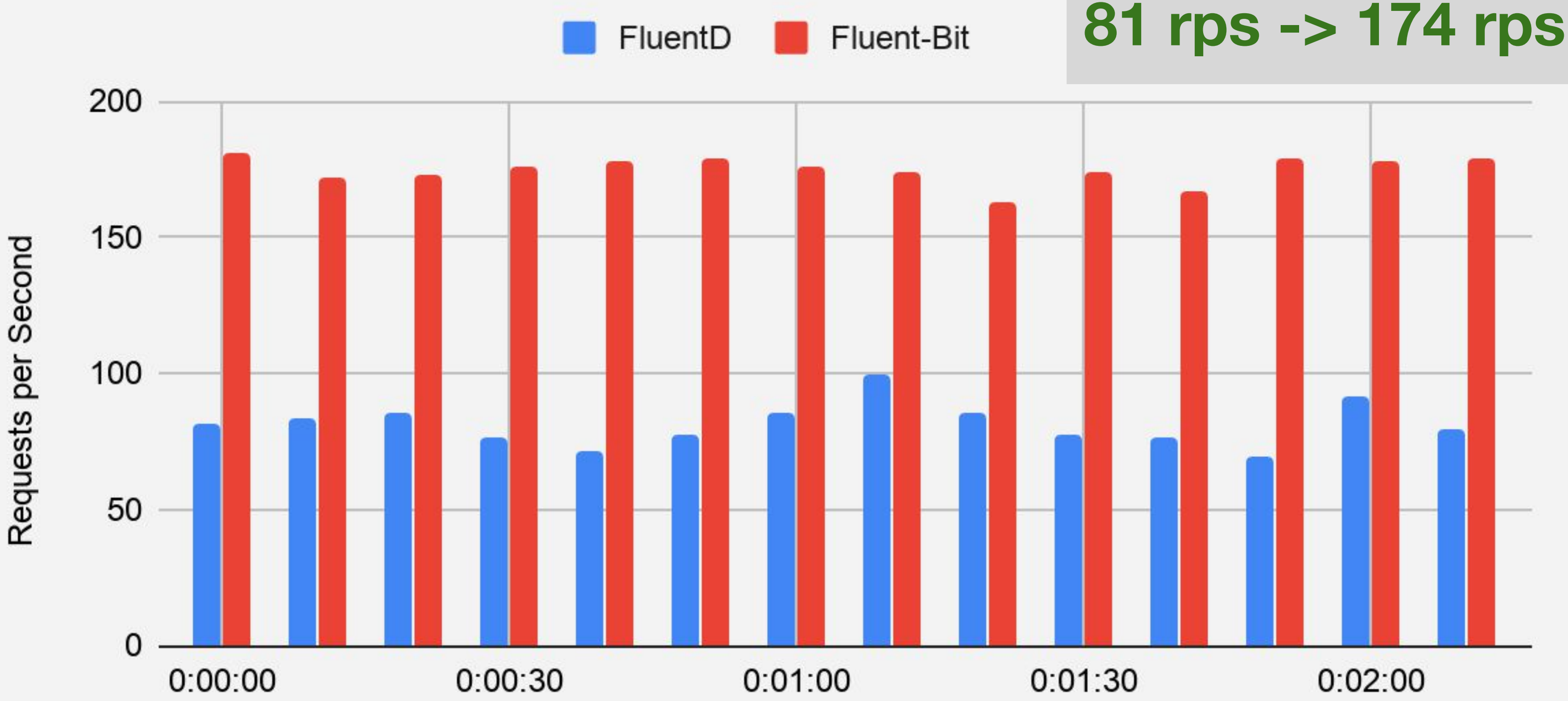
```
   PID USER      PR  NI    VIRT    RES    SHR S  %CPU  %MEM     TIME+ COMMAND
  1702 root      20   0  492392 114860  12088 S  64.9   3.0   9:37.36 ruby
156342 root      20   0 1212528   1.1g      4 S  18.1  31.0   2:09.33 GoFeoAmdAnalyze
  1008 root      20   0  813508  84708  26616 S   4.2   2.2   8:57.25 dockerd
```



```
   PID USER      PR  NI    VIRT    RES    SHR S  %CPU  %MEM     TIME+ COMMAND
130111 root      20   0 1073412 982432      4 S  49.5  25.9   0:25.54 GoFeoAmdAnalyze
130565 root      20   0  305132  41024  17664 S  23.6   1.1   0:12.54 fluent-bit
   999 root      20   0  813252  86336  26572 S  11.3   2.3  13:31.20 dockerd
130094 root      20   0  716104   6120   5232 S   6.0   0.2   0:03.55 containerd-shim
  1055 root      20   0  922516 108396  62900 S   2.3   2.9  23:48.48 kubelet
```

Experiment : 100% more workload per node

81 rps -> 174 rps



Go-ing upstream

Stackdriver makes sense...

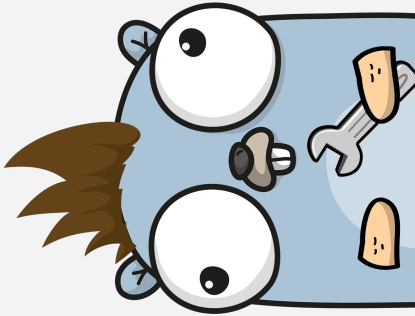
Fluent-bit >= x3 boost for K8S logging

This is a Golang meeting!

Fluent-bit has support for GO plugins

Can Go plugin perform on par with C?

… 1 weekend later



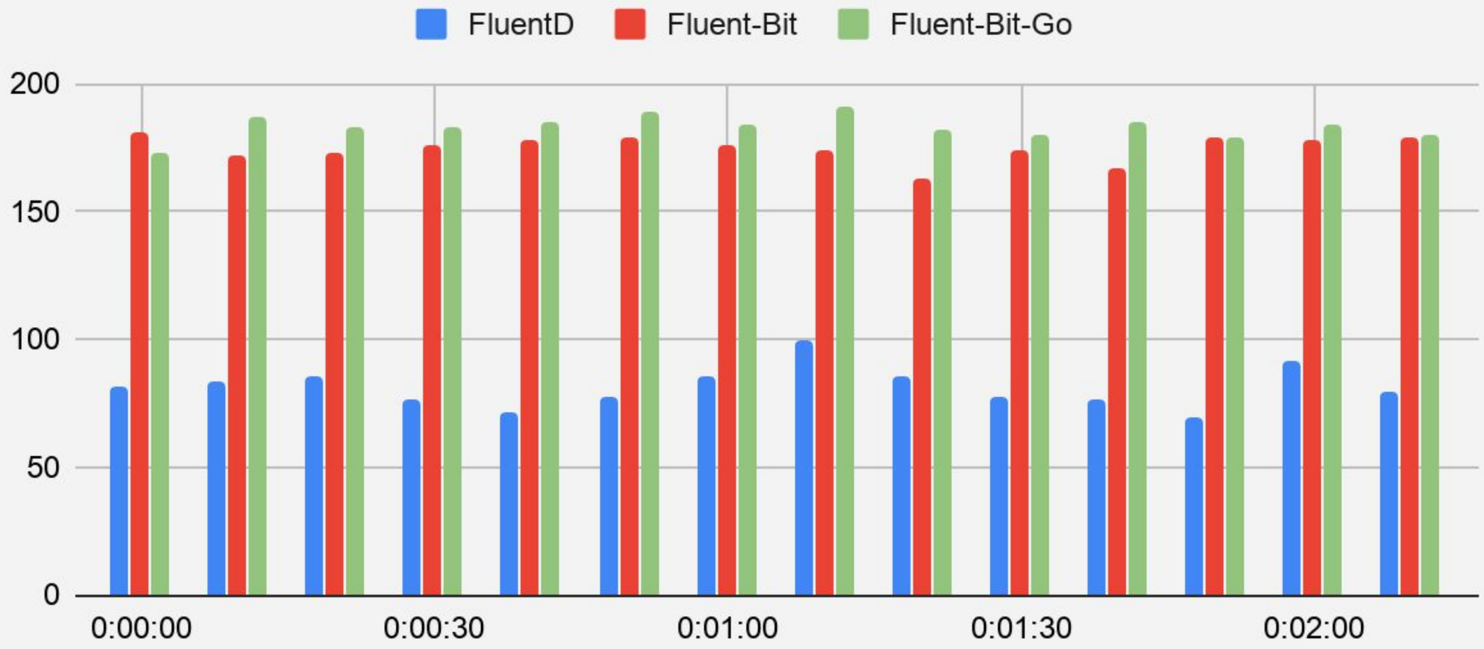
Golang fluent-bit stackdriver plugin

GRPC stream over HTTP/2

MSGPACK->MAP->JSON->PB

PoC: https://github.com/urtho/fluent-bit-out-gostackdriver

Here we go :)


FluentD
Fluent-Bit
Fluent-Bit-Go
200
150
100
50
0
0:00:00
0:00:30
0:01:00
0:01:30
0:02:00

What’s left

More metadata

Unit tests

GKE 1.15 / 1.16 / workload id. / hybrid

Manual + Medium article