Aug 29, 2023

Customer 360 Best Implementation Practices

- Matt Boardman, Principal Solution Architect, IPS
- Deepak Khetan, Solution Architect, IPS
- Emily Sabanegh, Senior Consultant, IPS

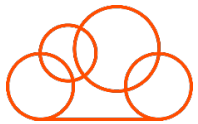
Informatica

Housekeeping Tips

- Today's Webinar is scheduled for 1 hour
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the Q&A option and we will respond at the end of the presentation
- The webinar is being recorded and will be available on our Success Portal - where you can also download the slide deck for the presentation. The link to the recording will be emailed as well.
- Please take time to complete the post-webinar survey and provide your feedback and suggestions for upcoming topics.

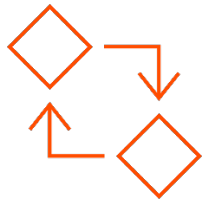
© Informatica. Proprietary and Confidential.

Feature Rich Success Portal



Bootstrap trial and POC Customers

Enriched Customer Onboarding experience



Product Learning Paths and Weekly Expert Sessions

Informatica Concierge

Tailored training and content recommendations

© Informatica. Proprietary and Confidential.

Informatica

More Information

Success Portal

https://success.informatica.com

Communities & Support

https://network.informatica.com

Documentation

https://docs.informatica.com

University

https://www.informatica.com/in/services-and-training/informatica-university.html

© Informatica. Proprietary and Confidential.

Informatica

Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

Informatica

Informatica

8.29.2023

Customer 360: MarTech Best Practices

Matt Boardman
Principal Solution Architect

Deepak Khetan
Solution Architect

Emily Sabanegh
Senior Consultant

Agenda

1. Introduction
2. Business Problem
3. Implementation Architecture
4. Integration Patterns
5. Data Model + Matching
6. Q&A

© Informatica. Proprietary and Confidential.

Today’s Presenters

Matthew Boardman

Principal Solution Architect

Deepak Khetan

Solution Architect

Emily Sabanegh

Senior Consultant

© Informatica. Proprietary and Confidential.

Informatica

Customer 360 SaaS

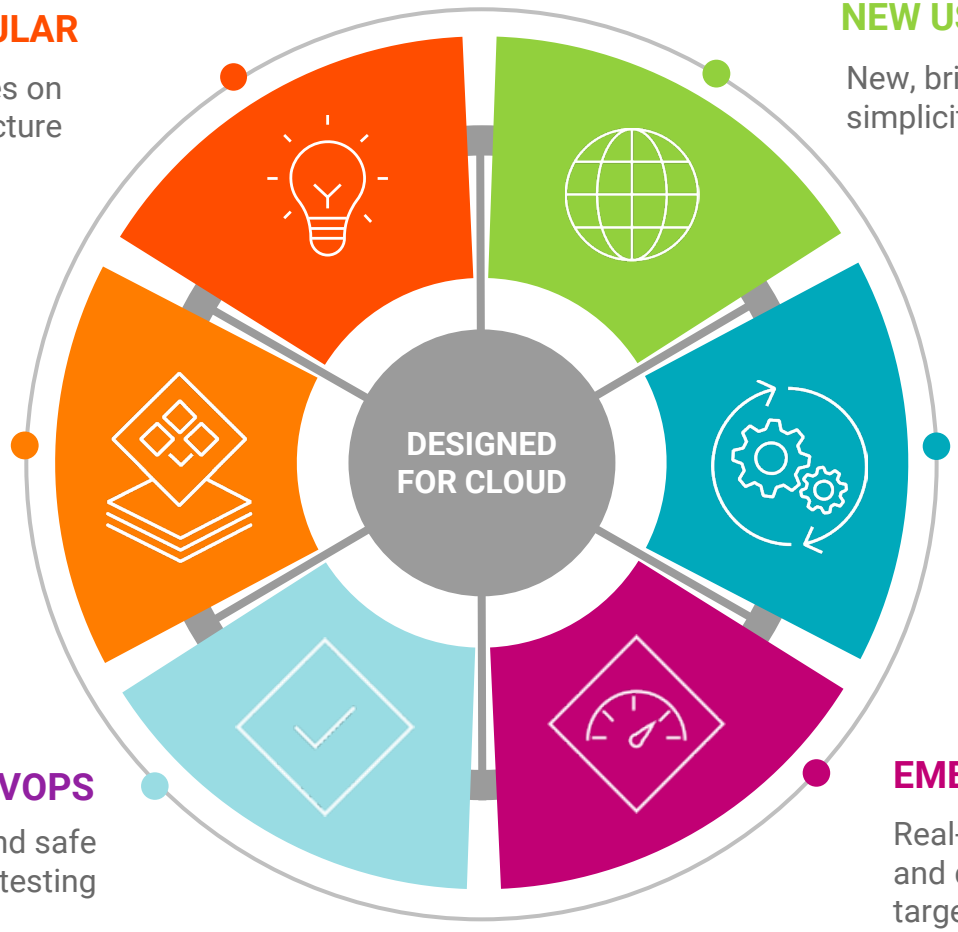
Prepackaged SaaS offering built cloud-native data management platform for greater productivity

FLEXIBLE & MODULAR

Modular, API-driven, shared microservices on market leading scalable IPaaS architecture

NEW USER EXPERIENCE

New, brighter, sleeker visual design preconfigured for simplicity



SIMPLIFIED PRICING AND PACKAGING

Transparency and ease of transactions

BUILT-IN SMARTS & ORCHESTRATION

AI training, content aware data types, orchestrated tasks across MDM, DQ and Data Governance, pre-integrated DaaS

SIGNIFICANT INVESTMENTS IN DEVOPS

Continuous innovation, monthly releases and safe upgrades, backed by rigorous automated testing

EMBEDDED DATA INTEGRATION AT ANY LATENCY

Real-time, streaming and batch integration, with drag and drop workflows and data flows for sources and targets

© Informatica. Proprietary and Confidential.

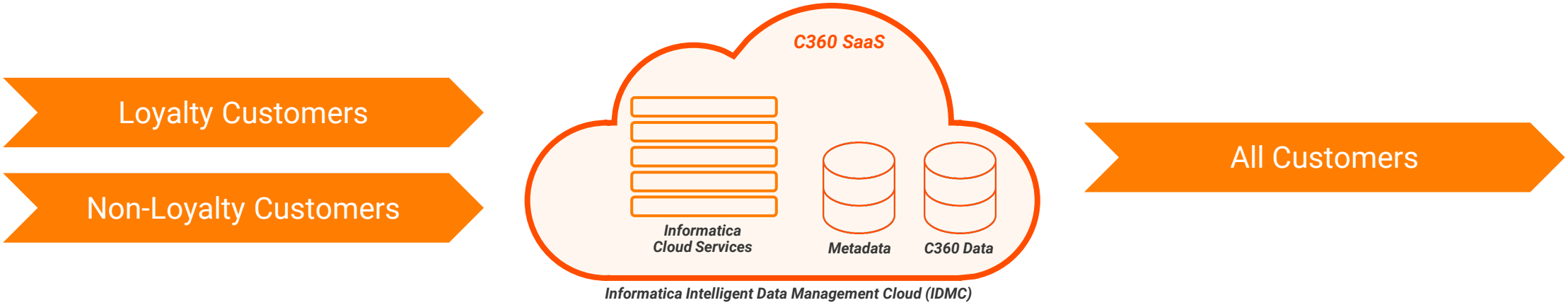
Informatica

Business Problem

Use Case: Large Customer Volume

Customer: Global Fortune 500 Company

Business Problem – Identity Resolution for Individuals



Large Customer Volume (100s of millions)

- Loyalty Program – de-dupe multiple accounts
- Non-Loyalty – identify unique individuals
- High rate of change: 5M – 10M records per day
- High initial load volume: 300M – 400M records
- Each incoming record matches against 500M+
- Applications: Get vs. Search vs. Search Match

Implications

- Near Real-time response returning customer identifier
- Daily match & merge – results within 12 hours
- High degree of automation – "Low Touch"
- Data Profiling → Unique Customer Scenarios
- Cleanse noise words, Address Validation, Standardization
- "Chunking Process" breaks up initial load

© Informatica. Proprietary and Confidential.

Implementation Architecture

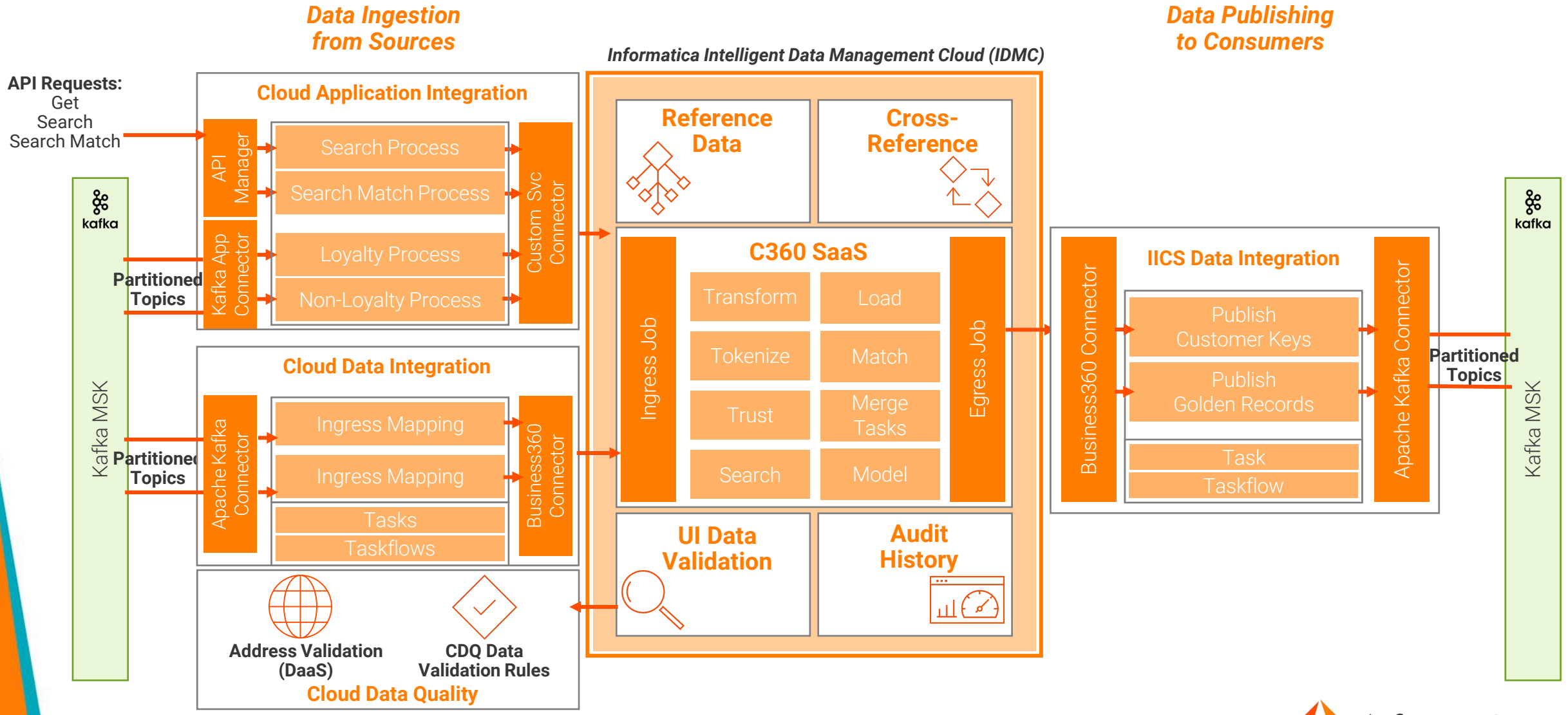
Informatica MDM SaaS (C360)

Informatica CAI

Informatica CDI

Informatica CDQ

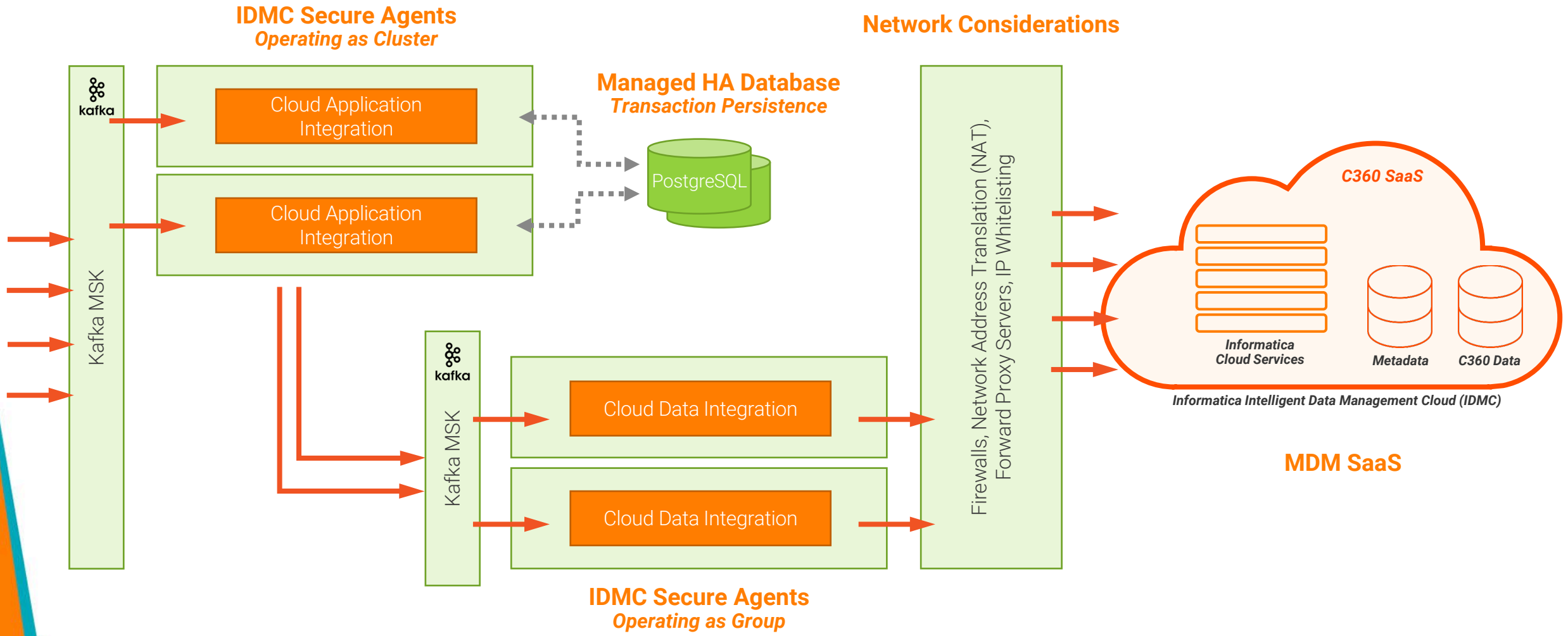
MDM View of MarTech Ecosystem



© Informatica. Proprietary and Confidential.

Informatica

C360 Supporting Infrastructure



© Informatica. Proprietary and Confidential.

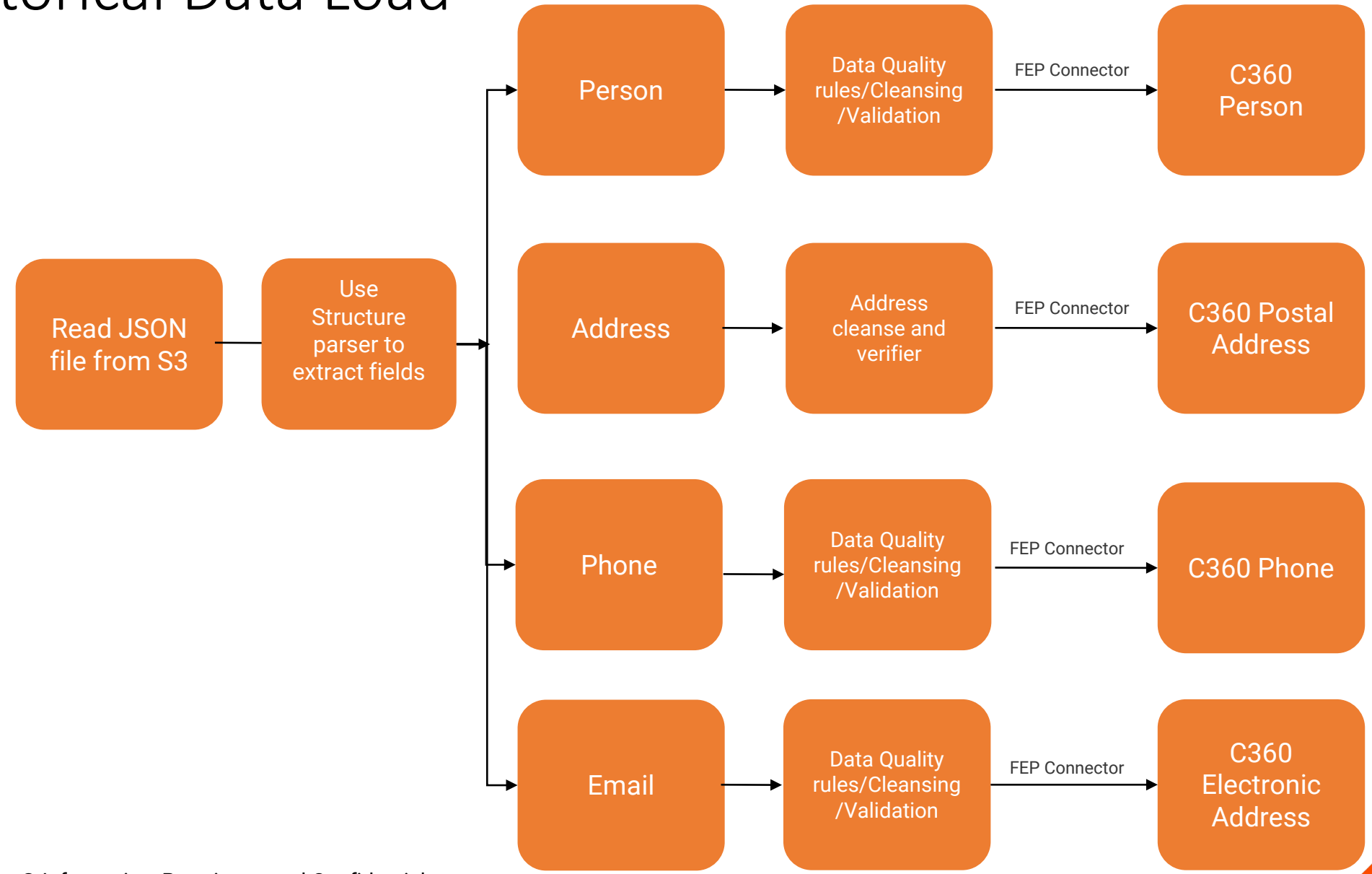
Informatica

Integration Patterns

Historical Load

Incremental Near Real-Time Load

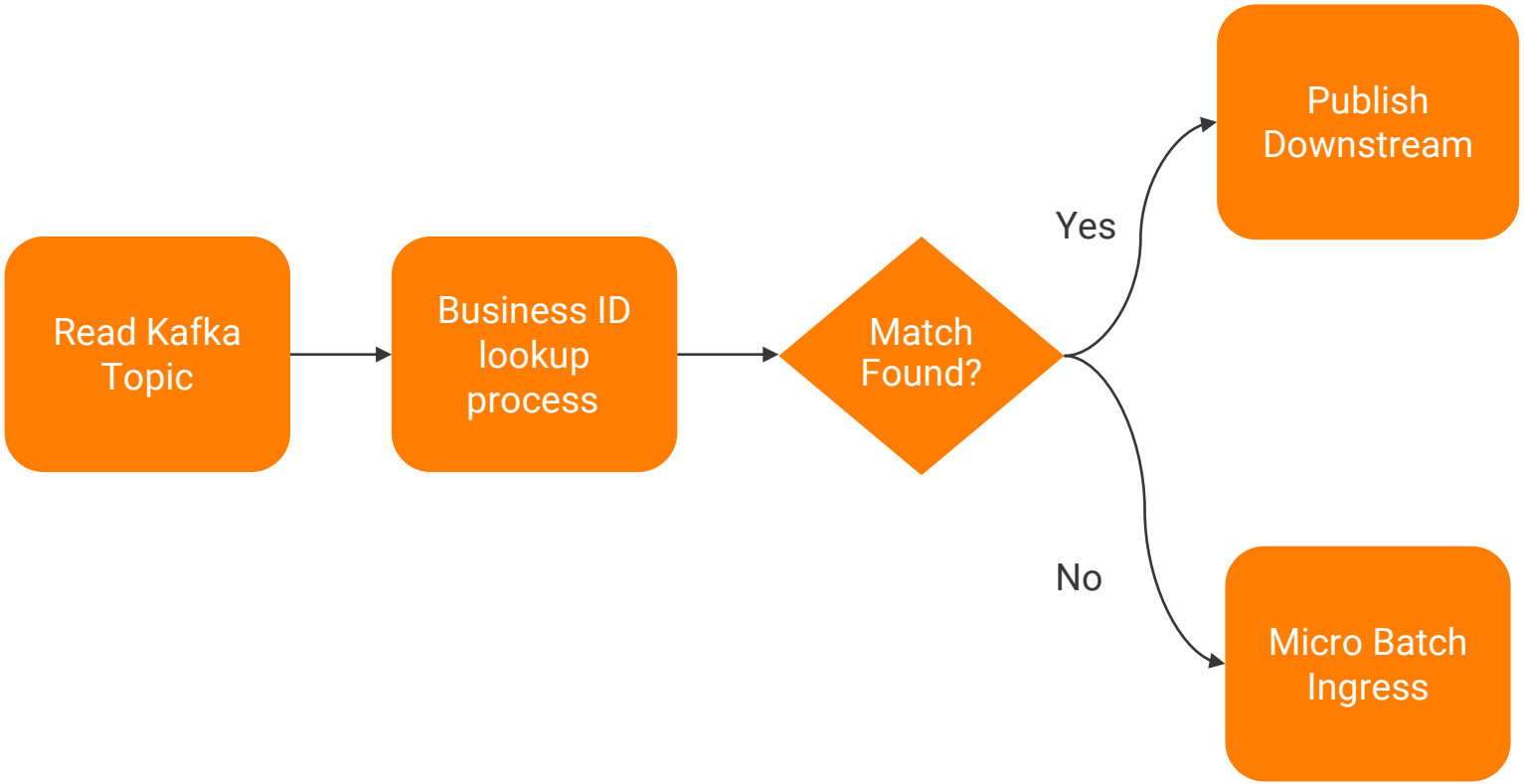
Historical Data Load



© Informatica. Proprietary and Confidential.

Informatica

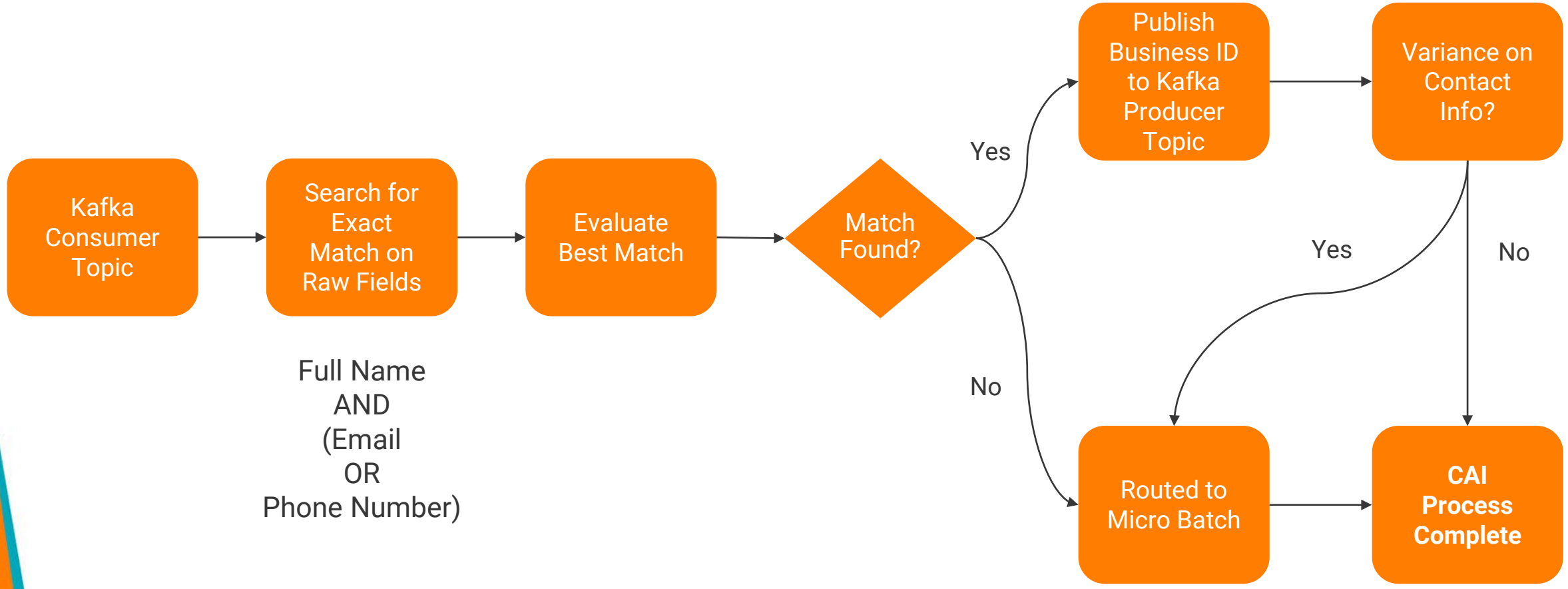
Incremental Data Load Overview



© Informatica. Proprietary and Confidential.

Informatica

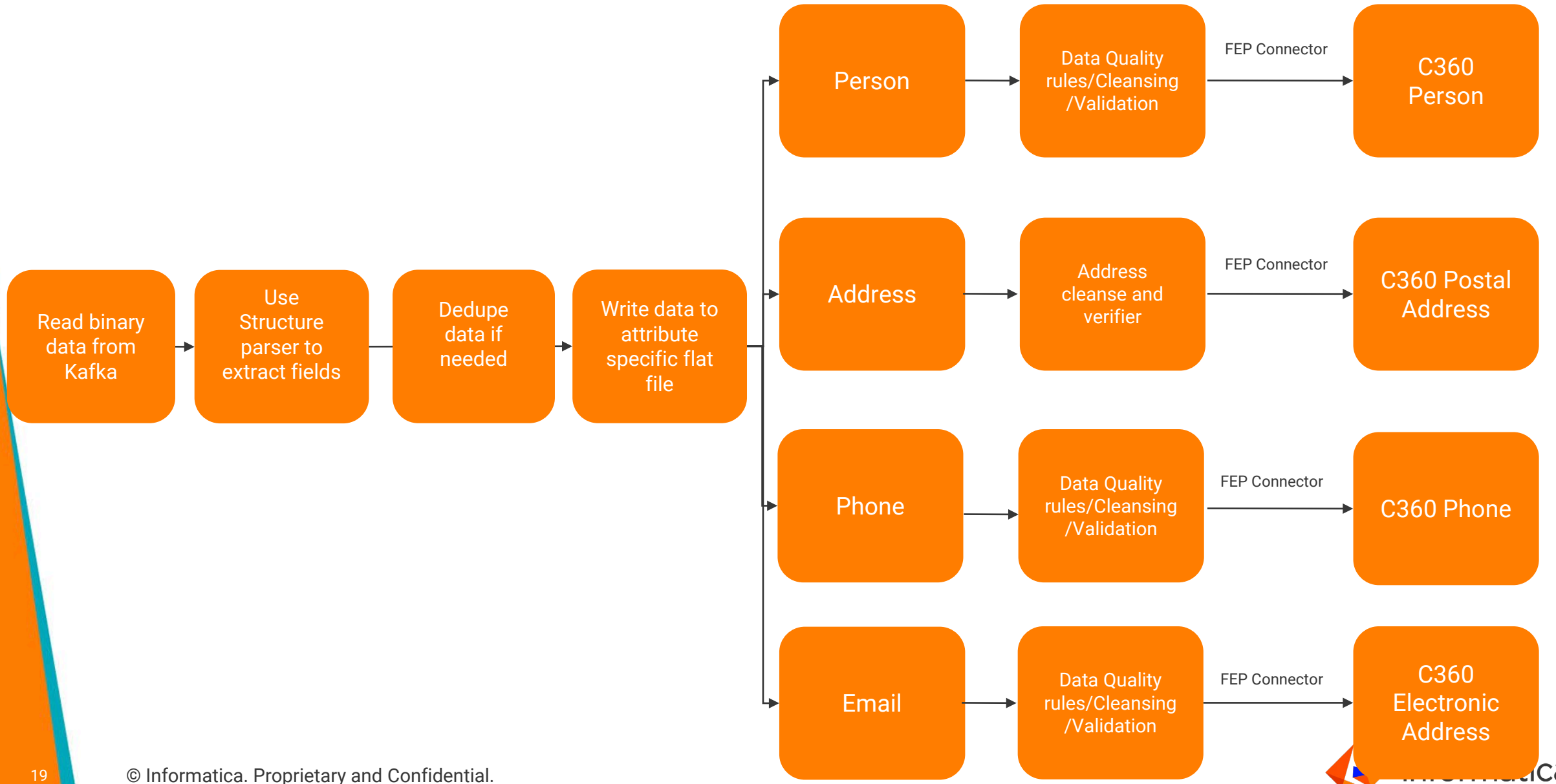
Business ID Lookup – Cloud Application Integration



© Informatica. Proprietary and Confidential.

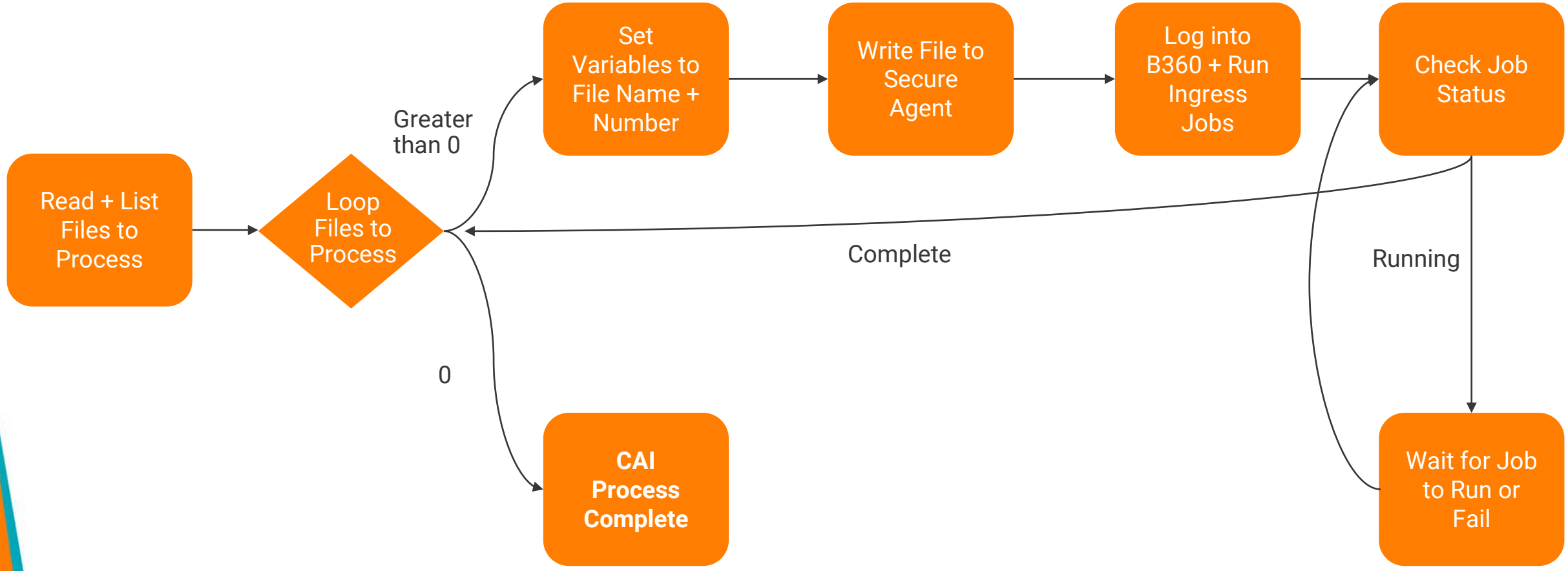
Informatica

Incremental Micro Batch



© Informatica. Proprietary and Confidential.

Chunking Process – Cloud Application Integration



© Informatica. Proprietary and Confidential.

Informatica

Data Modeling and Matching

Data Models

Match & Merge

Data Modeling in C360 SaaS

Data Model

- Business Entity – Person or Organization
- Child Field Groups – Address, Phone, Email, Loyalty Account, Payment Type, Segmentation, Status, Demographics
- C360 Extensions – Healthcare, Insurance, others

Other Considerations

- Data Quality Rules – character lengths, allowed values, noise words, capitalization, date ranges, advanced CDQ rules
- Survivorship – most trusted source systems
- Match Rules – same algorithm, different implementation
- Code Value translation – Reference 360 code lists, cross-walks
- Hierarchies – consider enrichment, e.g. Dunn & Bradstreet, LexisNexis, Bloomberg, Statistica, Informatica DaaS
- Relationships – householding, personal / work accounts

© Informatica. Proprietary and Confidential.

Matching in C360 SaaS

Considerations

- **Match Partitions** – Partition by Country, State
- **Multiple Rulesets** – Daily, Weekends, Real-Time
- **Match Rules** – Exact before Fuzzy, High Hits First
- **Exact Field Segmentation** – Act as Filters
- **Null Matching** *(New!)* – Null-Null, Null-Not Null
- **Globalization**: Phone, Address standardization

Informatica Assistance

- Match report is limited to 100K pairs, but **GCS** can extract a high-volume **match report**
- **CSM** can provide specialists: "**Ask an Expert**"
- **IPS** can help with **full life-cycle**: user requirements, data profiling, determining match rules, testing through multiple iterations, tuning & optimization

© Informatica. Proprietary and Confidential.

Informatica

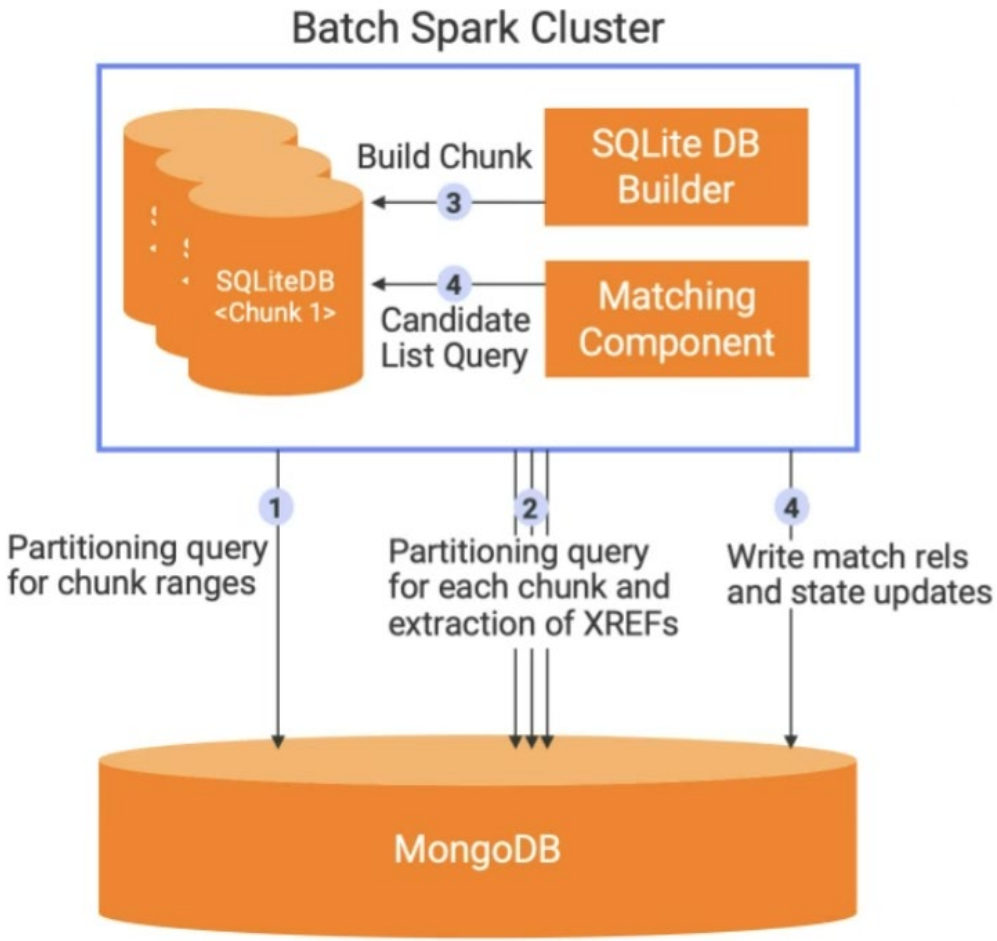
Example Match Ruleset

Attribute		Rule 1	Rule 2	Rule 3	Rule 4
Match Criterion		Person_Name	Person_Name	Person_Name	Resident
Merge Strategy		Automated	Automated	Automated	Automated
Match Level		(Typical)	(Conservative)	(Conservative)	(Conservative)
Threshold		95-100	97-100	97-100	97-100
		Strongly similar name and exactly matching identifier	Similar name and same email	Similar name and same phone number	Strongly similar name, address
Person	First Name				
	Last Name				
	Full Name	Fuzzy	Fuzzy	Fuzzy	Fuzzy
	Date of Birth				
Alternative Identifier	Alt Identifier Type	Exact			
	Alt Identifier Value	Exact			
Address	Address Type				
	Address Line 1				Fuzzy
	Address Line 2				Optional (Fuzzy)
	City				
	State				Exact (Null-Null)
	Postal Code				Optional (Fuzzy)
	Country Code				Exact
Phone	Phone Type			Exact	
	Phone Num			Exact	
Email	Email Type				
	Email Address		Exact		

© Informatica. Proprietary and Confidential.

Informatica

Example: Analyzing Match Performance



Step	Time	Opportunity
1 – Partitioning Query	60 mins	Reduce to ~10 mins
Match Chunk "C1" – *Match C1 to C1*		
Spark Job Setup	15 mins	
2 – Partitioning Query	70 mins	Reduce to ~10 mins
2 – Data Extract	10 mins	
3 – SQLite DB Building	45 mins	Can optimize further, benefit TBD
4 – Matching	15 mins	
Match Chunk "C2" – *Match C2 to C2*		
Spark Job Setup	15 mins	
2 – Partitioning Query	15 mins	Reduce to ~10 mins
2 – Data Extract	10 mins	
3 – SQLite DB Building	45 mins	Can optimize further, benefit TBD
4 – Matching	15 mins	
Match Chunk "C3" – *Match C3 to C3*		
Spark Job Setup	15 mins	
2 – Partitioning Query	30 mins	Reduce to ~10 mins.
2 – Data Extract	10 mins	
3 – SQLite DB Building	45 mins	
4 – Matching	30 mins	
Match C1 to C2, C1 to C3, C2 to C1, C2 to C3, C3 to C1, C3 to C2		
Spark Job Setup	65 mins	
4 – Matching	100 mins	Reduce by 60 minutes
Total	**10.5 hours**	**7.5 hours**

© Informatica. Proprietary and Confidential.

Informatica

Engineering Roadmap

Near-Term Roadmap

- Incremental Match Support
- Support for Extra-Large Chunks
- Near Real-Time Load, Merge API Optimizations
- Error reporting, Match reporting
- Radius matching

Strategic Roadmap

- Match Tokenization during Source XRef creation
- Batch and Near Real-Time Data Quality
- Offload transient Data to Apache Iceberg
- AI-driven Indexing and Blocking

© Informatica. Proprietary and Confidential.

Thank you

mboardman@informatica.com

dkhetan@informatica.com

esabanegh@informatica.com

Informatica

Appendix