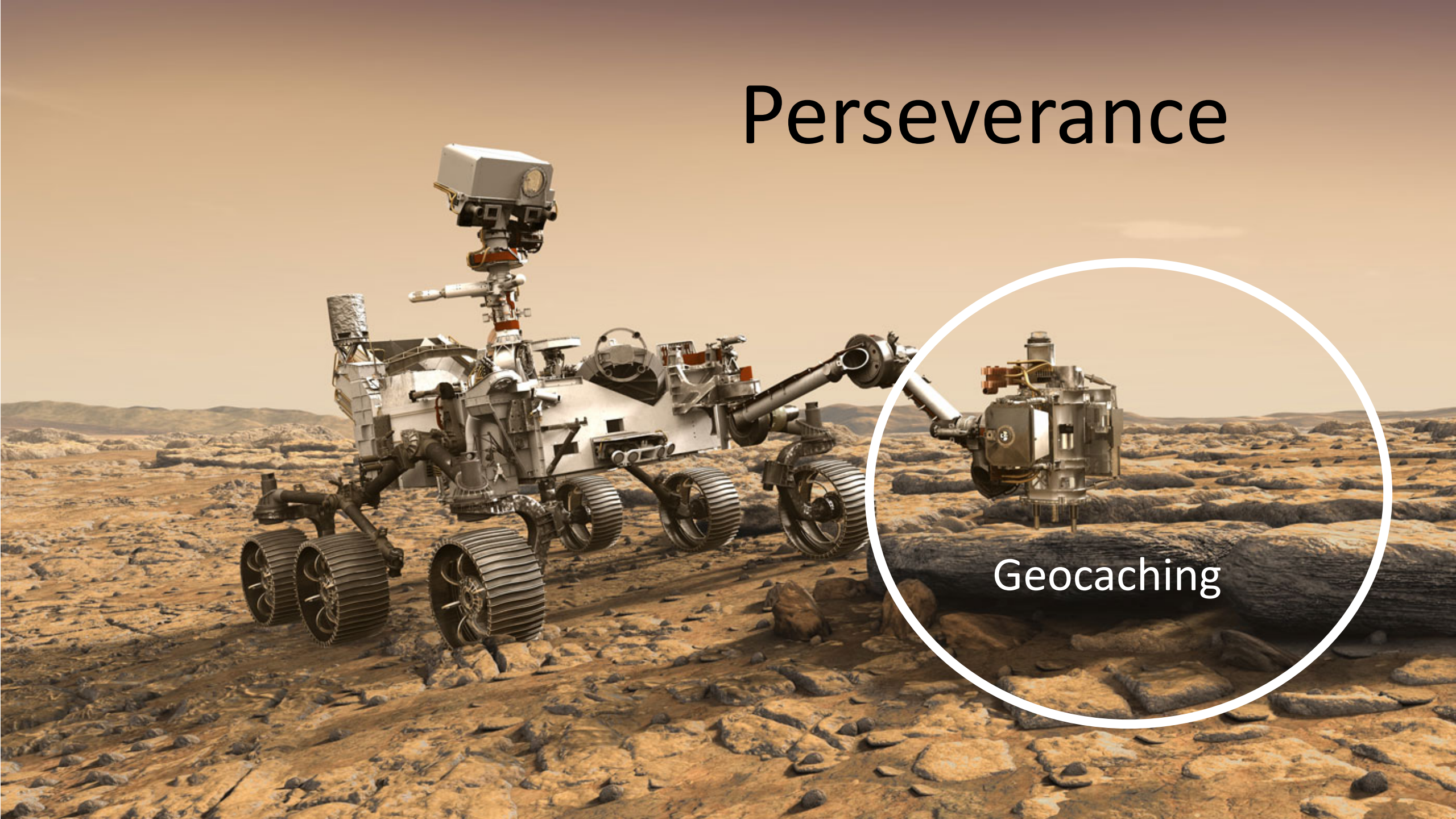
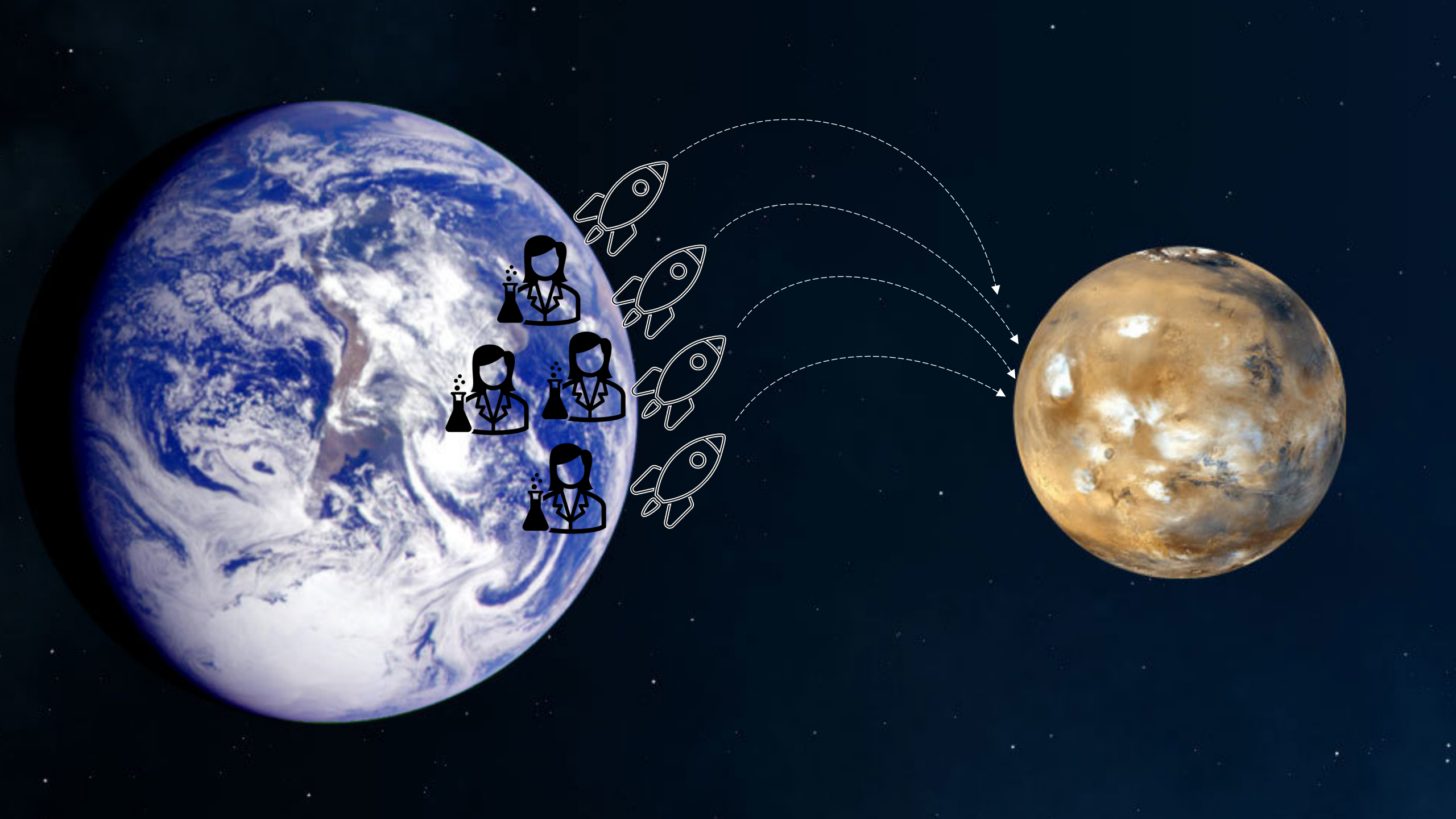
Production Experiences from Computation Reuse At Microsoft

Alekh Jindal, Shi Qiao, Hiren Patel,
Abhishek Roy, Jyoti Leeka, Brandon Haynes

Perseverance





Microsoft

Shared Data

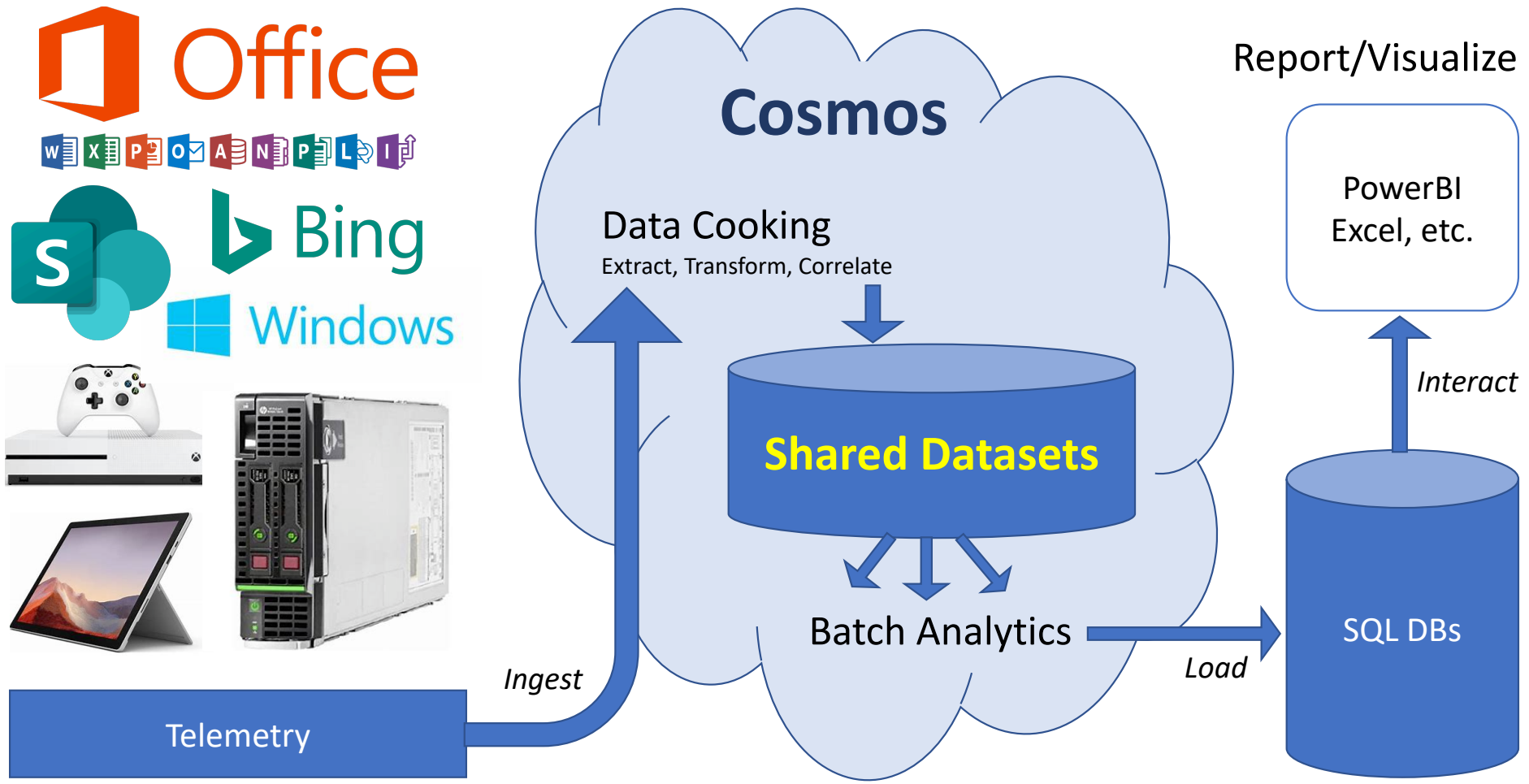
67 Million jobs
4.3 Billion sub-computations
2.5 Thousand users
776 Clusters

75% **sub-computations overlap!**

Cosmos Big Data Platform

Enterprise Data Pattern

Data Cooking



Shared Data Analytics

Average sales per customer in Asia

```
SELECT CustomerId, AVG(Price*Quantity)
FROM Sales
JOIN Customer
WHERE MktSegment='Asia'
GROUP BY CustomerId
```

Average discount per part brand in Asia

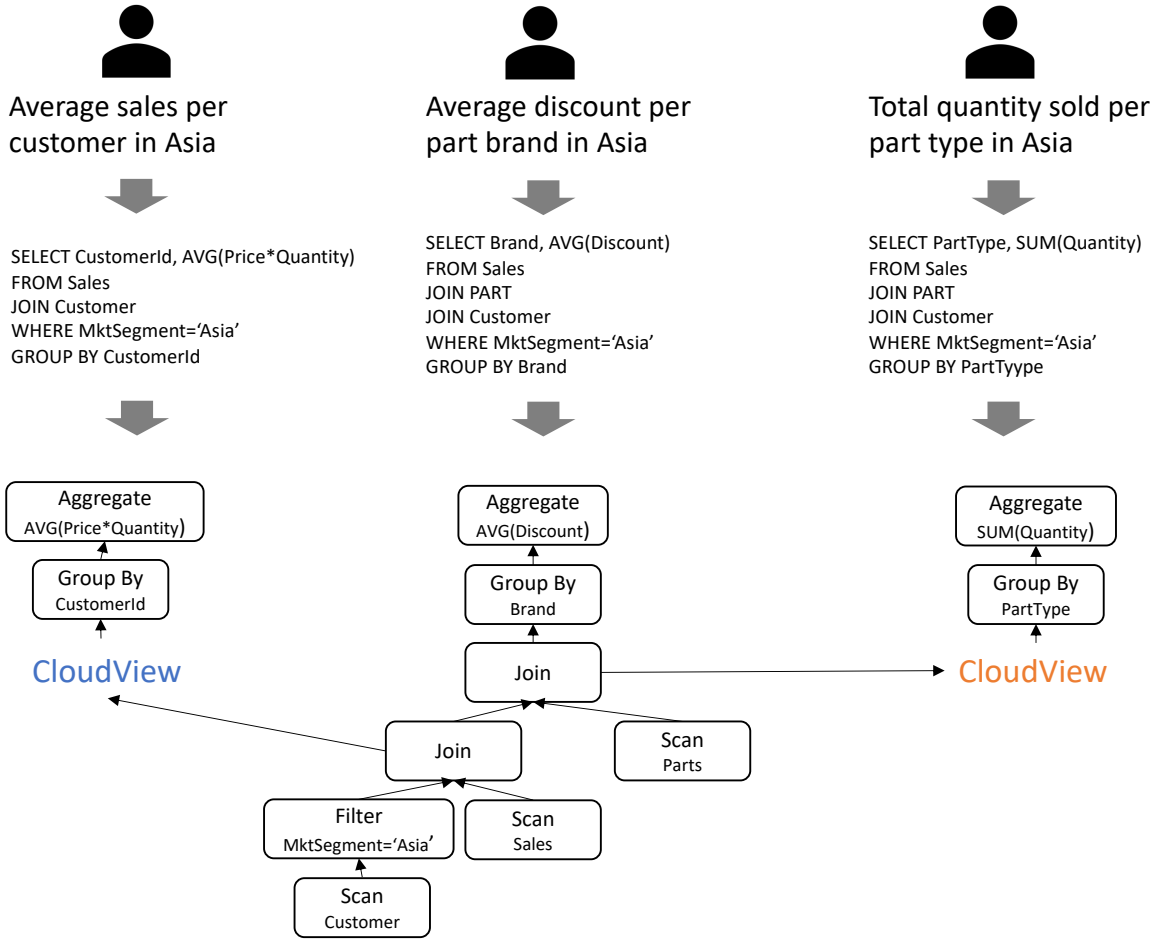
```
SELECT Brand, AVG(Discount)
FROM Sales
JOIN PART
JOIN Customer
WHERE MktSegment='Asia'
GROUP BY Brand
```

Total quantity sold per part type in Asia

```
SELECT PartType, SUM(Quantity)
FROM Sales
JOIN PART
JOIN Customer
WHERE MktSegment='Asia'
GROUP BY PartTyype
```

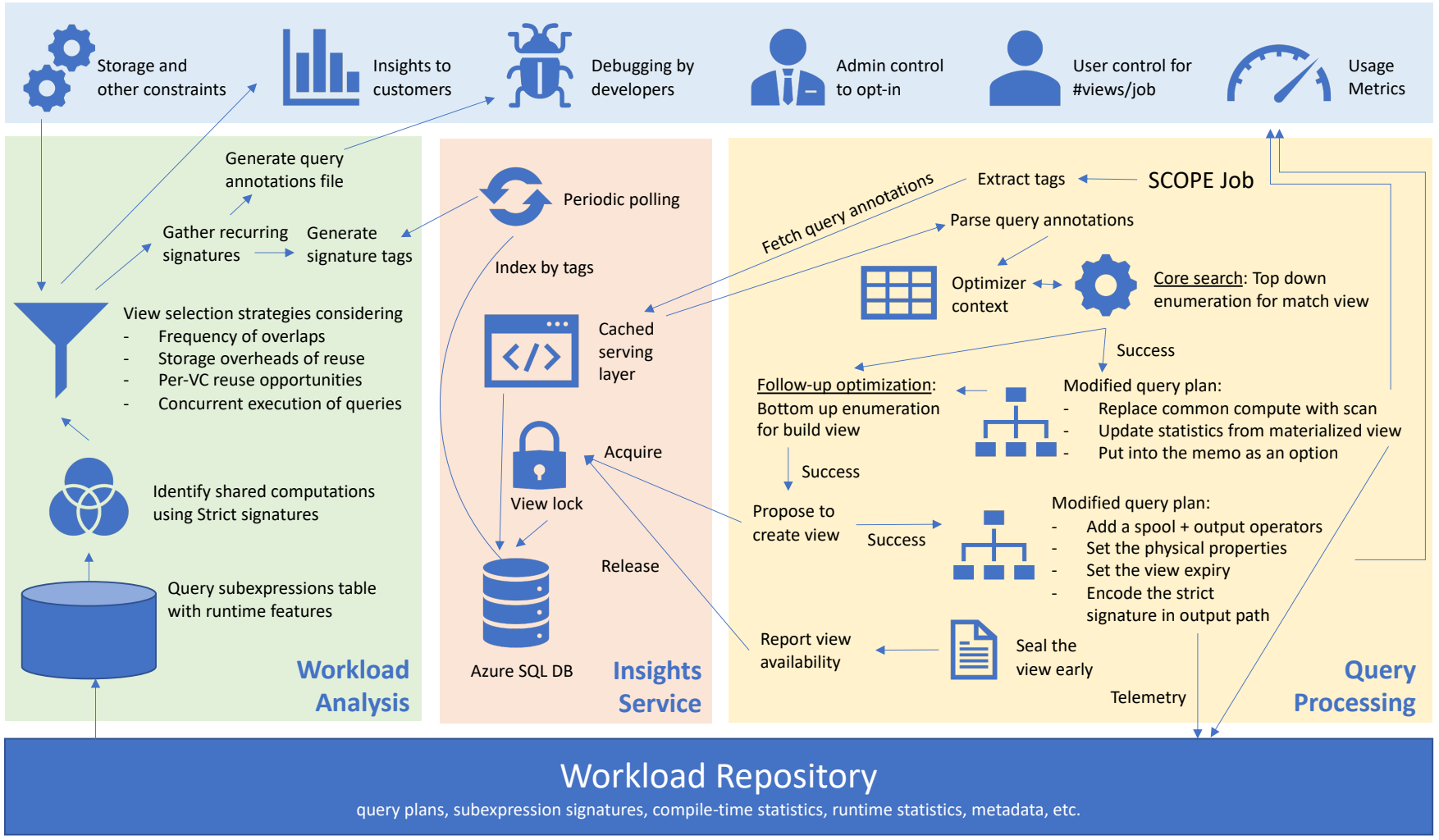
>50% datasets shared
>10% datasets shared by >16 consumers

CloudViews : Augmented Data Cooking



- Design choices
  - Preserving query boundaries
  - Online materialization
  - Just-in-time views
  - Accurate cost estimates
  - Scalable view selection
  - Lightweight view matching
- Limitations
  - Exact match only
  - Concurrent queries
  - Not maintained
  - No DDL in the catalog
  - User expectations

End to end Architecture



- Massive cloud workloads
- Scalable workload analysis
- Feedback & coordination service
- Automatic materialize and reuse
- Multiple personas and control

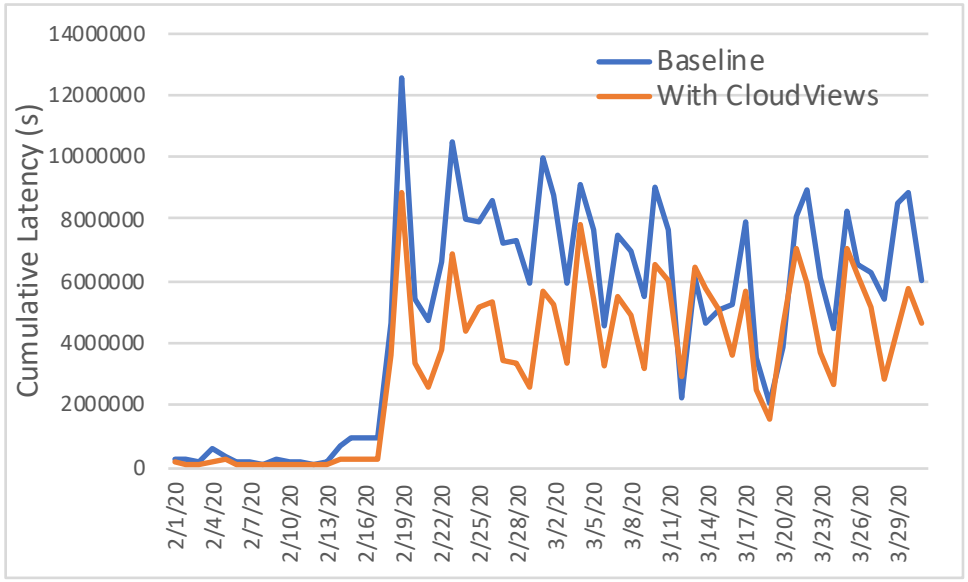
Production Impact

Impact Summary

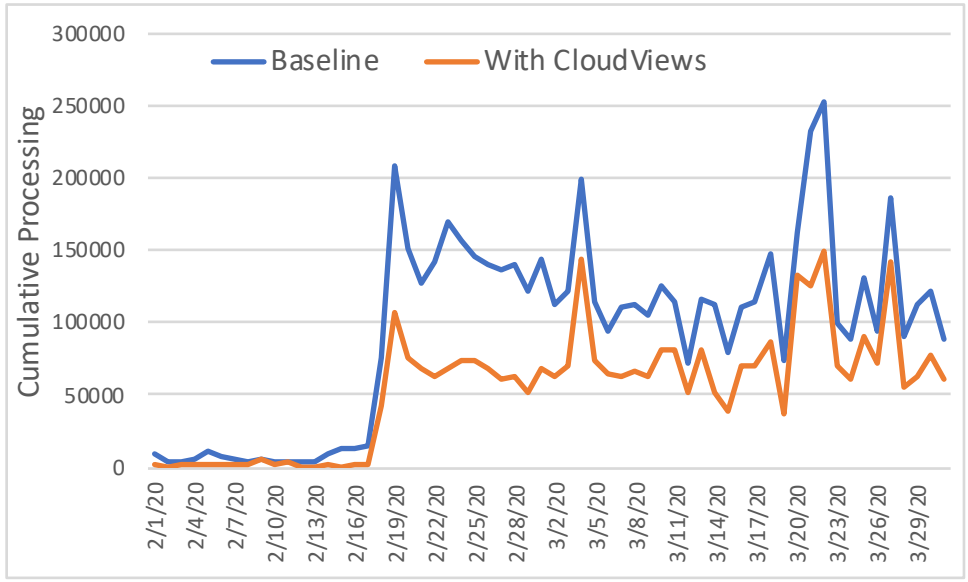
Feb-March 2020

Metric	Value
# Jobs	257,068
# Pipelines	619
# Virtual Clusters	21
# SCOPE Runtime Versions	12
# Views Created	58,060
# Views Reused	344,966
% latency improvement	33.97%
% processing time improvement	38.96%
% bonus processing time improvement	45.01%
% container count improvement	35.76%
% input size improvement	36.38%
% data read improvement	38.84%
% queuing length improvement	12.87%

Impact: user experience

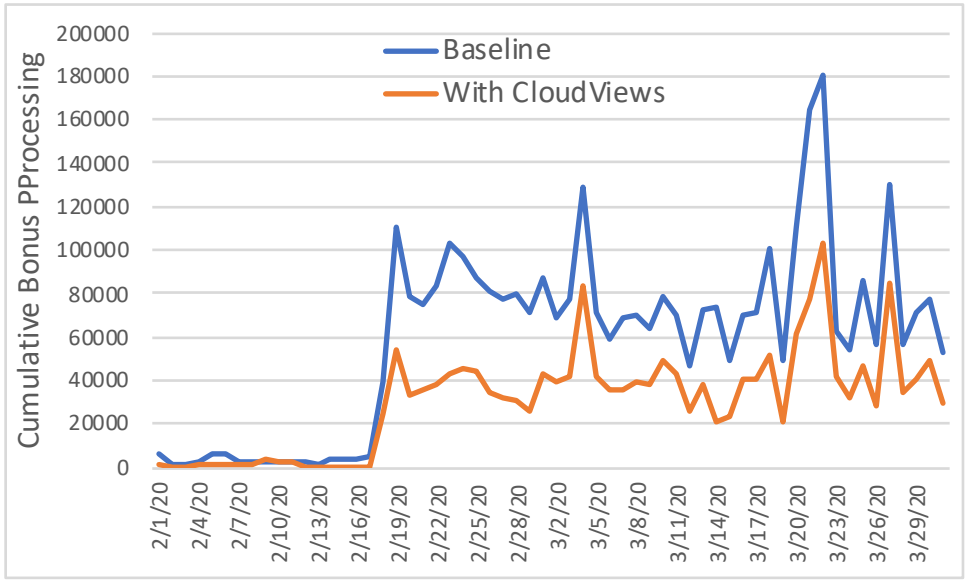


Latency: 15% median improvement

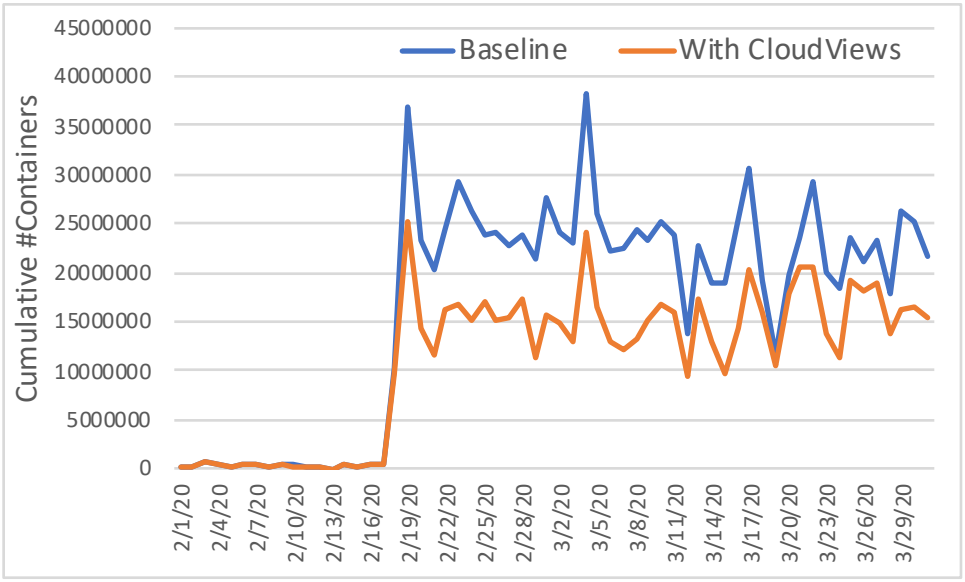


Processing time: 39% overall improvement

Impact: cluster efficiency

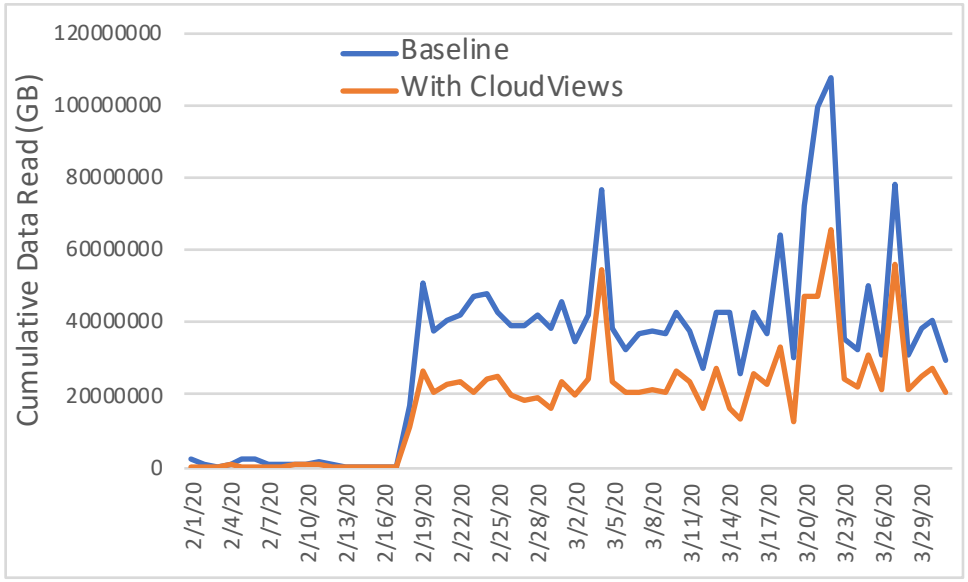


Opportunistic Processing: 45% overall reduction

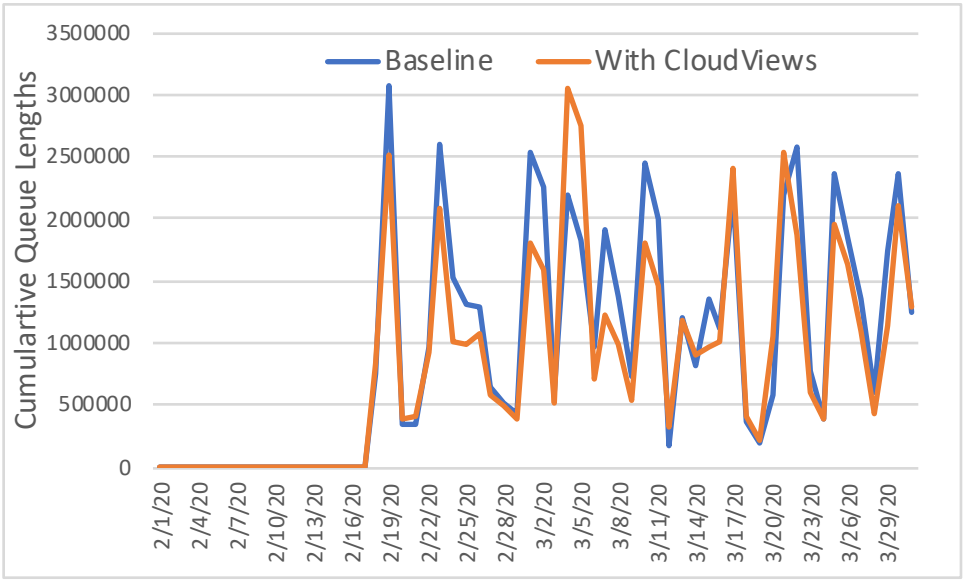


Containers: 36% overall reduction

Impact: non-obvious



Data read: 39% overall reduction



Queuing: 13% overall reduction

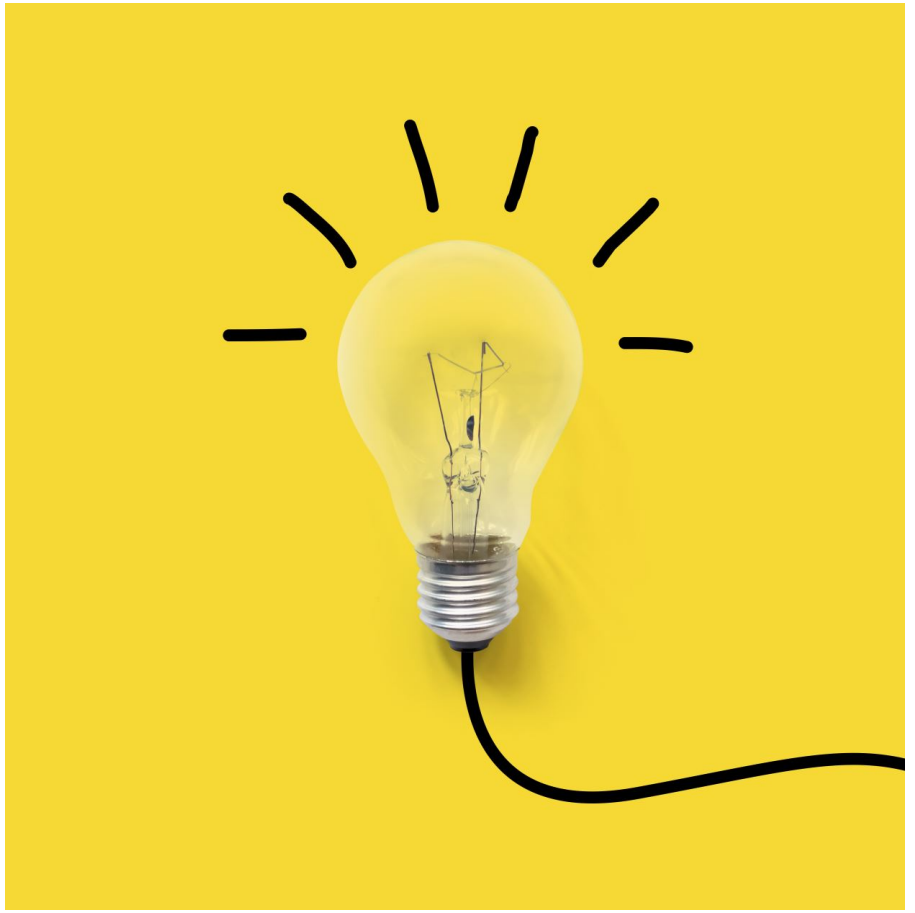
Lessons Learned

Operational Challenges

- Schedule-aware views
- Per-customer view selection
- Signature correctness
- Impact of changed signatures
- Other dependencies
- Handling GDPR requirements
- Opt-in vs Opt-out
- Multi-level control
- Measuring impact

Looking Back & Forth

- From research to production
- Towards broader workload optimization
- Generalized reuse
- Reuse in concurrent queries
- Reuse in other engines
- Other applications of reuse
  - Checkpointing
  - Pipeline optimization
  - Sampling
  - Bit-vector filtering



Microsoft

Summary

- Big data systems => democratized scalable processing
- Data engineers often have repeated computations in their pipelines
- Challenge: automatically detect and reuse repeated computations in serverless infrastructures
- Opportunity: large workloads to learn and improve upon
- We have built CloudViews for automatic computation reuse in Cosmos; valuable production lessons learned
- Hot topic in industry: Microsoft, Snowflake, BigQuery, Redshift, Alibaba, Oracle, others