



# Ventum Consulting

Hadoop im Unternehmenseinsatz, aber sicher

Nürnberg, November 2015



Results, no excuses.

**A** | Ventum Consulting

**B** | 5 Säulen der Sicherheit

**C** | Frameworks

**D** | Fazit

## Geschichte

---

**2005** gegründet

Von Beginn an unabhängig,  
inhabergeführt und erfolgreich.

## Team

---

**> 100** Mitarbeiter

Verbunden durch unseren Anspruch  
**Results**, no excuses.

## Standorte

---

**International**

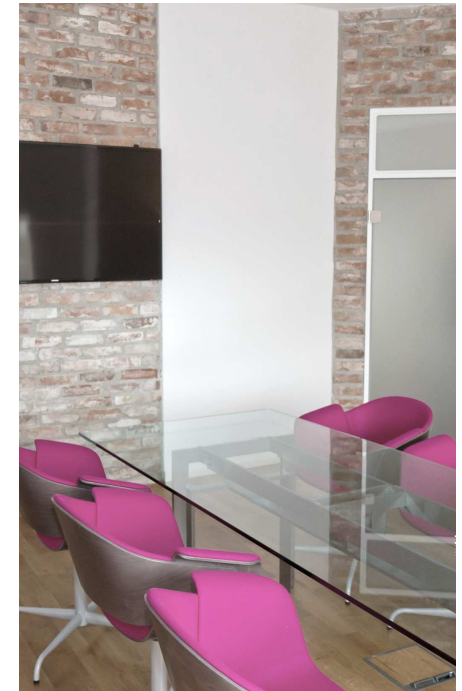
**München** | Wien | Zürich | Foshan

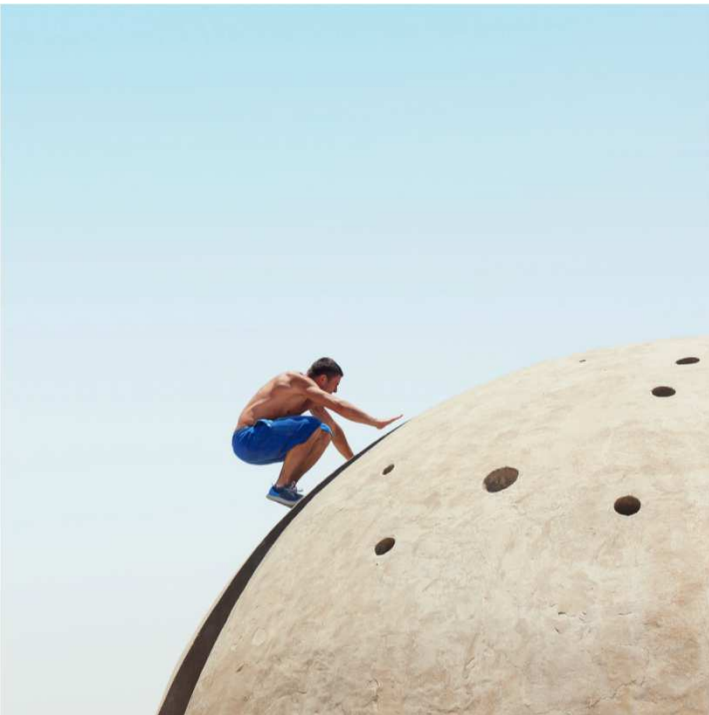
## Branchenfokus

---

**Automotive,  
Insurance, TelCo**

**Etablierter** Partner für komplexe  
Projekte





# B | 5 Säulen der Sicherheit

# 5 Säulen der Sicherheit

- Authentifizierung
- Autorisierung
- Auditing
- Integrität
- Vertraulichkeit

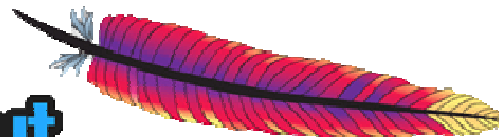
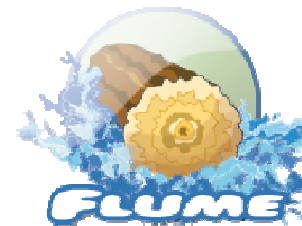
# Sicherheit bei Hadoop out of the Box



# Zum Superuser in 5 Sekunden

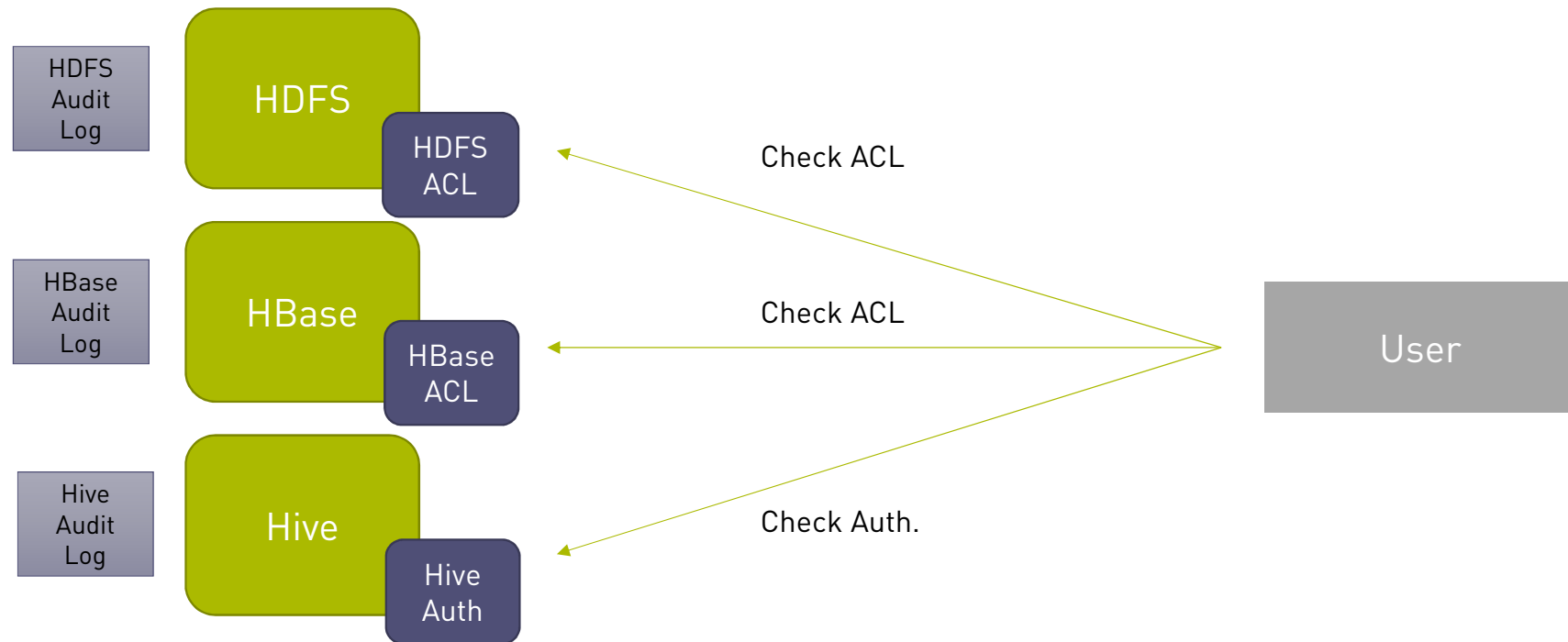
```
[student@ip-10-0-0-38 ~]$ whoami
student
[student@ip-10-0-0-38 ~]$ hadoop fs -ls /user/hive
ls: Permission denied: user=student, access=READ_EXECUTE, inode="/user/hive":
e:hdfs:drwx-----
[student@ip-10-0-0-38 ~]$ HADOOP_USER_NAME=hive hadoop fs -ls /user/hive
Found 2 items
drwxr-xr-x  - hive hdfs          0 2015-11-11 05:13 /user/hive/.hiveJars
drwx-----  - hive hdfs          0 2015-11-11 05:15 /user/hive/.staging
[student@ip-10-0-0-38 ~]$ █
```

# Hadoop ist ein Ökosystem





# User interagieren bei Hadoop mit mehreren Tools



# Authentifizierung

- Basic Security bietet keine praktische Sicherheit
- Kerberos zwingend notwendig
- SSL Verbindungen zum Cluster

# Autorisierung

- Berechtigungsvergabe je Tool notwendig
- Keine übergreifende Vergabe möglich
- Transparent Data Encryption

Wer hat was gemacht?



# Auditing

- Logging und Auditing muss auf Toolebene erfolgen
- Kein übergreifendes Auditing möglich

# Integrität

- Datenverschlüsselung für Verbindungen zum Cluster
- Verschlüsselte Verbindungen innerhalb des Clusters notwendig?

# Vertraulichkeit

- Datenverschlüsselung
  - Festplattenverschlüsselung
  - Transparent Data Encryption
- Transportverschlüsselung



# C Frameworks

## Apache Knox

- Absicherung des Clusters nach außen
- Kapselung der Cluster-Interna
- LDAP und SSO Anbindung
- Schnittstelle zu Rest & HTTP Services des Clusters
- Autorisierung und Audit der Zugriffe

Authentifizierung

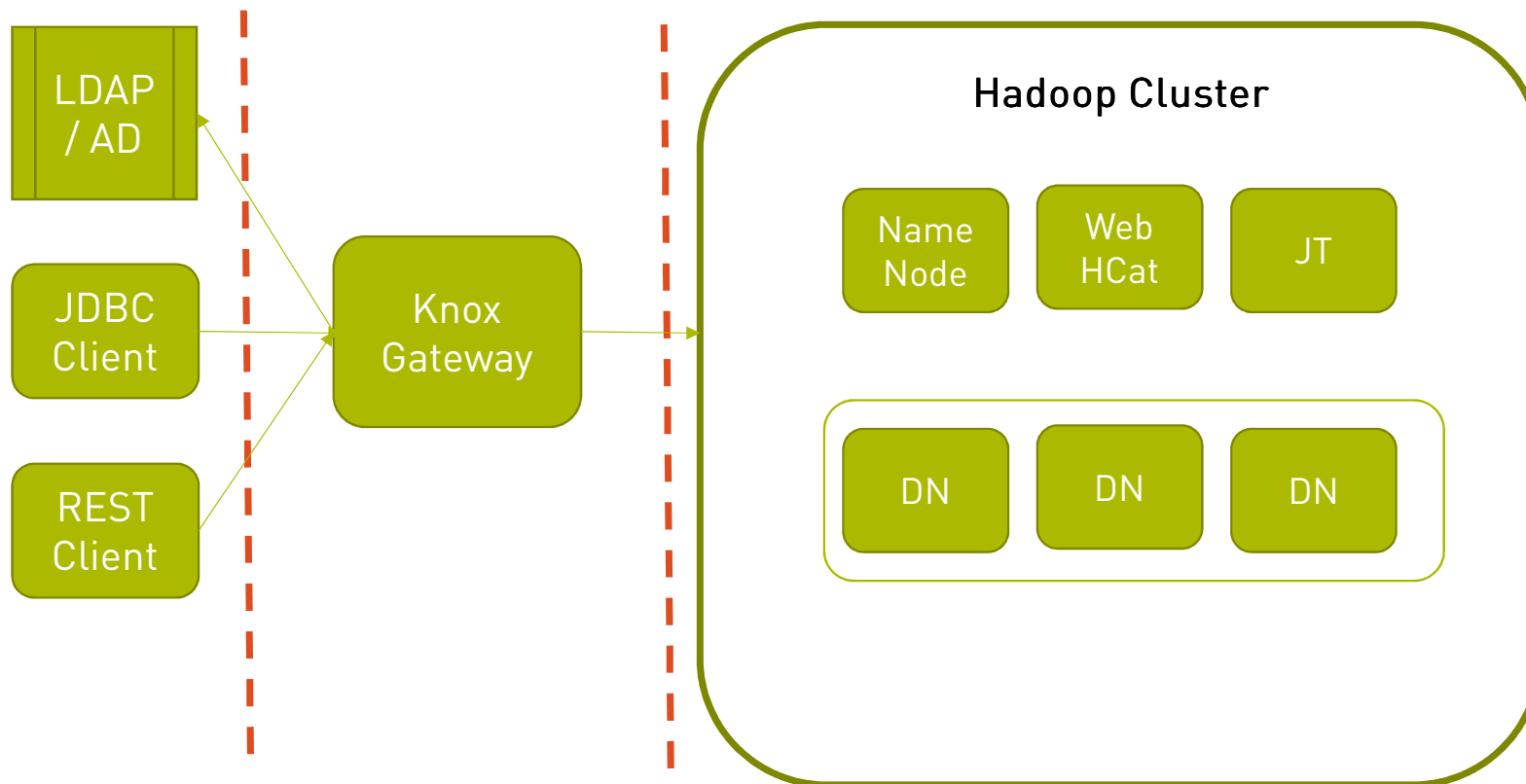
Autorisierung

Integrität

Audit



# Architektur Apache Knox



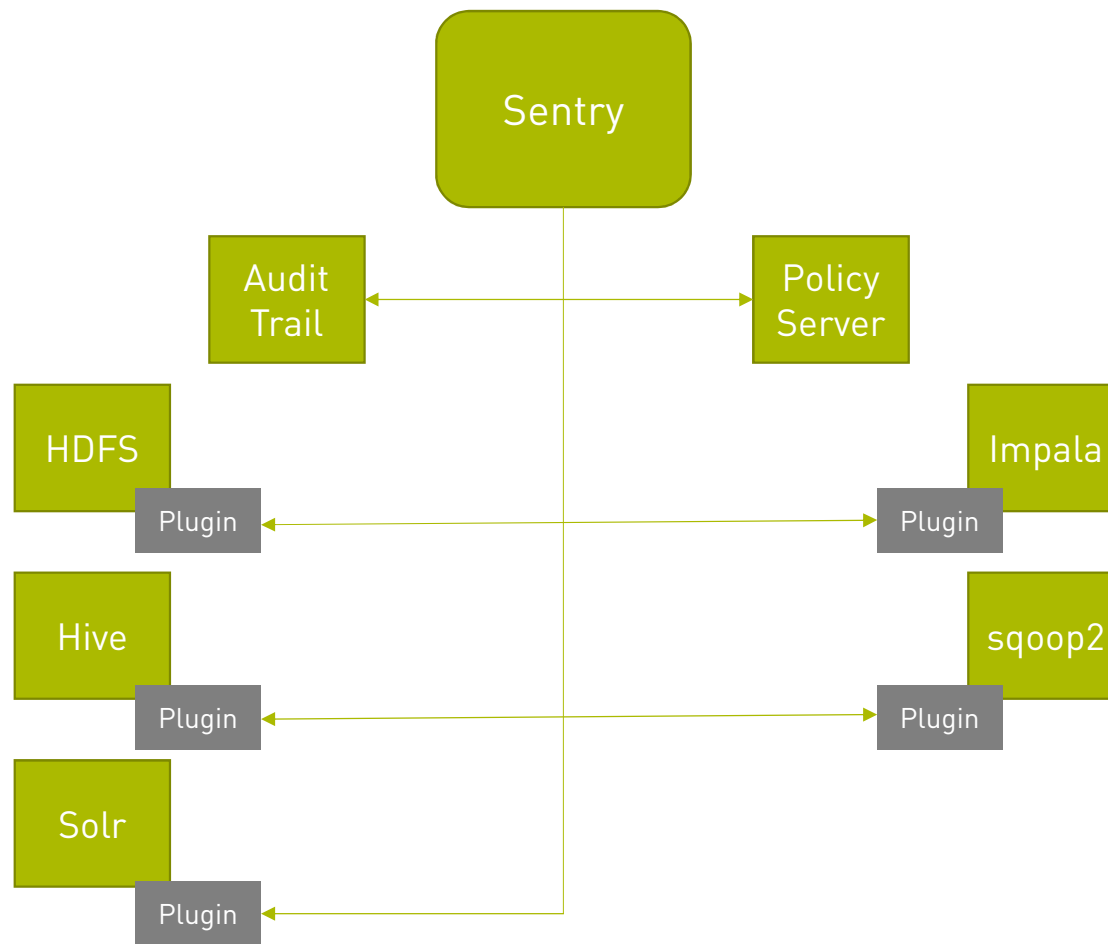
# Apache Sentry

- Mandantenfähige Vergabe granularer Rechte auf Rollenbasis
  - CRUD Rechte
  - LDAP Integration
- Unterstützt
  - HDFS, Hive
  - Solr , Impala

Autorisierung

Audit

# Architektur Apache Sentry



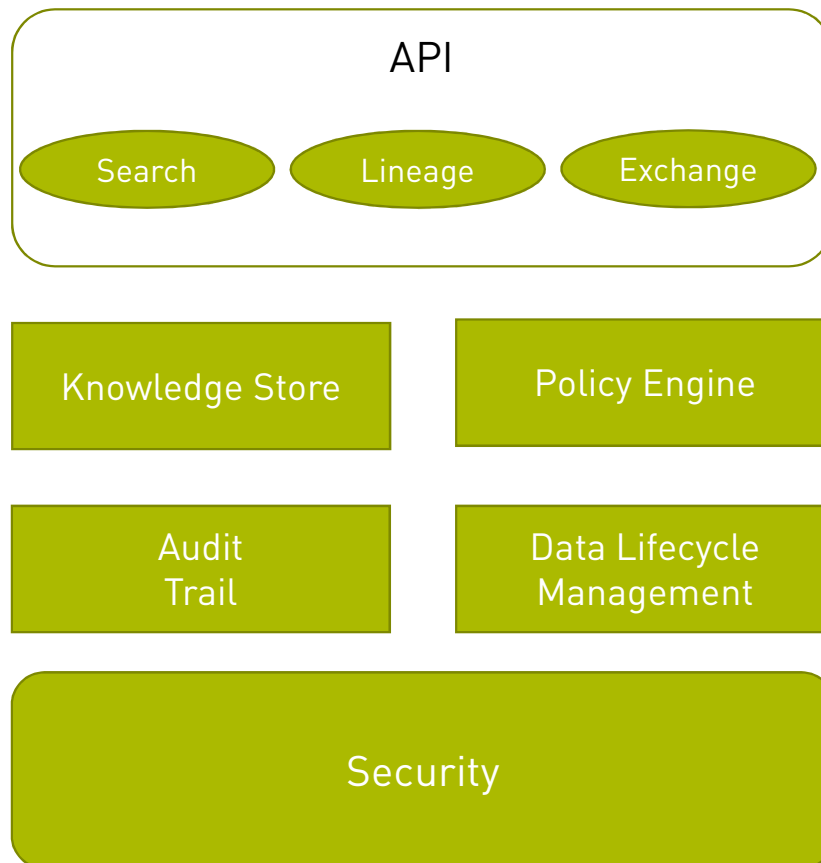
# Apache Atlas

- Data Governance Framework
- Data Lineage
- Zentrales Auditing
- Attribute Based Access Control
- Metadatenmanagement
- Datenklassifikation

Autorisierung

Audit

# Architektur Apache Atlas



# Apache Ranger

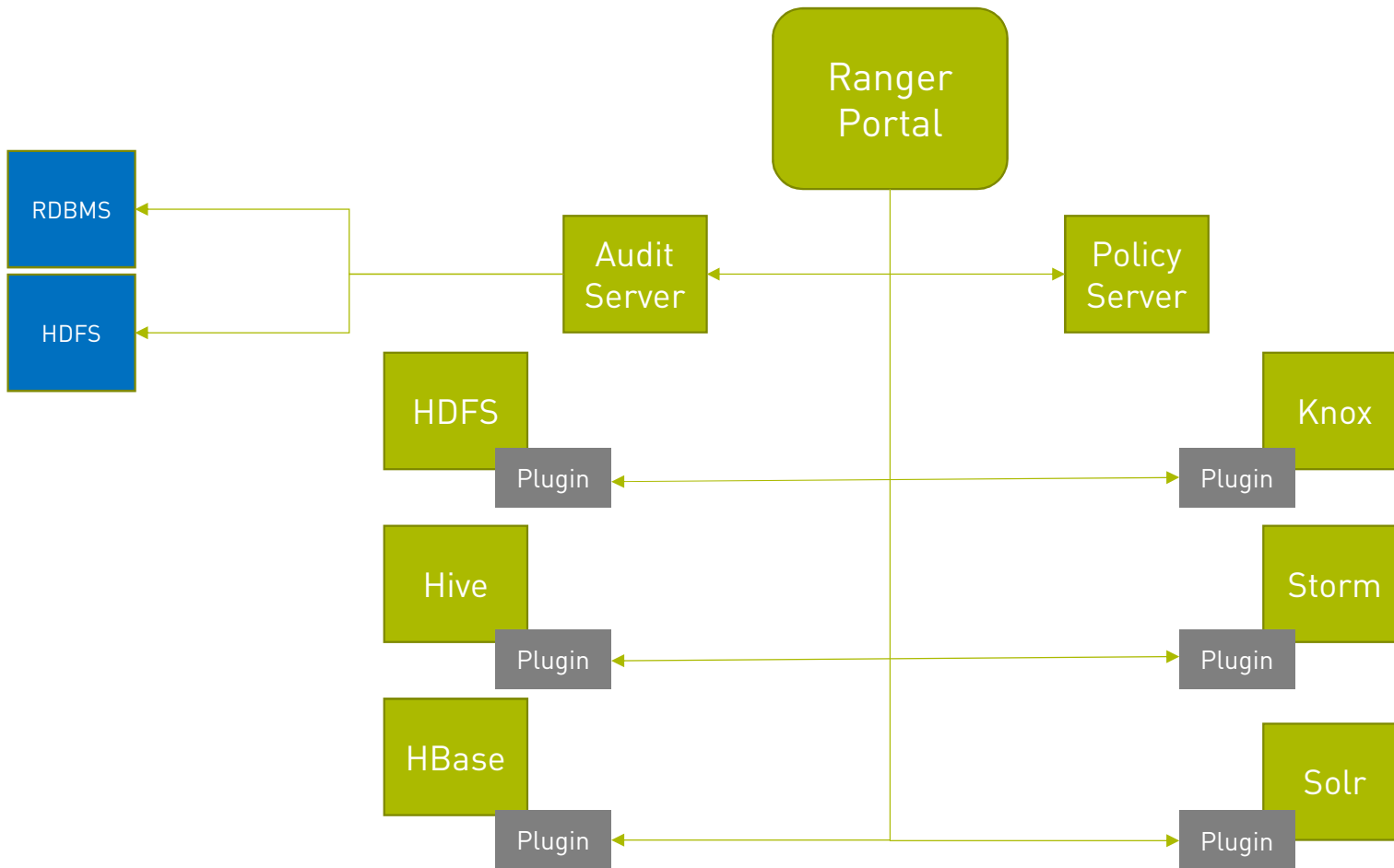
- Zentrales Policy Management und Auditing
- Plugins für zahlreiche Tools
- Zentrale Webkonsole
- Transparent Data Encryption (Key Management)

Autorisierung

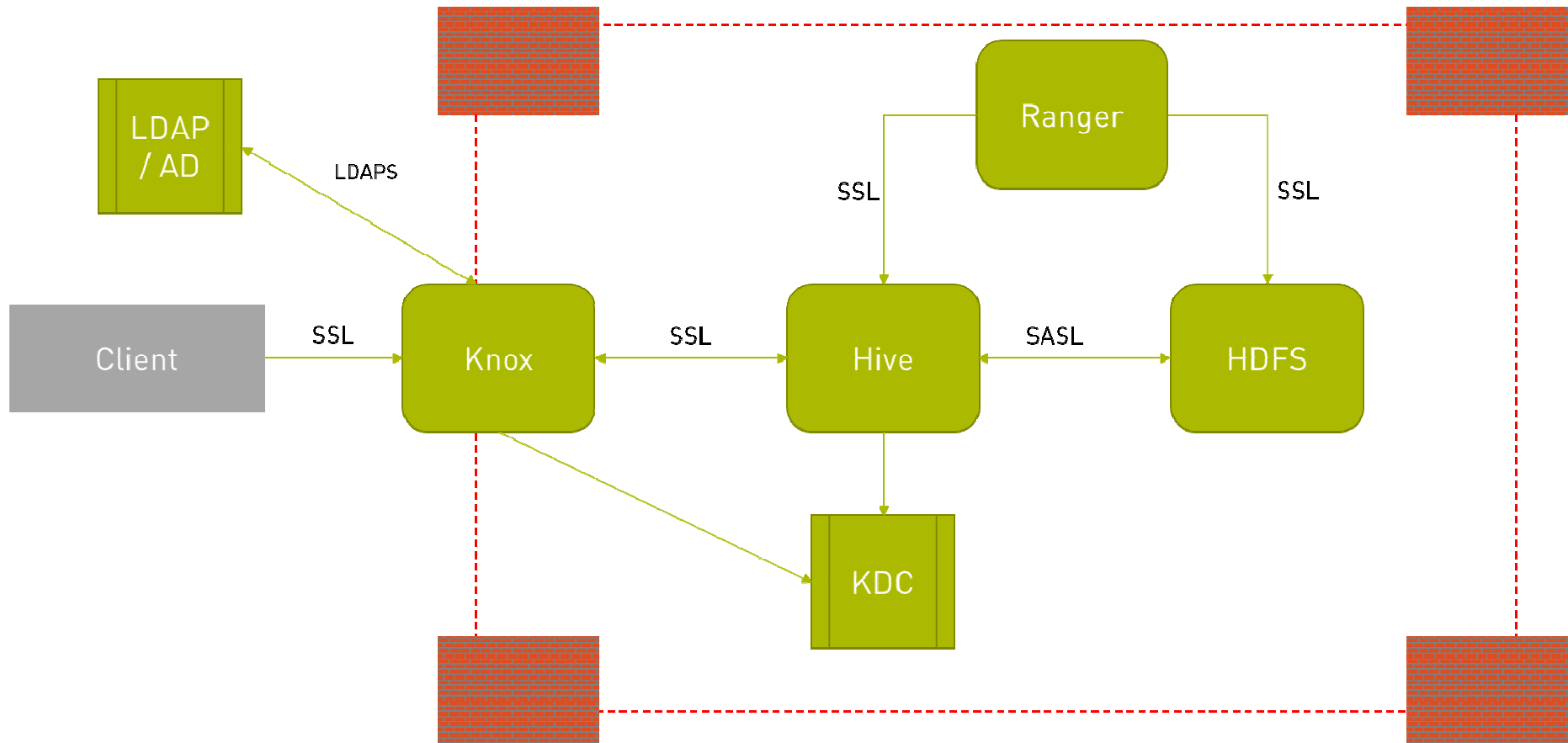
Audit

Vertraulichkeit

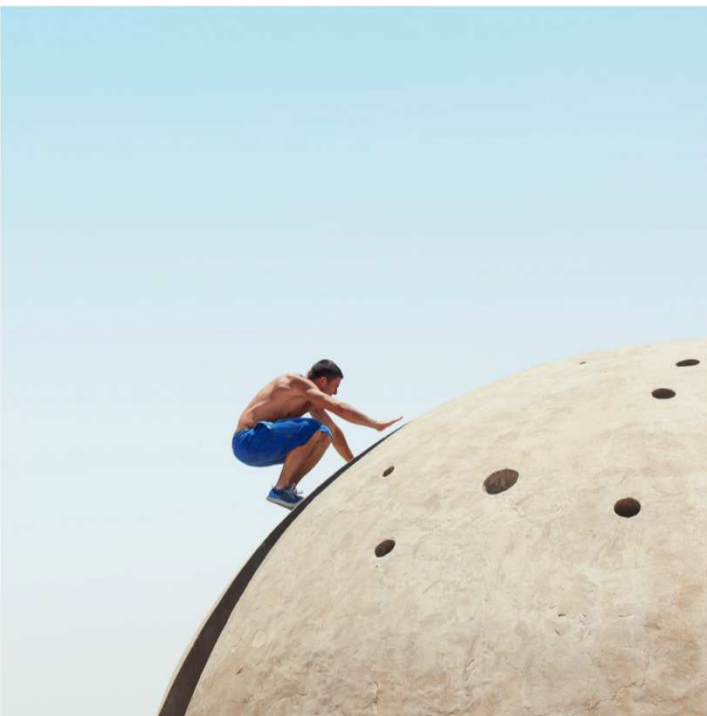
# Architektur Apache Ranger



# Gesamtarchitektur eines sicheren Hadoop-Clusters







# D | Fazit

## Fazit

- Bedeutung von Security bei Hadoop Distributionen erkannt
- Frameworks erleichtern Setup und Administration
- Stetige Weiterentwicklung
- Security muss bereits vor dem Aufbau einer Hadoop Infrastruktur geplant werden



**Herzlichen Dank**  
für Ihre Aufmerksamkeit!

**Ventum Consulting GmbH & Co. KG**

Infanteriestr. 11a  
D-80797 München

Telefon +49 89 1222 1964 2  
Fax +49 89 1222 1964 25

[www.ventum-consulting.com](http://www.ventum-consulting.com)



**Results, no excuses.**