

# FusionInsight HD

## 集成设计

[www.huawei.com](http://www.huawei.com)





# 目标

- 学完本课程后，您将能够：
  - 熟悉集群组网设计
  - 掌握集群节点部署规划
  - 掌握磁盘规划
  - 熟悉集群设计案例



# 目录

1. 简介
2. 集群组网设计
3. 集群节点部署规划
4. 磁盘规划
5. 配置规划工具使用
6. 设计案例

# 安装要求

约束项	描述
OS	<ul style="list-style-type: none"><li>• <b>SUSE 11 SP1/SP2/SP3 for AMD64 &amp; Intel64</b></li><li>• <b>RedHat-6.4-x86_64,RedHat-6.5-x86_64,RedHat-6.6-x86_64</b></li><li>• <b>CentOS-6.4,CentOS-6.5,CentOS-6.6</b></li></ul>
浏览器	<ul style="list-style-type: none"><li>• <b>Internet Explorer 9</b>标准模式</li><li>• <b>Google Chrome 21</b>及以上版本</li></ul>
业务平面带宽	建议采用 <b>10GE</b> 网络（根据典型的流量模型评估结果，为了让节点间的交换带宽不成为系统性能的瓶颈）
管理平面带宽	建议管理节点使用 <b>10GE</b> ，其他节点使用 <b>GE</b> （根据管理平面业务模型进行评估）
IP规划	当采用双平面隔离组网时，集群中每个节点有两个 <b>IP</b> 地址，包括业务平面 <b>IP</b> 地址和管理平面 <b>IP</b> 地址
硬盘分区规则	当集群节点规模较小时（小于 <b>30</b> 节点），控制节点的 <b>IO</b> 负载一般不高，控制节点节点的元数据可采用共盘分区的模式；节点规格较大（大于 <b>30</b> 节点）建议采用分盘的模式

# 集群评估

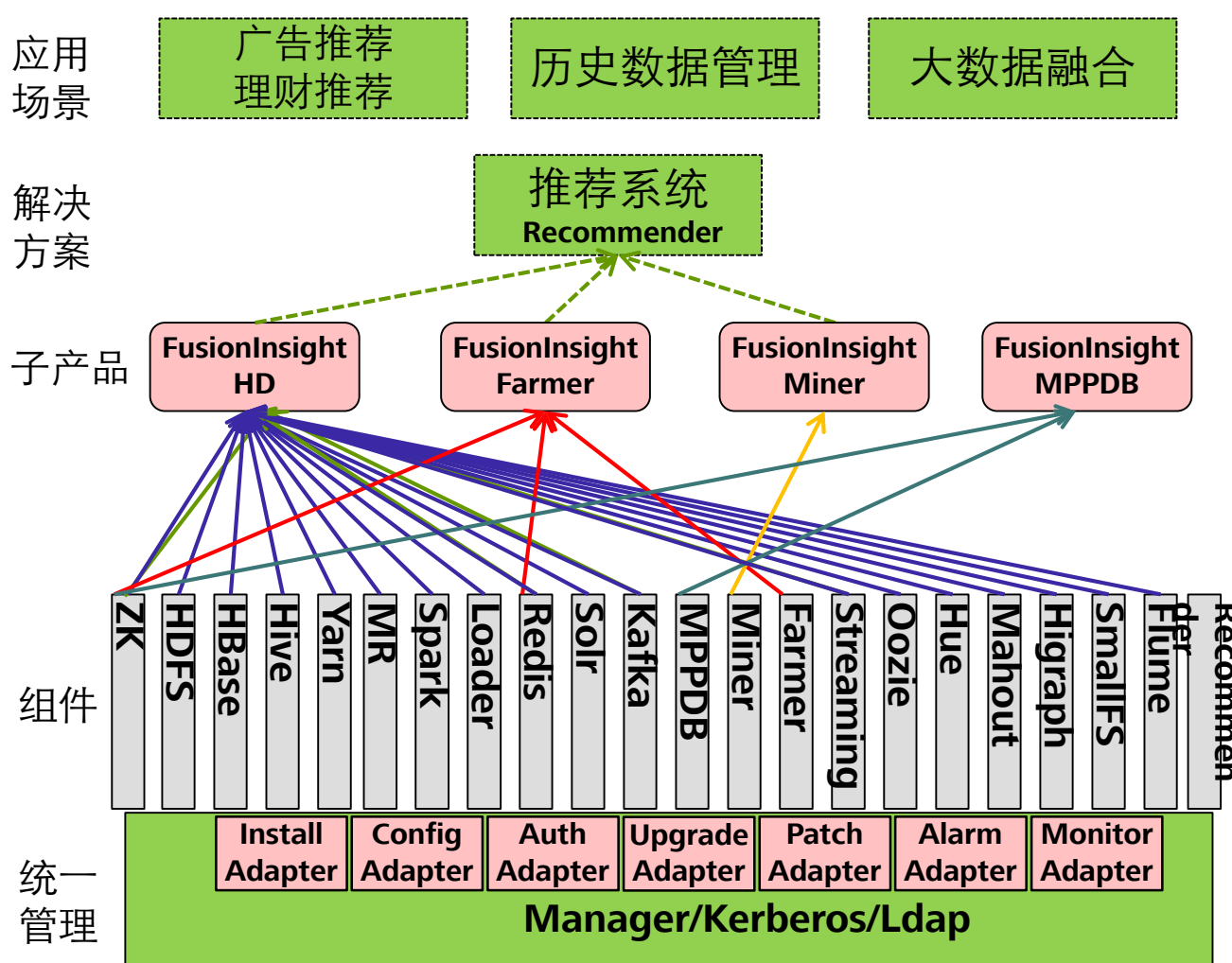
- 1.大数据业务类型众多，例如以存储为主或者以计算为主，同时**Hadoop**集群业务处理复杂，涉及因素众多，没法给出一个通用的计算公式；
- 2.本业务模型是在一系列贴近实际业务的假设下，给出的一个理论业务模型，如下几种典型场景供参考，根据数据量确定集群规模：
  - **6节点Mini业务模型：**
    - 系统并发**Map**总个数：**60**个
    - **HBase** 客户端并发个数：**10**个(单个客户端发送写请求个数**50**次/秒，每次**50**个记录数)
  - **200节点业务模型：**
    - 系统并发**Map**总个数：**3900**个
    - **HBase** 客户端并发个数：**100**个(场景同上)
  - **500节点业务模型：**
    - 系统并发**Map**总个数：**9900**个
    - **HBase** 客户端并发个数：**200**个(场景同上)



# 目录

1. 简介
- 2. 集群组网设计**
3. 集群节点部署规划
4. 磁盘规划
5. 配置规划工具使用
6. 设计案例

# 解决方案介绍



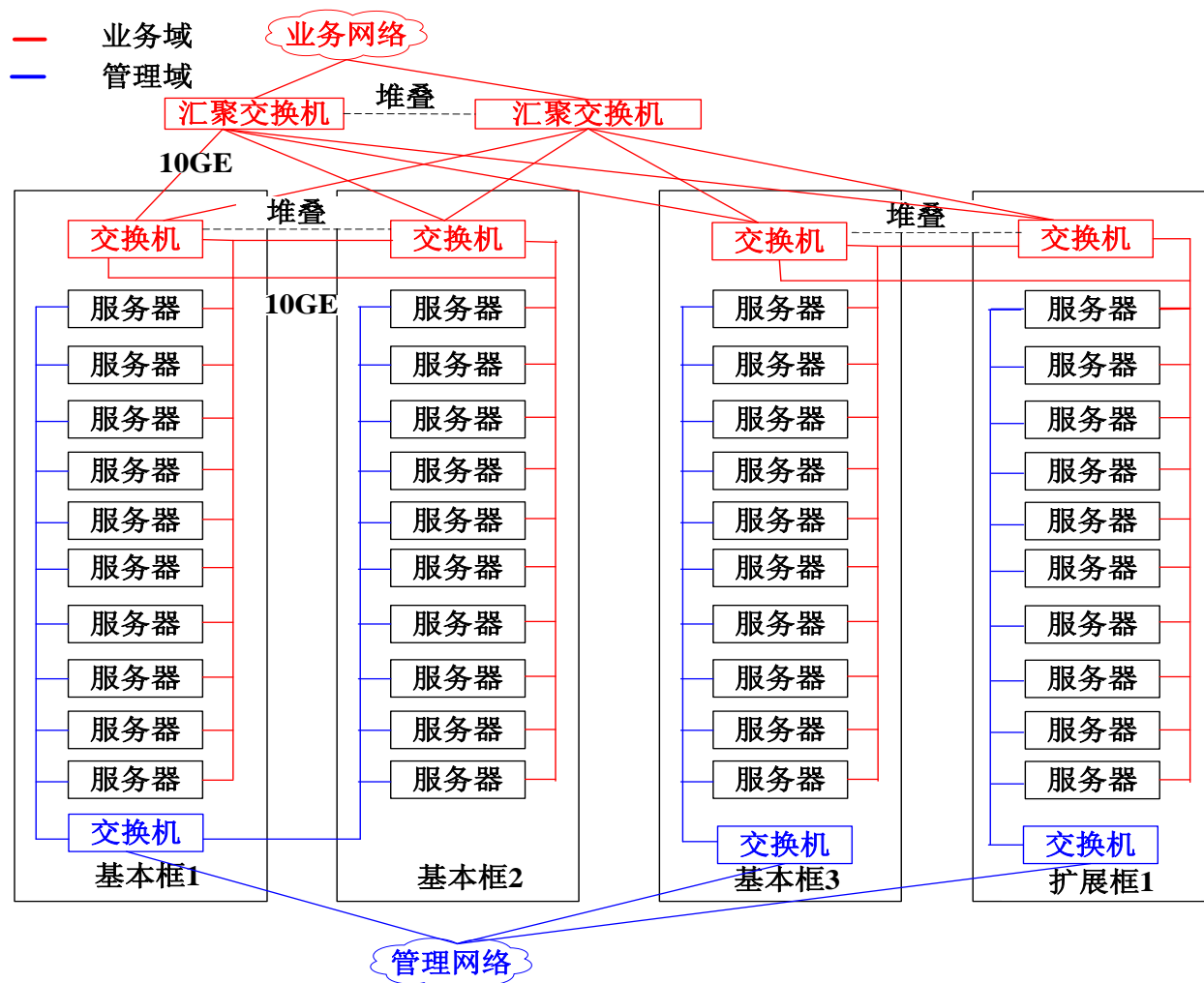
**FusionInsight HD:**  
FusionInsight的Hadoop是指广义的hadoop生态圈，提供商用Hadoop基础平台。

**FusionInsight Farmer :**  
提供数据服务框架，大数据实时应用使能器，支撑企业快速开发基于大数据平台的应用。

**FusionInsight Miner :**  
提供数据挖掘服务集，基于分布式内存计算的数据分析平台。

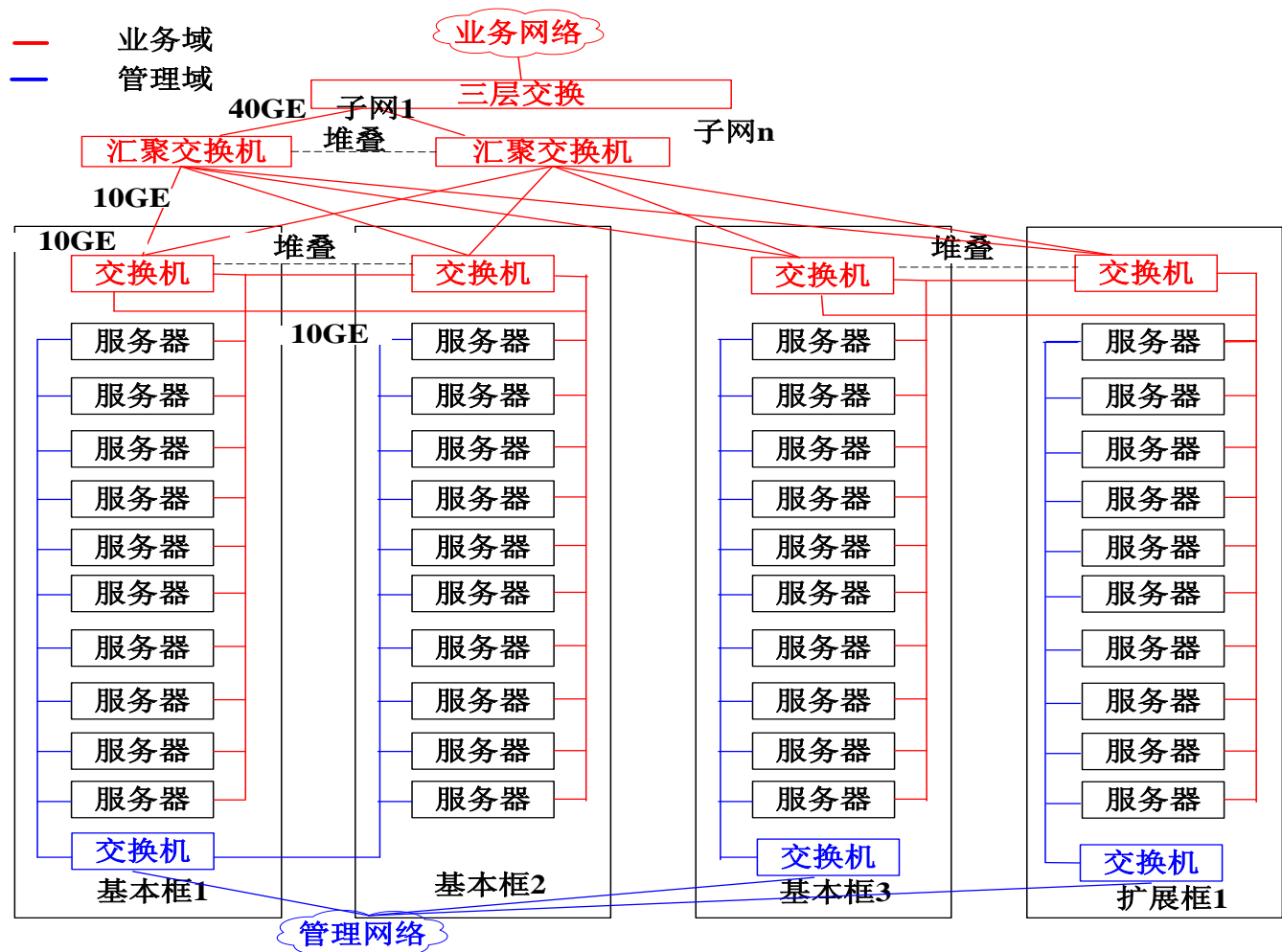
**FusionInsight MPPDB :**  
提供独立部署的通用MPP数据库，用于性能较高的交互分析场景。

# 集群组网设计——二层组网





# 集群组网设计——三层组网

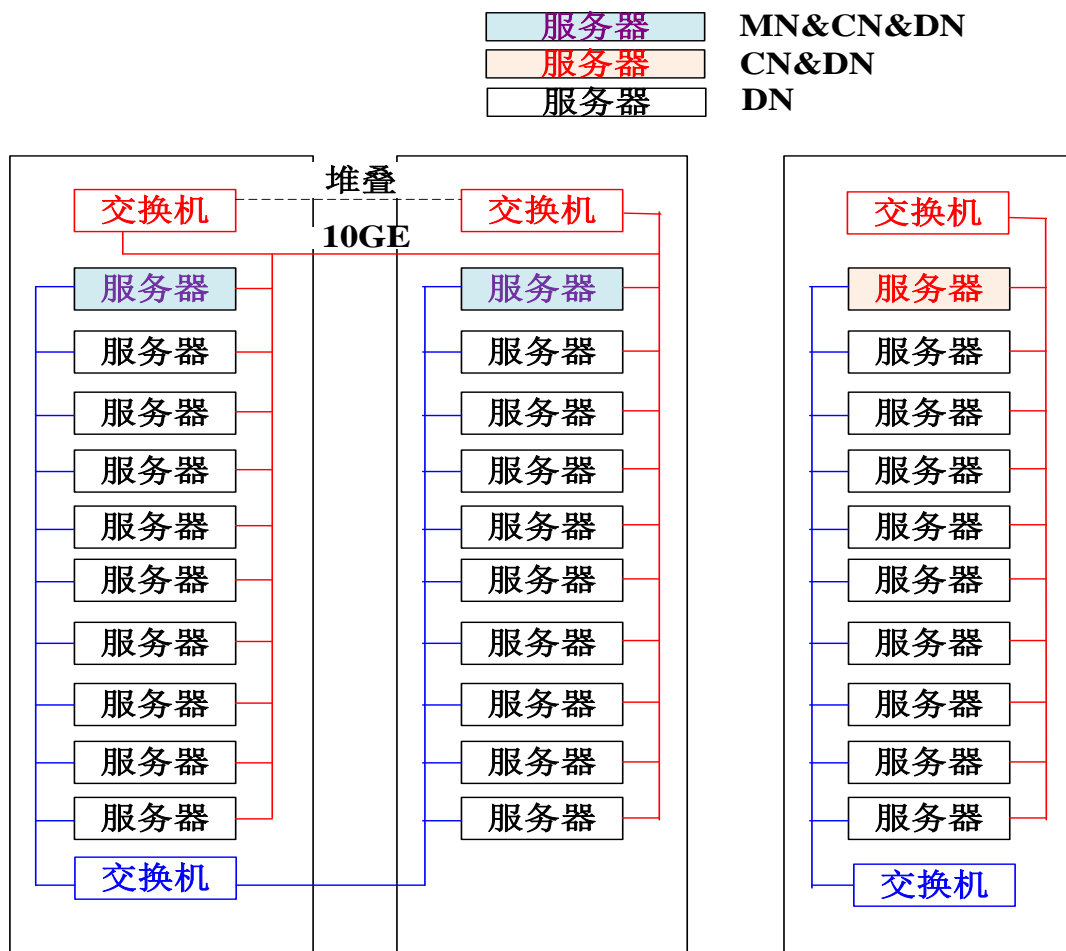




# 目录

1. 简介
2. 集群组网设计
- 3. 集群节点部署规划**
4. 磁盘规划
5. 配置规划工具使用
6. 设计案例

# 管理节点/控制节点/数据节点合一部署



- 适用场景：
  - 节点数小于30的集群
- 部署要点：
  - 管理节点&控制节点&数据节点合一部署
- 说明：
  - 如节点数满足要求，建议将数据节点单独部署。如节点数不满足单独部署，必须使用该场景时，需要使用双平面组网方式。

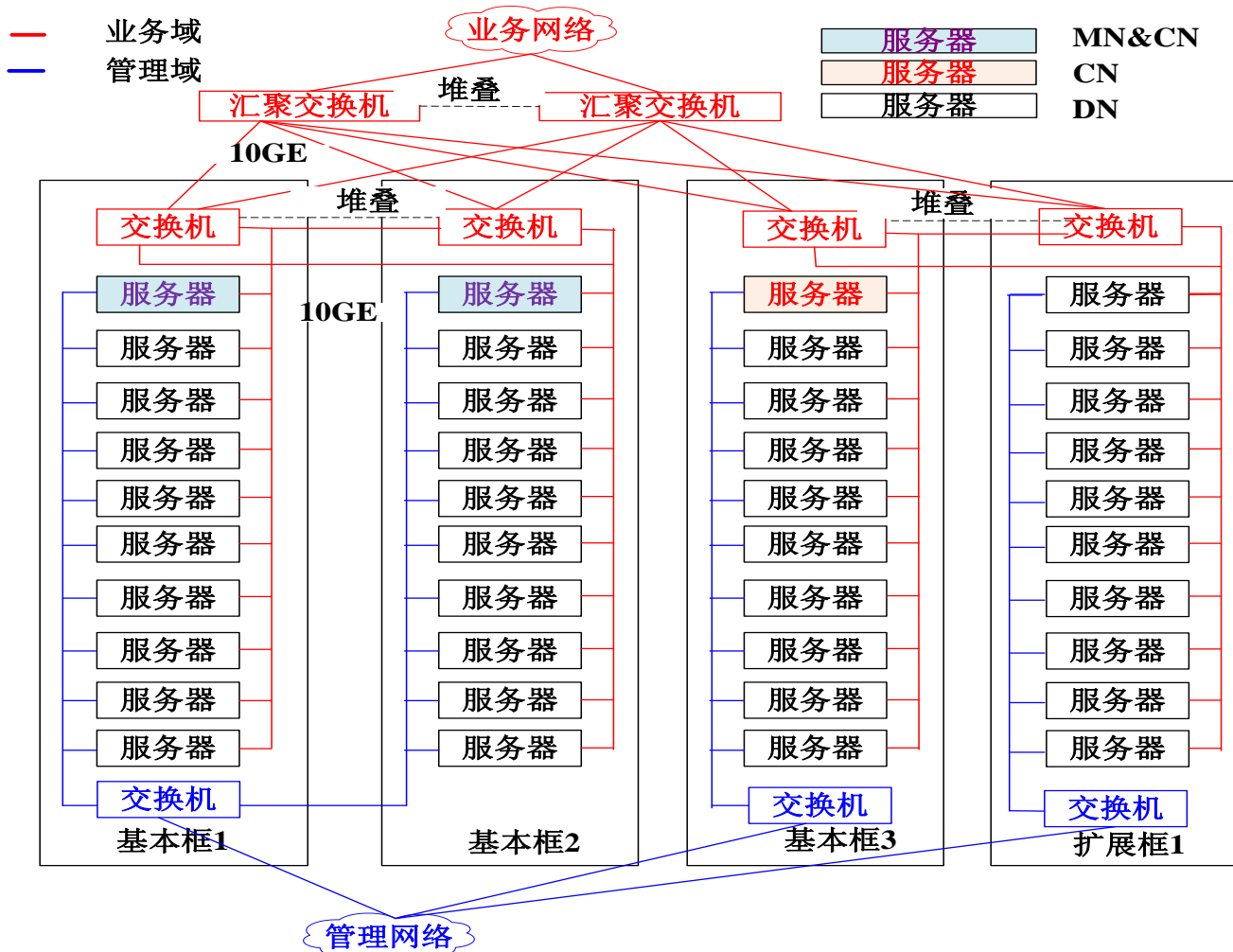
# 管理节点/控制节点/数据节点合部署 (一)

组件	角色	主机1 MN&CN&DN	主机2 MN&CN&DN	主机3 CN&DN	主机4 DN	主机5 DN	主机6 DN	主机7 DN	主机8 DN
OMSServer	OMSServer	Y	Y						
LdapServer	SlapdServer	Y	Y						
KrbServer	KerberosServer	Y	Y						
	KerberosAdmin	Y	Y						
ZooKeeper	quorumpeer	Y	Y	Y					
HDFS	Zkfc		Y	Y					
	NameNode		Y	Y					
	DataNode	Y	Y	Y	Y	Y	Y	Y	Y
	JournalNode	Y	Y	Y					
Yarn	NodeManager	Y	Y	Y	Y	Y	Y	Y	Y
	ResourceManager		Y	Y					
Mapreduce	JobHistoryServer			Y					
DBService	DBServer	Y		Y					
Hue	Hue	Y		Y					
Loader	LoaderServer	Y	Y						
Spark	JDBCServer		Y	Y					
	JobHistory		Y	Y					
	SparkResource	Y	Y	Y	Y	Y	Y	Y	Y
Hive	MetaStore	Y	Y						
	WebHCat			Y					
	HiveServer	Y	Y						

# 管理节点/控制节点/数据节点合一部署 (二)

组件	角色	主机1 MN&C N&DN	主机2 MN&C N&DN	主机3 CN&DN	主机4 DN	主机5 DN	主机6 DN	主机7 DN	主机8 DN
HBase	HMaster		Y	Y					
	RegionServer	Y	Y	Y	Y	Y	Y	Y	Y
	ThriftServer	Y	Y	Y					
SmallFS	FGCServer	Y	Y						
FTP-Server	FTP-Server		Y	Y					
Flume	Flume	Y	Y	Y	Y	Y	Y	Y	Y
	MonitorServer	Y	Y						
Kafka	Broker	Y	Y	Y	Y	Y	Y	Y	Y
Streaming	Nimbus	Y	Y						
	Supervisor	Y	Y	Y	Y	Y	Y	Y	Y
	Logviewer	Y	Y	Y	Y	Y	Y	Y	Y
	UI	Y	Y						
Metadata	MetadataServer			Y					
Oozie	oozie		Y	Y					
Solr	SolrServerAdmin	Y	Y						
	SolrServer1	Y	Y	Y	Y	Y	Y	Y	Y
	SolrServer2	Y	Y	Y	Y	Y	Y	Y	Y
	SolrServer3	Y	Y	Y	Y	Y	Y	Y	Y
	SolrServer4	Y	Y	Y	Y	Y	Y	Y	Y
	SolrServer5	Y	Y	Y	Y	Y	Y	Y	Y

# 管理节点/控制节点合一部署



# 管理节点/控制节点合部署（一）

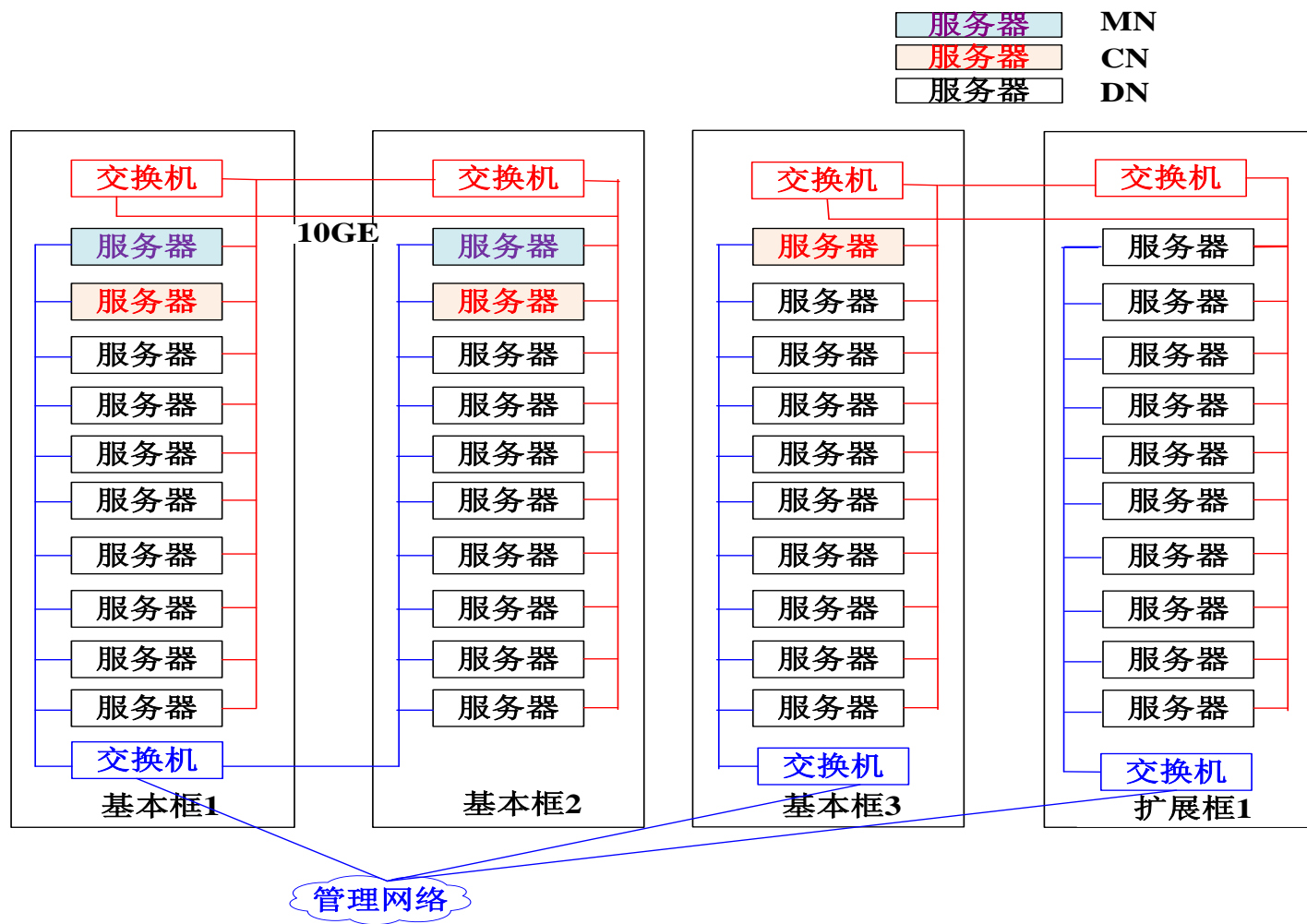
组件	角色	主机1 MN&CN	主机2 MN&CN	主机3 CN	主机4 DN	主机5 DN	主机6 DN	主机7 DN	主机8 DN
OMSServer	OMSServer	Y	Y						
LdapServer	SlapdServer	Y	Y						
KrbServer	KerberosServer	Y	Y						
	KerberosAdmin	Y	Y						
ZooKeeper	quorumpeer	Y	Y	Y					
HDFS	Zkfc		Y	Y					
	NameNode		Y	Y					
	DataNode				Y	Y	Y	Y	Y
	JournalNode	Y	Y	Y					
Yarn	NodeManager				Y	Y	Y	Y	Y
	ResourceManager		Y	Y					
Mapreduce	JobHistoryServer			Y					
DBService	DBServer	Y		Y					
Hue	Hue	Y		Y					
Loader	LoaderServer				Y	Y			
Spark	JDBCServer		Y	Y					
	JobHistory		Y	Y					
	SparkResource				Y	Y	Y	Y	Y
Hive	MetaStore	Y	Y						
	WebHCat			Y					
	HiveServer	Y	Y						

# 管理节点/控制节点合部署（二）

组件	角色	主机1 MN&C N	主机2 MN&C N	主机3 CN	主机4 DN	主机5 DN	主机6 DN	主机7 DN	主机8 DN
HBase	HMaster		Y	Y					
	RegionServer				Y	Y	Y	Y	Y
	ThriftServer	Y	Y	Y					
SmallFS	FGCServer	Y	Y						
FTP-Server	FTP-Server		Y	Y					
Flume	Flume				Y	Y	Y	Y	Y
	MonitorServer	Y	Y						
Kafka	Broker				Y	Y	Y	Y	Y
Streaming	Nimbus	Y	Y						
	Supervisor				Y	Y	Y	Y	Y
	Logviewer				Y	Y	Y	Y	Y
	UI	Y	Y						
Metadata	MetadataServer			Y					
Oozie	oozie		Y	Y					
Solr	SolrServerAdmin				Y	Y			
	SolrServer1				Y	Y	Y	Y	Y
	SolrServer2				Y	Y	Y	Y	Y
	SolrServer3				Y	Y	Y	Y	Y
	SolrServer4				Y	Y	Y	Y	Y
	SolrServer5				Y	Y	Y	Y	Y



# 管理节点/控制节点/数据节点分设部署



# 管理节点/控制节点/数据节点分设部署

组件	角色	主机1 MN	主机2 MN	主机3 CN	主机4 CN	主机5 CN	主机6 DN	主机7 DN	主机8 DN
OMSServer	OMSServer	Y	Y						
LdapServer	SlapdServer			Y	Y				
KrbServer	KerberosServer			Y	Y				
	KerberosAdmin			Y	Y				
ZooKeeper	quorumpeer			Y	Y	Y			
HDFS	Zkfc				Y	Y			
	NameNode				Y	Y			
	DataNode						Y	Y	Y
	JournalNode			Y	Y	Y			
Yarn	NodeManager						Y	Y	Y
	ResourceManager				Y	Y			
Mapreduce	JobHistoryServer					Y			
DBService	DBServer			Y		Y			
Hue	Hue			Y		Y			
Loader	LoaderServer						Y	Y	
Spark	JDBCServer				Y	Y			
	JobHistory				Y	Y			
	SparkResource						Y	Y	Y
Hive	MetaStore			Y	Y				
	WebHCat					Y			
	HiveServer			Y	Y				

# 管理节点/控制节点/数据节点分设部署

组件	角色	主机1 MN	主机2 MN	主机3 CN	主机4 CN	主机5 CN	主机6 DN	主机7 DN	主机8 DN
HBase	HMaster				Y	Y			
	RegionServer						Y	Y	Y
	ThriftServer			Y	Y	Y			
SmallFS	FGCServer			Y	Y				
FTP-Server	FTP-Server				Y	Y			
Flume	Flume						Y	Y	Y
	MonitorServer			Y	Y				
Kafka	Broker						Y	Y	Y
Streaming	Nimbus			Y	Y				
	Supervisor						Y	Y	Y
	Logviewer						Y	Y	Y
	UI			Y	Y				
Metadata	MetadataServer					Y			
Oozie	oozie				Y	Y			
Solr	SolrServerAdmin						Y	Y	
	SolrServer1						Y	Y	Y
	SolrServer2						Y	Y	Y
	SolrServer3						Y	Y	Y
	SolrServer4						Y	Y	Y
	SolrServer5						Y	Y	Y



# 目录

1. 简介
2. 集群组网设计
3. 集群节点部署规划
- 4. 磁盘规划**
5. 配置规划工具使用
6. 设计案例

# 节点OS磁盘规划

Raid	用途	磁盘容量 GB	分区目录	分区容量 GB	用途
Raid1	所有节点的OS盘	>=600	/	10	操作系统根分区，包括了所有目录（除去下面指定的目录）
			/tmp	10	临时文件存放目录
			/var	10	操作系统运行时目录
			/var/log	50	日志存放目录
			/srv/BigData	60	作为下面的/srv/BigData/*的母路径。
			/opt	all	程序存放目录

# 元数据共盘规划

Raid	用途	分区目录	分区容量 GB	用途
Raid1	管理节点和控制节点的元数据盘	/srv/BigData/dbdata_om	>=200	OMS的数据目录
		/srv/BigData/dbdata_service	>=100	组件数据库服务数据目录
		/srv/BigData/zookeeper	>=100	存放ZK数据
		/srv/BigData/journalnode	>=100	存放journalnode数据
		/srv/BigData/LocalBackup	>=100	集群OMS数据、元数据本地备份的数据目录
		/srv/BigData/streaming	>=100	存放Streaming Nimbus的元数据
		/srv/BigData/namenode	>=100	存放NameNode数据

# 管理节点独立部署

Raid	用途	磁盘容量 GB	分区目录	分区 容量 GB	用途
Raid 1	管理节点元数据盘	>=600	/srv/BigData/dbdata_om	all	OMS的数据目录
Raid 1		>=600	/srv/BigData/LocalBackup	all	集群OMS数据、元数据本地备份的数据目录

说明：管理节点OS盘规划同17页

# 控制节点独立部署

Raid	用途	磁盘容量 GB	分区目录	分区 容量 GB	用途
Raid1	控制节点元数据盘	>=600	/srv/BigData/dbdata_service	all	组件数据库服务数据目录
Raid1		>=600	/srv/BigData/zookeeper	all	存放ZK数据
Raid1		>=600	/srv/BigData/journalnode	all	存放journalnode数据
Raid1		>=600	/srv/BigData/streaming	all	存放Streaming Nimbus的元数据
Raid1		>=600	/srv/BigData/namenode	all	存放NameNode数据

说明：控制节点OS盘规划同17页



# 数据节点独立部署

Raid	用途	磁盘容量 GB	分区目录	分区 容量 GB	用途
Raid0 /non- Raid	数据节点	>=600	/srv/BigData/hadoop/data1~N	all	存放 <b>DataNode</b> 数据以及 <b>MR</b> 中间数据
		>=600	/srv/BigData/kafka/data1~N	all	存放 <b>Kafka Broker</b> 数据
		>=600	/srv/BigData/redis/Redis_1~32	all	存放 <b>Redis</b> 数据库数据，允许多个分区放在一个磁盘上，每块盘最大 <b>5</b> 个分区
		>=600	/srv/BigData/solr/solrserver1~5	all	存放 <b>Solr</b> 的元数据和索引数据， <b>Solr</b> 的索引数据放在本地时，会规划该分区
		>=600	/srv/BigData/solr/solrserveradmin	all	存放 <b>Solr</b> 的元数据和索引数据， <b>Solr</b> 的索引数据放在本地时，会规划该分区
		>=600	/srv/BigData/streaming_data	all	存放 <b>Streaming Supervisor</b> 及 <b>workers</b> 的数据

说明：数据节点**OS**盘规划同**17**页



# 目录

1. 简介
2. 集群组网设计
3. 集群节点部署规划
4. 集群磁盘规划
- 5. 配置规划工具使用**
6. 设计案例

# LLD配置规划工具使用简介

- 基础配置

配置套餐需选择正确，参考“节点部署规划”章节，配置完成后点击左上角“下一步”按钮，下同

- 选择服务

选择需要安装的服务，服务可以根据依赖关系自动选择

- IP规划与进程部署

配置管理IP与业务IP，进程部署一般保持默认即可

- 节点信息

根据提示信息填写节点的CPU，内存，磁盘，主机名信息

- 浮动IP

根据提示信息配置OM的浮动IP，以及Hue、DBservice、Loader、Solr等组件的浮动IP（如果没有选择安装这些组件，则不需要配置）

- 磁盘配置

元数据盘与数据盘的数目之和不能超过总的磁盘数(包括OS盘)

- 实例参数配置

配置实例参数，所有实例参数可根据当前配置自动生成，一般保持默认即可全部配置完成之后，在“磁盘配置”页面上，点击“生成配置文件”按钮

# 基础配置

配置	内容	备注
产品类型	HD	LLD可支持HD、Stream等四种解决方案，该篇只选择HD即可
集群名称	Test_01	——
版本号	FusionInsight V100R002C60SPC XXX	版本号不能修改，LLD版本号要和HD安装包版本号匹配，否则安装失败
安装模式	Sec	Sec：安全版本中各组件启用kerberos认证，目前只支持Sec模式。 NoSec：非安全版本中各组件不使用kerberos认证。
软件安装 路径	/opt/huawei/Bigdata	——
数据存放 路径	/srv/BigData	HDFS数据和各组件元数据的存放路径。

# 基础配置

配置	内容	备注
升级模式	<b>normal</b>	后续版本的升级模式，取值为“ <b>normal</b> ”和“ <b>advanced</b> ”： <ul style="list-style-type: none"><li>• <b>normal</b>（推荐）：只需在主备管理节点使用<b>root</b>用户，其他节点使用运行用户（<b>omm</b>）即可完成集群的升级操作。</li><li>• <b>advanced</b>（谨慎使用）：所有节点必须使用<b>root</b>用户进行集群升级操作。使用该模式必须确保集群所有节点均可通过<b>root</b>用户远程登录，且密码一致。</li></ul>
用户名	<b>root</b>	各节点操作系统中统一的用户名，使用 <b>root</b> 安装时，请选择 <b>root</b> ；使用非 <b>root</b> 安装时，请选择 <b>omm</b> ，并且需要在安装 <b>OMS</b> 之前执行安装包中的 <b>preset</b> 工具
<b>NTP</b> 服务器地址	——	多个 <b>NTP</b> 服务器IP之间用“,”分隔，如 <b>192.168.1.1</b> ， <b>192.168.2.1</b> <b>NTP</b> 服务器地址配置为空时，系统默认使用主管理节点作为 <b>NTP</b> 服务器
第三方 <b>SSO Server</b> 的IP地址	——	当采用单点登录时，此处填写第三方 <b>SSO Server</b> 的IP地址、端口

# 基础配置

配置	内容	备注
OS类型	suse-11.3	预安装时根据OS类型补齐相应版本的RPM包
OS镜像挂载目录	/media/	——
配置套餐	套餐3: MN&CN&DN	——
是否自定义套餐	No	如果默认的三种部署套餐无法满足用户需求，可以选择Yes，然后在项目SA/SE指导下进行集群部署规划（自定义的情况下，配置套餐无法变更）
集群节点数量	3	——
数据分区Raid类型	nonraid	——
元数据分区Raid类型	nonraid	——

# 基础配置

配置	内容	备注
输出配置文件路径	D:\Test	存放本地PC中，可使用WinSCP上传至OMS节点的 安装目录，该路径不能有空格。
配置文件上传路径	—	格式： <b>host:&lt;软件包解压路径 &gt;/FusionInsight/software</b> ，路径中不能有空格。 如： <b>192.168.1.1:/opt/FusionInsight/software</b> 。 可以填写多个路径，用换行分隔，例如： <b>192.168.1.1:/opt/FusionInsight/software</b> <b>192.168.1.2:/opt/FusionInsight/software</b>
时区	(UTC+08:00)北京， 重庆，香港特别行 政区，乌鲁木齐	

# IP规划与进程部署

类型		机架名称		管理IP		业务IP		OMSSe rver	LdapSer ver	KrbServer		ZooKe eper
								OMSSe rver	SlapdSe rver	Kerberos Server	KerberosA dmin	quoru mpeer
MN&CN& DN		/default/rack 0		192.168.1.10 1		192.168.2.1 01		Y	Y	Y	Y	Y
MN&CN& DN		/default/rack 0		192.168.1.10 2		192.168.2.1 02		Y	Y	Y	Y	Y
CN&DN		/default/rack		192.168.1.10		192.168.2.1						Y
HDFS				Yarn		Mapreduce		DBServ ice	Hu e	Loader	Spark	
Zkf c	Name Node	DataN ode	JournalN ode	NodeMa nager	ResourceMa nager	JobHistorySe rver	DBServ er	Hu e	LoaderSer ver	JDBCSe rver	JobHisto ry	SparkRes ource
		Y	Y	Y			Y	Y	Y			Y
Y	Y	Y	Y	Y	Y				Y	Y	Y	Y
Y	Y	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y



# IP规划与进程部署

Hive			HBase			SmallFS	FTP-Server	Flume		Kafka	Streaming			
MetaStore	WebHcat	HiveServer	HMaster	RegionServer	ThriftServer	FGCServer	FTP-Server	Flume	MonitorServer	Broker	Nimbus	Supervisor	Logviewer	UI
Y		Y		Y	Y	Y		Y	Y	Y	Y	Y	Y	Y
Y		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Y		Y	Y	Y		Y	Y		Y		Y	Y	

Metadata	Oozie	Solr							Redis
MetadataServer	oozie	SolrServerAdmin	HBaseIndexer	SolrServer1	SolrServer2	SolrServer3	SolrServer4	SolrServer5	Redis
		Y	Y	Y	Y	Y	Y	Y	Y
	Y	Y	Y	Y	Y	Y	Y	Y	Y
Y	Y		Y	Y	Y	Y	Y	Y	Y

# 节点信息

	节点上CPU的逻辑核数，查看逻辑CPU个数的命令 <code>cat /proc/cpuinfo   grep "processor"   sort -u   wc -l</code>	节点上的物理内存，大小请结合服务的实际部署进行规划 查询方式： <code>free -g</code> 在Mem行total列的数据就是节点的内存大小	节点上raid之后所有磁盘的数量，包括OS盘。 获取方式： <code>fdisk -l   grep "Disk /dev/"   wc -l</code>	除OS盘外，准备使用的节点里磁盘空间最小的磁盘的容量，考虑到算法上的误差，建议这里的配置应不超过实际磁盘大小的90%，如：磁盘大小2000G，这里应配置为1800G。 获取方式： <code>fdisk -l   grep "Disk /dev/"</code> 取除OS盘之外的最小值，然后乘以0.9。	主机名信息为选填信息，用于管理主机名和IP的映射关系，不影响安装结果。只能由数字“0-9”、字母“a-z A-Z”、中划线“-”组成，不能有下列线“_” 获取方式： 执行命令 <code>hostname</code>
管理IP	VCores	内存(G)	主机逻辑磁盘数量	最小数据盘容量(G)	主机名
192.168.1.101	32	128	16	1600	host1
192.168.1.102	32	128	16	1600	host2
192.168.1.103	32	128	16	1600	host3

- 推荐的服务器配置为：**CPU：Vcores 16**以上，内存：**64G**以上，磁盘数：**12**块磁盘以上，数据盘容量：**1000G**以上
- 该界面会随着“基础配置”界面的节点数量的变化自动增删节点，不用在该页面中手工增删节点；
- 新增节点的默认配置以第一个节点的内容为模板，只需手工修改个别配置即可；

# 浮动IP

类型	配置项	浮动IP	接口	子网掩码	网关	配置说明
OMSServer	om_float_ip	192.168.3.100	bond0:oms	255.255.0.0	192.168.0.1	配置为管理平面使用的IP地址；
OMWebService	ws_float_ip	192.168.3.101	bond0:web	255.255.0.0	192.168.0.1	配置为用户可访问的IP地址，可以是管理平面的地址，可以和OMSServer配置相同； 注意事项： 当和OMSServer配置为相同的浮动IP时，接口、掩码、网关也必须一样； 当和OMSServer配置为不同的浮动IP时，接口要配置为不同的名称；
LdapServer	LDAP_FLOAT_IP	192.168.3.102			192.168.0.1	需要配置LdapServer的浮动IP，配置为管理平面IP地址、网关，必须填写；
DBService	dbservice.floatip	192.168.3.103			192.168.0.1	DBService的浮动IP，配置为业务平面IP地址，在没有安装DBService组件时，该配置项不会生效，可不填写；

## 浮动IP(续)

类型	配置项	浮动IP	接口	子网掩码	网关	配置说明
Hue	HUE_FLOAT_IP	192.168.3.104	-	-	-	<b>Hue</b> 的浮动IP，配置为业务平面IP地址，在没有安装 <b>Hue</b> 组件时，该配置项不会生效，可不填写；
Loader	loader.float.ip	192.168.3.105	-	-	-	<b>Loader</b> 的浮动IP，配置为业务平面IP地址，在没有安装 <b>Loader</b> 组件时，该配置项不会生效，可不填写；
Solr	SOLRHA_SERVER_FLOAT_IP	192.168.3.106	-	-	-	<b>Solr</b> 的 <b>HAServer</b> 的浮动IP，配置为业务平面IP地址，在没有安装 <b>Solr</b> 组件时，该配置项不会生效，可不填写；

# 磁盘配置—OS盘

<b>OS盘（单位：GB）</b> <b>建议OS盘容量大于600GB</b>						
1.要求用户按照标明的 <b>RAID</b> 方式进行服务器 <b>RAID</b> 配置。 2.为保持系统的可靠运行，推荐按要求配置 <b>RAID1</b> ，如不做 <b>RAID1</b> 配置也可正常安装集群。						
操作系统根分区，包括了所有目录，最少要求10G	临时文件存放目录，最少要求10G	操作系统运行时目录，最少要求10G	日志存放目录，需根据集群规模和日志保留时间来配置 以200节点集群为例进行说明： 200节点集群需要保留15天日志，建议容量不小于200G 200节点集群需要保留30天日志，建议容量不小于400G	数据存放路径所在分区，大小必须满足配置要求，最少要求60G	软件安装路径，该路径所在分区大小必须满足配置要求，建议不小于300G	
管理IP	/	/tmp	/var	/var/log	/srv/BigData	/opt/huawei/Bigdata
192.168.1.101	10	10	10	50	60	300
192.168.1.102	10	10	10	50	60	300
192.168.1.103	10	10	10	50	60	300

- OS盘上的分区是OS安装时创建的，集群安装过程（Preinstall）中不会创建这些分区，只对这些分区大小进行检查；
- 软件安装路径不一定是分区，只需要安装路径所在分区满足配置大小即可，如：  
/opt/huawei/Bigdata所在的分区可以是/opt，只要/opt分区大于300G即可

# 磁盘配置—元数据盘

管理IP	管理元数据盘						
	元数据盘数	元数据分区					
		nonraid					
192.168.1.101	6	/srv/BigData/str	/srv/BigData/beam	/srv/BigData/data_service	/srv/BigData/zookeeper	/srv/BigData/zo	/srv/BigData/db
192.168.1.102	6	/srv/BigData/str	/srv/BigData/beam	/srv/BigData/zookeeper	/srv/BigData/na	/srv/BigData/na	/srv/BigData/db
192.168.1.103	4	/srv/BigData/db	/srv/BigData/data_service	/srv/BigData/zookeeper	/srv/BigData/na	/srv/BigData/na	/srv/BigData/jo

- 元数据盘一共有**7**种元数据分区，均衡分配在控制节点上，建议使用**Raid1**；
- 配置规划工具默认是各元数据分区独占磁盘，可以减少元数据盘数，这时元数据盘需要并区（多个元数据分区放在一块磁盘上，生产环境中不推荐）
- 元数据分区的规划和安装的服务有关，如果没有安装相应的服务，配置规划工具不会规划相应的元数据分区；
- 并区场景下，元数据分区的大小是根据分区最小要求的比例进行分配；

# 磁盘配置—数据盘

	HDFS		Streaming		Kafka		Solr		Redis	
	数据盘数	数据分区	数据盘数	数据分区	数据盘数	数据分区	数据盘数	数据分区	数据盘数	数据分区
管理IP		nonraid		nonraid		nonraid	Stored_On_HDFS	nonraid		nonraid
192.168.1.101	6	/srv/BigData/ha doop/data1~6	1	/srv/BigData/ 1streaming_da ta	1	/srv/BigData/k afka/data1~1	0		1	/srv/BigData 1/redis/Redis _1~30
192.168.1.102	6	/srv/BigData/ha doop/data1~6	1	/srv/BigData/ 1streaming_da ta	1	/srv/BigData/k afka/data1~1	0		1	/srv/BigData 1/redis/Redis _1~30
192.168.1.103	8	/srv/BigData/ha doop/data1~8	1	/srv/BigData/ 1streaming_da ta	1	/srv/BigData/k afka/data1~1	0		1	/srv/BigData 1/redis/Redis _1~30

- **HD**产品中数据盘一共有**5**种，存放在数据节点上，由于数据存放时，有多个副本，建议使用**Raid0**，提高数据读写性能；
- 每个数据分区必须独立占用一块磁盘，不能和其他数据分区并区（**Redis**除外，**Redis**是多个**Redis**分区共用磁盘）；
- 在**Preinstall**时会创建这些分区，如果节点磁盘数和实际不符，**Preinstall**会失败；

# 实例参数配置

该界面默认无需配置，如需配置请在项目SA/SE指导下进行。

	HDFS			Yarn						Kafka
	DataNode		NameNode	NodeManager					Resource Manager	Broker
管理IP	dfs.datanode.data.dir	GC_OPTS	GC_OPTS	yarn.node.manager.resource.memory-mb	yarn.nodemanager.resource.cpu-vcores	yarn.nodemanager.local-dirs	yarn.nodemanager.log-dirs	RES_CPU_PERCENTAGE	yarn.scheduler.maximum-allocation-mb	log.dirs
192.168.1.101	/srv/BigData/hadoop/data1~6/dn	-Xms2G -Xmx4G		98304	48	/srv/BigData/hadoop/data1~6/nm/local-dir	/srv/BigData/hadoop/data1~6/nm/containerlogs	100		/srv/BigData/kafka/data1~1/kafka-logs
192.168.1.102	/srv/BigData/hadoop/data1~6/dn	-Xms2G -Xmx4G	-Xms2G -Xmx4G	98304	48	/srv/BigData/hadoop/data1~6/nm/local-dir	/srv/BigData/hadoop/data1~6/nm/containerlogs	100	6144	/srv/BigData/kafka/data1~1/kafka-logs
192.168.1.103	/srv/BigData/hadoop/data1~8/dn	-Xms2G -Xmx4G	-Xms2G -Xmx4G	98304	48	/srv/BigData/hadoop/data1~8/nm/local-dir	/srv/BigData/hadoop/data1~8/nm/containerlogs	100	6144	/srv/BigData/kafka/data1~1/kafka-logs





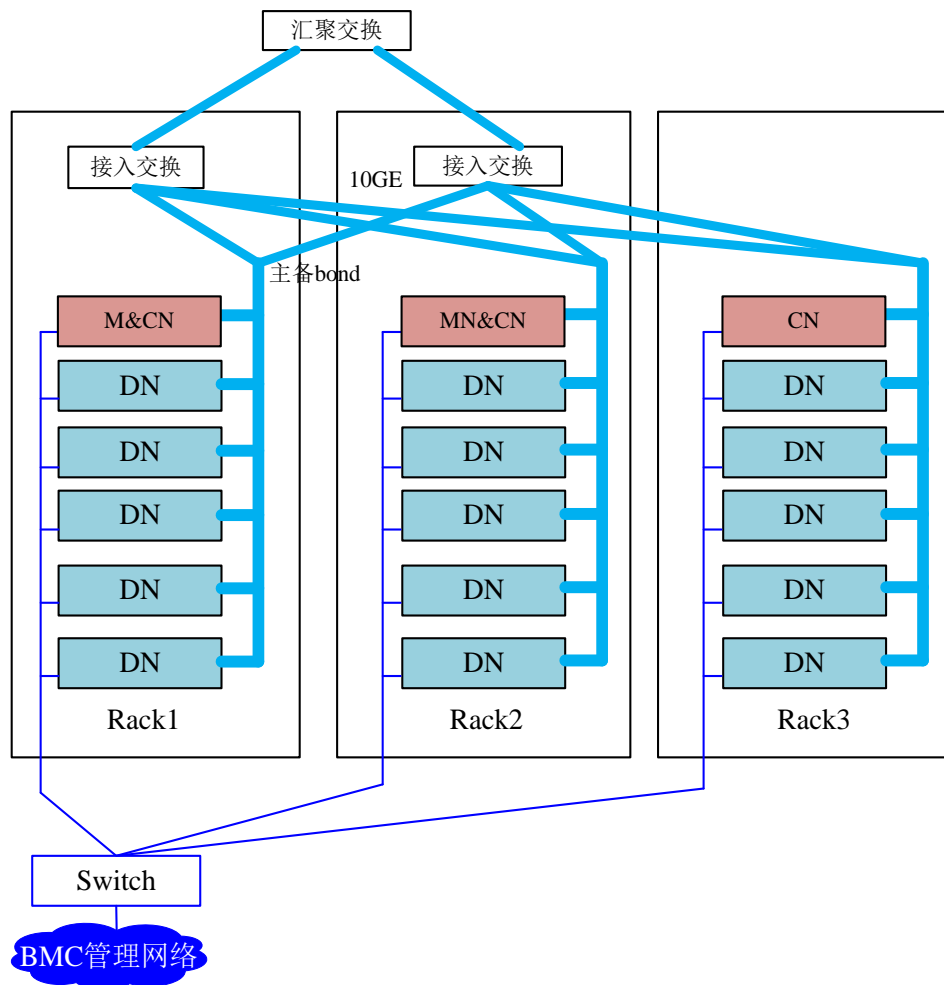
# 目录

1. 简介
2. 集群组网设计
3. 集群节点部署规划
4. 集群磁盘规划
5. 配置规划工具使用
6. 设计案例

# 需求

- **XX**银行历史明细项目
  - 用途：历史交易明细查询
  - 数据量：**20 TB~40TB**
  - 节点数：**20**
  - 可靠性与性能优先

# 组网设计



## 组网要点:

- 集群内二层交换，及集群节点在一个子网里
- 接入交换间无堆叠（硬件约束）
- 主机主备**10GE bond**，主备子端口分别连接到两交换机
- 所有主机缺省主用端口连到同一接入交换机
- 接入交换机与汇聚交换的级联带宽**40GB**

# 磁盘规划——管理节点&控制节点

Raid	用途	磁盘容量GB	分区目录	分区容量GB	用途
Raid1	OS盘	600	/	10	操作系统根分区，包括所有目录（除去下面指定的目录）
			/tmp	10	临时文件存放目录
			/var	10	操作系统运行时目录
			/var/log	50	日志存放目录
			/srv/BigData	60	作为下面的/srv/BigData/*的母路径。
			/opt	all	程序存放目录
Raid1	管理元数据盘	900	/srv/BigData/dbdata_om	all	OMS的数据目录
Raid1		900	/srv/BigData/dbdata_service	all	组件数据库服务数据目录
Raid1		900	/srv/BigData/zookeeper	all	存放ZK数据

# 磁盘规划——管理节点&控制节点

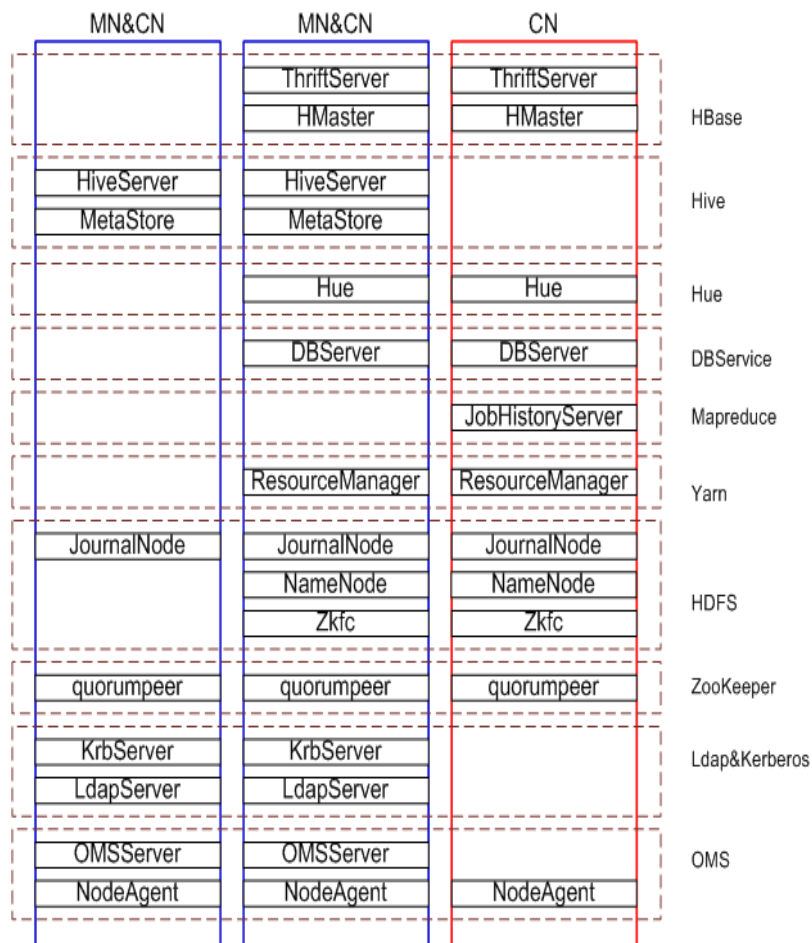
Raid	用途	磁盘容量GB	分区目录	分区容量GB	用途
Raid 1	管理元数据盘	900	/srv/BigData/journalnode	all	存放JournalNode数据
Raid 1		900	/srv/BigData/LocalBackup	all	集群OMS数据、元数据本地备份的数据目录
Raid 1		900	/srv/BigData/namenode	all	存放NameNode数据

# 磁盘规划——数据节点

Raid	用途	磁盘容量GB	分区目录	分区容量GB	用途
Raid 1	OS盘	600	/	10	操作系统根分区，包括了所有目录（除去下面指定的目录）
			/tmp	10	临时文件存放目录
			/var	10	操作系统运行时目录
			/var/log	50	日志存放目录
			/srv/BigData	60	作为下面的/srv/BigData/*的母路径。
			/opt	all	程序存放目录
non-Raid	数据盘	900	/srv/BigData/hadoop/data1~12	all	存放DataNode数据以及MR中间数据

注：单节点采用**900GB\*12**，总HDFS容量= **900GB\*12(盘)\*17(节点)/3(备份) = 60TB**，满足用户使用的存储需求；

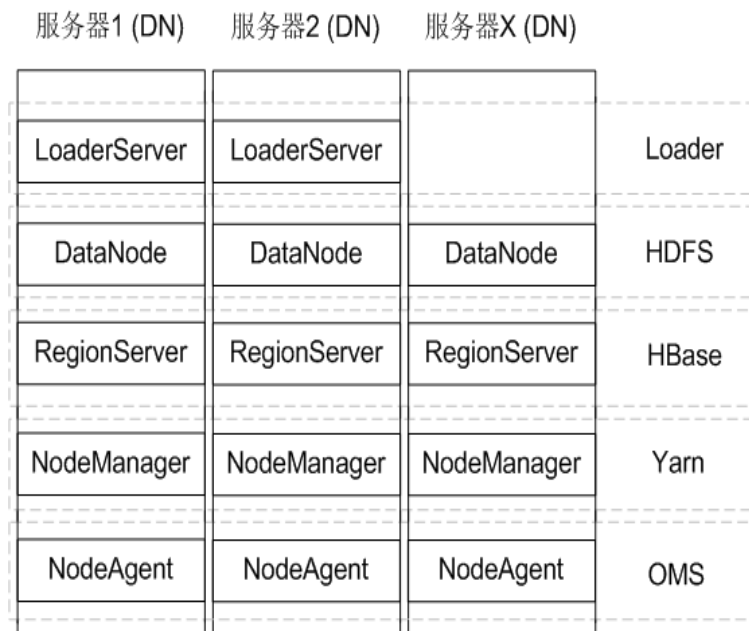
# 进程部署



管理节点&控制节点

## 部署要点:

- 管理节点与控制节点合布
- 数据节点独立部署



数据节点

# 专业术语（一）

组件名称	组件描述
Redis	一个开源的、高性能的 <b>key-value</b> 分布式存储数据库，支持丰富的数据类型，弥补了 <b>memcached</b> 这类 <b>key-value</b> 存储的不足，满足实时的高并发需求。
Solr	一个高性能，基于 <b>Lucene</b> 的全文检索服务器。 <b>Solr</b> 对 <b>Lucene</b> 进行了扩展，提供了比 <b>Lucene</b> 更为丰富的查询语言，同时实现了可配置、可扩展，并对查询性能进行了优化，并且提供了一个完善的功能管理界面，是一款非常优秀的全文检索引擎。
Oozie	提供了对开源 <b>Hadoop</b> 组件的任务编排、执行的功能。以 <b>Java Web</b> 应用程序的形式运行在 <b>Java servlet</b> 容器（如： <b>Tomcat</b> ）中，并使用数据库来存储 工作流定义、当前运行的工作流实例（含实例的状态和变量）。
Metadata	为 <b>FusionInsight HD</b> 数据仓库类型的组件对外提供元数据的抽取能力，并且可以由人工为每个元数据进行标签设定，用于后向的数据分析、搜索等扩展功能。
Streaming	华为 <b>FusionInsight Streaming</b> 是一个基于 <b>Apache Storm</b> 的分布式、实时计算系统。
Kafka	一个分布式的、分区的、多副本的实时消息发布-订阅系统。提供可扩展、高吞吐、低延迟、高可靠的消息分发服务。



## 专业术语（二）

组件名称	组件描述
Flume	一个分布式、可靠和高可用的海量日志聚合系统，支持在系统中定制各类数据发送方，用于收集数据；同时， <b>Flume</b> 提供对数据进行简单处理，并写入各种数据接受方（可定制）的能力。
FTP-Server	通过通用的 <b>FTP</b> 客户端、传输协议提供对 <b>HDFS</b> 文件系统进行基本的操作，例如：文件上传、文件下载、目录查看、目录创建、目录删除、文件权限修改等。
SmallFS	提供小文件后台合并功能，能够自动发现系统中的小文件(通过文件大小阈值判断)，在闲时进行合并，并把元数据外置到第三方 <b>KV</b> ( <b>Key Value</b> ) 系统来降低 <b>NameNode</b> 压力，同时提供新的 <b>FileSystem</b> 接口，让用户能够透明的对这些小文件进行访问。
HBase	提供海量数据存储功能，是一种构建在 <b>HDFS</b> 之上的分布式、面向列的存储系统。
Hive	建立在 <b>Hadoop</b> 基础上的开源的数据仓库，提供类似 <b>SQL</b> 的 <b>Hive QL</b> 语言操作结构化数据存储服务和基本的数据分析服务。
KrbServer	用于网络的身份鉴别，存储钥匙并进行票据发放。
LdapServer	提供用户和用户组的信息存储。

## 专业术语（三）

组件名称	组件描述
<b>Spark</b>	基于内存进行计算的分布式计算框架。
<b>Loader</b>	实现 <b>FusionInsight HD</b> 与关系型数据库、文件系统之间交换数据和文件的数据加载工具；同时提供 <b>REST API</b> 接口，供第三方调度平台调用。
<b>Hue</b>	提供了开源 <b>Hadoop</b> 组件的 <b>WebUI</b> ，可以通过浏览器操作 <b>HDFS</b> 的目录和文件，调用 <b>Oozie</b> 来创建、监控和编排工作流，可操作 <b>Loader</b> 组件，查看 <b>ZooKeeper</b> 集群情况。
<b>DBService</b>	<b>DBService</b> 是一个通用的数据存储组件，它提供了其他组件的数据存储、查询、删除及其他功能。
<b>Mapreduce</b>	提供快速并行处理大量数据的能力，是一种分布式数据处理模式和执行环境。
<b>Yarn</b>	<b>Hadoop 2.0</b> 中的资源管理系统，它是一个通用的资源模块，可以为各类应用程序进行资源管理和调度。
<b>HDFS</b>	<b>Hadoop</b> 分布式文件系统（ <b>Hadoop Distributed File System</b> ），提供高吞吐量的数据访问，适合大规模数据集方面的应用。
<b>ZooKeeper</b>	提供分布式、高可用性的协调服务能力。帮助系统避免单点故障，从而建立可靠的应用程序。

## 思考题

- 1、**FusionInsight** 共有几种节点类型，可分为几个平面？
- 2、管理节点，控制节点，数据节点有哪些部署组合方法？
- 3、每个节点的系统盘分区是如何规划的？



## 习题

- 判断题：

- 1、管理节点，控制节点，数据节点都有元数据盘。
- 2、建议管理节点使用**10GE**，其他节点使用**GE**（根据管理平面业务模型进行评估）
- 3、集群组网可以使用二层组网也可以使用三层组网。
- 4、节点数小于**30**时，可以使用管理节点，控制节点，数据节点合统一部署。

- 单选题

- 1.下列选项不属于节点系统盘分区的是（）

**A./ B./opt C./srv/BigData D./srv/BigData/zookeeper**



## 习题

- 多选题

1.使用**LLD** 生成的配置文件有哪些 ( )

**A .preinstall**工具的配置文件(**preinstall.ini**和**hostX.ini**)

**B .precheck**的配置文件(**checkNodes.config**)

**C .OMS**安装的配置文件(**xx.xx.xx.xx.ini**)

**D .集群安装模板配置**(**installTemplate.xml**)

## 本章总结

- 本章首先介绍了安装**FusionInsight**的要求细节，比如操作系统、浏览器、系统分区要求等。
- 其次介绍了集群组网设计方法，可以使用二层和三层组网。还有集群节点部署方案和磁盘规划。
- 再次介绍了配置工具的使用方法。
- 最后和大家一起学习了一个设计案例。

# Thank you

[www.huawei.com](http://www.huawei.com)