System Design Decisions

wins and losses we had in recent Cloud Native projects

chris.chen@cevo.com.au
June 2022

https://cevo.atlassian.net/wiki/spaces/DP/pages/1648918532/Brownbag+Schedule

System Design Decisions

~~wins and losses~~ we had in recent Cloud Native projects

war stories

suggested by Jason Keane

chris.chen@cevo.com.au

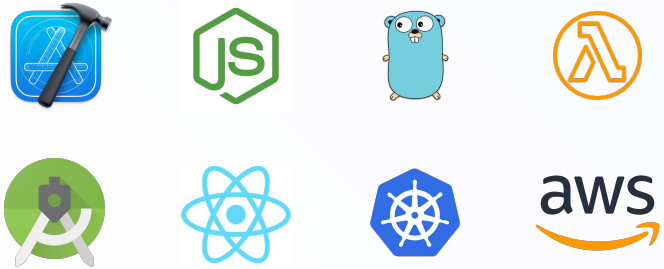
June 2022

https://cevo.atlassian.net/wiki/spaces/DP/pages/1648918532/Brownbag+Schedule

About me

cevo

your friendly neighbourhood jack of all trades



clients since Mar 2021

CENTRE for SOCIAL IMPACT

DAVID JONES

SERVICE

What this talk is about

fancy system architectures

awesome AWS services

sophisticated database schemas

technical deep-dive

❌

it actually works?

wish I had thought of that …

crap, what I've done!

anecdotes and self-reflection

✅

Stories

1. API is a no-brainer, let’s do it
2. Idempotence, what a strange word
3. How much integration testing is enough
4. Go-live is not the finish line

Stories

1. **API is a no-brainer, let's do it**
2. Idempotence, what a strange word
3. How much integration testing is enough
4. Go-live is not the finish line

Send data to gov systems

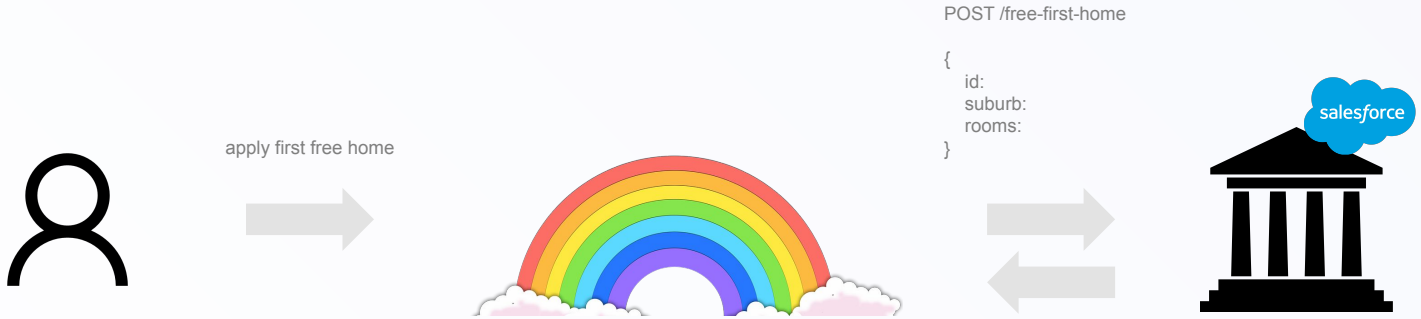
Send data to gov systems - File-based

- 🙂 loosely coupled
- 😞 limited data formats
- 😞 inefficient
- 😞 no real time

Send data to gov systems - Salesforce API

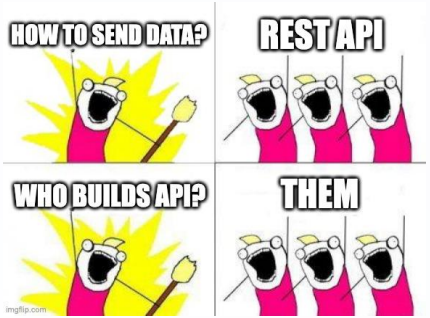
- 🙂 real time
- 🙂 richer data formats
- 😞 tightly coupled
- 😞 domain knowledge

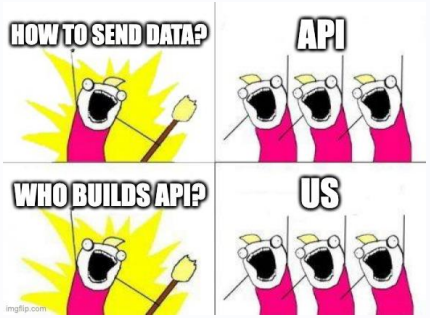
Send data to gov systems - REST API

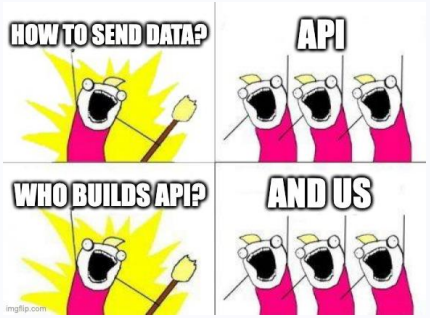


- 🙂 real time
- 🙂 richer data formats
- 🙂 loosely coupled
- 🙂 send & receive

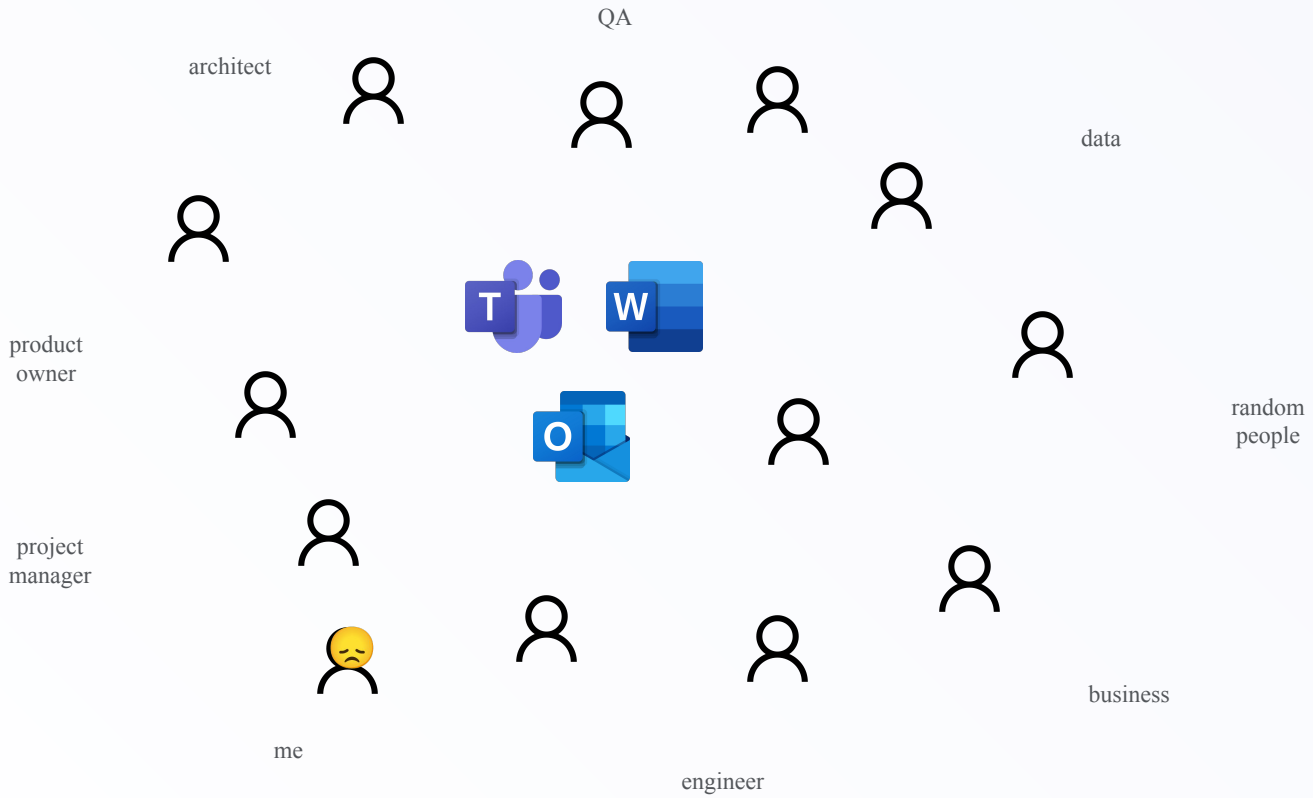
Everyone is onboard







Communication overhead escalated



Communication overhead escalated

258 DJPR GEMS – Service Victoria Integration API

Solution Architecture Document (Ver. 1.0)

re: SAD (Ver. 1.0)

fw: SAD (Ver. 1.0)

…

What we've asked for

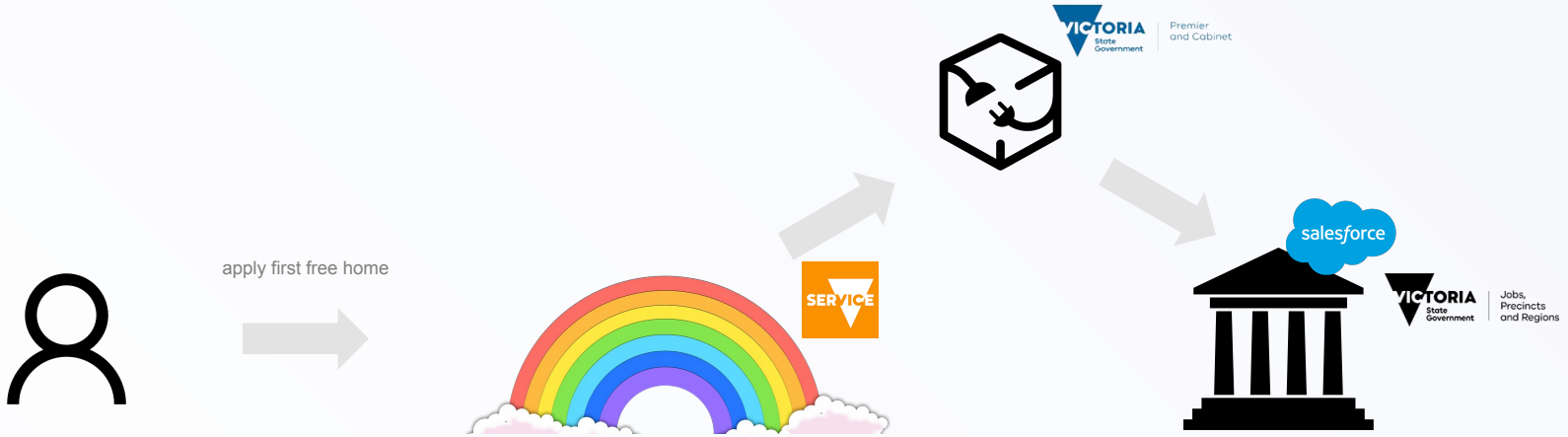
apply first free home

POST /free-first-home

```
{
   id:
   suburb:
   rooms:
}
```

- 🙂 real time
- 🙂 richer data formats
- 🙂 loosely coupled
- 🙂 send/receive

What we've got



- 🙂 real time
- 🙂 richer data formats
- 🙂 loosely coupled
- 🙂 send/receive

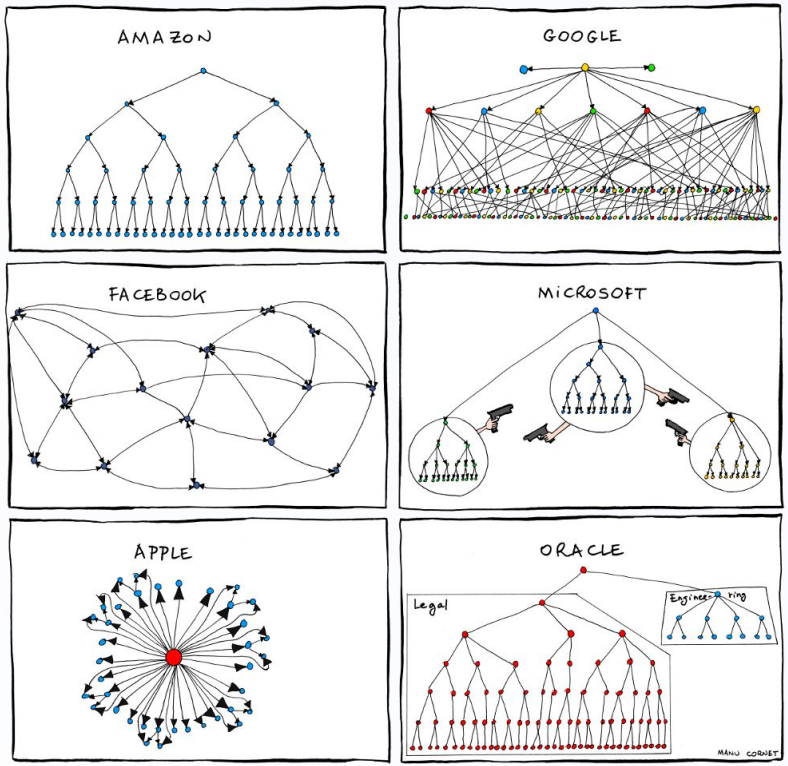
- 😞 **communication overhead**
- 😞 **testing/release coordination**
- 😞 **API reliability/performance**
- 😞 **eyebrow-raising**

Conway’s Law

Any organization that designs a system will produce a design whose structure is a copy of the organization's communication structure.

— Melvin E. Conway

https://en.wikipedia.org/wiki/Conway%27s_law



Stories

1. **API is a no-brainer, let's do it**
   a. what out of our control is a risk
   b. communication tax compounds
   c. once a one-way door decision is made, execute

2. Idempotence, what a strange word

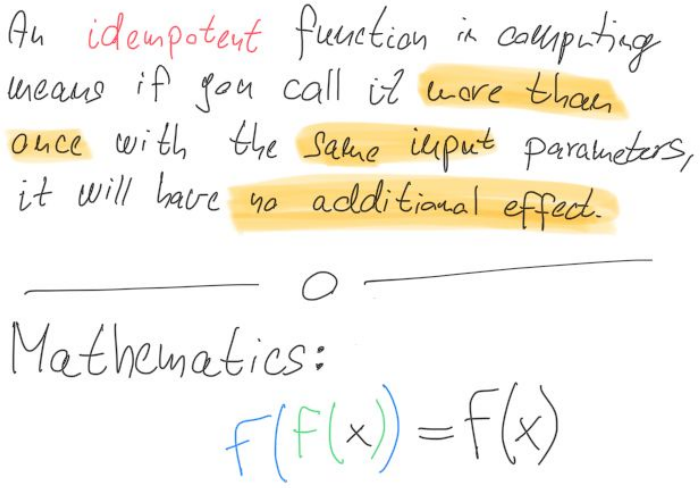
3. How much integration testing is enough

4. Go-live is not the finish line

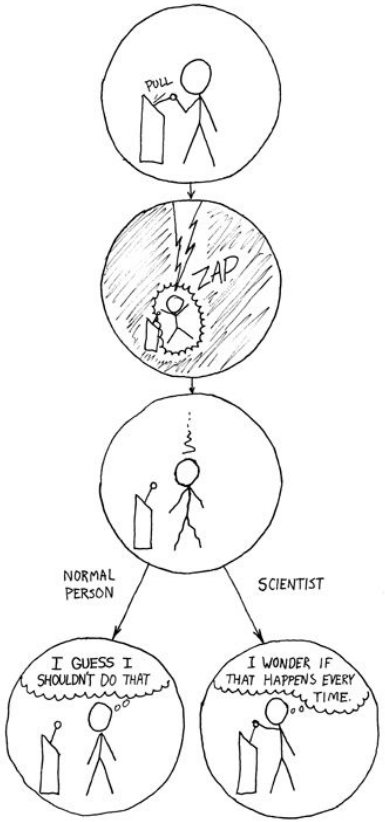
Stories

1. API is a no-brainer, let’s do it
2. **Idempotence, what a strange word**
3. How much integration testing is enough
4. Go-live is not the finish line

Idempotence, what a strange word



Idempotence, what a strange word

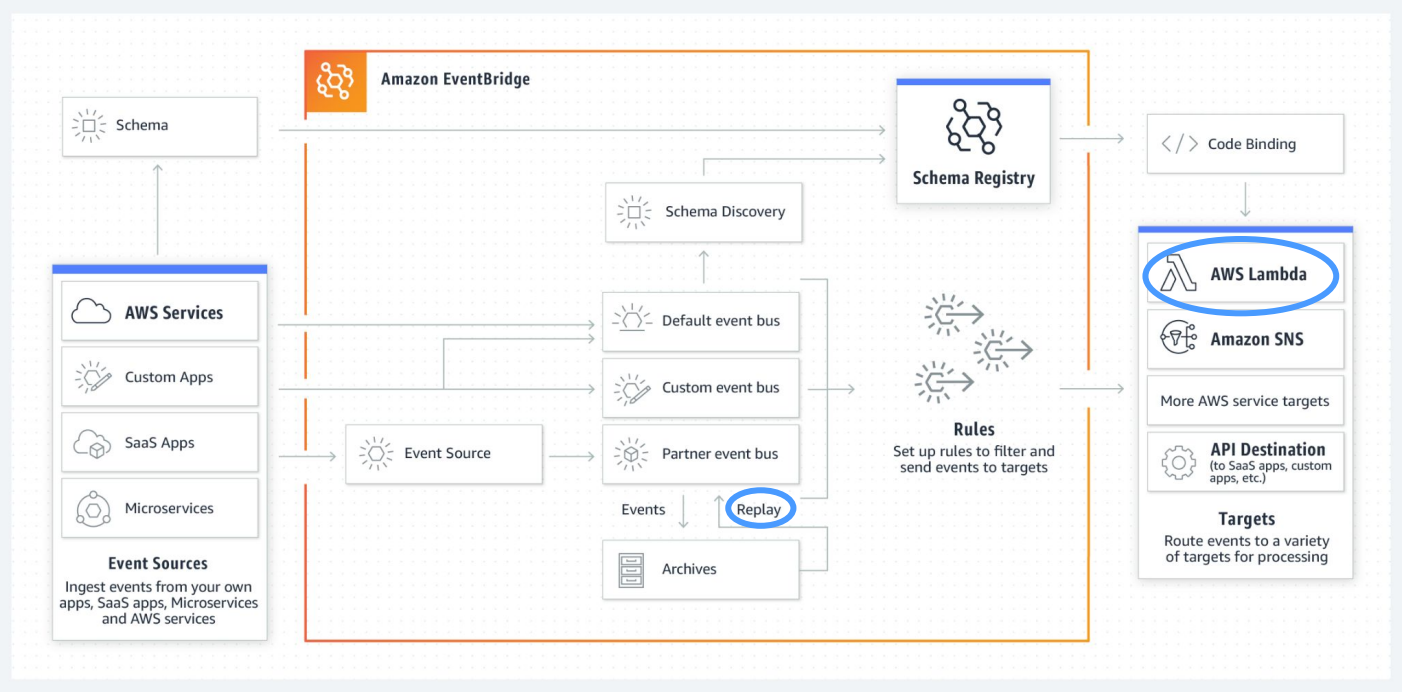


Send data to gov systems

It does more than that

What if something goes wrong?

Retry - AWS EventBridge or alike

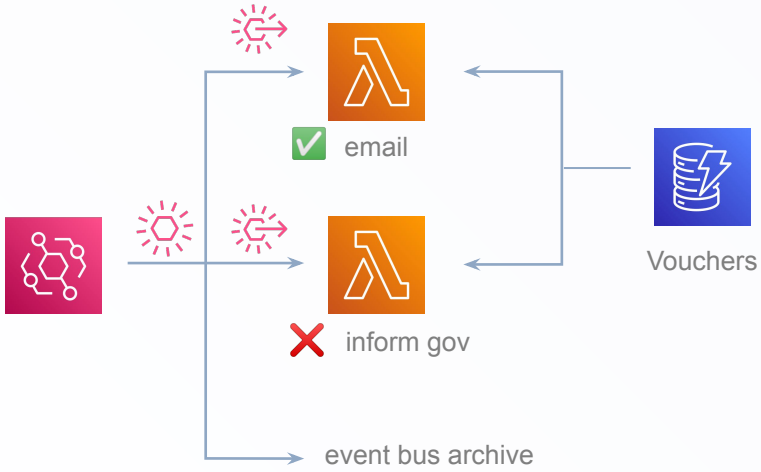


https://aws.amazon.com/eventbridge/

Without idempotence

With idempotence

EventBridge – archive & replay

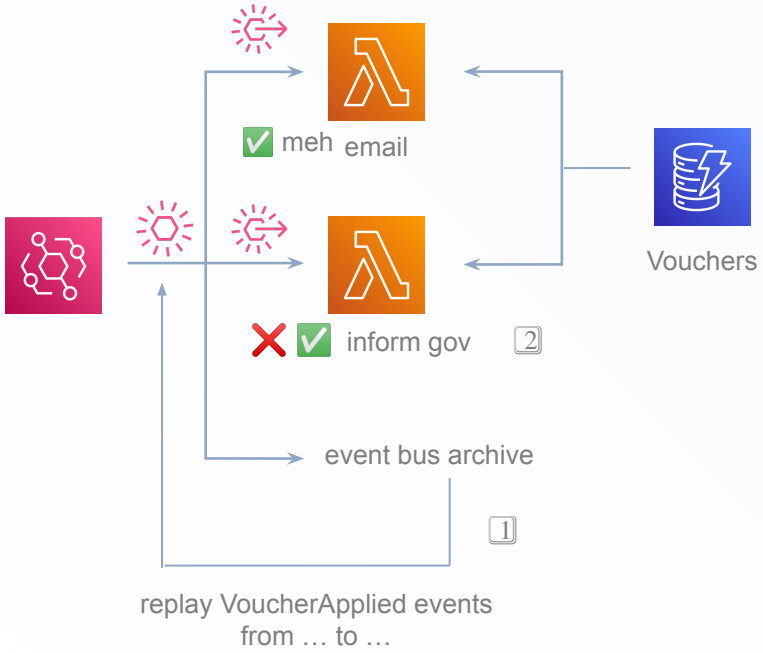


ID	…	statusEvents

1	…	[
		{
		"status": "email_sent",
		"timestamp": "2022-02-01T14:48:00.000Z"
		}
		]

…	…	…

EventBridge – replay & idempotence



ID	…	statusEvents

1	…	[
		{
		"status": "email_sent",
		"timestamp": "2022-02-01T14:48:00.000Z"
		},
		{
		"status": "gov_informed",
		"timestamp": "2022-02-01T14:50:00.000Z"
		}
		] 3

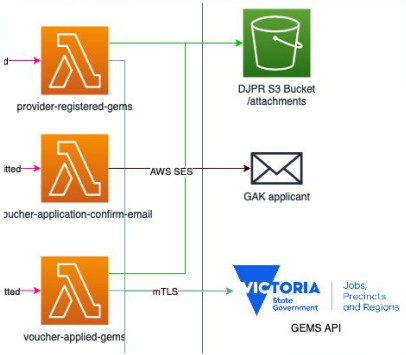
…	…	…

It actually works!

It also works in an unexpected way

It also works in an unexpected way

API not available: Glue ETL job daily run, exporting and transferring CSV



API available: live switch to transferring data over API, not data duplication

ID	…	statusEvents

1	…	[
		{
		"status": "email_sent",
		"timestamp": "2022-02-01T14:48:00.000Z"
		},
		{
		"status": "gov_informed",
		"timestamp": "2022-02-01T14:50:00.000Z"
		}
		]

…	…	…

Wish I have thought of

ID	…	statusEvents

1	…	[
		…
		{
		"status": "gov_informed",
		"timestamp": "2022-02-01T14:50:00.000Z"
		}
		]

…	…	…

ID	…	statusEvents

1	…	[
		…
		{
		"status": "informing_gov",
		"timestamp": "2022-02-01T14:50:00.000Z"
		},
		{
		"status": "api_timeout",
		"timestamp": "2022-02-01T14:51:00.000Z"
		},
		{
		"status": "informing_gov",
		"timestamp": "2022-02-01T14:52:30.000Z"
		},
		{
		"status": "gov_informed",
		"timestamp": "2022-02-01T14:52:35.000Z"
		}
		]

…	…	…

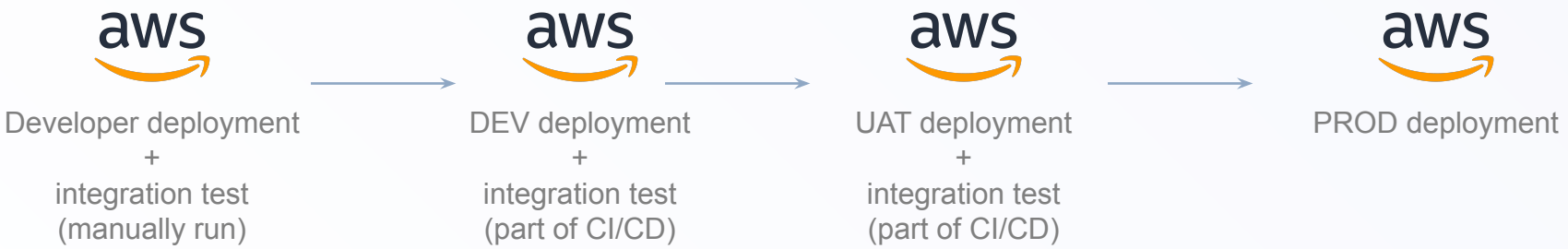
Stories

1. API is a no-brainer, let's do it
2. **Idempotence, what a strange word**
   a. anticipate system failures
   b. design fine-grained state machines
3. How much integration testing is enough
4. Go-live is not the finish line

Stories

1. API is a no-brainer, let’s do it
2. Idempotence, what a strange word
3. **How much integration testing is enough**
4. Go-live is not the finish line

Deploy and test with real things



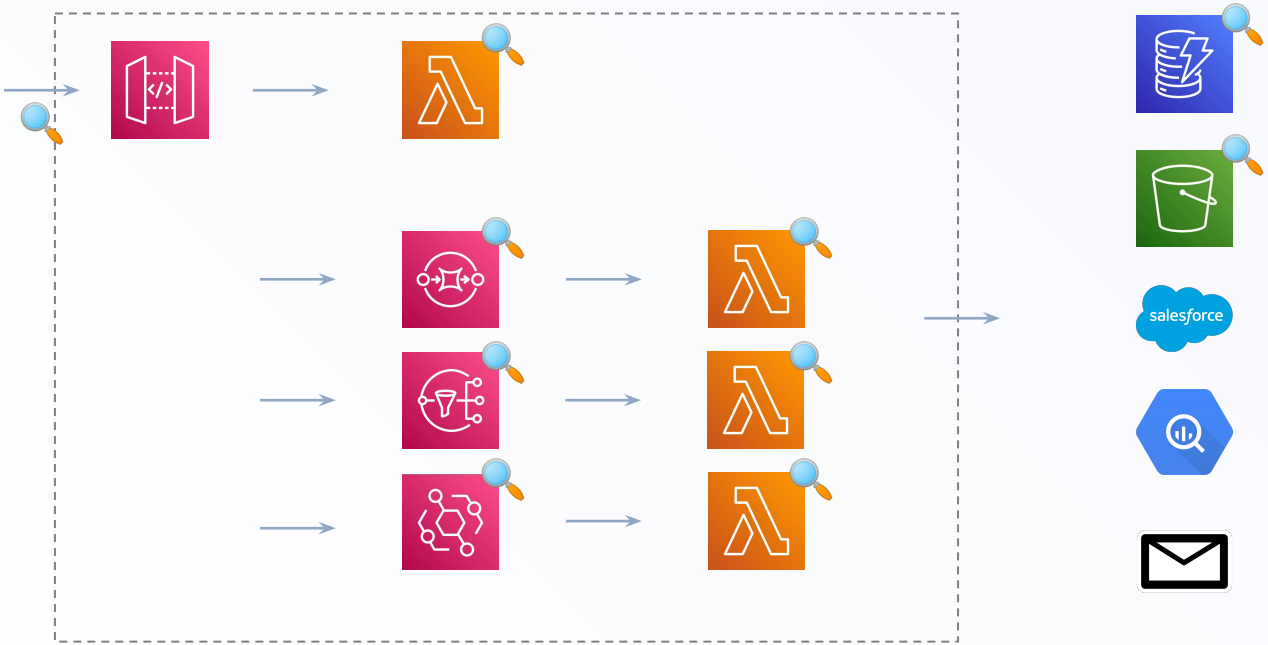
🔍 race condition

🔍 network latency

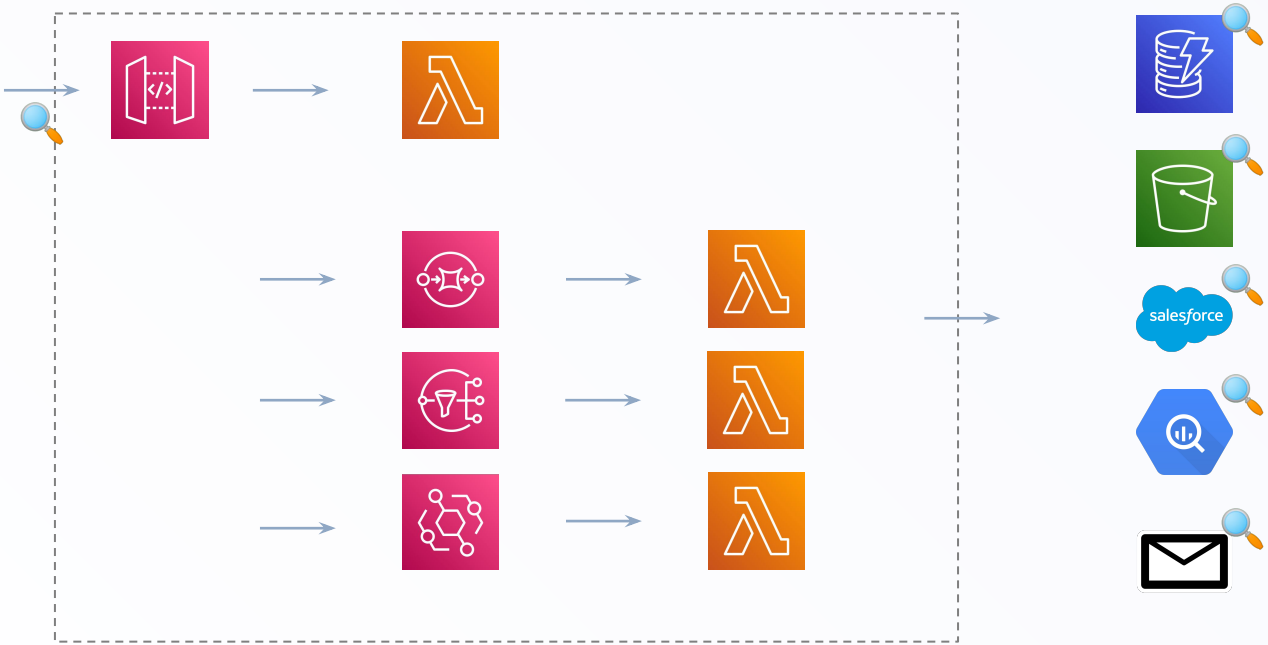
🔍 API throttling

🙂 PROD deploy on Friday

Too much testing in side the box



~~Think~~ Test outside the box



and I still broke the PROD 🙈

cevo

Sharath Menon 17 Mar at 13:55
@Evan Bruhn - looks like I am not able to upload an a
UAT

43 replies

Evan Bruhn 3 months ago
OK, checking...

Sharath Menon 3 months ago
not working for provider registration as well

Evan Bruhn 3 months ago
@Daniel Kim @Chris Chen @Antony Dabonde the f
returning a 403

Sharath Menon 3 months ago
in prod as well

Evan Bruhn 3 months ago
Yikes

Antony Dabonde 3 months ago
@Chris Chen WAF related?

Daniel Kim 3 months ago
He is just off a bit

Sharath Menon 3 months ago
has WAF been setup for provider registration in prod
app wide setup?

Daniel Kim 3 months ago
Yes GAK wide setup

Daniel Kim 3 months ago
Looking now to disable

Sharath Menon 17 Mar at 13:55
@Evan Bruhn - looks like I am not able to upload an attachment either in UAT

43 replies

Evan Bruhn 3 months ago
OK, checking...

Sharath Menon 3 months ago
not working for provider registration as well

Evan Bruhn 3 months ago
@Daniel Kim @Chris Chen @Antony Dabonde the file upload API is returning a 403

Sharath Menon 3 months ago
in prod as well

Evan Bruhn 3 months ago
Yikes

Antony Dabonde 3 months ago
@Chris Chen WAF related?

Daniel Kim 3 months ago
He is just off a bit

Sharath Menon 3 months ago
has WAF been setup for provider registration in prod as well or is it a GAK app wide setup?

Daniel Kim 3 months ago
Yes GAK wide setup

Daniel Kim 3 months ago
Looking now to disable

Daniel Kim 3 months ago
Can we directly disable in AWS console?

Sharath Menon 3 months ago
but are we sure that WAF would have caused this?

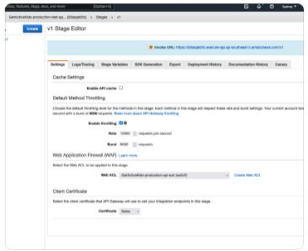
Antony Dabonde 3 months ago
Don't know regarding how to disable. I'd need to lookup/look at how Chris implemented.

Unsure if WAF would have caused this either. Just seems suspicious.

Evan Bruhn 3 months ago
For when you're back @Chris Chen, it's the `limit-by-common-rules` rule that's being tripped here

Daniel Kim 3 months ago
Looks like we can from here

Screen Shot 2022-03-17 at 2.25.14 pm.png



Daniel Kim 3 months ago
anyway once DEV deploy success then I will merge to master

Evan Bruhn 3 months ago
yay, UAT is working again

Daniel Kim 3 months ago
Now deploying on DEMO as well. we can test actually.

Daniel Kim 3 months ago
Cool

Daniel Kim 3 months ago
I will merge to master

Sharath Menon 3 months ago
nice work guys! & good hunch @Antony Dabonde 👍🏽

Chris Chen 3 months ago
thanks!

Chris Chen 3 months ago
blocked by `AWS#AWSManagedRulesCommonRuleSet#SizeRestrictions_BODY`

Chris Chen 3 months ago

Daniel Kim 3 months ago
@Sharath deployed to production

Stories

1. API is a no-brainer, let's do it

2. Idempotence, what a strange word

3. **How much integration testing is enough**
   a. focus on upstream/downstream impacts
   b. focus on critical paths
   c. Murphy's law, i.e.: don't deploy before going out for lunch

4. Go-live is not the finish line

Stories

1. API is a no-brainer, let’s do it
2. Idempotence, what a strange word
3. How much integration testing is enough
4. **Go-live is not the finish line**

Don’t forget your blanket

design, build, launch

monitor, observe, discover

Monitor, observe, discover

- 🙂 Logs, system/business metrics on DataDog
- 😞 Dead-letter Queue (DLQ) was done, dusted, and forgotten
- 😞 Most issues were discovered by the users

Monitor, observe, discover

🙂 Logs, system/business metrics on DataDog

😞 Dead-letter Queue (DLQ) was done, dusted, and forgotten
- DLQ monitoring and alerting
- Mitigation: manual vs automated

😞 Most issues were discovered by the users

Monitor, observe, discover

🙂 Logs, system/business metrics on DataDog

😞 Dead-letter Queue (DLQ) was done, dusted, and forgotten
- DLQ monitoring and alerting
- Mitigation: manual vs automated

😞 Most issues were discovered by the users
- Data analytics pipeline
- Data anomalies discovery: manual vs automated
- MTTR (mean time to recovery, repair, respond, or resolve)

Recap

1. API is a no-brainer, let's do it
    a. what out of our control is a risk
    b. communication tax compounds
    c. once a one-way door decision is made, execute

2. Idempotence, what a strange word
    a. anticipate system failures
    b. design fine-grained state machines

3. How much integration testing is enough
    a. focus on upstream/downstream impacts
    b. focus on critical paths
    c. Murphy's law, i.e.: don't deploy before going out for lunch

4. Go-live is not the finish line
    a. don't forget monitor, observe, and discover

cevo