

INC-2496 High IO Wait correlation Events & Processes by Resource Consumption

 Eric Swanson

Analysis

This morning (5/29) at 10:22 AM Eastern a steady write volume of ≥ 1 GiB per minute was seen on:

- `host:i-0bef875283ff49024`


which correlates with disk latency seen on

- `device:nvme0n1`

and Max EBS Write times > 3 minutes seen on

- `attachment-device:/dev/sda1`
- `volume-id:vol-027070cc7fb5dbeda`

Additional hosts with k8s Liveness Probe Failures


- `i-01986e7bee90235fd`
- `i-0ecfe517efc14a6d1`
 - Log errors for [Filter pkg_resources deprecation warning for error log alert](#)  **MERGED**

Between 10:33:20 10:33:40 is where the 1 minute load began to spike

Hosts in High IO Wait alert scope

- Most frequently alerting host: `i-047d8572c4c652cf6`

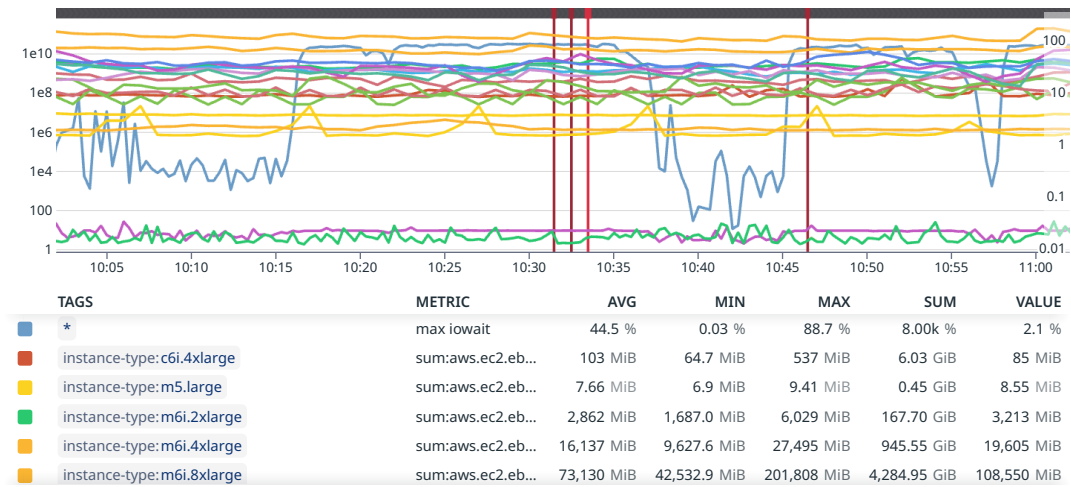
Links / Views

- [🔗 MT/MC Production - High IO wait on k8s worker node](#)
- [🔗 Python Command Processes - Saved View](#)
- [🔗 Processes by Command and Namespace - Saved View](#)
 - move and expand time window to `May 29, 10:00 am – May 29, 11:00 am`
- [Host Metrics Dashboard: i-0bef875283ff49024](#)
- [🔗 APM Traces: api-gateway http://gb428.us1.dbt.com/api/v2/accounts/22484/jobs/647267/run/](#)
- [🔗 Process Outlier pid:7](#)
- [dispatcher: start trace as soon as message is received](#)  **MERGED**
 - `raise `` ``GracefulExit``(``"detached completed run"``)`
- <https://docs.datadoghq.com/infrastructure/process/?tab=helm#queries>

Disk + CPU Wait Correlation

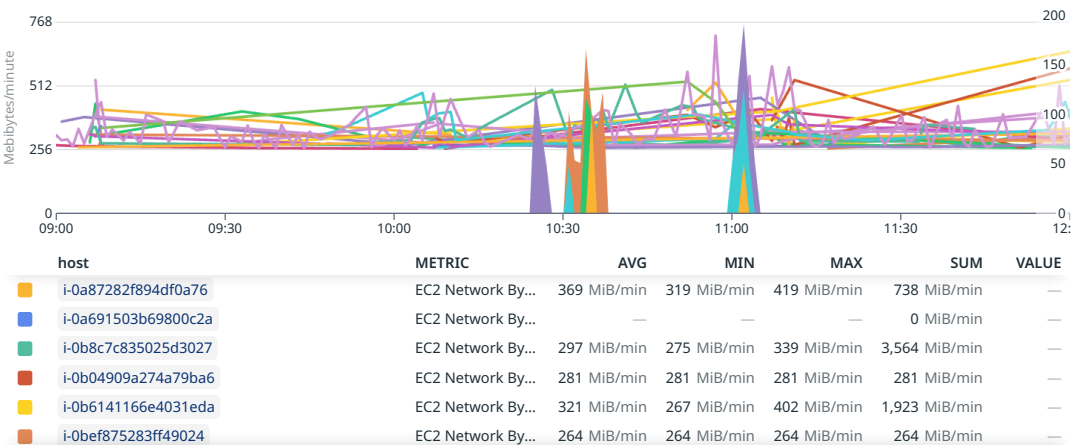
Log Scale CPU + Disk IO Wait Investigation

1h 10:02 am - 11:02 am



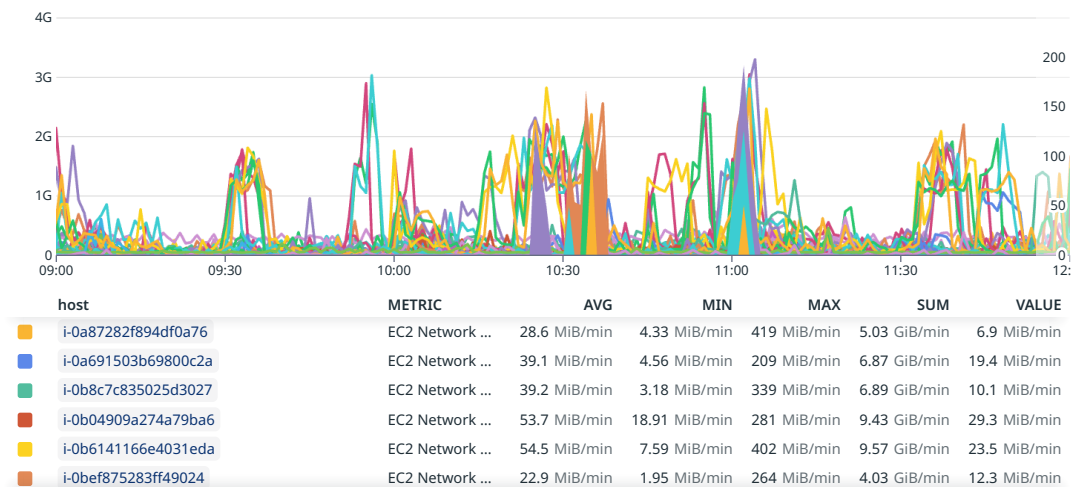
System Load 1 + Network Bytes Written Correlation by Host for Process start times

3h 9:00 am - 12:00 pm



Log Scale CPU + Disk IO Wait Investigation

3h 9:00 am - 12:00 pm



Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Top 10 Scribe Process by Open File Count - Timeseries for duration

15m 10:20 am - 10:35 am

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

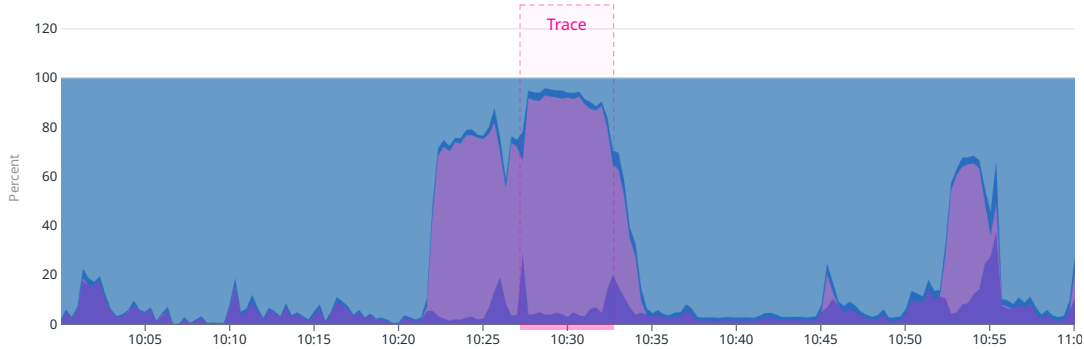
Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

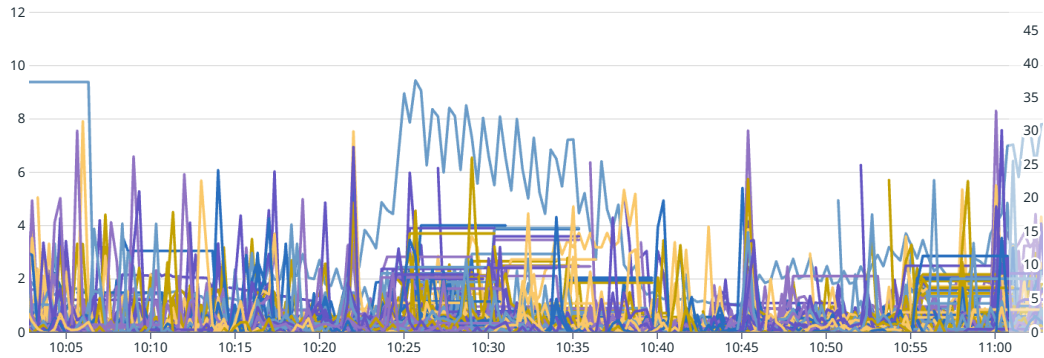
CPU usage (%) with orc-dispatch trace overlay

1h 10:00 am - 11:00 am



Load Averages 1-5-15

1h 10:02 am - 11:02 am



Process Metric Exported Widgets

write-heavy Process Investigation - # Threads

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Top write-heavy Process Investigation - git process metrics

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Top write-heavy dbt Process metrics

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Top write-heavy scribe Process metrics

6m 10:28 am - 10:34 am

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Most Recent Impacting Processes & Events

May 29, 10:00 am – May 29, 1:00 pm

Top 10 dbt-run processes + OOM Kill & High IO Wait

1h 9:55 am - 10:55 am

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

Top 50 User CPU % non-dbt-run processes + OOM Kill & High IO Wait

40m 9:55 am - 10:35 am

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

Top 50 User CPU % non-dbt-run processes + High IO Wait

40m 9:55 am - 10:35 am

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

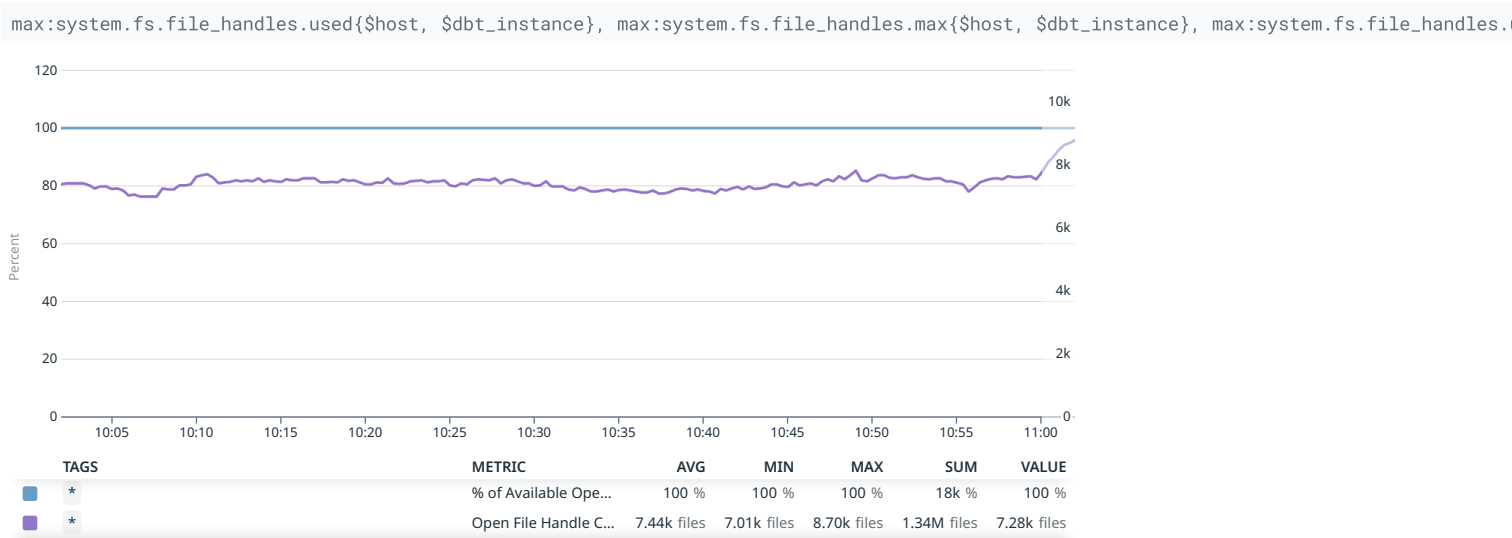
High IO Wait + Nonzero Exit & Error Events

Invalid Timeframe

Process query must end within last 36h

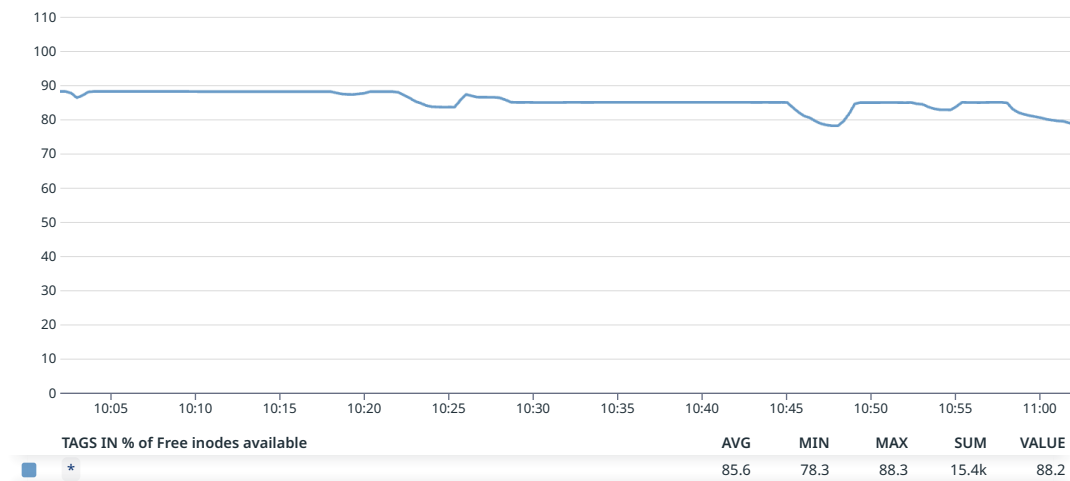
[Set Global Time to Past 1 Week](#)

File Descriptors



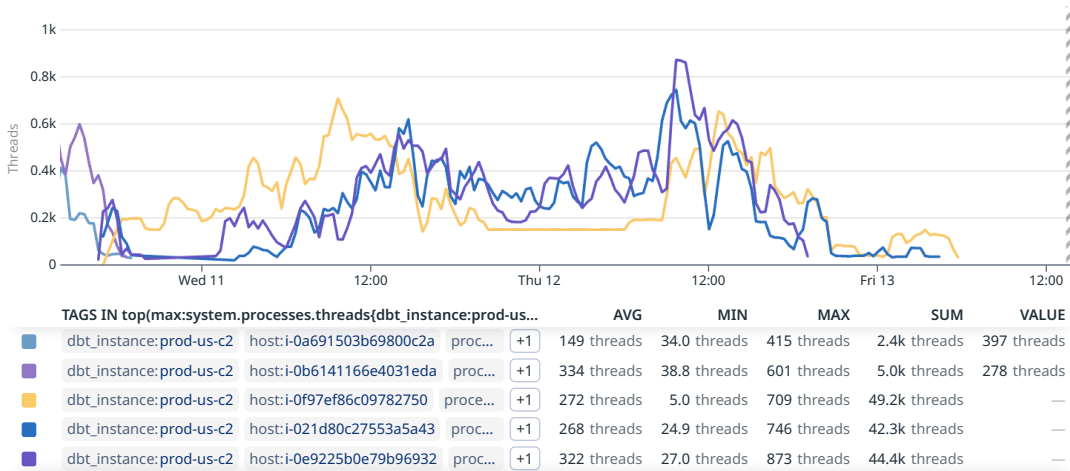
max:system.fs.inodes.used{\$host, \$dbt_instance}, max:system.fs.inodes.free{\$host, \$dbt_instance}

1h 10:02 am - 11:02 am



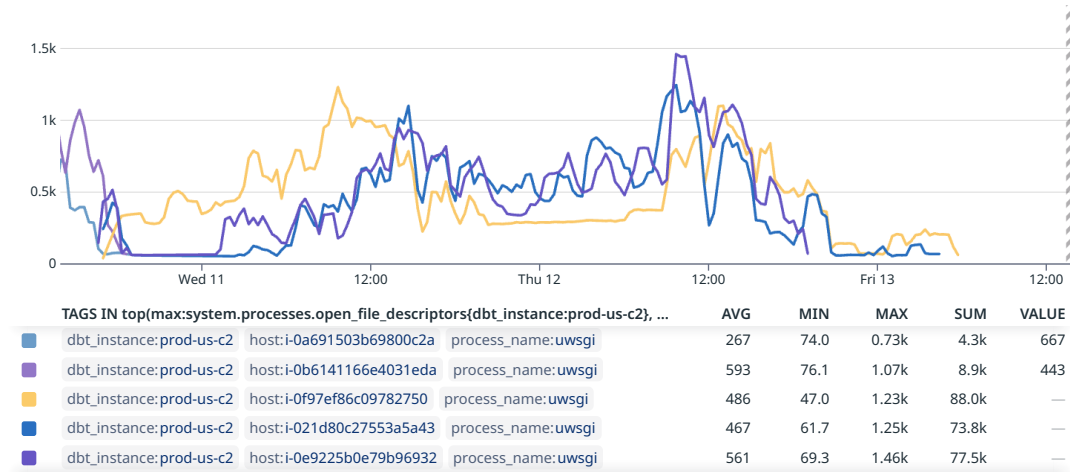
uwsgi Top 5 max # threads by host

3d Past 3 Days



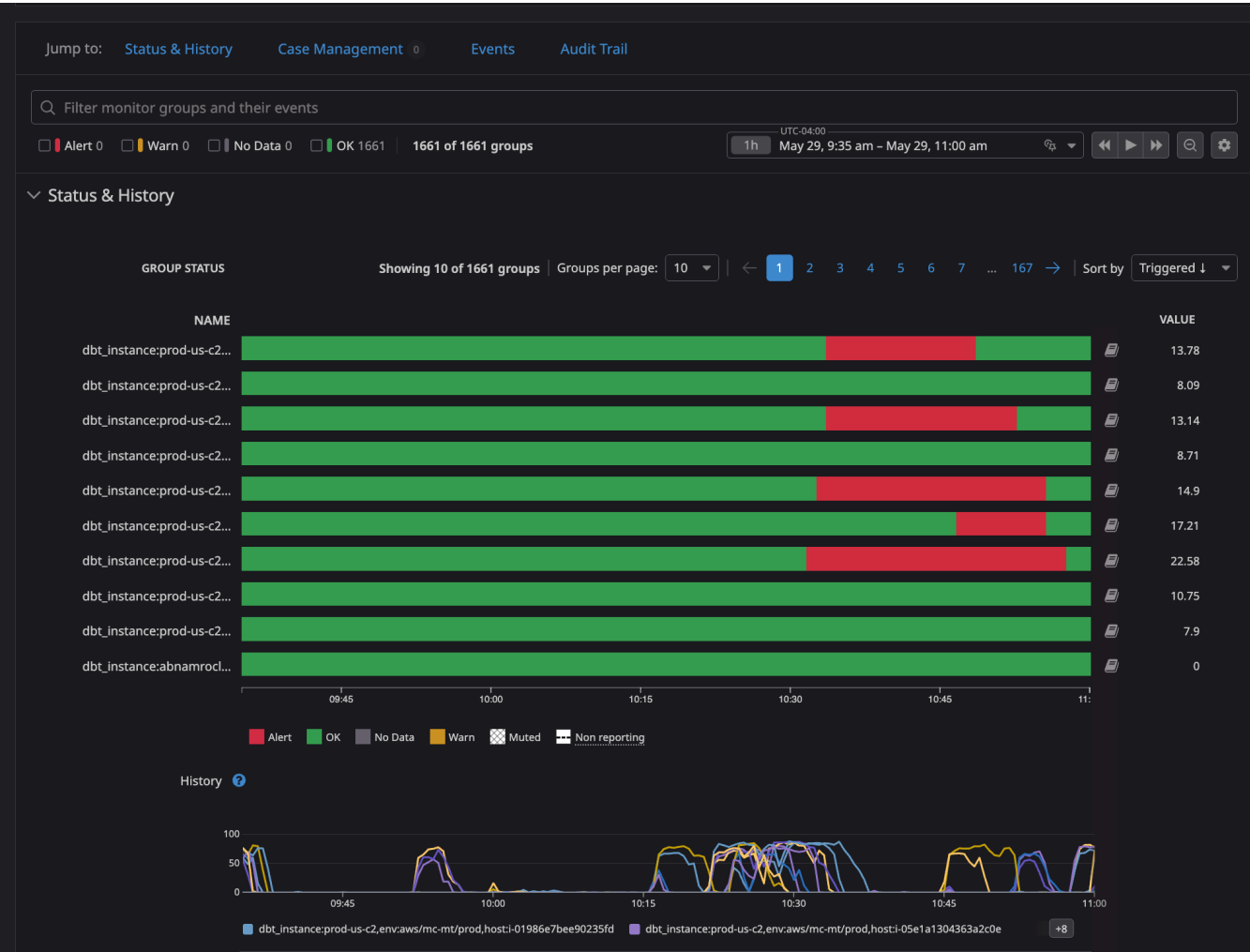
uwsgi Top 5 Max Open File Count processes by host

3d Past 3 Days



High IO Wait Monitor Details

Alert History Screenshot



Alert Events

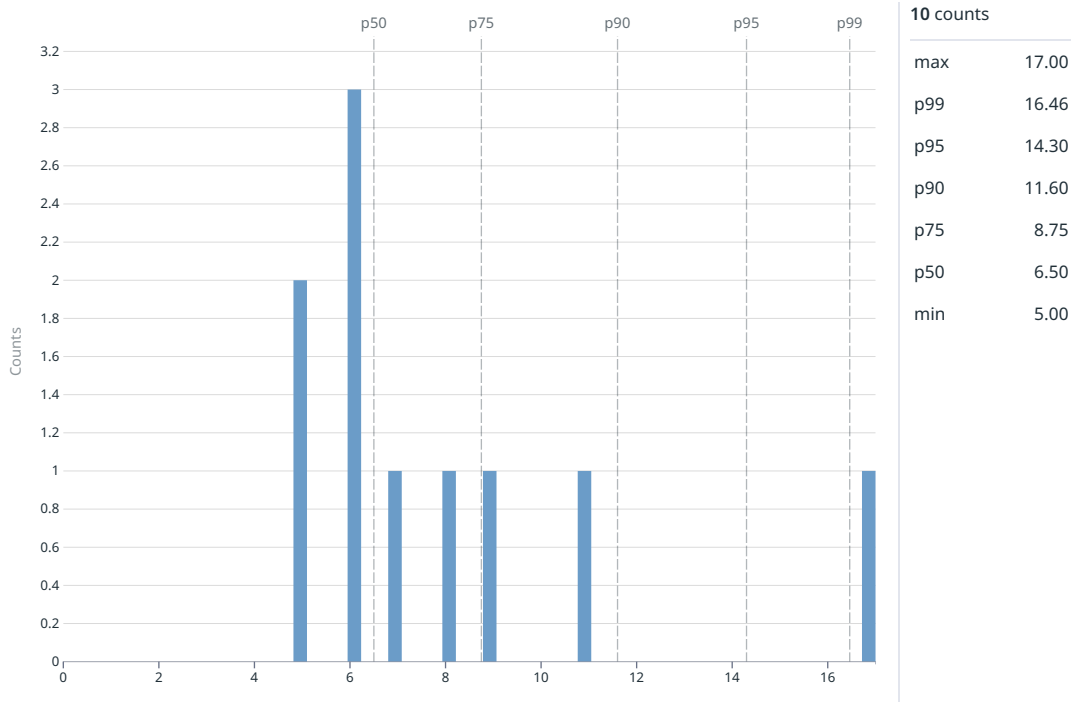
Count of High IO Wait Alert Events by Host

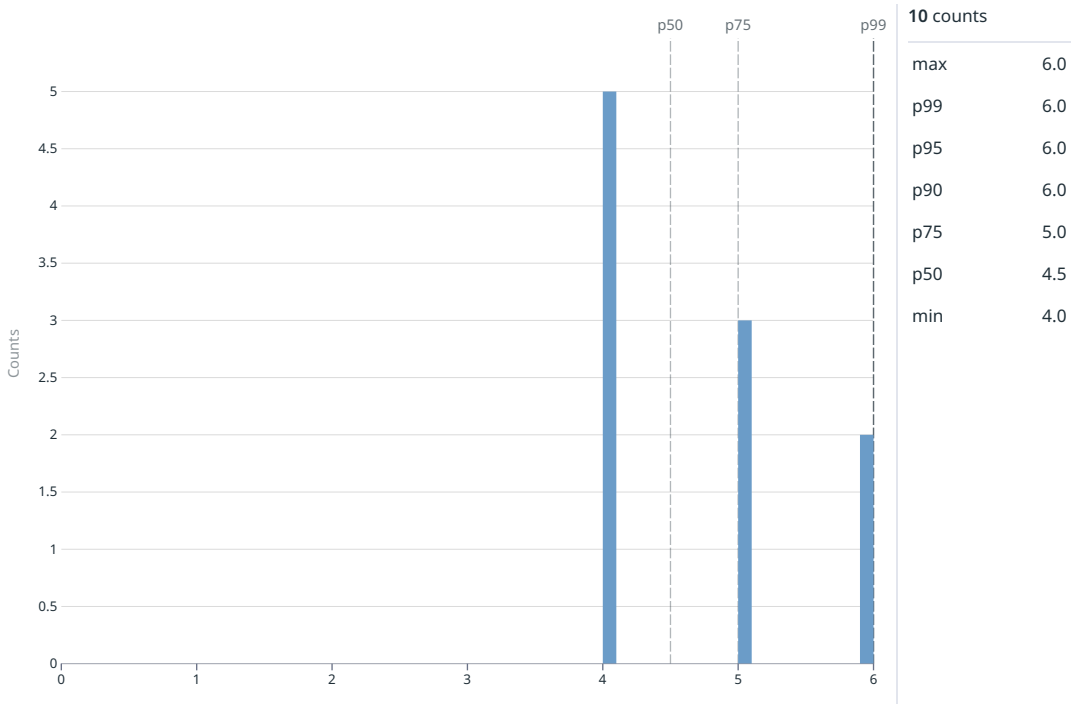
3mo Past 3 Months



Distribution of High IO Wait Alert Events by Host

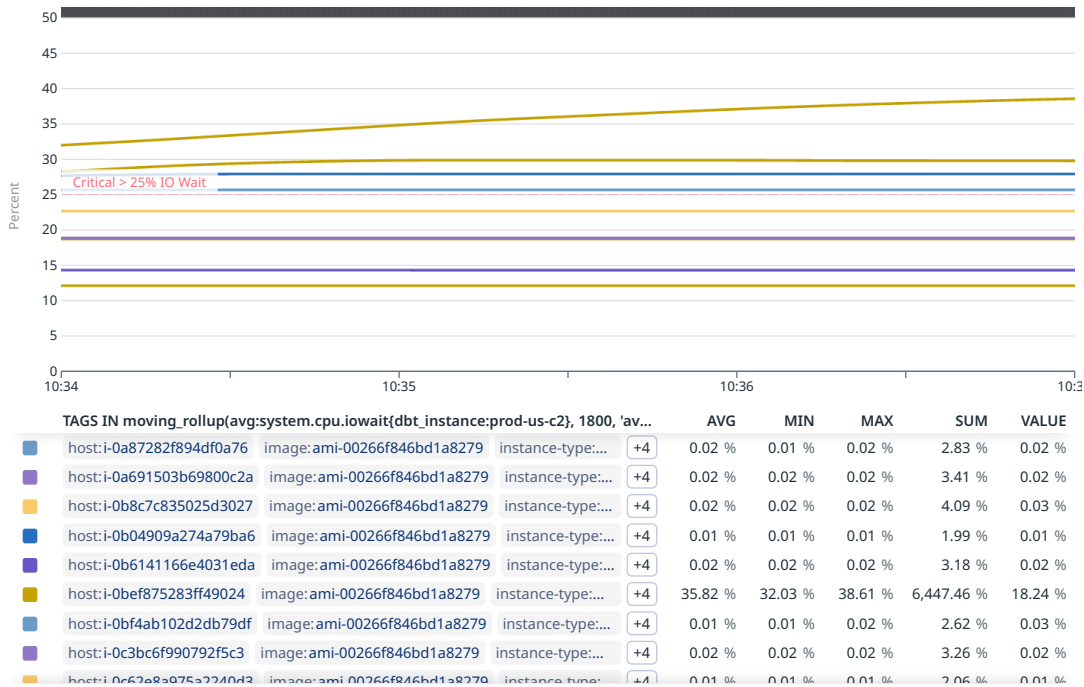
3mo Past 3 Months





Monitor Metrics

MT/MC Production - High IO wait on k8s worker node - 30 minute rolling avg w/ 25% Threshold



a

Recent Process Metrics

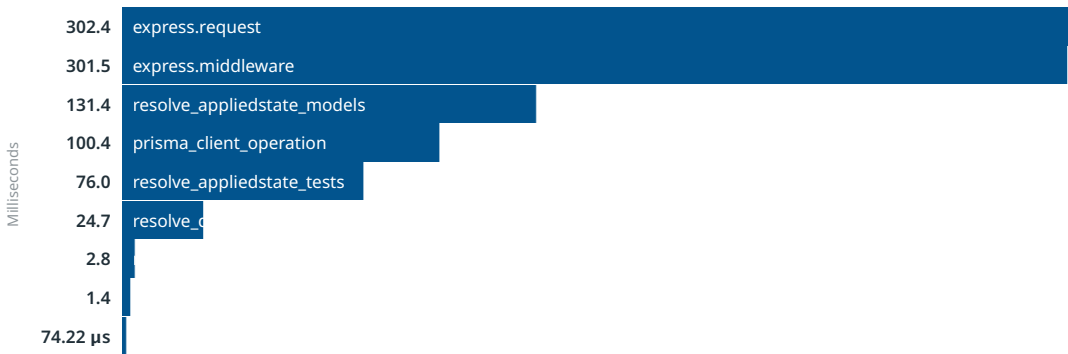
Argocd Monitor Metrics

Host Metrics

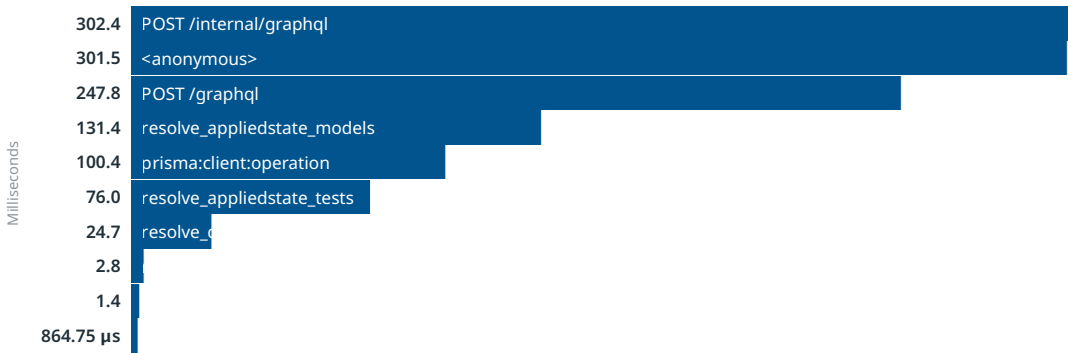
Historic Metrics

codex-api

Median of @duration by @service, dbt_instance over "dbt_instance:prod-us-c2 @process_id:7"



Median of @duration by @service, dbt_instance over "dbt_instance:prod-us-c2 @process_id:7"



Process TreeMaps: (min 15m window required)

Process User Cpu % Treemap for command:*

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Top 200 Process User Cpu % Treemap for host:* - All Pods and Commands

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Process Open File Count Treemap for command:*

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Top 200 Process Open File Count Treemap for host:* - all pods and commands

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

agent process: Top 10 open file count in dbt_instance:prod-us-c2

2h

 10:00 am - 11:30 am

Invalid Timeframe

Process query must end within last 36h

[Edit Local Time](#)

Process # Threads Treemap for command:*

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Top 200 Process # Threads Treemap for host:* - all commands and pods

Invalid Timeframe

Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

Process RSS Memory % Treemap for host:* command:* dbt-run pods

Invalid Timeframe

Process query must end within last 36h

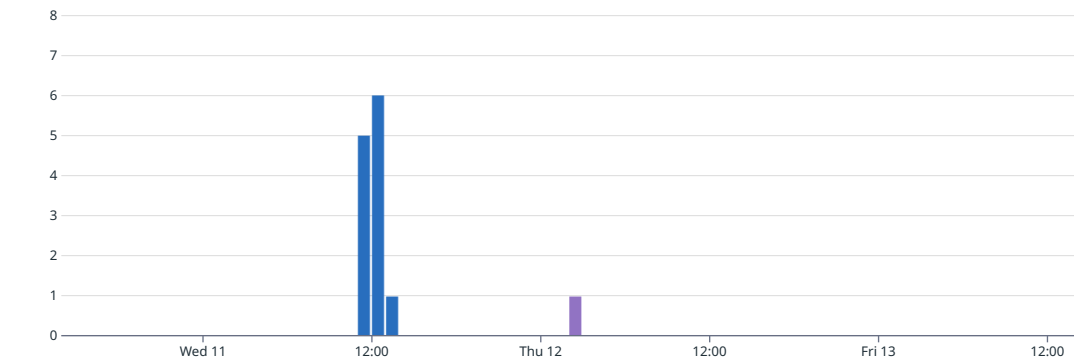
[Set Global Time to Past 1 Week](#)

Invalid Timeframe

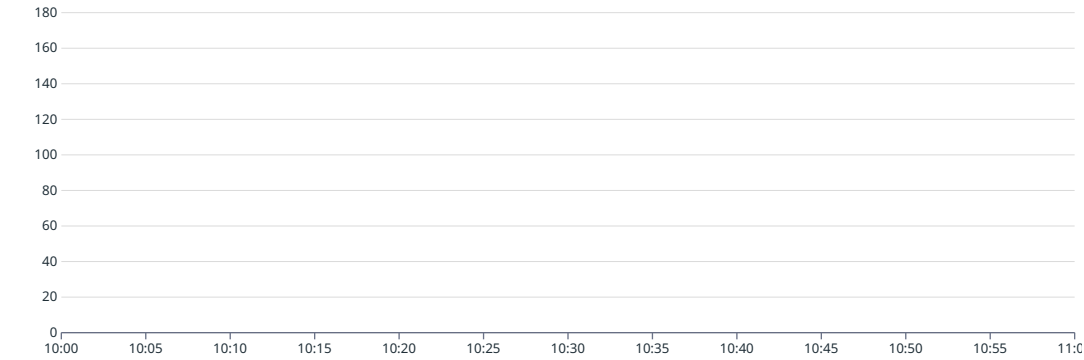
Process query must end within last 36h

[Set Global Time to Past 1 Week](#)

dbt_instance:prod-us-c2 "CODE_DEFERRAL_ARTIFACT_DOWNLOAD_EXCEPTION"  3d Past 3 Days

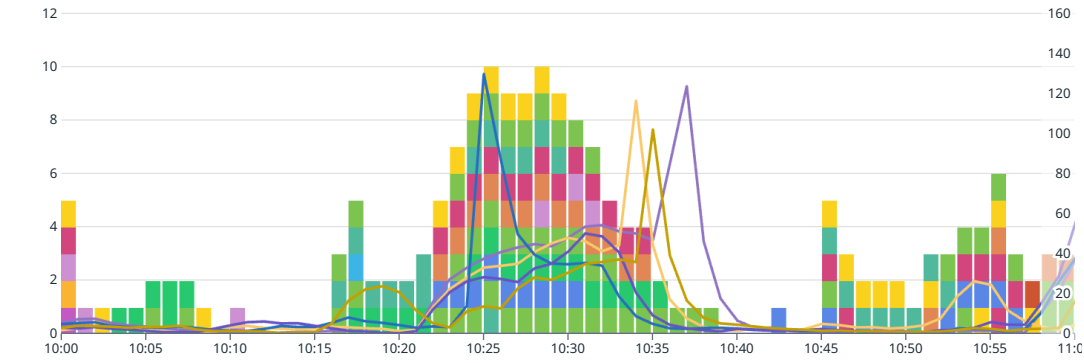


dbt_instance:prod-us-c2 @account_id:70437463654464 -status:info  1h 10:00 am - 11:00 am



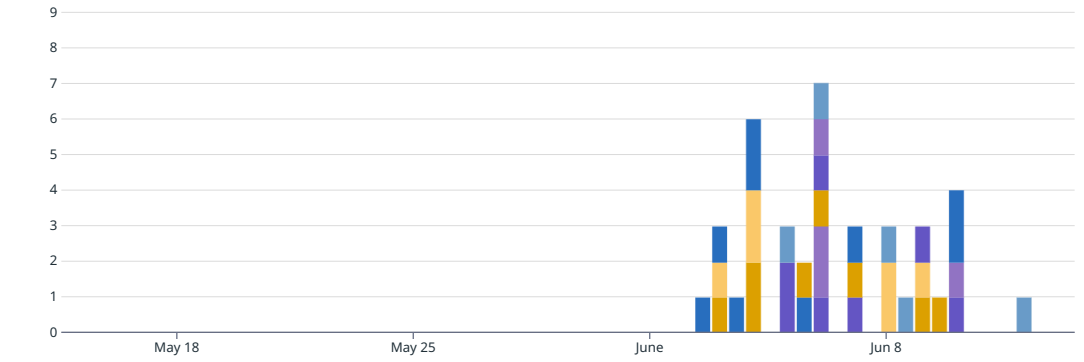
Top Logs for @account_id:70437463654464 with Top System Load 1

1h 10:00 am - 11:00 am

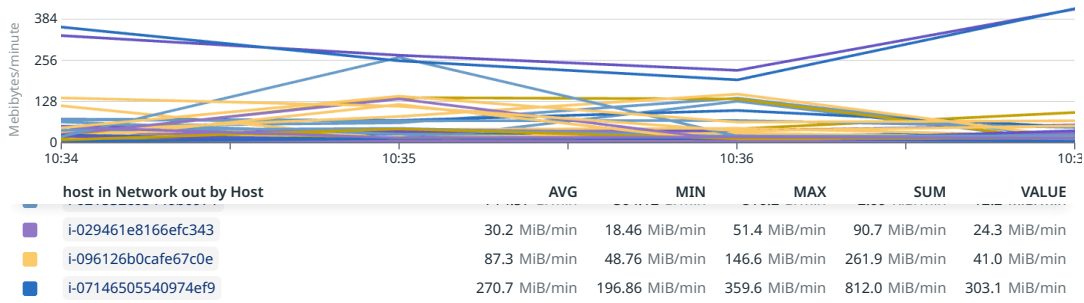


dbt_instance:prod-us-c2 "is too long and would be truncated"

1mo Past 1 Month



EC2: Outbound traffic



dbt_instance:prod-us-c2 accounts/70437463654464 status:error

14d Past 14 Days

COUNT	~VOLUME	@ACCOUNT_ID	MESSAGE
1.59K			yyyy/MM/dd HH:mm:ss [error] [302-304]#[302-304]: * [3-1115145] *.dbt-cloud.svc.cluster.local could not be resolved (3: Host not found), client: XXX.XXX.XXX.XXX, server: , request: "** * HTTP/1.1", host: "ke963.us1.dbt.com", referrer: ""
36			Internal Server Error: /api/*/accounts/70437463654464/*/
			yyyy/MM/dd HH:mm:ss [error] 303#303: *[8-1079852] upstream timed out (110: Operation timed out) while reading response header from upstream client:

	COUNT	~VOLUME	STATUS	@HTTPURL_DETAILS.PATH	
