|  |  |
| --- | --- |
| 产品名称 | 密级 |
|  | 机密 |
| 产品版本 | 共页 |
|  |

BI大数据平台系统安全性设计说明书

|  |  |  |  |
| --- | --- | --- | --- |
| 拟制 | 李勇l00164768 | 日期 | 2016/4/14 |
| 审核 | 产文00294294、陈小华 00322652 | 日期 | 2016/5/3、2016/5/16、2016/5/17 |
| 批准 | 产文00294294、陈小华 00322652 | 日期 | 2016/5/18 |



华为技术有限公司

版权所有 侵权必究

修订记录

|  |  |  |  |
| --- | --- | --- | --- |
| 日期 | 修订版本 | 修改描述 | 作者 |
| 2016/4/14 | V0.1 | 初始版本 | 李勇00164768 |
| 2016/4/24 | V0.3 | 把hive、hdfs和gaussdb合为一个FusionInsight组件 | 李勇00164768 |
| 2016/5/6 | V0.4 | 去除NFS相关内容 | 李勇00164768 |
| 2016/5/11 | V0.5 | 去掉产品人员 | 李勇00164768 |
|  |  |  |  |

目 录

[1 简介 7](#_Toc449347602)

[1.1 范围 7](#_Toc449347603)

[1.2 特性概述 7](#_Toc449347604)

[1.3 假设和约束 7](#_Toc449347605)

[2 Low Level威胁分析 7](#_Toc449347606)

[2.1 Low Level数据流图 7](#_Toc449347607)

[2.2 业务场景及信任边界说明 8](#_Toc449347608)

[2.2.1 业务推送数据 9](#_Toc449347609)

[2.2.2 交互式查询 9](#_Toc449347610)

[2.2.3 TCC任务调度 9](#_Toc449347611)

[2.2.4 计算结果回传 9](#_Toc449347612)

[2.3 外部交互方分析 9](#_Toc449347613)

[2.3.1 业务服务器 9](#_Toc449347614)

[2.3.2 TCC 10](#_Toc449347615)

[2.3.3 产品人员 11](#_Toc449347616)

[2.3.4 数据开发人员 12](#_Toc449347617)

[2.4 数据流分析 13](#_Toc449347618)

[2.4.1 业务服务器⇒/NFS/DataIn 13](#_Toc449347619)

[业务服务器⇒/NFS/DataIn 13](#_Toc449347620)

[2.4.2 /NFS/DataIn⇒FusionInsight 14](#_Toc449347621)

[2.4.3 FusionInsight⇒DataOut 16](#_Toc449347622)

[2.4.4 FusionInsight⇔执行网关 17](#_Toc449347623)

[2.4.5 DataOut⇒数据开发人员 18](#_Toc449347624)

[2.4.6 产品人员⇔执行网关 19](#_Toc449347625)

[产品人员⇔执行网关 19](#_Toc449347626)

[2.5 处理过程分析 21](#_Toc449347627)

[2.5.1 DataPush 21](#_Toc449347628)

[2.5.2 DataLoad 23](#_Toc449347629)

[2.5.3 FusionInsight 25](#_Toc449347630)

[2.5.4 执行网关 27](#_Toc449347631)

[2.6 数据存储分析 29](#_Toc449347632)

[2.6.1 /NFS/DataIn 29](#_Toc449347633)

[2.6.2 Share 30](#_Toc449347634)

[2.6.3 /NFS/DataOut 31](#_Toc449347635)

[2.7 隐私设计规范Checklist自检 33](#_Toc449347636)

[2.8 缺陷特性汇总 33](#_Toc449347637)

[3 安全及隐私保护设计 33](#_Toc449347638)

[3.1 安全需求设计 33](#_Toc449347639)

[3.1.1 秘钥管理 33](#_Toc449347640)

[3.1.2 传输安全设计 33](#_Toc449347641)

[3.1.3 日志安全设计 33](#_Toc449347642)

[3.1.4 用户权限设计 34](#_Toc449347643)

[3.2 隐私保护设计 34](#_Toc449347644)

[3.2.1 个人数据列表 34](#_Toc449347645)

[3.2.2 隐私数据保护设计 34](#_Toc449347646)

[4 安全及隐私保护设计检视 34](#_Toc449347647)

[4.1 检视对象与策略 34](#_Toc449347648)

[4.2 结果评价 34](#_Toc449347649)

[4.3 问题分析 34](#_Toc449347650)

[4.4 安全及隐私保护设计检视Checklist 35](#_Toc449347651)

[5 备注和附录 35](#_Toc449347652)

表目录

表1 子系统接口列表 53

Keywords 关键词：篡改;仿冒;泄露;安全;BI

Abstract 摘要：本文主要描述了BI大数据平台的架构、数据流图、威胁建模、威胁分析以及安全设计等方面

缩略语清单：

| 缩略语 | 英文全名 | 中文解释 |
| --- | --- | --- |
| BI | Business Intelligence | 商业智能 |
| TCC | Task Control Center | 任务调度中心 |
| FI | FusionInsight | 融合洞察 |
|  |  |  |
|  |  |  |

# 简介

## 范围

本文描述了BI大数据平台系统的安全设计和实现方案，覆盖了数据推送、数据入库、执行网关、批处理计算以及文件服务器等子模块。数据推送工具和数据入库工具使用开源的Talend开发。文件服务器使用操作系统自带的网络共享文件系统NFS。批处理组件使用公司内部已经通过安全红线测试的FusionInsight产品。

## 特性概述

外部的业务服务器周期性的把产生的日志文件以及数据库数据等推送到本系统，在本系统内部

，对数据进行清洗、转换、入库、计算以及结果输出等操作通过外部的任务调度系统进行周期性的调度执行，计算结果通过定时任务推送到报表、画像等数据开发人员处。

## 假设和约束

本文档主要是面对业务的设计、开发和测试人员，并假定上述人员已对BI数据平台的特性和使用方法基本了解。

文档的主要依据是安全威胁分析思维导图及安全红线要求。该系统访问时，可以在公司内网、公司外网等网络环境；

# Low Level威胁分析

## Low Level数据流图



数据流图元素及说明如下表所示：

|  |  |  |
| --- | --- | --- |
| **元素** | **符号** | **描述** |
| 外部交互方 |  | 能驱动系统业务，但不受系统控制的人和物(如用户，管理员，第三方系统等).通常表示目标系统的输入/输出 |
| 处理过程 |  | 一个过程执行一个任务时的逻辑表示，例如Web Server 、ftp server、LMT server |
| 数据存储 |  | 数据存储表示文件、数据库、注册表项、内存等 |
| 数据流 |  | 数据在系统中的移动方向，如网络通讯、共享内存、函数调用等。 |
| 信任边界 |  | 当数据流穿越不同的信任级别(区域)时，就存在信任边界，例如从用户态到内核态，从客户端到服务端等。 |

## 业务场景及信任边界说明

BI大数据平台信任边界主要涉及业务推送数据、TCC任务调度、数据回传以及产品人员进行交互式数据查询。

### 业务推送数据

业务服务器通过配置linux定时任务，定时的推送文件、数据库数据等到BI平台内部文件服务器的/NFS/DataIn目录下，进行临时的数据存储。在进行数据推送时，使用SSH协议进行数据传输，信任度高。



### TCC任务调度

开发人员使用TCC执行调度工作，定期的执行推送、入库以及计算任务。TCC通过调用部署在网关机上的脚本，驱动计算任务在FusionInsight的Hive等组件上执行。TCC与网关机的通信采用ssh方式，信任度较高。



网关机上的执行壳在执行过程中产生的日志实时返回到TCC所在服务器，写入到TCC所在的mysql中。存在一定安全风险。

### 计算结果回传

通过FusionInsight计算出的结果，临时存储在数据通道的DataOut目录，通过TCC配置周期任务，执行数据推送任务，把计算结果回传到数据开发人员处，进行报表、画像等数据产品的开发。数据回传通过SSH协议发送，走公司内部专用网络，信任度较高。

### 运维人员监控NFS状态

## NFS对运维人员端口，运维人员可以通过监控端口的状态来判断NFS服务器的状态，运维人员通过登录zabbix用户管理界面登录zabbix，查看服务器状态。外部交互方分析

### 业务服务器

|  |  |  |
| --- | --- | --- |
| 元素名称 | 业务服务器 | |
| 元素概述 | 各个业务的服务器，对外提供服务，对内以数据推送的方式和数据平台交互，主要为BI数据平台提供数据来源。 | |
| 威胁 | SR | |
| 仿冒（S） | 风险（Risk） | **风险等级：低**  **影响：高**。  1.系统被仿冒后，会推送错误的数据到BI平台，错误的数据部分会被清洗，即使部分数据传输到数据通道，只会造成数据无法入库。  2.被仿冒后不断往BI内部推送数据，造成BI文件服务器宕机，或者推送错误数据，导致产生错误的计算结果。  **可能性评估 低** |
| 已有消减措施 | 1.通过白名单机制，保证非法IP无法完成数据推送  2.已被授权进行数据推送的用户，生成SSH协议的公钥和私钥，公钥存储在BI平台内部的数据通道上，推送的用户在进行数据推送时使用私钥进行认证。生成密钥对的口令通过加密方式存储在配置文件中。 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 抵赖（R） | 风险（Risk） | 风险：中  影响：中。业务侧否认进行了错误的配置，推送超量数据或者重复数据或者推送失败，导致数据通道磁盘满载，无法接收其他业务数据或者导致数据入库任务无法执行。  可能性：低 |
| 已有消减措施 | 1.通过操作系统日志可以记录用户的操作记录 |
| 建议消减措施 | 根据公司审计日志规范修改工具审计日志格式 |
| 产品落地计划 | NA |

### TCC

|  |  |  |
| --- | --- | --- |
| 元素名称 | TCC | |
| 元素概述 | Task Control Center，自研的任务调度和控制系统。开发人员通过web页面配置任务的类型、执行周期等。TCC通过执行配置再网关机上的脚本执行各类任务。 | |
| 威胁 | SR | |
| 仿冒（S） | 风险（Risk） | **风险等级：低**  **影响：高。**如果TCC系统被仿冒，可能会执行需要大量资源的任务或者执行获取用户隐私的任务。  **可能性评估：低** |
| 已有消减措施 | 1.TCC部署在BI内部机房，有专用网络保护  2.TCC和执行网关通过SSH进行身份认证（认证方式为公钥认证），SSH协议版本号为V2.0 ，算法包括aes128-ctr,aes192-ctr,aes256-ctr |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 抵赖（R） | 风险（Risk） | **风险等级：中**  **影响：中。**开发人员否认自己的操作，有日志记录，抵赖影响较小。  **可能性评估：小** |
| 已有消减措施 | 2. TCC日志格式为：2016-05-16 15:31:48.083 INFO - - info() Thread-76730 [listerner[com.huawei.platform.tcc.service.impl.TccServiceImpl@3d62ba]](mailto:listerner[com.huawei.platform.tcc.service.impl.TccServiceImpl@3d62ba]) needs 2 ms to process Event[UPDATE\_JOB\_IDS,[[Ljava.lang.Object;@747174b2](mailto:[Ljava.lang.Object;@747174b2),[com.huawei.platform.tcc.SSH.ResultProcess@4d74f5d5]!](mailto:com.huawei.platform.tcc.SSH.ResultProcess@4d74f5d5]!) 不满足公司审计日志要求 |
| 建议消减措施 | TCC审计日志应该由执行网关机上的壳程序输出，从而可以在BI内部追踪用户的操作。修改TCC审计日志为满足公司审计要求的格式。 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |

### 文件服务器

|  |  |  |
| --- | --- | --- |
| 元素名称 | 文件服务器 | |
| 元素概述 | 数据开发人员通过使用Hive计算的结果，进行报表、群体用户画像等数据产的开发 | |
| 威胁 | SR | |
| 仿冒（S） | 风险（Risk） | **风险等级：中**  **影响：高**  数据开发人员被仿冒后，可能导致大量用户隐私数据被泄露，影响较大  **可能性评估：可能性较小**  1.VPN控制特定的开发人员能否访问内网机房  2.操作系统用户权限控制，有SSH协议做认证  3.DataPush工具提供日志记录 |
| 已有消减措施 | 1.安全  1.1VPN控制特定的开发人员能否访问内网机房  1.2.操作系统用户权限控制，有SSH协议做认证(补充协议等信息)  1.3.DataPush工具提供日志记录  2.隐私 |
| 建议消减措施 | 实验室继承办公室的意见：  1) BI具有回传数据到中国的能力  2) 什么数据通过BI回传需要走备案通过，通过后才能回传  3) 在回传前数据还留着欧洲  4） 在后续版本中完善回传方案 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |
| 抵赖（R） | 风险（Risk） | **风险等级：低**  **影响：中。**  **可能性评估：低** |
| 已有消减措施 | 1. datapush日志记录文件推送行为   2. 操作系统日志记录OS用户行为 |
| 建议消减措施 |  |
| 产品落地计划 |  |

### 运维人员

|  |  |  |
| --- | --- | --- |
| 元素名称 | 运维人员 | |
| 元素概述 | 运维人员通过NFSzabbix监控NFS端口服务的状态 | |
| 威胁 | SR | |
| 仿冒（S） | 风险（Risk） | **风险等级：低**  **影响：低**  运维人员被仿冒后，只能观察到NFS服务状态，无法获取到存储在NFS上的数据  **可能性评估：可能性较小**  VPN控制特定的开发人员能否访问内网机房 |
| 已有消减措施 | 1.zabbix有密码强度校验  2.运维人员只能监控状态，不能修改 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 抵赖（R） | 风险（Risk） | **风险等级：低**  **影响：低。**  **可能性评估：低** |
| 已有消减措施 | Zabbix有审计日志 |
| 建议消减措施 |  |
| 产品落地计划 |  |

## 数据流分析

### 业务服务器⇒NFS

|  |  |  |
| --- | --- | --- |
| 元素名称 | 业务服务器⇒NFS | |
| 元素概述 | 业务服务器通过配置定时的数据推送任务，对数据进行初步的清理和转换，推送数据到数据通道的/NFS/DataIn目录下，等待数据入BI数据仓库 | |
| 威胁 | TID | |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：中。**数据被篡改会导致数据无法入库或者计算任务执行异常，或者得到错误的计算结果。  **可能性评估：**  接入华为机房对推送数据流出口进行劫持篡改数据流可能性较小 |
| 已有消减措施 | 推送时使用SSH进行认证和加密传输， SSH协议版本号为V2.0 ，算法包括aes128-ctr,aes192-ctr,aes256-ctr |
| 建议消减措施 | NA |
| 产品落地计划 | NA |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：大**  此数据流的数据如果泄露，可能会泄露用户隐私数据，影响较大  **可能性评估：** |
| 已有消减措施 | ssh协议传输，保证数据传输安全，SSH协议版本号为V2.0 ，算法包括aes128-ctr,aes192-ctr,aes256-ctr |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：中**  **影响：**中。数据推送工具异常会导致数据无法推送到数据通道，周期任务无法执行  **可能性评估：中**  1.推送工具所在业务服务器异常，到时数据无法推送  2.业务服务器和数据通道之间的网络异常  3.数据通道文件服务器异常 |
| 已有消减措施 | 1.数据推送异常有告警机制，运维人员可快速得到通知进行处理  2.数据重新推送机制，数据推送失败会再次推送 |
| 建议消减措施 | NFS需要负载均衡机制和高可用机制 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |

### NFS ⇒FusionInsight

|  |  |  |
| --- | --- | --- |
| 元素名称 | NFS ⇒FusionInsight | |
| 元素概述 | 临时数据通道的数据通过入库工具，被加载到FusionInsight的Hive表在HDFS上的目录下 | |
| 威胁 | TID | |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：低**  数据被篡改会导致计算任务执行失败或者结果不正确。  **可能性评估：低**  1.数据从数据通道加载到HDFS上，使用https等安全的通信协议  2.加载数据的用户身份通过kerbros进行认证 |
| 已有消减措施 | 1.需要VPN权限才能访问部门内部机房  2.机房服务器需要主管审批才能访问  3.FusionInsight内部的数据加密传输机制保证 |
| 建议消减措施 | NA |
| 产品落地计划 | NA |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：高**。信息泄露会涉及到用户的隐私数据，影响较大  **可能性评估：低** |
| 已有消减措施 | 1.机房服务器需要主管审批才能访问  2.FusionInsight内部的数据加密传输机制保证 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：低**  **影响：低**。数据无法加载到数据仓库，导致周期任务无法执行。  **可能性评估：中**  1.DataLoad工具所在服务器异常  2.FusionInsight HDFS或者hive组件无法访问  3.Dataload工具到FusionInsight网络异常 |
| 已有消减措施 | 1.数据入库异常有告警机制，运维人员可快速得到通知进行处理 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |

### FusionInsight⇒NFS

|  |  |  |
| --- | --- | --- |
| 元素名称 | FusionInsight⇒NFS | |
| 元素概述 | FusionInsight中的数据仓库计算结果导出到临时数据通道的DataOut目录，等待推送回国内的数据开发人员处 | |
| 威胁 | TID | |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：中**。国内数据开发人员看不到本周期的计算结果，无法进行报表开发或者画像开发；或者看到错误的计算结果，影响产品决策。  **可能性评估：**可能性较小 |
| 已有消减措施 | 1.机房服务器需要主管审批才能访问  2.FusionInsight内部的数据加密传输机制保证 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：高**  信息泄露会涉及到用户的隐私数据，影响较大  **可能性评估：**低 |
| 已有消减措施 | 1.机房服务器需要主管审批才能访问  2.FusionInsight内部的数据加密传输机制保证 |
| 建议消减措施 | NA |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：中**  **影响：中。**国内数据开发人员看不到本周期的计算结果，无法进行报表开发或者画像开发  **可能性评估：高**  1.FusionInsight组件服务异常  2.数据通道服务器异常  3.数据通道和FusionInsight之间网络异常 |
| 已有消减措施 | 1.FusionInsight提供告警机制，可及时通知运维和开发人员进行处理  2.机房网络、主机以及端口提供监控告警机制 |
| 建议消减措施 | NFS需要负载均衡机制和高可用机制 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |

### FusionInsight⇔执行网关

|  |  |  |
| --- | --- | --- |
| 元素名称 | FusionInsight⇔执行网关 | |
| 元素概述 | 执行网关机提交Hive脚本到FusionInsight中的Hive组件执行计算任务，Hive执行结果或者日志返回网关机 | |
| 威胁 | TID | |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：中**  脚本被篡改，会导致危险的计算任务被执行，影响较大  **可能性评估：低** |
| 已有消减措施 | 1.网关机所在机房的接入需要VPN权限  2.FusionInsight内部机制保证数据安全传输 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：高**  1.泄露hive表结构信息  2．执行结果返回执行网关的数据可能泄露用户隐私数据  **可能性评估：中**  通过网络嗅探抓包分析 |
| 已有消减措施 | 1．脚本通过FusionInsight内部安全加密机制提交到hive  2. 网关机访问权限审批机制 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：低**  **影响：中**。无法提交任务到Hive执行计算任务，导致报表等数据产品不能正常使用和开发。  **可能性评估：高**  1.FusionInsight组件异常  2.执行网关所在的文件服务器异常  3.执行网关和FusionInsight之间的网络异常 |
| 已有消减措施 | 1.运维监控服务器状态  2.FusionInsight提供Hive服务状态告警 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |

### NFS⇒文件服务器

|  |  |  |
| --- | --- | --- |
| 元素名称 | DataOut⇒数据开发人员 | |
| 元素概述 | 数据通道上的DataOut目录的数据通过定时的数据推送任务推送回国内的数据开发人员处，进行数据产品的开发和测试。 | |
| 威胁 | TID | |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：低**。计算结果被篡改，会导致数据开发人员获取到错误的计算结果，影响数据产品的开发  **可能性评估：低**  1.数据通道服务器的访问权限需要审批  2.ssh协议保证数据安全传输（协议版本号和算法） |
| 已有消减措施 | 1.数据通道生成SSH协议密钥对，公钥存储在数据开发人员所在的服务器，通过建立安全的ssh通道进行数据传输。SSH协议版本号为V2.0 ，算法包括aes128-ctr,aes192-ctr,aes256-ctr |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：大。**如果计算结果中包含了用户隐私数据，则影响较大  **可能性评估：** |
| 已有消减措施 | 1.数据通道生成SSH协议密钥对，公钥存储在数据开发人员所在的服务器，通过建立安全的ssh通道进行数据传输。SSH协议版本号为V2.0 ，算法包括aes128-ctr,aes192-ctr,aes256-ctr |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：低**  **影响：中**  **可能性评估：高**  服务器或者服务器之间的网络异常，可能导致数据无法推送到数据开发人员处 |
| 已有消减措施 | 1.运维监控服务器状态和机房网络状态，及时通知开发人员进行处理 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |

### NFS⇒运维人员

|  |  |  |
| --- | --- | --- |
| 元素名称 | NFS⇒运维人员 | |
| 元素概述 | NFS运维人员通过监控NFS服务器的端口来判断NFS服务是否正常 | |
| 威胁 | TID | |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：低**  **可能性评估：低** |
| 已有消减措施 | 1.具有VPN权限才能访问部门内部机房  2.zabbix安全机制 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：高。**涉及用户隐私的数据可能被泄露  **可能性评估：**  数据可能通过NFS的端口泄露 |
| 已有消减措施 | 1.具有VPN权限才能访问部门内部机房  2.NFS端口不对外开放，由运维人员管理 |
| 建议消减措施 | NFS需要确认zabbix安全传输协议（确认使用tls1.2版本安全传输协议） |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：低**  **影响：中**  **可能性评估：中**  1.NFS组件异常  2.NFS服务器网络异常 |
| 已有消减措施 | 1.运维监控服务器和网络状态 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |

## 处理过程分析

### DataPush

|  |  |  |
| --- | --- | --- |
| 元素名称 | DataPush | |
| 元素概述 | 把FusionInsight的计算结果从数据通道的DataOut目录推送到数据开发人员 | |
| 威胁 | STRIDE | |
| 仿冒（S） | 风险（Risk） | **风险等级：低**  **影响：中**  1.仿冒DataPush，可能导致业务服务器推送不符合要求的数据  **可能性评估：低** |
| 已有消减措施 | 1.VPN权限审批机制保证特定用户才能访问BI内部机房  2.通过IP地址白名单机制保证不再列表内的IP无法推送数据  3.ssh双向认证机制保证ssh服务端和ssh客户端双向认证 |
| 建议消减措施 |  |
| 产品落地计划 |  |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：高可能性评估：**  DataPush的配置文件被篡改，可能性较大 |
| 已有消减措施 | 1.申请VPN权限才能访问BI内部机房  2.业务服务器访问权限控制  3.ssh密钥进行访问认证和加密数据传输  4.代码通过统一的平台进行编译、打包和发布（VMP） |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 抵赖（R） | 风险（Risk） | **风险等级：低**  **影响：高。**  **可能性评估：**中 |
| 已有消减措施 | 1.datapush业务日志记录推送行为  2.操作系统日志记录用户执行行为 |
| 建议消减措施 |  |
| 产品落地计划 |  |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：高**  如果计算结果中包含了用户个人数据，那么用户数据会泄露，影响较大  **可能性评估：**  获取服务器权限，修改DataPush配置文件，可能性较小 |
| 已有消减措施 | 1.VPN权限访问机房  2.业务服务器访问权限控制 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：中**  **影响：中**  计算结果数据无法推送到国内数据开发人员处，影响数据产品的开发和测试  **可能性评估：**  DataPush所在服务器或者网络故障。 |
| 已有消减措施 | 1.运维对机房内部的服务器和网络进行监控，发现异常及时通知开发人员进行处理  2. DataPush具有高可用设计，NFS具有单机宕机风险 |
| 建议消减措施 | NFS具有单机宕机风险，需要实现NFS高可用设计 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |
| 特权提升（E） | 风险（Risk） | **风险等级：低**  **影响：高**  1.运行DataPush的进程获取root权限后，可能导致数据被推送到不该推送服务器，导致用户隐私数据泄露  2.不涉及数据库权限提升  **可能性评估：低** |
| 已有消减措施 | 1.datapush使用普通用户运行(hadoop-NJ)  2获取root权限需要安全部门主管审批 |
| 建议消减措施 | 1. 无 |
| 产品落地计划 |  |

### 执行网关

|  |  |  |
| --- | --- | --- |
| 元素名称 | 执行网关 | |
| 元素概述 | TCC通过执行存储在执行网关上的脚本，向数据通道、FusionInsight提交数据推送或者计算任务。 | |
| 威胁 | STRIDE | |
| 仿冒（S） | 风险（Risk） | **风险等级：低**  **影响：高**  **被仿冒后，会执行有风险的任务，导致用户隐私数据泄露**  **可能性评估：**  可能性较小 |
| 已有消减措施 | 1.VPN权限访问国内机房  2.（TCC认证执行网关，需要确认），是否SSH协议？需要加上ssh客户端认证ssh服务端功能（需要配置）（记录到缺陷列表）（把服务端的公钥保存在客户端） |
| 建议消减措施 |  |
| 产品落地计划 |  |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：中**  篡改网关机上的脚本，可能会提交有风险的计算任务，把用户隐私数据输出到有风险的目录，导致用户隐私数据泄露  可能性：低 |
| 已有消减措施 | 1.VPN权限访问国内机房  2.ssh协议保证用户身份安全和数据传输安全，SSH协议版本号为V2.0 ，算法包括aes128-ctr,aes192-ctr,aes256-ctr |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 抵赖（R） | 风险（Risk） | **风险等级：中**  **影响：高**  **可能性评估：中** |
| 已有消减措施 |  |
| 建议消减措施 | TCC审计日志应该由执行网关机上的壳程序输出，从而可以在BI内部追踪用户的操作 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：高**  存储在网关机上的执行脚本泄露，会导致hive表信息泄露  **可能性评估：**  可能性较小 |
| 已有消减措施 | 1.VPN权限访问国内机房  2.OS用户权限审批 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：低**  **影响：中**  各项任务无法提交，影响报表、画像等产品依赖的数据  **可能性评估：**  1.执行网关所在服务器异常  2.执行网关和FusionInsight之间的网络异常  3.多机部署，有可靠性保障 |
| 已有消减措施 | 运维监控BI内部机房服务器和网络的运行状态，如果有异常，及时进行处理 |
| 建议消减措施 |  |
| 产品落地计划 |  |
| 特权提升（E） | 风险（Risk） | **风险等级：低**  **影响：中**  **可能性评估：**  1.执行任务的客户端端进程获取root权限后，可能执行不属于该用户执行的脚本任务，导致计算周期混乱，计算结果出错  2.执行任务的客户端进程操作HIVE的权限提升后，可能导致客户端访问该用户本没有权限的表和数据 |
| 已有消减措施 | FusionInsight实现了基于服务器的kerbros认证，即使获得root权限，因为无法通过服务端认证，导致无法连接FusionInsight |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |

### NFS

|  |  |  |
| --- | --- | --- |
| 元素名称 | NFS | |
| 元素概述 | 数据通道配置再在NFS上，作为临时数据存储的地方 | |
| 威胁 | STRIDE | |
| 仿冒（S） | 风险（Risk） | **风险等级：中**  **影响：中**  **可能性评估：低** |
| 已有消减措施 |  |
| 建议消减措施 | 客户端挂载NFS，需要增加NFS客户端对NFS服务端的认证 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：低**  数据被篡改会导致数据无法入库或者hive任务计算出错误的结果 |
| 已有消减措施 |  |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 抵赖（R） | 风险（Risk） | **风险等级：高**  **影响：高**  **可能性评估：操作系统日志可以追溯用户对NFS的操作** |
| 已有消减措施 |  |
| 建议消减措施 | 启动NFS审计日志功能，对用户操作进行监控 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：高**  存储在NFS上的数据泄露，会导致用户隐私数据泄露  **可能性评估：**  可能性较小 |
| 已有消减措施 | 1.VPN权限访问国内机房  2.OS用户权限审批 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：中**  **影响：中**  数据无法完成推动，结果文件无法推送回国内  **可能性评估：**   1. NFS所在的服务器或者网络出现异常 2. 可靠性保证 |
| 已有消减措施 | 运维监控BI内部机房服务器和网络的运行状态，如果有异常，及时进行处理 |
| 建议消减措施 | NFS具有单机宕机风险，需要实现NFS高可用设计 |
| 产品落地计划 |  |
| 特权提升（E） | 风险（Risk） | **风险等级：低**  **影响：中**  **可能性评估：NFS运行操作系统账号** |
| 已有消减措施 | 壳程序脚本权限（550） |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |

## 数据存储分析

### /NFS/DataIn

|  |  |  |
| --- | --- | --- |
| 元素名称 | /NFS/DataIn | |
| 元素概述 | 业务服务器推送数据后，数据在数据通道上的临时存放路径 | |
| 威胁 | TRID | |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：中**  数据被篡改后，执行数据入库任务报错或者入库错误的数据，导致开发人员获取不到数据或者获取到错误的计算结果  **可能性评估：**  1.具有VPN权限才能访问BI内部机房  2.访问数据通道的服务器权限，需要审批 |
| 已有消减措施 | 1. 华为专有机房，只有华为运维人员使用vpn登录到本地系统才能访问（统一由运维作为入口，如果统一由运维操作，可以去掉vpn相关内容）  2. 只有运行对应业务的用户才能访问，其他linux用户无法访问(文件，目录，都是什么权限，针对篡改进行权限设置说明，去掉写权限) |
| 建议消减措施 | 1.数据通道datain目录的数据定期清理，减少数据通道数据量，降低NFS服务器磁盘空间使用  2.加入数据质量监控，防止数据量猛增导致NFS服务不可用 |
| 产品落地计划 | NA |
| 抵赖（R） | 风险（Risk） | **风险等级：低**  **影响：低**  **可能性：低** |
| 已有消减措施 |  |
| 建议消减措施 | NA |
| 产品落地计划 | NA |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：高**  存放了用户的隐私数据，一旦泄露，影响较大  **可能性评估：**  vpn账号和对应的linux用户密码泄露导致恶意用户通过ssh接入服务器查看文件信息  可能非常小 |
| 已有消减措施 | 参考篡改已有消减措施 |
| 建议消减措施 | **拷贝不需要读权限，可以去掉** |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：中**  **影响：中**  数据无法推送，导致数据无法入库、计算和回传，影响数据开发人员进行数据产品开发  **可能性评估：高**  数据通道服务器异常或者网络异常 |
| 已有消减措施 | 运维监控BI内部机房服务器和网络的运行状态，如果有异常，及时进行处理 |
| 建议消减措施 | 1.数据通道datain目录的数据定期清理，减少数据通道数据量，降低NFS服务器磁盘空间使用  2.加入数据质量监控，防止数据量猛增导致NFS服务不可用 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |

### /NFS/Share

|  |  |  |
| --- | --- | --- |
| 元素名称 | /NFS/Share | |
| 元素概述 | 执行网关节点存储的执行脚本和执行工具 | |
| 威胁 | TRID | |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：中**  篡改后导致各类任务无法执行或者执行的结果不正确  **可能性评估：**  可能性较小 |
| 已有消减措施 | 1. 华为专有机房，只有华为运维人员使用vpn登录才能操作  2. 只有运行对应业务的用户才能访问，其他linux用户无法访问 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 抵赖（R） | 风险（Risk） | **不分析：可接受的风险** |
| 已有消减措施 | NA |
| 建议消减措施 | NA |
| 产品落地计划 | NA |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：低**  任务执行的相关信息会丢失，但是用户数据不会丢失。影响较小  **可能性评估：**  可能非常小 |
| 已有消减措施 | 1. 华为专有机房，只有华为运维人员使用vpn登录才能操作  2. 只有运行对应业务的用户才能访问，其他linux用户无法访问 |
| 建议消减措施 | 无 |
| 产品落地计划 | NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：中**  **影响：中**  无法执行各类任务  **可能性评估：高**  Share目录所在的服务器或者网络环境异常 |
| 已有消减措施 | NFS具有单机宕机风险，需要实现NFS高可用设计 |
| 建议消减措施 | 无 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |

### /NFS/DataOut

|  |  |  |
| --- | --- | --- |
| 元素名称 | /NFS/DataOut | |
| 元素概述 | 存储hive的计算结果，通过启动推送任务把结果回传到数据开发人员 | |
| 威胁 | TRID | |
| 篡改（T） | 风险（Risk） | **风险等级：低**  **影响：中**  待入库的数据一旦被篡改，将导致业务流程异常或数据异常，数据无法入库或者  **可能性评估：**  可能性非常低 |
| 已有消减措施 | 1. 华为专有机房，只有华为运维人员使用vpn登录才能操作  2. 只有运行对应业务的用户才能访问，其他linux用户无法访问 |
| 建议消减措施 | 1. 无 |
| 产品落地计划 | 1. NA |
| 抵赖（R） | 风险（Risk） | **风险等级：中**  **影响：中**  **可能性：中** |
| 已有消减措施 | NA |
| 建议消减措施 | 启动NFS审计日志功能，对用户操作进行监控 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |
| 信息泄漏（I） | 风险（Risk） | **风险等级：低**  **影响：高。**DataOut中如果存储了用户隐私数据，将导致隐私数据泄露，影响很大。  **可能性评估：**  可能性非常低 |
| 已有消减措施 | 1. 华为专有机房，只有华为运维人员使用vpn登录才能操作  2. 只有运行对应业务的用户才能访问，其他linux用户无法访问 |
| 建议消减措施 | 1. 无 |
| 产品落地计划 | 1. NA |
| 拒绝服务（D） | 风险（Risk） | **风险等级：中**  **影响：中。**导致数据无法推送到开发人员侧，无法完成数据产品开发。  **可能性评估：**  数据通道所在的服务器或者网络出现异常 |
| 已有消减措施 |  |
| 建议消减措施 | 1. 1.数据通道dataout目录的数据定期清理，减少数据通道数据量，降低NFS服务器磁盘空间使用 2. 2.加入数据质量监控，防止数据量猛增导致NFS服务不可用 |
| 产品落地计划 | 20160810-DSPT 1.0.3.400 |

## 隐私设计规范Checklist自检



## 缺陷特性汇总

本章节将以上分析出的缺陷及消减措施汇总，便于跟踪落实。



# 安全及隐私保护设计

## 安全需求设计

### 秘钥管理

1. 根秘钥实现规则

根秘钥有两部分组成，一部分保存在配置文件中，一部分保存在代码中，二者拼接形成根秘钥；

### 传输安全设计

所有涉及与周边部件的交互必须确保传输过程是安全的。

1. 本系统中，与周边华为部件的交互，采用ssh进行传输。
2. 与周边华为部件交互时，需要做身份认证。ssh为通过公钥私钥配对的方式进行身份认证
3. 与周边华为部件的交互，必须使用内部ip进行通信，同时要配允许访问的ip白名单列表；
4. 本系统与产品人员进行交互时，使用SSH协议进行产品人员的身份认证。

### 日志安全设计

1. 非必要场景，禁止打印用户敏感信息（IMEI、手机号、email）；
2. 日志否有有做访问权限控制（没有强制要求）不要放到portal

### 用户权限设计

该系统的用户权限管理分两类，一类是基于操作系统的用户权限管理，该权限管理的粒度精确到目录。另一类是基于FusionInsight中kerbros的权限管理，该类管理是基于角色的权限控制，权限的控制粒度，对hdfs精确到目录，对hive精确到表。

基于操作系统的权限管理信息存储在操作系统内部。基于kerbros的权限管理，权限信息存储在kerbors服务内部。

## 隐私保护设计

### 个人数据列表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 操作场景 | 收集目的 | 个人数据 | 个人数据处理过程 | 保护措施 | 数据存留期 |
| 数据推送，数据回推 | 批处理计算 | 业务用户的手机号、imei号、email | 在Hive中进行计算 | 传输通道加密；  数据存储加密 | 长期 |

### 隐私数据保护设计

参考秘钥管理和传输安全设计。

# 安全及隐私保护设计检视

## 检视对象与策略

**检视人：**

王光明 00257730，李勇 00164768，罗向龙 00201279, 周川祥 00190491，张政勇 00215119

**检视时间:**

2016.4.10~2016.4.25

**检视策略：**

本次检视主要针对datapush、dataload以及壳工具等BI常用工具在日志安全性、密钥和口令安全性以及NFS 安全性进行了全量检视。

## 结果评价

**本次检视供发现安全设计问题4个，已在当前版本中进行了修改。**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *高风险问题数* | *当前版本已整改* | *计划其它版本整改* | *无整改计划* | *已备案* |
| *0* | *0* | *4* | *0* | *1* |

所发现的问题都属于违反公司安全设计规范的问题，如密钥的管理和日志的管理等。对于红线，产品基本上都满足了相应要求，但对于各领域的安全设计规范，开发人员在开发过程还缺少相应的遵从。后续在开发过程中需要开发人员熟悉相应安全设计规范，尤其是在会话管理、身份认证管理上，需要加强对中低级别问题的重视和修复*。*

## 问题分析

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 编号 | 技术领域 | 自检问题以及风险详细描述 | 风险等级 | 改进方案/建议 | 计划整改落地版本 |
| 1 | 日志管理 | 壳程序中使用system.out打印日志到控制台 | 低 | 使用专用日志库重定向日志到文件中 |  |
| 2 | 密钥管理 | AES128加密使用EBC模式不符合公司密钥管理要求 | 中 | 使用CBC模式的ACE128加密算法对口令进行加密 |  |
| 3 | 密钥管理 | 密钥因子未分开存储 | 中 | 密钥因子要分开存储 |  |
| 4 | 权限管理 | 壳程序存储目录权限太大 | 低 | 最小化壳程序目录权限，只允许本用户具有读写权限 |  |
|  |  |  |  |  |  |

## 安全及隐私保护设计检视Checklist

（附上安全及隐私保护设计检视结果）



# 备注和附录

不涉及

### 

参考资料清单：

《产品网络安全红线落地解读及指导V2\_0.xls》

《消费者BG隐私设计规范Checklist-V1.0.xlsx》

《密码算法应用规范\_V1.0.docx》