IOT & Big Data: The Future
Information Processing Architecture

Dr. Michael Faden
Dirk Weise
Oder:

Was mach’ ich nur mit den ganzen Daten?
Agenda

1. Introduction
2. Terms and Definitions
3. Reference Architecture
4. Architecture Building Blocks
5. Architecture Patterns
6. Solution Building Blocks
1. Introduction
2. Terms and Definitions
3. Reference Architecture
4. Architecture Building Blocks
5. Architecture Patterns
6. Solution Building Blocks
**Discussions about Big Data**

- **Is our use case Big Data or not?**
- **What shall we do with all that Sensor Data?**
- **Do we have to use Hadoop or NoSQL for Big Data?**
- **Do we miss out if we don’t do Big Data?**

---

2014 © Trivadis
.. an Architecture View
25.06.2015
Understand Drivers for Big Data Undertakings

A business scenario bridges three drivers end to end:

- Opportunity – There is something one could do.
- Capability – There is something we can do.
- Business demand – There is something we need to do.
**Business Value:** Investment in Big Data, be it in terms of personnel or technology, has to be rectified by convincing business demand.

**Integration:** Big Data sources have to be integrated with the enterprise architecture – semantically, technologically and procedurally.

**Governance:** Big Data must not dilute the enterprise information model. Business processes have to be aligned to data source properties (e.g. topicality, reliability).
Business Value: Investment in Big Data, be it in terms of personnel or technology, has to be rectified by convincing business demand.

Integration: Big Data sources have to be integrated with the enterprise architecture – semantically, technologically and procedurally.

Governance: Big Data must not dilute the enterprise information model. Business processes have to be aligned to data source properties (e.g. topicality, reliability).
..how to avoid this?
Separate Concepts from Implementation

Required capability
- Conceptual – requirements capture and high-level solution
- Independent from technology, organisation and processes

Provided capability
- Concrete – detailed solution and implementation
- Specific to technology, organisation and processes
Agenda

1. Introduction
2. Terms and Definitions
3. Reference Architecture
4. Architecture Building Blocks
5. Architecture Patterns
6. Solution Building Blocks
Terminology

- **Architecture**
  - Fundamentals of a system
  - Context, structure, abstraction, process

- **Architecture View**
  - System seen from the perspective of specific concerns
  - Comprehensible, comprehensive

- **Building Block**
  - Potentially reusable component of an architecture
  - Relevant, interrelated, fractal

ISO/IEC 42010
Systems and software engineering – Architecture description

TOGAF Version 9.1
Definitions
Our Approach

• Reference architecture. The canvas for our architecture
• ABBs. Commonly found requirements
• SBBs. Commonly applied solutions
• Views. Tailored to stakeholders
• Patterns. Commonly found requirement/solution profiles
• Method. Guideline to populate the canvas including consistency checks
Agenda

1. Introduction
2. Terms and Definitions
3. Reference Architecture
4. Architecture Building Blocks
5. Architecture Patterns
6. Solution Building Blocks
Reference Architecture

Governance

Data Sources → Data Ingestion → Data Factory & Managed Enterprise Information → Information Provisioning → Information Query & Visualisation

Discovery Lab

Organisation
Agenda

1. Introduction
2. Terms and Definitions
3. Reference Architecture
4. Architecture Building Blocks
5. Architecture Patterns
6. Solution Building Blocks
Agenda

1. Introduction
2. Terms and Definitions
3. Reference Architecture
4. Architecture Building Blocks
5. Architecture Patterns
6. Solution Building Blocks
Architecture Patterns – Schema on Write

1. Data Sources
   - Data Source Connectivity & Capturing
   - Raw Data at Rest

2. Data Ingestion
   - Raw Data in Motion
   - Factory
   - Managed Information

3. Data Factory & Managed Enterprise Information
   - Managed Information

4. Information Provisioning

5. Information Query & Visualisation

6. Discovery Lab

7. Organisation

---

2014 © Trivadis
25.06.2015
Architecture Patterns – Classic BI

Data Sources
- Structured Data Sources
- Master & Reference Data Sources

Data Ingestion
- Data Source Connectivity & Capturing

Data Factory & Managed Enterprise Information
- Factory
- Managed Information

Information Provisioning
- Access & Performance Layer
- Query (pull)

Information Query & Visualisation
- Prebuilt & Ad-hoc BI Assets

Discovery Lab

Governance
- Metadata Management
- Master Data Management
- Information Quality & Accountability
- Security
- Legal Compliance

Organisation
- BI Competence Centre
- IT Operations
- Business Stakeholders
Architecture Patterns – Classic BI (in its own words)
Architecture Patterns – Schema on Read

Data Sources

Data Ingestion

Data Factor & Managed Enterprise Information

Governance

Information Provisioning

Information Query & Visualisation

Data Source Connectivity & Capturing

Raw Data in Motion

Discovery Lab

Raw Data at Rest

Data Factory & Managed Enterprise Information

Organisation
Architecture Patterns – Lambda Architecture

Data Sources
- Data Ingestion
  - Data Source Connectivity & Capturing
- Data Ingestion
  - Raw Data at Rest
  - Raw Data in Motion
- Data Factory & Managed Enterprise Information
  - Managed Information
- Data Factory & Managed Information
  - Managed Informations
- Information Provisioning
  - Access & Performance Layer
- Information Provisioning
  - Discovery Lab
- Organisation

Governance

2014 © Trivadis
Architecture Patterns – Discovery Lab

Data Sources

- Data Ingestion
  - Data Source Connectivity & Capturing

Data Factory & Managed Enterprise Information

- Managed Information Layer

Information Provisioning

Information Query & Visualisation

- Discovery Lab Sandboxes
  - Advanced Analysis & Data Science Tools

Integration with various stakeholders:

- BI Competence Centre
- IT Operations
- Business Stakeholders

Raw Data at Rest → Raw Data in Motion

Governance
Agenda

1. Introduction
2. Terms and Definitions
3. Reference Architecture
4. Architecture Building Blocks
5. Architecture Patterns
6. Solution Building Blocks
Solutions – Data Ingestion

**Architecture**

- Data Source Connectivity & Capturing
- Raw Data at Rest
- Raw Data in Motion
- Data Factory

**Solutions**

**An RDBMS Way (synchronous)**

- DB Upsert
- DB Triggers
- ... (Staging Tables)

**An MS Way (asynchronous)**

- EventHub
- SQLServer
- StreamInsight

Similar:
- CDC
- ESB
- Kafka & Storm
Solutions – Data Factory (Integration)

Architecture

Solutions

To RDBMS

To Hadoop

To NoSQL

Data Store (RDBMS)

Data Store (Hadoop)

Data Store (NoSQL)

DB Link

JDBC

SQOOP, JDBC

Hive, Pig, Drill

SPARQL

Data Store (RDBMS)

Data Store (Hadoop)

Data Store (NoSQL)

SPARQL

SPARQL

SPARQL

2014 © Trivadis
25.06.2015
Solutions – Data Factory (Interpretation)

Architecture

- Raw Data
- MEI
- Parsing
- NLP
- OCR
- Normalisation

Solutions

Schema Application

- Raw Data
- MEI
- Matching Algorithms
- Reference Data
- Operational Data

Identity Resolution

- Raw Data
- MEI
- Cleansing Algorithms
- Reference Data

Data Cleansing

- Raw Data
- MEI
Solution Building Block

Data Sources

Governance

Data Ingestion

Azure HDInsight

Power BI

ML Studio

Information Query & Visualisation

Information Provisioning

Data Factory & Managed Enterprise Information

Organisation

2014 © Trivadis

.. an Architecture View

25.06.2015
Questions & Answers

Dr. Michael Faden
Senior Consultant
Tel.: +49 162 291 47 74
michael.faden@trivadis.com

Dirk Weise
Senior Consultant
Tel.: +41 79 909 72 48
dirk.weise@trivadis.com