

ClearView

Technology Overview



Data Platform

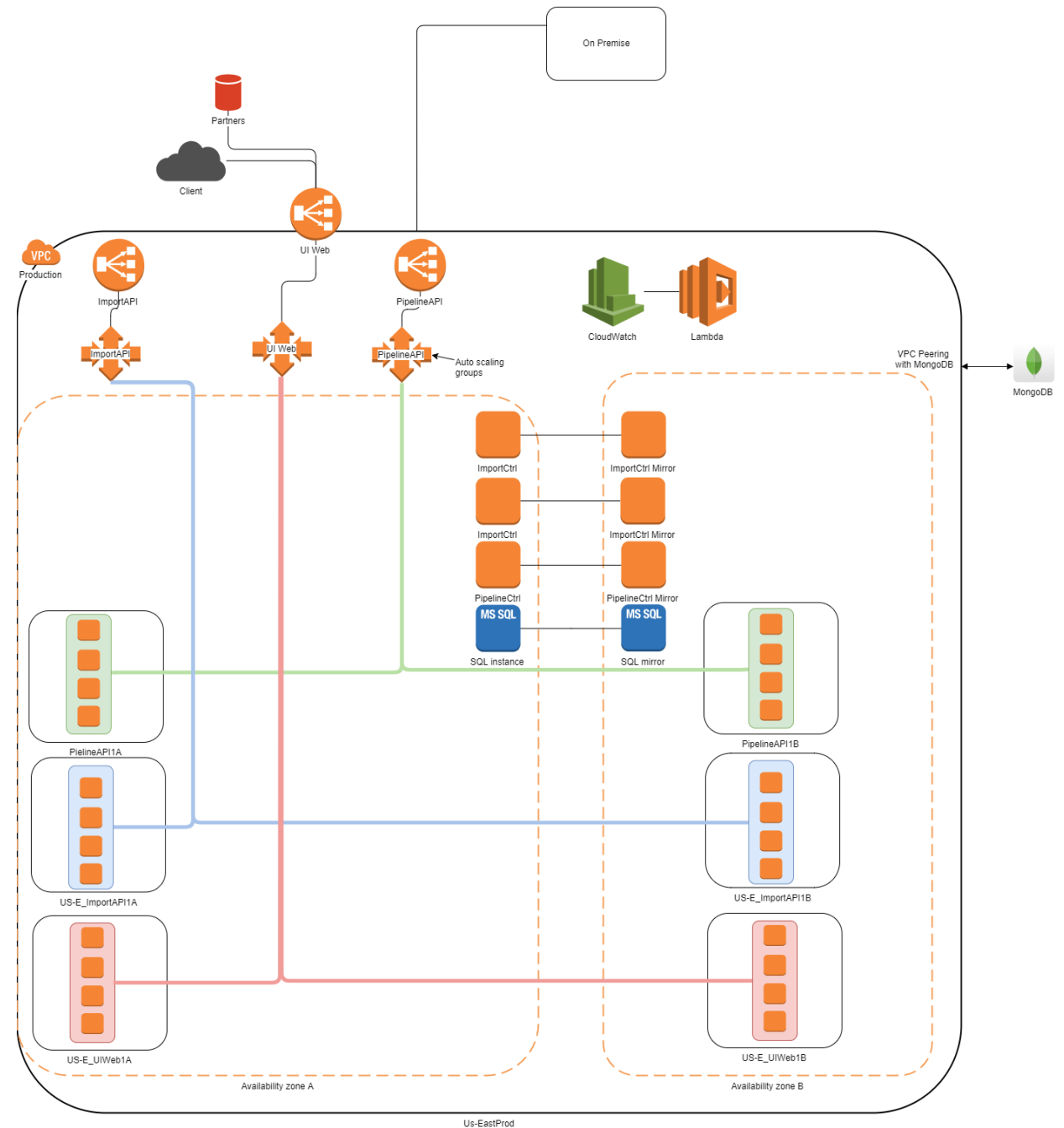
AWS Environment

Server Clusters

- UI/Data Visualization cluster, Pipeline cluster, and integration cluster
- Virginia, Frankfurt, Sydney
- CodeDeploy for Continuous Integration

Redundancy

- Autoscaling instances across multiple Availability Zones
- No single point of failure
- MongoDB replica sets across AZs



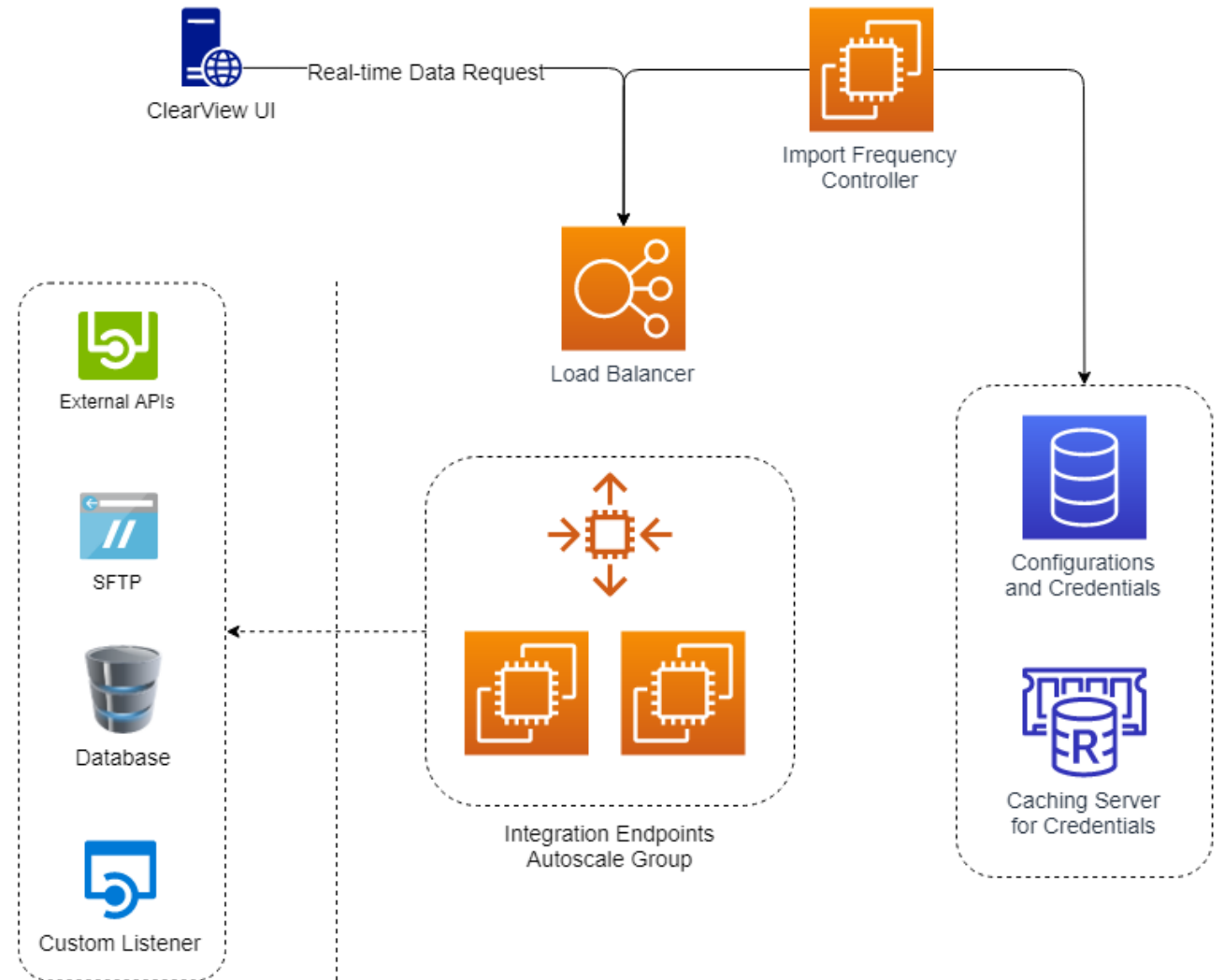
Data Ingestion

Data Import Process

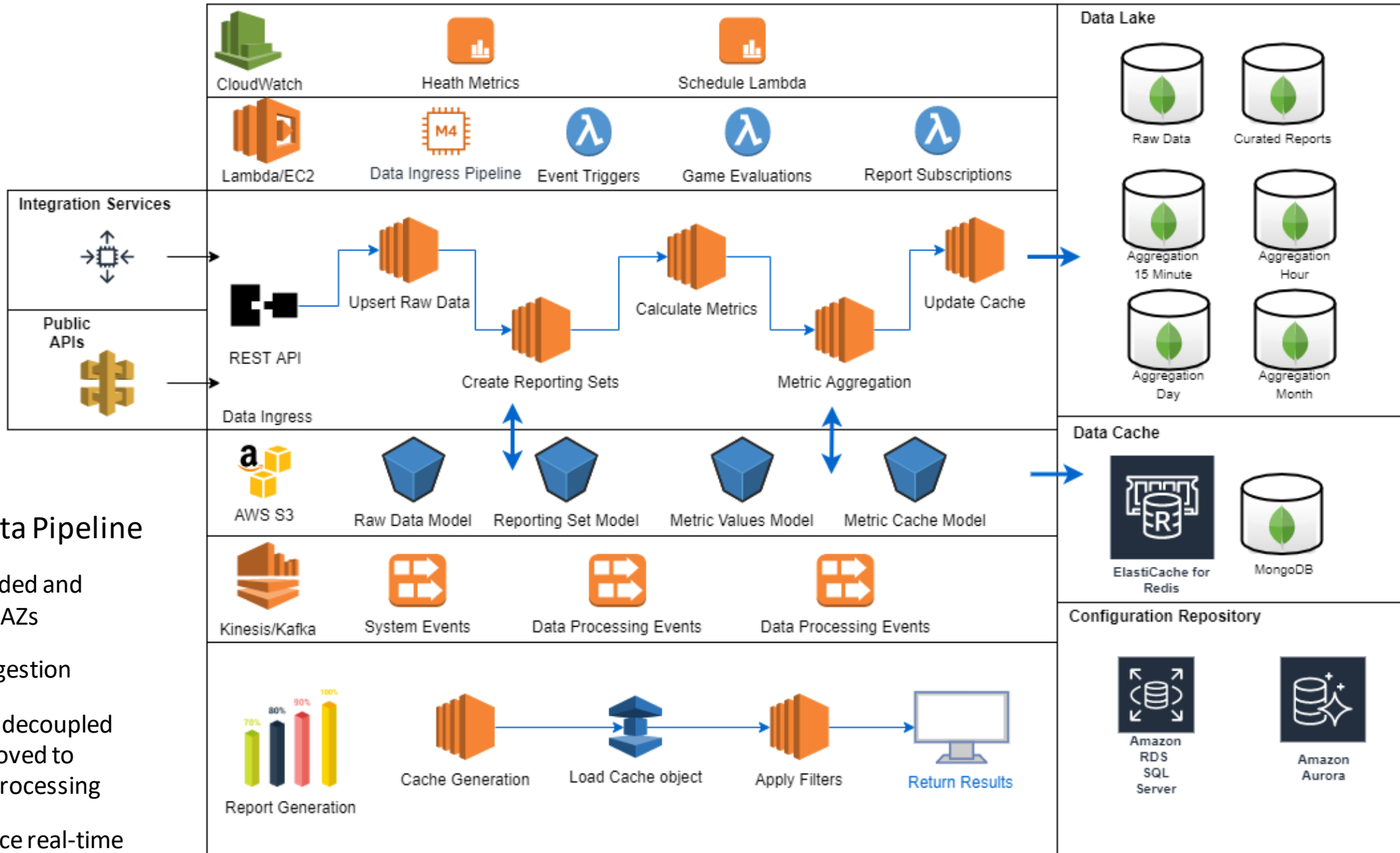
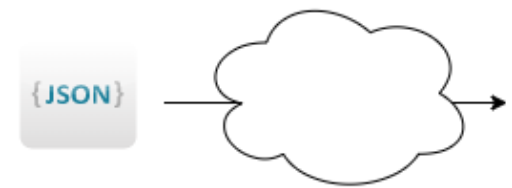
Deployment Diagram

Data Import Pipeline

- Database Driven Configuration
- API-based Integrations
- Real-time endpoints
- Autoscale Clusters
- High-Availability
- Fully configurable through front-end UI tools



Data Pipeline Architecture



Database Server and Data Pipeline

- MongoDB DBMS sharded and clustered across AWS AZs
- Queue-driven data ingestion
- Data transformations decoupled from database and moved to Pipeline multi-stage processing
- Caching layer to surface real-time data visualizations



Data Ingestion Pipeline

Definitions

- All data definitions configurable through the UI
- Definitions stored in JSON

Data Feeds

- Raw data matches customer's format
- Primary Key identified to support Upserts

Reporting Sets

- Supports Joins on data feeds
- Identifies hierarchy and reporting dates
- Can be pulled as reports

Metrics

- Configured based on Reporting Sets
- Stored in multiple aggregation levels
- Updates caching layer
- Enables events on data update

Platform Improvements

System Performance

- MongoDB and AWS Auto Scale Clusters provide consistent performance under load
- Excess capacity on clusters and individual servers is eliminated – better load balancing
- Improved caching layer and database technologies cut report load times significantly (from 20+ seconds for a Quarter's data to 1-2 seconds)

High Availability

- All web servers are clustered across multiple Availability Zones
- MongoDB replicates all data across multiple servers. Reports can be read from any server.
- Full redundancy even in the event of an entire data center outage

Data Improvements

- All integrations can be configured and monitored through the front-end web application
- Data stays in the customer's format, simplifying custom integrations by customers
- One data record can generate 10-20 metrics, as opposed to 1 or 2 currently