Polyglot & Hybrid Persistence Architectures for Big Data Analytics

Infocom World Conference, Athens
26 November 2019
Hybrid Data Persistence Architectures
Challenges

- Design
- Deployment
- Querying
- Monitoring & Analytics
- Evolution
Polystore Design

- Develop a tailored modelling language (TyphonML)
- Support for capturing availability, consistency & partitioning requirements
- Supporting modelling tools
- Extensibility to accommodate future data stores
Polystore Deployment

- Develop a language for modelling hybrid polystore deployments (TyphonDL)
- Develop automated TyphonML→TyphonDL transformations
- Develop automated transformations of TyphonDL models to installation & configuration scripts
Polystore Querying

- Unified language for querying data in polystores
- Compilation to native querying facilities
- Static analysis, lazy loading and prefetching

```
find c:Customer where
  c.orders.totalAmount.sum() > 1000
  and
  c.reviews.comment.responses.size() > 20
```
Polystore Monitoring & Analytics

- Record data access and update events
- Publish data events to distributed ledger
- Facilitate subscription-based data analytics & text mining pipelines
Polystore Evolution

- Support for schema evolution
- Support for internal data migration
- Query migration tools
- Data usage monitoring tools to recommend evolutions
Technology Stack

- Java as the main programming language
- Eclipse as the base modelling & development platform
  - TyphonML, TyphonDL, TyphonQL editors & tools
- Docker for container management
- OpenAPI for REST APIs
- Open-source software under the Eclipse Public License 2.0
OTE’s Contribution in Typhon

- **Data for Predictive Maintenance and QoS**
  - Technology Requirements Designation
  - Big Data Provision
  - Typhon’s Technology Deployment/Evaluation

- **Dissemination and Communication Activities**
  - Publication of project news in internal networks
The Use Case ...

Why Predictive Maintenance and QoS?

- Minimize Maintenance Costs
- High Availability
- Service Reliability

AI-based Operations’ Ecosystem
The System
(Use Case Description – CRM System Functionality)

- **Call Centers**: Manage all incoming calls & e-services for all sales.

- **Order Management**: Keep track of order progress & provide accordingly updates to supply chain.

- **Campaign Management**: Organize campaigns over the phone, SMS and email for various business needs.

- **Fault to Repair Management**: Monitor faults until final resolution of customer service’s interruption.
Use Case Objectives

Descriptive Reporting for Prediction Accuracy

Identification of:

- Quantitative Variables
- Categorical Variables
- Correlations
- Anomalies

Predictive Maintenance for improved QoS

Implementation of:

- Prediction Models
- Training/Testing
- Evaluation
Our Approach ... in *Steps*

- **Monitoring Tools**
  - Dynatrace
  - eG Enterprise
  - Event Viewer Logs

- **Big Data**

- **Big Data (after TYPHON integration)**

- **Analysis**
  - Anomalies
  - Correlations
  - Descriptive Model
  - Predictive Model

- **Predictive Maintenance**

We are here

We want to go here with TYPHON
Information Flow

Data Sources
- Human-generated Data
- Synthetic Data
- Agent Data
- Machine Data
- Wire Data

Collect
- beats
- dynatrace

Preprocess
- Kafka
  - TopicName
  - TopicName
  - TopicName

Store/
Process/
Analyze
- HDFS
- Typhon
Thank you!