

JKDB 备份恢复模块介绍

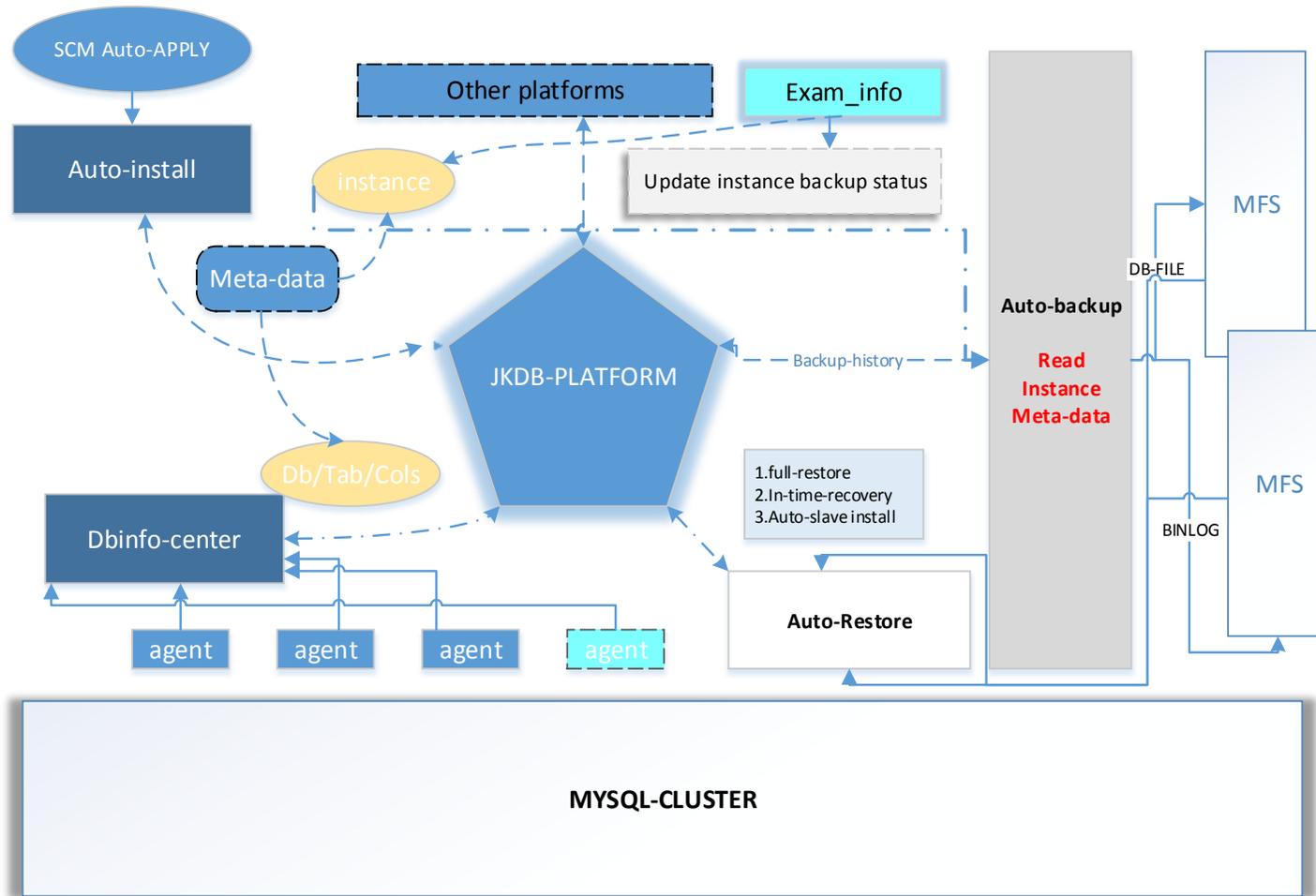
前言

JKDB (jkdb.pa.jk-ent.com) 作为 DBA 部门的整体数据管理平台，承载着整个公司的关系型数据以及缓存数据的管理与维护工作。

JKDB 平台整体上是多个功能模块的一个集合，包括了备份恢复，实例管理，装机需求，应用管理，元数据管理，容量预测，自助 workflow 等功能模块。对外 JKDB 通过权限控制，可以对业务部门暴露元数据信息，实时监控，实例状态等信息。对内 JKDB 为 DBA 部门提供了可视化的操作界面，极大的方便了日常运维与管理工作的，JKDB 作为一个容器，将诸多核心模块串联起来，最大程度的满足了流程及安全审计的需求。

此篇着重介绍 JKDB 的备份恢复模块功能以及相关灾难应急手段。

备份模块整体架构:



JKDB 备份恢复模块主要包含了以下几个功能：

1. 自动化备份（基于 JKDB 元数据信息的自动化备份）
2. 可调整化备份（可以任意的调整备份频率，备份周期以及备份节点）
3. 自动化恢复（主要包含自动备库搭建，基于备份文件的整体恢复以及基于任意时间点的恢复）
4. 备份日志审计（主要包括备份历史日志，当前备份任务展示）
5. 备份失败告警（主要依赖检验中心的检测）
6. 分布式备份存储（基于 MooseFS 的分布式文件系统）

自动化备份

基于元数据的自动化备份，所有的备份任务均基于元数据中心，元数据中心初始化由统一的安装配置模块以及 DBINFO 数据中心模块负责。

元数据主要包括 实例信息，实例角色，实例所属业务信息，实例 HA 配置信息，实例域名信息等。另外元数据还包括细粒度的信息 包括数据库分布信息，数据字段信息，表信息等。

自动化任务通过任务脚本触发，调用 XtraBackup tool 对 slave 节点做物理级别的复制，同时通过流式压缩，将备份完成的 tar 包文件传输至备份存储平台（两台 MooseFS 文件系统，一套存放 FULL-BACKUP 备份，一套存放 BINLOG 备份）。

MySQL实例信息管理

新增 删除 IP地址 分组名称 机房位置 角色 域名 监控状态

分组名称	角色	机房	IP	端口	域名	监控状态	元数据采集	操作
tddl03	主库	sh	10.129.80.71	3316	tddl03-m1.db.pajkdc.com	开启	开启	✎ ⓧ
yunzhisheng	主库	sh	10.129.80.72	3306	yunzhisheng-m1.db.pajkdc.com	开启	开启	✎ ⓧ
detection	主库	sh	10.129.80.73	3306	detection-m1.db.pajkdc.com	开启	开启	✎ ⓧ
snscenter	主库	sh	10.129.80.73	3307	snscenter-m1.db.pajkdc.com	开启	开启	✎ ⓧ
tddl00	主库	sh	10.129.80.74	3313	tddl00-m1.db.pajkdc.com	开启	开启	✎ ⓧ
tddl01	主库	sh	10.129.80.74	3314	tddl01-m1.db.pajkdc.com	开启	开启	✎ ⓧ
item	主库	sh	10.129.80.75	3307	item-m1.db.pajkdc.com	开启	开启	✎ ⓧ
point	主库	sh	10.129.80.75	3306	point-m1.db.pajkdc.com	开启	开启	✎ ⓧ
friendsoutput	主库	sh	10.129.80.76	3306	friendsoutput-m1.db.pajkdc.com	开启	开启	✎ ⓧ
promotion	主库	sh	10.129.80.76	3307	promotion-m1.db.pajkdc.com	开启	开启	✎ ⓧ
accounts00	主库	sh	10.129.80.77	3313	accounts00-m1.db.pajkdc.com	开启	开启	✎ ⓧ
accounts01	主库	sh	10.129.80.77	3314	accounts01-m1.db.pajkdc.com	开启	开启	✎ ⓧ
accounts02	主库	sh	10.129.80.77	3315	accounts02-m1.db.pajkdc.com	开启	开启	✎ ⓧ
accounts03	主库	sh	10.129.80.77	3316	accounts03-m1.db.pajkdc.com	开启	开启	✎ ⓧ
accounts04	主库	sh	10.129.80.77	3317	accounts04-m1.db.pajkdc.com	开启	开启	✎ ⓧ
silver	主库	sh	10.129.80.77	3306	silver-m1.db.pajkdc.com	开启	开启	✎ ⓧ
jkshop	主库	sh	10.129.80.78	3306	jkshop-m1.db.pajkdc.com	开启	开启	✎ ⓧ

Showing 51 to 100 of 477 rows 50 records per page

整个实例信息包括监控开启状态，元数据采集状态，备份状态等。备份策略分为（备份频率，备份保留天数，是否备份）默认我们只备份 ROLE=2（备库），并且 Enable 备份属性的实例。备份保留天数为备份文件在本地的保留天数（所有的备份均存在分布式文件系统中）。

主
数据中心
装机平台
配置平台
备份平台
监控平台
容量平台
服务平台
消息平台
线上
liuyang315

MySQL备份策略

IP	端口	分组名称	机房	是否备份	正在备份	最近备份时间	备份频率(天)	备份保留天数(天)	下次备份时间	角色	操作
10.129.128.88	3306	dns	sh	是	否	2017-05-10 19:00:39	1	10	2017-05-11 19:00:00	从库	Q
10.129.80.104	3307	httpdns	sh	是	否	2017-05-11 13:00:49	1	10	2017-05-12 13:00:00	从库	Q
10.129.80.104	3306	clouder	sh	是	否	2017-05-11 12:09:25	1	10	2017-05-12 12:02:00	从库	Q
10.129.80.58	3307	hc	sh	是	否	2017-05-10 19:19:54	2	5	2017-05-12 16:45:00	从库	Q
10.129.80.58	3308	jsumo	sh	是	否	2017-05-10 15:29:26	1	10	2017-05-11 15:10:00	从库	Q
10.129.80.58	3306	sims	sh	是	否	2017-05-10 22:45:43	1	5	2017-05-11 15:10:00	从库	Q
10.129.80.69	3313	oscar	sh	是	否	2017-05-10 19:00:29	1	10	2017-05-11 18:59:00	从库	Q
10.129.12.70	3306	love	sh	是	否	2017-05-10 18:01:41	3	10	2017-05-13 18:00:00	从库	Q
10.129.12.71	3306	jufeng	sh	是	否	2017-05-11 12:07:54	2	10	2017-05-13 11:53:00	从库	Q
10.129.80.117	3313	tcorder00	sh	是	否	2017-05-10 12:17:07	2	10	2017-05-12 12:00:00	从库	Q
10.129.80.117	3314	tcorder01	sh	是	否	2017-05-10 12:26:44	2	10	2017-05-12 12:10:00	从库	Q
10.129.80.117	3315	tcorder02	sh	是	否	2017-05-10 12:32:22	2	10	2017-05-12 12:15:00	从库	Q
10.129.80.117	3316	tcorder03	sh	是	否	2017-05-10 12:46:55	2	10	2017-05-12 12:30:00	从库	Q
10.129.80.117	3317	tcorder04	sh	是	否	2017-05-10 13:01:56	2	10	2017-05-12 12:45:00	从库	Q
10.129.80.117	3318	tcorder05	sh	是	否	2017-05-10 13:16:53	2	10	2017-05-12 13:00:00	从库	Q
10.129.80.117	3319	tcorder06	sh	是	否	2017-05-10 13:32:00	2	10	2017-05-12 13:15:00	从库	Q
10.129.80.117	3320	tcorder07	sh	是	否	2017-05-10 13:47:21	2	10	2017-05-12 13:30:00	从库	Q

Showing 1 to 50 of 233 rows
 records per page
1
2
3
4
5

自动化恢复

自动化恢复也是基于元数据的恢复模块（分为自动化恢复以及基于时间点的恢复）。

普通的自动化恢复分为以下三个功能：

1. 只做备份恢复，仅仅恢复一个完整的备份集
2. 将恢复的备份加入到集群中成为一个 slave 节点
3. 将恢复的备份加入到集群中成为一个级联状态的 slave 节点

XtraBackup Slave Reocver

😊 功能说明：恢复数据至任意备份文件，或者使用备份文件在任意主机恢复出新的备库。

备库实例	请检索
备份文件	请选择
恢复主机	请检索
恢复端口	请选择
恢复方式	<input checked="" type="radio"/> 只恢复数据，不创建Slave <input type="radio"/> 恢复为Master的Slave <input type="radio"/> 恢复为Slave的Slave
开启同步	<input checked="" type="radio"/> 开启 <input type="radio"/> 不开启

✓ 提交任务



基于时间点的恢复可以恢复数据库至 3 个月内的任意时间点。

通过我们的 BINLOG 订阅系统，自动解析 LOG 从而生成恢复语句逻辑恢复至任意指定时间点。

基于时间点的恢复从根本上是依赖自动化备份任务的全备集合（通过从分布式备份存储集群上拉取历史备份集合，结合相关 BINLOG 完成逻辑恢复）。

xtrabackup Binlog roll forward

😊 功能说明：通过XtraBackup+Binlog Server将数据恢复至任意一个指定的时间点。使用流程如下：选择备份实例->选择备份文件->选择恢复目标主机和端口->选择恢复的时间点(例如：2017-01-01 13:00:00)->提交任务。

备份主机IP	请检索
备份文件	请选择
恢复主机	请检索
恢复端口	请选择
恢复时间点	恢复时间点
✓ 提交任务	



备份历史以及监控

所有的备份记录均以 LOG 形式留存以供审计需要，当然我们也会定期的对线上的备份数据集做相关恢复验证工作（定期恢复指定 GROUP 的备份集，定期恢复相关时间点的备份以验证备份可用性）。

对当前的备份任务，我们会通过校验中心做相关预警工作。



MySQL当前备份任务

IP	端口	备份状态	备份信息	备份开始时间	备份结束时间	文件传输开始时间	文件传输结束时间	任务更新时间
10.129.128.48	3306	xtrabackup failed	xtrabackup failed ,please check log	2016-11-28 18:54:02	-	-	-	2016-11-28 18:54:02
10.129.128.52	3306	xtrabackup failed	xtrabackup failed ,please check log	2016-11-28 18:54:02	-	-	-	2016-11-28 18:54:02
10.129.80.14	3306	xtrabackup failed	xtrabackup failed ,please check log	2016-12-26 13:00:19	-	-	-	2016-12-26 14:12:07
10.255.12.18	3306	xtrabackup failed	xtrabackup failed ,please check log	2017-04-27 14:41:28	-	-	-	2017-04-27 14:43:54
10.255.128.17	3306	xtrabackup failed	xtrabackup failed ,please check log	2017-04-27 14:41:28	-	-	-	2017-04-27 14:43:54
10.129.80.47	3306	waiting for xtrabackup	start xtrabackup	2017-05-11 11:56:54	-	-	-	2017-05-11 11:56:54
10.129.80.27	3307	waiting for xtrabackup	start xtrabackup	2017-05-11 12:00:55	-	-	-	2017-05-11 12:00:55
10.129.80.37	3308	waiting for xtrabackup	start xtrabackup	2017-05-11 12:00:55	-	-	-	2017-05-11 12:00:55
10.129.80.48	3318	waiting for xtrabackup	start xtrabackup	2017-05-11 12:15:56	-	-	-	2017-05-11 12:15:56
10.129.80.48	3319	waiting for xtrabackup	start xtrabackup	2017-05-11 12:35:59	-	-	-	2017-05-11 12:35:59
10.129.80.48	3320	waiting for xtrabackup	start xtrabackup	2017-05-11 13:15:04	-	-	-	2017-05-11 13:15:04
10.129.80.42	3315	3	xtrabackup success scp ing	2017-05-11 13:40:07	2017-05-11 15:02:45	2017-05-11 15:02:45	-	2017-05-11 15:02:45
10.129.80.43	3320	waiting for xtrabackup	start xtrabackup	2017-05-11 13:45:07	-	-	-	2017-05-11 13:45:07
10.129.80.48	3321	waiting for xtrabackup	start xtrabackup	2017-05-11 13:45:07	-	-	-	2017-05-11 13:45:07
10.129.80.138	3321	waiting for xtrabackup	start xtrabackup	2017-05-11 13:46:07	-	-	-	2017-05-11 13:46:07
10.129.80.42	3316	waiting for xtrabackup	start xtrabackup	2017-05-11 14:00:09	-	-	-	2017-05-11 14:00:09
10.129.80.43	3321	waiting for xtrabackup	start xtrabackup	2017-05-11 14:00:09	-	-	-	2017-05-11 14:00:09

Showing 1 to 24 of 24 rows



校验中心模块会抓取备份失败的信息，并以邮件、短信的形式告知。

校验类型	校验异常信息	IP地址	端口	分组名称	校验时间	
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>		待处理 ▾
check_instance	group:dental slave readonly is not enabled	10.129.80.45	3307	dental	2016-10-27 13:56:47	响应 忽略
check_instance	group:livediablo slave readonly is not enabled	10.129.80.38	3306	livediablo	2016-10-27 13:56:46	响应 忽略
check_instance	group:yunzhisheng slave readonly is not enabled	10.129.80.72	3306	yunzhisheng	2016-10-27 13:56:44	响应 忽略
check_machine	hostnames is inconsistent with real ip	10.129.80.119	10086	dbagroup	2016-10-27 13:56:43	响应 忽略
check_instance	group:dental slave readonly is not enabled	10.129.80.90	3307	dental	2016-10-27 13:56:39	响应 忽略
check_instance	group:livediablo slave readonly is not enabled	10.129.80.84	3306	livediablo	2016-10-27 13:56:37	响应 忽略
check_instance	group:openfalcon slave readonly is not enabled	10.129.80.95	3306	openfalcon	2016-10-27 13:56:37	响应 忽略
check_instance	group:webcast slave readonly is not enabled	10.129.80.39	3306	webcast	2016-10-27 13:56:34	响应 忽略
check_machine	hostnames is inconsistent with real ip	10.128.6.102	10086	dbagroup	2016-10-27 13:56:33	响应 忽略
check_machine	hostnames is inconsistent with real ip	10.129.80.55	10086	dbagroup	2016-10-27 13:56:33	响应 忽略
check_instance	group:jufeng slave readonly is not enabled	10.129.12.71	3306	jufeng	2016-10-27 13:56:33	响应 忽略

backup | 当前文件夹

全部 未读 按日期 最新 ↓

上个月

- db_alarm@service.jk.cn
exam_info 2017/4/13
ip port group_name
- db_alarm@service.jk.cn
exam_info 2017/4/7
ip port group_name

更早

- db_alarm@service.jk.cn
exam_info 2017/3/29
ip port group_name
- db_alarm@service.jk.cn
exam_info 2017/3/28
ip port group_name
- db_alarm@service.jk.cn
exam_info 2017/3/28
ip port group_name
- db_alarm@service.jk.cn
exam_info 2017/3/28
ip port group_name
- db_alarm@service...
exam_info 2017/3/28
ip port group_name
- db_alarm@service.jk.cn
exam_info 2017/3/27
ip port group_name

答复 全部答复 转发



2017/3/29 (周三) 17:56
db_alarm@service.jk.cn
exam_info

收件人 DBA团队邮件组

ip	port	group_name	info
10.255.128.17	3306	devdb	group:devdb mha thread not exist
10.255.128.17	3306	devdb	group:devdb dbmon is not running
10.255.128.17	3306	devdb	group:devdb slave backup is not enabled
10.255.128.17	3306	devdb	group:devdb slave failed to backup
10.255.128.17	10086	dbagroup	sec is down
10.255.132.19	3306	devdb	group:devdb mha thread not exist
10.255.132.19	3306	devdb	group:devdb dbmon is not running
10.255.132.19	10086	dbagroup	sec is down

灾难应急

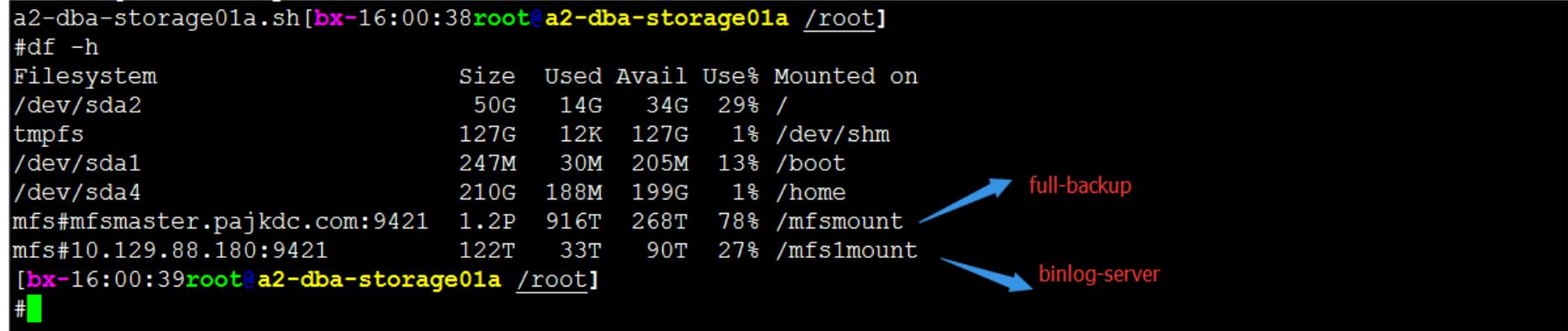
首先基于应用级别的，我们依赖自身的 HA 模块可以秒级的切换至健康的节点(对前端应用无感知), 通过自动恢复模块可以快速的搭建新的主备集群。

未来基于机房级别的容灾手段:

1. 远程灾备机房 HA 容灾 (同城、跨省、跨国机房)
2. 远程备份文件传输至 IDC 进行自动化恢复 (结合 DPDK 等技术, 基于整个实例级别的 FULL-BACKUP , IN-TIME-RECOVERY 恢复)。

最后放一张我们备份集群的截图:

```
a2-dba-storage01a.sh[bx-16:00:38root@a2-dba-storage01a /root]
#df -h
Filesystem                Size  Used Avail Use% Mounted on
/dev/sda2                  50G   14G   34G   29% /
tmpfs                      127G   12K  127G    1% /dev/shm
/dev/sda1                  247M   30M  205M   13% /boot
/dev/sda4                  210G  188M  199G    1% /home
mfs#mfsmaster.pajkdc.com:9421 1.2P  916T  268T   78% /mfsmount
mfs#10.129.88.180:9421      122T   33T   90T   27% /mfs1mount
[bx-16:00:39root@a2-dba-storage01a /root]
#
```



full-backup

binlog-server