Kerberos安全认证系统

www.huawei.com

HUAWEI

目标

- 学完本课程后，您将能够:
  - 熟悉**FusionInsight**产品中**Kerberos&Ldap**产品软件架构组成；
  - 熟悉**FusionInsight**产品中**Kerberos&Ldap**产品主要功能特性；
  - 了解**FusionInsight**产品中**Kerberos&Ldap**产品典型应用案例。

目录

Kerberos介绍

- **Kerberos**这一名词来源于希腊神话“三个头的狗——地狱之门守护者”，后来沿用作为安全认证的概念，该系统设计上采用客户端/服务器结构与**DES**、**AES**等加密技术，并且能够进行相互认证，即客户端和服务器端均可对对方进行身份认证。
- 可以用于防止窃听、防止**replay**攻击、保护数据完整性等场合，是一种应用对称密钥体制进行密钥管理的系统。

Kerberos基本概念

- 票据授权票据(**TGT Ticket-Granting Ticket**)：用于应用程序与**KDC(Key Distribution Center** 密钥分发中心)服务器建立安全会话的票据，票据授权票据存在有效期，当票据授权票据失效后，应用侧需要重新建立与**KDC**服务器的安全会话。会话有效期为**24**小时，不可配置。
- 服务票据(**ST Service Ticket**)：用于应用程序与服务端建立安全会话的票据，服务票据存在有效期，当服务票据失效后，应用侧需要重新建立与服务端的安全会话。默认有效期为**5**分钟，不可配置。

Ldap介绍

- **Ldap (Lightweight Directory Access Protocol)**，轻量目录访问协议，提供被称为目录服务的信息服务，特别是基于**x.500**（构成全球分布式的目录服务系统的协议）的目录服务。
- **Ldap**运行在**TCP/IP**或其他面向连接的传输服务之上。
- **Ldap**同时是一个**IETF**标准跟踪协议，在“轻量级目录访问协议**(Ldap)**技术规范路线图”**RFC4510**中被指定。
- **Ldap**软件来源于**OpenLDAP**项目，该项目是一个由志愿者组成的团队。

Ldap文件结构

- **LDAP**信息模型是基于条目来组织的（如图**1.1**），一个条目是一个属性的集合，有一个全球唯一的识别名（**DN, Domain Name**）。
- **DN**用于标识条目。每个条目的属性有一个类型和一个或多个值。该类型通常是可记忆的字符串，如“**cn**”就是标识通用名称，或者“电子邮件”就是电子邮件地址。该类型值的语法依赖于属性类型。 例如，一个 **cn** 属性可以包含一个值 **Marix Wong**。 一个 **mail** 属性可以包含值 “**marix@example.com**”

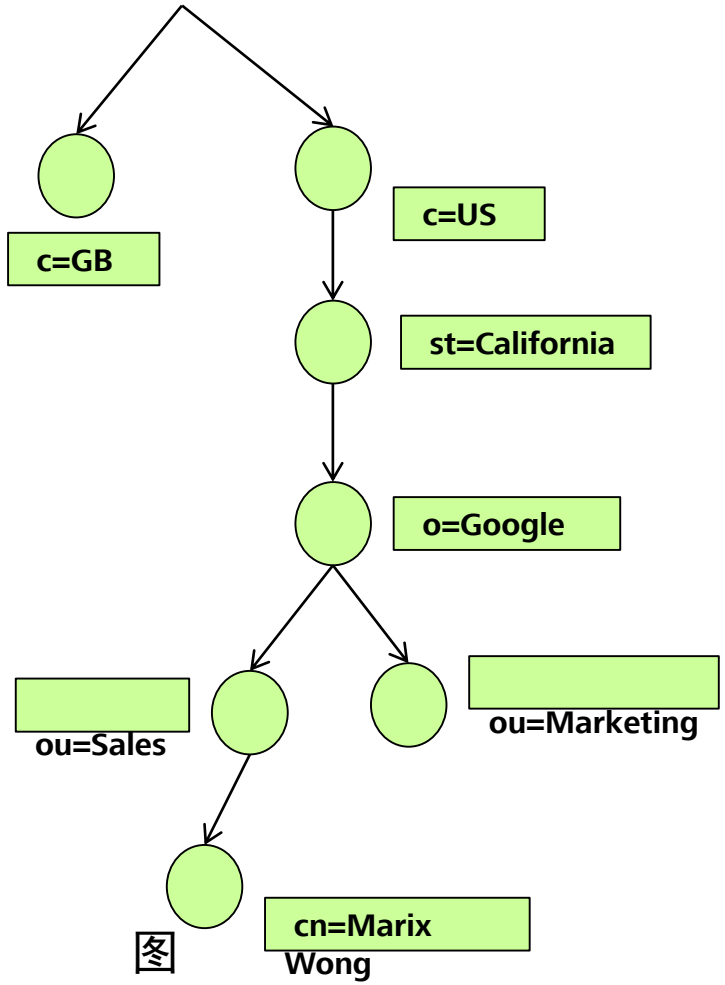


图 1.1

Ldap文件结构（续）

- 树也可以根据互联网域名组织。这种命名方式目前在业界非常普遍，因为它允许使用**DNS**为目录服务定位 。如图**1.2**所示的**Ldap**目录树中使用基于域的命名。

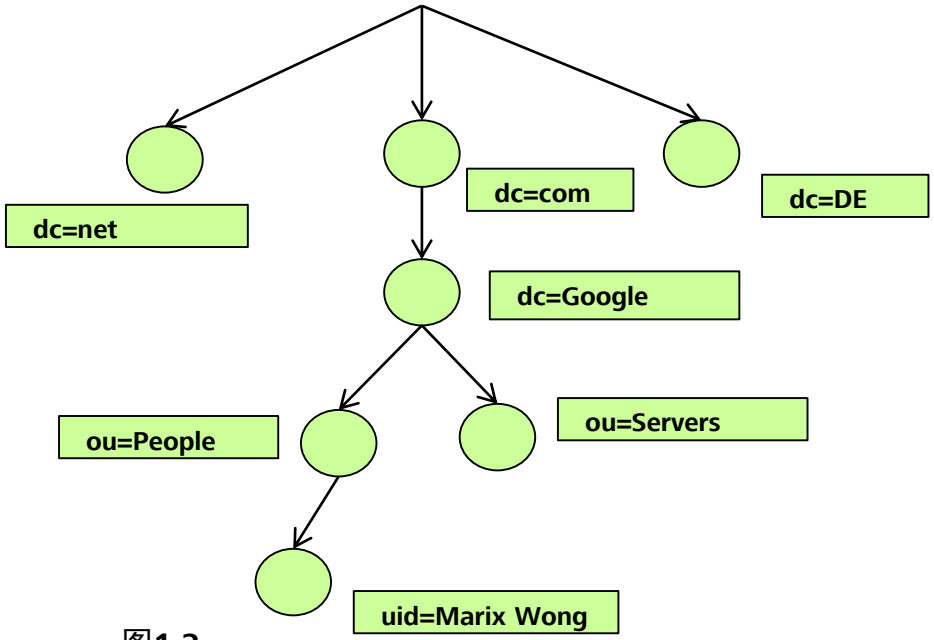


图1.2

用户属性

- 用户是作为管理**FusionInsight**的一个基本实体，又称之为帐号。用户包含如下的附属信息：用户名，密码，用户类别，用户归属组，用户归属角色，用户电话号码，用户邮箱，用户基本描述。
- 其中用户类别分为人机用户和机机用户，做出这样的区分主要是从用户使用的场景考虑的（**API**接口使用，还是界面交互使用）。其中创建人机用户类别的用户时，需要输入指定的密码，而创建机机用户类别的帐号时，不需要指定密码，由**Kerberos**通过安全算法生成随机密钥。

用户属性（续）

- 用户在**FusionInsight**系统中又分为默认用户，非默认用户，默认用户指的是**FusionInsight**系统安装，运行所需要的用户，对客户不可见，完全是系统自运行所需要而由系统自行创建，例如**HDFS**业务运行所依赖的**hdfs**用户。
- 非默认用户指的是由客户创建或者二次开发应用创建，对客户可见，在用户管理系统的界面上可以直接看到该用户及附属信息，例如通过用户管理界面创建一个**test**用户，此时可以看到该用户的附属信息。

用户属性（续）

- 从安全认证的角度去划分用户的类别，又可以将用户分为服务端用户和客户端用户。
- 服务端用户指的是集群某个服务的运行用户，例如**HDFS**服务运行依赖的**hdfs/hadoop.hadoop.com**用户。
- 客户端用户指的是进行安全认证的客户端应用程序运行使用的用户，例如某个应用程序需要访问**HDFS**服务，那么该业务流程需要先进行**Kerberos**认证，那么此时就需要使用一个客户端用户去进行**Kerberos**认证，例如该用户为**test**，即为客户端用户。

用户属性（续）

- 用户属性还有一个关键的因素，即用户密码策略，密码策略指定了该用户的密码属性信息，包含最短密码长度，密码组成的字符种类，密码校验失败的锁定时长等信息。
- 对于人机用户和机机用户的密码策略是有区别的，人机用户的密码策略包含了密码的过期时间，默认为**90**天，而机机用户由于没有显示的密码（有系统随机生成），因此其密码策略不包含过期时间，即永久有效。

用户属性（续）

- **keytab**文件是由服务端**Kerberos**生成，内容为该用户的密码，并且通过**AES-128/AES-256**加密生成的密文。
- **Keytab**由**FusionInsight**服务器生成，主要由两个场景使用：
  - 一个是集群内某个服务需要使用**keytab**进行运行时安全认证使用，例如**hdfs**启动依赖**keytab**文件，启动前进行安全认证。
  - 一个是客户进行二次开发的应用程序需要使用**keytab**进行集群业务访问前的安全认证，例如查询**hdfs**目录或者**put**数据前，需要使用该**keytab**通过**kerberos**认证。
- **keytab**的安全需由使用者确保安全，避免泄露给其他人。

Ldap用户信息介绍

- **FusionInsight**中**Ldap**用户信息主要是提供给应用服务获取对应的**gid**信息（即用户组信息）进行权限的识别，具体格式：

  uid=20002(admin) gid=10000(supergroup)

  groups=10000(supergroup),8003(System_administrator_186)

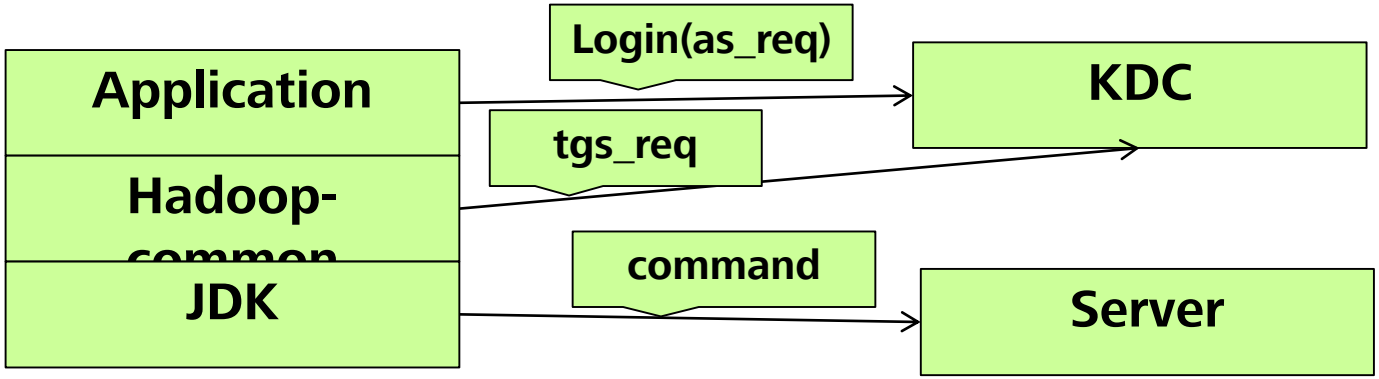
其中**uid**为用户的唯一识别码，**gid**为用户组的唯一识别码，例如上面的**10000,8003,**均为用户组的识别码，**uid**和**gid**均由**FusionInsight**的用户管理系统分配。

如果**OS**内有某个用户与**FusinInsight**内部创建的用户重名，会造成权限覆盖现象。

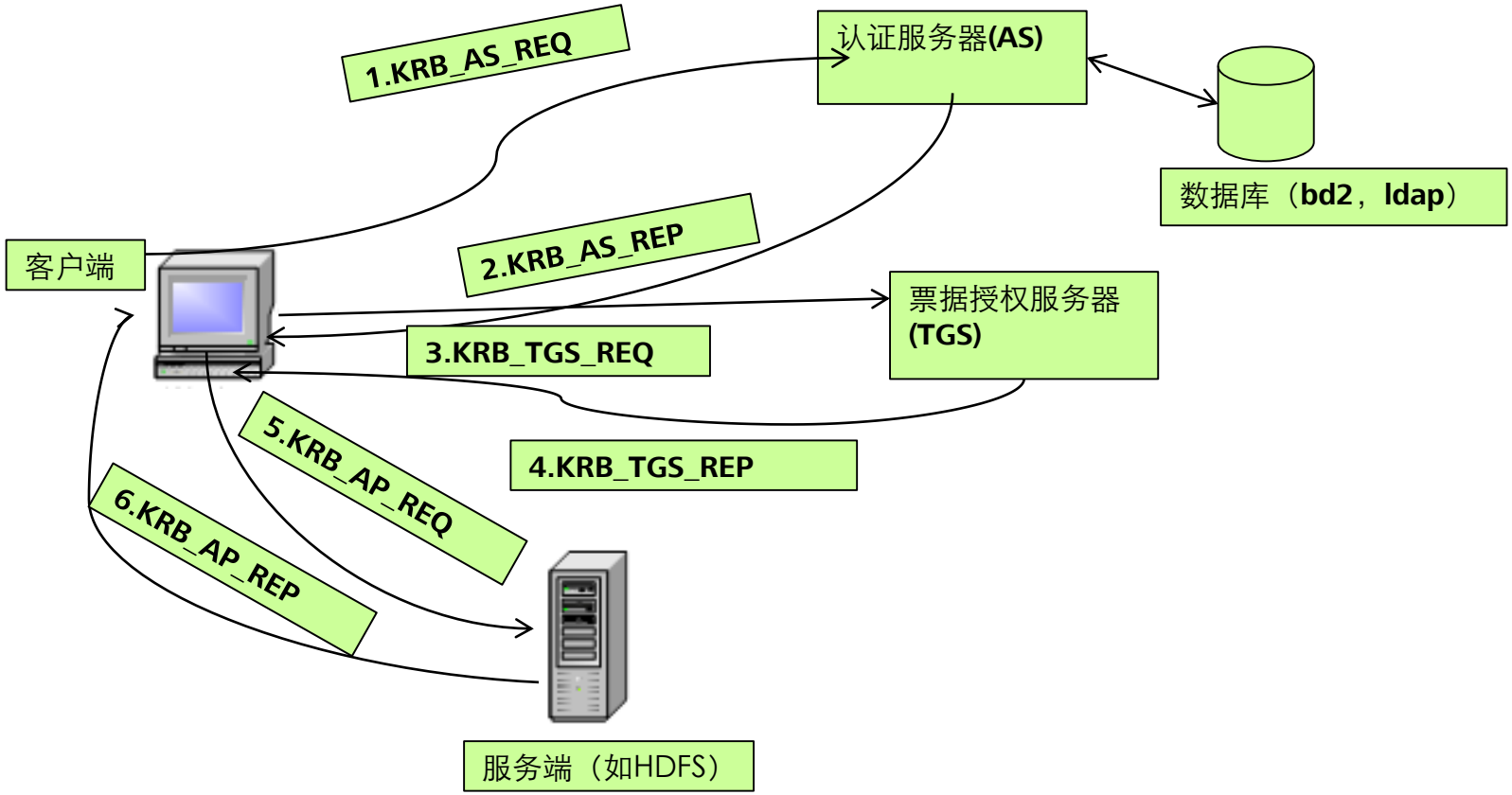
目录

Kerberos认证应用场景

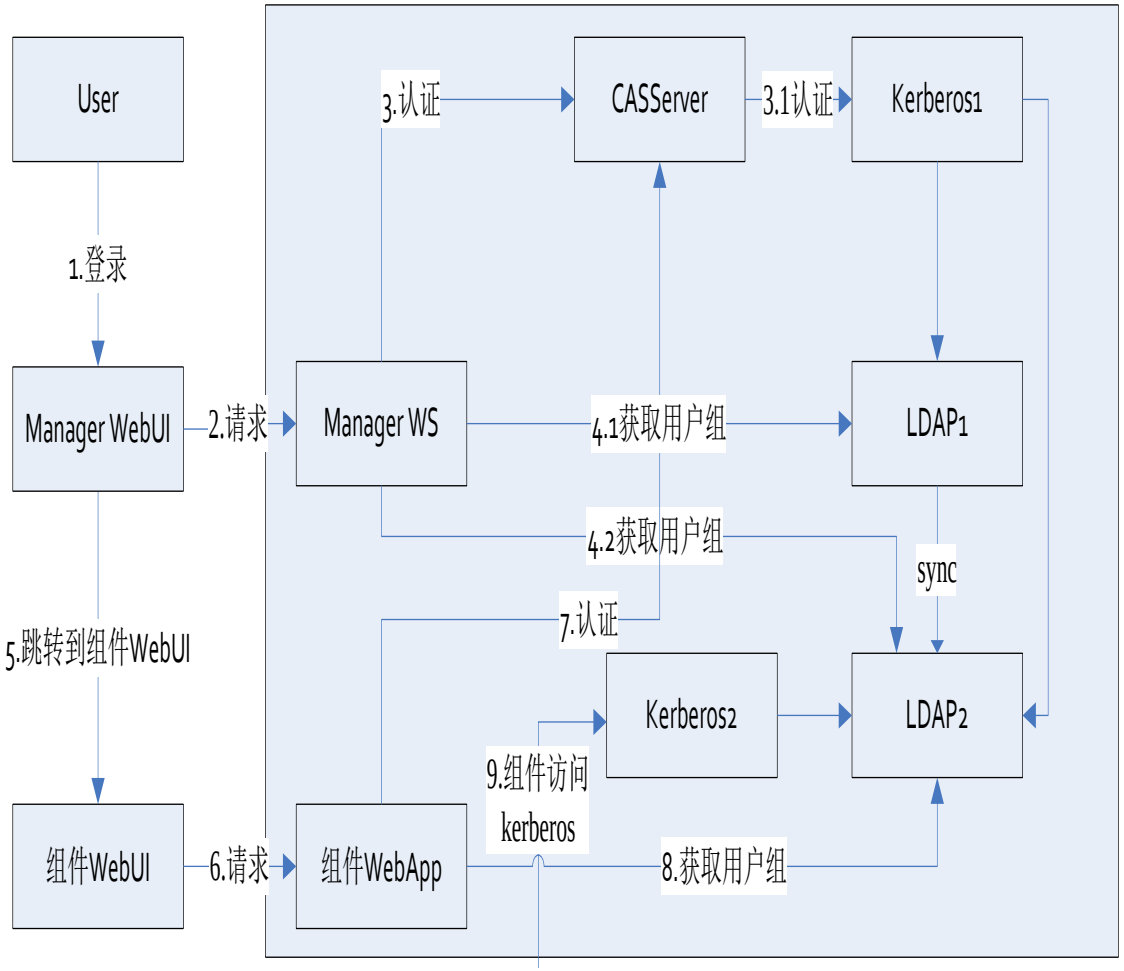
- **Kerberos**认证在**FusionInsight**中主要用于应用程序需要访问集群中某一个组件资源场景，在安全模式集群中，**FusionInsight**集群中的任意资源，如**HDFS**，**HBase**等服务，访问这些服务之前均需要通过**Kerberos**认证，建立安全会话链接，如下逻辑结构图。



Kerberos认证原理



FusionInsight Kerberos架构原理

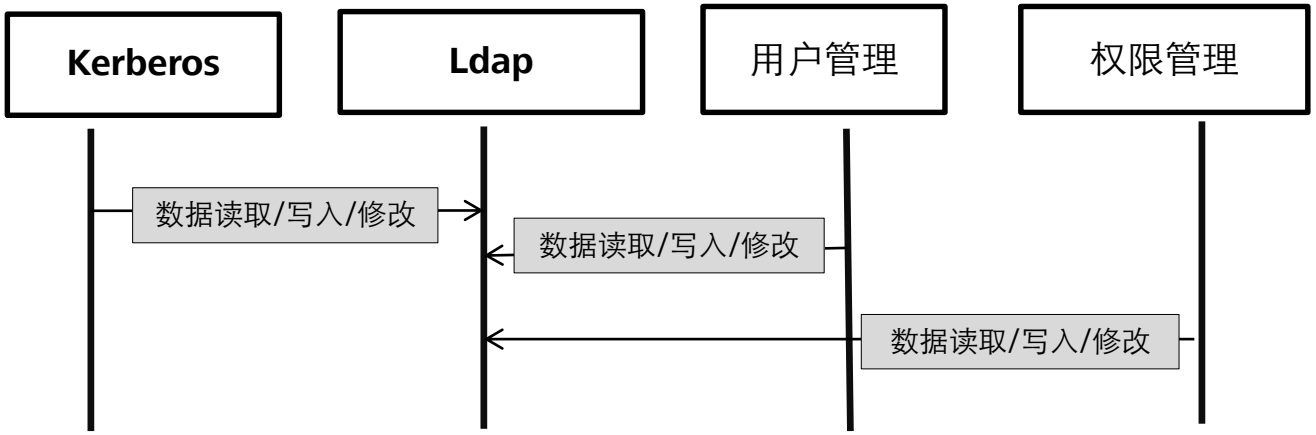


Step 1、2、3、4:
登录**Manager WebUI**的流程
Step 5、6、7、8:
登录组件UI的流程
Step9 :组件间访问安全认证

Kerberos1对**Ldap**中数据的操作方式：访问**Ldap1**（主备两个实例）和**Ldap2** （主备两个实例），是采用负荷分担访问，数据的写操作只能在**Ldap2**（主实例）上。数据的读操作可以在**Ldap1**或者**Ldap2**上。

Kerberos2对**Ldap**中数据的操作方式：只能访问**Ldap2**（包含主备两个实例），数据的写操作只能在**Ldap2**（主实例）

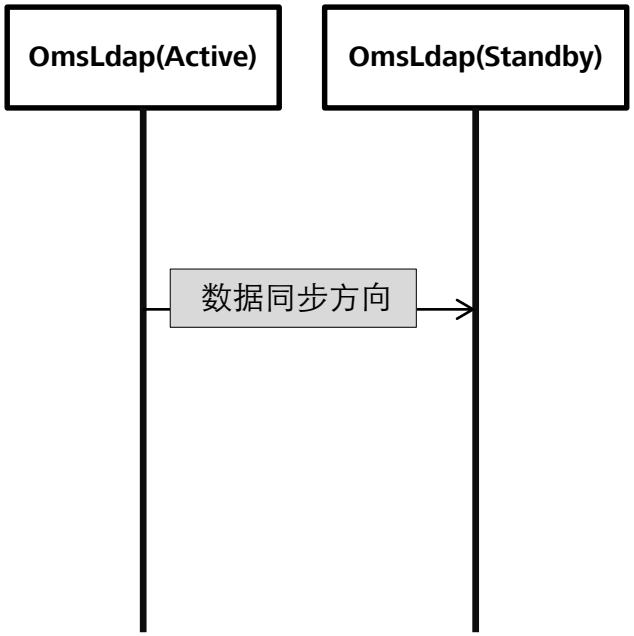
Kerberos与Ldap业务交互原理



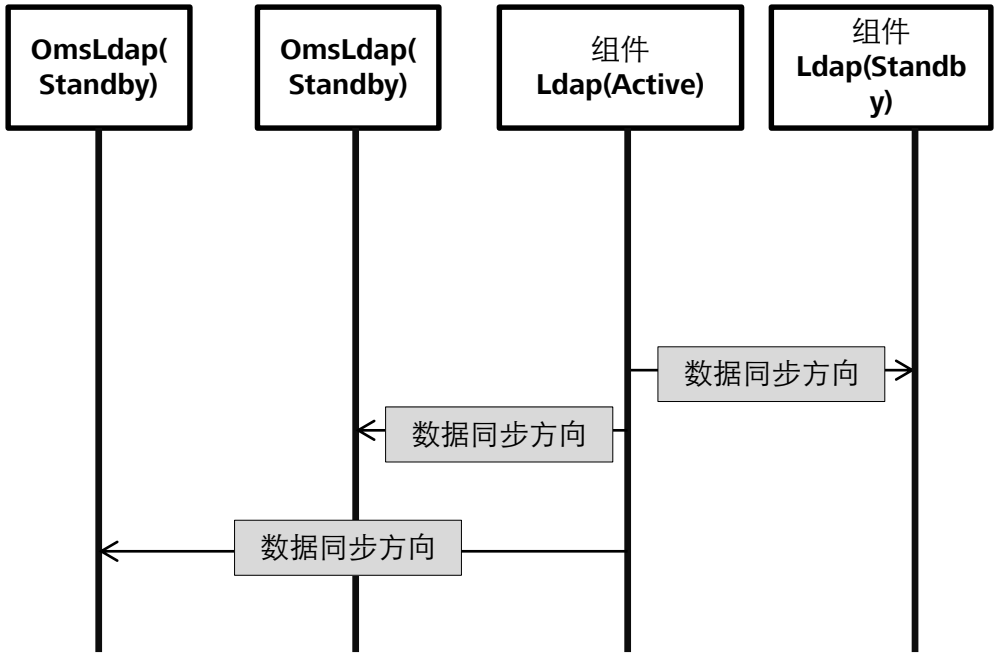
- **Kerberos**作为认证服务器中心，向集群内所有服务以及客户的二次开发应用提供统一的认证服务。
- **Ldap**作为用户数据存储中心，存储了集群内用户的信息，包含密码，附属信息等。
- 统一认证的过程中，**Kerberos**的所有数据，包含用户的密码，用户的附属信息（例如用户归属组信息）均需要从**Ldap**获取。
- 每一次的认证业务，**Kerberos**均需要从**Ldap**中获取用户信息。

Ldap数据同步原理

安装集群前**OmsLdap**数据同步



安装集群后**Ldap**数据同步



安装集群前数据同步方向：主**OmsLdap---->**备**OmsLdap**

安装集群后数据同步方向：主组件**Ldap---->**备组件**Ldap**&备**OmsLdap**&备**OmsLdap**

目录

用户存储

- **Ldap**作为**FusionInsight**的基础组件主要提供用户数据存储功能，其中包含集群内的默认用户（例如**admin**用户），客户创建的用户（例如通过**UI**界面创建一个用户）。
- 存储了用户的两个关键信息，分别是**Kerberos**信息和**Ldap**信息。
  - **Kerberos**信息主要包含用户的用户名，密码信息，对**Kerberos**提供的提供认证查询功能，即密码的校验功能。
  - **Ldap**信息主要包含用户的附属信息，例如：用户类别，用户组信息，角色信息，邮箱等。对外提供鉴权功能，即权限的识别，例如该用户是否具有访问**HDFS**某个文件目录的权限。

集群内服务认证

- **Kerberos**服务在**FusionInsight**集群中是一个基础组件模块，在安全模式下，所有的业务组件都需要依赖**Kerberos**服务（组件），业务组件（如**HDFS**）在对外提供业务的过程中，均需要先通过**Kerberos**的认证服务，如果无法通过**Kerberos**认证，将无法获取该业务组件的任何业务服务。
- **LdapServer**作为**Kerberos**的数据存储模块，与其他所有的业务组件（除**Kerberos**服务）均不直接交互。
- **FusionInsight** 安全模式集群内任意服务间的相互访问都是基于**Kerberos**安全方案架构实现，集群内某个服务（如**HDFS**）**prestart**的时候，会预先去**Kerberos**中获取该服务对应的服务名称**sessionkey**（**keytab**，主要是提供给应用程序进行身份认证使用），在后续任意其他服务（如**YARN**）需要去**HDFS**中执行增/删/改/查数据时，必须获取到对应的**TGT**和**ST**，用于本次的安全访问。

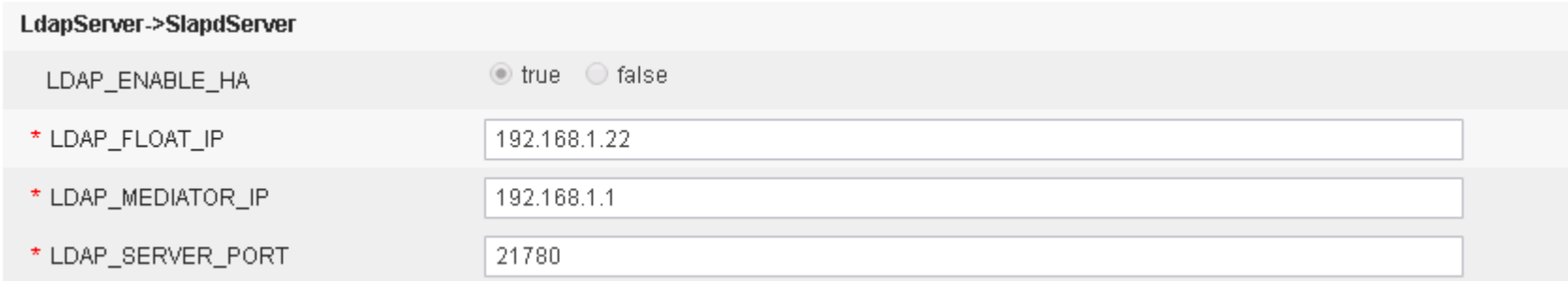
集群内服务认证（续）

二次开发认证

- **FusionInsight** 提供了二次开发接口，用于客户或者上层业务产品集成**FusionInsight**作为二次开发使用，二次开发过程中，安全模式集群提供了特定的二次开发认证接口，用于二次开发过程中的安全访问。

Ldap HA机制

- FusionInsight中集成的Ldap服务提供了HA机制，提升Ldap的高可靠性，具体配置如下：



跨集群互信特性

- FusionInsight提供两个集群之间的互信特性，用于实现集群之间的数据读、写。

跨集群互信特性（参数说明）

default_realm:本端集群的域名，唯一标识该集群的域信息。

password_suffix:密码后缀，用于配置跨集群后的人机账号初始化密码后缀，默认值为**Admin@123**。配置互信之后，系统的人机账号密码均会初始化，通过默认的密码前缀加上用户的特性密码后缀组成安全的初始化密码，默认的密码前缀为**Admin@123**。

peer_kdc_services:标识对端集群的**KDC**监听**IP**和端口，例如**192.168.1.5:21731**。

peer_realm：标识对端集群的域名。

跨集群互信特性的配置步骤，请参见产品文档的《管理员指南》的“配置跨集群互信”章节。 跨集群互信的原理，请参考本**ppt**末尾的附件。

目录

常用角色部署方式

- **Kerberos**服务角色：**KerberosServer**、**KerberosAdmin**
  - **KerberosServer**对外提供认证功能，**KerberosAdmin**对外提供用户管理（用户的增，删，改）功能。
- **Kerberos**服务采用负荷分担模式部署，安装的时候，需将**Kerberos**服务选择到集群内两个控制节点。
- **LdapServer**服务角色：**SlapdServer**
- **LdapServer**服务采用主备模式部署，安装的时候，需将**LdapServer**服务选择到集群内两个控制节点。

考虑性能最优化，建议所有集群中**LdapServer**都与**KrbServer**部署在相同节点上。

参数介绍

配置项	配置含义
KADMIN_PORT	kadmin服务提供用户管理的端口
Kdc_ports	kdc服务实例的端口
Kdc_timout	kdc提供认证服务的超时时长
KPASSWD_PORT	kadmin提供密码管理的端口
LDAP_OPTION_TIMEOUT	Kerberos对后端LDAP数据库进行连接的超时时长，连接操作时间超过该值，Kerberos返回失败。
LDAP_SEARCH_TIMEOUT	Kerberos对后端LDAP数据库进行操作的查询时长，查询操作时间超过该值，Kerberos返回失败。
max_retries	JDK进程连接KDC进行认证的最大次数，如果连接次数超过设定值，返回失败。

常用命令

命令	命令含义
ldapsearch	系统自带的**Ldap**客户端命令工具，查询Ldap中的用户信息
ldapadd	系统自带的**Ldap**客户端命令工具，向Ldap中的添加用户信息
ldapdelete	系统自带的**Ldap**客户端命令工具，删除Ldap中的用户信息
kinit	**Kerberos**用户身份认证，只有通过身份认证的用户才能执行**FusionInsight HD**各组件的**shell**命令，完成组件的维护任务
kdestroy	**Kerberos**用户身份注销，完成组件任务后使用
kadmin	切换致**kerberos admin**用户，拥有**admin**权限，该用户可以获取、修改**kerberos** 用户信息
kpasswd	修改**Kerberos**用户密码
klist	列出当前通过身份认证的**Kerberos**用户

习题

1. 以下对于**Kerberos**描述正确的是（）

（**a**）**Kerberos**组件分为**KerberosServer**和**KerberosAdmin**两个进程。

（**b**）**Kerberos**组件在**FusionInsight**上面部署的个数可以随意选择，只要不超过集群节点个数即可。

（**c**）**Kerberos**组件部署采用主备方式部署。

（**d**）**Kerberos**组件后端数据库**FusionInsight**采用的是**PostgreSQL**数据库存储认证数据。

2. 有关**Kerberos**认证流程，描述不正确的有（）

（**a**）**Kinit –kt hdfs.keytab hdfs**命令只能用于认证机机账户。

（**b**）**Kdestroy**命令默认会删除**/tmp**下面的所有**tgt**信息。

（**c**）**Kadmin**用户可以重置所有非**admin**用户的密码。

（**d**）**Kerberos**认证当前支持命令行和**api**两种方式认证。

习题

3. Kerberos不支持的功能有（）

（**a**）跨集群互信。

（**b**）掉电自动选择**KDC**。

（**c**）**Tgt**永久可信。

（**d**）多**ldap**数据连接。

思考题

1. Kerberos作为安全模式下的基础组件，哪些服务（组件）都需要与Kerberos进行交互？都分别在服务的什么流程中会涉及到?
2. 通过客户端执行kinit命令认证方式和调用二次开发的接口（例如hadoop提供的login接口）认证，这两种认证方式有何差异?

本章总结

- 本章介绍了**Kerberos**安全认证系统，通过从协议流程上讲解了认证的基本流程，从**FusionInsight**集成部署方面，讲解了如何产品化，以及产品化过程中新增的一些特性。
- 通过本章节的学习，可以提升对**FusionInsight**产品中安全认证的理解，提高自身对产品的可维护能力。

更多信息

- **MIT Doc官方网址：http://web.mit.edu/kerberos/krb5-latest/doc/index.html**
- **OpenLdap官方网址：http://www.openldap.org**

Thank you

www.huawei.com