The Hidden Cost of Data in Salesforce

Tahsin Zulkarnine | NTT Data

Reservation at Global Scale

- A global reservation platform
- Millions of travelers booking across multiple brands
- Data Cloud chosen to unify and activate reservation data
- Two months in **90%** of Annual Data

Budget Depleted!!!

Tahsin Zulkarnine

Sr. Digital Solution Architect,

NTT Data

tahsinz@gmail.com

tahsin.zulkarnine@nttdata.com

True North Dreamin

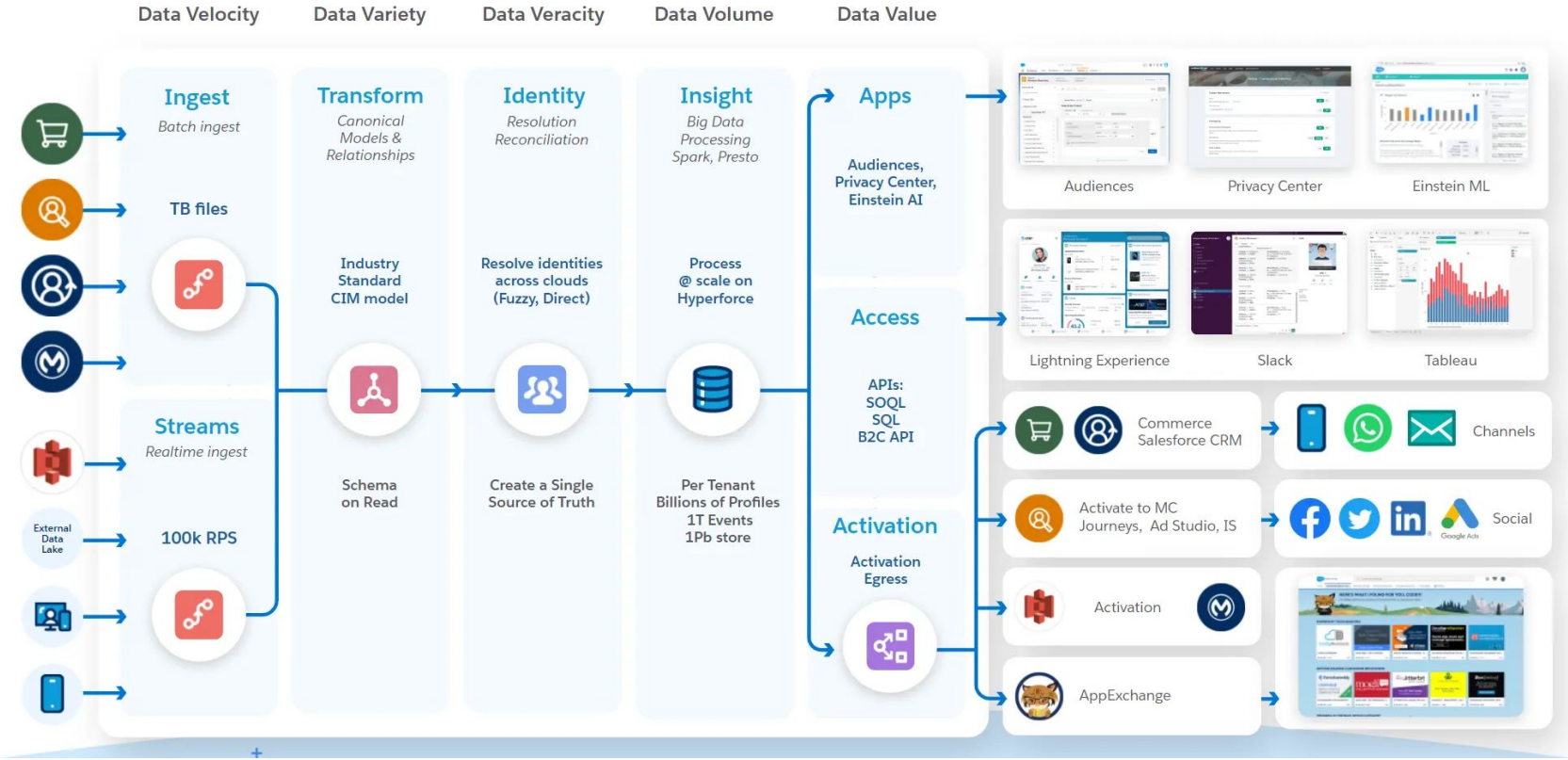
Agenda

01	/	How Data 360 Cost Works	2’
02	/	The Hidden Drivers of Spend	4’
03	/	Eliminating Unnecessary Cost	3’
04	/	Design Decision That drives Cost	5’
05	/	Governance to ROI	3’
06	/	Q&A	3’

True North Dreamin

CDP - Data drives value

Data 360



True North Dreamin

How Data 360 Consumption Is Calculated

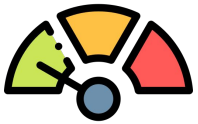
#	Operation	Units of Measure	Credits *	Description/Notes
1	Data Ingestion	1M rows processed	~ 2,000 credits	Ingesting structured records into Data Cloud
2	Data Transforms	1M rows processed	~ 400 credits	Transforming / harmonizing data within Data Cloud
3	Segment	1M rows processed	~ 20 credits	When building a segment, data is scanned / filtered
4	Activation	1M rows processed	~ 10 credits	Activating a segment (pushing data / making it available)
5	Query	1M rows processed	~ 2 credits	Query processing / analytics scans
6	Streaming	1M events	~ 800 credits	When using real-time or streaming pipelines
7	Calculated Insights (CI)	1M rows processed	10 credits	Executes aggregations or logic across Data Cloud Objects

*Credit Multiplier termed in public available rates

Estimated Cost Models

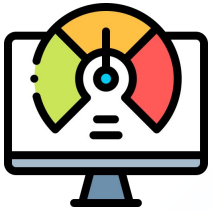
What Data 360 Can Cost at Scale

Low Consumption



- Batch-oriented workloads
- Limited real-time queries
- Simple segmentation
- Smaller customer datasets
- **Low six figures annually**

Medium Consumption

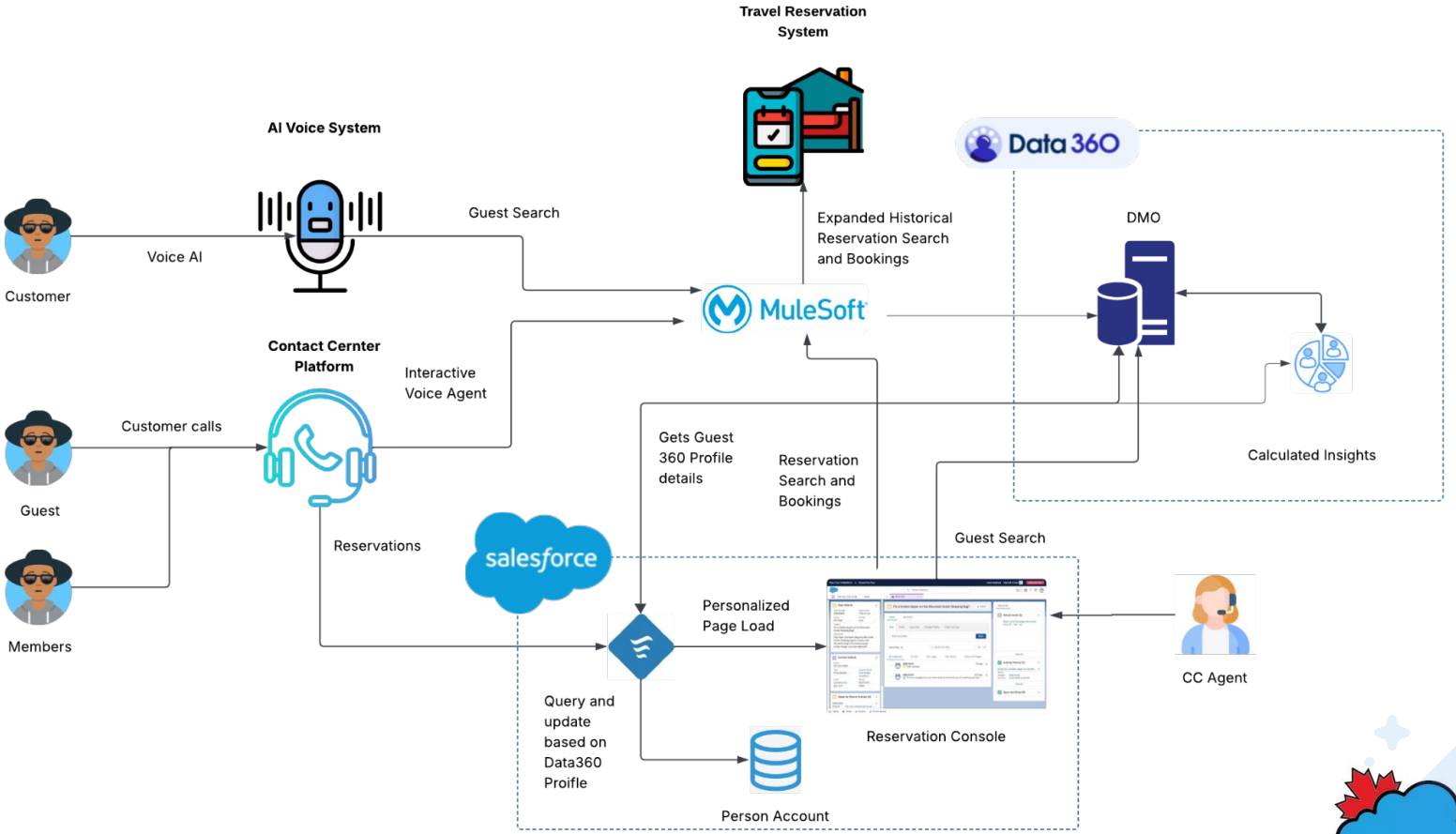


- Multi-channel customer interactions
- Moderate real-time usage
- Frequent profile lookups
- Multiple integrations
- **Mid six to low seven figures annually**

High Consumption

- High-frequency real-time queries
- Voice AI / contact center integrations
- Large-scale profile access
- Complex processing workloads
- **Multi-million dollar annual range**

Use Case: A High-Frequency Reservation System



True North Dreamin

Production Consumption Trend

Illustrative trend of daily credit consumption across key phases

Query Consumption
90%
of total consumption

Data Transform
4% – 5%
of total consumption

Profile Unification
1% – 3%
of total consumption

Data Ingestion
1% – 2%
of total consumption

Activation / Other
1%
of total consumption

Daily Credit Consumption (Illustrative)

Stacked by usage type

Show breakdown

The Hidden Drivers of Spend

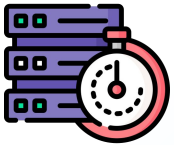
Broad Searches

Fuzzy Matching. Large Scans. Low Selectivity

Inefficient Queries

Duplicate Lookups. Repeated Fetches. Excessive Refreshes

Real-time Overuse



Always on Processing. Live Calculations. Frequent Triggers

Poor Data Efficiency

Incomplete Fields. Redundant Attributes

Cost Optimization Levers

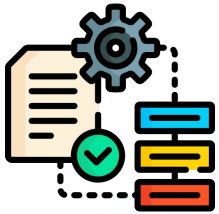
Scan Reduction

Reduce Full table scan.
Secondary Indexes

Query Precision

Exact Matching. Field Normalization. Cardinality Optimization

Execution Control



Lazy Loading. Precomputation. Call Reduction

Workload Shift

Query Consolidation

Non-critical Workload to Off-platform

Optimization Has Limits

Why Optimization Isn’t Enough

What Optimization Improves

Query Efficiency

- Faster Execution
- Lower Scan Volume
- Improved Filtering

Performance Gains

- Reduced Latency
- Better Response Times
- Incremental Savings

What Optimization Doesn’t Solve

Scaling Challenges

- High Query Frequency
- Real-Time Dependency
- Repeated Interactions

Structural Limitations

- Consumption Still Scales
- Cost Pressure Returns
- Platform Dependency Remains

Modernization Strategies

#	Options	Details	Pros	Cons	Recommendation
1	Databricks-Centric Architecture	• Sources → Databricks (standardization, cleanup and enrichment) • Curated data → Data 360 (unify profiles) • Unified data → back to Databricks (query/analytics) • Downstream systems query → Databricks	• Reduces significant query-driven load • Handles large joins & transformations efficiently • Scales beyond Data Cloud limits • Enables more systems to consume Data 360 without increased query cost	• Additional Databricks and Mulesoft cost • Integration complexity (bi-directional flow) • Cross-platform governance	✔
2	Hybrid Model	• Hot data → Data Cloud • Cold data → Databricks • Queries split across both (80/20 Rule)	• Partial load reduction • Keeps real-time access	• Query-heavy patterns remain • Complex routing logic • Limited cost savings	?
3	Service Cloud-Based	• Store key attributes in Person Account • Periodic sync from Data Cloud • UI served from Service Cloud	• Faster UI response • Fewer real-time queries (partial)	• Data duplication and sync overhead • Limited scalability • Still depends on Data Cloud Procession	✘

What Scales Better

Recommended Direction

High Consumption Model

Data Cloud Handles

- Real-time lookups
- Heavy query execution
- Complex processing
- High-frequency interactions

Result

- Multi-million dollar annual consumption profile

Modernized Consumption Model

Data Cloud Focuses On

- Identity resolution
- Customer segmentation
- Activation workflows

External Systems Handle

- Heavy processing
- Frequent queries
- Complex computations

Result

- **Consumption profile reduced significantly**

Business Impact and ROI

Consumption Reduction

40–70% Reduction

Lower Query Volume

Reduced Processing Load

Architectural Efficiency

Distributed Workloads

Predictable Scaling

Reduced Platform Dependency

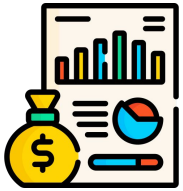
Performance Improvements

Faster Response Times

Lower Processing Latency

Improved System Scalability

Financial Impact



Seven-Figure Savings

Lower Consumption Tier

Reduced Annual Spend

Key Takeaways

- High interaction frequency drives consumption growth
- Real-world usage changes cost behavior
- Architecture redesign unlocks significant savings
- Strategic workload shift reduce annual spend
- Consumption visibility drives long-term savings

salesforce

True North Dreamin

Q & A

True North Dreamin