

安装手册 V8.6.2.X



www.gbase.cn

#### GBase 8a MPP Cluster 安装手册, 南大通用数据技术股份有限公司

GBase 版权所有©2004-2019, 保留所有权利。

版权声明

本文档所涉及的软件著作权、版权和知识产权已依法进行了相关注册、登记,由南大通 用数据技术股份有限公司合法拥有,受《中华人民共和国著作权法》、《计算机软件保护条例》、 《知识产权保护条例》和相关国际版权条约、法律、法规以及其它知识产权法律和条约的保 护。未经授权许可,不得非法使用。

免责声明

本文档包含的南大通用公司的版权信息由南大通用公司合法拥有,受法律的保护,南大 通用公司对本文档可能涉及到的非南大通用公司的信息不承担任何责任。在法律允许的范围 内,您可以查阅,并仅能够在《中华人民共和国著作权法》规定的合法范围内复制和打印本 文档。任何单位和个人未经南大通用公司书面授权许可,不得使用、修改、再发布本文档的 任何部分和内容,否则将视为侵权,南大通用公司具有依法追究其责任的权利。

本文档中包含的信息如有更新, 恕不另行通知。您对本文档的任何问题, 可直接向南大 通用数据技术股份有限公司告知或查询。

未经本公司明确授予的任何权利均予保留。

通讯方式

南大通用数据技术股份有限公司

天津华苑产业区海泰发展六道6号海泰绿色产业基地J座(300384)

电话: 400-013-9696 邮箱: info@gbase.cn

商标声明

**GBASE**是南大通用数据技术股份有限公司向中华人民共和国国家商标局申请注册的注册商标,注册商标专用权由南大通用公司合法拥有,受法律保护。未经南大通用公司书面许可,任何单位及个人不得以任何方式或理由对该商标的任何部分进行使用、复制、修改、传播、抄录或与其它产品捆绑使用销售。凡侵犯南大通用公司商标权的,南大通用公司将依法追究其法律责任。



目录

前言	<u>.</u>	
	手册简グ	▶1
	公约	
1	系统要求	₹3
	1.1	硬件配置 3
	1.2	硬件的物理部署和网络规划3
	1.3	配置磁盘 RAID5
	1.4	安装操作系统 17
	1.5	多网卡绑定设置20
	1.6	操作系统环境检查及配置23
	1.6	.1 各节点操作系统版本保持一致23
	1.6	.2 各节点 root 用户的密码保持一致
	1.6	.3 各节点操作系统中软件包23
	1.6	.4 各节点的端口是否占用25
	1.6	.5 各节点防火墙状态32
	1.6	. 6   各节点 SELINUX 状态 36
	1.6	.7 各节点的系统时间保持一致37
	1.6	.8 各节点 ssh 服务开启状态
	1.6	.9 各节点修改操作系统参数39
		1.6.9.1 内核参数配置40
		1.6.9.2 修改虚存配置 43
		1.6.9.3 修改 I/O 调度方式 44
		1.6.9.4 关闭 transparent_hugepage 参数
		1.6.9.5 磁盘调度策略算法 46
		1.6.9.6 修改内存配置 47
		1.6.9.7 Swap 分区
		1.6.9.8 关闭 CPU 超频 47
	1.6	.10 各节点修改主机名配置文件
2	安装集郡	¥
	2.1	部署限制及建议49
	2.2	获取安装文件 49
	2.3	修改安装参数文件50
	2.4	执行安装54
	2.4	.1 一键安装
	2.5	采集指纹 58



	2.6	获取 license 文件	59
	2.7	导入 license 文件	60
	2.8	获取 license 状态	61
	2.9	配置 distribution	64
	2.10	集群的初始化	65
3	启停与登	禄	67
	3.1	集群服务的启停	67
	3.2	集群登录与退出	67
4	卸载集郡	É	70
	4.1	执行卸载脚本	70
5	升级、回	]退集群	71
	5.1	V8.5.1.2集群升级到 V8.6.X.X 集群	71
	5.1.	1 升级前的准备	71
		5.1.1.1 查看现有集群数据库信息	71
		5.1.1.2 恢复 FEVENTLOG	72
		5.1.1.3 停止集群服务	73
	5.1.	2 具体升级步骤	74
	5.2	V8.6.X.X 集群间版本升级	75
	5.3	从 V8.6.X.X 集群回退到 V8.5.1.2 集群	77
	5.4	V8.6.X.X 集群间版本回退	78
6	增加、凋	心少集群节点	80
	6.1	增加集群 data 节点	80
	6.2	增加集群 coordinator 节点	80
	6.3	减少集群 data 节点	81
	6.4	减少集群 coordinator 节点	82
7	安装C3.		83
	7.1	安装步骤	83
	7.2	安装后的验证	84
8	FAQ		86
	8.1	集群服务不能正常启动	86
	8.2	gcadmin 命令的使用问题	86
	8.3	是否可以指定用户登录的默认数据库	86
	8.4	将 V8.5.1.2 版本集群卸载后安装新集群失败	87
	8.5	add nodes 阶段由于异常导致失败	87

GBASE

# 前言

## 手册简介

GBase 8a MPP Cluster 安装手册(Linux RHEL6)主要为用户详细介绍了集群的安装、卸载的内容。

第一章主要介绍了安装集群的系统要求,包括,集群硬件物理设备的部署 原则、网络规划,多网卡绑定,以及安装集群的硬件配置和操作系统的要求等 相关内容。

第二章主要介绍了如何安装 GBase 8a MPP Cluster,安装后的初始化操作,以及集群的登录与退出。

第三章主要介绍了集群的启停操作、登录、退出的内容

第四章主要介绍了如何卸载 GBase 8a MPP Cluster 的内容。

第五章主要介绍了如何在非集群节点上安装客户端的内容

第六章主要介绍了如何升级、回退集群的内容。

第七章主要介绍了如何增加、减少集群节点的内容。

第八章主要介绍了如何安装 C3 工具的内容。

第九章主要以 FAQ 的形式,为用户解答一些常见的集群安装、使用时遇到的问题,解决办法。

#### 公约

下面的文本约定用于本文档:

约定	说明
加粗字体	表示文档标题
大写英文 (例如: SELECT)	表示GBase 8a MPP Cluster 关键字



约定	说明
等宽字体	表示代码示例
•••	省略号指明被省略的内容。



1 系统要求

### 1.1 硬件配置

建议使用如下配置的机器:

硬件	最低配置	推荐配置
处理器	1×2核 2.0GHz	2×4核 2.0GHz
内存	4GB	128GB或更多
硬盘	SATA 7200 rpm, 100GB	SAS 10k rpm, 1TB
光驱	CD-ROM	CD-ROM

#### 1.2 硬件的物理部署和网络规划

一般来说,在实际项目中,首先需要考虑的就是硬件设备的物理部署,一般包括,机柜,电源,主机,硬盘,网卡,交换机等物理设备,另外,物理设备之间的通讯,也需要依靠网络的部署。因此,建议用户在安装操作系统和 GBase 8a MPP Cluster 产品前,需要考虑如何进行物理设备的部署和网络规划。。

我们以下图为例,说明该如何进行合理的物理部署和网络规划:



在上面的图中,包含两个机柜(机柜的电源要保证各自独立供电),机柜 中摆放 GBase 8a MPP Cluster 产品的节点主机和应用服务器,他们之间的网络 通过交换机进行通讯,为了保证网络的高效运行,实际项目中的业务网络也需 要连接到这两个互备的交换机上。

下面说明硬件物理部署和网络规划的原则:

- 电源的高可用:两个机柜的电源是独立的,互不影响的,这样遵循了 主机供电高可用的原则;
- 交换机高可用:每个机柜上配置了一台交换机,两个机柜上共包含2 台交换机,并且这两台交换机之间是互备的关系,当其中一台交换机 发生故障,另一台交换机立即提供服务,遵循了交换机高可用的原则;
- ▶ 节点机器的高可用性: 图中用虚线标识出来的 20 台主机,用于部署 GBase 8a MPP Cluster 产品。

#### 1.3 配置磁盘 RAID

将本地磁盘配置为 RAID,操作系统会将多块物理磁盘视之为一个大硬盘, 并具有容错及冗余的功能。磁盘 RAID 还可以支持容量扩展,只需要加入新的硬 盘并执行一些简单的命令,系统便可以实时利用这新加的容量。

GBase 8a MPP Cluster 推荐将主机的本地磁盘设置为 RAID5。RAID5 工作 模式,至少需要 3 块完全相同的物理磁盘。

不同厂商的硬件服务器设置略有不同。本手册以型号为 DELL R710/R720 的服务器为例,讲述 RAID5 设置的步骤。

▶ 开机启动系统,如下图:



▶ 准备进入 RAID 设置界面,

当服务器开机启动信息出现如下界面时,按下 CTRL+R。



F2 = System Setup F10 = Lifecycle Controller F11 = BIOS Boot Manager F12 = PXE Boot Two 2.60 GHz Eight-core Processors, Bus Speed:8.00 GT/s, L2/L3 Cache:2 MB/20 MB System running at 2.60 GHz System Memory Size: 128.0 GB, System Memory Speed: 1333 MHz, Voltage: 1.35V Dell Serial ATA AHCI BIOS Version 1.0.2 Copyright (c) 1988-2012 Dell Inc. Port E: TSSTcorp DVD+/-RW SN-208BB Broadcom NetXtreme Ethernet Boot Agent Copyright (C) 2000-2012 Broadcom Corporation All rights reserved. Press Ctrl-S to enter Configuration Menu PowerEdge Expandable RAID Controller BIOS Copyright(c) 2011 LSI Corporation Press (Ctrl>(R) to Run Configuration Utility HA -0 (Bus 3 Dev 0) PERC H710P Mini FW package: 21.0.2-0001 -

之后会出现如下界面,表示成功进入 RAID 设置界面:

依照下图,首先需要为用户介绍几个名词解释:

- 1、 Disk Group: 磁盘组, 这里相当于阵列。
- 2、 VD(Virtual Disk): 虚拟磁盘, 虚拟磁盘可以不使用阵列的全部容量, 也就是说一个磁盘组可以分为多个 VD。
- 3、 PD(Physical Disk): 物理磁盘。



▶ 先删除保留的虚拟磁盘,为重新创建做准备。



选择要删除的虚拟磁盘,按下F2,出现如下界面:



使用上下键选择"Delete VD",如下界面:



按回车确认,出现如下界面:





使用左右箭头按键选择"YES"按钮,按回车,删除成功,出现如下界面:

PERC H710P Mini BIDS Configuration Utility 4.00-0014					
UD Mgmt PD Mgmt Ctrl Mgmt Properties					
Uirtual Disk Management           [-] PERC H710P Mini (Bus 0x03, Dev 0x00)           [-] Disk Group: 0, RAID 0           [-] Uirtual Disk           [-] Uirtual Disks           [-] Uirtual Disks           [] [] [] 0, 837.75 GB           [] [] [] 0, 837.75 GB					
-[+] Physical Disks -[+] Total Free Capacity: 0.00 GB - Hot spares -[-] Unconfigured Physical Disks - 00:01:01: Ready: 837.75 GB	Uperation: None CacheCade Capable: Yes Jisk Group 0: Virtual Disks: 1				
<ul> <li>00:01:02: Ready: 837.75 GB</li> <li>00:01:03: Ready: 837.75 GB</li> <li>00:01:04: Ready: 837.75 GB</li> <li>00:01:05: Ready: 837.75 GB</li> <li>00:01:06: Ready: 837.75 GB</li> <li>00:01:06: Ready: 837.75 GB</li> <li>00:01:06: Ready: 837.75 GB</li> </ul>	Physical Disks: 1 Free Cap.: 0.00 GB Free Areas: 0				
F1-Help F2-Openations F5-Refresh Ctrl-N-Next Page Ctrl	P-Ppeu Page F12-Ct lp				

▶ 成功删除保留的虚拟磁盘后,需要重新创建虚拟磁盘

把光标放在界面中蓝色选中的位置:

![](_page_12_Picture_0.jpeg)

PERC H710P Mini BIOS Configuration Utility 4.00-0014 UD Mgmt PD Mgmt Ctrl Mgmt Properties			
Uirtual Disk Management — [-] PERC H710P Mini (Bus 0x03, Dev 0x00) [-] Disk Group: 0, RAID 0 [-] Uirtual Disks	Controller: Disk Groups: 1		
L ID: 0, 837.75 GB -[+] Physical Disks -[+] Total Free Capacity: 0.00 GB Hot swares	Virtual Disks: 1 Physical Disks: 8		
-[-] Unconfigured Physical Disks - 00:01:01: Ready: 837.75 GB - 00:01:02: Ready: 837.75 GB			
└── 00:01:07: Ready: 837.75 GB			
F1-Help F2-Operations F5-Refresh Ctrl-N-Next Page Ctrl	-P-Prev Page F12-Ctlr		

按下F2, 弹出如下界面:

PERC H710P Mini BIOS Configuration Utility 4.00-0014 UD Mgmt PD Mgmt Ctrl Mgmt Properties					
Virtual Disk Management					
L-J PERC H710P Mini (Bus 0x03, Dev 0x00) L-[-] Disk Group: 0. BAID 0					
-[-] Virtual Disks - ID: 0, 837.75 GB	Clear Config				
-I+J Physical Disks -[+] Total Free Capacity: 0.00 GB - Hot suares	Foreign Config				
L-1 Unconfigured Physical Disks 00:01:01: Ready: 837.75 GB	Manage Preserved Cache				
	Security Key Management				
	treate CacheCade Virtual Disk				
00:01:07: Ready: 837.75 GB					
F1-Help F2-Operations F5-Refresh Ctrl-N-	Next Page Ctrl-P-Prev Page F12-Ctlr				

选择"Create New VD",按回车,出现如下界面:

![](_page_13_Picture_0.jpeg)

PERC H710P Mini BIDS Configuration Utility 4.00-0014 UD Mamt PD Mamt Ctrl Mamt Properties					
Uirtual Disk Management					
	- Create New VD				
RAID Level : RAID-0	Basic Settings VD Size: GB				
PD per Span :N/A	UD Name:				
🖵 Physical Disks ————	[] Advanced Settings				
Disk ID Size #	Strip OK				
[ 100:01:01 837.75 GB -	▲ Element Size: 64KB				
L J00:01:02 837.75 GB					
[ 100:01:03 037.75 GB -	Read Policy : Adaptive R CANCEL				
[ ]00:01:05 837.75 GB					
[ ]00:01:06 837.75 GB					
	Write Policy: Write Back				
Secure VD:	[ ] Force WB with no battery				
	[] Initialize				
No	[ ] Configure HotSpare				
F1-Holp F12-Ctly					
ri-neip riz-cuir					

弹出"选择 RAID Level"菜单项,如下界面所示:

PERC H710P Mini BIOS Configuration Utility 4.00-0014					
Uirtual Disk Management					
	Create New VD				
RAID Level : RAID-0 RAID-1	UD Size:				
RAID-5 PD per Span : RAID-6	VD Name:				
Physical D RAID-10 Disk ID RAID-50 #	I 1 Advanced Settings OK				
[ ]00:01:01 B [ ]00:01:02 832.25 GB	Element Size: 64KB				
[ 100:01:03 837.75 GB [ 100:01:04 837.75 GB [ 100:01:05 837.75 GB	Read Policy : Adaptive R CANCEL				
L J00:01:06 837.75 GB	Write Policy: Write Back				
Secure VD:	[ ] Force WB with no battery [ ] Initialize				
	L J Configure HotSpare				
F1-Help F12-Ctlr					

在菜单选项中选择 "RAID-5" 方式, 如下界面所示:

![](_page_14_Picture_0.jpeg)

PERC H710P Mini BIDS Configuration Utility 4.00-0014					
Virtual Disk Management					
	Create New VD				
RAID Level : RAID-0 RAID-1	Basic Settings VD Size: GB				
RAID-5 PD per Span : RAID-6	UD Name:				
Physical D Disk ID [ 100:01:01 RAID-50 H B A	I Advanced Settings				
[ 100:01:02 837.75 GB [ 100:01:03 837.75 GB [ 100:01:04 837.75 GB [ 100:01:05 837.75 GB [ 100:01:05 837.75 GB	Read Policy : Adaptive R CANCEL				
1 100.01.00 031.13 00	Write Policy: Write Back				
Secure UD: No	[ ] Force WB with no battery [ ] Initialize [ ] Configure HotSpare				
F1-Help F12-Ctlr					

按回车确认,出现如下界面:

PERC H710P Mini BIOS Configuration Utility 4.00-0014 VD Mgmt PD Mgmt Ctrl Mgmt Properties				
	Irti	ial Disk Hanagement		
		- Basic Settings		
RAID Level : RAID-5		VD Size: GB		
PD per Span 'NZA		VD Name:		
- Physical Disks		-[] Advanced Settings		
Disk ID Size	#	Strin	пк	
[ 100:01:01 837.75 GB	<u> </u>	Element Size: 64KB		
[ 100:01:02 837.75 GB				
[ 100:01:03 837.75 GB				
[ ]00:01:04 837.75 GB		Read Policy : Adaptive R	CANCEL	
[ 100:01:05 837.75 GB	U			
[ ]00:01:06 837.75 GB				
		Write Policy: Write Back		
Secure VD:		[ ] Force WB with no battery		
		[ ] Initialize		
No		[ ] Configure HotSpare		
P				
F1-Help F12-Ctlr				

RAID 设置完毕后,选择 RAID 所使用的磁盘。使用空格键进行选择,图中 红色框中[]内的 X 表示选中的磁盘,每个 RAID5 的 VD Size 总和不能超过 16T; VD Name 为逻辑磁盘的名称,非必填项,用户可以自行决定是否需要填写。如 下界面所示:

![](_page_15_Picture_0.jpeg)

PERC H710P Mini VD Mgmt PD Mgmt Ctrl Mgmt	BIO Pro	5 Configuration Utility 4.00-0 operties	014
-	Uirt	Ial Disk Management —	
		- Bacic Settings	
RAID Level : RAID-5		VD Size: 5026.50 GB	
PD nen Snan 1824		VD Name:	
- Physical Disks		- 1 Advanced Settings	
Disk ID Size	#	Strin	лк
[X100:01:02 837.75 GB	01	Element Size: 64KB	
[X]00:01:03 837.75 GB	02		
[X]00:01:04 837.75 GB	03		
[X]00:01:05 837.75 GB	04	Read Policy : Adaptive R	CANCEL
[X100:01:06 837.75 GB	05		
LXJ00:01:07 837.75 GB	06	Write Policy: Write Back	
Secure UD:		[ ] Force WB with no hatteru	
		[] Initialize	
No		[ ] Configure HotSpare	
F1-Help F12-Ctlr			

按 TAB 键,将光标移至"OK"处,按回车确认,显示如下所示:

PERC H710P Mini BJ VD Mgmt PD Mgmt Ctrl Mgmt F	DS Configuration Utility 4.00-0014 Properties
Uir	tual Disk Management
	Create New VD
	Basic Settings
RAID Level : RAID-5	VD Size: 5026.50 GB
	UD Name I
PD per Span (NZA	VD Hame:
- Physical Disks	[ ] Advanced Settings
Disk ID Size #	Strip OK
[X]00:01:02 837.75 GB 01	▲ Element Size: 64KB
[X]00:01:03 837.75 GB 02	
[X100:01:04 837.75 GB 03	
[X]00:01:05 837.75 GB 04	Read Policy : Adaptive R CANCEL
[X]00:01:06 837.75 GB 05	
[X]00:01:07 837.75 GB 06	
	Write Policy: Write Back
Secure III:	[ ] Force UP with no battery
Secure VD.	[] I Initialize
No	[ ] Configure HotSnare
	L' I COM IGNO HOTOPATO
F1-Help F12-Ctlr	

按回车,继续确认,界面如下:

![](_page_16_Picture_0.jpeg)

PERC H VD Mgmt PD Mgmt C RAID Level : R	Y10P Mini BIDS Configuration Utility 4.06-0014       trl Mgmt Properties       Uirtual Disk Management       Create New UD       Basic Settings       AID-5       UD Size:       5026.50
PD per Span :N/A	It is recommended that all newly
Physical Disk	created logical drives be initialized
Disk ID S	unless you are attempting to recreate
[X100:01:02 83	a previous configuration and recover
[X100:01:04 83	data as initialization is a
[X100:01:04 83	destructive process.
[X100:01:06 83	
[X100:01:06 83	CANCEL
[X100:01:07 83	
Secure VD:	I 1 Configure HotSpare

按回车,确认完成,设置成功。

▶ 初始化磁盘

选择要初始化的磁盘,如下界面:

![](_page_16_Picture_6.jpeg)

按下 F2, 出现如下界面:

![](_page_17_Picture_0.jpeg)

![](_page_17_Picture_2.jpeg)

按下右方向键,选择"Fast Init",如下界面所示:

![](_page_17_Picture_4.jpeg)

按回车确认,出现如下界面:

![](_page_18_Picture_0.jpeg)

![](_page_18_Picture_2.jpeg)

选择"YES",按回车确认,出现如下界面:

![](_page_18_Picture_4.jpeg)

选择"YES",按回车确认,出现如下界面:

![](_page_19_Picture_0.jpeg)

PERC H710P Mini BIOS Configuration Utili VD Mgmt PD Mgmt Ctrl Mgmt Properties	ty 4.00-0014
[-] PERC H710P Mini (Bus 0x03, Dev 0x00) -[-] Disk Group: 0, RAID 0 -[-] Uirtual Disks -[-] Uirtual Disks -[+] Physica -[+] Thetal F	Uirtual Disk 0: State: Optimal RAID Level: 0 ion: None
Hot spare -[-] Disk Group -[-] Uirtual -[-] Physica -[+] Total F Hot spare -[-] OK OK	roup 0: 1 Disks: 1 al Disks: 1 ap.: 0.00 GB reas: 0

按回车,完成磁盘初始化工作。

▶ 保存并退出 RAID 设置:

![](_page_19_Picture_5.jpeg)

选择"OK",按回车,界面如下:

![](_page_20_Picture_1.jpeg)

![](_page_20_Picture_2.jpeg)

按下 CTRL+Alt+Delete, 重启系统。

至此, RAID5 设置成功。

# 1.4 安装操作系统

GBase 8a MPP Cluster 产品,提供了Linux 平台的安装包文件,支持的Linux 版本如下:

操作系统名称	版本	架构
Red Hat Linux	6. 1 <sup>~</sup> 6. 8, 7. 0 <sup>~</sup> 7. 3	X86_64bit
Cent OS	6. 1 <sup>~</sup> 6. 8, 7. 0 <sup>~</sup> 7. 3	X86_64bit
SUSE	11	X86_64bit
中标麒麟	6.0	X86_64bit

对于操作系统的安装要求和建议如下:

- ▶ 建议安装系统时,选择"软件开发工作站"方式安装;
- ▶ 安装操作系统后,需要确认是否能够正常执行 killall 命令,以保证 节点替换和扩容操作能够正常执行;

![](_page_21_Picture_1.jpeg)

- ▶ 针对 RHEL6. x 建议在安装系统或格式化硬盘时,按 ext4 格式进行格式 化; 针对 RHEL7 系统建议按 xfs 格式进行格式化;
- ▶ 检查 CPU 情况:

可以使用 grep 命令检查 CPU 信息如下:

# grep 'model	name'/proc/cpuinfo		
model name	: Intel(R) Core(TM)2 Duo CPU	E7500	@ 2.93GHz
model name	: Intel(R) Core(TM)2 Duo CPU	E7500	@ 2.93GHz

▶ 检查内存大小

可以使用 grep 命令检查内存信息如下:

▶ 检查硬盘

安装系统时,需要把划分的逻辑卷 mount 到 /opt 目录,因为 GBase 8a MPP Cluster 默认会安装在这个目录下,需要使用的空间比较大。

可以使用 df 命令检查硬盘信息, 显示信息如下:

# df -h					
Filesystem	Size	Used	Avail	Use% Mounted on	
/dev/mapper/vg_lv_r	oot				
	50G	6.6G	41G	a 15% /	
tmpfs	32G	44M	32G	1% /dev/shm	
/dev/sda1	485M	37M	423M	1 8% /boot	
/dev/sdb1	3.OT	743G	2.1T	27% /opt	

注意: /opt 的 mount 配置需要写在/etc/fstab 中而不是

/etc/rc.d/rc.local中,这样可以确保在 8a 集群启动前/opt 完成 mount,否则有可能导致 8a 集群启动失败。

具体配置方式如下:

假设要 mount 的分区为/dev/sdb1, 如果是 ext4 文件系统, 在/etc/fstab 文件尾部增加:

![](_page_22_Picture_0.jpeg)

/dev/sdb1	/opt	ext4	defaults, noatime,
nodiratime, nobarrier	0 0		

如果时 xfs 文件系统, 在/etc/fstab 文件尾部增加:

/dev/sdb1 /opt xfs defaults, noatime, n odiratime, nobarrier 0 0

▶ 要求安装集群的各节点在同一网段。

按如下方法修改网卡的配置文件(此处为单网卡的设置,本手册<u>1.5</u>章节 讲述多网卡绑定的方法):

#	vi /etc/sysconfig/network-scripts/ifcfg-eth0
D	EVICE=eth0
H	WADDR=00:0C:29:2B:93:C1
N	M_CONTROLLED=yes
01	NBOOT=yes
I	PADDR=192. 168. 8. 1
В	00TPR0T0=none
N	ETMASK=255. 255. 0. 0
T	YPE=Ethernet
IJ	PV6INIT=no
U	SERCTL=no

确认 IP 设置是否正确:

# servic	e network restart
# ifconf	ig
eth0	Link encap:Ethernet HWaddr 00:0C:29:2B:93:C1
	inet addr:192.168.8.1 Bcast:192.168.255.255 Mask:255.255.0.0
	inet6 addr: fe80::20c:29ff:fe2b:93c1/64 Scope:Link
	UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
	RX packets:38240 errors:0 dropped:0 overruns:0 frame:0
	TX packets:341 errors:0 dropped:0 overruns:0 carrier:0
	collisions:0 txqueuelen:1000
	RX bytes:3568806 (3.4 MiB) TX bytes:41599 (40.6 KiB)

### 1.5 多网卡绑定设置

网卡的绑定和操作系统有关,因此,本手册按照 Redhat6.2 操作系统为例, 说明一个双网卡绑定的案例,多网卡绑定与此相似。

GBase 8a 产品支持外网和内网的方式,所谓外网,是可以供应用访问的网络,一般使用千兆网络即可;内网是用于 GBase 8a 产品内部数据交换和信息发送使用的网络,一般使用万兆或者以上的网络;外网和内网同时也可以组成一个网络。手册中的示例是假设主机中包含四块网卡,其中两块绑定,用于外网;其余两块绑定,用于内网。

注意: 在进行多网卡设置时,每个网卡的 IP 地址都需要配置在不同的网段,以避免节点之间同步出现问题。

以下为网卡绑定的方式:

- ▶ 识别网卡名称和对应的设备位置。
- ➢ 安装操作系统完毕后,操作系统会识别出主机的网卡名称,例如一台 主机包含两块网卡,接□名称分别为 eth1, eth2。
- ▶ 使用 ethtool -p ethl 命令查看相应设备名称对应的设备位置,执行 该命令后,与 ethl 相对应的网卡接□旁边的指示灯将会闪烁,这样就 能查看 ethl 网卡对应的位置;其余网卡位置的判定同此。
- ▶ 确定绑定方式:

eth1 和 eth2 网卡绑定为 bond0。

➤ 进行 bond0 的配置

切换目录到/etc/sysconfig/network-scripts下,使用 root 用户进行如下操作;

新建 ifcfg-bond0 文件,编写内容如下:

# cd /etc/sysconfig/network-scripts

# vi ifcfg-bond0

![](_page_24_Picture_0.jpeg)

DEVICE=bond0 ONB00T=yes BONDING\_0PTS="mode=1 miimon=100" BO0TPR0T0=static TYPE=Ethernet USERCTL=no IPADDR=192.168.103.88 NETMASK=255.255.0 NM\_CONTROLLED=no

修改 ifcfg-eth1 文件,修改内容如下:

<pre># cd /etc/sysconfig/network-scripts</pre>
# vi ifcfg-eth1
DEVICE=eth1
ONBOOT=yes
BOOTPROTO=none
TYPE=Ethernet
MASTER=bond0
SLAVE=yes
USERCTL=no
NM_CONTROLLED=no

修改 ifcfg-eth2 文件,修改内容如下:

<pre># cd /etc/sysconfig/network-scripts</pre>
# vi ifcfg-eth2
DEVICE=eth2
ONBOOT=yes
B00TPR0T0=none
TYPE=Ethernet
MASTER=bond0
SLAVE=yes
USERCTL=no
NM CONTROLLED=no

![](_page_25_Picture_1.jpeg)

▶ 加载绑定模块

在/etc/modprobe.d/dist.conf这个文件的末尾加入如下内容:

alias bond0 bonding

▶ 将绑定设置为开机自动加载方式

在/etc/rc.d/rc.local 这个文件的倒数第2行加入如下内容:

ifenslave bond0 eth1 eth2

ifup bond0

▶ 重启操作系统,验证配置正确性

重启操作系统后,在 root 用户下使用 if config 命令验证配置是否正确, 验证如下情况:

A) bond0 虚拟网卡能被操作系统识别,并且配置的 IP 地址正确无误。

B) eth1, eth2 网卡不显示 IP 地址。

C) bond0的 MAC 地址与 eth1, eth2 的其中一个相同 (bond0 与哪个网卡的 MAC 地址相同,则那个网卡为现在启用的网卡)。

使用 if config 命令验证上面三点:

# ifconfi	# ifconfig									
bond0	Link encap:Ethernet HWaddr 90:B1:1C:2F:FB:BF									
	inet addr:192.168.103.88 Bcast:192.168.255.255									
Mask:255.255.	255. 0									
	inet6 addr: fe80::92b1:1cff:fe2f:fbbf/64 Scope:Link									
	UP BROADCAST RUNNING MASTER MULTICAST MTU:1500 Metric:1									
	RX packets:735202 errors:0 dropped:0 overruns:0 frame:0									
	TX packets:13 errors:0 dropped:0 overruns:0 carrier:0									
	collisions:0 txqueuelen:0									
	RX bytes:562057464 (536.0 MiB) TX bytes:1016 (1016.0 b)									
eth1	Link encap:Ethernet HWaddr 90:B1:1C:2F:FB:BF									
	UP BROADCAST RUNNING SLAVE MULTICAST MTU:1500 Metric:1									

![](_page_26_Picture_0.jpeg)

RX packets:735202 errors:0 dropped:0 overruns:0 frame:0 TX packets:13 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:562057464 (536.0 MiB) TX bytes:1016 (1016.0 b) Interrupt:34

# 1.6 操作系统环境检查及配置

安装 GBase 8a MPP Cluster 产品,需要各个节点符合如下要求,否则集群 将不能正确安装。

# 1.6.1 各节点操作系统版本保持一致

请检查各节点中操作系统版本的一致性,查看 Linux 系统版本的命令如下:

# lsb_releas	e —a					
LSB						
Version: :core	e-4.0-amd64:core-4.0-noarch:graphics-4.0-amd64:graphics-4.0-noa					
rch:printing-4.0-amd64:printing-4.0-noarch						
Distributor	ID: RedHatEnterpriseServer					
Description:	Red Hat Enterprise Linux Server release 6.2 (Santiago)					
Release:	6.2					
Codename:	Santiago					

# 1.6.2 各节点 root 用户的密码保持一致

请检查各节点中 root 用户的密码的一致性,请根据实际情况进行修改。

# 1.6.3 各节点操作系统中软件包

![](_page_27_Picture_1.jpeg)

请使用 root 用户执行 **rpm** -qa | grep 命令检查系统中是否安装了必须的 软件包。GBase8a MPP Cluster 运行时依赖的 RPM 包列表保存在安装包根目录 下的 dependRpms 文件中。

示例:

```
# rpm -qa | grep corosync
```

corosync-1.4.6-2.el6.x86\_64 corosynclib-1.4.6-2.el6.x86\_64

1) Redhat 操作系统需要检查是否安装如下安装包:

libstdc++

cyrus-sas1-lib zlib nspr libidn libuuid ncurses-libs nss-util open1dap libgcc nss glibc nss-softokn-freebl 2) SUSE 操作系统需要检查是否安装如下安装包: libstdc++43 glibc libgcc43 libncurses5 libuuid1 c3-5.1.2-1.noarch

libntlm-1.3-0.6.el7.x86\_64

GBase8a MPP Cluster 安装包会在系统上安装如下 RPM 包:

pexpect

![](_page_28_Picture_0.jpeg)

monit
c3
corosync

corosynclib

要求:

1) 用户环境不能有 corosync 和 corosynclib rpm 包,如果有用户需要手动卸载,或者由安装脚本自动卸载。

2) 用户可以选择自己安装 pexpect、monit 和 c3,或者由安装脚本自动安装。安装包内包含 rpm 版本可能与系统光盘中的 rpm 版本不一致。如果用户对这些 rpm 版本有要求,建议由用户手动安装。

3) 如果用户环境缺少 dependRpms 文件中列举的 rpm 包或者缺少 pexpect、 monit、c3、corosync 和 corosynclib 依赖的 rpm 包,安装将会报错,用户需 要补齐 rpm 包后再重新安装。

4) c3 需要依赖 rsync 和 perl rpm 包。

## 1.6.4 各节点的端口是否占用

需要检查各个节点上,各服务使用的默认端口是否被占用。

在 root 用户下,使用 lsof -i:*prot\_name* 命令检查端口是否被占用,如下所示:

# 1sof -i:5258							
COMMAND PID USER	FD	TYPE DEVICE	SIZE/OFF	NODE NAME			
gclusterd 1791 gbase	14u	IPv4 12511	0t0	TCP *:52	58 (LISTEN)		

执行 1sof -i:5258 命令后,显示结果告诉用户 PID=1719 的进程正在使用 该端口号,如果该端口号与下表列出的端口号有冲突,则可以在 root 用户下, 使用 kill -9 *pid\_value* 命令来终止发生冲突的进程。

GBase 8a MPP Cluster 各服务使用的默认端口如下,如果用户的机器上安装的应用已经占用了下表中的端口号,而用户又不想停用目前已使用的应用,

![](_page_29_Picture_0.jpeg)

则用户安装完 GBase 8a MPP Cluster 产品后,可以修改下表中集群默认使用的端口号,具体方法参见表格下方的说明:

所属段	端口名称	含义	默认值	修改方式
client	port	客户端连接集群服 务默认端口	5258	需要同对外提供服务端 □值一致
gbased	port	集群对外提供服务 端□	5258	要求所有节点一致
client	port	客户端连接单机服 务默认端口	5050	gnode 对外接口,需要和 gcluster_gnode_port — 致
gbased	port	单机对外连接端口	5050	要求所有节点一致
无	recover_mo nit_port	监控收集信息端口	6268	要求所有节点一致
无	SERVER_POR T 同步连接端口		5288	要求所有节点一致
config/ port	port	日志收集工具对外 服务端口	6957	要求所有节点一致
config/ port	port	日志收集工具对外 服务端口	6957	要求所有节点一致
totem	mcastport	gcware 广播端口	5493	要求所有节点一致
	gcluster_p ort	检测 gclusterd 状 态端口	5258	要求和 gcluster 对外服 务端口一致
	gnode_port	检测 gbased 状态端 口	5050	要求和 gnode 对外服务 端口一致
gcware	syncserver _port	检测 syncserver 状 态端□	5288	要求和 synctool 配置端 □一致
	node_ssh_p ort	检测系统是否在线 端□	22	需要和 ssh 服务端口配 置一致
	singlecast port	gcware 持久化一致 性端□	5678	所有节点一致

说明如下:

▶ gcluster 端□ 5258 可以更改,在集群的所有节点上进行如下操作:

# GBASE

 使用 gbase 用户修改/opt/gcluster/config/gbase\_8a\_gcluster.cnf 文件中相关内容,具体修改参数如下:

```
[client]
port=5258
socket=/tmp/gcluster_5258.sock
```

```
[gbased]
socket=/tmp/gcluster_5258.sock
port=5258
```

2、 下面需要将 5258 修改为新的端口号,具体如下;

```
# su - gbase
$ vi /opt/gcluster/config/gbase_8a_gcluster.cnf
[client]
port=5259
socket=/tmp/gcluster_5259.sock
...
[gbased]
...
socket=/tmp/gcluster_5259.sock
...
port=5259
...
```

3、 各个节点上的配置文件修改后,使用 root 用户,重新启动 gcware 服务。

```
# su - gbase
$ su - root
容码:
# service gcware restart
Signaling GCMETARECOVER (gcmetarecover) to terminate: [ OK ]
Waiting for gcmetarecover services to unload: [ OK ]
Signaling GCDATARECOVER (gcdatarecover) to terminate: [ OK ]
Waiting for gcdatarecover services to unload: [ OK ]
Signaling GCSYNC (gc_sync_server) to terminate: [ OK ]
```

![](_page_31_Picture_1.jpeg)

Waiting for gc\_sync\_server services to unload:[ OK ]
Signaling GCLUSTERD to terminate: [ OK ]
Waiting for gclusterd services to unload:.. [ OK ]
Waiting for gbased services to unload:. [ OK ]
Signaling GCWARE (gcware) to terminate: [ OK ]
Waiting for gcware services to unload:[ OK ]
Starting GCWARE (gcwexec): [ OK ]
Starting GCLUSTERD : [ OK ]
Starting GCLUSTERD : [ OK ]
Starting GCMETARECOVER : [ OK ]
Starting GCDATARECOVER : [ OK ]

4、 使用 root 用户, 验证新端口设置成功。

```
$ su - root
密码:
# lsof -i:5259
COMMAND PID USER FD TYPE DEVICE SIZE/OFF NODE NAME
gclusterd 4780 gbase 14u IPv4 23640 0t0 TCP *:5259 (LISTEN)
```

- ▶ gnode 端口 5050 可以更改 ,在集群的所有节点上进行如下操作:
- 使用 gbase 用户修改/opt/gcluster/config/gbase\_8a\_gcluster.cnf
   文件中相关内容,具体修改的参数如下:

[gbased]

```
...
gcluster_gnode_port = 5050
```

. . .

2、 下面需要将 5050 修改为新的端口号,具体如下;

```
# su - gbase
$ vi /opt/gcluster/config/gbase_8a_gcluster.cnf
[gbased]
...
gcluster_gnode_port = 5051
```

![](_page_32_Picture_0.jpeg)

. . .

```
    使用 gbase 用户修改/opt/gnode/config/gbase_8a_gbase.cnf 文件中
相关内容,具体修改的参数如下:
```

```
[client]
port=5050
socket=/tmp/gbase_8a_5050.sock
[gbased]
...
socket=/tmp/gbase_8a_5050.sock
...
port=5050
...
```

4、下面需要将 5050 修改为新的端口号,具体如下;

```
# su - gbase
$ vi /opt/gnode/config/gbase_8a_gbase.cnf
[client]
port=5051
socket=/tmp/gbase_8a_5051.sock
```

```
[gbased]
...
socket=/tmp/gbase_8a_5051.sock
...
port=5051
...
```

5、 各个节点上的配置文件修改后,使用 root 用户,重新启动 gcware 服务。

```
$ su - root
密码:
# service gcware restart
Signaling GCMETARECOVER (gcmetarecover) to terminate: [ 0K ]
```

![](_page_33_Picture_1.jpeg)

Waiting for gcmetarecover services to unload: [ OK ]
Signaling GCDATARECOVER (gcdatarecover) to terminate: [ OK ]
Waiting for gcdatarecover services to unload: [ OK ]
Signaling GCSYNC (gc_sync_server) to terminate: [ OK ]
Waiting for gc_sync_server services to unload:[ OK ]
Signaling GCLUSTERD to terminate: [ OK ]
Waiting for gclusterd services to unload: [ OK ]
Signaling GBASED to terminate: [ OK ]
Waiting for gbased services to unload:. [ OK ]
Signaling GCWARE (gcware) to terminate: [ OK ]
Waiting for gcware services to unload: [ OK ]
Starting GCWARE (gcwexec): [ OK ]
Starting GBASED : [ OK ]
Starting GCLUSTERD : [ OK ]
Starting GCSYNC : [ OK ]
Starting GCMETARECOVER : [ OK ]
Starting GCDATARECOVER : [ OK ]

6、 使用 root 用户, 验证新端口设置成功。

\$ su - root		
密码:		
# lsof -i:5051		
COMMAND PID USER	FD	TYPE DEVICE SIZE/OFF NODE NAME
gbased 3452 gbase	5u	IPv4 17730 0t0 TCP *:ita-agent (LISTEN)

- ▶ syncserver 端口 5288 可以更改,在集群的所有节点上进行如下操作
- 1、 使用 gbase 用户在/opt/gnode/config 下创建 synctool. conf 文件
- 2、 在 synctool. conf 增加如下内容:

```
# su - gbase
$ cd /opt/gnode/config/
$ vi synctool.conf
SERVER_PORT=5287
LOG_PATH=/opt/gcluster/log/gcluster
LOG_LEVEL=0
```

![](_page_34_Picture_0.jpeg)

DISABLE=0 SERVER\_PARALLEL=4

SYNC\_LEVEL=1

其中:

SERVER\_PORT 是端口号,默认 5288,在此填写指定的端口;

LOG\_PATH 是 log 的目录, 默认是/opt/gcluster/log/gcluster

LOG\_LEVEL 是 log 级别, 取值 0, 1, 2, 3, 4 分别代表 ERROR, WARNING, INTERFACE, INFO, DEBUG;

SERVER\_PARALLEL 是 server 端读取 I/0 并行线程数, 默认是 4;

SYNC\_LEVEL 是否同步刷入磁盘,默认是 1。

3、 修改完后,在各个节点机器中,使用 root 用户,重新启动 gcware 服务生效。

\$ su - root

#### 密码:

#### # service gcware restart

Signaling GCMETARECOVER (gcmetarecover) to terminate: [ OK ] Waiting for gcmetarecover services to unload: [ OK ] Signaling GCDATARECOVER (gcdatarecover) to terminate: [ OK ] Waiting for gcdatarecover services to unload: [ OK ] Signaling GCSYNC (gc\_sync\_server) to terminate: [ OK ] Waiting for gc sync server services to unload: [ OK ] Signaling GCLUSTERD to terminate: [ OK ] Waiting for gclusterd services to unload:..[ OK ] Signaling GBASED to terminate: [ OK ] Waiting for gbased services to unload: [ OK ] Signaling GCWARE (gcware) to terminate: [ OK ] Waiting for gcware services to unload: [ OK ] Starting GCWARE (gcwexec): [ OK ] Starting GBASED : [ OK ] Starting GCLUSTERD : [ OK ] Starting GCSYNC : [ OK ] Starting GCMETARECOVER : [ OK ]

![](_page_35_Picture_1.jpeg)

	Starting GCDATARECOVER : [ OK ]										
	4、 使用	roo	t用户,	验证	新端	□是否	设置成功	0			
# 1sof -i:5287											
	COMMAND	PID	USER	FD	TYPE	DEVICE	SIZE/OFF	NODE	NAME		
	gc_sync_s	3876	gbase	4u	IPv4	20103	0t0	TCP	*:5287	(LISTEN)	

# 1.6.5 各节点防火墙状态

以 RHEL6 操作系统为例,执行如下命令查看各节点的防火墙状态:

//显示如下信息表示防火墙是关闭状态:										
# service iptables status										
iptables: Firewall is not running.										
<pre># service ip6tables status</pre>	<pre># service ip6tables status</pre>									
ip6tables: Firewall is not running.										
//显示如下信息表示防火墙是开启状态:	//显示如下信息表示防火墙是开启状态:									
<pre># service iptables status</pre>										
Table: nat										
Chain PREROUTING (policy ACCEPT)										
num target prot opt source	destination									
Chain POSTROUTING (policy ACCEPT)										
num target prot opt source	destination									
1 MASQUERADE tcp 192.168.122.0/24	!192.168.122.0/24	masq								
ports: 1024-65535										
2 MASQUERADE udp 192.168.122.0/24	! <b>192. 168. 122. 0</b> /24	masq								
ports: 1024-65535										
3 MASQUERADE all 192.168.122.0/24	! <b>192.</b> 168. 122. 0/24									
Chain OUTPUT (policy ACCEPT)										
num target prot opt source	destination									
Table: mangle										


Cha	in PREROUTIN	NG (policy	ACCEPT)		
nun	target	prot opt	source	destination	
Cha	in INPUT (po	olicy ACCE	PT)		
nun	ı target	prot opt	source	destination	
Cha	in FORWARD (	(policy AC	CEPT)		
num	target	prot opt	source	destination	
Cha	in OUTPUT (p	policy ACC	EPT)		
num	ı target	prot opt	source	destination	
Cha	in POSTROUTI	LNG (polic	y ACCEPT)		
num	target	prot opt	source	destination	
1	CHECKSUM	udp	0. 0. 0. 0/0	0.0.0.0/0	udp dpt:68
CHECKSU	M fill				
Tab	le: filter				
Cha	in INPUT (po	olicy ACCE	PT)		
num	ı target	prot opt	source	destination	
1	ACCEPT	udp	0. 0. 0. 0/0	0.0.0.0/0	udp dpt:53
2	ACCEPT	tcp	0. 0. 0. 0/0	0.0.0.0/0	tcp dpt:53
3	ACCEPT	udp	0. 0. 0. 0/0	0.0.0.0/0	udp dpt:67
4	ACCEPT	tcp	0. 0. 0. 0/0	0.0.0.0/0	tcp dpt:67
		<i>(</i>	)		
Cha	in FORWARD (	(policy AC	CEPT)		
num	l target	prot opt	source	destination	
1	ACCEPT	all	0. 0. 0. 0/0	192.168.122.0/24	state
RELATED	, ESTABLISHED	)			
2	ACCEPT	all	192. 168. 122. 0/24	0.0.0.0/0	
3	ACCEPT	all	0. 0. 0. 0/0	0.0.0.0/0	
4	REJECT	all	0. 0. 0. 0/0	0.0.0.0/0	
reject-	with icmp-po	ort-unreac	hable		
5	REJECT	all	0. 0. 0. 0/0	0.0.0.0/0	
reject-	with icmp-po	ort-unreac	hable		
Cha	in OUTPUT (p	policy ACC	EPT)		



	num	target	prot opt	source	destination	
	# se:	rvice ip6ta	bles stat	us		
	Table	e: filter				
	Chai	n INPUT (po	licy ACCE	PT)		
	num	target	prot opt	source	destination	
	1	ACCEPT	all	::/0	::/0	state
REL	ATED, I	ESTABLISHED				
	2	ACCEPT	icmpv6	::/0	::/0	
	3	ACCEPT	all	::/0	::/0	
	4	ACCEPT	tcp	::/0	::/0	state NEW
tcp	dpt:	22				
	5	REJECT	all	::/0	::/0	
rej	ect-w	ith icmp6-a	dm-prohib:	ited		
	Chai	n FORWARD (	policy ACC	CEPT)		
	num	target	prot opt	source	destination	
	1	REJECT	all	::/0	::/0	
rej	ect-w	ith icmp6-a	dm-prohib:	ited		
	Chai	n OUTPUT (p	olicy ACCH	EPT)		
	num	target	prot opt	source	destination	

在各节点执行如下命令查看防火墙是否在开机时自动启动:

<pre># chkconfig —]</pre>	list ipta	bles					
iptables	0:关闭	1:关闭	2:关闭	3:关闭	4:关闭	5:关闭	6:关闭
<pre># chkconfig —]</pre>	list ip6t	ables					
ip6tables	0:关闭	1:关闭	2:关闭	3:关闭	4:关闭	5:关闭	6:关闭

建议用户关闭防火墙,具体命令如下:

```
# service iptables stop
iptables: Flushing firewall rules: [ OK ]
iptables: Setting chains to policy ACCEPT: filter [ OK ]
iptables: Unloading modules: [ OK ]
```

## GBASE

```
# chkconfig iptables off
说明: 如果此命令执行无效,可执行如下命令:
# chkconfig iptables off --level 2345
# service ip6tables stop
ip6tables: Flushing firewall rules: [ OK ]
ip6tables: Setting chains to policy ACCEPT: filter [ OK ]
ip6tables: Unloading modules: [ OK ]
# chkconfig ip6tables off
说明: 如果此命令执行无效,可执行如下命令:
# chkconfig ip6tables off --level 2345
注: 针对 RHEL7 操作系统而言,防火墙的相关命令如下:
```

systemctl stop firewalld.service (重启恢复) systemctl disable firewalld.service (永久关闭) systemctl status firewalld.service (查询状态)

如果防火墙为开启状态,则需要为 GBase 8a MPP Cluster 开放如下端口:

端口名称	默认值
port	5258
port	5258
port	5050
port	5050
recover_monit_port	6268
SERVER_PORT	5288
port	6957
port	6957
mcastport	5493
gcluster_port	5258
gnode_port	5050
syncserver_port	5288
node_ssh_port	22
singlecastport	5678



```
需要在每个节点上执行以下命令:
     设置默认规则。
1
# iptables -A INPUT -j DROP
# iptables -A FORWARD -j ACCEPT
2
     开放 gcluster 端口。
# iptables -I INPUT -p tcp -dport 5258 -j ACCEPT
# iptables -I OUTPUT -p tcp --dport 5258 -j ACCEPT
# iptables -I INPUT -p tcp -- sport 5258 -j ACCEPT
# iptables -I OUTPUT -p tcp --sport 5258 -j ACCEPT
3
     开放 gnode 端口。
# iptables -I INPUT -p tcp -- dport 5050 -j ACCEPT
# iptables -I OUTPUT -p tcp --dport 5050 -j ACCEPT
# iptables -I INPUT -p tcp -- sport 5050 -j ACCEPT
# iptables -I OUTPUT -p tcp -- sport 5050 -j ACCEPT
4
     开放 syncserver 端口。
# iptables -I INPUT -p tcp -- dport 5288 -j ACCEPT
# iptables -I OUTPUT -p tcp --dport 5288 -j ACCEPT
# iptables -I INPUT -p tcp -- sport 5288 -j ACCEPT
# iptables -I OUTPUT -p tcp --sport 5288 -j ACCEPT
5
     开放 gcware 组播诵信端口
# iptables -I INPUT -p udp -- dport 5493 -j ACCEPT
# iptables -I OUTPUT -p udp --dport 5493 -j ACCEPT
#iptables -I INPUT -p udp --sport 5493 -j ACCEPT
#iptables -I INPUT -p udp --dport 5493 -j ACCEPT
```

注意:如果端口号因被占用发生变更,请同步修改以上命令中的 sport 或 dport 对应的参数值。

## 1.6.6 各节点 SELINUX 状态

要求各节点禁用 SELINUX, 禁用 SELINUX 的方法如下(需重启系统)。



修改/etc/sysconfig/selinux 文件,将 SELINUX 设置为 disabled:



#### 重启操作系统后,可以使用如下方法确认 SELINUX 是否被禁用:

# sestatus		
SELinux status:	disabled	

## 1.6.7 各节点的系统时间保持一致

要求各节点统一系统时间,推荐使用 ntp 自动同步各节点时间。

以 192.168.181.35 时钟同步服务器为例说明如何同步客户端各节点时间, 以下操作在 root 用户下进行。

1、 检查是否安装 ntp 软件包 (大多数时候 ntp 已经安装好)

```
# rpm -qa | grep ntp
ntpdate-4. 2. 4p8-2. e16. x86_64
ntp-4. 2. 4p8-2. e16. x86_64
```

2、 修改/etc/ntp.conf 文件

注释掉

# vi /etc/ntp.conf

南大通用数据技术股份有限公司



```
# service ntpd start
```

正在启动 ntpd: [确定]

```
# chkconfig —level 345 ntpd on
```

使用上面的2条命令后,首先第一条命令启动ntpd服务,然后第二条命令 设定开机后自动运行。

```
# service ntpd start
关闭 ntpd: [确定]
启动 ntpd: [确定]
```

#### 4、 查看 nptd 的状态

```
# ntpstat
unsynchronised
time server re-starting
polling server every 64 s
```

#### 5、 客户端校准本地时钟

客户端机器不能运行 ntpd 服务,可以使用如下操作,首先停止客户端的 ntpd 服务,然后设定开机后也不会自动运行。



# service ntpd stop

关闭 ntpd: [确定]

# chkconfig ntpd off

#### # ntpdate 192.168.181.35

12 Nov 11:07:28 ntpdate[1809]: no server suitable for synchronization found

当客户端运行 ntpdate 192.168.181.35 后,出现上面的提示信息时,是由 于通常 server 端的 ntpd 服务刚刚运行,通常需要等待 3-5 分钟后才能生效。 因此,等待 5 分钟后,再次运行同步时钟命令,显示结果正常。具体如下:

#### # ntpdate 192.168.181.35

12 Nov 09:50:44 ntpdate[2500]: step time server 192.168.181.35 offset -2392.295293 sec

至此,经过上述步骤,各节点就会和192.168.181.35进行时钟同步了。

## 1.6.8 各节点 ssh 服务开启状态

要求各节点的 ssh 服务是正常开启状态。确认各节点 ssh 服务是否正常开 启的方法为,在 root 用户下执行 chkconfig sshd on 命令,这样,每次开机 后, ssh 服务会自动加载运行。

# chkconfig sshd on

在 root 用户下执行 chkconfig --list sshd 命令, 查看是否将 ssh 服务 开启, 加粗显示为 0N, 代表 ssh 服务开启:

**# chkconfig —list sshd** sshd 0:关闭 1:关闭 **2:启用 3:启用 4:启用 5:启用** 6:关闭

## 1.6.9 各节点修改操作系统参数

【涉及机器】:所有服务器节点



系统安装前,请您使用 root 用户进行如下参数的修改,重启后生效。

#### 1.6.9.1 内核参数配置

Sysct1 是一个允许您改变正在运行中的 Linux 系统的接口。它包含一些 TCP/IP 堆栈和虚拟内存系统的高级选项。

sysctl.conf 文件内调整参数已在安装脚本中自动调整并生效,调整的参数及相关描述如下:

配置参数	值	描述
kernel.core_uses_pid	1	控制核心转储是否附加 PID 的核心文件。用于调试多线程应用程序。
net.ipv4.tcp_syncookies	1	控制 TCPsyncookies 的使用。
net.ipv4.ip_local_reser ved_ports	5050, 525 8, 5288, 6 666	防止数据库需要使用的端口被其他程 序强行占用。需要根据实际修改的端 口号进行配置
vm.vfs_cache_pressure	1024	表示内核回收用于 directory 和 inodecache 内存的倾向;缺省值 100 表示内核将根据 pagecache 和 swapcache,把 directory 和 inodecache 保持在一个合理的百分 比;降低该值低于 100,将导致内核 倾向于保留 directory 和 inodecache;增加该值超过 100,将 导致内核倾向于回收 directory 和 inodecache。
vm.min_free_kbytes	2097152	表示强制 LinuxVM 最低保留多少空闲 内存 (Kbytes)。建议设为物理内存的 1/10 的大小。不超过 4G
vm.swappiness	1	设置为1(缺省是60),代表在有空闲 物理内存的情况下尽量不要使用 swap。
vm.overcommit_memory	0	表示检查是否有足够的内存可用,如 果是,允许分配;如果内存不够,拒



		绝该请求,并返回一个错误给应用程 序。
vm.zone_reclaim_mode	0	当某个 NUMA 节点可用内存不足时:如 果为 0 的话,系统会倾向于从其他节 点分配内存。
net.core.netdev_max_bac klog	262144	每个网络接口接收数据包的速率比内 核处理这些包的速率快时,允许送到 队列的数据包的最大数目
net.core.rmem_default	8388608	接收套接字缓冲区大小的默认值(以 字节为单位)
net.core.rmem_max	16777216	接收套接字缓冲区大小的最大值(以 字节为单位)
net.core.somaxconn	65535	用来限制监听(LISTEN)队列最大数据 包的数量,超过这个数量就会导致链 接超时或者触发重传机制
net.core.wmem_default	8388608	发送套接字缓冲区大小的默认值(以 字节为单位)
net.core.wmem_max	16777216	发送套接字缓冲区大小的最大值(以 字节为单位)
net.ipv4.tcp_fin_timeou t	1	如果套接字由本端要求关闭,这个 参数决定了它保持在 FIN-WAIT-2 状 态的时间。对端可以出错并永远不关 闭连接,甚至意外当机。缺省值是 60 秒
net.ipv4.tcp_max_orphan s	3276800	系统中最多有多少个 TCP 套接字不被 关联到任何一个用户文件句柄上。如 果超过这个数字, 孤儿连接将即刻被 复位并打印出警告信息。这个限制仅 仅是为了防止简单的 DoS 攻击, 不能 过分依靠它或者人为地减小这个值, 更应该增加这个值(如果增加了内存 之后)。
net.ipv4.tcp_max_syn_ba cklog	262144	表示那些尚未收到客户端确认信息的 连接(SYN 消息)队列的长度,默认 为1024,加大队列长度为262144,可 以容纳更多等待连接的网络连接数。

## GBASE

net.ipv4.tcp_max_tw_buc kets	20000	系统在同时所处理的最大 timewait sockets 数目。如果超过此数的话, time-wait socket 会被立即砍除并 且显示警告信息。之所以要设定这个 限制,纯粹为了抵御那些简单的 DoS 攻击,千万不要人为的降低这个限制, 不过,如果网络条件需要比默认值更 多,则可以提高它(或许还要增加内 存)
	94500000	第一个值是内存使用的下限。第二个
	91500000	值是内存压刀模式开始对缓冲区使用
net.1pv4.tcp_mem	0	应用压力的上限。第三个值是内存上 第二十六个
	92700000	限。任这个层次上可以将报义去并,
	0	
net.ipv4.tcp_rmem	4096 87380 4194304	TCP 接收缓冲区,3个字段分别是 min,default,max。Min:为 TCP socket 预留用于接收缓冲的内存数量,即使 在内存出现紧张情况下 TCP socket 都 至少会有这么多数量的内存用于接收 缓冲。
net.ipv4.tcp_timestamps	0	时间戳可以避免序列号的卷绕。一个 1Gbps 的链路肯定会遇到以前用过的 序列号。时间戳能够让内核接受这种 "异常"的数据包。
net.ipv4.tcp_tw_recycle	1	开启 TCP 连接中 TIME-WAIT sockets 的快速回收,默认为 0,表示关闭
net.ipv4.tcp_wmem	4096 16384 4194304	TCP 发送缓冲区,3 个字段分别是 min, default, max。Min:为 TCP socket 预留用于发送缓冲的内存最小值。每 个 TCP socket 都可以使用它。
net.ipv4.tcp_tw_reuses	1	表示是否允许重新应用处于 TIME-WAIT 状态的 socket 用于新的 TCP 连接。
net.ipv4.tcp_sack	1	表示是否启用有选择的应答 (Selective Acknowledgment),这可 以通过有选择地应答乱序接收到的报 文来提高性能(这样可以让发送者只



		发送丢失的报文段);(对于广域网通 信来说)这个选项应该启用,但是这 会增加对 CPU 的占用
net.ipv4.tcp_window_sca ling	1	支持更大的 TCP 窗□. 如果 TCP 窗□ 最大超过 65535(64KB), 必须设置该 数值为 1

#### 1.6.9.2 修改虚存配置

安装数据库之前,请您将 virtual memory 配置为 unlimited,有2种 方法能完成虚存的修改:

#### 【方法一】

通过修改配置文件/etc/security/limits.conf,添加如下两行,操作 系统重启后生效:

- \* soft as unlimited
- \* hard as unlimited

#### 【方法二】

执行以下命令,修改当前系统配置,操作系统重启后需重新设置:

#### # ulimit -H -v unlimited

#### # ulimit -S -v unlimited

- -H 设定资源的硬性限制,也就是管理员所设下的限制。
- -S 设定资源的弹性限制。
- -v 〈虚拟内存大小〉 指定可使用的虚拟内存上限,单位为 KB。



#### 1.6.9.3 修改 I/0 调度方式

安装前请您修改 I/O 调度方式,设置系统使用 deadline 调度器,有 2 种方法能完成 I/O 调度方式的修改:

#### 【方法一】

修改 grub 启动参数,操作系统启动时 deadline 全局生效。编辑 /boot/grub/menu.lst 文件,kernel 参数添加内容"elevator=deadline":

```
# vi /boot/grub/menu.lst
   # grub.conf generated by anaconda
    #
   # Note that you do not have to rerun grub after making changes to this file
   # NOTICE: You have a /boot partition. This means that
               all kernel and initrd paths are relative to /boot/, eg.
    #
               root (hd0, 0)
    #
               kernel /vmlinuz-version ro root=/dev/sda3
    #
               initrd /initrd-[generic-]version.img
    #
   #boot=/dev/sda
    default=0
    timeout=5
   splashimage=(hd0, 0)/grub/splash.xpm.gz
   hiddenmenu
   title Red Hat Enterprise Linux (2.6.32-220.el6.x86 64)
            root (hd0, 0)
            kernel /vmlinuz-2.6.32-220.el6.x86 64 ro
root=UUID=ff665fa7-53b1-48a1-b8b3-0fe90d9abd42 rd_NO_LUKS rd_NO_LVM
LANG=en_US.UTF-8 rd_NO_MD quiet SYSFONT=latarcyrheb-sun16 rhgb crashkernel=auto
KEYBOARDTYPE=pc KEYTABLE=us rd_NO_DM elevator=deadline
            initrd /initramfs-2.6.32-220.el6.x86_64.img
```

【方法二】



执行如下命令,修改当前系统配置,操作系统重启后需重新设置:

# echo deadline > /sys/block/XXX/queue/scheduler

其中 XXX 表示要修改的物理磁盘,一般应设置第二块磁盘 sdb。可使 用如下语句查看修改结果:

# cat /sys/block/sdb/queue/scheduler
noop anticipatory [deadline] cfq

## 1.6.9.4 关闭 transparent\_hugepage 参数

为保证数据库高效运行,需要关闭 transparent\_hugepage 参数,具体操作如下:

【方法一】注:本方法在操作系统重新启动后,需要重新设置。

首先需要查询操作系统版本,由于版本的不同,具体的参数位置名称 会略有不同。

# cat /etc/redhat-release

Red Hat Enterprise Linux Server release 6.2 (Santiago)

关闭 transparent\_hugepage 参数。

# cat /sys/kernel/mm/redhat\_transparent\_hugepage/enabled

[always] never

# echo "never" > /sys/kernel/mm/redhat\_transparent\_hugepage/enabled

# cat /sys/kernel/mm/redhat\_transparent\_hugepage/enabled

always [never]

【方法二】本命令需要重启系统后生效。(此方法生效后,操作系统 再进行重启就不需重新设置本参数了。)

编辑/boot/grub/menu.lst, 在"kenel"行加入"root=LABEL=/

transparent\_huge"关闭 transparent\_hugepage 参数。

# vi /boot/grub/menu.lst



```
# grub.conf generated by anaconda
#
# Note that you do not have to rerun grub after making changes to this file
# NOTICE: You have a /boot partition. This means that
#
           all kernel and initrd paths are relative to /boot/, eg.
           root (hd0, 0)
#
           kernel /vmlinuz-version ro root=/dev/sda3
±
           initrd /initrd-[generic-]version.img
#
#boot=/dev/sda
default=0
timeout=5
splashimage=(hd0, 0)/grub/splash.xpm.gz
hiddenmenu
title Red Hat Enterprise Linux (2.6.32-220.el6.x86_64)
        root (hd0,0)
        kernel /vmlinuz-2.6.32-220.el6.x86_64 ro
root=UUID=8140f03b-d7f4-41eb-9d34-911945cfc013 rd NO LUKS rd NO LVM
LANG=en US. UTF-8 rd NO MD quiet SYSFONT=latarcyrheb-sun16 rhgb crashkernel=auto
KEYBOARDTYPE=pc KEYTABLE=us rd NO DM root=LABEL=/ transparent huge
page=never
        initrd /initramfs-2.6.32-220.el6.x86 64.img
```

重新启动服务器后,查询参数,可以看到 transparent\_hugepage 参数 已经关闭。

# cat /sys/kernel/mm/redhat\_transparent\_hugepage/enabled
always [never]

## 1.6.9.5 磁盘调度策略算法

磁盘调度策略算法不能为CFQ(但包含 swap space, log files, Linux system files 的 drivers 仍然可以使用该方式),建议 deadline (适用于 disk)或者 loop (适用于 SSD)。

例如执行如下命令:

#### echo deadline > /sys/block/sda/queue/scheduler

需要根据/opt/目录所在的实际设备进行设置,由于大部分场景使用 disk,因此建议为 deadline。

#### 1.6.9.6 修改内存配置

安装数据库之前,请您将 max memory size 配置为 unlimited,有 2 种方法能完成内存的修改:

#### 【方法一】

通过修改配置文件/etc/security/limits.conf,添加如下内容,操作 系统重启后生效:

\* rss as unlimited

【方法二】

执行以下命令,修改当前系统配置,操作系统重启后需重新设置:

#### # ulimit -m unlimited

-m 〈内存大小〉 指定可使用的内存上限,单位为 KB。

#### 1.6.9.7 Swap 分区

Swap 分区大小设置为和物理内存大小相同,建议 swap 文件和数据文件放到不同的磁盘上。

#### 1.6.9.8 关闭 CPU 超频



可在 BIOS 中进行设置关闭 CPU 超频。

开机自检画面,按 DEL 键进入 BIOS 设置界面。选择设置中的 Load Default optimal (读取默认设置),确认后退出。

## 1.6.10各节点修改主机名配置文件

安装集群前,需要修改主机名配置文件/etc/hosts,在127.0.0.1 后对应上 hostname。

示例: hostname为 gc110,具体设置如下:

# # vi /etc/hosts 127.0.0.1 localhost localhost.localdomain localhost4 localhost4.localdomain4 gc110 ::1 localhost localhost.localdomain localhost6 localhost6.localdomain6 ~

## 2 安装集群

## 2.1 部署限制及建议

V8.6.2.X 版本支持非对称部署方式, Coordinator 节点支持最多部署 64 个,数据节点支持部署 300 个以上。当数据节点不超过 64 节点时,可以采用对 称部署方式。在非对称部署方式下考虑管理节点高可用的因素,建议部署至少 3 台(允许 1 个管理节点出现故障)或者 5 台(允许 2 个管理节点出现故障), 因此一般建议在节点数在 5 台以下(包括 5 台)的情况下,采用对称部署方式, 5 台以上的采用非对称部署方式。

注:针对 Redhat7.x 或 SUSE12,需要修改配置文件 /etc/systemd/system.conf中,设置参数 DefaultTasksMax=infinity。

## 2.2 获取安装文件

将安装光盘放入光驱,使用 mount 命令将光驱挂载到 Linux 文件系统。加载光驱命令如下(加载光驱需要 root 用户):

<pre># mkdir /mnt/cdrom</pre>
<pre># mount /dev/cdrom /mnt/cdrom</pre>
将光驱中的安装包的压缩文件 GBase8a_MPP_Cluster- <product< th=""></product<>
version>- <build version="">-<os and="" archtecture="" version="">.tar.bz2 (如:</os></build>
GBase8a_MPP_Cluster-8.6.1.1_build6.1-redhat6.2-x86_64.tar.bz2)复制到
文件系统的某个目录下。进入该目录(下文将以安装包放在了/root 目录下为

例),在命令行模式下使用 tar 命令进行解压。解压命令如下:

# tar xfj GBase8a\_MPP\_Cluster-8.6.1.1\_build6.1-redhat6.2-x86\_64.tar.bz2 解压后,将会在解压目录下生成 gcinstall 目录。



注:在安装或升级前,需确认安装包和平台、操作系统的匹配性,不可使 用跨平台和操作系统的安装包进行安装或升级。

## 2.3 修改安装参数文件

用户在安装集群前,需要根据实际情况,对安装参数文件进行修改。安装 参数文件存在于解压安装包文件后的gcinstall/这个路径里,安装参数文件的 文件名为:demo.options。

用户参照 2.1 章节介绍的操作方法,进入解压后的 gcinstall 目录,根据 实际的集群环境修改安装参数文件 demo.options,具体命令如下:

```
# cd gcinstall
# vi demo.options
installPrefix= /opt
coordinateHost =172.16.83.11,172.16.83.12,172.16.83.13
dataHost = 172. 16. 83. 11, 172. 16. 83. 12, 172. 16. 83. 13
#existCoordinateHost =
#existDataHost =
loginUser= root
loginUserPwd = '111111'
#loginUserPwdFile = loginUserPwd.json
dbaUser = gbase
dbaGroup = gbase
dbaPwd = gbase
rootPwd = 111111
#rootPwdFile = rootPwd.json
dbRootPwd = ''
#mcastAddr = 226.94.1.39
mcastPort = 5493
```

安装参数说明如下:

```
installPrefix: 指定安装目录。
```

coordinateHost : 所有 coordinator 节点 IP 列表, IP 地址之间用"," 分隔。 dataHost: 所有 data 节点 IP 列表, IP 地址之间用","分隔。

existCoordinateHost: 所有已存在的 coordinator 节点 IP 列表, IP 地址 之间用","分隔。

existDataHost: 所有已存在的 data 节点 IP 列表, IP 地址之间用","分隔。

loginUser:通过该用户 ssh 到集群各节点,再 su 切换到 root 用户执行命令,该参数必须使用。默认用户为 root。

loginUserPwdFile 表示支持用户 loginUser 在集群节点不同密码的方式; loginUserPwd 表示用户 loginUser 在集群节点密码相同的方式。两者不能兼用, 否则报错,可以根据密码的异同选择参数使用,必须使用其一。默认密码均为 111111。

dbaUser: 集群运行时使用的操作系统用户名。配置 dbaUser 与 installPrefix 执行目录相同,需要保证该用户已创建。

dbaGroup: 集群运行时使用的操作系统用户所属的组名。

dbaPwd: 集群运行时使用的操作系统用户的密码。

rootPwdFile, 该参数支持 root 用户在多节点不同密码方式, 与参数 rootPwd 不能同时使用, 否则报错。默认密码 111111。

dbRootPwd:数据库 root 用户密码。安装集群时无需设置,仅在集群升级、 扩容、节点替换时使用生效。

mcastAddr: 组播地址, 缺省值为 226.94.1.39, 组播方式会引起网络风暴, 因此不建议使用此方式, 如果将 mcastAddr 参数删除, 则会按照 UDPU 方式进行 安装, 我们建议用户按此方式安装。

mcastPort:组播端口,缺省值为5493,如果按照UDPU方式安装,则此参数无效。

rootPwd:安装集群节点机器中 linux 操作系统中 root 用户的统一的密码。

characterSet:系统支持指定字符集安装,默认值为utf8,取值范围【utf8,gbk】。同时需保证,客户端设置的字符集,以及操作系统设置的字符集,必须与其保持统一。

sshPort:用于指定安装集群的机器的 ssh port, 要求所有节点 ssh port 一致。默认不在 demo.options 显示。默认值为 22。

coordinateHostNodeID: 支持对 IPV6 地址的输入,当集群节点为 IPV4 地 址,则该参数无效,可不用设置;当集群节点为 IPV6 地址,则用户必须手动设 置该参数,该参数为若干个正整数,需要与 coordinateHost 中的 IP ——对应, 用逗号隔开。例如 coordinateHostNodeID = 1,2,3

注: 不支持 IPV4、IPV6 的混合模式。

```
# cd gcinstall
   # vi demo.options
   installPrefix= /opt
   coordinateHost = 2001:DB8:0:23:8:800:200C:5, 2001:DB8:0:23:8:800:200C:6,
2001:DB8:0:23:8:800:200C:7
   coordinateHostNodeID = 1, 2, 3
    dataHost = 2001:DB8:0:23:8:800:200C:5, 2001:DB8:0:23:8:800:200C:6,
2001:DB8:0:23:8:800:200C:7
   #existCoordinateHost =
   #existDataHost =
   loginUser= root
   loginUserPwd = '111111'
   #loginUserPwdFile = loginUserPwd.json
   dbaUser = gbase
   dbaGroup = gbase
   dbaPwd = gbase
   rootPwd = 111111
   #rootPwdFile = rootPwd.json
   dbRootPwd = ''
    #mcastAddr = 226,94.1.39
   mcastPort = 5493
```

用户可根据实际情况修改以上参数。

GBASE

注意:

- 以上涉及密码的参数,如果设置值包含特殊符号,需要使用''包围符, 无需转义;
- 建议 coordinator 集群节点数不超过 64 个, data 集群节点数不超过 300 个;
- 3) 参数 loginUser, dbaUser, dbaGroup 使用规范:
  - ▶ 支持单引号、双引号、包围符;
  - ▶ 参数值仅支持"字母,数字以及下划线,但是不能以数字开头";

修改结果举例如下:

示例 1: 全新安装

```
installPrefix= /opt
coordinateHost =172.16.83.11,172.16.83.12,172.16.83.13
dataHost = 172.16.83.11,172.16.83.12,172.16.83.13
#existCoordinateHost =
#existDataHost =
dbaUser = gbase
dbaGroup = gbase
dbaGroup = gbase
rootPwd = 111111
dbRootPwd = ''
#mcastAddr = 226.94.1.39
mcastPort = 5493
```

示例 2: 增加 coordinator 和 data 节点

```
installPrefix= /opt
coordinateHost = 172.16.83.14
dataHost = 172.16.83.14
existCoordinateHost = 172.16.83.11,172.16.83.12,172.16.83.13
existDataHost = 172.16.83.11,172.16.83.12,172.16.83.13
dbaUser = gbase
dbaGroup = gbase
dbaPwd = gbase
```





```
rootPwd = 111111
dbRootPwd = ''
#mcastAddr = 226.94.1.39
mcastPort = 5493
```

## 2.4 执行安装

成功修改安装参数文件后,执行安装脚本 gcinstall.py,安装执行中,具体命令如下:

# .,	/gcinstall.py	silent=demo.options	
------	---------------	---------------------	--

参数说明:

1、gcinstall.py 命令可选参数: passwordInputMode

参数含义:指定密码获取的方式,通过不同的参数实现不同的获取方式。 若指定该参数,则 demo. options 中的密码不必再修改。

默认值: file

取值范围: [file, pwdsame, pwddiff]

- file:表示从文件或命令行参数获取,和原有的方式一致,该方式下, 文件中的密码是明文的;
- pwdsame:表示从终端由用户输入密码,并且所有节点的密码一致情况 下使用该参数,对于不同用户密码只输入一次,适用于所有节点都用 这个密码;
- pwddiff:表示从终端由用户输入密码,并且节点间的密码不一致情况下使用该参数,对于不同用户密码每个节点分别输入一次,适用于不同节点使用不同的密码;

2**、**-a

参数含义:安装时跳过与用户交互提示信息。

示例:

./gcinstall.py ---silent=demo.options -a

Do you accept the above licence agreement ([Y,y]/[N,n])?(忽略, 继续安装)

注: 必须在 coordinator 节点上执行 gc install.py 进行安装。

安装成功后,会在每个节点上创建一个操作系统的用户 gbase,其密码默 认是 gbase。GBase 8a MPP Cluster 的主要软件模块都运行在 gbase 用户下。

每个 coordinator 节点的/opt 下会创建 gcluster 子目录, 每个 data 节点的/opt 下会创建 gnode 子目录, GBase 8a MPP Cluster 的数据库用户, 默认为 root, 密码为空。

安装成功后,通过 gcadmin 查看集群状况,如下:

[gbase@node1 ;	gbase@node1 gcinstall-8611]\$ gcadmin					
CLUSTER STATE	LUSTER STATE: ACTIVED					
CLUSTER MODE:		NORMAL				
+=====================================	=== ASE	E COORDINATOR C	ELUSTER I	======================================	 I	
-======================================						
NodeName		IpAddress	gcware	gcluster	DataState	
coordinator1		172. 16. 83. 11	OPEN	OPEN	0	
coordinator2		172. 16. 83. 12	OPEN	OPEN	0	
coordinator3		172. 16. 83. 13	OPEN	OPEN	0	
+======================================					+	
	GBA	ASE DATA CLUSTE	R INFORM	ATION ==========	 +======+	
NodeName	NodeName   IpAddress  gnode  syncserver  DataState					
node1   172.16.83.11   OPEN   OPEN   0						
++		+	+	+	+	





node2	172.16.83.12   OF	PEN   C	OPEN	0	
+	172.16.83.13   OF	PEN   C	DPEN	0	-

## 2.4.1 一键安装

1、修改安装参数文件 demo. options

```
# cd gcinstall
# vi demo.options
installPrefix= /opt
coordinateHost =172.16.83.11,172.16.83.12,172.16.83.13
dataHost = 172. 16. 83. 11, 172. 16. 83. 12, 172. 16. 83. 13
#existCoordinateHost =
#existDataHost =
loginUser= root
loginUserPwd = '111111'
#loginUserPwdFile = loginUserPwd.json
dbaUser = gbase
dbaGroup = gbase
dbaPwd = gbase
rootPwd = 111111
#rootPwdFile = rootPwd.json
dbRootPwd = ''
#mcastAddr = 226.94.1.39
mcastPort = 5493
updateConf = cluster.conf
```

#### 2、配置 cluster. conf 文件

```
[init]
distibution = gcadmin distribution gcChangeInfo.xml
initnodedatamap = 1
[gcluster]
XXXX = *
```



[gbase] XXXX= \* [gcrecover] XXXX = \* [corosync] XXXX = \*

参数说明:

该配置文件支持修改集群层、单机层、gcrecover、corosync 配置文件, 以及初始化。集群在安装过程中,根据用户预先配置信息,在集群安装成功后, 作对应修改。

[init]: 创建 distribution 和初始化集群。

若成员 distribution 和 initnodedatamap,不存在或写错,不执行;

若 distribution 对应值写错或当前集群执行报错,仅记日志;

initnodedatamap 的执行依赖 distribution 是否创建成功;

initnodedatamap 值为 1, 表示初始化, 其他, 不执行;

[gcluster]

对应为集群层配置文件 gbase\_8a\_gcluster. cnf;

指定配置参数对应为 gbase\_8a\_gcluster. cnf 中[gbased]成员;

若指定配置参数无效和写错,导致集群 gclusterd 服务启动失败,集群回滚;

[gbase]

对应为集群层配置文件 gbase\_8a\_gbase.cnf;

指定配置参数对应为 gbase\_8a\_gbase. cnf 中[gbased]成员;

若指定配置参数无效和写错,导致集群 gbased 服务启动失败,集群回滚; [gcrecover]

对应为集群层配置文件 gc\_recover. cnf;

指定配置参数须为 gcrecover 支持参数,写错或对应参数值无效(或不在 取值范围内(即取默认值)),gcrecover 不会影响该功能使用。

[corosync]

对应为/etc/corosync/corosync.conf 配置文件;

corosync 配置文件包含 totem, logger, gcware 三个模块, 其中 gcware 包含集群所有端口信息, 己有功能 (可通过修改 gcinstall 目录下文件 extendCfg.xml 实现) 已支持集群端口修改, 因此不支持该模块参数的修改;

totem支持参数为 token, token\_retransmits\_before\_loss\_const, send\_join, join, consensus;

logger 支持参数为 fileline: off, to\_stderr: no, to\_file: yes, to\_syslog: no, debug: off, timestamp: on;

其它参数不支持,配置无效,不修改;

参数值无效,配置无效,不修改;

3、执行一键安装

./gcinstall.py --silent=demo.options -a

## 2.5 采集指纹

注: 若您安装的 NOLICENSE 版本集群,请忽略此小节。

如需进行集群 LICENSE 认证,首先,需要使用指纹采集工具 gethostsid 进行集群指纹采集,具体语法如下:

- a) gethostsid -n 192.168.10.1, 192.168.10.2, 192.168.10.3 <-f /path/name> <-u username> <-p password>
- b) gethostsid--hosts=demo.hosts <-f /path/name> <-u
  username> <-p password>

GBASE

其中:

a)用于采集单个节点或少数几个节点服务器指纹的使用场景

b)用于采集多个节点服务器指纹的使用场景

-f 指定生成的指纹文件路径

-u 指定集群安装的用户名,目前只支持使用 root 用户采集指纹

-p 指定集群安装用户的密码

注意:

1) 对于 a) 指定多个 ip 地址时, 通过 "," 作为多个 ip 地址间的分隔符。

2) 对于 b) demo. hosts 文件中包含需要采集的 ip 列表

```
#vim demo.hosts
```

Hosts = 192.168.10.1, 192.168.10.2, 192.168.10.3

文件中, 多个 ip 地址时, 通过","作为多个 ip 地址间的分隔符。

4) 若限制 ssh 到 root 权限,建议使用 dbaUser 及其 pwd,即 ./gethostid
 -n\* ... -ugbase -ggbase;

否则,建议使用 root 用户,即./gethostid -n\*... -uroot -p 111111。

示例:使用 root 用户进行如下操作:

\$ gethostsid -n 10.10.10.10.10.10.2,10.10.10.3 -u root -p 111111 -f hostsfinger.txt

其中, 10.10.10.1,10.10.10.2,10.10.10.3 为集群节点 IP。

或者采用下面的操作:

\$ gethostsid --hosts=mpp.hosts -u root -p 1111111 -f hostsfinger.txt

其中 mpp.hosts 的内容如下:

Hosts=10. 10. 10. 1, 10. 10. 10. 2, 10. 10. 10. 3

## 2.6 获取 license 文件

注: 若您安装的 NOLICENSE 版本集群,请忽略此小节。

用户将采集的指纹文件发送给数据库厂商(请将您的指纹文件发到 <u>license@gbase.cn</u>中,同时,需要在 mail 中附上安装 GBase 8a 产品所使用的 操作系统名称及版本),数据库厂商最终为用户提供试用版或者商用版 license 文件。

注意:确保指纹文件里包含了全部集群节点的指纹信息,否则可能会导致 部分节点无法启动。

## 2.7 导入 license 文件

注: 若您安装的 NOLICENSE 版本集群,请忽略此小节。

用户使用 license 导入工具将 license 文件导入到集群的每个节点,具体 语法如下:

a) License -n 192. 168. 10. 1, 192. 168. 10. 2, 192. 168. 10. 3

 $\langle -f \ /path/name.lic \rangle \langle -u \ username \rangle \langle -p \ password \rangle$ 

b) License --hosts=demo.hosts <-f /path/name.lic> <-u
username> <-p password>

其中:

- a) 用于导入单个节点或少数几个节点 License 的使用场景;
- b) 用于导入多个节点 License 的使用场景;
- -f 待导入的 License 文件路径及文件名称
- -u 指定集群安装的用户名,目前仅支持 DBA 用户
- -p 指定集群安装用户的密码

注意:



对于 a) 指定多个 ip 地址时,通过","作为多个 ip 地址间的分隔符。
 对于 b) demo. hosts 文件中包含需要采集的 ip 列表:

#### #vim demo.hosts

Hosts = 192.168.10.1, 192.168.10.2, 192.168.10.3

文件中, 多个 ip 地址时, 通过","作为多个 ip 地址间的分隔符

示例:使用 gbase 用户进行如下操作:

\$ License -n 10.10.10.1, 10.10.2, 10.10.10.3 -u gbase -p gbase -f
/home/gbase/licensetools/mpp.lic

或者

\$ License --hosts=mpp.hosts -u gbase -p gbase -f /home/gbase/licensetools/mpp.lic

其中, "mpp.lic"为获得的license文件。

- 2) 保证 license 导入工具有 license 文件的读权限,否则会导致操作失败。
- 3) -u 参数指定的用户必须为集群安装使用的 DBA 用户。

## 2.8 获取 license 状态

注: 若您安装的 NOLICENSE 版本集群,请忽略此小节。

用户使用 license 检测工具获取 license 状态信息,包括集群节点是否全 部具备 license 文件,license 为试用版还是商用版,如果是试用版 license, 需要显示 license 到期日期,具体语法如下:

a) chkLicense [-n 192.168.10.1,192.168.10.2,192.168.10.3] -u username -p password

b) chkLicense ---hosts=demo.hosts -f/path

/filename -u username -p password

其中:

a)用于查看指定一个或少数几个节点的 License 信息,在终端上输出显示; -n 指定需要检查的节点的 IP 地址,该参数可省略,如省略则 默认检查本节点 License 信息;

b) 用于查看--hosts=demo.hosts 中指定的多个节点的 License 信息,并 将结果汇总到-f/path /filename;

-f 指定 License 检查结果存放的文件及路径

-u 指定集群安装的用户名,目前仅支持 DBA 用户

-p 指定集群安装用户的密码

注意:

1) 对于 a)指定多个 ip 地址时,通过","作为多个 ip 地址间的分隔符。

2) 对于 b)demo.hosts 文件中包含需要采集的 ip 列表

#vim demo.hosts

Hosts = 192.168.10.1, 192.168.10.2, 192.168.10.3

文件中,多个 ip 地址时,通过","作为多个 ip 地址间的分隔符。

示例一:使用 gbase 用户进行如下操作:

\$ chkLicense -n 10.10.10.1,10.10.2,10.10.10.3 -u gbase -p gbase //其中,10.10.10.1,10.10.10.2,10.10.3 为集群节点 IP。如果为商用版 license, 输出 如下 =====10.10.10.1===== is\_exist:yes version:business is\_valid:yes =====10.10.10.2===== is\_exist:yes version:business is\_valid:yes =====10.10.10.3===== is\_exist:yes version:business



is_valid:yes					
Successful node nums: 3					
=====10. 10. 10. 1======					
is_exist:yes					
version:trial					
expire_time:20161230					
is_valid:yes					
=====10. 10. 10. 2=====					
is_exist:yes					
version:trial					
expire_time:20161230					
is_valid:yes					
=====10. 10. 10. 3=====					
is_exist:yes					
version:trial					
expire_time:20161230					
is_valid:yes					
Successful node nums: 3					

其中, is\_exist 用于标识 license 文件是否存在, 如果为 yes 代表存在, no 为不存在;

version 用于标识 license 类型,即是试用版还是商用版,如果为 trial 代表为试用版 license,如果为 business 代表为 商用版 license;

is\_valid 用于标识 license 是否有效,如果为 yes 代表 license 有效, no 代表 license 失效;

expire\_time 用于标识试用版 license 的到期日期,只在检测试用版 license 时才会显示;



示例二:使用 gbase 用户进行如下操作:

\$ chkLicense --hosts=mpp.hosts -u gbase -p gbase -f licensestatus.txt //其中 mpp.hosts 的内容如下: //Hosts=10. 10. 10. 1, 10. 10. 10. 2, 10. 10. 10. 3 //在这种情况下,屏幕输出为 \_\_\_\_\_ Successful node nums: 3 输出文件中的内容与示例一信息输出类似,在商用版 license 的情况下, licensestatus.txt 的内容如下: ====10. 10. 10. 1===== is exist:yes version:business is valid:yes ====10. 10. 10. 2===== is exist:yes version:business is\_valid:yes ====10. 10. 10. 3===== is exist:yes version:business is valid: yes

注:完成以上操作后,执行 service gcware restart。

## 2.9 配置 distribution

切换为 gbase 用户,在安装包根目录下配置 distribution。以上述安装为例,配置 distribution 的操作如下:

\$ gcadmin distribution gcChangeInfo.xml p 3 d 1 gcadmin generate distribution ...

NOTE: node [172.16.83.13] is coordinator node, it shall be data node too NOTE: node [172.16.83.12] is coordinator node, it shall be data node too



NOTE: node [172.16.83.11] is coordinator node, it shall be data node too gcadmin generate distribution successful

说明: gcadmin distribution 是 gcadmin 的子命令,详细说明请参考管理 员手册中 gcadmin 命令说明。

生成的 distribution 如下所示:

\$ g	cadmin	showdistribution
------	--------	------------------

Distribution ID: 1   State: new   Total segment num: 9							
Primary	Segment Node IP	Segment ID		Duplicate Segment node	IP		
172.	16.83.13	1		172.16.83.12	 		
172.	16.83.12	2		172.16.83.11			
172.	16.83.11	3		172.16.83.13			
172.	16.83.13	4		172.16.83.11			
172.	16.83.12	5		172.16.83.13			
172.	16.83.11	6		172.16.83.12			
172.	16.83.13	7		172.16.83.12			
172.	16.83.12	8		172.16.83.11			
172.	16.83.11	9		172.16.83.13			

## 2.10 集群的初始化

集群安装完毕,并且建立 distribution 之后,在首次执行 SQL 命令之前, 需要对数据库系统做初始化操作,才能正确执行所有的 SQL 操作。如果不做初 始化操作,执行数据库操作时会提示不能创建或执行。

1、 首先使用 root 帐号登录集群。

```
$ gccli -uroot -p
Enter password:
GBase client 8.6.1.1 build 49382. Copyright (c) 2004-2015, GBase. All Rights
```



#### Reserved.

gbase>

2、 执行初始化命令 initnodedatamap 命令。

#### gbase> initnodedatamap;

Query OK, O rows affected

注意:此命令只需执行一次,如果重复执行,系统会报告如下信息:

#### gbase> initnodedatamap;

ERROR 1707 (HY000): gcluster command error: (GBA-02CO-0004) nodedatamap is already initialized.

至此, GBase 8a MPP Cluster 安装结束。

GBASE

# 3 启停与登录

## 3.1 集群服务的启停

成功安装集群后,用户可以在每个节点上使用操作系统的 root 用户执行以下命令启停 gcware 服务:

▶ 启动服务:

# service gcware start

▶ 停止服务:

# service gcware stop

▶ 重启服务:

# service gcware restart

其中, coordinator 节点针对的服务有

gclusterd, gcmonit.corosync, gcrecover; data 节点针对的服务有 gbased, syncserver, gcmonit。

## 3.2 集群登录与退出

1、 集群登录:

在任何节点上,在命令行状态,用户必须使用 gbase 用户,通过 gbase 客户端工具连接数据库。命令格式如下:

gccli -u<username> -p<password> [-h<ipaddress>] [-P<portname>]
[-D<databasename>] [--nice\_time\_format] [-c]

参数:

username: -u 的参数值, 连接数据库的用户名称, 该名称默认为 root。

password: -p 的参数值,连接数据库的密码,该密码默认为空,首次安装并登录集群后,需要修改 root 用户的 password。

ipaddress: -h 的参数值,登录集群节点的 IP 地址,默认 127.0.0.1,可选参数;如果指定多个 IP 地址,则启动 gccli 的高可用功能, IP 地址之间采用","分隔,例如 192.168.100.10,192.168.100.11,192.168.100.12。

在指定多个集群节点 IP 的情况下,gccli 首先尝试与第一个节点 IP 建立 连接,如果无法建立连接,则尝试与下一个节点 IP 建立连接。以此类推,如果 配置的多个集群节点 IP 均无法建立连接,则 gccli 无法建立与集群的连接,本 次登录失败。

portname: -P的参数值,集群使用的端□号,默认 5258,可选参数。

databasename:-D的参数值,指定登录时默认的数据库(数据库必须存在),可选参数。

--nice\_time\_format: 指定时间的显示格式,使用该参数,格式显示为'n hour n min nnn.nn sec',不使用该参数,格式显示为'hh:mm:ss.ff'。可选参 数。nice\_time\_format 可以这样使用:

- 1) gccli -uroot --nice\_time\_format
- 2) gccli -uroot --nice\_time\_format = 1
- 3) gccli -uroot --nice\_time\_format = 0

其中1和2等价, 3等价于不使用 nice\_time\_format。

-c:使用此参数,用于使用 hint 优化方式。可选参数。

示例 1: 默认登录。此时操作用时精度为秒,格式为"Elapsed: \*\*:\*\*:\*\*.\*\*"。

#### \$ gccli -uroot -p

Enter password:


GBase client 8.6.1.1 build 25776. Copyright (c) 2004-2013, GBase. All Rights Reserved.

gbase> USE gbase;

Query OK, 0 rows affected (Elapsed: 00:00:00.03)

gbase>

示例 2: 使用--nice\_time\_format 登录, 时间精度为毫秒, 格式为 "\*\*.\*\*

sec" 。

\$ gccli -uroot -p -nice\_time\_format

#### Enter password:

GBase client 8.6.1.1 build 25776. Copyright (c) 2004-2013, GBase. All Rights Reserved.

gbase> USE gbase;

Query OK, 0 rows affected (0.00 sec)

gbase>

2、 退出登录

直接输入\q即可退出登录。

gbase> \q

Bye

# 4 卸载集群

### 4.1 执行卸载脚本

当用户需要卸载 GBase 8a MPP Cluster 时,首先需要停止集群各节点服务, 然后进入 gcinstall 目录,再进行如下操作。

Ħ	./unInstall.py	silent=demo.options	
---	----------------	---------------------	--

说明:

unInstall.py 命令可选参数: passwordInputMode

参数含义:指定密码获取的方式,通过不同的参数实现不同的获取方式。 若指定该参数,则 demo. options 中的密码不必再修改。

默认值: file

取值范围: [file, pwdsame, pwddiff]

- file:表示从文件或命令行参数获取,和原有的方式一致,该方式下, 文件中的密码是明文的;
- pwdsame:表示从终端由用户输入密码,并且所有节点的密码一致情况 下使用该参数,对于不同用户密码只输入一次,适用于所有节点都用 这个密码;
- pwddiff:表示从终端由用户输入密码,并且节点间的密码不一致情况下使用该参数,对于不同用户密码每个节点分别输入一次,适用于不同节点使用不同的密码;

如果因为当前集群环境被破坏,导致执行 unInstall.py 报错,可以使用" --Force"参数卸载集群,如下所示:

# ./unInstall.py --silent=demo.options --Force



# 5 升级、回退集群

注:

1) 系统不支持 LICENSE 版本与 NOLICENSE 版本之间的集群升级与回退;

2)禁止并行执行升级,安装,卸载等命令,包括在同一个节点不同的 session 和不同节点执行;

3) 升级过程中必须保持原有节点在线;

4) 升级前需保证集群已完成初始化;

# 5.1 V8.5.1.2 集群升级到 V8.6.X.X 集群

### 5.1.1 升级前的准备

#### 5.1.1.1 查看现有集群数据库信息

首先查看集群状态:

# gcadmin										
CLUSTER STATE: ACTIVED										
CLUSTER MODE: NORMAL										
GCLUSTER INFORMATION										
rowid	nodename	IpAddress	sgname	dpname	gcware	gnode	gcluster	syncserver	datastate	nodestate
1	node1	192.168.151.205	sg01	n1	online	OPEN	OPEN	OPEN	[0]	[0]
2	node2	192.168.151.206	sg01	n2 n2	Online	OPEN	OPEN	OPEN	[0]	[0]
3	node3	192.168.151.207	sg02	n3	Online	OPEN	OPEN	OPEN	[0]	[0]
4	node4	192.168.151.208	sg02	n4	Online	OPEN	OPEN	OPEN	[0]	[0]

必须保证所有节点的状态正常,才可以进行下一步操作,否则需要解决故 障后才能继续操作。

然后进入数据库进行操作:

# su - gbase



GBase 8a MPP Cluster 安装手册

\$ gccli -uroot -p								
Enter password:								
GBase client 8.6.1.1 build 31465. Copyright (c) 2004-2013, GBase. All Rights								
Reserved.								
//查看现有集群版本								
gbase> SHOW VARIABLES LI	KE '%VERS	ION';						
+	+	-+						
Variable_name	Value							
+	+	-+						
gcluster_hash_version	1							
protocol_version	10							
version	8.6.1.1							
+	+	-+						
3 rows in set								

## 5.1.1.2 恢复 FEVENTLOG

第一步,通过 gcadmin 命令查看集群状态,根据 datastate 值判断是否还 有未恢复的日志。若所有节点 datastate 值为 0,则可以直接进行 6.1.2章节 所描述步骤。

\$	gcadmin		
	CLUSTER STATE: CLUSTER MODE:	ACTIVED NORMAL	
+=	 	+	GCLUSTER INFORMATION
 += ==		+	



+	-++-	+	+-	+		+	+
rowic	d   nodename c   syncserven	+   IpAddress   r   datastate   node	+ sgname   estate	dpname	gcware	gnode	
+	++   node1   OPEN	+++++++	+ sg01   	nl	Online	+ OPEN	+
+   2 OPEN	++   node2   OPEN	+++++++	+ sg01   	n2	Online	OPEN	
+   3 OPEN	++   node3   OPEN	++++++	+ sg01   	n3	Online	+ OPEN	
+	++-	++++++	+	+		+	+

第二步:如果有节点 datastate 值不为 0,则需要恢复这些节点的日志。 首先需等待 (大概 30min)自动恢复,若自动恢复不成功,则需要手工进行日 志恢复。

第三步:完成日志恢复后,最后需检查一下是否所有节点的 FEVENTLOG 已 经恢复成功。

#### 5.1.1.3 停止集群服务

执行集群的升级操作前,首先需要停止集群各个节点上的 gcware 服务。停止 gcware 服务,需要在各个节点机器上使用 root 用户进行。

具体操作如下(其他各个节点都要进行一次该操作):

GBase 8a MPP Cluster 安装手册



\$ su - root Password: # service gcware stop Stopping GCMonit success! Signaling GCMETARECOVER (gcmetarecover) to terminate: [ OK ] Waiting for gcmetarecover services to unload: [ OK ] Signaling GCDATARECOVER (gcdatarecover) to terminate: [ OK ] Waiting for gcdatarecover services to unload: [ OK ] Signaling GCSYNC (gc\_sync\_server) to terminate: [ OK ] Waiting for gc\_sync\_server services to unload: [ OK ] Signaling GCLUSTERD to terminate: [ OK ] Waiting for gclusterd services to unload:...[ OK ] Signaling GBASED to terminate: [ OK ] Waiting for gbased services to unload: [ OK ] Signaling GCWARE (gcware) to terminate: [ OK ] Waiting for gcware services to unload: [ OK ]

## 5.1.2 具体升级步骤

V8.5.1.2集群是对称部署方式,映射到V8.6.X.X集群的概念就是每一个物理节点上同时安装了 coordinator 和 data 集群节点。从V8.5.1.2升级到V8.6.X.X集群时,需要设置所有的集群节点都是 coordinator 节点和 data 节点。

具体升级步骤如下:

1、 停止所有集群节点。

2、 解压 V8.6.X.X 集群安装包,并切换到 gcinstall 目录下。

3、修改 demo. options 配置文件。设置 coordinateHost 和 dataHost 为所 有集群节点 IP。如果 V8.5.1.2 使用组播模式,那么还需要设置 existCoordinateHost 和 existDataHost 为所有集群节点 IP。如下所示:

installPrefix= /opt coordinateHost =172.16.83.11,172.16.83.12,172.16.83.13 dataHost = 172.16.83.11,172.16.83.12,172.16.83.13



#existCoordinateHost =
#existDataHost =
dbaUser = gbase
dbaGroup = gbase
dbaPwd = gbase
rootPwd = 111111
#mcastAddr = 226.94.1.39
mcastPort = 5493

4、使用 dbaUser 用户执行./gcinstall.py --silent=demo.options -U
进行集群升级。如果升级成功集群会自动启动,如果升级失败集群回退到
V8.5.1.2版本。

注: gcinstall.py 命令可选参数: passwordInputMode

参数含义:指定密码获取的方式,通过不同的参数实现不同的获取方式。 若指定该参数,则 demo. options 中的密码不必再修改。

默认值: file

取值范围: [file, pwdsame, pwddiff]

- file:表示从文件或命令行参数获取,和原有的方式一致,该方式下, 文件中的密码是明文的;
- pwdsame:表示从终端由用户输入密码,并且所有节点的密码一致情况 下使用该参数,对于不同用户密码只输入一次,适用于所有节点都用 这个密码;
- pwddiff: 表示从终端由用户输入密码,并且节点间的密码不一致情况下使用该参数,对于不同用户密码每个节点分别输入一次,适用于不同节点使用不同的密码;

## 5.2 V8.6.X.X 集群间版本升级



V8.6.X.X 集群之间的版本升级时需要设置 coordinateHost 和 dataHost 为所有已存在的 coordinator 和 data 集群节点 IP。升级操作需要在一个已存在的 coordinator 节点上进行。

若升级前集群在 IPV6 网络环境下,则升级时需设置 coordinateHostNode ID 与原集群一致,且 IP 类型与原集群一致。

必须保证所有节点的状态正常,且已完成集群初始化,才可以进行下一步 操作,否则需要解决故障后才能继续操作。

具体升级步骤如下:

1、 停止所有集群节点服务。

2、 解压 V8.6.X.X 集群安装包,并切换到 gcinstal1 目录下。

3、 修改 demo. options 配置文件。

installPrefix= /opt coordinateHost =172.16.83.11,172.16.83.12,172.16.83.13 dataHost = 172.16.83.11,172.16.83.12,172.16.83.13 #existCoordinateHost = #existDataHost = dbaUser = gbase dbaGroup = gbase dbaGroup = gbase rootPwd = 111111 #mcastAddr = 226.94.1.39 mcastPort = 5493

4、使用 dbaUser 用户执行./gcinstall.py --silent=demo.options -U 进行集群升级。如果升级成功集群会自动启动,如果升级失败集群回退到老版本。

注: gcinstall.py 命令可选参数: passwordInputMode

参数含义:指定密码获取的方式,通过不同的参数实现不同的获取方式。 若指定该参数,则 demo. options 中的密码不必再修改。

默认值: file

取值范围: [file, pwdsame, pwddiff]

- file:表示从文件或命令行参数获取,和原有的方式一致,该方式下, 文件中的密码是明文的;
- pwdsame:表示从终端由用户输入密码,并且所有节点的密码一致情况 下使用该参数,对于不同用户密码只输入一次,适用于所有节点都用 这个密码;
- pwddiff:表示从终端由用户输入密码,并且节点间的密码不一致情况下使用该参数,对于不同用户密码每个节点分别输入一次,适用于不同节点使用不同的密码;

#### 5.3从 V8.6.X.X 集群回退到 V8.5.1.2 集群

我们不建议把集群从 V8.6.X.X 版本回退到 V8.5.1.2 版本,而且多数场景 下这种回退是无效的,并且会导致数据错误。

但是如果满足如下条件,可以把 V8.6.X.X 集群成功回退到 V8.5.1.2 集群。

1、 当前集群是从 V8.5.1.2 版本升级到 V8.6.X.X 版本。

2、从 V8.5.1.2 升级到 V8.6.X.X 时使用的 V8.6.X.X 安装包解压目录仍然 存在。从 V8.5.1.2 版本升级到 V8.6.X.X 版本时会在安装包根目录 下.systablebak 中暂时保留 V8.5.1.2 版本的系统用户数据文件。

3、 从 V8.5.1.2 版本升级到 V8.6.X.X 版本后,没有执行过 DDL 操作、扩 容操作、生成新 distribution 操作,没有生成新的 FEVENTLOG。

4、从 V8.5.1.2 升级到 V8.6.X.X 时升级工具对 V8.5.1.2 生成的备份文件 还存在。V8.5.1.2 升级到 V8.6.X.X 时备份文件名包含 federal 关键字,如: gcluster\_backup\_federal\_20151110162202.tar.bz2。

注: V8.5.1.2版本包含 TableSpace 功能数据,升级后产生过新的数据, 是不能进行回退的。



首先要停掉集群所有节点服务,然后必须在 gbase 用户下执行回退集群版本的命令,如下所示:

python Restore.py

backupFile=/home/gbase/gcluster_	_backup_	_federal_	_20151110162202.	tar.	bz2
silent=demo.options					

注: Restore.py 命令可选参数: passwordInputMode

参数含义:指定密码获取的方式,通过不同的参数实现不同的获取方式。 若指定该参数,则 demo. options 中的密码不必再修改。

默认值: file

取值范围: [file, pwdsame, pwddiff]

- file:表示从文件或命令行参数获取,和原有的方式一致,该方式下, 文件中的密码是明文的;
- pwdsame:表示从终端由用户输入密码,并且所有节点的密码一致情况 下使用该参数,对于不同用户密码只输入一次,适用于所有节点都用 这个密码;
- pwddiff:表示从终端由用户输入密码,并且节点间的密码不一致情况下使用该参数,对于不同用户密码每个节点分别输入一次,适用于不同节点使用不同的密码;

#### 5.4 V8.6.X.X 集群间版本回退

集群回退步骤如下:

- 1,停止集群所有节点服务。
- 2, 执行集群版本回退命令。

执行回退命令:

python Restore.py



--backupFile=/home/gbase/gcluster\_backup\_20151110162202.tar.bz2

--silent=demo.options

注: Restore.py 命令可选参数: passwordInputMode

参数含义:指定密码获取的方式,通过不同的参数实现不同的获取方式。 若指定该参数,则 demo. options 中的密码不必再修改。

默认值: file

取值范围: [file, pwdsame, pwddiff]

- 1) file:表示从文件或命令行参数获取,和原有的方式一致,该方式下, 文件中的密码是明文的;
- pwdsame:表示从终端由用户输入密码,并且所有节点的密码一致情况 下使用该参数,对于不同用户密码只输入一次,适用于所有节点都用 这个密码;
- pwddiff:表示从终端由用户输入密码,并且节点间的密码不一致情况下使用该参数,对于不同用户密码每个节点分别输入一次,适用于不同节点使用不同的密码;



# 6 增加、减少集群节点

所有增加减少集群节点的操作必须在已存在的 coordinator 节点上使用 gbase 用户操作。

# 6.1 增加集群 data 节点

增加集群 data 节点的步骤如下:

1、 修改 demo. options,设置 dataHost 参数为需要增加的 data 节点所在的机器 IP,设置 existCoordinateHost 为已存在的 coordinator 节点所在机器 IP,设置 existDataHost 为已存在的 data 节点所在机器 IP。如下所示:

```
installPrefix= /opt
#coordinateHost =172.16.83.11,172.16.83.12,172.16.83.13
dataHost = 172.16.83.14
existCoordinateHost = 172.16.83.11,172.16.83.12,172.16.83.13
existDataHost = 172.16.83.11,172.16.83.12,172.16.83.13
dbaUser = gbase
dbaGroup = gbase
dbaPwd = gbase
rootPwd = 111111
#mcastAddr = 226.94.1.39
mcastPort = 5493
```

2、 执行集群节点安装脚本。

./gcinstall.py --silent=demo.options

## 6.2 增加集群 coordinator 节点

增加集群 coordinator 节点的步骤如下:

1, 停止所有集群节点服务。

### GBASE

2, 修改 demo. options, 设置 coordinateHost 参数为需要增加的 coordinator 节点所在的机器 IP, 设置 existCoordinateHost 为已存在的 coordinator 节点所在机器 IP,设置 existDataHost 为已存在的 data 节点所在 机器 IP。如下所示:

```
installPrefix= /opt
coordinateHost =172.16.83.14
#dataHost = 172.16.83.11,172.16.83.12,172.16.83.13
existCoordinateHost = 172.16.83.11,172.16.83.12,172.16.83.13
existDataHost = 172.16.83.11,172.16.83.12,172.16.83.13
dbaUser = gbase
dbaGroup = gbase
dbaGroup = gbase
rootPwd = 111111
#mcastAddr = 226.94.1.39
mcastPort = 5493
```

3, 执行集群节点安装脚本。

./gcinstall.py --silent=demo.options

# 6.3 减少集群 data 节点

减少集群 data 节点的步骤如下:

- 1, 停止将要卸载的 data 节点服务。
- 2, 修改 demo. options,设置 dataHost 为将要卸载 data 节点所在机器 IP。

如下所示:

installPrefix= /opt
#coordinateHost =
dataHost = 172.16.83.14
#existCoordinateHost = 172.16.83.11,172.16.83.12,172.16.83.13
#existDataHost = 172.16.83.11,172.16.83.12,172.16.83.13
dbaUser = gbase
dbaGroup = gbase





```
dbaPwd = gbase
rootPwd = 111111
#mcastAddr = 226.94.1.39
mcastPort = 5493
```

3, 执行集群节点卸载脚本。

./unInstall.py --silent=demo.options

# 6.4 减少集群 coordinator 节点

减少集群 coordinator 节点的步骤如下:

1, 停止所有节点集群服务。

2, 修改 demo. options,设置 coordinateHost 为将为将要卸载 coordinator 节点所在机器 IP。如下所示:

```
installPrefix= /opt
coordinateHost = 172.16.83.14
#dataHost =
#existCoordinateHost = 172.16.83.11,172.16.83.12,172.16.83.13
#existDataHost = 172.16.83.11,172.16.83.12,172.16.83.13
dbaUser = gbase
dbaGroup = gbase
dbaGroup = gbase
rootPwd = 111111
#mcastAddr = 226.94.1.39
mcastPort = 5493
```

3, 执行集群节点卸载脚本。

./unInstall.py --silent=demo.options

### GBASE

# 7 安装 C3

当集群节点机器很多时,例如20个节点时,对于集群的一些维护工作需要 在每个节点上都执行一次相关操作,这样十分耗时,而且繁琐。因此,我们在 集群安装包中,提供了C3工具的安装脚本,使用C3工具,可以减少集群节点 的维护操作,仅在装有C3工具的节点机器上输入一次维护命令,便可在所有集 群节点上执行,方便用户操作。

#### 7.1 安装步骤

- 确认 GBase 8a MPP Cluster 安装无误。解压集群的安装包,并切换到 gcinstall 目录下。
- 2、 配置集群节点之间的互信

配置 root 用户互信

```
python SetSSHTrust.py —host=172.16.83.11,172.16.83.12,172.16.83.13
--userName=root --userPwd=111111
```

配置 gbase 用户互信

python SetSSHTrust.py —host=172.16.83.11,172.16.83.12,172.16.83.13 --userName=gbase —userPwd=gbase

3、修改 c3 的配置文件

coor 中配置所有 coordinator 集群节点。Data 中配置所有 data 集群节点。

```
cluster coor {
    172.16.83.11:127.0.0.1 #head node
    dead remove-index-00
    172.16.83.11
    172.16.83.12
    172.16.83.13
}
cluster data {
```



127. 0. 0. 1:127. 0. 0. 1 #head node dead remove-index-00 172. 16. 83. 11 172. 16. 83. 12 172. 16. 83. 13

# 7.2 安装后的验证

C3 工具安装完毕后,可以运行一个操作命令来验证 C3 工具是否可以使用。 我们在 root 用户下,执行 service gcware status 命令做为验证示例:

# cexecs "service gcware status" ----- 172. 16. 83. 11-----corosync (pid 4328) is running... gclusterd (pid 4516) is running... gcrecover (pid 4578) is running... gbased (pid 4440) is running... gc\_sync\_server (pid 4484) is running... ----- 172. 16. 83. 12----corosync (pid 4599) is running... gclusterd (pid 6801) is running... gcrecover (pid 6855) is running... gbased (pid 4691) is running... gc\_sync\_server (pid 6775) is running... ----- 172. 16. 83. 13-----corosync (pid 8397) is running... gclusterd (pid 10599) is running... gcrecover (pid 10652) is running... gbased (pid 8489) is running... gc\_sync\_server (pid 10573) is running...

可以看到,在每个节点上都有输出结果,C3工具可以正常使用。

更多关于 C3 命令的介绍,可以参见《GBase 8a MPP Cluster 管理员手册》 中 2.3 C3 辅助工具的相关介绍。



# 8 FAQ

# 8.1 集群服务不能正常启动

问题:

安装集群后,集群服务不能正常启动。

答复:

通过 chkconfig --list | grep iptables 命令,确认防火墙是关闭的。

通过 sestatus 命令,确认 SELINUX 状态是 disabled。

如果防火墙是打开状态或者 selinux 开启, 那么请关闭防火墙或者 selinux 后重启集群。

## 8.2 gcadmin 命令的使用问题

问题:

安装完集群后,在操作系统的 root 用户下能使用 gcadmin 命令,而在操作系统的 gbase 用户下不能使用 gcadmin。

答复:

需要使用操作系统的 root 用户进行手动修改/var/lib/gcware 目录权限, 如下三个步骤:

\$ chown gbase:gbase /var/lib/gcware

- \$ chmod 750 /var/lib/gcware
- \$ 11 /var/lib/gcware

drwxr-x---. 2 gbase gbase 4096 Jul 4 10:24 gcware

# 8.3 是否可以指定用户登录的默认数据库



问题:

是否可以指定用户登录的默认数据库?

答复:

登录集群时是可以指定登录后默认的数据库,使用-D参数即可。

\$ gccli -uroot -p -Dtest				
Enter password:				
GBase client 8.6.1.1 build 257	776. Copyright	(c) 2004-2013	, GBase.	All Rights
Reserved.				
gbase> SELECT DATABASE() FROM	t;			
++				
DATABASE ()				
++				
test				
++				
1 row in set				

## 8.4 将 V8.5.1.2 版本集群卸载后安装新集群失败

问题:

原来集群上安装过 V8.5.1.2 版本集群,已经卸载,安装大规模集群报错。

答复:

如果环境上安装过原对称集群,需要先删除/etc/corosync/corosync.conf 文件。

## 8.5 add nodes 阶段由于异常导致失败

问题:

安装执行成功,但是 add nodes 阶段由于异常导致失败。



#### 答复:

安装集群成功后,在 add nodes 时由于集群锁住等原因导致 add nodes 执行失败时,建议排除问题后,手工执行 gcadmin addnodes dataHost.xml 命令。





微博二维码 微信二维码

