

Slides at



Benni De Jagere

**POWER BI PREMIUM
PRACTICAL TIPS TO MAKE THE MOST OF IT**

kapacity Fellowwind Initiate



KONICA MINOLTA



R E L A T E I T

[stoltze][it]unit^{it}

WHO AM I



Power BI CAT

dataMinds.be Member



@BenniDeJagere



/bennidejagere



/bennidejagere



/bennidejagere



#SayNoToPieCharts

OUR TEAM activities

We engage with strategic Power BI customers and the sales professionals supporting them to showcase and provide technical guidance, help drive product adoption and consumption and influence our engineering roadmap.

BDM SHOWCASE

We own demo resources for CxO-level conversations

FIELD SUPPORT

We support technical escalations & training opportunities

PATTERNS & PRACTICES

We curate content on cross-cutting technology topics

ENGINEERING FEEDBACK

We influence the products roadmap via feedback rhythms

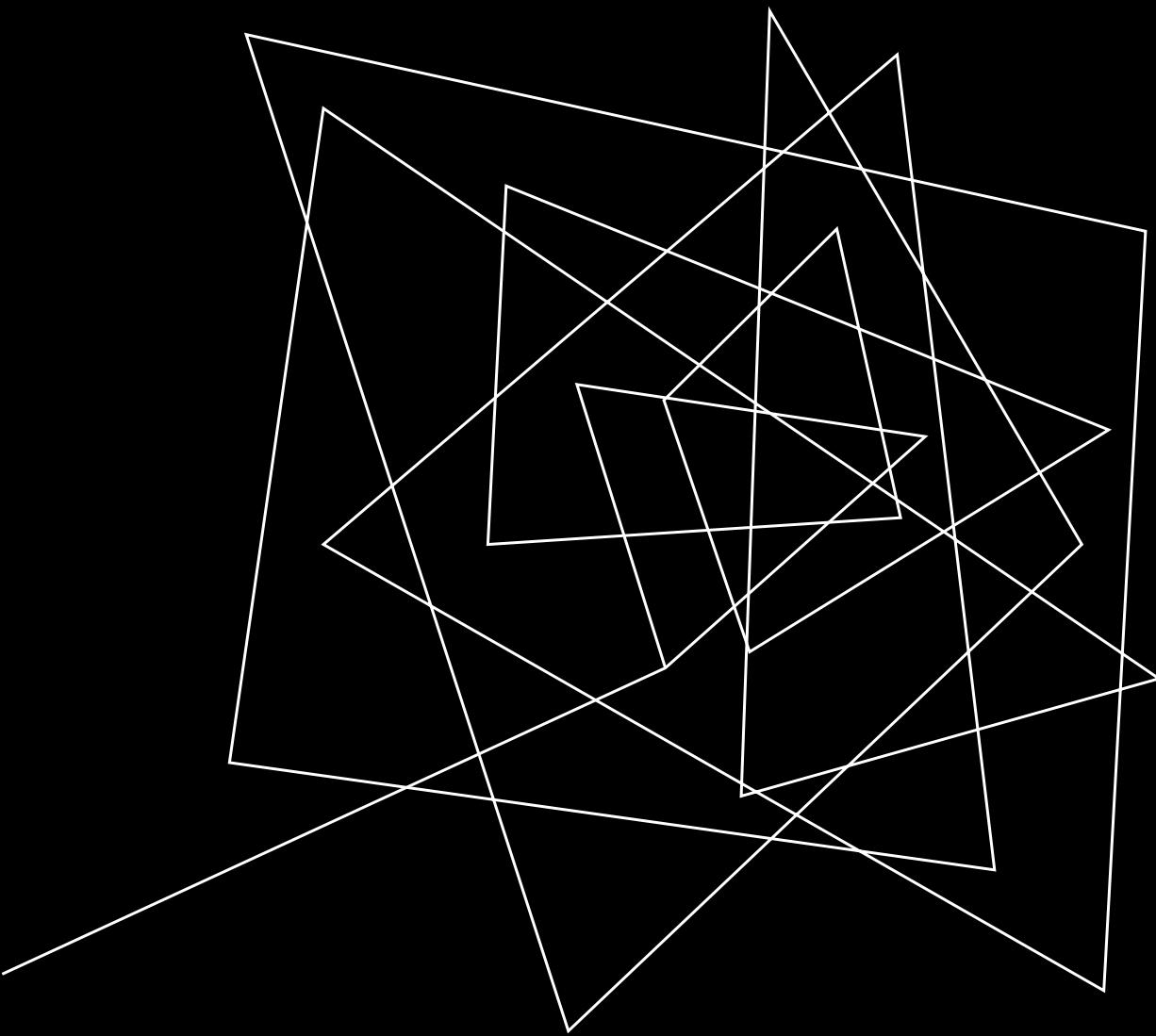
OBJECTIVES

Set the scene on Power BI Premium Gen 2 +
PPU

Usable tips to assist daily tasks

Convince proactive monitoring is key

NOTE: Power BI Premium Gen 1 is not covered, nor is Multi-Geo



LET'S START!

POWER BI PREMIUM?

Dedicated Reserved Capacity for your organisation

Provides a richer feature set

Some features are Premium First, Pro later (or not ..)

Support for:

Data residency by region (Multi-Geo)

Customer-managed encryption keys for data at rest (BYOK)

Users can access artefacts with a free Power BI Subscription

NOT Easy-mode for performance & optimisation

RESOURCE OVERVIEW PER CAPACITY

Capacity SKUs	Total v-cores	Backend v-cores	Frontend v-cores	RAM (GB) ^{1, 2}	DirectQuery/Live connection (per second) ^{1, 2}	Max memory per query [GB] ^{1, 2}	Model refresh parallelism ²
EM1/A1	1	0.5	0.5	3	3.75	1	5
EM2/A2	2	1	1	5	7.5	2	10
EM3/A3	4	2	2	10	15	2	20
P1/A4	8	4	4	25	30	6	40
P2/A5	16	8	8	50	60	6	80
P3/A6	32	16	16	100	120	10	160
P4/A7 ³	64	32	32	200	240	10	320
P5/A8 ³	128	64	64	400	480	10	640

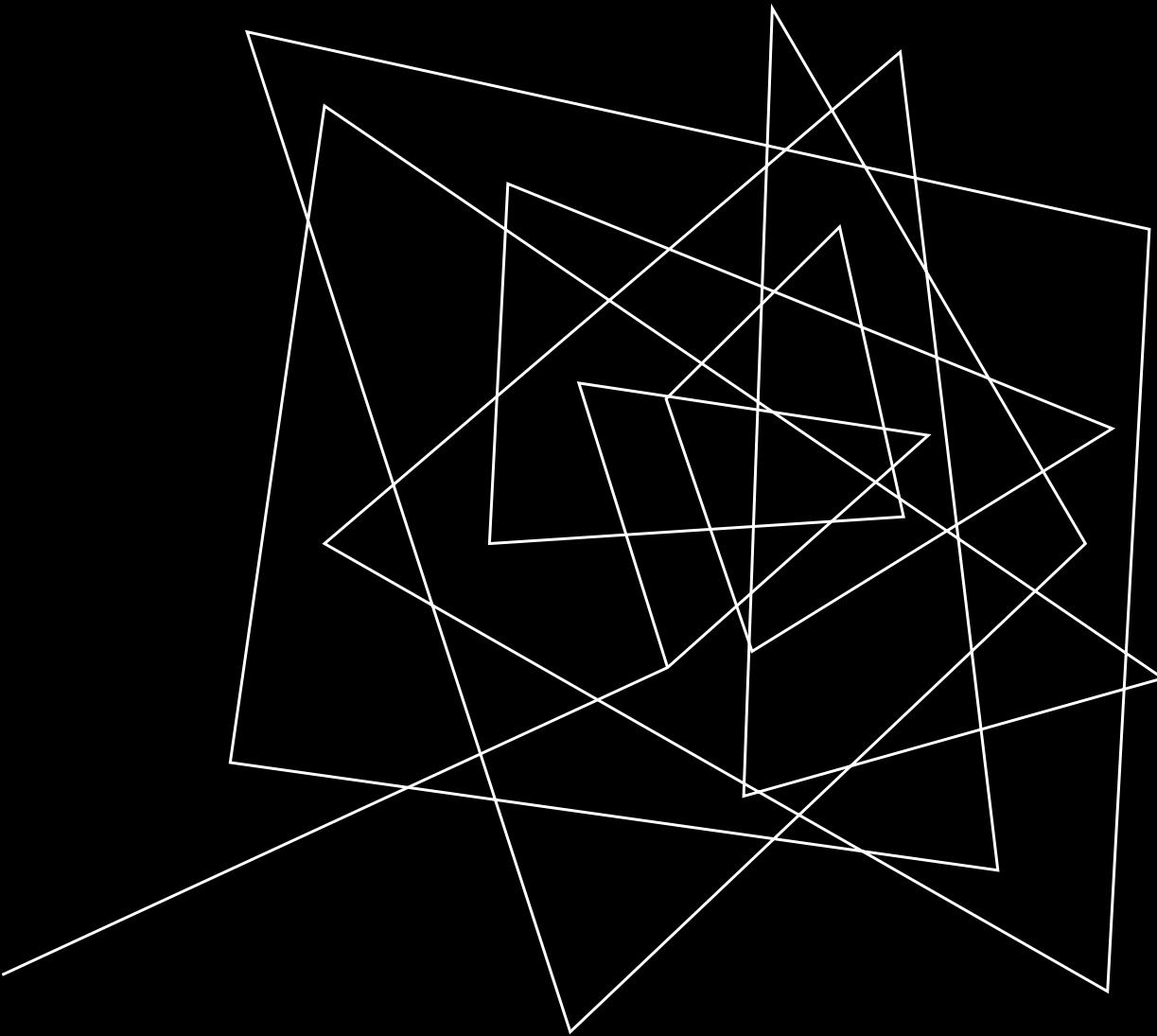
RESOURCE OVERVIEW PER CAPACITY

Capacity SKUs	Total v-cores	Backend v-cores	Frontend v-cores	RAM (GB) ^{1, 2}	DirectQuery/Live connection (per second) ^{1, 2}	Max memory per query [GB] ^{1, 2}	Model refresh parallelism ²
---------------	---------------	-----------------	------------------	--------------------------	--	---	--

¹ The Power BI Premium Utilization and Metrics app doesn't currently expose these metrics.

² These limits only apply to dataset workloads.

³ SKUs greater than 100 GB are not available in all regions. To request using these SKUs in regions where they're not available, contact your Microsoft account manager.



HOW DOES IT
WORK?

POWER BI PREMIUM GEN 2



Flexibility to license by capacity **AND** by user



Up to 16X performance boost



Scale capacity when needed with Autoscale



Consistent and reliable cost-management

KEY CONCEPTS

A few key concepts to remember:

Under the covers, everyone operates on “P3” shared nodes

Goodbye Dedicated Capacity, hello Reserved Capacity!

Usage metrics shift from memory & cpu usage to cpu usage

Usage metrics now return 14 days by default

ADVANTAGES

Organisational Noisy Neighbours are limited

Load balancing will block out resources for your process

Cross-workload resource contention is eliminated

Resource limitations shift to:

Total capacity CPU Throughput

Memory limit per artefact equivalent to your capacity sku limit

BACKEND V-CORES

Backend v-cores are implemented on regional clusters

Shared (yet secured) by all tenants in that region

Clusters have resource groups, allocated to handle workloads

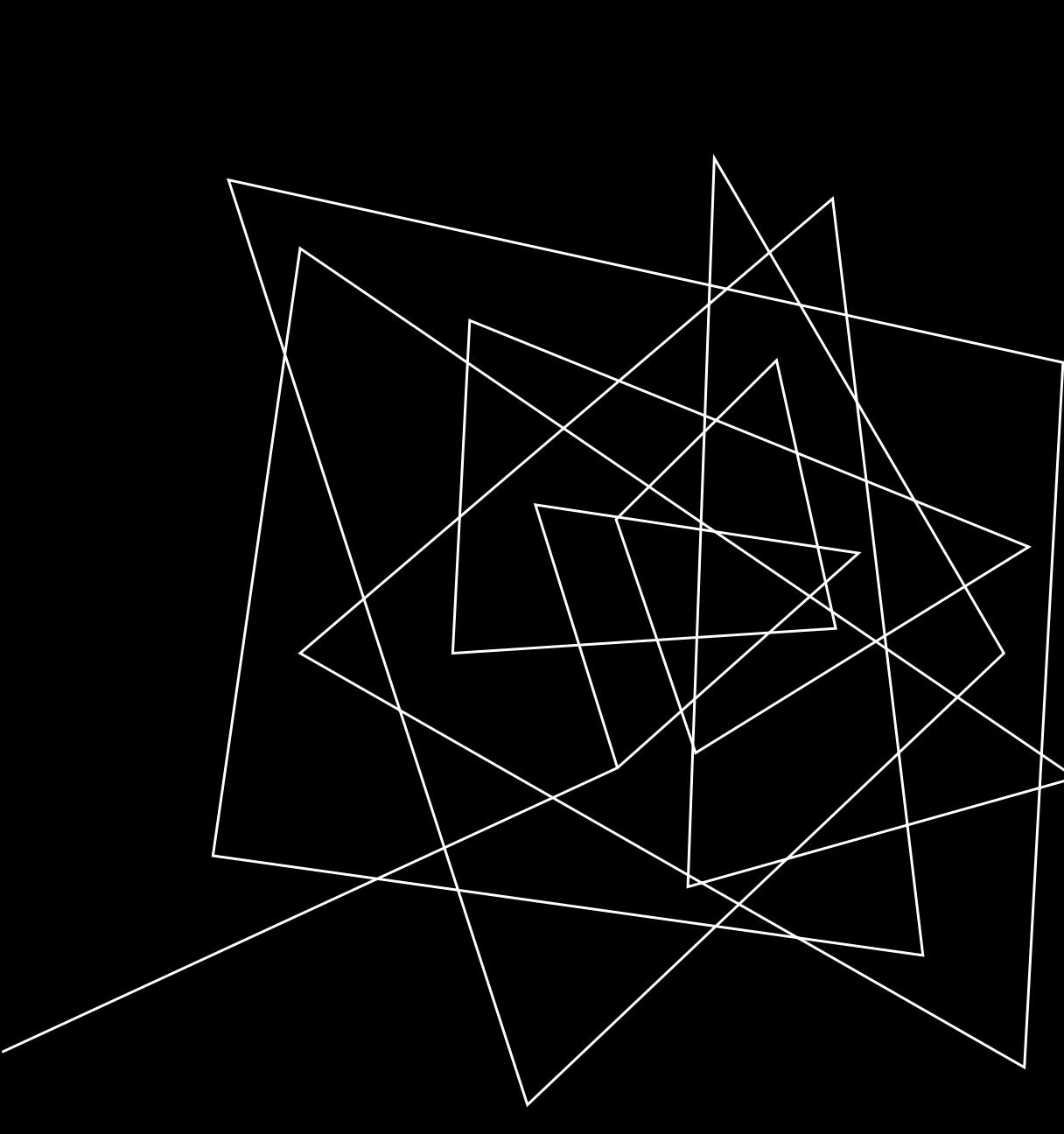
STORAGE

Contents are stored on the organisational storage layer

Limited to 100TB per organisation, across all workloads

Loaded into shared compute nodes by request

Power BI handles load balancing



LOAD EVALUATION

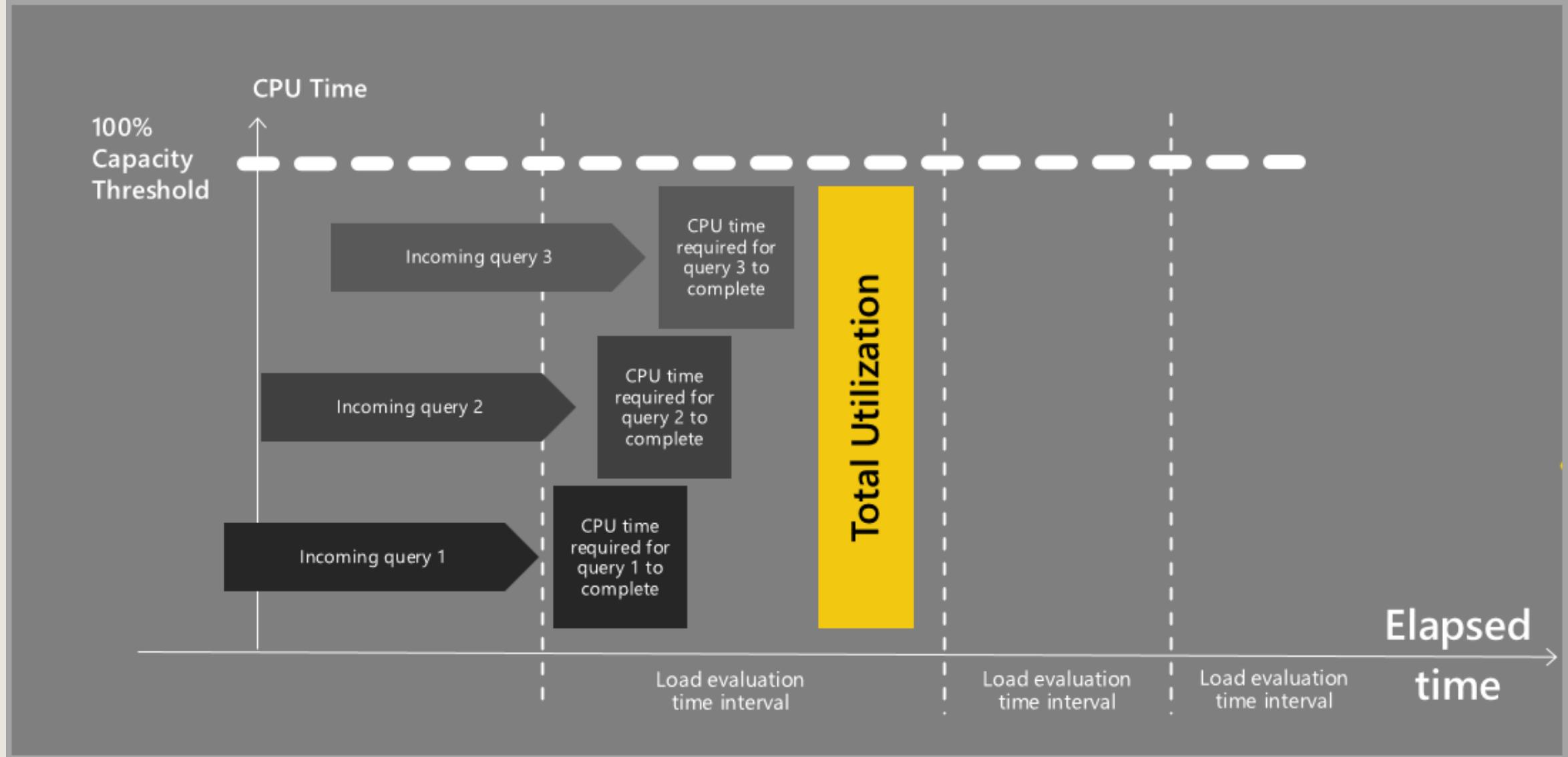
LOAD EVALUATION

Based on your P-sku, you have an ‘allowance’ per cycle
ie. P1 has 4 backend v-cores, for 30 seconds each.

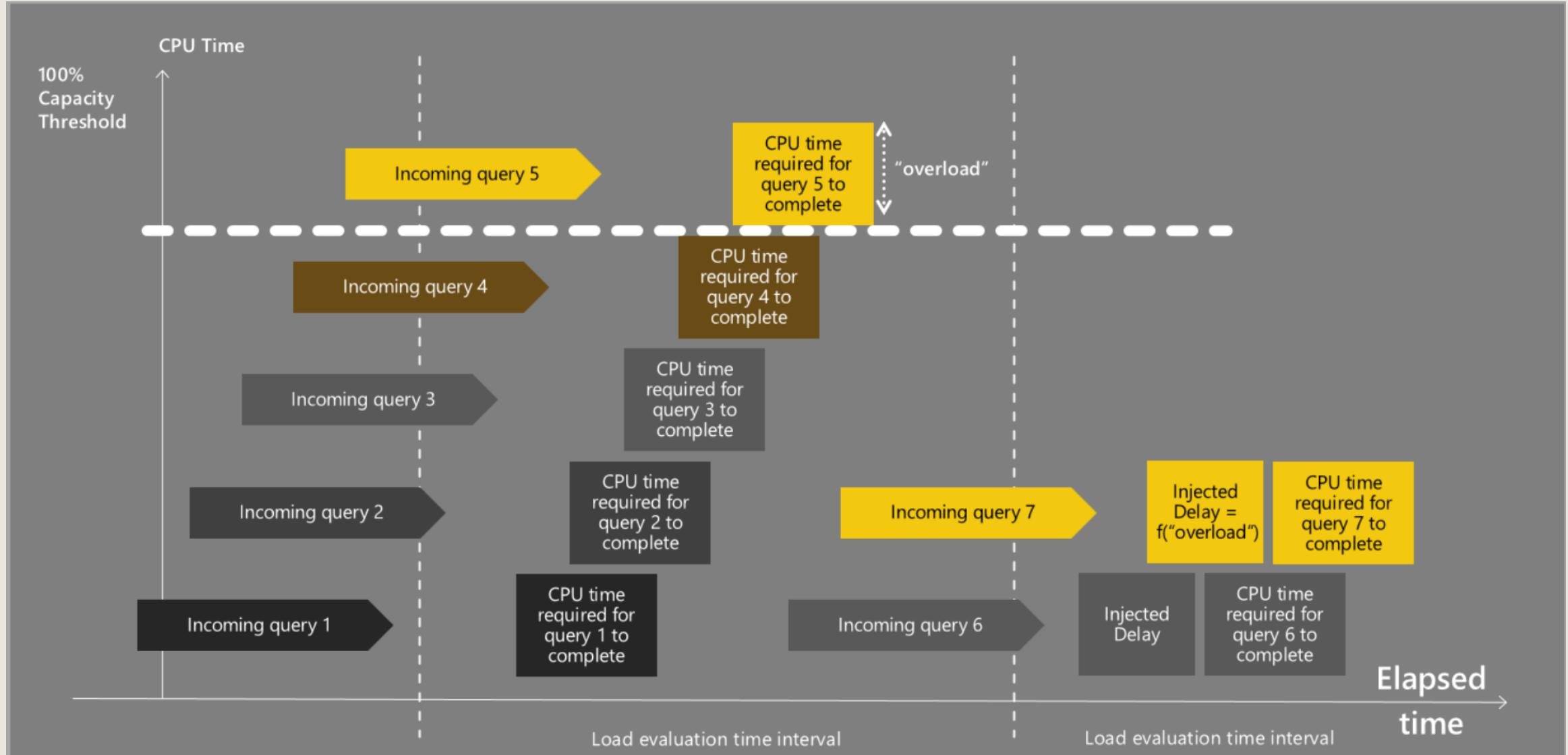
Every 30 seconds, Power BI evaluates throughput
Slow operations (Background) are spread out across 24 hours
Fast operations (Interactive) are spread out across the next
10 minutes (subject to change)

Aggregation of these two determines load per cycle
Keep in mind evictions still exist

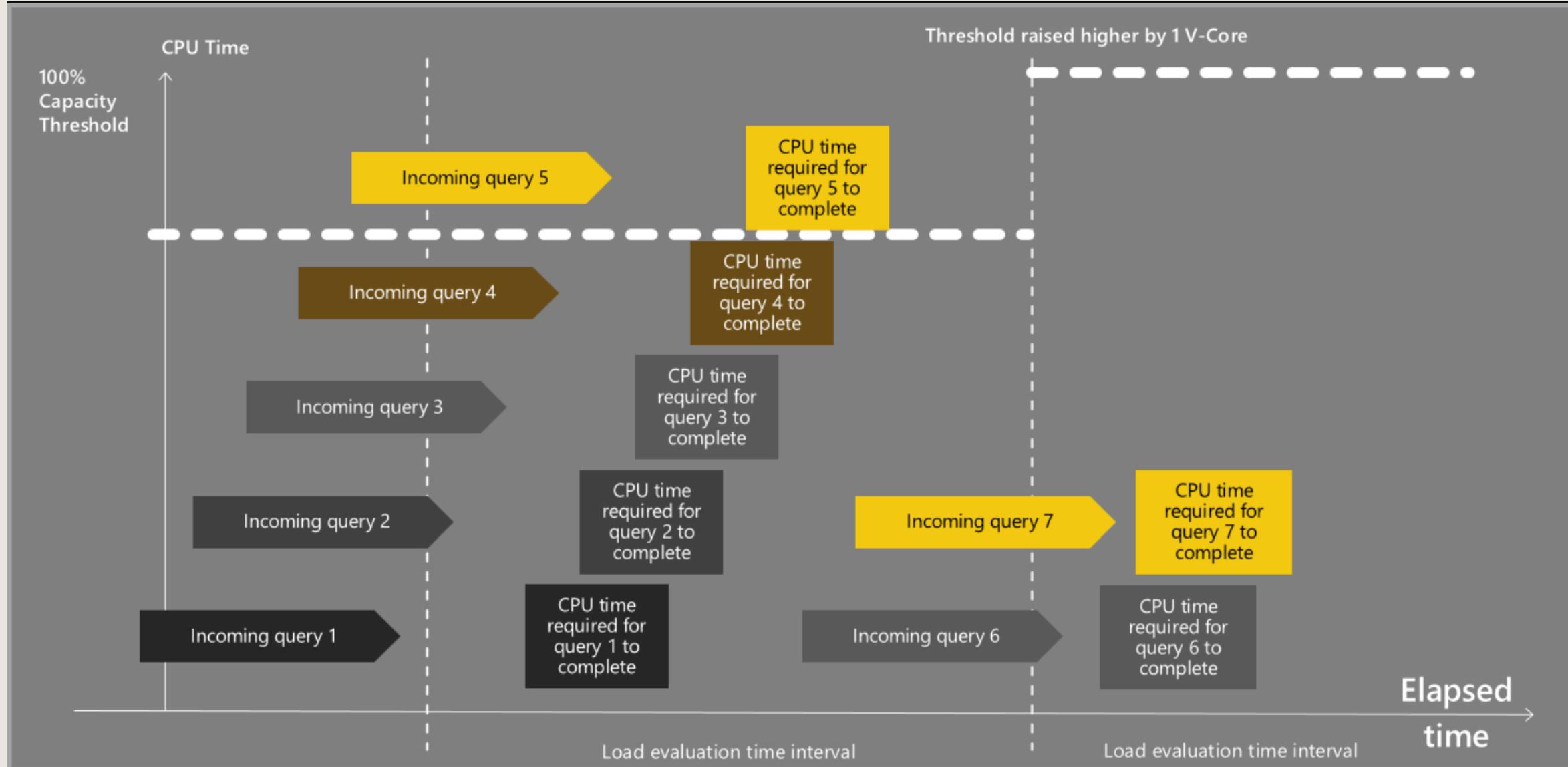
LOAD EVALUATION



LOAD EVALUATION



AUTOSCALE



AUTOSCALE

AutoScale can be configured per P Sku

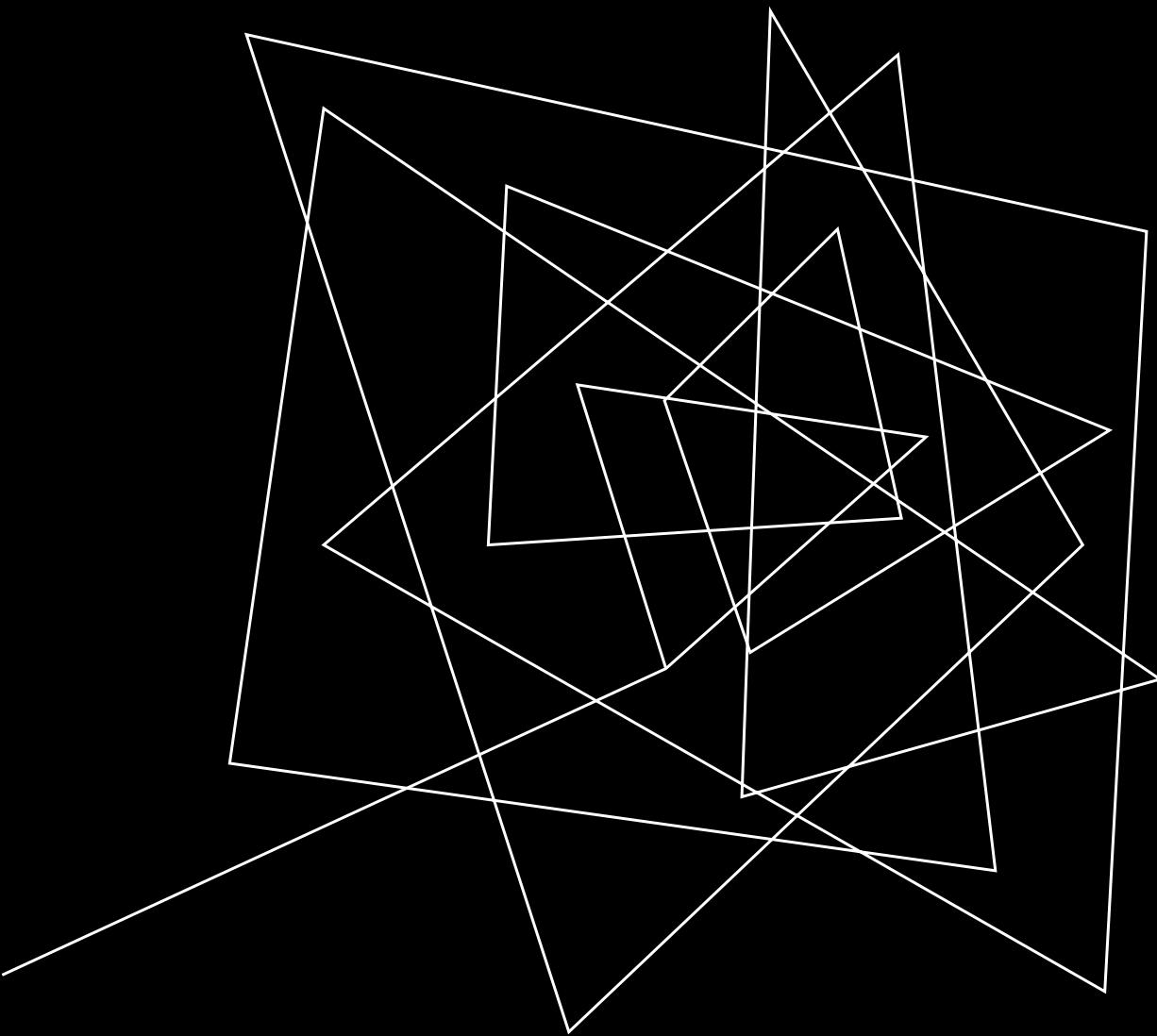
Reaching a load evaluation limit, AutoScale can kick in

For 24 hours, you'll receive an extra v-core

Up to maximum amount of v-cores you have for your capacity

Billed through your Azure Subscription

TIP: Set up reactive / proactive budget control



METRICS APP

PROACTIVE MONITORING

Premium Capacity Usage and Resource Utilization

Overview Evidence Refresh Help

Capacity Name: PRD1

By date: Sat 22, Sun 23, Mon 24, Tue 25, Wed 26

CPU (s) Duration (s) Users Artifact Size (GB) Overloaded minutes Performance delta

Artifacts (14 days)

OperationName	CPU (s)	Duration (s)	Users	Artifact Size (GB)	Overloaded minutes	Performance delta
+ Query	672	485	2			
+ XMLA Read Operation	35	38	2			
Total	707	523	3	1.28		57

Dataflow

Dataset

PaginatedReport

Artifacts (14 days)

Artifact	CPU (s)	Duration (s)	Users	Artifact Size (GB)	Overloaded minutes	Performance delta
	832,775	509,105	10	1.41		
	321,289	133,426	5	2.73	1.0	
	143,883	40,713	1	9.37		
	108,913	44,141	2	3.17		
	100,900	65,557	76	2.19		
	96,741	26,527	9	1.66	1.0	
Total	4,339,463	6,123,145	406	11.72		

Minute: 0, 22 Jan, 29 Jan

Artifact size: Peak (blue dot), SKU limit (red dotted line), Dataset refresh Limit (yellow dotted line)

Active artifacts: Line chart showing artifact count over time.

Active users: Line chart showing user count over time.

Cores: Line chart showing core usage over time.

Performance profile: Bar chart showing performance levels (Fast, Moderate, Slow) over time.

PROACTIVE MONITORING

Premium Capacity Usage and Resource Utilization

Overview Evidence Refresh Help

Capacity Name
PRD1

Pick a capacity to review in the Capacity Name picker. All visuals on the page will refresh each time a capacity is picked. Your selection will apply to other pages as well. Learn how to use this page by clicking the "info" button.

?

...

CPU Duration Operations Users

Background CPU Interactive CPU CPU Limit

By date

0.2M

0.1M

0.0M

Sat 22 Sun 23 Mon 24 Tue 25 Wed 26 Thu 27 Fri 28 Sat 29 Sun 30 Mon 31 Tue 1 Wed 2 Thu 3 Fri 4

TimePoint Mon 31 4:02:00 PM

Interactive CPU 735.43

Right-click to drill through

Overloaded minutes per hour

0 100 Minutes

22 Jan 29 Jan

Artifacts (14 days)

Dataflow Dataset PaginatedReport

Artifact CPU (s) Duration (s) Users Artifact Size (GB) Overloaded minutes Performance delta

Artifact	CPU (s)	Duration (s)	Users	Artifact Size (GB)	Overloaded minutes	Performance delta
	57,078	28,739	76	2.19		
	37,509	10,557	16	1.78		
	29,057	26,409	18	9.76	3.0	
	24,075	9,659	43	1.05		

Weekly trendlines

These visuals provide a quick look at high-level trends. Each data point in these visuals shows the value for a rolling seven-day period.

CPU

Active artifacts

Active users

Cores

Artifact size

Peak SKU limit Dataset refresh Limit

0 50 100 GB

23 Jan 25 Jan 27 Jan 29 Jan

Performance profile

Fast Moderate Slow

Detailed description: The screenshot displays a complex monitoring interface. At the top, a dropdown menu 'Capacity Name' is set to 'PRD1'. A note below it says: 'Pick a capacity to review in the Capacity Name picker. All visuals on the page will refresh each time a capacity is picked. Your selection will apply to other pages as well. Learn how to use this page by clicking the "info" button.' Below this are navigation links: Overview, Evidence, Refresh, and Help. A large red box highlights the 'By date' section, which contains a stacked bar chart showing CPU usage from Saturday 22 to Friday 4. An inset provides a detailed view of the 'Interactive CPU' usage for Monday 31 at 4:02:00 PM, showing a value of 735.43. Another red box highlights the 'Overloaded minutes per hour' chart, which shows a sharp spike on January 29. To the left, a table lists 'Artifacts (14 days)' with columns for CPU (s), Duration (s), Users, Artifact Size (GB), Overloaded minutes, and Performance delta. On the right, there are five trendline charts: 'CPU' (blue line), 'Active artifacts' (blue line), 'Active users' (blue line), 'Cores' (blue line), and 'Artifact size' (blue line, red dotted line for SKU limit, yellow dotted line for dataset refresh limit). A legend at the bottom defines performance profiles: Fast (green), Moderate (yellow), and Slow (red).

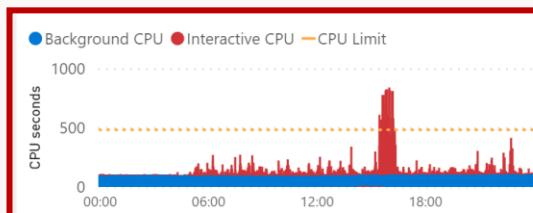
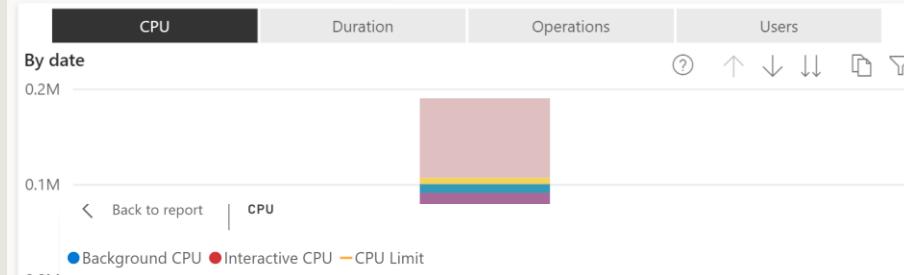
PROACTIVE MONITORING

Premium Capacity Usage and Resource Utilization

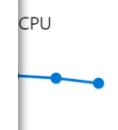
Overview Evidence Refresh Help

Capacity Name
PRD1

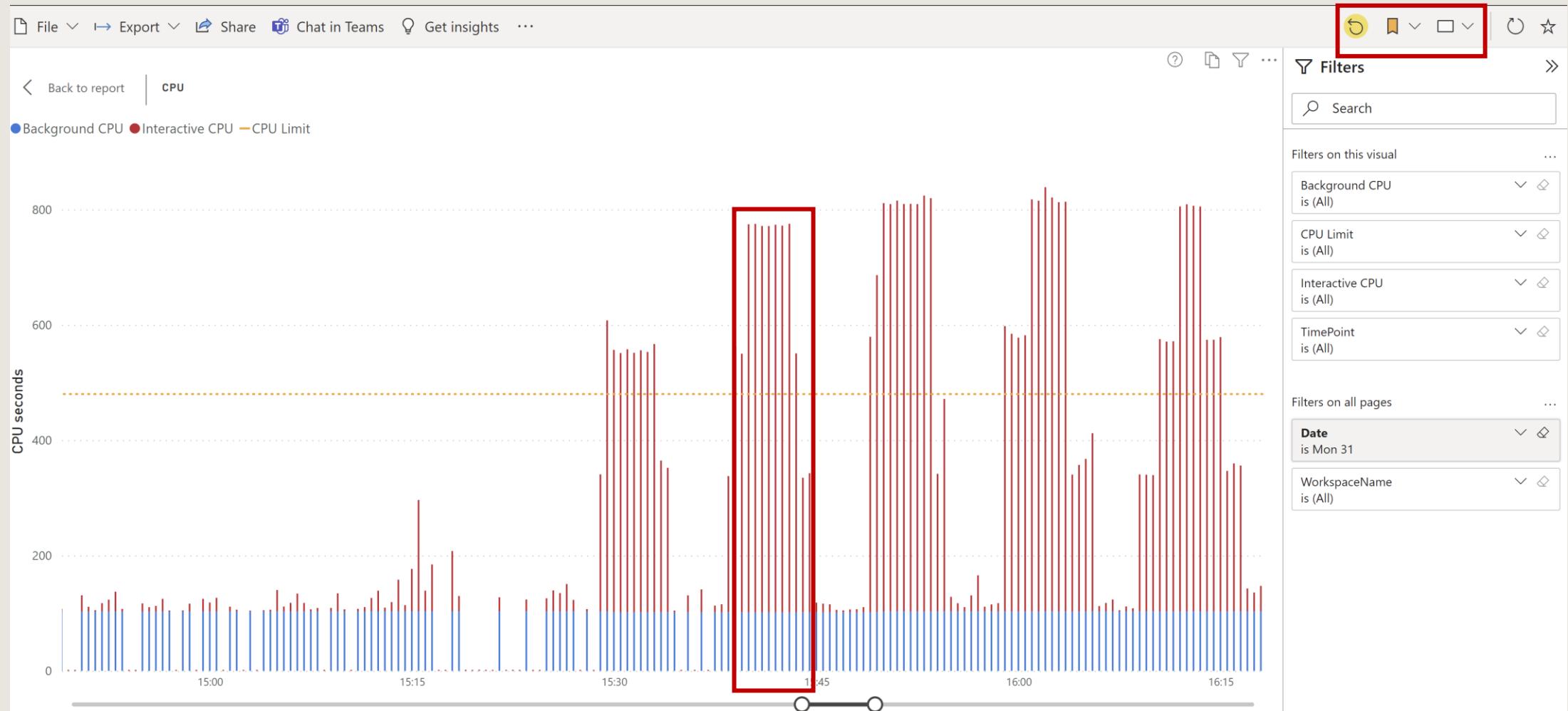
Pick a capacity to review in the Capacity Name picker. All visuals on the page will refresh each time a capacity is picked. Your selection will apply to other pages as well. Learn how to use this page by clicking the "info" button.



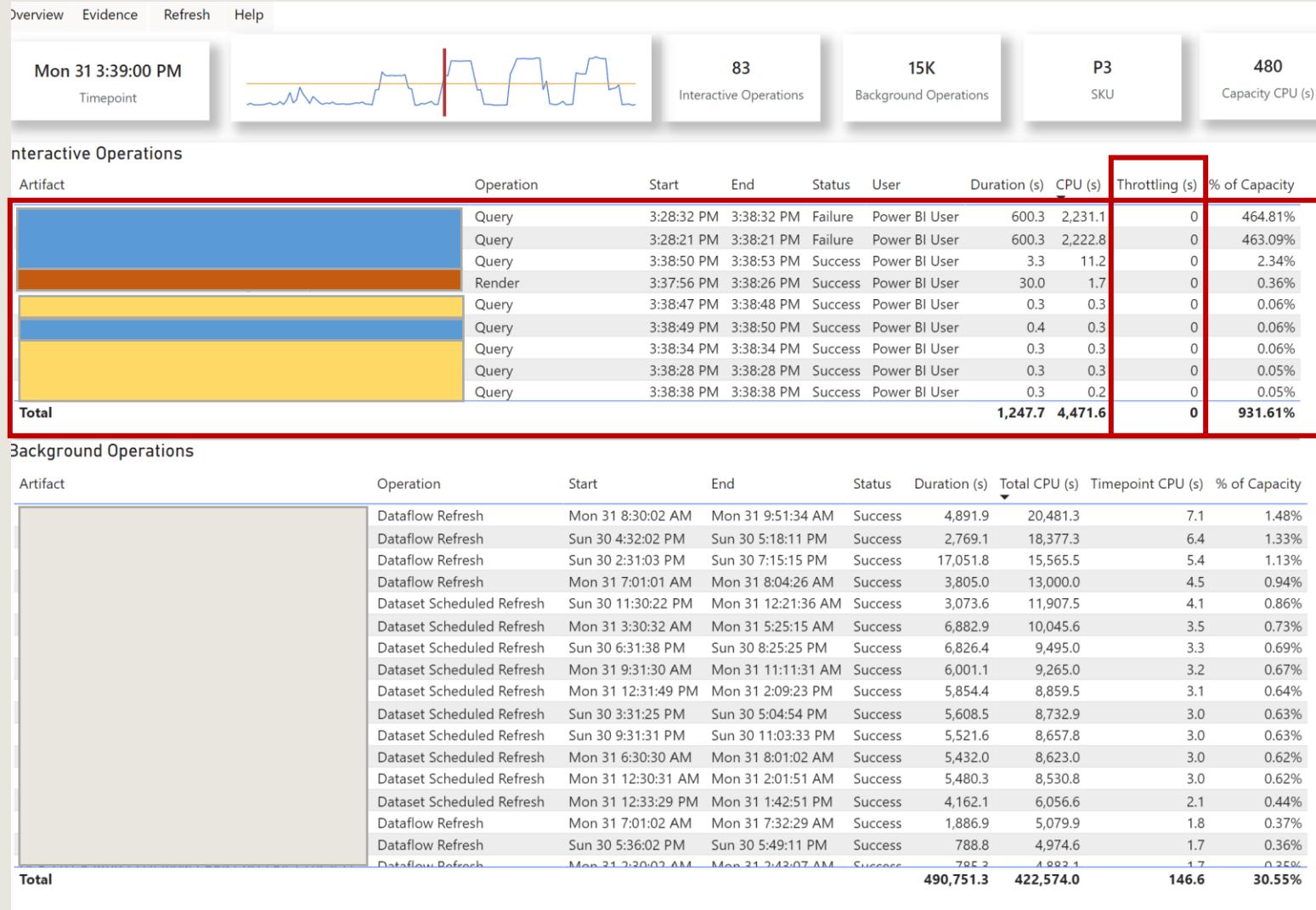
Weekly trendlines
These visuals provide a quick look at high-level trends. Each data point in these visuals shows the value for a rolling seven-day period.



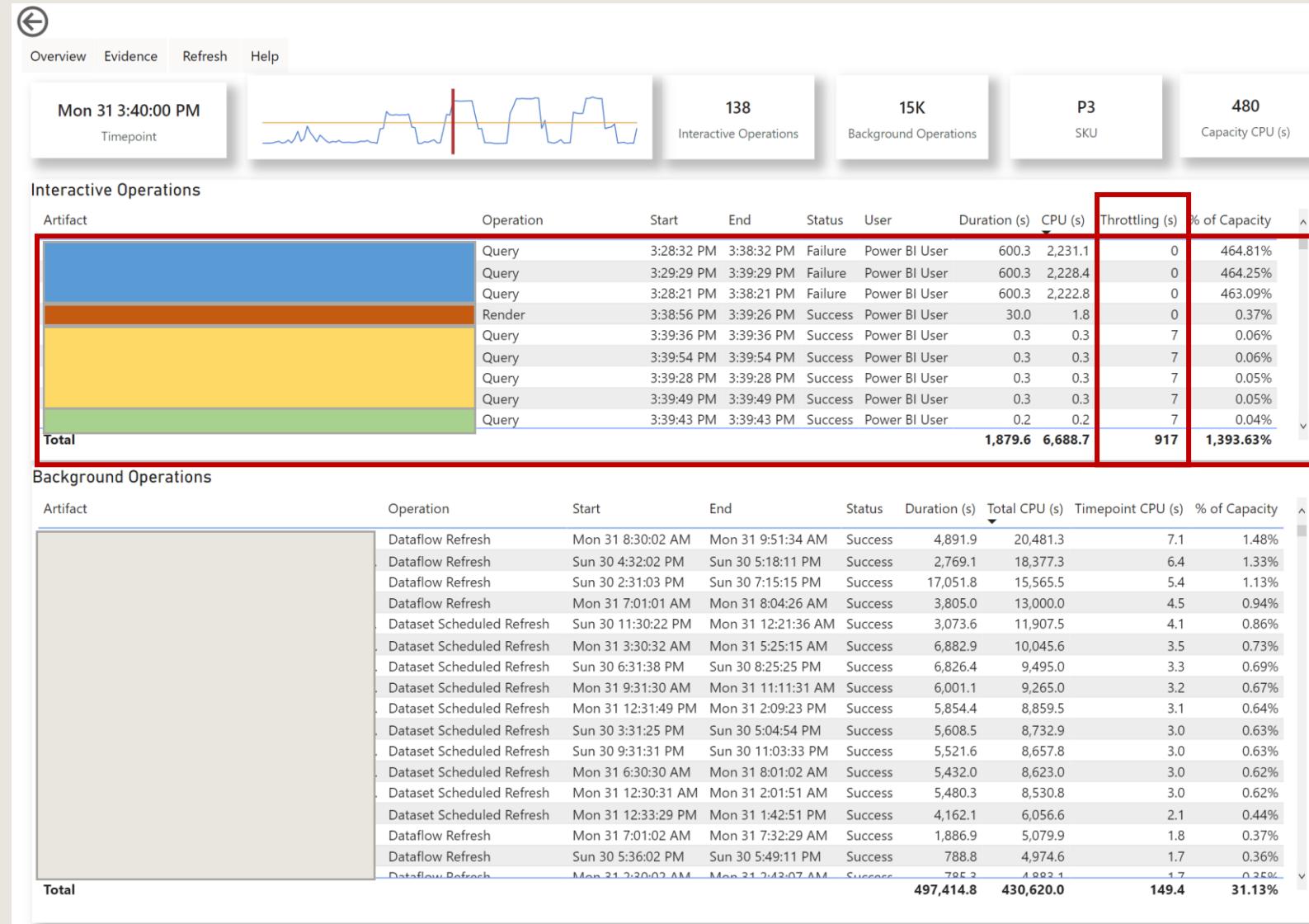
TIMEPOINT ANALYSIS



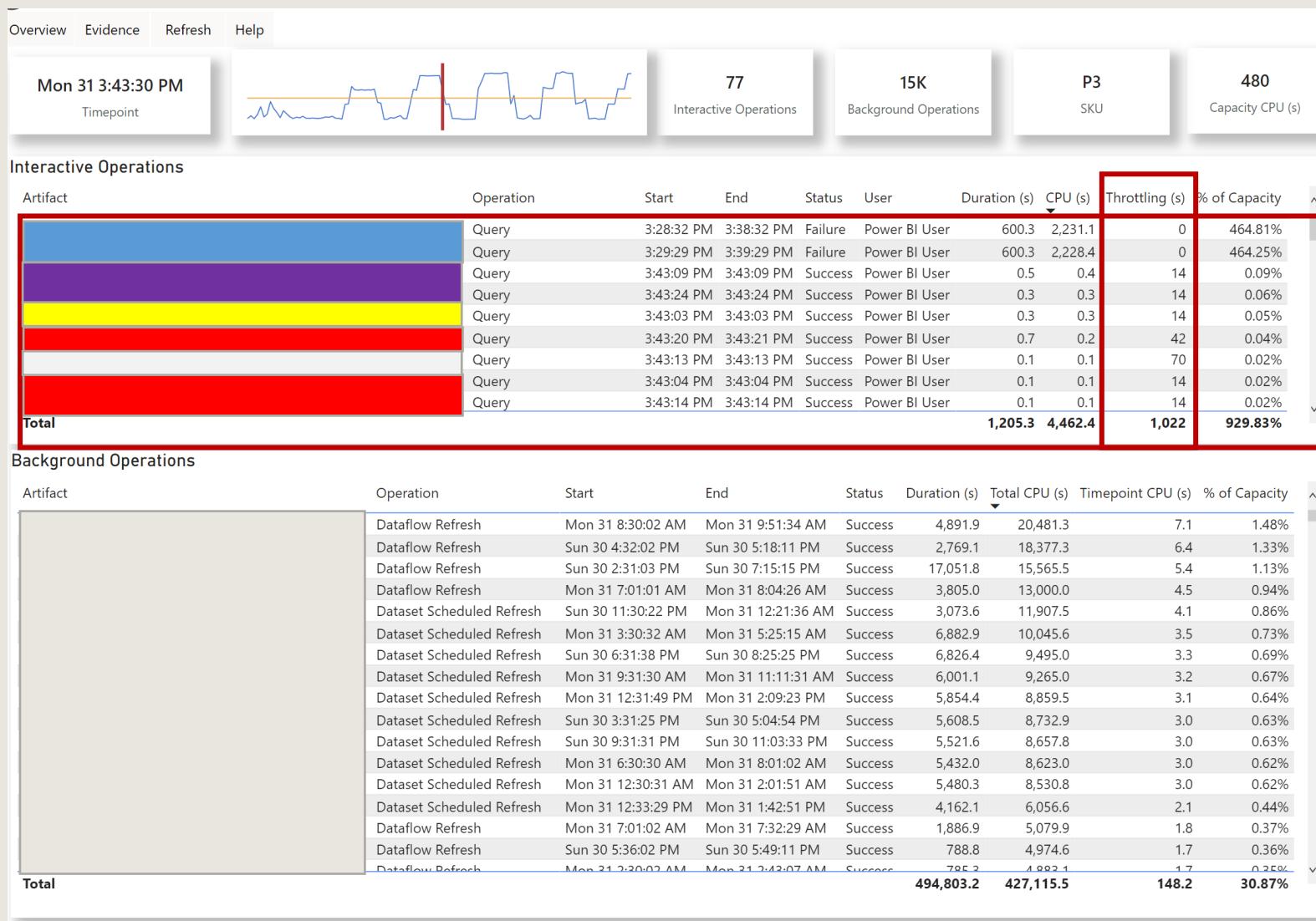
TIMEPOINT ANALYSIS



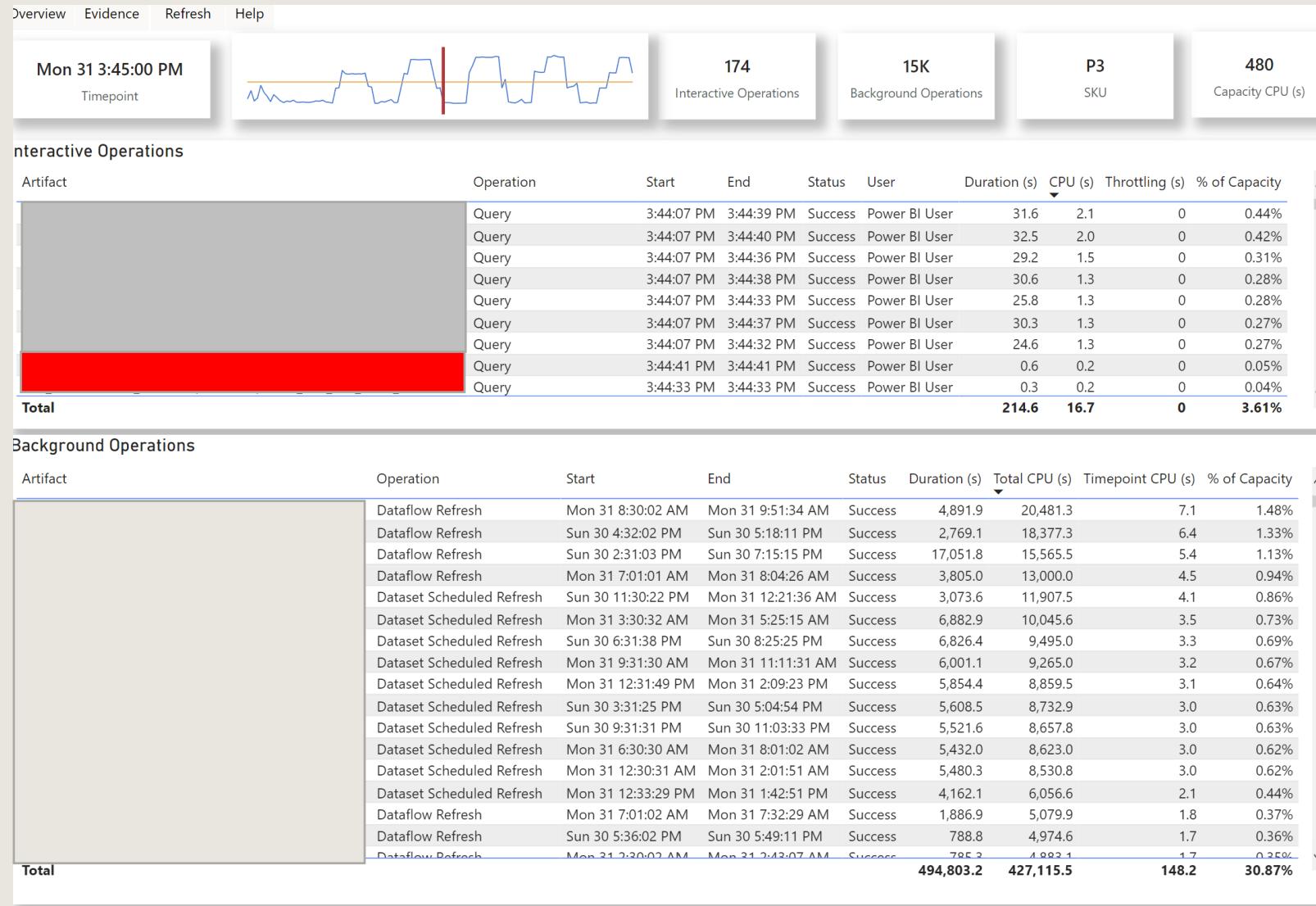
TIMEPOINT ANALYSIS



TIMEPOINT ANALYSIS



TIMEPOINT ANALYSIS



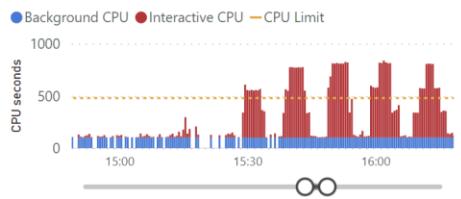
PROACTIVE MONITORING

Premium Capacity Usage and Resource Utilization

Overview Evidence Refresh Help

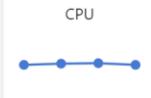
Capacity Name
PRD1

Pick a capacity to review in the Capacity Name picker. All visuals on the page will refresh each time a capacity is picked. Your selection will apply to other pages as well. Learn how to use this page by clicking the "info" button.



Weekly trendlines

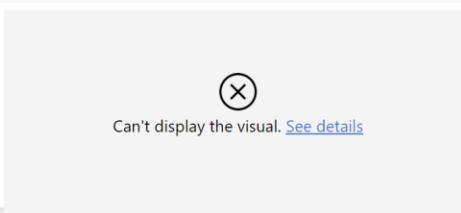
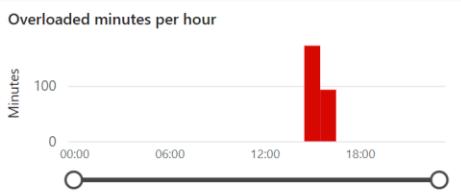
These visuals provide a quick look at high-level trends. Each data point in these visuals shows the value for a rolling seven-day period.

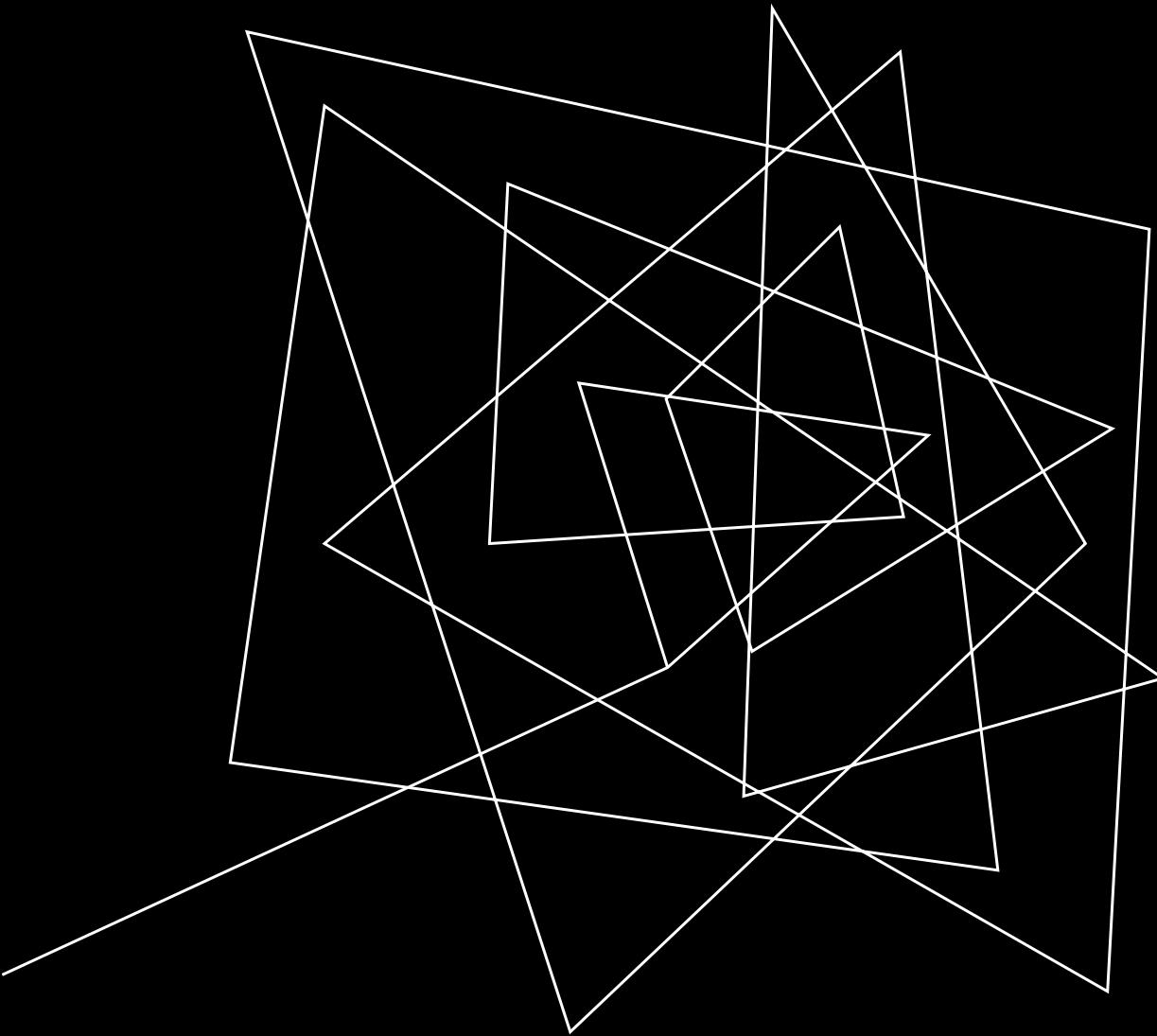


Dataflow Dataset PaginatedRe ?

Artifacts (1 days)

Artifact	CPU (s)	Duration (s)	Users	Artifact Size (GB)	Overloaded minutes	Performance delta
	503	1,248	323	0.62	27.5	
	94	108	5	1.15	8.0	
	26	43	17	0.11	7.0	
	18	89	31	0.08	7.0	
	139	146	14	0.68	6.0	
	107	130	34	0.69	6.0	
Total	40,340	69,637	323	11.90		





BYOLA?!

BRING YOUR OWN LOG ANALYTICS

Azure Monitor Integration

Flight recorder for AS Engine events

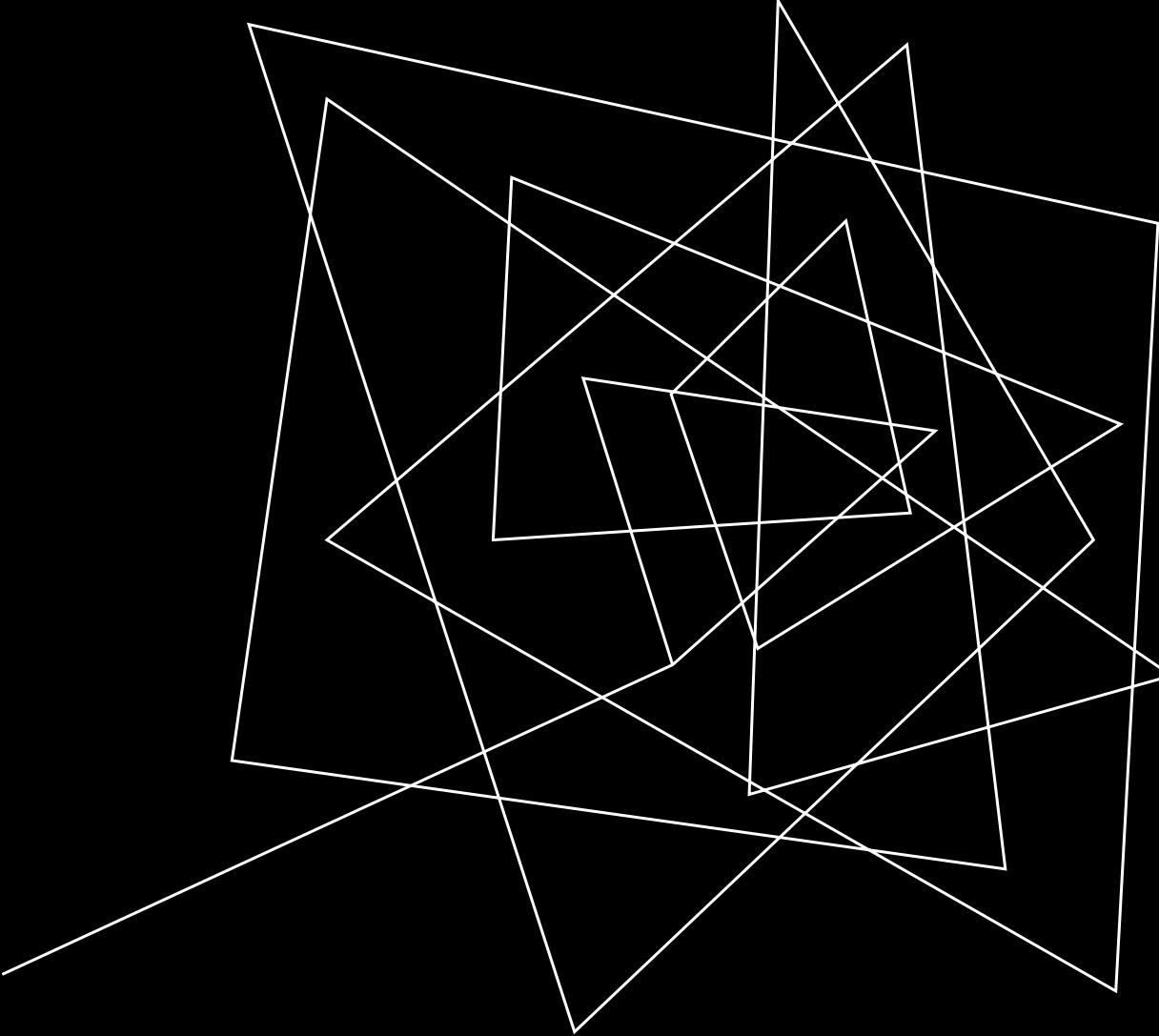
Can help you obtain:

Query Details, Users, Application

Workspace, Report, Visual, ..

Use this for deep dive investigations

[Chris Webbs' series on BYOLA](#)



CAPACITY SETTINGS

CAPACITY SETTINGS

Consider your notification strategy

TIP: Balance information vs. alert overload

Notifications

Get notified when you're close to exceeding your available capacity (which includes base and Autoscale v-cores).

Send notifications when

- You're using % of your available capacity
- You've exceeded your available capacity and might experience slowdowns
- An Autoscale v-core has been added
- You've reached your Autoscale maximum

Send notifications to

- Capacity admins
- These contacts:

Enter email addresses

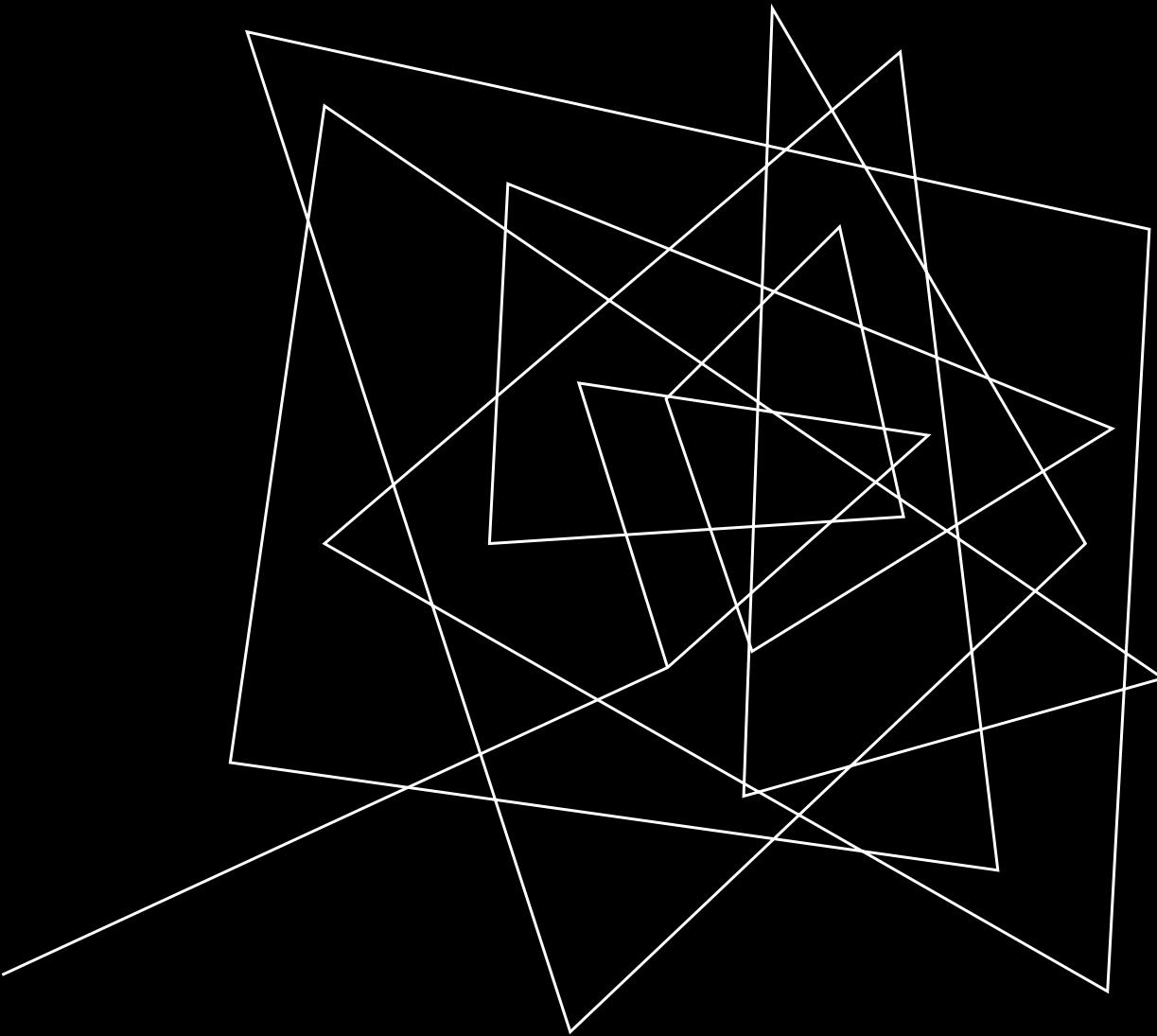
CAPACITY WORKLOAD SETTINGS

Not all Gen 1 workload settings were transferred

Check the default capacity settings

TIP: Don't play around with these on a production instance

AI	Allow usage from Power BI Desktop
<input checked="" type="checkbox"/> On	
DATASETS	
Query Memory Limit (%)	<input type="text" value="0"/>
Query Timeout (seconds)	<input type="text" value="3600"/>
Max Intermediate Row Count	<input type="text" value="1000000"/>
Max Result Row Count	<input type="text" value="2147483647"/>
Max Offline Dataset Size (GB)	<input type="text" value="0"/>
Automatic page refresh	<input checked="" type="checkbox"/> On
Minimum refresh interval	<input type="text" value="1"/> Seconds ▾
Change detection measure	<input checked="" type="checkbox"/> On
Minimum execution interval	<input type="text" value="30"/> Seconds ▾
XMLA Endpoint	<input type="text" value="ReadWrite"/>



IT'S ALL ABOUT
THE €€€

POWER BI PRO LICENSES

TIP: Consumers don't need Power BI Pro Subscriptions

Users will not need a Power BI Pro Subscription when :

They are a viewer in an App Workspace (v2)

A member of an App Workspace App

Users will need a Power BI Pro Subscription when :

They are an Admin, Member, Contributor in any Workspace

Sharing content with other users in the organisation

Consuming content that is not hosted on Power BI Premium

POWER BI EMBEDDED (A SKU)

A Sku has close to feature parity

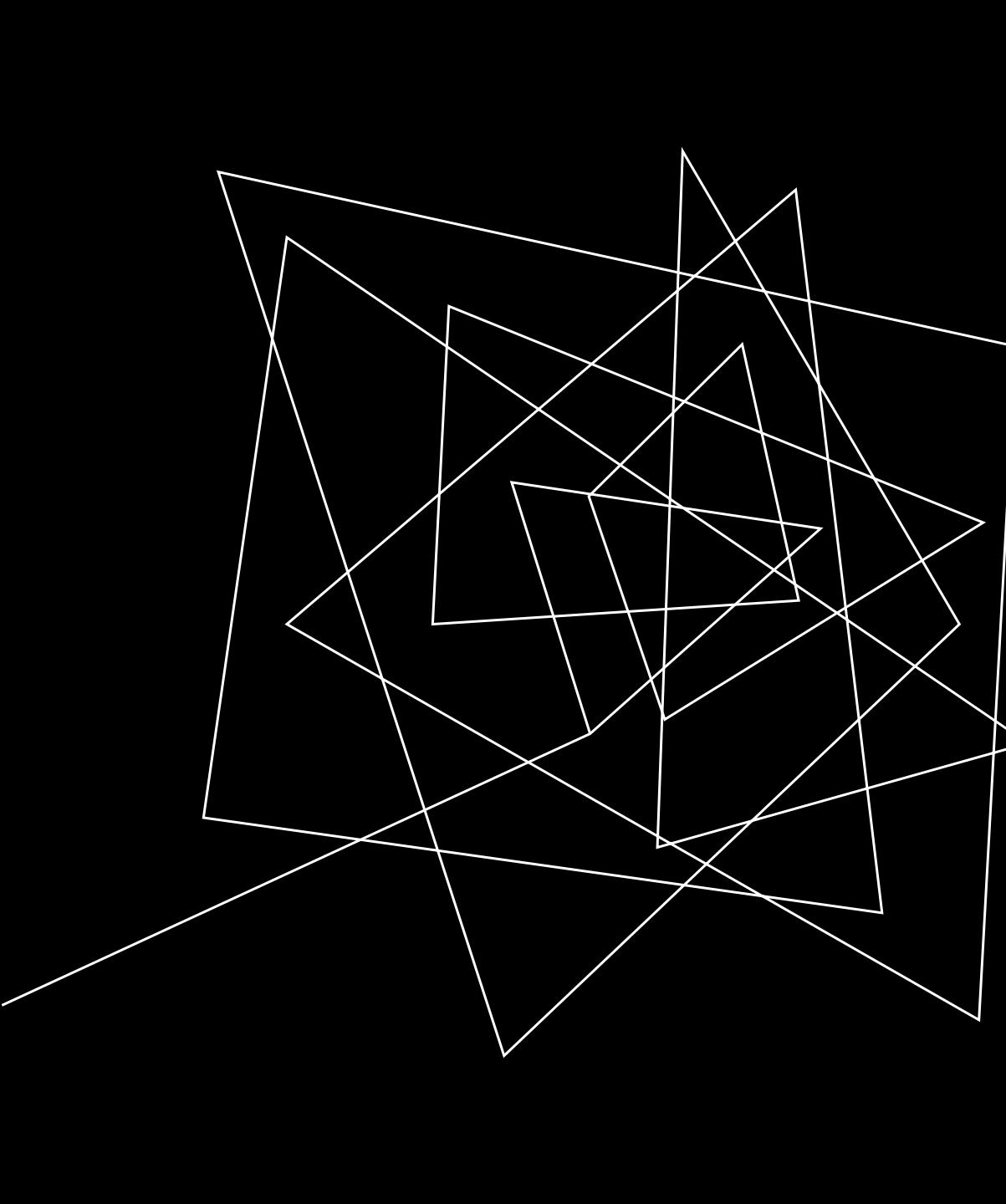
Billed by the minute to Azure Billing

Don't forget to set up Billing Alerts, Policies, and Automation

TIP: You can spin up A SKU to host certain 'Premium Features' when needed, to reduce usage on the P SKU

Ie. Use an A SKU for the 'DEV' and 'ACC' workspace of Deployment Pipelines

A Sku can not be used for organisational production usage



**POWER BI
PREMIUM PER
USER**

PREMIUM PER USER (PPU)

PPU is a subscription based model, based on Premium Gen 2

Introduced at Ignite (H2 2020)

General Availability on April 2nd 2021

Pricing:

- 10\$ as an add-on to Power BI Pro / Office 365 E5
- 20\$ as a standalone subscription

PREMIUM PER USER (PPU)

Access to Premium features on a per user basis

Every tenant has a PPU Capacity (P3 equivalent)

Workspaces are assigned to the Capacity

PPU Content requires every consumer to have a subscription

Aimed at smaller organisations, or offloading scenarios

PREMIUM PER USER (PPU)

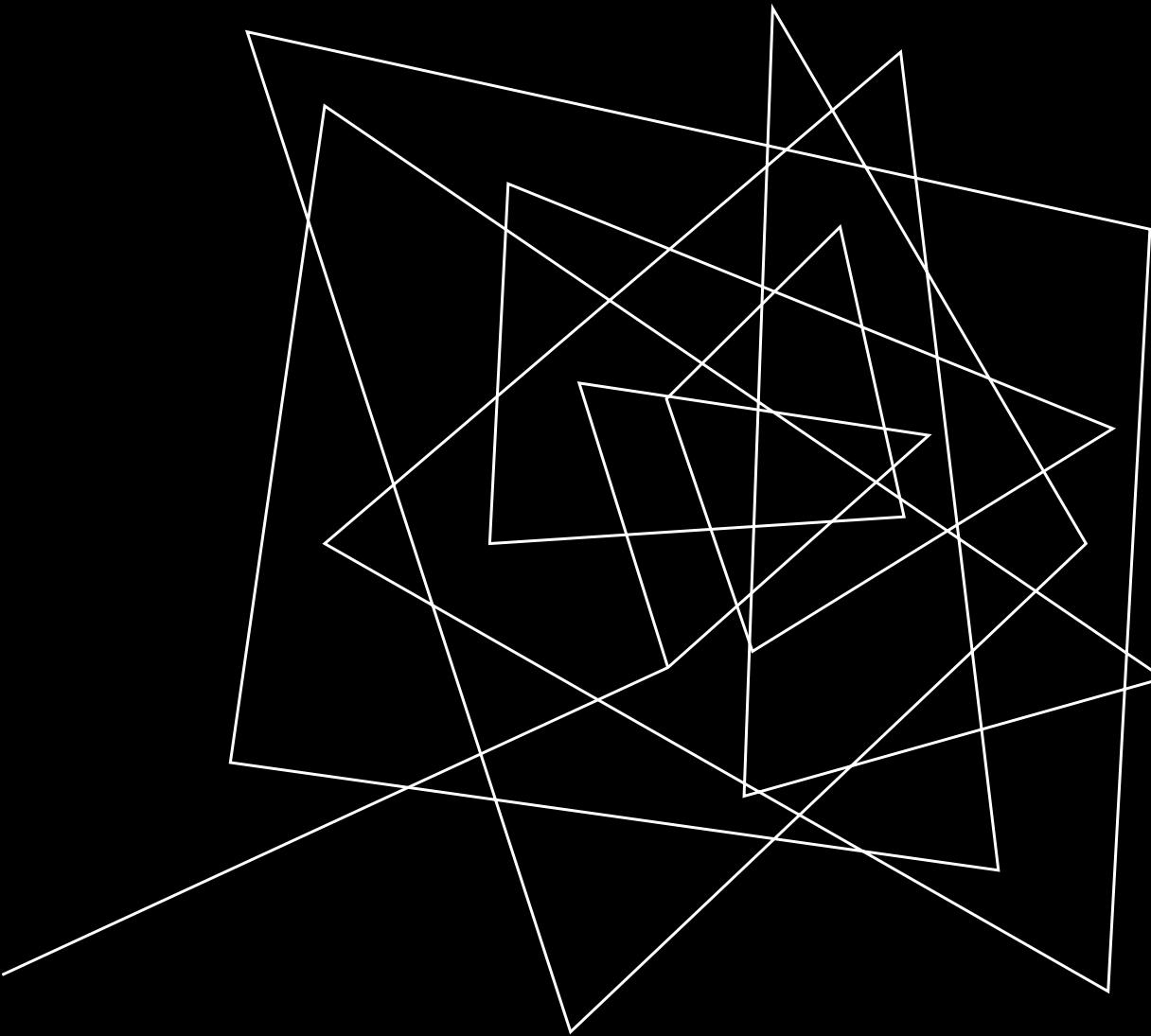
Admin portal

- Usage metrics
- Users
- Premium Per User
- Audit logs
- Tenant settings
- Capacity settings
- Refresh summary
- Embed Codes
- Organizational visuals
- Azure connections (preview)
- Workspaces
- Custom branding
- Protection metrics
- Featured content

Premium Per User

- ▲ Auto Refresh
 - Automatic page refresh
 - On
 - Minimum refresh interval
 - 5 Minutes ▾
 - Change detection measure
 - On
 - Minimum execution interval
 - 30 Seconds ▾
- ▲ Dataset workload settings
 - XMLA Endpoint
 - Read Write ▾

Apply **Discard**



TAKEAWAYS

TAKEAWAYS

Check (and adjust) the Capacity Workload Settings

Premium Capacity Metrics App v2 is your Ops Base

Evictions are perfectly normal, don't stress

Set up auditing to have a complete overview of what's happening

Consider offload non-production workloads to A Sku / PPU

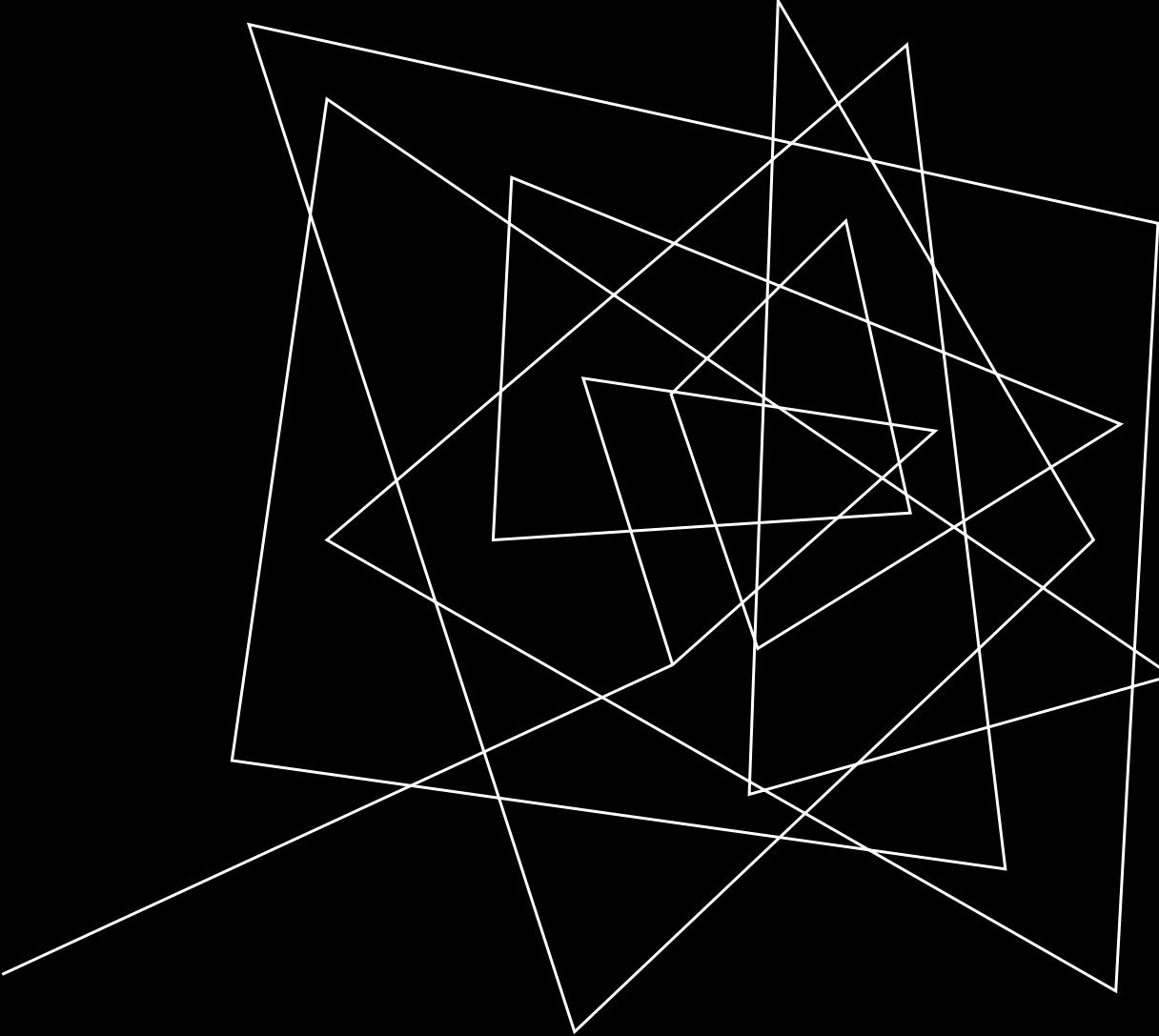
TAKEAWAYS

Understand the shift in usage metrics (Memory -> CPU)

When using Gen 2 AutoScale, set up consumption limits

Usage Patterns are still key, even though the need is less pressing

Leverage BYOLA for in-depth analysis



REFERENCES

RESOURCES

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-gen2-what-is>

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-transition-gen1-to-gen2>

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-architecture>

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-concepts>

RESOURCES

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-capacity-manage-gen2>

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-auto-scale>

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-install-gen2-app?tabs=1st>

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-gen2-metrics-app>

RESOURCES

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-gen2-faq>

<https://docs.microsoft.com/en-us/power-bi/admin/service-premium-per-user-faq>

RESOURCES

[Do we need more Power BI Premium Capacity?](#)

[Power BI Premium Gen 2 Updates \(January 2022\)](#)

[Making Sense of the Power BI Premium Gen 2 Metrics App](#)

[The Power BI Premium Video you need to watch .. probably \(Oct 2021\)](#)

WHO AM I



Power BI CAT

dataMinds.be Member



@BenniDeJagere



/bennidejagere



/bennidejagere



/bennidejagere



#SayNoToPieCharts



SLIDES CAN BE FOUND AT :

[https://github.com/BenniDeJagere/Presentations
/{Year}/{Date}_{Event}](https://github.com/BenniDeJagere/Presentations/{Year}/{Date}_{Event})

