|  |  |  |
| --- | --- | --- |
| 华为技术有限公司  Huawei Technologies Co. Ltd. | 版本 | 密级 |
| V1.00 | 内部公开 |
| 名称 | 总页数 |
| 大数据平台白皮书 |  |

大数据平台白皮书

|  |  |  |  |
| --- | --- | --- | --- |
| 拟制 |  | 日期 |  |
| 审核 |  | 日期 |  |
| 批准 |  | 日期 |  |



华为技术有限公司

Huawei Technologies Co., Ltd.

版权所有 侵权必究

All rights reserved

修订记录

|  |  |  |  |
| --- | --- | --- | --- |
| 日期 | 修订版本 | 描述 | 作者 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

目 录

1 总述 1

2 平台使用规范 1

2.1 1

3 平台安全规范 1

4 平台编码规范 1

5 平台自研工具 1

6 UDF 1

7 平台维优案例 1

7.1 现网开发问题定位 1

7.2 现网维优经验 1

8 平台维优工具集 1

9 第三方工具 1

9.1 FI文档 1

9.1.1 FI产品文档 1

9.1.2 FI维护宝典 1

9.2 InforSight文档 1

9.3 MFS 1

**大数据平台白皮书**

关 键 词

摘 要

缩略语清单

| 缩略语 | 英文全名 | 中文解释 |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

# 总述

本文档统一整理归档大数据平台使用和维优中的各种文档，便于大家查阅和按规范执行。主要包括平台使用规范、平台安全、平台编码规范、平台维优案例、平台维优工具集等等。

# 平台使用规范



# 平台安全规范

# 平台编码规范

请参考模型组编写的《DataOne编码规范》。



为了保证文档的准确和完整，请向模型组获取该文档的最新版本。

# 平台自研工具

平台组自研工具主要包括以下工具：

1. TCC
2. 壳工具
3. 自助查询工具
4. Data\_Agent
   1. Flume
   2. DataPush
5. Data\_Gateway
   1. DataLoad
   2. HIBI-Export
   3. 壳工具HIBI-ExecuteShell
   4. EndLogDownloader
   5. DataPipeLine（数据冷备）

平台工具的相关文档可以在平台组发布的版本的VMP中获取。Bigdata platform 2.0.4.300的VMP文档路径如下：

\\10.1.90.32\IT\_VMP\_SZX\_green\_2\_F\BL22528813\_BigData Platform 2.0.4.300\ReleaseDoc\

# UDF

UDF函数相关文档可以在平台组发布的版本的VMP中获取。Bigdata platform 2.0.4.300的VMP文档路径如下：

\\10.1.90.32\IT\_VMP\_SZX\_green\_2\_F\BL22528813\_BigData Platform 2.0.4.300\ReleaseDoc\BigData Platform\_2.0.4.300\_版本说明书\Common\_UDF

# 平台维优案例

## 现网开发问题定位

## 现网维优经验



# 平台维优工具集

# 第三方工具

## FI文档

FI在support网站上提供了大量的相关文档。地址如下：<http://support.huawei.com/enterprise/productNewOffering?idAbsPath=7919749|7919788|19942925|21110924&pid=21110924>

### FI产品文档

FI提供的文档中，最重要的是产品文档。用户可以从FI的support网站下载对应版本的产品文档。

FusionInsight HD产品文档由以下内容组成：

* 描述类文档基于各种维度提供产品的描述类信息。
* 过程类文档在操作和维护方面提供说明。
* 参考类文档在操作的过程需要参考信息时提供附加的信息。

| **文档类型** | **文档名称** | **内容简介** | **读者对象** |
| --- | --- | --- | --- |
| 描述类 | 产品描述 | 本文档对FusionInsight HD产品进行详细描述，包含产品介绍、系统架构介绍、部署方案、企业级增强特性介绍等信息。 | * 规划工程师 * 运维工程师 |
| 过程类 | 软件安装 | 本文档介绍FusionInsight HD的部署方案及软件安装的具体操作过程。 | 安装工程师 |
| 管理员指南 | 本文档提供详细指导，供系统管理员进行日常健康检查、集群管理、用户管理、日志管理、备份与恢复、告警处理、故障处理等操作。 | 运维工程师 |
| 故障管理 | 本文档介绍日常维护过程中，业务出现故障时收集故障信息、处理故障的操作。 | * 运维工程师 * 技术支持工程师 |
| 应用开发指南 | 本文档提供应用开发的样例及开发规范，专供需要在FusionInsight HD上进行业务应用开发的用户使用，主要适用于具备Java开发经验的开发人员。 | 软件开发工程师 |
| 参考类 | 术语 | 本文档提供FusionInsight HD产品在整个使用阶段中涉及到的基本概念信息。 | * 规划工程师 * 安装工程师 * 运维工程师 * 软件开发工程师 |

### FI维护宝典

FI维护宝典详细的描述了FusionInsight HD在安装、维护、二次开发、升级等过程中场景的故障场景以及具体的操作指导。可以在下面的URL中获取最新的FI维护宝典。

http://support.huawei.com/ehedex/hdx.do?docid=DOC1000126368&lang=zh&clientWidth=1664&browseTime=1504513424022

该URL目前只有华为员工才可以访问，合作方员工可是使用离线版本的文档。

要使用该离线版本文档，需要下载华为电子文档桌面管理软件HedEx Lite。HedEx Lite无需安装，您只需拷贝至计算机中直接双击Startup.exe就可使用（如果是压缩包，需解压后才能使用）。

## InforSight文档

InforSight，智能、开放、高效、易用的大数据分析挖掘平台，通过对各类海量数据信息进行实时和非实时的分析和挖掘，帮助企业从海量数据信息中获取到真正的价值，及时洞察和决策新的机会与风险。

InforSight基于当前流行的Hadoop/Spark平台，采用分布式存储和并行计算技术，具备高性能内存计算能力；兼容通用的Python/R数据分析环境，数据科学家可使用自己擅长语言进行数据挖掘；提供了分布式算法（Java、R、Python三种语言）和可视化能力，预置业界主流挖掘算法，帮助用户有效、快速完成数据挖掘。

在大数据平台版本发布路径中可以获取InforSight的相关文档。Bigdata platform 2.0.4.300的VMP文档路径如下：

\\10.1.90.32\IT\_VMP\_SZX\_green\_2\_F\BL22528813\_BigData Platform 2.0.4.300\ReleaseDoc\BigData Platform\_2.0.4.300\_版本升级指导\DataSight InforSight\_V100R003C00B011

## MFS

在大数据平台版本发布路径中可以获取MFS的相关文档。Bigdata platform 2.0.4.300的VMP文档路径如下：

\\10.1.90.32\IT\_VMP\_SZX\_green\_2\_F\BL22528813\_BigData Platform 2.0.4.300\ReleaseDoc\BigData Platform\_2.0.4.300\_版本升级指导\MFS

MFS的使用可以参考MFS的文档：