



influx | **days**

2015-06-11T20:46:02Z

InfluxEnterprise Architectural Patterns

Dean Sheehan
Snr Director Sales Engineering

28:07.580664347Z

1985-10-25T19:28:07.5

What we will be covering

- ✓ Enterprise Overview
- ✓ Other Features
- ✓ Ingestion & Query Rates
- ✓ Deployment Examples
- ✓ Replications Patterns
- ✓ General Advice

2015-06-11T20:46:02Z



Why InfluxEnterprise?

1985-10-25T19:20:07.5

Signs You're Ready for InfluxEnterprise

1. The sales team starts calling you on weekends
2. Data recording and availability matters
3. Vertical scaling not providing further benefit
4. Sprawling number of single node deployments
5. Increasing throughput causing write drops errors
6. Your CPU average is $\geq 70\%$

InfluxEnterprise

- Open Source Core
- High Availability
- Horizontal Scalability
- Enterprise Security
- Support from InfluxData
- OnPremise/Cloud Deployment Options

2015-06-11T20:46:02Z



What Problem Are You
Trying to Solve?

1985-10-25T19:20:07.5

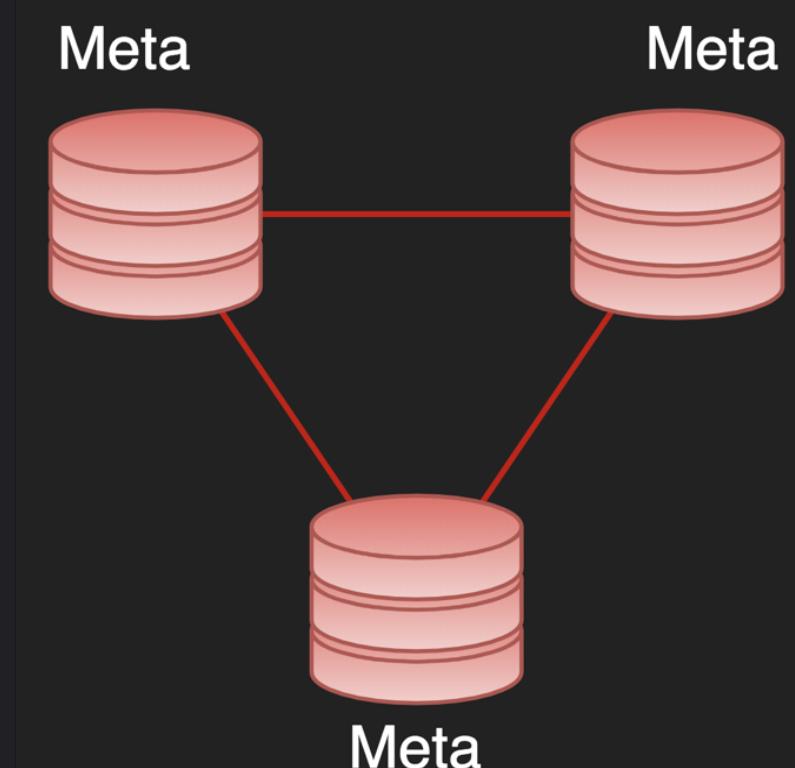
What are you dealing with?

- Metrics
- Events
- Log Data
- Sensors
- Apps
- Servers
- Long-Term Storage
- Vendor Replacement
- Time-Series Alerts
- Visualization
- Network Data
- Custom Solution
- Real-Time Analytics
- Virtualization Monitoring
- Managed Service (InfluxCloud)

InfluxEnterprise Overview

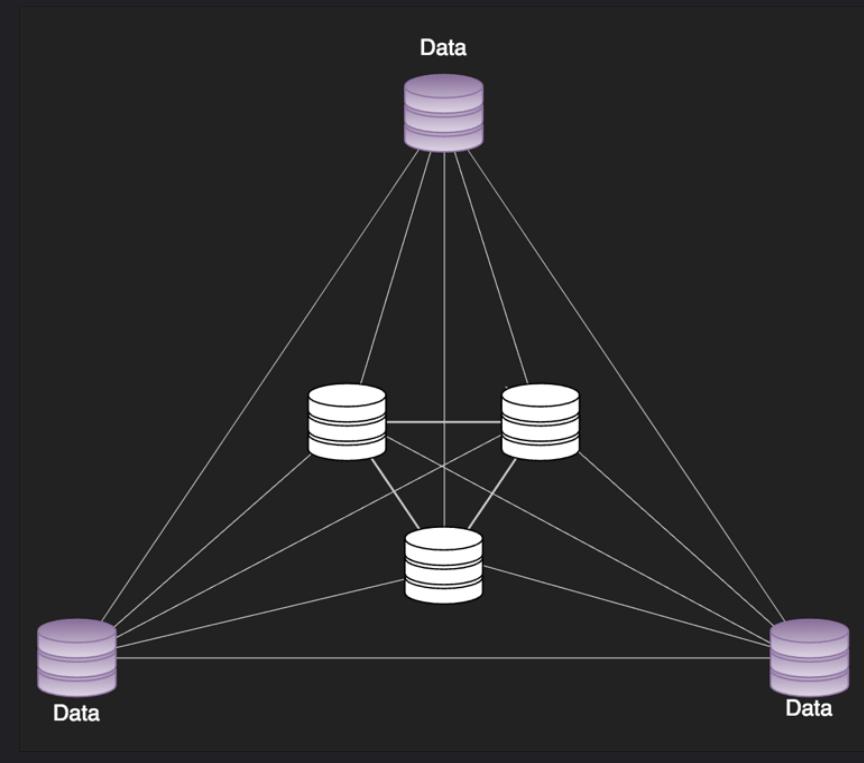
InfluxEnterprise Cluster Architecture: Meta Nodes

- Keep state consistent across the cluster
 - Users
 - Databases
 - Continuous Queries
 - Retention Policies
 - Shard Metadata
 - Cluster Membership - Servers
- 3 Meta nodes for High Availability
 - Uses Raft consensus algorithm
 - Quorum needs an odd number of nodes
 - $(n/2)+1$ nodes available for consensus

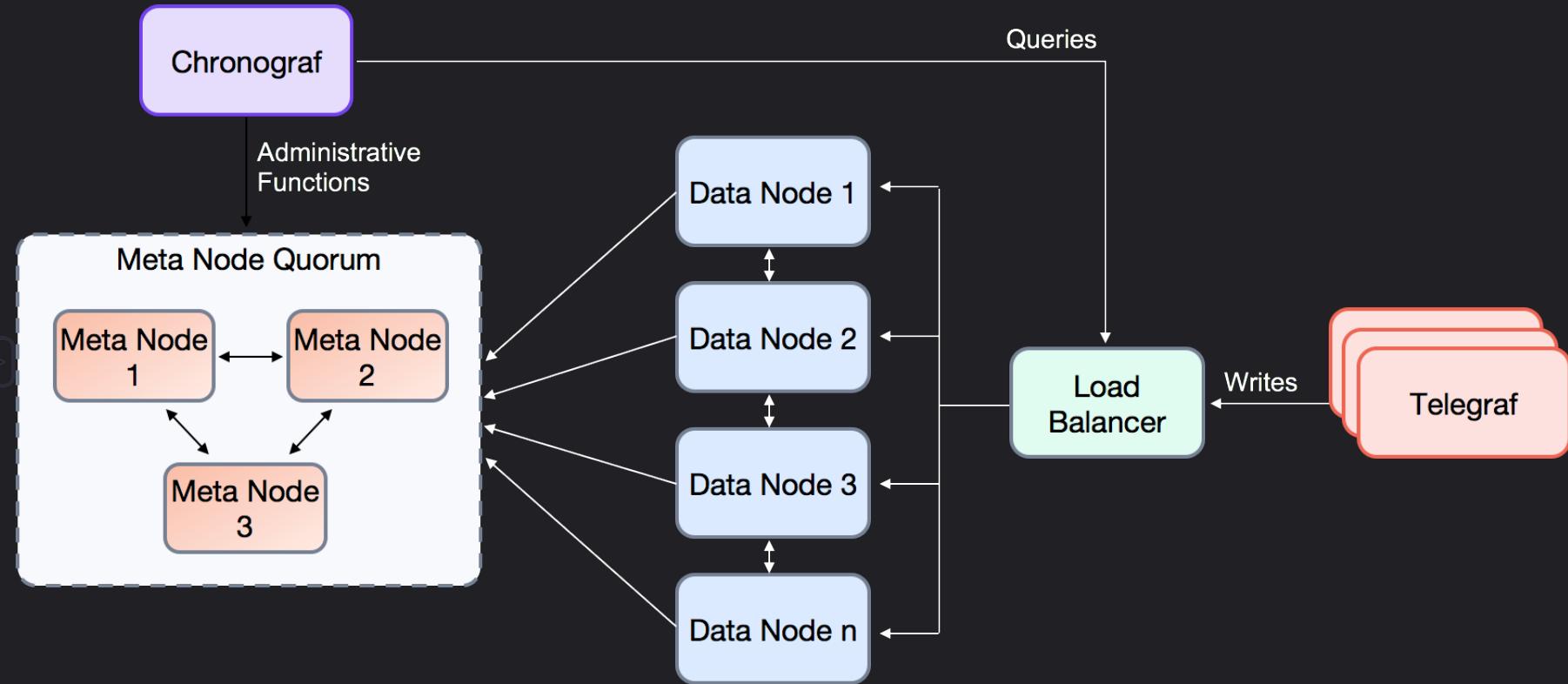


InfluxEnterprise Clustering: Data Nodes

- Data nodes hold all the raw time series data and metadata including:
 - Measurements
 - Tag keys and values
 - Field keys and values
- Do not participate in consensus
- 2 Data Nodes for High Availability RF = 2
- # of Data Nodes divisible by RF
- On disk the data is organized by:
 - <database>/<retention_policy>/<shard_id>
- By default the parent directory is:
 - /var/lib/influxdb/data



InfluxEnterprise Cluster Architecture



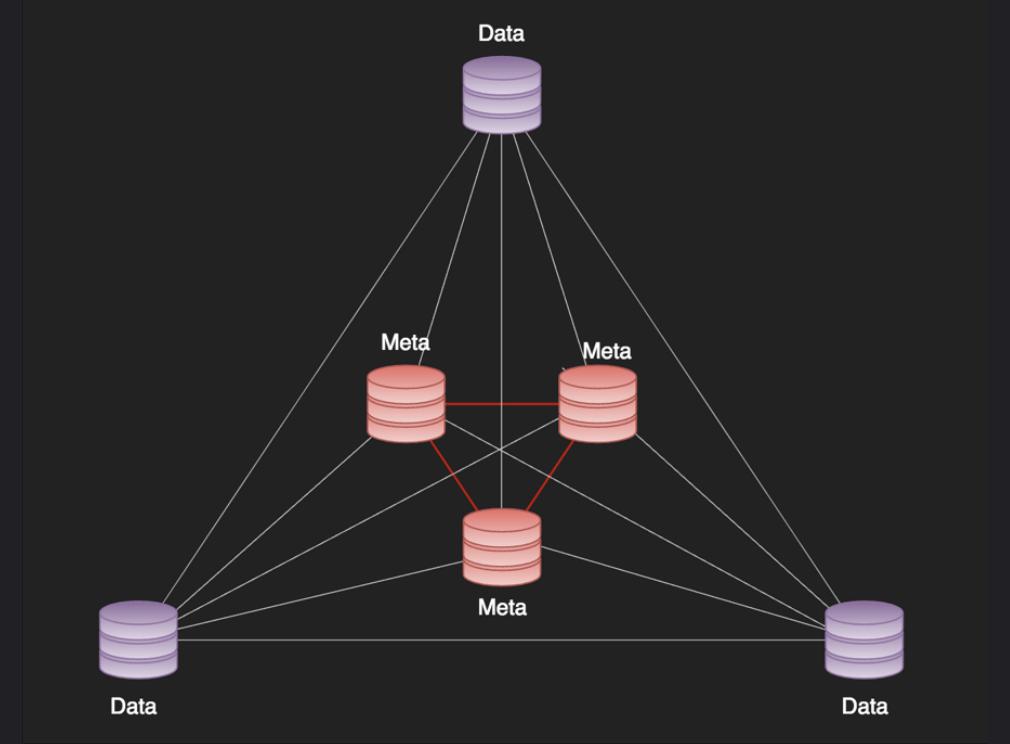
InfluxEnterprise Cluster Architecture

Meta Nodes

- CPU: 1-2 cores
- RAM: 512 MB – 1 GB
- Disk: 1 HDD of any size
- Can be run in a VM or container

Data Nodes

- CPU: 8+ cores
- RAM: 64+ GB
- Disk: SSD drives for WAL, HH and Data
 > 1000 IOPS recommended
- Network: 10 Gbps NIC



Features

Security

- **LDAP Support**
 - Enterprise customers can configure the database to use LDAP as a backing authentication source for users, roles and permissions.
 - Connection between DB and LDAP server secured once connected
- **Fine-grained authorization**
 - Used to control access at a measurement or series level (compared to limiting access at the database level)
 - Enable authentication in your configuration file
 - Create users through the query API
 - Grant users explicit read and/or write privileges
 - Set restrictions which define a combination of database, measurement, and tags which cannot be accessed without an explicit grant

Eventual Consistency

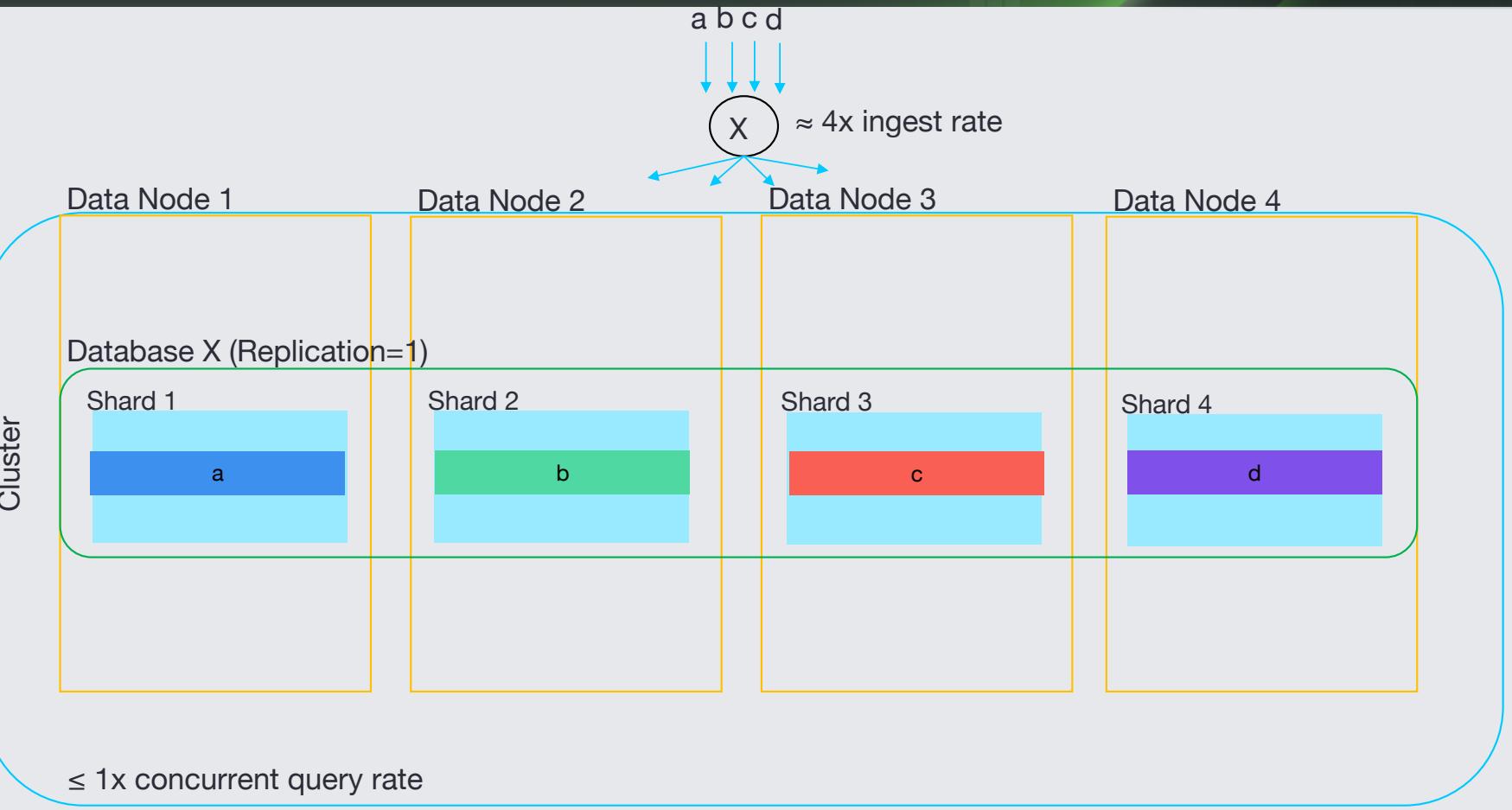
- Hinted-Handoff Queue
 - Queue inbound points destined to land on other nodes in the cluster which may currently be down
 - Stored by node and shard (10GB - default)
- Anti-Entropy Service
 - Expands on capabilities to detect and copy full shards
 - Now allows for detection and repair of inconsistent shards

Backup and Restore

- Useful for: Disaster recovery, Debugging, Restoring clusters to a consistent state
- What it does: Creates a copy of the metastore and the shard data
 - Backup is compressed and is not human readable
 - Export is not compressed but is human readable
 - OSS and InfluxEnterprise ARE NOW compatible – aka portable
 - Full or partial backup options
 - Move data into a new database (with new Retention Policies, etc)

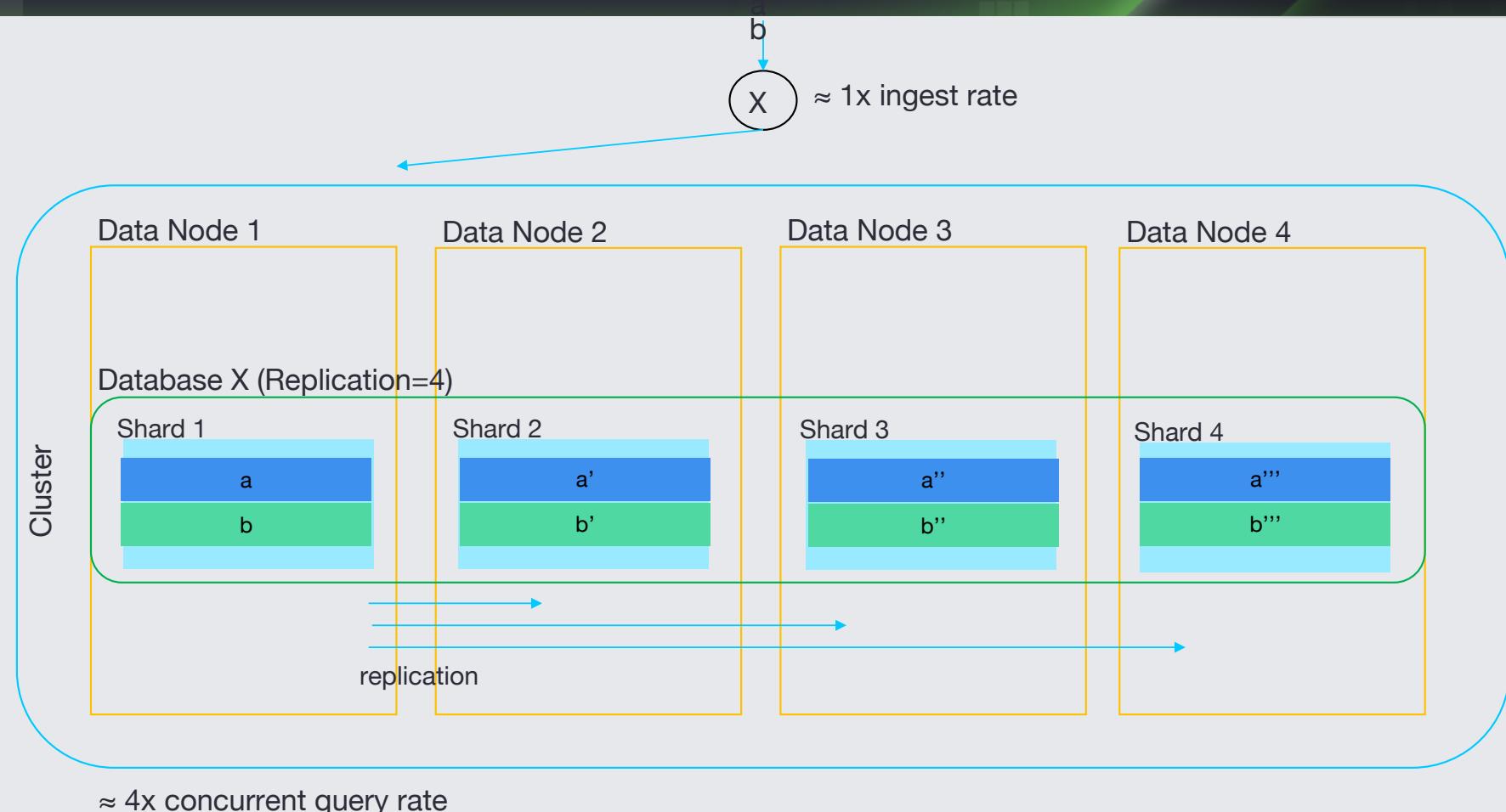
Ingestion & Query Rates

2015-06-11T20:46:02Z

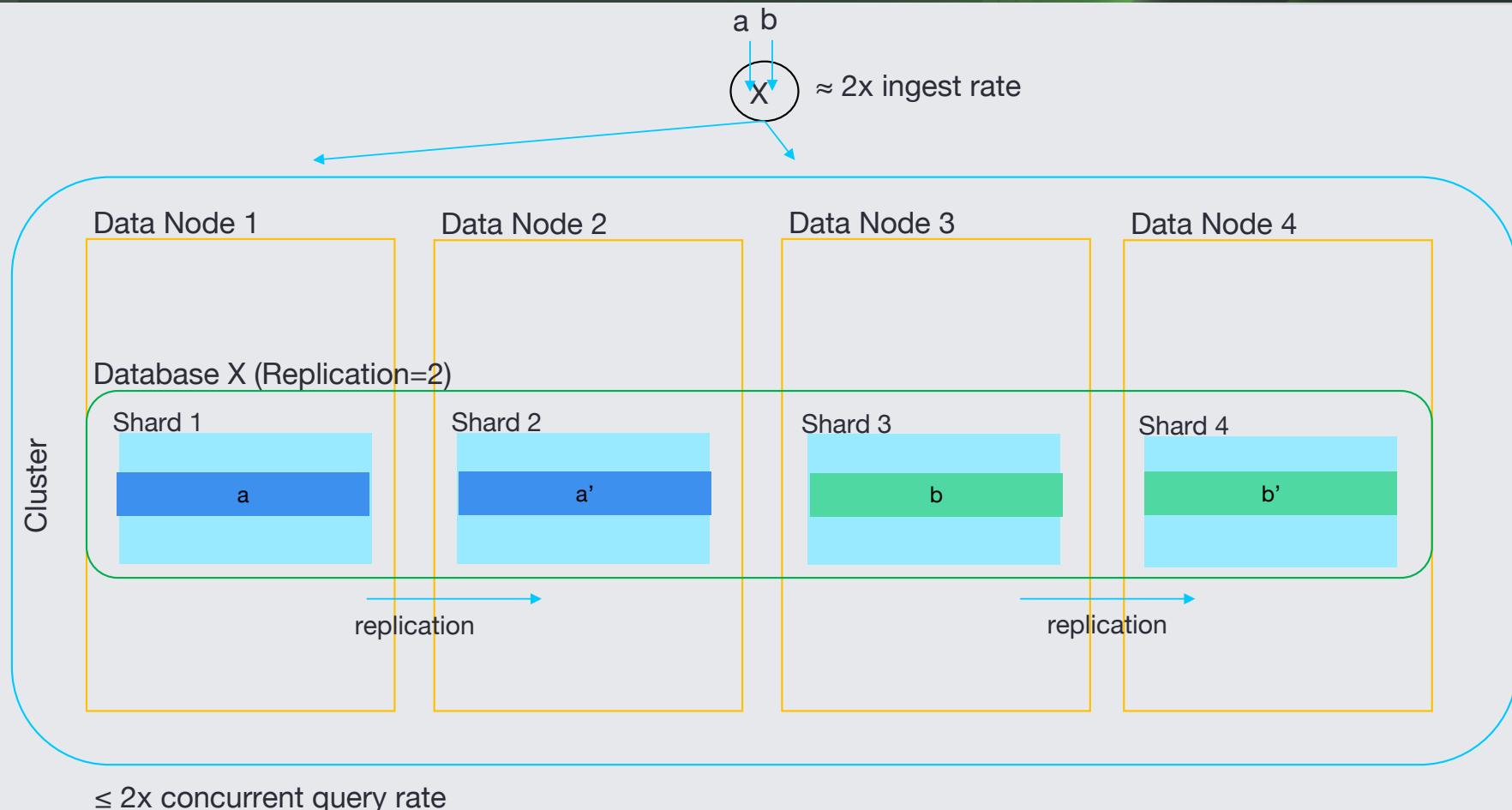


1985-10-25T19:20:07.5

2015-06-11T20:46:02Z



2015-06-11T20:46:02Z



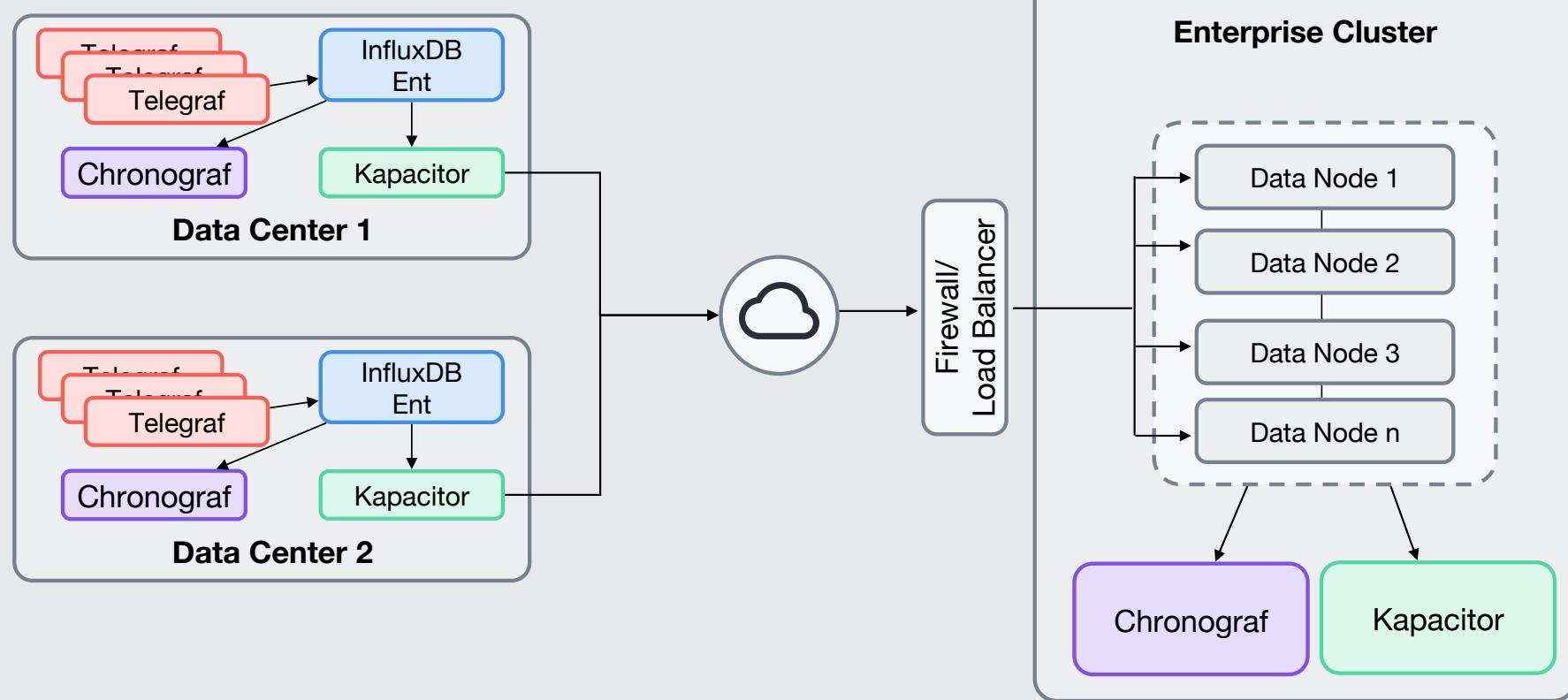
Deployment Examples

2015-06-11T20:46:02Z



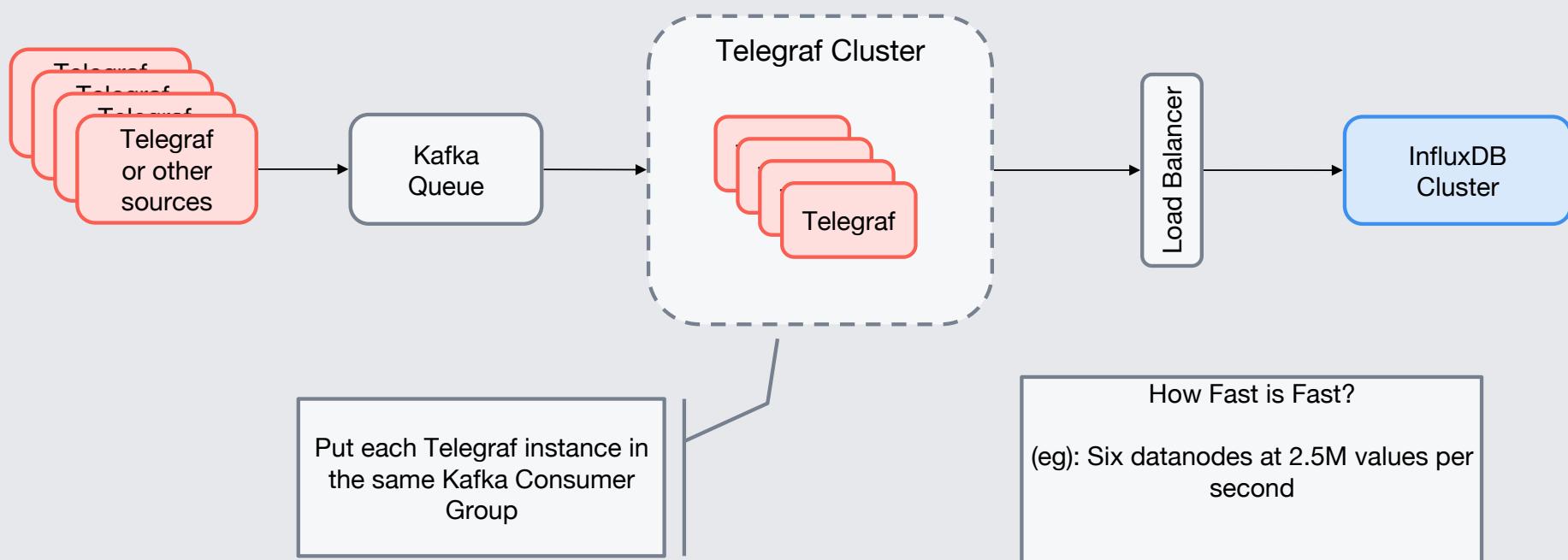
How does InfluxEnterprise Fit?

Example 1: Mothership

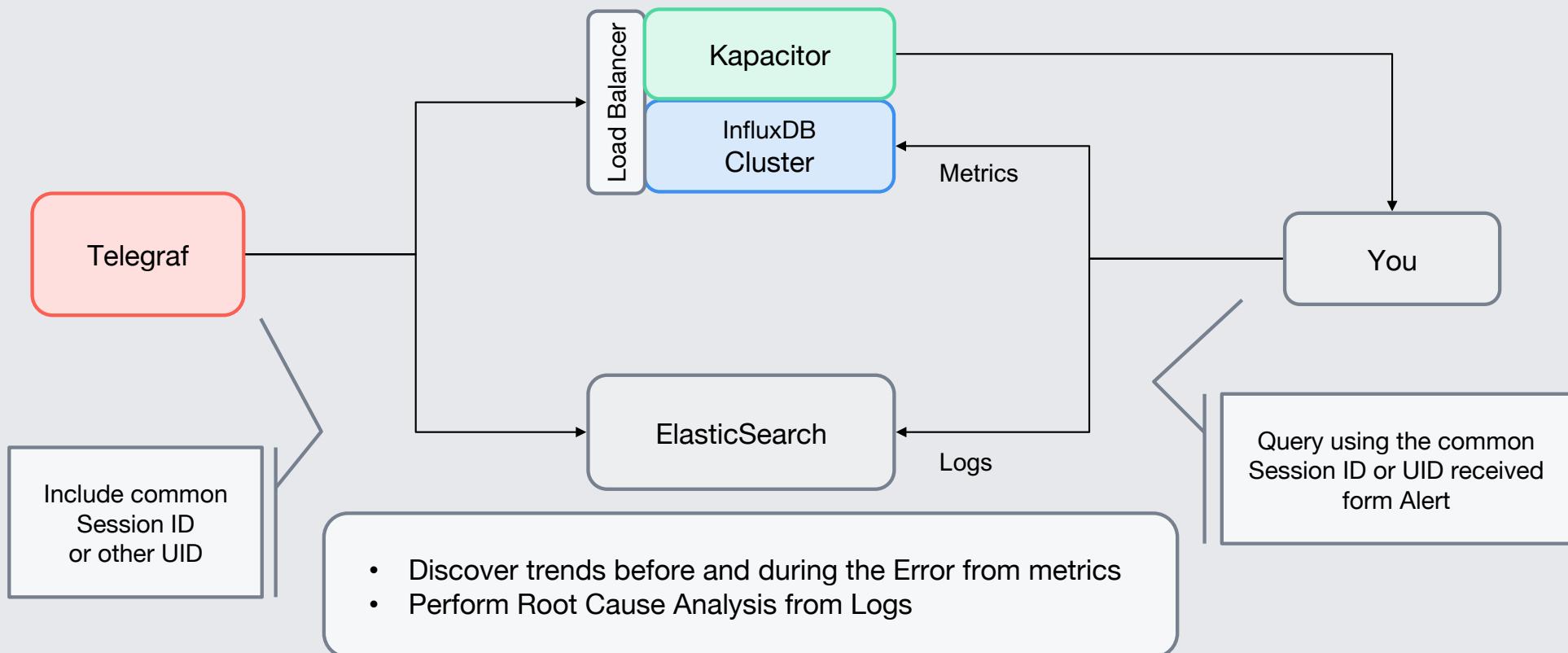


1985-10-25T19:20:07.5

Example 2: Durable Data Ingest



Example 3: Influx with ElasticSearch



Replication Patterns

2015-06-11T20:46:02Z



How Are You Using InfluxDB?

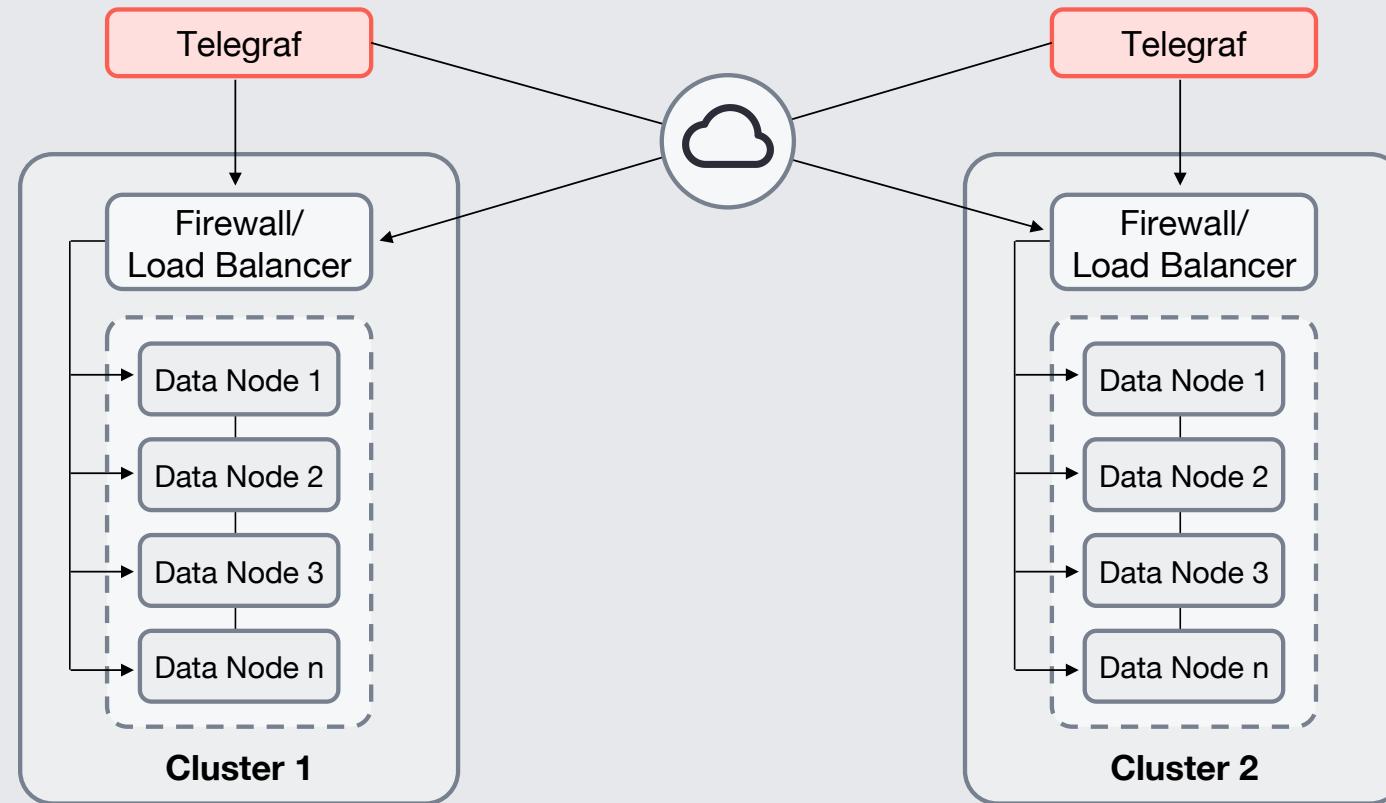
1985-10-25T19:20:07.5

Data Replication

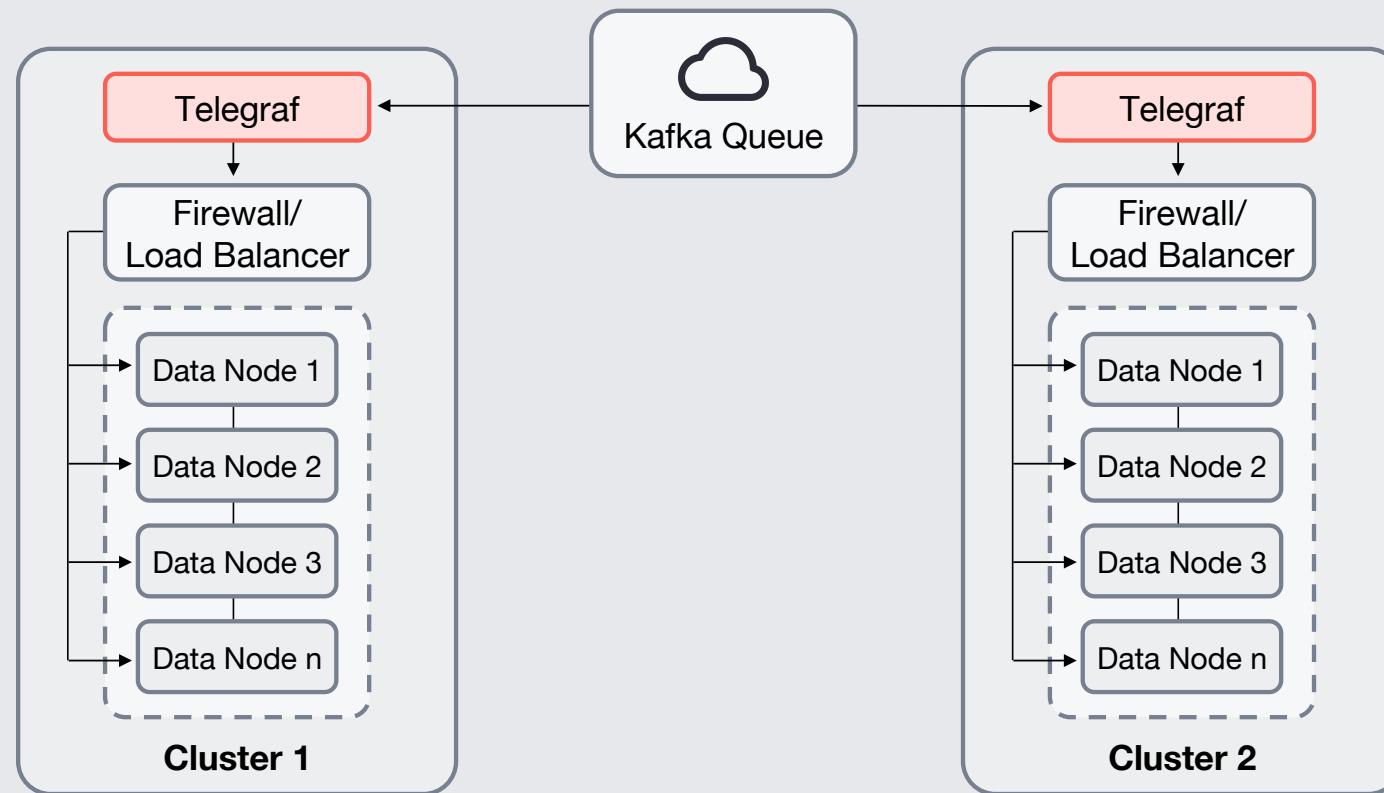
Generally there are two types of data that we care about replicating:

- New Data – Data which is coming from our raw sources
- Derived Data – The output of a SELECT INTO, or TICK script

Replication of New Data – Pattern 1

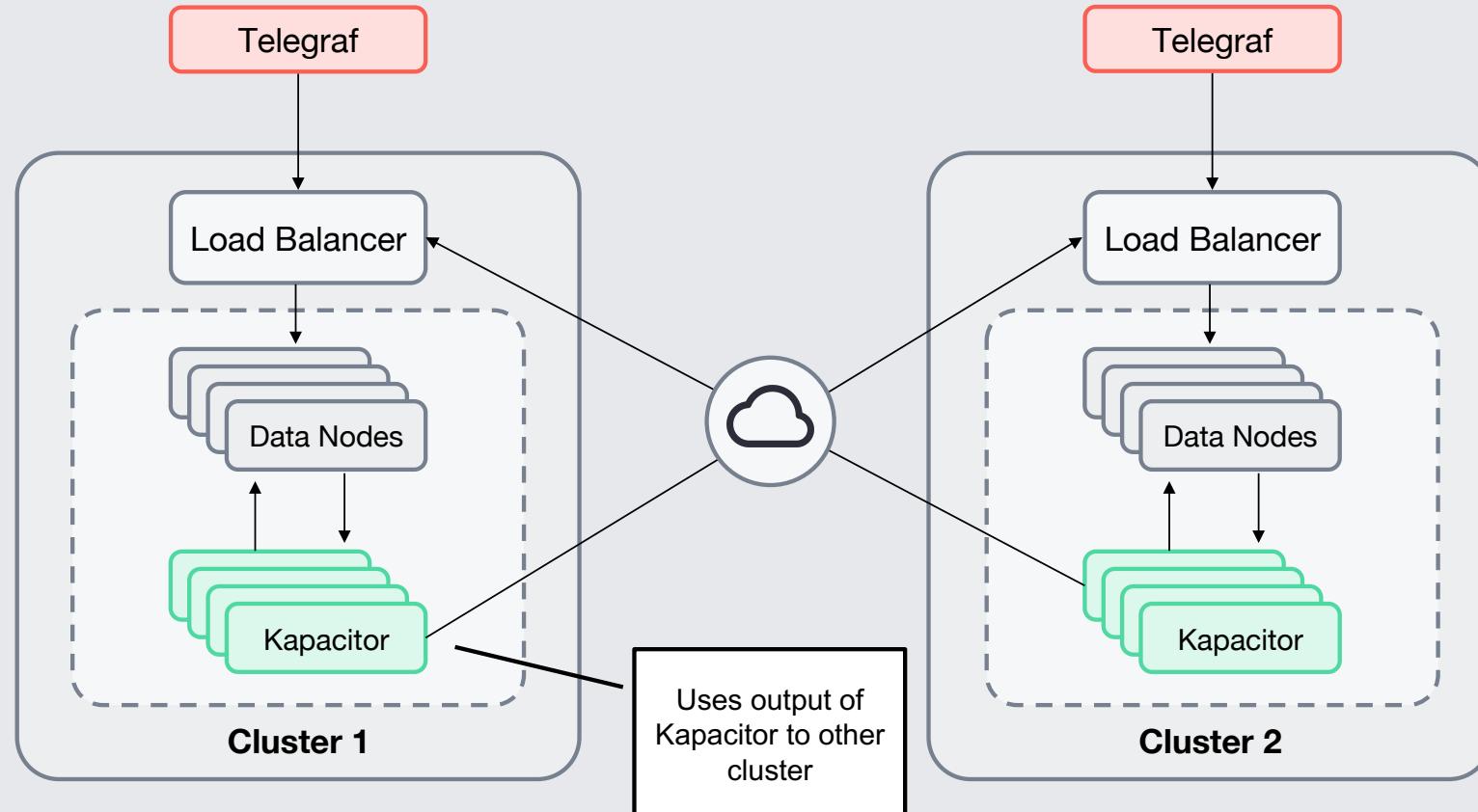


Replication of New Data – Pattern 2



1985-10-25T19:20:07.5

Replication of Derived Data – Pattern 3



General Advice

General Cluster Advice

- Batch your writes!
- The number of data nodes should be a multiple of your replication factor
- Use a single node of InfluxDB to monitor your cluster
- Put a load balancer in front of each of your data nodes
- Higher replication factors result in higher query concurrency, but higher write latency.
- Use Fine Grained Authorization instead of multiple databases



influx | **days**

Thank You!



influx | days