

深信服数据库管理平台 DMP-技术主打PPT

目录

一、DMP整体技术框架

二、高可靠

1. 集群架构部署
2. 自愈能力
3. 备份保护
4. 监控体系
5. 底层可靠

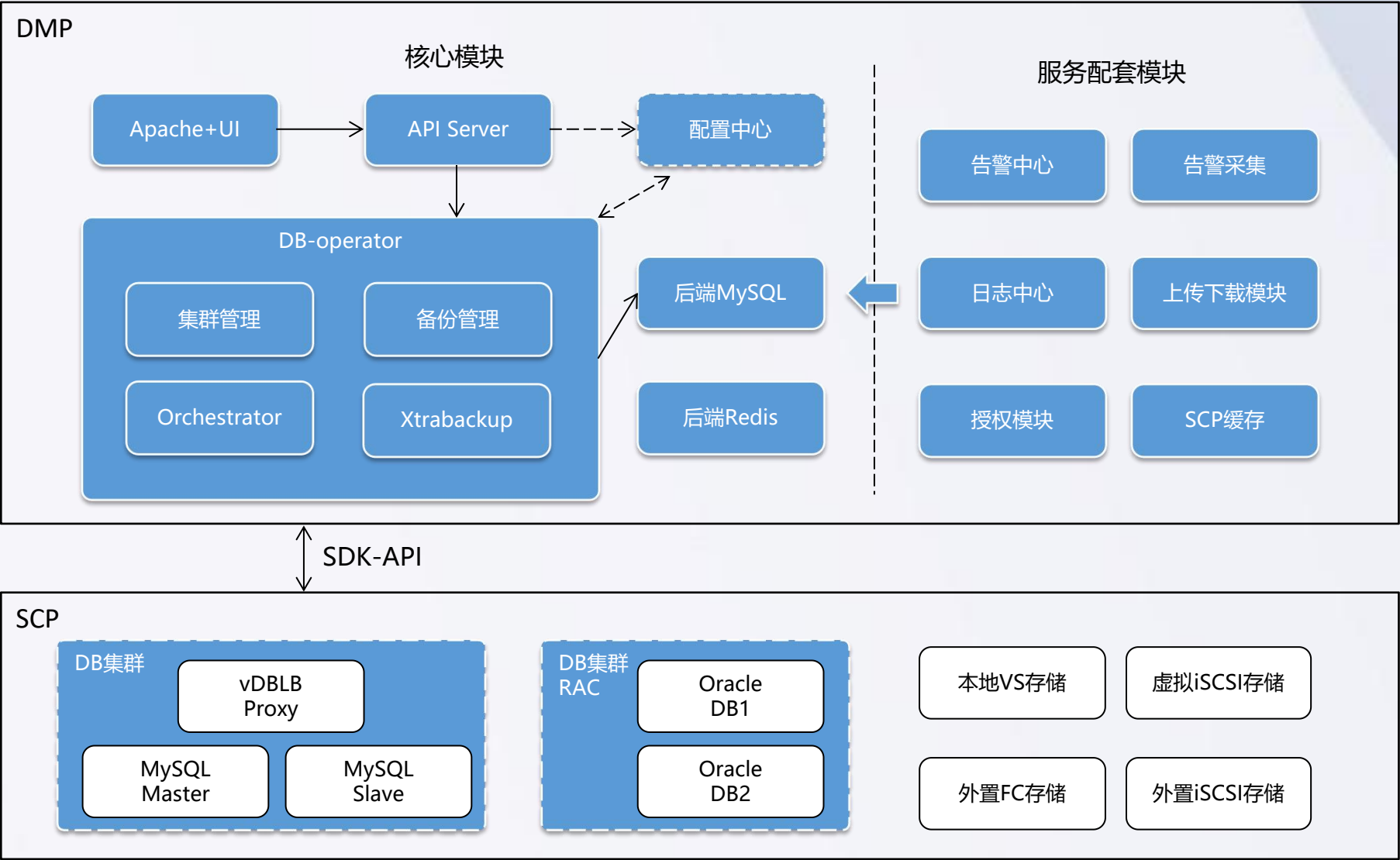
三、高性能

1. 最佳实践调优
2. 存储层优化
3. 性能测试报告
4. SQL级性能诊断

四、易运维

1. 业务快速上线
 - Oracle DTS
 - 自动化部署
2. 生命周期管理
3. 租户自服务

DMP整体技术框架



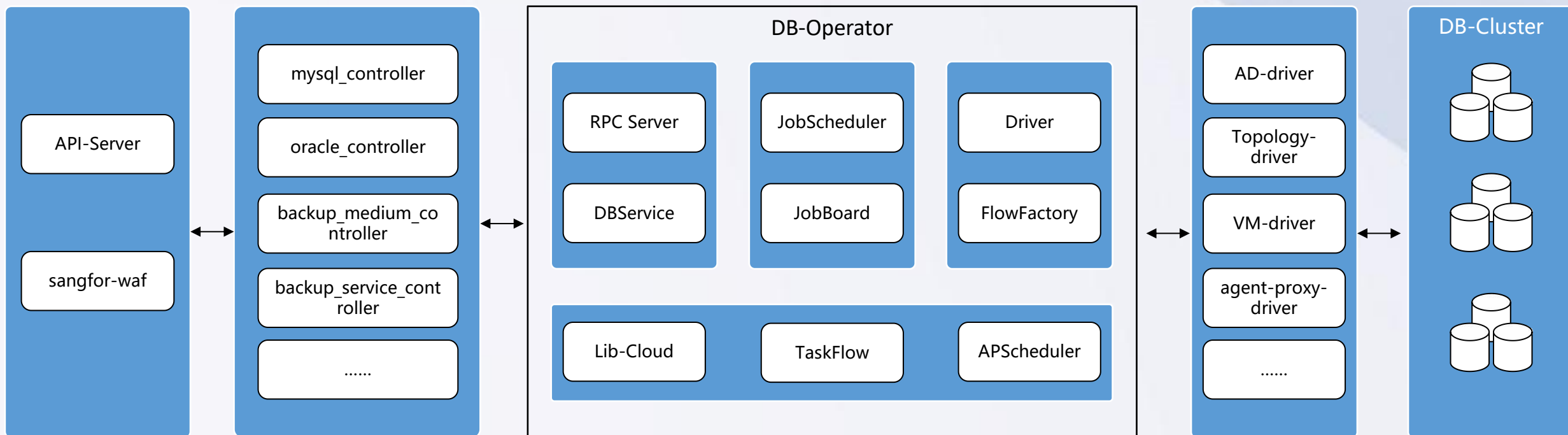
DMP平台的底层架构从上到下可以分为：

- 核心模块（UI，API-Server，配置中心，DB-Operator，后端MySQL等）、
- 服务配套模块（告警、日志、授权、缓存等），

与SCP通过SDK-API交互完成DBVM的各项操作（包括创建、配置变更、挂载备份盘等）

DMP整体技术框架——DB-Operator

- API-Server 主要负责初步的参数校验和用户配置的编辑;
- DB-Operator是DMP管理系统中负责任务调度和集群状态管理的关键模块;
- 管理平台通过各个服务的controller调用DB-Operator提供的API, 包括MySQL、Oracle的创建, 备份任务的执行, 系统管理等等;
- DB-Operator收到请求后, 通过Driver生成不同的 workflows 执行对应任务和状态管控。



DMP整体技术框架——其他模块

模块		描述
API-Server		主要负责初步的参数校验和用户配置的编辑。一旦配置变更，其会通过调用后端对应的服务执行；同时API-Server还负责通过外部（SCP）完成用户认证并向后提供一套多租户的体系
集群管理	Orchestrator	集群管理中，负责MySQL数据库集群故障自愈前置检测的关键模块。其通过连续的定时任务对目标MySQL数据库实例进行探测以获取健康状态，为后续的故障自愈提供判断基础。自愈模块详细技术细节见后续章节内容
	Jaguar	集群管理中，负责Oracle数据库集群故障自愈前置检测的关键模块，与上述Orchestrator发挥作用的方式类似。自愈模块详细技术细节见后续章节内容
备份管理	Xtrabackup	备份管理中，负责MySQL数据库备份的关键模块。由此模块来执行MySQL数据库集群具体的备份任务，包含全备/增备，并保障数据库的事务一致性。备份详细技术细节见后续章节内容
	RMAN	备份管理中，负责Oracle数据库备份任务的关键模块。由此模块来执行Oracle数据库集群的具体备份任务。备份详细技术细节见后续章节内容
后端MySQL		平台后端模块，用以存储DMP自身的各项数据的数据库。
告警中心		平台后端基础模块，用以对平台各项告警内容收集、存放、外发。
日志中心		平台后端基础模块，对平台各项操作、任务、工作流的执行进行记录，提供了操作日志、服务日志、任务中心等模块的基础。
授权模块		平台后端基础模块，管理MySQL、Oracle等各基础模块的授权认证及校验服务
SDK-API		由SCP提供，DMP平台与SCP的各项交互、认证，均通过SDK-API执行

目录

一、DMP整体技术框架

二、高可靠

1. 集群架构部署
2. 自愈能力
3. 备份保护
4. 监控体系
5. 底层可靠

三、高性能

1. 最佳实践调优
2. 存储层优化
3. 性能测试报告
4. SQL级性能诊断

四、易运维

1. 业务快速上线
 - Oracle DTS
 - 自动化部署
2. 生命周期管理
3. 租户自服务

集群架构部署——与传统模式对比

数据库上线速度提升20倍

传统部署数据库方式耗时耗力，且容易出错



手工创建耗时一般以天为单位，至少1天

- 1、步骤繁琐，操作门槛高
- 2、涉及配置项太多，容易出错，一旦出错需要检查，消耗很多处理时间
- 3、配置不当会导致数据库运行处于不良状态

STEP1: 准备服务器机器，下载相关的操作系统和数据库镜像包以及相关补丁

STEP2: 安装操作系统

STEP3: 准备存储，规划ASM磁盘

STEP4: 手动安装数据库，包括设置IP、上传安装包、上传安装脚本、解压安装脚本、设置字符集等

STEP5: 设置相关配置项，设置参照文档修改配置文件、修改脚本权限、执行安装脚本

STEP6: 优化相关参数，包括修改大页内存、SGA、PGA、redolog大小等等

.....以及一旦出错之后冗长的检查过程

一键化、规范化部署数据库，大大加快业务上线效率



实测采用DMP自动化部署数据库耗时:

➤ Oracle 单机: 10min

➤ Oracle RAC: 50min

- 1、部署向导总共4步，输入配置信息即可，创建全自动化
- 2、自动提供数据库最佳实践配置

STEP1: 配置节点规格、网络等基本信息



STEP2: 配置数据库字符集、账号密码等



STEP3: 配置备份策略



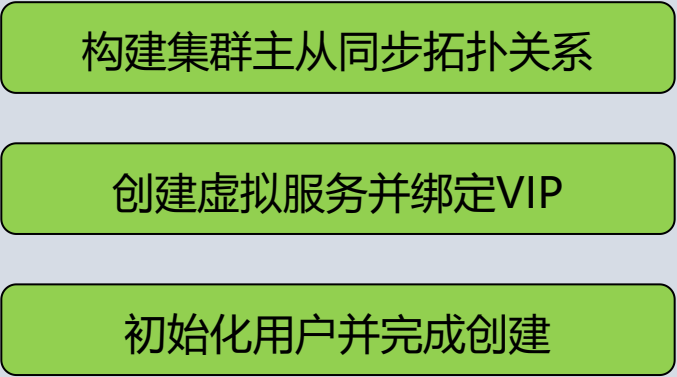
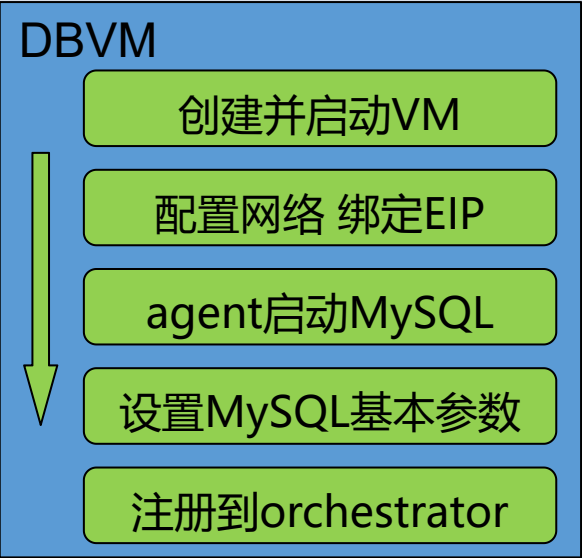
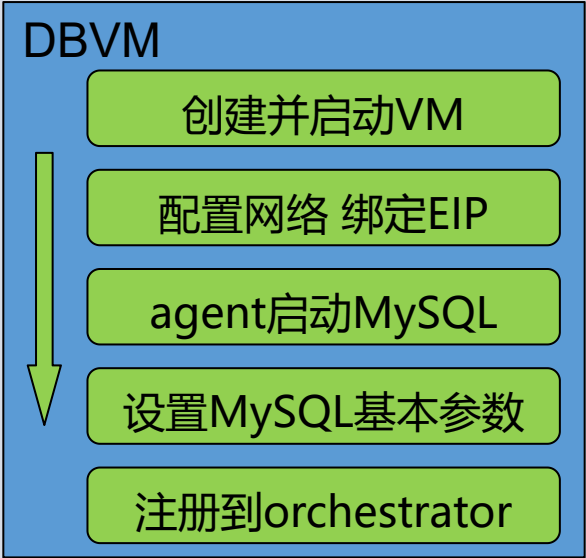
STEP4: 确认信息，开始创建数据库

一套RAC50分钟完成部署

集群架构部署——MySQL主从

- MySQL集群部署的技术方案实现

db-operator



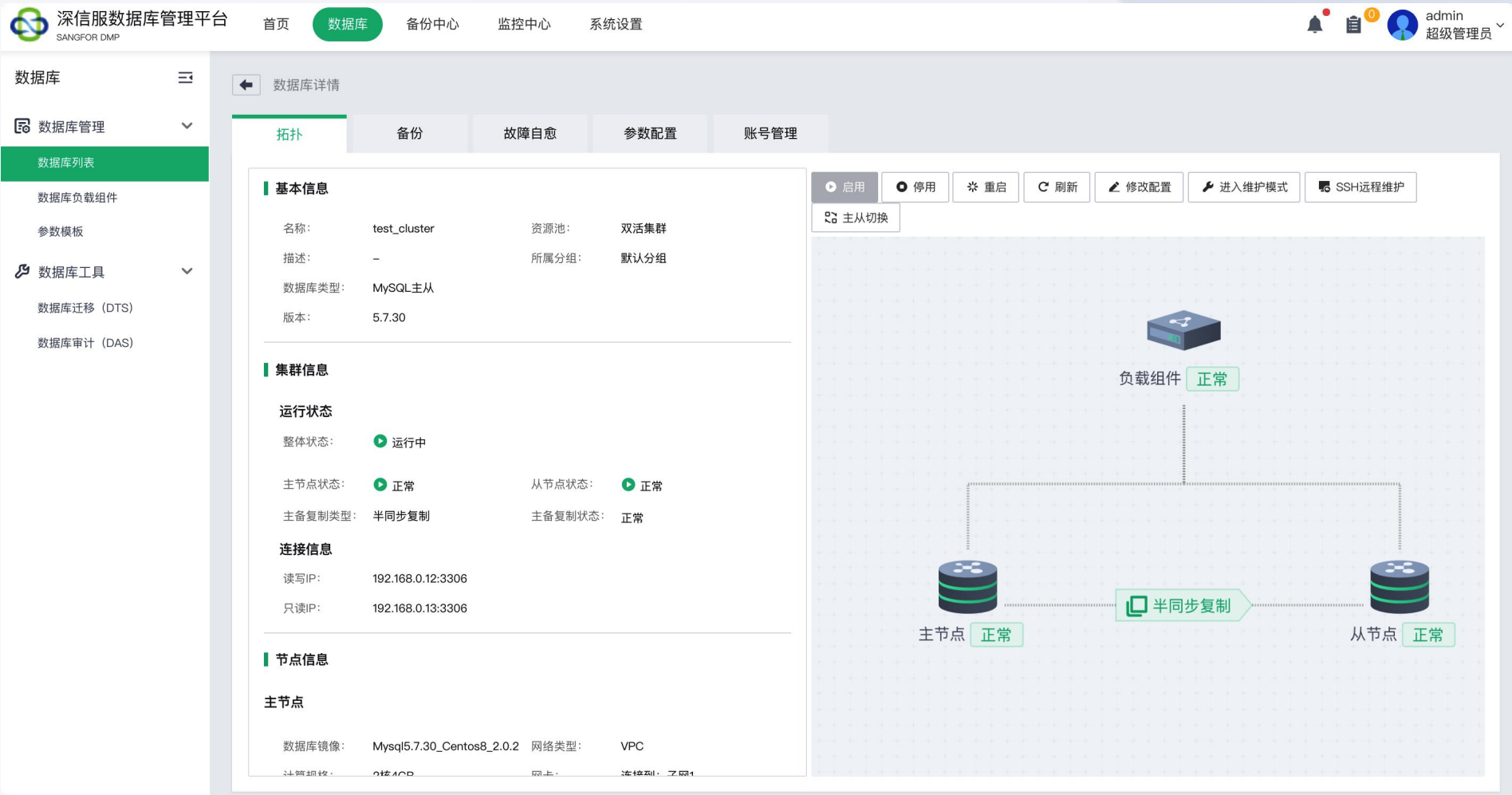
整体创建流程如下：

- 1、DMP提供的数据库镜像包含db-agent组件可直接用于数据库自动化部署。
- 2、默认参数即为最佳实践配置，包含innodb_flush_log_at_trx_commit、sync_binlog设置“双1”来确保数据安全
- 3、整个自动化部署过程，调用API执行虚拟机创建，数据库启动，注册到orchestrator，构建主从拓扑关系，均由db-operator自动进行调度管理与控制

其中，orchestrator会负责DB集群健康状态检查，与自愈模块进行交互，确保集群始终处于可用状态。

集群架构部署——MySQL主从可视化管理

- 拓扑可视化：屏蔽底层复杂度，可以直接查看简化版数据库集群的拓扑详情，集群节点的基本配置信息，并支持对数据库进行开关机、重启、修改配置等基本操作，清晰直观



集群架构部署——Oracle RAC

- RAC集群部署的技术方案实现

db-operator



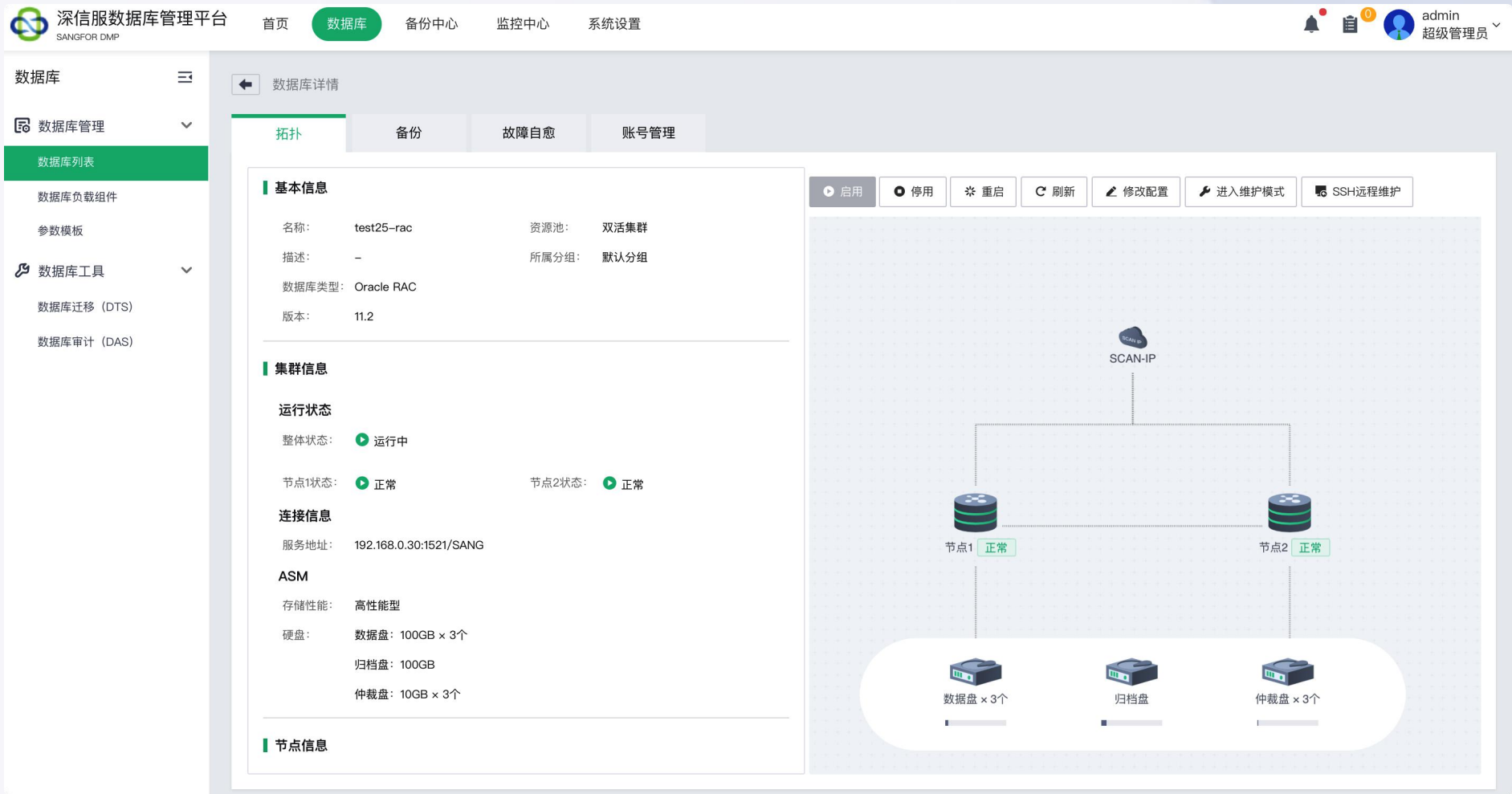
RAC是Oracle高可用的一种部署模式，传统方式部署非常复杂，涉及ASM、Grid Infrastructure等等系统和部署组件，需要手动执行上百个步骤，而现在DMP已将此过程实现自动化：

- 整体创建流程如下：
- 1、由DMP提供的数据库镜像创建DBVM
 - 2、DMP申请集群VIP、SCAN VIP、共享盘等集群配置资源
 - 3、给DBVM节点挂载共享盘，更新集群信息
 - 4、开始执行数据库创建并构建RAC

全程只需要用户指定规格、磁盘容量、网卡等基本信息，便无需额外的手动操作。

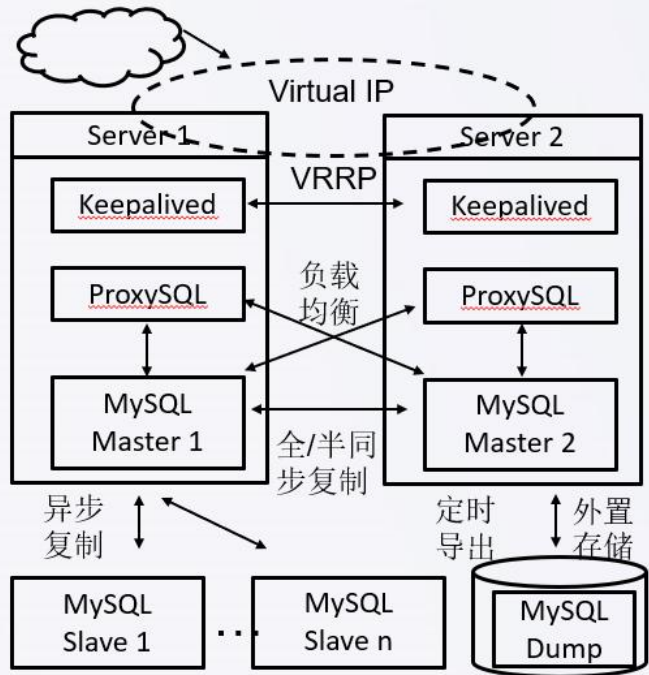
集群架构部署——Oracle RAC可视化管理

- 拓扑可视化：屏蔽底层复杂度，可以直接查看简化版数据库集群的拓扑详情，数据库访问服务地址，集群节点的基本配置信息，并支持对数据库进行开关机、重启、修改配置等基本操作，支持集群节点规格的纵向扩容，覆盖全生命周期



自愈能力——背景知识

MySQL自建高可用架构依然存在弊端

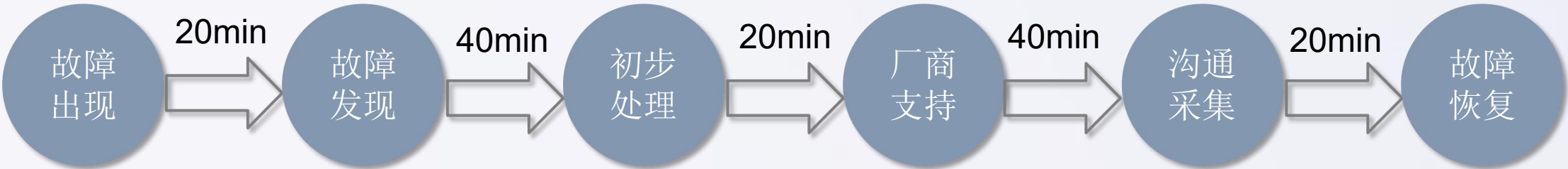


两节点无法防脑裂?
Relay Log 未消费完就发生了切换?
服务出错挂起?
同步降级?
缺乏实时监控?
恢复时间长导致风险敞口大?
故障切换后恢复正确的从库日志同步?

同步延迟?
网络抖动?
数据库假死?
半同步的最后一个事务要回滚?
网络隔离?
VIP摘除失败?
备份数据失效?

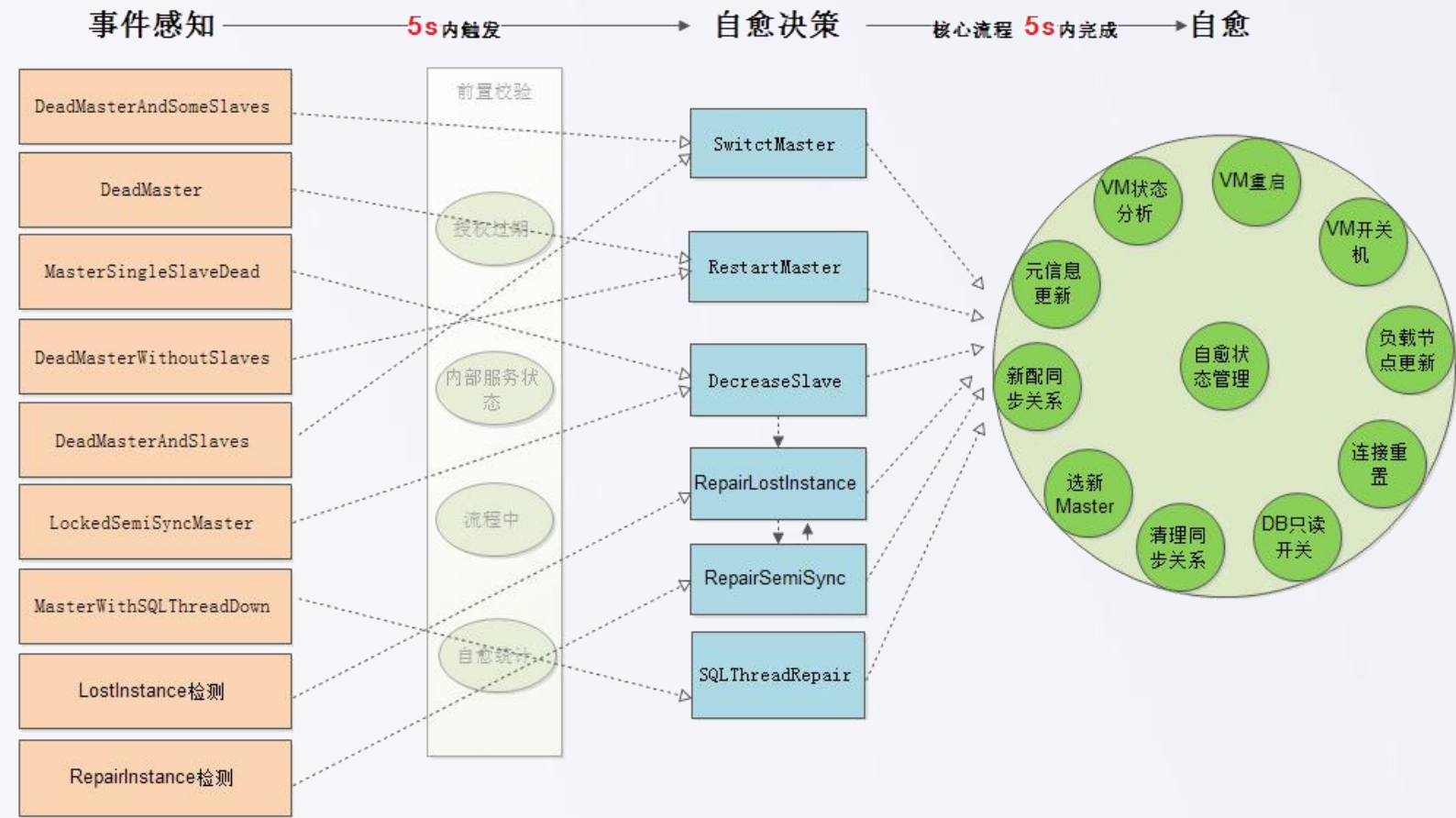
涉及层面多、操作易错、技术门槛要求高，故障处理时间很长，对业务影响大

能否有简单、高效、高性价比的方案，实现高可靠，故障可快速恢复？



自愈能力——MySQL实现

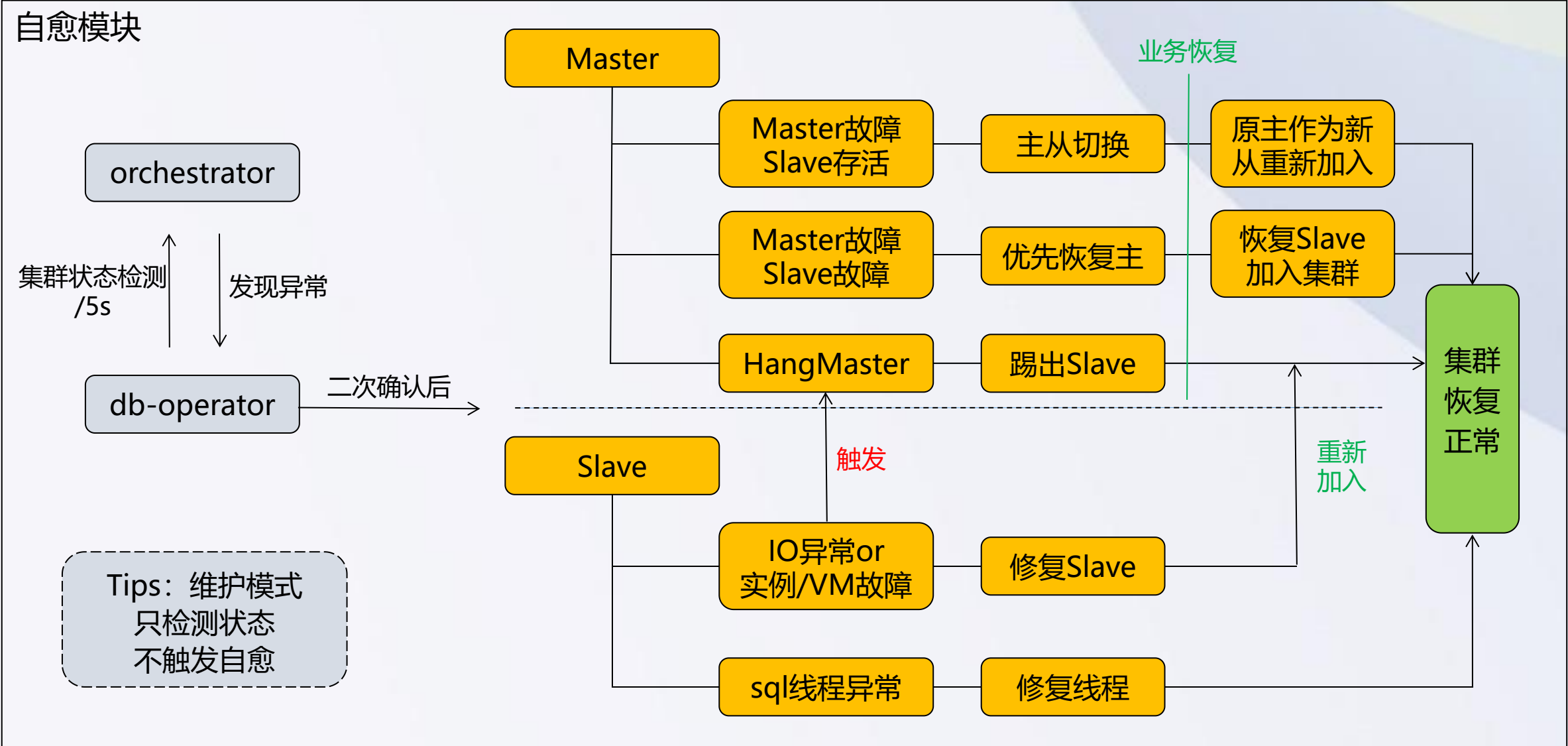
- 自动化运维的技术方案实现（集群自愈）



- 1、自动化运维处理的前提条件是能够识别问题，平台针对MySQL 数据库集群详细分析了各类异常状态，进行了统一定义，涵盖集群中单个、多个节点以及节点间的可能状态
- 2、在有了状态定义基础上，平台会在5s周期内轮询监控数据库集群状态，在发现异常之后对当前状态所需的修复操作进行决策判断
- 3、基于上述步骤，执行特定的自愈动作。所遵循的原则是：最大程度保障数据安全；最大程度降低业务影响

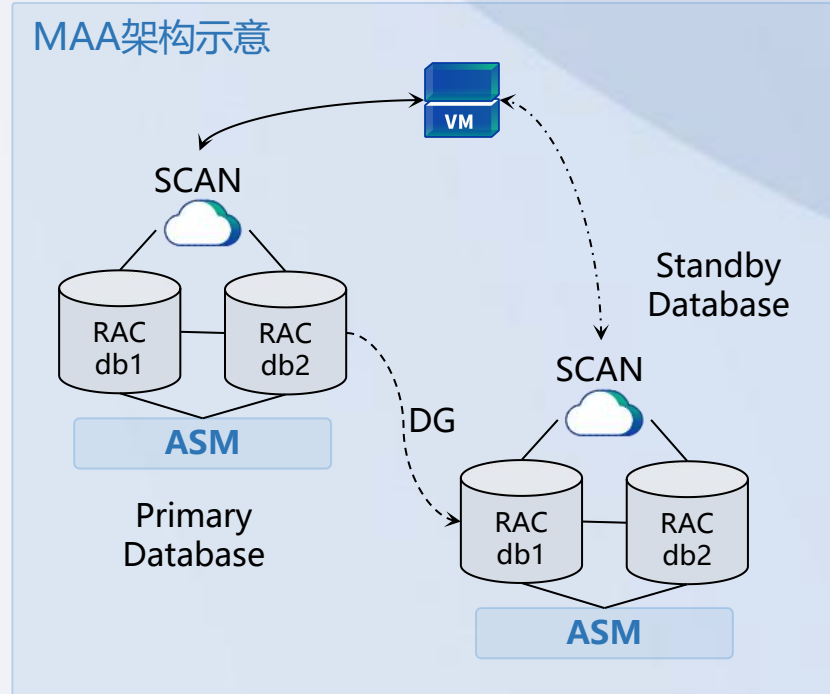
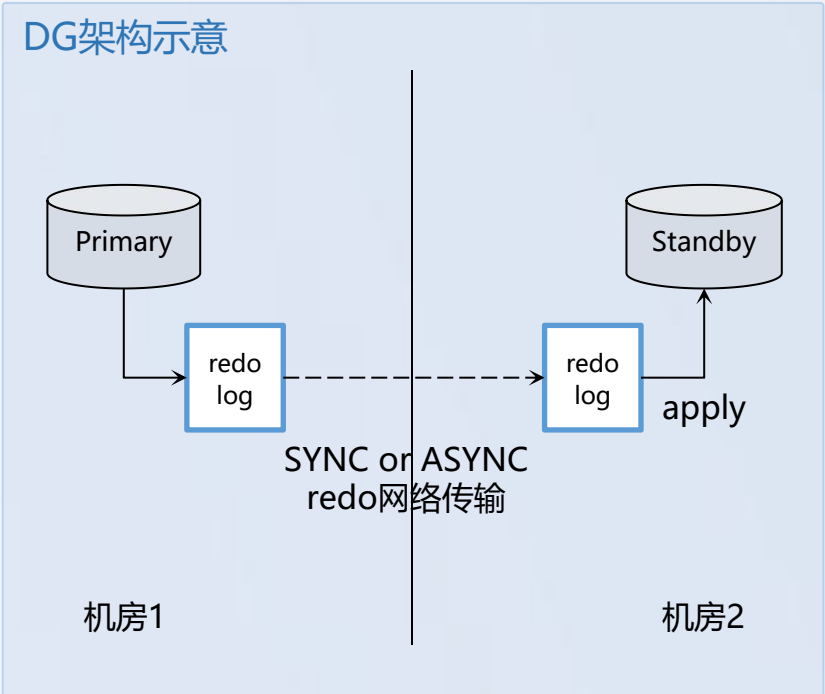
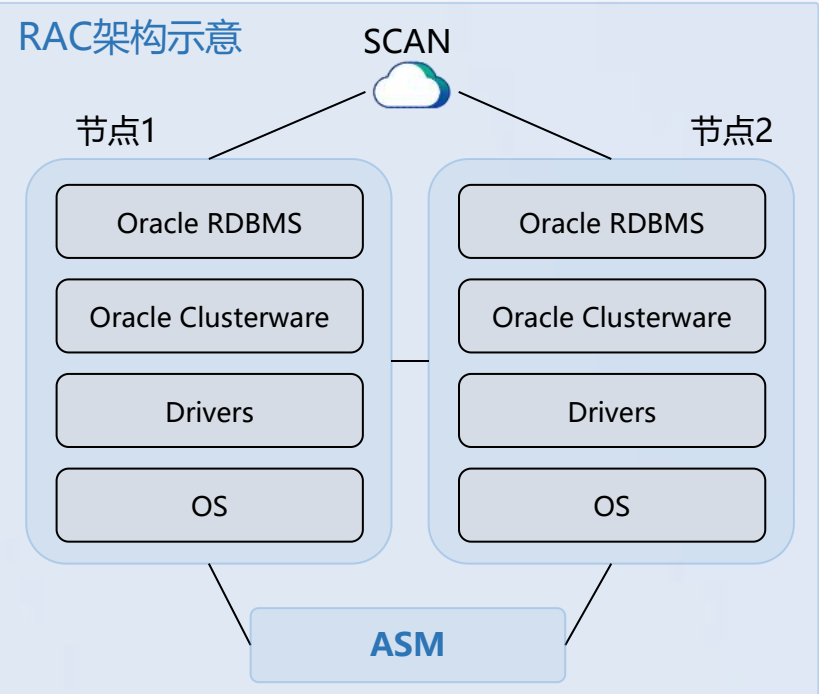
自愈能力——MySQL实现

- 自动化运维的技术方案实现（集群自愈）



自愈能力——背景知识

高可用方案	描述
RAC (Real Application Cluster)	Oracle高可用性的一种，多个Oracle服务器组成一个共享的Cache，而这些Oracle服务器共享一个基于网络的存储。这个系统可以容忍单机/或是多机失败（ 当此类问题出现，故障节点属于弃置状态 ）。不过系统内部的多个节点需要高速网络互连，基本上也就是要全部东西放在在一个机房内，或者说一个数据中心内。如果机房出故障，那也无法自动应对。
DG (DataGuard)	相比于RAC，更适用于容灾场景，即跨机房跨地部署模式。一个机房放生产库(primary)，另外其他机房部署备库(standby)的数据库。备库分物理的和逻辑的。物理的standby数据库主要用于primary失败后做切换。而逻辑的standby数据库则在平时可以分担primary数据库的读负载。属于高可用解决方案而非性能提升方案。
MAA (Maximum Availability Architecture)	其实不是独立的第三种，而是前面两种的结合，来提供最高的可用性。每个机房内部署RAC集群，多个机房间用Data Guard同步。当然实施成本也最高

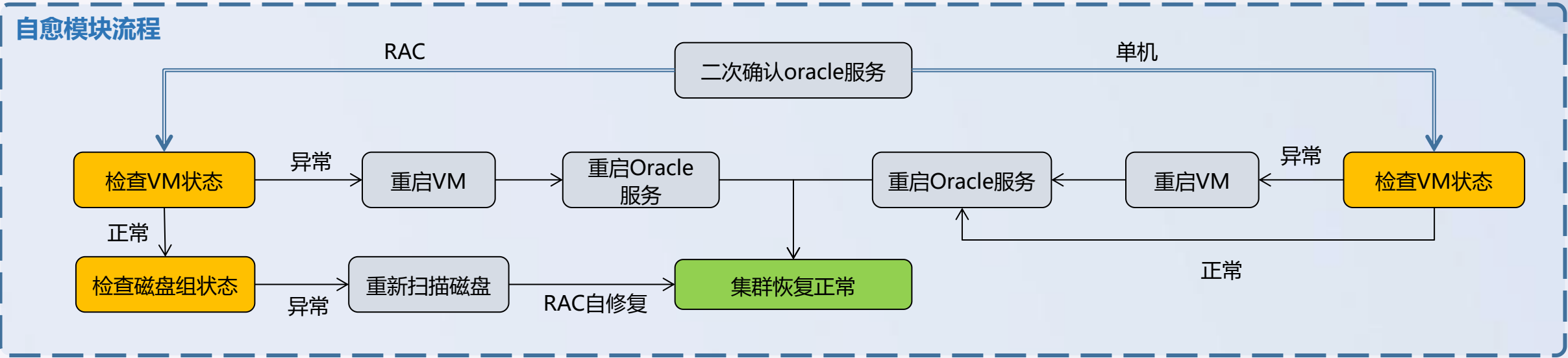
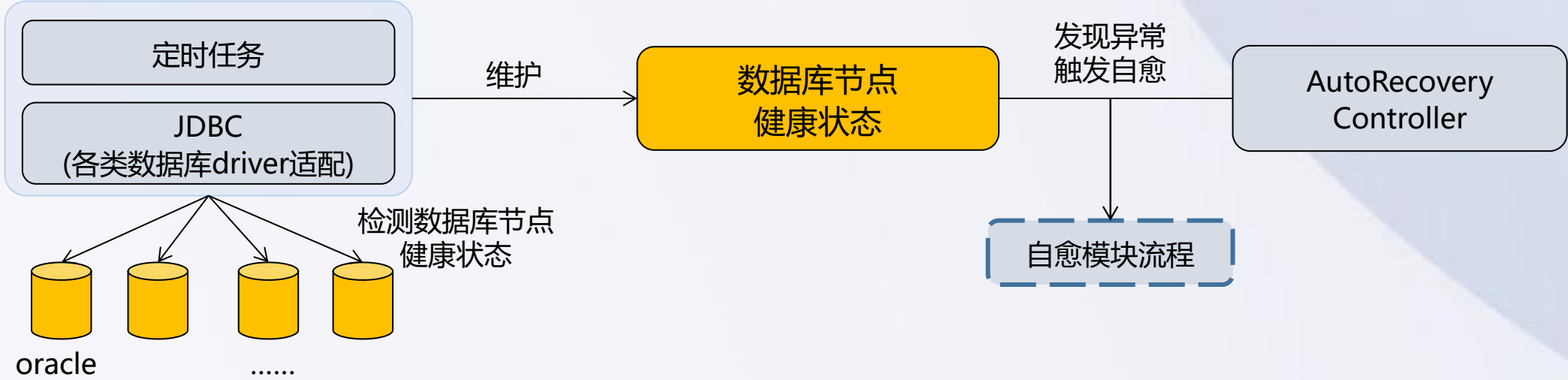


自愈能力——Oracle实现

- 自动化运维的技术方案实现

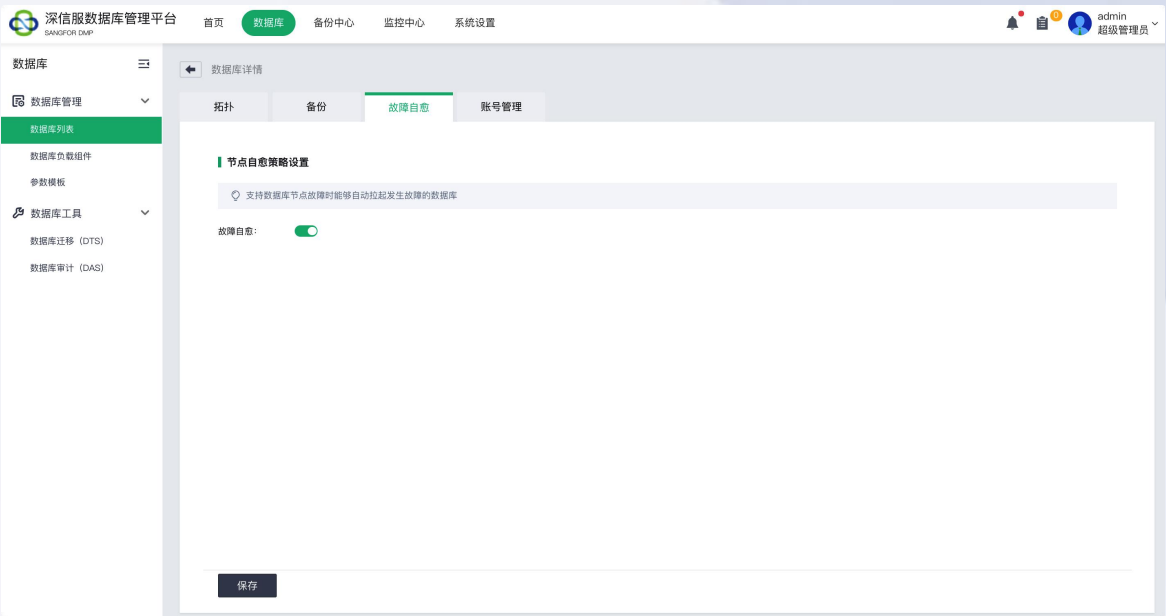
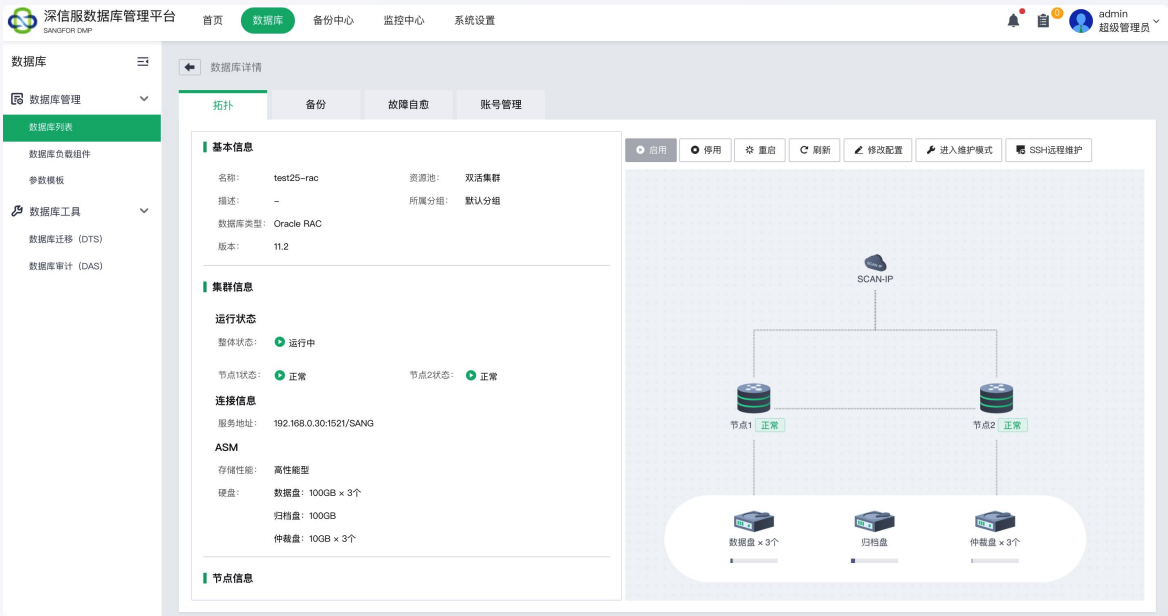
Jaguar (自研多活数据库通用故障检测模块)

Tips: 维护模式下
只检测状态
不触发自愈



自愈能力——产品功能

- 包含主从切换和故障自愈
 - 部署MySQL主从架构的数据库集群，默认开启故障自动主从切换和故障自愈；
 - 部署MySQL单机和Oracle 单机、RAC，默认开启故障自愈
 - 在开启维护模式时，切换和自愈均处于关闭状态，平台不做自动运维，允许用户自行对数据库、数据库云主机进行运维操作；也方便技服人员进行排障



备份保护

常见的数据库备份方式对比

备份方式	描述	优势	劣势
VM级备份	针对VM进行备份，只能适用于单机版数据库，集群架构不能采用此方式进行备份	简单，成本低	备份可靠性无法保障
数据库脚本备份	在数据库层手动写备份执行脚本+备份软件，通过脚本控制备份软件进行备份操作	相比VM层备份，可以做到数据库级的备份保护	1、脚本撰写需要依赖DBA，门槛较高； 2、无法对备份文件的有效性进行及时验证
RDS备份	针对数据库的事务级一致性备份，可以设定备份策略进行自动备份，也支持手动全量备份	1、平台自带，无需第三方软件，无需编写脚本； 2、事务级一致性保障； 3、可随时备份恢复演练；	暂时支持的最小备份粒度为小时级（后续会逐步支持分钟级、秒级）
第三方备份软件	通过外置设备或软件对数据库进行备份，支持RPO粒度较广，实现方式有逻辑/物理文件拷贝等，相对比较成熟	1、备份功能完整性较高； 2、可以支持较低的RPO粒度；	1、通常比较贵，成本高； 2、备份恢复演练不方便；

基于上述横向对比可见，RDS备份是相对最优的选项，DMP将RDS备份方案作为内置能力提供

备份保护——数据库备份背景知识（事务）

事务介绍：

简单地说，**事务就是指逻辑上的一组SQL语句操作**，组成这组操作的各个SQL语句，执行时**要么全成功要么全失败**。

例如：你给我转账5块钱，流程如下

a.从你银行卡取出5块钱，剩余计算 $money-5$

b.把上面5块钱打入我的账户上，我收到5块，剩余计算 $money+5$ 。

上述转账的过程，对应的sql语句为：

update 你_account set money=money-5 where name='你';

update 我_account set money=money+5 where name='我';

上述的两条SQL操作，在事务中的操作就是要么都执行，要么都不执行，不然钱就对不上了。

这就是事务的原子性(Atomicity)。

事务的四大特性ACID：

1.**原子性(Atomicity)** 事务是一个不可分割的单位，事务中的所有SQL等操作要么都发生，要么都不发生。

2.**一致性(Consistency)** 事务发生前和发生后，数据的完整性必须保持一致。

3.**隔离性(Isolation)** 当并发访问数据库时，一个正在执行的事务在执行完毕前，对于其他的会话是不可见的，多个并发事务之间的数据是相互隔离的。也就是其他人的操作在这个事务的执行过程中是看不到这个事务的执行结果的，也就是他们拿到的是这个事务执行之前的内容，等这个事务执行完才能拿到新的数据。

4.**持久性(Durability)** 一个事务一旦被提交，它对数据库中的数据改变就是永久性的。如果出了错误，事务也不允撤销，只能通过'补偿性事务'。

备份保护——数据库备份背景知识（MySQL）

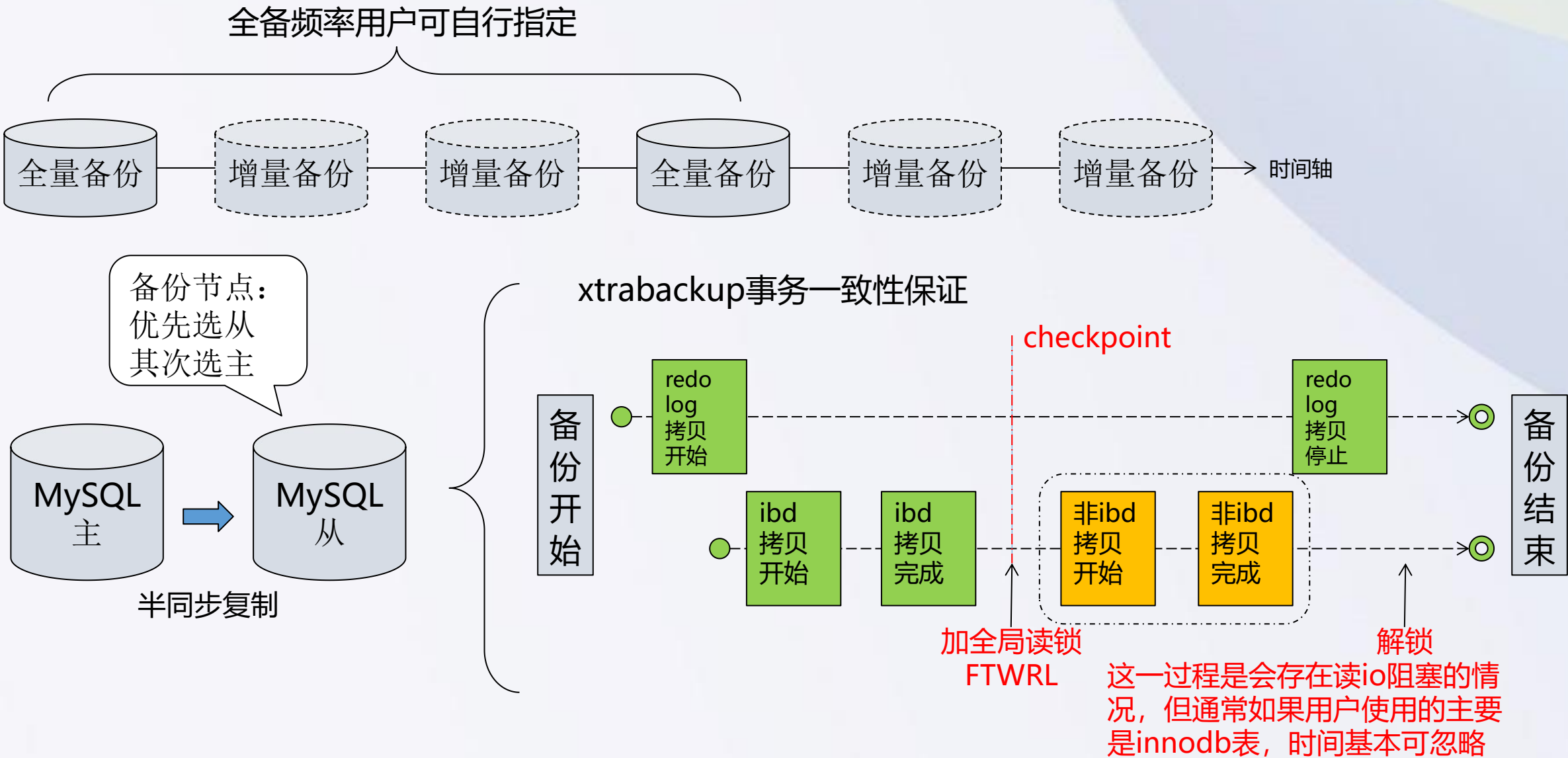
一般来说，物理备份恢复速度比较快，占用空间比较大，逻辑备份速度比较慢，占用空间比较小。

分类	实现方式	特性描述
逻辑备份	mysqldump	mysqldump是采用SQL级别的备份机制，他将数据表导出成SQL脚本文件，是最常用的逻辑备份方法。
物理备份	xtrabackup	mysql开源备份工具，主要特点是： 1) 在线热备份，可以备份innodb和myisam； 2) 支持流备份，可以备份到disk、tape和remote host； 3) 可以利用lsn和基础备份目录来进行增量备份； 4) 支持记录slave上的master log和master position信息； 5) 支持多进程同时热备份；
	LVM	热备、支持所有基于本地磁盘的存储引擎、快速备份、低开销、容易保持完整性、快速恢复等
	cp+tar	使用直接拷贝数据库文件的方式进行打包备份，需要注意的是执行步骤：锁表、备份、解表。 恢复也比较简单，直接拷贝到之前的数据库文件的存放目录即可。
	mysql主从复制	mysql的复制是指将主数据库的DDL和DML操作通过二进制文件（bin-log）传送到从服务器上，然后在从服务器上对这些日志做重新执行的操作，从而使得从服务器和主服务器保持数据的同步。

最终选用
xtrabackup方案。速度快，
可靠性高，有
一致性保障

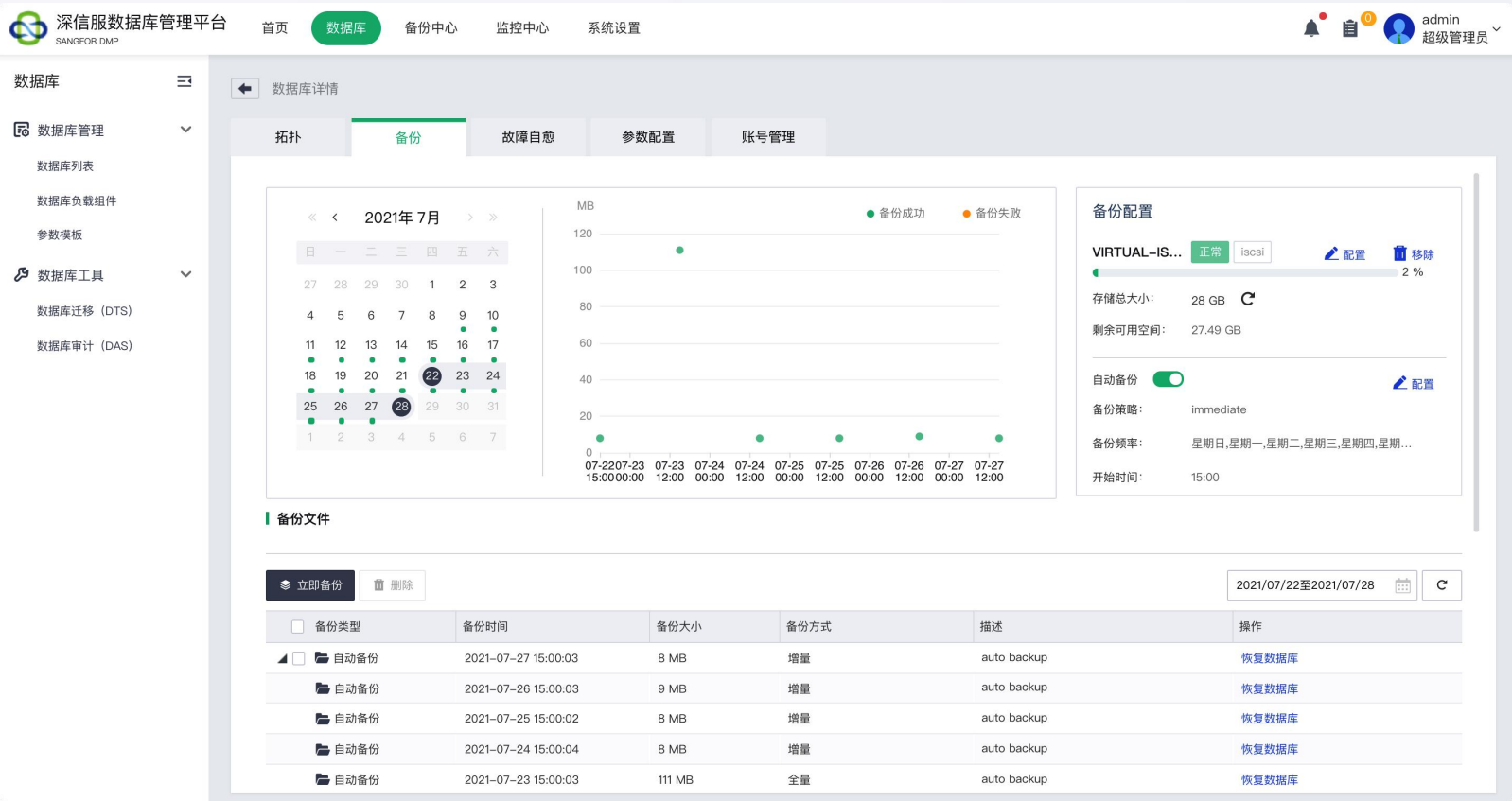
备份保护——MySQL数据保护

- 数据保护（数据库一致性备份）的技术方案实现



备份保护——MySQL数据保护

- 数据库备份：以集群维度，支持数据库事务一致性备份
 - 备份策略：只需要自定义备份时间，并选择全备频率，平台会在备份时间触发备份任务，首次执行备份以及达成全备频率间隔条件时，均为全量备份，其余在备份时间均触发增量备份；
- 备份管理：支持在DMP对所有DB集群的备份文件、备份策略进行统一管理



备份设置

备份策略: immediate

备份时间: 星期日,星期一,星期二,星期三,星期四,星期五,星期六 15:00

全量备份频率: 每两周一次

备份保留时长: 30天

确定 取消

立即备份

描述: 手动备份

备份位置: VIRTUAL-IS-DISK_SANGFOR_1SANGFOR_ASAN-010-

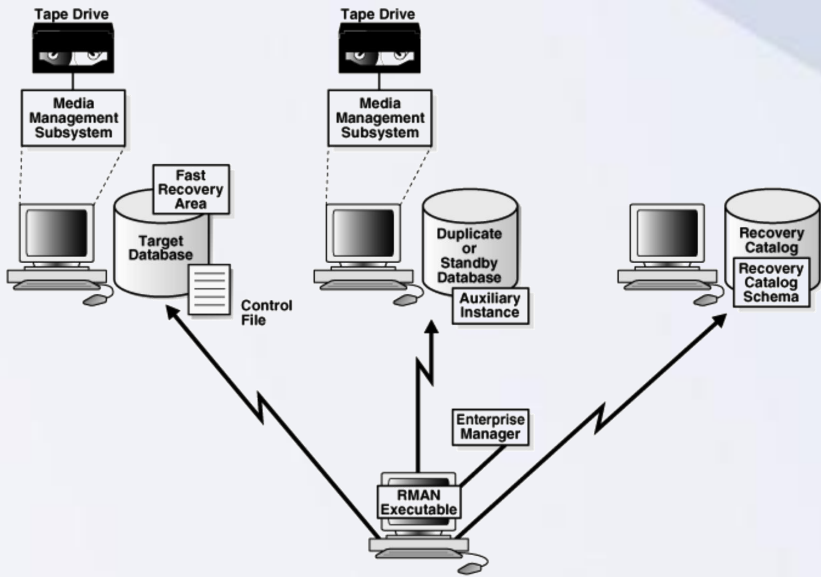
确定 取消

备份保护——数据库备份背景知识（Oracle）

备份方法	描述
导出/导入	利用Export可将数据从数据库中提取出来，利用Import则可将提取出来的数据送回到Oracle数据库中去。
冷备份	冷备份发生在数据库已经正常关闭的情况下，当正常关闭时会提供给我们一个完整的数据库。冷备份时将关键性文件拷贝到另外的位置的一种说法。对于备份Oracle信息而言，冷备份是最快和最安全的方法，但备份过程必须关闭数据库。
热备份	热备份是在数据库运行的情况下，采用archivelog mode方式备份数据库的方法。例如，如果你有昨天夜里的一个冷备份而且又有今天的热备份文件，在发生问题时，就可以利用这些资料恢复更多的信息。热备份要求数据库在Archivelog方式下操作，并需要大量的归档空间。

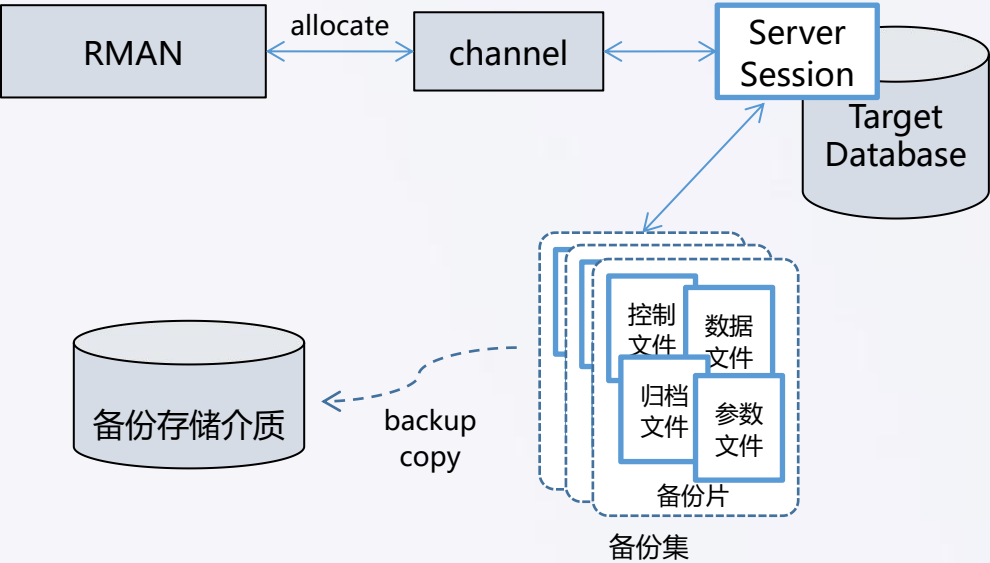
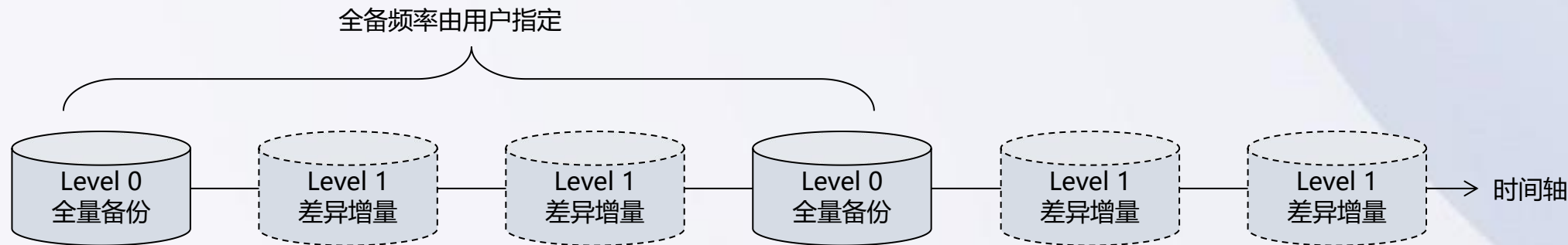
RMAN（Recovery Management）

- RMAN是随Oracle服务器软件一同安装的工具软件，属于物理备份；
- 它可以用来备份和恢复**数据库文件、归档日志和控制文件、参数文件**，用来执行完全或不完全的数据库恢复；
- 支持**全备份和增量备份**，它会检测和报告损坏的数据块，并在备份时进行跳过；
- 其使用Oracle特有的二进制压缩模式，与传统备份的压缩方式相比，能够最大程度地压缩数据块中的一些典型数据。



备份保护——Oracle数据保护

- 数据保护（数据库一致性备份）的技术方案实现



- RMAN需数据库开启归档模式
- RMAN备份对象包括数据文件、控制文件、归档文件、参数文件，备份完成会清理归档
- RMAN 保证不备份fractured block(即SCN不一致的块)
- 在 RMAN 备份期间，数据库Server Session会读取每个数据块，并通过比较块的头尾来检查它是否已断裂(fractured)。如果一个块被破坏，则会重新读取该块

备份保护——Oracle数据保护

- 数据库备份：以集群维度，支持数据库事务一致性备份
 - 备份策略：只需要自定义备份时间，并选择全备频率，平台会在备份时间触发备份任务，首次执行备份以及达成全备频率间隔条件时，均为全量备份，其余在备份时间均触发增量备份；
- 备份管理：支持在DMP对所有DB集群的备份文件、备份策略进行统一管理

深信服数据库管理平台

SANGFOR DMP

首页

数据库

备份中心

监控中心

系统设置

admin

超级管理员

数据库

数据库管理

数据库列表

数据库负载组件

参数模板

数据库工具

数据库迁移 (DTS)

数据库审计 (DAS)

数据库详情

拓扑

备份

故障自愈

账号管理

<<

<

2021年 7月

>

>>

日

一

二

三

四

五

六

27

28

29

30

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

1

2

3

4

5

6

7

MB

● 备份成功

● 备份失败

1,200

1,000

800

600

400

200

0

07-22 15:00

07-23 00:00

07-23 12:00

07-24 00:00

07-24 12:00

07-25 00:00

07-25 12:00

07-26 00:00

07-26 12:00

07-27 00:00

07-27 12:00

备份配置

VIRTUAL-IS-DISK_SANGF...

正常iscsi配置移除

存储总大小: 28 GB

剩余可用空间: 24.65 GB

自动备份: immediate

备份策略: 星期日,星期一,星期二,星期三,星期四,星期五,星期六

备份频率: 15:00

备份文件

立即备份

删除

2021/07/22至2021/07/28

备份类型	备份时间	备份大小	备份方式	描述	操作
自动备份	2021-07-26 15:00:04	247 MB	增量	auto backup	恢复数据库
自动备份	2021-07-25 15:00:02	281 MB	增量	auto backup	恢复数据库
自动备份	2021-07-24 15:00:06	1.09 GB	增量	auto backup	恢复数据库

新增备份策略

✕

基本信息

策略名称:

请输入

描述:

请输入

用户:

rds1

备份方式

备份时间:

星期日 ×

00:00

全量备份频率:

☒ 每两周一次 ☐ 每四周一次

备份保留时长:

30天

确定

取消

备份保护——数据库快速恢复

- 基于备份点的快速恢复，支持恢复为全新的数据库集群
 - 单机架构的备份可恢复为单机；MySQL主从可恢复为主从集群；RAC的备份可恢复为RAC

恢复数据库

基本信息

恢复方式：

生成全新数据库（单机）

恢复点：

2021-07-27 15:00:03

新数据库名称：

test_db_restored_(2021-07-27_15-00-03)

描述：

请输入描述

恢复到资源池：

双活集群

恢复到分组：

所有/双活集群/rds/默认分组

节点信息

数据库镜像：

Mysql5.7.30_Centos8_2.0.2

计算规格：

2核4GB

存储性能：

性能型

高性能型

容量型

硬盘：

磁盘1：

100

GB

磁盘2：

100

GB

确定

取消

恢复数据库

基本信息

恢复方式：

生成全新数据库（主从）

恢复点：

2021-07-27 15:00:03

新数据库名称：

test_cluster_restored_(2021-07-27_15-00-03)

描述：

请输入描述

恢复到资源池：

双活集群

恢复到分组：

所有/双活集群/rds/默认分组

节点信息

数据库镜像：

Mysql5.7.30_Centos8_2.0.2

计算规格：

2核4GB

存储性能：

性能型

高性能型

容量型

硬盘：

磁盘1：

100

GB

磁盘2：

100

GB

确定

取消

恢复数据库

基本信息

恢复方式：

生成全新数据库（Oracle RAC）

恢复点：

2021-07-26 15:00:04

新数据库名称：

test25-rac-restored-(2021-07-26-15-00-04)

描述：

请输入描述

恢复到资源池：

双活集群

恢复到分组：

所有/双活集群/rds/默认分组

节点信息

数据库镜像：

Oracle11.2_Centos7.9_2.0.2

计算规格：

4核8GB

存储性能：

性能型

高性能型

容量型

硬盘：

数据盘、归档盘和仲裁盘的冗余策略均为外部冗余。外部冗余即，磁盘组中的有效磁盘空间是其所有磁盘空间的总和，数据冗余由底层磁盘副本机制保障。

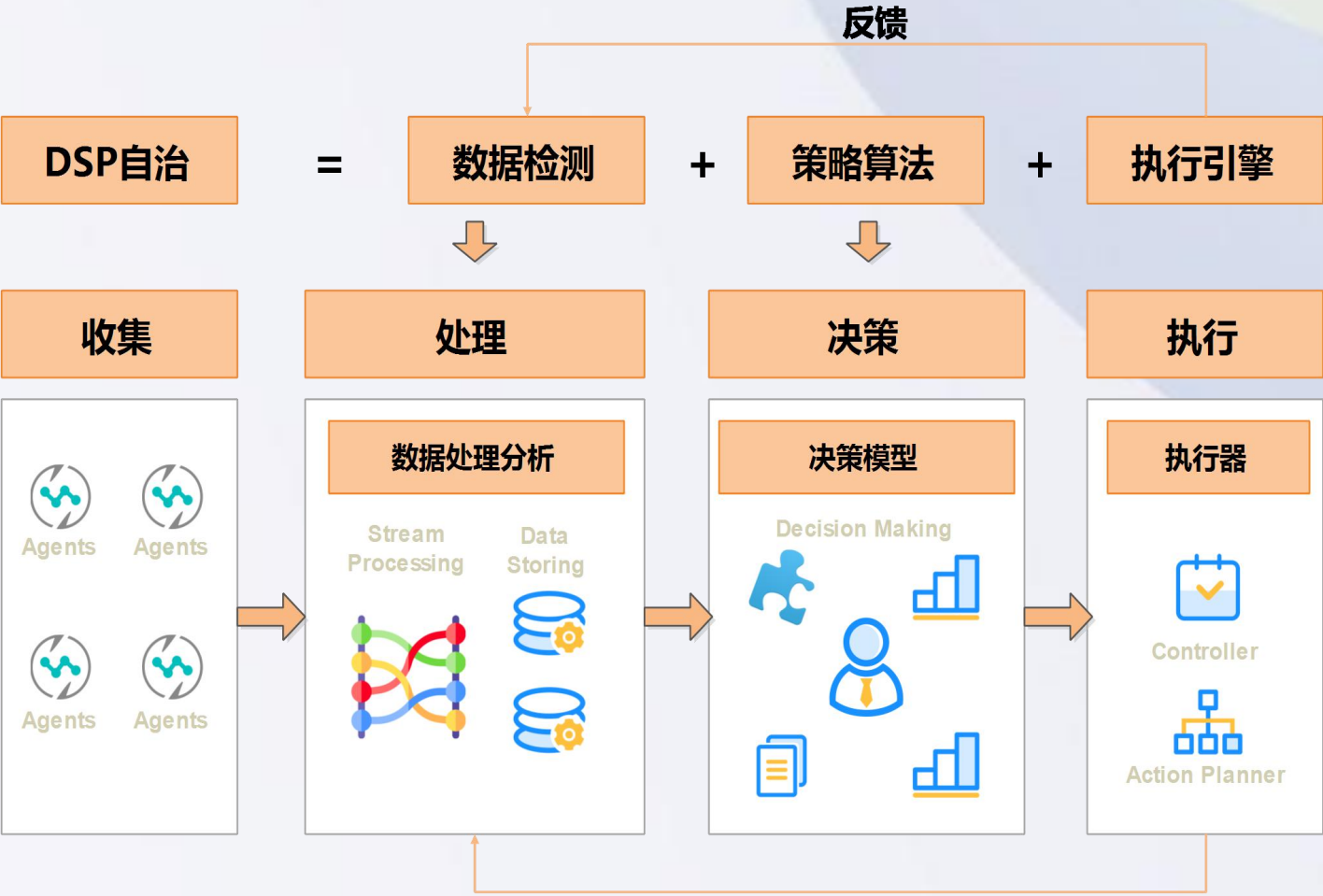
系统盘

确定

取消

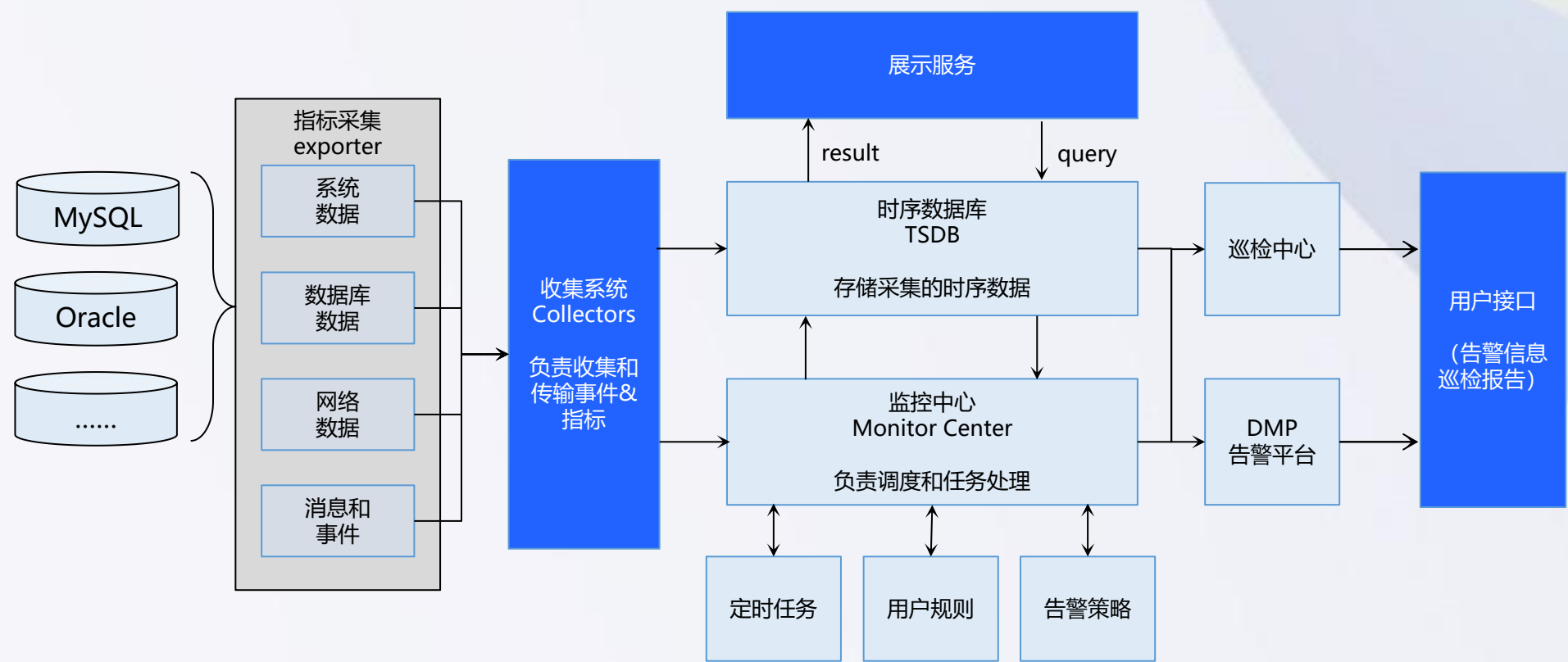
全方位监控和分析体系——核心理念

- 数据驱动
- 专家经验+机器学习
- 诊断流程闭环
- 持续进行
- 规模优化



Q4开始会逐步提供数据库自治相关的功能

全方位监控和分析体系——架构设计

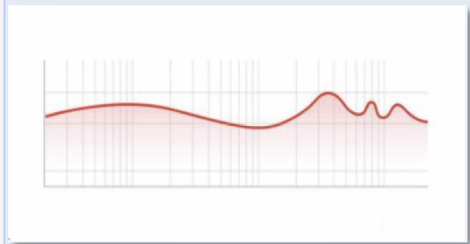


- 监控巡检系统的整体架构如上图所示，
- 1、首先由指标采集组件对不同数据库的各项相关数据（系统数据、网络数据、数据库数据、消息和事件）进行采集，、
 - 2、采集的数据经由收集系统Collector进行汇总整理，
 - 3、存储至时序数据库TSDB。
 - 4、再通过数据处理模块对数据进行实时处理，分别由巡检中心生成巡检报告、告警系统触发告警、展示服务提供UI呈现

全方位监控和分析体系——功能概述

- ▲ 无需人工参与的自动监控和巡检机制
- ▲ 提供有效的处置建议，故障分析有据可依

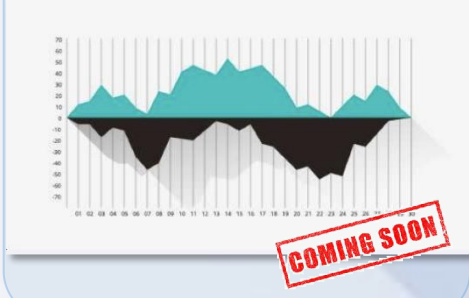
异常告警



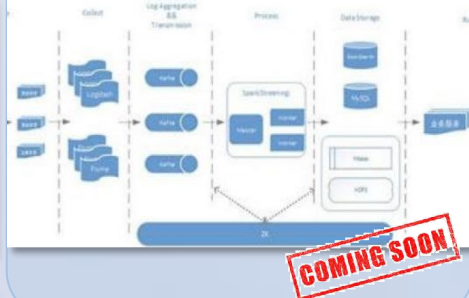
健康巡检



智能调参



SQL优化



产品截图

深信服数据库管理平台

SANFIDOR DMP

首页

数据库

备份中心

监控中心

系统设置

admin

超级管理员

监控中心

性能趋势

健康巡检

告警日志

告警通知

健康巡检

手动巡检

立即巡检

自动巡检

未开启

每日凌晨2点自动巡检数据库实例，跟踪数据库运行状态。

巡检设置

2021-07-20 00:00:00 ~ 2021-07-26 23:59:59

更多筛选

搜索数据库名称

数据库名称	数据库类型	巡检类型	开始时间	巡检范围	健康等级	资源池	用户	巡检报告
test25-rac	Oracle RAC	手动巡检	2021-07-26 19:20:11 至 2021-07-26 19:20:38	2021-07-25 19:19:49-2021-07-26 19:19:00	健康	双活集群	dmp_test	下载
test25	Oracle单机	手动巡检	2021-07-26 19:19:56 至 2021-07-26 19:20:11	2021-07-25 19:19:49-2021-07-26 19:19:00	健康	双活集群	dmp_test	下载
test25-rac	Oracle RAC	手动巡检	2021-07-26 19:19:56 至 2021-07-26 19:20:26	2021-07-25 19:19:49-2021-07-26 19:19:00	高危	双活集群	rds	下载
test_cluster...	MySQL主从	手动巡检	2021-07-20 17:46:35 至 2021-07-20 17:47:20	2021-07-19 17:46:14-2021-07-20 17:46:00	健康	双活集群	rds	下载

显示第 1 到 20 条记录，共 4 条记录 共 4 页 < 1 > 每页 20 条 前往 1 页

深信服数据库管理平台

SANFIDOR DMP

首页

数据库

备份中心

监控中心

系统设置

admin

超级管理员

监控中心

性能趋势

健康巡检

告警日志

告警通知

MySQL

Oracle

数据库:

rds1

210-oracle-single

192.168.0.3 (节点1)

时间: 最近一小时

刷新

监控类型:

资源监控

性能监控

CPU和内存使用率 (%)

100
80
60
40
20
0

2021-07-26 19:32:24 2021-07-26 19:36:36 2021-07-26 19:40:48 2021-07-26 19:45:00 2021-07-26 19:49:12 2021-07-26 19:53:24 2021-07-26 19:57:36 2021-07-26 20:01:48 2021-07-26 20:06:00 2021-07-26 20:10:12 2021-07-26 20:14:24 2021-07-26 20:18:36 2021-07-26 20:22:48 2021-07-26 20:27:00 2021-07-26 20:32:17

CPU使用率 内存使用率

磁盘空间 (MB)

120000
100000
80000
60000
40000
20000
0

2021-07-26 19:32:24 2021-07-26 19:36:36 2021-07-26 19:40:48 2021-07-26 19:45:00 2021-07-26 19:49:12 2021-07-26 19:53:24 2021-07-26 19:57:36 2021-07-26 20:01:48 2021-07-26 20:06:00 2021-07-26 20:10:12 2021-07-26 20:14:24 2021-07-26 20:18:36 2021-07-26 20:22:48 2021-07-26 20:27:00 2021-07-26 20:32:17

数据盘总容量 数据盘可用空间 非数据盘总容量 非数据盘可用空间

IOPS (次/秒)

12
10
8
6
4
2
0

2021-07-26 19:32:24 2021-07-26 19:36:36 2021-07-26 19:40:48 2021-07-26 19:45:00 2021-07-26 19:49:12 2021-07-26 19:53:24 2021-07-26 19:57:36 2021-07-26 20:01:48 2021-07-26 20:06:00 2021-07-26 20:10:12 2021-07-26 20:14:24 2021-07-26 20:18:36 2021-07-26 20:22:48 2021-07-26 20:27:00 2021-07-26 20:32:17

底层可靠保障——分布式存储

数据校验（避免静默损坏）

- 有些类型的存储错误在一些存储系统中完全未报告和未检测到。它们会导致向应用程序提供损坏的数据，而不会发出警告、记录、错误消息或任何类型的通知。
- 深信服HCI端到端校验方案使用XXhash64算法，相比目前业界普遍使用的CRC-32和Adler-32算法，具有速度更快、冲突率更低的特点。

SSD 寿命可用承诺

- 支持磨损均衡，保证SSD寿命使用均匀，避免SSD短时间内寿命到期
- 支持寿命保护，寿命达到最低阈值时，不再写入数据

SSD vs HDD

(Solid State Drive)



(Hard Disk Drive)



磁盘亚健康

- 亚健康硬盘，典型的表现为：硬盘变慢、卡顿、拥塞、寿命即将用尽，坏道数过多等等。
- 当硬盘出现卡顿、变慢、拥塞等异常情况时，可能会影响整体的性能，此时通过亚健康盘隔离和读写切源保证性能的连续性

坏道预测

- 基于AI算法，可以精准地预测出环境中的机械硬盘的坏道的发生情况，根据实际环境的测试数据，7天内坏道的预测准确率高达95%以上

底层可靠保障——数据多副本机制

1. 全局参与，多并发

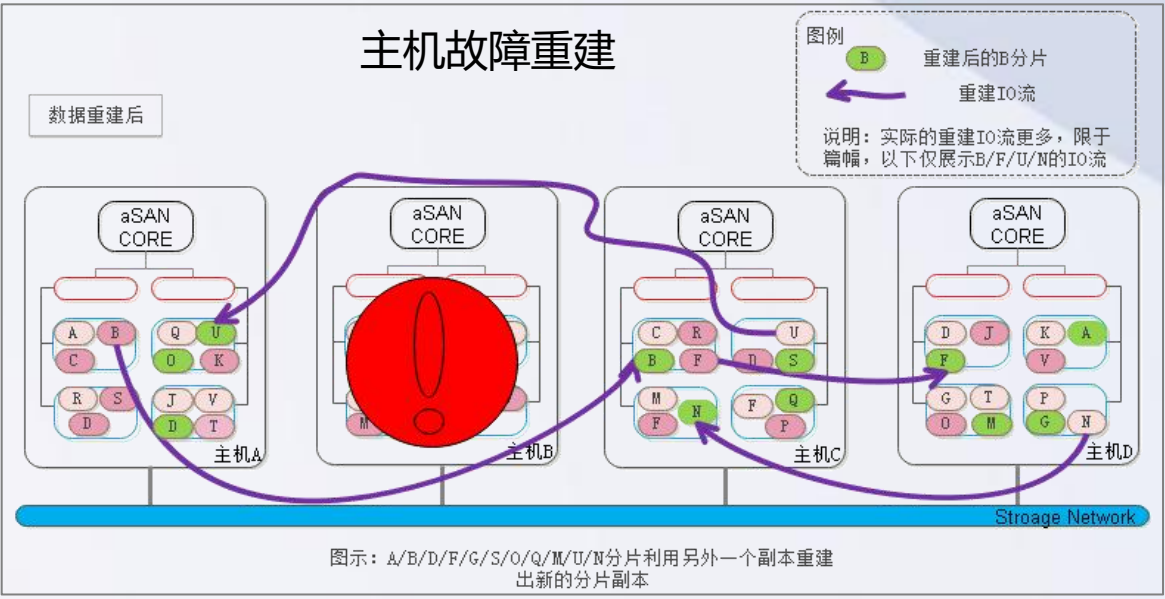
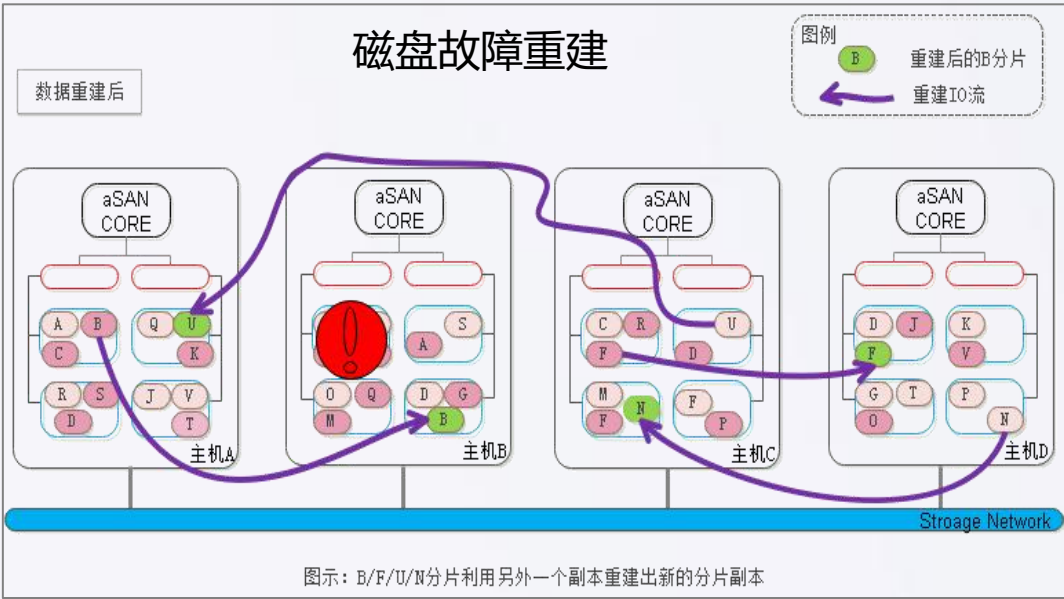
数据重建的I/O是多并发式的，即从多个源端磁盘读取，往多个目的端磁盘写入的，实现了数据的快速重建，可达**每TB/30min以上**

2. 智能重建

数据在重建过程中会占用一部分的存储网络带宽和磁盘性能，那么重建程序可以感知到上层业务的情况并以此智能地调整重建所占用的I/O，在保障业务正常运行的前提下快速重建数据。

3. 分级重建

重建任务发起后，系统默认会优先重建对用户重要的虚拟机数据进行重建，并且可以由用户自定义优先级。在存储卷可用于重建数据的空间资源紧缺时，通过分级重建可以优先保障用户重要的数据。



目录

一、DMP整体技术框架

二、高可靠

1. 集群架构部署
2. 自愈能力
3. 备份保护
4. 监控体系
5. 底层可靠

三、高性能

1. 最佳实践调优
2. 存储层优化
3. 性能测试数据
4. SQL级性能诊断

四、易运维

1. 业务快速上线
 - Oracle DTS
 - 自动化部署
2. 生命周期管理
3. 租户自服务

最佳实践调优

将经过经验调优的数据库最佳实践配置直接产品化，保障自动部署的数据库即具备最佳参数，既保障部署数据库的性能，同时显著提升部署效率；从下层VM到上层数据库均有覆盖

MySQL

VM 配置：

重要虚拟机,应用NUMA调度,使用Host CPU 高性能时钟,使用大页内存,使用预分配磁盘,...

数据库配置：

innodb_**_io_threads,innodb_log_buffer_size,i
nnodb_flush_method,...

高可用调优配置：

master_info_repository,
relay_log_info_repository,
slave_preserve_commit_order, ...

安全配置：

safe-user-create, local-infile,
allow-suspicious-udfs, ...

Oracle

VM 配置：

重要虚拟机,应用NUMA调度,使用Host CPU 高性能时钟,使用大页内存,使用预分配磁盘,...

数据库功能配置：

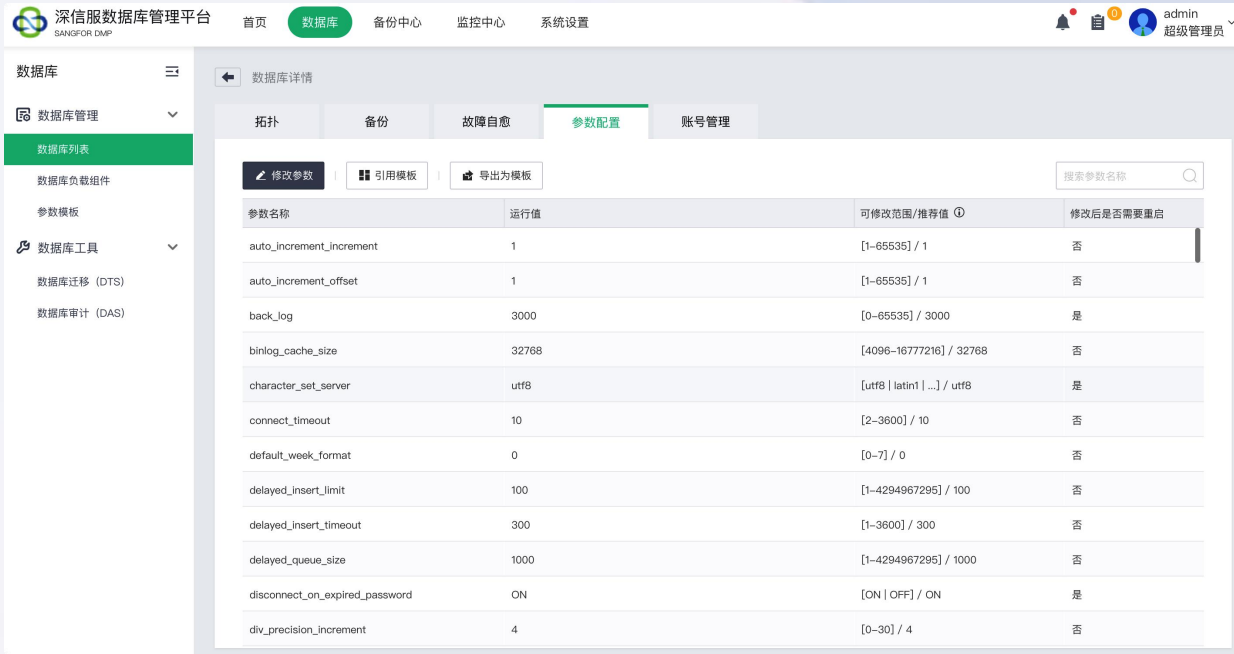
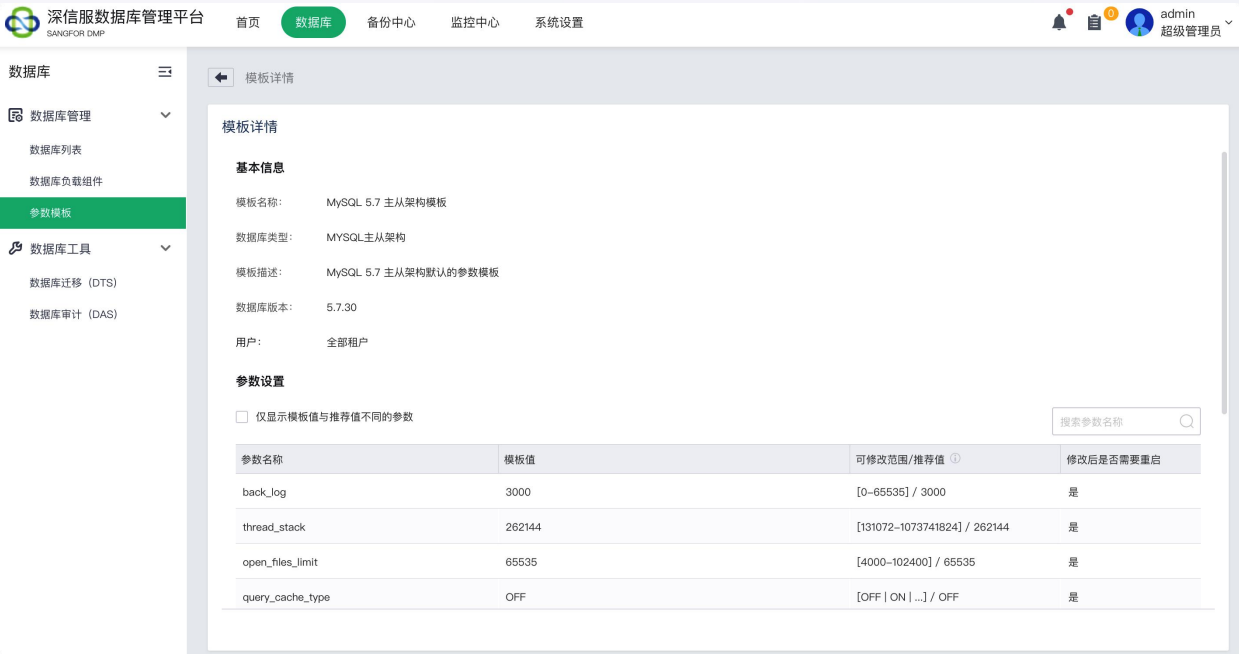
max_dump_file_size,
standby_file_management standby,
db_create_file_dest Oracle-managed,
_gc_bypass_readers、_gc_policy_time、
_gc_read_mostly_locking,

性能调优配置：

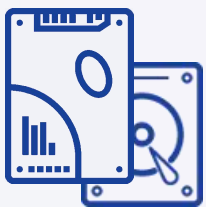
_use_adaptive_log_file_sync,
archive_lag_target,
db_file_multiblock_read_count, db_files,
open_cursors, optimizer_index_caching,
optimizer_index_cost_adj,
session_cached_cursors,

最佳实践调优——产品实现

- 默认的参数模板，对于各层面默认参数的修改，从虚拟机层、网络层、数据库层多个维度，提高了部署效率，同时保证数据库运行的性能、稳定和安全；
- 支持可视化参数调整；
- 实验环境实测自动创建耗时：
 - Oracle单机：10min
 - Oracle RAC：50min
 - MySQL 单机：2min
 - MySQL 主从：5min

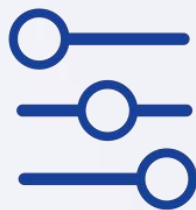


存储层六大措施提升随机IO性能



高效分层
SSD+HDD

SSD作为缓存层提供极致的随机IO性能



智能条带化

通过条带化，充分利用集群中所有的磁盘组，
让所有的磁盘组同时进行读写操作



IO本地化

基于本地聚合副本，实现本地读，
避免跨网络读取数据带来的延时问题



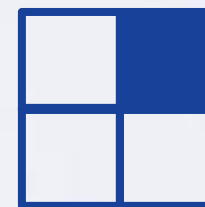
数据最优分布

散列副本充分利用集群的存储空间，
实现数据最优分布



磁盘分卷

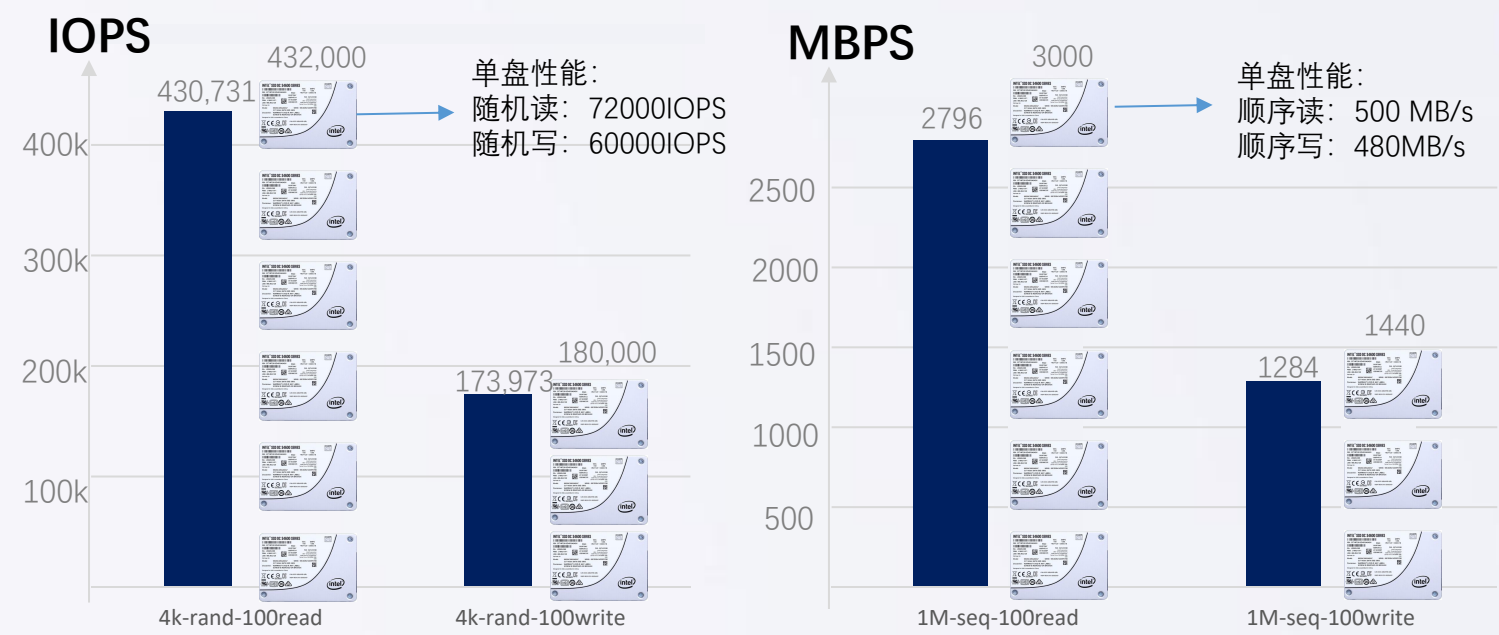
在混闪配置下，构建全闪存的存储空间
提供高性价比的解决方案



热力图算法

基于热力图算法对冷热数据做分层，
以最近访问时间和访问次数作为
数据冷热的评判标志

充分挖掘SSD随机IO读写性能



集群规模	3台服务器
单节点配置:	
CPU	2 × Intel Xeon E5-2680 v4 @ 2.40GHz
RAM	128GB
SSD	2 × Intel S4600 480GB
HDD	6 × 1TB
Nic	2 × 10GbE
超融合软件版本:	
aCloud	5.8.7R1
测试用例	10VM3DISKs

充分发挥出物理硬件的性能**90%以上!**

Oracle性能测试——测试环境

硬件配置

服务器	型号	aServer-P-2300
	CPU	Intel(R) Xeon(R) Gold 6132 CPU @ 2.60GHz
	内存	224 GB (32GB * 7)
	系统盘	1 * 128GB SSD
	缓存盘	2 * TOSHIBA KHK61RSE 960G
	数据盘	6 * 4TB SATA HDD
	网口	6 * GE电口 + 2 * 10GE光口
	数量	3台
交换机	万兆交换机	2台
	千兆交换机	2台

软件配置

HCI	版本	HCI6.3.0R1
SCP	版本	SCP6.3.80/SCP6.3.70
	规格	8C16G
DMP	版本	DMP2.1.0
	规格	8C16G
DBVM	Oracle镜像	Oracle11.2_Centos7.9_2.1.0.vma
	MySQL镜像	MySQL5.7.30_Centos8_2.1.0.vma

测试模型

模型	条目	用户注册	更新用户信息	浏览商品	订购商品	处理订单	浏览订单
交易模型1		1%	1%	85%	5%	3%	5%
交易模型2		10%	10%	35%	35%	5%	5%

Oracle性能测试——测试数据

测试结果

规格	压测并发（用户数） 交易模型1	TPM
Oracle 单机 16C64G	1000用户	474276
	1200用户	509204
	1400用户	512114
	压测并发（用户数） 交易模型2	TPM
	400用户	162027
	500用户	202550
	600用户	241096

规格	压测并发（用户数） 交易模型1	TPM
Oracle 单机 32C128G	2600用户	663259
	2800用户	660505
	压测并发（用户数） 交易模型2	TPM
	400用户	160268
	600用户	241091
	800用户	315924

规格	压测并发（用户数） 交易模型1	TPM
Oracle RAC 16C64G	2400用户	834952
	2600用户	841366
	2800用户	852833
	压测并发（用户数） 交易模型2	TPM
	1400用户	456723
	1600用户	472115
	1800用户	469946

规格	压测并发（用户数） 交易模型1	TPM
Oracle RAC 32C128G	3400用户	1182117
	压测并发（用户数） 交易模型2	TPM
	1200用户	669971
	1400用户	703906
	1600用户	756245
	1800用户	815566

由上述结果可见，RAC32C128G 峰值可达约120W TPM

提供SQL级别洞察能力，提供性能诊断（Q4发布）

深入的关联分析，从故障发现到问题源头的告警下钻

智能化关联分析

某时间点慢

点击对应时间点
即找到问题语句

语句执行时间分析
千行过程哪行慢

硬件消耗多少

其他语句阻塞多少

语句如何调优

CPU/内存/IO
卡在什么资源

被什么语句阻塞

调优建议和脚本

从外到内，从现象到本质，层层剖析



硬件：服务器、存储、网络

环境：OS、DB版本及配置

运行：Linux、Oracle层面的信息

行为：指标、语句、会话

确认：问题的根源

目录

一、DMP整体技术框架

二、高可靠

1. 集群架构部署
2. 自愈能力
3. 备份保护
4. 监控体系
5. 底层可靠

三、高性能

1. 最佳实践调优
2. 存储层优化
3. 性能测试数据
4. SQL级性能诊断

四、易运维

1. **业务快速上线**
 - Oracle DTS
 - 自动化部署
2. **生命周期管理**
3. **租户自服务**

业务快速上线——Oracle数据库迁移背景知识

迁移场景需求分析

硬件系统升级	升级存储、磁盘、CPU，以满足日益增长的业务性能需求。
数据库版本升级	旧版本数据库功能受限，需要满足更复杂的业务需求，需要升级最新版本的数据库。
数据同步	数据迁移以迁移整个数据库为目的，数据同步以维持不同业务或者不同数据库之间的数据持续性流动。进行数据同步的基础首先要完成对数据的全量迁移。
数据库上云	为了更好地管理数据库、节约数据服务成本，将数据库迁移到公有云或者私有云平台。
数据多活灾备	为了保障节点的安全性和稳定性，解决某一节点或者实例的因为出现数据出错导致不可用，数据复制服务推出灾备场景，为用户业务连续性提供数据库的同步保障。数据多活灾备一般分为两地三中心、两地四中心灾备架构。



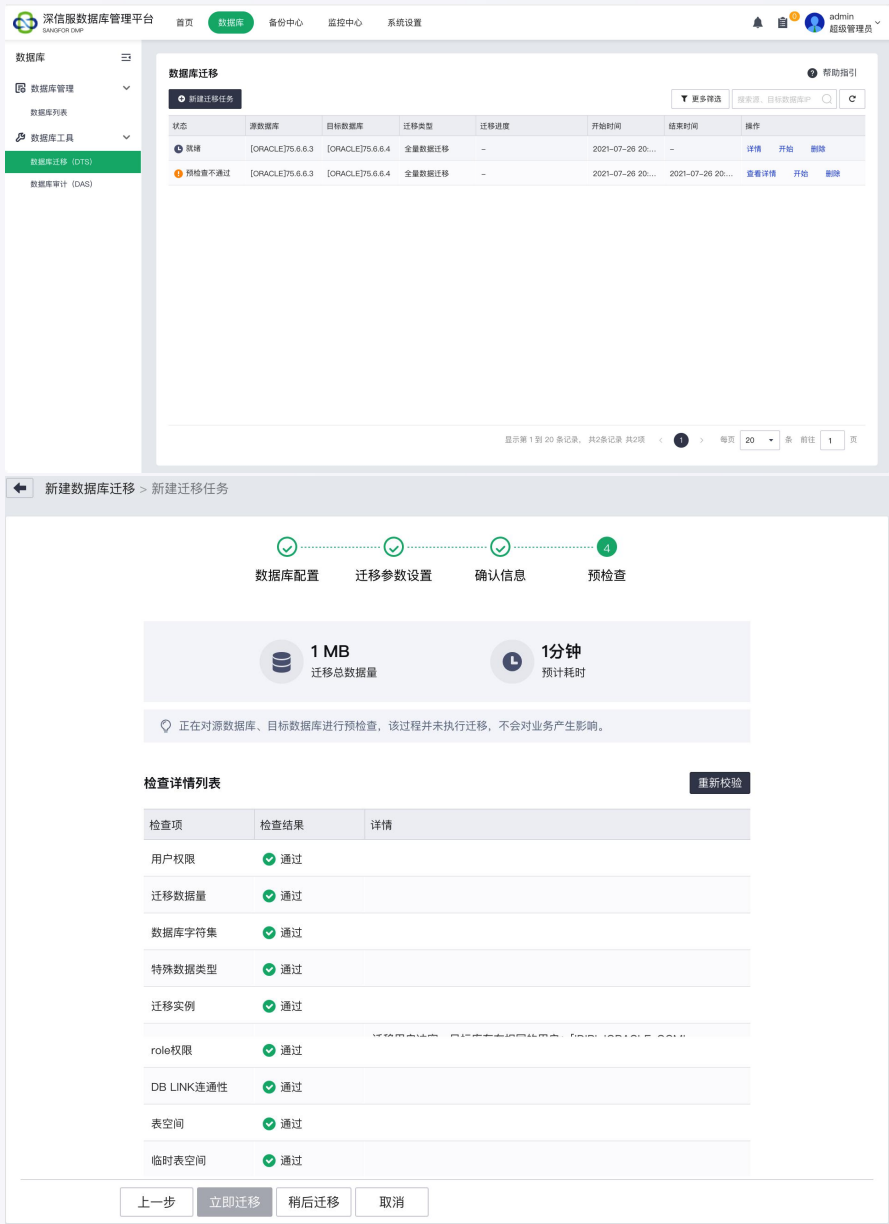
迁移过程注意事项

迁移前	需要了解数据库版本，数据库架构，表空间状态，表空间数量，表空间容量，主外键，字符集，特殊字符类型如Blob、Clob，检查数据库中视图、索引、序列、存储过程、同义词、定时脚本以及物化视图、db_link等对象是否存在及其状态等信息；
迁移中	需要注意源端与目标端的网络是否出现抖动、是否有中断的情况发生。 迁移注意预留充足磁盘空间。
迁移后	业务验证，一般都需要咨询业务人员或者开发人员；应用割接方式：使用配置管理库，使用DNS方式，直接改连接串，无论哪种方式，均有短时的业务中断

可以看到，Oracle迁移需要考虑的关联项非常多，操作也非常复杂，需要**依赖很深的专业知识，实施难度巨大**。

DMP将迁移任务统一考量，在**平台内置了DTS迁移工具**，简化迁移任务，方便用户进行业务上线。

业务快速上线——Oracle DTS实现



DTS实现逻辑



DTS (Data Transmission Service)

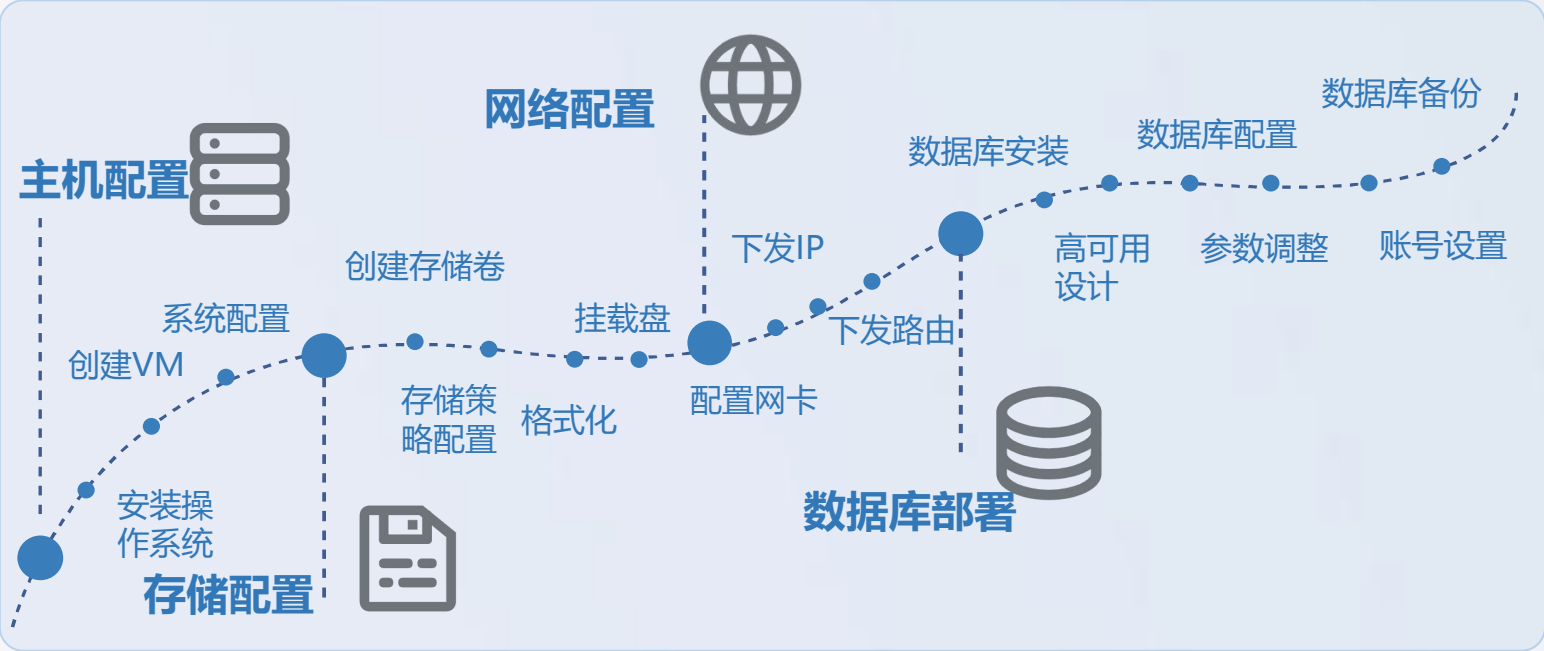
DTS选用**数据泵**作为基础工具进行整合开发

- 可支持Oracle数据库自动化在线迁移
- 支持全量迁移和结构迁移两种迁移类型
- 支持网络传输DB-link、本地转储两种方式进行迁移

极大**简化迁移步骤**复杂度:

- 用户只需从页面进行相关配置即可操作数据库迁移
- 待迁移任务完成后, 需进行业务确认, 方可按需进行业务割接

业务快速上线——数据库传统部署模式



传统的数据库部署模式包含

- 硬件设备主机、存储、
- 网络搭建、配置、
- 数据库安装、参数配置等步骤，过程复杂，往往要耗费数天，严重拖延业务上线进展

人为操作往往是问题的根源

一个“锁表”损失800万，运维被判5年半

51CTO · 2020-05-08 阅读数 1229

近日，云头条发布的“一个违规操作、损失 800 万、被判五年半：运维夏某某致判瘫痪 2 个小时，判破坏计算机信息系统罪”一文引发了技术圈的热议。

一个致命的Redis命令，导致公司损失400万!!

慕课网 已认证的官方帐号

302 人赞同了该文章

最近安全事故频发啊，前几天发生了“[顺丰高级运维工程师的删库事件](#)”，今天工程师在线执行了Redis危险命令导致某公司损失400万..

一个数据库脚本bug导致Salesforce史上最严重宕机

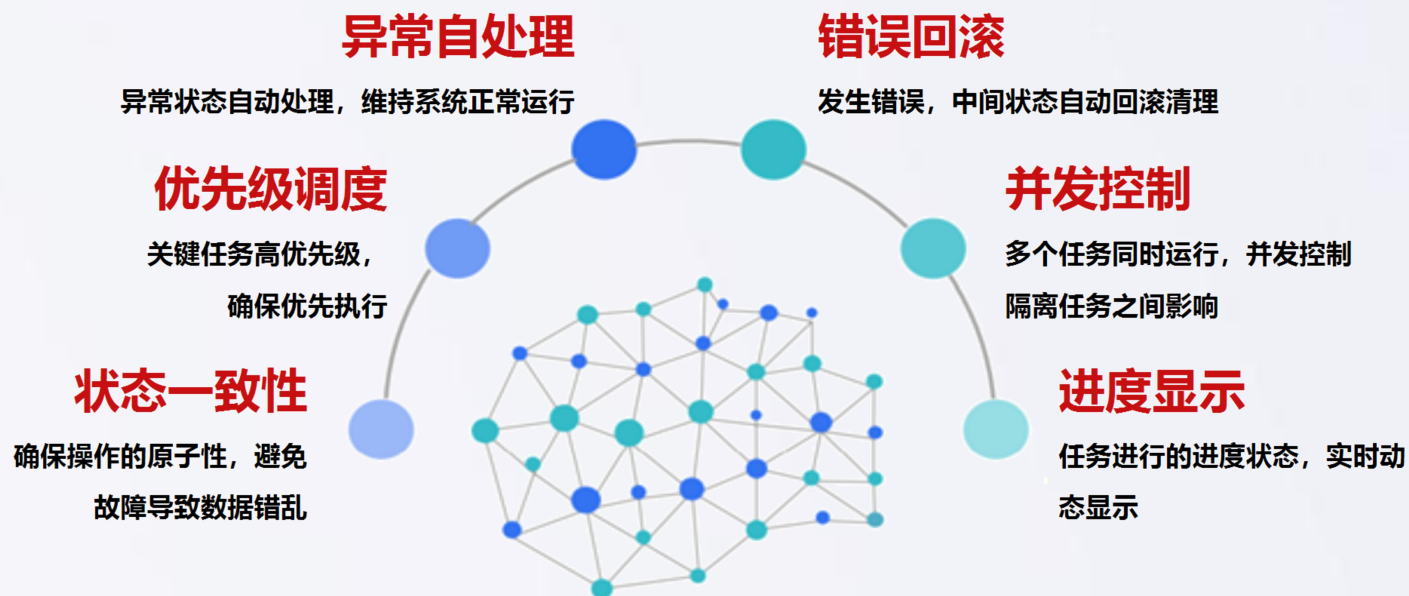
作者: CashCat 星期一, 五月 20, 2019 动态, 数据安全 无评论

这个周末很多Salesforce客户的员工们发现无法继续享受996福报，因为Salesforce，宕机了。

业务快速上线——数据库自动化部署框架

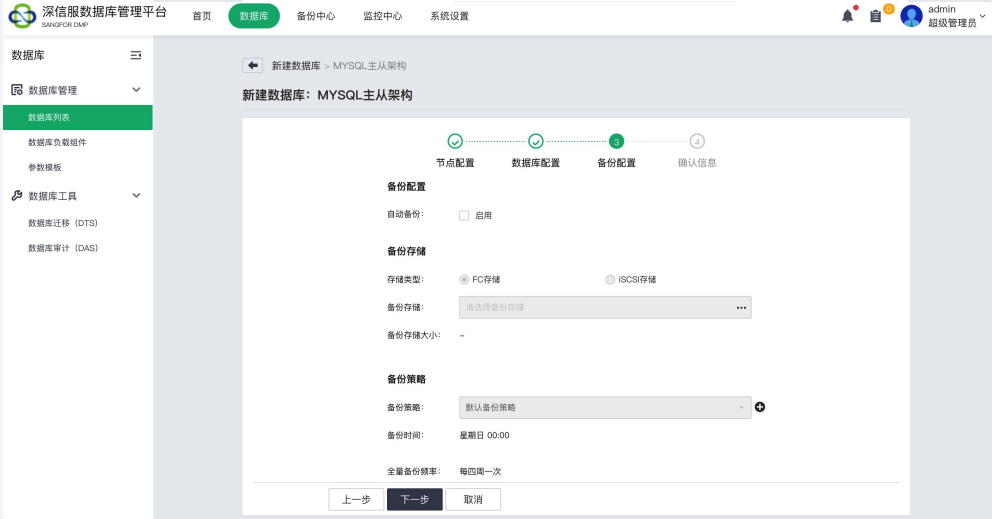
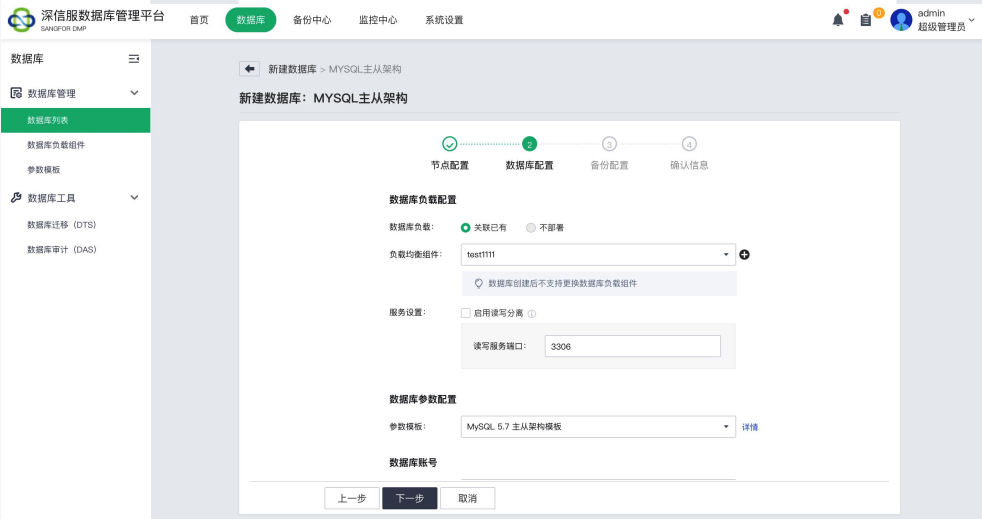
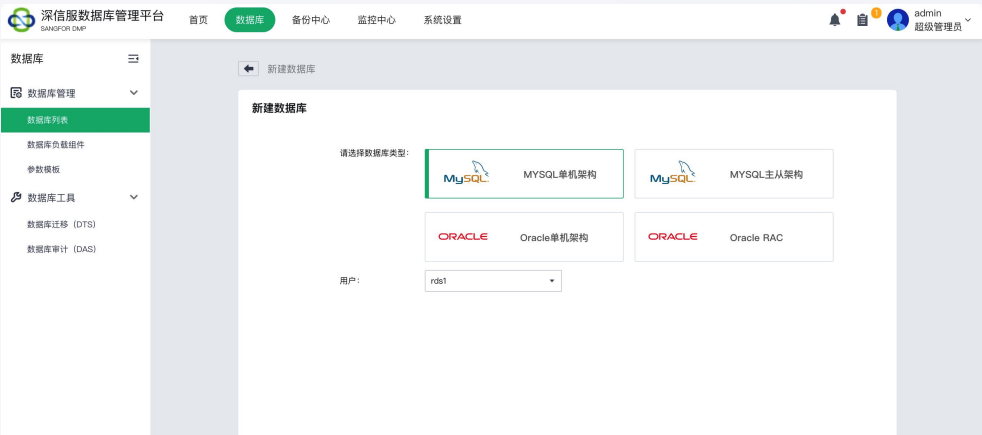
- **挑战：**在创建特别是集群的创建过程中涉及的步骤**多达二十多步**（创建 VM，创建盘，格式化盘，挂载盘，创建双网卡，下发 IP，下发路由等等等等），这些步骤不仅需要**严格的时序同步**，而且更重要的是**每一步都可能出错**需要回滚。一旦状态残留就可能埋下不可预测的风险
- **解决：**基于 openstack 的 taskflow 组件设计并实现了高可靠的工作流执行和回滚框架 dmp-framework。

workflow引擎核心特性



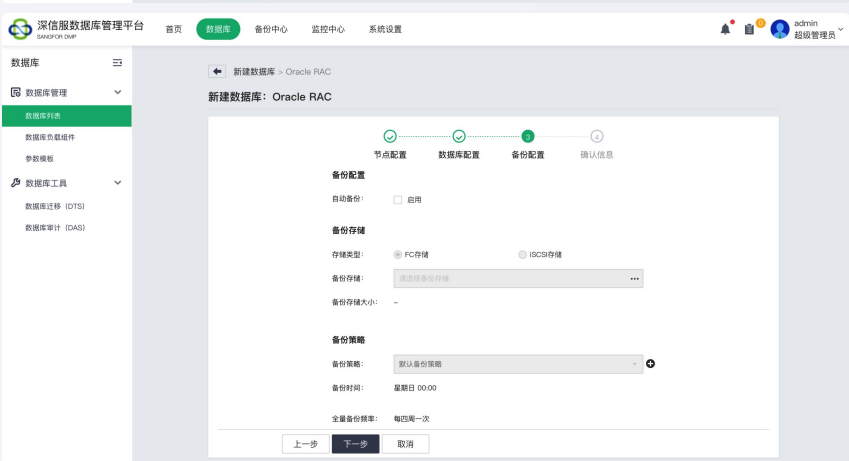
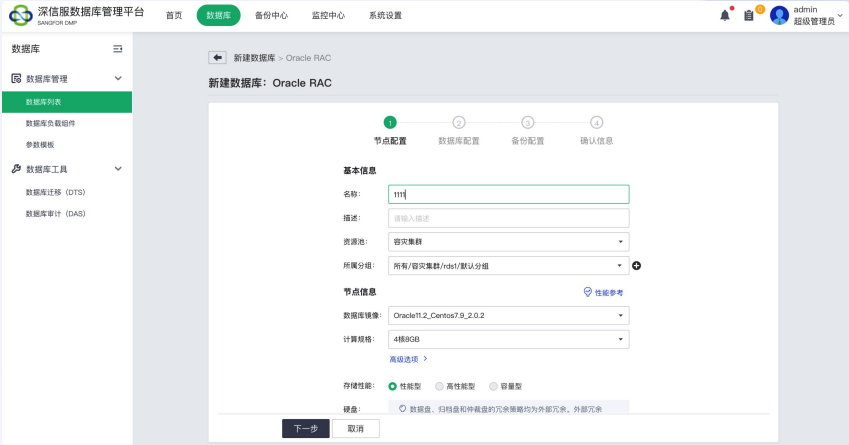
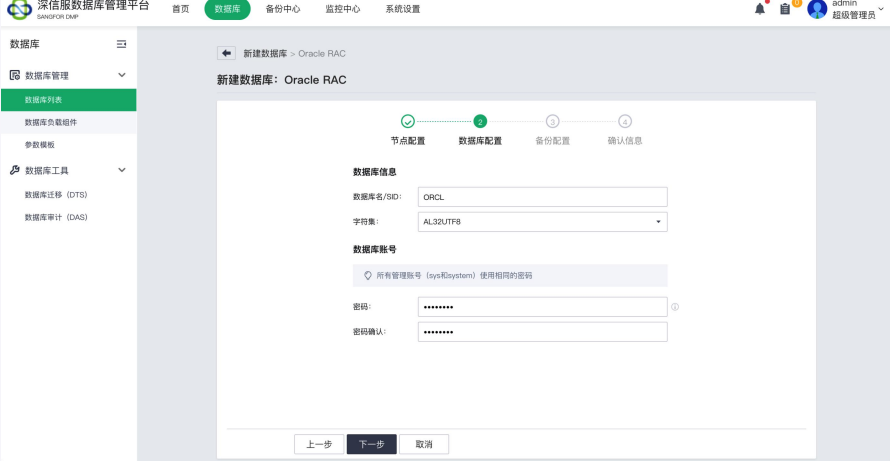
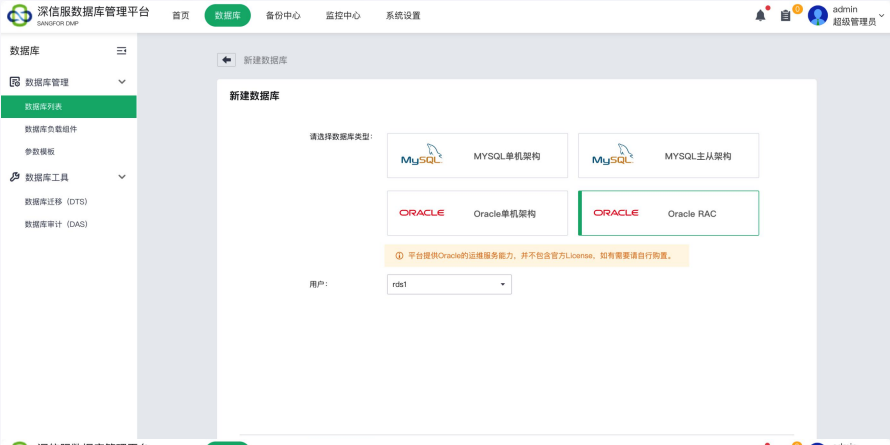
业务快速上线——MySQL自动化部署

- 向导式部署数据库集群：可选MySQL单机、一主一从架构
 - 第一步：数据库基本信息和节点配置（主从架构的主从节点配置统一）
 - 第二步：数据库负载和参数、账号密码设置
 - 第三步：勾选开启备份策略
- 可在列表统一查看数据库集群

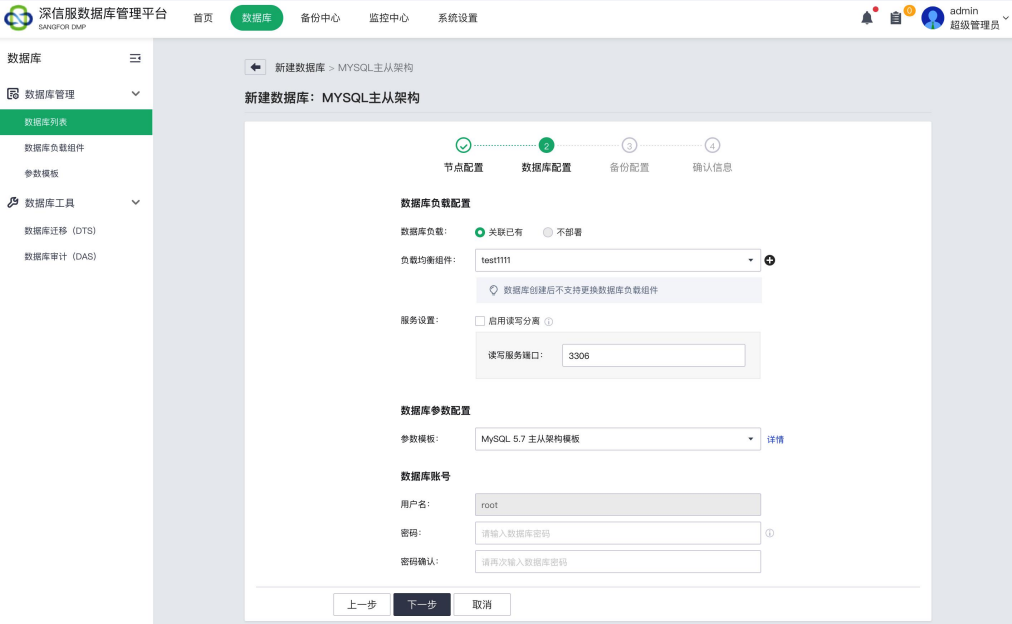


业务快速上线——Oracle自动化部署

- 向导式自动化部署：可选Oracle单机、RAC（2节点）
 - 第一步：数据库基本信息和节点规格、SGA、PGA配置（RAC节点配置统一，RAC自动完成ASM配置，VIP、SCANIP均自动生成）
 - 第二步：数据库SID、字符集、管理账号配置
 - 第三步：可选开启备份策略、配置备份存储
- 可在列表统一查看数据库集群

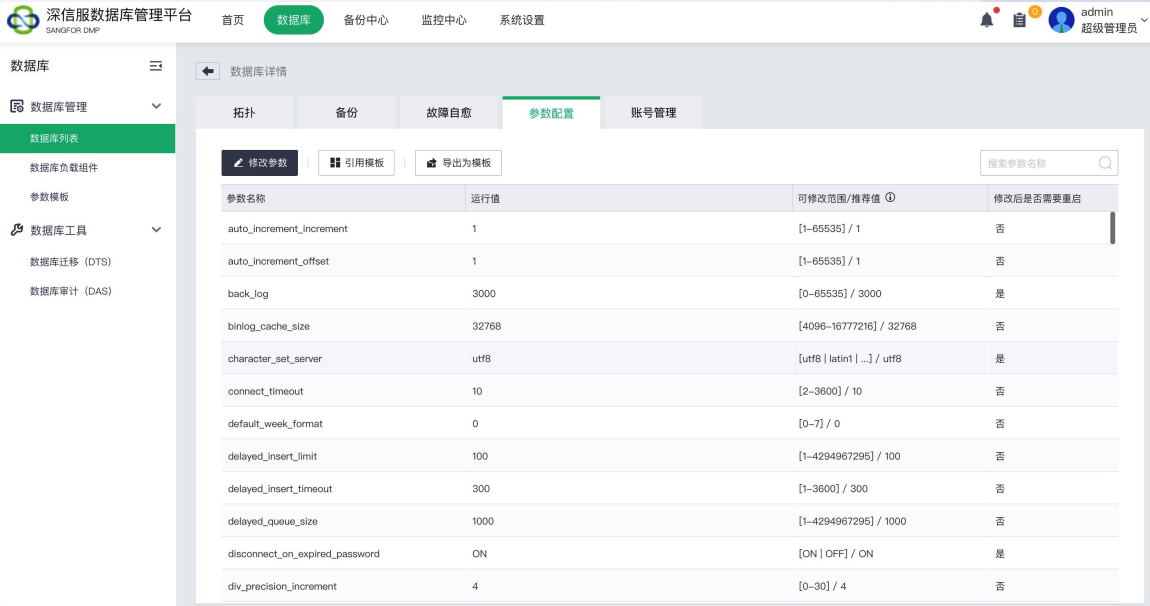
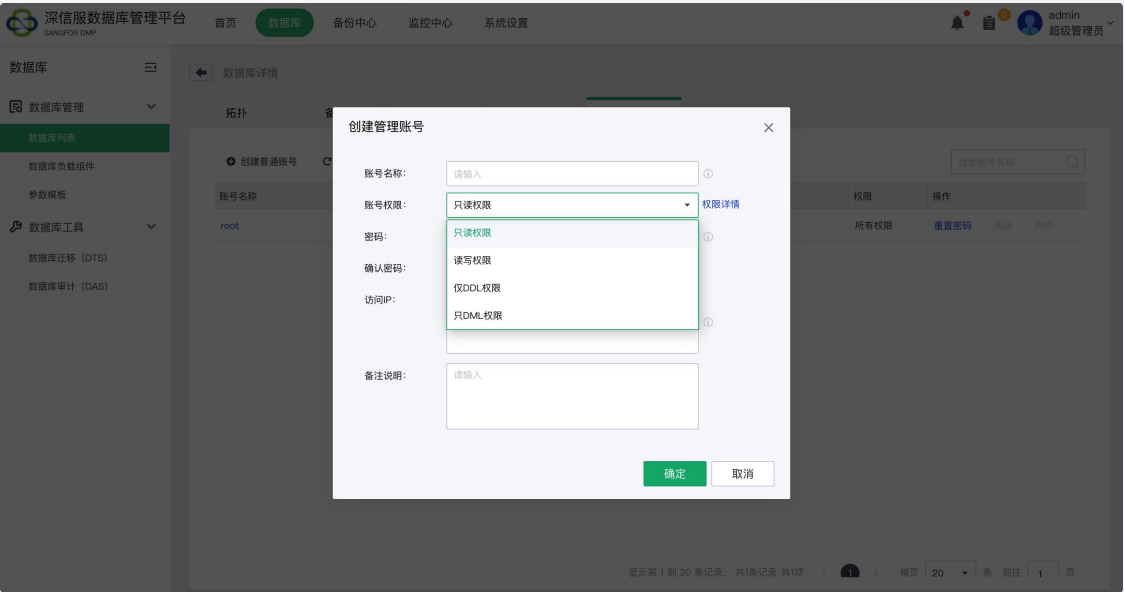


数据库生命周期管理——MySQL参数及账号管理

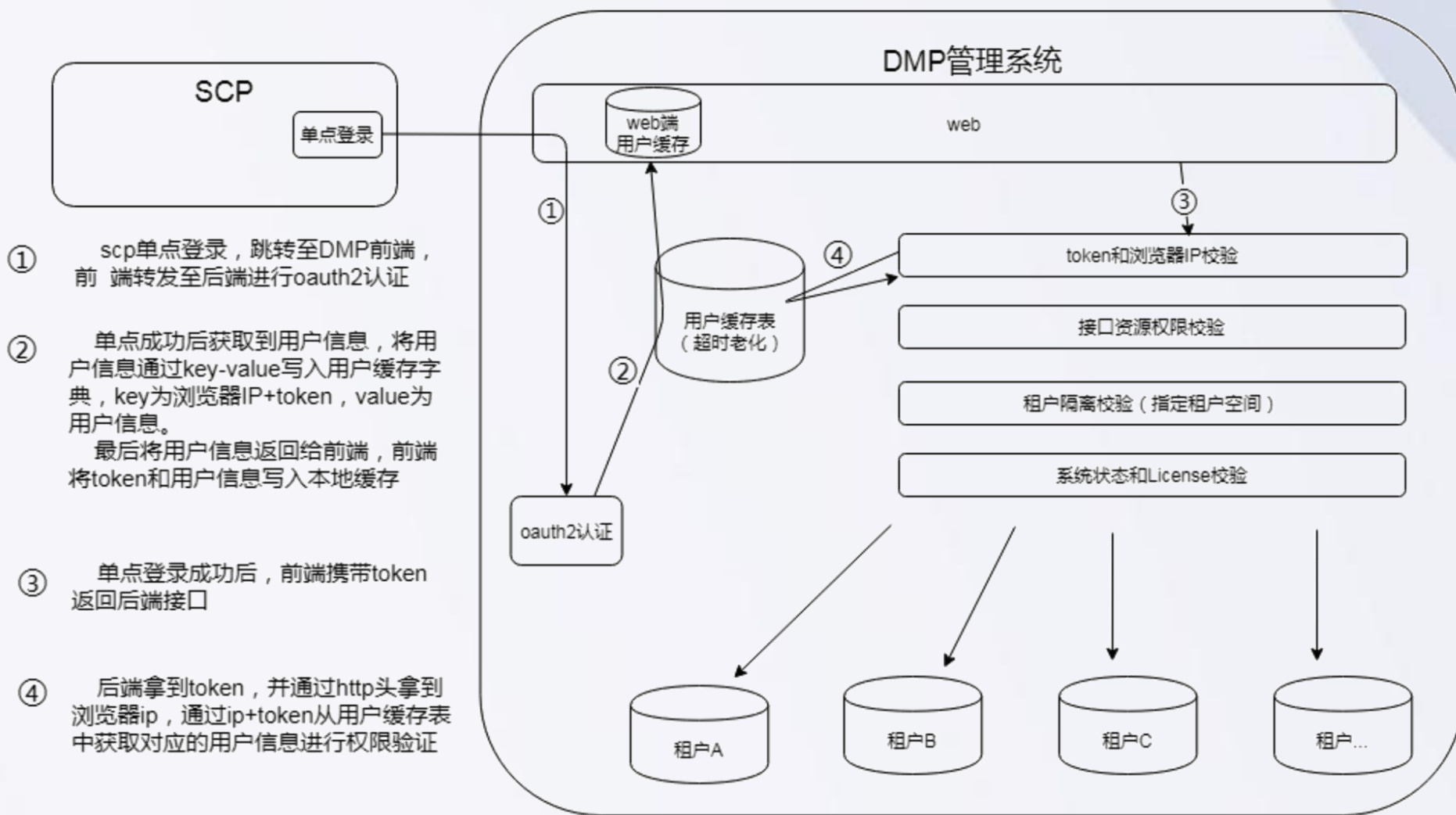


支持数据库内部参数及账号管理

- 预置调优过的参数模板，可在部署阶段直接选用
- 支持在平台页面调整内部参数，共计160+项，无需登录数据库即可调参；
- 在部署阶段即创建管理账号
- 数据库内部的账号，分为管理账号和普通账号，均可统一在平台进行管控，可自行定义普通账号权限，包括只读、读写、DDL、DML



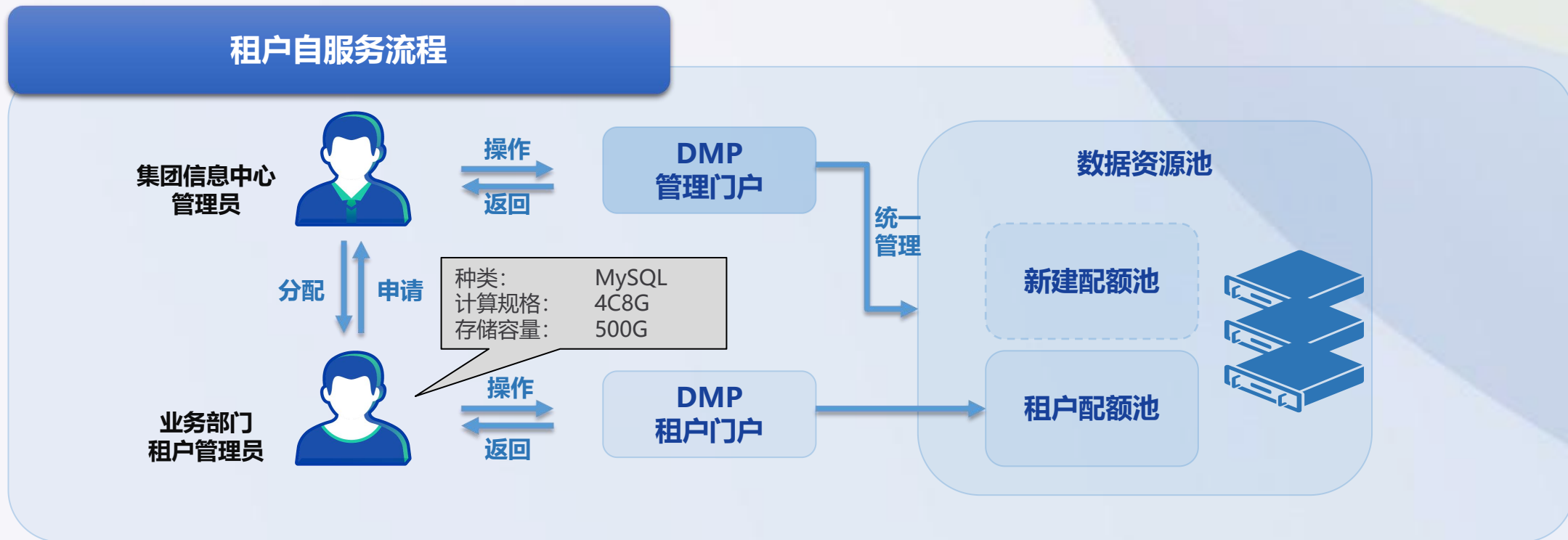
租户自服务——云平台多租户账号体系打通



DMP实现了与SCP账号体系的打通，可以支持：

- SCP单点登录
- 多租户账号同步
- 授权配额同步校验
- SCP租户账号可以直接通过SCP跳转登录至DMP；
- 租户在SCP所分配的RDS配额可以直接在DMP进行查看、使用；

租户自服务——云平台多租户账号应用场景



- 借助租户自助服务，实现数据库服务化交付，**标准化管理**各个子单位的数据库
- 自助式自动化完成数据库上线、运维管理，无需依赖运维人员，**减少人为误操作带来的隐患**

云平台存在多租户的使用场景，DMP也配套了与之匹配的租户账号体系，方便在多租户场景下使用DMP进行数据库相关运维管理操作。

谢谢
Thanks