

中国移动广东公司 2015 年度 Hadoop 平台项目测试规范

The Test Specification for
Hadoop Platform

版本号：2.0.1



IBM（中国）有限公司全球服务部

目	录	
前	言	II
1.	文档说明	1
2.	项目概述	1
3.	术语、定义和缩略语	1
3.1.	术语/定义	1
3.2.	缩略语	2
4.	测试环境	2
4.1.	测试环境组网拓扑	3
4.2.	硬件配置要求	4
5.	数据模型	4
5.1.	数据表模型	4
5.1.1.	语音主叫话单	4
5.1.2.	语音被叫话单	10
5.1.3.	PS 域用户业务表	16
5.1.4.	PS 域 WAP/HTTP 事件话单	17
5.2.	维度表模型	19
5.2.1.	小区维表	19
5.2.2.	业务维表	20
5.2.3.	终端类型维表	20
5.3.	接口数据格式及大小	20
6.	测试内容	21
6.1.	总体约束	22
6.2.	平台规划能力测试	25
6.3.	运维优化能力测试	26
6.3.1.	平台管理能力测试	26
6.3.2.	故障处理能力测试	31
6.3.3.	调优能力测试	35
6.4.	模型管控能力测试	43
6.4.1.	离线批量复杂分析	43
6.4.2.	即席查询	50
6.4.3.	简单键值查询	54
6.4.4.	准实时流计算	56

前 言

本标准的目的是对制定对中国移动广东公司网管中心Hadoop平台运营相关能力的测试标准。

1. 文档说明

本规范为中国移动广东公司网络管理中心 Hadoop 平台项目测试规范。

本规范是中国移动广东公司进行 Hadoop 平台相关厂商招标的测试方法和测试依据，具体内容包括测试环境、测试数据以及测试内容等三个方面。

2. 项目概述

大数据技术的出现为海量数据的处理提供了高性价比的解决方案。由于采用 x86 通用服务器作为硬件集群，通过并行计算和分布式存储实现系统的扩展性，通过冗余备份实现系统的可靠性，以 Hadoop 为代表的大数据平台在大型系统建设成本、扩展性方面具有强大的优势，这种平台已成为大型互联网公司的基础平台，国内外运营商也开始逐步引入大数据技术建设相关系统。

由于大数据发展时间相对较短，很多技术/产品都以开源形式出现，系统的可靠性、稳定性、可维护性尚不能达到传统运营商的要求，在关键系统中引入大数据技术初期离不开专业厂商支持。

为了系统评估各种技术的可用性、可靠性以及支持厂家的技术水平，本项目拟为中国移动广东公司网络管理中心提供支撑，提供 Hadoop 平台运营相关各项能力的测试规范，其中包括平台规划能力、运维优化能力以及模型管控能力以满足网络管理中心对现有 Hadoop 平台运营的各类应用场景的需求；与此同时，项目同时也会兼顾对未来平台架构的演进规划，以便为后续平台扩展打好基础。

3. 术语、定义和缩略语

下列术语、定义和缩略语适用于本标准。

3.1. 术语/定义

表3-1 术语/定义

术语/定义	解释
X86集群	指符合X86架构的PC服务器集群
SAN	存储区域网络（Storage Area Network，简称SAN）。采用光纤通道（Fiber Channel，简称FC）技术，通过光纤通道交换机连接存储阵列和服务器主

	机，建立专用于数据存储的区域网络。
SVC	SAN卷控制器（SAN Volume Controller，简称SVC）。是一个存储虚拟化系统，可以对存储资源实现单点控制，从而有助于支持提高的业务应用程序可用性和更高的资源利用率。
接口机集群	主要完成日志数据获取功能，将合并后的大文件加载至Hadoop集群
Hadoop集群	包括日志处理和网页分析功能，底层基于X86集群
获取层	实现集中化经分Hadoop平台从各数据源系统中采集相关基础源数据的功能
数据处理层	实现面向互联网分析所需的网页分析数据处理与计算和面向流量清单查询的原始网络日志的处理

3.2. 缩略语

表3-2 缩略语

缩写	英文描述	中文描述
ETL	Extraction, Transformation, Loading	抽取、转换和加载（是数据获取过程）
BASS	Business Analyse Support System	经营分析系统
HDFS	Hadoop Distributed File System	Hadoop分布式文件系统
Map Reduce	Map Reduce	Hadoop并行计算框架
HBase	Hadoop Database	Hadoop分布式数据库
ZooKeeper	ZooKeeper	Hadoop分布式同步软件
HadoopManager	HadoopManager	Hadoop配置管理软件
UTF	Unicode Transformation Format	Unicode传输格式
URL	UniformResourceLocator	统一资源定位符
TB	Terabyte	万亿字节
MSISDN	Mobile Subscriber International ISDN/PSTN number	唯一能识别移动用户的号码

4. 测试环境

Hadoop 测试平台测试环境主要包括数据发布部分（数据发布服务器、SAN、SVC、后端存储设备）和参测厂商集群部分（应用服务器）。

4.1. 测试环境组网拓扑

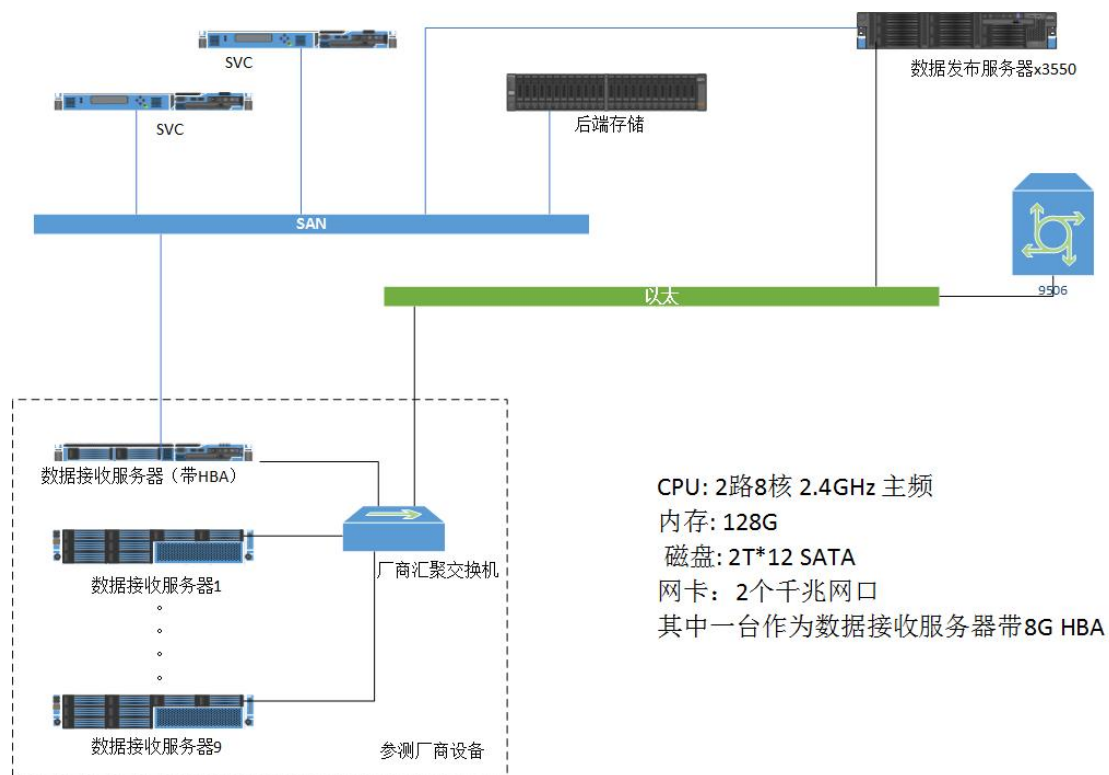


图 1 Hadoop 平台系统测试环境组网拓扑

Hadoop 平台系统测试环境组网拓扑中各个节点类型和说明见下表。

表4-1 测试环境各节点类型表

设备域		配置和功能说明
数据发布部分	数据发布服务器	<ul style="list-style-type: none">● 由两台 IBM x3550的通用X86架构服务器构成。● 功能：<ul style="list-style-type: none">■ 分发测试数据到参测厂商集群。■ 调用测试厂商程序，收集测试结果。
	SVC	<ul style="list-style-type: none">● 由两台 IBM SVC构成。● 功能：对存储资源实现控制。
	SAN	<ul style="list-style-type: none">● 由一台 IBM SAN构成。● 功能：建立专用于数据存储的区域网络。
	后端存储设备	<ul style="list-style-type: none">● 由一台 IBM XIV存储器组成，总容量100T。● 功能：存放测试数据。
参测厂商集群部分	应用服务器	<ul style="list-style-type: none">● 由10台中端内置大容量硬盘的通用X86架构服务器构成。其中一台带8G HBA作为数据接收服务器。● 功能：参测厂商Hadoop平台机器。

4.2. 硬件配置要求

测试拓扑：10 台 x86 架构服务器 + 1 台千兆以太网交换机

表4-2 x86架构国产服务器（应用服务器）配置信息

机器数量	10台，其中一台需带8G HBA
处理器	2*8核*2.4GHz（不能高于该标准）
内存	128GB
硬盘	不少于12个硬盘，总容量不少于24T
千兆网卡	不超过两块

表4-3 千兆以太网交换机配置信息

设备数量	1
端口类型	1000M
端口数量	24

5. 数据模型

数据模型主要描述本次测试所采用的接口文件数据表及其结构、数据表的使用及在测试过程中对数据表的各项相关操作等内容。

5.1. 数据表模型

本次测试中用到的数据表如下：

数据表	中文名称	字段数目	数据规模（亿条）
AIU_MOC	语音主叫话单	99	400
AIU_MTC	语音被叫话单	102	400
PS_USERSERVICE	PS域用户业务表	36	400
PS_WAPHTTP	PS域WAP/HTTP事件话单	42	400

5.1.1. 语音主叫话单

序号	中文描述	英文描述	字段类型	备注
1	数据创建时间	CREATETIME	NUMBER(10)	相对 1970 年的秒数。（时间为 UTC

	(s)			格式，下同)
2	开始时间 (s)	STARTTIME	NUMBER(10)	相对 1970 年的秒数。
3	起始毫秒	MILLISEC	NUMBER(3)	开始时间的毫秒部分
4	业务状态	SRVSTAT	NUMBER(1)	0、 成功 1、 失败 2、 切出存在 alerting / connect 消息， 则为成功；否则为失败。如果发生 了切出流程，则状态为切出。
5	单据状态	CDRSTAT	NUMBER(1)	0、 正常 1、 超时 2、 不完整首先判断是否存在任意一 个呼叫定时器超时，是则记为 1 (超时)；如果没有定时器超时， 没有收到 SCM_CR，或者 SCM_CC 则记为 2 (不完整)； 收到 SCM_CREF 或者 (SCM_RLSD, SCM_RLC) 均记 为 0 (正常)
6	网络标识	NI	NUMBER(1)	0:国际主用、 1:国际备用、 2:国内主用、 3:国内备用
7	BSC/RNC 信令点	OPC	NUMBER(8)	
8	MSC 信令点	DPC	NUMBER(8)	
9	链路类型	LINKTYPE	NUMBER(5)	基于 SCTP 的链路填写为：0 基于 TDM, ATM 的链路填写为：1 IP 链路 (0)，SS7 链路 (1)
10	接入网类型	ACCESS_TY P E	NUMBER(1)	0:2G 1:3G
11	业务类型	SRVTYPE	NUMBER(1)	0: 语音呼叫 1: 视频呼叫 2: 数据业务
12	信道类型	CHTYPE	NUMBER(3)	1—SDCCH 4--8 全速率 TCH 8--1 全速率 TCH 9--1 半速率 TCH 10--2 全速率 TCH 11--3 全速率 TCH 12--4 全速率 TCH 13--5 全速率 TCH 14--6 全速率 TCH 15--7 全速率 TCH

13	CIC 系统号	CICSYSNO	NUMBER(4)	E1 号 可空
14	CIC 时隙号	CICTSNO	NUMBER(3)	E1 内时隙号 可空
15	主叫 IMSI	IMSI	VARCHAR2(20)	
16	主叫 IMEI	IMEI	VARCHAR2(16)	
17	主叫 TMSI	TMSI	VARCHAR2(20)	
18	拨打号码	CALLEDNO	VARCHAR2(32)	主叫实际拨打的号码，可能含前缀
19	被叫号码类型	CLDTYPE	NUMBER(1)	
20	主叫号码	CALLERNO	VARCHAR2(32)	
21	移动国家码	MCC	CHAR(3)	
22	移动网络码	MNC	CHAR(3)	
23	终结小区的位置区编码	LastLAC	VARCHAR2(10)	
24	终结小区的小区标识	LastCI	VARCHAR2(10)	
25	业务请求接受时间	CM_SRVACP_TIME	NUMBER(10)	与开始时间的时间差值（毫秒），可空
26	鉴权请求时间	AUTH_REQ_TIME	NUMBER(10)	同上
27	鉴权响应时间	AUTH_RSP_TIME	NUMBER(10)	同上
28	身份识别请求时间	IDENTITY_REQ_TIME	NUMBER(10)	同上
29	身份识别响应时间	IDENTITY_RSP_TIME	NUMBER(10)	同上
30	加密模式请求时间	CIPH_REQ_TIME	NUMBER(10)	同上
31	加密模式请求响应时间	CIPH_RSP_TIME	NUMBER(10)	同上
32	呼叫建立时间	SETUP_TIME	NUMBER(10)	同上
33	呼叫进行时间	CALL_PROC_TIME	NUMBER(10)	同上
34	指配请求时间 (RAB 指配请求时间)	ASSN_TIME	NUMBER(10)	同上
35	指配完成时间 (RAB 指配完成时间)	ASSN_CMPT_TIME	NUMBER(10)	同上

36	振铃时间 (ms)	ALERT_TIME	NUMBER(10)	同上
37	应答时间 (ms)	ANSWER_TIME	NUMBER(10)	同上
38	断连时间	DISCONN_TIME	NUMBER(10)	同上
39	释放时间	REL_TIME	NUMBER(10)	同上
40	释放完成时间	RELCMP_TIME	NUMBER(10)	同上
41	清除时间	CLR_CMD_TIME	NUMBER(10)	同上
42	清除完成时间	CLR_CMP_TIME	NUMBER(10)	同上
43	结束时间 (ms)	END_TIME	NUMBER(10)	同上
44	结束位置子协议类型	PPD	NUMBER(2)	
45	结束位置	REL_PHASE	NUMBER(3)	结束消息的前一条消息
46	结束消息子协议类型	PD	NUMBER(2)	MS、BSC、RNC 到 MSC 方向： 0 BSSMAP 1 RANAP 2 DTAPMM 3 DTAPCC 4 DTAPRR 5 DTAPSS 6 SMSRP 14 SMSCP MSC 到 MS、BSC、RNC 方向： 7 BSSMAP 8 RANAP 9 DTAPMM 10 DTAPCC 11 DTAPRR 12 DTAPSS 13 SMSRP 15 SMSCP 16 SCCP 98 未采集到的消息 99 UNKNOWN
47	结束消息	FIRFAILMSG	NUMBER(3)	第一拆线消息
48	结束原因	CAUSE	NUMBER(3)	对应结束阶段的结束原因值
49	业务请求模式	Request Mode	NUMBER(1)	0:采用 IMSI 起呼 1:采用 TMSI 起呼

50	CDR 标识	CDRID	varchar2 (20)	
51	呼叫进展时间 (ms)	PROGRESS_T IME	NUMBER(10)	
52	AbisCDR 标识	ABISCDRID	varchar(20)	A+ABIS 原始单据(或 ABIS 原始单据) 中的 ABIS MR ID
53	A 口标识	ACDRID	varchar(20)	关联 A 口单据时, 填写 A 口单据的 CDRID
54	ABIS 标识	ABIS	numeric(1)	0: A 1: Abis 2: A+Abis 3: A+Abis+MR 4: Abis+MR 5: MR
55	被叫号码	CALLEDUSRN 0	VARCHAR2(2 4)	格式化后的用户号码
56	起始小区的位 置区编码	FirstLAC	char(4)	
57	起始小区的小 区标识	FirstCI	char(4)	
58	第一小区载波 编号	FirstTEI	numeric(3)	
59	第一小区信道 号码	Firstchann el	numeric(3)	
60	上行信号强度 总和	SumRxlevlU L	Bigint (8BYTE)	
61	下行信号强度 总和	SumRxlevlD L	Bigint (8BYTE)	
62	上行信号质量 总和	SumRxqualU L	Bigint (8BYTE)	
63	下行信号质量 总和	SumRxqualD L	Bigint (8BYTE)	
64	上行功控强度 总和	SumBspwr	Bigint (8BYTE)	
65	下行功控强度 总和	SumMspwr	Bigint (8BYTE)	
66	TA 总和	SumTA	Bigint (8BYTE)	
67	呼叫 MR 总数	TotalMRno	Int (4BYTE)	
68	MR 总数	TOTALMRCOU NT	Int (4BYTE)	
69	全速率 MR 总数	TotalFullR	Int (4BYTE)	

		ateMRno		
70	半速率 MR 总数	TotalHalfRateMRno	Int (4BYTE)	
71	0 级上行信号质量次数	SumRxqual0	Int (4BYTE)	针对 TCH 信道
72	1 级上行信号质量次数	SumRxqual1	Int (4BYTE)	
73	2 级上行信号质量次数	SumRxqual2	Int (4BYTE)	
74	3 级上行信号质量次数	SumRxqual3	Int (4BYTE)	
75	4 级上行信号质量次数	SumRxqual4	Int (4BYTE)	
76	5 级上行信号质量次数	SumRxqual5	Int (4BYTE)	
77	6 级上行信号质量次数	SumRxqual6	Int (4BYTE)	
78	7 级上行信号质量次数	SumRxqual7	Int (4BYTE)	
79	0 级下行信号质量次数	SumRxqualDL0	Int (4BYTE)	
80	1 级下行信号质量次数	SumRxqualDL1	Int (4BYTE)	
81	2 级下行信号质量次数	SumRxqualDL2	Int (4BYTE)	
82	3 级下行信号质量次数	SumRxqualDL3	Int (4BYTE)	
83	4 级下行信号质量次数	SumRxqualDL4	Int (4BYTE)	
84	5 级下行信号质量次数	SumRxqualDL5	Int (4BYTE)	
85	6 级下行信号质量次数	SumRxqualDL6	Int (4BYTE)	
86	7 级下行信号质量次数	SumRxqualDL7	Int (4BYTE)	
87	TA 为 0 或 1 的次数	TA01Count	Int (4BYTE)	
88	Rxlevel Down 大于-85 次数	Rxlevel_Down	Int (4BYTE)	Rxlevel Down 大于-85 次数
89	弱覆盖 MR 个数	Weak_coverage	Int (4BYTE)	DL_RXLEVEL<-90 and TA<2 次数
90	过覆盖 MR 个数	Over_coverage	Int (4BYTE)	DL_RXLEVEL>-80 and TA>1 次数

91	上下行不平衡 MR 个数	Imbalance	Int (4BYTE)	abs(DL_RXLEVEL-UL_RXLEVEL)>=15 次数
92	干扰 MR 个数	Interference	Int (4BYTE)	DL_RXLEVEL>-80 and DL_RXQUAL>4 次数
93	立即指配命令时刻	Imm_Ass_Co md_Time	numeric(10)	
94	立即指配拒绝时刻	Imm_Ass_Re j_Time	numeric(10)	
95	立即指配建立成功时刻	Imm_Ass_Co mp_Time	numeric(10)	
96	指配命令时间	Ass_Comd_T ime	numeric(10)	
97	掉话 Connection_Fa ilure 时间	Connection _Failure_T ime	numeric(10)	
98	SCCPID	SCCP 连接标 识	Varchar(20)	同一个 SCCP 连接上所有 A 接口单据填写相同的 SCCPID，华为探针中填写同一个 SCCP 连接中首个 CDR 的 CDR 标识
99	CSFBIND	CSFB 指示	numeric(2, 0)	1: CS fallback mobile terminating call 2: CS fallback mobile originating call 其余情况为 0

5.1.2. 语音被叫话单

序号	中文描述	英文描述	字段类型	备注
1	数据创建时间 (s)	CREATETIME	NUMBER(10)	相对 1970 年的秒数。(时间为 UTC 格式，下同)
2	开始时间 (s)	STARTTIME	NUMBER(10)	开始时间：相对 1970 年的秒数。
3	起始毫秒	MILLISEC	NUMBER(3)	开始时间的 ms 部分
4	业务状态	SRVSTAT	NUMBER(1)	0、成功 1、失败 存在 alerting/connect 消息，则为成功；否则为失败。 如果发生了切出流程，则状态为切出
5	单据状态	CDRSTAT	NUMBER(1)	0、正常 1、超时 2、不完整 首先判断是否存在任意一个呼叫定时器超时，是则记为 1（超时）；如果没有定时器超时，没有收到 SCM_CR，或

				者 SCM_CC 则记为 2（不完整）；收到 SCM_CREF 或者（SCM_RLSD，SCM_RLC）均记为 0（正常）
6	网络标识	NI	NUMBER(1)	网络标识： 0:国际主用、 1:国际备用、 2:国内主用、 3:国内备用
7	BSC/RNC 信令点	OPC	NUMBER(8)	
8	MSC 信令点	DPC	NUMBER(8)	
9	链路类型	LINKTYPE	NUMBER(5)	基于 SCTP 的链路填写为：0 基于 TDM, ATM 的链路填写为：1 IP 链路（0），SS7 链路（1）
10	接入网类型	ACCESS_TYPE	NUMBER(1)	0:2G 1:3G
11	业务类型	SRVTYPE	NUMBER(1)	0:语音呼叫 1:视频呼叫 2:数据业务
12	信道类型	CHTYPE	NUMBER(3)	1—SDCCH 4—8 全速率 TCH 8--1 全速率 TCH 9--1 半速率 TCH 10--2 全速率 TCH 11--3 全速率 TCH 12--4 全速率 TCH 13--5 全速率 TCH 14--6 全速率 TCH 15--7 全速率 TCH
13	CIC 系统号	CICSYSNO	NUMBER(4)	E1 号 可空
14	CIC 时隙号	CICTSNO	NUMBER(3)	E1 内时隙号 可空
15	被叫 IMSI	IMSI	VARCHAR2(20)	
16	被叫 IMEI	IMEI	VARCHAR2(16)	
17	被叫 TMSI	TMSI	VARCHAR2(20)	
18	被叫号码	CALLEDNO	VARCHAR2(32)	
19	主叫号码	CALLERNO	VARCHAR2(32)	
20	移动国家码	MCC	CHAR(3)	
21	移动网络码	MNC	CHAR(3)	
22	终结小区的位	LastLAC	VARCHAR2(1)	

	置区编码		0)	
23	终结小区的小区标识	LastCI	VARCHAR2(10)	
24	第一次寻呼时间 (ms)	FIRPAGING_TIME	NUMBER(10)	与开始时间的时间差值 (毫秒), 可空
25	第二次寻呼时间 (ms)	SECPAGING_TIME	NUMBER(10)	同上
26	第三次寻呼时间 (ms)	THIRDPAGING_TIME	NUMBER(10)	同上
27	第四次寻呼时间 (ms)	FOURTHPAGING_TIME	NUMBER(10)	同上
28	第五次寻呼时间 (ms)	FIFTHPAGING_TIME	NUMBER(10)	同上
29	寻呼响应时间 (ms)	PAGINGRSP_TIME	NUMBER(10)	同上
30	鉴权请求时间	AUTH_REQ_TIME	NUMBER(10)	同上
31	鉴权响应时间	AUTH_RSP_TIME	NUMBER(10)	同上
32	身份识别请求时间	IDENTITY_REQ_TIME	NUMBER(10)	同上
33	身份识别响应时间	IDENTITY_RSP_TIME	NUMBER(10)	同上
34	加密模式请求时间	CIPH_REQ_TIME	NUMBER(10)	同上
35	加密模式请求响应时间	CIPH_RSP_TIME	NUMBER(10)	同上
36	呼叫建立时间	SETUP_TIME	NUMBER(10)	同上
37	呼叫确认时间	CALL_PROC_TIME	NUMBER(10)	同上
38	指配请求时间 (RAB 指配请求时间)	ASSN_TIME	NUMBER(10)	同上
39	指配完成时间 (RAB 指配完成时间)	ASSN_CMPT_TIME	NUMBER(10)	同上
40	振铃时间 (ms)	ALERT_TIME	NUMBER(10)	同上
41	应答时间 (ms)	ANSWER_TIME	NUMBER(10)	同上
42	断连时间	DISCONN_TIME	NUMBER(10)	同上
43	释放时间	REL_TIME	NUMBER(10)	同上
44	释放完成时间	RELCMP_TIME	NUMBER(10)	同上

		E		
45	清除时间	CLR_CMD_TIME	NUMBER(10)	同上
46	清除完成时间	CLR_CMP_TIME	NUMBER(10)	同上
47	结束时间 (ms)	END_TIME	NUMBER(10)	同上
48	结束位置子协议类型	PPD	NUMBER(2)	
49	结束位置	REL_PHASE	NUMBER(3)	结束消息的前一条消息
50	结束消息子协议类型	PD	NUMBER(2)	MS、BSC、RNC 到 MSC 方向： 0 BSSMAP 1 RANAP 2 DTAPMM 3 DTAPCC 4 DTAPRR 5 DTAPSS 6 SMSRP 14 SMSCP MSC 到 MS、BSC、RNC 方向： 7 BSSMAP 8 RANAP 9 DTAPMM 10 DTAPCC 11 DTAPRR 12 DTAPSS 13 SMSRP 15 SMSCP 16 SCCP 98 未采集到的消息 99 UNKNOW
51	结束消息	FIRFAILMSG	NUMBER(3)	第一拆线消息
52	结束原因	CAUSE	NUMBER(3)	对应结束阶段的结束原因值
53	寻呼模式	Paging Mode	NUMBER(1)	0:采用 IMSI 寻呼 1:采用 TMSI 起呼
54	CDR 标识	CDRID	varchar2 (20)	
55	AbisCDR 标识	ABISCDRID	varchar(20)	A+ABIS 原始单据(或 ABIS 原始单据)中的 ABIS MR ID
56	A 口标识	ACDRID	varchar(20)	关联 A 口单据时, 填写 A 口单据的 CDRID

57	ABIS 标识	ABIS	numeric(1)	0: A 1: Abis 2: A+Abis 3: A+Abis+MR 4: Abis+MR 5: MR
58	起始小区的位置区编码	FirstLAC	char(4)	
59	起始小区的小区标识	FirstCI	char(4)	
60	第一小区载波编号	FirstTEI	numeric(2)	
61	第一小区信道号码	Firstchannel	numeric(2)	
62	上行信号强度总和	SumRxlevelUL	Bigint (8BYTE)	
63	下行信号强度总和	SumRxlevelDL	Bigint (8BYTE)	
64	上行信号质量总和	SumRxqualUL	Bigint (8BYTE)	
65	下行信号质量总和	SumRxqualDL	Bigint (8BYTE)	
66	上行功控强度总和	SumBspwr	Bigint (8BYTE)	
67	下行功控强度总和	SumMspwr	Bigint (8BYTE)	
68	TA 总和	SumTA	Bigint (8BYTE)	
69	呼叫 MR 总数	TotalMRno	Int(4BYTE)	
70	MR 总数	TOTALMRCOUNT	Int(4BYTE)	
71	全速率 MR 总数	TotalFullRateMRno	Int(4BYTE)	
72	半速率 MR 总数	TotalHalfRateMRno	Int(4BYTE)	
73	0 级上行信号质量次数	SumRxquaUL0	Int(4BYTE)	
74	1 级上行信号质量次数	SumRxquaUL1	Int(4BYTE)	
75	2 级上行信号质量次数	SumRxquaUL2	Int(4BYTE)	
76	3 级上行信号质量次数	SumRxquaUL3	Int(4BYTE)	
77	4 级上行信号质	SumRxquaUL	Int(4BYTE)	

	量次数	4		
78	5 级上行信号质量次数	SumRxquaUL 5	Int (4BYTE)	
79	6 级上行信号质量次数	SumRxquaUL 6	Int (4BYTE)	
80	7 级上行信号质量次数	SumRxquaUL 7	Int (4BYTE)	
81	0 级下行信号质量次数	SumRxquaDL 0	Int (4BYTE)	
82	1 级下行信号质量次数	SumRxquaDL 1	Int (4BYTE)	
83	2 级下行信号质量次数	SumRxquaDL 2	Int (4BYTE)	
84	3 级下行信号质量次数	SumRxquaDL 3	Int (4BYTE)	
85	4 级下行信号质量次数	SumRxquaDL 4	Int (4BYTE)	
86	5 级下行信号质量次数	SumRxquaDL 5	Int (4BYTE)	
87	6 级下行信号质量次数	SumRxquaDL 6	Int (4BYTE)	
88	7 级下行信号质量次数	SumRxquaDL 7	Int (4BYTE)	
89	TA 为 0 或 1 的次数	TA01Count	Int (4BYTE)	
90	Rxlevel Down 大于-85 次数	Rxlevel_Down	Int (4BYTE)	Rxlevel Down 大于-85 次数
91	弱覆盖 MR 个数	Weak_coverage	Int (4BYTE)	DL_RXLEVEL<-90 and TA<2 次数
92	过覆盖 MR 个数	Over_coverage	Int (4BYTE)	DL_RXLEVEL>-80 and TA>1 次数
93	上下行不平衡 MR 个数	Imbalance	Int (4BYTE)	abs (DL_RXLEVEL-UL_RXLEVEL)>=15 次数
94	干扰 MR 个数	Interference	Int (4BYTE)	DL_RXLEVEL>-80 and DL_RXQUAL>4 次数
95	立即指配命令时刻	Imm_Ass_Co md_Time	numeric(10)	
96	立即指配拒绝时刻	Imm_Ass_Re j_Time	numeric(10)	
97	立即指配建立成功时刻	Imm_Ass_Co mp_Time	numeric(10)	
98	指配命令时间	Ass_Comd_T ime	numeric(10)	

99	掉话 Connection_Failure_Time	Connection_Failure_Time	numeric(10)	
100	掉话时间	CLEAR_REQ_TIME	numeric(10)	
101	SCCPID	SCCP 连接标识	Varchar(20)	同一个 SCCP 连接上所有 A 接口单据填写相同的 SCCPID，华为探针中填写同一个 SCCP 连接中首个 CDR 的 CDR 标识
102	CSFBIND	CSFB 指示	numeric(2,0)	1: CS fallback mobile terminating call 2: CS fallback mobile originating call 其余情况为 0

5.1.3. PS 域用户业务表

序号	英文描述	中文描述	类型	备注
1	CREATETIME	数据创建时间	int	相对 1970 年的秒数。 (时间为 UTC 格式)
2	first_time	本片开始时间	dateTime	本片开始时间
3	end_time	本片结束时间	dateTime	本片结束时间
4	calling	用户号码	string	用户号码
5	imsi	国际移动用户识别码	string	国际移动用户识别码
6	user_ip	用户 IP 地址	unsignedInt	用户 IP 地址
7	imei	国际移动身份识别	string	国际移动身份识别
8	mcc	国家代号	int	国家代号
9	mnc	网络代号	int	网络代号
10	lac	位置区号	int	位置区号
11	rac	路由区标识	int	路由区标识
12	cell_id	小区识别码	int	小区识别码
13	sgsn_sg_ip	SGSN 信令传输 IP	int	SGSN 信令传输 IP
14	ggsn_sg_ip	GGSN 信令传输 IP	int	GGSN 信令传输 IP
15	sgsn_data_ip	SGSN 数据传输 IP	int	SGSN 数据传输 IP
16	ggsn_data_ip	GGSN 数据传输 IP	int	GGSN 数据传输 IP
17	apn	访问点名	string	访问点名
18	rat	2G/3G 网络标识	int	取值： 1 3G 2 2G

19	service_type	子业务	int	例如, QQ, 迅雷等
20	service_group	业务类型	int	例如, IM, P2P 等
21	up_packets	上行总数据包数	int	单位: 个
22	down_packets	下行总数据包数	int	单位: 个
23	up_bytes	上行流量	int	上行总流量 单位: Byte
24	down_bytes	下行流量	int	下行总流量 单位: Byte
25	up_speed	上行速率	float	上行速率, 单位: kbps
26	down_speed	下行速率	float	下行速率, 单位: kbps
27	trans_time	数据传输的时长	int	单位为 0.001ms
28	is_end	是否最后一片	int	是否最后一片
29	user_port	用户端口	int	用户端口
30	proto_type	协议: TCP 6 UDP 17	int	协议: TCP 6 UDP 17
31	dest_ip	目标 IP	int	目标 IP
32	dest_port	目标端口	int	目标端口
33	CDR_ID	CDR ID	LongInt	信令共享平台针对应用层订阅所生成的 CDR 标识
34	service_type_sdk	sdk 子业务	int	例如, QQ, 迅雷等
35	service_group_sdk	sdk 业务类型	int	例如, IM, P2P 等
36	is_http	HTTP 标识	int	0 非 HTTP; 1 HTTP

5.1.4. PS 域 WAP/HTTP 事件话单

序号	英文描述	中文描述	类型	备注
1	CREATETIME	数据创建时间	int	相对 1970 年的秒数。 (时间为 UTC 格式)
2	start_time	开始时间	dateTime	采集第一条消息的时间
3	cdr_type	CDR 类型	unsignedByte	21: GET(Gn 接口) 22: POST(Gn 接口) 6007: GET (GN) 6008: POST (GN)
4	apn	APN	string	访问点
5	lac	位置区号	int	位置区号
6	rac	路由区标识	int	路由区标识
7	cell_id	小区识别码	int	小区识别码

8	rat	2G/3G 网络标识	int	2G/3G 网络标识	
9	imei	终端标识号	string	终端标识号	
10	source_ip	源 IP 地址	unsignedInt	SGSN_IP	
11	dest_ip	目的 IP 地址	unsignedInt	GGSN_IP	
12	user_ip	用户 IP 地址	unsignedInt	用户 IP 地址	
13	wapgw_ip	WAP 网关 IP 地址	unsignedInt	WAP 网关 IP 地址	
14	host_ip	Host IP 地址	unsignedInt	Host IP 地址	
15	imsi	国际移动用户识别码	string	国际移动用户识别码	
16	calling	用户号码	string	用户号码	
17	url	URL	string	URL, 512Byte	
18	host	HOST	string	HOST	
19	x_online_host	x_online_host	string	x_online_host	
20	app_type	应用类型	string	应用类型	
21	result	过程结果	int	响应码	
22	user_agent	终端浏览器类型	string	终端类型, 256Byte	
23	req_num	请求次数	unsignedByte	请求次数 同一次请求重发了 n-1 次	
24	content_packets	内容分片包数	unsignedShort	wtp 协议时才有效	
25	resend_packets	重传包数	unsignedShort	wtp 协议时才有效	
26	lost_packets	丢失包数	unsignedShort	wtp 协议时才有效	
27	up_bytes	上行消息字节数	int	上行消息字节数 单位: Byte	
28	down_bytes	下行消息字节数	int	下行消息字节数 单位: Byte	
29	version	WAP 版本号	unsignedByte	表示协议版本号, 取值包括	
				码值	中文解释
				0	HTTP
				1	WAP1.x
				2	HTTP1.0
				3	HTTP1.1
30	gtp_ver	GTP 版本	unsignedByte		
31	teid_sgsn_data	SGSN 数据面 TEID	unsignedInt		
32	teid_ggsn_data	GGSN 数据面 TEID	unsignedInt		
33	res_delay	响应时延	int	单位为 0.01ms;	

34	duration	CDR 持续时间	int	单位为 0.001ms;
35	CDR_ID	CDR ID	LongInt	信令共享平台针对应用层订阅所生成的 CDR 标识
36	up_packets	上行总数据包数	int	上行数据包数 单位：个
37	down_packets	下行总数据包数	int	下行数据包数 单位：个
38	service_type_sdk	sdk 子业务	int	例如，QQ，迅雷等
39	service_group_sdk	sdk 业务类型	int	例如，IM，P2P 等
40	Keyword	内容关键字	char	内容关键字
41	AppID_sdk	APP_ID	Int32	APP 代号
42	KeyWordTypeID_sdk	关键字类型	Int32	

5.2. 维度表模型

本次测试中用到的维度表如下：

维度表	中文名称	字段数目	数据规模（条）
DIM_CELL	小区维表	6	327,007
DIM_SERVICE_SDK	业务类型维表	4	2,898
DIM_IMEI	终端类型维表	3	451,460

5.2.1. 小区维表

序号	英文描述	中文描述	字段类型	备注
1	lac	位置区号	int	位置区号，与数据表 PS_USERSERVICE 和 PS_WAPHTTP 的字段 lac 关联
2	cell_id	小区识别码	int	小区识别码，与数据表 PS_USERSERVICE 和 PS_WAPHTTP 的字段 cell_id 关联
3	province_name	省	string	例如：广东省
4	city_name	市	string	例如：广州市
5	town_name	区	string	例如：天河区
6	sac_name	小区名	string	例如：广州地铁 APM 线妇儿中心

5.2.2. 业务维表

序号	英文描述	中文描述	字段类型	备注
1	service_group_sdk	业务类型	int	与数据表 PS_USERSERVICE 和 PS_WAPHTTP 的字段 service_group_sdk 关联
2	service_type_sdk	子业务	int	与数据表 PS_USERSERVICE 和 PS_WAPHTTP 的字段 service_type_sdk 关联
3	service_group_sdk_name	业务类型名称	string	例如：即时通信
4	service_type_sdk_name	子业务名称	string	例如：中国移动飞聊

5.2.3. 终端类型维表

序号	英文描述	中文描述	字段类型	备注
1	imei	imei 前 8 位	int	imei 前 8 位
2	firm	设备厂商	string	例如：苹果
3	type	设备型号	string	例如：PHONE 6

5.3. 接口数据格式及大小

按照尽量贴近实际生产环境的原则设计测试环境和测试数据。

说明要求如下：

- 接口文件以“,”进行分割，用回车符为行分割，不带表头。
- 文件编码格式统一为 UTF-8。
- 单个文件大小不超过 128MB。
- 各数据表的文件名说明如下：

表名	文件名	例子
AIU_MO C	aiu-moc-cdr- <yyyyMMddHHmm_id_yy yyMMdd#_yyyyMMddHHmmss#_ip>.dat	aiu-moc-cdr-201502142350-00001_201502 14#20150214235701#_192.168.35.199.dat
AIU_MT C	aiu-mtc-cdr- <yyyyMMddHHmm_id_yy yyMMdd#_yyyyMMddHHmmss#_ip>.dat	aiu-mtc-cdr-201502142350-00001_201502 14#20150214235701#_192.168.35.199.dat

PS_USE RSE CE	user_service_CDR_<yyyyMMddHHmm_ id_yyyyMMdd#_yyyyMMddHHmmss#_ip >.dat	user_service_CDR_201502141530_00073_2 0150214#20150214165005#_192.168.35.19 6.dat
PS_WAP HTTP	wap_http_CDR_<yyyyMMddHHmm_id_y yyyMMdd#_yyyyMMddHHmmss#_ip>.dat	wap_http_CDR_201502141540_00002_20150 214#20150214165704#_192.168.35.196.dat

6. 测试内容

网管中心Hadoop平台测试内容分为平台规划能力测试、运维和优化能力测试、以及模型管控能力测试，满足网管中心对现有Hadoop平台运营的各类应用场景的需求。

平台规划能力测试将以网管中心实际的业务需求作为输入，考察参测厂商对于Hadoop平台的规划能力，以应对未来网管中心平台扩容以及共享新架构的规划工作。

运维和优化能力测试将参考目前Hadoop平台的日常运维工作场景，考察参测厂商对于Hadoop平台的日常管理能力、平台故障处理能力、以及平台性能相关优化能力。

模型管控能力测试将参考目前Hadoop平台的核心业务场景，考察参测厂商对于业务的理解能力以及应用开发能力，核心业务场景包括离线批量复杂分析、即席查询、简单键值查询以及准实时流计算。

根据实际场景下的不同业务需求，对各测试用例赋予不同分值。各测试用例中分值均按照100分计算。测试用例总表及相应分值占比如表6-1所示。

表 6-1 测试用例分值占比

测试类型	测试内容	
平台规划能力测试	业务场景理解	
	物理节点及网络规划	
	平台架构设计	
	数据处理流程设计	
	平台运维支持方案	
	平台风险点预估和应对机制	
	平台安全性管控	
	平台建设成本和时间点规划	
运维和调优能力测试	平台管理能力测试	部署能力测试
		配置管理测试

			任务调度测试
			安全管理测试
		故障处理能力测试	监控管理测试
			告警管理测试
			故障恢复测试
	调优能力测试		加载性能调优
			查询性能调优
			数据处理性能调优
模型管控能力测试	离线批量复杂分析		用户业务与流量分析
			网络故障分析
	即席查询		CDR 智能查询
	简单键值查询		用户上网详单查询
	准实时流计算		重点区域保障计算

6.1. 总体约束

参照Hadoop平台产品必须支持Linux操作系统，在初始化及相关软件部署过程中，应具备批量安装部署工具，支持各节点的集中化统一部署。

要求所有Hadoop集群节点（10节点）共用一个HDFS文件系统。

要求Hadoop平台基于Apache Hadoop 2.0以上版本开发，且能够持续跟进开源版本功能点进行迭代更新，保持向下兼容。

在测试报告中记录性能测试场景中的环境，包括如下：

管理节点数量、副本数量、基于开源 Hadoop的版本、总可用数据空间（用户可使用部分）等。

搜集上述信息，并在测试之前需要填写在表6-2中。

表 6-2 测试信息收集表

硬件搭建方式	管理节点数	是否做 RAID
基于开源Hadoop版本		
数据装载前占用空间容量(GB)		
数据装载前可用空间容量(GB)		
对Hadoop源码做了哪些方面的改进	1、... 2、...	

需要根据实际情况补充记录表格。测试进程中，对测试脚本执行前后文件空间的变化进行对比监控；测试过程中如出现系统意外中断、重启等现象，需进行记录。

场景测试以及调优能力测试全部使用RESTful接口方式进行，返回数据的目录以及表名严格遵守该文档中的定义，我们提供以下类型接口：

1. StartAPI:启动厂商对应的测试场景的执行程序
2. StatusAPI:检查程序是否完成
3. SelectAPI:检查结果正确性，返回数据格式如下：

`{"size":,"data":["一条数据，用逗号隔开",""]}`

时间格式约束：除了流式处理精确到分钟外，其它用例全部精确到天，具体格式见接口描述。

时间区间:StartAPI参数以及Select参数采用闭区间[starttime,endtime],流式处理统计口径采用前开后闭区间(0,5](5,15]

脏数据定义：字段数与数据表约定的不符，为脏数据，可以丢弃；在一个测试用例中，如果一条数据中涉及分析的相关字段为空或非法值（如：空格），则认为该条数据是脏数据，可以丢弃。

重复数据定义：当两条数据完全一致时，认为是重复数据，保留一条即可。

计算数值的精度：浮点数保留小数点后两位。

数据集的准备和使用约束：

不同测试用例数据将会复用，一共准备两份测试数据，用于不同测试用例数据切换。

1. 数据分发

- ✓ 厂商数据接口服务器准备一个10T的分区：file:///data
- ✓ 将所需要的测试数据及维表数据发送到厂商数据接口服务器file:///data下，数据分为三个目录：
 - * file:///data/test1: 测试数据1
 - * file:///data/test2: 测试数据2
 - * file:///data/COMMONS: 维表数据

其中test1和tes2目录结构如下：<Table_Name>/<City_NUM>/<Day>/，commons目录结构如下：<Table_Name>/

2. 数据上传

- ✓ 测试数据上传到厂商HDFS上；
- ✓ 厂商准备脚本upload-testdata.sh，将file:///data下全部文件上传到HDFS的hdfs:///data目录下（上传时要注意数据文件名的#号）。

3. 数据切换

- ✓ 厂商准备脚本，用于重命名HDFS的测试数据目录。
- ✓ 所有依赖于HDFS目录的测试用例，有统一的数据输入目录：
hdfs:///data/HDPTEST，执行具体测试用例时会将该所需的测试数据集放到该目录下，为了避免数据拷贝的时间开销，统一采用数据移动方法来进行。
 - * rename1.sh: 将hdfs:///data/test1重命令为hdfs:///data/HDPTEST;
 - * rename2.sh: 将hdfs:///data/HDPTEST重命令为hdfs:///data/test1;
 - * rename3.sh: 将hdfs:///data/test2重命令为hdfs:///data/HDPTEST;
 - * rename4.sh: 将hdfs:///data/HDPTEST重命令为hdfs:///data/test2;
- ✓ 运行测试用例前，根据实际数据集需要，调用rename1.sh或rename3.sh，将备用测试数据目录切换成HDPTEST。
- ✓ 测试用例结束后，调用rename2.sh或rename4.sh，将HDPTEST切换回备用测试数据目录。

4. 装载数据库

- ✓ 厂商准备脚本和程序，用于将数据装载到HBase中（仅适用于简单键值查询测试用例）。

测试结果存档：

为了保证测试的公平和可回放，统一对测试结果进行存档，具体包括：

1. 测试检查点的截图存档；
2. 测试脚本统计的时间指标和结果正确性的存档；
3. 测试用例输出结果表的存档，厂商需要配合准备相应的脚本或程序将结果表导出成文件。

6.2. 平台规划能力测试

平台规划能力测试是为了应对网管中心未来平台建设、扩容的需求，考察厂商对于Hadoop平台整体的规划能力。平台规划的业务场景的如下：

数据类型和规模：网络管理中心负责的移动信令平台的全量数据，核心数据包括用户业务 / 语音 / 短信 / 上网的流量详单数据，每天数据量为100T。

业务需求：平台要求能支持信令各层的数据分析应用，包括解码层的网络性能和质量监控、共享层的用户活跃度分析、流失预警和新增用户挖掘，以及应用层的各种业务分析；平台的核心业务场景能覆盖主要的四大类：离线批量复杂分析、即席查询、简单键值查询以及准实时流计算。

平台性能指标：平台要求能支持如下性能指标：

- 对于按照小时为单位的数据分析需求，最多不能超过 2 小时完成
- 对于按照天为单位的数据分析需求，最多不能超过 7 小时完成
- 对于实时数据分析和监控需求，延时不能超过 15 分钟完成

参测厂商需要根据上述数据类型和规模、业务需求以及需要达到的性能指标，给出一份完整的平台架构规划方案，其中核心需要包括如下部分：

- 1、 业务场景理解：深入分析平台在各个数据层可能面对的业务场景和核心业务需求，作为整个平台规划的业务输入。此处主要目的是考察厂商对于移动业务的熟悉程度以及业务场景的理解能力。
- 2、 物理节点及网络的规划：根据对应的业务需求规划所需部署的物理节点和网络拓扑结构。此处主要目的是考察厂商的物理节点及网络规划的合理性。
- 3、 平台架构设计：根据对应的业务需求规划整个Hadoop平台架构，其中包括的组件以及各自的功能，以及数据存储方案。此处主要目的是考察厂商的平台架构设计的合理性以及对主要业务场景的覆盖能力。
- 4、 数据处理流程的设计：根据对应的业务需求规划整个核心数据处理流程，其中包括每个核心业务场景的数据处理环节以及对应的输入输出。此处主要目的是考察厂商是否能根据业务场景做合理的数据处理流程规划，以及整体在数据处理流程上设计的合理性。
- 5、 平台运维支持方案：提供平台需要进行运维的主要事项以及对应的支撑方案。此处主要目的是考察厂商在运维支持方面的经验以及方案的合理和完备性。

6、平台风险点的预估和应对机制：分析平台可能存在的风险点以及对应的处理机制。

此处主要目的是考察厂商是否能够对未来的风险性进行预估和防范，同时能否具备合理的处理和应对机制。

7、平台安全性管控机制：提供平台在数据、系统和服务方面的安全管控机制。

8、平台建设的成本和时间点规划合理性：提供对整个平台建设成本以及建设周期的详细规划。此处主要考察厂商在平台建设成本方面的控制能力以及工期进度方面的把握能力。

该项测试要求的输出结果为文档形式，最终的评分标准请参考表 6-1。

6.3. 运维优化能力测试

运维优化测试是为了考察厂商对Hadoop平台的运维和优化能力，包括三个部分：平台管理能力、故障处理能力测试以及调优能力测试。

6.3.1. 平台管理能力测试

平台管理能力测试着眼于基础能力的考察，在Hadoop集群不出意外情况的前提下，厂商必须具备的能力，包括部署能力、配置管理能力、任务调度能力以及安全管理能力。

6.3.1.1. 部署能力测试

项目：	平台管理能力测试	分项目：	部署能力测试
用例编号：	HadoopTEST_FUNC_MGM	版本：	
用例分数：	100分	适用场景：	通用
测试目的：	验证对Hadoop平台是否具备自动安装功能和支持扩容		
预置条件：	1. 按照测试组网图搭建测试环境； 2. 所有节点还没安装 Hadoop 平台，已经安装的需要删除； 3. 所有节点安装了自动化部署工具； 4. 厂商提供部署的计划，用于结果检查。		
测试步骤：	1. 选择 8 个节点准备进行集群安装； 2. 自动化的以向导的方式安装集群； 3. 安装完成后，查看集群系统状态； 4. 再对集群进行扩容 2 个节点。		

预期结果:	1. 支持所有组件的自动化部署; 2. 安装部署操作以向导方式进行; 3. 支持多个节点的自动化安装部署任务; 4. 安装完成后查询系统所有组件(至少包括 HDFS, YARN, ZooKeeper, HBase, Spark) 的状态正常, 整个集群必须支持 HDFS 读写操作、HBase 的读写、MapReduce 任务的运行和 Spark 任务的运行; 5. 支持集群系统的扩容多个节点, 扩容的节点必须包含 HDFS 组件、YARN 组件、HBase 组件、MapReduce 组件和 Spark 组件。		
测试结果:	1) 是否支持产品所有组件的自动化部署: 是 (), 否 (), 截图 3-1-1-1; 2) 是否提供向导方式安装集群: 是 (), 否 (), 截图 3-1-1-2; 3) 是否支持多个节点的自动化部署: 是 (), 否 (), 截图 3-1-1-3; 4) 查询集群所有组件的状态是否正常: 是 (), 否 (), 截图截图 3-1-1-4-[对应的检查点]-[id]; 5) 是否支持扩容多个节点: 是 (), 否 (), 截图 3-1-1-5 -[对应的监控页面]; 6) 从开始集群安装到部署成功的时间。 <div style="margin-left: 40px;">安装 8 个节点</div> <div style="margin-left: 40px;">开始:</div> <div style="margin-left: 40px;">结束:</div> <div style="margin-left: 40px;">耗时:</div> <div style="margin-left: 40px;">扩容 2 个节点</div> <div style="margin-left: 40px;">开始:</div> <div style="margin-left: 40px;">结束:</div> <div style="margin-left: 40px;">耗时:</div>		
评分标准:	1、未能在规定时间 5 小时内成功部署, 该用例不得分; 2、支持产品所有组件的自动化部署, 得该用例总分的 20%; 3、能提供向导方式安装集群, 得该用例总分的 20%; 4、支持多个节点的自动化部署, 得该用例总分的 20%; 5、系统全部组件状态正常, 得该用例总分的 20%; 6、支持扩容多个节点且在 2 小时内完成, 得该用例总分的 20%。		
备注:	该测试用例完成后, 集群不能再增加新的组件或节点, 只能修改集群的配置。		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

6.3.1.2. 配置管理测试

项目：	平台管理能力测试	分项目：	配置管理测试
用例编号：	HadoopTEST_FUNC_CONF	版本：	
用例分数：	100分	适用场景：	通用
测试目的：	测试Hadoop系统的配置功能，包括对HDFS、MapReduce、HBase和Zookeeper的配置、查看和修改功能。		
预置条件：	厂商完成了 Hadoop 平台的部署。		
测试步骤：	1、登陆 Hadoop 管理系统控制台，任意选择查看 YARN，HDFS、MapReduce、HBase 的各 1 项配置信息；登陆到 Web 管理页面对比验证 4 项配置信息是否与控制台一致。 2、通过 Web 管理页面分别修改 YARN、HDFS、MapReduce、HBase 的各 1 项配置信息。随机抽选 3 个节点，通过 Hadoop 管理系统控制台，验证上述 4 项配置信息是否自动同步到所有节点。 3、通过 Web 管理页面修改其中任意 3 个节点的配置信息,通过 Hadoop 管理系统控制台，验证这 3 个节点的配置项修改是否生效。		
预期结果：	1、通过 Web 管理页面和控制台能查看系统的配置信息，并且 Web 管理页面和控制台查看到的信息一致； 2、可以通过 Web 管理页面修改系统配置并自动同步到所有节点； 3、可以实现通过 Web 管理页面修改部分节点的配置信息。		
测试结果：	1、登陆 Hadoop 管理系统控制台，任意选择查看 HDFS、MapReduce、HBase 和 Zookeeper 的各 1 项配置信息；登陆到 Web 管理页面对比验证 4 项配置信息是否与控制台一致： YARN：是（ ），否（ ），截图 3-1-2-1-[对应的配置信息]-[id:1/2] HDFS：是（ ），否（ ），同上 MapReduce：是（ ），否（ ），同上 HBase：是（ ），否（ ），同上 2、通过 Web 管理页面分别修改 HDFS、MapReduce、HBase 和 YARN 的各 1 项配置信息后，随机抽选 3 个节点，通过 Hadoop 管理系统控制台，验证上述 4 项配置信息是否自动同步到所有节点： YARN：是（ ），否（ ），截图 3-1-2-2-[对应的配置信息]-[id:1/2] HDFS：是（ ），否（ ），同上 MapReduce：是（ ），否（ ），同上 HBase：是（ ），否（ ），同上 3、通过 Web 管理页面修改其中任意 3 个节点的配置信息,通过 Hadoop 管理系统控制台，验证这 3 个节点的配置项修改是否生效： 节点 1：是（ ），否（ ），3-1-2-3-[r/节点] 节点 2：是（ ），否（ ），同上 节点 3：是（ ），否（ ），同上		

评分标准：	<p>1、 通过Hadoop管理系统控制台和Web管理页面查看配置信息，且查看到的信息一致，该项分数占该项用例的40%。每项不一致扣10%；</p> <p>2、 通过Web管理页面修改YARN、HDFS、MapReduce和HBase 4项配置信息，并且控制台验证4项配置信息成功同步至全部集群，该项分数占该项用例的30%。每一项验证不一致扣7.5%；</p> <p>3、通过Web管理修改指定的3个节点的配置信息，且通过控制台验证3个节点配置生效，该项分数占该项用例的30%。每一台修改失败扣10%。</p>		
备注：	在部署能力测试之后测试该项。		
测试执行人员：			
测试人员签字：		测试时间：	
测试审核员签字：		厂家配合人员签字：	

6.3.1.3. 任务调度测试

项目：	平台管理能力测试	分项目：	任务调度
用例编号：	HadoopTEST_FUNC_SCHEDULE	版本：	
用例得分：	100分	适用场景：	通用
测试目的：	测试多用户作业时，平台是否可以按照配置的用户资源享有量来进行任务调度。		
预置条件：	<p>1. 厂商完成了 Hadoop 平台的部署；</p> <p>2. 创建两个队列，队列 1 配置享有集群 80%的资源，队列 2 配置享有集群 20%的资源；</p> <p>3. 创建一个用户，并将其与上述两个队列对应；</p> <p>4. 准备作业 A（6.3.3.2 JOIN 操作用例）。</p>		
测试步骤：	<p>1. 使用创建的用户分别向队列 1 和 2 中提交作业 A；</p> <p>2. 查看两个任务分配到的资源数量的比例，应该为 4：1。</p>		
预期结果：	平台支持按资源配置进行任务调度。		

测试结果:	是否支持按资源配置进行任务调度: 是 (), 否 (), 截图3-1-3		
评分标准:	支持按资源配置进行任务调度得满分, 否则得零分。		
备注:			
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

6.3.1.4. 安全管理测试

项目:	平台管理能力测试	分项目:	安全管理测试
用例编号:	HadoopTEST_FUNC_SAFE	版本:	
用例分数:	100分	适用场景:	通用
测试目的:	验证Hadoop的系统对用户访问权限控制功能		
预置条件:	1. 厂商完成了 Hadoop 平台的部署; 2. 创建两个用户组 GROUP1 和 GROUP2, 创建两个普通用户 USER1 和 USER2 分别属于两个普通用户组。		
测试步骤:	1. 以用户 USER1 上传测试文件 text1, 并将文件权限改为 640; 2. 使用用户 USER2 读取测试文件 text1, 查看是否可以下载操作; 3. 以用户 USER1 提交测试作业 job1; 4. 使用用户 USER2 尝试 kill 掉 job1。		
预期结果:	1. 用户 USER2 访问 text1 文件失败; 2. USER1 的作业不能被 USER2 杀掉 (kill)。		
测试结果:	1、 USER2 能否访问 USER1 创建的 text1 文件: 能 (), 不能 (), 截图 3-1-4-1 2、 USER2 能否 kill 掉 USER1 提交的 job1: 能 (), 不能 (), 截图 3-1-4-2		
评分标准:	1、 USER2无法访问USER1的text1文件, 则可得该项用例50%分数; 2、 USER2 无法killUSER1的job1, 则可得该项用例的50%分数。		

备注：			
测试执行人员：			
测试人员签字：		测试时间：	
测试审核员签字：		厂家配合人员签字：	

6.3.2. 故障处理能力测试

故障处理能力测试着眼于应急能力的考察，当集群或者业务系统出现异常的情况下，厂商是否具备提前预知并且预备相应的解决方案，包括监控管理方案、异常告警方案以及故障恢复能力等。

6.3.2.1. 监控管理测试

项目：	故障处理能力测试	分项目：	监控管理测试
用例编号：	HadoopTEST_FUNC_MON	版本：	
用例分数：	100分	适用场景：	通用
测试目的：	验证 Hadoop 管理系统对 Hadoop 系统性能进行图形化监控管理的功能，验证的指标： 1、主机节点：主机名称、空闲 CPU 百分比、用户空间占用 CPU 百分比、内核空间占用 CPU 百分比、缓存内存大小、空闲内存大小、共享内存大小、交换分区总量、磁盘总大小、剩余磁盘空间、进程总数、每分钟的系统平均负载、每 5 分钟的系统平均负载、每 15 分钟的系统平均负载、网络入口带宽速度、网络出口带宽速度； 2、HDFS：HDFS 文件系统块总数、总大小、文件总数、剩余量、损坏块、需复制块； 3、Mapreduce：任务运行情况，任务占资源情况； 4、HBASE：表结构信息，RegionServer 的请求次数以及 RegionServer Regions 数目； 5、YARN：Current Applications、Finished Applications、NodeManagers		
预置条件：	1、厂商完成了 Hadoop 平台的部署； 2、执行 6.3.3.2 JOIN 操作用例，产生系统性能历史数据信息。		

测试步骤:	1、 登录性能监测页面或打开图形化监控客户端； 2、 实时监测以图表方式显示的 30 分钟内的系统性能变化情况； 3、 查看服务性能历史数据；
预期结果:	操作成功，系统未有任何错误提示。 所有监控指标均能正确显示。
测试结果:	1、记录主机节点是否显示以下信息：截图 3-2-1-1-[id] [1] 主机名称：是（ ） 否（ ） [2] 空闲 CPU 百分比：是（ ） 否（ ） [3] 用户空间占用 CPU 百分比：是（ ） 否（ ） [4] 内核空间占用 CPU 百分比：是（ ） 否（ ） [5] 缓存内存大小：是（ ） 否（ ） [6] 空闲内存大小：是（ ） 否（ ） [7] 共享内存大小：是（ ） 否（ ） [8] 交换分区使用大小：是（ ） 否（ ） [9] 磁盘总大小：是（ ） 否（ ） [10] 总剩余磁盘空间：是（ ） 否（ ） [11] 进程总数：是（ ） 否（ ） [12] 每分钟的系统平均负载：是（ ） 否（ ） [13] 每 5 分钟的系统平均负载：是（ ） 否（ ） [14] 每 15 分钟的系统平均负载：是（ ） 否（ ） [15] 网络入口带宽速度：是（ ） 否（ ） [16] 网络出口带宽速度：是（ ） 否（ ） 2、HDFS 需监控信息，截图 3-2-1-2-[id] [1] HDFS 文件系统块总数：是（ ） 否（ ） [2] 总大小：是（ ） 否（ ） [3] 文件总数：是（ ） 否（ ） [4] 剩余量：是（ ） 否（ ） [5] 损坏块：是（ ） 否（ ） [6] 需复制块：是（ ） 否（ ） 3、MapReduce 需监控信息，截图 3-2-1-3-[id] [1] MR 任务运行情况：是（ ） 否（ ） [2] MR 任务占资源情况：是（ ） 否（ ） 4、HBASE 需监控信息，截图 3-2-1-4-[id] [1] 表结构信息：是（ ） 否（ ） [2] RegionServer 的请求次数：是（ ） 否（ ） [3] RegionServer Regions 数：是（ ） 否（ ） 5、YARN 需监控信息，截图 3-2-1-5-[id] [1] Current Applications：是（ ） 否（ ） [2] Finished Applications：是（ ） 否（ ） [3] NodeManagers：是（ ） 否（ ）
评分标准:	根据 Hadoop 系统状态监控能力的完备性进行评分。每缺失一项扣 4%，扣完为止。

备注：			
测试执行人员：			
测试人员签字：		测试时间：	
测试审核员签字：		厂家配合人员签字：	

6.3.2.2. 告警管理测试

项目：	故障处理能力测试	分项目：	告警管理测试															
用例编号：	HadoopTEST_FUNC_ALA	版本：																
用例分数：	100分	适用场景：	通用															
测试目的：	验证Hadoop管理系统系统对Hadoop系统故障异常信息进行管理的功能, 需验证告警指标： [1] 主机状态 [2] 存储块状态 [3] NameNode 进程状态 [4] DataNode 进程状态 [5] ZKSERVER 进程状态 [6] HMASTER 进程状态 [7] REGIONSERVER 进程状态 [8] HDFS 服务状态 [9] YARN 服务状态																	
预置条件：	1、厂商完成了 Hadoop 平台的部署； 2、Hadoop 平台运行正常； 3、登录故障管理页面（厂商提前提供）。																	
测试步骤：	1、按照测试步骤进行操作，产生故障信息； 2、查看是否能对故障进行分级；采用以下测试方法, 看是否能够查看故障以及分级状态，具体测试方法如下： <table><tr><th>编号</th><th>指标名称</th><th>测试方法</th></tr><tr><td>1</td><td>主机状态</td><td>后台拔掉一台主机的网线</td></tr><tr><td>2</td><td>存储块状态</td><td>后台卸载一块数据盘</td></tr><tr><td>3</td><td>NameNode 进程状态</td><td>后台停止 NameNode 进程</td></tr><tr><td>4</td><td>DataNode 进程状态</td><td>后台停止 DataNode 进程</td></tr></table>			编号	指标名称	测试方法	1	主机状态	后台拔掉一台主机的网线	2	存储块状态	后台卸载一块数据盘	3	NameNode 进程状态	后台停止 NameNode 进程	4	DataNode 进程状态	后台停止 DataNode 进程
编号	指标名称	测试方法																
1	主机状态	后台拔掉一台主机的网线																
2	存储块状态	后台卸载一块数据盘																
3	NameNode 进程状态	后台停止 NameNode 进程																
4	DataNode 进程状态	后台停止 DataNode 进程																

	5	ZKSERVER 进程状态	后台停止 ZKSERVER 进程
	6	HMASTER 进程状态	后台停止 HMASTER 进程
	7	REGIONSERVER 进程状态	后台停止 REGIONSERVER 进程
	8	HDFS 服务状态	后台停止 NameNode 或 DataNode 进程
	9	YARN 服务状态	后台停止 MR/YARN 服务
3、查看在指定时间（60 秒）内是否返回告警信息（在界面高亮显示）； 4、查看/清除系统在指定时间范围内故障信息。			
预期结果:	1、针对每一个故障，均可以产生告警； 2、告警时间不超过 60 秒		
测试结果:	具备以下告警指标监控：截图 3-2-2-[id] 1) 主机状态：是（ ），否（ ）； 2) 存储块状态：是（ ），否（ ）； 3) NameNode 进程状态：是（ ），否（ ）； 4) DataNode 进程状态：是（ ），否（ ）； 5) ZKSERVER 进程状态：是（ ），否（ ）； 6) HMASTER 进程状态：是（ ），否（ ）； 7) REGIONSERVER 进程状态：是（ ），否（ ）； 8) HDFS 服务状态：是（ ），否（ ）； 9) YARN 服务状态：是（ ），否（ ）；		
评分标准:	具备告警功能，告警中各类信息完备，且告警时间不超过 60 秒，该用例得满分；否则，每缺失或者超时一项，扣 12%，扣完为止。		
备注:			
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

6.3.2.3. 故障恢复测试

项目:	故障处理能力测试	分项目:	故障恢复测试
用例编号:	HDPTEST_FUNC_RECOVERY	版本:	

用例得分:	100分	适用场景:	通用
测试目的:	测试在系统服务、进程、节点出现故障的情况下，系统的健壮性以及 对故障排查和恢复的能力。		
预置条件:	厂商完成了 Hadoop 平台的部署；		
测试步骤:	1、 往各类计算集群提交系统任务； 2、 在系统有任务（YARN/HBase/HDFS/Spark）在运行的情况下随机选择停止一项服务、进程或断开某个机器网络，引发故障。故障场景按难度分为低级 / 中级 / 高级三大类，该测试将从三个级别的故障场景中各随机抽取一个。 3、 记录任务是否中断，若未中断则记录任务运行时间；若中断则记录修复手段以及恢复所消耗的时间，然后重新提交任务并记录任务是否可以正常运行和到执行完毕所用的时间。 4、 任务执行完成，核对执行结果。		
预期结果:	各类计算集群可以正常执行完任务，并输出正确的结果。		
测试得分:			
评分标准:	1、 低级故障得分占 20%，中级故障得分占 30%，高级故障得分占 50%； 2、 每个故障级别条件下，任务可不中断并正确执行完成，得该级别的满分；在规定时间内经修复后重新提交可正确执行完成，得该级别的 50%分数(高级故障依然得满分)；若在规定时间内无法修复或执行结果错误，则不得分。		
备注:	1、 每个组件需要运行的任务： <ul style="list-style-type: none"> ● YARN：用6.3.3.3数据处理性能调优测试用例程序。 ● HBase：用6.3.3.1加载性能调优测试用例程序。 ● Spark：将6.3.3.3测试用例改成用Spark运行。 		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

6.3.3. 调优能力测试

调优能力测试着眼于满足基本需求之后，在性能上是否具备精益求精能力的考察。调优是一项高级能力的考察，一方面可以提升业务系统响应速度，另一方面可以为后期添加更加

丰富的功能提供可扩展的空间。这方面的测试包括数据加载性能调优、数据查询性能调优以及数据处理性能调优等。

6.3.3.1. 加载性能调优测试

项目：	调优能力测试	分项目：	加载性能调优
用例编号：	HDPTEST_FUNC_LOAD_HBASE	版本：	
用例得分：	100分	适用场景：	通用
测试目的：	测试以最优方式将本地文件系统上的文件加载到 HBase（明确到 HBase）的速度。		
预置条件：	1、厂商完成了 Hadoop 平台的部署； 2、需要加载的数据（PS_USERSERVICE 表）在厂商装有 HBA 卡的机器的本地文件系统。		
测试步骤：	1、通过 RESTful 接口启动厂商的程序。 2、任务执行完成后，通过查询接口核对执行结果。截图 3-3-1		
预期结果：	可以在规定时间内执行完调优任务，并输出正确的结果。		
评分标准：	测试执行过程出错或执行结果不正确，该测试用例不得分；执行结果正确的用时最短的产品得满分，用时最长的厂商得基础分，其他厂商得分按时长差进行加权计算；允许有 1 小时的人工干预，但会扣除 20% 的得分。		
备注：	<ul style="list-style-type: none"> StartAPI：启动厂商的数据装载程序（从本地文件系统到 HBase） <ul style="list-style-type: none"> http://host/HDPTEST_FUNC_LOAD_HBASE/start.json 参数 无 json 返回值 <ul style="list-style-type: none"> ret: 返回结果，ok 表示正常启动，error 表示启动失败 msg: 启动失败时说明原因 StatusAPI：检查程序是否完成 <ul style="list-style-type: none"> http://host/HDPTEST_FUNC_LOAD_HBASE/status.json 参数 <ul style="list-style-type: none"> 无 json 返回值 <ul style="list-style-type: none"> status: 运行状态，running 表示正在运行，finish 表示已结束。 SelectAPI：用于结果检查 <ul style="list-style-type: none"> http://host/HDPTEST_FUNC_LOAD_HBASE/select.json 		

	<ul style="list-style-type: none"> ○ 参数 <ul style="list-style-type: none"> ▪ calling ▪ starttime: 开始时间, 比如 yyyyMMdd ▪ endtime: 结束时间, 比如 yyyyMMdd ○ json 返回值 <ul style="list-style-type: none"> ▪ size: 返回条数 ▪ data: 数据列表 		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

6.3.3.2. 查询性能调优测试

项目:	调优能力测试	分项目:	查询性能调优
用例编号:	HDPTEST_FUNC_QUERY_JOIN	版本:	
用例得分	100分	适用场景	通用
测试目的:	测试两张表进行 JOIN 操作并写入新表（文件）的时候的性能表现。		
预置条件:	1、厂商完成了Hadoop平台的部署； 2、需要执行JOIN操作的数据已经在HDFS上。		
测试步骤:	1、通过 RESTful 接口启动厂商的程序； 2、任务执行完成后，通过查询接口核对执行结果。截图 3-3-2-1		
预期结果:	可以在规定时间内执行完调优任务，并输出正确的结果。		
测试得分:			
评分标准:	测试执行过程出错或执行结果不正确，该测试用例不得分；执行结果正确的用时最短的产品得满分，用时最长的厂商得基础分，其他厂商得分按时长差进行加权计算；允许有 1 小时的人工干预，但会扣除 20%的得分。		
备注:	<ul style="list-style-type: none"> • JOIN操作 <ul style="list-style-type: none"> • 数据表 PS_USERSERVICE 与维表 DIM_SERVICE_SDK, DIM_CELL 进行 LEFT JOIN 操作。 JOIN 条件是翻译 PS_USERSERVICE 和 DIM_SERVICE_SDK 的业务类型（service_group_sdk 和 service_type_sdk 共同确定）和小 		

	<p>区地址（cell_id 和 lac 共同确定）两个字段的內容。</p> <ul style="list-style-type: none"> JOIN 输出表字段： PS_USERSERVICE.*,DIM_SERVICE_SDK.service_group_sdk_name,DIM_SERVICE_SDK.service_type_sdk_name,DIM_CELL.sac_name。 StartAPI：启动厂商的 JOIN 操作程序，将 JOIN 结果输出到一张新的表 <ul style="list-style-type: none"> http://host/HDPTTEST_FUNC_QUERY_JOIN/start.json 参数 <ul style="list-style-type: none"> 无 json 返回值 <ul style="list-style-type: none"> ret：返回结果，ok 表示正常启动，error 表示启动失败 msg：启动失败时说明原因 StatusAPI：检查程序是否完成 <ul style="list-style-type: none"> http://host/HDPTTEST_FUNC_QUERY_JOIN/status.json 参数 <ul style="list-style-type: none"> 无 json 返回值 <ul style="list-style-type: none"> status：运行状态，running 表示正在运行，finish 表示已结束。 SelectAPI：用于结果检查 <ul style="list-style-type: none"> http://host/HDPTTEST_FUNC_QUERY_JOIN/select.json 参数 <ul style="list-style-type: none"> calling starttime：开始时间，比如 yyyyMMdd endtime：结束时间，比如 yyyyMMdd json 返回值 <ul style="list-style-type: none"> size：返回条数 data：数据列表 		
测试执行人员：			
测试人员签字：		测试时间：	
测试审核员签字：		厂家配合人员签字：	

项目：	调优能力	分项目：	查询性能调优
用例编号：	HDPTEST_FUNC_QUERY_SUBQUERY	版本：	
用例得分	100分	适用场景	通用
测试目的：	SUBQUERY 性能测试，测试子查询操作的性能		
预置条件：	1、厂商完成了 Hadoop 平台的部署； 2、需要执行 SubQuery 操作的数据已经在 HDFS 上。		
测试步骤：	1、通过 RESTful 接口启动厂商的程序。 2、任务执行完成后，通过查询接口核对执行结果。截图 3-3-2-2		
预期结果：	可以在规定时间内执行完调优任务，并输出正确的结果。		
评分标准	测试执行过程出错或执行结果不正确，该测试用例不得分；执行结果正确的用时最短的产品得满分，用时最长的厂商得基础分，其他厂商得分按时长差进行加权计算；允许有 1 小时的人工干预，但会扣除 20% 的得分。		
备注：	<ul style="list-style-type: none"> SubQuery 操作 <ul style="list-style-type: none"> 执行 6.4.2.1 中的查询二 StartAPI：启动厂商的 SubQuery 程序，输出结果到一张新的表 <ul style="list-style-type: none"> http://host/HDPTEST_FUNC_QUERY_SUBQUERY/start.json 参数 <ul style="list-style-type: none"> count：通话阈值 starttime endtime json 返回值 <ul style="list-style-type: none"> ret：返回结果，ok 表示正常启动，error 表示启动失败 msg：启动失败时说明原因 StatusAPI：检查程序是否完成 <ul style="list-style-type: none"> http://host/HDPTEST_FUNC_QUERY_SUBQUERY/status.json 参数 <ul style="list-style-type: none"> 无 json 返回值 <ul style="list-style-type: none"> status：运行状态，running 表示正在运行，finish 表示已结束。 SelectAPI：用于结果检查 		

	<ul style="list-style-type: none"> • http://host/HDPTTEST_FUNC_QUERY_SUBQUERY/select.json • 参数 <ul style="list-style-type: none"> ○ calling ○ starttime: 开始时间, 比如 yyyyMMdd ○ endtime: 结束时间, 比如 yyyyMMdd • json 返回值 <ul style="list-style-type: none"> ○ size: 返回条数 ○ data: 数据列表, 每条对应原始数据一条。 		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

项目:	调优能力	分项目:	查询性能调优
用例编号:	HDPTTEST_FUNC_QUERY_CONCURRENCE	版本:	
用例得分:	100分	适用场景:	通用
测试目的:	多任务并发查询测试, 使用相同条件做并发查询, 验证查询效率		
预置条件:	1、厂商完成了 Hadoop 平台的部署; 2、需要进行高并发查询的数据已经装载到 HBase 中。		
测试步骤:	1、通过 RESTful 接口查询结果正确性; 截图 3-3-2-3-1 2、使用 Jmeter 进行高并发测试, 并发度为梯度测试, 绘制性能曲线图。截图 3-3-2-3-2-[并发数]		
预期结果:	高并发查询时不会报错。		
测试得分:			
评分标准:	Jmeter 查询响应结果不报错时, 按并发度和 90% 查询响应平均时间排序, 第一名得满分, 第二名获 60% 得分, 第三名得 30% 得分, 第四名得 0 分; 如果高并发查询报错得 0 分; 允许有 1 小时的人工干预, 但会扣除 20% 的得分。		
备注:	<ul style="list-style-type: none"> • 高并发查询测试 <ul style="list-style-type: none"> ○ 执行 6.4.3.1 的查询一。 • SelectAPI <ul style="list-style-type: none"> ○ http://host/HDPTTEST_FUNC_QUERY_CONCURRENCE/select.json ○ 参数 		

	<ul style="list-style-type: none"> ▪ calling ▪ starttime: 开始时间, 比如 yyyyMMdd ▪ endtime: 结束时间, 比如 yyyyMMdd ○ json 返回值 <ul style="list-style-type: none"> ▪ size: 返回条数 ▪ data: 数据列表, 每条对应原始数据一条。 		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

6.3.3.3. 数据处理性能调优测试

项目:	调优能力测试	分项目:	数据处理性能调优
用例编号:	HDPTEST_FUNC_PROCESS	版本:	
用例得分:	100分	适用场景:	通用
测试目的:	选定 MapReduce 计算框架, 在限定机器和可用总内存的前提下, 测试参数调节和性能调优的能力。		
预置条件:	1、完成 Hadoop 集群的部署。 2、需要进行处理的数据已经在 HDFS 上。		
测试步骤:	1、通过 RESTful 接口启动厂商提供的查询程序, 并将结果写到指定的 HDFS 目录 2、任务执行完成后, 通过查询接口核对执行结果。截图 3-3-3		
预期结果:	完成指定任务的数据处理速度。		
测试得分:			
评分标准:	测试执行过程出错或执行结果不正确, 该测试用例不得分; 执行结果正确的用时最短的产品得满分, 用时最长的产品得基础分, 其他产品得分按时长差进行加权计算; 允许有 1 小时的人工干预, 但会扣除 20% 的得分。		
备注:	<ul style="list-style-type: none"> • 查询内容: 对全部<PS_WAPHTTP>的日志文件进行 (数据量至少为 10TB), 根据电话号码(calling)进行日志条数频次统计(count), 输出频次在指定范围内的电话号码和相应的频次。 		

	<ul style="list-style-type: none"> StartAPI: 执行厂商的查询程序 <ul style="list-style-type: none"> http://host/HDPTTEST_FUNC_PROCESS/start.json 参数 <ul style="list-style-type: none"> up : 频次上限 down : 频次下限 (其中 up > down > 0) out: 结果输出目录 json 返回值 <ul style="list-style-type: none"> ret: 返回结果, ok 表示正常启动, error 表示启动失败 msg: 启动失败时说明原因 StatusAPI: 检查程序是否完成 <ul style="list-style-type: none"> http://host/HDPTTEST_FUNC_PROCESS/status.json 参数 <ul style="list-style-type: none"> 无 json 返回值 <ul style="list-style-type: none"> status: 运行状态, running 表示正在运行, finish 表示已结束。 SelectAPI <ul style="list-style-type: none"> http://host/HDPTTEST_FUNC_PROCESS/select.json 参数 <ul style="list-style-type: none"> calling json 返回值 <ul style="list-style-type: none"> size: 返回条数 data: 数据列表, 每条包含电话号码和频次。 		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

6.4. 模型管控能力测试

模型管控能力测试主要考察参测厂商对于业务的理解能力以及应用开发能力。本节将分别列出四个核心业务场景，包括离线批量复杂分析、即席查询、简单键值查询以及准实时流计算，每个场景下至少有一个具体的业务案例，结合案例设计测试案例及测试标准。

6.4.1. 离线批量复杂分析

6.4.1.1. 用户业务与流量分析

6.4.1.1.1. 业务需求说明

分析特定业务的用户行为，掌握该业务群体的业务使用粘性及共性，研究自有业务与竞争业务的用户差异及业务差异，为业务营销和业务应用提供精细维度剖析。主要从业务区域、用户数、流量等方面对比分析自有业务与竞争业务特点，分析共同用户在使用次数、流量、接入方式、使用时间、使用位置等方面的特点。

对高流量用户的流量行为进行定义分析，研究高流量用户的行为特征、业务偏好及其与网络资源的关系，掌握其流量产生来源和发展趋势。输出高流量用户名单，在流量、时间、终端、业务及位置等方面进行行为分析，得到用户行为特征。

6.4.1.1.2. 输入输出说明

输入：

- 1) 数据表：PS_USERSERVICE, PS_WAPHTTP
- 2) 接入维表：DIM_CELL, DIM_SERVICE_SDK, DIM_IMEI。

输出：

表一、业务分类表：对 PS_USERSERVICE 按时间、业务类型和子业务进行统计。

序号	指标名称	单位	字段或算法	备注
1	时间		first_time	统计维度，以天为时间粒度， 输出格式为 yyyyMMdd
2	业务类型		service_group_sdk_name	统计维度，PS_USERSERVICE

	名称			与 DIM_SERVICE_SDK 的 service_group_sdk 字段关联得出
3	子业务名称		service_type_sdk_name	统计维度，PS_USERSERVICE 与 DIM_SERVICE_SDK 的 service_type_sdk 字段关联得出
4	总用户数	人次	imsi 标识唯一用户	当天业务的总用户数
5	访问次数	次		当天业务的总访问数
6	总流量	Byte	$\Sigma(\text{up_bytes} + \text{down_bytes})$	当天业务的总流量

表二、URL 分类表，对 PS_WAPHTTP 按时间、业务类型和子业务进行统计。

序号	指标名称	单位	字段或算法	备注
1	时间		first_time	统计维度，以天为时间粒度
2	业务类型名称		service_group_sdk_name	统计维度，PS_WAPHTTP 与 DIM_SERVICE_SDK 的 service_group_sdk 字段关联得出
3	子业务名称		service_type_sdk_name	统计维度，PS_WAPHTTP 与 DIM_SERVICE_SDK 的 service_type_sdk 字段关联得出
4	总用户数	人次	imsi 标识唯一用户	当天 URL 类型的总用户数
5	访问次数	次		当天 URL 类型的总访问数
6	总流量	Byte	$\Sigma(\text{up_bytes} + \text{down_bytes})$	当天 URL 类型的总流量

表三、用户业务分类下钻表：对 PS_USERSERVICE 按时间、用户号码、IMEI 和 IMSI 进行统计。

序号	指标名称	单位	字段或算法	备注
----	------	----	-------	----

1	时间		first_time	统计维度，以天为时间粒度
2	用户号码		calling	统计维度
3	IMEI		imei	统计维度
4	IMSI		imsi	统计维度
5	业务类型名称		service_group_sdk_name	统计维度，PS_USERSERVICE 与 DIM_SERVICE_SDK 的 service_group_sdk 字段关联得出
6	子业务名称		service_type_sdk_name	统计维度，PS_USERSERVICE 与 DIM_SERVICE_SDK 的 service_type_sdk 字段关联得出
7	访问次数	次		当天该用户的总访问数
8	总流量	Byte	$\Sigma(\text{up_bytes} + \text{down_bytes})$	当天该用户的总流量

表四、用户 URL 分类下钻表：对 PS_WAPHTTP 按时间、用户号码、IMEI 和 IMSI 进行统计。

序号	指标名称	单位	字段或算法	备注
1	时间		first_time	统计维度，以天为时间粒度
2	用户号码		calling	统计维度
3	IMEI		imei	统计维度
4	IMSI		imsi	统计维度
5	业务类型名称		service_group_sdk	统计维度，PS_WAPHTTP 与 DIM_SERVICE_SDK 的 service_group_sdk 字段关联得出
6	子业务名称		service_type_sdk_name	统计维度，PS_WAPHTTP 与 DIM_SERVICE_SDK 的 service_type_sdk 字段关联

				得出
7	访问次数	次		当天该用户的总访问数
8	总流量	Byte	$\Sigma (\text{up_bytes} + \text{down_bytes})$	当天该用户的总流量

算法说明

统计所在地区的业务分类以及 url 分类的用户数、总流量、访问次数。并下钻用户使用的业务类型、访问次数和流量。

6.4.1.1.3. 测试用例情况汇总

项目：	离线批量复杂分析	分项目：	用户业务与流量分析
用例编号：	HDPTTEST_FUNC_BATCH1	版本：	
用例得分：	100分	适用场景：	通用
测试目的：	从业务区域、用户数、流量等方面对比分析自有业务与竞争业务特点，分析共同用户在使用次数、流量、接入方式、使用时间、使用位置等方面的特点。		
预置条件：	输入数据表：PS_USERSERVICE、PS_WAPHTTP 接入维表：DIM_CELL, DIM_SERVICE_SDK, DIM_IMEI 输入数据目录：hdfs:///data/HDPTTEST 结果输出到 4 张表中： 业务分类表：PS_USERSERVICE_BATCH1_SUMMARY URL 分类表：PS_WAPHTTP_BATCH1_SUMMARY 用户业务分类下钻表：PS_USERSERVICE_BATCH1_DETAILS 用户 URL 分类下钻表：PS_WAPHTTP_BATCH1_DETAILS		
测试步骤：	1、通过RESTful接口启动厂商程序； 2、任务执行完成后，通过相对应的查询接口检查结果的正确性。截图 4-1-1		
预期结果：	结果正确。		
测试得分：			
评分标准：	测试执行过程出错或执行结果不正确，该测试用例不得分；执行结果正确的用时最短的产品得满分，用时最长的厂商得基础分（20%），其他厂商得分按时长差进行加权计算；允许有 1 小时的人工干预，但会扣除 20%的得分。		
备注：	<ul style="list-style-type: none"> StartAPI：执行离线批量复杂分析，输出结果到数据表 <ul style="list-style-type: none"> http://host/HDPTTEST_FUNC_BATCH1/start.json 参数 		

	<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ 无 • json 返回值 <ul style="list-style-type: none"> ○ ret: 返回结果, ok 表示正常启动, error 表示启动失败 ○ msg: 启动失败时说明原因 • StatusAPI: 检查程序是否完成 <ul style="list-style-type: none"> • http://host/HDPTTEST_FUNC_BATCH1/status.json • 参数 <ul style="list-style-type: none"> ○ 无 • json 返回值 <ul style="list-style-type: none"> ○ status: 运行状态, running 表示正在运行, finish 表示已结束 (四张表全部统计结束)。 • SelectAPI: <ul style="list-style-type: none"> • 表一查询: http://host/HDPTTEST_FUNC_BATCH1/select1.json • 参数 <ul style="list-style-type: none"> ○ first_time: 时间 ○ service_group_sdk_name: 业务类型名称 ○ service_type_sdk_name: 子业务类型名称 • json 返回值 <ul style="list-style-type: none"> ○ size: 返回条数 ○ data: 数据列表 • 表二查询: http://host/HDPTTEST_FUNC_BATCH1/select2.json • 参数与返回值与表一一样。 • 表三查询: http://host/HDPTTEST_FUNC_BATCH1/select3.json • 参数 <ul style="list-style-type: none"> ○ first_time: 时间 ○ calling: 用户号码 • json 返回值 <ul style="list-style-type: none"> ○ size: 返回条数 ○ data: 数据列表 • 表四查询: http://host/HDPTTEST_FUNC_BATCH1/select4.json • 参数与返回值与表三一样。 		
测试执行人员:			
测试人员签字:		测试时间:	

测试审核员签字：		厂家配合人员 签字：	
----------	--	---------------	--

6.4.1.2. 网络故障分析

6.4.1.2.1. 业务需求说明

用户在进行数据业务使用的过程中，用户终端与网络的交互需经历附着网络、创建 PDP 上下文、DNS 查询、连接网关/SP、业务请求、位置（路由区）更新、PDP 上下文更新等一系列的网络及业务事件。这些事件在网络中需要经过终端、基站、BSC/RNC、SGSN、DNS 服务器、GGSN、网关、外网等多重网元设备。在这些交互过程中，往往由于用户终端或网元的问题，导致用户数据业务使用中出现失败、上网慢的情况。本功能主要为针对在 Gn 接口出现的用户在数据业务使用过程中的故障情况进行分析，帮助网络维护人员一览现网中的故障情况，分析故障的影响面，并快速定位故障原因及集中的网元。

6.4.1.2.2. 输入输出说明

输入：

1) 输入数据表：PS_WAPHTTP

2) 接入维表：无。

输出：

业务分类表：对 PS_WAPHTTP 按时间统计各种 host 的非 200 ~ 399 的 result 次数及、用户数。

序号	指标名称	单位	字段或算法	备注
1	时间		first_time	统计维度，以天为时间粒度
2	host		host	统计维度
3	result		result	统计维度, 非 200 ~ 399
4	用户数	人次	imsi 标识唯一用户	业务的总用户数
5	访问次数	次		业务的总访问数

算法说明：

本功能主要为对网络中的故障进行统计分析，分析故障的影响面，并定位故障原因及集中的网元。故障的影响面分析统计逻辑：对 PS_WAPHTTP 按时间统计各种 host 的非 200 ~ 399 的 result 次数及、用户数。

6.4.1.2.3. 测试用例情况汇总

项目：	开发能力测试	分项目	离线批量复杂分析
用例编号：	HDPTTEST_FUNC_BATCH2	版本：	
用例得分：	100分	适用场景：	通用
测试目的：	为对网络中的故障进行统计分析，分析故障的影响面，并定位故障原因及集中的网元。		
预置条件：	输入数据表：PS_WAPHTTP 接入维表：无 输入数据目录：hdfs:///data/HDPTTEST 输出表：PS_WAPHTTP_BATCH2_FALUT_SUMMARY		
测试步骤：	1、通过RESTful接口启动厂商的程序； 2、任务执行完成后，通过查询接口检查结果。截图4-1-2		
预期结果：	结果正确。		
测试得分：			
评分标准：	测试执行过程出错或执行结果不正确，该测试用例不得分；执行结果正确的用时最短的产品得满分，用时最长的产品得基础分，其他产品得分按时长差进行加权计算；允许有 1 小时的人工干预，但会扣除 20% 的得分。		
备注：	<ul style="list-style-type: none"> StartAPI：执行厂商的网络故障分析，输出结果到新的表 <ul style="list-style-type: none"> http://host/HDPTTEST_FUNC_BATCH2/start.json 参数 <ul style="list-style-type: none"> 无 json 返回值 <ul style="list-style-type: none"> ret：返回结果，ok 表示正常启动，error 表示启动失败 msg：启动失败时说明原因 StatusAPI：检查程序是否完成* <ul style="list-style-type: none"> http://host/HDPTTEST_FUNC_BATCH2/status.json 参数 <ul style="list-style-type: none"> 无 		

	<ul style="list-style-type: none"> • json 返回值 <ul style="list-style-type: none"> ◦ status: 运行状态, running 表示正在运行, finish 表示已结束。 • SelectAPI <ul style="list-style-type: none"> • http://host/HDPTEST_FUNC_BATCH2/select.json • 参数 <ul style="list-style-type: none"> ◦ first_time ◦ host • json 返回值 <ul style="list-style-type: none"> ◦ size: 返回条数 ◦ data: 数据列表, 每条对应一条结果数据 		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

6.4.2. 即席查询

6.4.2.1. CDR 智能查询

6.4.2.1.1. 业务需求说明

CDR 智能查询可以根据自户的需求,灵活的设置查询条件;可以通过界面选择或输入 SQL 查询语句方式,完成一系列复杂的包括信令协议表和基础数据表等多表查询功能,可针对常用的业务类型、用户、小区、错误码、终端等值得关注的内容进行相关记录筛选查询,并可实现在此基础上进行二次查询;筛选的内容可以进行通用输出,或输出成报表呈现。

6.4.2.1.2. 输入输出说明

输入:

1) 输入数据表: AIU_MOC、PS_USERSERVICE、PS_WAPHTTP

2) 接入维表：无

输出：

查询一：根据时间范围（start_time）和内容关键词（Keyword，模糊匹配），查询 PS_WAPHTTP，输出符合条件的 PS_WAPHTTP 所有字段；

查询二：根据时间范围（STARTTIME），统计 AIU_MOC 表中每个用户（CALLERNO）的通话频次（frequency），筛选出通话次数超过一定阈值的用户，然后在 PS_USERSERVICE 表中查询这些用户（calling）在同一时间段（first_time）内使用的业务（service_group_sdk，service_type_sdk），输出字段包括 first_time, calling, frequency(通话频次), service_group_sdk, service_type_sdk。

查询三：根据时间范围（first_time），从 PS_USERSERVICE 中筛选出使用过指定业务（service_group_sdk，service_type_sdk）的用户（calling），在 AIU_MOC 表中查询这些用户（CALLERNO）在同一时间段（STARTTIME）的 CDR 清单。

6.4.2.1.3. 算法说明

详见本文档 6.4.2.1.2 章节，输出指标的算法描述。

6.4.2.1.4. 测试用例情况汇总

项目：	即席查询	分项目：	CDR智能查询
用例编号：	HDPTEST_FUNC_CDR	版本：	
用例得分：	100分	适用场景：	通用
测试目的：	根据自户的需求，灵活的设置查询条件，对基础数据表等多表查询功能。		
预置条件：	输入数据表：AIU_MOC、AIU_MTC、PS_USERSERVICE、PS_WAPHTTP 接入维表：DIM_CELL, DIM_SERVICE_SDK, DIM_IMEI 输入数据目录：hdfs:///data/HDPTEST 输出数据表： 查询一输出 PS_WAPHTTP_CDR1_RESULT 查询二输出		

	<p>PS_WAPHTTP_CDR2_RESULT</p> <p>查询三输出</p> <p>PS_WAPHTTP_CDR3_RESULT</p>
测试步骤:	<p>1、通过RESTful接口启动厂商的程序；</p> <p>2、执行程序完成后，通过相应接口检查结果。截图4-2-1</p>
预期结果:	结果正确。
测试得分:	
评分标准:	<p>测试执行过程出错或执行结果不正确，该测试用例不得分；执行结果正确的用时最短的产品得满分，用时最长的产品得基础分，其他产品得分按时长差进行加权计算；允许有 1 小时的人工干预，但会扣除 20% 的得分。</p>
备注:	<ul style="list-style-type: none"> StartAPI：执行厂商的 CDR 查询，输出结果到新的数据表 <ul style="list-style-type: none"> http://host/HDPTEST_FUNC_CDR/start.json 参数 <ul style="list-style-type: none"> q1_starttime：开始时间，比如 yyyyMMdd q1_endtime：结束时间，比如 yyyyMMdd q1_keyword：内容关键词 q2_starttime：开始时间，比如 yyyyMMdd q2_endtime：结束时间，比如 yyyyMMdd q2_count：通话阈值 q3_starttime：开始时间，比如 yyyyMMdd q3_endtime：结束时间，比如 yyyyMMdd q3_service_group_sdk q3_service_type_sdk json 返回值 <ul style="list-style-type: none"> ret：返回结果，ok 表示正常启动，error 表示启动失败 msg：启动失败时说明原因 StatusAPI：检查程序是否完成 <ul style="list-style-type: none"> http://host/HDPTEST_FUNC_CDR/status.json 参数 <ul style="list-style-type: none"> 无 json 返回值 <ul style="list-style-type: none"> status：运行状态，running 表示正在运行，finish 表示已结束。

	<ul style="list-style-type: none"> • SelectAPI <ul style="list-style-type: none"> • 查询一: http://host/HDPTTEST_FUNC_CDR/select1.json • 参数 <ul style="list-style-type: none"> ◦ starttime: 开始时间, 比如 yyyyMMdd ◦ endtime: 结束时间, 比如 yyyyMMdd • json 返回值 <ul style="list-style-type: none"> ◦ size: 返回条数 ◦ data: 数据列表, 每条对应一条结果数据。 • 查询二: http://host/HDPTTEST_FUNC_CDR/select2.json • 参数 <ul style="list-style-type: none"> ◦ starttime: 开始时间, 比如 yyyyMMdd ◦ endtime: 结束时间, 比如 yyyyMMdd ◦ calling • json 返回值 <ul style="list-style-type: none"> ◦ size: 返回条数 ◦ data: 数据列表, 每条对应一条结果数据。 • 查询三: http://host/HDPTTEST_FUNC_CDR/select3.json • 参数 <ul style="list-style-type: none"> ◦ starttime: 开始时间, 比如 yyyyMMdd ◦ endtime: 结束时间, 比如 yyyyMMdd • json 返回值 <ul style="list-style-type: none"> ◦ size: 返回条数 ◦ data: 数据列表, 每条对应一条结果数据。 		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员 签字:	

6.4.3. 简单键值查询

6.4.3.1. 用户上网详单查询

6.4.3.1.1. 业务需求说明

以时间条件、用户号码、小区等固定条件查询用户话单记录。

测试数据的范围不少于 100 亿以上的记录集合，时间跨度不小于 15 日。

6.4.3.1.2. 输入输出说明

输入：

- 1) 输入数据表：AIU_MOC、PS_USERSERVIC 和 PS_WAPHTTP
- 2) 接入维表：无

输出：

查询一：根据时间范围（STARTTIME）和主叫号码（CALLERNO），查询 AIU_MOC，输出符合条件的 AIU_MOC 所有字段。

查询二：根据时间范围（first_time）和用户号码（calling），查询 PS_USERSERVICE，输出符合条件的 PS_USERSERVICE 所有字段。

查询三：根据时间范围（start_time）和用户号码（calling），查询 PS_WAPHTTP，输出符合条件的 PS_WAPHTTP 所有字段。

6.4.3.1.3. 算法说明

简单键值查询。

6.4.3.1.4. 测试用例情况汇总

项目：	简单键值查询	分项目：	用户详单上网查询
用例编号：	HDPTEST_FUNC_KV	版本：	
用例得分：	100分	适用场景：	通用
测试目的：	以时间条件、用户号码查询通话记录。		

预置条件:	输入数据表: AIU_MOC、PS_USERSERVIC 和 PS_WAPHTTP 输入数据目录: 无, 数据已经提前装载到 HBase 中 输出表: AIU_MOC_KV_RESULT PS_USERSERVICE_KV_RESULT PS_WAPHTTP_KV_RESULT		
测试步骤:	1、通过 RESTful 接口启动厂商的程序; 2、执行程序完成后, 通过相应接口检查结果, 截图 4-3-1;		
预期结果:	结果正确。		
测试得分:			
评分标准:	测试执行过程出错或执行结果不正确, 该测试用例不得分; 执行结果正确的用时最短的产品得满分, 用时最长的产品得基础分, 其他产品得分按时长差进行加权计算。允许有 1 小时的人工干预, 但会扣除 20% 的得分。		
备注:	<ul style="list-style-type: none"> 测试数据的范围不少于 100 亿以上的记录集合, 时间跨度不小于 15 日。 测试时, 使用多次查询的平均响应时间进行度量。 SelectAPI: 执行简单键值查询, 返回查询结果, 记录查询时间。 <ul style="list-style-type: none"> 查询一: http://host/HDPTTEST_FUNC_KV/select1.json 参数 <ul style="list-style-type: none"> calling starttime: 开始时间, 比如 yyyyMMdd endtime: 结束时间, 比如 yyyyMMdd json 返回值 <ul style="list-style-type: none"> size: 返回条数 data: 数据列表, 每条对应一条结果数据。 查询二/三同上 		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	

6.4.4. 准实时流计算

6.4.4.1. 重点区域保障指标计算

6.4.4.1.1. 业务需求说明

针对重点区域保障，需要能够准实时地分析关键指标和超过阈值的告警，以便能够进行有针对性的保障。

6.4.4.1.2. 输入输出说明

输入：

- 1) 输入数据表：PS_USERSERVICE 和 PS_WAPHTTP
- 2) 接入维表：DIM_CELL, DIM_SERVICE_SDK, DIM_IMEI。

输出：

浮点数的精度需要到小数点后两位。

指标一： 网络流量 LAC 指标_5 和 15 分钟

数据源：PS_USERSERVICE

输出指标：

序号	指标名称	单位	字段或算法	备注
1	时间	TIMESTAMP	first_time	统计的起始时间
2	LAC		lac	统计维度
3	总流量	MB	$\Sigma (\text{up_bytes} + \text{down_bytes})$	
4	上行流量	MB	$\Sigma \text{up_bytes}$	
5	下行流量	MB	$\Sigma \text{down_bytes}$	
6	上行平均速率	kbps	上行流量/传输时长	
7	下行平均速率	kbps	下行流量/传输时长	
8	传输时长	秒	$\Sigma \text{trans_time} / 100000$	

指标二： 网络流量 RAT 指标_5 和 15 分钟

数据源：PS_USERSERVICE

输出指标：

序号	指标名称	单位	字段或算法	备注
1	时间	TIMESTAMP	first_time	统计的起始时间
2	RAT		rat	统计维度
3	总流量	MB	Σ (up_bytes+ down_bytes)	
4	上行流量	MB	Σ up_bytes	
5	下行流量	MB	Σ down_bytes	
6	上行平均速率	kbps	上行流量/传输时长	
7	下行平均速率	kbps	下行流量/传输时长	
8	传输时长	秒	Σ trans_time/100000	

指标三： 业务类型 LAC 指标_5 和 15 分钟

数据源：PS_USERSERVICE

输出指标：

序号	指标名称	单位	字段或算法	备注
1	时间	TIMESTAMP	first_time	统计的起始时间
2	LAC		lac	统计维度
3	业务类型		service_group_sdk	统计维度
4	流量	MB	Σ (up_bytes+ down_bytes)	
5	流量占比		流量/总流量（不设维度的流量总和）	总流量不显示
6	用户数		按 imsi 去重计数	

指标四： 业务类型 RAT 指标_5 和 15 分钟

数据源：PS_USERSERVICE

输出指标：

序号	指标名称	单位	字段或算法	备注
1	时间	TIMESTAMP	first_time	统计的起始时间

2	RAT		rat	统计维度
3	业务类型		service_group_sdk	统计维度
4	流量	MB	Σ (up_bytes+ down_bytes)	
5	流量占比		流量/总流量（不设维度的流量总和）	总流量不显示
6	用户数		按 imsi 去重计数	

指标五：异常流量用户 5 和 15 分钟

统计条件：2 以每 5/15 分钟为粒度统计流量大于 45MB/135MB 的每用户；

统计列表：时间、用户 calling、终端 IMEI、APN、小区（CI）、终端类型（需终端配置表映射）、网络制式（2\3G）、sum(上行流量)、sum(下行流量)、总流量；

数据源：PS_USERSERVICE

输出指标：

序号	指标名称	单位	字段或算法	备注
1	时间	TIMESTAMP	first_time	统计的起始时间
2	calling		calling	统计维度
3	终端 IMEI		imei	统计维度
4	APN		apn	统计维度
5	小区名		(lac+cell_id)=> sac_name	根据小区维表映射成小区名，统计维度
6	终端类型		imei=> firm	根据终端类型维表映射成终端类型，统计维度
7	网络制式		rat	统计维度
8	sum(上行流量)	MB	Σ (up_bytes)	
9	sum(下行流量)	MB	Σ (up_bytes)	
10	总流量	MB	Σ (up_bytes+ down_bytes)	

6.4.4.1.3. 算法说明

详见本文档 6.4.4.1.2 章节，输出指标的算法描述。

6.4.4.1.4. 测试用例情况汇总

项目：	准实时流计算	分项目：	重点区域保障计算
用例编号：	HDPTEST_FUNC_INSTANT	版本：	
用例得分：	100分	适用场景：	通用
测试目的：	针对重点区域保障，需要能够准实时地分析关键指标，以便能够进行有针对性的保障。		
预置条件：	输入数据表：PS_USERSERVICE 接入维表：DIM_CELL, DIM_SERVICE_SDK, DIM_IMEI 数据输入方法：通过 Socket 通信方式提供数据 输出表： PS_USERSERVICE_INSTANT_LAC_BYTES_5min, PS_USERSERVICE_INSTANT_LAC_BYTES_15min PS_USERSERVICE_INSTANT_RAT_BYTES_5min, PS_USERSERVICE_INSTANT_RAT_BYTES_15min PS_USERSERVICE_INSTANT_LAC_SERVICE_TYPE_5min, PS_USERSERVICE_INSTANT_LAC_SERVICE_TYPE_15min PS_USERSERVICE_INSTANT_RAT_SERVICE_TYPE_5min, PS_USERSERVICE_INSTANT_RAT_SERVICE_TYPE_15min PS_USERSERVICE_INSTANT_BYTES_ABNORMAL_USER_5min, PS_USERSERVICE_INSTANT_BYTES_ABNORMAL_USER_15min		
测试步骤：	1. 厂商准备好实时流处理程序，连接我方提供的socket server； 2. 通过脚本让socket server开始传输数据，持续时间约为1小时，最后一条数据结束后，socket server端的socket自动close，传输数据结束；截图socket传输状态4-4-1-1 3. 等待厂商的实时分析程序结束，输出到指定的表中，然后通过相应接口检查结果。截图4-4-1-2		
预期结果：	结果正确。		
测试得分：			
评分标准：	5/15 分钟统计各占该用例的 50%。每个指标测试执行过程出错或执行结果不正确，该指标不得分；执行结果正确的平均计算延迟最短的产品得满分，平均计算延迟最长的产品得基础分，其他产品得分按时平		

	均计算延迟进行加权计算。允许有 1 小时的人工干预，但会扣除 20% 的得分。
备注：	<ul style="list-style-type: none"> 测试指标 <ul style="list-style-type: none"> 统计每 5 分钟（数据量的）计算平均延迟 统计每 15 分钟（数据量的）计算平均延迟 测试数据 <ul style="list-style-type: none"> 测试 socket server 5 分钟可以传输数据量的上限，记做 N 条。 准备约 12*N 条数据（每 5 分钟约 N 条），修改数据字段中时间戳，保证有序。 测试细节 <ul style="list-style-type: none"> 准备 socket server 发送第 i 个 5/15 分钟区间数据前，记录当前时间戳 T5_i_start 或 T15_i_start； 当 status.json 对第 i 个 5/15 分钟区间计算返回 finish 状态时，记录时间戳 T5_i_end 或 T15_i_end，并立即调用 select.json 接口进行结果检查； 结果计算 <p>5 分钟平均延迟 = $\text{avg}(\text{sum}(T5_i_end - T5_i_start)) - 5\text{min}$</p> <p>15 分钟平均延迟 = $\text{avg}(\text{sum}(T15_i_end - T15_i_start)) - 15\text{min}$</p> StartAPI：启动厂商的流式计算程序，输出结果到数据表 <ul style="list-style-type: none"> http://host/HDPTEST_FUNC_INSTANT/start.json 参数 <ul style="list-style-type: none"> 无 json 返回值 <ul style="list-style-type: none"> ret：返回结果，ok 表示正常启动，error 表示启动失败 msg：启动失败时说明原因 StatusAPI：检查程序是否完成 <ul style="list-style-type: none"> http://host/HDPTEST_FUNC_INSTANT/status.json 参数 <ul style="list-style-type: none"> i：标识第 i 个 5/15 分钟区间，5 分钟对应取值为 1-12，15 分钟对应取值为 1-4。 type：统计时间类型，取值为 5 或 15。 json 返回值 <ul style="list-style-type: none"> status：运行状态，running 表示正在运行，finish

	<p>表示已结束。</p> <ul style="list-style-type: none"> • SelectAPI <ul style="list-style-type: none"> • 指标一: http://host/HDPTEST_FUNC_INSTANT/select1.json • 参数 <ul style="list-style-type: none"> ◦ starttime: 起始时间, 比如 yyyyMMddhhmm ◦ endtime: 结束时间, 比如 yyyyMMddhhmm ◦ lac • json 返回值 <ul style="list-style-type: none"> ◦ size: 返回条数 ◦ data: 数据列表, 每条对应一条结果数据。 • 指标二: http://host/HDPTEST_FUNC_INSTANT/select3.json • 参数 <ul style="list-style-type: none"> ◦ starttime: 起始时间, 比如 yyyyMMddhhmm ◦ endtime: 结束时间, 比如 yyyyMMddhhmm ◦ rat • json 返回值 <ul style="list-style-type: none"> ◦ size: 返回条数 ◦ data: 数据列表, 每条对应一条结果数据。 • 指标三/四参考指标一/二 • 指标五: http://host/HDPTEST_FUNC_INSTANT/select9.json • 参数 <ul style="list-style-type: none"> ◦ starttime: 起始时间, 比如 yyyyMMddhhmm ◦ endtime: 结束时间, 比如 yyyyMMddhhmm ◦ calling • json 返回值 <ul style="list-style-type: none"> ◦ size: 返回条数 ◦ data: 数据列表, 每条对应一条结果数据。 		
测试执行人员:			
测试人员签字:		测试时间:	
测试审核员签字:		厂家配合人员签字:	