Rebuilding the Crossword

with Google Cloud

O'Reilly Software Architecture Conference - NYC 2018

JP Robinson

whoami

Principal Software Engineer - Games
- NYT since 2011
- Using & evangelizing Go since 2013
  - "Using Go at The New York Times" - 2015
  - OSS: Gizmo, openapi2proto
- Google Cloud Platform
  - OSS: Marvin, drone-gae, spannerr
- Member, NYT Architecture Review Board

The NYT Crossword

- Started in 1942, daily in 1950
- Went digital in 1996, in-house in 2014
- Now available on Android, iOS, Web, Windows
- Daily (Free!) Mini crossword
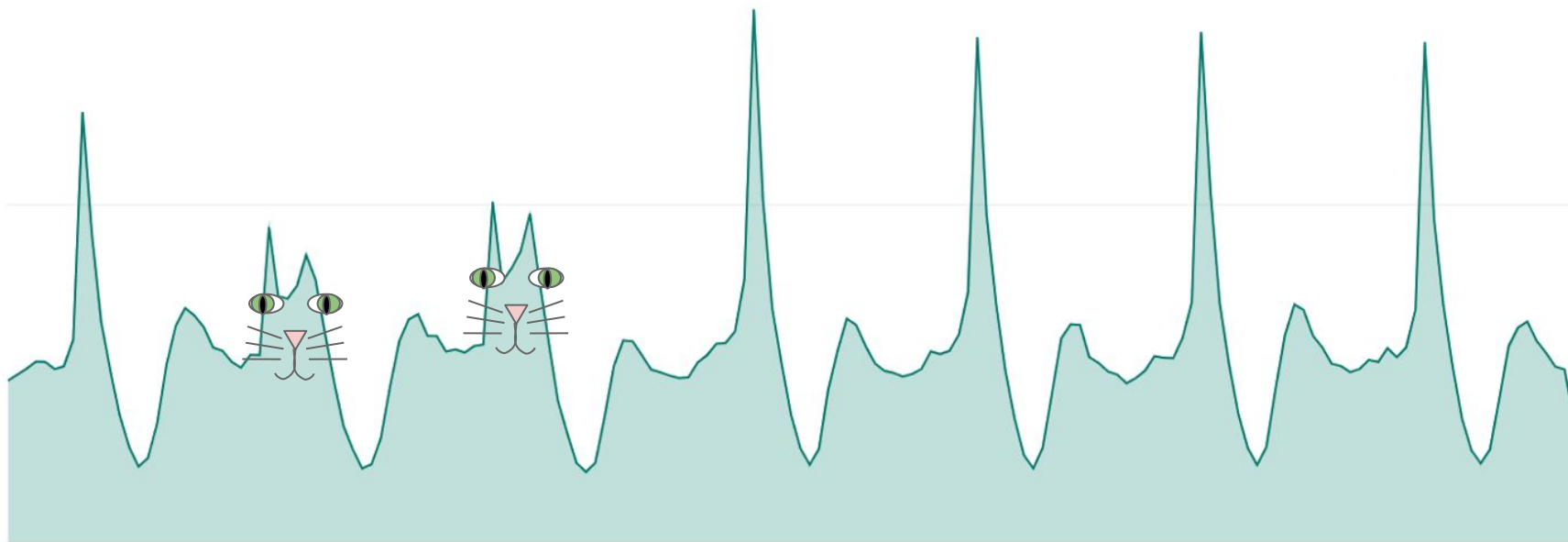- 300,000+ paid subscribers
- New puzzles are published nightly

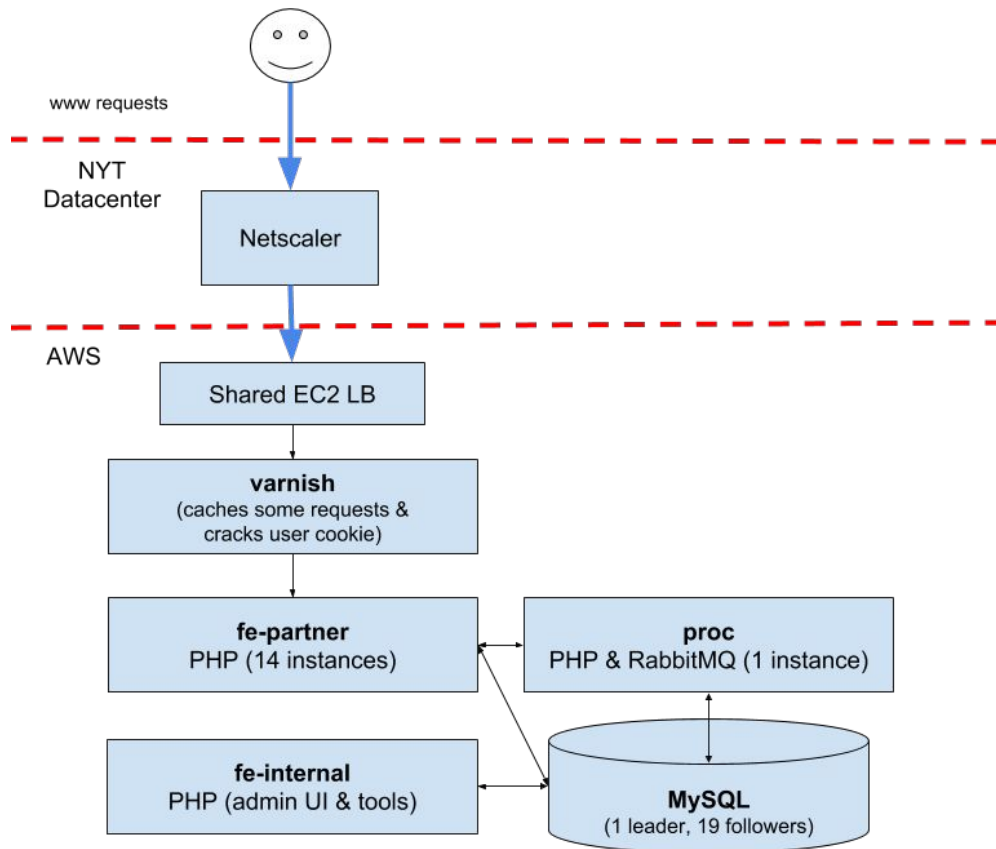# The NYT Crossword Spike

Last 7 Days of Traffic

NYT's "Legacy" Infrastructure

Pre-2016:

- Multiple data centers
  - Core news product and internal services
- AWS
  - Search, Blogs, Interactive Articles, Cooking, Games
  - Internal console (Nimbul) for accessing services
- Mix of Puppet, Ansible and custom tooling
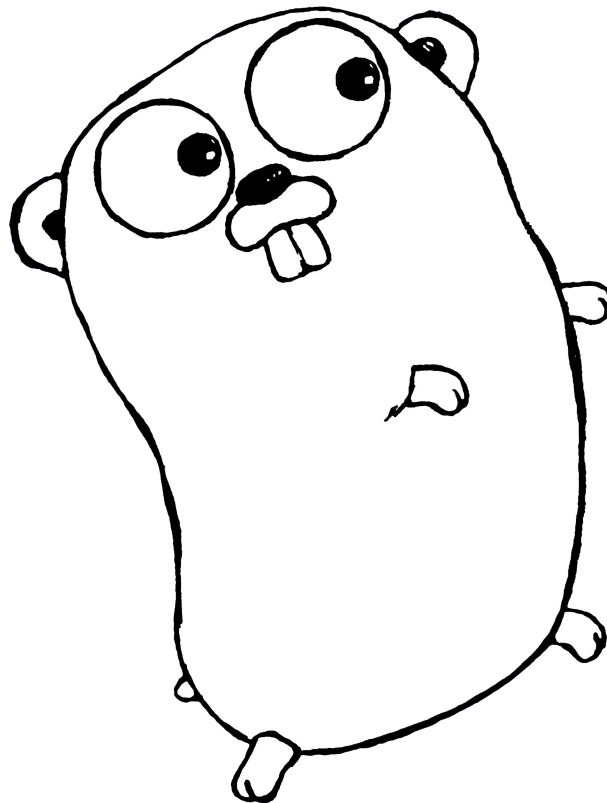
The "Legacy" Architecture

The "Legacy" Architecture

- Always scaled for peak traffic
- Adding VMs a manual process
- Deployments take 30-60+ minutes
- No log aggregation, minimal monitoring
- Internal PHP framework
- Several SPOF
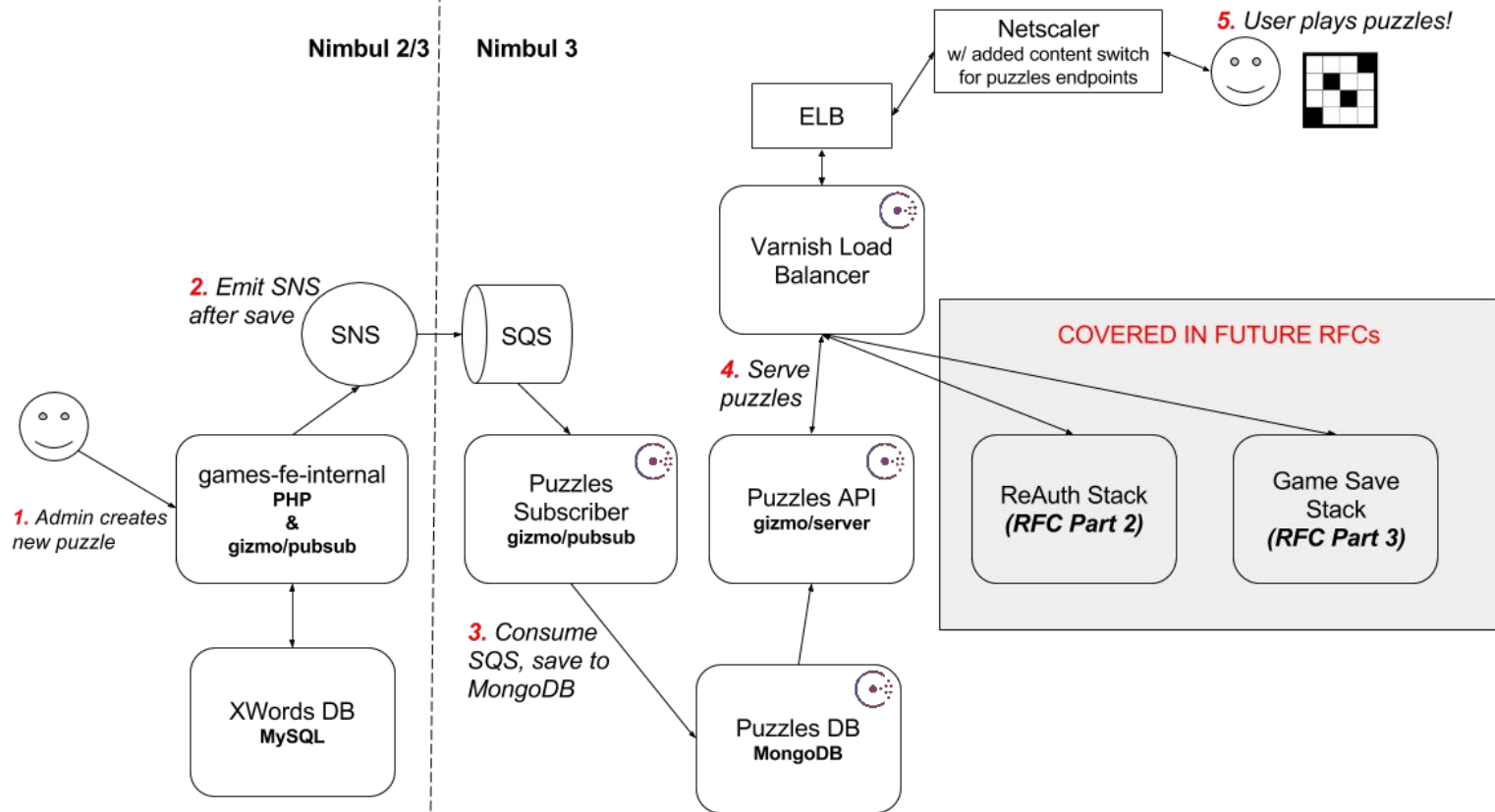
Deciding on Go

- Easy to read
- Easy to test
- Easy to deploy
- Efficient, fast and fun
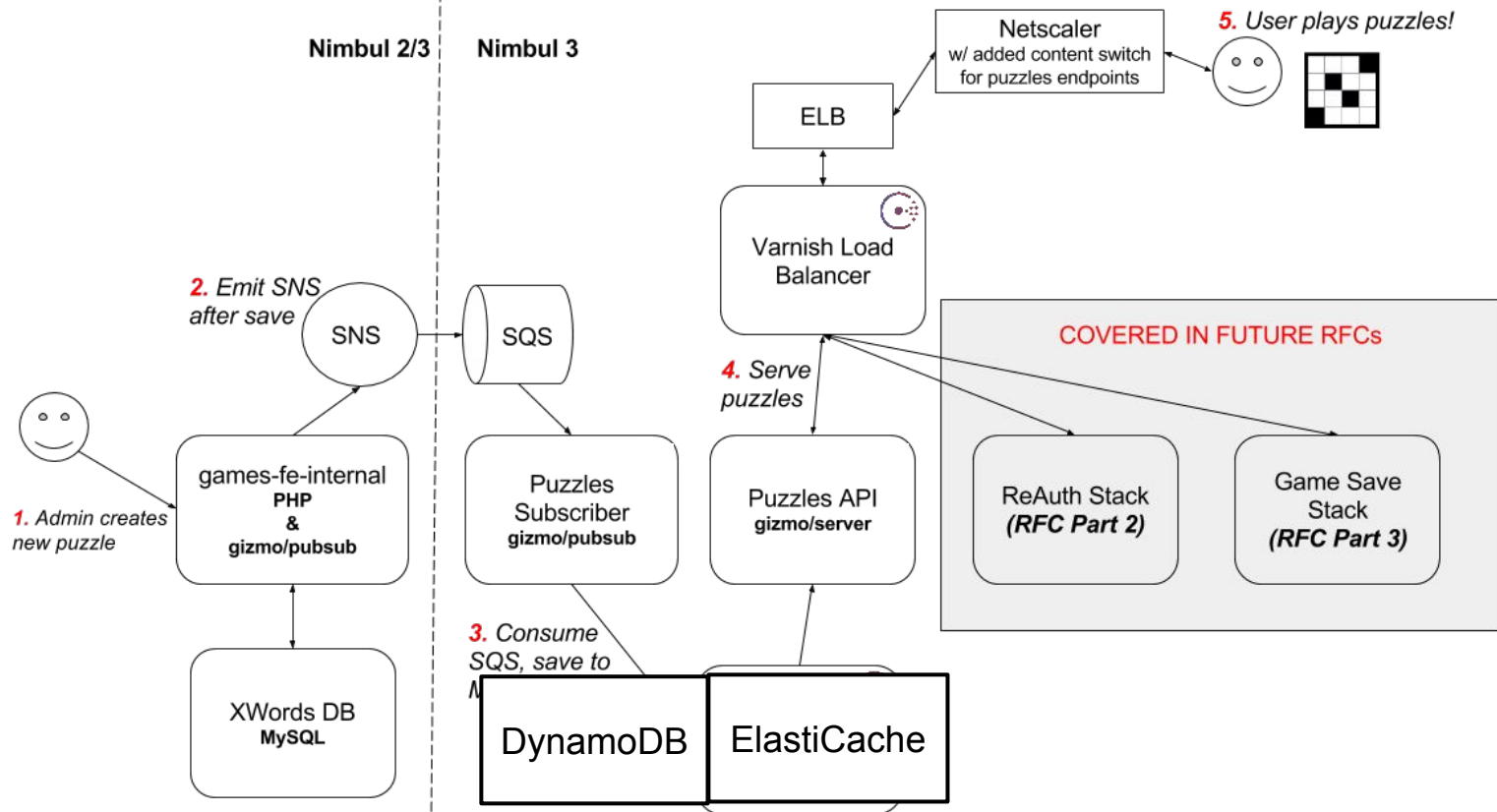- 9+ NYT teams using it

Deciding on Microservices

- Wanted to avoid any possible SPOF
  - Whatever happens, *always* serve the puzzle
- Needed to prototype and move quickly
  - Try new technology and swap out implementations
- Needed a temporary polyglot system
  - We can't rewrite it all at once!

Initial Architecture in AWS

Nimbul 2/3 | Nimbul 3

5. User plays puzzles!

Netscaler
w/ added content switch
for puzzles endpoints

ELB

Varnish Load Balancer

2. Emit SNS after save

SNS

SQS

4. Serve puzzles

COVERED IN FUTURE RFCs

1. Admin creates new puzzle

games-fe-internal
PHP
&
gizmo/pubsub

Puzzles Subscriber
gizmo/pubsub

Puzzles API
gizmo/server

ReAuth Stack
(RFC Part 2)

Game Save Stack
(RFC Part 3)

3. Consume SQS, save to MongoDB

XWords DB
MySQL

Puzzles DB
MongoDB

Initial Architecture in AWS

Nimbul 2/3   Nimbul 3

Netscaler
w/ added content switch
for puzzles endpoints

5. *User plays puzzles!*

ELB

2. *Emit SNS after save*

SNS → SQS

Varnish Load Balancer

4. *Serve puzzles*

1. *Admin creates new puzzle*

games-fe-internal
PHP
&
gizmo/pubsub

Puzzles Subscriber
gizmo/pubsub

Puzzles API
gizmo/server

COVERED IN FUTURE RFCs

ReAuth Stack
(RFC Part 2)

Game Save Stack
(RFC Part 3)

3. *Consume SQS, save to M...*

XWords DB
MySQL

DynamoDB | ElastiCache

Problems with AWS (for NYT)

- Alerting & Observability
- Service Discovery
- 0 downtime deployments
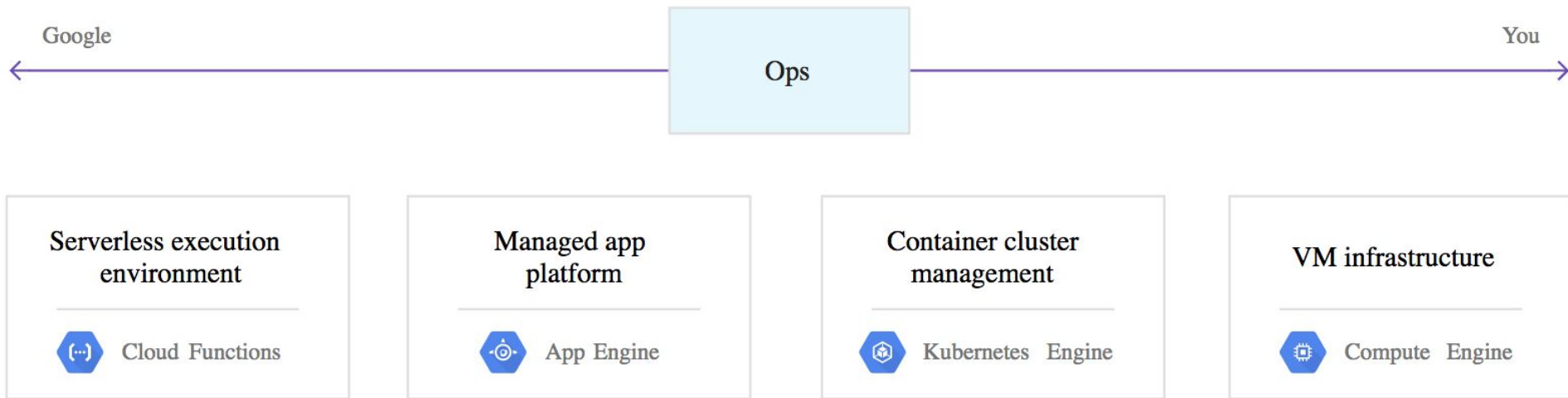- Local development

New CTO, New Cloud Vision

- Nick Rockwell joined Nov. 2015
- Moved everything to the cloud
- Introduced Fastly, saved 💰
- Expressed interest in Google Cloud Platform
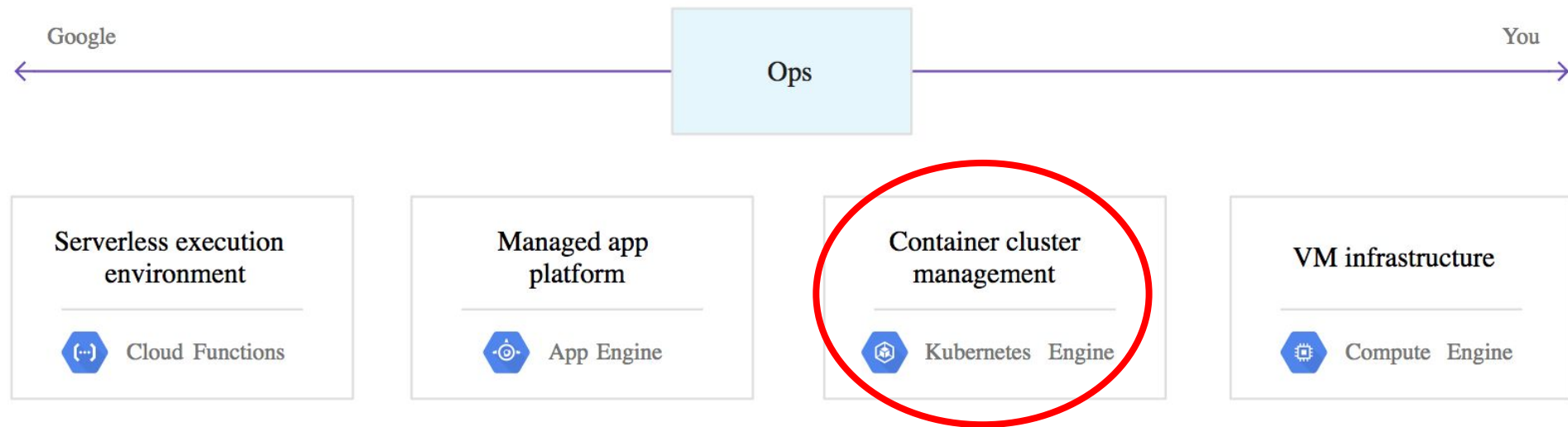- *"Serverless"*

Serverless, According to Nick

...serverless for me means any platform that provides utility while completely abstracting scaling and reliability.

- Nick Rockwell, CTO - NYTimes

Google's Compute Spectrum

Google's Compute Spectrum

Google ←——→ You

Ops

Serverless execution environment	Managed app platform	Container cluster management	VM infrastructure
Cloud Functions	App Engine	Kubernetes Engine	Compute Engine

Google Container Engine

scratch images, cool!

```
FROM scratch

ADD ca-certificates.crt /etc/ssl/certs/ca-certificates.crt
ADD zoneinfo.zip /usr/local/go/lib/time/zoneinfo.zip

ADD server /server

ENTRYPOINT ["/server"]
```

Google Container Registry

GCR is nice and simple!

```
gcloud docker [--authorize-only, -a] [--docker-host=DOCKER_HOST]
    [--server=SERVER,[SERVER,...], -s SERVER,[SERVER,...];
        default="gcr.io,us.gcr.io,eu.gcr.io,asia.gcr.io,b.gcr.io,
        bucket.gcr.io,l.gcr.io,launcher.gcr.io,appengine.gcr.io,
        us-mirror.gcr.io,eu-mirror.gcr.io,asia-mirror.gcr.io,mirror.gcr.
    [GCLOUD_WIDE_FLAG ...] [-- DOCKER_ARGS ...]
```

Google Container Engine

Illustrated

The ^Children's Guide to
Kubernetes

Written by: Matt Butcher   Illustrated by: Bailey Beougher

Google Container Engine

See a few
Kelsey Hightower
talks...

Google Container Engine!

OK, we have a service!!

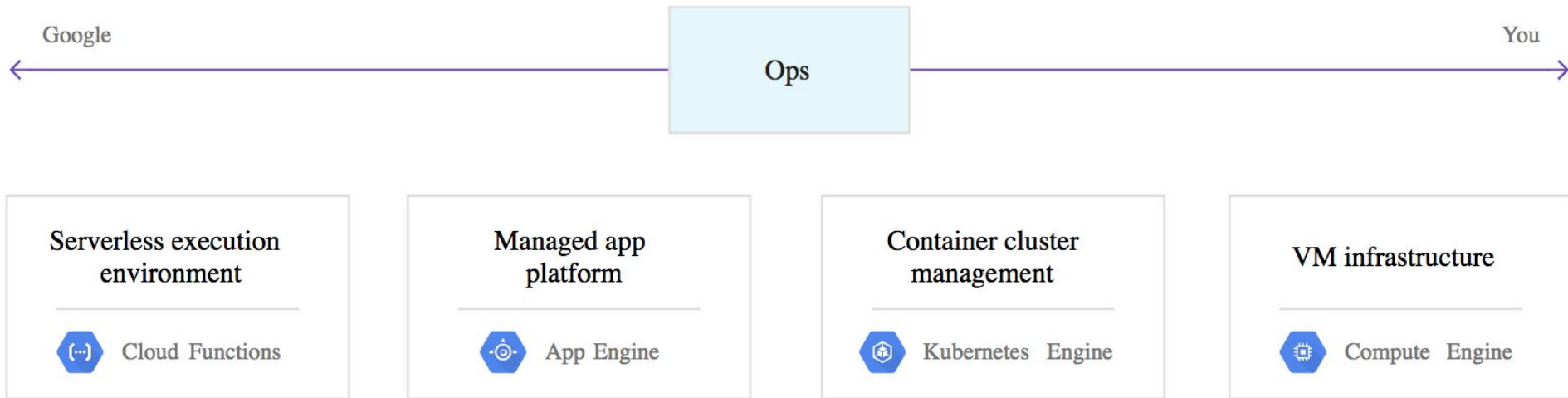...but something's missing

Google Container Engine?

We have no way of automating SSL??

Bummer.

Google's Compute Spectrum

@NYTDevs | developers.nytimes.com

Google's Compute Spectrum

Google   You

| Ops |

Serverless execution environment	Managed app platform	Container cluster management	VM infrastructure
Cloud Functions	App Engine	Kubernetes Engine	Compute Engine

App Engine "Managed VMs"

I can reuse my container
...or it builds it for me!

```
└ gcloud app deploy -h
Usage: gcloud app deploy [DEPLOYABLES ...] [optional flags]
  optional flags may be    --bucket | --help | --image-url | --promote |
                           --stop-previous-version | --version | -h
```

App Engine "Managed VMs"

DNS and SSL for free!!

```
https://SERVICE_ID-dot-MY_PROJECT_ID.appspot.com
```

App Engine "Managed VMs"

It autoscales!

```
automatic_scaling:
  min_num_instances: 1
  max_num_instances: 15
  cool_down_period_sec: 180
  cpu_utilization:
    target_utilization: 0.6
```

App Engine "Managed VMs"

It's not GA?! (2016)

...OK, let's revisit it later.

(**note**: it is GA *now*)

Google's Compute Spectrum

Google   You

Ops

Serverless execution environment	Managed app platform	Container cluster management	VM infrastructure
Cloud Functions	App Engine	Kubernetes Engine	Compute Engine
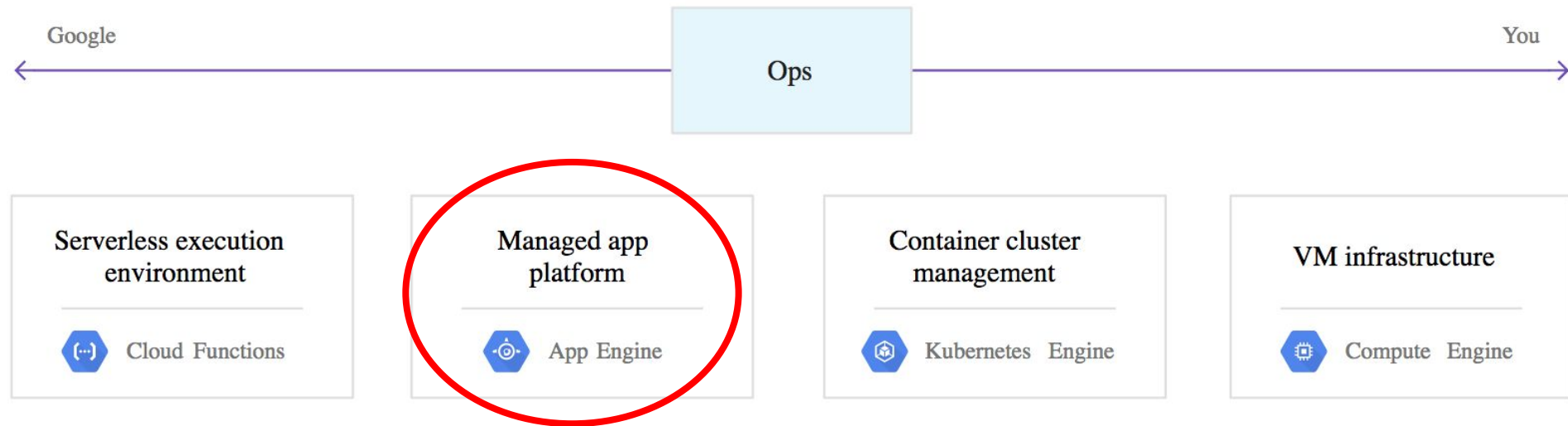
App Engine Standard

DNS and SSL!

```
https://SERVICE_ID-dot-MY_PROJECT_ID.appspot.com
```

App Engine Standard
Combined Access & App Logs!!


```
λ  13:48:40.346  PUT   200      268 B    114 ms  Firefox 52  /svc/crosswords/v2/game.json?regi

0.1.0.40 - - [23/Jun/2017:13:48:40 -0400] "PUT /svc/crosswords/v2/game.json?regi_id=56254915
6/25 "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:52.0) Gecko/20100101 Firefox/52.0 AppEngine-Goo
prd.appspot.com" ms=114 cpu_ms=46 cpm_usd=2.9951e-8 loading_request=0 instance=00c61b117cc28
8 trace_id=-
```

```
▼ {
    ▶ protoPayload: {…}
      insertId: "594d54790002b2e671eedcc0"
    ▶ httpRequest: {…}
    ▶ resource: {…}
      timestamp: "2017-06-23T17:48:40.346263Z"
      severity: "DEBUG"
    ▶ labels: {…}
      logName: "projects/nyt-games-prd/logs/appengine.googleapis.com%2Frequest_log"
    ▶ operation: {…}
      receiveTimestamp: "2017-06-23T17:48:41.178909151Z"
  }
```

```
λ  13:48:40.350  incoming commit: n7aroa
λ  13:48:40.350  incoming update_timestamp: 1498239820
λ  13:48:40.350  looking up progress for 51945265-13423
λ  13:48:40.385  querying for newer commits that are also reset with key of 51945265-13423
λ  13:48:40.390  skipping getall and applying to existing progress
λ  13:48:40.395  adding commit: n7aroa
```

App Engine Standard
Tracing!
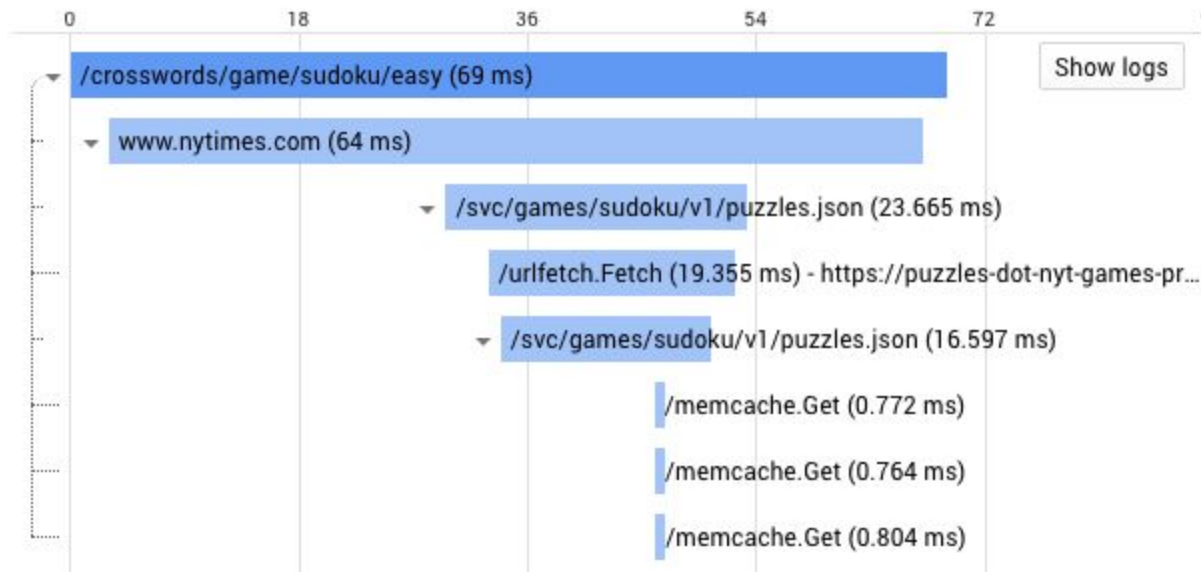
App Engine Standard

Scales to 100s of instances in seconds?!
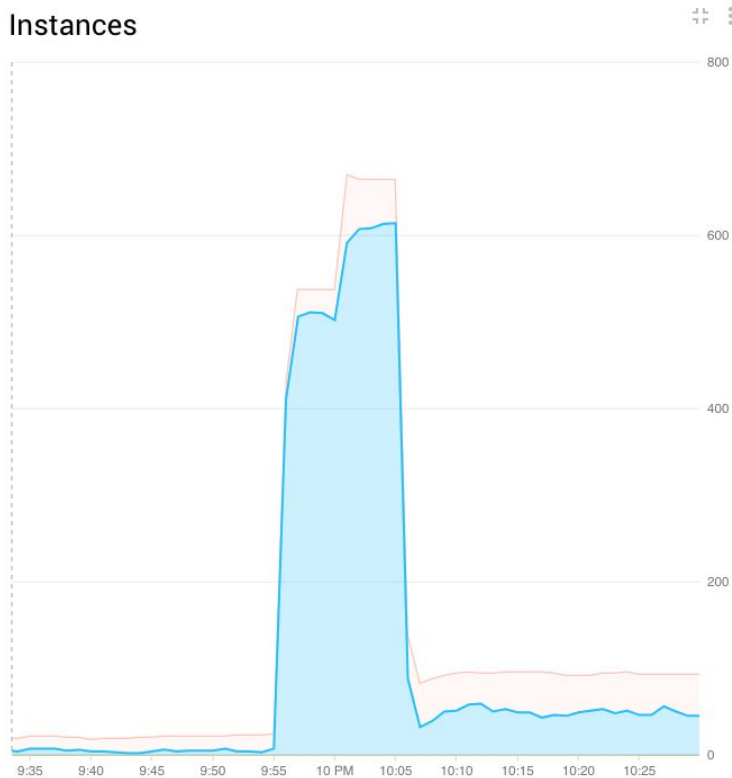
App Engine Standard

Simple user and service level security built in??

```
handlers:
- url: /youraccount/.*
    script: _go_app
    login: required
    secure: always
```

App Engine Standard

Asserting identity to other App Engine apps

If you want to determine the identity of the App Engine app that is making a request to your App Engine app, you can use the request header `X-Appengine-Inbound-Appid`. This header is added to the request by the URLFetch service and is not user modifiable, so it safely indicates the requesting application's ID, if present.

App Engine Standard

Local Development

```
    goapp serve -h
usage: serve [serve flags] [application_dir | package | yaml_files...]

Serve launches your application on a local development App Engine server.
```

Time to Propose a GCP Solution

We wrote an internal RFC:

- 4 step process
- Most concerns were over lock in
- SLOs were made
- Agreed this will be a PoC

Step 1: The Proxy
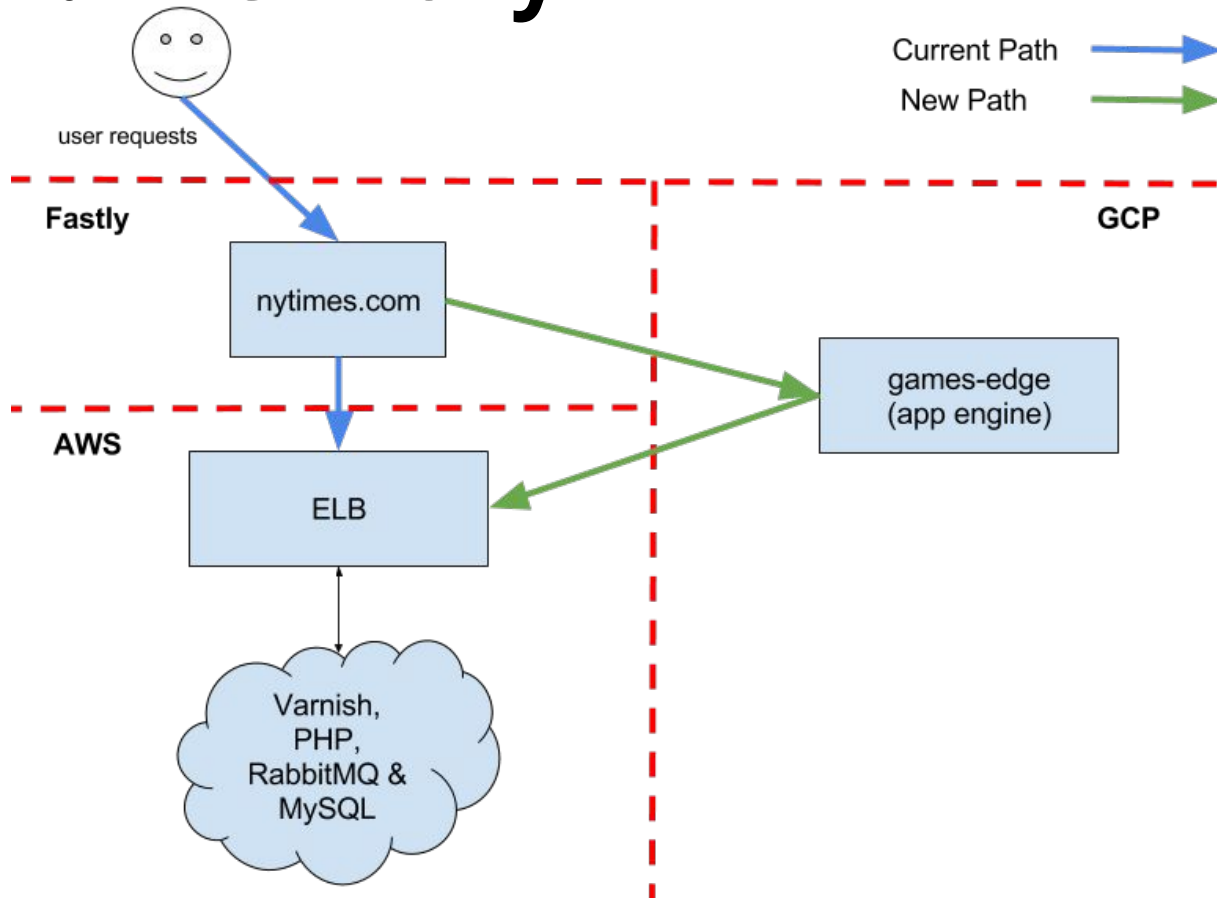
Step 1: Problems

- Large headers
  - Too many third party cookies from web
  - Fix: Only pass what we need
- Autoscaling couldn't *quite* keep up
  - *"Request was aborted after waiting too long to attempt to service your request."*
  - Fix: sprinkle a little cron on it

Scheduled Autoscaling

```yaml
- description: nightly 10pm scale up event
  url: /_ah/push-handlers/scale/up
  schedule: every day 21:55
  timezone: America/New_York
  target: default

- description: nightly scale down after 10pm scale up event
  url: /_ah/push-handlers/scale/down
  schedule: every day 22:05
  timezone: America/New_York
  target: default
```

Scheduled Autoscaling

```
- url: /_ah/push-handlers/*
  script: _go_app
  login: admin

env_variables:
  IDLE_INSTANCES_UP: 400
  IDLE_INSTANCES_DOWN: 2
```
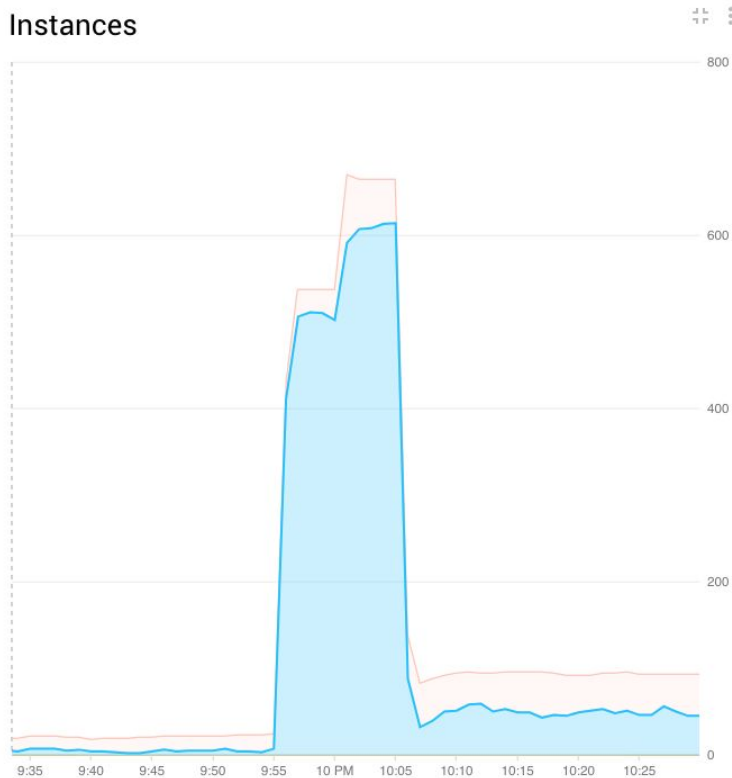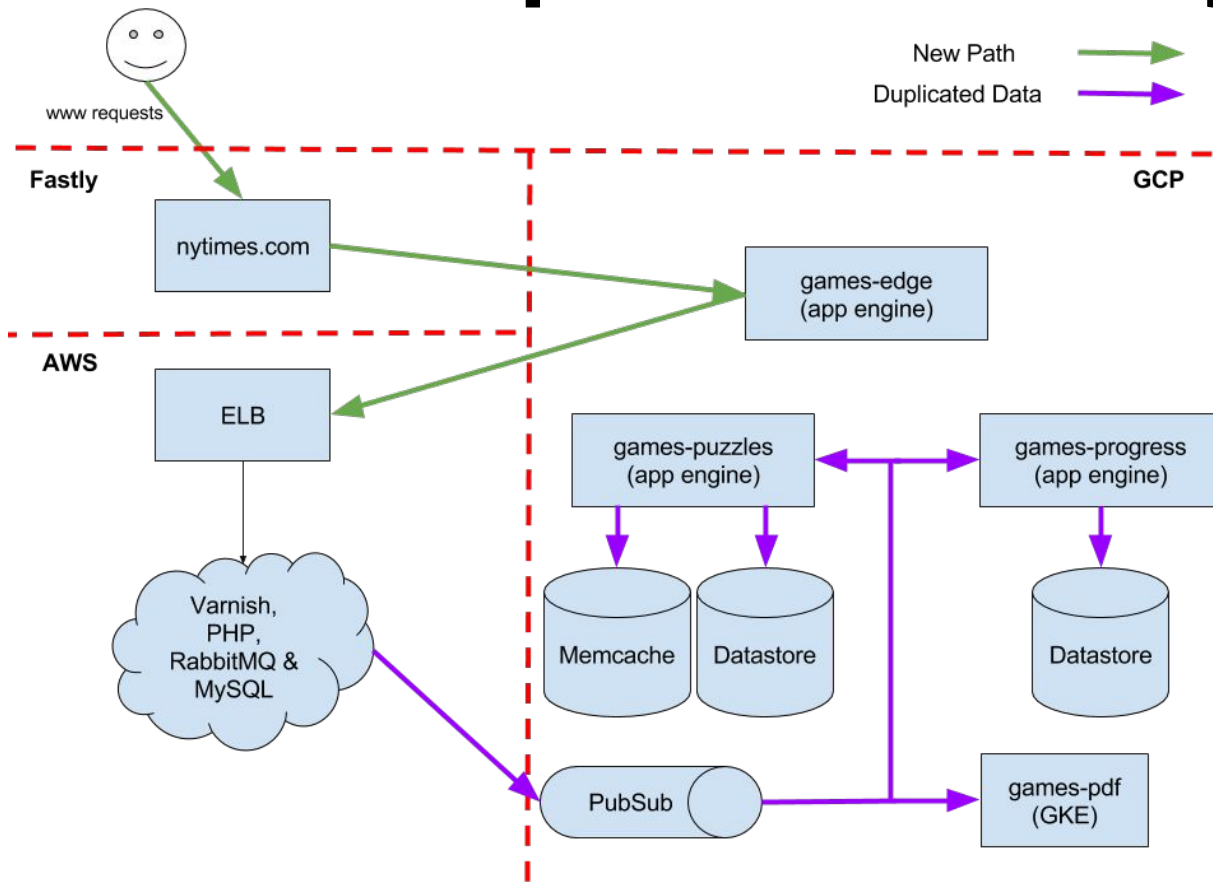
Scheduled Autoscaling

```go
svc := appengine.NewAppsServicesVersionsService(clnt)
call := svc.Patch(appengin.AppID(ctx), svcName, vrsn, &appengine.Version{
        AutomaticScaling: &appengine.AutomaticScaling{
                MaxIdleInstances: idleInstances + 5,
                MinIdleInstances: idleInstances,
        },
}).UpdateMask("automaticScaling.min_idle_instances,automaticScaling.max_idle_instances")
```
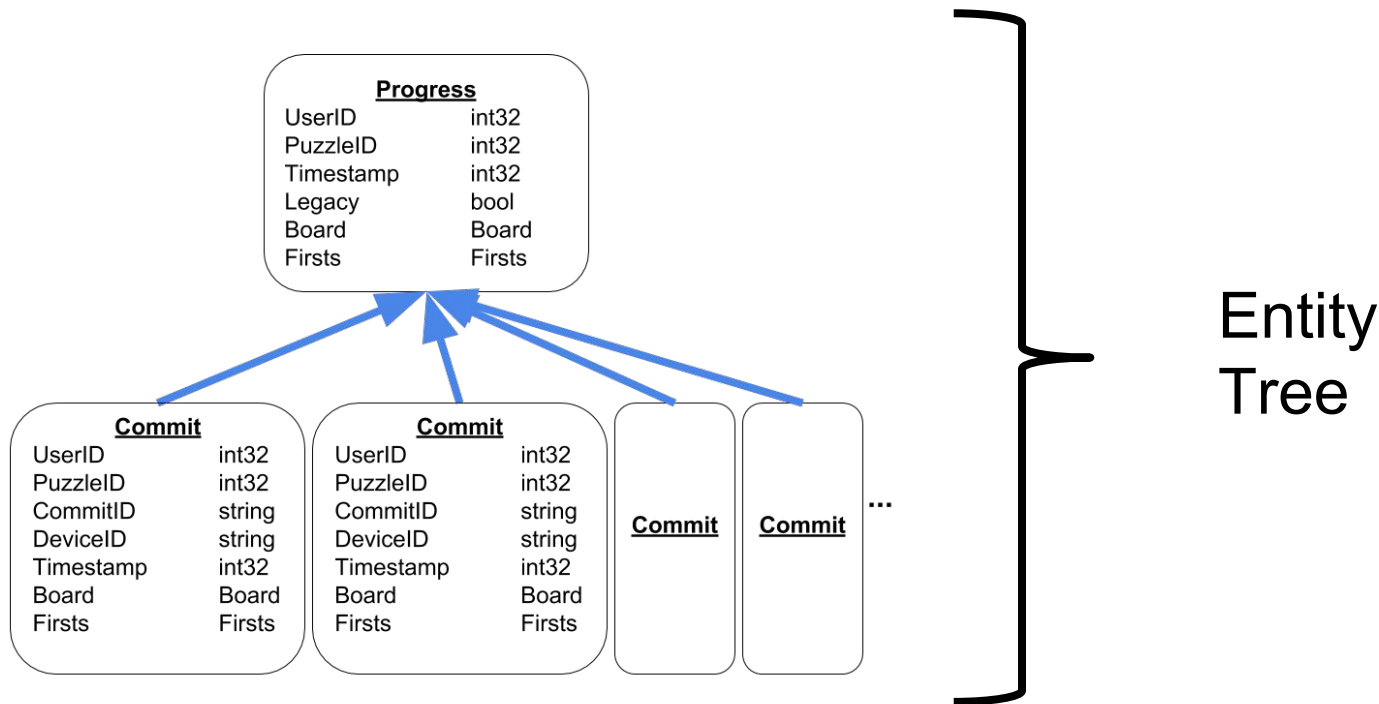
Scheduled Autoscaling
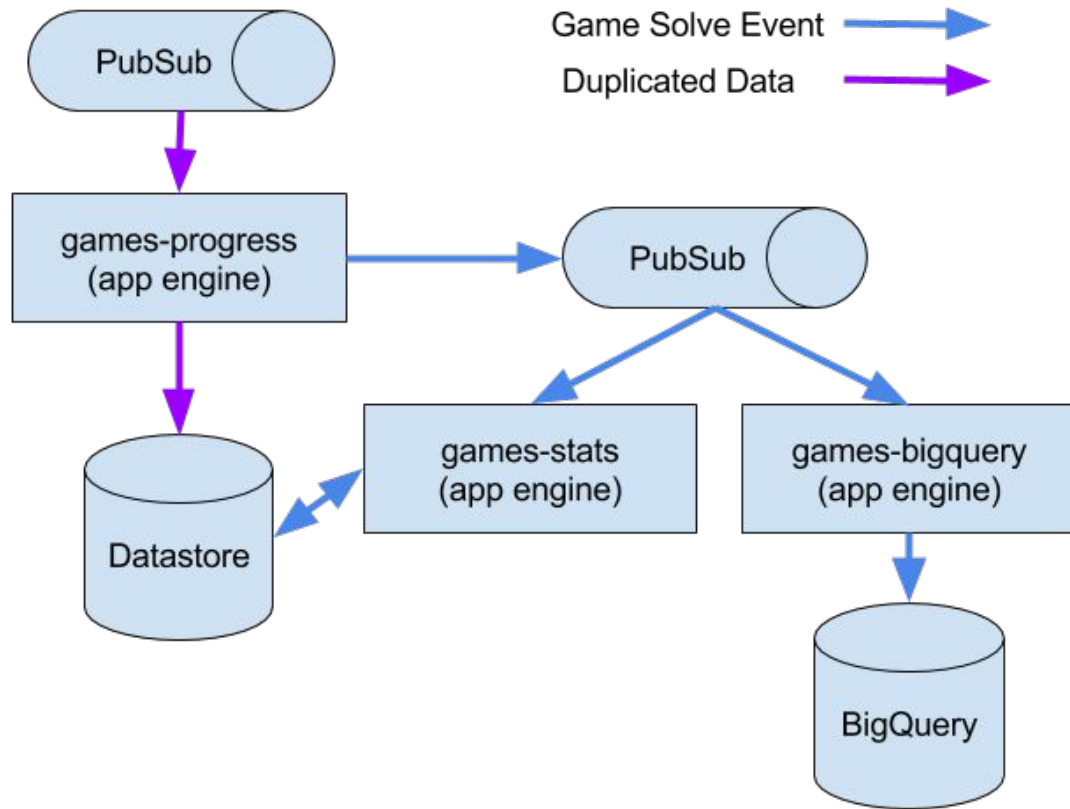
100s of instances in seconds!!!

Step 2: New Endpoints + Data Sync

Step 2: Using Datastore

Progress

UserID	int32
PuzzleID	int32
Timestamp	int32
Legacy	bool
Board	Board
Firsts	Firsts

Commit

UserID	int32
PuzzleID	int32
CommitID	string
DeviceID	string
Timestamp	int32
Board	Board
Firsts	Firsts

Commit

UserID	int32
PuzzleID	int32
CommitID	string
DeviceID	string
Timestamp	int32
Board	Board
Firsts	Firsts

Commit

Commit

...

Entity Tree

Step 2: Data Sync

Step 2: Problems

- How to shrink a PDF?
  - We need ghostscript!
  - Fix: GKE and PubSub
- Whoa! Datastore got pricey!
  - Too many reads when replaying data
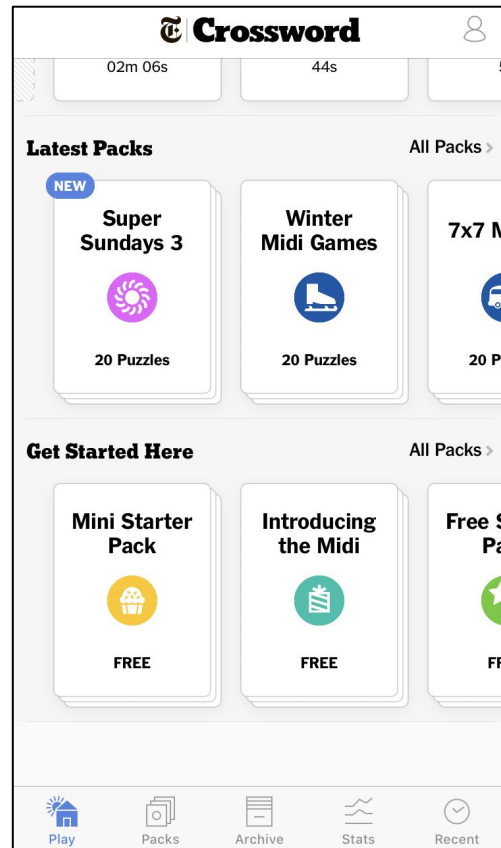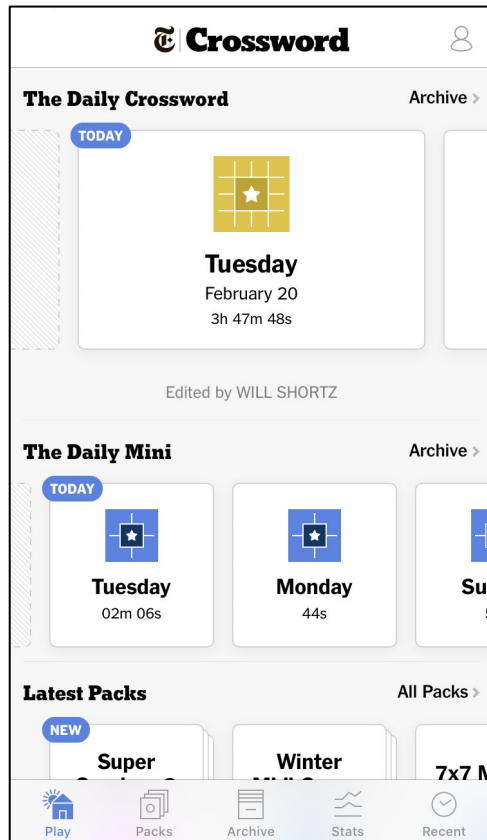 - Fix: add a lock around stats calculations
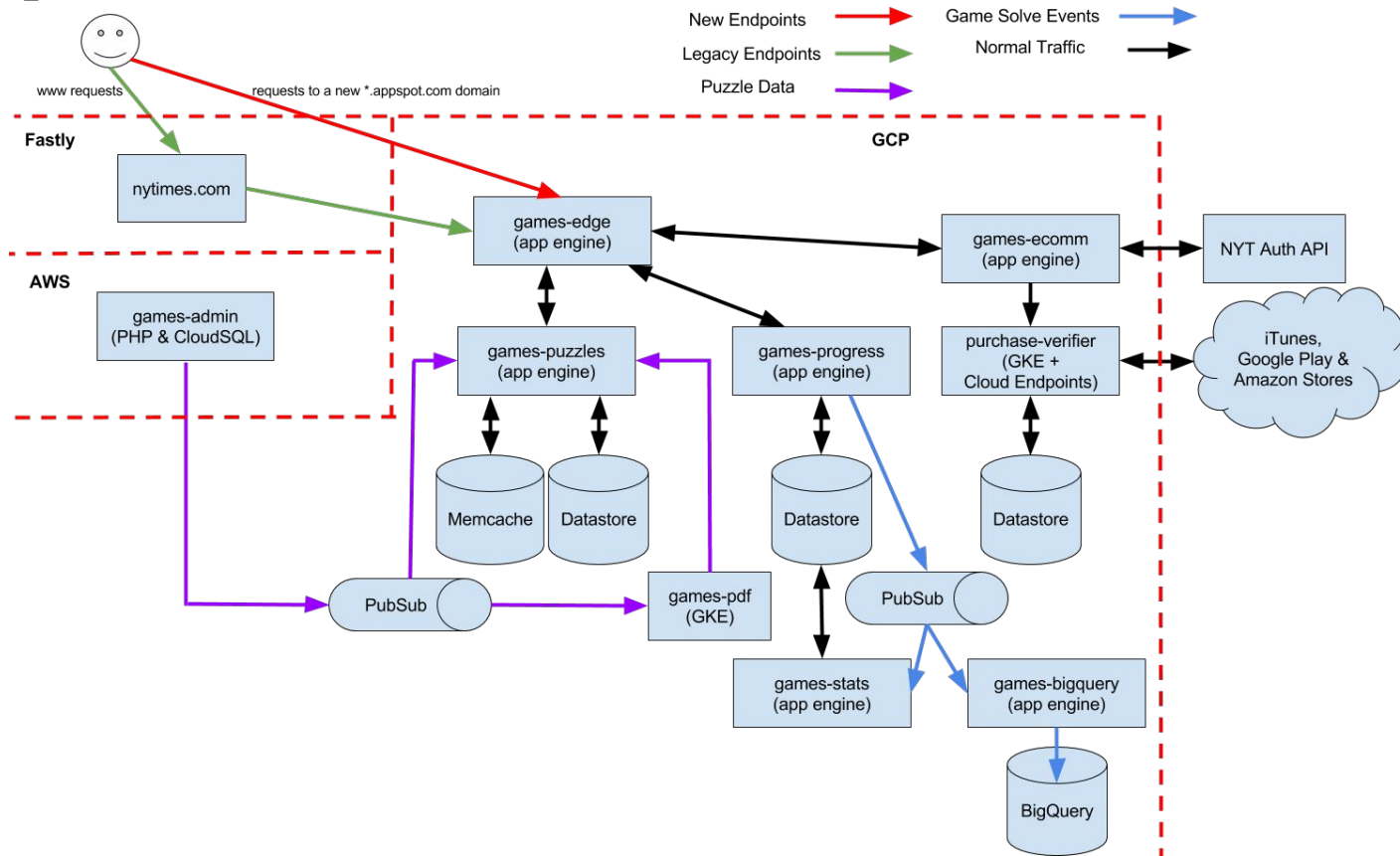
Step 3: Making GCP Live!
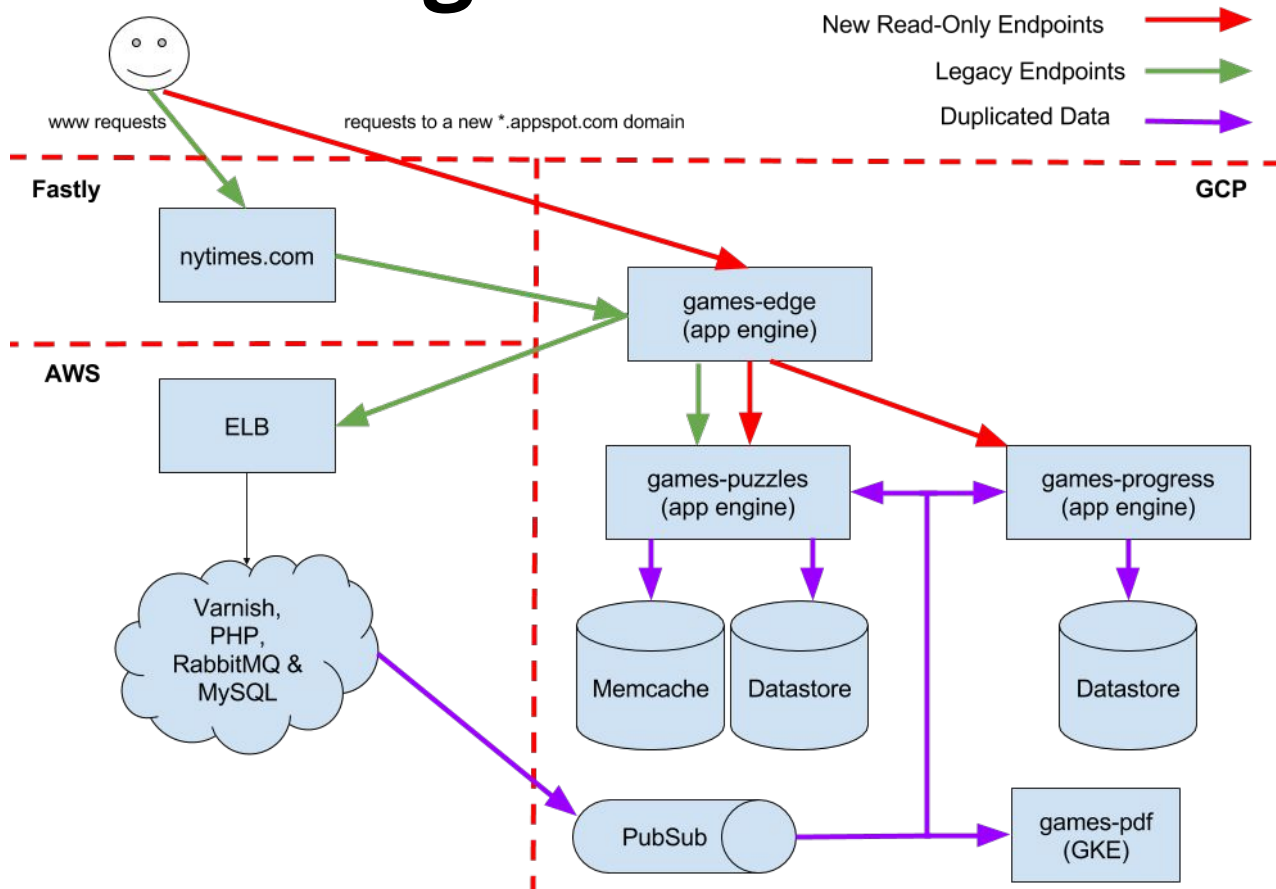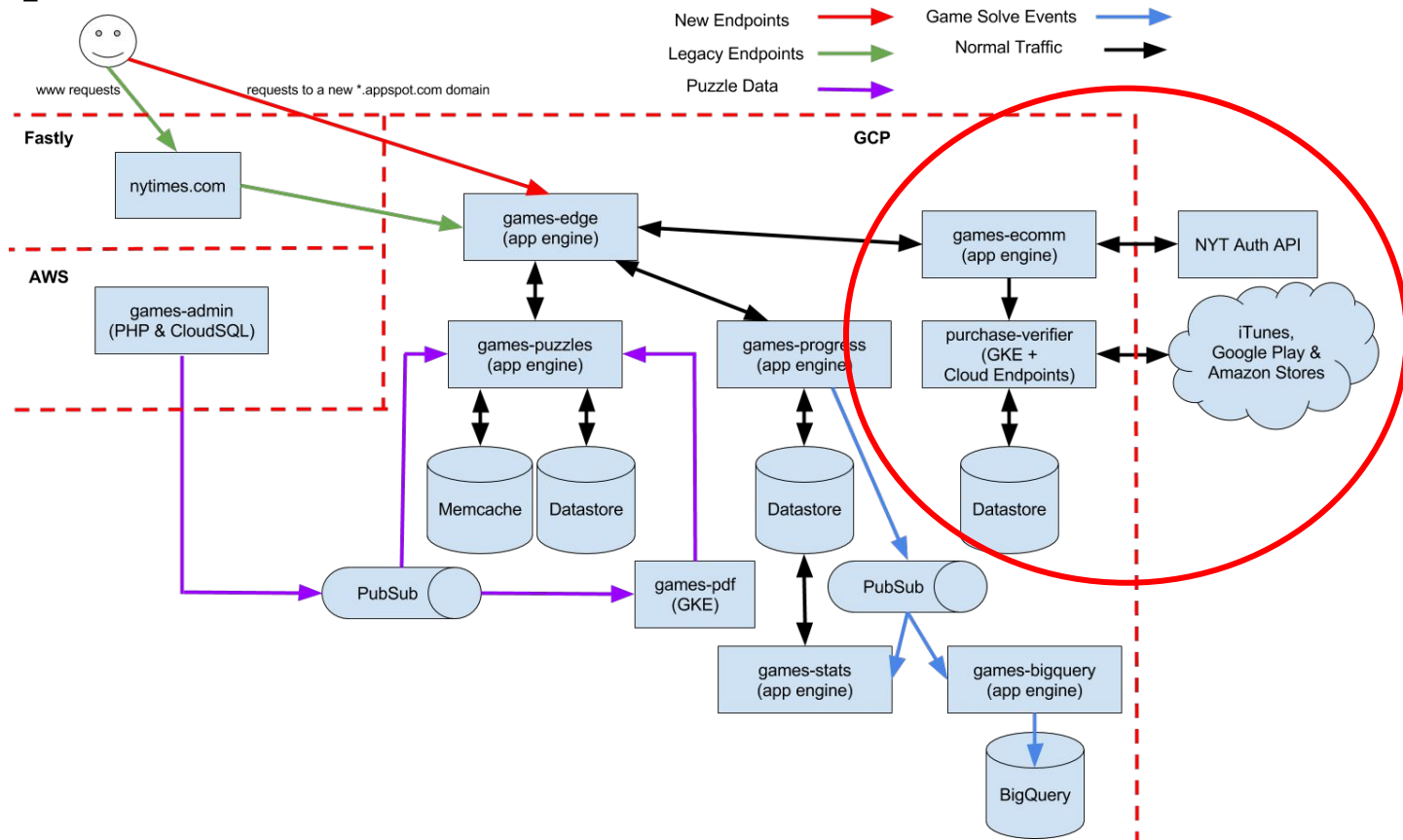
Step 3: The "Server View Model"

New iOS home screen!

Step 4: Final Piece of the Puzzle

Step 4: Final Piece of the Puzzle

Step 4: OpenAPI, gRPC & Endpoints

- Service specs written in Open API
- Converted to Protobuf w/ openapi2proto
- Protos uploaded to Cloud Endpoints
- CE converts JSON to gRPC for us

The Results

- Started December 2016
  - ~10MM requests/day
- Completed September 2017
  - ~30MM+ requests/day
- Cloud costs cut in half
  - 💰💰💰💰
- 6 game prototypes publicly tested
  - Coming soon: Spelling Bee!

What's Next?

Multiplayer Crosswords!
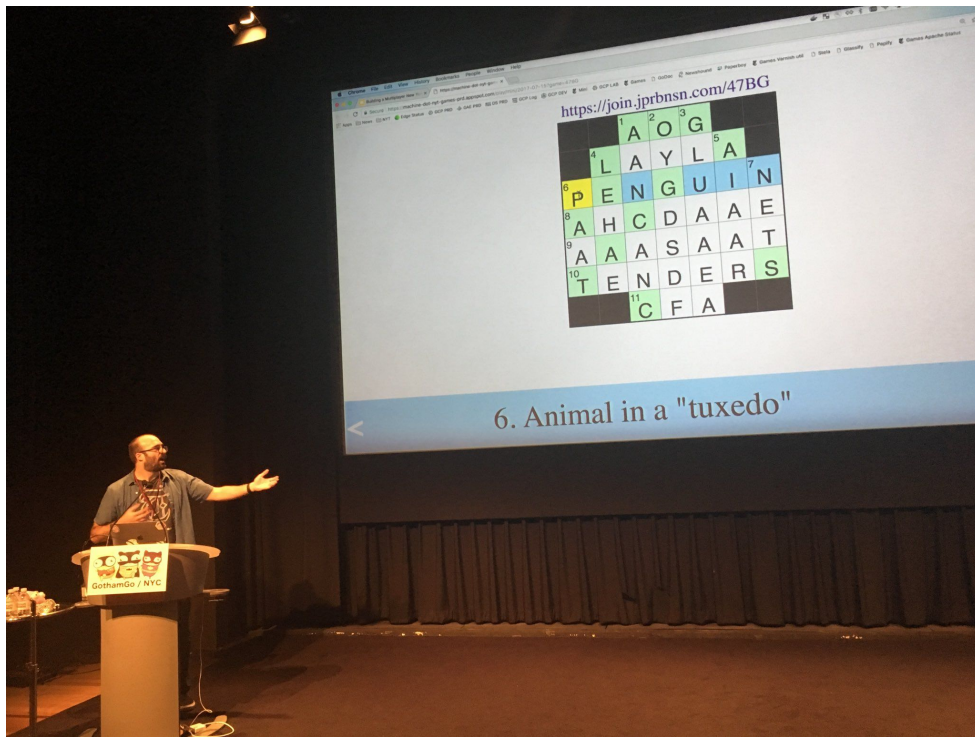
What's Next?

JP Robinson
@jprbnsn

been playing around with @googlecloud's #spanner, #appengine and #golang to manage (new!) social leaderboards for the @NYT_Crossword this week.

great dev experience so far, but I should probably reach out to our design team 😳

mini - 2017-12-22

Name	Timer	# Revealed
JP	22	0
Ben	26	0

10:54 AM - 22 Dec 2017

Thanks!

Questions?