

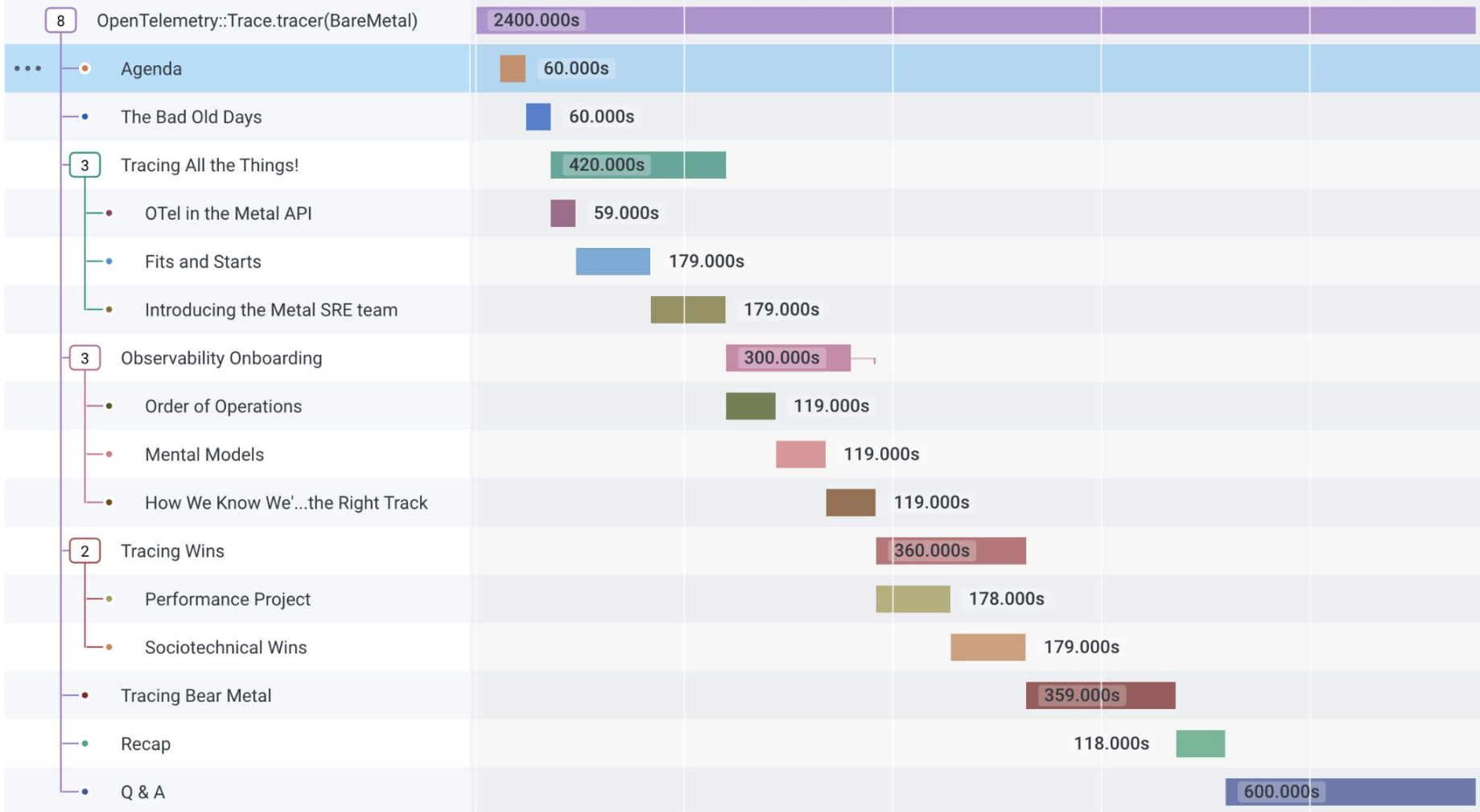


EQUINIX

Tracing Bare Metal with OpenTelemetry

Amy Tobey, Sr. Principal Engineer
Shelby Spees, SRE





The Bad Old Days

long incidents, inconclusive findings

learned helplessness

rhyming incidents

team siloization & heroes

low autonomy

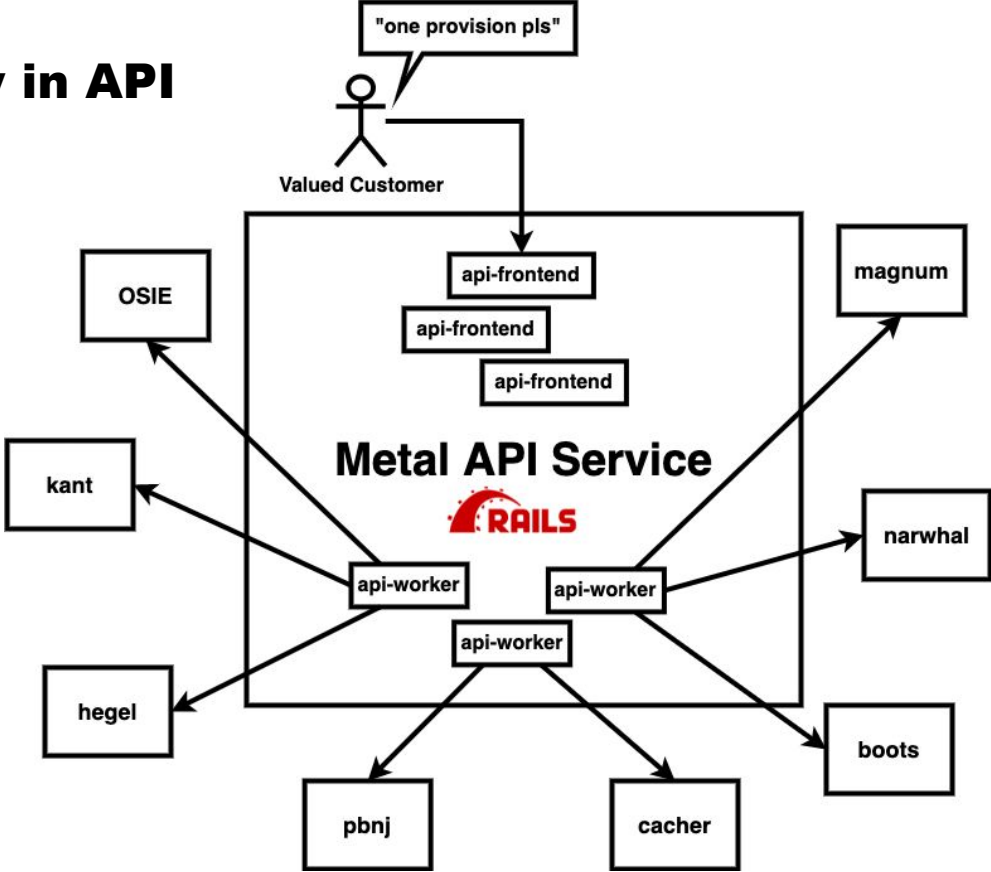
~~Simply~~

~~Restart~~

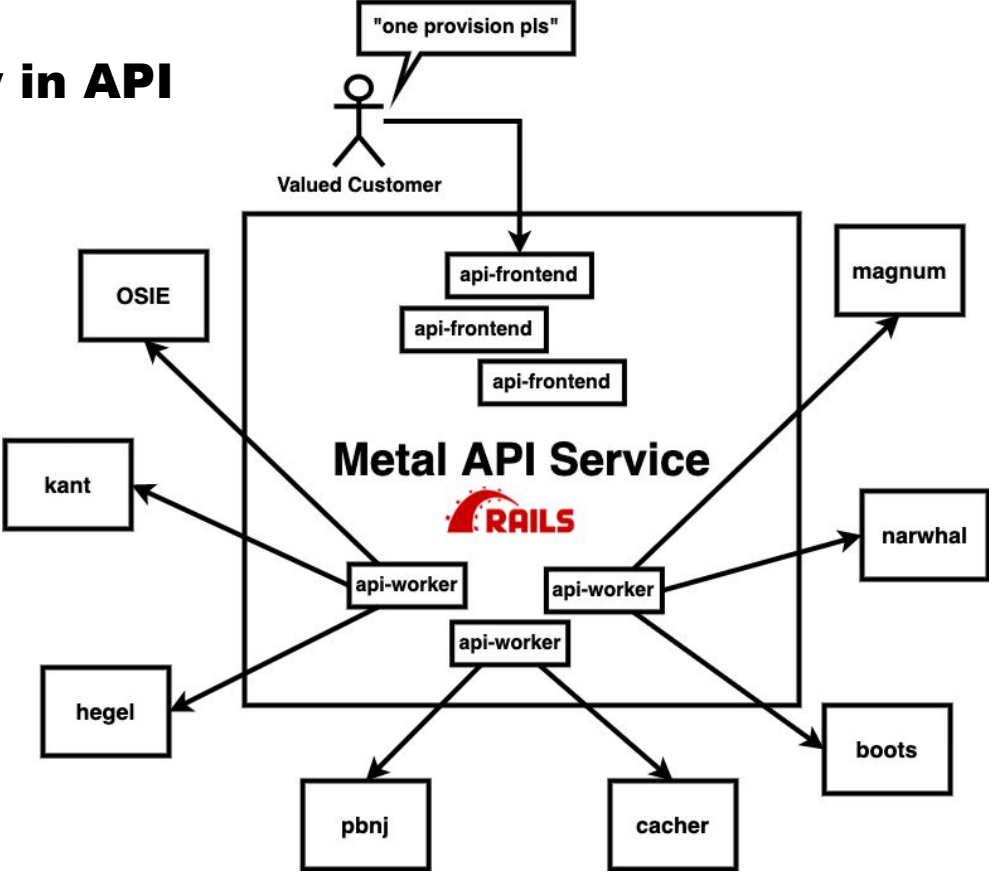
~~Everything~~

Tracing All The Things!

OpenTelemetry in API



OpenTelemetry in API



Fits and Starts

-  [SRE-168] Reenable OTEL tracing for jobs (Reverts PR 7182) ✓

#7212 by merged on Jan 7 • Approved

-  Disabling otel for activejob ✓

#7182 by was merged on Dec 17, 2021 • Approved

-  Revert "Re-enable ActiveRecord OTEL span integration" ●

#6666 by was merged on Jun 2, 2021 • Approved

-  [ENG-15475] Re-enable ActiveRecord OTEL span integration ●

#6665 by was merged on Jun 2, 2021 • Approved

-  Disable ActiveRecord OTEL span integration ●

#6648 by was merged on May 26, 2021 • Approved



atobey 8:40 AM

otel code is executing, just the sending part is disabled

-  Disable OTEL ✓

#363 by was merged on Jun 4, 2021 • Approved

-  [ENG-15475] Re-enable OTEL Tracing in API codebase ✓

#360 by was merged on Jun 2, 2021 • Approved

-  revert otel enabled ✓

#333 by was merged on May 6, 2021 • Approved

December 2021: Disabled OTel in API Service

HEATMAP(LOG10(duration_ms))



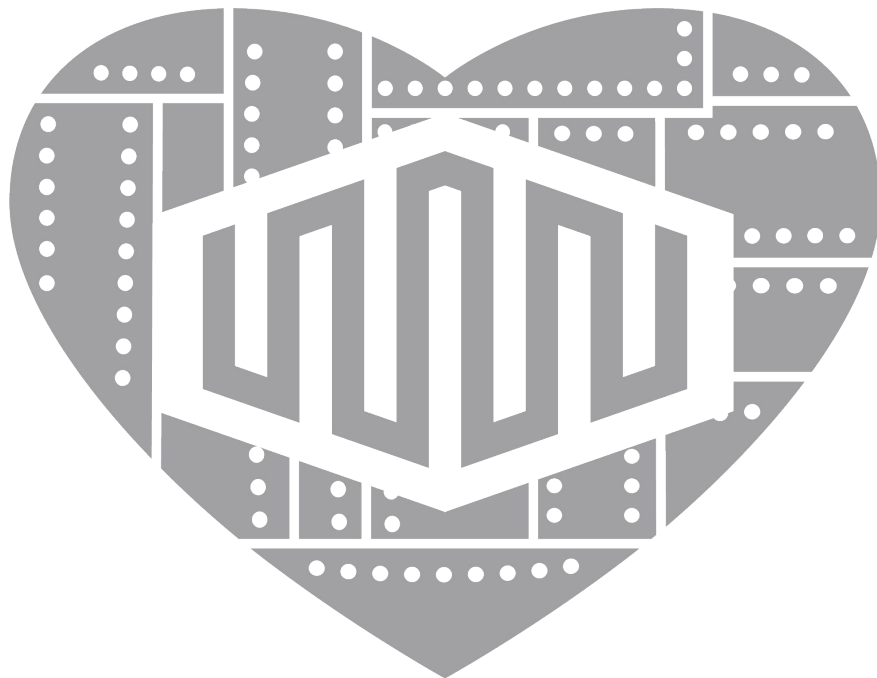
COUNT



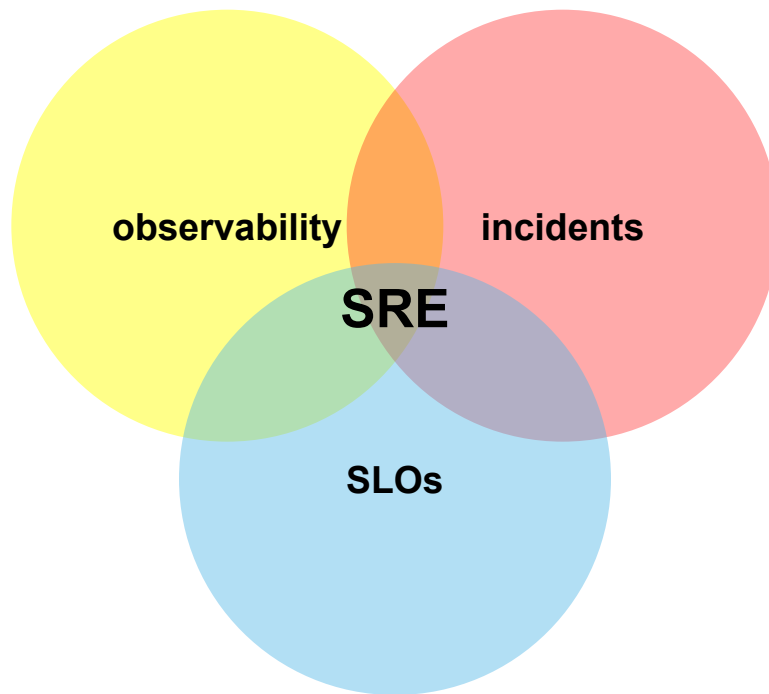
Blog: "OpenTelemetry and the Ghost of Provisioning Past"



One SRE Wasn't Enough, We Needed a Team



The Very Model of a Modern SRE Program



Observability Onboarding

Order of Operations

- (always) add plumbing first!
- (probably) OpenTelemetry Collector
- (maybe) tracing initializers
- (almost certainly) propagation



Blog: "OpenTelemetry: What's a Collector and Why Would I Want One?"

Trace Propagation

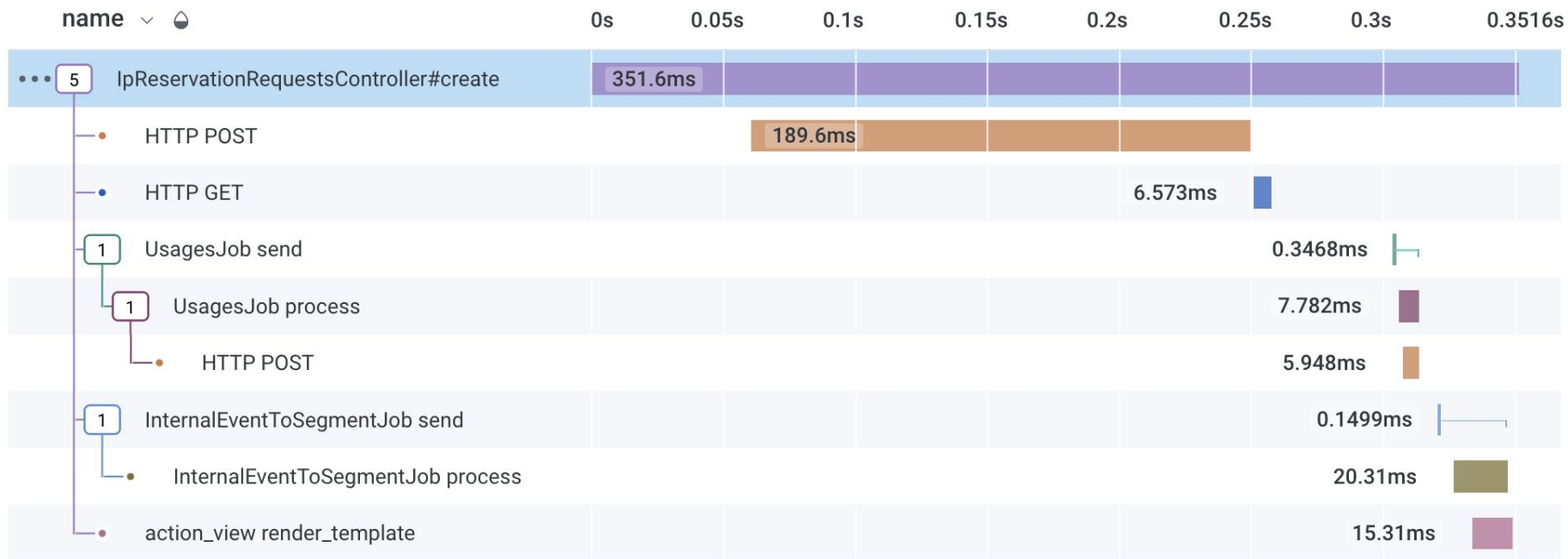
traceparent=00-7cf6adc22888fb0b798bfa180a2ba21f-43d43ae150985538-01

tracestate=vendor=stuff,keys=values

baggage=keys=values,out-of-band=signaling

00-7cf6adc22888fb0b798bfa180a2ba21f-43d43ae150985538-01

Auto-instrumentation



Mental Models

Pattern matching requires experience, but experience doesn't scale.

Code tracing > Mental tracing

- data outside of people's brains
- visual reference
- linkable and shareable



Custom Instrumentation

What attributes to add?

- focus on business-critical domain data
- for Equinix Metal: instance, server, operating system, etc.
- reflect what's in the code

fine print: ask your Telemetry vendor if high Cardinality is for you!

`str` instance.os

VMware ESXi 6.5

`str` instance.os_distro

vmware

`str` instance.plan

m3.large.x86

`str` server.arch

x86_64

`str` server.id

`str` server.state

in_use






How We Know We're on the Right Track








developers asking for access to tracing tools & pairing time

PRs being merged to add custom instrumentation

people (and SRE) sharing links to traces and query results during incidents

people exclaiming "the system does WHAT!?"

  12:46 PM
 hello, I'm here with more Honeycomb questions
 1 

  12:52 PM
replied to a thread:  [About to turn on OTEL for PBNJ](#) 
Landed and deployed! Let the  traces flow!
 4 

Send archetype feature value to honeycomb

 committed 2 days ago ✓

 liked your page

 **Observability and Tracing**

Engineering

Tracing Wins

February 2022: Performance Issues in Customer Portal

Long-running issues needed to be addressed

Name	St...	Ty...	I...	Size	Time	Waterfall
<input type="checkbox"/> ips?state=all&&include=assignments,assignments.assigned_to&exclude=	200	fe...	d...	439 B	8.52 s	
<input type="checkbox"/> ips?&include=assignments,assignments.assigned_to&exclude=	200	fe...	d...	430 B	7.84 s	
<input type="checkbox"/> plans?include=labels&deployment_type=on_demand	200	fe...	d...	411 B	7.26 s	
<input type="checkbox"/> metros	200	fe...	d...	517 B	1.57 s	
<input type="checkbox"/> projects?page=1&per_page=200&include=transfers&exc...d,updated_at,customdata,event_alert_configuration	200	fe...	d...	413 B	1.38 s	
<input type="checkbox"/> organizations?per_page=25	200	fe...	d...	9.3 kB	1.18 s	
<input type="checkbox"/> virtual-networks?&include=instances,internet_gateway,facility,devices&exclude=	200	fe...	d...	320 B	949 ms	
<input type="checkbox"/> capacity	200	fe...	d...	420 B	922 ms	
<input type="checkbox"/> user?include=emails,number_of_devices,features,messenger_hash	200	fe...	d...	2.4 kB	650 ms	
<input type="checkbox"/> facilities?include=labels,address&sort_by=name&sort_direction=asc	200	fe...	d...	291 B	526 ms	
<input type="checkbox"/> operating-systems?include=labels,default_operating_system_version_id	200	fe...	d...	413 B	479 ms	
<input type="checkbox"/> memberships?&include=user&exclude=&per_page=100	200	fe...	d...	423 B	398 ms	
<input type="checkbox"/> memberships?&include=user&exclude=&per_page=100	200	fe...	d...	411 B	272 ms	
<input type="checkbox"/> metros	200	fe...	d...	507 B	259 ms	
<input type="checkbox"/> ssh-keys	200	fe...	d...	1.6 kB	239 ms	

23,000 spans in one trace???

Scale problems:

- Rails (and other ORMs) tend to introduce N+1 queries
- tons of data for large customers, internal Metal orgs

 **SRE-26** some API endpoints are querying Postgres more than 20,000 times per request

relates to

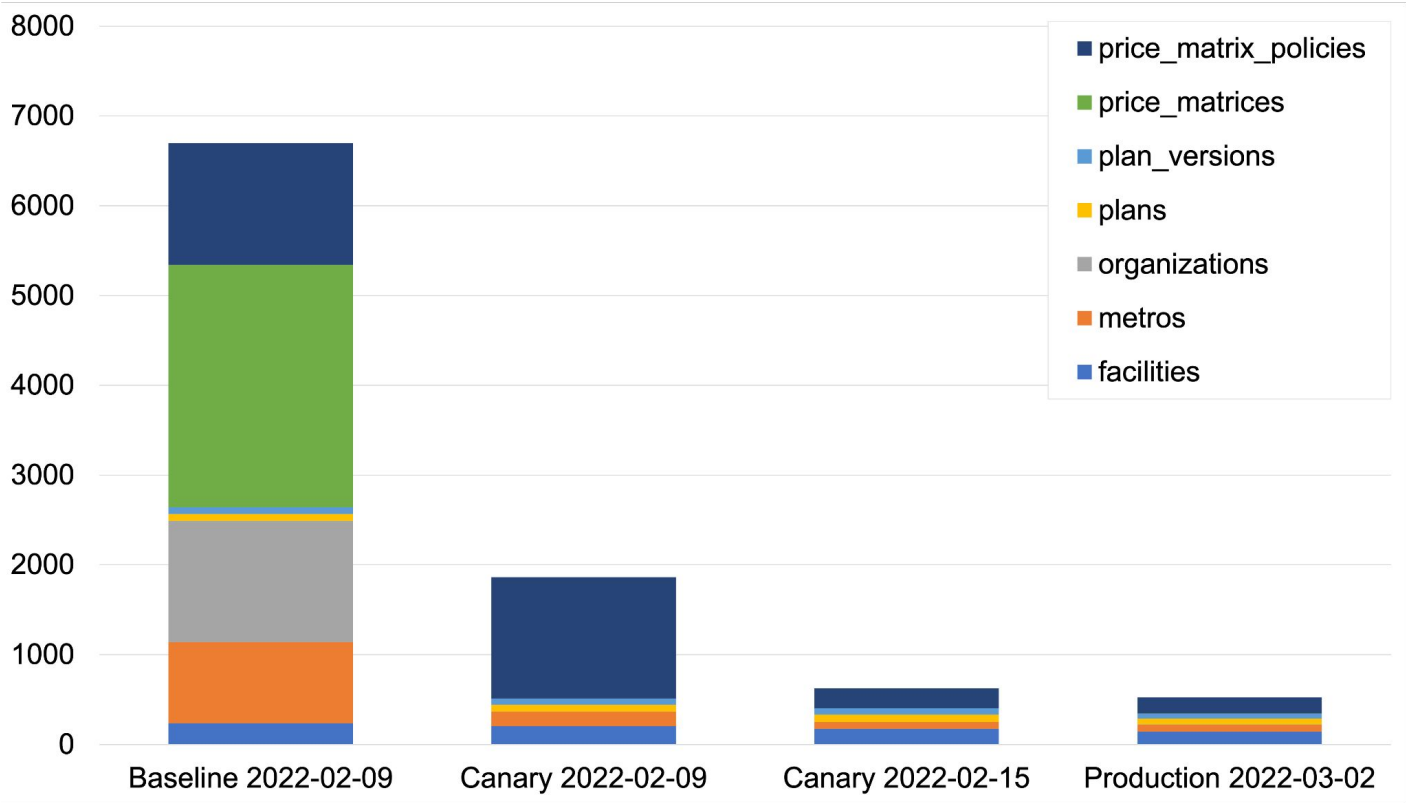
 ENG-7615	Some API Endpoints are Excessively Chatty with Storage Resources (DB, Memcache)			COMPLETE 
 ENG-15474	Limit the number of DB queries per API endpoint to << 50			OPEN 
 ENG-17851	Fetching plans executes thousands of queries per request			COMPLETE 
 ENG-18001	Fetching IP reservations executes too many queries			COMPLETE 

How to minimize the blast radius?

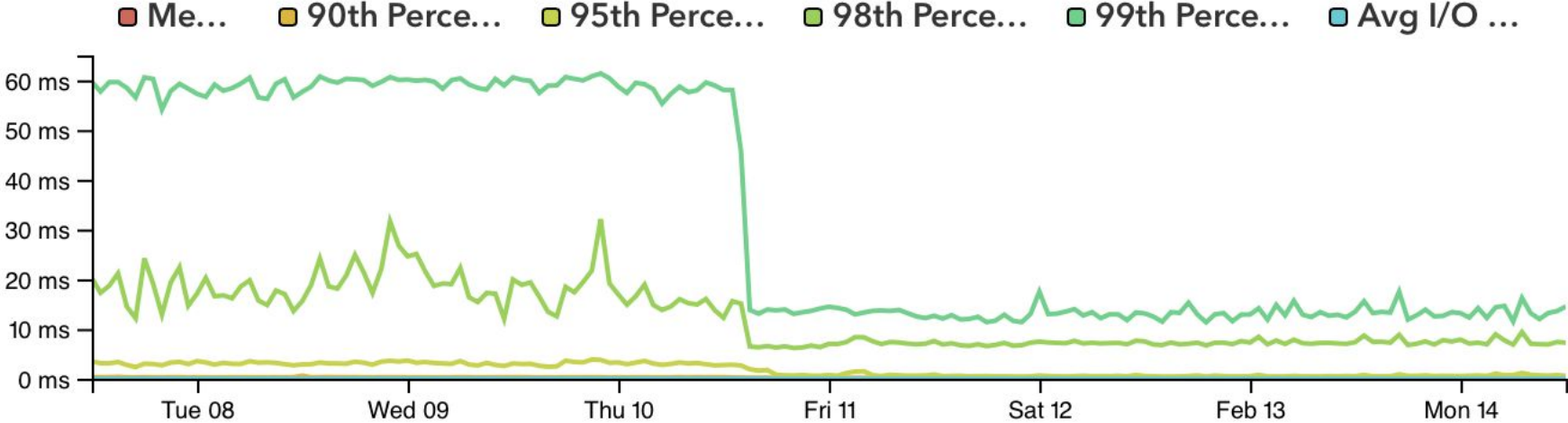
```
def with_active_record_tracing
  # LaunchDarkly requires a user, so only enable ActiveRecord tracing on requests with a user.
  return yield unless current_user && LaunchDarklyClient.instance

  begin
    user_record = { key: current_user.id }
    Trace::ActiveRecordSubscriber.enable(
      LaunchDarklyClient.instance.variation(ACTIVE_RECORD_TRACING_KEY, user_record, false)
    )
  end
  yield
end
ensure
  # restore thread-local state
  Trace::ActiveRecordSubscriber.enable(false)
end
end
```

Number of ActiveRecord Lookups Per Request to /plans



Database optimizations



API Performance: /ips

End of January:

P50(duration_ms)



P95(duration_ms)

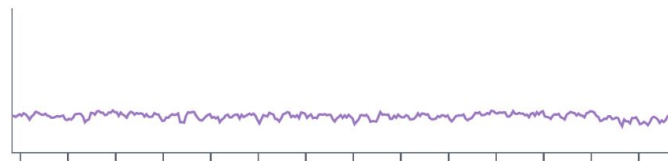


P99(duration_ms)



End of February:

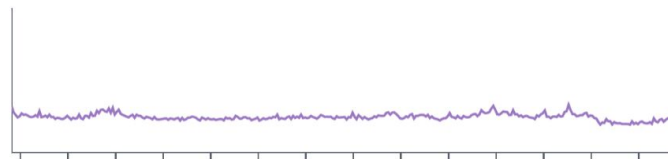
P50(duration_ms)



P95(duration_ms)



P99(duration_ms)



Sociotechnical wins

Teams want tracing for new services

More engineers are canary testing 🐣

Iterating on SLOs



spees 11:55 AM

the following SLOs will now send an alert here as well as in [#sre-bots](#):

- [Customer provisions](#)
- [Customer deprovisions](#)
- [API Latency](#)
- [API Error Rate](#)
- [Authentication](#)

(edited)



2



1



● honeycomb for archetype service



Wednesday, February 23, 2022

3:00 PM to 3:30 PM (30 minutes)



<https://equinix.zoom.us/>

● Otel in github.com/tinkerbell/dhcp



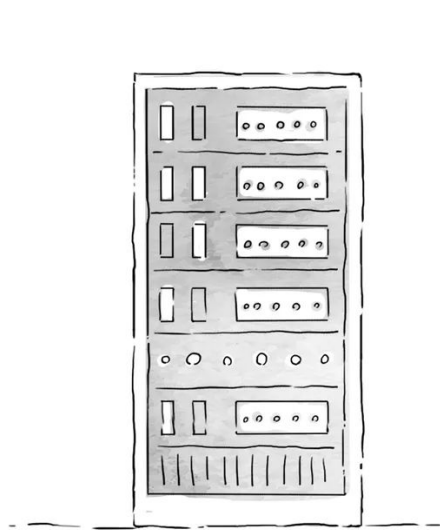
Thursday, February 24, 2022

4:00 PM to 5:00 PM (1 hour)



<https://equinix.zoom.us/>

Tracing Bear Bare Metal

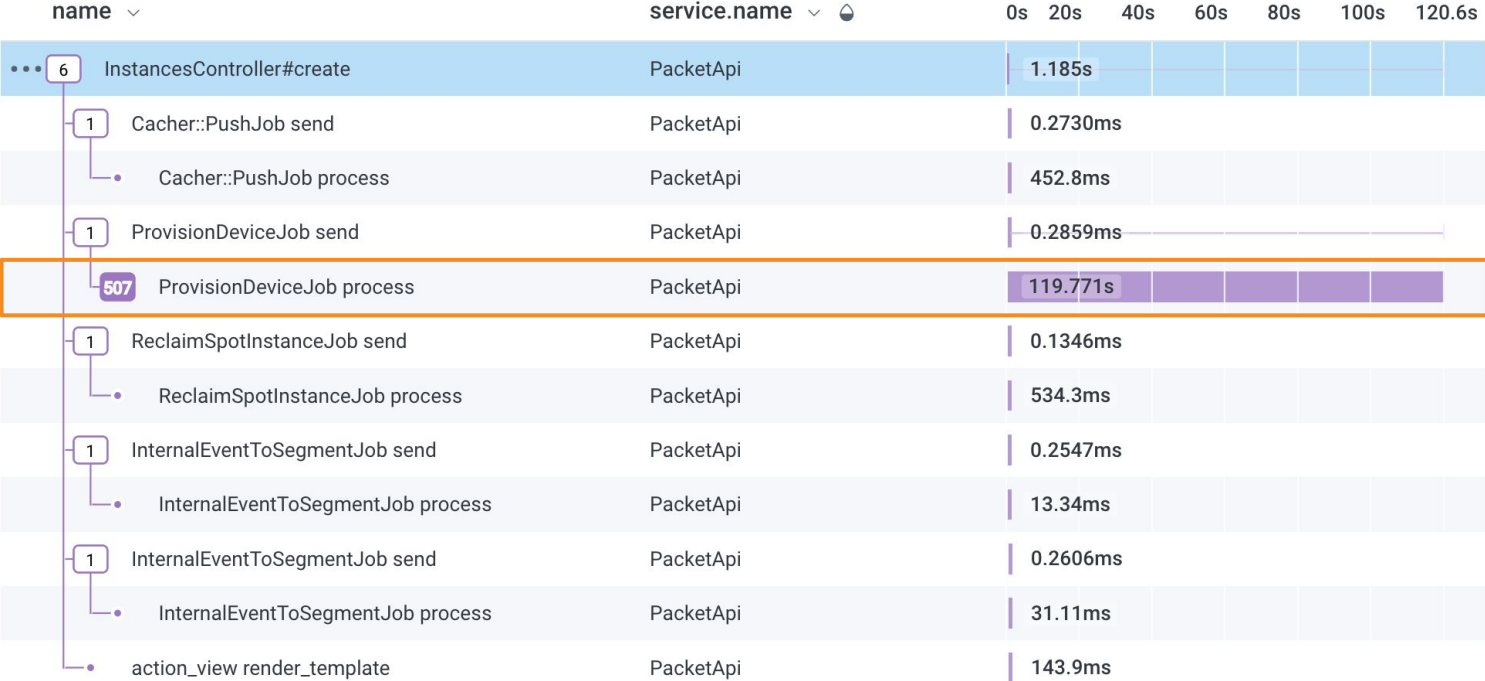


“Bare Metal”



“Bear Metal”

Provisioning a New Instance



Inside the ProvisionDeviceJob

137	ProvisionDeviceJob process	PacketApi	119.771s
	• HTTP POST	PacketApi	7.789ms
1	github.com.tinkerbell...v1.Machine/BootDevice	PacketApi	300.4ms
1	github.com.tinkerbe...Machine/BootDevice	pbnj	54.7µs
	• client.SetBootDevice	pbnj	3.015s
1	github.com.tinkerbell.pbnj.api.v1.Task/Status	PacketApi	300.9ms
	• github.com.tinkerbel...nj.api.v1.Task/Status	pbnj	76.9µs
1	github.com.tinkerbell.pbnj.api.v1.Task/Status	PacketApi	75.11ms
	• github.com.tinkerbel...nj.api.v1.Task/Status	pbnj	74.1µs
1	github.com.tinkerbell.pbnj.api.v1.Task/Status	PacketApi	75.15ms
	• github.com.tinkerbel...nj.api.v1.Task/Status	pbnj	78.5µs
1	github.com.tinkerbell.pbnj.api.v1.Task/Status	PacketApi	75.88ms
	• github.com.tinkerbel...nj.api.v1.Task/Status	pbnj	85.1µs

name	service.name	http.url	0s	20s	40s	60s	80s	100s	120.6s
1 HTTP POST	PacketApi	https://narwhal.....packet.net/ports	314.3ms						
5 PortsController#create	narwhal		10.20ms						
• postgresql.Switch_Load	narwhal		0.1891ms						
1 build switch tasks	narwhal		0.9875ms						
10 perform task: create_port	narwhal		7.291s						
• juniper.edit_config	narwhal		15.65ms						
• juniper.edit_config	narwhal		13.58ms						
• juniper.edit_config	narwhal		8.534ms						
• juniper.edit_config	narwhal		11.39ms						
• juniper.edit_config	narwhal		6.673ms						
• juniper.edit_config	narwhal		6.234ms						
• juniper.edit_config	narwhal		10.23ms						
• juniper.edit_config	narwhal		14.38ms						
• juniper.edit_config	narwhal		83.71ms						
• juniper.edit_config	narwhal		16.20ms						
• postgresql.begin	narwhal		0.1356ms						
• postgresql.Task_Create.insert	narwhal		0.3422ms						
• postgresql.commit	narwhal		1.676ms						

name	service.name	http.url	0s	20s	40s	60s	80s	100s	120.6s	
1 HTTP POST	PacketApi	https://narwhal.....packet.net/acls	313.6ms							
5 AclsController#create	narwhal		9.123ms							
• postgresql.Switch_Load	narwhal		0.2218ms							
1 build switch tasks	narwhal		1.040ms							
2 perform task: create_acl	narwhal		108.6ms							
• juniper.edit_config	narwhal		47.24ms							
• juniper.edit_config	narwhal		61.15ms							
• postgresql.begin	narwhal		0.1419ms							
• postgresql.Task_Create.insert	narwhal		0.3605ms							
• postgresql.commit	narwhal		1.515ms							

"Is it done yet?"

name	service.name	http.url	0s	20s	40s	60s	80s	100s	120.6s
1 HTTP GET	PacketApi	https://narwhal...10-28a6def2697e	310.8ms						
2 TasksController#show	narwhal		5.499ms						
• postgresql.Task_Load	narwhal		0.2170ms						
• postgresql.Switch_Load	narwhal		0.1442ms						
1 HTTP GET	PacketApi	https://narwhal...ba-759364c5b7a6	309.5ms						
2 TasksController#show	narwhal		5.594ms						
• postgresql.Task_Load	narwhal		0.2257ms						
• postgresql.Switch_Load	narwhal		0.1390ms						

add_port_acl

name	service.name	http.url	0s20s	40s	60s	80s	100s	120.6s
1 HTTP POST	PacketApi	https://narwhal...-1%2F0%2F4/acls			313.1ms			
5 PortAclsController#create	narwhal				8.152ms			
• postgresql.Switch_Load	narwhal				0.1728ms			
1 build switch tasks	narwhal				0.8454ms			
2 perform task: add_port_acl	narwhal				52.89ms			
• juniper.edit_config	narwhal				52.62ms			
• juniper.edit_config	narwhal				29.64ms			
• postgresql.begin	narwhal				0.1130ms			
• postgresql.Task_Create.insert	narwhal				0.3265ms			
• postgresql.commit	narwhal				1.405ms			

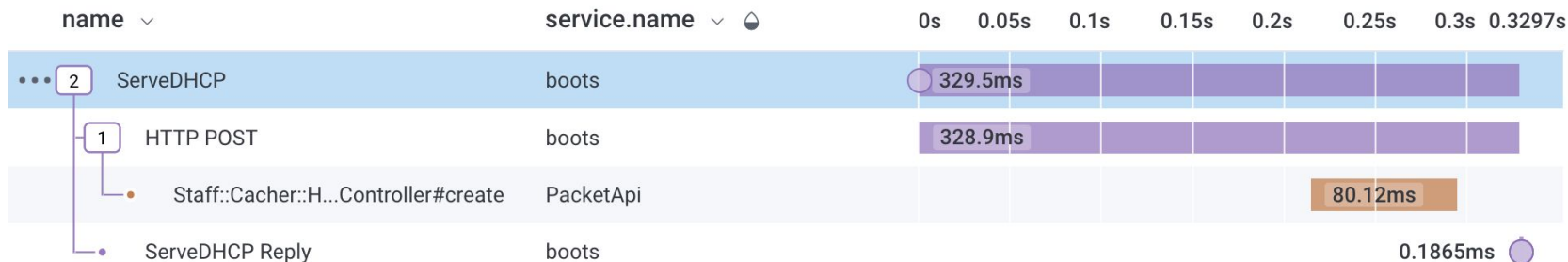
Wrapping up

name	service.name	http.url	0s	20s	40s	60s	80s	100s	120.6s
• HTTP POST	PacketApi	https://kant.edg.../device-metadata						411.9ms	
1 InternalEventToSegmentJob send	PacketApi							0.3327ms	
• InternalEventToSegmentJob process	PacketApi							2.004ms	
1 InternalEventToSegmentJob send	PacketApi							0.2914ms	
• InternalEventToSegmentJob process	PacketApi							1.855ms	
1 ReclaimSpotInstanceJob send	PacketApi							0.1346ms	
• ReclaimSpotInstanceJob process	PacketApi							534.3ms	
1 InternalEventToSegmentJob send	PacketApi							0.2547ms	
• InternalEventToSegmentJob process	PacketApi							13.34ms	
1 InternalEventToSegmentJob send	PacketApi							0.2606ms	
• InternalEventToSegmentJob process	PacketApi							31.11ms	
• action_view render_template	PacketApi							143.9ms	

end of
ProvisionDeviceJob

end of
InstancesController#create

Boots calling PacketAPI



```
tracer := otel.Tracer("DHCP")
ctx, span := tracer.Start(context.Background(), "DHCP Reply",
    trace.WithAttributes(attribute.String("MAC", mac.String())),
    trace.WithAttributes(attribute.String("IP", gi.String())),
    trace.WithAttributes(attribute.String("MessageType", req.GetMessageType().String())),
    trace.WithAttributes(attribute.String("CircuitID", circuitID)),
)
```

Recap Slide!

SRE team is responsible for enablement

Success relies on cultural adoption

Cultural adoption relies on success!

Observability is the enemy of helplessness

Incident management lets you observe your observability

It's feedback loops all the way down

SLOs are easier with customer-centric trace data

span.End()

Reach out!

[@MissAmyTobey](#)

[@shelbyspees](#)

OTel on Equinix Labs

[equinix-labs/otel-cli](#)

[equinix-labs/otel-init-go](#)

OSS Bare Metal Provisioning

[tinkerbelle.org](#)

We're hiring! [careers.equinix.com](#)

Read the blog! [metal.equinix.com/blog](#)