

KOMMUNIKATION ZUR LAMBDA ARCHITEKTUR

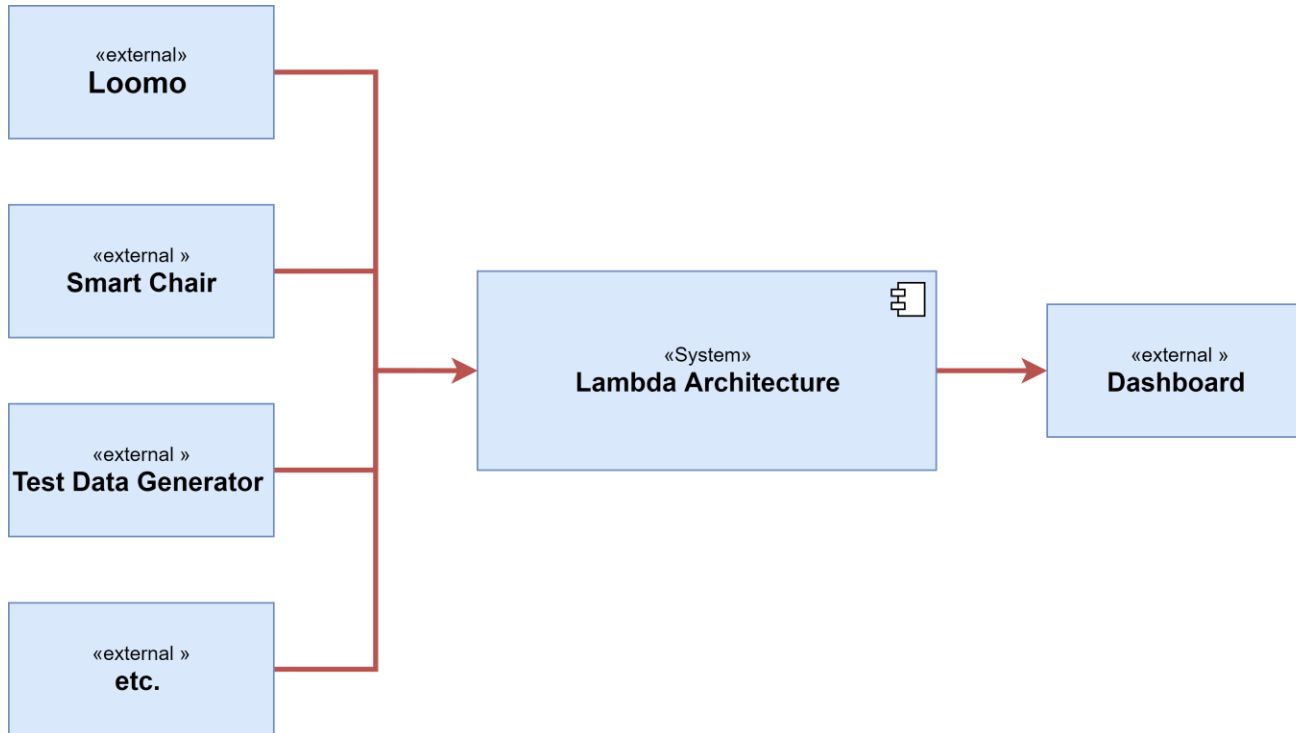
Mike Wüstenberg

Was ist eine Lambda
Architektur?

Datenaufnahme über Apache
Kafka

Datenformat der Nachrichten

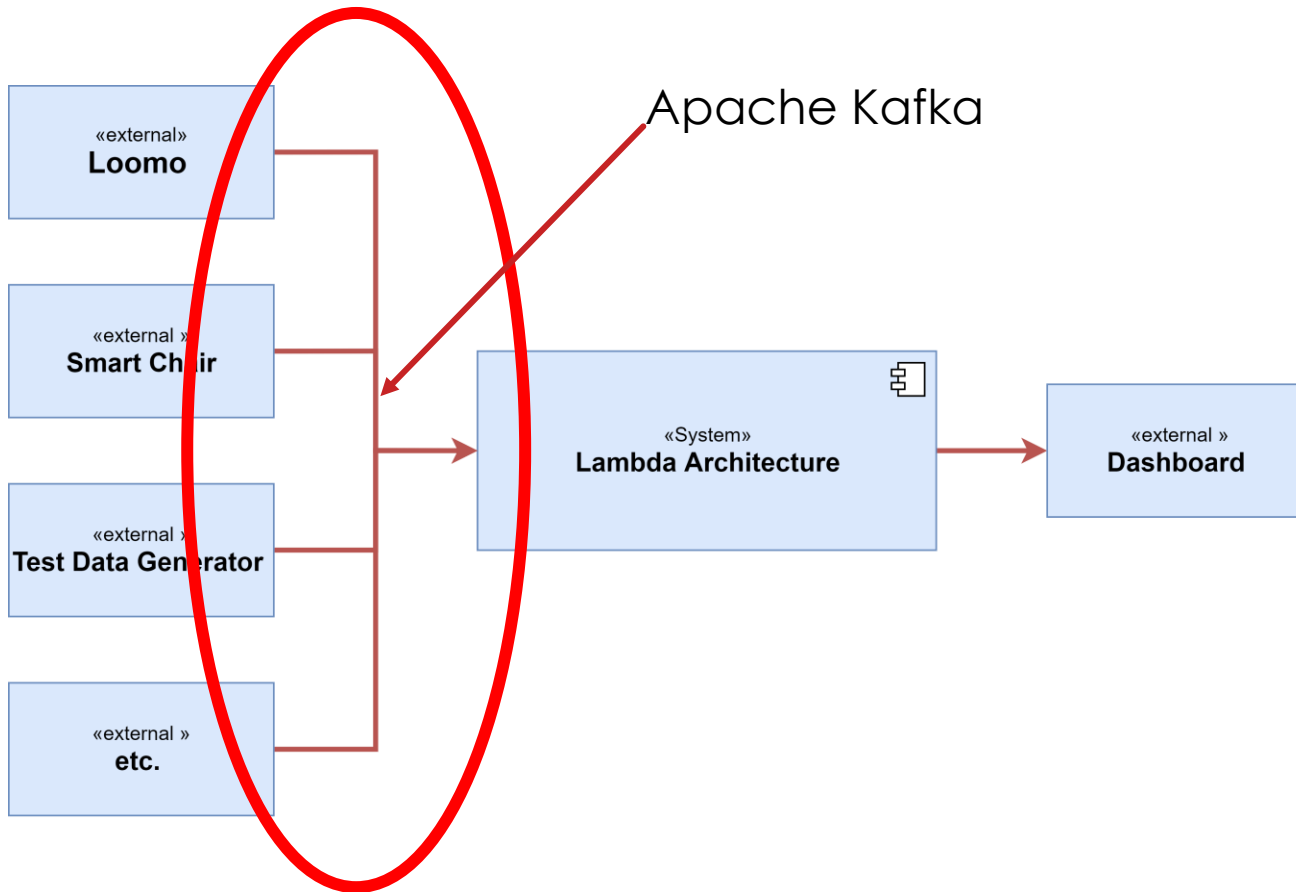
ÜBERBLICK



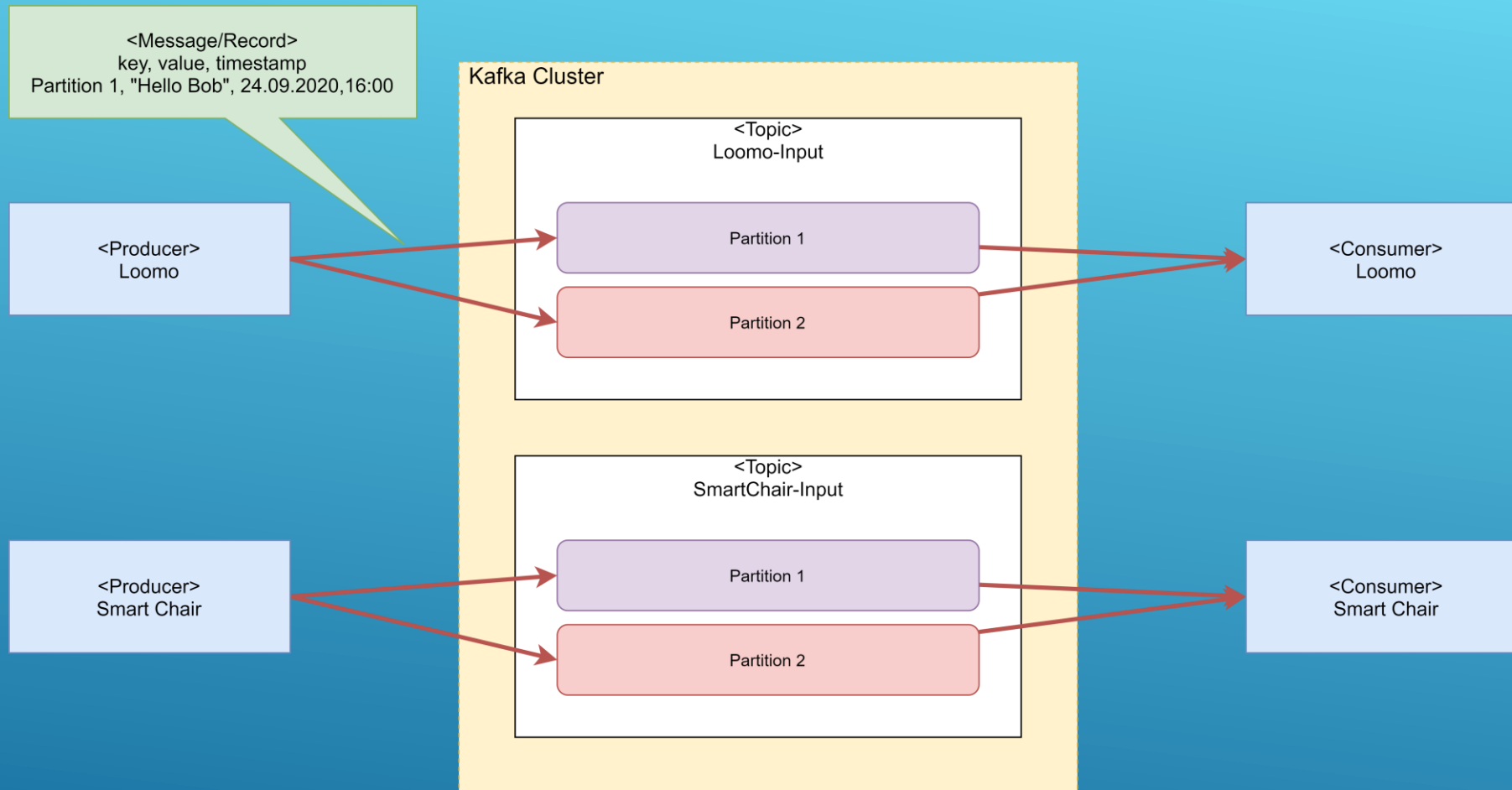
LAMBDA ARCHITEKTUR

- Datenverarbeitungs Architektur
- Sehr große Datenmengen
- Scalability
- Fault tolerance
- Kombiniert Batch und Streaming Technology

DATENAUFNAHME ÜBER APACHE KAFKA



- Verarbeitung von Datenströmen
- Datenübertragung von Nachrichten(Records)
- Verbindung zur Lambda Architektur
- Clients in unterschiedlichen Sprachen
 - C/C++
 - Python
 - Java
 - Erlang



- Viele Topics in einem Kafka Cluster
- 1 Topic pro Producer
- Partition erlaubt lesen/schreiben von mehreren Verbrauchern
- Key: Partition Auswahl
- Value: Daten
- Timestamp

APACHE KAFKA TOPICS

```
implementation group: 'org.apache.kafka', name: 'kafka-clients', version: '2.5.0'  
implementation group: 'org.apache.avro', name: 'avro', version: '1.10.0'
```

```
private static Producer<Long, TestData> createProducer() {  
final Properties props = new Properties();  
    props.put(ProducerConfig.BOOTSTRAP_SERVERS_CONFIG, 192.168.80.139:29092);  
    props.put(ProducerConfig.CLIENT_ID_CONFIG, data-generator-1);  
    props.put(ProducerConfig.KEY_SERIALIZER_CLASS_CONFIG, LongSerializer.class.getName());  
    props.put(ProducerConfig.VALUE_SERIALIZER_CLASS_CONFIG, AvroSerializer.class.getName());  
    return new KafkaProducer<>(props); }  
}
```

```
final Producer<Long, TestData> producer = createProducer();  
try {  
    final ProducerRecord<Long, TestData> record =  
        new ProducerRecord<>(TOPIC, testDataMessage);  
  
    final RecordMetadata metadata = producer.send(record  
} catch (ProducerFencedException |  
    OutOfOrderSequenceException | AuthorizationException e) {  
    producer.close();  
} finally {  
    producer.flush();  
    producer.close();  
}
```

- Apache Avro
 - Serialisierungs Framework
 - Kompaktes Binärformat
 - Schema definiert in JSON
 - Typisierung und Pflichtfeldern
 - API
 - Java
 - C/C++
 - C#
 - Python

```
{ "namespace": "producer.avro",  
  "type": "record",  
  "name": "TestData",  
  "version": "1",  
  "fields": [  
    { "name": "sendereName", "type": "string",  
      "doc": "Name of the sender" },  
    { "name": "senderType", "type": "string",  
      "doc": "Type of the sender" },  
    { "name": "message", "type": "string",  
      "doc": "Type of the sender" }  
  ]  
}
```

DATENFORMAT DER NACHRICHTEN

- Funktioniert das auf allen Geräten?
- Fehlen noch wichtige Informationen?

FRAGEN?

- <https://avro.apache.org/docs/1.10.0/index.html>
- <https://kafka.apache.org/documentation/#introduction>
- Big Data -PRINCIPLES AND BEST PRACTICES OF SCALABLE REAL-TIME DATA SYSTEMS - NATHAN MARZ with JAMES WARREN

QUELLEN