Valmet Industrial Internet & Snowflake - learnings

Valmet

Leading process technologies, automation and services for the pulp, paper and energy industries

This is Valmet

Services

Mill and plant improvements, roll and workshop services, parts and fabrics, and life-cycle services

Automation

Supplies and develops automation and information management systems, applications and services

Pulp and Energy

Technologies and solutions for pulp production, power generation, and biomass conversion

Paper

Technologies and solutions for board, tissue, and paper

Global customer base

Process technology, services and automation

Valmet’s unique offering differentiates the company from its competitors

Paper
- Board, paper and tissue production lines
- Rebuilds
- Stand-alone products

Pulp
- Wood and pulp handling
- Fiber processing
- Recovery

Energy
- Heat and power generation
- Air emission control
- Biofuels

Services
- Spare parts and components
- Maintenance and shutdown services
- Outsourcing services
- Production consumables
- Process support and optimization

Automation
- Distributed Control Systems (DCS)
- Quality Management Systems (QMS)
- Analyzers and measurements
- Industrial Internet solutions

Valmet Industrial Internet

A meaningful dialogue with data brings tangible results

Valmet Industrial Internet

Dialogue with data:

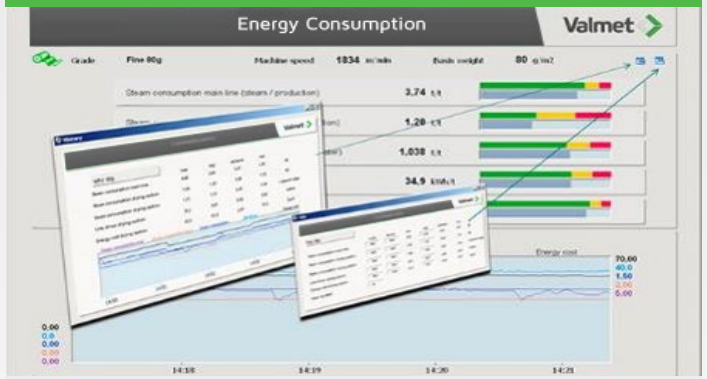
- Combining process and business data from different mill or plant systems
- Leveraging advanced analytics and Valmet's know-how to create new data driven applications
- Providing applications for operator assistance and new set points for the automation system

Results

- Reduced raw material and energy cost
- Reduced downtime and unplanned stops
- Improved product quality

Key elements of Valmet Industrial Internet

Industrial Internet applications



From analytical applications for reliability and performance to Advanced Process Controls, information management and process simulators

Valmet Performance Center

Provides remote support, monitoring and data analysis and access to Valmet's expert network

Valmet Customer Portal

A digital, personalized collaboration space between you and Valmet

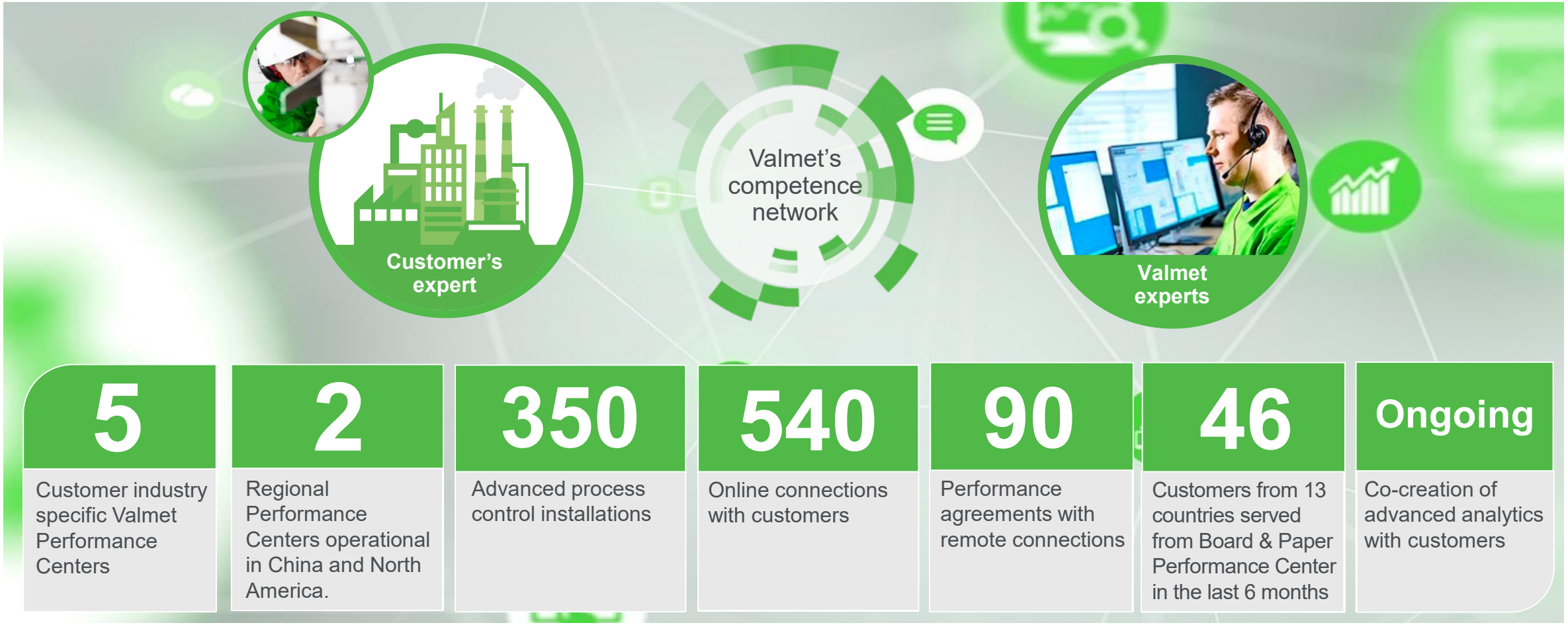
Intelligent machines and automation

A solid data source for Industrial Internet solutions

Solution ecosystem

Brings leading industry players and innovative start-ups together to co-create new value-adding data driven services

Today, customers are extensively utilizing our Industrial Internet capabilities



5	2	350	540	90	46	Ongoing
Customer industry specific Valmet Performance Centers	Regional Performance Centers operational in China and North America.	Advanced process control installations	Online connections with customers	Performance agreements with remote connections	Customers from 13 countries served from Board & Paper Performance Center in the last 6 months	Co-creation of advanced analytics with customers

Industrial Internet case example:
Paper strength prediction combined with advanced process controls

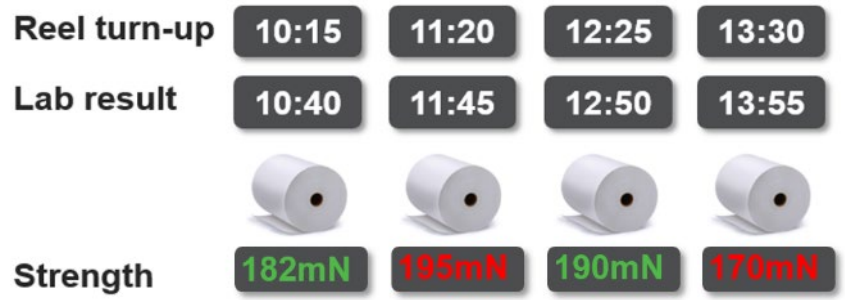
Challenge: In the paper production process, there are still several quality variables, which can't be measured and controlled until the product is manufactured.

Solution:

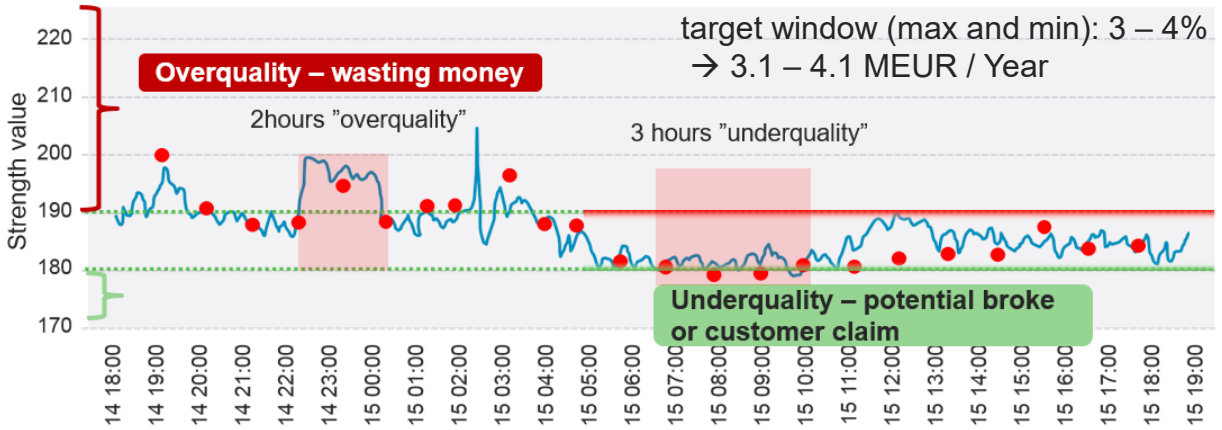
- A decision support application for the operator to control stock preparation based on predicted paper strength level to minimize raw material cost.
- Remote service is in key role to maintain the application.

Example results from magazine paper:

- Real time information on paper strength level has enabled operator to control blending to allow 1-2% savings in kraft consumption (~1M€ per year)



Kraft amount difference inside the target window (max and min): 3 – 4%
→ 3.1 – 4.1 MEUR / Year



Valmet

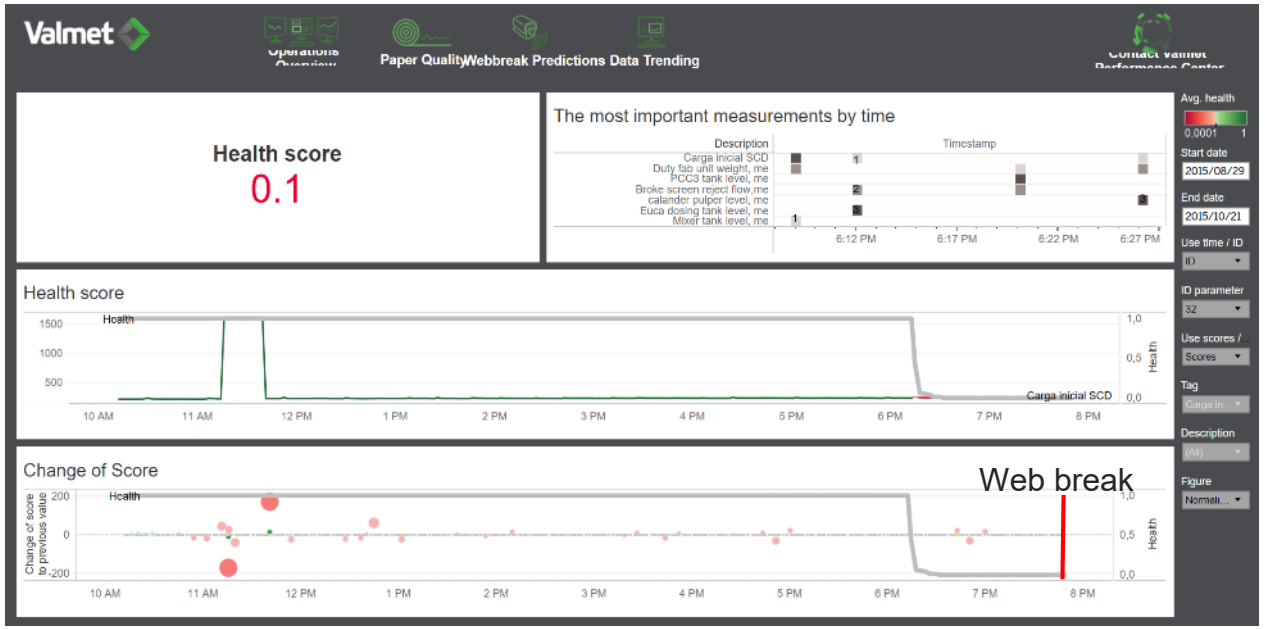
Industrial Internet case example: New performance-based services relying on data analytics and remote services

Challenge:

- Web breaks are causing a lot of unplanned operational downtime in the paper industry.
- In most cases, the operator does not know the real root cause
- There is information value, if you can predict them, but monetary value comes via preventing them happening.

Solution:

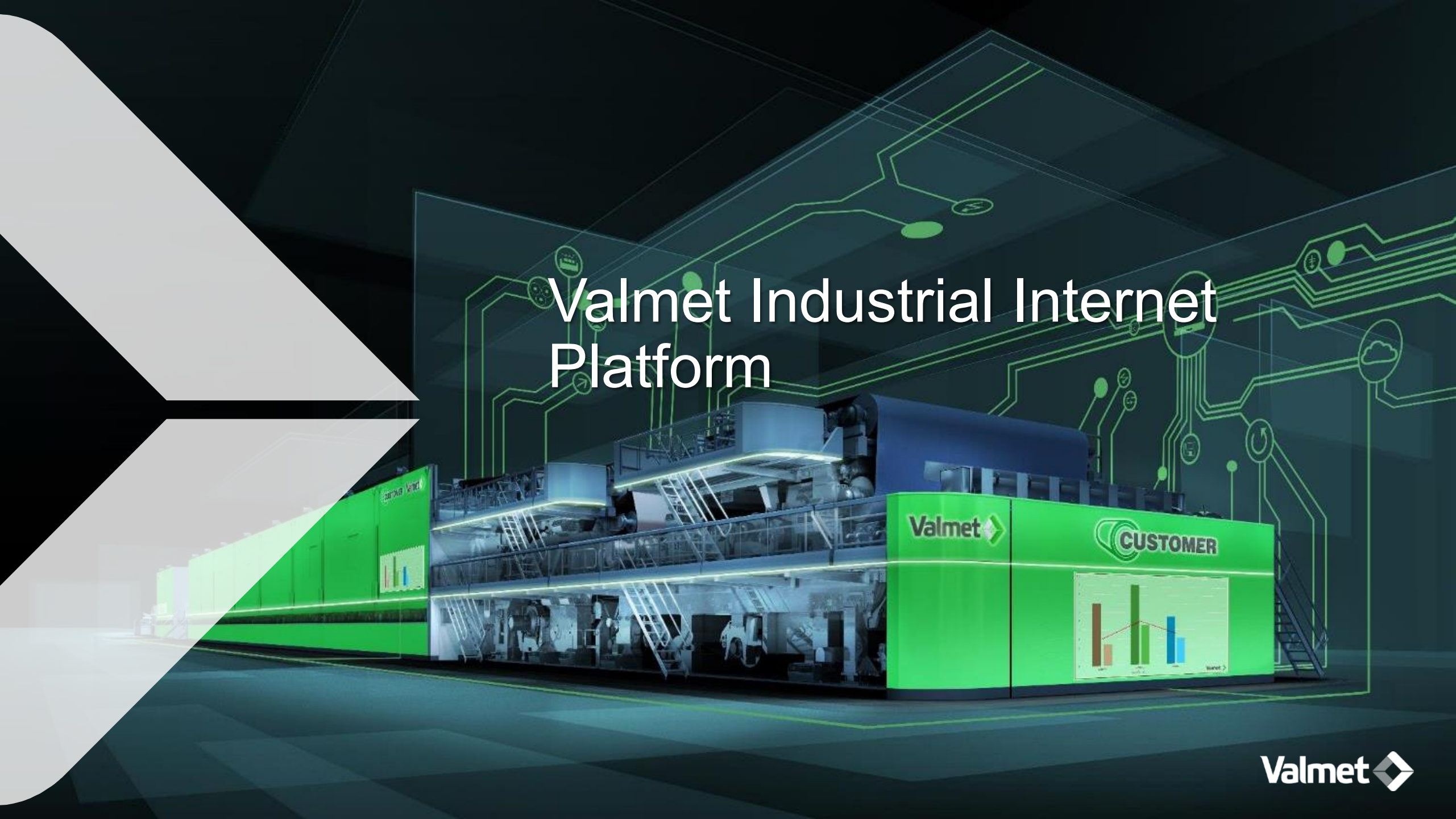
- An analytical application that predicts upcoming sheet breaks in the process and communicates with operators through specific user interface
- The application also illustrates the root cause (contributing variables) for predicted web break



Results:

- Mill 1. 50% web break capture rate (2h in advance)
- Mill 2. 62% web break capture rate (2h in advance)

Valmet Industrial Internet Platform

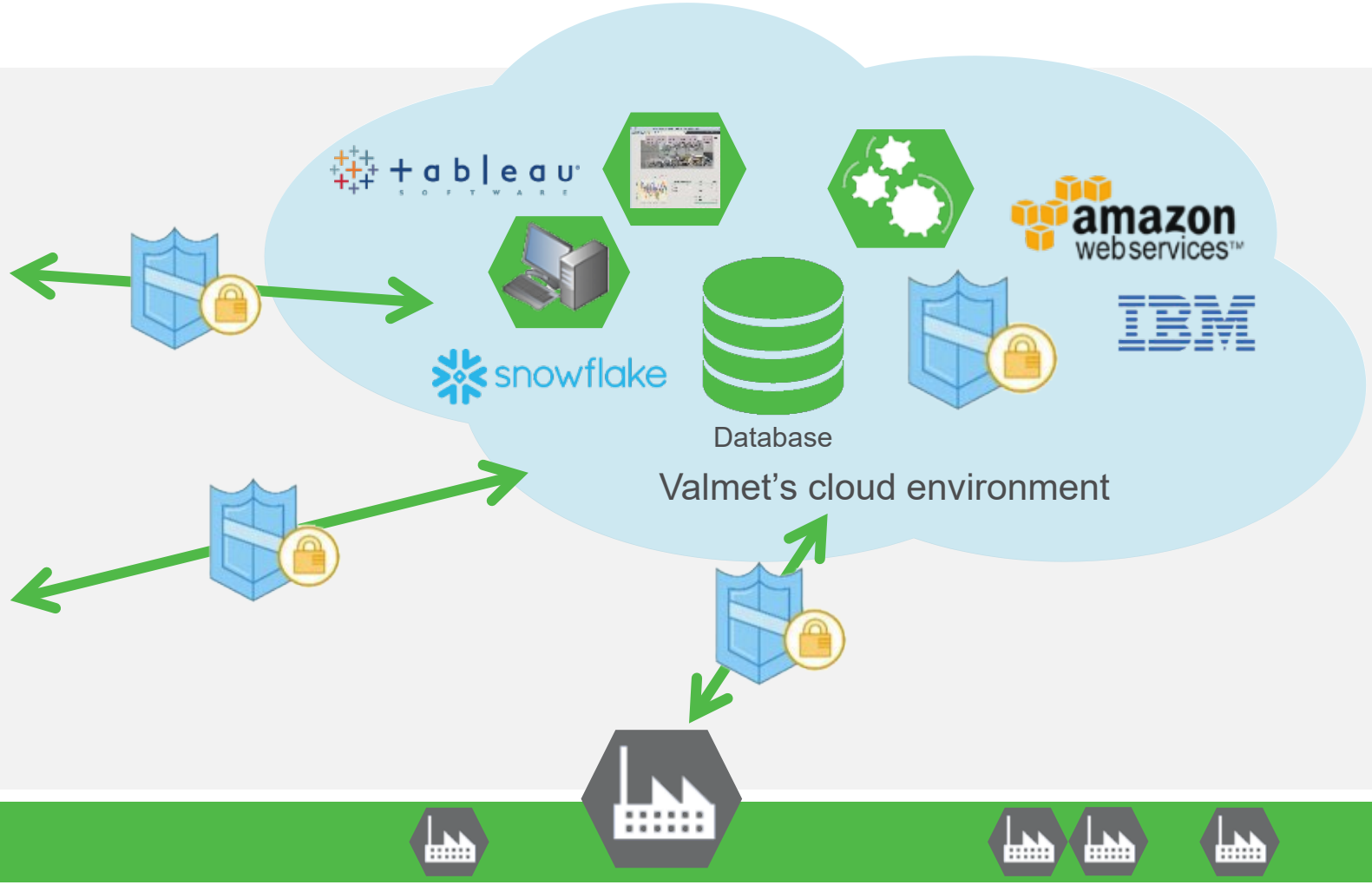


Valmet Industrial Internet platform components

based on certified and secure world-leading technologies

Valmet Customer Portal

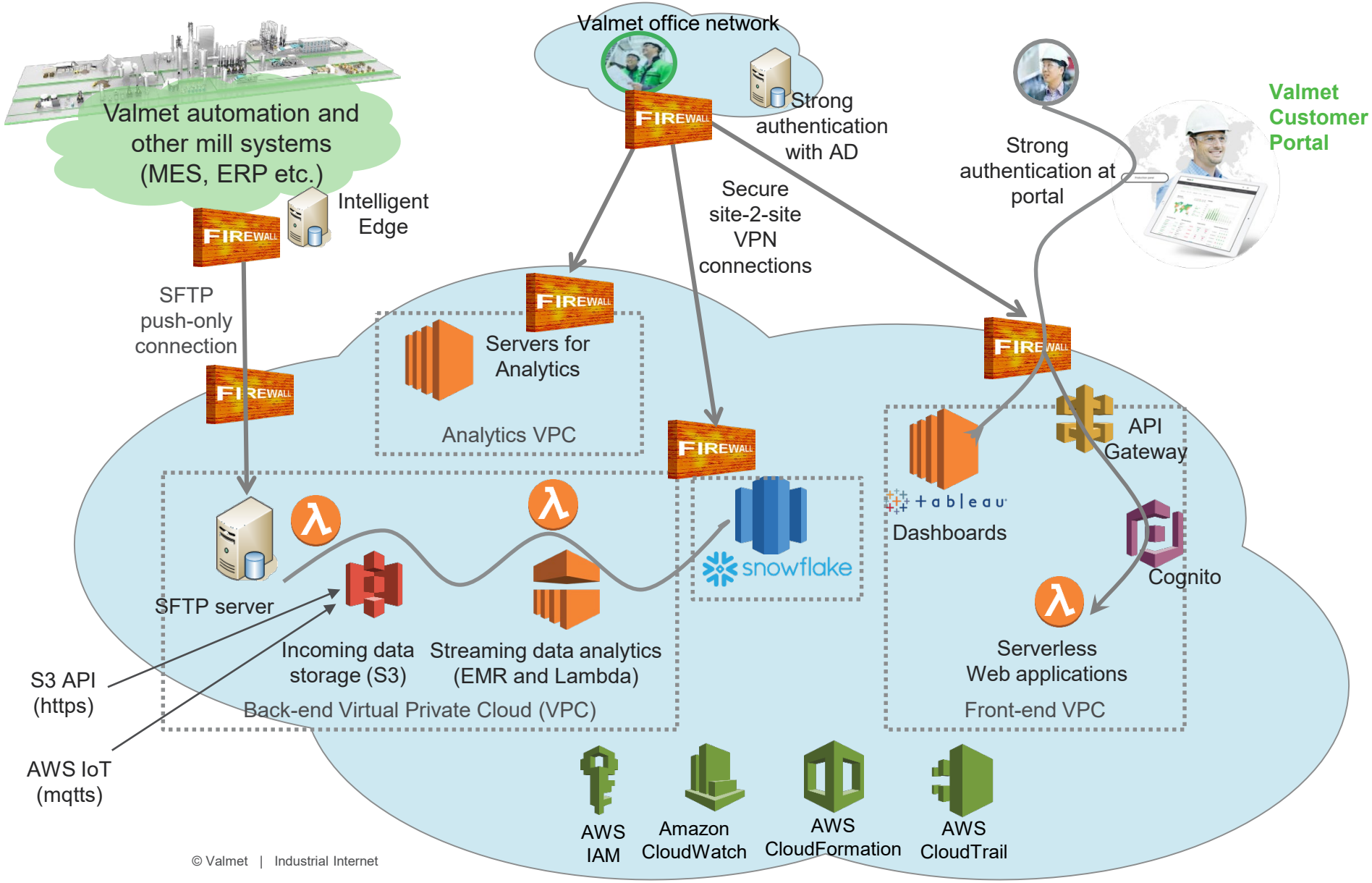
Valmet Performance Center



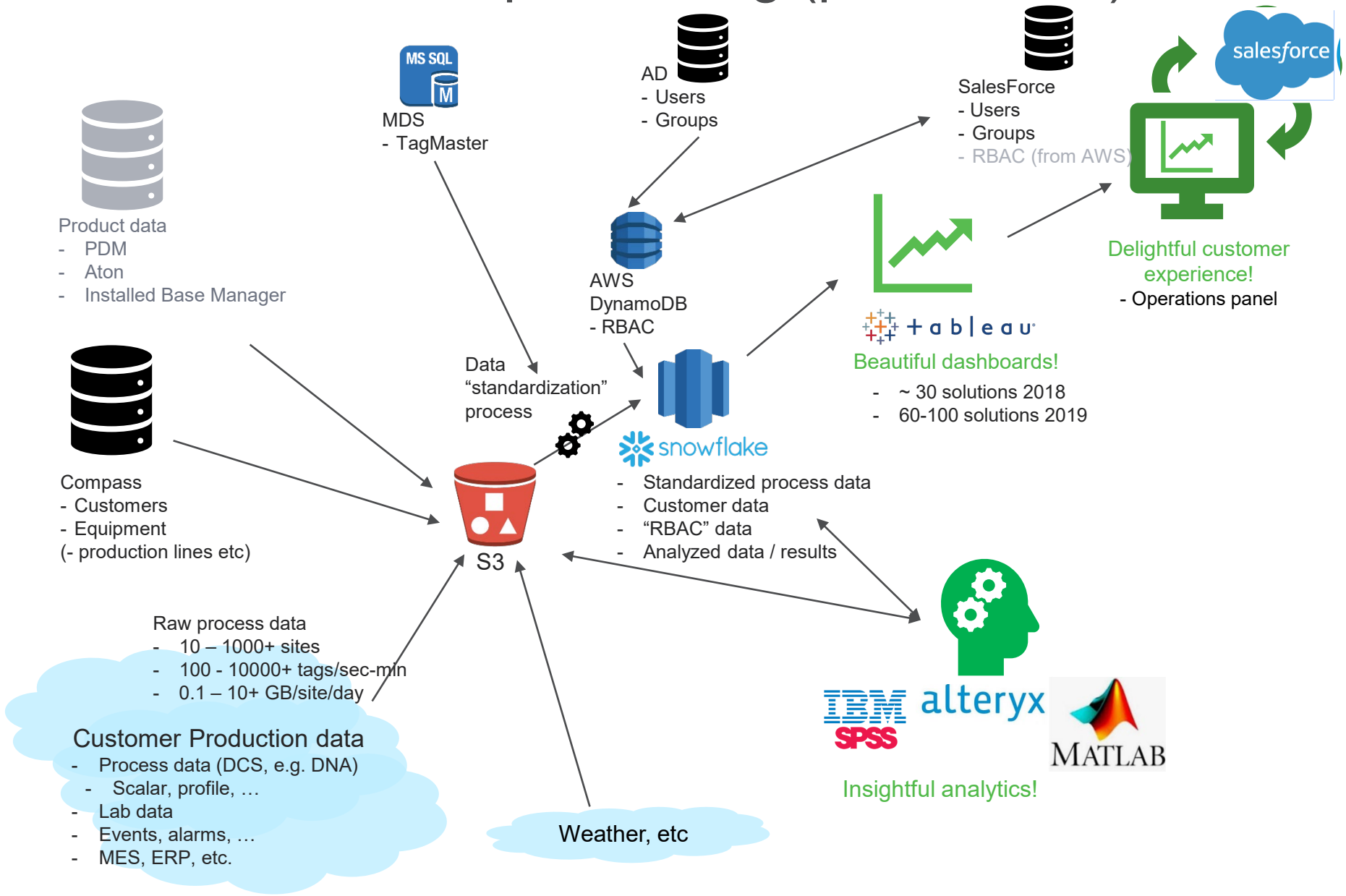
Valmet DNA

Valmet

Valmet Industrial Internet in Cloud

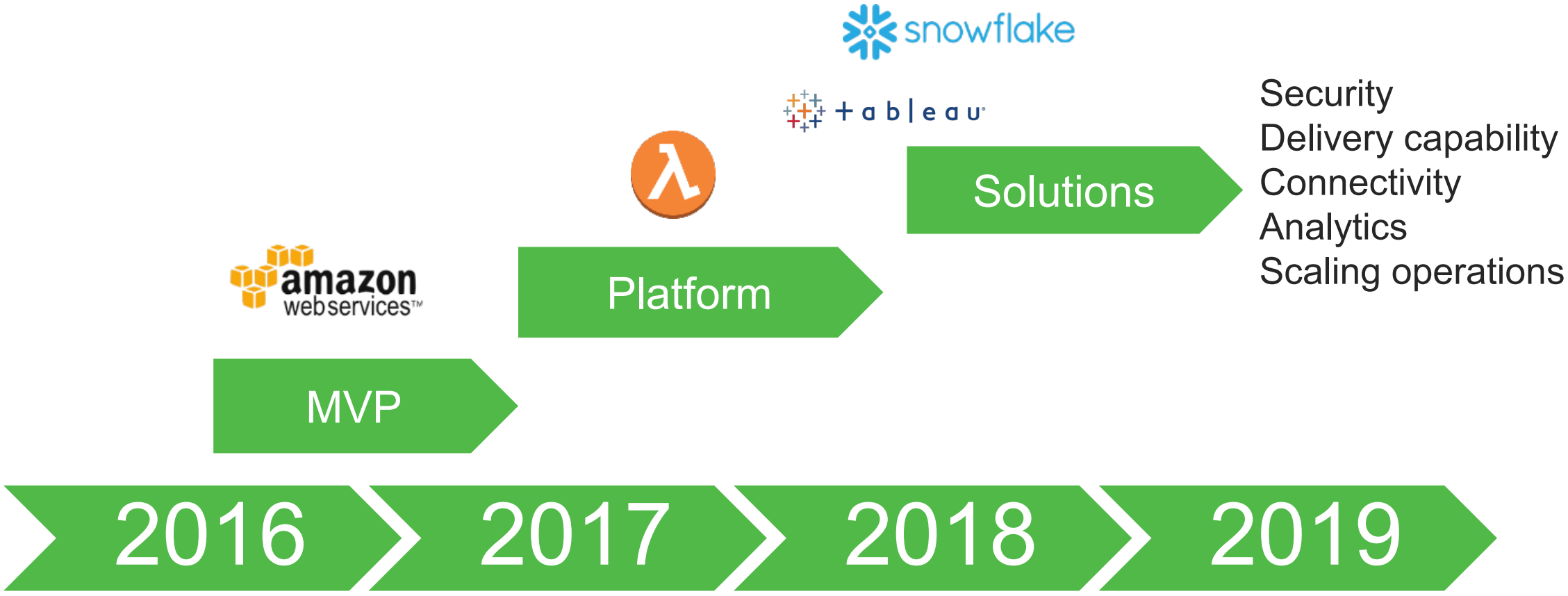


Data sources and processing (partial view)



Challenge & Snowflake

Timeline for Valmet Industrial Internet platform evolution



Our challenge late 2017 when building the platform

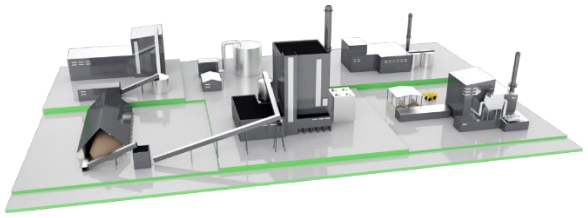
The setup in MVP / first platform version was complex and poorly performing although the amount of data was fairly limited (less than 200 GB)

With then selected technologies we had significant delays in querying data for visualization (up to 1 minute)

After analysis we decided to renew both data visualization and database technologies we use > cost of switching still small in the early implementation

Valmet

Selecting new database tech: Data as a basis of selection



Scalar data from one line from one mill / year

- 900m rows
- 75 GB

Profile data from one line from one mill / year

- 100m rows
- 15 GB

Data from 200 production lines would mean about **18 TB of compressed data / year**

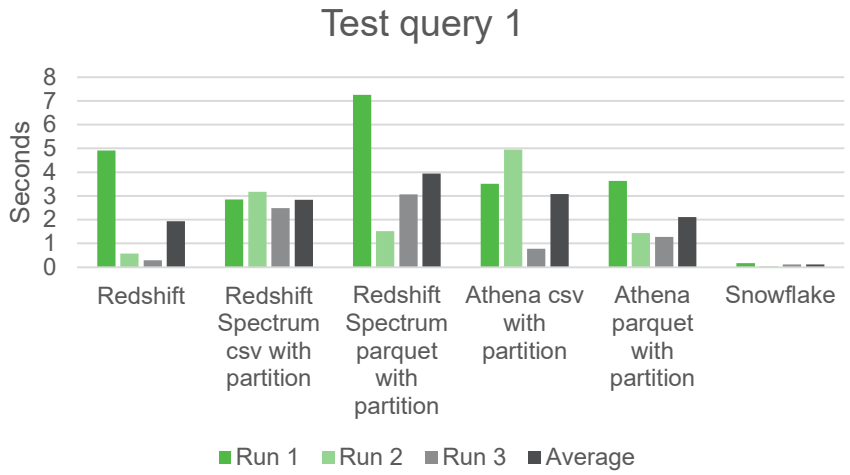
Potentially hundreds of concurrent simultaneous queries from thousands of end users

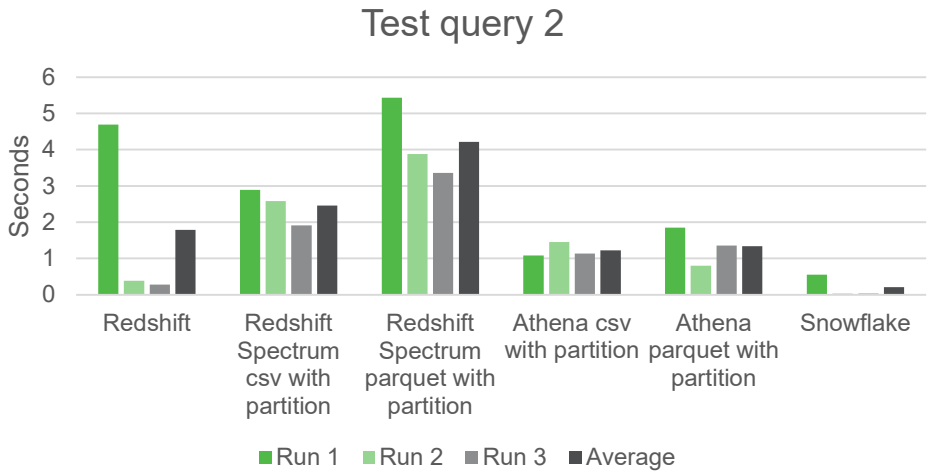
Most of the data comes from sensors. Multiple sources for data and data coming in constantly.
Many data consumers. Query performance should not be affected by the number of parallel queries.
Only small portion of the data comes from internal systems.

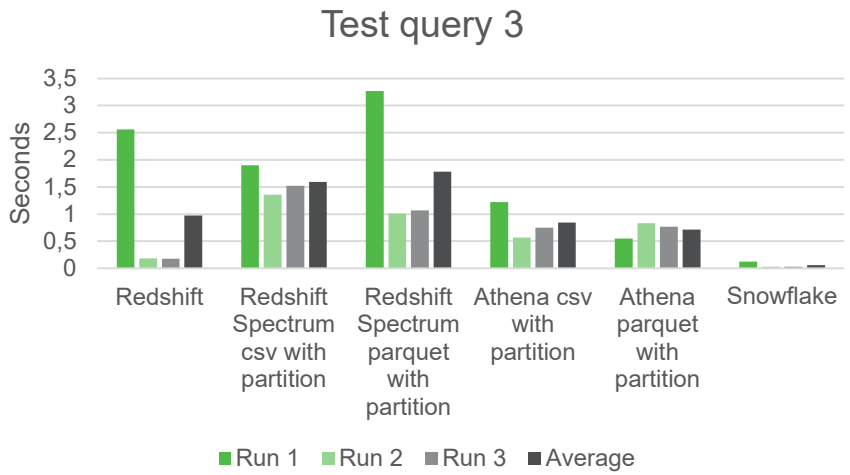
Valmet

Testing performance: Data Platform / Queries

Query performance: Snowflake vs. Athena vs. Redshift vs. Redshift Spectrum



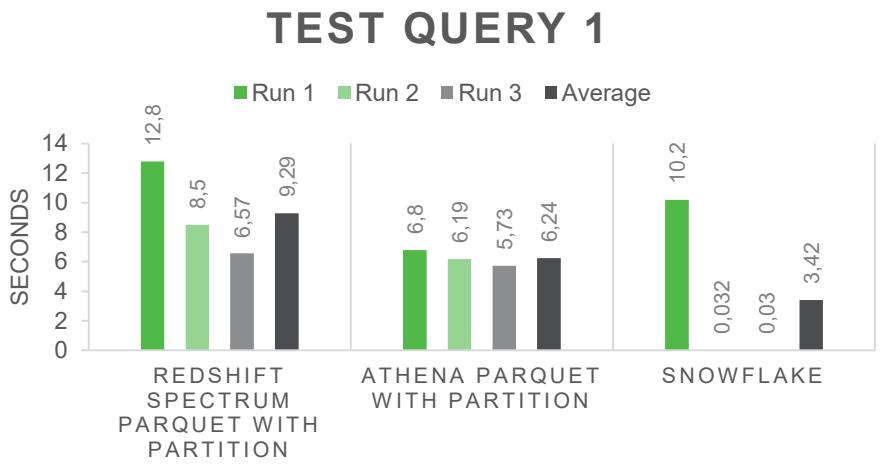


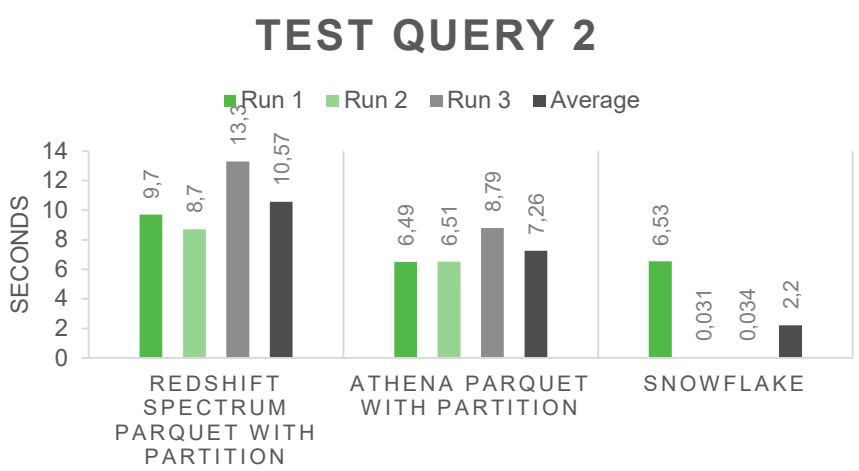


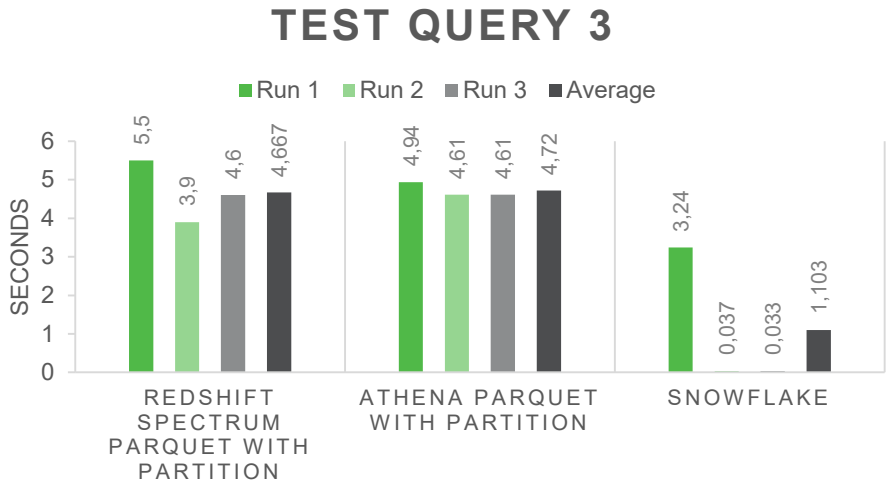
- Test case:
  - Snowflake database size X-Small (Snowflake)
  - Redshift cluster with 4 nodes (dc2.large) (Redshift, Redshift Spectrum)
  - Athena is serverless
  - Test table with 11 857 626 rows.

Testing performance: Data Platform / Queries

Query performance: Snowflake vs. Athena vs. Redshift Spectrum



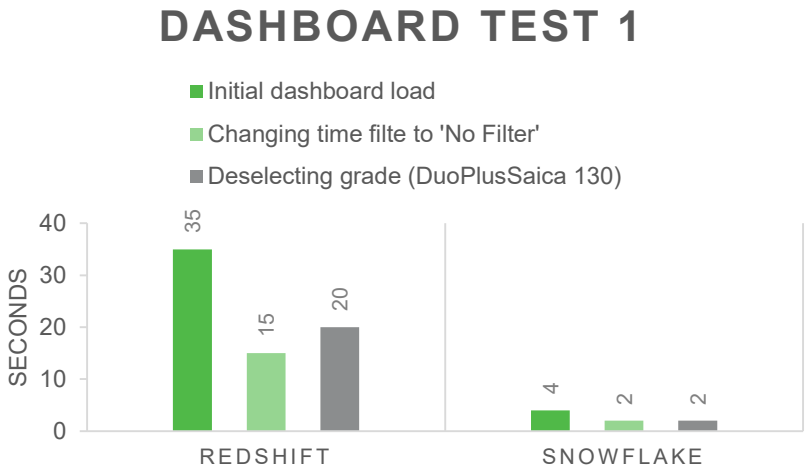


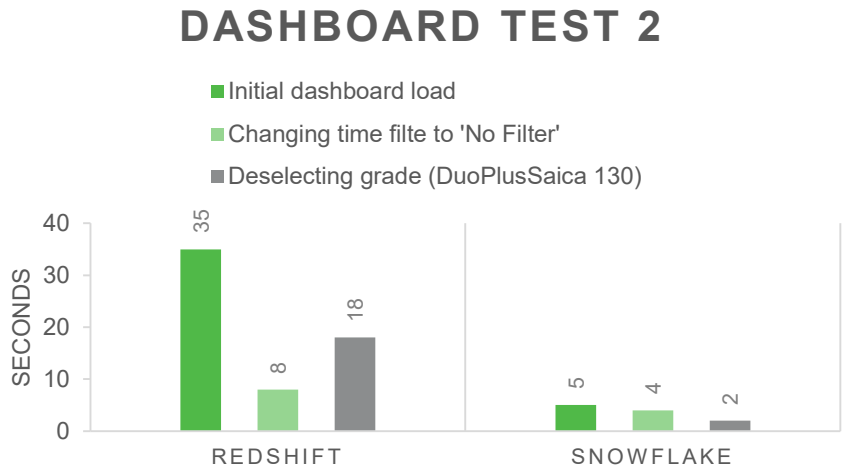


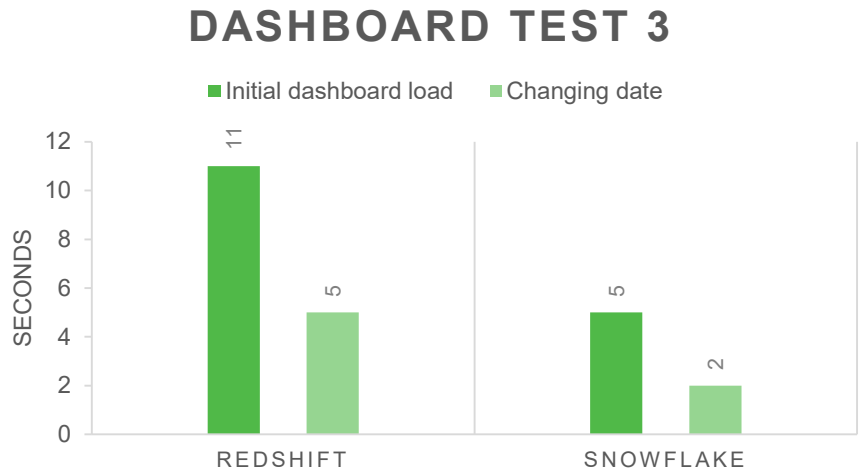
- Test case:
  - Snowflake database size X-Small (Snowflake)
  - Redshift cluster with 4 nodes (dc2.large) (Redshift, Redshift Spectrum)
  - Athena is serverless
  - Tested with table with 6 881 759 701 rows.
  - Redshift run out of disk space during tests.

Data platform / Queries

Pre-Tableau dashboard performance (QCS KPI)







- Test cases:
  - Redshift 4 node cluster (dc2.large)
  - Snowflake XS warehouse
  - Test Case 1: Paper and board KPI old / Home page
  - Test Case 2: Paper and boar KPI old / Production
  - Test Case 3: QCS / Copy of IQ MD Basis Weight Reel Detailed Trend. Utilizing Google chart

Why Snowflake in our case

Concurrency

- Snowflake separates storage from the data computation.
- Different user groups can have their own virtual warehouses and not get affected by other user groups queries
- Data integration to database can also be separated to different virtual warehouses (even to the same table), so the integration won't affect the consumption.

Performance

- Snowflake is the fastest storage tested in the most important test cases.
- For some use cases there are faster solutions in the market, but Snowflake was the best fit for us in terms of overall performance

Support for heterogeneous data

- Snowflake implements a "schema-on-read" functionality allowing semi-structured data such as JSON, XML, and AVRO to be loaded directly into a traditional relational table.
- The semi-structured data can be queried using SQL without worrying about the order in which objects appear.

Scalability

- Possibility of auto-scaling, multi-cluster warehousing to seamlessly increase compute resources during peak load.
- Data remains fully accessible during scaling.
- No need for separate platform for hot / warm / cold data

Cost savings

- No need for additional / separate clusters for loading and reading
- No need for additional / separate clusters for dev and test
- No need for engineering resources to tune performance
- Storage costs only $25 / TB / month

Valmet

Valmet
FORWARD