Demystifying Capacity Metrics

Daan Humblé

Agenda

- What are capacities

- How do capacities work

- Capacity Metrics App

- Tips & Tricks
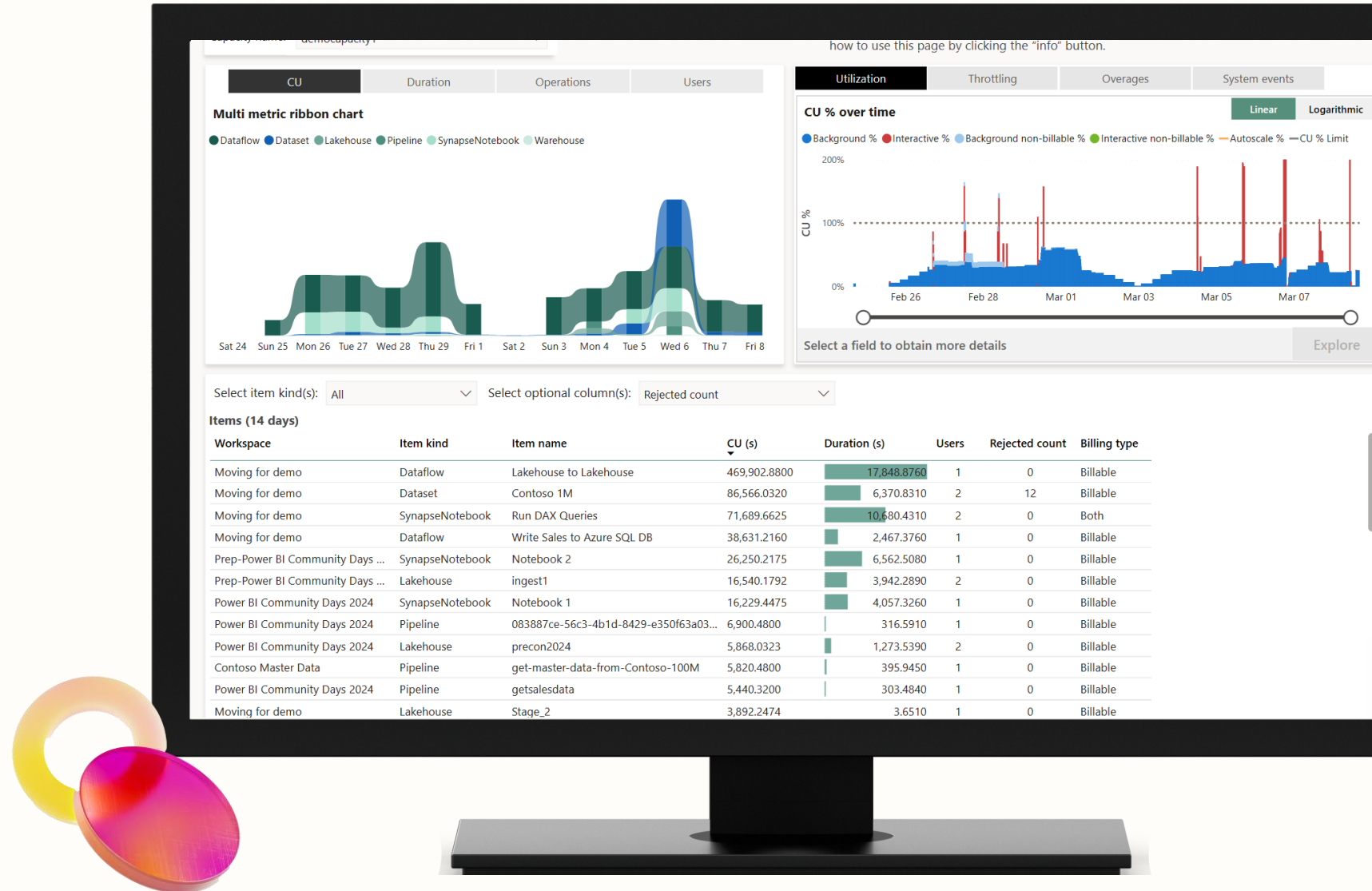
Daan Humblé

Cloud Solution Architect Power BI & Fabric

linkedin.com/in/daanhumble/

Microsoft Fabric

Intelligent data platform

Data Factory	Data Engineering	Data Warehouse	Data Science	Real Time Analytics	Power BI	Data Activator

Powered by AI with Copilot in Microsoft Fabric

Serverless
Compute

Unified
Storage

Single Data Foundation

Microsoft Fabric

Intelligent data platform

| Data Factory | Data Engineering | Data Warehouse | Data Science | Real Time Analytics | Power BI | Data Activator |

Powered by AI with Copilot in Microsoft Fabric

Serverless Compute

Unified Storage

Single Data Foundation

Intelligent data platform

Data Factory	Data Engineering	Data Warehouse	Data Science	Real Time Analytics	Power BI	Data Activator

Powered by AI with Copilot in Microsoft Fabric

Serverless Compute

A shared pool of capacity that powers all capabilities in Microsoft Fabric.

Unified Storage

OneLake

A single place to store all data.

Single Data Foundation

Microsoft Fabric capacity

SKU	vCore
P1	8
P2	16
P3	32
P4	64
P5	128

Microsoft Fabric capacity

SKU	vCore
P1	8
P2	16
P3	32
P4	64
P5	128

SKU	CU
F 64	64
F 128	128
F 256	256
F 512	512
F 1024	1024

Microsoft Fabric capacity

SKU	vCore
P1	8
P2	16
P3	32
P4	64
P5	128

SKU	CU
F 64	64
F 128	128
F 256	256
F 512	512
F 1024	1024

Additional Fabric SKUs

F SKU	Capacity unit (CU)
F 2	2
F 4	4
F 8	8
F 16	16
F 32	32

Microsoft Fabric capacity

A shared capacity that provides a set of capacity units (CUs) powering the compute for Data Warehouse, Data Integration, Data Science, Data Engineering, Real-Time Analytics, Power BI, Data Activator, and Copilot.

Simplicity of purchasing	**Flexibility in use**	**Cost savings**	**Value optimization**	**Transparent monitoring**
No need to purchase separate compute for each workload	With no pre-allocation required, the same set of CUs can be used for any workload	All CUs are pooled together and are not locked to an idle workload. Smoothing on compute usage also allows you to size capacity closer to the average usage than the peaks	Adjust compute capacity by scaling up and down	Centralized dashboard to monitor usage and costs in Microsoft Fabric utilization and metrics app

Compute Power

Capacity units (CUs) = Compute power

Capacity units (CUs) are units of measure that represent a pool of compute power needed. Compute power is required to run queries, jobs, or tasks.

SKU	CU
F 64	64

F64 gives you
64 CUs per second

Consumption

Capacity units (CUs) = Compute power

Capacity units (CUs) are units of measure that represent a pool of compute power needed. Compute power is required to run queries, jobs, or tasks.

CU Consumption

The CU consumption is highly correlated to the underlying compute effort needed for the tasks performed by the capability during the processing time.

Each capability, such as Power BI, Spark, Data Warehouse, with the associated queries, jobs, or tasks has a unique consumption rate.

The inner-workings of your capacity
Two types of activities

Interactive

(DAX) Queries
XMLA Endpoint

Background

Dataset refreshes
Dataflow refreshes

Lakehouse
Notebooks
Real Time Analytics
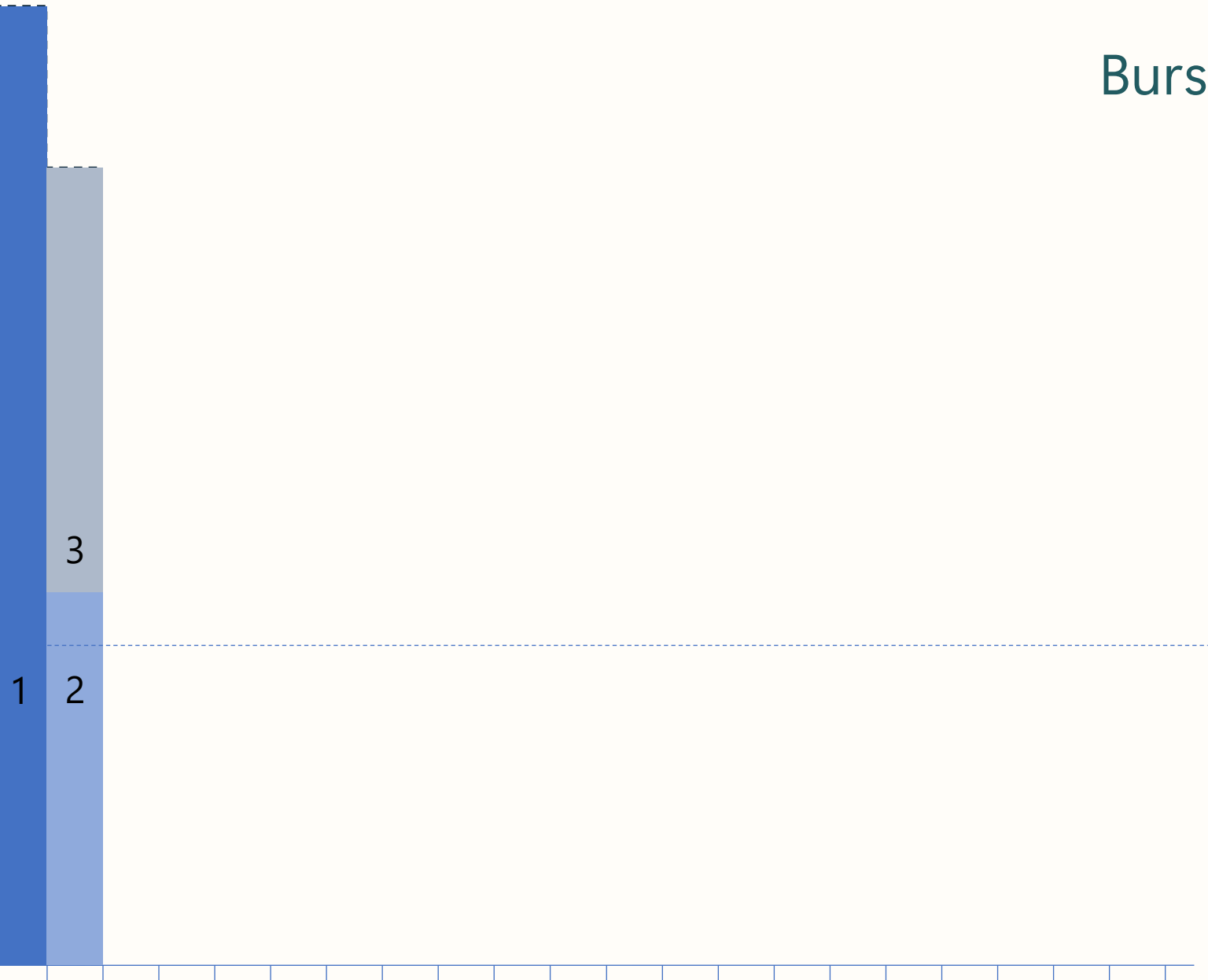Copilot

The inner-workings of your capacity

Bursting
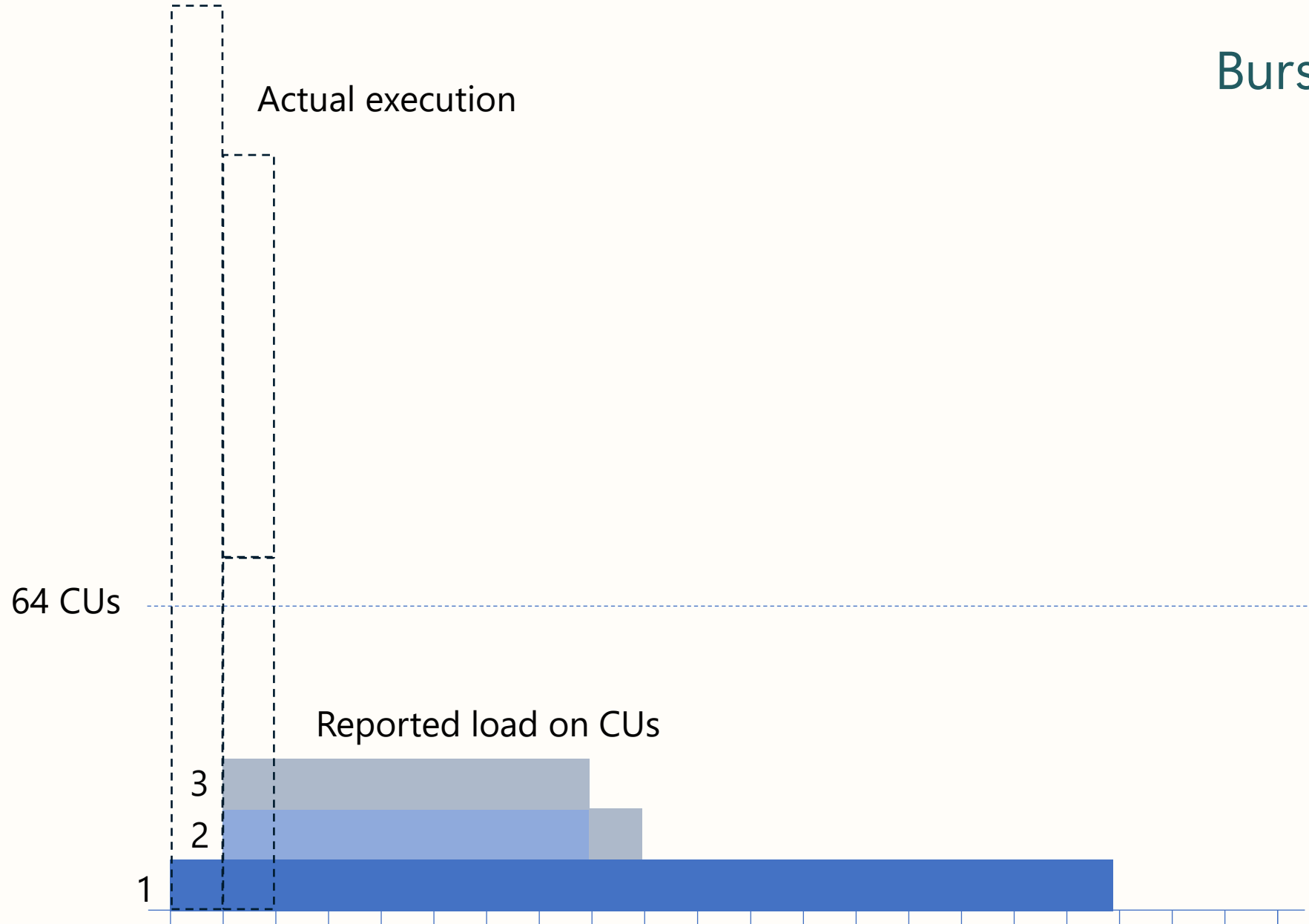
Smoothing

Carry Forward

Overages

Carry Forward

64 CUs

Carry Forward

Usage above limits "carries forward"

64 CUs
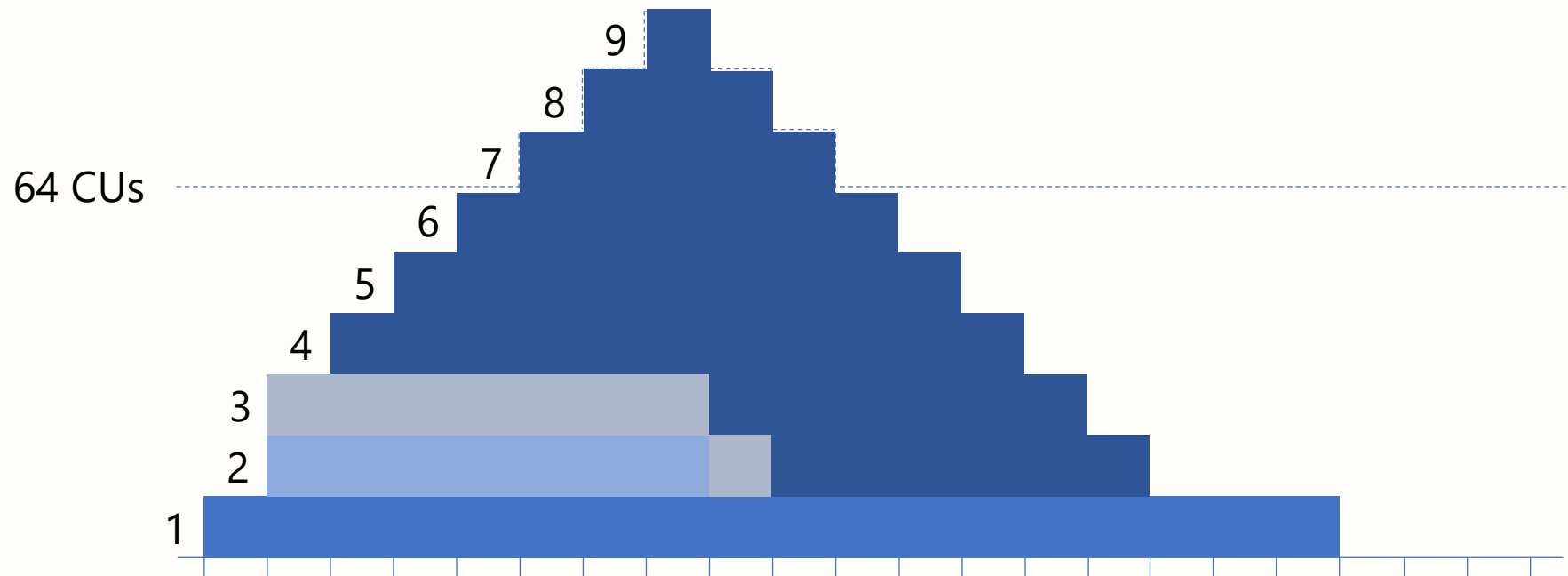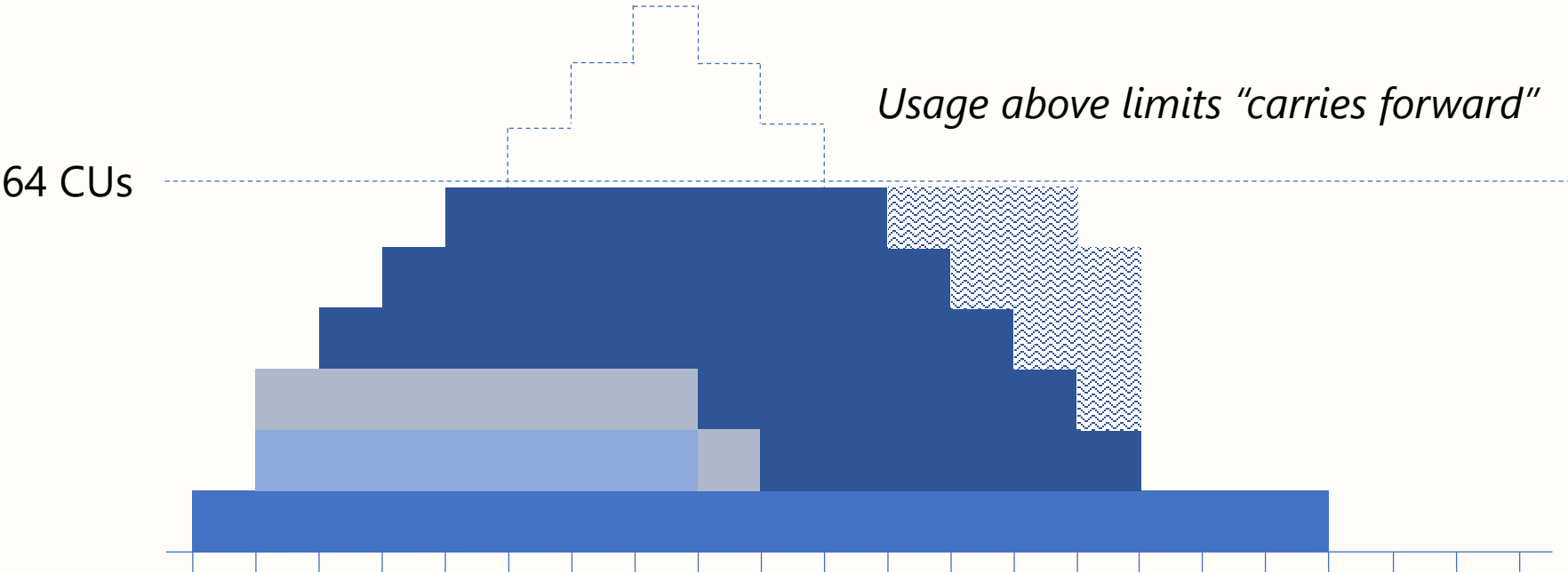
Up to 10 minutes

Overages

What happens if there isn't enough?

Approaching the limit

Warning & emails to capacity admins

Exceeding the limit

10 minutes future capacity = interactive delay

60 minutes future capacity = interactive rejection

24 hours future capacity = background rejection

Remedy an overload

Reduce consumption

Upgrade

Wait

Pause the capacity

Unable to load model due to reaching capacity limits

×

Unable to open this report because your organization's compute capacity has exceeded its limits. Try again later.

Please check the technical details for more information. If you contact support, please provide these details.

See details ⌄

Learn more about capacity limits

Close

Demo

Demo

F2 vs F64

```
EVALUATE
ROW (
    "x", COUNTROWS (
        CROSSJOIN (
            SELECTCOLUMNS ( CALENDAR ( 1, 10000 ), "Num1", INT([Date]) ),
            SELECTCOLUMNS ( CALENDAR ( 1, 10000 ), "Num2", INT([Date]) )
        )
    )
)
```

Demo (results)

Cache	Statistic	TotalDuration	SE Duration
Cold	Average	7.057,00	0,00
Cold	StdDev	145,94	0,00
Cold	Min	6.844,00	0,00
Cold	Max	7.237,00	0,00
Warm	Average	6.991,80	0,00
Warm	StdDev	406,63	0,00
Warm	Min	6.643,00	0,00
Warm	Max	7.563,00	0,00

F2

Cache	Statistic	TotalDuration	SE Duration
Cold	Average	6.884,80	0,00
Cold	StdDev	372,23	0,00
Cold	Min	6.594,00	0,00
Cold	Max	7.500,00	0,00
Warm	Average	7.325,40	0,00
Warm	StdDev	331,66	0,00
Warm	Min	7.047,00	0,00
Warm	Max	7.860,00	0,00

F64

Monitor your capacity in the capacity metrics app

- Tenant wide visibility into capacity usage for all Fabric workloads in one place

- Identify resource usage trends across artifacts and operations

- Monitor the impact of throttling to user experience to make scale-up decisions

- View preview usage in parallel with production workloads to make data-driven capacity sizing decisions
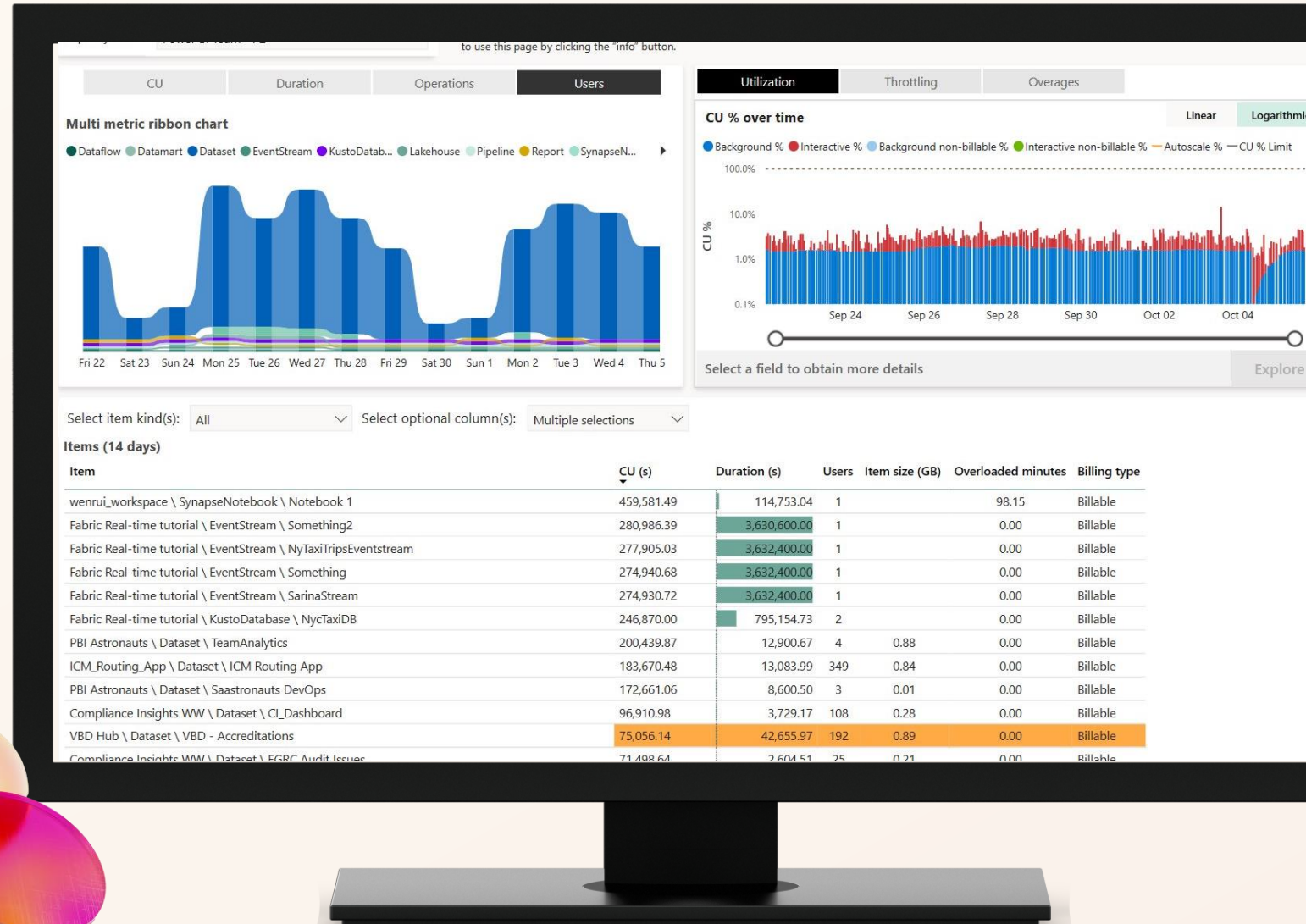
Investigate usage in the Timepoint-details

- Zoom in to analyze workload operations and artifacts with full fidelity down to 30 second granularity

- View the impact of long-running jobs against capacity limits with smoothing applied

- Realtime visibility of in progress operations

- Analyze user experience to efficiently scale and optimize

OneLake Storage

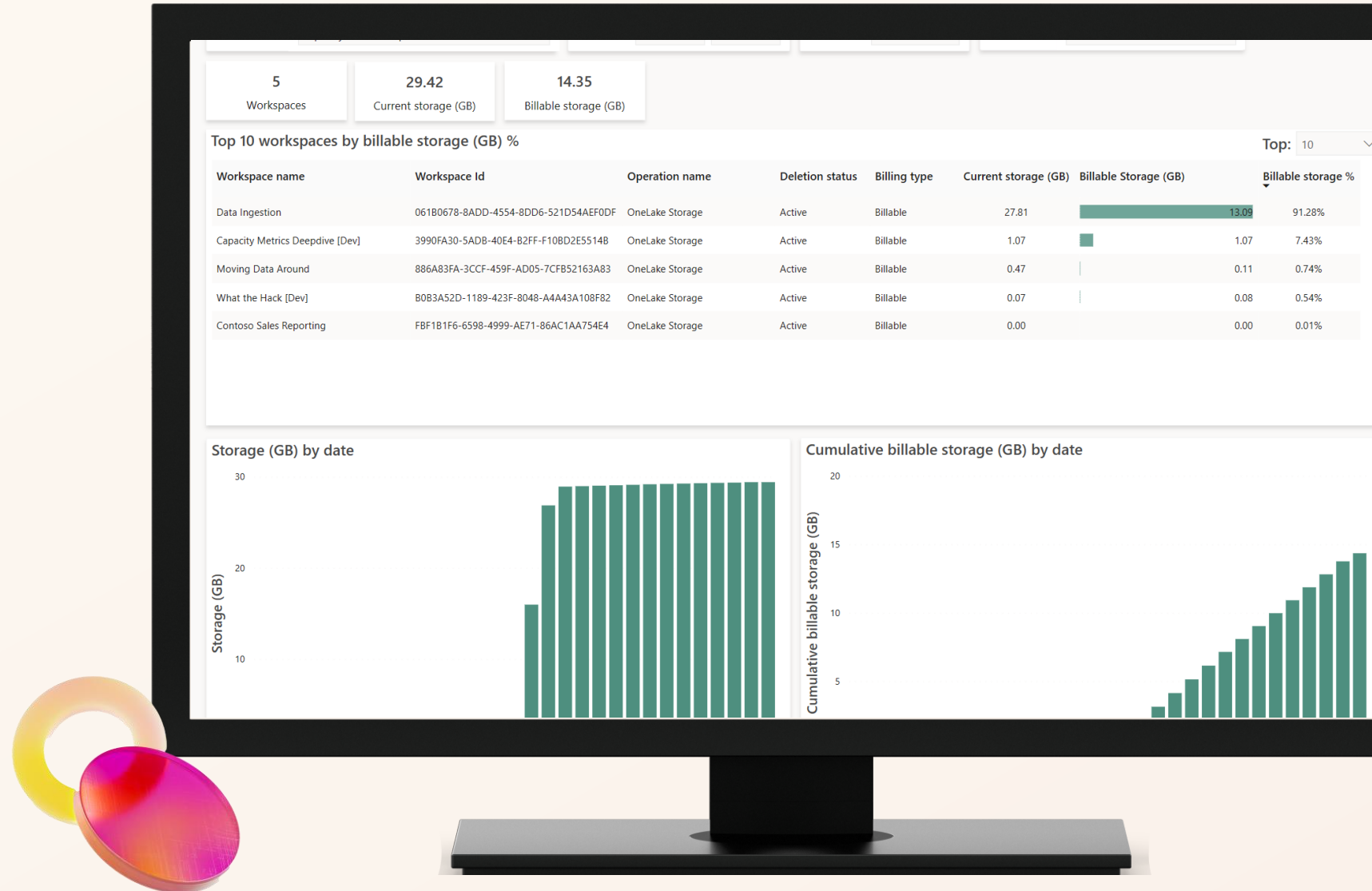|		
5 Workspaces	**29.42** Current storage (GB)	**14.35** Billable storage (GB)

Top 10 workspaces by billable storage (GB) %

Top: 10

Workspace name	Workspace Id	Operation name	Deletion status	Billing type	Current storage (GB)	Billable Storage (GB)	Billable storage %
Data Ingestion	061B0678-8ADD-4554-8DD6-521D54AEF0DF	OneLake Storage	Active	Billable	27.81	13.09	91.28%
Capacity Metrics Deepdive [Dev]	3990FA30-5ADB-40E4-B2FF-F10BD2E5514B	OneLake Storage	Active	Billable	1.07	1.07	7.43%
Moving Data Around	886A83FA-3CCF-459F-AD05-7CFB52163A83	OneLake Storage	Active	Billable	0.47	0.11	0.74%
What the Hack [Dev]	B0B3A52D-1189-423F-8048-A4A43A108F82	OneLake Storage	Active	Billable	0.07	0.08	0.54%
Contoso Sales Reporting	FBF1B1F6-6598-4999-AE71-86AC1AA754E4	OneLake Storage	Active	Billable	0.00	0.00	0.01%

Storage (GB) by date

Cumulative billable storage (GB) by date

Tips & Tricks

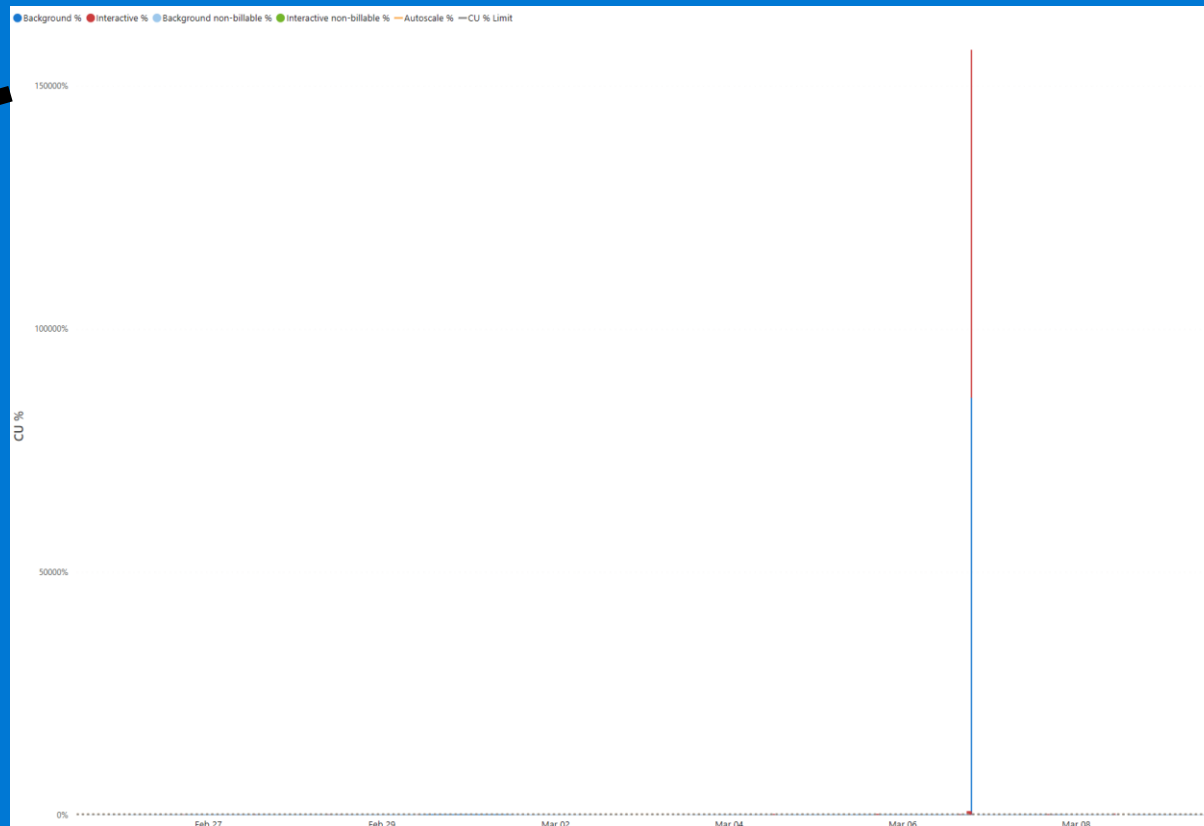# Remedy an overload
by reducing consumption

Interactive

Star Schema!

Improve DAX

Reduce number of visuals

Use aggregations

Consult heavy user

Background

Star Schema!

Reduce manual refreshes

Incremental refresh

Aggregations

Remedy an overload
by pausing the capacity

Guardrails

Notifications:

Power BI Workloads:
- Query Memory Limit (%)
- Query Timeout

Other workloads:
- Warehouse
- Spark pools

⊿ Notifications

Send notifications when
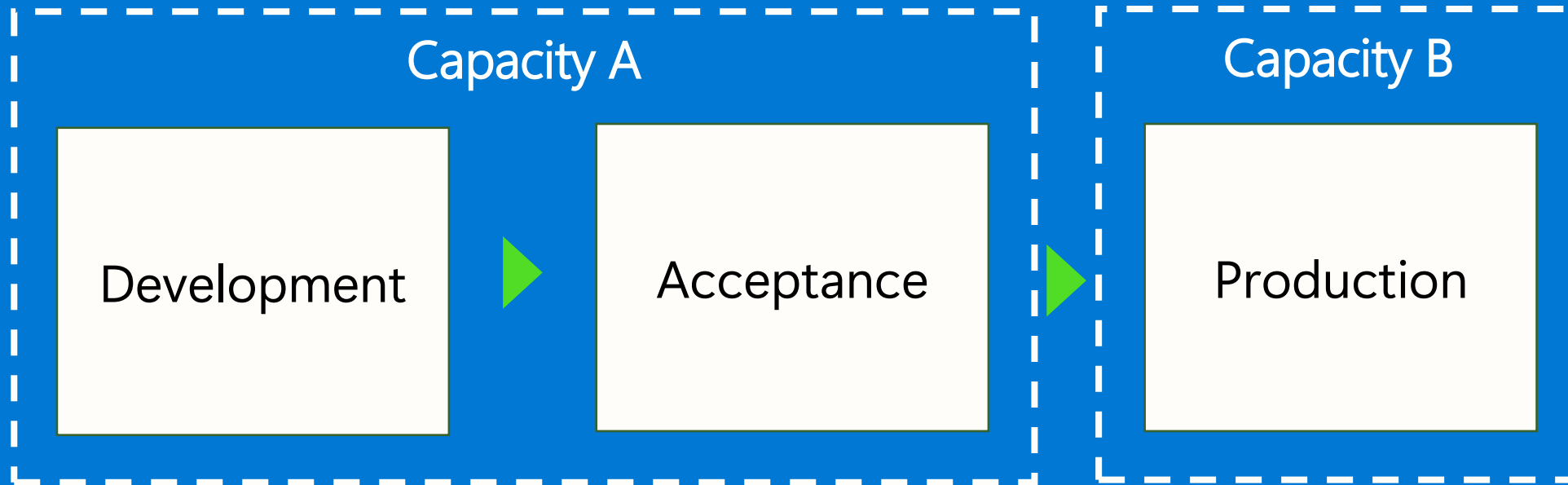
☐ You're using [____] % of your available capacity

☐ You've exceeded your available capacity and might experience slowdowns
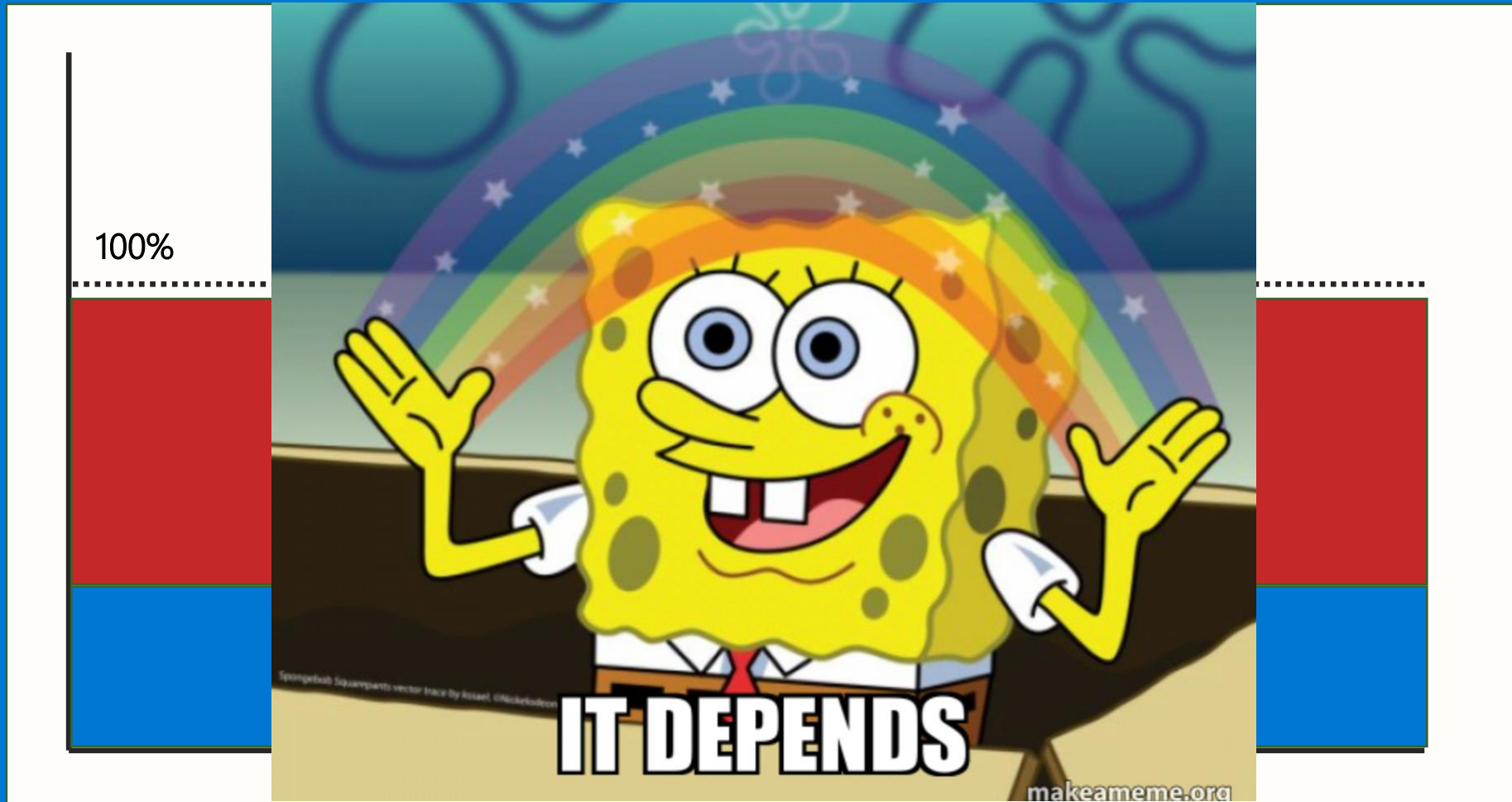
Send notifications to

☐ Capacity admins

☐ These contacts:

[Enter email addresses_____]

[Apply] [Discard]

Isolate heavy workloads

Capacity A

Capacity B

Development ▶ Acceptance ▶ Production

Ideal set-up*?

*) for Power BI usage

(Load) testing

Notifications | Fork 93 | Star 133

<> Code | ⊙ Issues 27 | ⊞ Pull requests 3 | ▷ Actions | ⊟ Projects | ⊙ Security | ⌁ Insights

Files

master

Go to file

> 📁 LoadTestingPowerShellTool
∨ 📁 RealisticLoadTestTool
  > 📁 DemoLoadTest1
  > 📁 DemoLoadTest2
  📄 PBIReport.json
  📄 PBIToken.json
  📄 Power BI Load Testing Demo.pbix
  📄 Readme.md
  📄 RealisticLoadTest.html
  📄 Run_Load_Test_Only.ps1
  📄 Setup_Load_Test.ps1
  📄 Update_Token_Only.ps1
> 📁 doc_images

PowerBI-Tools-For-Capacities / RealisticLoadTestTool / **Readme.md** ⧉

👤 **furmangg** adding layoutType to support MobilePortrait, MobileLandscape or Maste... ⋯   c292212 · 5 years ago | 🕘 History

Preview | Code | Blame   110 lines (94 loc) · 14.1 KB   Raw ⧉ ⬇ ☰

Power BI Realistic Load Test Tool

For "realistic user scenario" load testing to simulate a realistic set of user actions such as changing slicers, changing filters, clicking bookmarks, and simulating user "think time" where a user studies the report before clicking again, use this Realistic Load Test Tool.

Change Log

- 8/1/2019 - Initial check-in containing the following enhancements to the original LoadTestingPowerShellTool code:
  - Support for multiple slicers/filters and support for randomly cycling through filter combinations to achieve cold cache load tests
  - Support for text and integer report-level filters
  - Support for text and integer slicers
  - Support for specifying the report page to display

More detailed monitoring
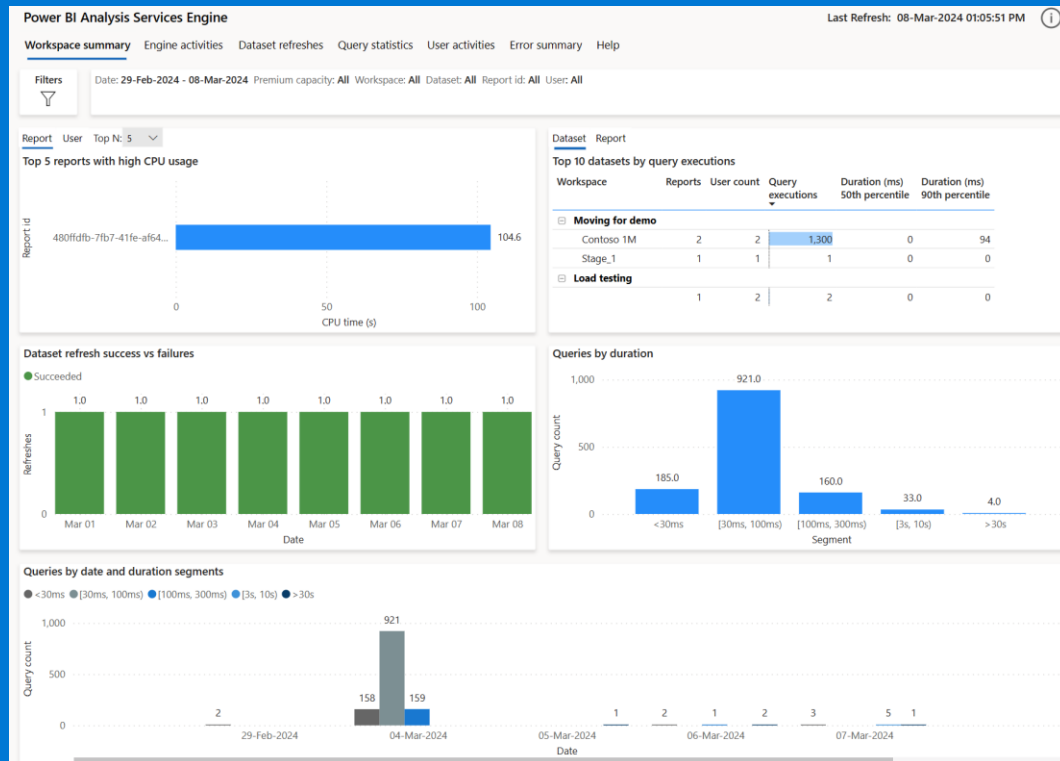
Log Analytics

GitHub - microsoft/PowerBI-LogAnalytics

PBI Monitor

GitHub - RuiRomano/pbimonitor

Wrap up

Plan for the average usage

Keep grip with the Metrics App

Optimize Optimize Optimize