



Problem

Access to multiple types of data from numerous sources and locations in either an event-based or time-scheduled manner places a large burden on enterprise data infrastructure and services.



Solution

Deploy a data integration platform in the edge node. A data integration platform essentially takes data sources from a number of supported source interfaces (file, database, object store, etc.), transforms it into a universal format, and then uses data services to provide varied consumer interface choices in order to consume the data. This already has widespread value to organizations that frequently need to integrate data between disparate applications (in one or more clouds). Used a different way, this platform also enables a data exchange. Data exchanges are groups of companies that are securely interconnected in the edge node for the purpose of accessing/sharing data (which is typically monetized). New data sources are valuable to data-oriented partners. As in analytical processing, more data sources directly translate to more experience (it could be IoT data, scientific data, medical trial data, etc.). Even if "translation" is not required and data is passed straight through, the other governance functions provide significant value and needed oversight in a dynamic, automated environment.



Constraints

1. There are hundreds of permutations of data transfer occurring. Some of these require governance items that may not be applied.
2. In a "trust nothing" environment, each action that changes the data should have some level of governance review, which is impractical for individual service implementers to do — anyone can call their service from an API (dynamic with no humans involved).
3. Random acts of data transformation (in some form or level) exist throughout the environment. Services like ETL (extract, transform and load) consume resources and lack macro coordination.
4. Data guardians experience difficulty governing the dynamic transformation, yet are still accountable for the data.



Steps

1. Deploy access (adapters/connectors), transformation and delivery services (adapters/connectors).
2. Wire each step to go through boundary control and inspection zone(s) (Security Blueprint*).
3. Apply event processing and policy enforcement.
4. Configure data profiling, data quality and operational processing.
5. Provide internal and external (productized) APIs for data integration (as a service).
6. Configure data repository (Step 1) for data staging and local private storage (Step 2) for caching/performance.
7. Apply metadata and master data, management and re-encryption for destination key management integration (Security Blueprint*).
8. Integrate with a data pipeline service and update provenance information.



Forces

- Data will soon replace traditional products in most firms as the most revenue-generating asset (digital economy).
- Universal methods of accessing, passing and transferring data between disparate systems, companies and networks will be a critical capability. As the need for exchange increases, it needs to be balanced with the mitigation of potential risks.
- The dependency on data sources will drive legal and service expectations (with monetary impact being almost guaranteed) in the event of unavailability or data loss.
- As data leaves a domain of control, where it will go, how it will be protected and what it will be used for may need to be policed.

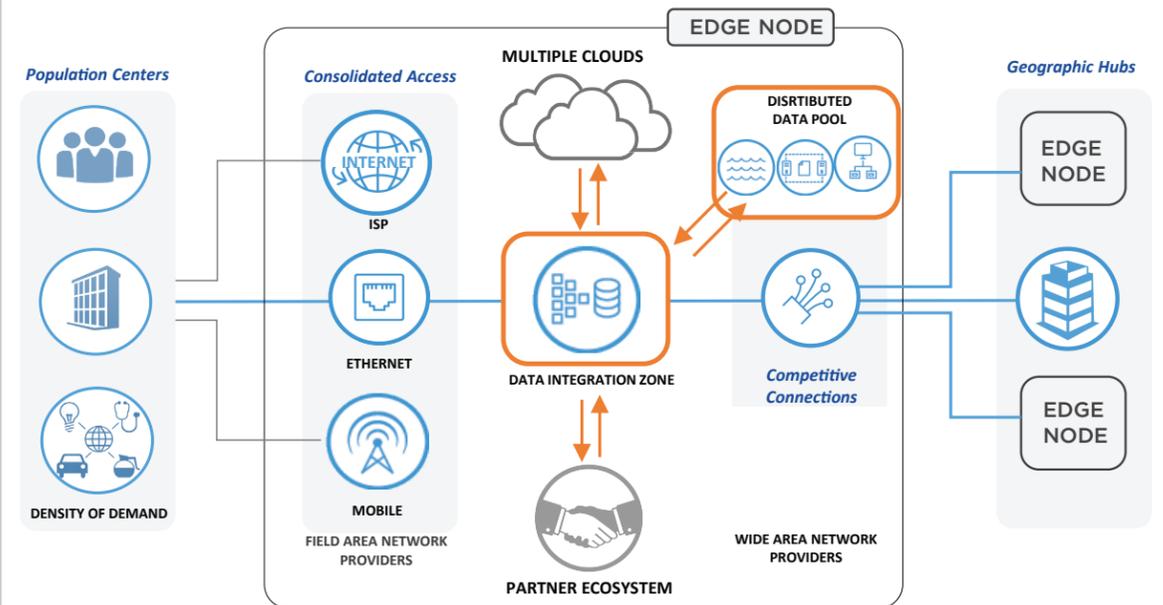


Results

- Uniform way to exchange data between applications, cloud services and business ecosystem partners (with or without compression).
- Delivering dynamically integrated real-time data and managed batch transfer/migrations.
- Full event processing of all data exchange points that go through a data exchange (internally and externally), forming a data events view and audit trail, with dependency analysis.
- Data encryption, masking, PII (personally identifiable information) alerting and leakage prevention can be applied. The exchange is another policy enforcement point.
- Detect data corruption and tampering with machine learning (Application Blueprint*).
- Data exchanges allow you to produce entirely new business models built on integrated data.



Reference View



* Security and Application Blueprints — IOAKB.com