GIGAOM

High-Performance Cloud Data Warehouse Performance Testing `v1.0`

Product Evaluation: Actian Avalanche, Amazon Redshift, Microsoft Azure Synapse, Google BigQuery, and Snowflake

WILLIAM MCKNIGHT AND JAKE DOLEZAL | OCT 13, 2020 - 11:29 AM CDT

TOPICS: **CLOUD FILE STORAGE** **CLOUD INFRASTRUCTURE** **DATA WAREHOUSE**

CREDIT: PETE LINFORTH

SPONSORED BY ACTIAN

High-Performance Cloud Data Warehouse Performance Testing

Product Evaluation: Actian Avalanche, Amazon Redshift, Microsoft Azure Synapse, Google BigQuery, and Snowflake

TABLE OF CONTENTS

1. Summary

Data-driven organizations rely on analytic databases to load, store, and analyze volumes of data at high speed to derive timely insights. Data volumes within modern organizations' information ecosystems are rapidly expanding—placing significant performance demands on legacy architectures. Today, to fully harness their data to gain competitive advantage, businesses need modern scalable architectures and high levels of performance and reliability to provide timely analytical insights. At the same time, many companies are migrating to fully managed cloud services. With these managed as-a-service deployment models, companies can leverage powerful data platforms without the technical debt and burden of finding talent to manage the resources and architecture in house. These models enable users to pay as they play and to stand up a fully functional analytical platform in the cloud with just a few clicks.

This report outlines the results from a GigaOm Analytic Field Test derived from the industry standard TPC Benchmark™ H (TPC-H)[1] to compare Actian Avalanche Cloud Data Warehouse, Amazon Redshift, Microsoft Azure Synapse, Google BigQuery, and Snowflake Data Warehouse. This test produced interesting results that reveal some of the performance characteristics of the five platforms.

Overall, the benchmark results were insightful in revealing query execution performance and some of the differentiators for Avalanche, Synapse, Snowflake, Amazon Redshift, and Google BigQuery. In our testing, Avalanche query response times on the 30TB TPC-H data set were overall 8.5 times faster than Snowflake in a test of 5 concurrent users. Furthermore, Avalanche was approximately 1.3 times faster than Redshift, 1.5 times faster than Synapse, and 7.9 times faster than BigQuery.

Examining price-performance, Avalanche ran the 5 concurrent user TPC-H queries roughly 6.4 times more cost effectively than Snowflake, as measured in cost per query per hour. Avalanche further proved 1.4 times better than Redshift, 1.3 times better than Synapse, and 12.4 times better than Google BigQuery in terms of cost per query per hour across the examined cluster classes.

---

1 - More can be learned about the TPC-H benchmark at http://www.tpc.org/tpch/

2. Platform Summary

Big data analytics platforms load, store, and analyze volumes of data at high speed, providing timely insights to businesses. Data-driven organizations leverage this data, for example, for advanced analysis to market new promotions, operational analytics to drive efficiency, or for predictive analytics to evaluate credit risk and detect fraud. Customers are leveraging a mix of relational analytical databases and data warehouses to gain analytic insights.

This report focuses on relational analytical databases in the cloud because deployments are at an all-time high and poised to expand dramatically. The cloud enables enterprises to differentiate and innovate with these database systems at a much more rapid pace than was ever possible before. The cloud is a disruptive technology, offering elastic scalability vis-à-vis on-premises deployments, enabling faster server deployment and application development, and allowing less costly storage. For these reasons and others, many companies have leveraged the cloud to maintain or gain competitive momentum.

This report compares Actian Avalanche, Amazon Redshift, Azure Synapse, Google BigQuery, and Snowflake—five relational analytical databases based on scale-out cloud data warehouses and columnar-based database architectures. Despite these similarities, there are some distinct differences in the five platforms.

Actian Avalanche

Actian Avalanche is the newest entrant to the cloud data warehouse world, offering a fully managed high-performance data warehouse, which was launched in early 2019. It is based on underlying technology, known as Vector (first released in 2010), which is an efficient implementation of modern analytical database concepts with high performance as the design point. The Actian patented Vector database engine utilizes a "cache-based execution model" along with vectorized processing of data. Avalanche performs "single instruction, multiple data" processes by leveraging the same operation on multiple data simultaneously and exploiting the parallelization capabilities of modern hardware. This innovation reduces overhead found in conventional "one-record-at-a-time processing" platforms. Additionally, the compressed column-oriented format uses a scan-optimized buffer manager. It is also the only ANSI-2011 SQL compliant warehouse among the ones tested in this report, which would imply an easier migration from other ANSI SQL compliant databases.

Actian Avalanche is offered as a fully managed service, which implies no administrative overhead for organizations. As of this writing, it is available on AWS, Azure, and Google Cloud, as well as on-prem.

The measure of Actian Avalanche compute power is known as Avalanche Units (AU), which you pay for on an hourly basis when the compute resources are being used. You can spin up warehouses of varying sizes from 4 AU to 128AU. Warehouse sizes can be scaled up and down as business needs change. Once the compute warehouse goes inactive, you no longer pay. At the time of this writing, Avalanche is priced at $1.99 per AU per hour. Storage cost is separate.

Amazon Redshift

Amazon Web Services Redshift was the first managed data warehouse cloud service and continues to get a high level of mindshare in this category. It indeed ticks all the table stakes boxes for a cloud analytic database. Amazon Redshift is a fit for organizations needing a data warehouse with little to no administrative overhead and a clear, consistent pricing model. Amazon Web Services supports most of the databases in this report and then some.

We tested the latest RA3 Redshift engine, which introduced a new managed storage layer that is an upgrade from the tighter coupled storage on the older DS2 and DC2 instance types. For Redshift, we paid an hourly rate for when the cluster was running, but it also has a pause feature to stop billing.

Azure Synapse Analytics

On Azure Synapse Analytics, formerly known as Azure SQL Data Warehouse, storage is separate from the compute Data Warehouse Unit (DWU). This enables Azure Synapse to scale columnar storage capacity and compute resources independently. This capability adjusts to various workload demands, offering potential cost savings when demand is low. Synapse can pause and resume compute billing, in which only storage is billed during the paused time. Synapse achieves good balance in both configurability and simplicity, in a way that is both easy to administer and flexible in handling almost any usage pattern.

With Synapse, you can scale the compute DWU on the fly. We paid an hourly rate for the time when our cluster was active, but there is also a separate data storage charge for the SQL database underneath the Synapse engine.

Google BigQuery

Google BigQuery is a managed service with some interesting distinctions. Google abstracts the details of the underlying hardware, database, and all configurations. BigQuery is a hands-off database without indexes or column constraints. Defragmentation and system tuning are not required. A serverless solution, Google Cloud manages the servers in a fully hands-off manner for the customer, dynamically allocating storage and compute resources. The customer does not define nodes or the capacity of the BigQuery instance. The provisioning of compute is particularly fast and seamless.

With BigQuery you pay for the amount of data you query and store. At the time of this writing, customers can pre-purchase flat-rate computation "slots" or units in increments of $10,000 per month per 500 compute units. However, Google recently introduced Flex Slots, which allow slot reservations as short as one minute that are billed by the hour. There is a separate charge for active storage of data.

Snowflake

As a cloud-only, fully managed solution, Snowflake has a clear separation between compute and

storage. For Snowflake on AWS, which is what we used for the queries, data is stored in AWS S3 and cached when queries are executed to bring the data in closer proximity to compute resources. Snowflake essentially offers two configuration "levers" — the size of the warehouse cluster and the number of clusters permitted to spin up to handle concurrency. Snowflake scales by cluster server count in powers of two (i.e., 1, 2, 4, 8, 16, and so on). If enabled, Snowflake will spin up additional clusters to handle multi-user concurrent query workloads. Snowflake would automatically spin down the additional clusters once demand has passed. If not enabled, it will place paused queries in a queue until resources free up.

For Snowflake, you pay an hourly rate for credits used, which is based on the size of the clusters and the number of clusters active when compute resources are being used. We paid $3.00 per credit-hour for the Enterprise tier. Once the compute warehouse goes inactive, you no longer pay. However, there is a separate charge for data storage.

Table 1. Platform Summary

	First Released	Version Tested	SQL Type
Actian Avalanche	2010 (as VectorWise)		ANSI-2011 Compliant
Snowflake	2014	4.27.5	Snowflake SQL
Amazon Redshift	2014	1.0.18004	PostgreSQL 8
Azure Synapse	2016 (Azure SQL Data Warehouse)*	10.0.15218.0	Transact-SQL
Google BigQuery	2010	Sept. 11, 2020 release	Standard SQL

Source GigaOm 2020   * SQL Server released 1989

3. Test Setup

The setup for this Field Test was informed by the TPC Benchmark™ H (TPC-H)[2] spec validation queries. This is not an official TPC benchmark. The queries were executed using the setup, environment, standards, and configurations described here. More can be learned about the TPC-H benchmark at http://www.tpc.org/tpch/

Benchmark Data

The data sets used in the benchmark were a workload derived from the well-recognized industry standard TPC Benchmark™ H (TPC-H). The tpc.org website describes the TPC-H benchmark thusly:

"The TPC-H is a decision support benchmark. It consists of a suite of business-oriented ad-hoc queries and concurrent data modifications. The queries and the data populating the database have been chosen to have broad industry-wide relevance. This benchmark illustrates decision support systems that examine large volumes of data, execute queries with a high degree of complexity, and give answers to critical business questions."

To show the data model, the diagram in **Figure 1** was taken from page 13 of the TPC-H Revision 2.17.3 specification document.
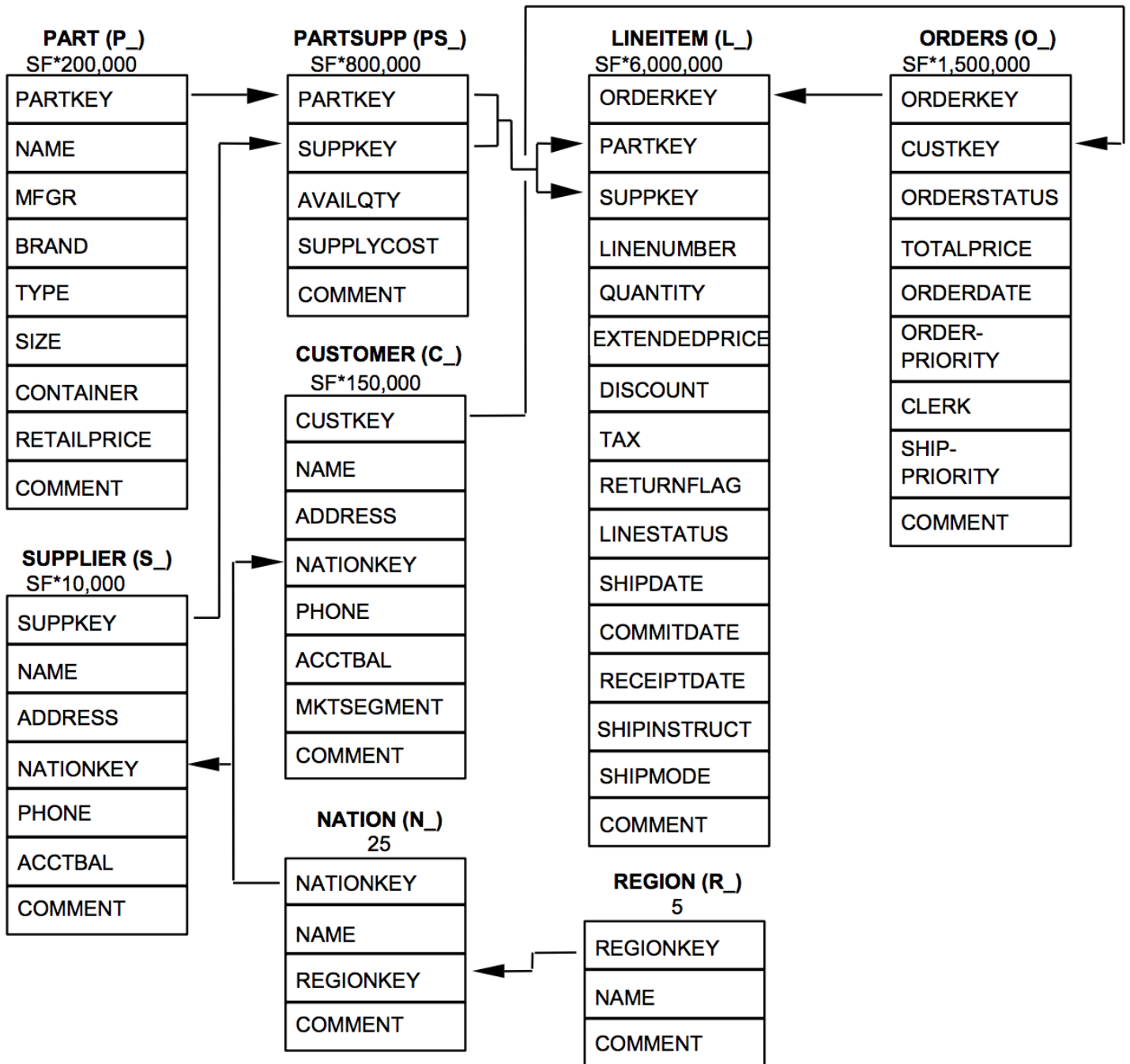
Figure 1. TPC-H Data Model

To provide an idea of the data volumes used in our benchmark, **Table 2** shows row counts of the database when loaded with 30TB of TPC-H data:

Table 2. TPC-H Database Row Count given 30TB

TPC-H Table	30TB Row Count
Customer	4,500,000,000
Line Item	180,000,000,000
Orders	45,000,000,000
Part	6,000,000,000
Supplier	300,000,000
Part Supp	24,000,000,000

Source GigaOm 2020

Cluster Environments

Our benchmark included the cluster environments detailed in **Table 3**:

Table 3. Cluster Environments

TPC-H 30TB	Avalanche	Snowflake	Redshift	Synapse	BigQuery
Cluster/Instance	N/A	3XLarge	ra3.16xlarge	DW15000c	N/A
Units	128 AUs	64 nodes	20 nodes	15,000 DWUs	5,000 Slots

Source GigaOm 2020

Queries

We sought to replicate the TPC-H Benchmark queries modified only by syntax differences required by the platforms. The benchmark is a fair representation of enterprise query needs. The TPC-H testing suite has 22 queries, which are described in **Table 4**.

GIGAOM

Table 4. TPC-H Query Parameters

QUERY #	DESCRIPTION	SUM	SUB-QUERY	JOIN	MIN/MAX	AVG	COUNT	TOP/LIMIT
1	Pricing Summary Report	✔				✔	✔	
2	Minimum Cost Supplier		✔	✔	✔			✔
3	Shipping Priority	✔		✔				✔
4	Order Priority Checking		✔	✔			✔	
5	Local Supplier Volume	✔		✔				
6	Forecasting Revenue Change	✔						
7	Volume Shipping	✔	✔	✔				
8	National Market Share	✔	✔	✔				
9	Product Type Profit Measure	✔	✔	✔				
10	Returned Item Reporting	✔		✔				✔
11	Important Stock Identification	✔	✔					
12	Shipping Modes and Order Priority	✔		✔				
13	Customer Distribution		✔	✔			✔	
14	Promotion Effect	✔		✔				
15	Top Supplier		✔	✔	✔			
16	Parts/Supplier Relationship		✔	✔			✔	
17	Small Quantity Order Revenue	✔	✔	✔		✔		
18	Large Volume Customer	✔	✔	✔				✔
19	Discounted Revenue	✔		✔				
20	Potential Part Promotion	✔	✔	✔				
21	Suppliers Who Kept Orders Waiting		✔	✔			✔	✔
22	Global Sales Opportunity	✔	✔	✔		✔	✔	

Source GigaOm 2020

4. Test Results

This section analyzes the query results from the fastest runs of the three sets of 22 TPC-H queries described in **Table 3**.

Tests with 1 User (No Concurrency)

TPC-H Query 1: "Pricing Summary Report" Execution Times

Execution Time in Seconds
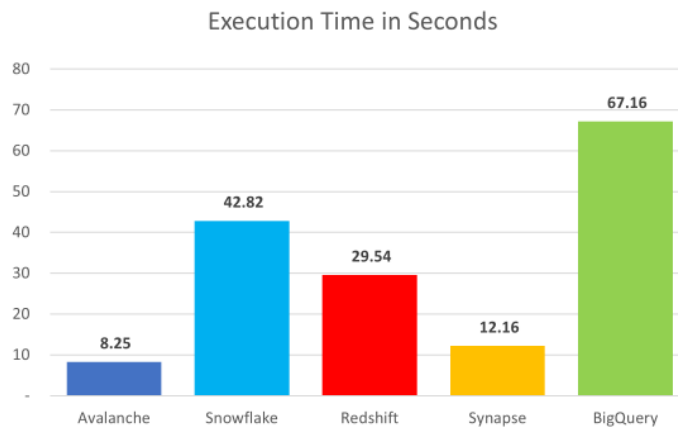
Query 1 is the only query that uses only SUM, AVERAGE, and COUNT. Avalanche was five times faster than Snowflake, while the closest competitor, Synapse, was 30% slower than Avalanche.

TPC-H Query 2: "Minimum Cost Supplier" Execution Times

Execution Time in Seconds
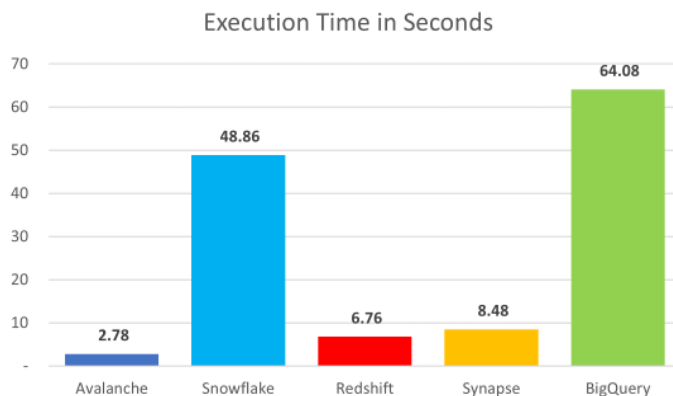
Query 2 was one of two queries that contained a MIN/MAX. Here Avalanche was the fastest by almost 2.5x over the second fastest challenger, Redshift.

Query 3: "Shipping Priority" Execution Times
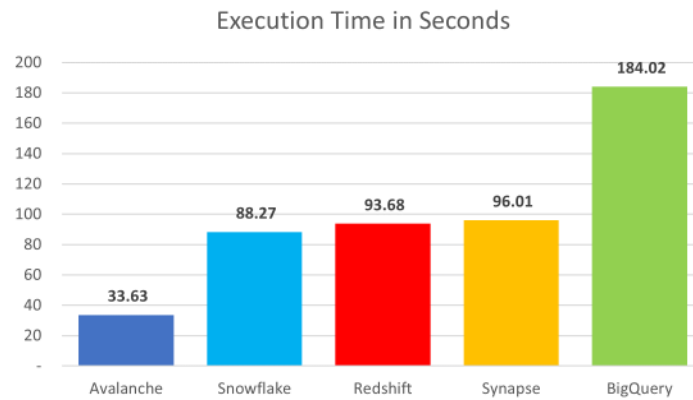
Avalanche was more than twice as fast as the tight bunching of Redshift, Snowflake and Synapse.

Query 4: "Order Priority Checking" Execution Times
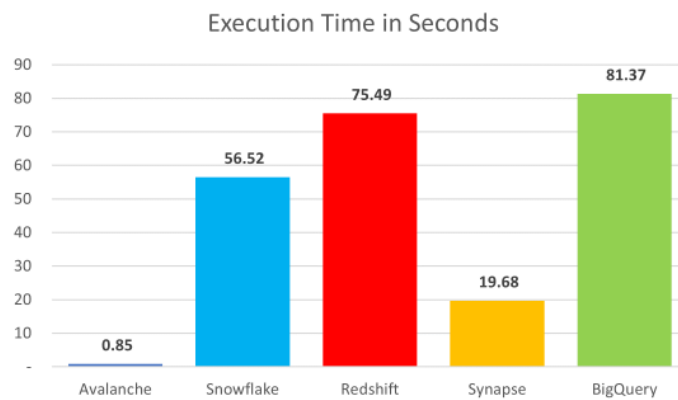
In Query 4, with a SUBQUERY and COUNT, we see the familiar pattern of Actian Avalanche outperforming the field, in this case executing more than 20 times faster than second place Synapse.

Query 5: "Local Supplier Volume" Execution Times

Execution Time in Seconds
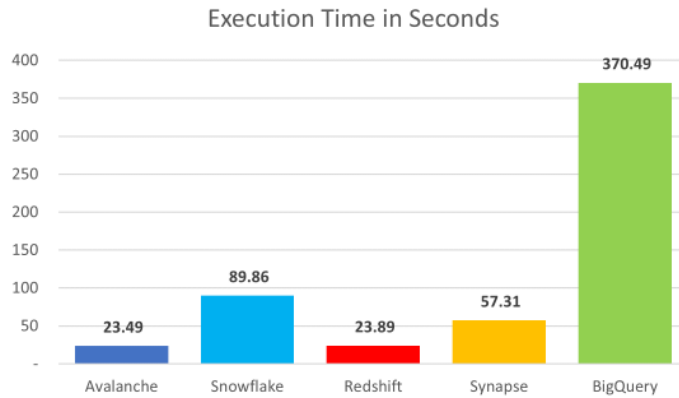
Query 5, which only employs a SUM aggregation, favored Avalanche narrowly over Redshift and by broader margins over Synapse, Snowflake, and BigQuery.

Query 6: "Forecasting Revenue Change" Execution Times

Execution Time in Seconds
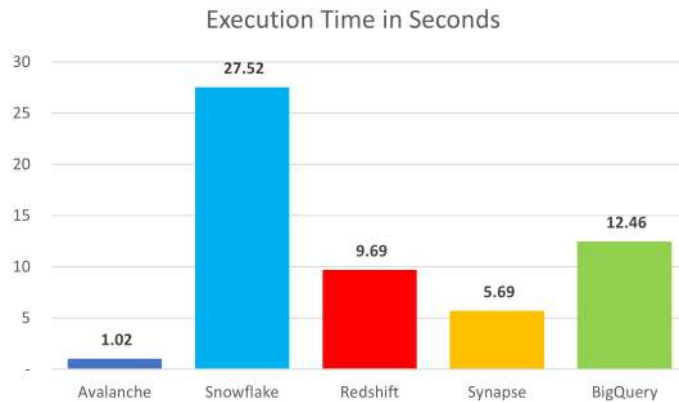
The simple SUM of Query 6 proved a very high relative performer for Avalanche, with a 5x performance margin over second-best Synapse and a 25x advantage over Snowflake.

Query 7: "Volume Shipping" Execution Times

Execution Time in Seconds

Avalanche	Snowflake	Redshift	Synapse	BigQuery
22.53	92.69	41.30	59.70	372.07
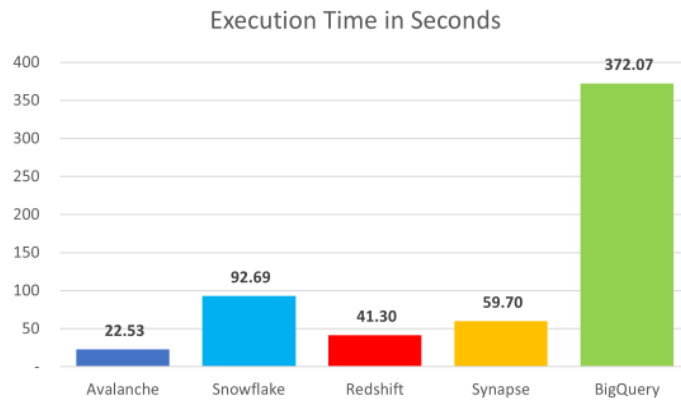
Again, we see the familiar pattern of an approximately 200% performance margin between Avalanche and the second-place option, in this case Redshift.

Query 8: "National Market Share" Execution Times

Execution Time in Seconds

Avalanche	Snowflake	Redshift	Synapse	BigQuery
7.51	102.50	39.31	45.79	336.13
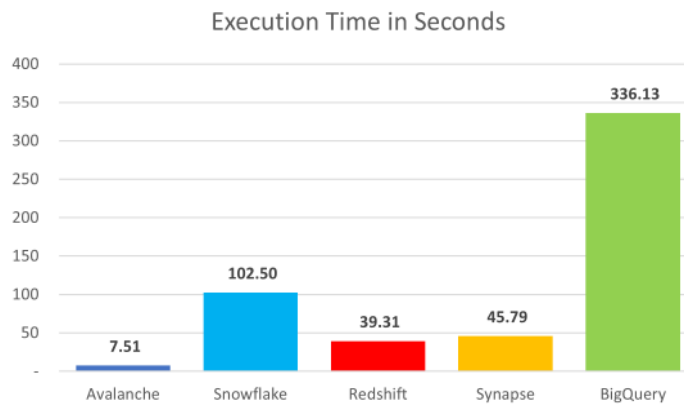
Query 8 performance shows an order of finish similar to Query 7, though with Avalanche here enjoying even larger margins than in the previous test.

Query 9: "Product Type Profit Measure" Execution Times

Execution Time in Seconds
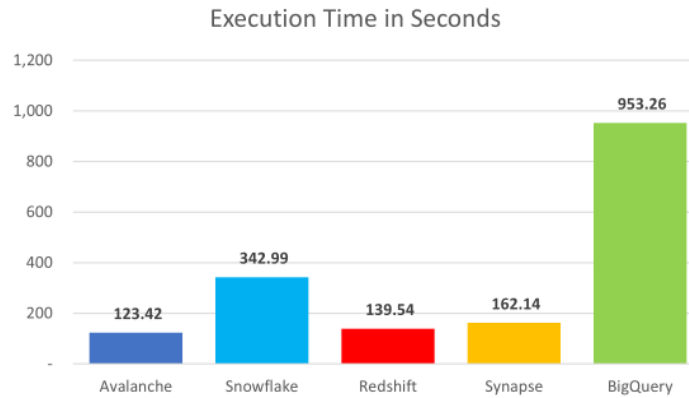
Avalanche narrowly outperformed Redshift and Synapse in Query 9 to lead the field.

Query 10: "Returned Item Reporting" Execution Times

Execution Time in Seconds
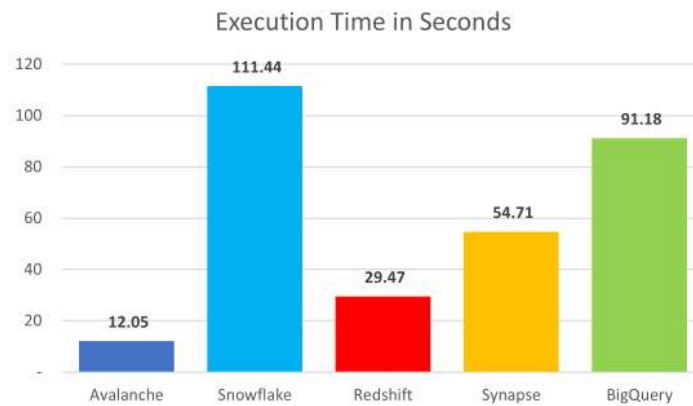
Query 10 uniquely has a SUM and a TOP/LIMIT. Here again Avalanche was the top performer.

Query 11: "Important Stock Identification" Execution Times

Execution Time in Seconds
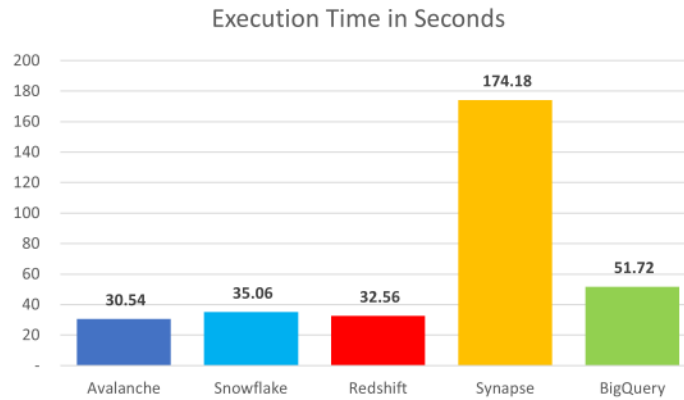
Query 11 is another SUB-SELECT and SUM operation and saw performance slightly favoring Avalanche over Redshift, Snowflake, and BigQuery.

Query 12: "Shipping Modes and Order Priority"

Execution Time in Seconds
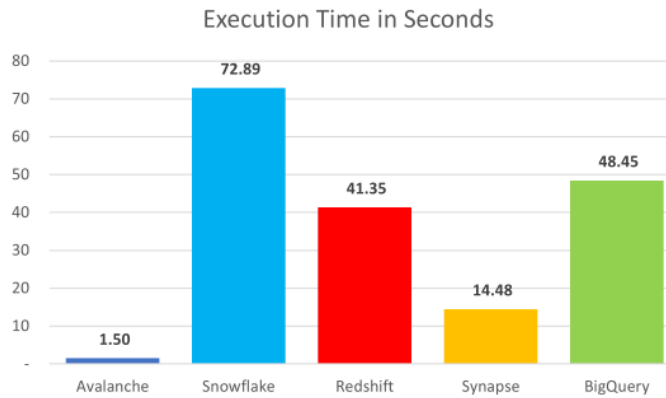
Query 12, another simple SUM, was also another plus performance query for Avalanche, which outperformed the second-fastest finisher Synapse by nearly a factor of 10.

Query 13: "Customer Distribution" Execution Times
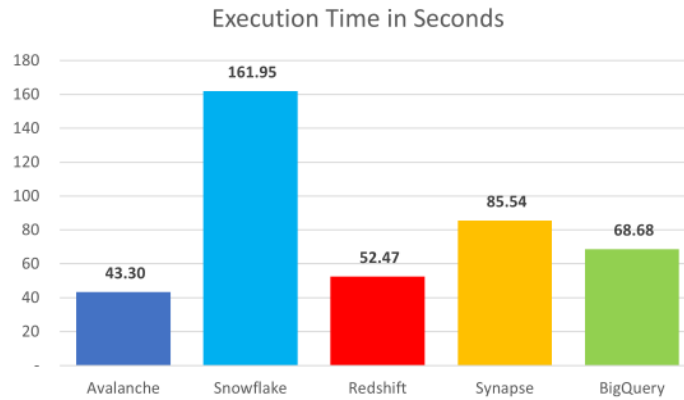
Query 13 is the only TPC-H query with an explicit JOIN. Avalanche outperformed the field, but Redshift was competitive with an execution time of 52.47 seconds.

Query 14: "Promotion Effect" Execution Times
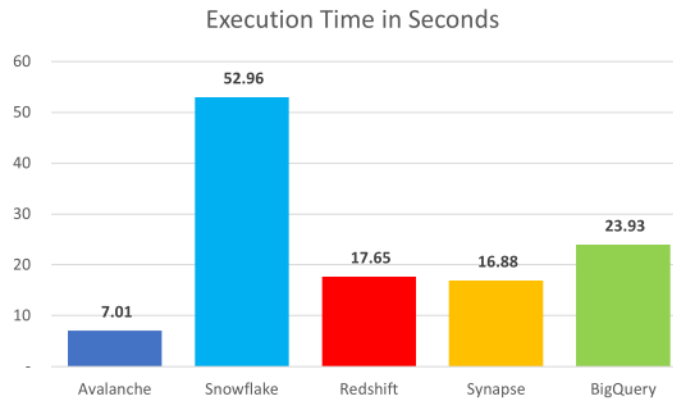
For Query 14, Actian Avalanche executed the test in an impressive 7.01 seconds, well ahead of its competition.

Query 15: "Top Supplier" Execution Times

Execution Time in Seconds
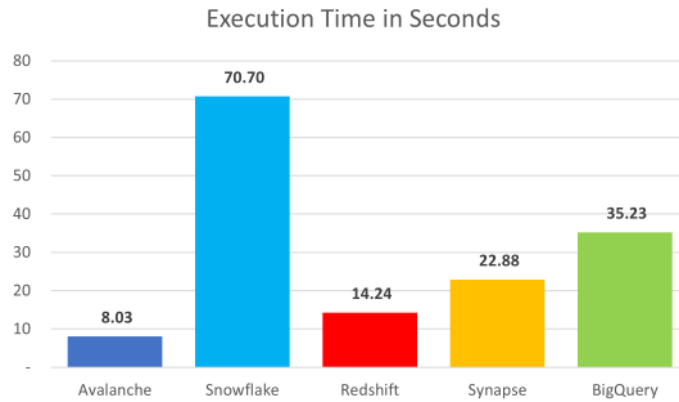
Avalanche dominated this query, outperforming second-place Redshift by 6 seconds.

Query 16: "Parts/ Supplier Relationship" Execution Times

Execution Time in Seconds
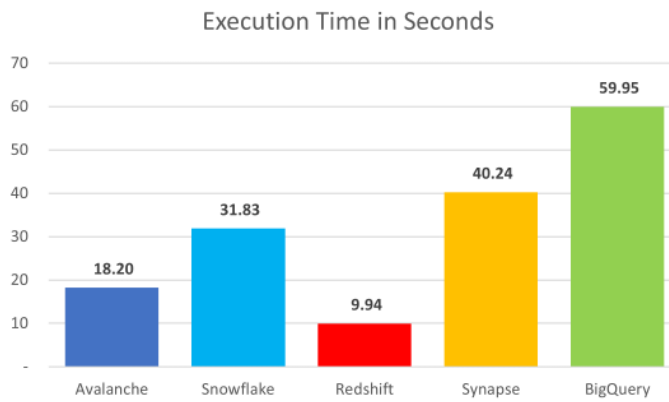
Redshift was nearly twice as fast as second-place Avalanche in Query 16.

Query 17: "Small Quantity Order Revenue" Execution Times

Execution Time in Seconds
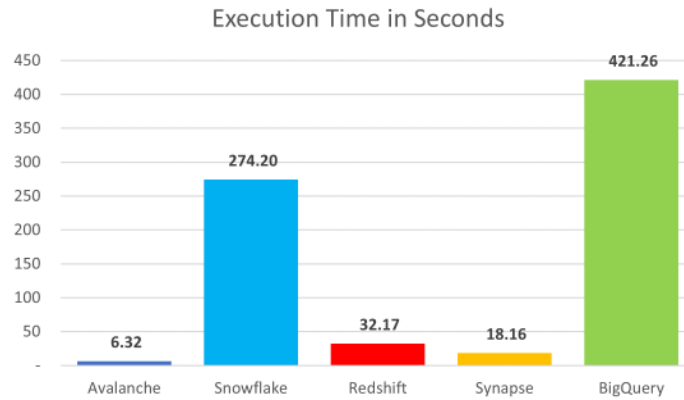
Query 17 yielded a substantial performance advantage for Avalanche.

Query 18: "Large Volume Customer" Execution Times

Execution Time in Seconds
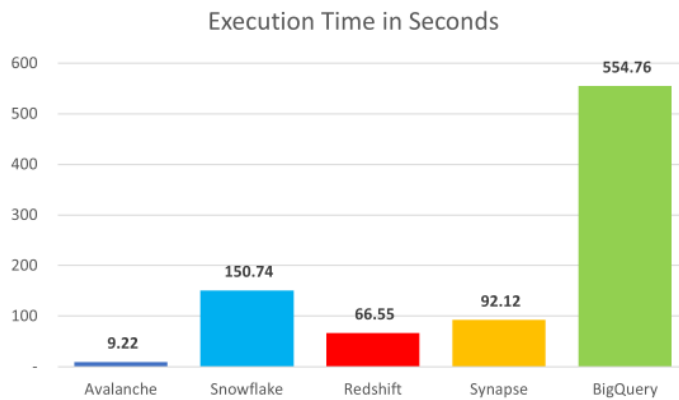
Avalanche shines here, with performance that is several times faster than the two closest competitors: Redshift and Synapse.

Query 19: "Discounted Revenue" Execution Times

Execution Time in Seconds
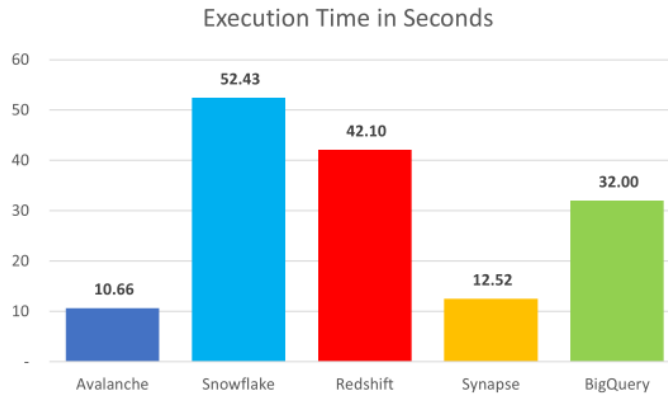
Query 19, a SUM, was another strong performer for Avalanche, with Synapse close behind.

Query 20: "Potential Part Promotion" Execution Times

Execution Time in Seconds
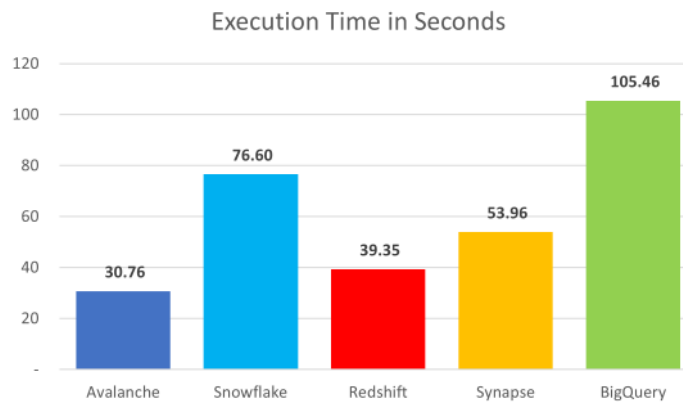
Query 20 showed Avalanche to be about 30 percent faster than second-place Redshift in an analogous configuration.

Query 21: "Suppliers Who Kept Orders Waiting" Execution Times
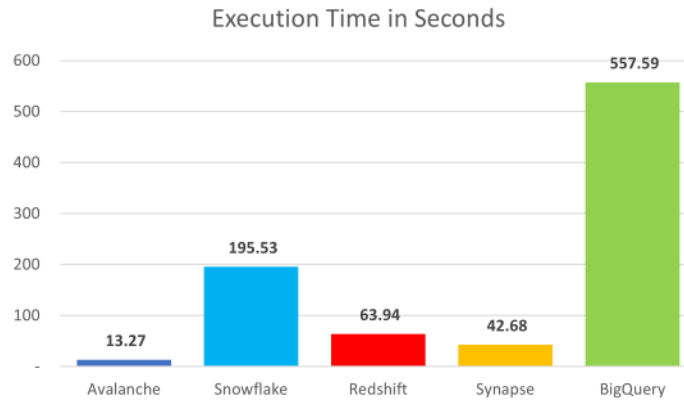
Query 21 shows the familiar pattern of high-performance differences between Avalanche and analogous configurations from its competitors.

Query 22: "Global Sales Opportunity" Execution Times
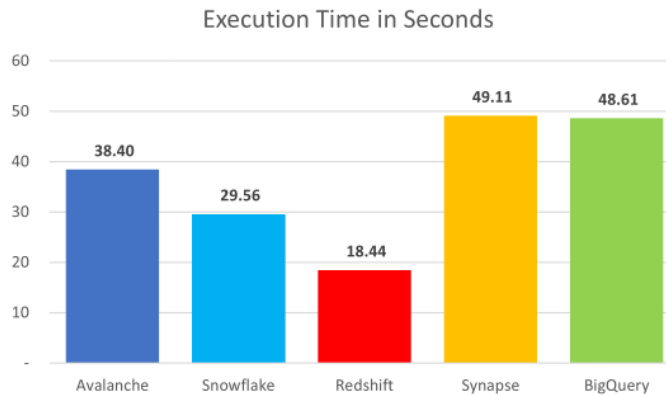
Redshift outperformed the field in Query 22, with Avalanche running third behind Snowflake.
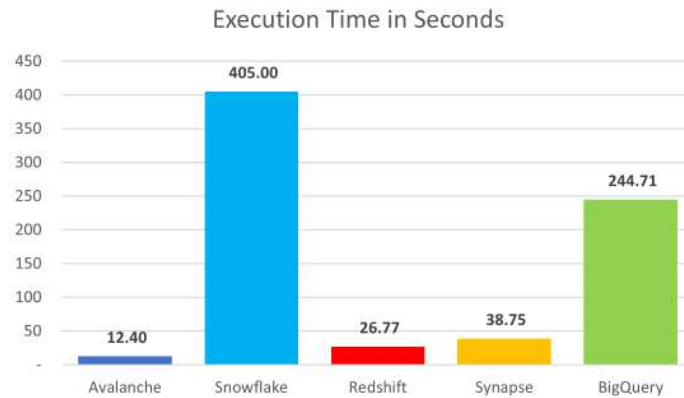
Tests with 5 Concurrent Users

*Query 1, 5 Users: "Pricing Summary Report" Execution Times*
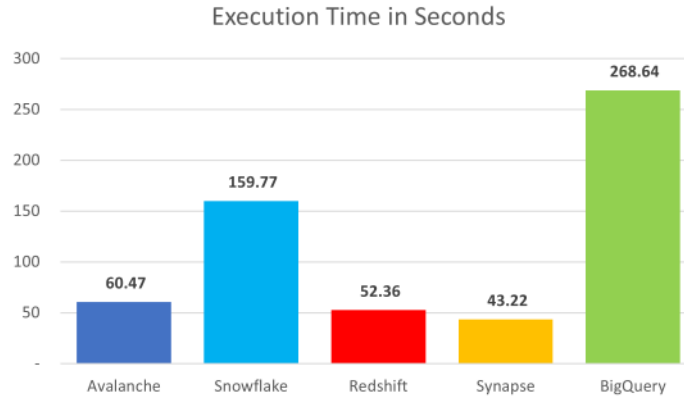
Query 1 is the only query that uses only SUM, AVERAGE, and COUNT. Synapse executed it in 43 seconds, with Redshift 9 seconds behind.

Query 2, 5 Users: "Minimum Cost Supplier" Execution Times
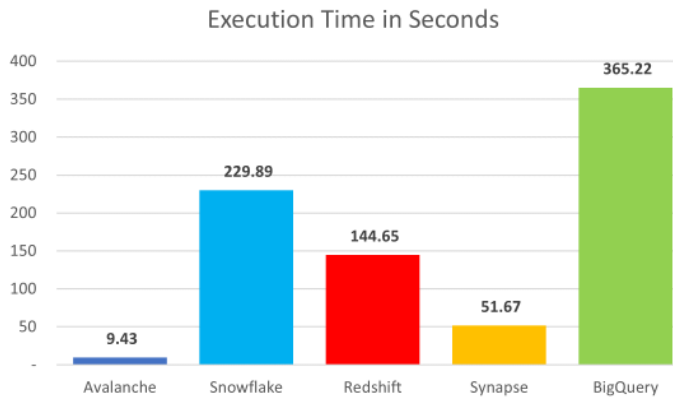

Query 2 was one of two queries that contained a MIN/MAX. Avalanche was more than twice as fast as the second-place Redshift.

*Query 3, 5 Users: "Shipping Priority" Execution Times*

Execution Time in Seconds
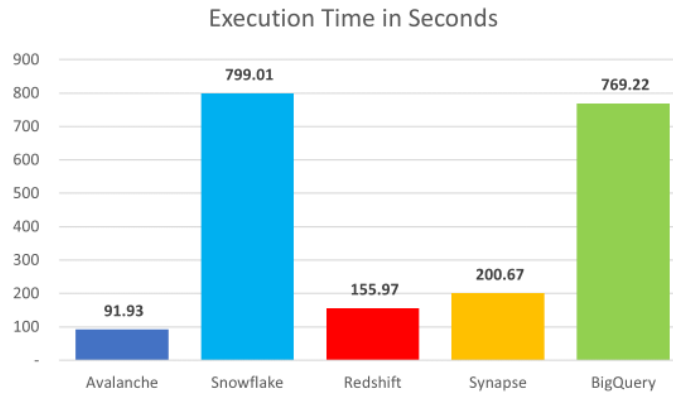
Avalanche again leads the field, with Redshift and Synapse following.

Query 4, 5 Users: "Order Priority Checking" Execution Times
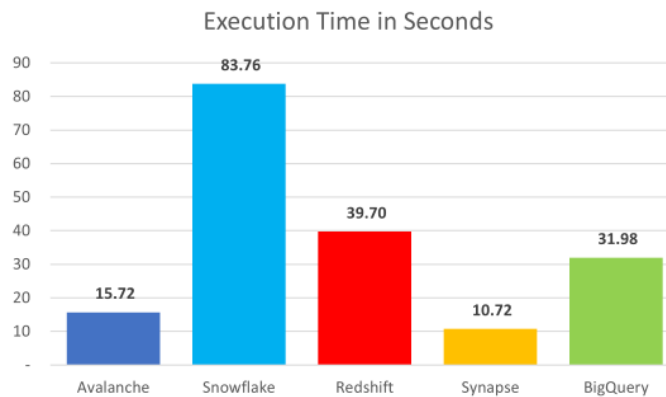
Execution Time in Seconds

In Query 4, with a SUBQUERY and COUNT, we see Avalanche being the fastest, this time by over five times ahead of second place Synapse.

Query 5, 5 Users: "Local Supplier Volume" Execution Times
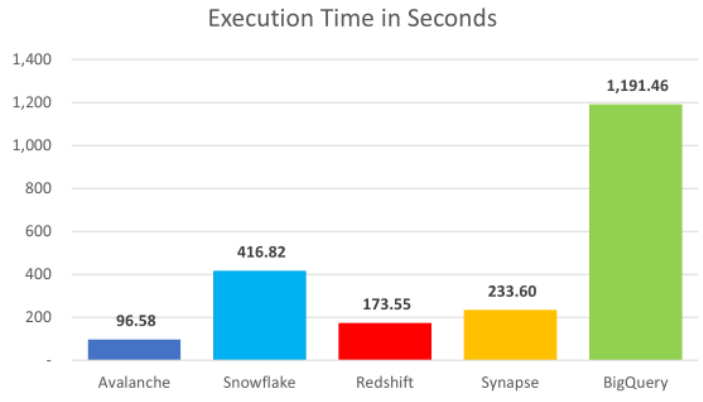
Execution Time in Seconds
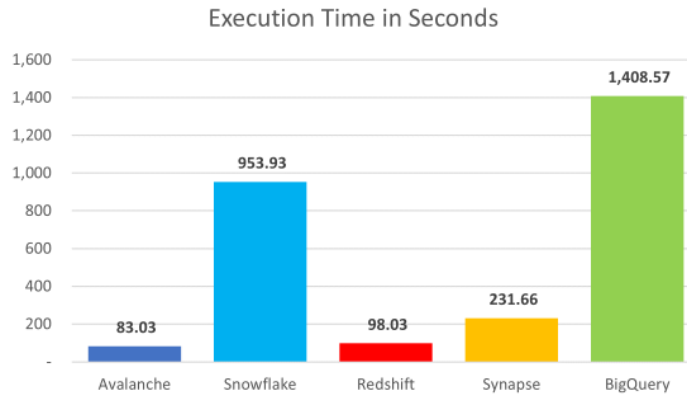
Query 5, which employs only a SUM aggregation, favored Avalanche slightly over Redshift.

Query 6, 5 Users: "Forecasting Revenue Change" Execution Times

Execution Time in Seconds

The simple SUM of Query 6 was a very high relative performer for Synapse, which was nearly 50% faster than second place Avalanche.

Query 7, 5 Users: "Volume Shipping" Execution Times

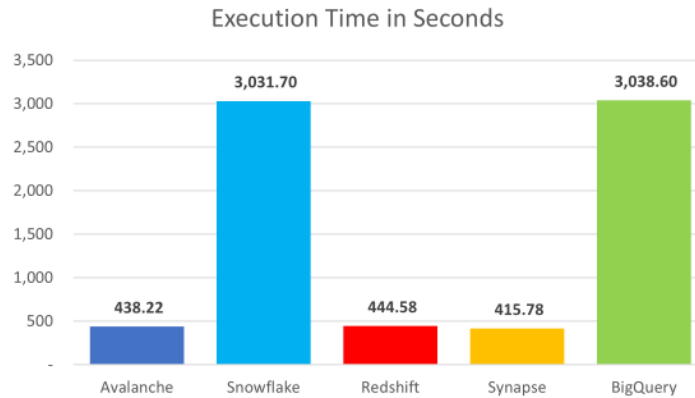
Execution Time in Seconds

In Query 7, Avalanche recorded 80% better performance than second place Redshift.

Query 8, 5 Users: "National Market Share" Execution Times
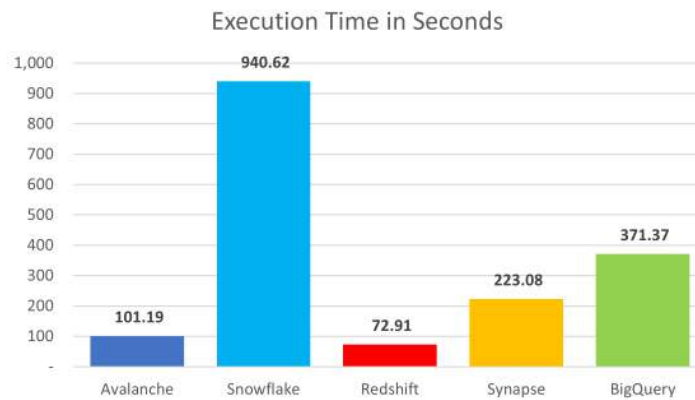

Execution Time in Seconds
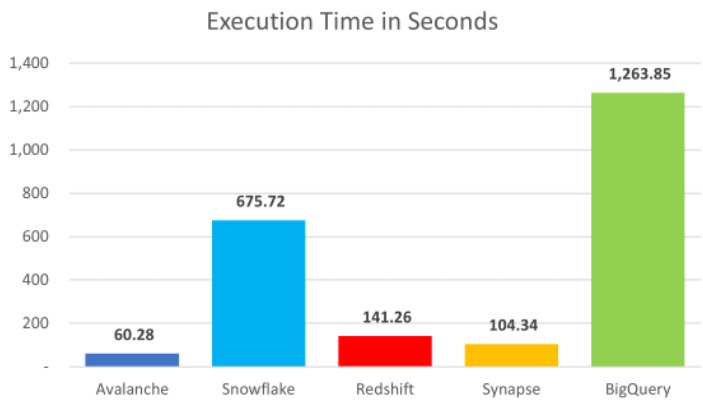
Query 8 performance proves similar to Query 7.

*Query 9, 5 Users: "Product Type Profit Measure" Execution Times*

Performance for Query 9 was very tight between Avalanche, Redshift, and Synapse.

Query 10, 5 Users: "Returned Item Reporting" Execution Times

Query 10 uniquely has a SUM and a TOP/LIMIT. Redshift here was the top performer, with Avalanche following.

Query 11, 5 Users: "Important Stock Identification" Execution Times

Execution Time in Seconds
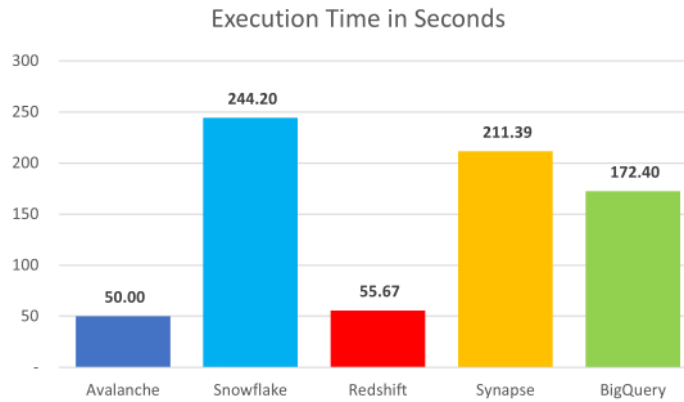
Query 11 (another SUB-SELECT and a SUM operation) saw performance slightly favoring Avalanche, with Redshift about 11% behind.

Query 12, 5 Users: "Shipping Modes and Order Priority"

Execution Time in Seconds
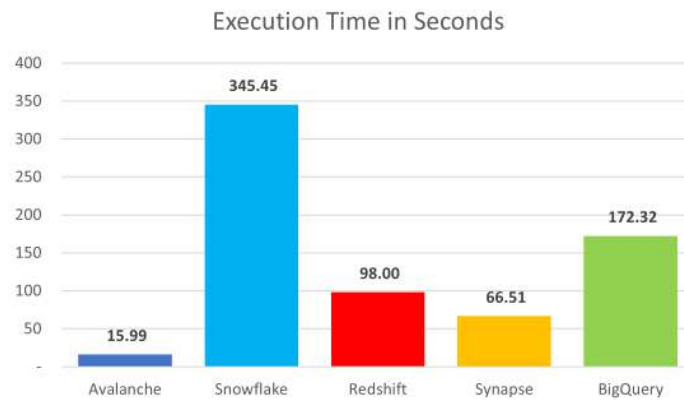
Query 12, another simple SUM, was also another plus performance query for Avalanche, executing more than 20 times faster than Snowflake.

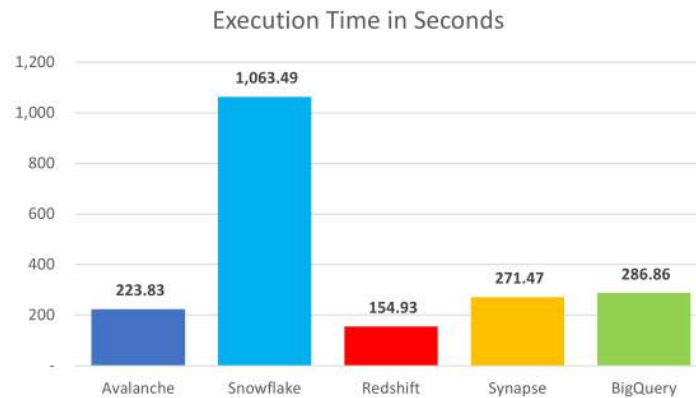*Query 13, 5 Users: "Customer Distribution" Execution Times*

Query 13 is the only TPC-H query with an explicit JOIN. Redshift outperformed the field, while second place was close between Avalanche and Synapse, with BigQuery in the mix.

Query 14, 5 Users: "Promotion Effect" Execution Times
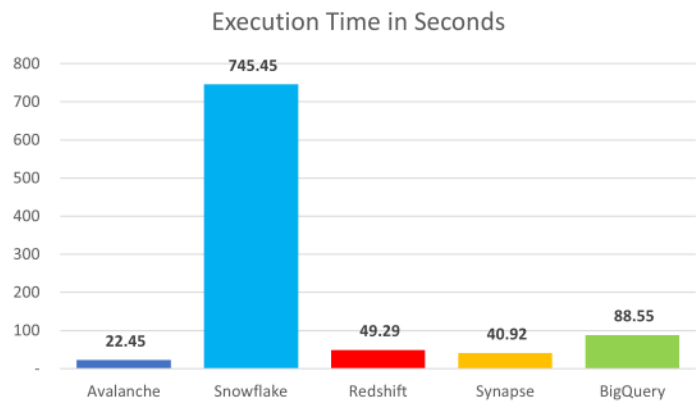
For Query 14, Actian Avalanche took an impressive 22.45 seconds to complete, 82% faster than second-place Synapse.

Query 15, 5 Users: "Top Supplier" Execution Times

Execution Time in Seconds
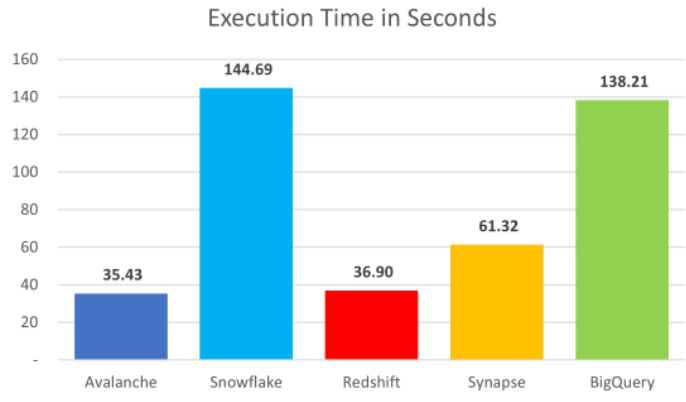
Avalanche and Redshift were neck and neck on this query.

Query 16, 5 Users: "Parts/ Supplier Relationship" Execution Times

Execution Time in Seconds
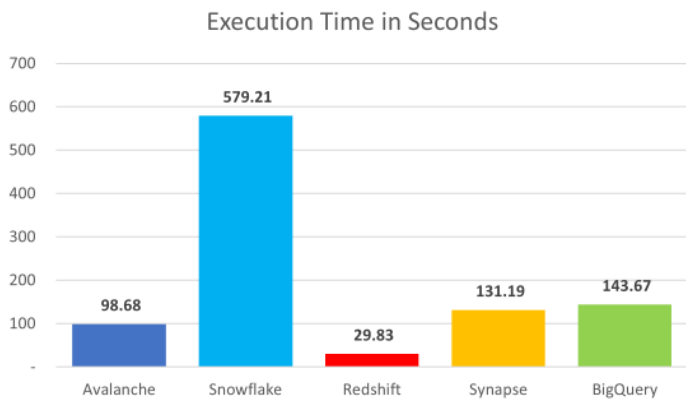
Redshift was more than three times as fast as second-place Avalanche in this query.

Query 17, 5 Users: "Small Quantity Order Revenue" Execution Times

Execution Time in Seconds
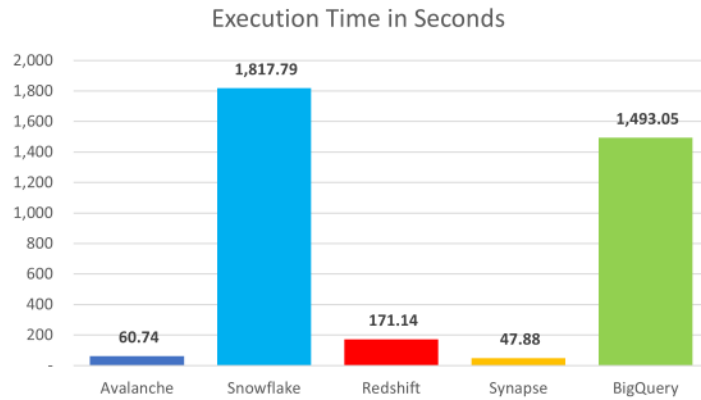
Query 17 yielded a performance advantage for Synapse, with Avalanche not far behind in second place.

Query 18, 5 Users: "Large Volume Customer" Execution Times

Execution Time in Seconds
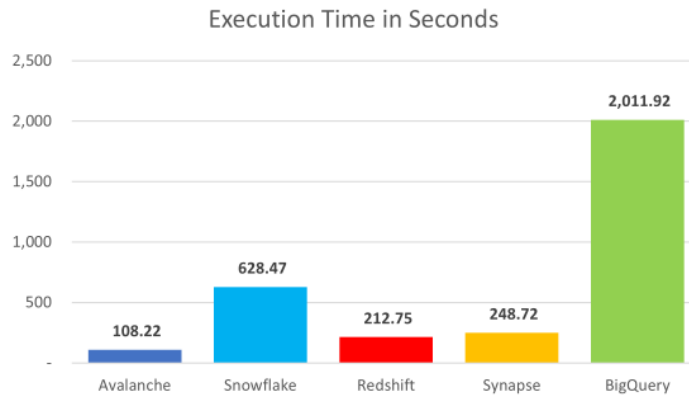
Avalanche shines here with big leads over both Redshift and Synapse.

Query 19, 5 Users: "Discounted Revenue" Execution Times

Execution Time in Seconds
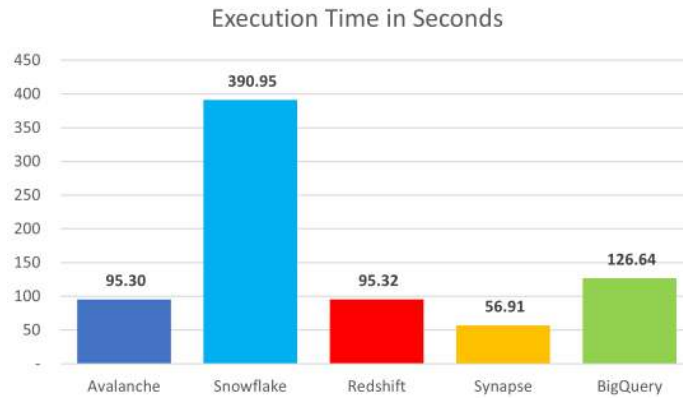
Query 19, a SUM, was a strong performer for Synapse, with Avalanche and Redshift in a dead heat for second.

Query 20, 5 Users: "Potential Part Promotion" Execution Times

Execution Time in Seconds
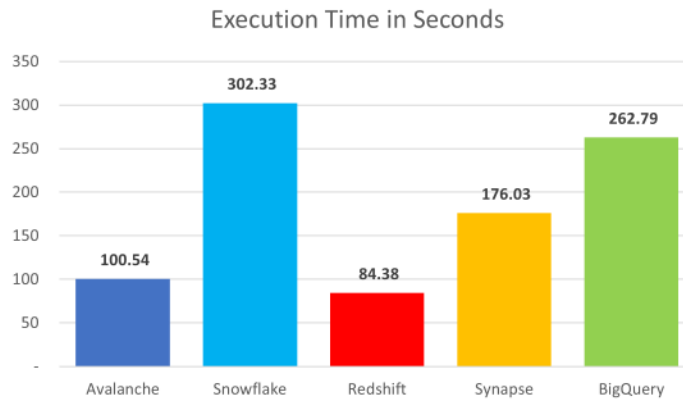
Query 20 showed Redshift to be faster in analogous configurations than Snowflake, Redshift, Synapse, and BigQuery.

Query 21, 5 Users: "Suppliers Who Kept Orders Waiting" Execution Times
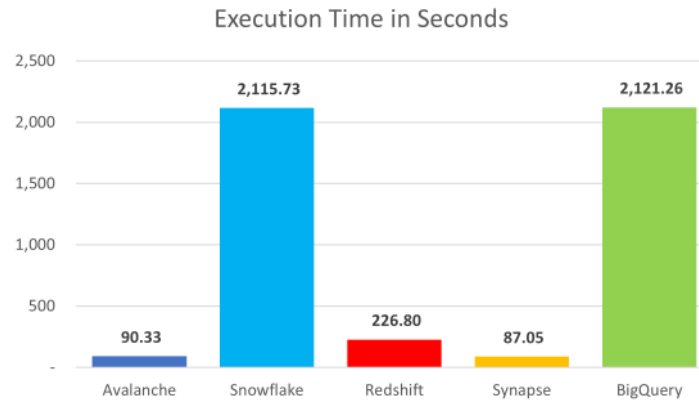
Query 21 shows Synapse with a narrow lead over second-place Avalanche.

Query 22: "Global Sales Opportunity" Execution Times
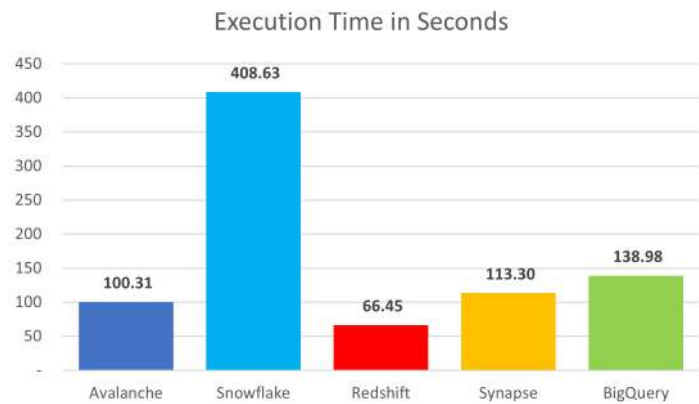
Redshift outperformed the field handily in Query 22, with Avalanche the runner up.

Aggregated One User Total Runtime

The following chart displays the aggregated elapsed time of all tests in the one user query thread. Overall, Actian Avalanche proved to be about twice as fast as runner-up Redshift.

Total Query Runtime - One User

One User Total Runtime
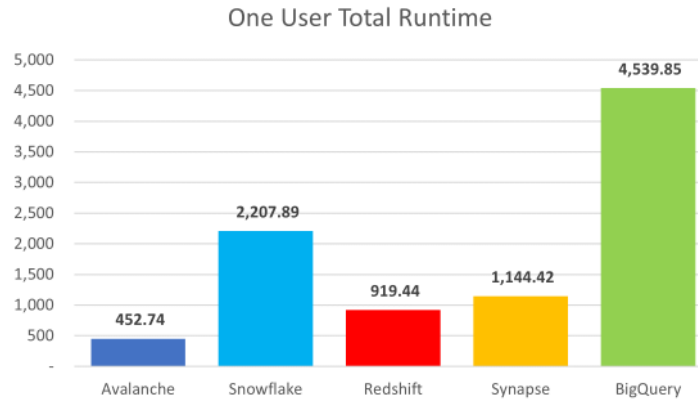
Aggregated 5-User Query Runtime

The following chart displays the query elapsed time for 5 users. Again, Avalanche enjoys a comfortable lead over second-place RedShift and third-place Synapse.

Total Query Runtime - 5 Users

Query Runtime
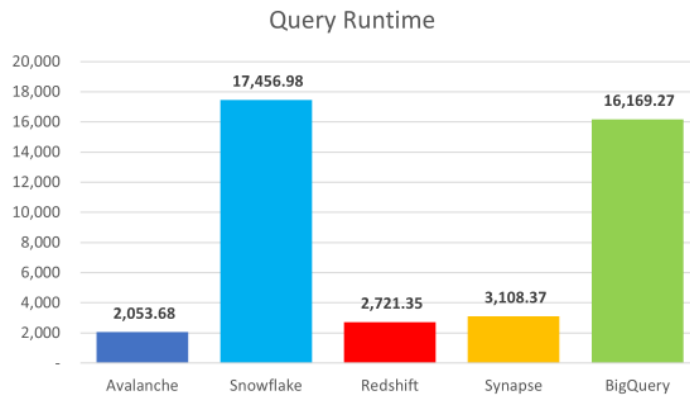
5. Price-Performance

System cost can be a difficult aspect to compare because vendor platforms vary in their pricing and licensing models. However, all three platforms have clear and consistent on-demand hourly cloud pricing that we can use to determine price per performance.

Actian Avalanche has a clear pricing model. For Avalanche software usage and the underlying platform, there is a $1.99 per Avalanche Unit (AU) per hour cost. Thus, with 128 AUs, we paid $254.72 per hour.

For Amazon Redshift, we simply paid a set dollar amount per hour by the instance class and node count that we configured. For example, considering the 20-node ra3.16xlarge configuration in the US East region (with rates at the time of the testing) we used for the 30TB test, we paid $13.04 per hour with 20 nodes for a total of $260.80 per hour.

Azure Synapse Analytics charges $0.0151 per Data Warehouse Unit (DWU) per hour. Since we used a DW15000c with 15,000 units, we paid $226.50 in the East US region.

Google BigQuery charges either monthly or hourly for Slot commitments. To ensure a like-for-like comparison, we used the Flex Slots hourly rate of $0.04 per slot per hour. With 5,000 slots, we paid $200.00 per hour in the US region.

Snowflake has several pricing tiers with varying levels of features and security capabilities. We used the Standard+ tier in US East on AWS for a cost of $3.00 per cluster node. We used a 3XLarge cluster with 64 nodes. Thus, we paid $192.00 per hour.

With the hourly cost of the configuration in hand, we calculated the price-per-performance using the following formula:

$$\frac{\text{Elapsed time of test (seconds) x Cost of platform (\$/hour)}}{3{,}600 \text{ (seconds/hour)}}$$

The elapsed time of the test is actually the duration of the slowest running thread of the concurrency test. For example, to complete a 60-user test, we have to wait until all 60 users complete all their queries. Thus, the slowest thread represents the elapsed time of the test from beginning to end.

Table 5 details the price-performance figures for the different tests, while the chart below visualizes the price-performance difference for five concurrent users.

Table 5. Price Performance

	Avalanche	Snowflake	Redshift	Synapse	BigQuery
Cluster	AU	3XLarge	ra3.16xlarge	DW15000c	Flex Slots
Units	128	64	20	15,000	5,000
Cost $/unit/hour	$1.99	$3.00	$13.04	$0.02	$0.04
$/hour	$254.72	$192.00	$260.80	$226.50	$200.00
Single User Total Time	453	2,208	920	1,144	4,540
Price Performance Single User	**$32.03**	**$117.75**	**$66.61**	**$72.00**	**$252.21**
Avalanche Comparison		3.7x	2.1x	2.2x	7.9x
5 Concurrent Users Total Slowest Time	2,054	17,458	2,721	3,108	16,169
Price Perf 5 Concurrent Users	**$145.31**	**$931.07**	**$197.15**	**$195.57**	**$1,796.59**
Avalanche Comparison		6.4x	1.4x	1.3x	12.4x

Price-Performance, 5 Concurrent Users

Price Performance, 5 Concurrent Users
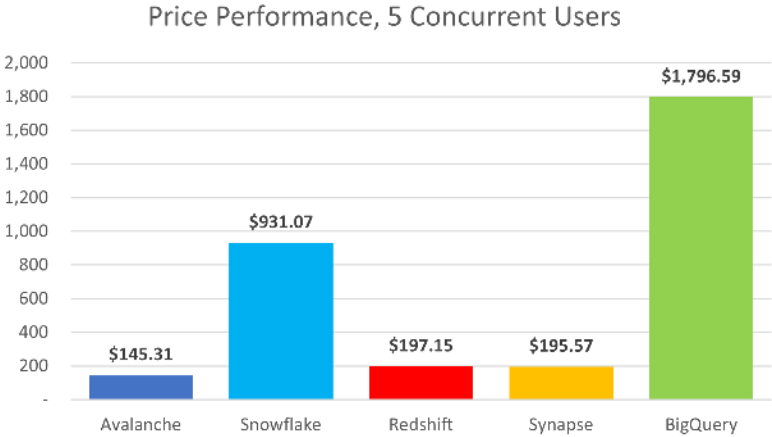
6. Conclusion

Cloud databases are a way for enterprises to avoid large capital expenditures, provision quickly, and provide performance at scale for advanced analytic queries. Relational databases with analytic capabilities continue to support the advanced analytic workloads of the organization with performance, scale, and concurrency. In a representative set of corporate-complex queries from the well-known TPC-H standard, Actian Avalanche consistently outperformed the competition.

Overall, the benchmark results were insightful in revealing query execution performance and some of the differentiators for Avalanche, Synapse, Snowflake Data Warehouse, Amazon Redshift, and Google BigQuery. Avalanche query response times on the GigaOm Analytic Field Test derived from the industry standard TPC Benchmark™ at 30TB were overall 8.5 times faster than Snowflake, approximately 1.3 times the performance of Redshift, 1.5 time the performance of Synapse, and 7.9 times faster than BigQuery in tests of five concurrent users.

Examining price-performance, Avalanche ran the 5 concurrent user TPC-H queries roughly 6.4 times more cost effectively than Snowflake, as measured in cost per query per hour. Avalanche further proved 1.4 times better than Redshift, 1.3 times better than Synapse, and 12.4 times better than Google BigQuery in terms of cost per query per hour across the examined cluster classes.

Price and performance are critical points of interest when it comes to selecting an analytics platform, because they ultimately impact total cost of ownership, value, and user satisfaction. Our analysis reveals Avalanche to be the industry leader on this criterion.

7. Disclaimer

Performance is an important criterion for data warehouse platform selection, but it is only one factor among many to consider in the overall procurement process. Other significant factors to consider in selection include: administration, integration, workload management, user interface, scalability, vendor, reliability, and numerous others. It is our experience that performance as a selection criterion is prone to change over time and can produce different conclusions based on varying workloads. A performance leader can likewise run up against the point of diminishing returns, while viable contenders are able to close the gap quickly. All of which is to say, performance testing is a useful tool in selecting a data warehouse platform, but it should be considered in the context of the overall solution and its value.

GigaOm runs all of its performance tests to strict ethical standards. The results of the report are the objective results of the application of queries to the simulations described in the report. The report clearly defines the selected criteria and process used to establish the field test. It also clearly states the data set sizes, the platforms, the queries, and other elements used. The reader is left to determine for themselves how to qualify the information for their individual needs. The report does not make any claim regarding third-party certification and presents the objective results received from the application of the process to the criteria as described in the report. The report strictly measures performance and does not purport to evaluate other factors that potential customers may find relevant when making a purchase decision.

This is a sponsored report. Actian chose the competitors, the test, and the Actian Avalanche configuration. GigaOm chose the most compatible configurations for the other tested platforms and ran the queries. Choosing compatible configurations is subject to judgment. Avalanche can be deployed on AWS, Azure, and Google Cloud, as well as on-premises. In this instance Avalanche was tested on AWS.

We have attempted to describe our decisions in this report.

In this writeup, all the information necessary to replicate this test is included. You are encouraged to compile your own representative queries, data sets, data sizes, and compatible configurations and test for yourself.

8. About Actian

Actian, the hybrid data management, analytics and integration company, delivers data as a competitive advantage to thousands of customers worldwide. Through the deployment of innovative hybrid data technologies and solutions Actian ensures that business critical systems can transact and integrate at their very best – on premise, in the cloud or both. Thousands of forward-thinking organizations around the globe trust Actian to help them solve the toughest data challenges to transform how they run their businesses, today and in the future. For more, visit **http://www.actian.com**.

9. About William McKnight

An Ernst & Young Entrepreneur of the Year Finalist and frequent best practices judge, William is a former Fortune 50 technology executive and database engineer. He provides Enterprise clients with action plans, architectures, strategies, and technology tool selection to manage information.

William McKnight is an Analyst for GigaOm Research who takes corporate information and turns it into a bottom-line producing asset. He's worked with companies like Dong Energy, France Telecom, Pfizer, Samba Bank, ScotiaBank, Teva Pharmaceuticals and Verizon — Many of the Global 2000 — and many others. William focuses on delivering business value and solving business problems utilizing proven, streamlined approaches in information management.

He is a frequent international keynote speaker and trainer. William has taught at Santa Clara University, UC-Berkeley and UC-Santa Cruz.

10. About Jake Dolezal

As a contributing Analyst at GigaOm, Jake Dolezal has two decades of experience in the Information Management field with expertise in analytics, data warehousing, master data management, data governance, business intelligence, statistics, data modeling and integration, and visualization. Jake has experience across a broad array of industries, including: healthcare, education, government, manufacturing, engineering, hospitality, and restaurants. He has a doctorate in information management from Syracuse University.

11. About GigaOm

GigaOm provides technical, operational, and business advice for IT's strategic digital enterprise and business initiatives. Enterprise business leaders, CIOs, and technology organizations partner with GigaOm for practical, actionable, strategic, and visionary advice for modernizing and transforming their business. GigaOm's advice empowers enterprises to successfully compete in an increasingly complicated business atmosphere that requires a solid understanding of constantly changing customer demands.

GigaOm works directly with enterprises both inside and outside of the IT organization to apply proven research and methodologies designed to avoid pitfalls and roadblocks while balancing risk and innovation. Research methodologies include but are not limited to adoption and benchmarking surveys, use cases, interviews, ROI/TCO, market landscapes, strategic trends, and technical benchmarks. Our analysts possess 20+ years of experience advising a spectrum of clients from early adopters to mainstream enterprises.

GigaOm's perspective is that of the unbiased enterprise practitioner. Through this perspective, GigaOm connects with engaged and loyal subscribers on a deep and meaningful level.

12. Copyright

© Knowingly, Inc. 2020 *"High-Performance Cloud Data Warehouse Performance Testing"* is a trademark of Knowingly, Inc.. For permission to reproduce this report, please contact sales@gigaom.com.