Academic cloud computing interoperability use cases

CloudWatch Concertation Meeting

Brussels, 13. March 2014

Wolfgang Ziegler
wolfgang.ziegler@scai.fraunhofer.de

Fraunhofer
SCAI

cloud4health project

- One of the 14 BMWi funded Trusted Cloud projects
- Cloud Services for Big Data Analysis in Medicine

- 5 partners including
  - 1 SME
    - Textmining, Coordination
  - 1 University
    - Data Provider, Clinical Bus Architecture
  - 1 Clinic
    - Data Provider, Clinical Bus Architecture, Transfer Database, Test of the Trusted Cloud
  - 1 Research Institute
    - Cloud-Expert, Text- und Data-Mining
  - 1 Service Provider
    - Data Protection, Legal Compliance

Trusted Cloud

Gefördert durch:
Bundesministerium
für Wirtschaft
und Technologie
aufgrund eines Beschlusses
des Deutschen Bundestages

averbis

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

RHÖN-KLINIKUM AKTIENGESELLSCHAFT

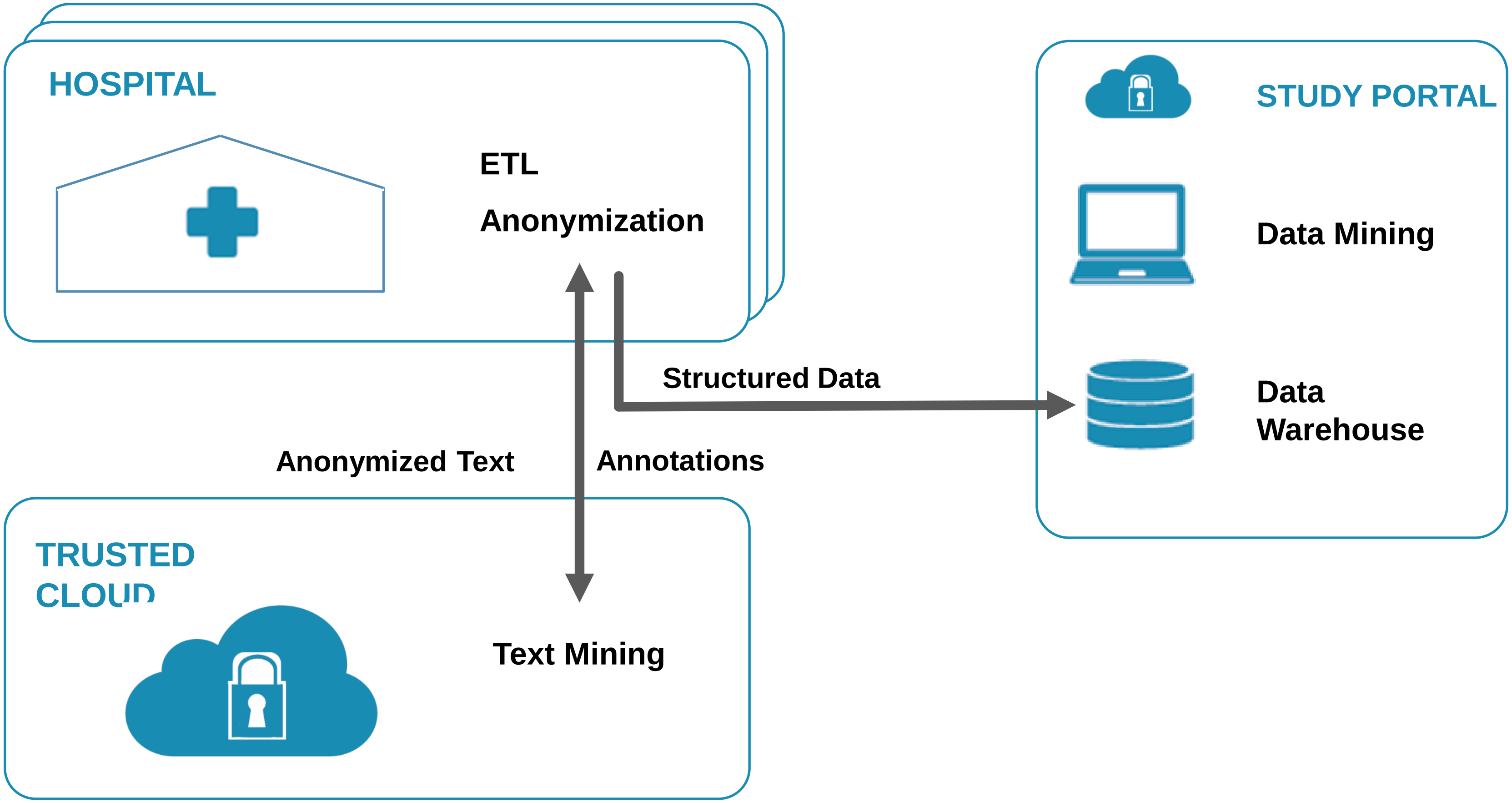
Fraunhofer SCAI

TMF

cloud4health objectives

- Determine the requirements of a trusted Cloud
  - Middleware and management layer
  - Security and data protection
  - IInterface for the researchers in clinics
- Implement a dynamically scalable framework for text mining
  - Capacity and performance determined by the individual studies
    - Number of instances and degree of parallelism
- Automated configuration and startup triggered by researcher in clinic
  - Mapping of study properties to infrastructure
- Dynamic encryption of data before sending to the Cloud, decryption in memory prior to processing
- No persistent copies of the data in the Cloud
  - Patient data kept in memory only
- Structured results in a standardised format for further analysis
  - ODM - Operational Data Model
    - Data model for archiving and exchanging Data and Metadata in the area of clinical research

cloud4health building blocks



HOSPITAL

ETL

Anonymization

Structured Data

Anonymized Text

Annotations

TRUSTED CLOUD

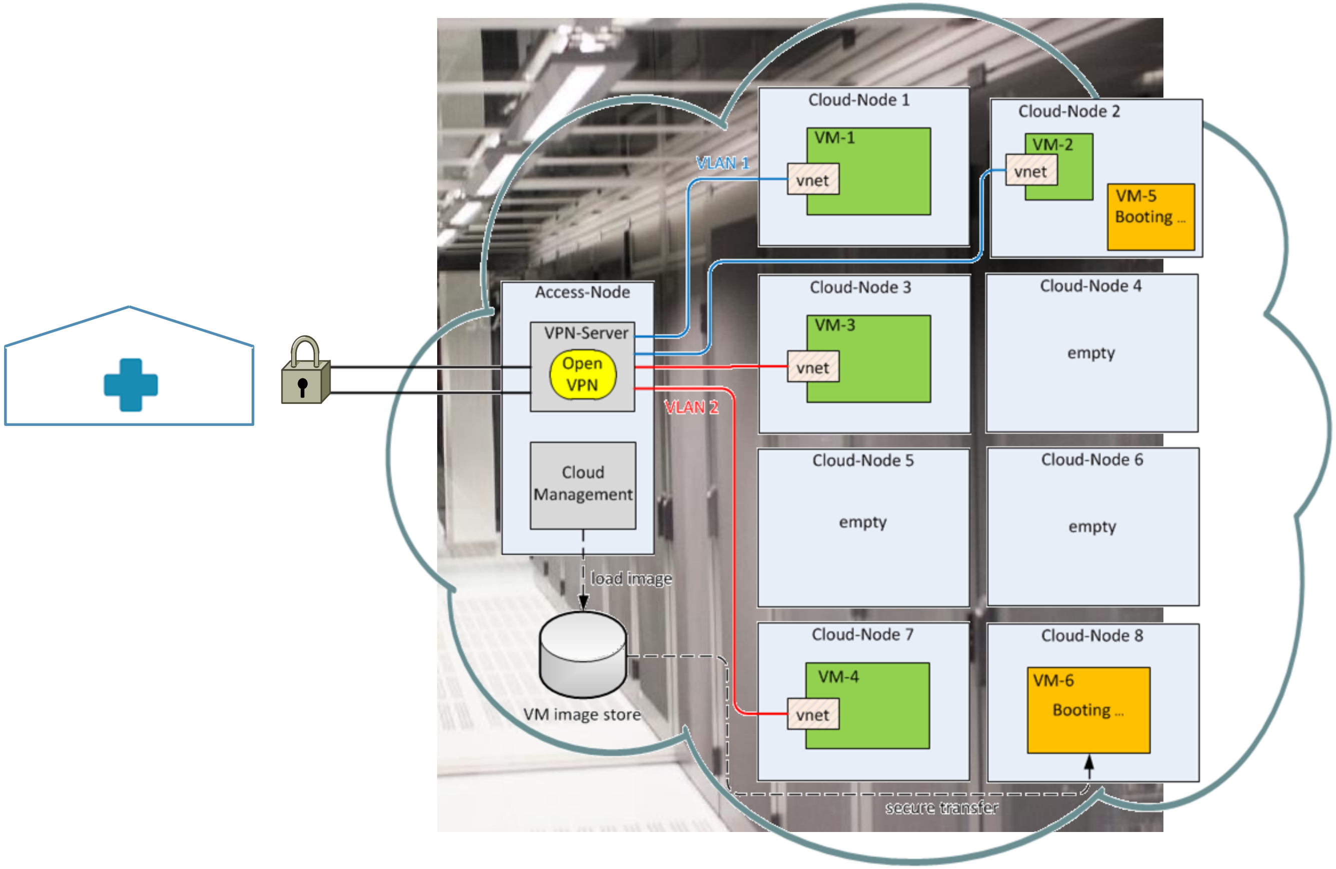
Text Mining

STUDY PORTAL

Data Mining

Data Warehouse

Cloud architecture



Access-Node

VPN-Server

Open VPN

Cloud Management

load image

VM image store

VLAN 1

VLAN 2

Cloud-Node 1

VM-1

vnet

Cloud-Node 2

VM-2

vnet

VM-5 Booting ...

Cloud-Node 3

VM-3

vnet

Cloud-Node 4

empty

Cloud-Node 5

empty

Cloud-Node 6

empty

Cloud-Node 7

VM-4

vnet

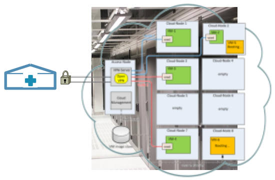
Cloud-Node 8

VM-6

Booting ...

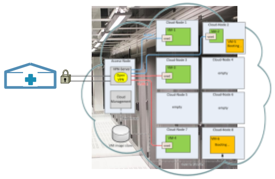
secure transfer

cloud4health interoperability requirements – trusted Cloud



- Dynamic set-up of Cloud infrastructure by clinics requires interoperable interfaces
- Same for study-dependant deployment of text mining services by the clinics
- Same for shutting down the entire infrastructure and secure deletion of all VMs
- Trusted Cloud computing requires more than a technical implementation
  - Well defined and agreed upon processes to assure data protection and legal compliance are equally important
- Current cloud4health prototype based on manually achieved agreements on processes as part of the Cloud service
  - Providing a blueprint for minimum requirements
- Manual agreements should be replaced by electronic Service Level Agreements between Cloud provider and clinics
  - Defining QoS, data protection and processes
- Entire cloud4health middleware deployable inside the clinics as private Cloud solution

cloud4health interoperability requirements – data protection



- Data access in clinics most often based on userid/password credentials today
  - Same authentication mechanisms used for Cloud and service management
  - Need for more secure, standardised authentication and authorisation, e.g. X.509 certificates
- Secure tunnel between clinic and Cloud
- Standardised processes for data encryption/decryption on the fly
  - E.g. based on a hybrid approach with shared keys and asymmetric keys
- Trade-off between key validity period and security
- Trustworthy and secure key management in the Cloud

Study Portal

- After text mining data is further processed in the hospitals with respect to the study goals
- Service point for multiple customers, operating a private Cloud for storing results of studies and executing further analysis
- Providing access to data resulting from previous text mining and analysis in the hospitals based on demand of customers
- Data is encrypted for each customer using X.509 PKI infrastructure
- Access through secure authentication and authorisation

Service Level Agreements – data protection and data placement

- Cloud customers need the possibility to define the protection of their data in Clouds as part of their dynamic electronic Service Level Agreements
- The OPTIMIS SLA enabled the customer to specify the level of protection
  - Geographical location of data storage and data processing
    - Restricting location e.g. to data centres in countries in the European Data Protection Area
    - Specifying encryption of data and the strength of the encryption
- Specifying procedures to be followed when the Cloud infrastructure is no longer needed (but before the end of the contract)
  - How to return the data
  - How to erase the stored data after returning
- SLAs requests defining data protection can be used to preselect Cloud providers
  - E.g. Service Manifest as developed in OPTIMIS

Service Level Agreements – certification

- SLAs should include relevant certifications of a data centre, e.g.
  - Conformance to ISO defined processes, e.g.
    - ISO 27001
  - Eco-efficiency certificates, e.g.
    - EnergyStar Rating
    - ISO14000
- Certification information needs
  - to be electronically accessible,
  - to have a limited lifetime based on the certification frequency
  - and should be signed by a trusted party
- SLAs requests including certification requirements can be used to preselect Cloud providers
  - E.g. eco-efficiency in the OPTIMIS Service Manifest

More requirements addressed in OPTIMIS SLAs

- Creation and negotiation of dynamic electronic SLAs must be based on standards to achieve interoperability and to empower the customer to understand and compare the offerings of different Cloud providers.
- Need for Standardized languages for expressing service description terms, service level objectives and KPIs to request and negotiate SLAs covering the same service levels from different providers prior to selecting a provider.
- Among other service terms not included in today's SLAs
  - the geographical location of a data centre, e.g. DPA, should be part of the SLA
  - also Standard Contractual Clauses, Binding Corporate Rules and IPR statement

As a consequence (but not realised in OPTIMIS)

- Need mechanisms allowing the customer to verify the geographical location of the resources provided at run-time
  - Electronic certification backed by a trusted party, similar to CAs for X.509 certificates
  - Heuristics for automated checks