

Optimizing Data Warehousing
Challenges with **Microsoft Fabric**

A Deep Dive into **Burstable Capacity, Smoothing, and Throttling Strategies**

Abstract

In the dynamic data management landscape, mastering the challenges of data warehousing has become a pivotal task for organizations seeking optimal efficiency while managing complexities such as unpredictable demand and controlling costs. This article explores how Microsoft Fabric provides a dynamic and scalable solution to overcome these challenges with innovative features, including burstable capacity, smoothing, and throttling. These features are instrumental in transforming the data warehousing paradigm, offering a comprehensive solution to enhance performance, scalability, and overall operational excellence. With a pay-as-you-go model and emphasis on elasticity and optimal performance, Microsoft Fabric enables enterprises, especially small and medium businesses (SMEs), to gain insights, facilitate decision-making, and streamline their business processes.

Burstable capacity, smoothing and throttling demystified

At its core, Microsoft Fabric, as a cloud-based platform, extends its burstable capacity, smoothing, and throttling features beyond the mentioned services, making them applicable to a broader spectrum, including real-time analytics, data science, data engineering, and power BI. Within this dynamic platform, Fabric capacity refers to a pool of resources whose size (or SKU) determines the computational power available. The Warehouse and SQL analytics endpoint brings forth crucial features such as burstable capacity and smoothing for data warehousing workloads. Additionally, it includes throttling policies designed to manage resource consumption and workload performance.

Microsoft Fabric Adaptive Data Warehousing Solutions

Data warehousing involves collecting, storing, and analyzing large volumes of data from various sources for organizations seeking insights and process optimization. Data warehousing enables organizations to gain insights, make decisions, and optimize their business processes.

That said, data warehousing also poses some challenges, including:



Addressing varying and unpredictable demands for data analysis



Ensuring peak performance and efficiency of data processing



Managing costs and complexity of data infrastructure

To tackle the challenges, Microsoft Fabric offers a dynamic and scalable solution that allows users to create and manage data warehouses via two services: **Warehouse** and **SQL analytics endpoint**. Operating with a shared pool of resources known as capacity, Microsoft Fabric helps organizations access data stored within its framework. The capacity determines the amount of computational power available for data processing, and can be purchased in different sizes or SKUs that provide a baseline number of capacity units (CUs), the unit of measure for computing resources in Microsoft Fabric.

Despite the baseline capacity's adequacy for routine data warehousing workloads, its limitations become evident in scenarios of heightened complexity, concurrency, and frequency in data analysis queries. Instances like sudden demand spikes during peak hours or resource-intensive complex queries may surpass the capabilities of baseline capacity. In such scenarios, Microsoft Fabric offers a robust solution to optimize data warehousing performance and efficiency by introducing burstable capacity, smoothing, and throttling functionalities.

Fabric Warehouse Capacity Dynamics

In Microsoft Fabric, warehouse, and SQL analytics endpoints offer burstable capacity, smoothing, and throttling. It is ideal for workloads demanding more resources than baseline capacity provides or facing periods of under/over-utilization.

01

Burstable Capacity

Functioning in the background, this feature adapts to the fluctuating demands of a workload, activating based on the workload's scale factor. It ensures optimal performance without surpassing the initially procured baseline capacity. This capability is particularly beneficial for data warehousing workloads with high demands or complex queries, enhancing response time and throughput, and managing unpredictable data analysis requirements.

02

Smoothing

This spreads capacity consumption over a more extended 24-hour period to reduce the impact of short-term demand spikes, especially during off-peak and peak hours. Smoothing can help simplify capacity management and avoid throttling when capacity consumption exceeds purchased capacity limits.

03

Throttling

After scaling up the capacity, this feature comes into play when a customer utilizes more CPU resources than purchased, limiting resource consumption and the performance of data warehousing workloads. Throttling is implemented to strike a balance between the performance and reliability of these workloads, preventing failures or degradation caused by resource exhaustion.

Burstable capacity, smoothing, and throttling in Microsoft Fabric provide a dynamic and scalable solution for SMEs, leveraging **the cloud-based platform's advantages:**



Pay-as-you-go

Pay only for utilized resources, cutting costs with burstable capacity and smoothing for increased flexibility.



Elasticity

Scale resources dynamically to match data analysis demands and ensure optimal flexibility without altering capacity size or configuration.



Performance Excellence

Enhance data processing performance with burstable capacity and smoothing, enabling faster, smoother execution of complex queries and preventing performance issues or failures from throttling.

Aligning Microsoft Fabric's Warehouse Capacity with Data Engineering Principles

The innovative trio of burstable capacity, smoothing, and throttling in Microsoft Fabric presents data warehousing solutions that seamlessly align with key principles and best practices of data engineering:



Data-Driven Approach

Encouraging SMEs to base decisions on data rather than assumptions, burstable capacity and smoothing drive accurate measurement and monitoring of data warehousing performance. This ensures optimized resource allocation, scaling up and down based on high- or low-demand phases.



Agility

Breaking free from static infrastructures, burstable capacity and smoothing enable SMEs to adapt to evolving business and user needs. This agility ensures the dynamic scaling of resources in response to unpredictable data analysis demands, all without needing capacity size or configuration alterations.



Quality Assurance

Emphasizing the importance of data quality and reliability, burstable capacity and smoothing mechanisms ensure consistent performance, including running complex queries efficiently, maintaining responsiveness, avoiding throttling or failures attributed to capacity limits, and safeguarding data analysis accuracy.

Use Cases

Harnessing Microsoft Fabric's Dynamic Performance Optimization

Consider a scenario where a data engineer manages a data warehouse supporting a business intelligence application that provides reports and dashboards to the sales and marketing teams. The organization has purchased an F16 capacity, which provides 16 CUs as the baseline capacity. The data warehouse has a daily load of 200 GB of data from various sources, and the data analysis queries vary in complexity and concurrency. The data engineer uses the Warehouse and SQL analytics endpoint to create and manage the data warehouse in Microsoft Fabric.

To optimize performance, the data engineer strategically deploys burstable capacity, smoothing, and throttling across different scenarios:

Smoothing During Off-Peak Hours

During reduced-demand periods, smoothing is employed to mitigate the impact of short-term demand spikes. For instance, when a batch job consumes **1600 CUs** in 10 minutes, smoothing over 24 hours ensures a consistent consumption rate of **66.7 CUs** per hour. This prevents exceeding the baseline of **16 CUs** at any given moment, preventing throttling or failures due to capacity limits.

Burstable Capacity in Peak Hours

In high-demand periods like midday or afternoon, Burstable capacity is activated to surpass the baseline and enhance performance. For a complex query requiring **24 CUs** for optimal execution, the data engineer utilizes Burstable capacity to employ **24 CUs**, improving response time and throughput. This enables meeting heightened demand for data analysis, delivering faster and smoother reports and dashboards.

Throttling During Very Peak Hours

During extremely demanding periods, such as flash sales or holiday seasons, throttling is implemented to restrict resource consumption and optimize performance. If a highly complex query necessitates **32 CUs** for optimal execution, the Fabric initiates throttling, limiting the query to use **16 CUs**. This action balances the data warehouse's performance and reliability, preventing failures or degradation due to resource exhaustion.

Decoding Microsoft Fabric: Advancing Data Warehousing Efficiency with Expert Insights

In data warehousing, integrating burstable capacity, smoothing, and throttling within Microsoft Fabric's Warehouse and SQL analytics endpoint presents a dynamic framework for optimizing resource utilization. These concepts are tied to the assigned capacity size or SKU, acting as a baseline for compute resources in the Microsoft Fabric workspace.

A crucial element in this framework is using a scale factor to calculate resources for data warehousing workloads. This scale factor, determined by the ratio of capacity units consumed to the workload's duration concerning baseline capacity units, ranges from 1x to 32x based on the capacity size. This scalable approach forms the foundation for cost-saving strategies and performance improvements.

In summary, Microsoft Fabric's dynamic trio of burstable capacity, smoothing, and throttling offers flexible and scalable solutions for data warehousing. Beyond enabling enhanced performance, these features align seamlessly with data engineering principles, embracing data-driven approaches and best practices for optimal resource management.

For instance, by employing an F16 capacity with a baseline of 16 CUs, a data engineer can leverage burstable capacity and smoothing effectively. The burstable capacity allows the utilization of up to 192 CUs (12x scale factor) for intricate queries while smoothing spreads this consumption over 24 hours. This strategic approach enables optimal performance at a cost equivalent to the baseline, exemplifying the efficiency gained through Microsoft Fabric.

Additionally, throttling becomes a pivotal tool in maintaining performance balance and reliability. By limiting resource consumption to the baseline capacity, a data engineer can prevent failures or degradation caused by resource exhaustion during periods of high demand.

About the Author



Raghavendra Koyalkondabanda

Principal-Data Engineering; Data & Analytics

Raghavendra is an expert in technology consulting leadership and has vast experience in Analytics. He has over two decades of experience in Data Engineering Solutions, Enterprise Data Warehousing, and Modernizing Big Data Solutions. He has extensive experience designing and implementing data analytics solutions for Fortune 500 clients across various industries. Raghavendra specializes in the Azure ecosystem's Data Lakehouse, Cloud Data Engineering, Data Quality, Business Intelligence, and Analytics. He is passionate about exploring new technologies and conducting research on Data Engineering, Data Science, and AI&ML.

Reference

- ▶ <https://learn.microsoft.com/en-us/fabric/data-warehouse/burstable-capacity>
- ▶ <https://learn.microsoft.com/en-us/fabric/data-warehouse/compute-capacity-smoothing-throttling>

LTIMindtree is a global technology consulting and digital solutions company that enables enterprises across industries to reimagine business models, accelerate innovation, and maximize growth by harnessing digital technologies. As a digital transformation partner to more than 700 clients, LTIMindtree brings extensive domain and technology expertise to help drive superior competitive differentiation, customer experiences, and business outcomes in a converging world. Powered by 81,000+ talented and entrepreneurial professionals across more than 30 countries, LTIMindtree — a Larsen & Toubro Group company — combines the industry-acclaimed strengths of erstwhile Larsen and Toubro Infotech and Mindtree in solving the most complex business challenges and delivering transformation at scale. For more information, please visit **www.ltimindtree.com**