Logs Workshop

Log Management & Analytics at scale!

Agenda

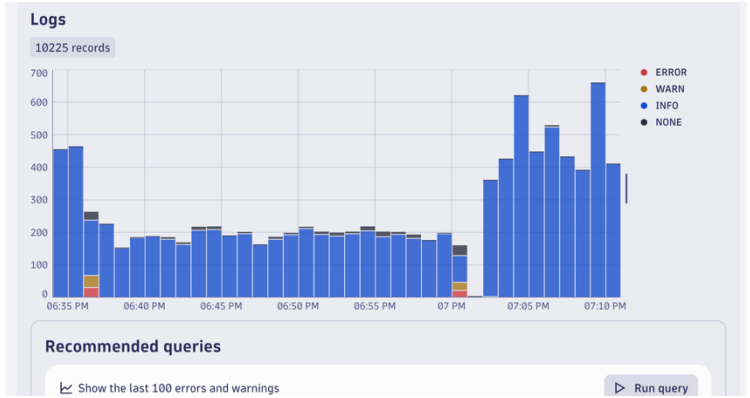
- Log Management and Analytics in Dynatrace provides instant answers to critical questions across:
  - Business
  - DevOps
  - SRE
  - Operations
  - Security

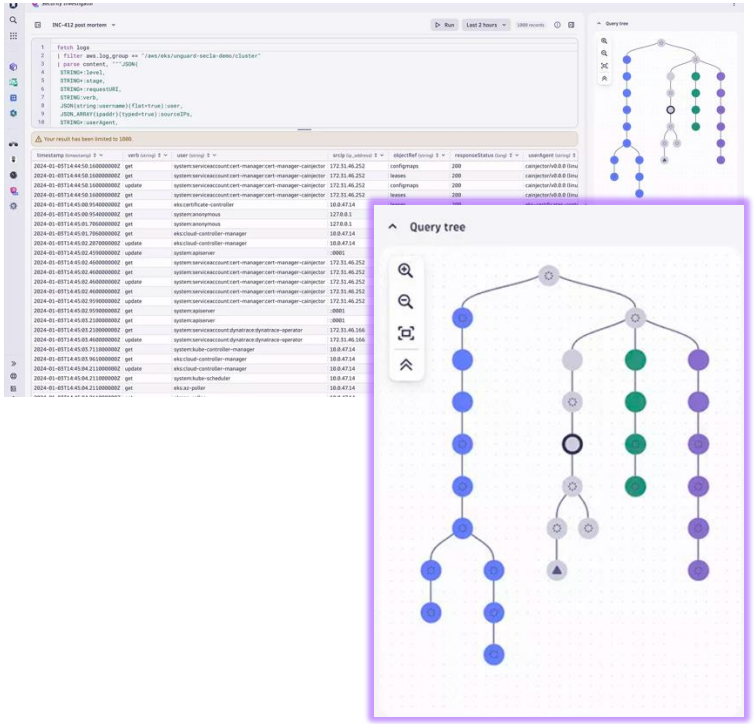
- This technical session will cover best practices and strategies for Log Management and Analytics in Dynatrace, including scalable implementation.
- Attendees will gain knowledge on managing enterprise logging in Dynatrace and leave with actionable insights for their own implementations.

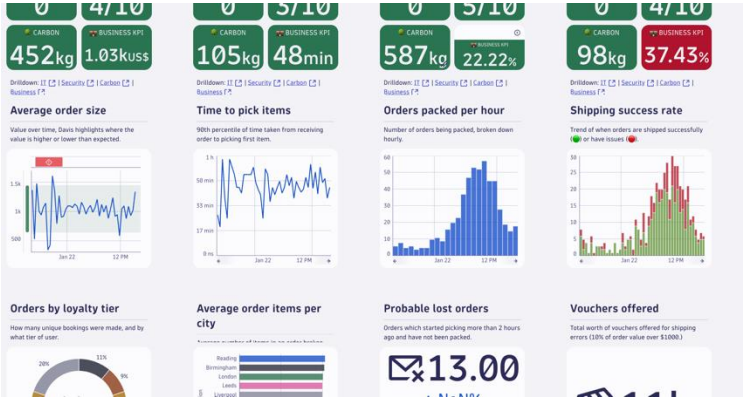
Agenda Items:

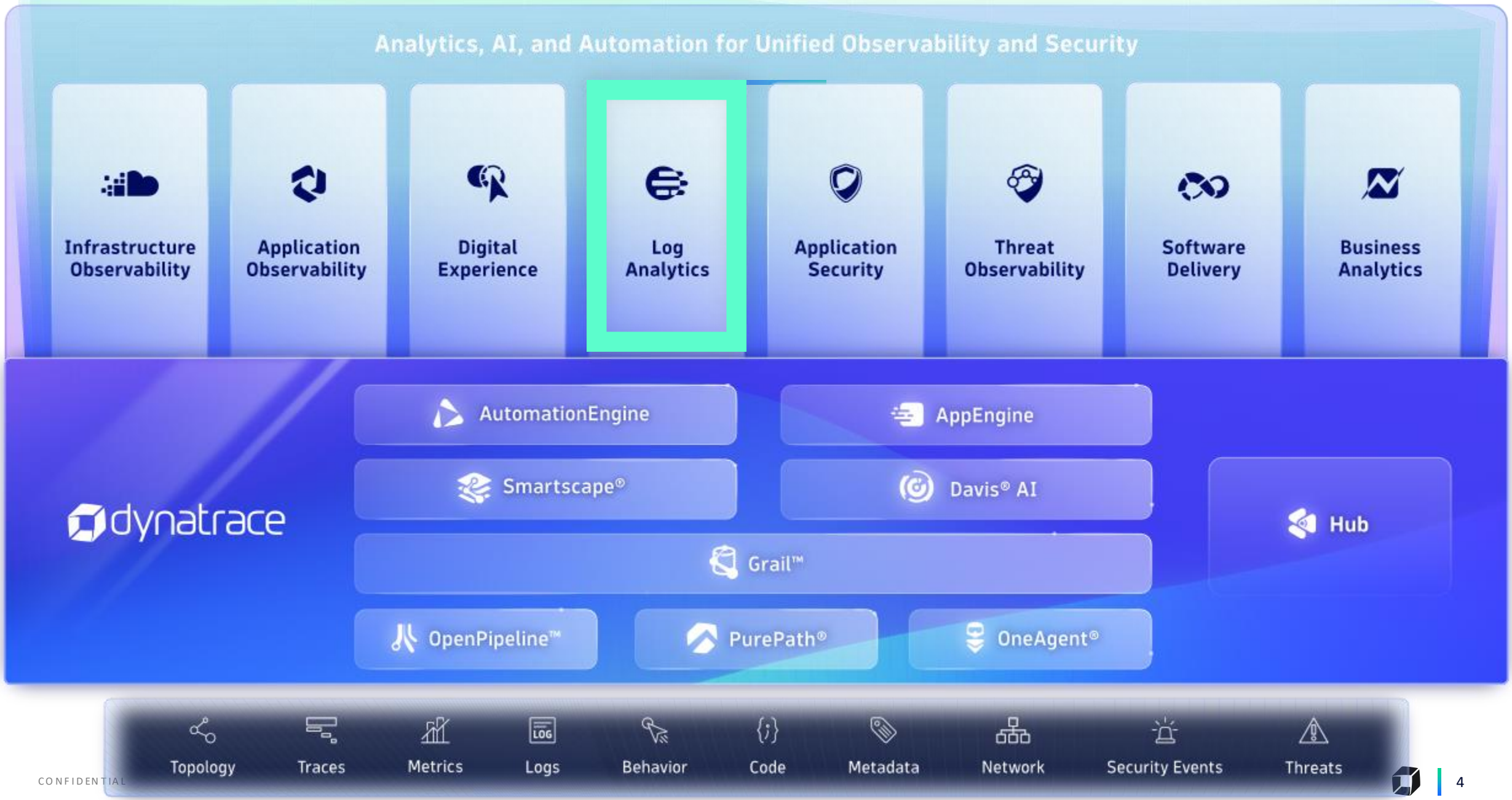
- Log Management and Analytics
- OpenPipeline
- Understanding Query behavior
- Buckets and Segments
- Permissions
- Logs to Metrics / Events
- Dashboarding
- Anomaly Detection
- Live Demo
  - Hands-on Training

Logs are a key data source for Observability, Security and Business Goals








Analytics, AI, and Automation for Unified Observability and Security
Infrastructure Observability
Application Observability
Digital Experience
Log Analytics
Application Security
Threat Observability
Software Delivery
Business Analytics
dynatrace
AutomationEngine
AppEngine
Smartscape®
Davis® AI
Hub
Grail™
OpenPipeline™
PurePath®
OneAgent®
Topology
Traces
Metrics
Logs
Behavior
Code
Metadata
Network
Security Events
Threats

Discussion Question

What are necessary capabilities for a Log Management and Analytics Solution?

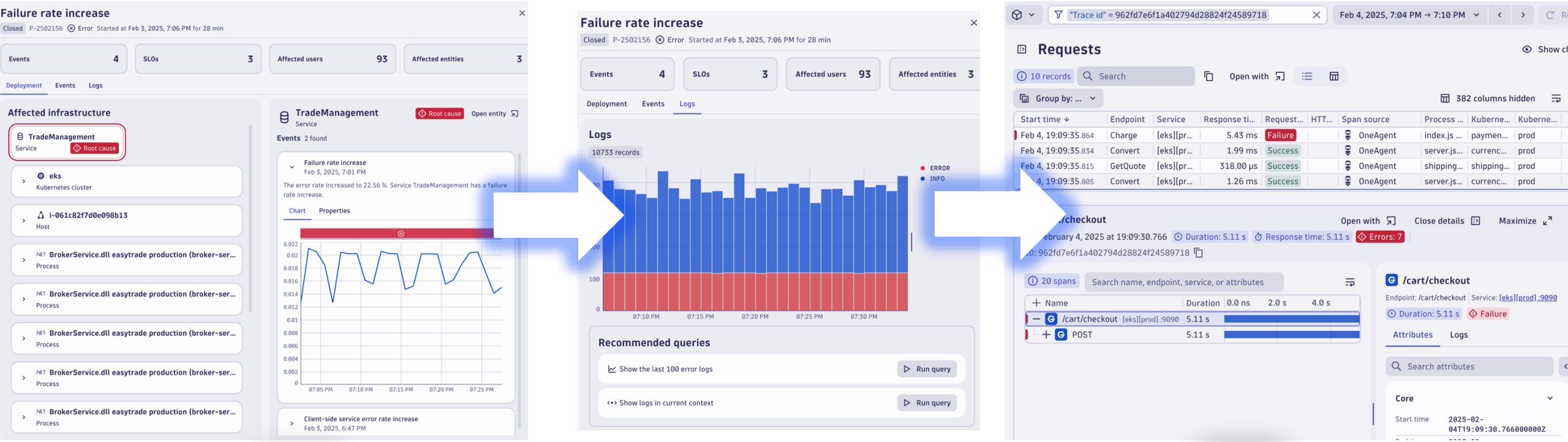
Discussion Question

What are necessary capabilities for a Log Management and Analytics Solution?

- Ingest logs from many sources
- Retain logs for selected timeframe
- Permissions and Access Controls
- Scalability
- Filtering
- Alerts from log data
- Automatic parsing for standard logs
- Deep dive analytics
- Data security

Easily get answers from Logs!

Logs in context reduces MTTR and provides instant answers



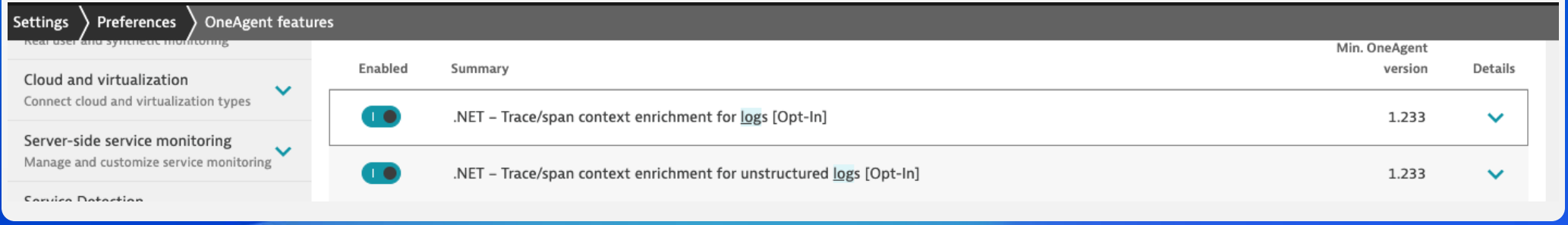
> We used to spend hours manually searching through metrics, logs, and traces to piece together insights about user experience. Now, this takes minutes or seconds.

Discussion Question

How does Dynatrace enrich logs with trace context?

Discussion Question

How does Dynatrace enrich logs with trace context?



Automatically for popular logging frameworks!
Just toggle it on in OneAgent Features

Dynatrace also provides guides for enrichment for other logging methods or for use with OpenTelemetry

Log Management & Analytics workflow

Fast insights without the painful data management

Log sources

- Kubernetes
- Technology
- Applications
- Host
- Process
- System

aws

Custom API integrations

- Log forwarders
- OpenTelemetry
- Logstash
- FluentD
- Fluent Bit

Log ingestion components

One Agents
Auto-detection

Log ingestion API
JSON, Syslog, OTLP

Open Pipeline

- Enrich logs with context
- Mask sensitive data
- Set Security Context
- Extract metrics and business events data
- Route to Grail Buckets
- Parse JSON for simplified analytics

Data Lakehouse

Grail™

Low cost ingest and retention

No storage tiers, schemas, or indexes to manage

End-to-end context

Answers

Davis® AI

Automated
- Entity discovery
- Contextualization
- Baselining
- Anomaly/issue detection

Log insights
- Integrated throughout the platform apps for easy troubleshooting

Simple analysis
- Easy filtering and log exploration
- Narrow by attributes and timeframe

DQL
Dynatrace Query Language

Advanced analytics
- Familiar query language
- Schema-on-read 10X faster parsing than RegEx

Discussion Question

What is OpenPipeline?

OpenPipeline is the data handling solution (pipeline) for Dynatrace to seamlessly ingest and process data.

Used for configuring processing, transforming, metric or data extraction, permissions and storage.

Discussion Question

What is OpenPipeline?

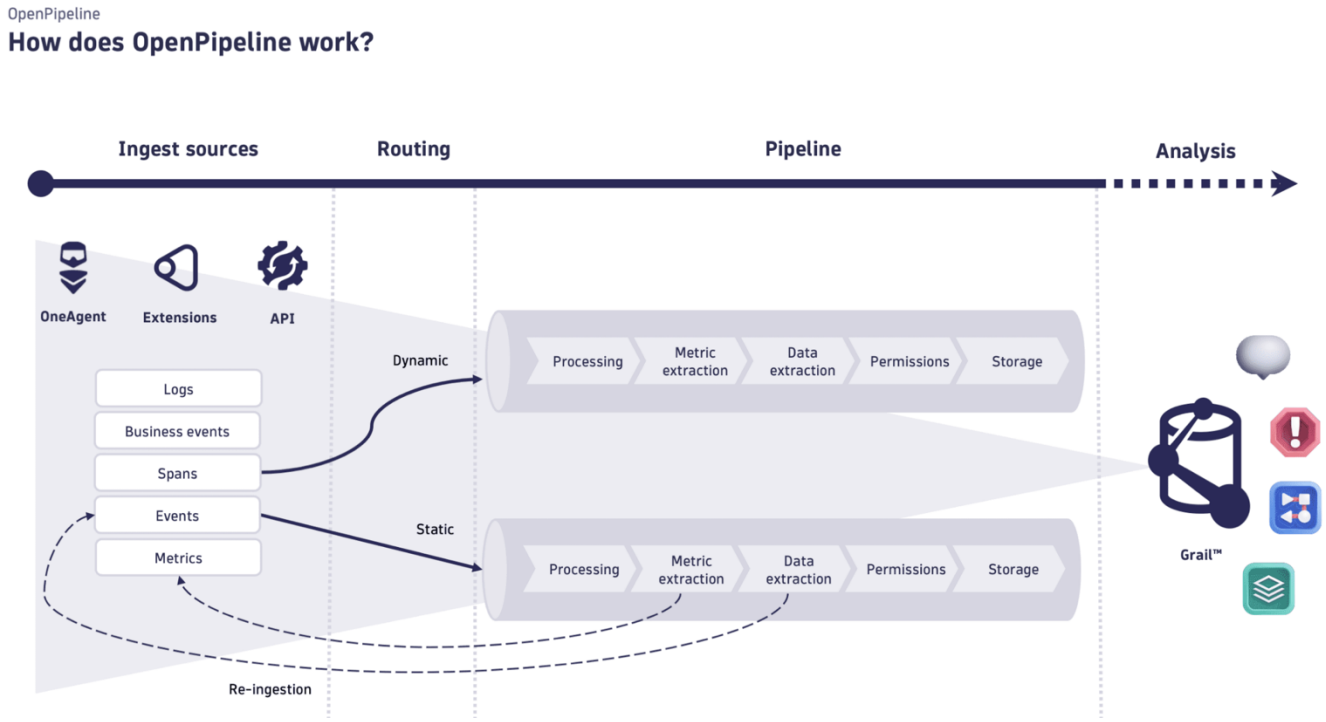
Unified solution for ingestion and processing of different data types, not just logs.

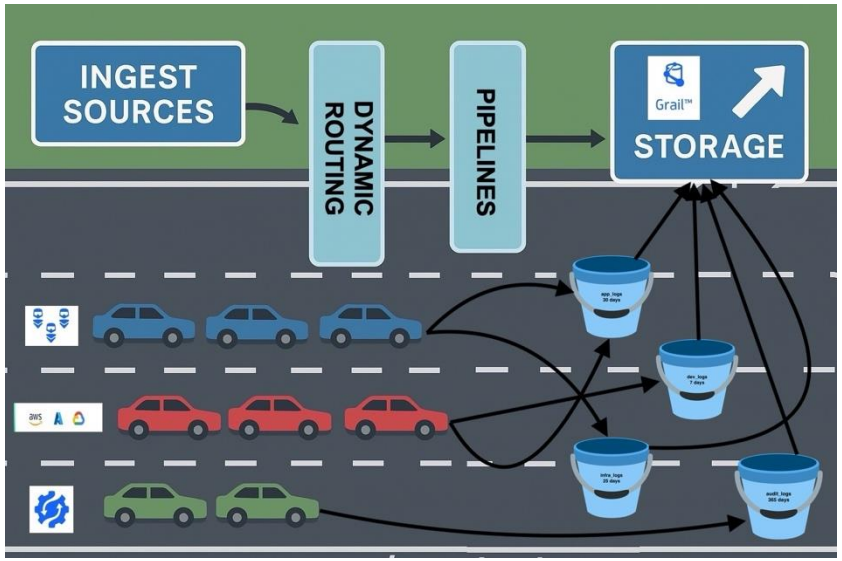
OpenPipeline

- OpenPipeline is used to normalize log records with processing rules, create metrics, create events, assign security context, and define storage to buckets

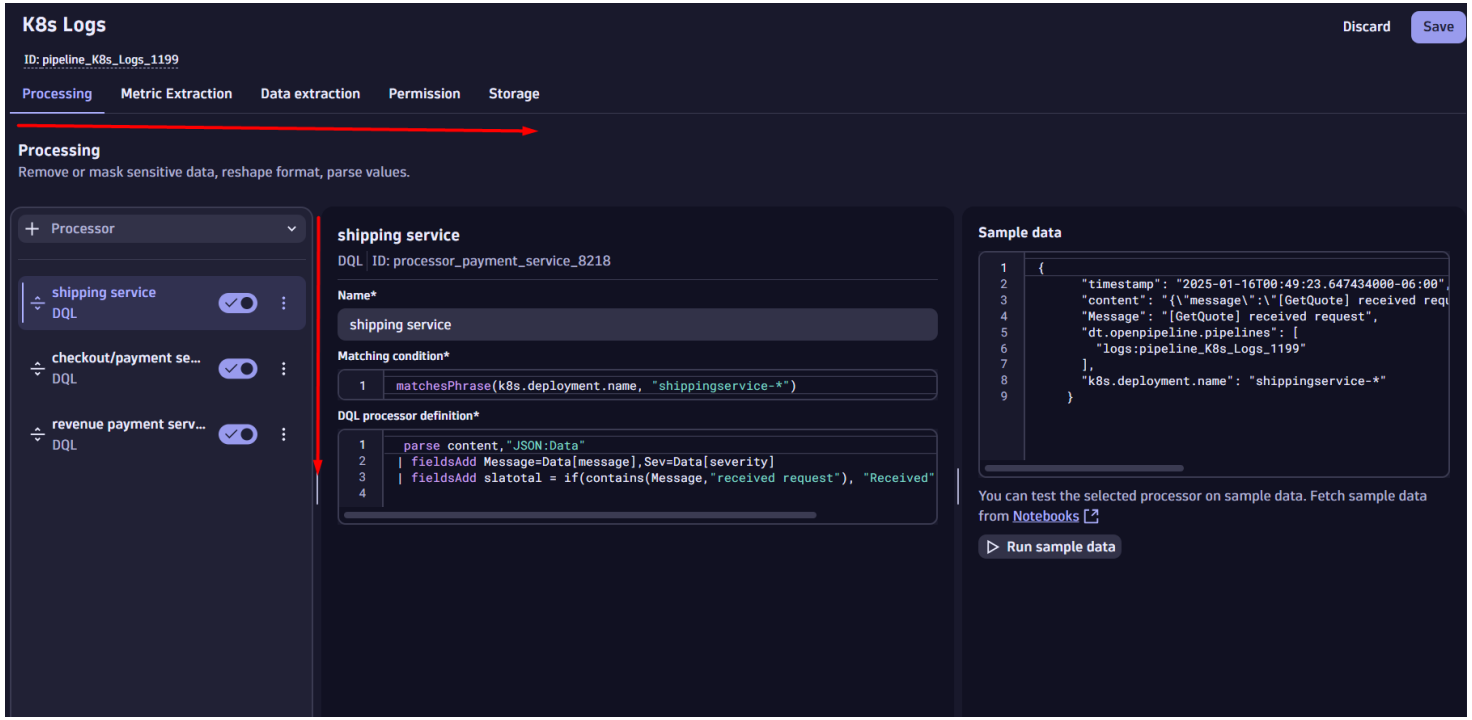
- Mapping incoming log records to specific buckets is done via OpenPipeline.

- OpenPipeline should be strategized in a similar manner to buckets.
  - 1:1 pipeline to bucket
  - 1:many pipeline to buckets





OpenPipeline

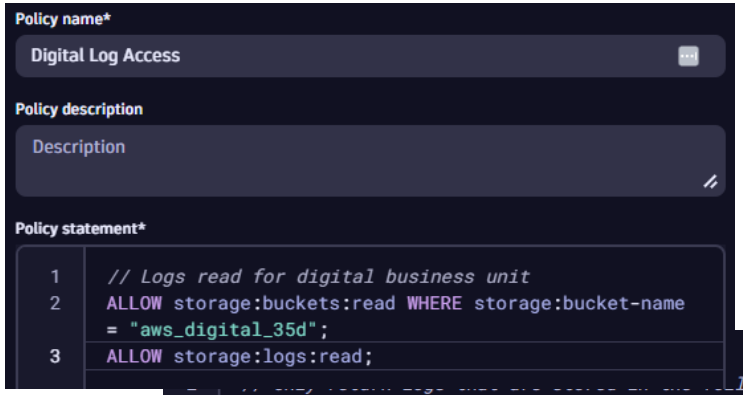


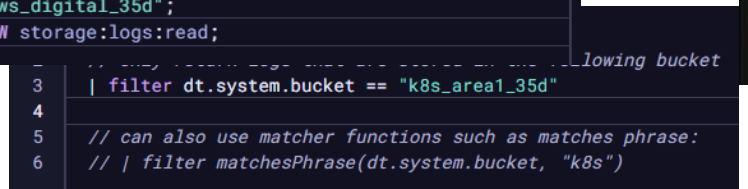
- OpenPipeline works in sequential mode, first vertical and then horizontal for each pipeline.
- Processing allows you to parse, add/remove/rename fields and drop record.
- Metric extraction allows you to create counter or value metric.

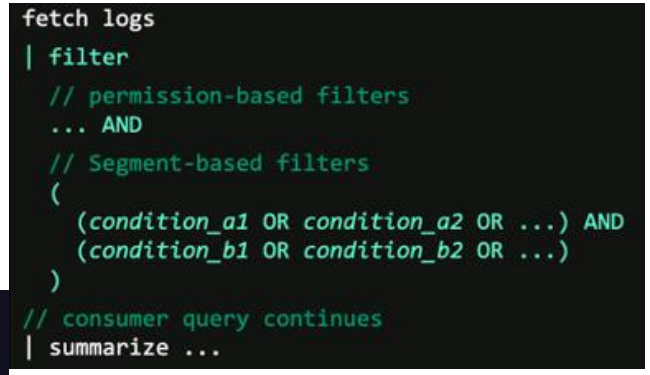
Scaling Log Analytics

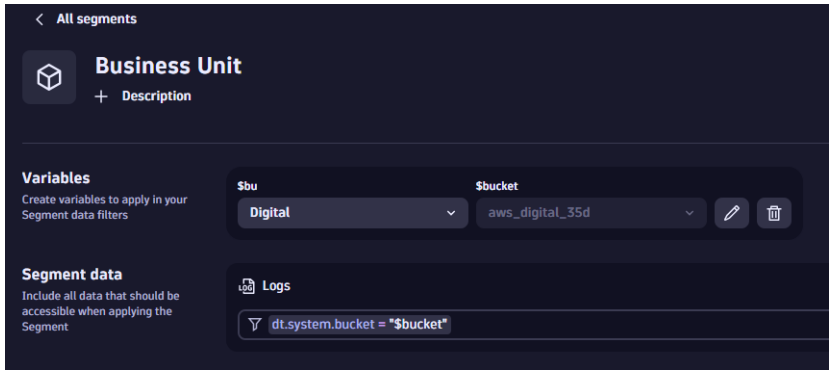
Understanding Query Behavior

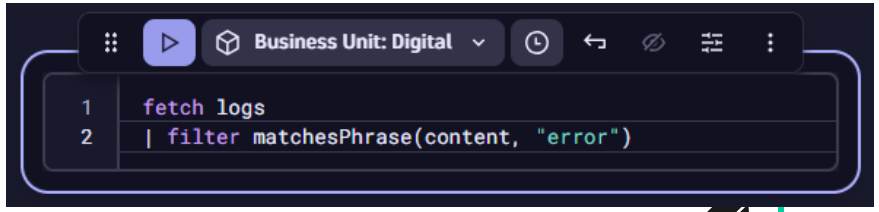
- DQL and in app queries **will scan all buckets** a user has permission to read by default.
- Dynatrace **automatically optimizes queries** based on filters
- To realize performance and cost optimizations from a bucket strategy, queries must be filtered to relevant buckets.
  - This can be accomplished in three ways:
    - Only give users **access to relevant buckets***
    - Apply **dt.system.bucket filter** into DQL query
    - Utilize **segments** for targeting appropriate buckets*











Discussion Question

What is a bucket in Dynatrace?

Buckets are a way to organize data in Dynatrace.

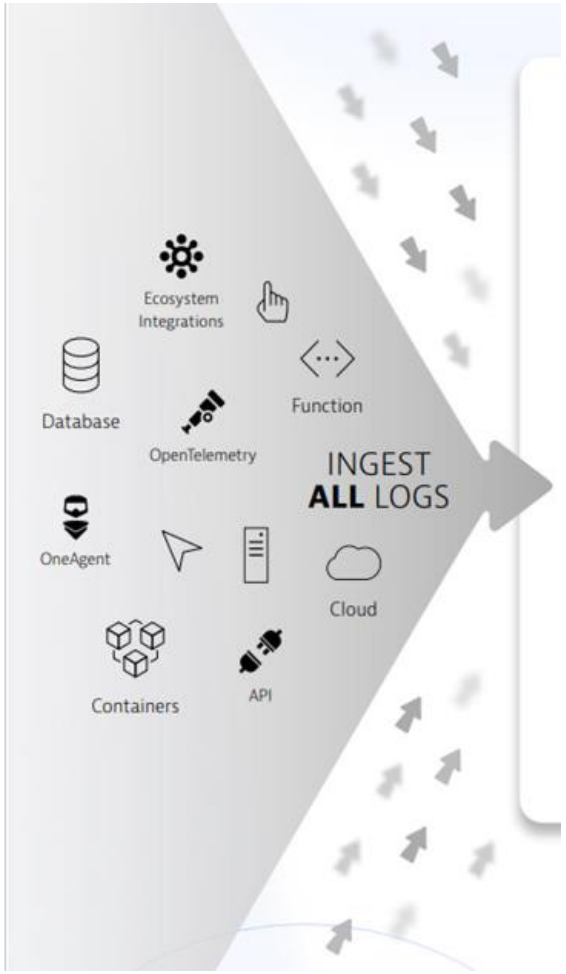
Buckets can be used for data retention, separating use cases, and access permissions.

Discussion Question

What is a bucket in Dynatrace?

Buckets can be used to improve query performance by reducing query execution time and the scope of data read.

Scaling Log Analytics



default_logs
retention: 35d
Size: 10tb/day

```
1  fetch logs
2  | filter isnotnull(k8s.deployment.name)
3  | summarize Count = count(), by: {k8s.deployment.name}
4  | sort Count desc
5  | limit 10
6
```

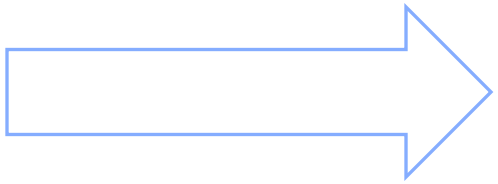
10 rec... Executed at: 1/12/2025, 22:09:23, Timeframe: 1/5/2025, 22:09:21 - 1/12/2025, 22:09:21, Scanned bytes: 500 GB

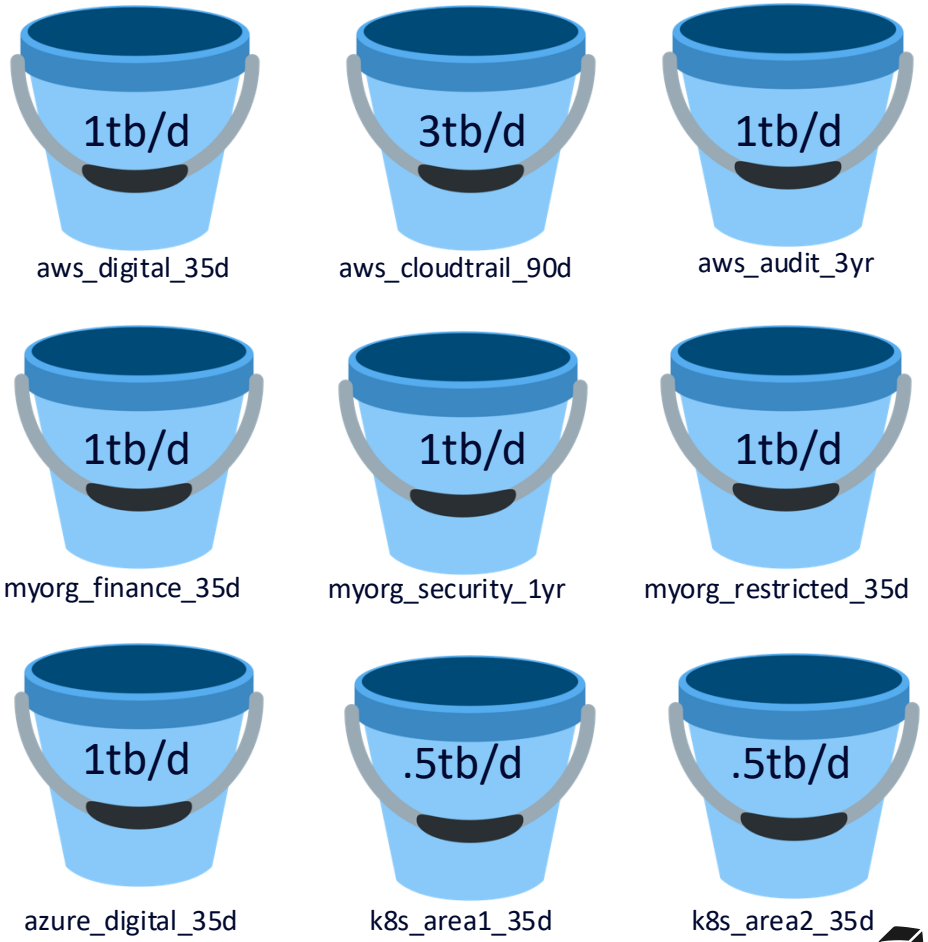
k8s.deployment.name	Count
unguard-malicious-load-generator-*	39,054,339
unguard-user-simulator-*	31,541,074
offerservice-*	14,599,336
calculationservice-*	13,817,824
ingress-nginx-controller-*	13,233,959
pvc-out-of-space-kaniko-big-image-push-*	10,944,380
currencyserviceproxy-*	9,513,482
frontendreverseproxy-*	5,521,804
astroshop-imageprovider-*	5,030,592
aggregator-service-*	4,461,697

Scaling Log Analytics

default_logs
retention: 35d
Size: 10tb/day

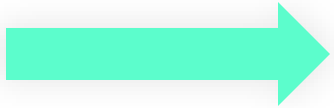
Partition log data
by scaling buckets





Define your bucket scaling strategy

- Developing a tailored data partitioning (bucket) strategy should gather the following information:
  - Ingest sources
  - Estimated Ingest volume
  - Retention desires
  - Compliance Requirements (audit)
  - Sensitive data & masking
  - Permission requirements



<u>Bucket strategy best practice:</u>

- Define a strict naming standard such as <provider>_<type>_<retention> or <provider>_<bu>_retention.
- Consider Business units or App Groupings
- Dedicate buckets to your largest ingest sources.
- Ingest of 1tb per day is ideal.
- Additional permissions such as restricted access can be built on top of the standard: aws_hr_restricted_35d & aws_hr_nonrestricted_35d

Additional bucket considerations

- 80 bucket limit per tenant by default
- Can be significantly increased based on ingest volume – contact support

- Buckets exist for all data types in Grail
- Logs, events, BizEvents, spans, RUM, Metrics etc..
- Strategy can be replicated for other high volume data types such as BizEvents or spans.

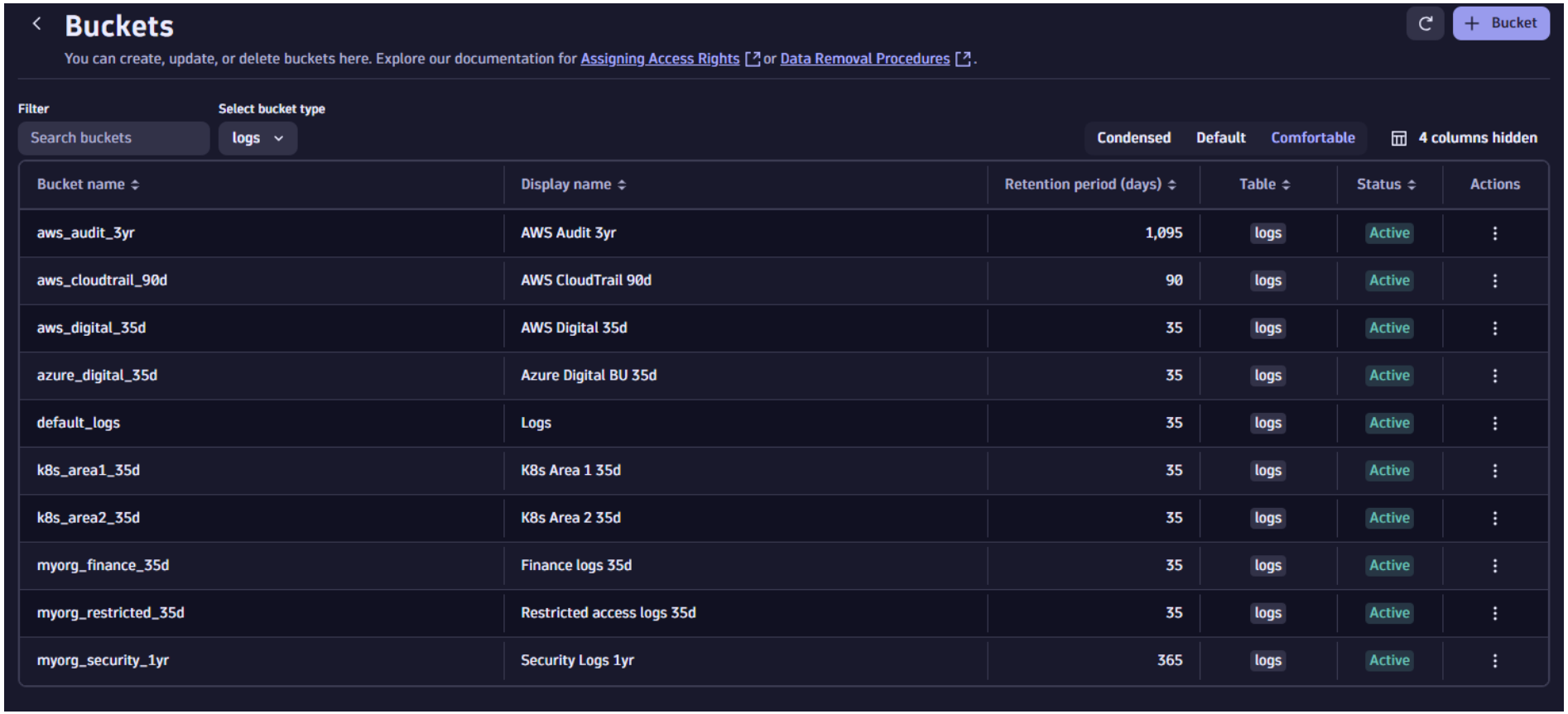
- More buckets = more setup / admin management – strike a balance

- Ensure bucket design aligns with available attributes on log records (host group, account / Sub-IDs, log source)

- Entity properties are not available for lookups to map records to buckets.

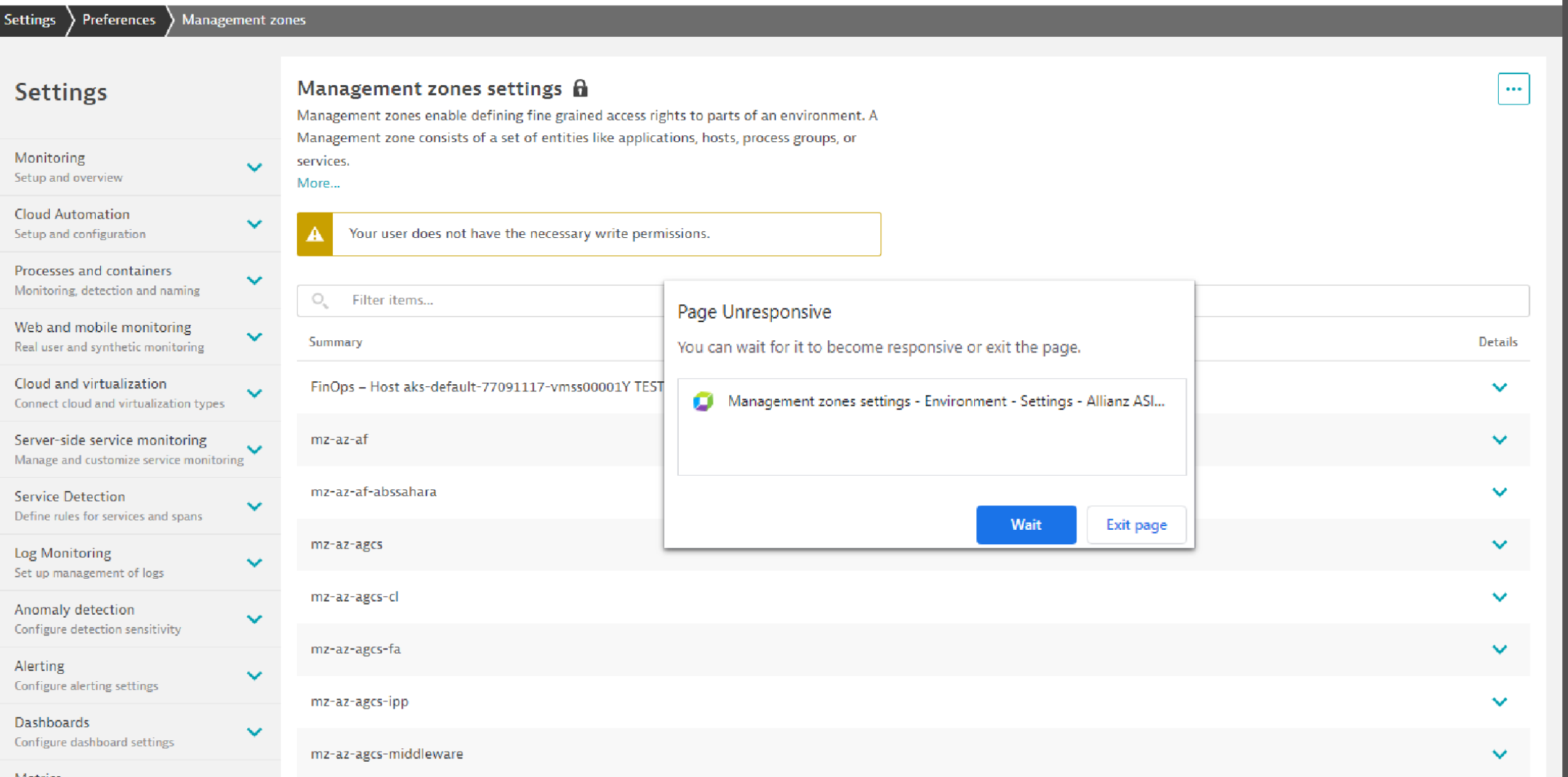
How to create buckets

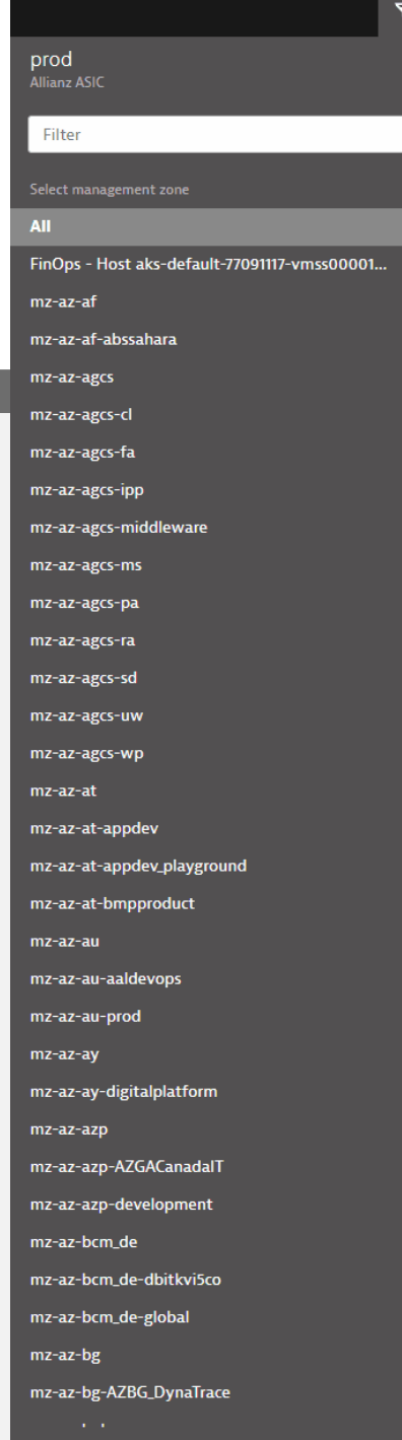
- Storage Management App or Platform API (/platform/storage/management/v1)



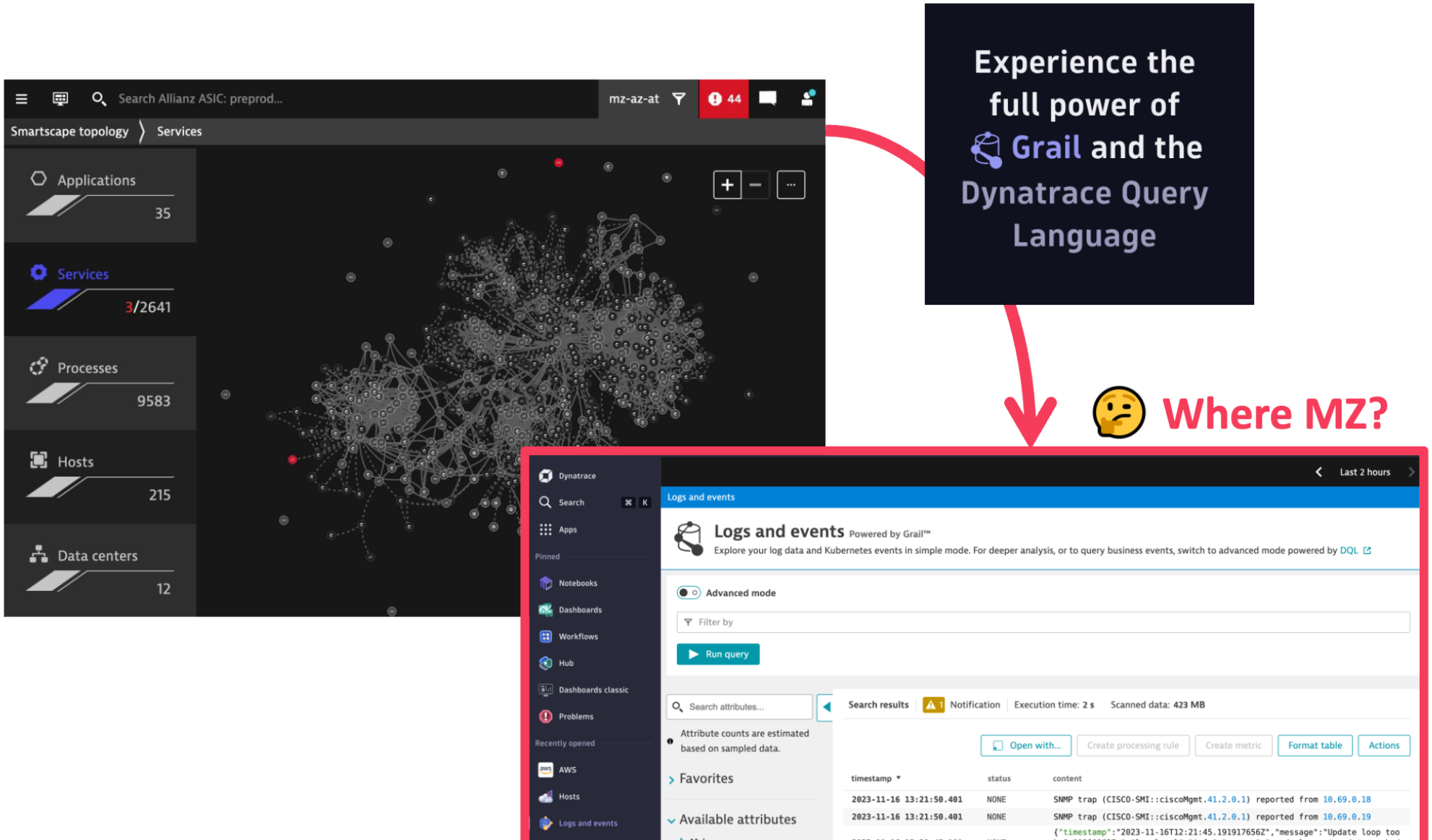
Segments

Why Segments?

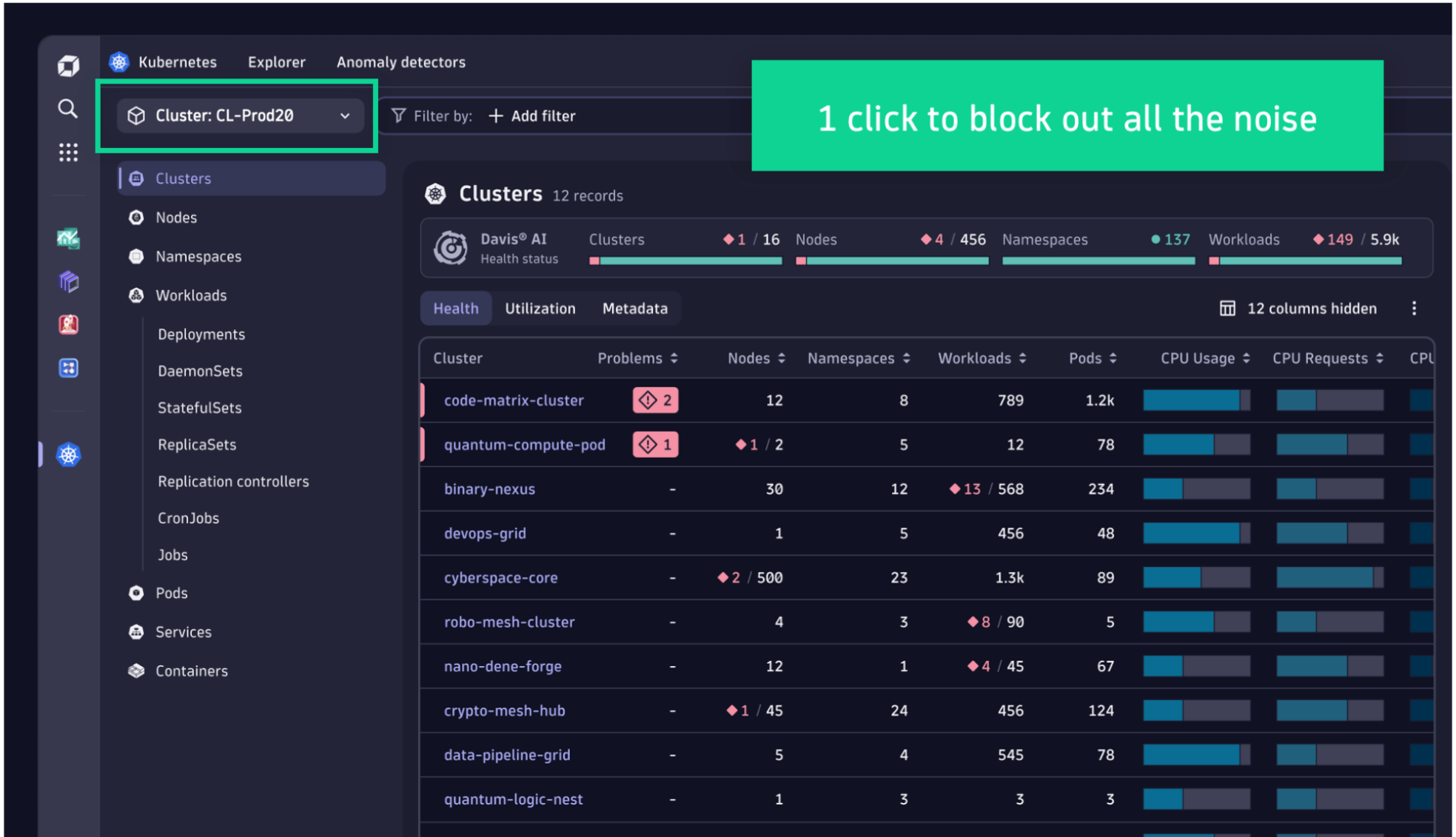




Why Segments?



Segments: Preconfigured Dynamic Filtering



Simple Segment

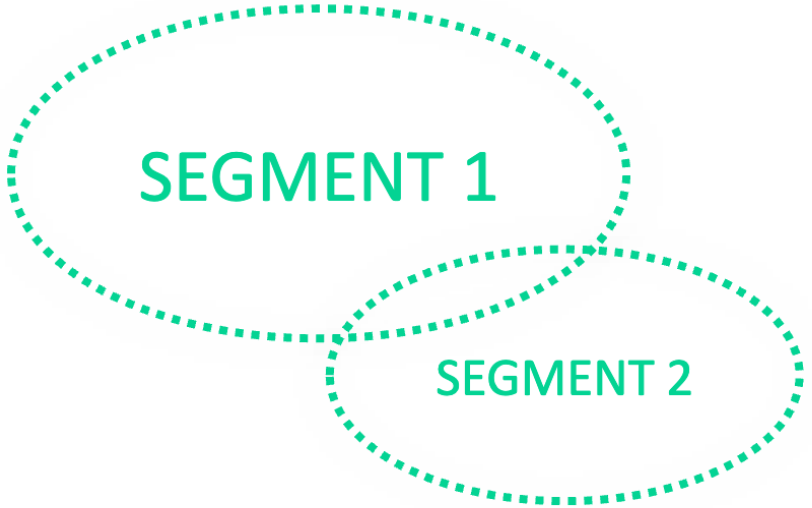
Show me entities, logs, events, etc. of my application

- Equivalent to single MZ

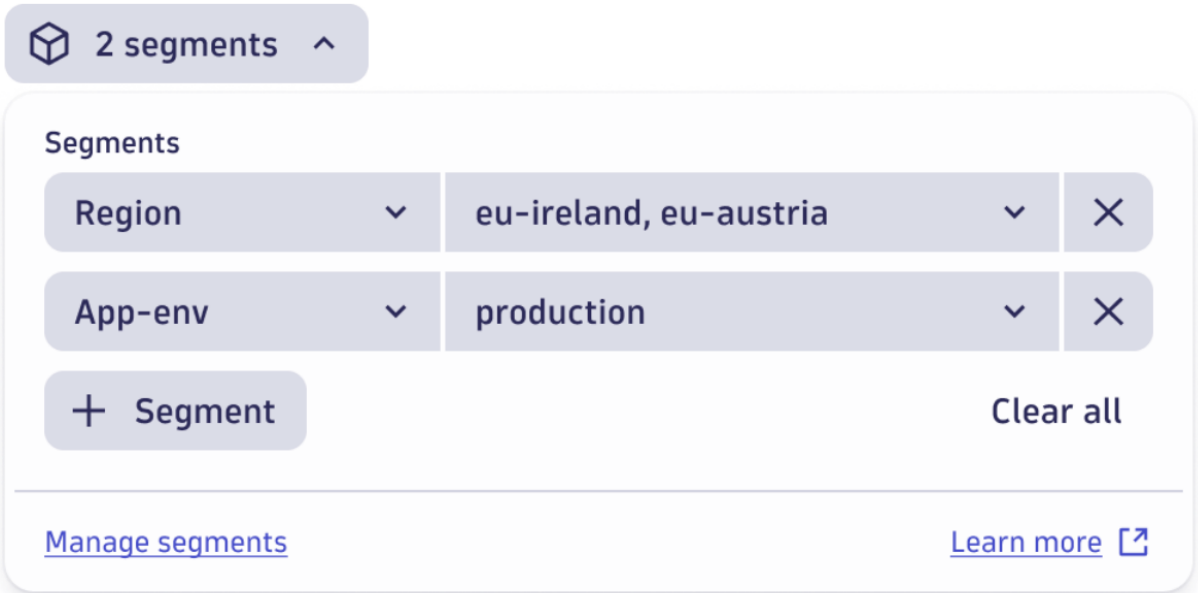
Multiple Segments

Show me data of my application from a specific infrastructure

- Intersection of multiple segments (eu-ireland OR eu-austria) AND production



How it all works



1) Segment IDs and values are passed to query API

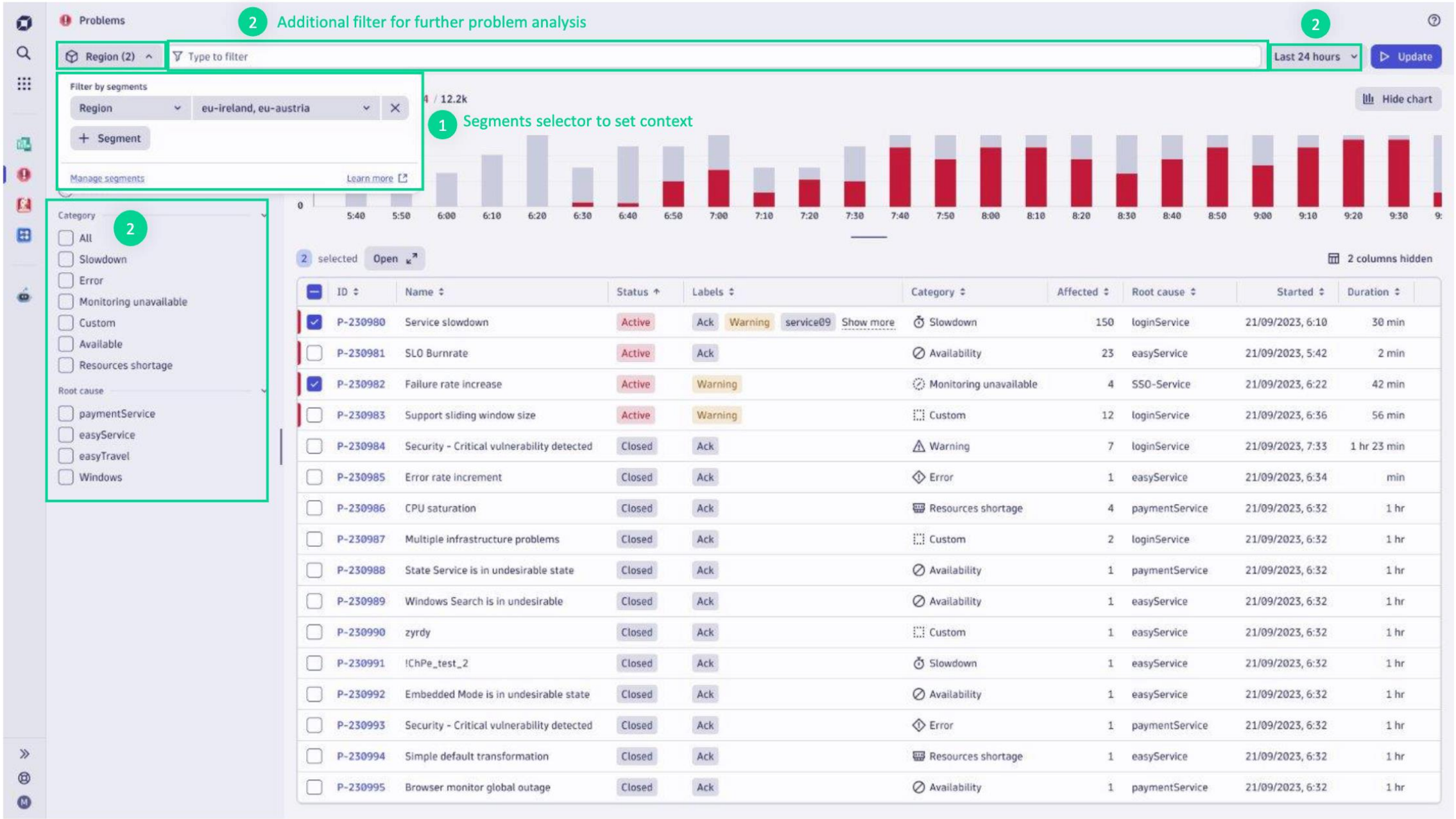
FiRmGMvWzPY(eu-ireland, eu-austria), HCkccp8aAIU(production)

2) Grail filters consumer queries with conditions defined in segments

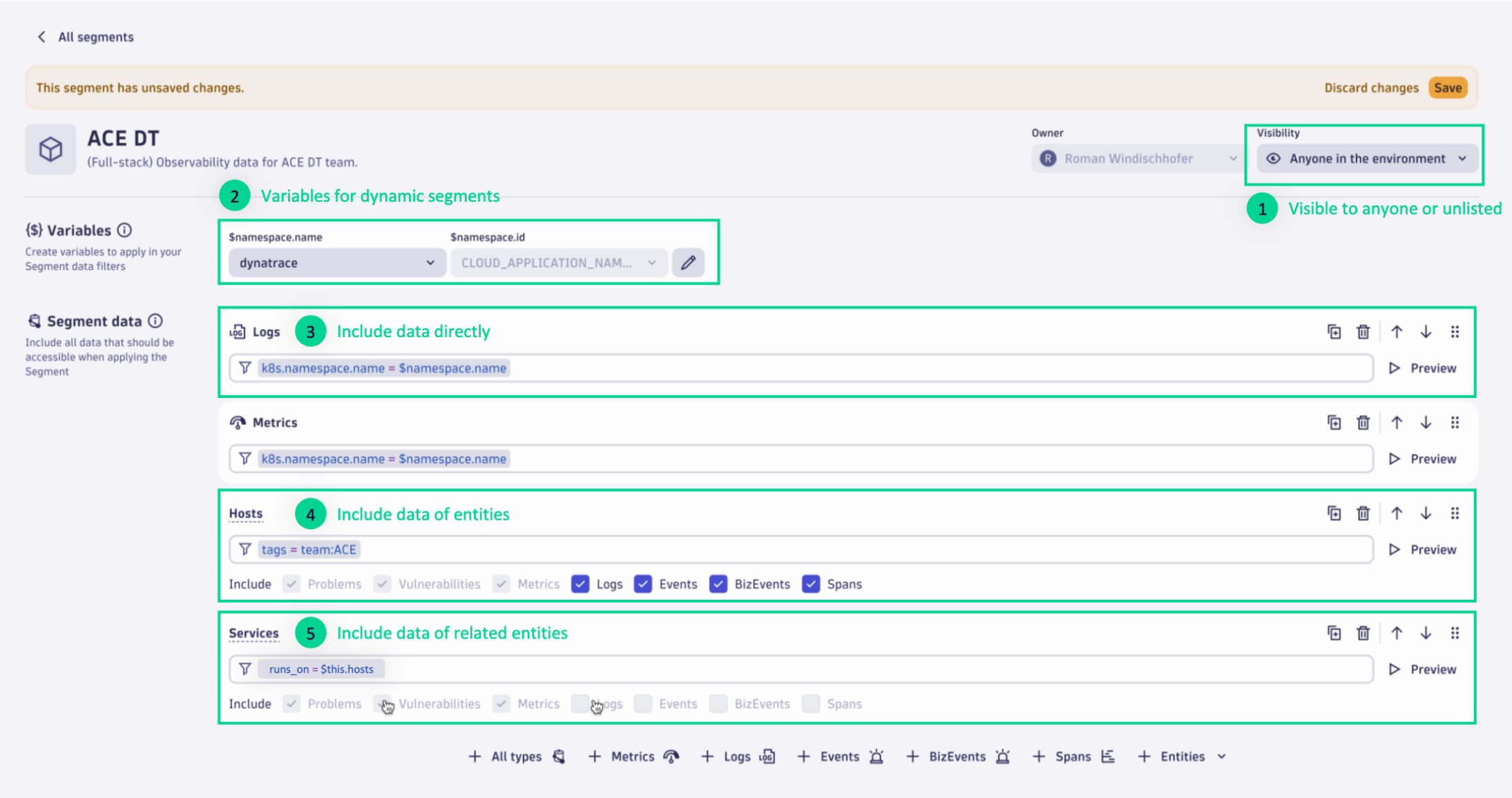
```
fetch logs
| filter
  // permission-based filters
  ... AND
  // Segment-based filters
  (
    (condition_a1 OR condition_a2 OR ...) AND
    (condition_b1 OR condition_b2 OR ...)
  )
// consumer query continues
| summarize ...
```

Transparent filter conditions

Segments and in-app filters


Problems
2
Additional filter for further problem analysis
Region (2)
Type to filter
2
Last 24 hours
Update
Filter by segments
Region
eu-ireland, eu-austria
Segment
Manage segments
Learn more
1
Segments selector to set context
4 / 12.2k
Hide chart
0
5:40
5:50
6:00
6:10
6:20
6:30
6:40
6:50
7:00
7:10
7:20
7:30
7:40
7:50
8:00
8:10
8:20
8:30
8:40
8:50
9:00
9:10
9:20
9:30
Category
2
All
Slowdown
Error
Monitoring unavailable
Custom
Available
Resources shortage
Root cause
paymentService
easyService
easyTravel
Windows
2 selected
Open
2 columns hidden
ID
Name
Status
Labels
Category
Affected
Root cause
Started
Duration
P-230980 Service slowdown Active Ack Warning service09 Show more Slowdown 150 loginService 21/09/2023, 6:10 30 min
P-230981 SLO Burnrate Active Ack Availability 23 easyService 21/09/2023, 5:42 2 min
P-230982 Failure rate increase Active Warning Monitoring unavailable 4 SSO-Service 21/09/2023, 6:22 42 min
P-230983 Support sliding window size Active Warning Custom 12 loginService 21/09/2023, 6:36 56 min
P-230984 Security - Critical vulnerability detected Closed Ack Warning 7 loginService 21/09/2023, 7:33 1 hr 23 min
P-230985 Error rate increment Closed Ack Error 1 easyService 21/09/2023, 6:34 min
P-230986 CPU saturation Closed Ack Resources shortage 4 paymentService 21/09/2023, 6:32 1 hr
P-230987 Multiple infrastructure problems Closed Ack Custom 2 loginService 21/09/2023, 6:32 1 hr
P-230988 State Service is in undesirable state Closed Ack Availability 1 paymentService 21/09/2023, 6:32 1 hr
P-230989 Windows Search is in undesirable Closed Ack Availability 1 easyService 21/09/2023, 6:32 1 hr
P-230990 zyrdy Closed Ack Custom 1 easyService 21/09/2023, 6:32 1 hr
P-230991 !ChPe_test_2 Closed Ack Slowdown 1 easyService 21/09/2023, 6:32 1 hr
P-230992 Embedded Mode is in undesirable state Closed Ack Availability 1 easyService 21/09/2023, 6:32 1 hr
P-230993 Security - Critical vulnerability detected Closed Ack Error 1 paymentService 21/09/2023, 6:32 1 hr
P-230994 Simple default transformation Closed Ack Resources shortage 1 easyService 21/09/2023, 6:32 1 hr
P-230995 Browser monitor global outage Closed Ack Availability 1 paymentService 21/09/2023, 6:32 1 hr

Building a Segment



20 include blocks per segment

1 filter for data of entities (logs, events, spans)

No limit for problems, vulnerabilities, metrics

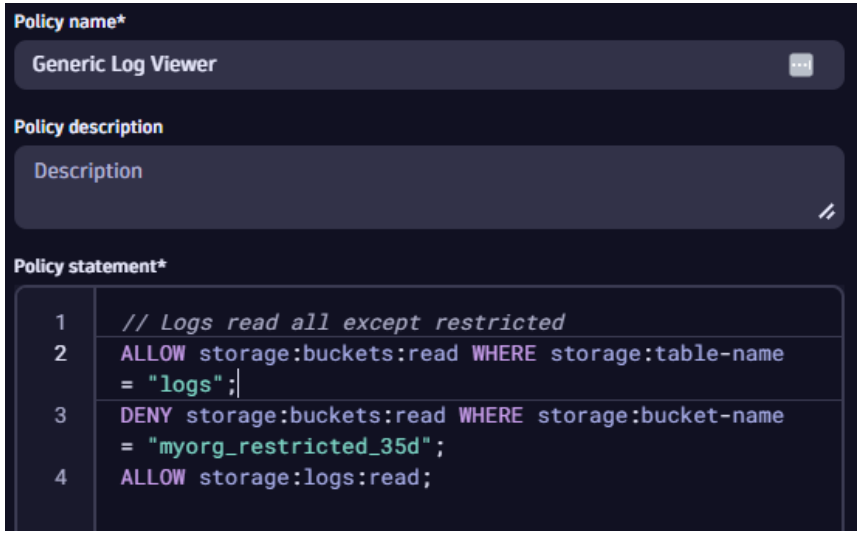
1 topology traversal step only

Permissions

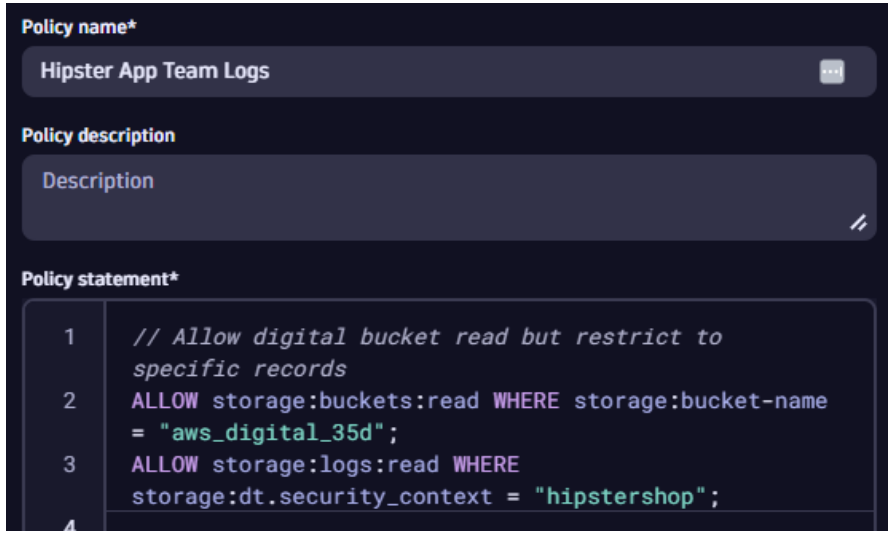
Permissions

- Controlling access to logs can be accomplished via IAM policies and defining bucket level and/or record level access.
- Bucket level control is done via bucket names.
- Record level control is done via fields or dt.security.context which must be set on the records themselves and defined within the policy.

All bucket access with deny



Digital bucket read & security context = hipstershop



Permissions: Supported Fields

Field name	IAM condition	Supported IAM tables
`event.kind`	`storage:event.kind`	`events`, `bizevents`, `system`
`event.type`	`storage:event.type`	`events`, `bizevents`, `system`
`event.provider`	`storage:event.provider`	`events`, `bizevents`, `system`
`k8s.namespace.name`	`storage:k8s.namespace.name`	`events`, `bizevents`, `logs`, `metrics`, `spans`
`k8s.cluster.name`	`storage:k8s.cluster.name`	`events`, `bizevents`, `logs`, `metrics`, `spans`
`host.name`	`storage:host.name`	`events`, `bizevents`, `logs`, `metrics`, `spans`
`dt.host_group.id`	`storage:dt.host_group.id`	`events`, `bizevents`, `logs`, `metrics`, `spans`
`metric.key`	`storage:metric.key`	`metrics`
`log.source`	`storage:log.source`	`logs`
`dt.security_context`	`storage:dt.security_context`	`events`, `bizevents`, `system`, `logs`, `metrics`, `spans`, `entities`
`gcp.project.id`	`storage:gcp.project.id`	`events`, `bizevents`, `logs`, `metrics`
`aws.account.id`	`storage:aws.account.id`	`events`, `bizevents`, `logs`, `metrics`
`azure.subscription`	`storage:azure.subscription`	`events`, `bizevents`, `logs`, `metrics`
`azure.resource.group`	`storage:azure.resource.group`	`events`, `bizevents`, `logs`, `metrics`

Discussion Question

Why would you want to create a metric from a log?

Query Performance

Cost efficient storage and query

Discussion Question

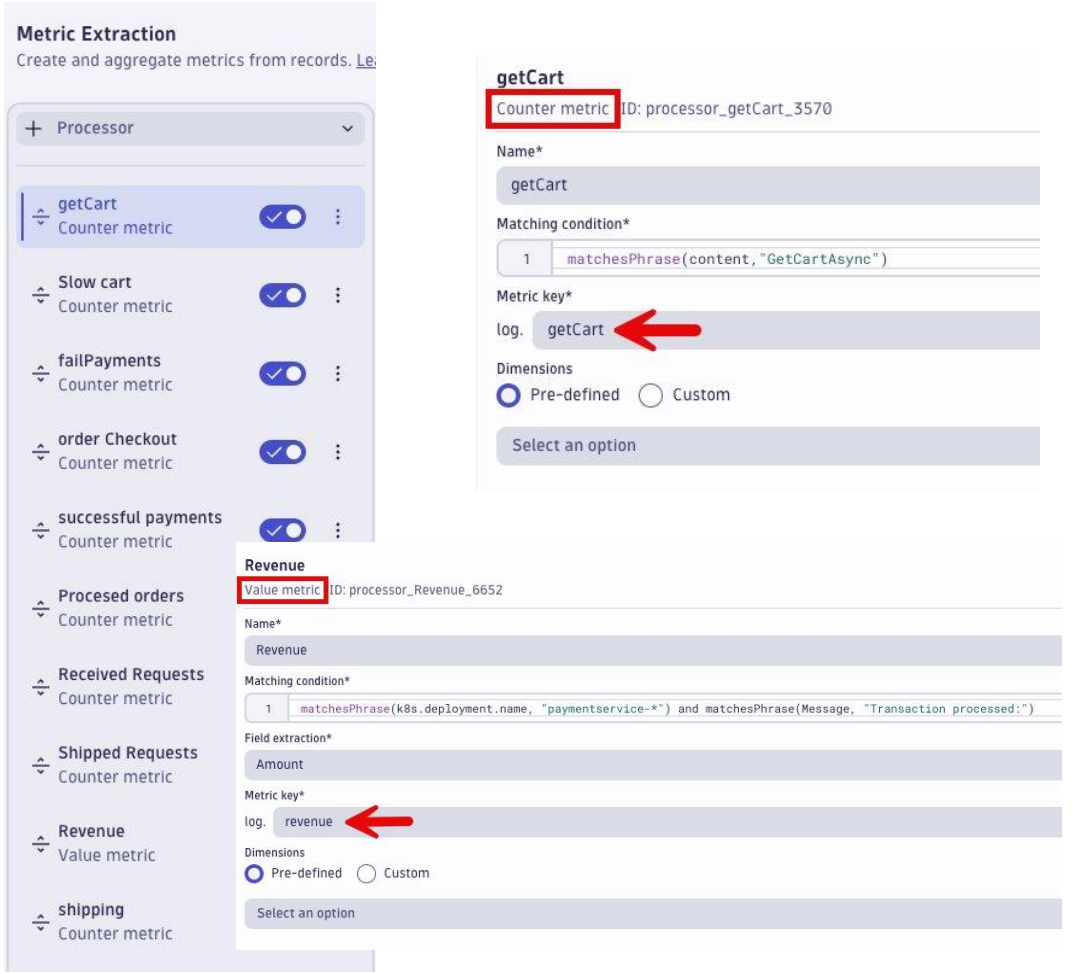
Why would you want to create a metric from a log?

Simplified alert creation

Why Logs to Metrics?

- **Enhanced Performance:**
  - Efficient storage
  - Faster query performance
  - Better scalability for handling of large volumes of data
- **Simplified Alerting:**
  - Easier to setup and manage Metrics-based alerts
  - Easier to generate dynamic thresholds
  - Easy to aggregate and summarize data

- A metric can be either a Counter or Value type.
- Counter metric - # of HTTP requests
- Value metric - Product quantity, Revenue



Log to Events

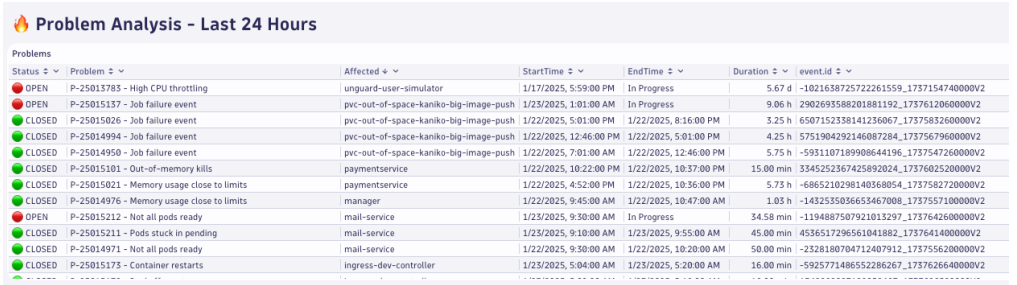
An event can be either Davis or Business event:

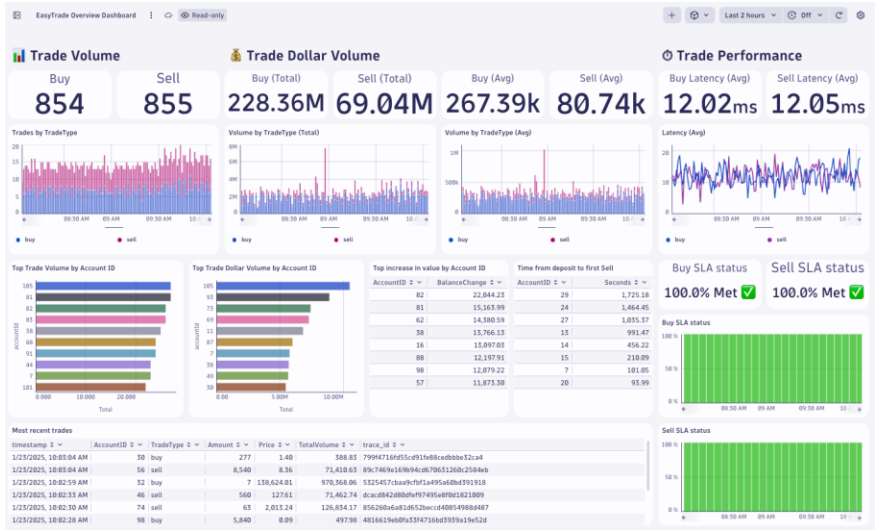
Davis Event

- Does this anomaly need to be reported?
- Is it an INFO Event that needs to be generated?

Business Event

- Is this data for reporting business data?
- Is this data for analytics for your Ops team?





Dashboarding

- Spot the differences between these two dashboards?

Use logs when:

Low frequency refresh

Involve a complex DQL

Joining different types of data

Analyzing short timeframes

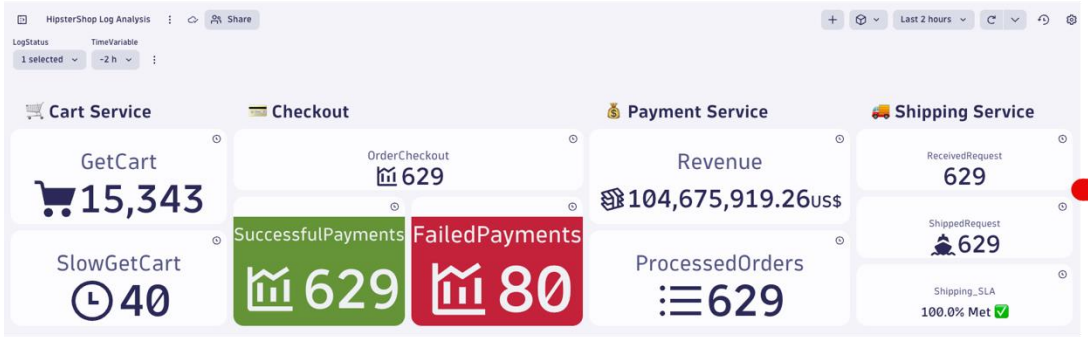
Use metrics when:

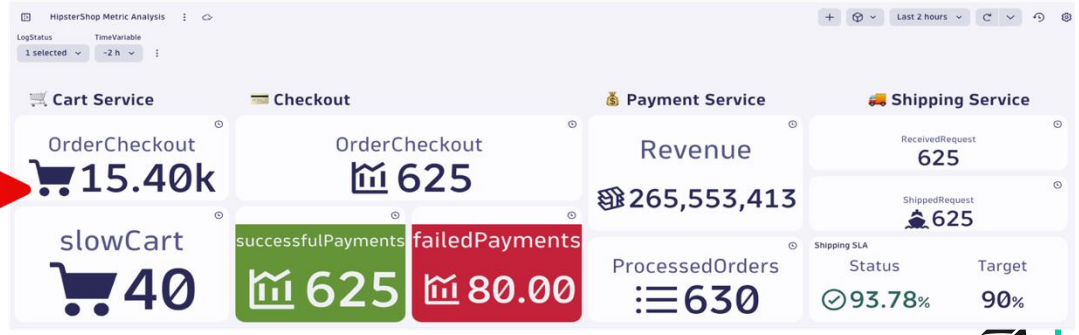
High frequency refresh

Trending over long timeframes (timeseries)

Optimize query cost

To create SLOs alerts

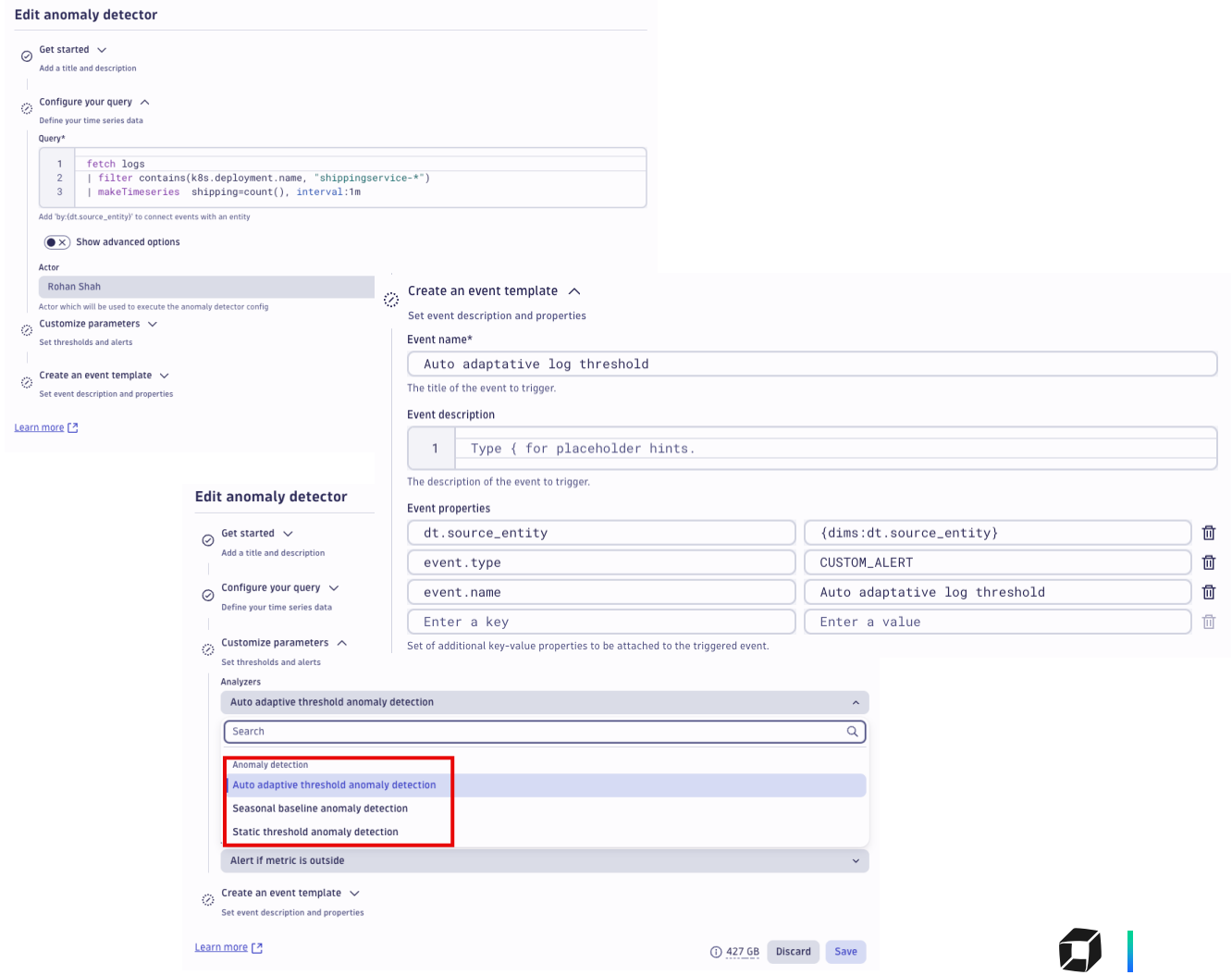




Alerting

- **Use logs when:**
  - Infrequent ingest of log records
  - Complex query
  - Detailed investigations

- **Use metrics when:**
  - Frequent data points
  - Dynamic thresholds
  - Performance tracking



Training Environment

1. Log out of any existing Dynatrace environments.

2. Login to the following tenant:

https://XXXXX.sprint.apps.dynatracelabs.com/

ID: XXXXX@protonmail.com
Password: Dynatrace123*

Live Demo & Hands-on Training

Questions?