

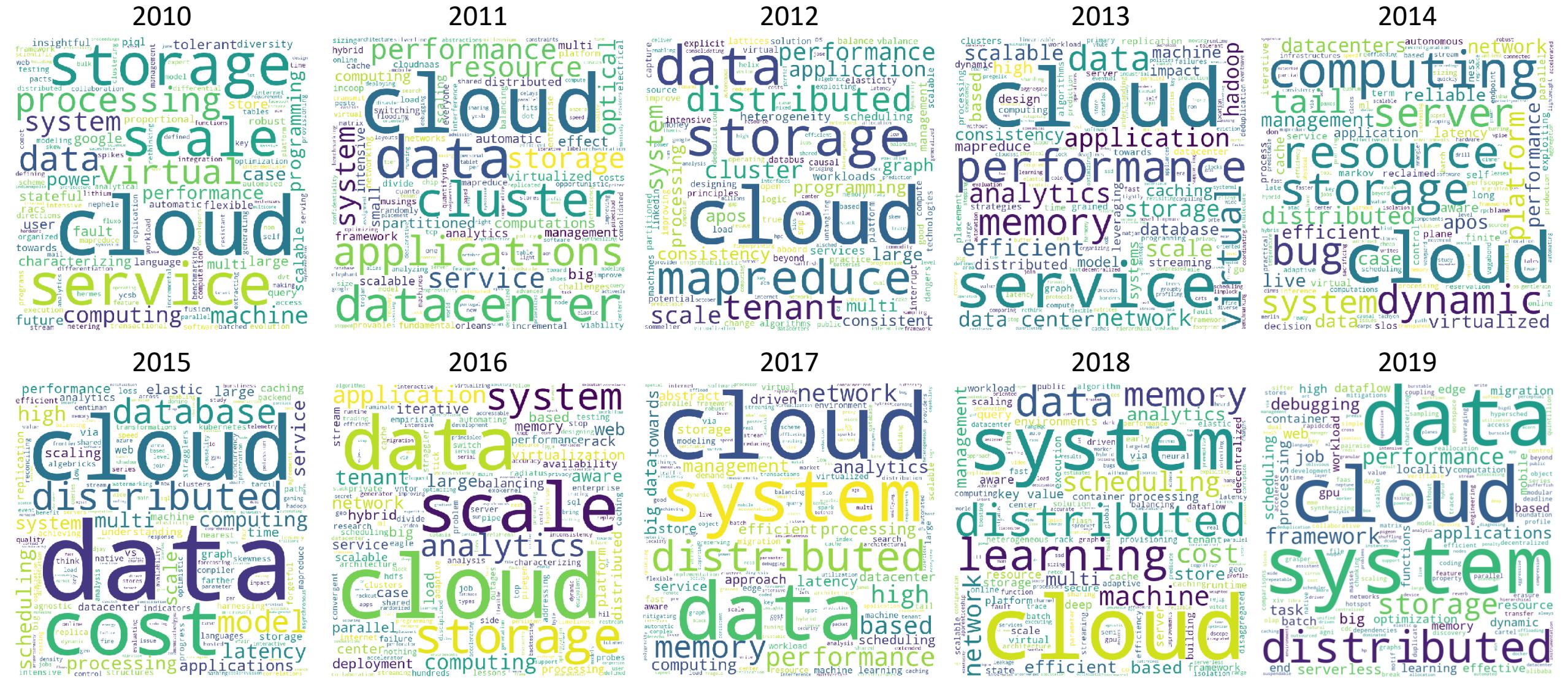


Data in the Cloud

Happy 10th ACM SoCC!

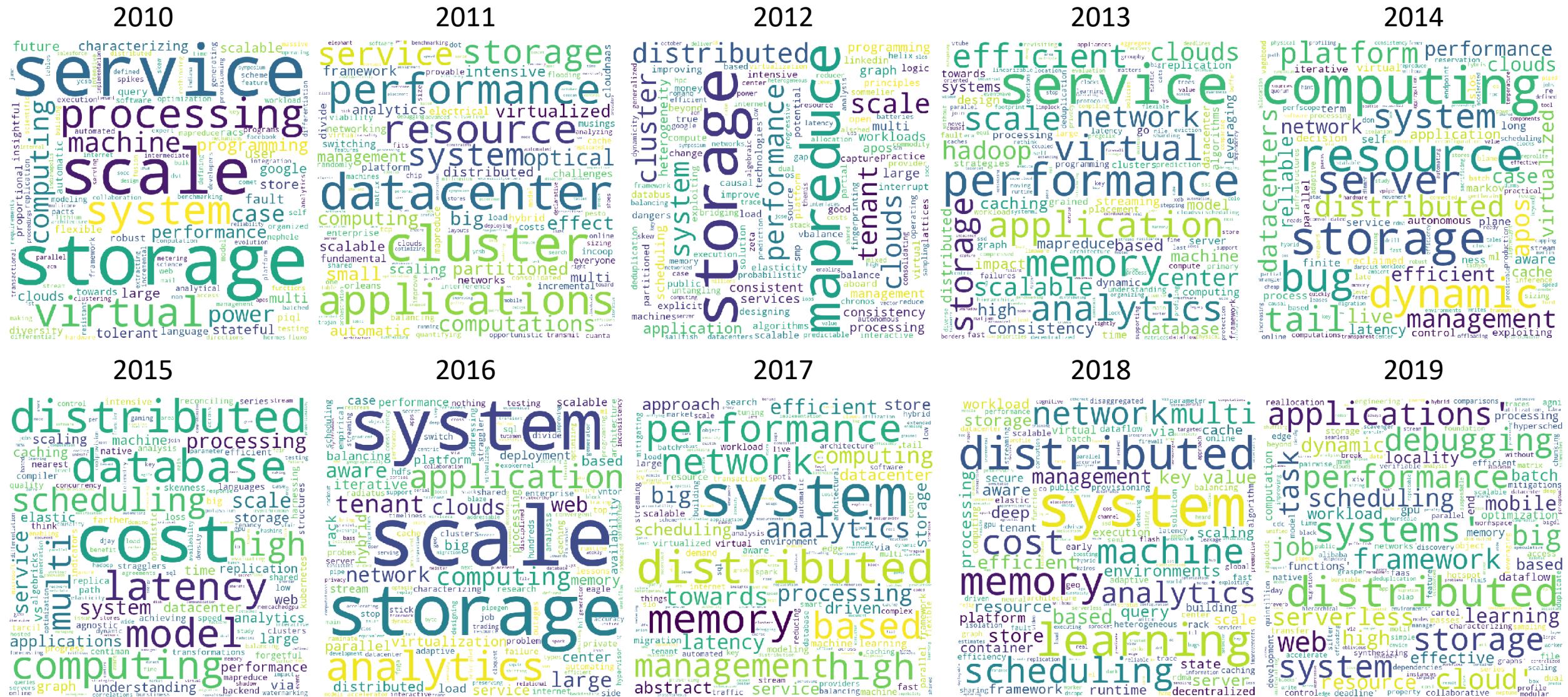
Raghu Ramakrishnan
CTO for Data, Technical Fellow

ACM SoCC Topics Over the Past 10 Years

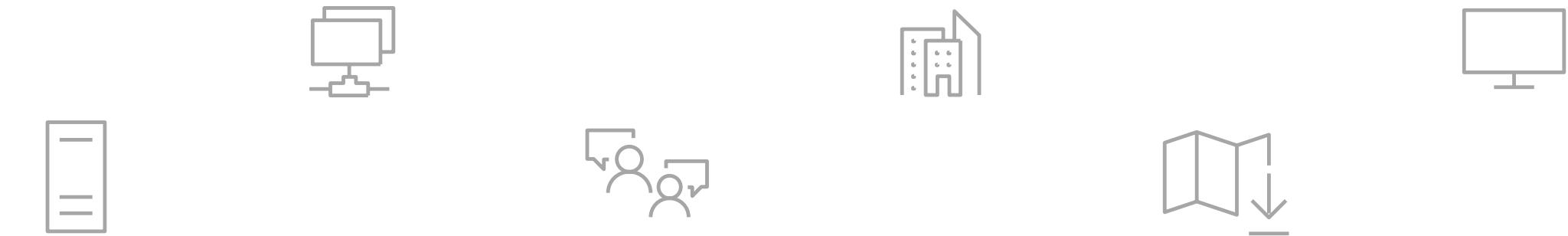


Word clouds courtesy Carlo Curino

ACM SoCC Topics After Filtering “data” and “cloud”



Word clouds courtesy Carlo Curino



Going Digital

1 million/hour
new devices coming
online by 2020

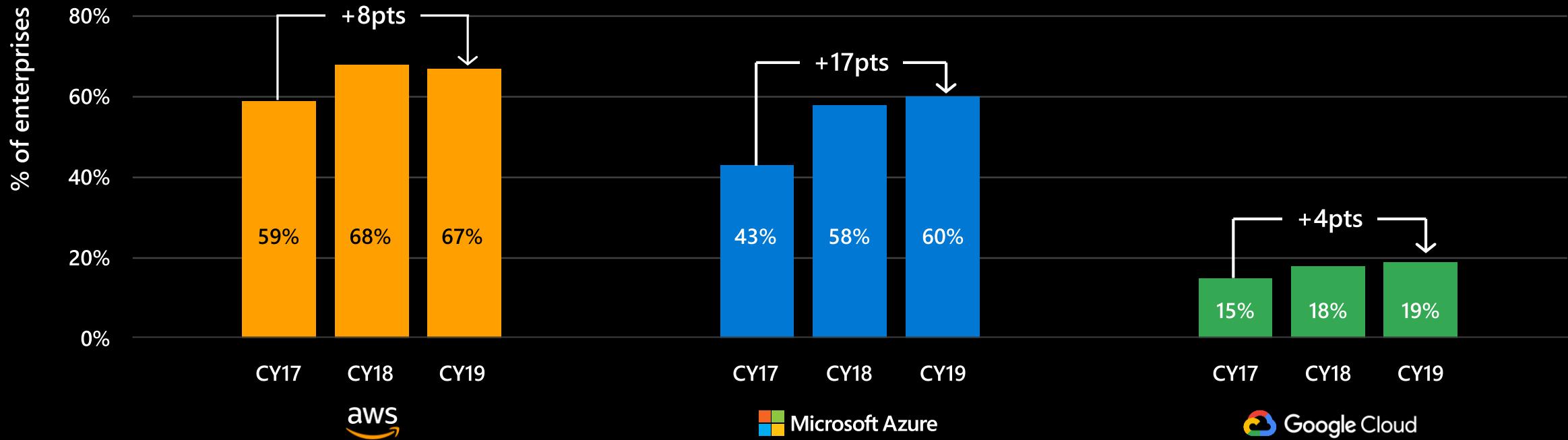
12 years average
age of S&P 500
corporations by 2020

60% computing
in the public cloud
by 2025



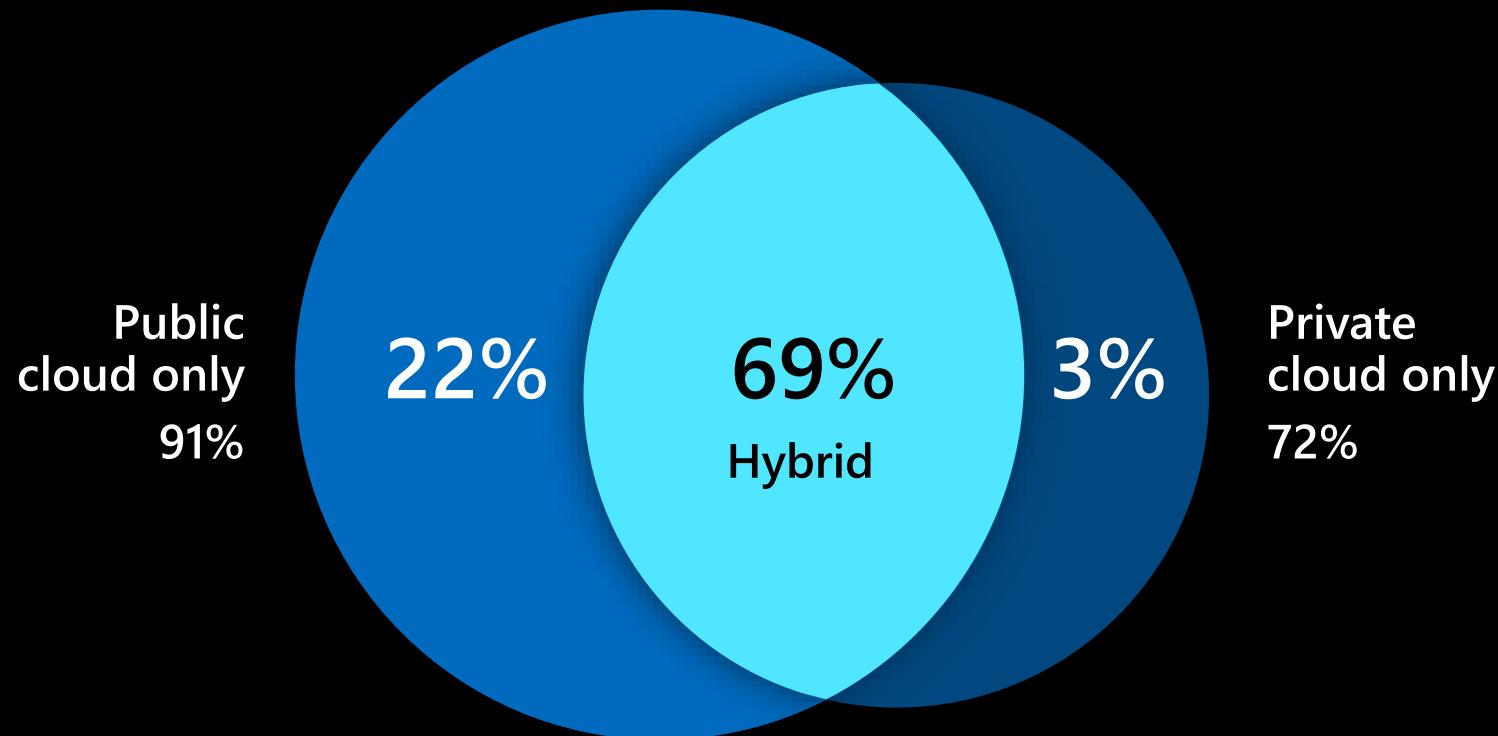
Cloud share of companies >1,000 employees

Rightscale Survey—Public Cloud Market Penetration



Source: RightScale 2019 State of the Cloud Report from Flexera

Hybrid is the prevalent strategy

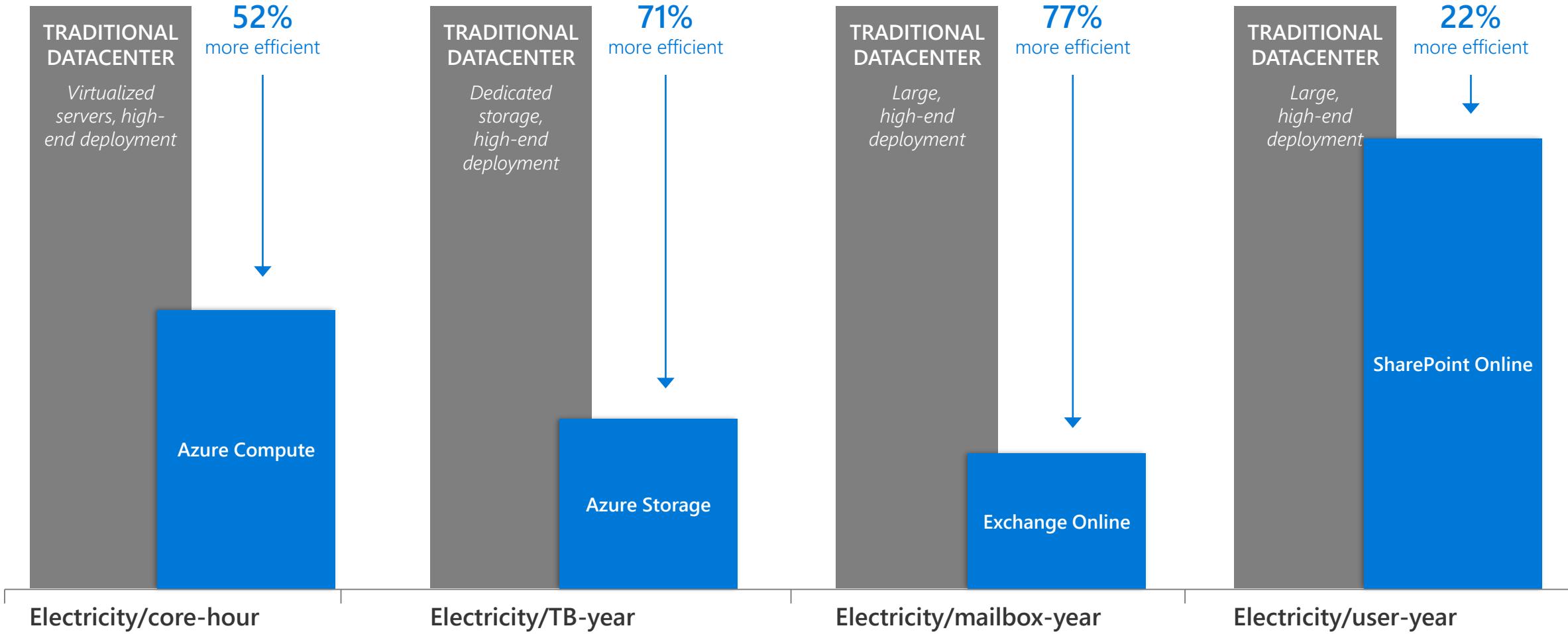


Source: RightScale 2019 State of the Cloud Report from Flexera

54
Azure regions



Carbon Footprint of Cloud Computing





AI in Operation & Optimization

IoT and big data platforms make it increasingly easy to optimize datacenters



IoT telemetry, analytics and ML optimization

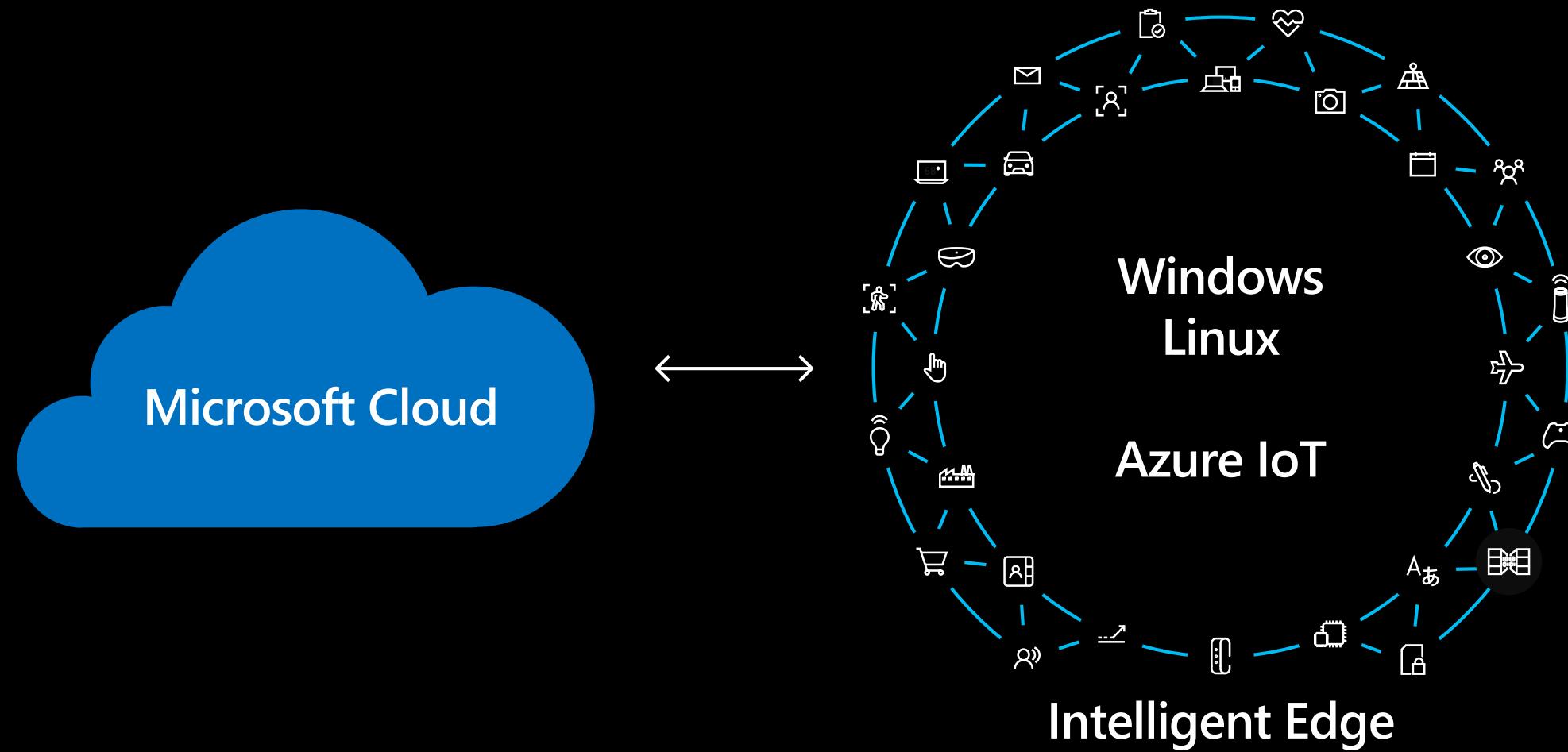


Predictive maintenance

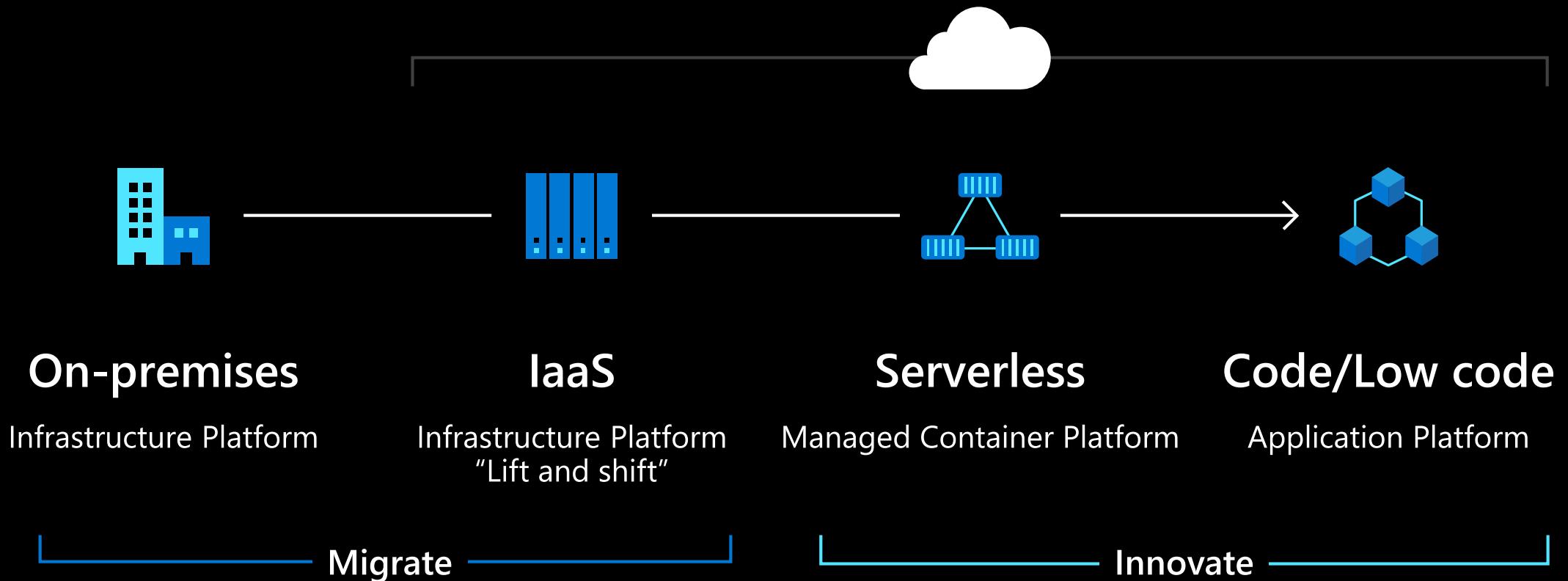


Capacity planning and workload placement

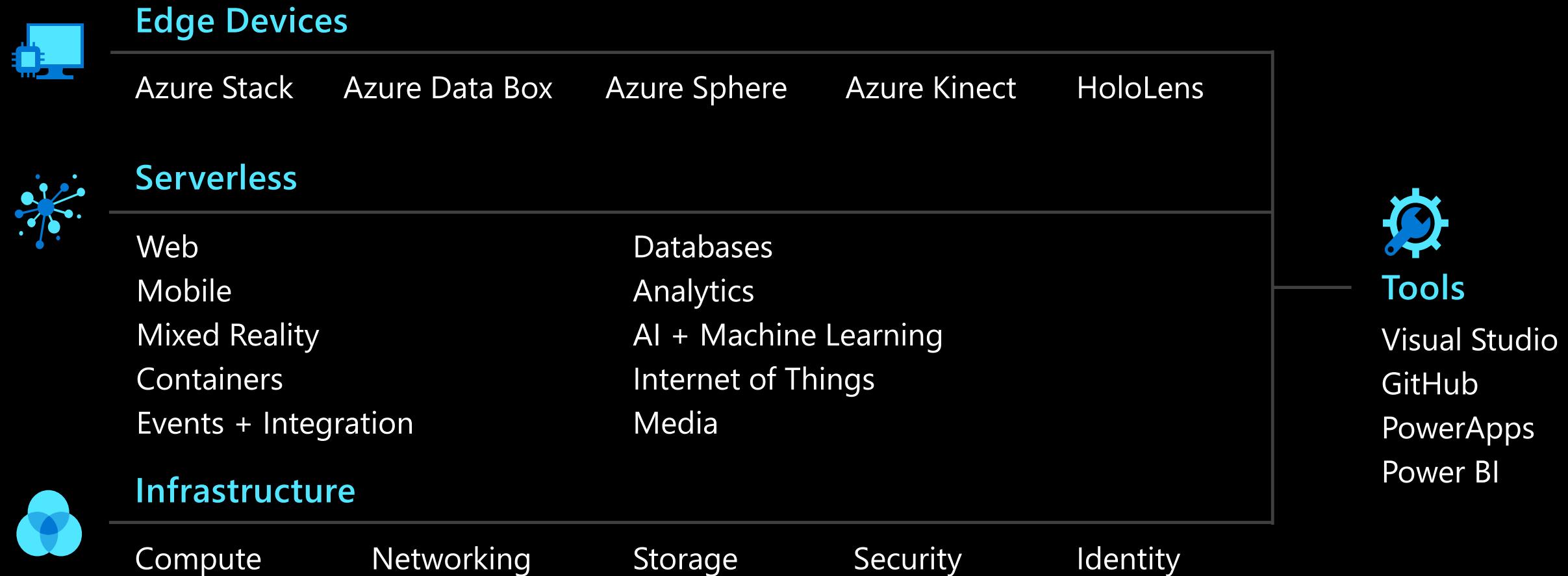




Journey to the cloud



A Public Cloud Ecosystem



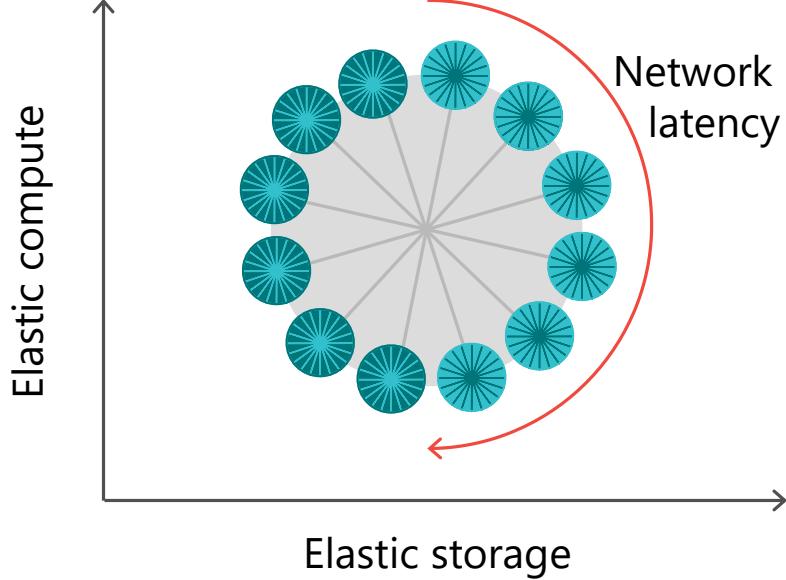
Ubiquitous Data

- What if you could see data describing everything you care about, all the time, in any detail, in “one place”?
- That “one place” is the continuum from cloud to edge
- All data interactions must be governed and tracked—security, privacy, compliance

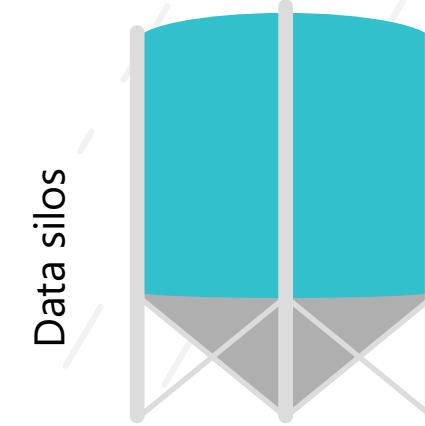
Cloud-Native Data

The Changing Landscape of Data

Cloud



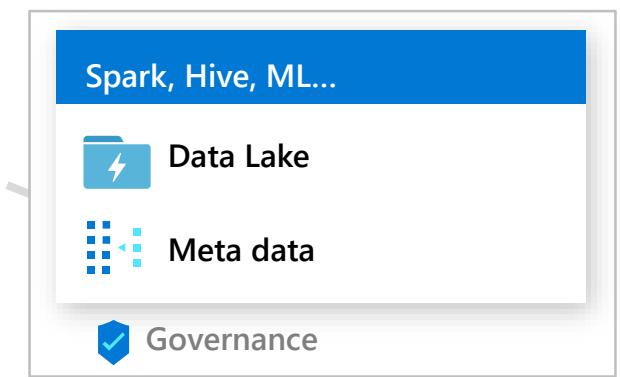
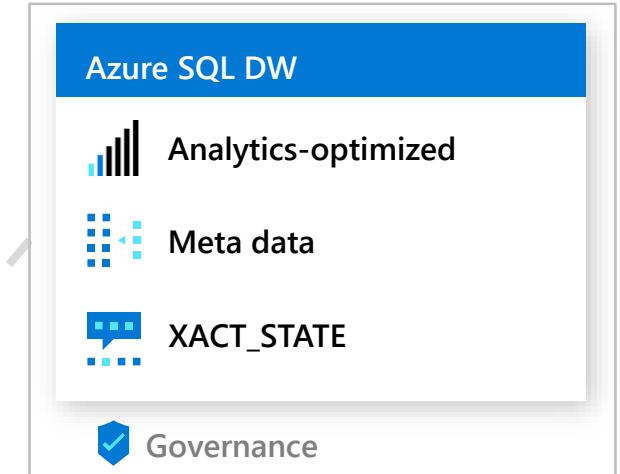
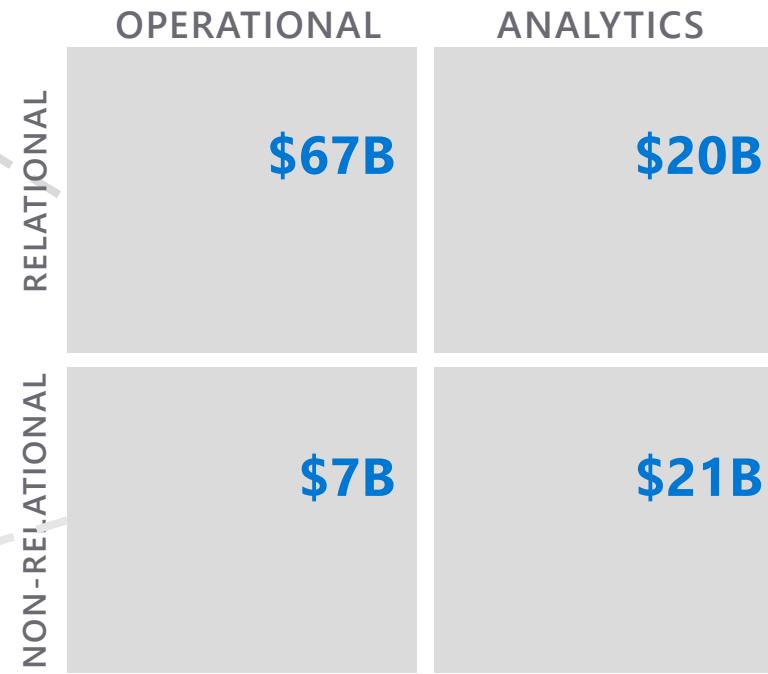
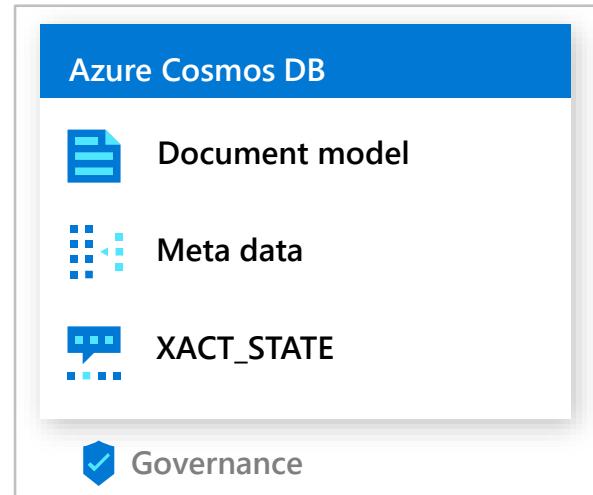
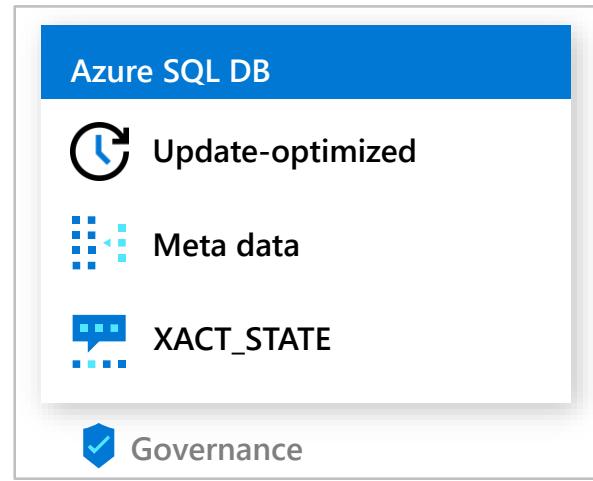
Analytic complexity



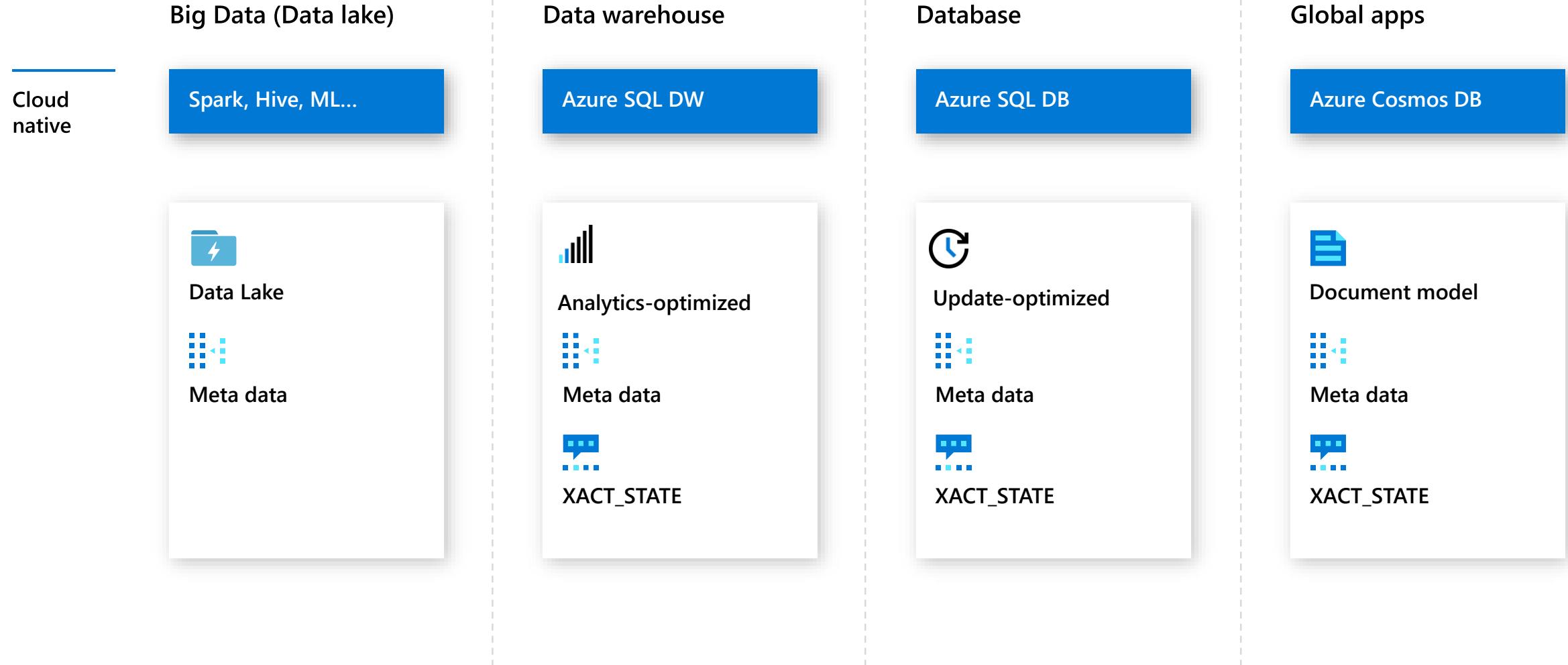
- Scale
- Heterogeneity
- Many engines
- Many workloads

OLTP challenges

- Size-of-data operations are slow
- Long recovery times are painful
- Independently scale storage vs. compute costs



Big Picture: Separation of Compute and State



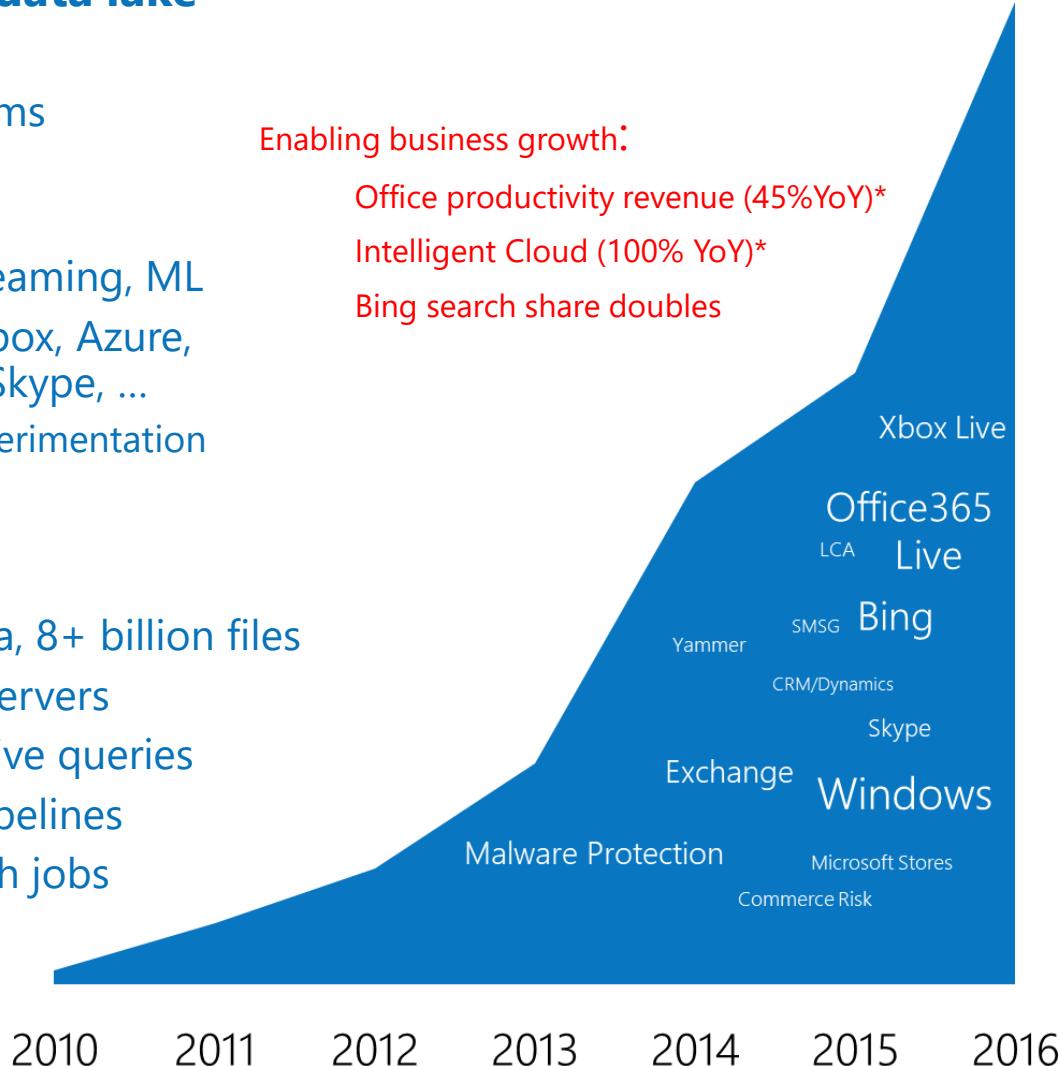
Microsoft's Internal Big Data Service

Microsoft's internal data lake

- A data lake for all teams @Microsoft
- Good developer tools
- Batch, Interactive, Streaming, ML
- Used across Office, Xbox, Azure, Windows, Ads, Bing, Skype, ...
- Production jobs and experimentation

By the numbers

- 9+ Exabytes of data, 8+ billion files
- 100Ks of physical servers
- Millions of interactive queries
- Huge streaming pipelines
- 100Ks of daily batch jobs
- 15K+ developers
- 300+ teams



Azure Data Lake Store

HDFS as a PaaS cloud service

- Microsoft's serverless Big Data platform
- **Fully aligned with Hadoop ecosystem** and standards, with full support for Hadoop tools and engines as well as unique Microsoft capabilities
- Migrated to ADLS
- **1P = 3P**

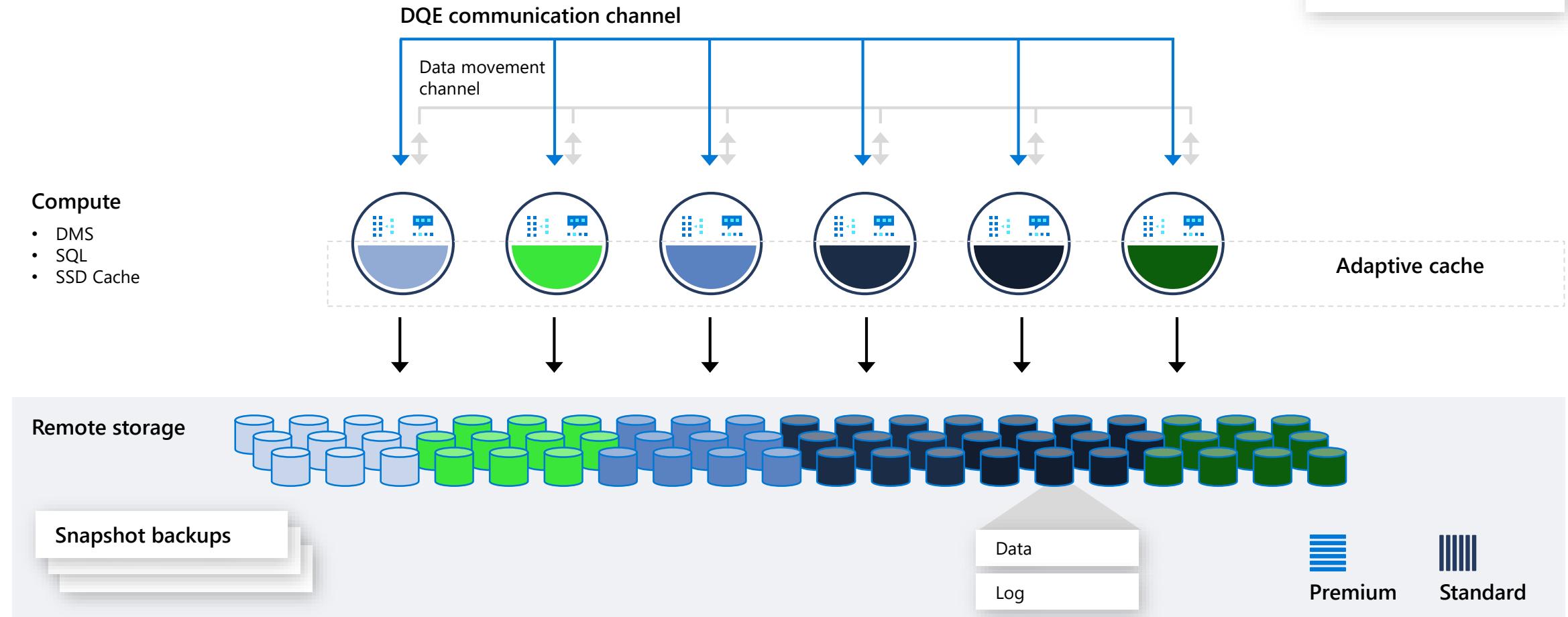
J. Zhou et. al., SCOPE: parallel databases meet MapReduce, VLDBJ 21(5)

R. Ramakrishnan et. Al., Azure Data Lake Store, SIGMOD 2017

Apache YARN Federation

Traditional MPP DW Architecture

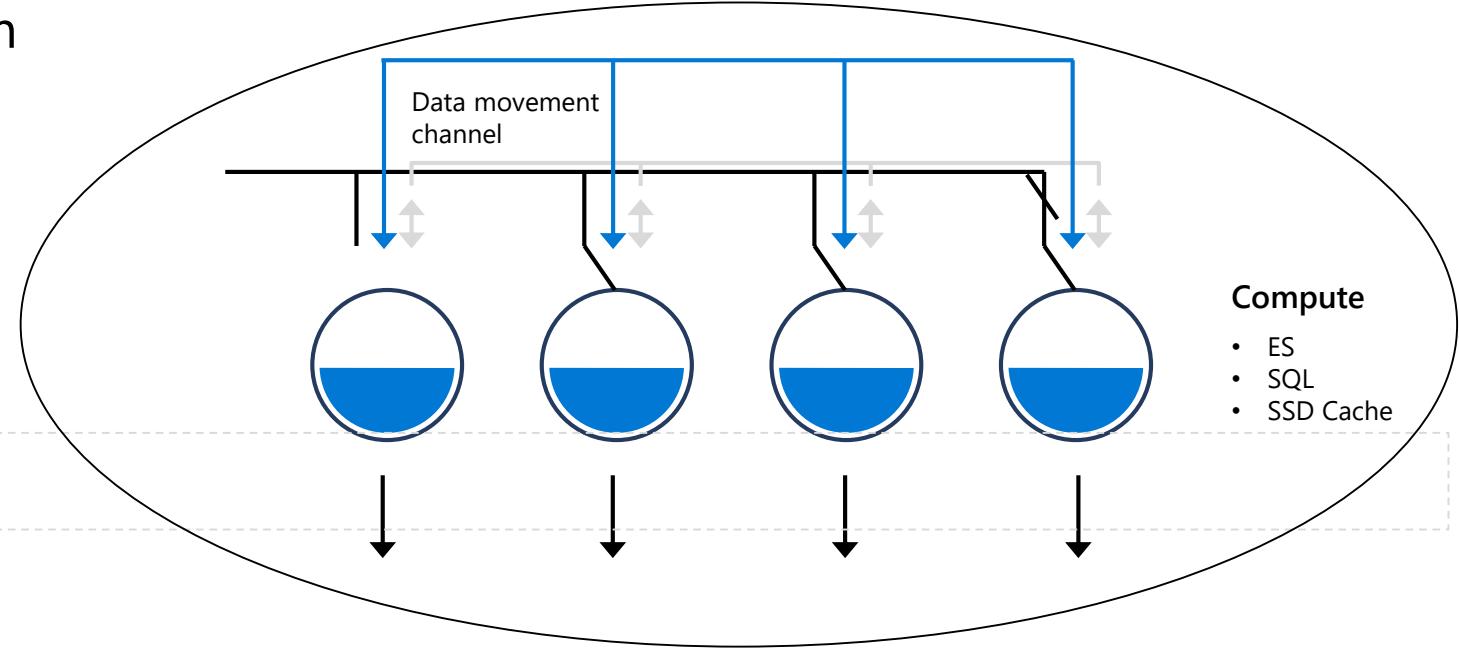
- Meta data
- Transactions



Cloud-Native Scale-Out, Data Heterogeneity

- Data and state separated from compute
- Fault-tolerant scale-out
- Online scaling
- Data heterogeneity

➤ Converge DW and Lake



Centralized services



Meta data



Transactions

Standard



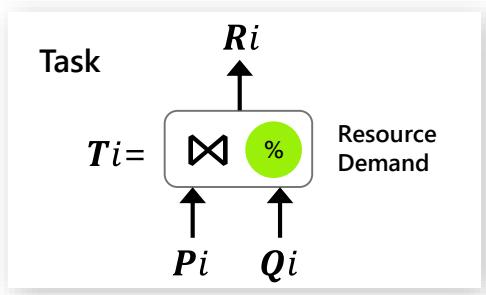
Remote storage

- Distribution-less

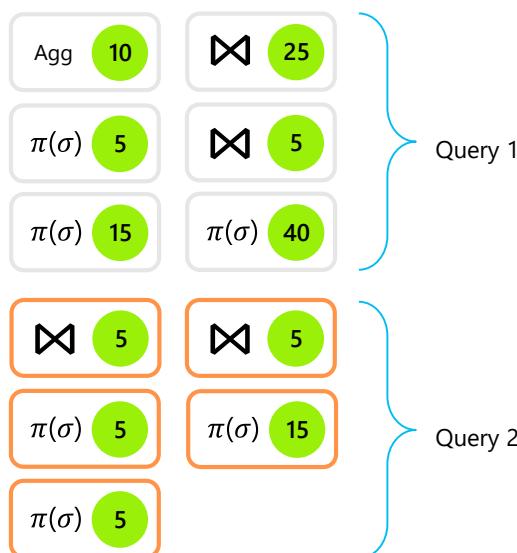
Columnar files

Polaris Concurrency – Workload Aware Scheduling

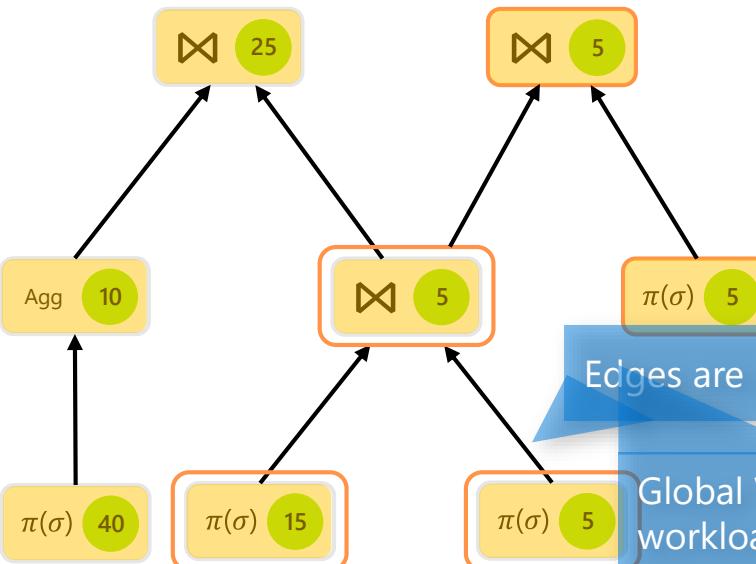
A next generation distributed query engine (blend massive scale batch QP with real-time streaming)



Workload Tasks



Workload Task Graph



Task-cost Driven Scheduling

Resource Aware Task Placement

State Machines:

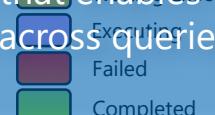
- Guarantees precedence constraints are satisfied
- Defines a formal model on how we recover from failures

State Machine Execution



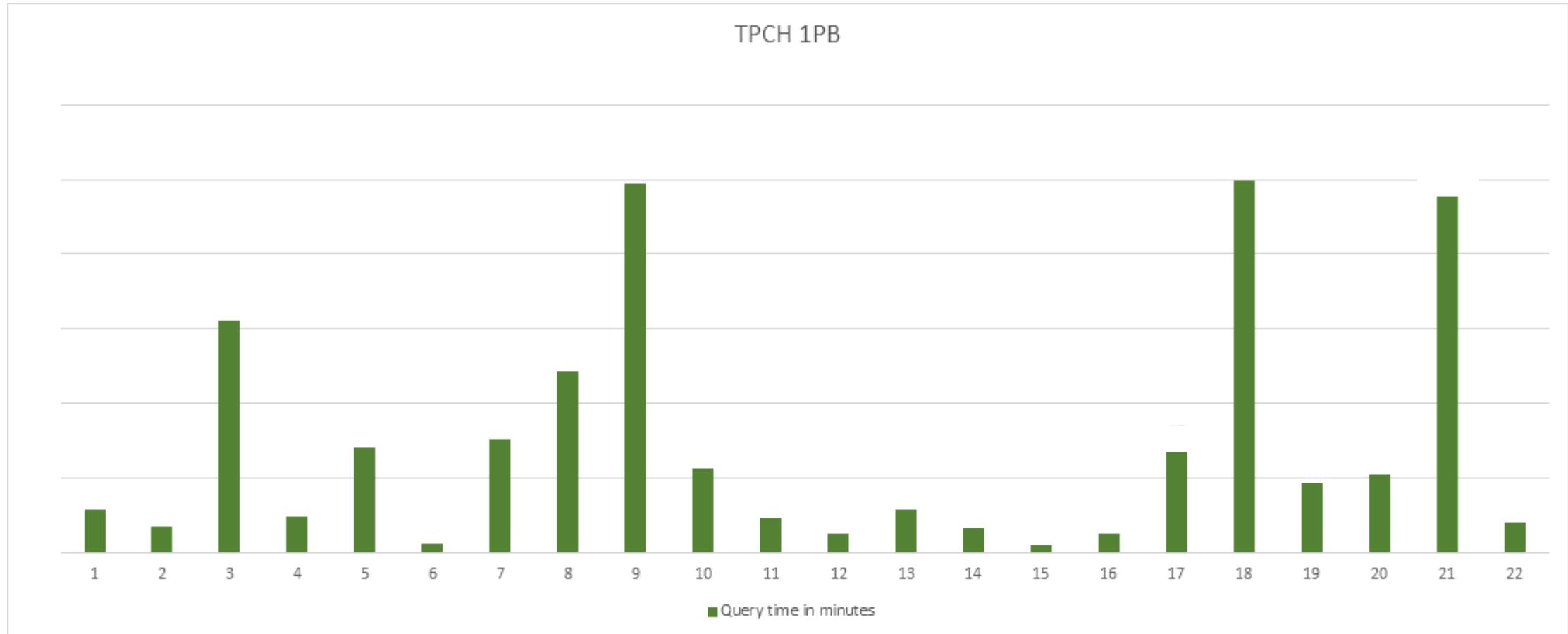
States

Global Workload Graph that enables for workload optimizations across queries



Scalability: All TPC-H Queries at 1PB Scale!

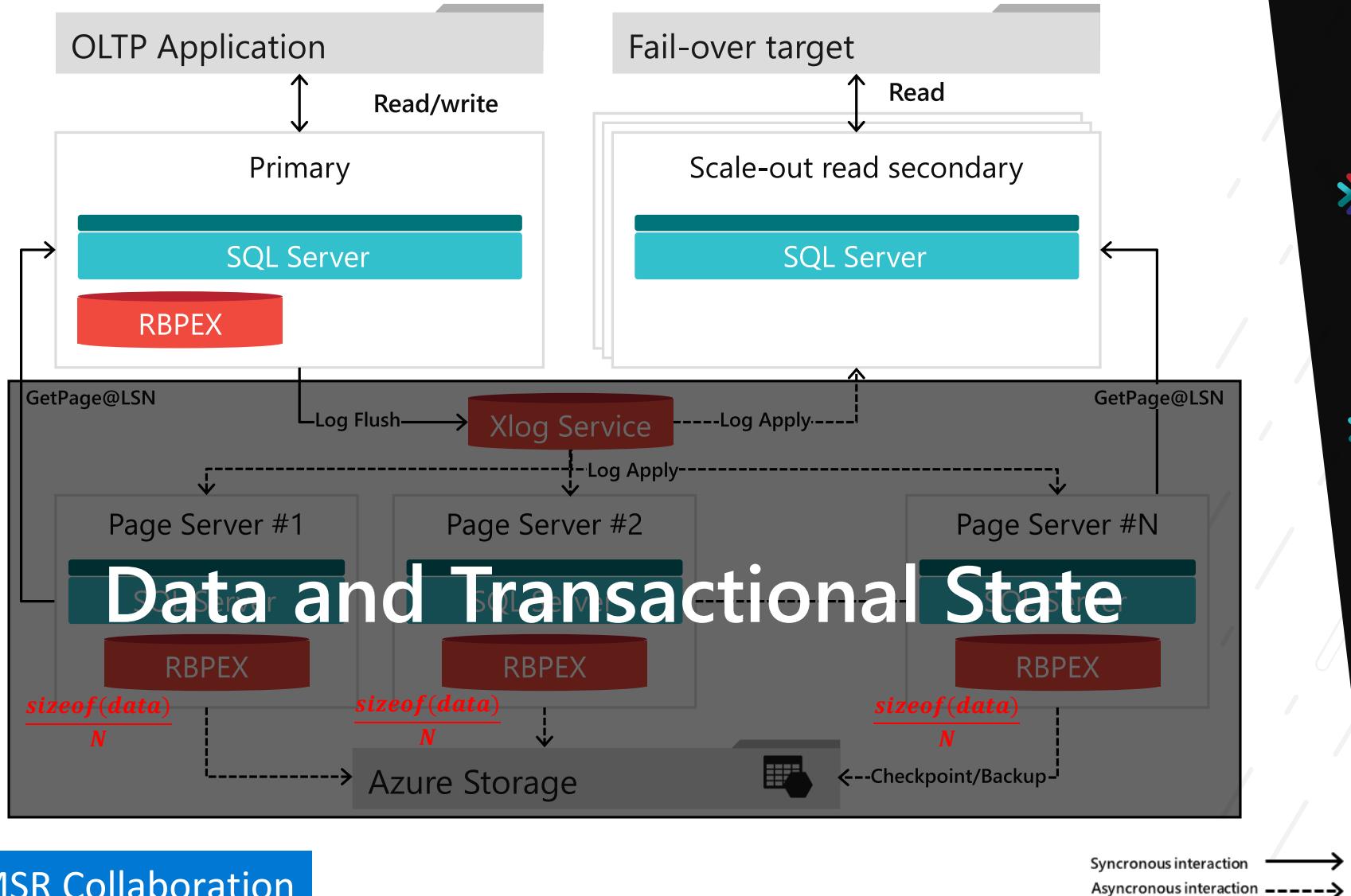
Elastic DQP – Unlimited Scale



<https://azure.microsoft.com/en-us/services/synapse-analytics/>

Socrates (Azure SQL DB HyperScale)

P. Antonopoulos, et. al., **Socrates: The New SQL Server in the Cloud**. ACM SIGMOD 2019



High level design choices

- Separate compute, storage, and log
- Partition at page server (low MTTR)
- Stateless compute with cache
- Primary compute orders xacts



Data flow

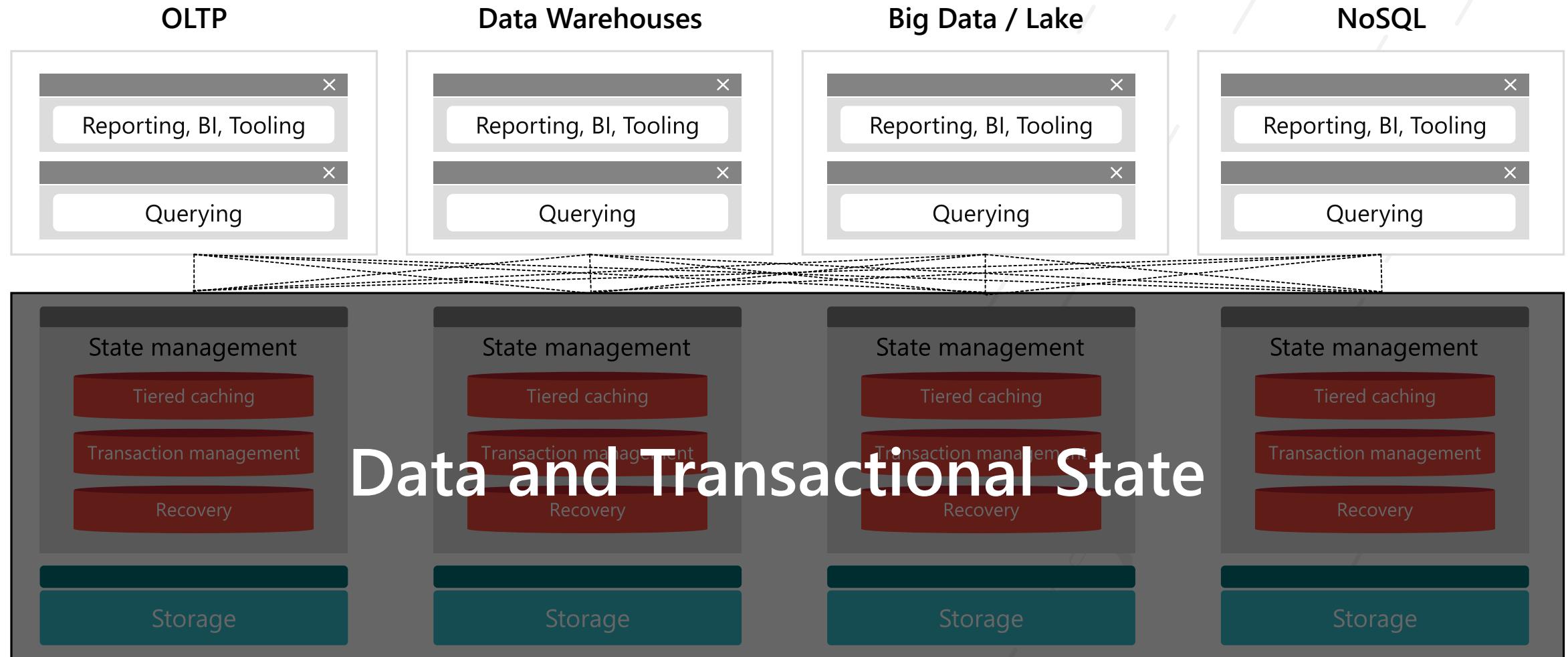
- Primary commits log
- Page servers apply log (for partition)
- Secondaries apply log (for cached pages)



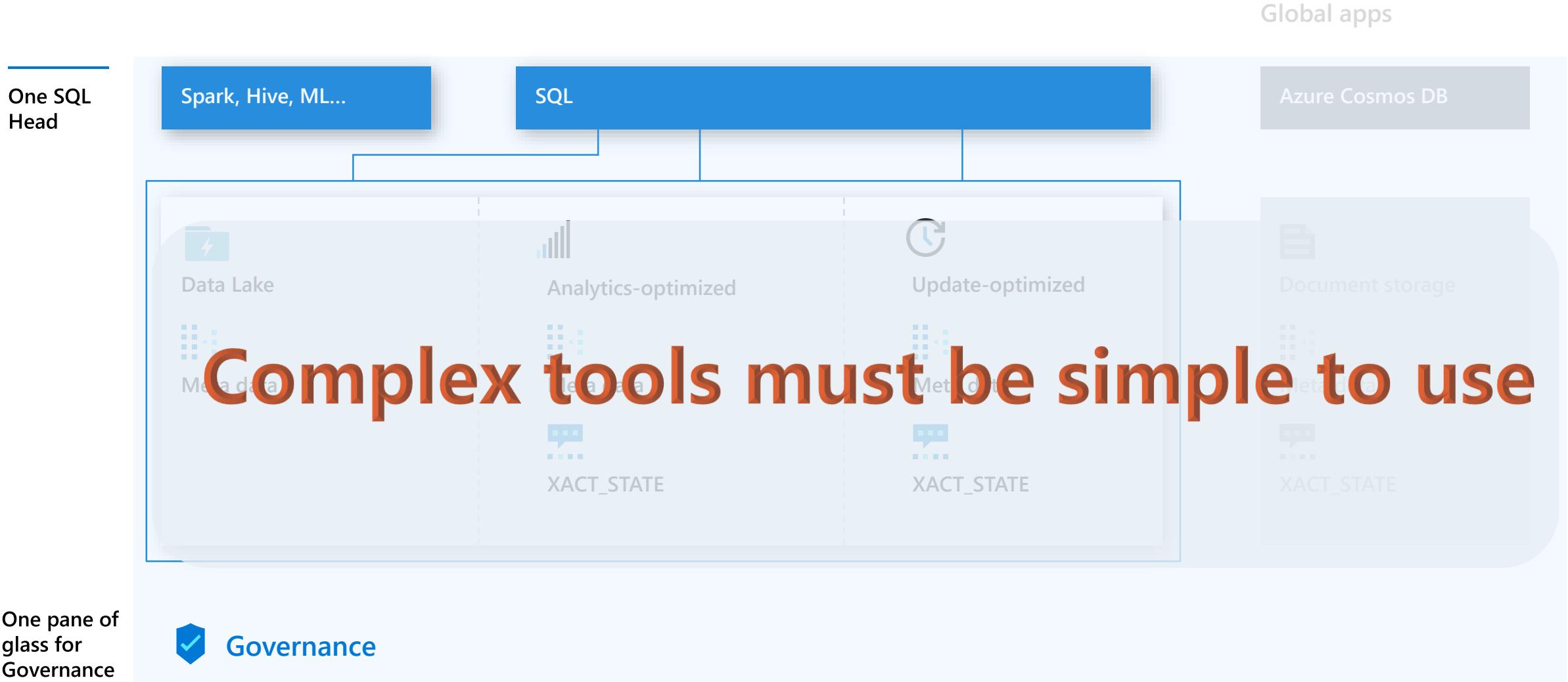
Advantages

- Constant time size-of-data operations (e.g., recovery)
- Flexible read replicas
- Cost vs. peak availability trade-offs

Cloud Native Data Architectures Emerging Across the Board

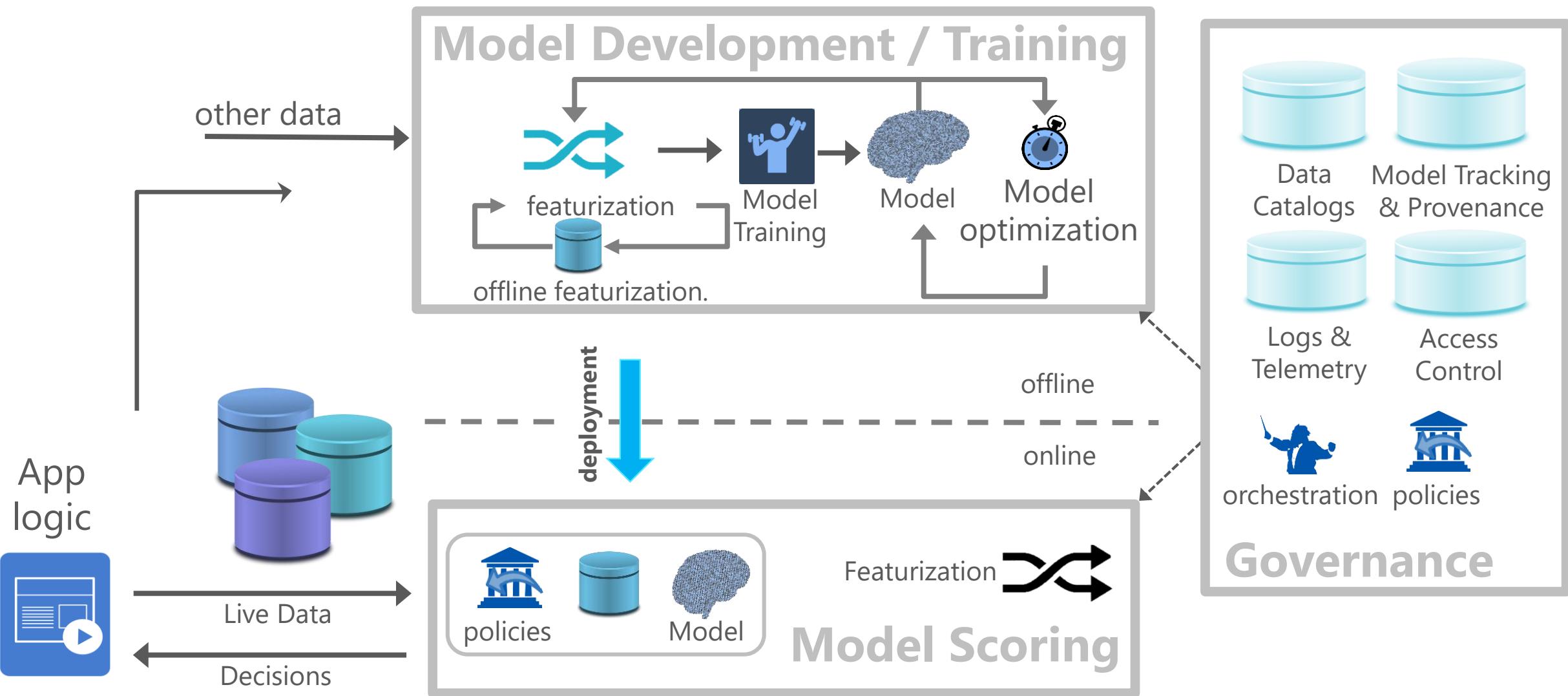


Unified Data Suite and Governance

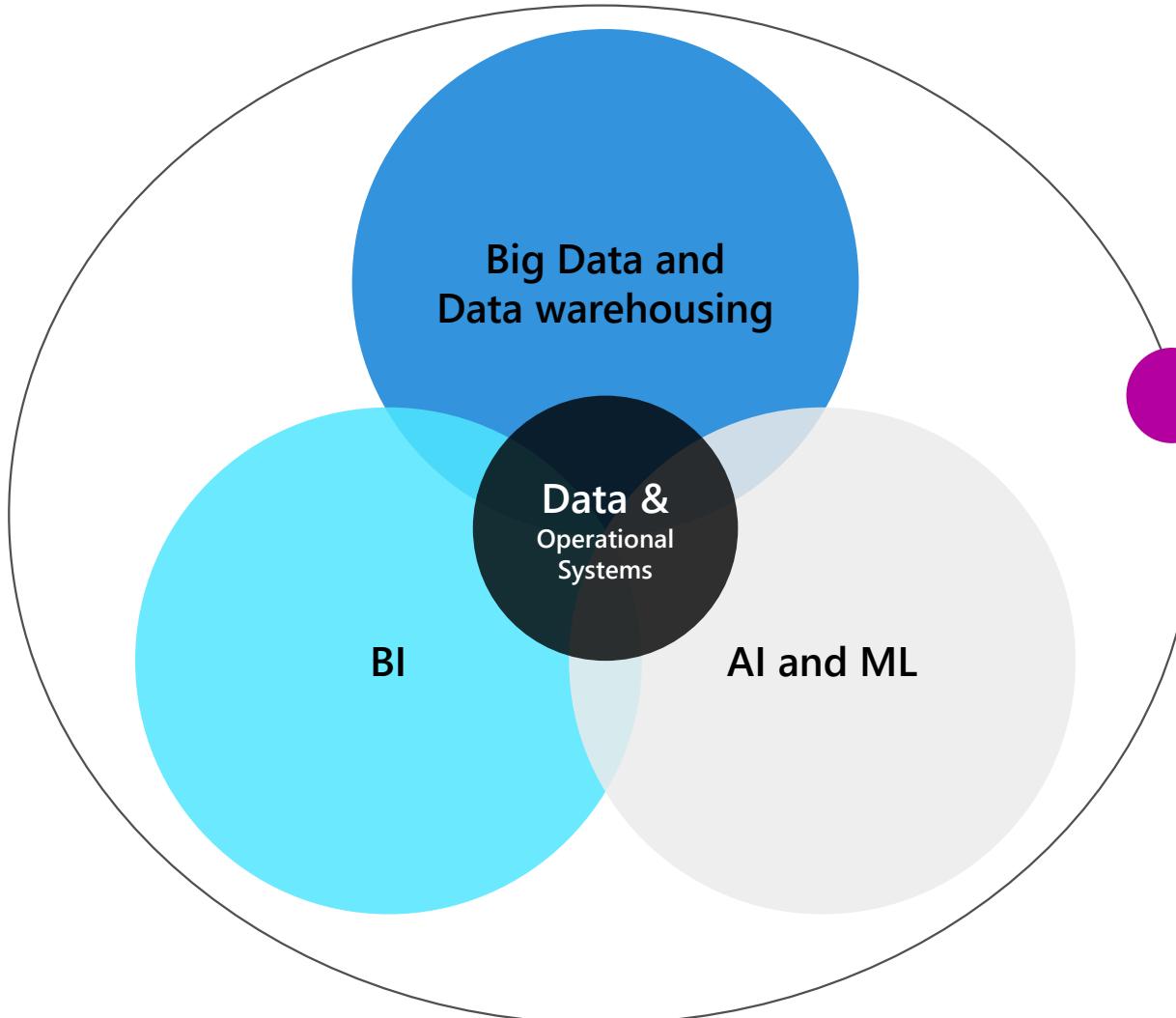


Big Picture: Must Simplify Usability and Governance

- Cloud
 - Elastic compute and storage is transformative
 - But compute-storage latency and bandwidth is key challenge
 - Edge blurs cloud/on-prem separation
- ML
 - An integral part of data processing, with a rapidly growing community of its own
- Implications for Data Management
 - Rethink what belongs in a “DBMS”—ML, data governance
 - Rethink data architectures from the ground up—OLTP/Analytics/HTAP



Unified Governance



A single pane of glass to...

Manage data lifecycle
(collect, clean, publish, discover, ...)

Ensure Data Quality & Correctness
Assess data compliance, privacy & protection

Author & manage data policy
(access, use, retention, location, sharing)

Across Cloud, Edge, On-Prem