

Enterprise User Management on a Kerberos Secured Oracle BDA

Erkan Ulgey
Garanti Technology
Istanbul

Keywords:

BigData, User Management, HUE, AD

Introduction

BigData technologies are evolving everyday. For administrators, managing this huge ecosystem and developing custom solutions for business needs are quite challenging. Oracle BigData Appliance (BDA) brings solutions in this area and make tasks easy for dba's while enjoying bigdata at the same time.

Enterprise level solutions need modeling of entire organization in order to satisfy security concerns. As an evolving and dynamic system of big data projects, existing users and groups logic should be sync with BDA. One can configure Active Directory (AD) or LDAP service for this purpose. Also, you can use Sentry tool to achieve group based authorization.

In this work, i will talk about following topics and issues

```
AD entegrated unix/hdfs enterprise user management
with Centrify and ACL authorization
AD entegrated sentry(hive/impala) authorization
AD entegrated Hue
```

I will also talk about how we established edgenode for restriction of SSH connectivity to BDA targets

With this work, you will accomplish an authorization plan for enterprise usermanagement and restrict access to your data.

Enterprise User Management and BigData

In this part, I will talk about what is Enterprise User Management and why we need it for our systems. I will talk about traditional methods like Oracle Internet Directory service, integration Active Directory and Oracle databases. And proxy users and so on..

After that, we will look at Hadoop side and compare it to our traditional systems. One can list those problems in following :

Problems of Integration to company
Where to locate
How to connect with other systems
Current Rol Modeling
Developer/user access methods to Hadoop
Commandline or GUI?
Complicated integration of tools
Not mature security mechanism on hadoop
No directory services like Oracle databases

After those problems, I will show how one can survive with those problems.

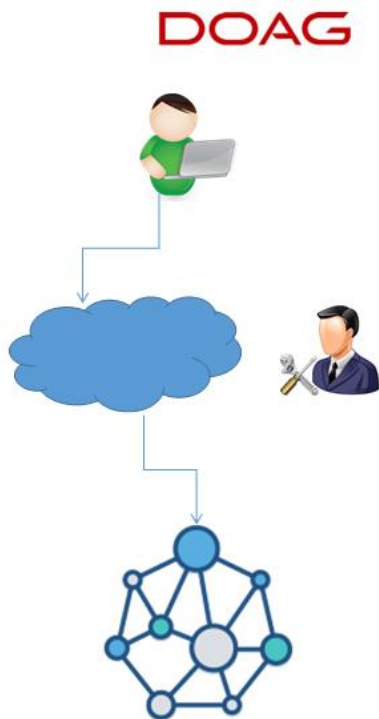


Illustration. 1: EUM and Big Data mystery

Establish Company role/grant modelling on Hadoop side

In this part, I will introduce my environment . Which tools or properties are used and enabled?
And then I will define an Active Directory tree to reflect a company rol modelling.
In next slides , we will add security to Hadoop and configure it with Sentry to maintain security.
I will give OS level commands to grant/revoke access on Hadoop and define ACL rules to limit access on Hadoop side.

Example grant command for HIVE:

```
grant select on database tmp1 to role role2;
```

Example grant command for ACL:

```
[EDGENODE]~$hadoop fs -getfacl /data/datapool/HDFSPATH/tmp1_dbpath
```

And as we all know, I will talk about bugs I faced during my implementation and I will design database and file directories on Hadoop side to reflect current datawarehouse environment . With this work, developers easily get used to work on bigdata.

HUE & LDAP Integration

In this part, I will talk about the main part of my presentation which is integrating HUE and LDAP to maintain user access to Hadoop. For this purpose, you will need to configure HUE to support LDAP access not to define new users for Hadoop.

After that we will filter users accessing Hadoop not to allow whole company to access there.

One can list action plan as following:

```
Define an Top AD Group  
Add other groups to TOP group  
Define LDAP filters  
Add Groups via HUE  
Add users via HUE  
Maintain user accessible menus on HUE
```

For this filter part, I will show some use cases with images and with bugs I faced during implementation. And I will make a comparison of those techniques at the end.

DOAG



Illustration. 2: HUE & Users

Establish an Edgenode

This part is a bonus part of my presentation and I will talk about how we kept out developers to work on directly on Hadoop . By creating an edgenode out of BDA nodes, we created a work area and configured to target Hadoop clusters.

Create a linux server
Download client configuration from Cloudera Manager
Apply and check configuration on edgenode
Establish commandline shortcuts

With this work, we aimed to maintain a development area for developers and users not their local systems and not to deal with configuration problems for every person. And here the main benefit is our bda nodes remain clean

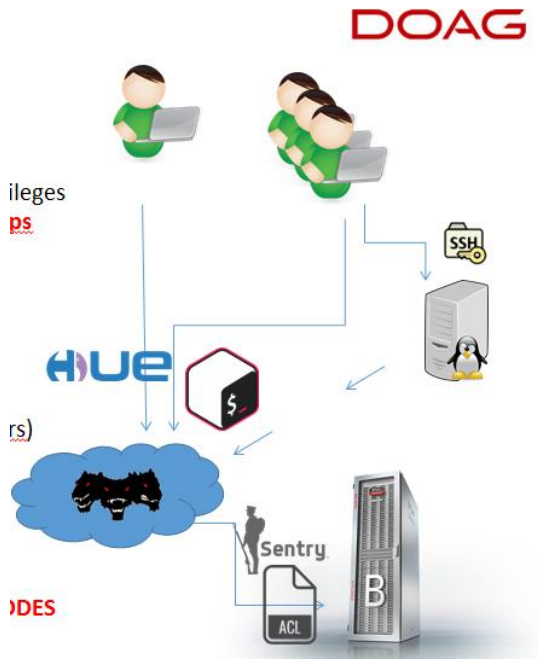


Illustration. 3: Role of an Edgenode

Contact address:

Erkan Ulgey
Garanti Technology
İstanbul

Phone: +905077280910
Fax: -
Email: ulgey.erkan@gmail.com
Internet: kamudba.blogspot.com.tr