

京东数据库备份系统进化之路

京东商城

张 洋

- **备份的重要性**
- **备份的用途**
- **备份系统演进**
- **遇到的挑战**
- **未来规划**

备份的重要性



事故

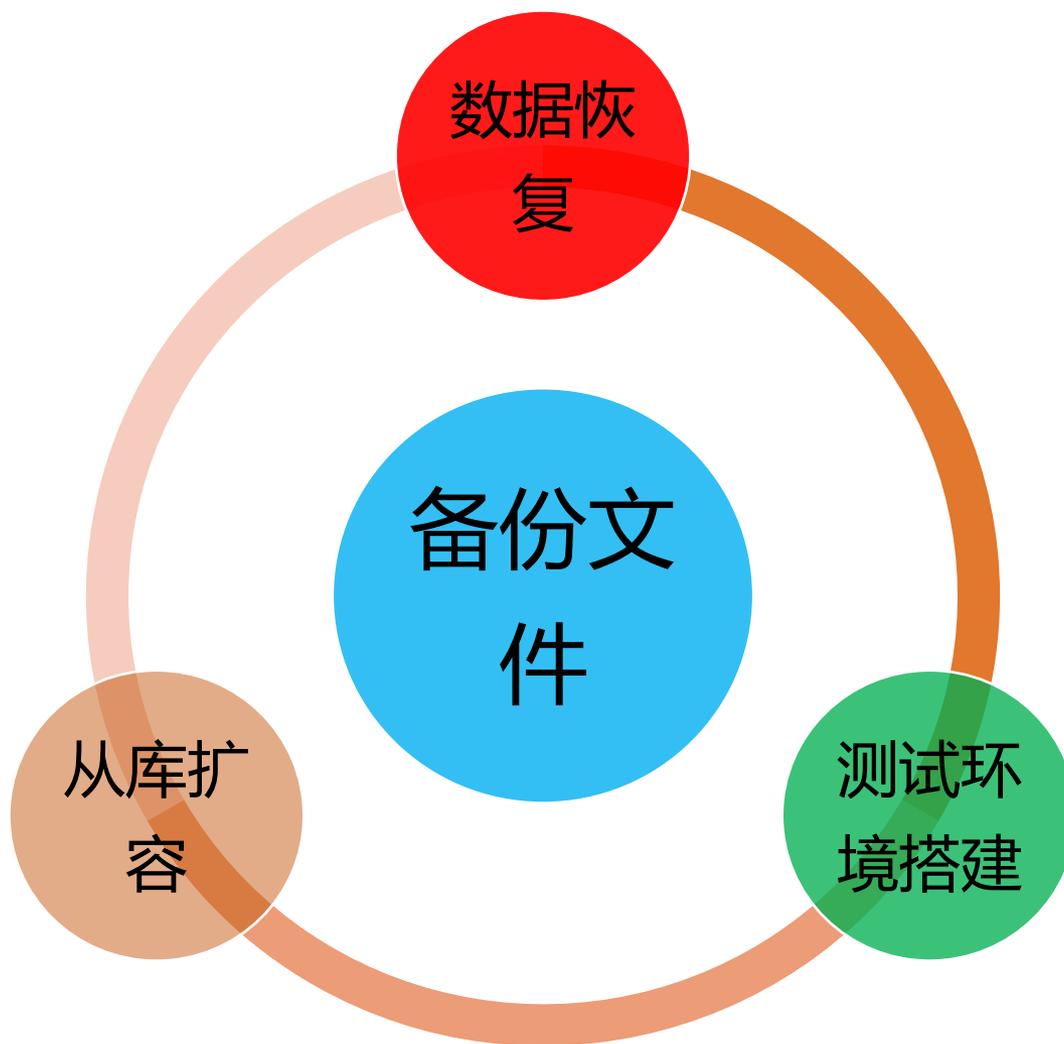
```
127.0  
ERROR  
hout  
127.0
```

table wit



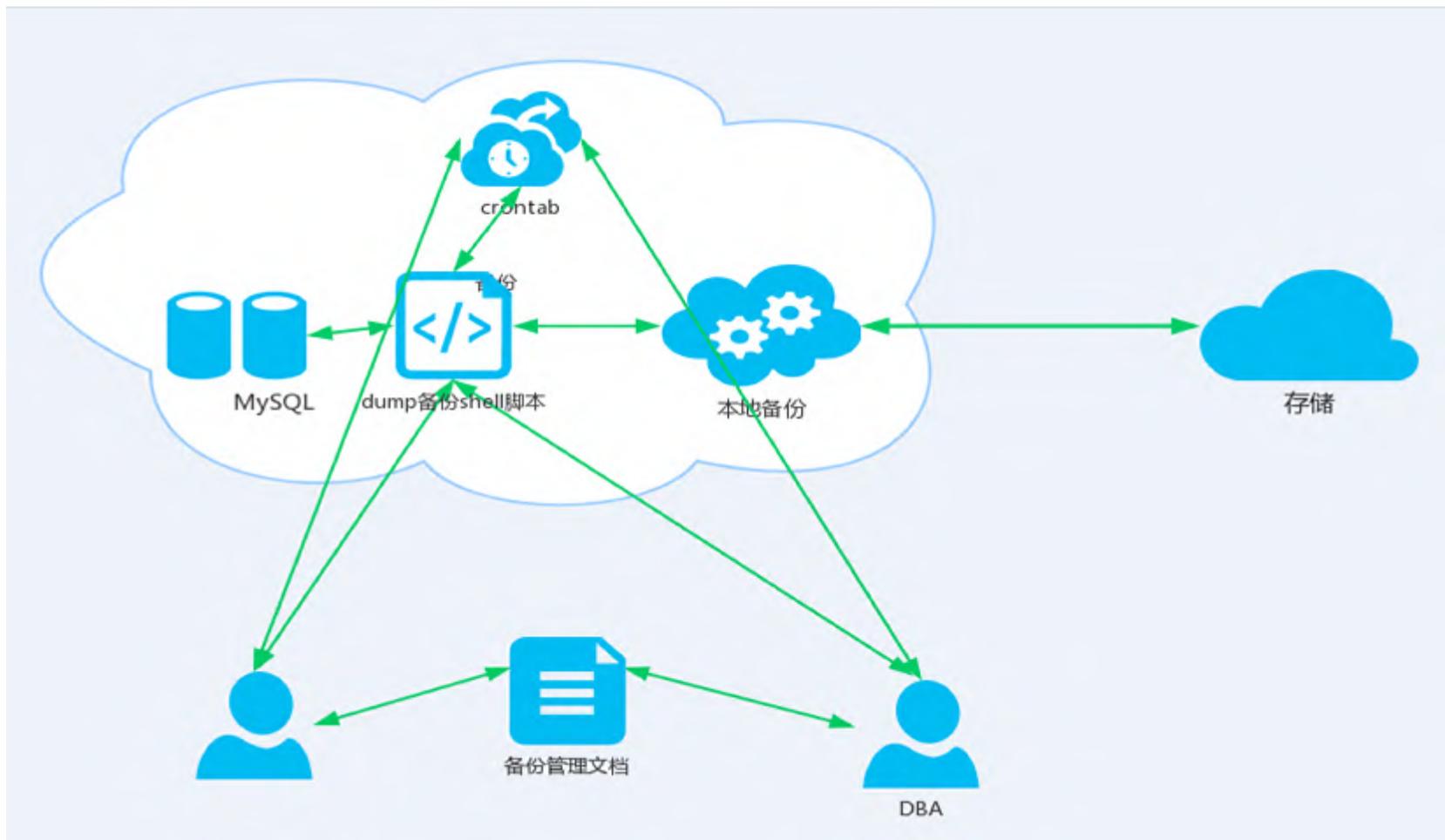
的

备份的用途



备份系统的演进

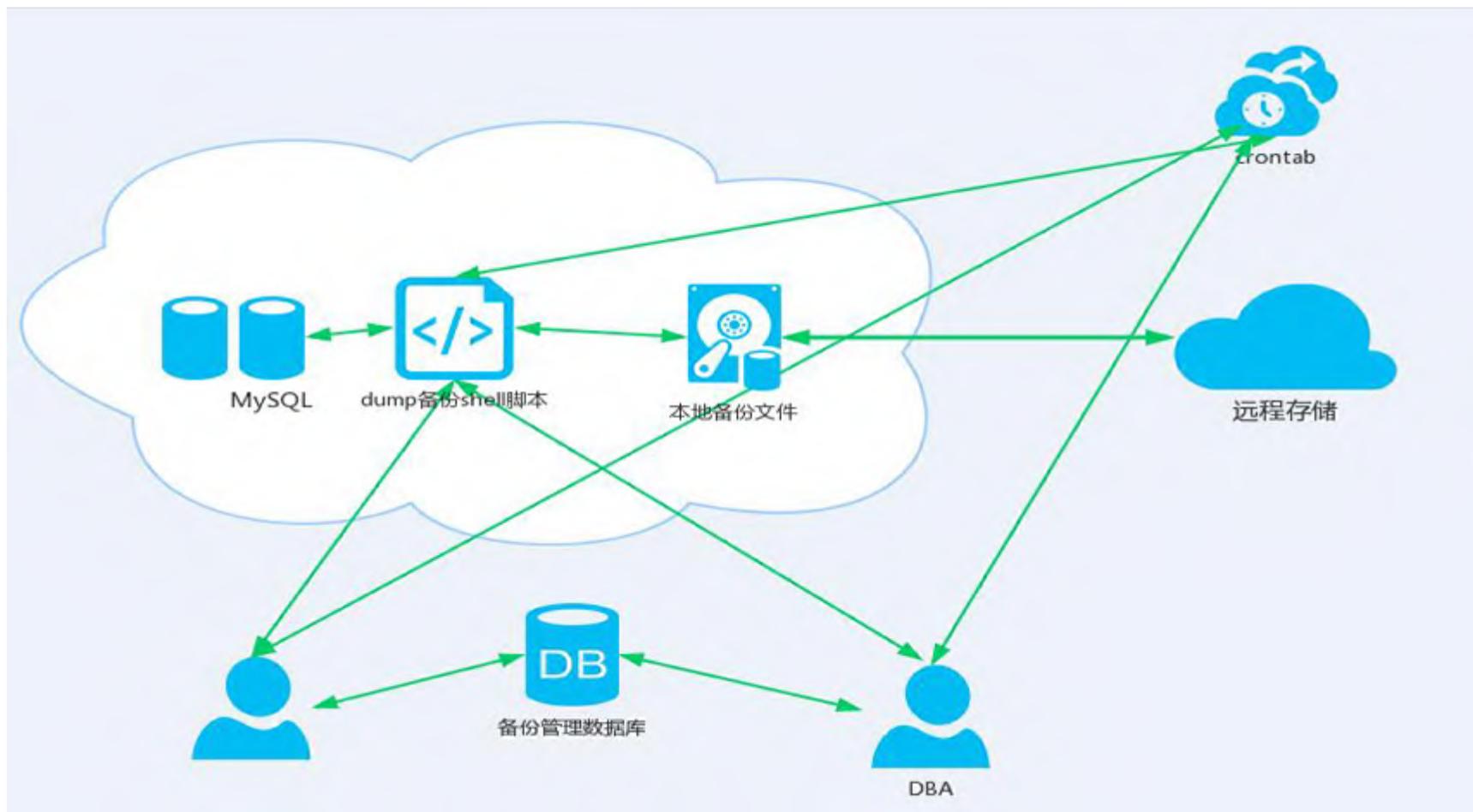
1.0版本



面临的问题

- 管理文档难以维护
- 备份信息较难获取
- 备份服务器及存储维护工作量较大

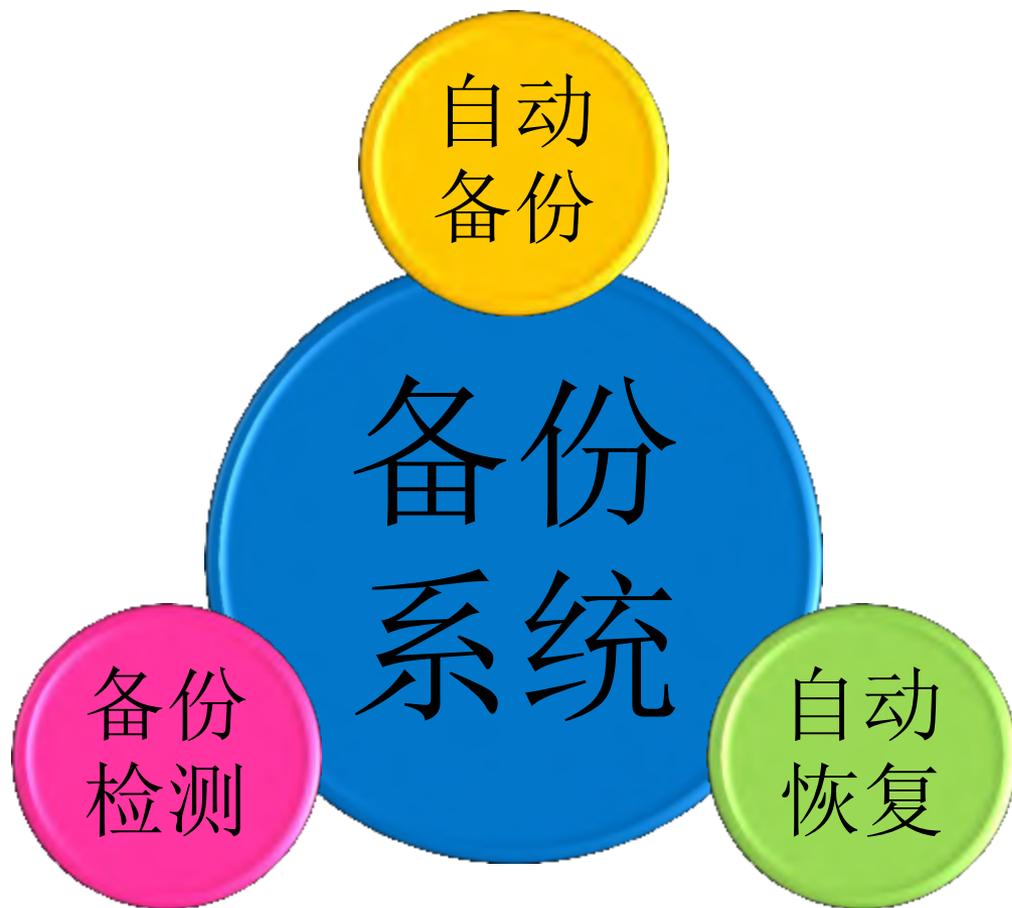
2.0版本



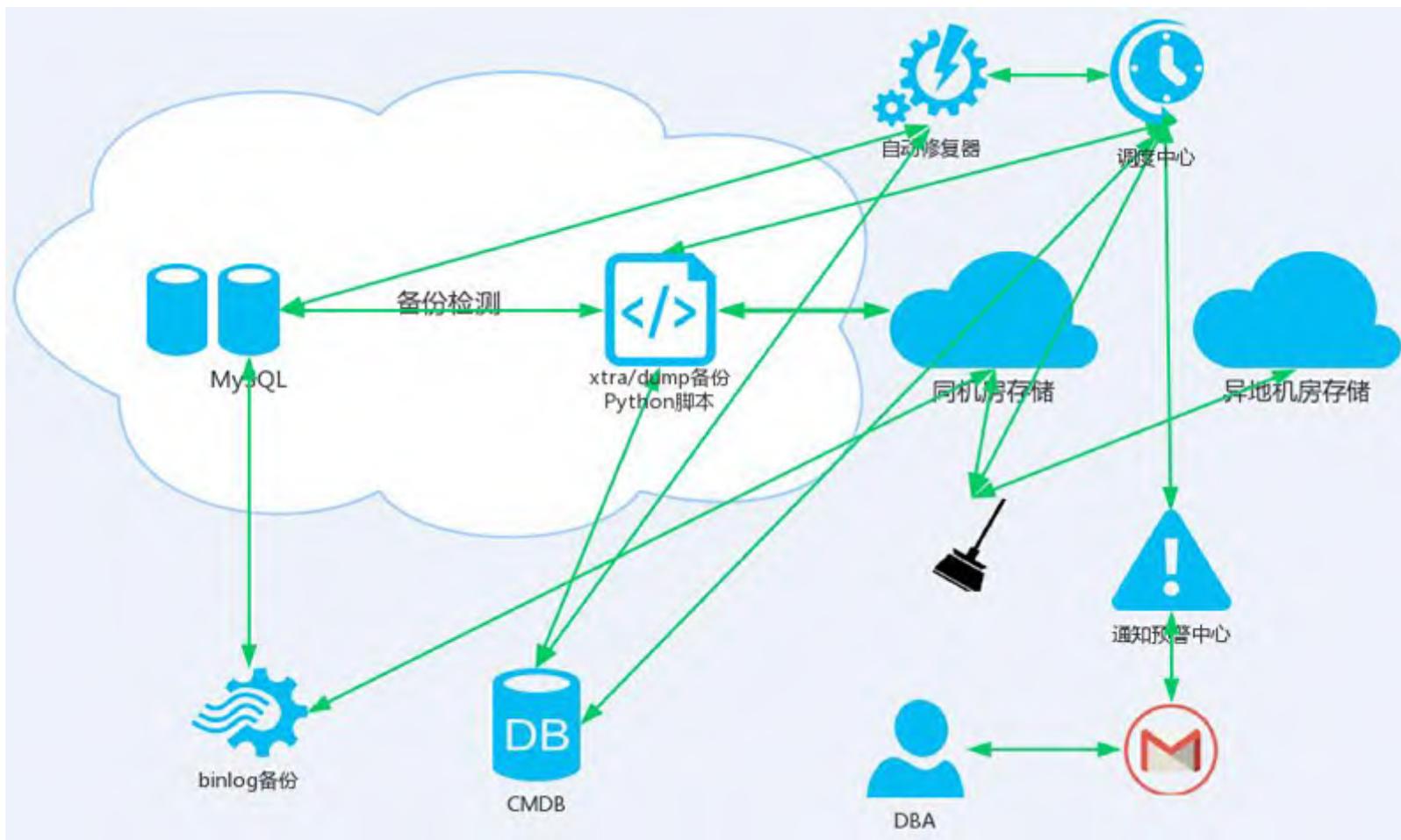
面临的问题

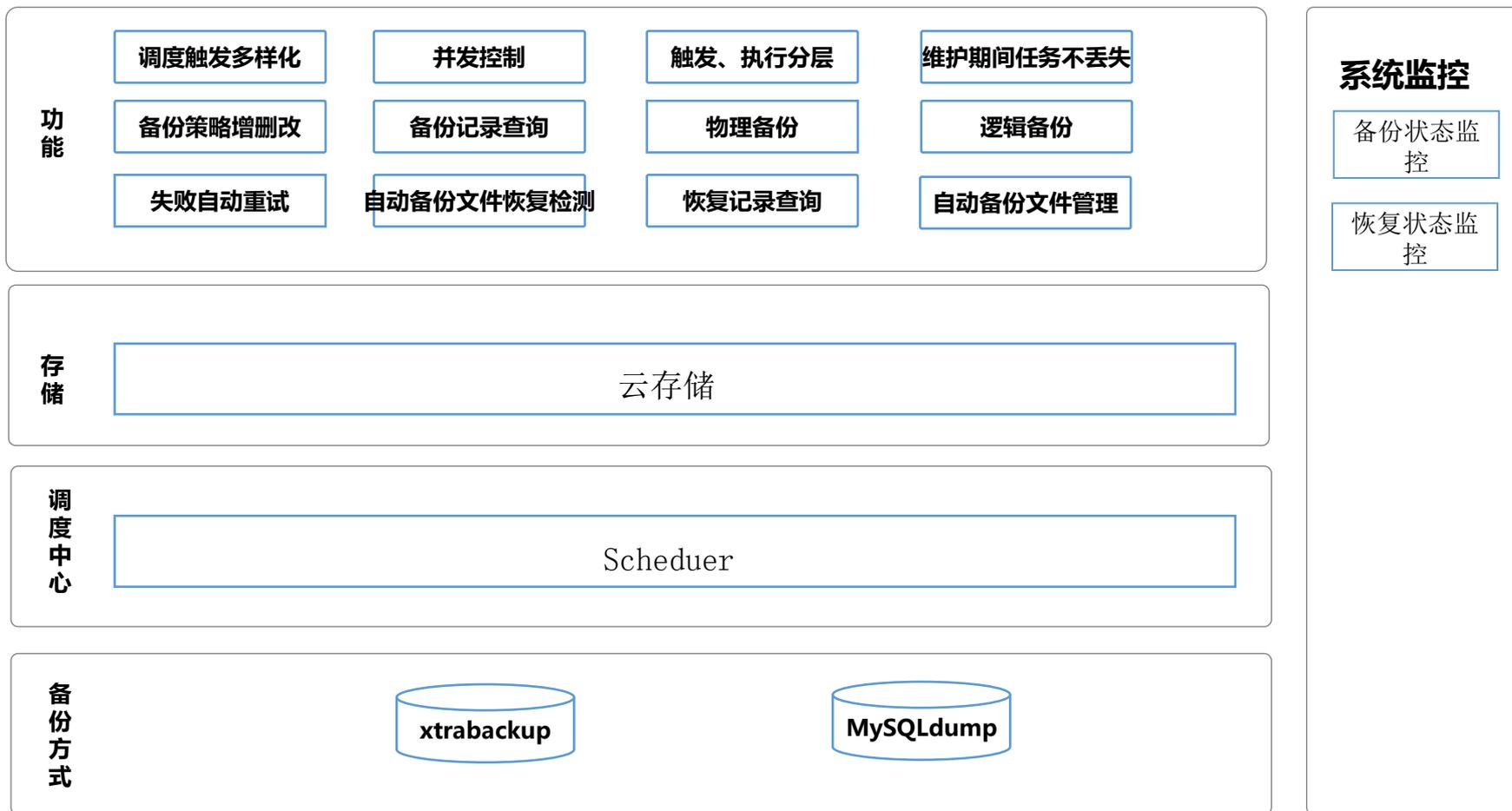
- **Crontab**调度时间太集中
- 调度服务器维护了
- **Shell**脚本无法捕捉异常
- 备份进程**hang**死
- 备份策略添加状态获取较麻烦
- **Dump**恢复时间太长
- 备份文件可用性不明确

3.0版本



自动备份





备份方式的选择

Dump逻辑备份

- 跨版本
- 备份文件小
- 内容可以编辑
- 与存储引擎无关



Xtra物理备份

- 备份速度快
- 对线上影响小
- 恢复速度快
- 数据一致性好



调度方式的对比



APScheduler

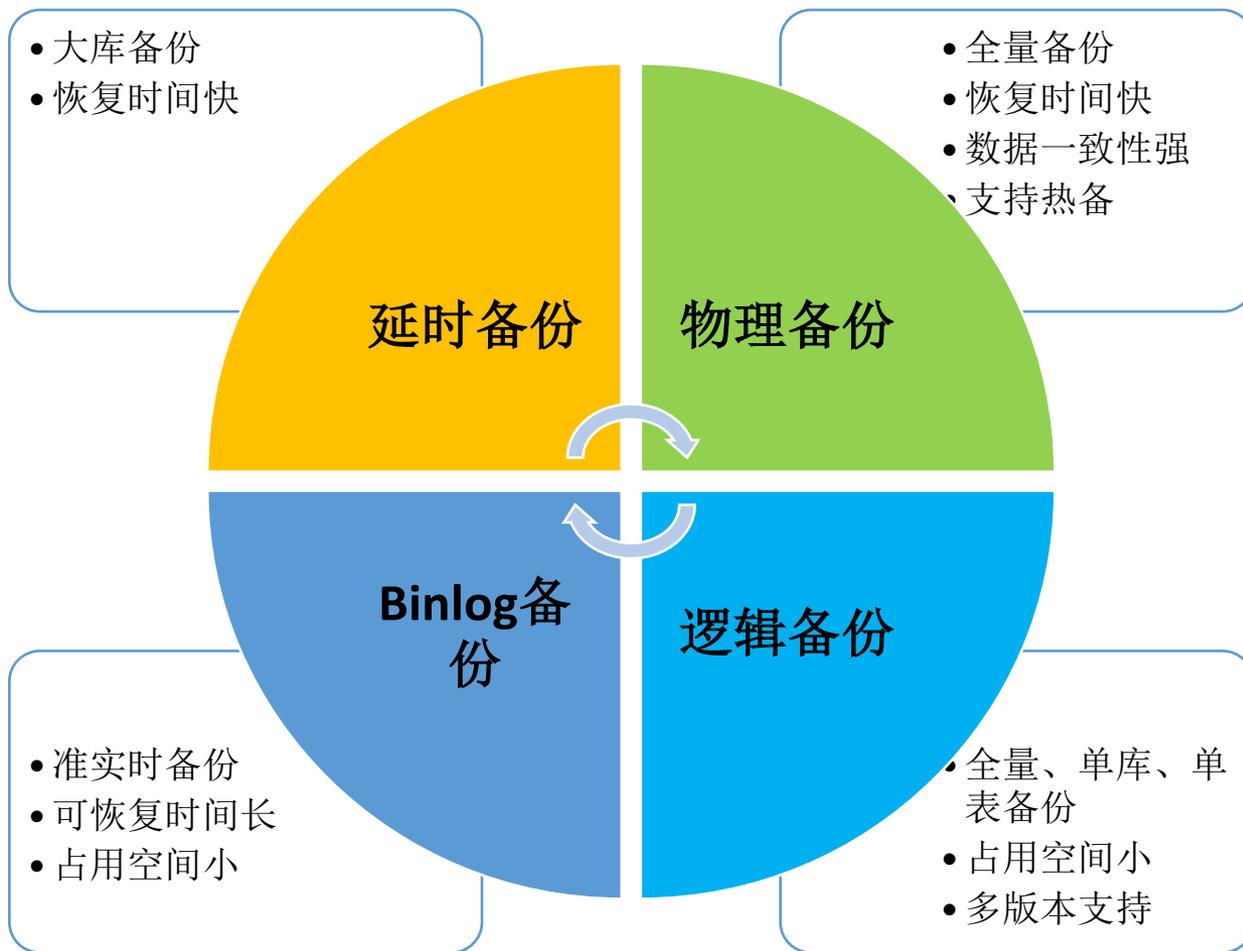
- 维护期任务不丢失
- 支持interval、crontab、date
- 数据库持久化
- 拥有线程池
- 管理简单

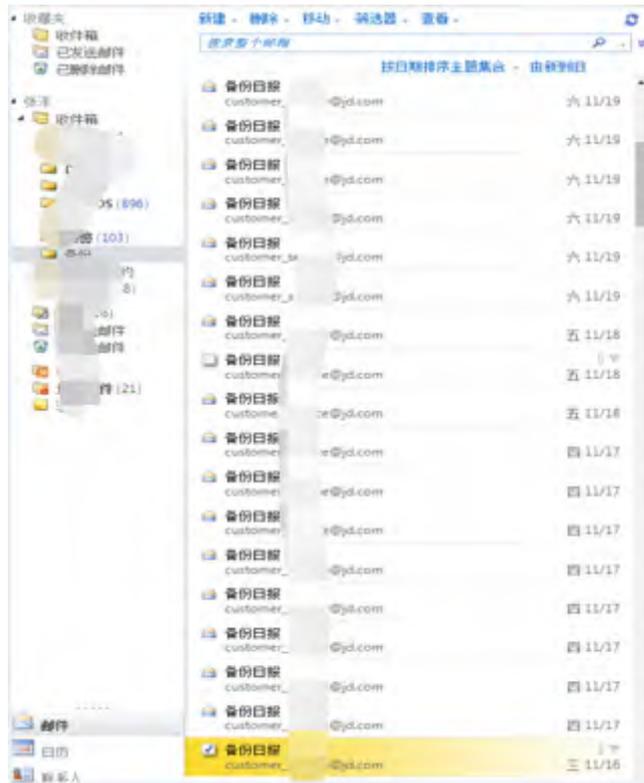


Crontab

- 维护期任务丢失
- 支持crontab
- 文件持久化
- 管理繁琐

备份方式多样化





备份日报

jd.com

备份汇总		
备份总数	备份成功数	备份成功率
359	359	100.000%

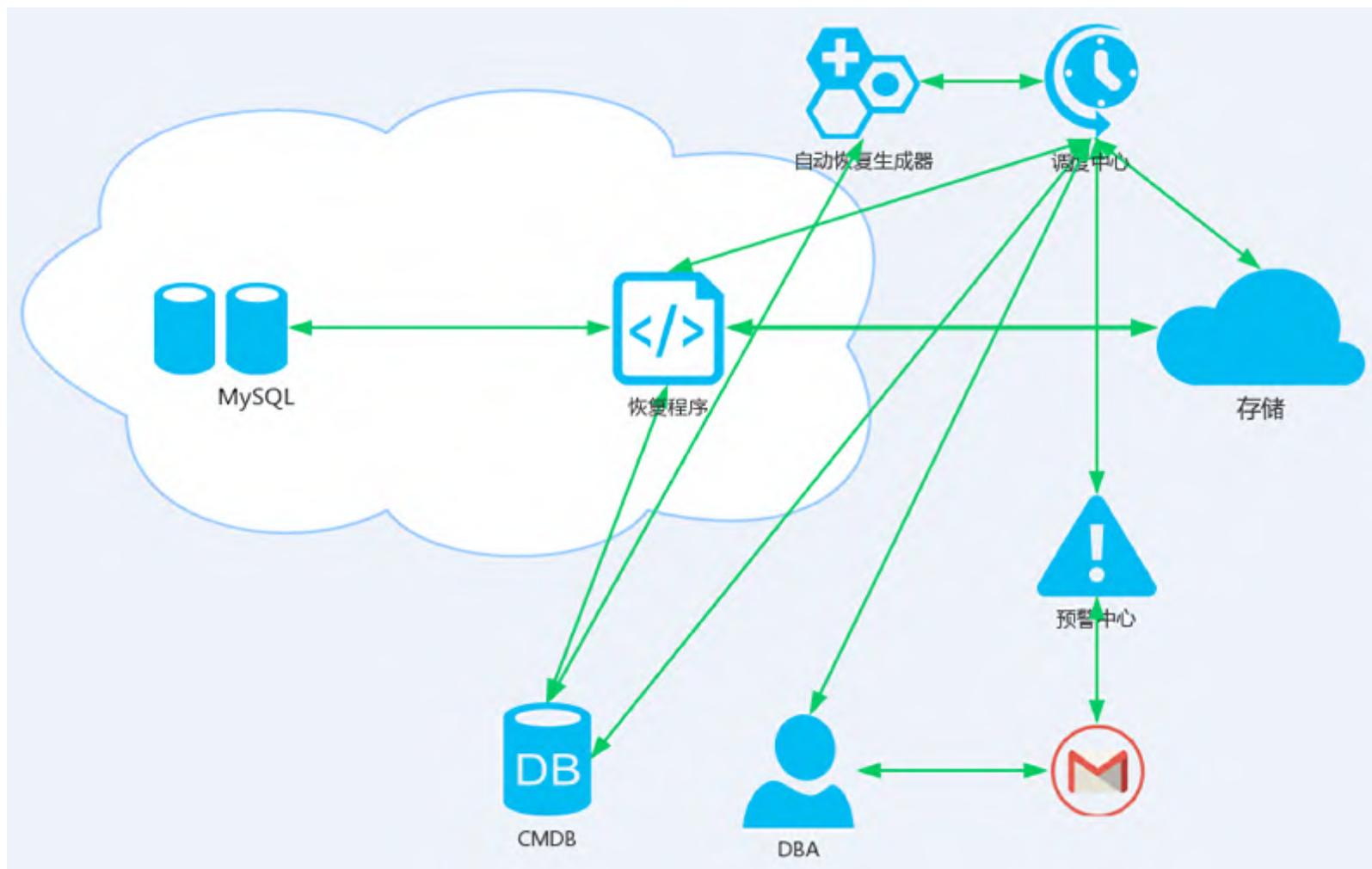
未启用备份策略列表				
项目名称	集群名称	Cluster ID	DBA	备份状态
转	159			None
目				None
目	1			None
控	loud_ump			None
控	loud_umpreport			None
控平台	port			None
控平台	5			None
数据库	406	06		None
数据库	2	37		None

7天内没有备份记录列表							
项目名称	集群名称	Cluster ID	备份机器IP	调度类型	调度参数	备份策略开始时间	DBA
中			my13	interval	(interval_time_type'days',interval_time_length:2)	2016-11-05 09:33:45	
自			my	interval	(interval_time_type'days',interval_time_length:2)	2016-11-07 16:21:29	
自			cd	interval	(interval_time_type'days',interval_time_length:2)	2016-11-12 16:48:00	ing

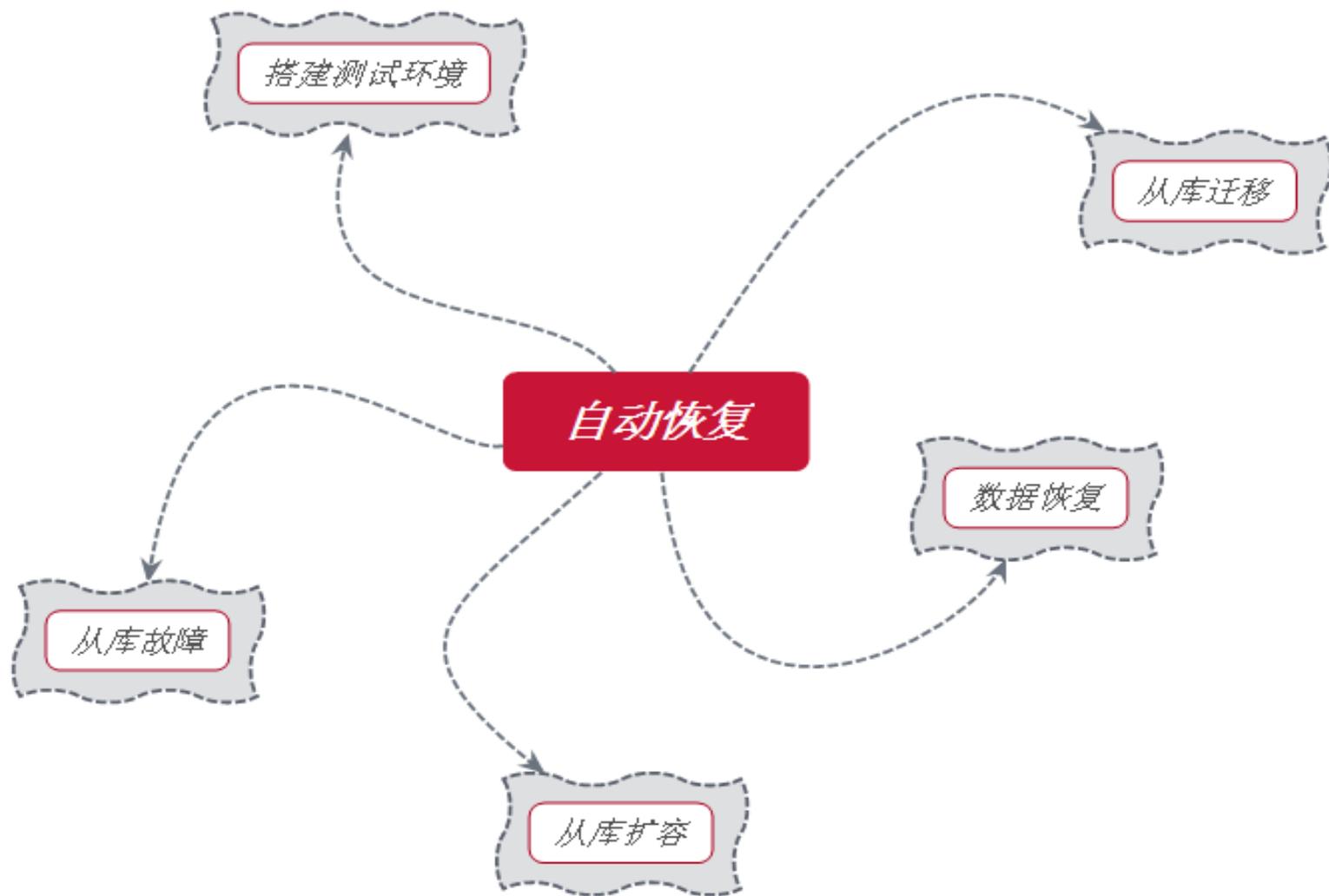
备份失败列表

项目名称	集群名称	Cluster ID	备份机器IP	备份信息	无备份天数	DBA
转	1384		222	xtra backup error	1	
票	004		22	xtra backup error	1	
去-承	去-承运商单	5	04	checksum failed	1	z'

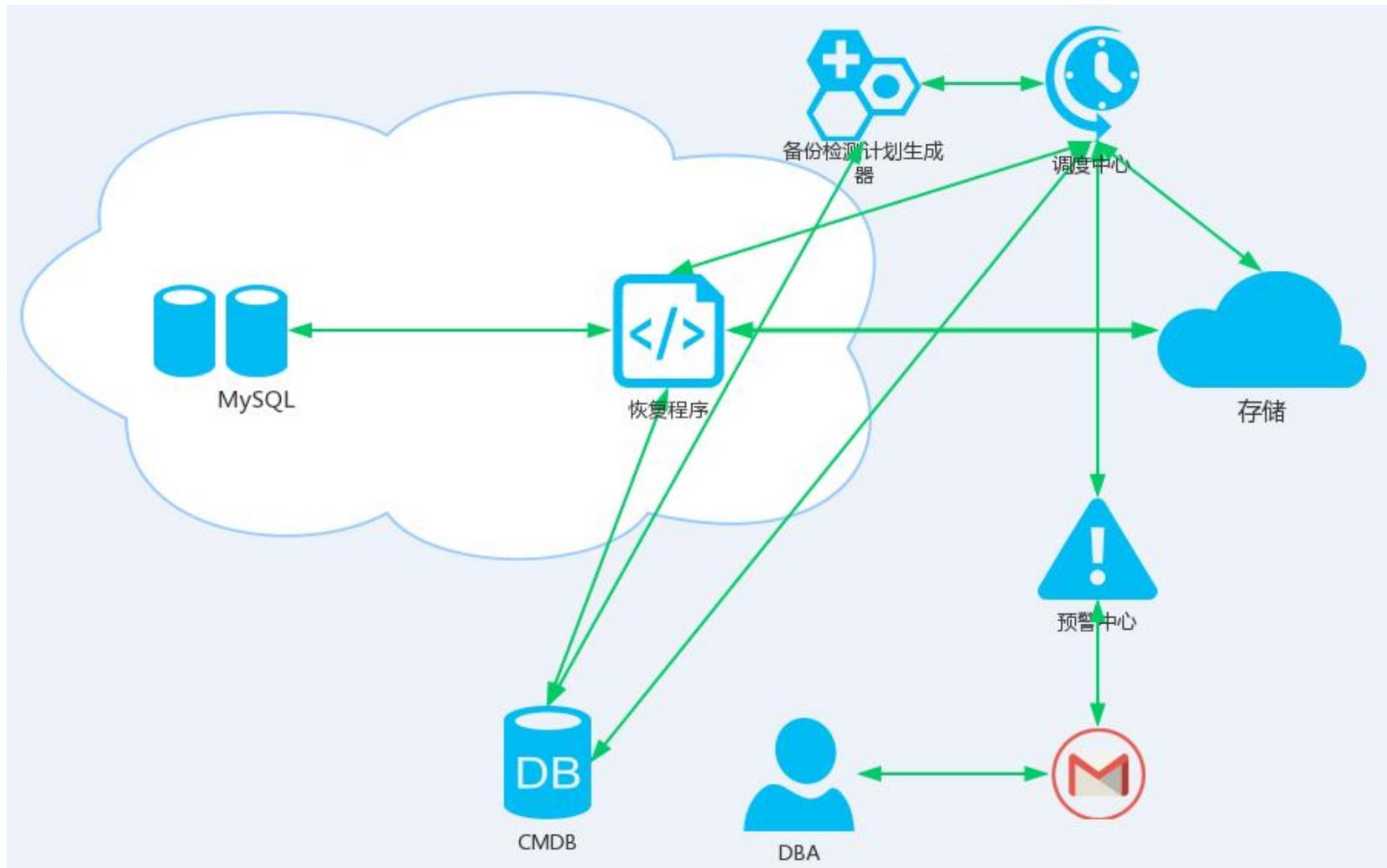
自动恢复



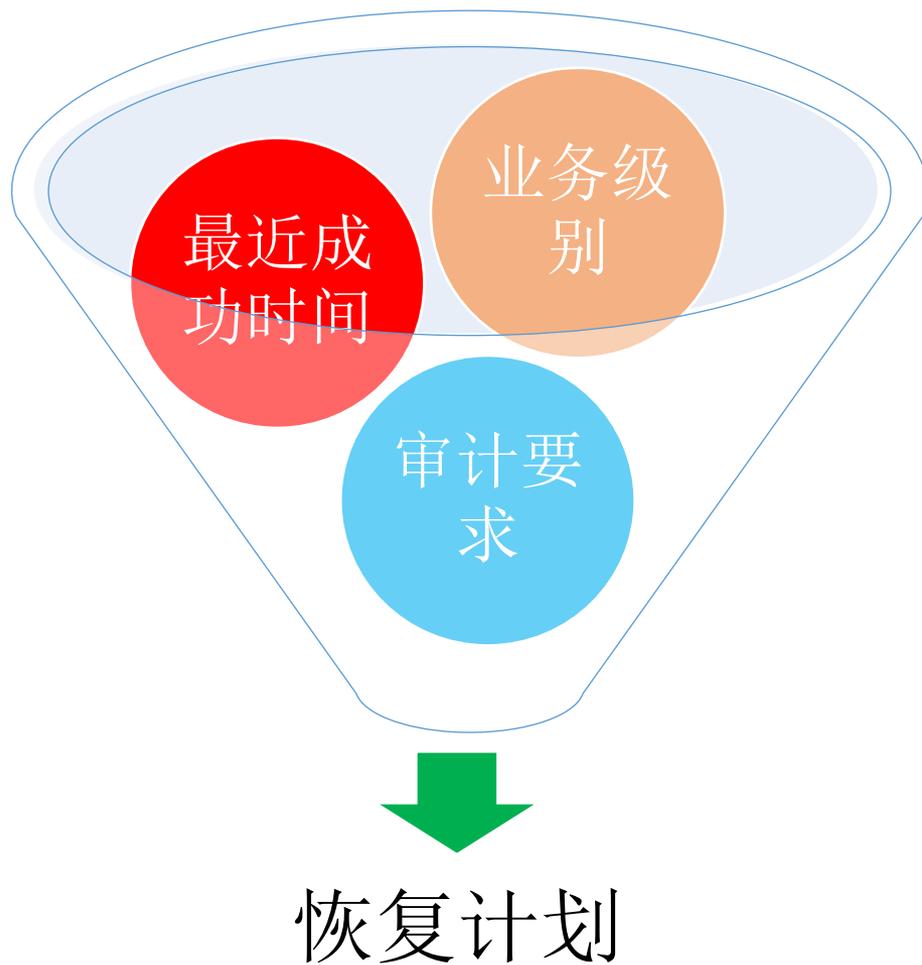
自动恢复



备份检测



备份检测



一些数据

备份集群**795**套

恢复检测**3246**次

备份**36037**次

备份文件总大小**1362T**

上线**60**天

项目**365**个

成功率**999**

- 存储未上云
- 有些老库redolog太小，**xtrabackup**不能成功
- 备份策略需要人工添加
- 大表**xtrabackup**容易失败
- 执行备份程序使用**ansible**，`python select.select()`
打开句柄数字达到**1024**后报
ValueError: filedescriptor out of range in select()

- 使用JSS云存储
- 规范化数据库配置
- 从库分类自动添加备份策略

谢谢！
Thank you!