

# 大型数据中心集中化建设案例分享



上海新炬网络技术有限公司 [WWW.SH5NC.CN](http://WWW.SH5NC.CN)

SHANGHAI NEW CENTURY NETWORK CO.LTD

SACC2012

**SACC**  
2012中国系统架构师大会  
SYSTEM ARCHITECT CONFERENCE CHINA 2012  
架构设计·自动化运维·云计算



# 自我介绍



杨志洪:

- 10年ORACLE服务经验
- 服务过的行业：电信、联通、移动、银行、保险、证券、基金、制造、政府
- 服务于国内最大的数据库第三方运维服务厂商**上海新炬**



联系方式:

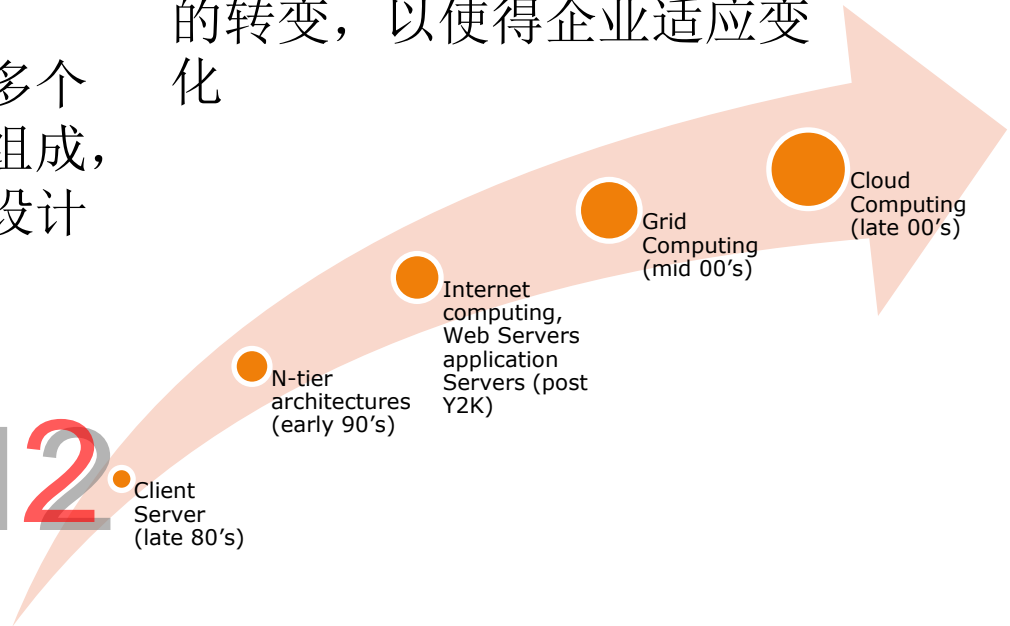
- [yangzhihong@shsnc.com](mailto:yangzhihong@shsnc.com)
- 18621566151
- <http://zhihong.org>
- <http://weibo.com/boypoo>



- 架构Architecture -
- “The fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution.” – ANSI/IEEE Std 1471-2000
- 架构是系统的基本结构，它由多个组件以及它们彼此间的关系而组成，并且在一定环境和原则下进行设计和演变。

- 过去30年间，IT技术的演进驱动着架构转型
- 企业不断通过技术创新来保持业务上对竞争对手的领先地位
- IT转型聚焦在新举措、新技术的采纳，推动IT技术组合的转变，以使得企业适应变化

SACC2012



数据库在整个应用架构中处在最核心部分

记录和提供客户行为和服务轨迹

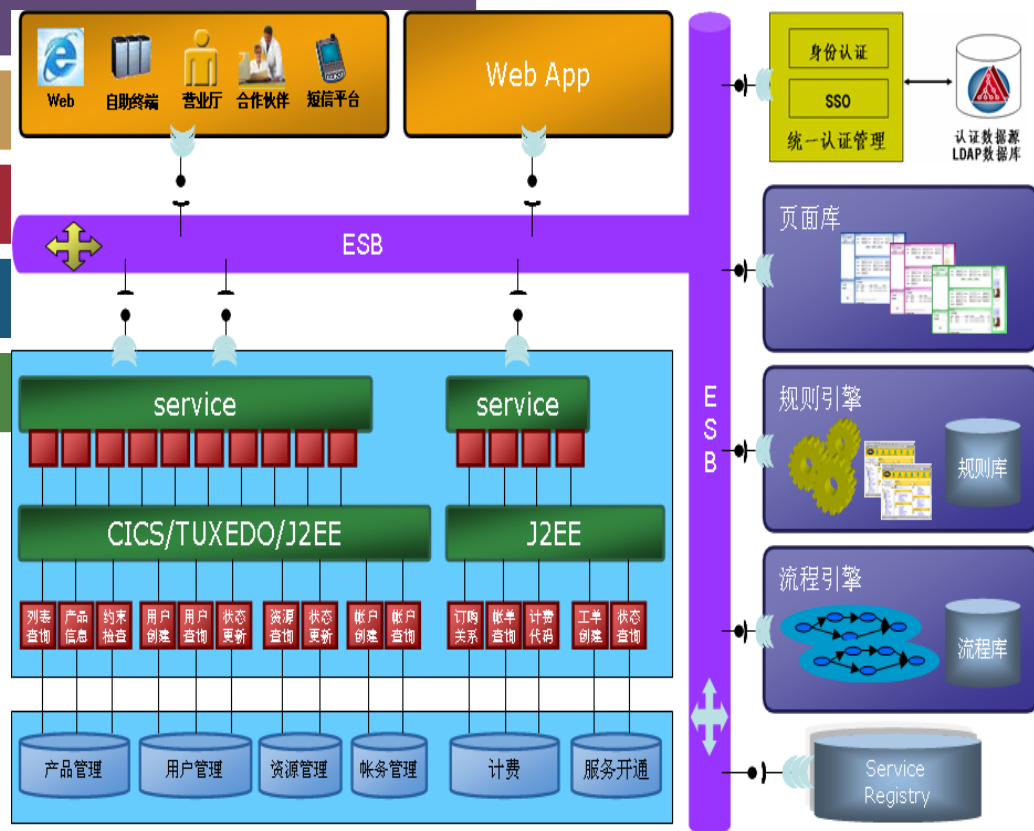
数据库担负一定的业务逻辑计算能力

安全、高效存储数据

高效展现业务数据的能力

提供信息安全保护并实现合规性

具备高可用确保业务连续性



系统集中的程度越高，系统故障影响业务的范围越大；数据损坏将导致大量的业务数据丢失





# 灾难恢复能力等级划分



国家标准《信息系统灾难恢复规范》（GB/T 20988-2007）

- 第1级 基本支持
- 第2级 备用场地支持
- 第3级 电子传输和部分设备支持
- 第4级 电子传输及完整设备支持
- 第5级 实时数据传输及完整设备支持
- 第6级 数据零丢失和远程集群支持

灾难恢复能力等级	RTO	RPO
1	2 天以上	1 天至 7 天
2	24 小时以后	1 天至 7 天
3	12 小时以上	数小时至 1 天
4	数小时至 2 天	数小时至 1 天
5	数分钟至 2 天	0 至 30 分钟
6	数分钟	0





- 风险分析

各种风险发生的概率及风险发生后对业务的影响程度

- 业务关键等级划分

关键业务/非关键业务

各项业务的容灾指标（RPO/RT0）

- 容灾策略

同城异址容灾/异地容灾

容灾层次：系统级、数据级、应用级和业务级

容灾范围：关键业务应急、全业务容灾

运营方式：主备中心、双中心、两地三中心

容灾规模：同级容灾、降级容灾

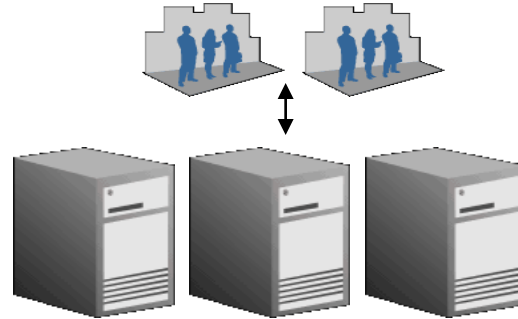
- 实施管理与演进路线

- 建设和使用成本

SACC2012

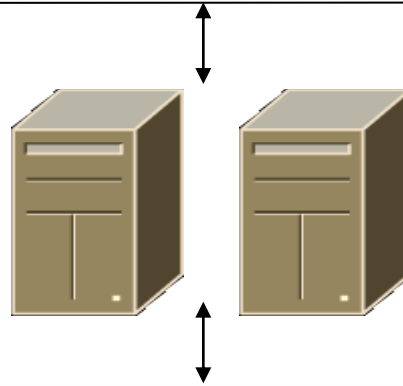
## 应用层

- 数据：业务应用程序
- 应用程序故障切换
- 多个数据库间保持一致



## 数据库层

- 数据：表、事务、模式
- 数据库故障切换
- 多个存储阵列间保持一致



## 存储层

- 数据：位、字节
- 存储阵列接管
- 存储阵列内保持一致





选择合适的技术达到最佳平衡的解决方案:

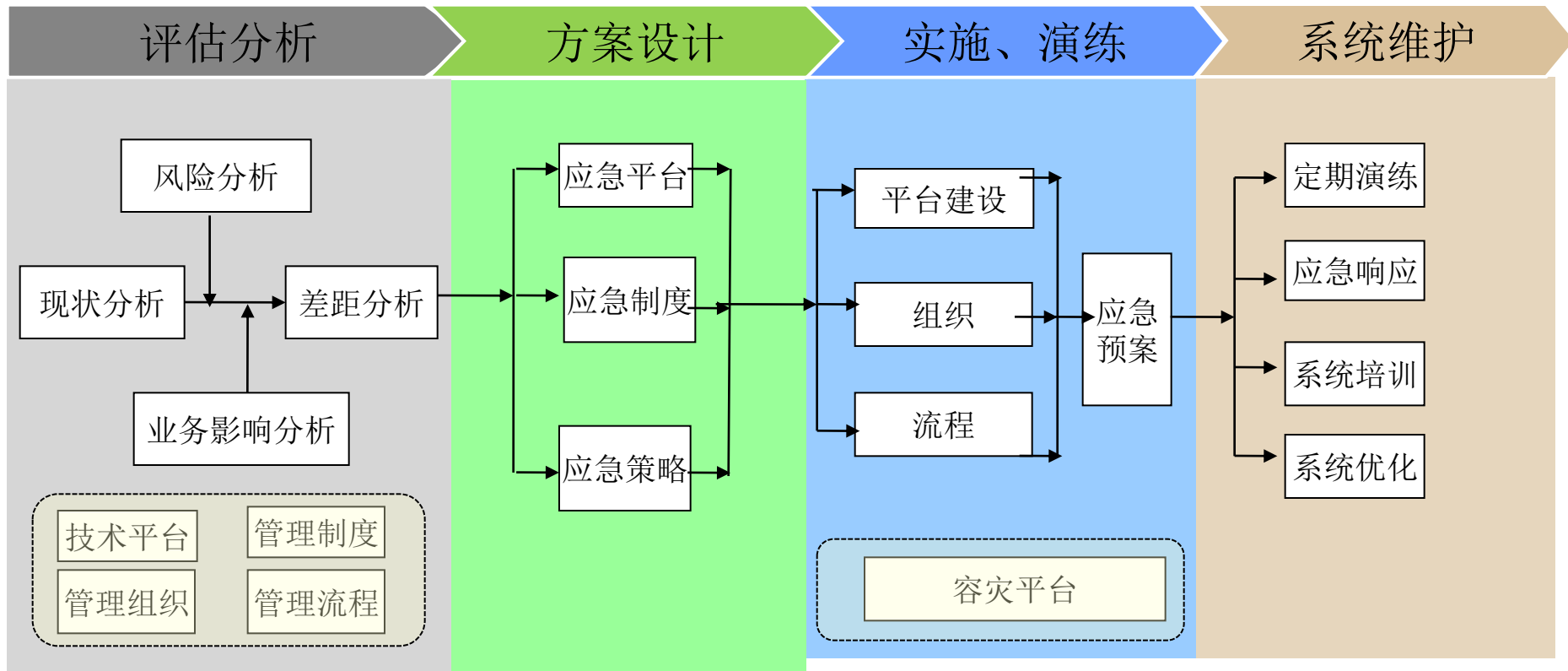
- 防止各种组件故障: 服务器故障、存储故障、站点故障、人为错误等
- 最大限度地缩短应用程序停机时间: RTO
- 最大限度地减少数据丢失: RPO
- 事务一致性
- 业务切换复杂程度
- 数据回录难易度
- 利用全部资源, 减少投资
- 降低运行时成本: 监控、管理、自动化程度、灾难恢复管理
- 是否支持多种技术混合

SACC2012



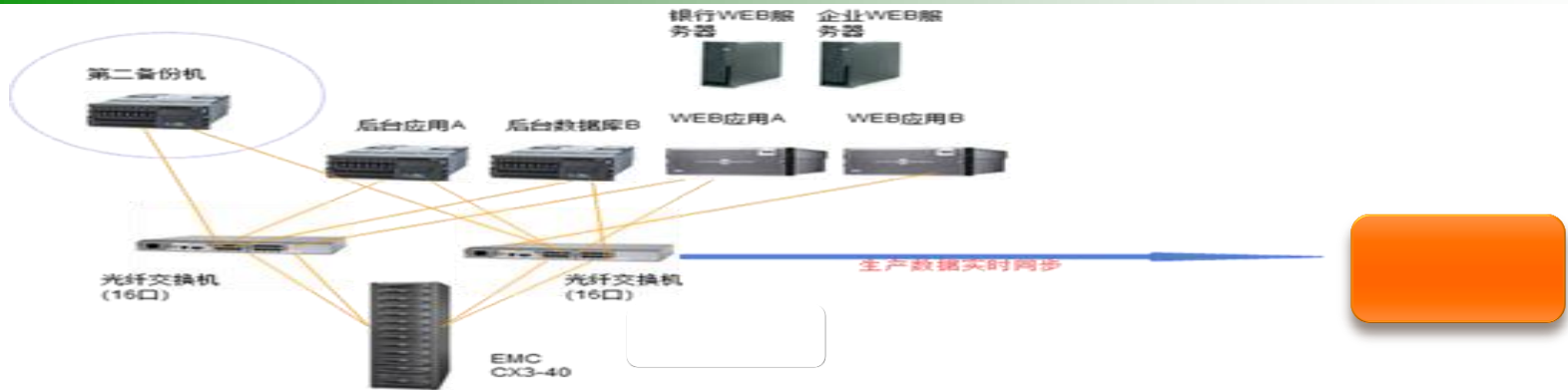
# 建设数据中心遵循的方法论：

“整体规划、分步实施、风险管理、定期演练”



标准化、流程化、文档化的项目管理

# 案例1:XX结算应急数据中心项目需求



同城系统拓扑图

▶根据《银行业信息系统灾难恢复管理规范》，灾备系统最低要求应该是第五个等级

▶根据结算中心的数据中心建设规划，分步实施远程异地的整体灾备系统，至少达到第5级的业务恢复能力，具备抵抗区域大灾难的能力。

存在  
问题

- ▶基于存储复制的数据级容灾，没有应用级容灾。
- ▶灾难恢复能力等级最多能达3级，不能满足金融系统的灾备要求。
- ▶容易发生因数据同步失败而影响灾备系统的可靠性

项目  
目标

实现最低第5级的业务恢复能力 (RTO≤15分钟和RPO≤10分钟)

分两阶段实施数据级、应用级同城容灾系统

制定对应的灾难恢复计划

第一步

实现同城异址的数据级别的灾备

第二步

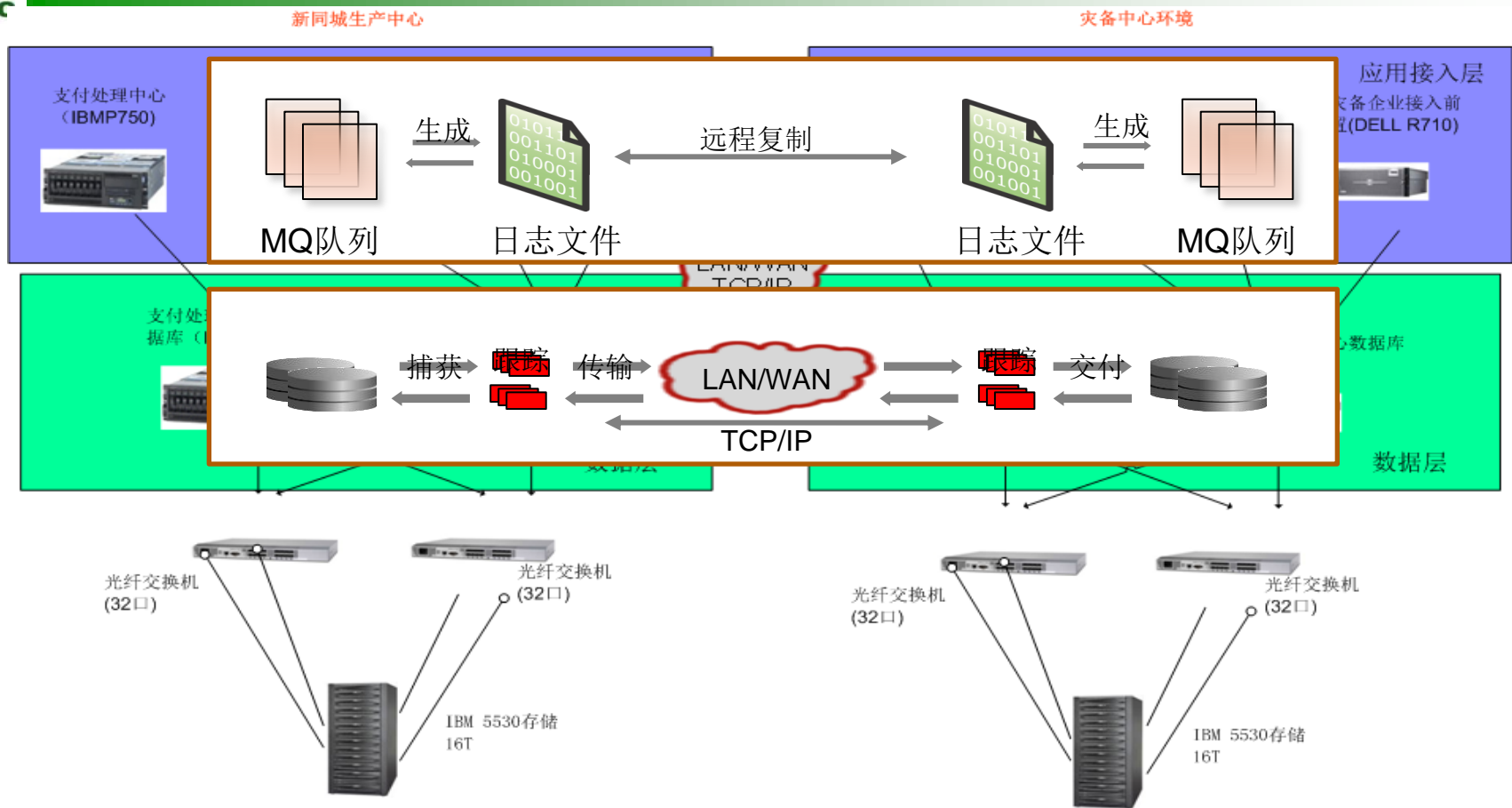
实现同城异址的应用级别的灾备

SACC2012

# 案例1:选择合适技术满足系统的数据同步要求

	模式	优势	局限性	适用场景
存储镜像	冷中心	独立于主机，不占用主机资源，异步模式对应用影响小； 模式简单	存储平台不独立，多中心存储设备硬件平台相同 单一的复制拓扑，不支持多对一模式 同步模式：性能影响大，网络要求高 异步模式：存在数据一致性问题 混合模式：需要较多的存储投资	容灾
逻辑卷镜像	冷中心	存储平台无关，源端和远端可以使用不同的存储设备和RAID保护级别 可以在IP网络上复制，网络要求不高 异步模式下应用响应时间不受影响，但RPO会被延长	长时间网络故障需要巨大日志文件空间 占用主机资源，IO等待较高 同步模式：性能影响大 异步模式：存在数据一致性问题	容灾
逻辑数据复制	热中心	日志解析，对主机（数据库服务器）负载较少 准同步模式，RPO短 主机无关，存储无关，数据库平台相关性弱 元数据对象独立于数据库,可以支持异构数据库 可以细化到具体的数据对象（Schema），可以选择性地进行复制 复制拓扑结构灵活 网络要求较低	只支持数据库同步，不支持其他类型数据同步 较多的数据层维护工作量 数据库支持有一定限制要求	应急/容灾
数据库复制	热中心	实现方式简单 应用透明，支持数据库所有特性 网络传输效率高 故障隔离，防止数据块损坏	同步模式：对应用性能有影响 异步模式：数据一致性问题 只支持数据库同步复制	容灾

# 案例1:XX结算应急数据中心项目解决方案



数据库层的灾备方案由Oracle GoldenGate产品实现生产中心与灾备中心数据实时双向的数据同步

应用层面, 由MQ日志文件远程复制技术实施生产中心和灾备中心应用数据同步

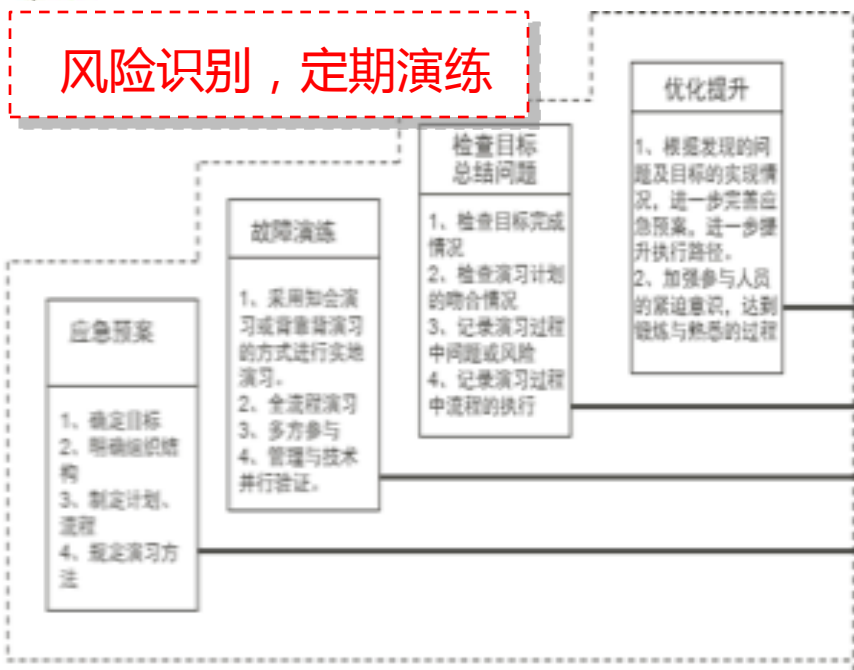
SACC2012



# 案例1:集中化应急系统恢复计划

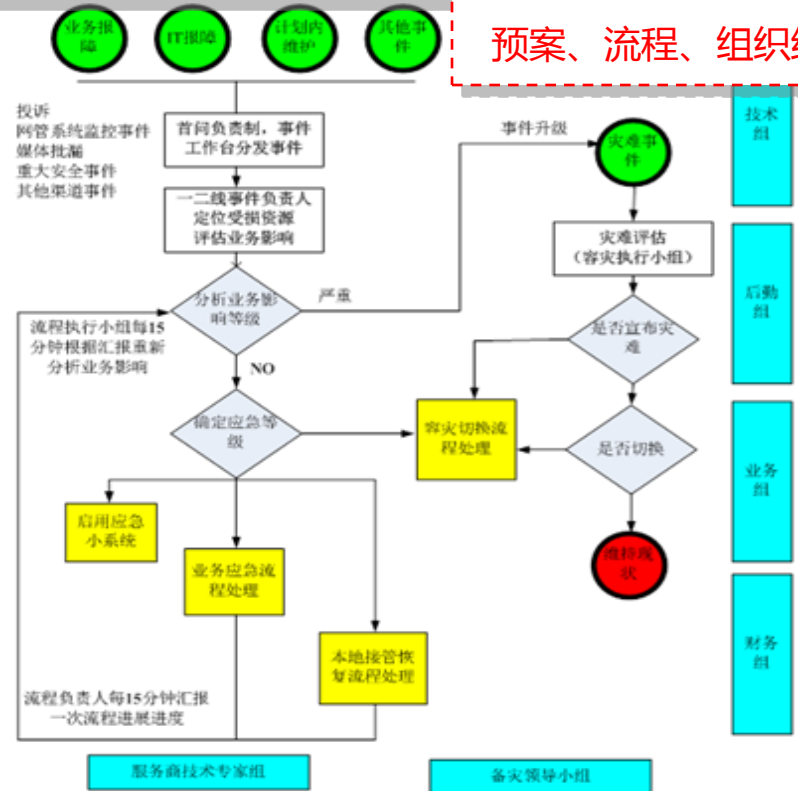


## 风险识别，定期演练



- 完善的应急预案、流程、组织结构是确保灾备系统发挥作用的必要手段。
- 帮助结算中心建立包括数据备份日常管理制度、备份数据保存制度、灾备切换流程、灾备系统变更管理流程、定期演练流程等管理流程。

## 预案、流程、组织结构



系统的建设不能只考虑到容灾系统建设，很显然灾难发生的机率是非常小的。如何建立一套机制来有效地防范日常的计划和计划外中断并合理利用灾备中心的资源，是整体灾备系统建设中的重要问题。



## 准实时复制

峰值交易吞吐量超过5000笔/秒，但业务高峰期生产系统数据可在10秒内复制到应急库。以最低的代价获得接近“两阶段提交”才具有的同步复制效果，RPO<10秒。

## 双活模式

应急数据库处于激活状态，可随时接管业务，数据级实现RTO=0，应用级采用一键切换方式，RTO<5分钟。

## 自动回录

应急期间在应急数据库产生的业务数据，应急结束后，通过GoldenGate反向复制功能自动回录到原生产端，减少系统还原工作复杂性及业务影响。

## 负载分离

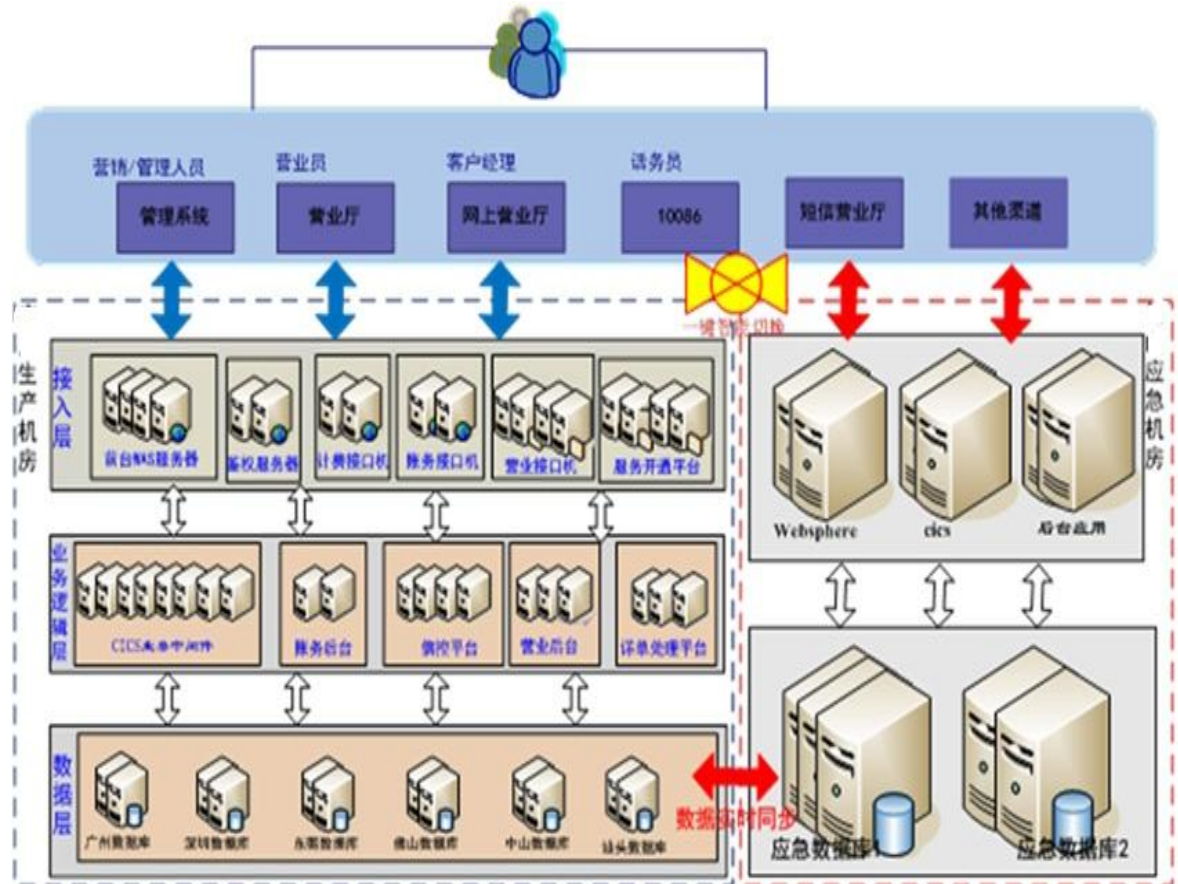
在支持关键业务应急功能的基础上，批量数据查询、统计等非关键业务转移至应急数据库，有效减轻生产端负载压力,目前已分担生产系统20%的查询压力

# SACC2012

# 案例2:XX移动核心业务应急数据中心系统建设

应急数据中心特点:

- 负载分离
- 双活数据中心
- 远距离低带宽
- 准实时数据同步
- 数据自动回录







## 案例2:建设成果

- 应急系统自2010年投入使用以来，在系统割接保障、故障应急方面发挥了重要作用，共启用应急系统8次，应急切换成功率100%。接管业务时累积长达共55小时。
- 渠道管理平台每个月固定时间在CRM系统抽取数据，然后分析，生成佣金数据。每当渠道平台系统在抽取数据时，生产系统数据库的负荷会增加30%左右，数据抽取需要时间非常长，导致2个系统之间产生了不良的影响。

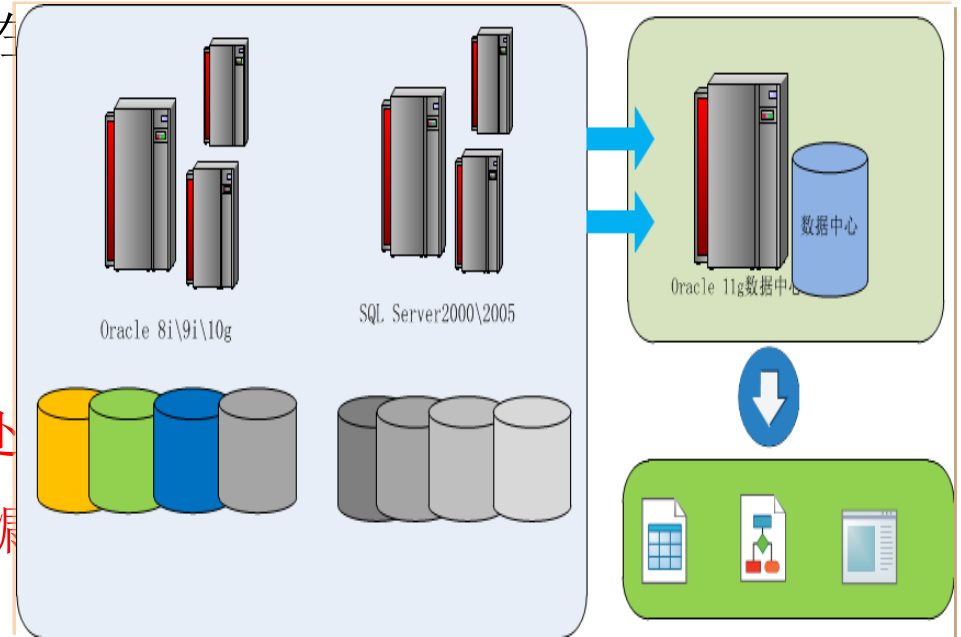
迁移应急系统前		迁移到应急系统后	
占用生产系统资源	执行时间	占用生产系统资源	执行时间
28%	30小时	0%	5小时

SACC2012

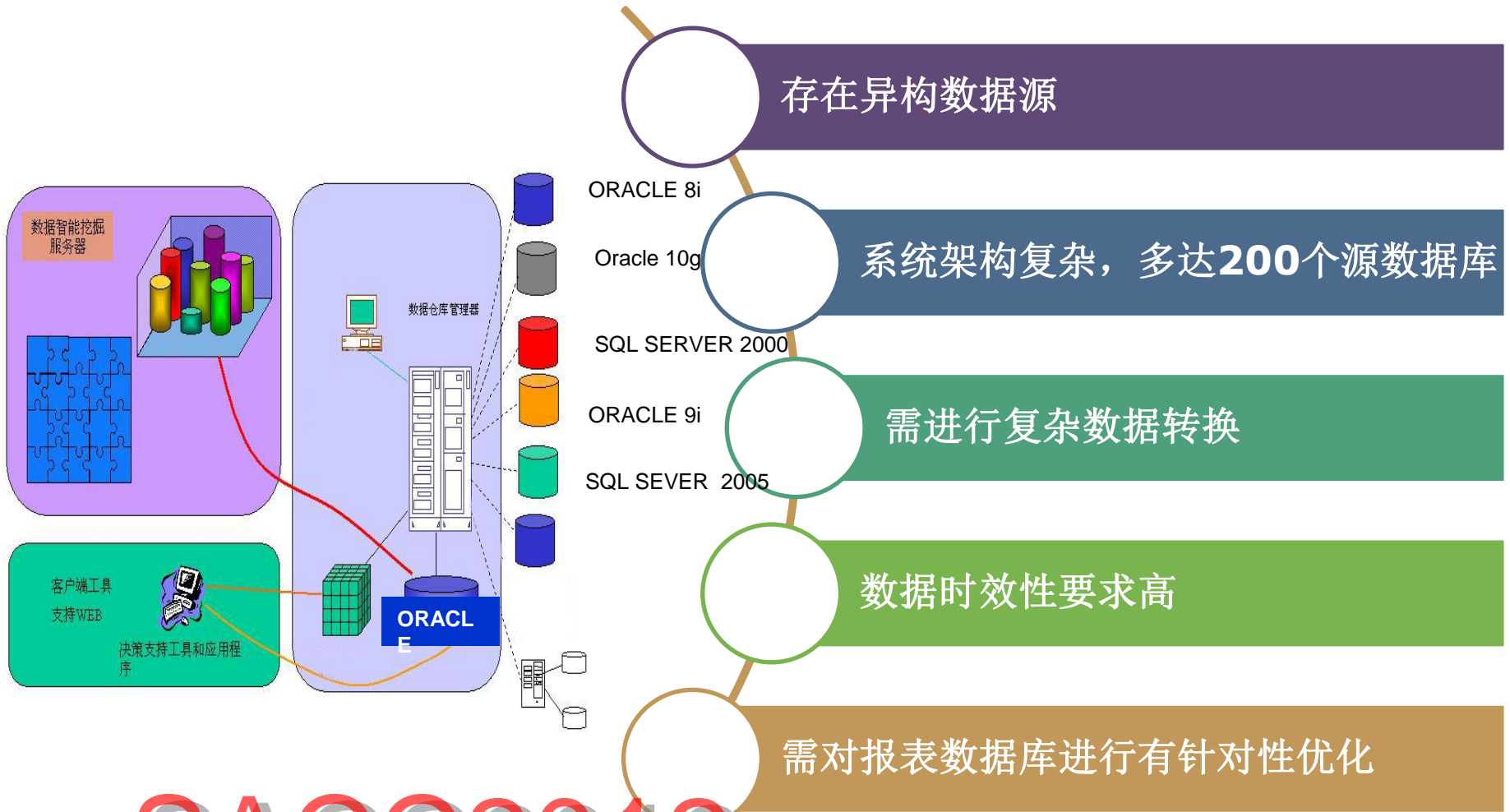
“故障能应急，平时能分流”是业务连续性保障建设的发展方向

# 案例3:XX公司数据集中化项目需求

- 两百多套内部流程系统，分散在不同版本的SQL SERVER(2000/2005)、Oracle(8i/9i/10g)数据库
- 遭遇困境1：同一个自然人需处理多个系统上的审批事项，遗漏问题严重
- 遭遇困境2：采用Oracle Portal工具统一展现各系统上的待办事项，数据库负载过高



# 案例3:数据集中化项目实施过程中面临诸多挑战



# SACC2012

以GoldenGate软件从各系统对应的数据库中采集数据，将信息同步到一个新的Oracle11g数据库里，Oracle portal工具直接从这个库获取信息

根据业务逻辑来设计源库以及目标库中需采集的表以及对应的目标表，并进行信息同步

借助Oracle eSSO工具实现单点登录功能，无需重复登录原有系统

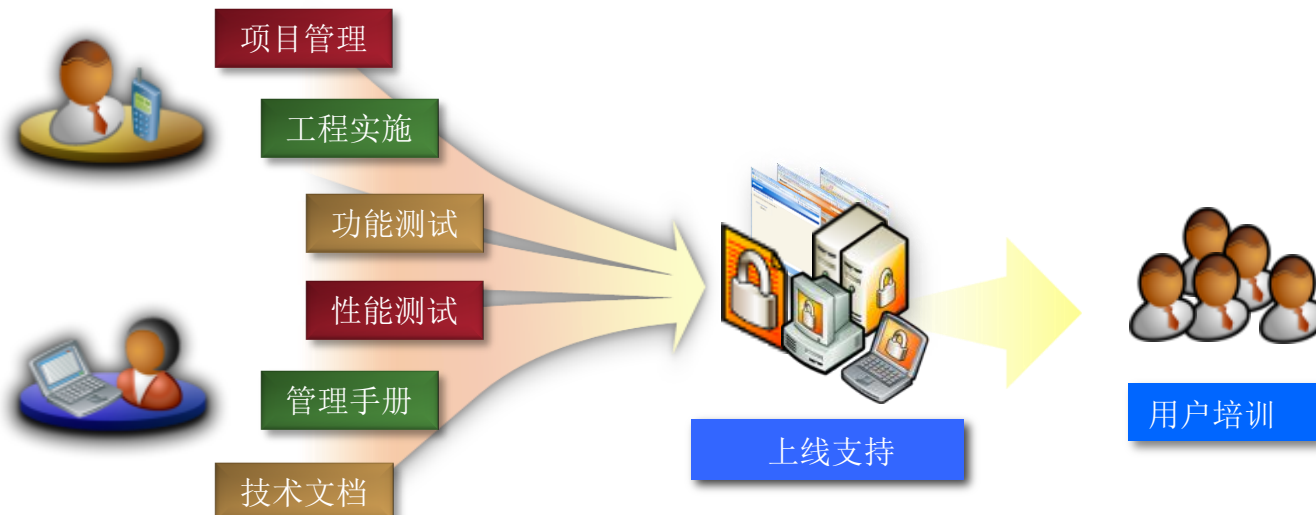
## 案例3:异构同步原理——同构同步原理

- 以GoldenGate抽取及投递进程从源库的日志文件中捕获变化信息，以队列文件的方式传输到目标端
- 在不影响源库运行的前提下，以rman、datapump、OGG Direct Load等方式将源库需采集的数据及结构克隆一份到目标库中
- 以GoldenGate复制进程将目标端队列文件翻译成SQL语句，并执行，以实现同步

SACC2012

- 捕获源库变化信息并传输到目标库的原理完全一样
- 源库数据及结构的克隆则比同构同步复杂，**OGG Direct Load**优势较大
- 异构目标库**GoldenGate**复制进程翻译**SQL**语句时，无法将目标库表结构当成源库表结构使用，相关信息需借助**defgen**工具提供
- 确保**匹配关系明确**且**主键信息齐全**

- 项目实施过程中，解决了涉及产品**BUG**、环境配置、业务逻辑、数据安全、应用程序等**5**方面的**10**多项技术难题，有效保证了项目的顺利完成。



# SACC2012

## 问题描述

- 项目过程中发现Oracle 8i版本的数据库不支持GoldenGate所必需的Supplemental logging功能，Oracle专家建议升级8i至更高版本的数据库。但应用环境的实际现状无法进行版本升级，严重影响项目进度。

## 我们的解决方案

- Oracle 8i确实没有Supplemental logging，但GoldenGate本身有建立触发器以记录主键字段的功能，我们建议将表主键字段也放到目标表，并将其作为MAP关系的主键，并提供异构表匹配关系总体解决方案，克服了因产品功能不足所带来的难题。



## 问题描述

- XX电信数据集市系统使用GoldenGate技术将CRM、SPS等各业务系统的数据集中到数据集市系统进行市场经营活动分析，自上线后2个月开始，发现存在数据同步丢失现象，业务数据无法准确同步到数据集市系统，严重影响数据分析的准确性。

## 我们的解决方案

- 多方位的故障分析，定位到数据丢失的根源是部分业务程序模块使用了GoldenGate不支持的事务模式，并提出解决方案。

# SACC2012

# 谢谢！

