

vinchin

# 聚焦企业关键业务 分布式数据库保护实战

云祺科技



# 目录

---

- PART 01 背景与挑战**
- PART 02 数据库介绍与与保护技术分享**
- PART 03 实操演示**
- PART 04 特性与优势总结**
- PART 05 应用场景**

# PART 01

## 背景与挑战



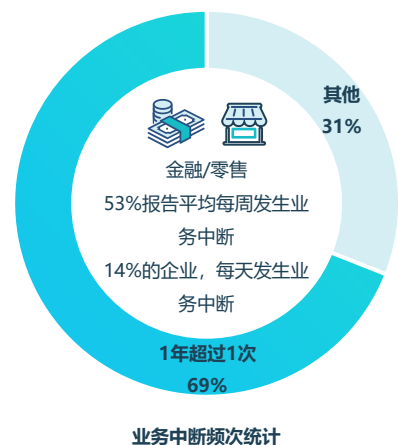
# 承载关键业务·安全不容有失



数据库作为信息系统的核心，通常起关键业务承载作用，不仅管理着大量数据，还负责向各类用户、关联系统提供高效、可靠的信息服务，因此确保数据库的安全性及其数据完整性对于业务平稳运行至关重要。尤其是金融、电信、政府等关键行业，对数据库的安全性、稳定性有着极高的要求。

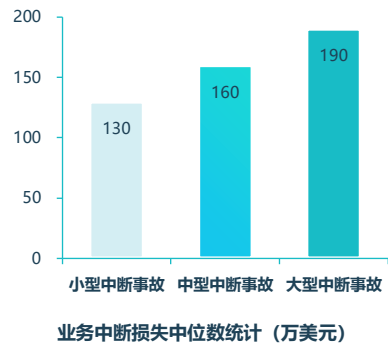
# 业务中断损失重大

## 多数企业每年遭遇业务中断



数据源: New Relic-2024 Observability Forecast

## 中断损失10万美元/hour



数据源: Cockroach Labs-The State of Resilience 2025

## 业务中断使企业面临多重损失

# 44%

### 生产经营损失

生产停滞、工作流程中止或进度缓慢、产品或服务订单交付能力受限等导致的直接收入损失

# 17%

### 恢复/重建成本

硬性的恢复资源与环境成本、无法预估的人力投入以及无法保证完整恢复的技术支持服务支出

# 39%

### 其他隐形损失

品牌声誉损失以及客户信任流失将造成长期财务影响, 违规罚款、甚至吊销营业执照以及面临法律风险

发生业务中断的企业中, 绝大部分企业营收都会受到影响, 其中34%的企业收入大幅下滑, 而12%的企业则因客户赔偿、订单锐减、高额罚款、吊销许可等导致停业或倒闭

## 业务中断使安全团队面临个人风险

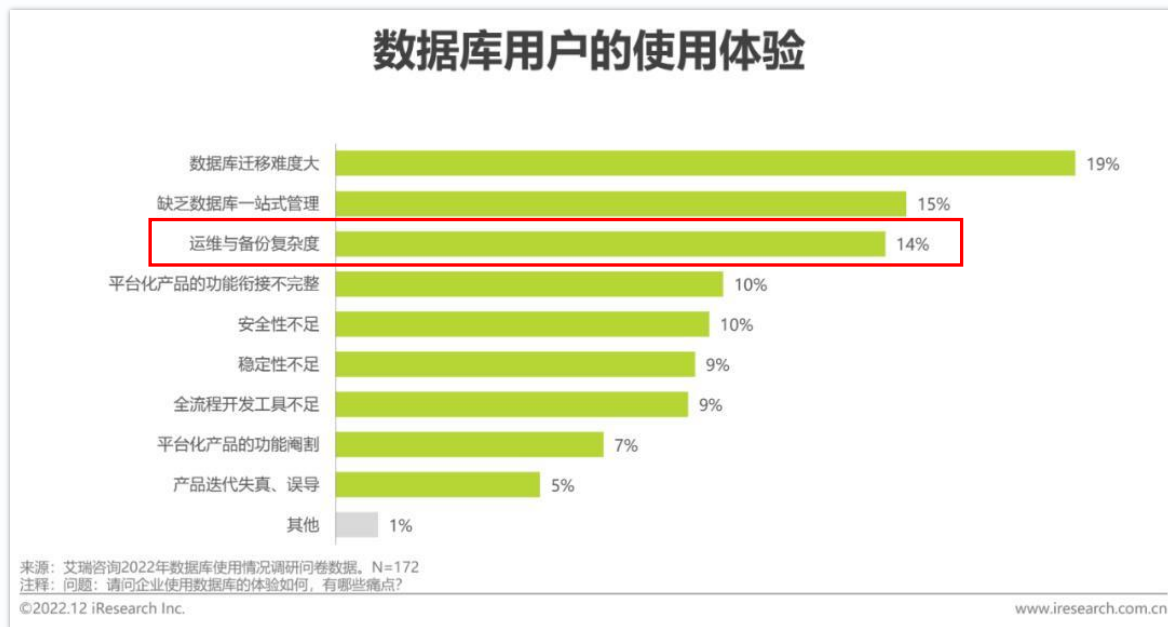
### 失职问责

### 精神压力

### 降职/解雇

### 法律责任

# 备份操作复杂·耗费大量成本



由于对数据一致性要求非常高，因此数据库备份通常比其他对象的备份具有更高的技术复杂度，企业不仅存在大量前期投资，更需要专业技术人员持续进行维护，耗费大量人力成本和时间成本

## 自带备份工具功能有限

数据库自带的备份工具虽然与数据库紧密集成，但只提供有限的备份功能，无法满足自定义的用户备份需求

## 脚本备份技术门槛高

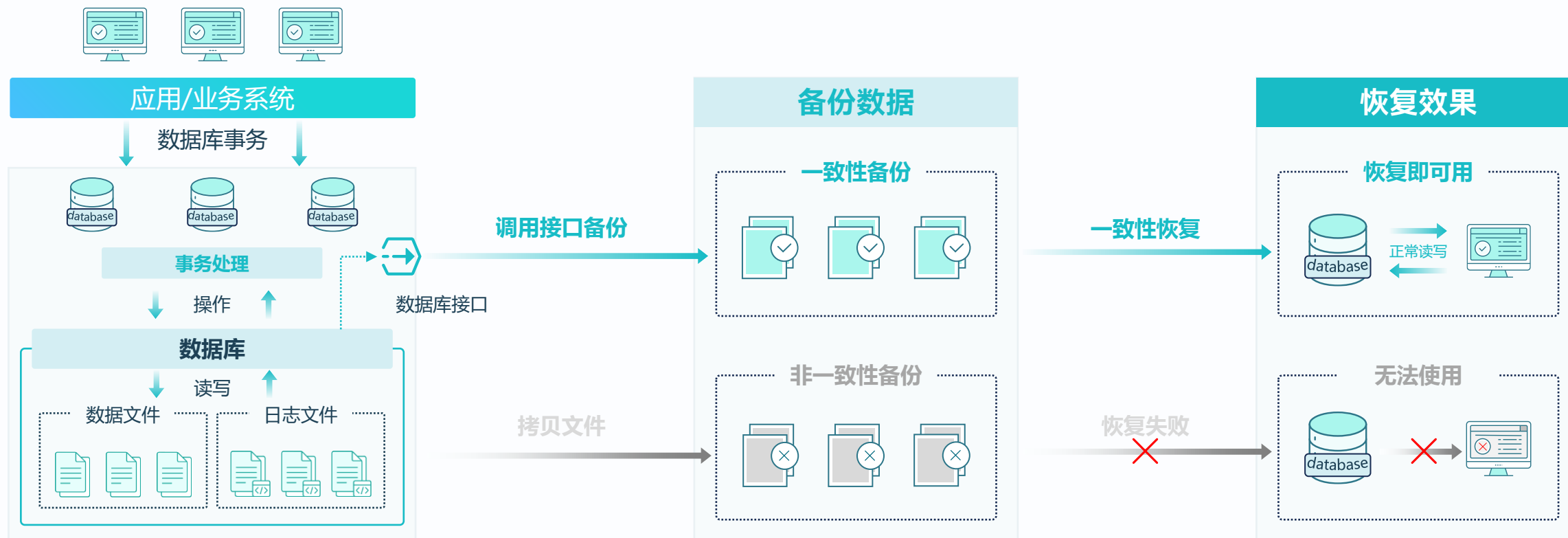
脚本备份缺乏图形化界面，且需要跟随系统环境变化频繁更新，对运维人员技术要求很高，还有可能因为脚本错误导致备份失效

## 占用生产系统资源

传统的手动备份和脚本备份都会在生产系统产生缓存文件，对于存储空间紧张的系统，大量备份可能导致系统性能受影响

**需要简单、易用的数据库备份！**

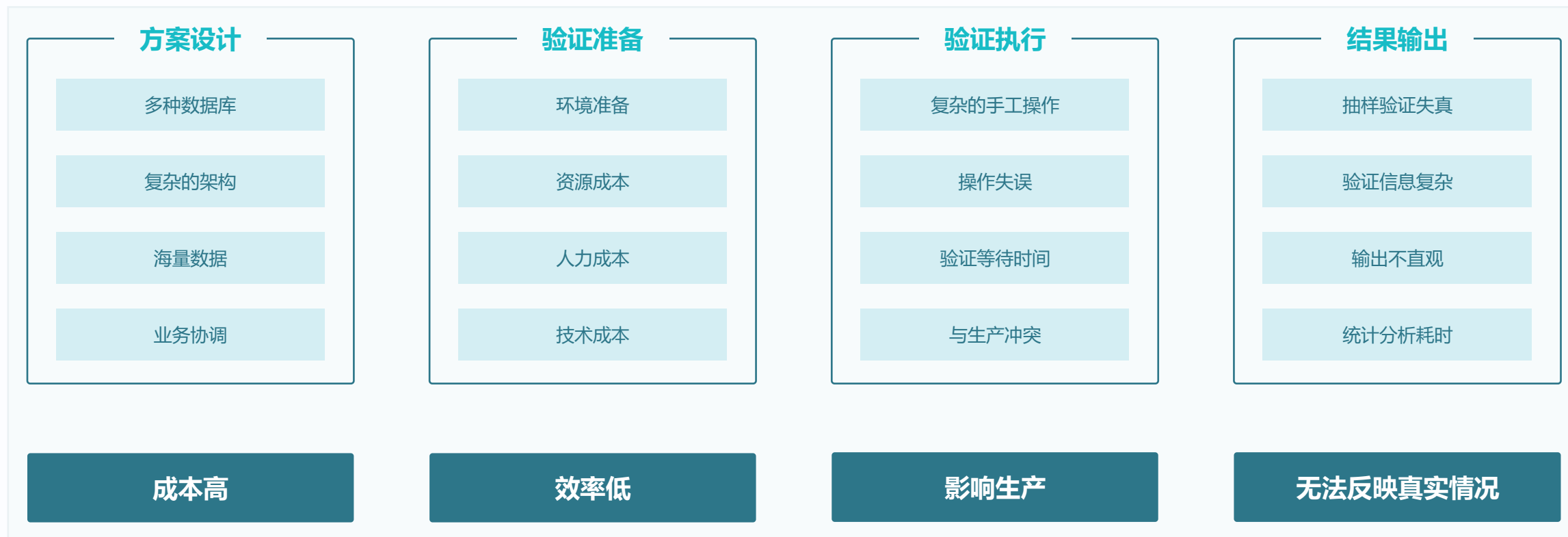
# 数据一致性要求高



**无法保证数据一致性的备份是无效备份!**

# 数据验证困难·恢复没有保障

备份是保障数据安全的“最后一道防线”，在遭受数据灾难时，能否使用备份数据成功恢复至关重要，未经验证之前，备份数据处于何种状态属于未知。因此，我们有必要通过验证来确认备份数据的完整性和可用性，而在实际的验证工作中，我们会遇到各种难题。



需要便捷的验证手段，确保恢复成功率！

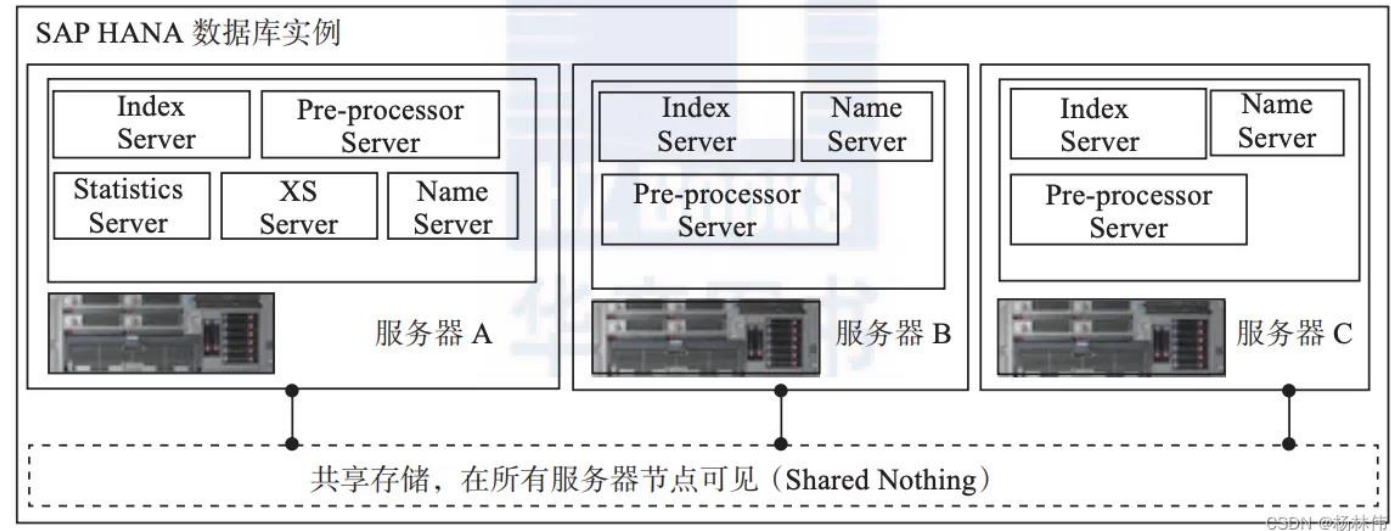
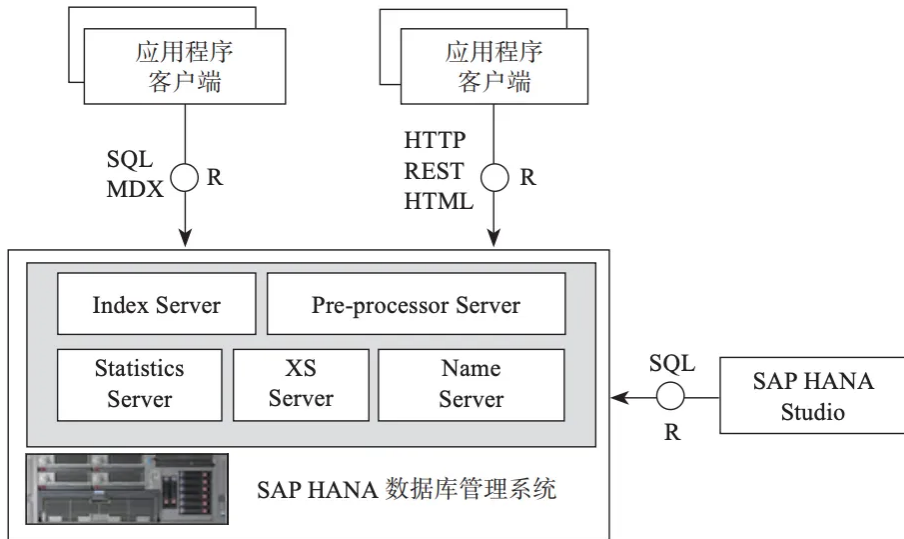
# PART 02

## 数据库介绍与保护技术分享



# SAP HANA数据库介绍

SAP HANA是一款内存数据库，将数据存储在内存中，可以几乎无延迟的处理海量数据与即时查询。通常作为大型企业的ERP以及其他业务应用的承载平台。



Name Server: 相当于整个 HANA 数据库系统环境中的“通信员”，通过 Name Server 可以知道当前 HANA 服务器的部署情况；

Index Server: 最核心的组件，承担着内存管理、事务管理、元数据管理器及权限认证、多版本并发控制 (MVCC) 等众多的管理工作；

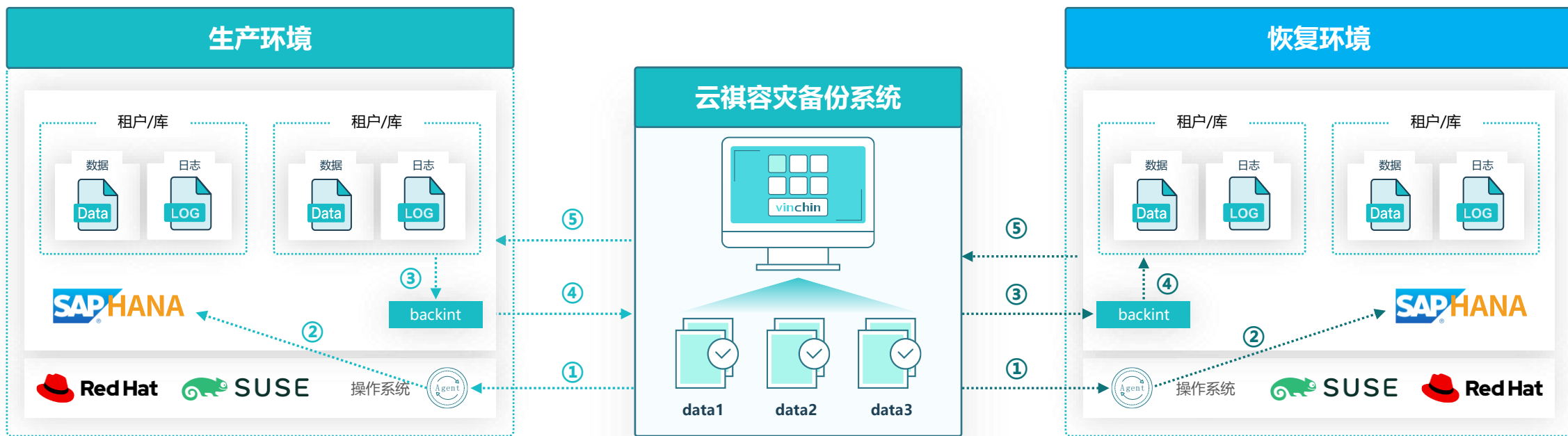
Statistics Server: 收集所有数据库组件运行的状态、执行效率和资源的消耗状态，还监控 HANA Studio 的访问并返回不同的信息给用户；

Pre-processor Server: 预处理器服务器

XS Engine : 将持久层的数据模型封装供外部使用，而且它具有对这些发布出去的服务进行搜索的功能，并且内置一个应用服务器；

Hdbdaemon: 以正确的顺序开始或停止其他进程

# SAP HANA备份恢复介绍



## 备份流程

- ①备份系统将备份配置（备份方式、备份策略等）传输到Agent；
- ②Agent连接HANA，按照备份设置执行备份命令并请求从backint接口获取对应数据；
- ③HANA解析备份命令后，将数据通过管道传输给backint，Agent创建传输通道并接收数据；
- ④Agent将从backint接收到的数据以数据流的方式，通过传输通道传输至备份系统；
- ⑤传输完成后，备份系统数据将备份成功的记录等信息返回给HANA，完成此次备份。

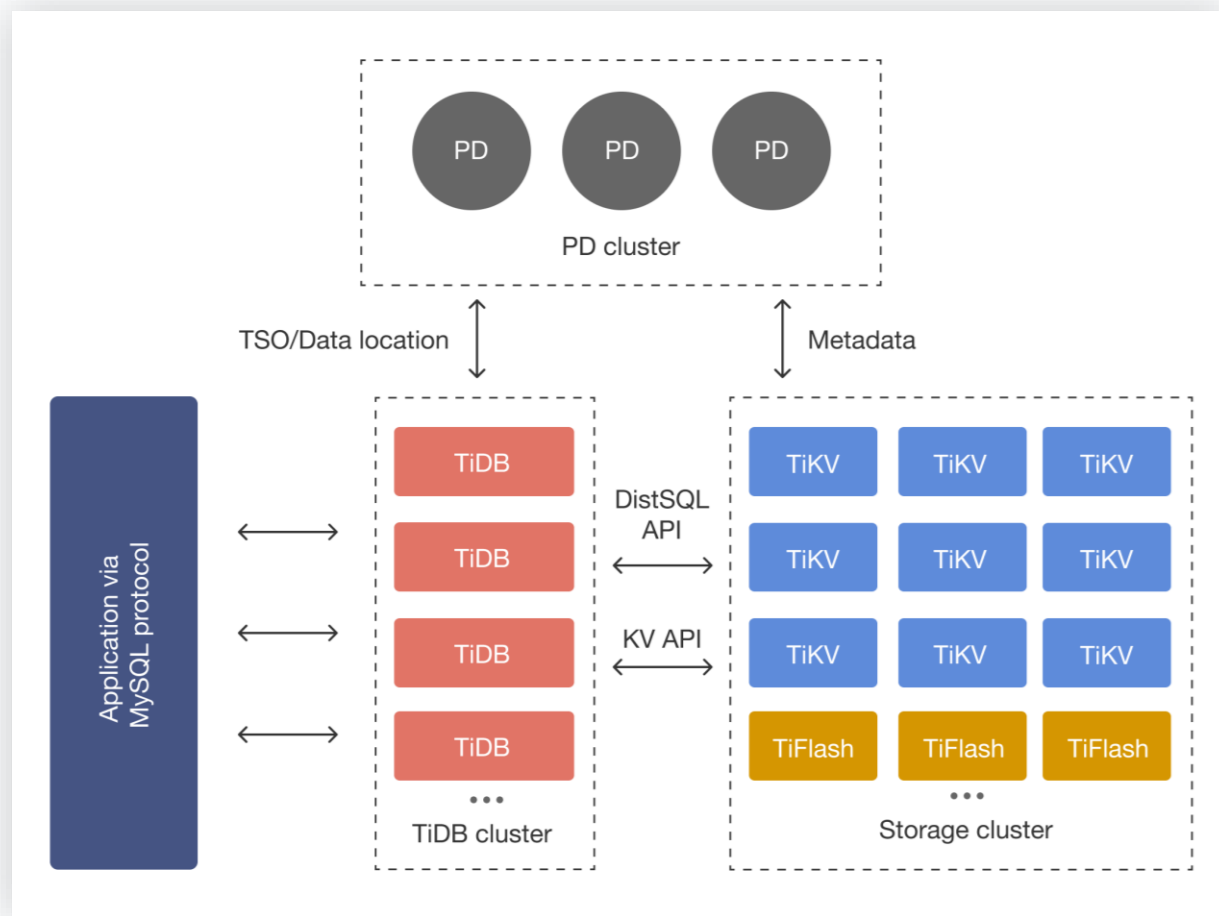
## 恢复流程

- ①备份系统发送恢复配置（如需要恢复的数据，恢复方式等）给Agent；
- ②Agent连接HANA并按照配置执行恢复命令，HANA解析恢复命令后请求获取对应备份数据；
- ③Agent创建传输通道，以数据流的方式将备份数据通过backint传递给HANA；
- ④HANA接收数据，并先恢复数据备份的内容，再按照顺序应用日志到指定时间；
- ⑤所有数据恢复完成后，HANA通过Agent向备份系统返回恢复成功的消息，完成此次恢复。

HANA的日志备份将一直持续，只要数据库触发了相关条件，云祺都会及时、自动对日志数据进行备份

# TiDB数据库介绍

TiDB 是 PingCAP 公司自主设计、研发的开源分布式关系型数据库，是一款同时支持在线事务处理与在线分析处理 (Hybrid Transactional and Analytical Processing, HTAP) 的融合型分布式数据库产品，具备水平扩容或者缩容、金融级高可用、实时 HTAP、云原生的分布式数据库、兼容 MySQL 协议和 MySQL 生态等重要特性。



## 主要组件与作用介绍

- 1、TiDB Server: 集群的计算节点。负责对外连接, 执行SQL等计算操作, 不存储数据
- 2、PD: 集群的管理节点。集群的元信息管理模块, 负责管理数据存储的分布情况、集群拓扑结构、进行数据调度等, 相当于“集群大脑”
- 3、TiKV Server: 集群的存储节点, 负责存储数据, 是一个分布式的存储引擎, 由一个超大的键值对Map管理存储的数据。支持多副本、高可用与故障转移。

各类节点可以部署在一台服务器上。

TiDB主要部署形式为集群, 有单集群、主备、多副本、共享等

TiDB 多用于银行、证券等金融行业, 海量数据、高并发OLTP、实时HTAP、数据汇聚与二次加工等场景

# TiDB数据库备份恢复介绍

TiDB 的事务的实现采用了 MVCC (多版本并发控制) 机制, 当新写入的数据覆盖旧的数据时, 旧的数据不会被替换掉, 而是与新写入的数据同时保留, 并以时间戳来区分版本。但旧数据不可能一直不清理, 因此它引入了Garbage Collection (GC) 机制来清理不再需要的旧数据, 其运行主要有以下特点:

- 1、GC机制按固定时间间隔来清理不再需要的旧数据, 默认配置是十分钟;
- 2、每次清理, GC会保留最近10分钟的数据, 也就是清理10分钟之前的旧数据;
- 3、清理前, TiDB会设置一个safe point (清理起点时间),  $\text{safe point} = \text{当前时间} - 10\text{分钟}$ ;
- 4、如果有长事务正在执行, 则会保证safe point不会超过该事务的执行时间;
- 5、如果这次的GC清理过程超过了10分钟, 那第二次GC清理则不开始运行。

**这会导致一个问题, 准备备份的时候, 可能要备份的数据已经被GC清理了, 没法做正常做备份**

## 解决方案:

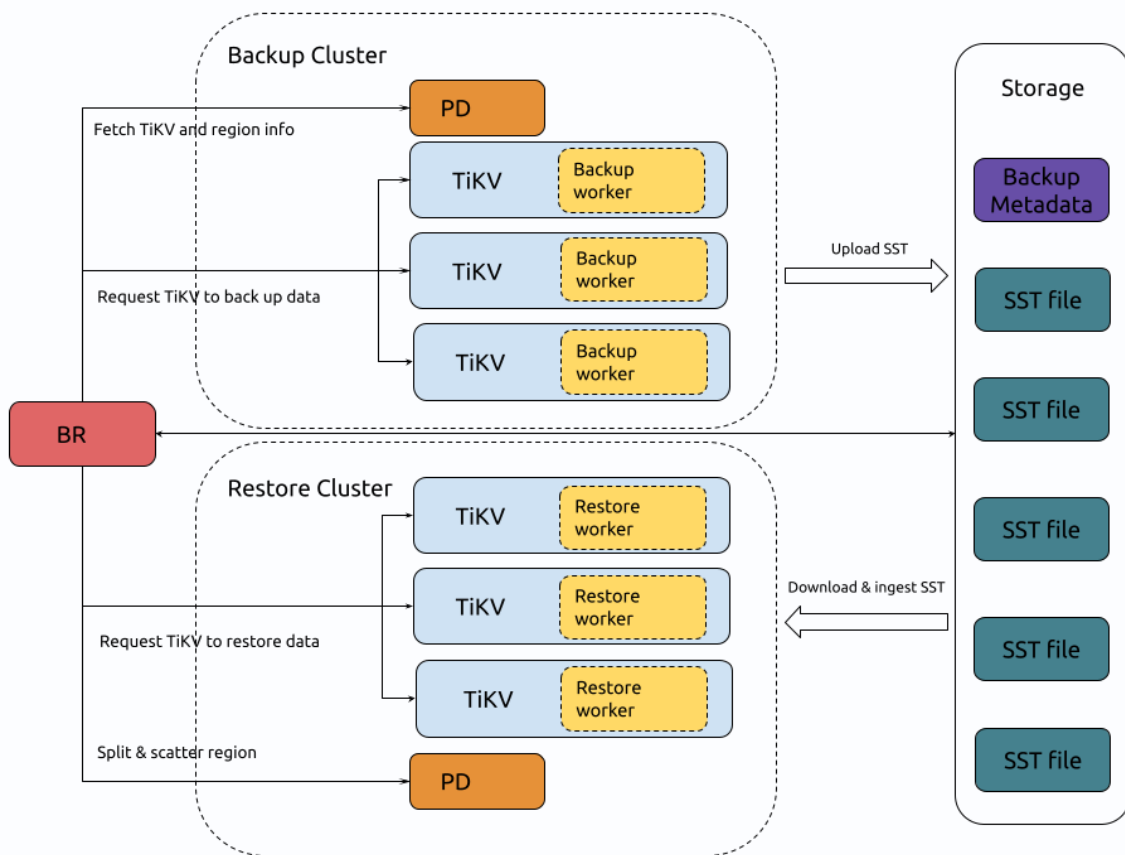
- 1、全量备份前配置 TiDB GC, 暂停其回收数据;
- 2、完成一次全量备份后, 持续做日志备份 (也叫持久化日志备份), 即使数据被回收, 通过日志也可以恢复。

# TiDB数据库备份恢复介绍

**安装部署:** 集群中的每个服务器上都要安装备份客户端, 因为需要关联集群

**备份粒度:** 集群实例级别的备份/恢复

**备份方式:** 对接TiDB的BR以及PITR进行备份/恢复



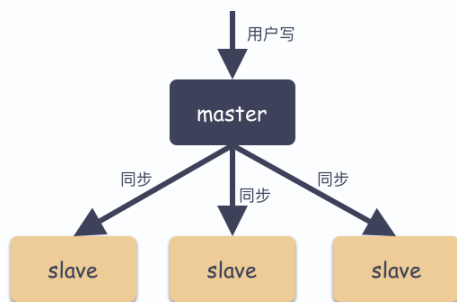
## 备份流程简介:

- 1、备份系统发送备份指令给备份客户端;
- 2、客户端向BR发起备份请求;
- 3、BR接收备份命令, 并获取数据库快照;
- 4、BR通过PD调度TiDB GC, 防止要备份的数据被 TiDB GC 机制回收;
- 5、BR访问PD, 获取TiKV的访问地址以及需要备份的数据数据分布位置;
- 6、BR发送备份请求给TiKV;
- 7、TiKV将需要备份的数据读取出来, 传输给备份存储;
- 8、BR从TiKV获取数据备份结果, 开始备份sechema等元信息, 传输到备份存储。

# MongoDB数据库介绍

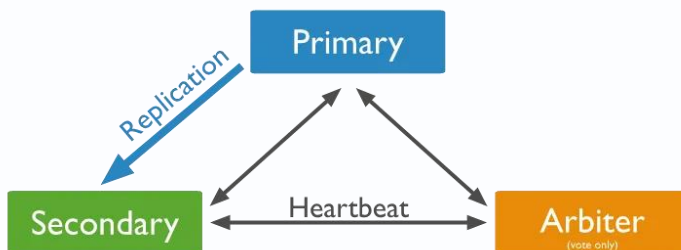
MongoDB是一款开源的文档型数据库，它以动态的模式存储类似JSON的BSON文档，能够处理大量的非结构化数据。其特点包括灵活的数据模型、高效的索引机制、支持数据复制和分片，以及易于使用的API。MongoDB适用于需要快速迭代和存储多样化数据格式的应用场景。主要以集群形态部署。

## 主从复制



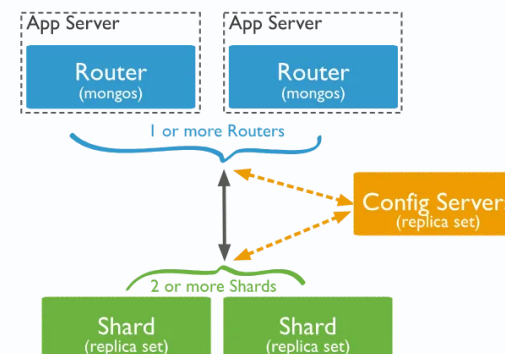
已经较少使用

## 副本集集群



使用较多

## 分片集群



海量非结构化数据处理场景

MongoDB 多用于大数据应用、社交媒体、手机APP、物联网、物流平台、直播平台等场景或行业

MongoDB 集群介绍: <https://developer.aliyun.com/article/1323982>

# PART 03

## 实操演示



# PART 04

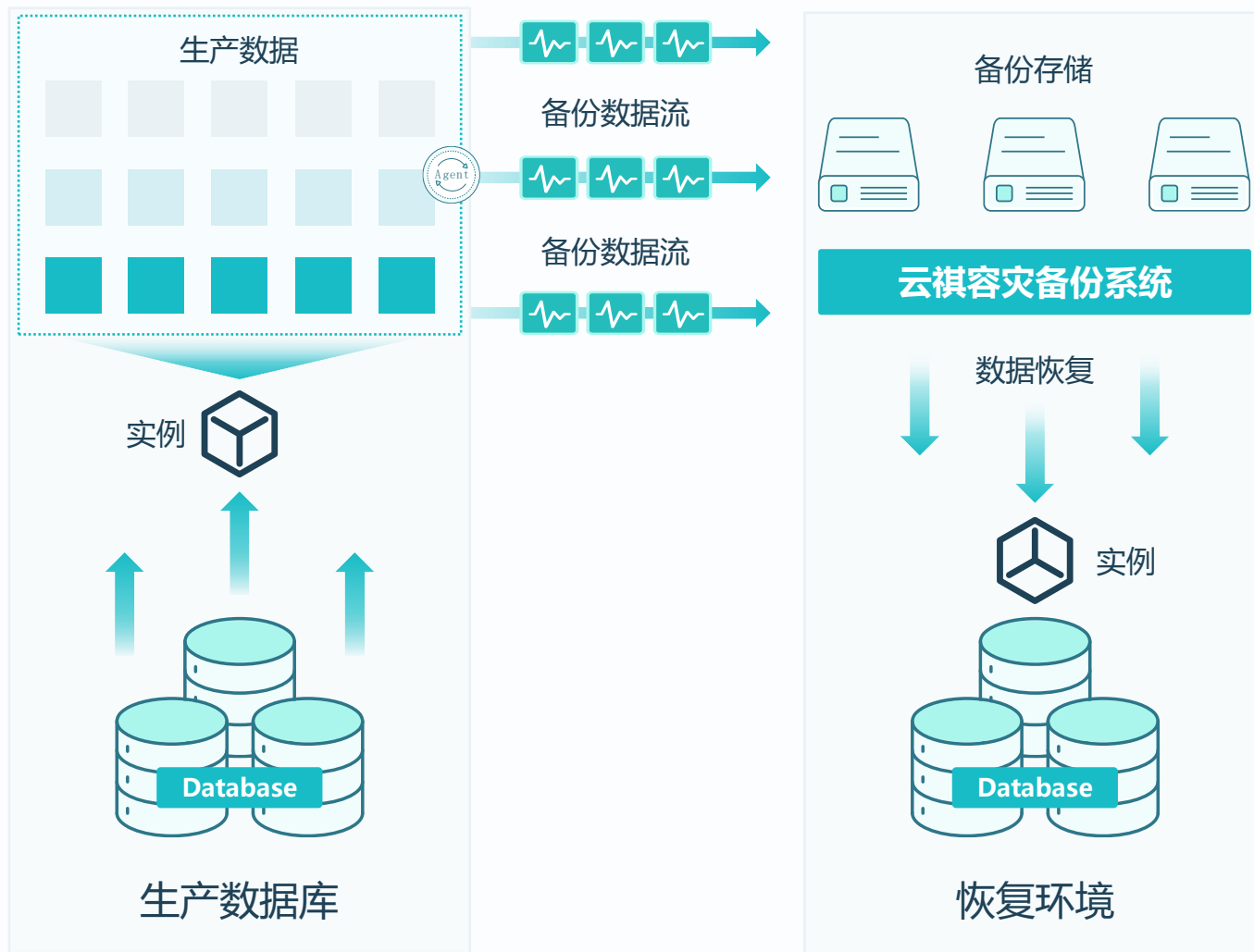
## 特性与优势总结



# 数据库保护生态



# 流式备份-节省存储空间，高效备份恢复



## 应用一致性保障

对接数据库标准接口进行数据备份，无需备份整个系统，减少需要传输的数据量并保证备份数据的强一致性

## 不占用生产存储空间

备份数据将以数据流的形式直接发送到云祺容灾备份系统，随读随传，无需在生产环境进行缓存，可有效减少对生产系统的性能压力

## 提升备份恢复效率

数据流通过Agent在云祺容灾备份系统和数据库之间自由流动，无需中转和额外的人工干预，更支持并发传输，相比传统方案效率得到大幅提升

# 可视化配置-降低操作复杂度



## 灵活易用·轻松备份

- 常规备份功能可在图形化界面进行启用/关闭，只需简单的点击或输入数值即可方便的进行设置，降低操作复杂度
- 可在界面直接对脚本进行编辑，设置自定义参数，无需登录数据库后台，更支持导入多个现成脚本，一键应用
- 支持将参数设置持久化生成到脚本模板中，新建备份可直接使用已有脚本，无需再次编辑，减少重复的参数配置操作



# 数据库自动恢复演练

**定期恢复演练：**自动恢复最新点到目标环境，用户可以手动验证数据

**自定义验证脚本：**用户可以根据自身需要配置验证脚本，实现全自动的验证

**可视化演练报告：**演练结束将生成报告供用户查看与下载

80.90演练恢复至80.185报告 b7e9885377806d0af0f79d83ec3de57b

## 80.90演练恢复至80.185报告

数据库类型: Oracle 🕒 恢复时间:2025-06-30 16:02:11

---

### 基本信息

恢复方式	备份时间点
不完全恢复	2025-06-30 15:48:09(增量备份)
源路径	恢复路径
192.168.80.90/orcl	192.168.80.185/orcl
恢复总大小	运行速度
16.31GB	84.78MB/s
传输开始时间	传输结束时间
2025-06-30 16:02:11	2025-06-30 16:05:28
运行时长	
00:03:17	

---

### 数据库状态

服务状态
正常

### SQL验证

脚本名称  
脚本\_0

脚本内容

```
-- SQL script
select instance_name,status from v$instance;
select name from v$tablespace;
```

脚本输出

```
INSTANCE_NAME  STATUS
-----
orcl           OPEN

NAME
-----
SYSaux
SYSTEM
UNDOTBS1
USERS
TEMP
SYSTEM
SYSaux
UNDOTBS1
TEMP
SYSTEM
SYSaux

NAME
-----
UNDOTBS1
TEMP
.....
```

每次验证开始前，将自动清理验证环境中上一次演练的数据

# PART 05

## 应用场景



# 本地备份与异地数据容灾



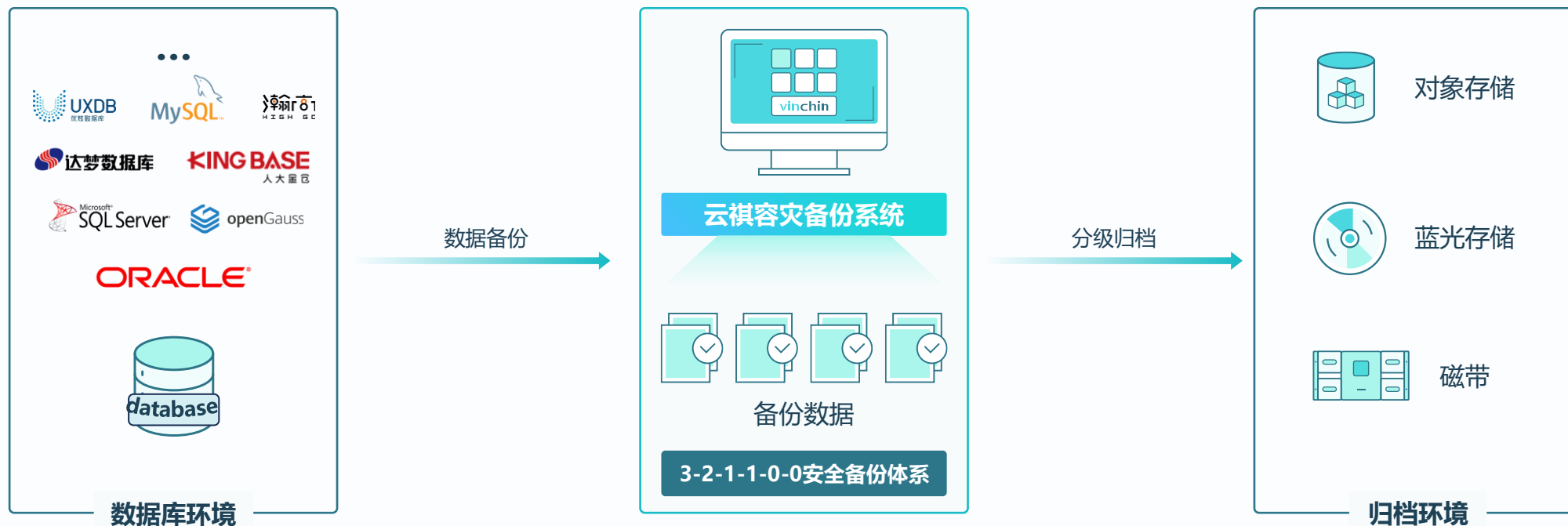
## 场景特点

- 本地数据库环境多样，需要统一进行备份，集中进行备份与灾难恢复管理
- 需要将数据在异地存放一份，以避免机房级灾难带来毁灭性数据损失以及满足相关合规要求

## 客户价值

- **统一灾备平台**：通过对接数据库标准接口实现各类数据库集中本地保护，统一数据管理、验证管理、作业编排。
- **窄带宽加速**：针对异地带宽有限、弱网连接等环境，通过源端压缩、远端增量、断网续传等技术实现传输效率和稳定性的双重提升

# 分级归档



## 多种存储支持

支持磁盘、磁带、蓝光以及AWS、华为云、Azure、腾讯云、阿里云、XSKY、深信服EDS等多种云存储和对象存储

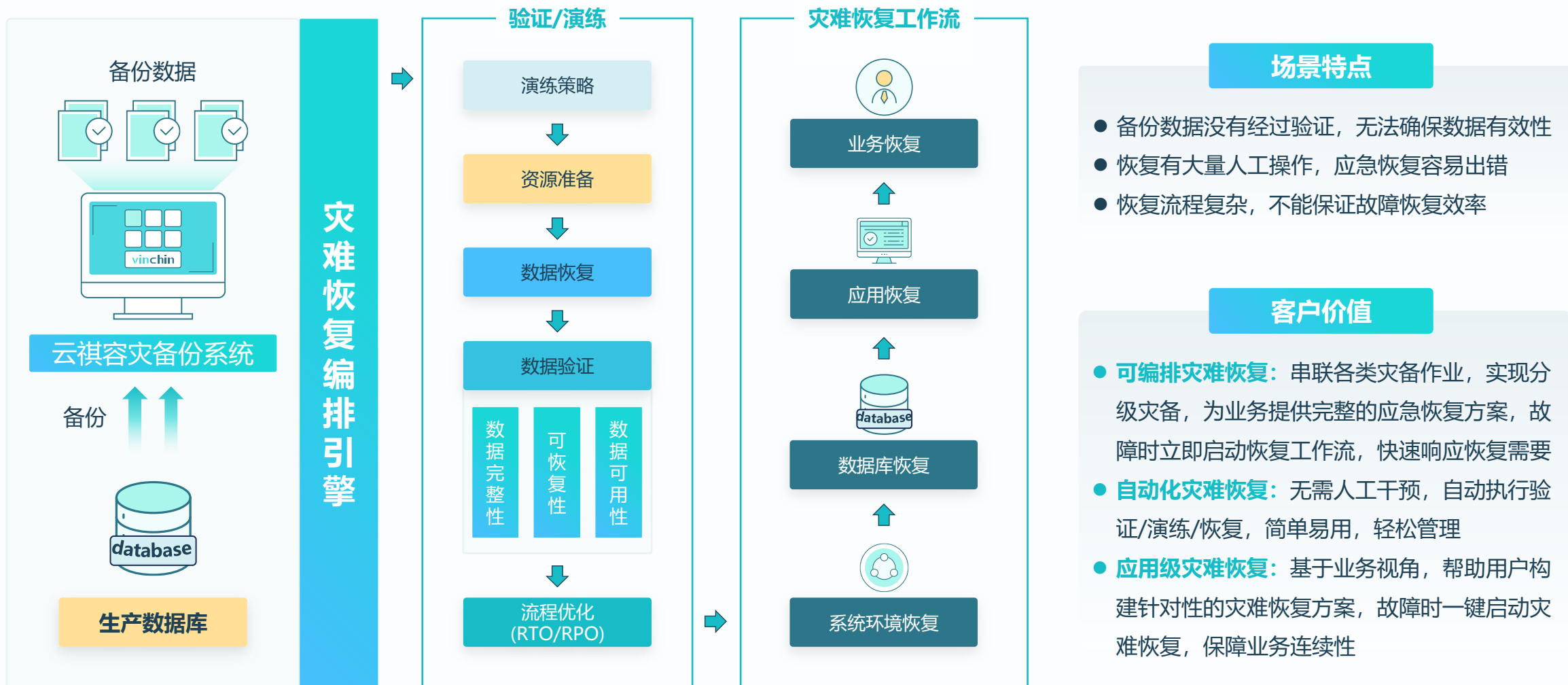
## 抵御勒索威胁

基于3-2-1-1-0-0安全备份体系中的不可变存储、强制数据保留等特性，抵御勒索病毒攻击，确保备份数据长期安全保存

## 满足合规要求

基于用户数据生命周期管理要求制定策略，实现热/温/冷分层归档，支持D2D2D、D2D2T、D2D2C等多种路径归档，满足安全合规要求

# 应用级灾难恢复



vinchin

# THANKS



云祺公众号



云祺视频号

