

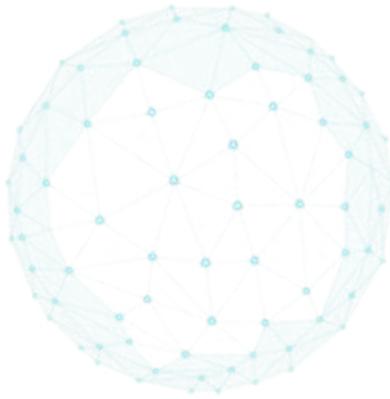
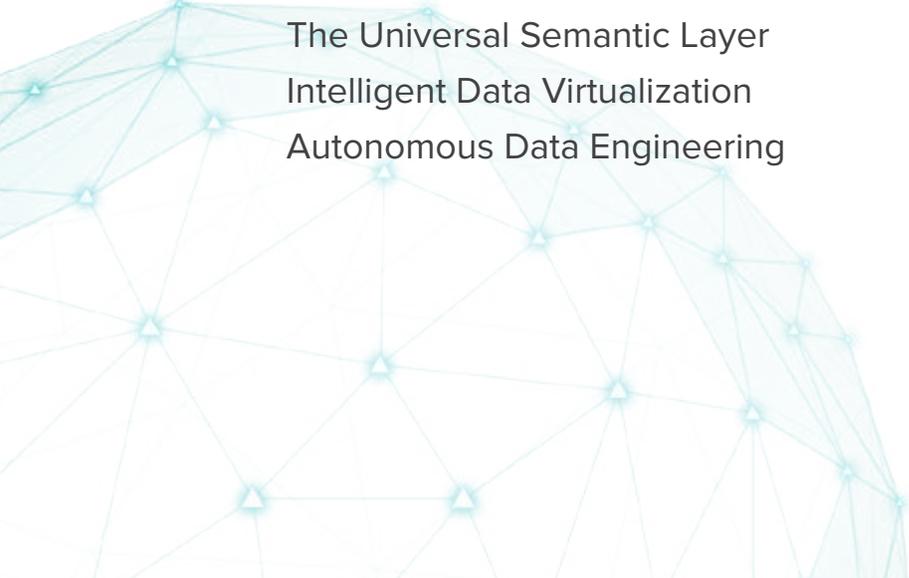
OLAP in the Cloud: Leaving SSAS for a Powerful Path Forward





TABLE OF CONTENTS

Introduction: Why You need to Solve Your OLAP Problem	1
Forecasting	1
Revenue Growth Management	1
Sales and Operations	1
Inventory Planning	2
The Current State of SSAS	3
Performance	3
Multidimensional Analysis	3
Support of Existing Technology	3
The Rise of Agile Analytics Architecture and the Limitations of SSAS	4
Need for Agility	4
Need to Support Extreme Data Volatility	4
Need for Cloud Support	5
Need for Centralized Data Governance	5
SSAS Functionality for Modern Data Cloud Architectures	7
OLAP in the Cloud	7
The Universal Semantic Layer	9
Intelligent Data Virtualization	9
Autonomous Data Engineering	10





Introduction: Why You need to Solve Your OLAP Problem

Data has tremendous value and it is everywhere. This very fact is driving the modern business to harness it in order to gain actionable insights and boost its bottom line. Data can be collected and utilized from a wide range of sources today. Understandably, the digital footprint left by online activity and events has reached enormous proportions, raising the need for capable and scalable solutions.

Online Analytical Processing (OLAP) is driving the new wave of big data across all sectors and industries today. While SQL Server Analysis Services (SSAS) has been effective in allowing for OLAP over the last decade, today's advanced data ecosystems require something more comprehensive for optimal results.

Traditional OLAP solutions as we know them have some severe limitations that are not allowing organizations to achieve their business goals. These include:

Forecasting

Predictive analytics is transforming the way businesses are approaching their infrastructure and budget planning. The ability to forecast performance and other metrics is made possible by breaking down historical data and blending it with machine learning, statistical modelling, and big data analytics.

This functionality goes far beyond specific sectors or geographical locations. As per a recent [Zion Market research report](#), the global market for predictive analytics is expected to cross the \$20 billion mark by 2022. This information can prove to be vital when it comes to optimizing resources and time consumption.

Revenue Growth Management

Today's dynamic markets require a proactive approach to gain scaling up (and down) capabilities. Revenue growth management can be optimized only when the business can use aggregate data to adjust marketing strategies and procedures, while also considering future business opportunities.

Sales and Operations

Sales is another department where reporting analytics is key to optimization and performance. Once implemented correctly, this data can be used to streamline sales cycles, optimize marketing campaigns, and target the right audiences for better conversion and retention metrics, while also optimizing nurturing and expansion opportunities.



The same applies for Operations teams, which gain access to real-time data to optimize processes and take productivity to a whole new level. Furthermore, having access to real-time reporting analytics can help establish a smooth and transparent operational pipeline in any kind of environment.

Inventory Planning

Supply chain volatility is leading a lot of financial bleeding, brand damage, and in many cases forcing businesses into bankruptcy. Only predictive reporting analytics can help foresee upcoming shortages or price hikes that need to be taken into consideration to maintain a healthy bottom line.

Reporting analytics is essentially transforming the way businesses are operating and managing their operations. These actionable insights are enabling more informed decisions and are shortening time to decision and the ability to respond quickly to global events and unexpected developments.

OLAP provides consistent information and calculations to help achieve the aforementioned business goals. It has indeed taken Business Intelligence (BI) to a whole new level with its multidimensional data representation, and has helped businesses perform efficient data analysis to boost their bottom line.

Did You Know?

Netflix saves over \$1 billion per year on customer retention with the use of big data.

Every person is generating roughly 1.7 megabytes of usable data every second in 2020.

Internet users are generating 2.5 quintillion bytes of data on a daily basis in 2020.



The Current State of SSAS

SQL Server Analysis Services (SSAS) is a multi-dimensional OLAP server as well as an analytics engine that allows businesses to “digest” large volumes of data. It is a part of the Microsoft SQL Server and helps perform analysis by using a wide range of dimensions. SSAS has become extremely popular over the years.

The smooth operation of SSAS hinges upon the correct implementation and utilization of the OLAP Cube. While building the Cube, you will have to choose between two processing orders - Parallel and Sequential. The latter can execute all tasks in one transaction or perform each one in a separate transaction.

Reasons for the popularity of SSAS, besides the well-documented simplicity, no-nonsense nature, and ease-of-use, are many. Here are just a few.

Performance

During cube processing, SSAS pre-calculates and stores aggregates of facts physically. For example, take Turnover by Year and Region aggregates. When the query is fired, SSAS doesn't calculate the outcome via underlying details (unlike T-SQL), but takes the values directly from the stored aggregates.

Furthermore, SSAS stores query results in a cache. It doesn't take more than a few milliseconds to return data, which has contributed to its popularity.

Multidimensional Analysis

SSAS allows you to slice, dice, and drill down to get accurate and insightful results. It must be noted that this hinges upon the identity of the tool or front end that is layered over the data, but in most cases you can navigate around the data to detect trends and spot patterns without too much hassle.

Data exploration is a key feature of cubes. This allows the end user to intuitively “explore” the data, without realising that they are actually analyzing it.

Support of Existing Technology

SQL Server Analysis Services allow you to work with existing technologies. This is a benefit that executives like because it requires less third-party integration. For example, Excel can be used to view data via Pivot Tables. Power BI Desktop can also be used in tandem with SSAS to create dynamic reports.

Furthermore, SSAS can be integrated in SharePoint in just a few minutes with zero programming and no additional tools. It's as simple as that.



The Rise of Agile Analytics Architecture and the Limitations of SSAS

Despite the aforementioned benefits of using the SSAS Cube, there are many challenges that adopters are facing today. To start, data is exploding in size, with the number of sources multiplying exponentially. As a result of these two developments alone, many new issues are surfacing.

Need for Agility

There is no denying the impressive business results that SQL Server Analysis Services can drive. It's definitely possible to generate impressive amounts of user generated analytical content. However, it's no longer possible to ignore the level of effort and attention to detail that goes into the whole process.

With businesses required to scale up (or down) on demand due to today's volatile market conditions, SSAS falls short when it comes to being agile in matching cloud data warehouse's scaling characteristics, or when the data needs to be surfaced at full fidelity (not just subsets and aggregates).

Additionally, the level of effort that can go into data re-formatting can become unbearably high. In some scenarios, it can take more than a month to rebuild the Cube and add a new dimension or metric. So while SSAS hits the mark in query performance and semantic consistency, it failed miserably at scaling.

Just as bad, the traditional architecture model requires an army of highly skilled data engineers to make it work and maintain it. Besides the need to invest in manpower and resources, additional data copies are needed. This eventually adds to costs and raises the risk and potential for discrepancies.

Need to Support Extreme Data Volatility

So you deployed your Multidimensional OLAP (MOLAP) and are getting good results. But it won't be long before your business requirements will change and so will the nature of the data you are collecting. The traditional SSAS Cube will require a lot of work to integrate the new sources and get back on track.

Yes, the MOLAP structure is optimized to maximize query performance and process "raw data". But at the end it isn't too versatile and can handle only pre-specified amounts of data at a time. As mentioned earlier, this simply isn't enough to accommodate the dynamic nature of today's big data space.



Furthermore, detailed data is key to create actionable insights that are accurate and effective. This requires access to atomic database values, despite the growing number of data sources that are piling up at any given moment. The summaries that the traditional SSAS Cube works with are now inadequate.

Need for Cloud Support

Sooner or later, your data will outgrow your traditional SSAS Cubing solution. This is because legacy on-premise architectures are resource heavy and can turn out to be lengthy on large data volumes. Add in other technical limitations like the inability to merge data between cubes and you have your hands full.

Most data warehouses and data lakes are in the cloud today. Traditional solutions are limited (best case scenarios) in their ability to hook up to these sources and fetch the required information. Getting this done requires days, if not weeks, of tedious manual labor, programming, and customization work.

Also, more and more businesses are shifting their entire ecosystems to the cloud, either completely or in some kind of hybrid method. It only makes sense that these massive infrastructure moves should be followed with the moving of SSAS Cubes and business logic to the cloud for optimal performance.

Traditional SSAS is now also being left behind by new query technologies like Apache Parquet, Optimized Row Columnar (ORC), and Avro, with the latter storing data in a row-based format (best for write-heavy transactional workloads). Parquet and ORC are better suited for read-heavy workloads.

These are on-the-wire formats that allow smooth data passing between Hadoop cluster nodes, versatile, flexible, and self-described ways to scale up.

All of the three aforementioned technologies are available for basically any Hadoop ecosystem product, regardless of the data model, processing framework, or programming language. It has become extremely important to have a solution that leverages the versatility of these technologies.

Need for Centralized Data Governance

Bad data is something that is commonly overlooked by businesses, who fail to take the inherited risks of SSAS into consideration. Analytics teams are having a tough time setting up and modifying data pipelines, something that makes it difficult to deliver consistent data streams for optimal results.



With multiple BI tools in place and integration taking up a lot of the team's bandwidth, it has become a challenge to execute smoothly. Receiving multiple answers to the same questions and emerging conflicting definitions has become a common occurrence. This causes damage on multiple levels.

Besides the lack of insights and remediation costs (time, money and resources), productivity is hampered, something that has a direct effect on the bottom line.

“Bad data cost us a lot more than \$15 million a year at Yahoo, and that was almost 10 years ago.”

Dave Marani, CEO, AtScale

The main reason behind these symptoms is the lack of data governance, a problem that is on the rise with multiplying extracts and data copies that are adding to the clutter. With data getting old and out of sync, only a centralized data governance solution can help establish (and maintain) key KPIs.



SSAS Functionality for Modern Data Cloud Architectures

It's safe to say that the traditional SSAS architecture is too rigid, fragile, does not scale well, and is not sustainable for today's dynamic and agile needs. AtScale delivers all of the power and functionality of SSAS that you love (OLAP, MDX connection to Excel) without the limitations you don't (lag times, extracts) - and at the scale you need to run analytics in the Cloud.

OLAP in the Cloud

Online Analytical Processing (OLAP) is still important and remains critical to many analytics teams. Its biggest advantage lies in its ability to quickly access shared multidimensional data or information. It also works on-demand. However there are many steps that need to be taken in order to get there. One of them is the physical moving of data.

AtScale works in the modern analytics stack and goes several steps further than SSAS, allowing the creation of virtual cubes that are more dynamic and efficient. Furthermore, AtScale doesn't require the manual uploading or migration of data to start working. Once you have configured and published the cube, you can start querying it immediately.

Cloud OLAP also doesn't require much maintenance as no re-tooling or integration is needed. Just make sure that your BI or ML tools are cloud-compatible and you are good to go. Needless to say, these advantages come into play especially for businesses that are experiencing hyper-growth with lots of data to work with.

It's also important to mention that all compatibility with Excel and other MDX and SQL speaking BI tools is native to AtScale.

In order to support Wayfair's fast-growing business, the Data Infrastructure team at Wayfair realized that they needed to shift away from their on-premises infrastructure and towards a cloud infrastructure. This presented a number of challenges - infrastructure optimization and migration logistics to name a few. [Read More](#)



“We wanted to move our fast-growing business to the cloud but didn’t want to lose the capabilities we had in our on-premise environment. AtScale helped us do that and was a major part of driving cloud adoption for our customers and data producers across our organization.”

Matt Hartwig, Associate Director, Data Infra Team, Wayfair



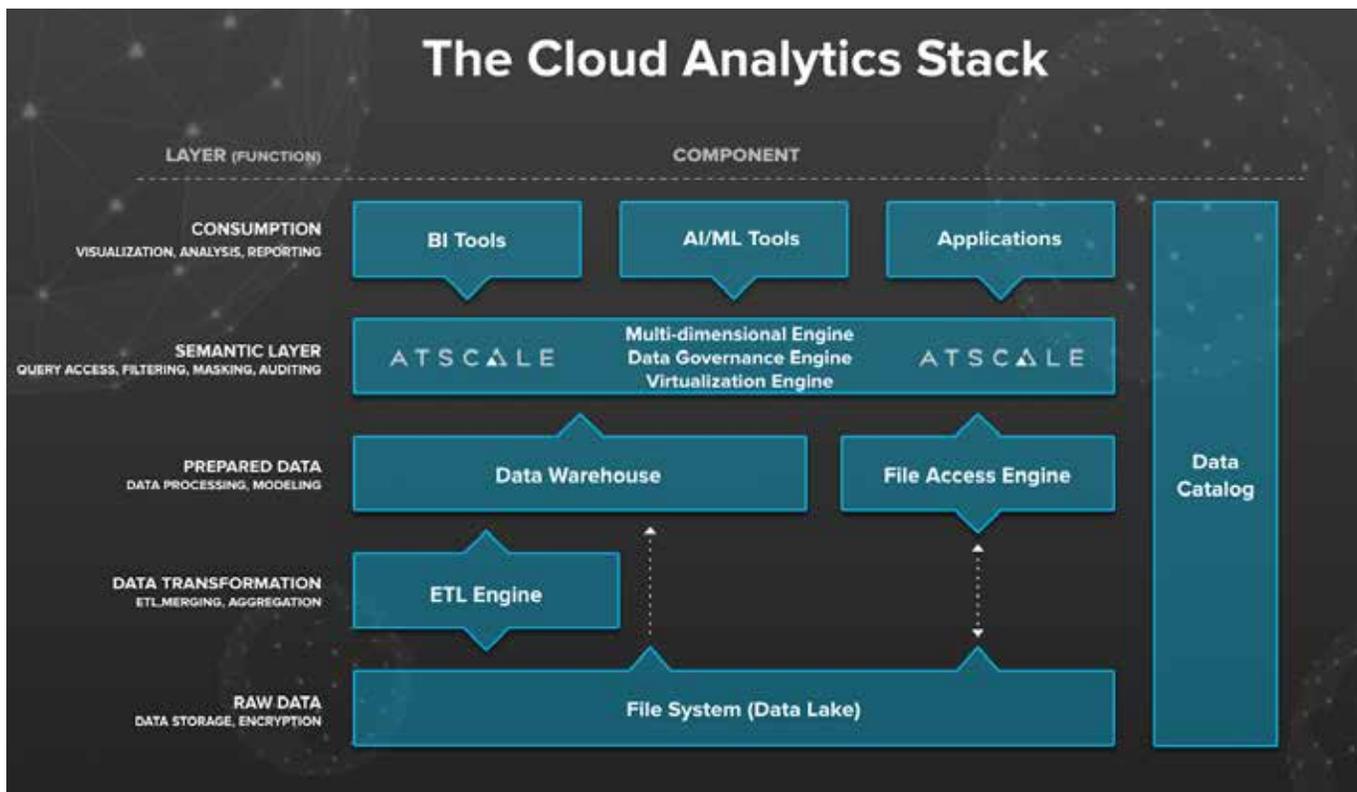
The Universal Semantic Layer™

In a nutshell, a universal semantic layer is a representation of corporate data that can be accessed by using common business terms. The semantic layer essentially categorizes complex technical data into easy and commonly used business terms. The main purpose of this layer is to simplify data utilization.

The semantic layer also promoted live connections, while not requiring data extracts to minimize SSAS Cube building and ETL chores, that are very time and resource consuming. This essentially breaks the ever-existing wall between the data and the end users who need to access it in real-time with minimal fuss.

There is also less pressure on the IT department since there is no client-side software to install or maintain. It’s easy to migrate users, without changing the MDX code of the BI tools of their choice (i.e - Excel, Power BI, Tableau, etc.). The move is fast, smooth, and doesn’t require any training or onboarding.

This type of service is basically a personal data warehouse for BI users of SQL tools that works as a dynamic OLAP cube for MDX-speakers.



Intelligent Data Virtualization™

OLAP in the Cloud opens up a whole new world of opportunities and new data avenues to explore, all in a user-friendly and straightforward manner. End users can now analyze data from multiple data lakes and data storehouses, at blazing fast speeds that were simply not available before.

It's important to understand that even moving to the cloud can become a costly affair if too many servers are being used or if your organization is running malformed queries that can pile up the bills if you are on a consumption-based pricing plan (or exhaust your resources if you are on a fixed one).

As mentioned earlier, intelligent data virtualization automates the sourcing, curation, and modeling of data from multiple data sources. This introduces more flexibility while adopting new platforms without disrupting their downstream data consumers or re-engineering their stack.

You can now understand and orchestrate complex data, without the technical and logical limitations of old on-premise SSAP solutions.



Autonomous Data Engineering™

Deciding on the right sources takes time, and so does collecting live data from them. Data engineers working with old SSAS solutions are wasting too much time preparing and moving data. Today's dynamic and evolving business model requires self-service data access via a centralized dashboard.

Autonomous data engineering can identify query patterns to create and manage intelligent aggregates, essentially mimicking data engineers. Besides the time and resource savings, autonomous solutions learn from user behavior and data relationships to speeden up the delivery of insightful insights.

With aggregates being built in real-time in response to user activity and queries being tweaked without human intervention, the data engineers can finally start focusing on projects to develop and enhance the BI posture of their organizations and prepare for the ever-evolving data landscape.

Less data transfer, regardless of the infrastructure you are using, has a positive effect on costs. While doing it in the cloud, predictability becomes extremely important. But even the biggest of migrations don't need to scare you if you are using a proven and tested "OLAP in the Cloud" solution.

Slowly but surely, the world is gravitating towards self-service BI in the Cloud. The reason is simple - the quality of your data is at risk due to mistakes being made with on-prem data storage, manpower limitations, time constraints, and lack of scalability. However, it is important to make the move properly to make sure that there are no down times or technical bottlenecks.

ABOUT ATSCALE

AtScale powers the analysis used by the Global 2000 to make million dollar business decisions. The company's Intelligent Data Virtualization™ platform provides Cloud OLAP, Autonomous Data Engineering™ and a Universal Semantic Layer™ for fast, accurate data-driven business intelligence and machine learning analysis at scale.

For more information, visit www.atscale.com.