



数据技术嘉年华

// Data Technology Carnival

开源 · 融合 · 数智化 — 引领数据技术发展 释放数据要素价值

DM openGauss PolarDB PostgreSQL MongoDB Hive HBase Teradata
OceanBase GreenPlumCassandra MariaDB DB2 SOLite

Memcached Sybase HANA

Aurora

MySQL SQL Server Redi
OSCAR Claims X-DB IBASE Haisql-Imemcach
skysTSDB Kingwom TrendDB Cedar DragonBas
Goldilocks DThinkADB
PDW HotDB Server OushuDB Gridsum ZETA
TalDB GeminiDB TDengine Argodb
MogDB Shentong Megawise TeleDB SinoDB
GreatDB KingDB LongDB ChronusDB RadonDB
UXDB CloudTrable TSDB HUABASE HighGoDB
HashData Huayisoft
ESgynDB AnalyticDB SequoiADB ArkDB
GoldenDB AIsQL CynosDB OpenBASE QuantumDB
Base Kingbase TimesTen
MySQL SQL Server RedisTSOL H2 LevelDB Percona
Oracle RedisDynamoDB Gbase Redshift CouchDB

Oracle MySQL SQL Server Redi

OSCAR Claims X-DB IBASE Haisql-Imemcach

skysTSDB Kingwom TrendDB Cedar DragonBas

Goldilocks DThinkADB

PDW HotDB Server OushuDB Gridsum ZETA

TalDB GeminiDB TDengine Argodb

MogDB Shentong Megawise TeleDB SinoDB

GreatDB KingDB LongDB ChronusDB RadonDB

UXDB CloudTrable TSDB HUABASE HighGoDB

HashData Huayisoft

ESgynDB AnalyticDB SequoiADB ArkDB

GoldenDB AIsQL CynosDB OpenBASE QuantumDB

Base Kingbase TimesTen

MySQL SQL Server RedisTSOL H2 LevelDB Percona

Oracle RedisDynamoDB Gbase Redshift CouchDB

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase

Aurora TDSQL H2 Memcached Sybase HANA

Cassandra MariaDB Hive HBase Teradat

PostgreSQL MongoDB DB2 SOLite

LevelDB Percona TBase Kingba

SinoDB DynamoDB Gbase Redshift Couch

GreenPlum DM openGauss Polar

TDDB Neely Informix OceanBase



中国DBA联盟
All China DBA Union



墨天轮

Memcached Sybase HANA
DM openGauss PolarDB PostgreSQL MongoDB DB2 SOLite
OceanBase GreenPlumCassandra MariaDB Hive



数据技术嘉年华

走好信创最后一公里

演讲人：韩锋



中国DBA联盟
All China DBA Union



墨天轮

DM openGauss PolarDB PostgreSQL MongoDB DB2 SOLite
OceanBase GreenPlumCassandra MariaDB Hive HBase Teradata

Memcached Sybase HANA

Aurora

MySQL SQL Server RedisTDSOL H2 LevelDB Percona
Oracle RedisDynamoDB Gbase Redshift CouchDB
TBase Kingbase TimeTen
GoldenDB AlisQL CynosDB OpenBase QuantumDB
ESgynDB AnalyticDB SequoiaDB ArkDB
UXDB CloudTrable TSDb HUABASE HighGoDB
HashData Huayisoft
GreatDB KingDB LongDB ChronusDB RadonDB
MogDB Shentong Megawise TeleDB SinoDB
Pala

Oracle MySQL SQL Server Redi
OSCAR Claims X-DB iBASE Haisql-Jmemcach
SkyTSDB Kingwew TrendDB Cedar DragonBase
Goldilocks DThinkADB
PDW HotDB Server OushuDB Gridsum ZETA
TaIDB GeminiIDB TDengine ArgonDB

PostgreSQL MongoDB DB2 SOLite

Aurora

MySQL SQL Server RedisTDSOL H2 LevelDB Percona

Oracle RedisDynamoDB Gbase Redshift CouchDB

TBase Kingbase TimeTen

GoldenDB AlisQL CynosDB OpenBase QuantumDB

ESgynDB AnalyticDB SequoiaDB ArkDB

UXDB CloudTrable TSDb HUABASE HighGoDB

HashData Huayisoft

GreatDB KingDB LongDB ChronusDB RadonDB

MogDB Shentong Megawise TeleDB SinoDB

Pala

PostgreSQL MongoDB DB2 SOLite

Aurora

MySQL SQL Server RedisTDSOL H2 LevelDB Percona

Oracle RedisDynamoDB Gbase Redshift CouchDB

TBase Kingbase TimeTen

GoldenDB AlisQL CynosDB OpenBase QuantumDB

ESgynDB AnalyticDB SequoiaDB ArkDB

UXDB CloudTrable TSDb HUABASE HighGoDB

HashData Huayisoft

GreatDB KingDB LongDB ChronusDB RadonDB

MogDB Shentong Megawise TeleDB SinoDB

Pala

PostgreSQL MongoDB DB2 SOLite

Aurora

MySQL SQL Server RedisTDSOL H2 LevelDB Percona

Oracle RedisDynamoDB Gbase Redshift CouchDB

TBase Kingbase TimeTen

GoldenDB AlisQL CynosDB OpenBase QuantumDB

ESgynDB AnalyticDB SequoiaDB ArkDB

UXDB CloudTrable TSDb HUABASE HighGoDB

HashData Huayisoft

GreatDB KingDB LongDB ChronusDB RadonDB

MogDB Shentong Megawise TeleDB SinoDB

Pala

PostgreSQL MongoDB DB2 SOLite

Aurora

MySQL SQL Server RedisTDSOL H2 LevelDB Percona

Oracle RedisDynamoDB Gbase Redshift CouchDB

TBase Kingbase TimeTen

GoldenDB AlisQL CynosDB OpenBase QuantumDB

ESgynDB AnalyticDB SequoiaDB ArkDB

UXDB CloudTrable TSDb HUABASE HighGoDB

HashData Huayisoft

GreatDB KingDB LongDB ChronusDB RadonDB

MogDB Shentong Megawise TeleDB SinoDB

Pala

PostgreSQL MongoDB DB2 SOLite

Aurora

MySQL SQL Server RedisTDSOL H2 LevelDB Percona

Oracle RedisDynamoDB Gbase Redshift CouchDB

TBase Kingbase TimeTen

GoldenDB AlisQL CynosDB OpenBase QuantumDB

ESgynDB AnalyticDB SequoiaDB ArkDB

UXDB CloudTrable TSDb HUABASE HighGoDB

HashData Huayisoft

GreatDB KingDB LongDB ChronusDB RadonDB

MogDB Shentong Megawise TeleDB SinoDB

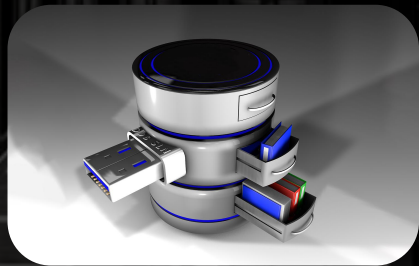
Pala

PostgreSQL MongoDB DB2 SOLite

Aurora

概述

随着近些年内外部形式剧变及企业自身发展诉求，国内企业愈发重视基础软件的自主可控。特别是对于某些涉及国计民生的重点行业，监管层面也提出了非常明确的指导意见。作为核心软件之一，数据库在其中扮演重要角色，且具有高复杂性。一是数据库自身复杂度就比较高；二是数据库技术发展迅猛，以分布式、多模、HTAP为代表新型数据库架构不断涌现。同时国内数据库发展活跃、厂商产品能力层次不齐，用户在选型、研发、迁移、使用上面临诸多痛点。特别是在整体改造的最后阶段，涉及将系统从原有技术栈迁移到新技术栈，这其中蕴含了较多工作及风险。本文尝试从信创改造角度出发，重点谈在改造中往往处于最后改造的数据库部分，即所谓信创改造“最后一公里”所面临的痛点问题及可能解决思路。





韩锋

CCIA（中国计算机协会）常务理事，前Oracle ACE，腾讯TVP，阿里云MVP，dbaplus等多家社群创始人或专家团成员。有着丰富的一线数据库架构、软件研发、产品设计、团队管理经验。曾担任多家公司首席DBA、数据库架构师等职。在云、电商、金融、互联网等行业均有涉猎，精通多种关系型数据库，对NoSQL及大数据相关技术也有涉足，实践经验丰富。曾著有数据库相关著作《SQL优化最佳实践》、《数据库高效优化》。

信创改造阶段

在信创改造过程，大致将其划分为四个阶段。

- 选型评估阶段

完成信创技术栈的选型，并开展相关技术栈培训，做好基础储备工作；同步完成资源、成本评估。

- 研发测试阶段

完成业务系统针对信创技术栈的改造及测试，其中涉及到较大的成本（人力、时间）的投入。

- 验证并行阶段

完成业务系统改造后，需针对新平台的功能、稳定性、可用性等方面进行验证。为保证替换平稳，推荐使用并行方式进行。

- 上线支持阶段

在系统已经得到充分验证后，将业务系统从原技术栈完全迁移到新技术栈。此阶段需重点解决迁移及出现问题的保障维护方面。

目录

CONTENTS

01

选型评估阶段

02

研发测试阶段

03

并行验证阶段

04

上线支持阶段



阶段：选型评估

● 迁移评估

对迁移工作做整体规划，制定原则，明确迁移范围、方式、是否停机、窗口期等。除技术外，其他如组织、管理、资源等，也在这一阶段一并考虑。迁移是个很复杂的过程，涉及的方方面面很多，尽量在项目之初就有个全面的掌握。



● 业务梳理

在迁移准备阶段，就对涉及的业务有个全面的梳理非常有必要。这里需要梳理的信息，非常宽泛。包括但不限于对业务系统涉及的软硬件环境、与数据库交互、业务系统间调用关系等。后续在做应用系统改造规划中，上述信息非常重要，其有助于评估工作难点、工作量等。

使用者	xx 部门/xx 系统
交互方式	<input type="checkbox"/> 人机交互 <input type="checkbox"/> 应用交互 <input type="checkbox"/> 数据交互
使用特征	<input type="checkbox"/> PC <input type="checkbox"/> APP <input type="checkbox"/> 其他
使用时长	使用时长: ____时____分 ~ ____时____分 高峰时段: ____时____分 ~ ____时____分 在线用户数: ____ 并发用户数: ____ 并发量大小: ____
重要级别	<input type="checkbox"/> A+ <input type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C
可用时长	<input type="checkbox"/> 7x24 <input type="checkbox"/> 5x8 <input type="checkbox"/> 5x12 <input type="checkbox"/> ____
关联业务	业务 1: 业务名称, 关联简述 业务 2: 业务名称, 关联简述 业务 3: 业务名称, 关联简述 ...
系统来源	<input type="checkbox"/> 自建 <input type="checkbox"/> 外采 <input type="checkbox"/> 合作 <input type="checkbox"/> 云 <input type="checkbox"/> 其他
基础技术栈	开发语言: <input type="checkbox"/> Java <input type="checkbox"/> C++ <input type="checkbox"/> Go <input type="checkbox"/> Python <input type="checkbox"/> 其他 <input type="checkbox"/> 混合
语言版本	
应用拓展	
交互方式	连接方式: <input type="checkbox"/> IP <input type="checkbox"/> VIP <input type="checkbox"/> DNS <input type="checkbox"/> OTHER 读写分离: <input type="checkbox"/> 否 <input type="checkbox"/> 是 芯片: _____
国产化需求	操作系统: _____ 中间件: _____ 其他组件: _____

数据引擎	# 平台/语言/中间件 单机数据引擎: _____ 集群数据引擎: _____ # 主从/分布式/云 半体/分片数据引擎: _____ 数据副本数: _____ 热数据引擎: _____
数据引擎	_____/月
数据特性	冷热分层: _____ 数据压缩: _____ 数据转储: _____ 数据热点: _____
计算引擎	<input type="checkbox"/> OLTP <input type="checkbox"/> OLAP <input type="checkbox"/> HTAP <input type="checkbox"/> ____ 说明: _____
计算指标	QPS: _____ TPS: _____ RT: _____ R/W: _____
日志引擎	GB/日
日志峰值	GB/分钟
数据同步	数据同步: _____ 一致性问题: <input type="checkbox"/> 强一致 <input type="checkbox"/> 弱一致 说明: _____
扩展性需求	接入扩展: <input type="checkbox"/> 水平 <input type="checkbox"/> 无 计算扩展: <input type="checkbox"/> 垂直 <input type="checkbox"/> 水平 <input type="checkbox"/> 混合 <input type="checkbox"/> 无 存储扩展: <input type="checkbox"/> 垂直 <input type="checkbox"/> 水平 <input type="checkbox"/> 混合 <input type="checkbox"/> 无
高可用需求	<input type="checkbox"/> 冷备 <input type="checkbox"/> 热备 <input type="checkbox"/> 双活 <input type="checkbox"/> 多活 <input type="checkbox"/> 无
高可用策略	策略: RPO=____ RTO=____ 策略: RPO=____ RTO=____
高可用方案	
数据库特性	<input type="checkbox"/> 特殊对象: _____ <input type="checkbox"/> 库内计算(如: 存储过程、函数、触发器等) <input type="checkbox"/> 其他特性: _____



阶段：选型评估

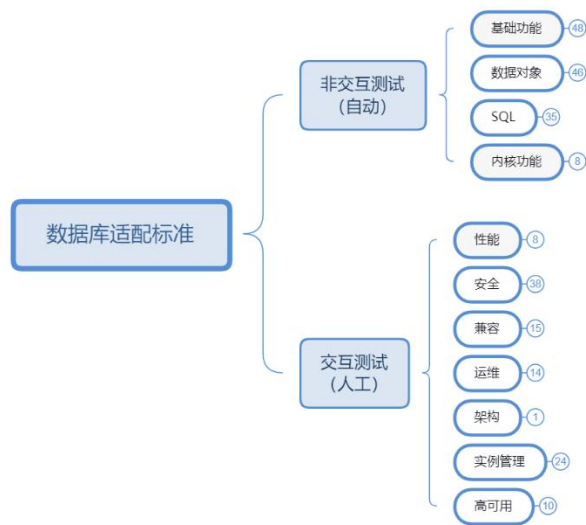
● 场景评估

收敛公司内对数据库场景需求，制定一个产品选型矩阵。根据要求对待选产品进行评估工作，包括初始调研、技术评估、数据库评测（功能、非功能、业务等）、适配性评估等。此项原则是在方案选型中保持自由度，不绑定厂商，随时可替换。

需求分类	适用场景					
	业务场景	数据规模	事务一致性	应用适配能力	典型负载	分析能力
事务类	联机交易 轻量数据分析	中	单库强一致 多库应用解决	高	高并发小数据量 事务性读写 / 小 数据量分析	弱
	联机交易 简单事务 大开发	大	应用层最终一致 性	低应用 逻辑有入侵	点查 / 点写 / 有 限的关联和分析 支持	弱
	联机交易 批量处理 实时分析 混合负载	大	强一致	高 对应用透明	点查 / 点写 / 有 限的关联和分析 支持	弱
分析类	批量处理 复杂分析 非实时查询	大	弱一致	低，应用逻辑入 侵	复杂分析查询	强
事务/分析混 合类	联机交易 批量处理 实时分析 混合负载 轻量级复杂分 析	大	强一致	高 对应用透明	高并发小数据量 事务性读写 / 复 杂分析	强

● 选型测试

对待选数据库产品进行全方位的评估，包括从功能、非功能、业务三个角度进行测试。需要建立企业自有的测试样板集。



阶段：选型评估

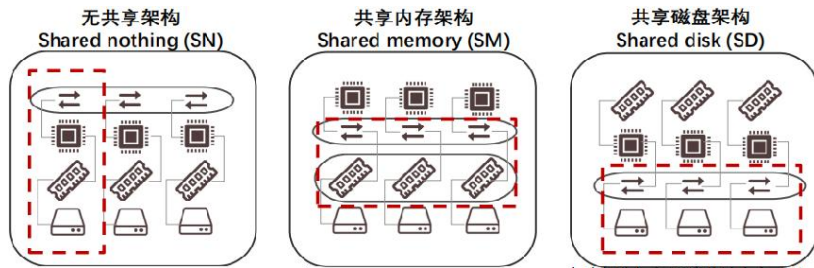
● 技术培训

对于新引入技术栈进行培训，包含对架构、研发、运维人员的培训。对研发人员，阐述架构设计、结构优化、SQL语句等与之前差异。对运维人员，侧重将这种新数据库融入到现有的运维体系中。特别是分布式架构数据库，与传统集中式数据库不同，其对于运维带来的挑战也更大。



● 方案制定

在明确选型后，需完成相关技术方案的制定。这里需结合业务发展、资源投入、基础环境、高可用等诸多因素进行考虑。此阶段输出成果，将直接影响后续的资源评估、改造及测试等。



阶段：选型评估

● 资源评估

资源评估通常是前置的，因为很多企业是采用预算制，需要提前很长周期做好后续的采购规划。因此将资源评估工作做了前置。上述资源，特指硬件及基础设施部分。

接入层

计算层

存储层

● 研发评估

包含对数据结构、语句及应用自身的改造评估。在结构上各数据库能力层次不齐，需进行调整甚至重构。建议减少复杂对象使用，只考虑表索引；其余视图、序列、触发器、存储过程、函数、包等都通过外部的等价设计来完成。目的是减少对数据库依赖。此外，分布式下还需考虑分片策略等。对 SQL 语句的评估，建议从复杂度、方言等进行评估。

对象使用情况		
对象类别	数量	
表	15492	
表(大表)	47	
表(分区表)	0	
字段(大对象)	16	
索引(B树)	6303	
索引(其他)	23	
视图	5	
触发器	4	
存储过程		
函数		
包		
SQL类别		
总SQL数	174	
超长SQL	64	
ANTI SQL	5	
Oracle Syntax SQL	52	
Join 3+ Table SQL	28	
SubQuery SQL	40	

阶段：选型评估

痛点

● 信创技术栈分散

尚未形成选型标准，选型困难。在生态兼容性上，有兼容MySQL、PG、Oracle、自有标准等多种形式。架构上包括单机、集中式、分布式等多种，包括以 NewSQL 为代表的产品受到关注。在部署平台上，包括私有部署及云端部署（含私有云带底座、公有云）等多种形式。

● 产品能力层次不齐

产品能力层次不齐，不同产品间功能差异明显，包括内核层面、周边生态层面及管理维护层面等多方面。无法面对统一的服务界面会很困难。此外，在产品部署形态上，云数据库产品成为很多企业的选择。但在云产品选择上，用户的自主权较差，存在与原有方式的管理差异。

数据库标准访问层

- 统一接入协议
- 统一语法规范

数据标准管理层

- 统一管理职能（如安全、审计）

数据标准基础层

- 统一资源配给（如云）

目录

CONTENTS

01

选型评估阶段

02

研发测试阶段

03

并行验证阶段

04

上线支持阶段



阶段：研发测试

● 工作排期

在研发开始前，需完成必要的排期工作。这也是很多信创项目的痛点之一，很难做到合理评估。建议采用“打样”方式，摸索内部研发能力，给出符合企业实际情况的排期。

工作量评估依据

1. 梳理分析 (M列)

关联系数

* 特殊SQL语句使用数量(E列)：特殊SQL改造量较大，建议按高成本核算。

* 使用特殊数据类型表的数量(F列)：使用特殊数据类型表改造成本较高，建议按高成本核算。

* 特殊对象数量(I列)：特殊对象是指目标库不兼容部分，需要进行对象级别改造适配，按对象数评估。对于存储过程、函数、触发器等按行数处理。

* 系统负载-DML次数(J列)：DML次数间接反映系统负载，影响整体改造成本，按100万、1000万、10000万分段来评估。

* SQL总数(K列)：SQL总数反映整理改造工作量。

* 选型(L列)：如目标库成本高，给出比例系数（对比MySQL）。

计算公式

工作量 = (特殊SQL语句使用数量 * 系数 + 使用特殊数据类型表的数量 * 系数 + 特殊对象数量 * 系数 + SQL总数 * 系数) * (1 + 系统负载系数) * 选型系数

2. 代码修改 (N列)

关联系数

* 平台相关(M列)：项目组需配合平台测试等带来的工作量。参照[梳理分析]部分得到的工作量，乘以此系数。

计算公式

工作量 = 梳理分析 * 系数

3. 单元测试 (O列)

关联系数

* 平台相关(M列)：项目组需配合改造测试等带来的工作量。参照[梳理分析]部分得到的工作量，乘以此系数。

计算公式

工作量 = 梳理分析 * 系数

4. 迁移工作量 (P列)

关联系数

* 数据量(H列)：整体迁移工作量。乘以此系数。

* 选型(L列)：目标库迁移成本高些，乘以此系数。

计算公式

工作量 = 数据量 * 数据量系数 * 选型系数

● 研发改造

对数据对象进行改造，有些简单映射即可，有些则需重构。有些对象（如复杂对象或重计算类对象），可从数据库端剥离，在上层等价实现。对语句修改，这项工作非常浩大。如用 ORM 工具还好，如有硬 code，改动较繁琐。有工具可辅助完成改造工作，但转换后仍需人工审核，要保证语义等价和执行效率等等。

SQL 开
发

SQL 查
询

SQL 审
核

诊断优化

SQL 转
换

对象转换



中国DBA联盟
All China DBA Union



墨天轮

阶段：研发测试

- 应用改造

配合迁移工作，应用也存在改造的工作量。例如有些在数据库端实现的逻辑，是需要改在应用端实现的；有些应用本身在新数据库架构下就需要进行适配性改造等等。

- 功能测试

数据库改造（对象+SQL）和应用改造均完成后，还需要一个关键步骤—功能测试。按业务功能拆分，针对每个独立功能，从应用侧角度进行测试验证，来保证上述改造的结果正确且性能符合预期。此处的功能测试，与后面的上线交割的功能测试不同，更强调是以小的应用单元为测试目标。这样便于随时修改，随时测试。

阶段：研发测试

痛点

● 系统评估难

在实际工作中，经常会面临一类问题就是旧有系统已无人了解或干脆是由第三方开发的。如在存量业务的改造中，缺乏有效的手段去收集、进而很难评估改造任务工作量。

● 改造成本高

原系统，对旧技术栈依赖严重，之前多采用 Oracle、DB2 为代表商业数据库，应用端依赖数据库方言、库内计算、生态工具等。新技术栈功能不足，需应用改造故存在较大工作。策略上，部分需改造适应新架构，部分需在异构平台（如大数据、缓存）或应用层解决。

● 评估结果难

对新产品兼容性存疑，仅通过语法兼容或少量改造，难保证语义一致，造成上线风险。缺乏有效的评测手段，针对语义等价（数据一致性）及性能等能做到评估。

系统评估工具

- 语句捕获
- 负载抓取、回访

辅助开发平台

- 类型语句转换

兼容增强平台

- 语法增强、改写

结果评估工具

- 语法语义评估
- 运行性能评估

目录

CONTENTS

01

选型评估阶段

02

研发测试阶段

03

并行验证阶段

04

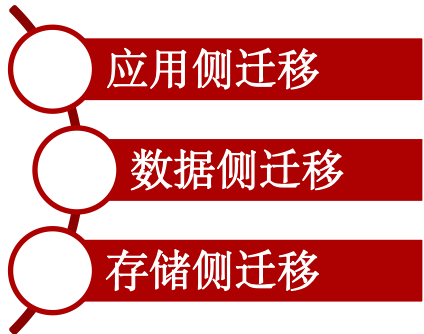
上线支持阶段



阶段：验证并行

● 迁移方案

确定迁移方案，包括数据迁移与应用迁移。针对前者主要取决于源数据库、物理环境、迁移窗口、是否并行、是否回退等诸多因素。从原理上分，可分为应用侧迁移、数据库侧迁移、存储侧迁移三种，各有优劣。个人建议，优先应用侧，原因是与业务结合紧密、同步验证容易、方便并行回退等等。但缺点在于，需要应用侧有修改工作，无法形成统一标准的方案。



● 数据迁移

结构迁移。一般可提前完成。结构确定后，即可完成。不与数据同步放在一起，为了便于问题排查分析。

全量迁移。一般是可离线、静态去做，通过备份恢复、导入导出等方式，将静态数据迁移过去。可不在停机窗口进行，并记录好位点。

增量迁移。从全量迁移位点开始，可采取停机或不停机的方式。一般通过追日志方式，追齐数据，并短时静默应用，完成数据最终达到一致的状态。

数据类型不兼容

在线迁移应用无感

迁移加工逻辑复杂

数据对比质量可控

阶段：验证并行

痛点

● 风险高、回退难

为了在验证阶段，验证系统是否工作正常，一般需要开发大量验证类的代码。这部分工作为了满足系统支持新旧技术栈及必要的对比等工作，但往往工作量巨大。如很多应用常见的数据双发逻辑，就是通过数据双写，同步验证两边执行结果。为达到这一诉求，不得不在原有业务逻辑上开发两套适应不同技术栈的代码。

● 验证难，无从下手

在验证阶段还有一个比较难的地方在于如何验证，最好的验证方式是带着真实流量的验证，但同时还需考虑风险问题。如果对业务访问做好精准的控制，按需求进行业务验证，且还需提供必要的退化能力保证安全。如常用的基于读写的分配、基于流量的分发（甚至基于业务特征的分发能力）。

数据双写平台

- 异构双写、数据强一致
- 异常退化、应用逃生
- 语法转换、减低成本

流量分发平台

- 比例、读写、白名单分发
- 拓扑感知、流量重分发
- 灰度切换、实时调整

目录

CONTENTS

01

选型评估阶段

02

研发测试阶段

03

并行验证阶段

04

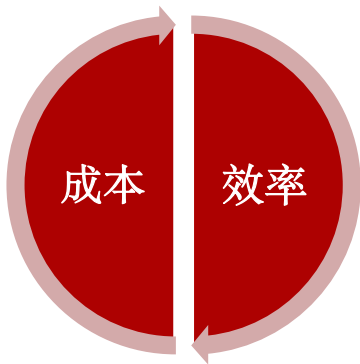
上线支持阶段



阶段：上线支持

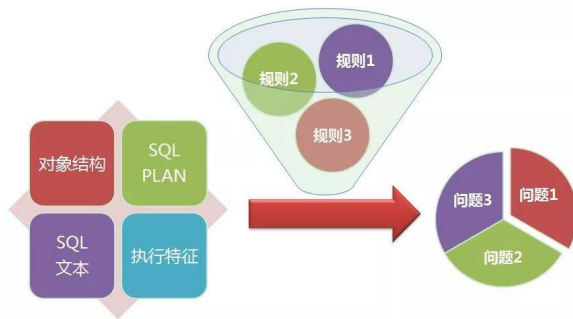
● 数据校验

在上线前需完成对数据迁移后的质量判断，这就引入数据校验的初衷。是对同步两边的数据是否一致做出判断，为是否正式切换的判断依据之一。难点在于：海量数据对比效率、异构数据比对及实时同步下同步下对比。



● 语句审核

在上线交割及日常保障中，还需支持 SQL 审核，是为了保证语句运行质量。SQL 审核细分，可分为事前、事中、事后审核，这里更多指事前审核部分。即在开发过程中，针对 SQL 运行情况给予评估判断，来保证上线后的质量可控。一般是通过预定义一组规则，完成对语句的审核。这一过程贯穿在整个开发过程中。



阶段：上线支持

痛点

● 窗口短、迁移难

在系统上线阶段，一个突出问题是迁移窗口期的问题，其普遍的上线窗口期很短。这就需要在较短的时间内能够完成数据库间（一般是异构）的数据的迁移工作，同时还需针对迁移后的数据提供质量对比，能够保证迁移数据是正确的。

● 不稳定、无逃生

系统上线后的稳定性问题，用户最为关心。作为新产品、新架构，很难保证上线后不出问题。虽可通过充测试、并行验证等手段尽量减少风险，但显然无法完全避免。较好方式是提供能力，根据可能出现的运行问题，通过一些手段可以尽量减少问题影响范围，恢复业务。

离在线迁移工具

- 离在线数据迁移
- 高吞吐、异构间
- 数据对比、人工补偿

流量治理平台

- 流量控制（标签、文本、用户名等）
- 熔断、限流、审计、追踪

系统逃生平台

- 异构逃生、数据一致
- 自动退化、人工恢复



走好信创“最后一公里”

- 前瞻布局
- 生态兼容
- 自主创新
- 广泛合作
- 风险评估
- 稳步推进

THANKS FOR WATCHING

[illegible]