

IBM 2010 *System z* 高峰论坛
暨新品发布

宏_聚天下 · 智_绘高远

IBM®



Infrastructure Architecture for Transactional, Operational & Analytical Data Process

David Heap

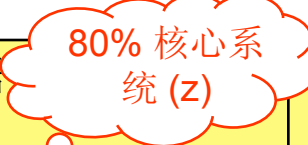
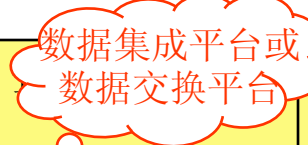

IBM大中华区系统与科技事业部高级IT咨询师

客天下 · 智高远

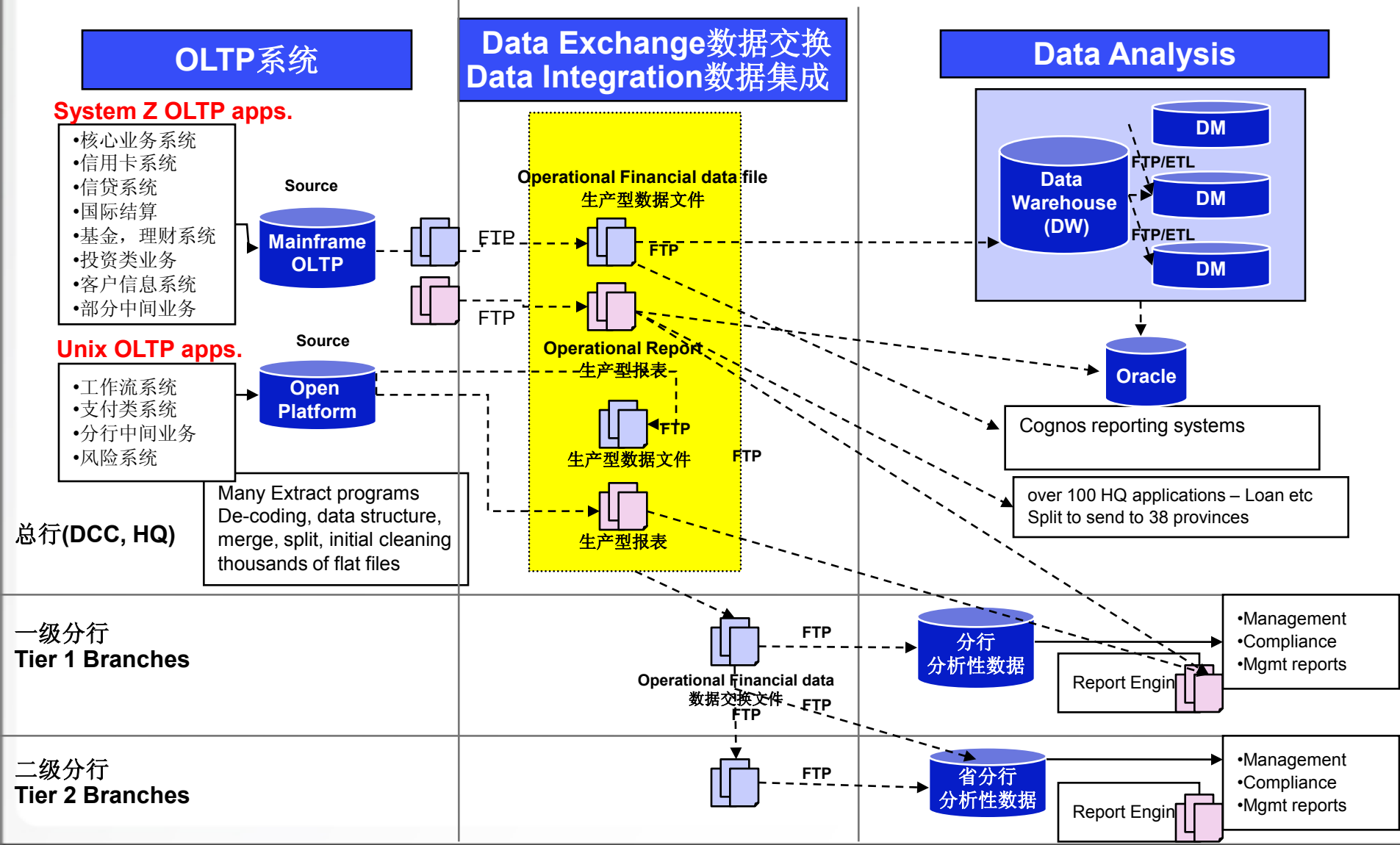
主题

- 当前大银行数据的基础架构全景及面临的挑战
- 大型银行数据集成基础架构设计指导原则
- 最优化的交易型、操作型、分析型数据处理流程基础架构
- 业务及技术价值
- 建议

交易型、操作型、分析型数据的特性比较

	交易型数据 	操作型性报表、 	分析性报表、 
使用对象	•生产一线人员和客户	•一线的生产和管理人员	•专业分析人员和企业中高层管理人员
系统功能	<ul style="list-style-type: none"> •记录交易流水 •记录帐务，客户，产品等业务关键信息的实时变更 •提供最实时的业务事实 	<ul style="list-style-type: none"> •数据整合，为生产系统提供数据共享； •客户数据展现：如：查询、报表（主要面向生产）、动态分析（主要面向生产分析）等； •数据质量审计和监控管理。 	<ul style="list-style-type: none"> •主要支持多个维度的主题分析； •趋势分析与问题发现； •分析型客户行为分析；风险；盈利性分析；资产负债管理
数据来源	•各个服务渠道业务的实时变更	•数据来源是生产系统；	<ul style="list-style-type: none"> •可以来自操作型平台， •操作型平台没有提供操作数据直接从源系统中提取。
数据架构	<ul style="list-style-type: none"> •按照交易的业务流程及功能要求组织数据 •提供高效的随机访问的能力，以随机修改为要 	<ul style="list-style-type: none"> •按生产视角主题域方式来组织数据； •主要提供细粒度运营数据，也存在部分粗粒度汇总数据 •提供操作数据的报表、查询。 	<ul style="list-style-type: none"> •按分析视角主题域方式来组织数据； •主要提供粗粒度汇总数据，且分析维度相对O更多、更复杂 •提供分析数据的展示和分析
数据存储时间	<ul style="list-style-type: none"> •每日大量的动态更新 •交易日志保持一日 	<ul style="list-style-type: none"> •汇总数据保存1年； •公用数据长久保存； •详细数据保存1—3个月； •其它整合层数据保存13月。 	<ul style="list-style-type: none"> •汇总数据保存5年； •公用数据长久保存； •详细数据至少在线保存2年； •其他整合层数据至少保存2年。
技术特征	<ul style="list-style-type: none"> •面向高效随机更改的数据库技术 •高效的OLTP数据类型和数据操作 	<ul style="list-style-type: none"> •数据库优化同时侧重索引和分区技术； •支持OLTP类型和OLAP类型的数据操作 	<ul style="list-style-type: none"> •数据库优化主要侧重分区技术； •支持OLAP类型的数据操作。

每日在交易型、操作型、分析型数据平台之间进行着大量的数据转换
每天300-400GB交易数据 (>1TB特殊日)，成千的操作型文件，上万的报表



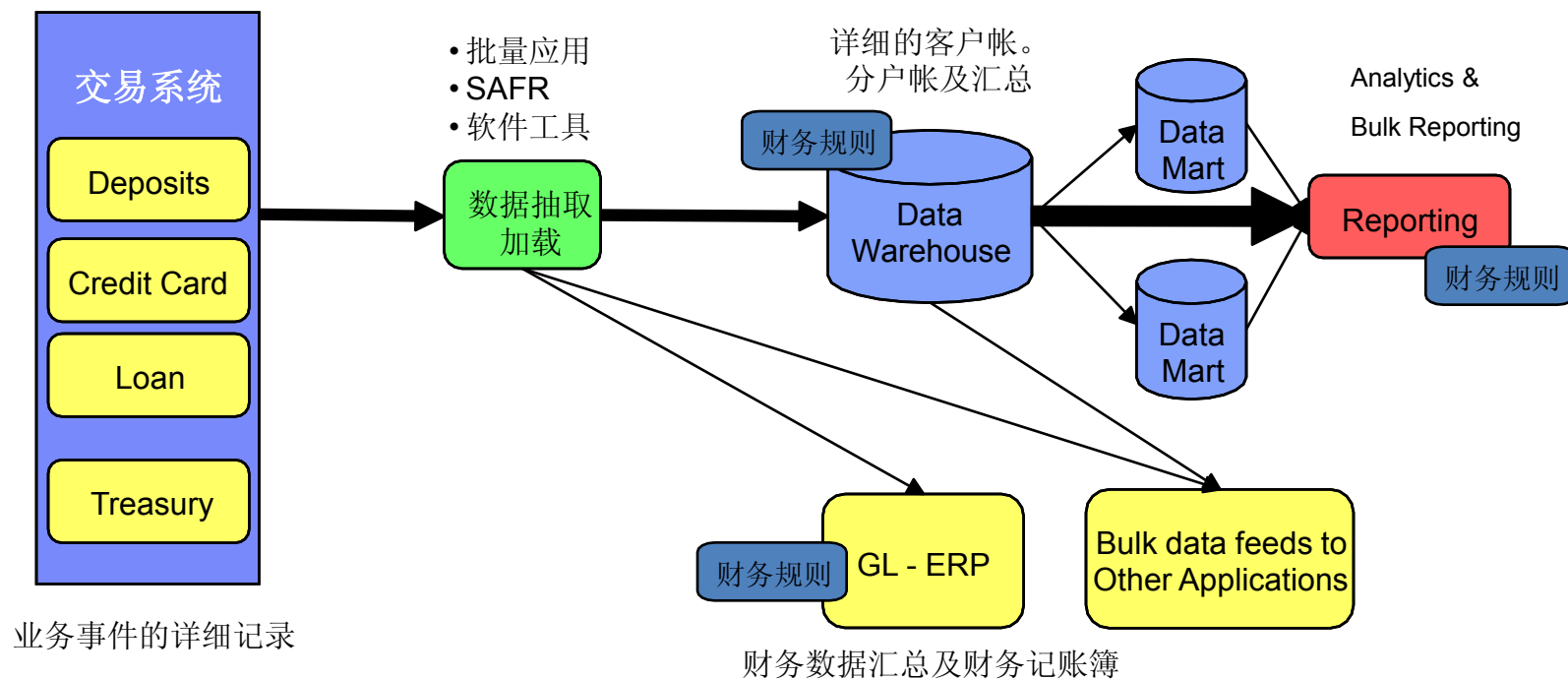
业务及技术驱动力

- 对于正确详细数据日益增加的需求
 - 分析报表动态的大量需求
 - 需要总帐汇总信息及详细的分户帐、客户帐信息来提供灵活的满足业务分析元素，以客户，产品，地区等等纬度分析
 - 复杂的平帐及调整
- 大规模的银行– 在现有z结构下，承受不了大数据库的分析型应用如：扫描汇总
 - 交易量过亿，账户数过4亿
 - 必须支持两级的数据服务30+省结构
- 数据可用性的延时，T+1服务水平的承诺难以达到
 - 需要跨多种异构平台收取源数据
 - 没有标准的数据抽取的处理流程和平台
 - 没有共享的集成数据平台，高冗余的数据存储、访问
- 需要在局域网及广域网进行大量数据的传输，因此造成一下后果
 - 对同一数据源进项重复的抽取
 - 数据的安全性没有保障，尤其跨广域网
 - 成本高-时间，带宽，系统资源

什么类型的数据及业务服务是z平台最能提供价值的

- 业务层面 - 面向业务的操作控制数据
 - 最重要的一个业务领域是银行的财务及管理会计。总帐及相关的操作型信息系统及数据。由于是基于会计准则，这些财务信息的数据模型相当成熟及稳定
 - 详细的客户帐及分户帐的数据量相当庞大，由于取于每天的业务交易数据，所以相对干净
- IT层面 - 操作型数据基础架构
 - 以上银行运营最关键的操作型业务数据**80%**起源于核心业务系统的交易型数据（如：存、贷、信用卡等），这些系统目前运行在**z**平台
 - 需要经过长批量的数据抽取、清洗、汇总及保留一定时期的存储以反映银行“财务唯一版本的事实”
 - 对于此类数据的业务功能服务需要满足动态的实时查询，大量的**T+1**批量报表，而报表的需求会容易变动
 - 并且这些操作型数据提供**BI**的基本数据，来形成分析型所需要的 **views, cubes**等等
- 因此以财务及管理会计数据为起步，是建立**z**操作性**BI**应用的一个有利的起点

让我们从一个简化了的典型操作型数据流程的现状分析



在哪里保存这些详细的财务事件数据？
需要保持多久？- 1 个月, 1 年, 5 年...
详细的财务事件数据的规模？
哪里存放准则？- 会计准则，业务及报表准则
当准则改变时，会发生什么？
报表的服务时间要求？- T+1, T+3, ME+8？

主题

- 当前大银行数据的基础架构全景及面临的挑战
- **大型银行数据集成基础架构设计指导原则**
- 最优化的交易型、操作型、分析型数据处理流程基础架构
- 业务及技术价值
- 建议

优化当前交易数据、操作数据、分析型数据的基础架构，需要遵照大型银行数据集成的基本原则

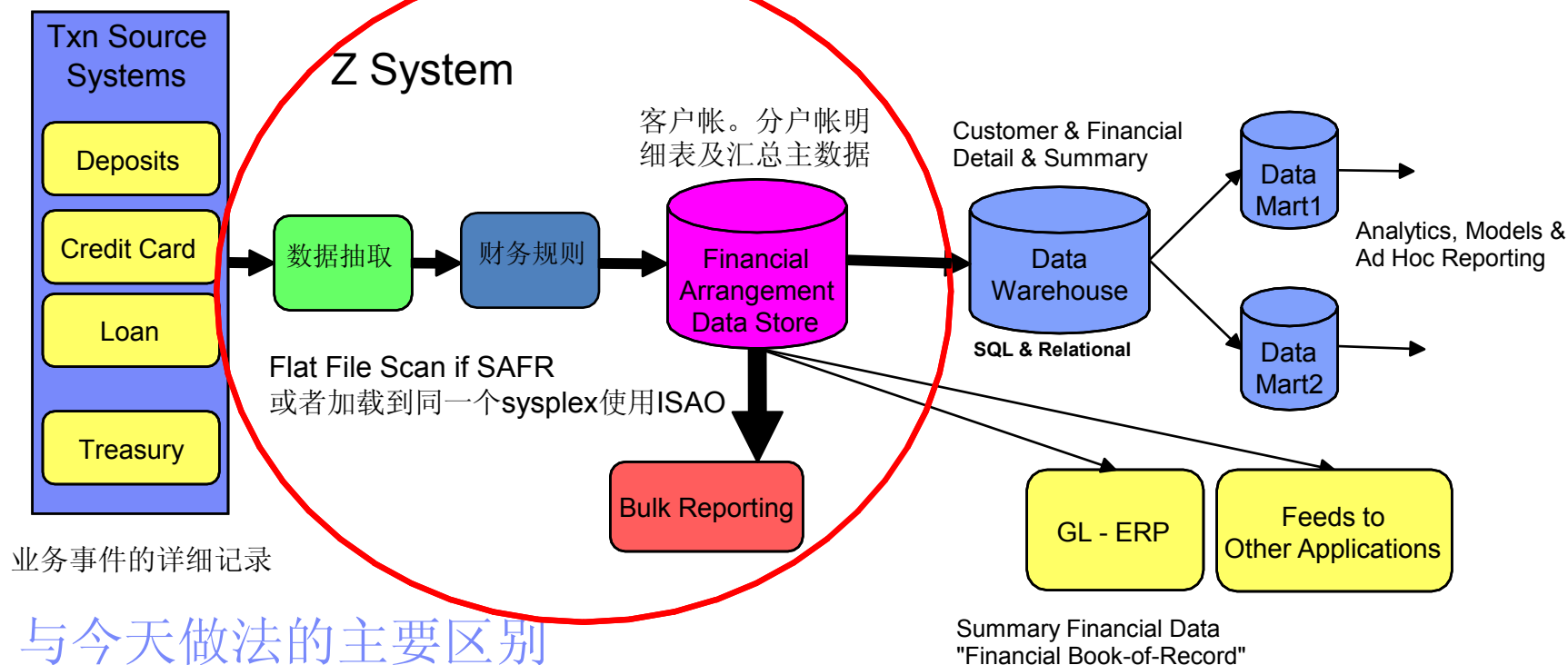
- 从原数据抽取财务数据仅一次
 - 保证数据来自于相同的时间段 Ensure all data is from same time-slot
 - 针对生产系统合并所有的抽取流程 Consolidate processing against production systems
- 集中数据的清洗及转型，集中规则的管理，基于规则的数据处理及报表
 - 保证对所有的数据实施相同的技术和业务规则
 - 集中规则从而简化业务规则的维护
 - 基于规则的报表生成而不是程序绑定从而提供报表的灵活度
- 保持最低层面的详细数据，靠近源数据
 - 保证数据的透明性 – drill-through from summary data
 - 保证汇总数据的一致性（所有的数据均从同样的原数据求和）
 - 减少数据的搬动

主题

- 当前大银行数据的基础架构全景及面临的挑战
- 大型银行数据集成基础架构设计指导原则
- **最优化的交易型、操作型、分析型数据处理流程基础架构**
- 业务及技术价值
- 建议

建立分户帐明细操作型数据存储

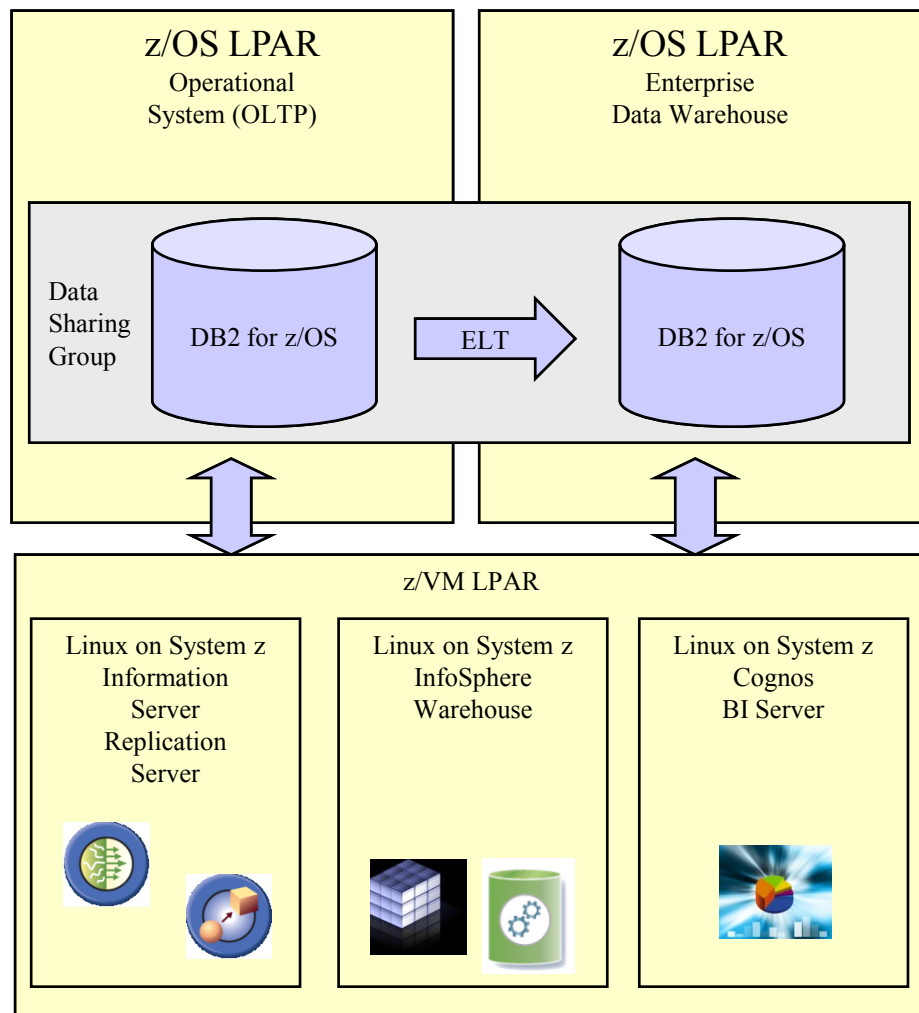
Detailed Financial Arrangement Data Store



与今天做法的主要区别

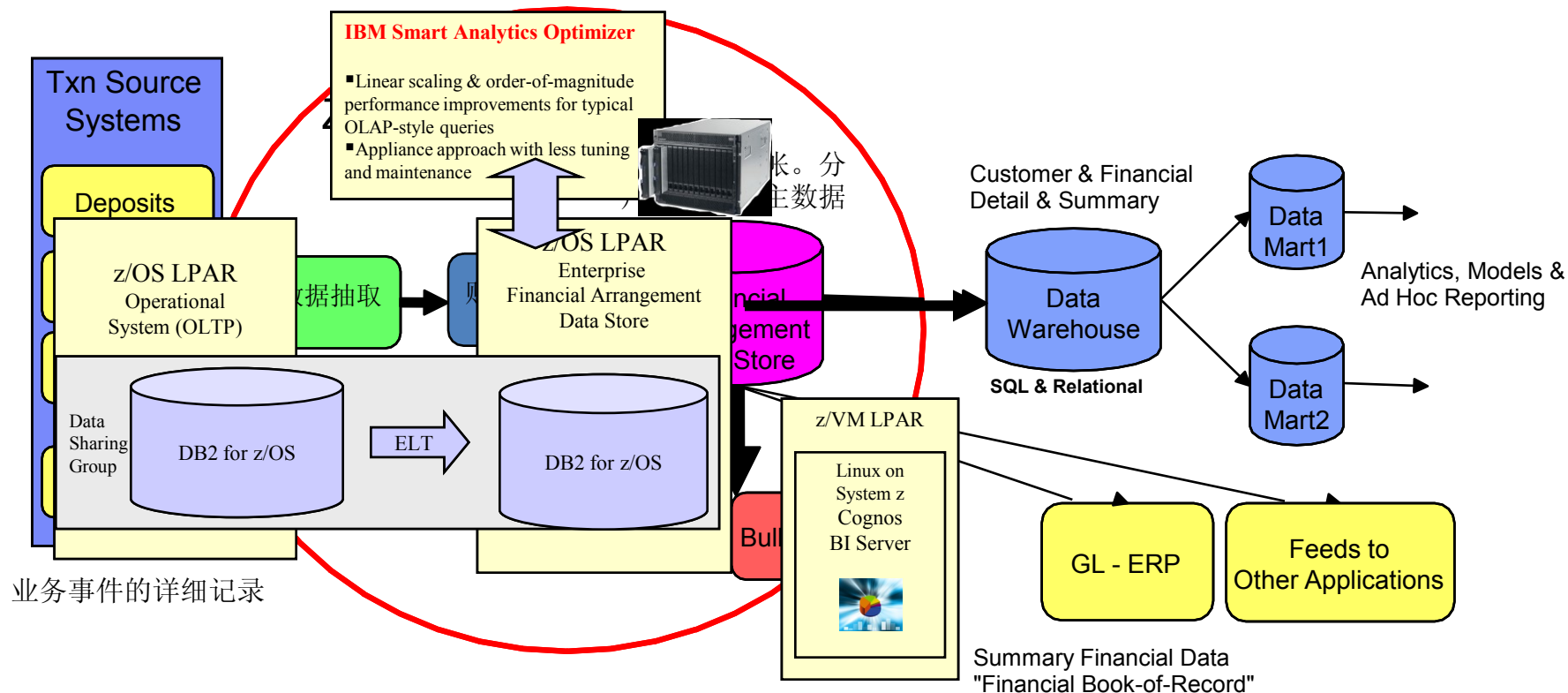
- 1) 集中会计规则在一个系统上
- 2) 在数据的处理流程的上游实施业务规则
- 3) 仅从原数据抽取财务数据一次
- 4) 将客户帐、分户帐等明细表及汇总表存储在一个集中的地方
- 5) 保持细粒度的详细数据 5 年
- 6) 为生成生产报表提供一致的详细数据

Z BI 部署架构



- 将关键的业务数据集中在**System z**
- 能够借助现有**system z**的高可用性，备份及监管流程环境及技能
- 在同一个数据共享的环境中有效的数据迁移 (不通过网络)
- 通过**cubing**服务(数据超市)及DB2的功能增强而提高性能及TCO
- 通过**Linux on System z with Information Server**完成复杂的数据转化及数据质量的控制

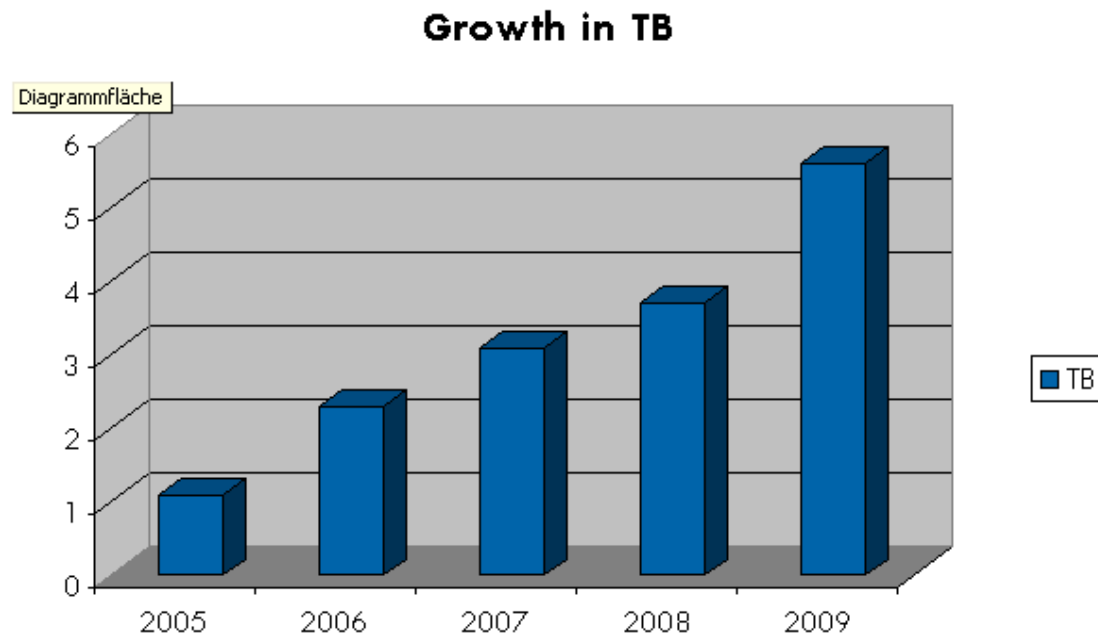
最优化的交易型、操作型、分析型数据处理流程基础架构



- 将从OLTP抽取的财务明细表装入MQT表。由于在同一个数据库共享组中，所以
 - ✓不转码
 - ✓不FTP
- 通过zLinux运行报表分析客户端，通过z的混合型平台运行复杂的查询、汇总等分析应用
- 自动享受2地3中心的高可用性及灾备体系
- 高时效性
- 减少数据的冗余

试验客户的实际数据

- 5 TB 的数据
 - 14 DB, 1568TS, 3517 IX, 1568 Tables, 53118 Columns
- 8000 夜数据仓库作业，4000个日分析作业
- 数据仓库的数据每日即时刷新，历史数据保持4年
- 最大表10亿条记录
- 数据增长如表所示



主题

- 当前大银行数据的基础架构全景及面临的挑战
- 大型银行数据集成基础架构设计指导原则
- 最优化的交易型、操作型、分析型数据处理流程基础架构
- **业务及技术价值**
- 建议

重新架构业务数据平台，运行OLAP on z的业务及技术价值

- 财务报表在z上的集中可以标准化，模块化处理财务报表流程，从而提高报表的性能，时效性
- 把交易流程及报表生产集中在一个平台上，可以减少数据的冗余影射，参照及平衡数据的维护开销
- zNext/ISAO具有数据仓库查询优化，大量的并行运行机制，庞大的数据在内存，降低z的运行成本，提高OLAP的能力
- ISAO有强大的线性可扩展性
- 无需DB2的繁杂的调优 (MQTs, aggregates, indexes, etc.)
- 保持传统System z及DB2的QOS，对于银行管理，监管生产报表可以纳入当前主机的灾备体系，提高业务的恢复范围

主题

- 当前大银行数据的基础架构全景及面临的挑战
- 大型银行数据集成基础架构设计指导原则
- 最优化的交易型、操作型、分析型数据处理流程基础架构
- 业务及技术价值
- **建议**

ISAO POC目标及计划

- 目标:
 - 验证DB2 on zOS加ISAO运行分析型应用在相应时间，CPU 使用率及应用的透明性等方面的能力，促使客户有足够的信心将分析型应用OLAP及操作型和分析型数据保留在主机上
- 计划:
 - 举办ISAO专题研讨会，详细介绍OLAP on DB2 for zOS+ISAO的价值
 - 选择适当的现有分析型应用(WAS,Cognos)
 - 分析确定适当的可以利用ISAO应用的数据模型
 - 设计测试场景
 - 比较测试结果，显示ISAO的性能价格比