Chapter 3
Content Centric Storage and Current Storage Systems

Michael C. Jaeger
Siemens AG, Corporate Technology, Germany

Uwe Hohenstein
Siemens AG, Corporate Technology, Germany

ABSTRACT
Content-centric storage represents an approach for handling large amounts of data. It is one of the innovations pursued by the VISION Cloud project. The goal of the VISION Cloud project is the development of an industry grade storage system using cloud technology. The envisaged use of the VISION Cloud involves the storage and management of millions of data items, potentially several hundreds of terabytes in size. On the one hand, the technical foundations must be capable of efficiently storing such an amount of data. On the other hand, the VISION Cloud must provide adequate means of an API for allowing the efficient navigation, search, and access for the right data item in this storage. For the latter purpose, VISION Cloud provides a data access layer, which is called “Content Centric Interface.” Applications can use this data access layer for accessing the VISION Cloud storage from a content-centric point of view, abstracted from actual storage representation. The content centric interface is different from existing cloud storage interfaces and is similar, from an architectural point of view, to object relational mapping frameworks for traditional applications with relational database systems.

INTRODUCTION
Content centric storage represents a novel research topic. When investigating major search engines at the start of the VISION Cloud project, we have found a few hits referring to content centric storage, mostly referring to research efforts such as the CIMPLE project (cf. Delaet and Joosen, 2009).

DOI: 10.4018/978-1-4666-3934-8.ch003

The term content centric storage forms an analogy to the term content centric networking or content centric routing. In the area of networking and routing, the idea of content “centricness” comes from the need to perform routing efficiently, even with a rising amount of data. Content centric networking / routing refers to the idea that the content is not (directly) routed from the logical location of a source to the logical location of a recipient, but an additional layer of content publishers and
Content Centric Storage and Current Storage Systems

content sinks is placed above and routing efforts are performed more efficiently on this layer.

Among other authors, Ted Nelson’s work was one of the first that have discussed this idea (cf. Nelson 1988). The goal is to implement a more efficient handling of content. Modeling the problem of getting content from A to B from a content-perspective enables a couple of useful mechanisms that improve the efficiency of the system, among them are (cf. Koponen 2007):

- The system can cache content that is subject to delivery to multiple consumers. The system can react more efficiently to congestion, as the awareness of the content will allow for a more suitable rerouting/exception handling than with network-based routing.
- Security functions can be implemented on a content level at distribution, which appears more suitable than at a central or isolated location.
- The system can be seen as a store for content rather than relying on the actual network or storage structure that holds the information.

A storage service appears similar to a network service, as the semantics of the operations of a storage service and a message-oriented middleware have strong similarities (cf. Gray 1995). A message-oriented middleware can be seen as storage service to when considering that both a network service and a storage service have similar data sink and data source semantics. Also, a message-oriented middleware must offer similar characteristics as database servers do, when referring to the ACID paradigm (atomicity, consistency, isolation and durability) in the database world.

Consequently, we see content centric storage as an approach to provide a storage service at the content level that shall enable similar advantages as with content centric networking or routing. In the VISION Cloud project, the front most goals are simplified and efficient storage access. The basic idea for the content centric access in VISION Cloud is to provide a persistence service at the content-level, hiding specialties of the implementation, such as the storage location, server addresses, etc.

BACKGROUND: REQUIREMENTS FROM VISION CLOUD PROJECT

Besides the research background for content centric storage, the VISION Cloud project has also engineered requirements for the content centric access to the data. In general, the requirements emerge from the intended application use cases of the project, which intends to deliver applications in four domains: healthcare, media production, business intelligence applications, and telecommunications services. We identify the following main areas for requirements:

- The so named “Big Data” capability, the goal of which is to actually enable the management of big data.
- The capability of working with metadata, also enabling richer data models from the application domain.
- Efficient search access and retrieval.
- Execution of tasks on the storage to save bandwidth and round-trip latency. This involves analysis or metadata handling in order to provide statistical data or other functionalities on metadata.

When we look at the application use cases in the VISION Cloud project, we have had a couple of very concrete requirements that have outlined what content centric storage actually means. In the following, we summarize the demands from the application use cases for such an access layer.
Related Content

A Manpower Allocation Model for Service Jobs
[www.igi-global.com/article/manpower-allocation-model-service-jobs/68968?camid=4v1a](www.igi-global.com/article/manpower-allocation-model-service-jobs/68968?camid=4v1a)

Does Technology Acceptance Change the Way from CRM to Customer Loyalty?: An Empirical Study on the Banking Industry
[www.igi-global.com/chapter/does-technology-acceptance-change-way/65943?camid=4v1a](www.igi-global.com/chapter/does-technology-acceptance-change-way/65943?camid=4v1a)

SLA Management in Storage Clouds
[www.igi-global.com/chapter/sla-management-storage-clouds/77432?camid=4v1a](www.igi-global.com/chapter/sla-management-storage-clouds/77432?camid=4v1a)

Emerging E-Services in Accounting: A Longitudinal Case Study
[www.igi-global.com/chapter/emerging-services-accounting/6416?camid=4v1a](www.igi-global.com/chapter/emerging-services-accounting/6416?camid=4v1a)