

数据仓库技术架构及方案

黄予辉

2008年12月13日

You've never seen your business like this before.



Agenda

- **Teradata**简介
- 架构设计原理
- 整体架构说明
- **ETL**架构说明



Teradata 公司概况

- Teradata Corporation – 2007年10月1日纽交所上市
 - > 企业级数据仓库全球领导者
 - EDW/ADW 数据库技术
 - 分析型解决方案
 - 咨询服务
 - > 自1999年开始，连续9年被Gartner评为数据仓库第一名
- 美国前10大上市软件公司
 - > S&P 500 标准普尔500成员
 - > 纽交所上市代码: “TDC”
 - > NYSE Arca Tech 100
- 世界级的客户遍布全球
 - > 超过 850 个世界级客户
 - > 超过2000个安装系统
- 全球员工超过5,500名



Teradata 市场份额

Teradata Top 10



**90% of Top 10
Global Telco Firms**



**70% of Top 10
Global Airlines**



**60% of the Top 10
Transportation
Logistic Firms**



**50% of Top 10
Global Retailers**



**40% of Top 10
Global Commercial
& Savings Banks**

FORTUNE Global Rankings, July 2006

- **Leading industries**
 - > **Banking/Financial Services**
 - > **Government**
 - > **Insurance & Healthcare**
 - > **Manufacturing**
 - > **Retail**
 - > **Telecommunications**
 - > **Transportation Logistics**
 - > **Travel**
- **World class customer list**
 - > **More than 850 customers**
 - > **More than 2,000 installations**
- **Global presence**
 - > **Over 100 countries**

Teradata 驱动世界级企业的可持续发展

Post



Retail



Communications



Financial



Travel



Insurance

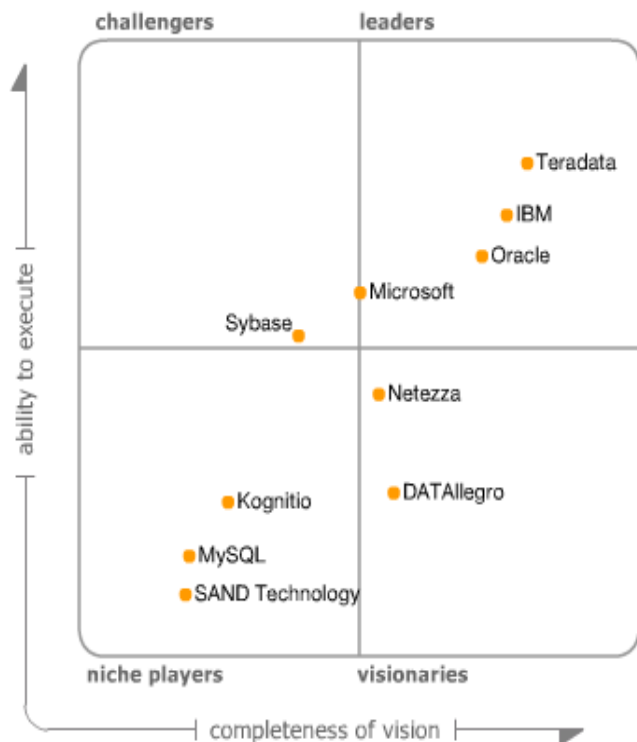


Manufacturing



Teradata 数据仓库技术的领导者

软件

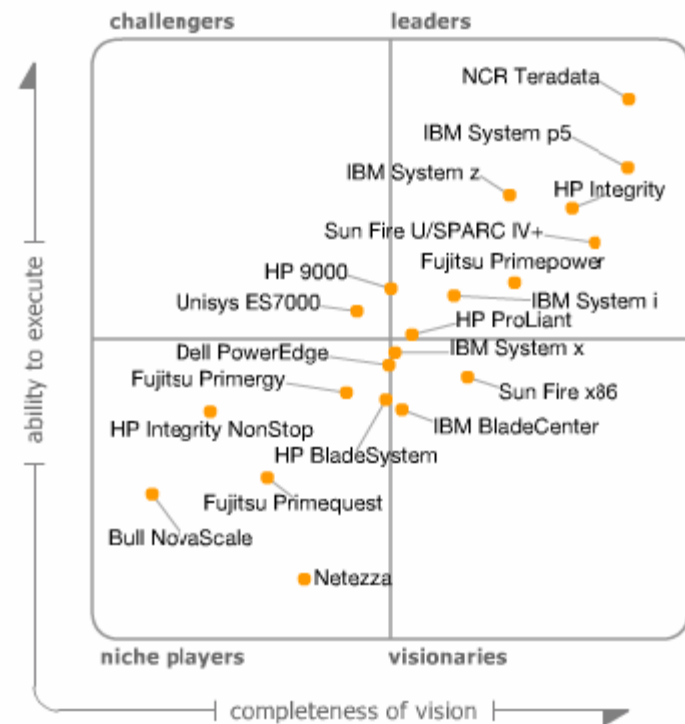


As of September 2006

Gartner Magic Quadrant for Data Warehouse DBMS, 2006

Feinberg & Beyer (9/2006)

硬件



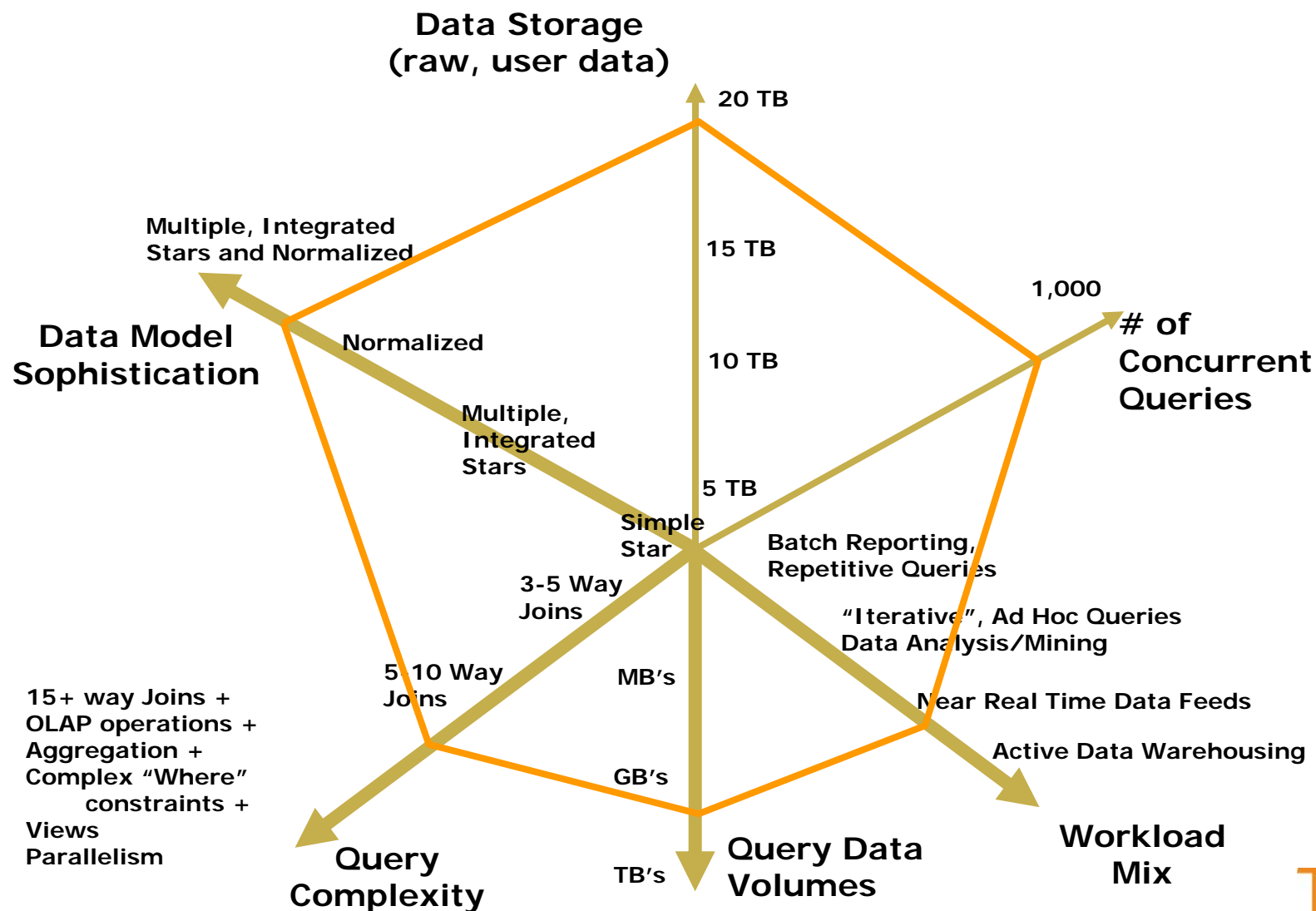
As of August 2006

Gartner Magic Quadrant for Data Warehouse DBMS Servers, 2006

Feinberg, Hardcastle, Butler, Dawson (8/25/2006)

The Magic Quadrant is copyrighted 9/12/06 by Gartner, Inc. and is reused with permission. The Magic Quadrant is a graphical representation of a marketplace at and for a specific time period. It depicts Gartner's analysis of how certain vendors measure against criteria for that marketplace, as defined by Gartner. Gartner does not endorse any vendor, product or service depicted in the Magic Quadrant, and does not advise technology users to select only those vendors placed in the "Leaders" quadrant. The Magic Quadrant is intended solely as a research tool, and is not meant to be a specific guide to action. Gartner disclaims all warranties, express or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

Teradata 系统扩展能力

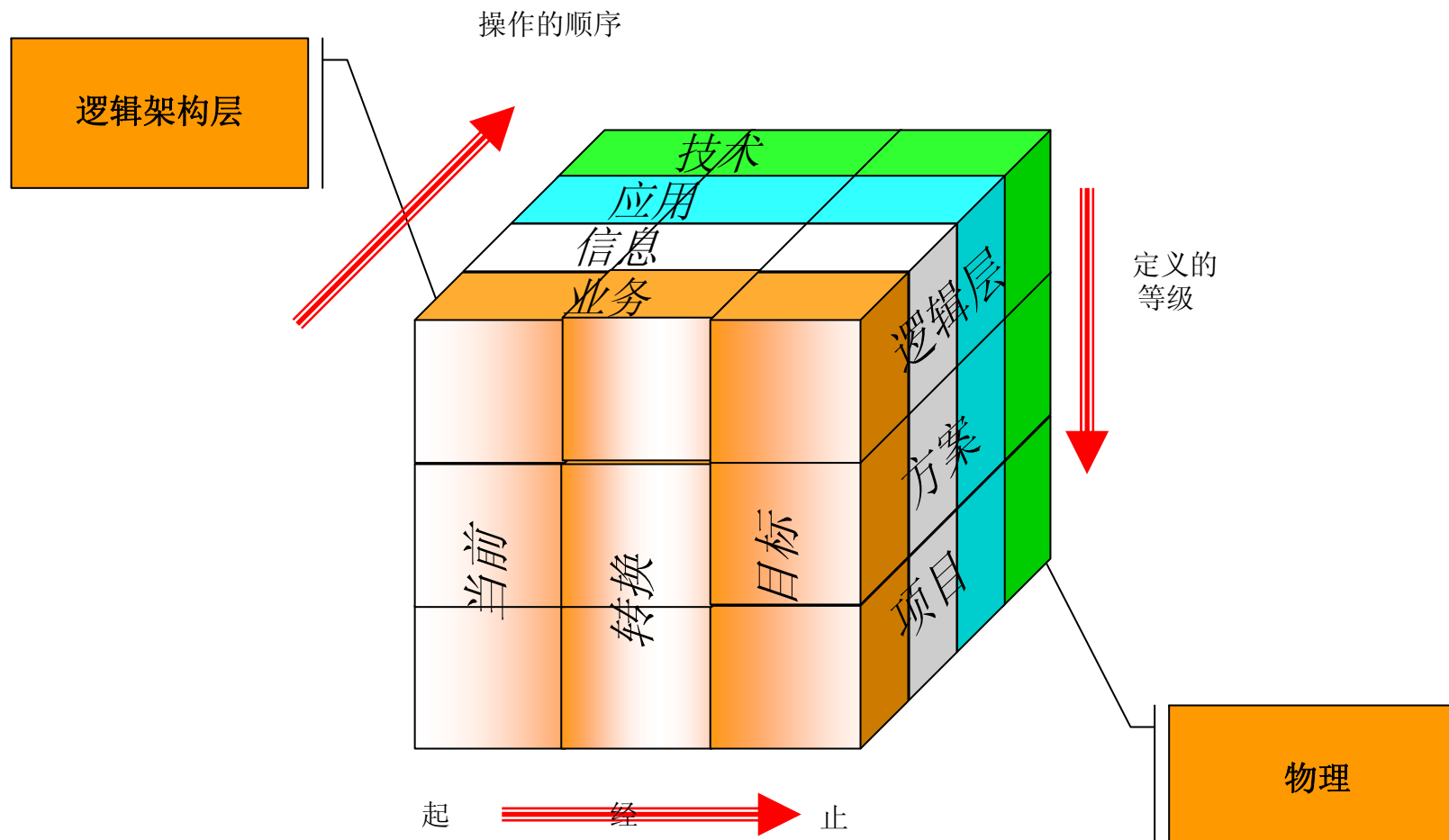


Agenda

- **Teradata**简介
- 架构设计原理
- 整体架构说明
- **ETL**架构说明



架构立方



成功由好的架构设计方法开始

- 业务

- > What is the business model, where is it going, how does it plan to get there?
- > The requirements. The business process.

- 信息

- > What data do we have and need to support the Business View? Information is also calculations and rules. Typically we see Logical & Physical data models here, all subject areas of the business.
- > The data is worked on by the applications, used by the business.

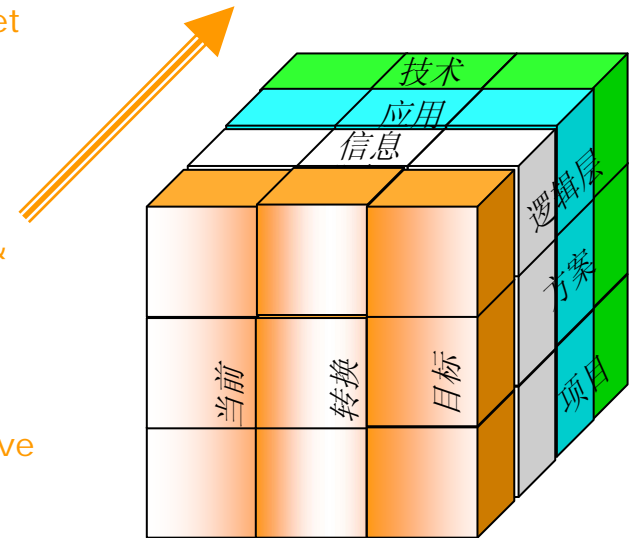
- 应用

- > What functions and interrelations of functions do the applications have and need? Sales, Marketing, Pricing, Manufacturing, Customer Management.
- > Works against information to support the Business View. **The applications work within the confines of the Information architecture, creating and consuming the data elements, rules and definitions of that architecture view.**

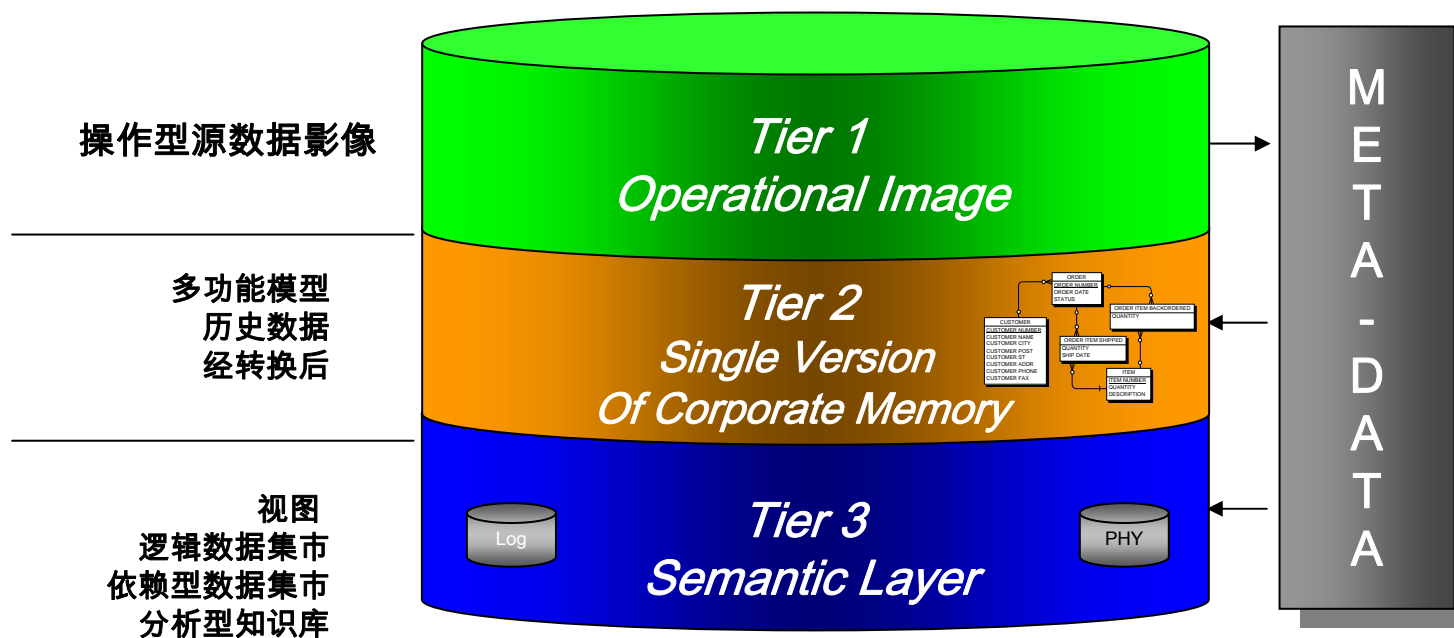
- 技术

- > The bit IT cares about most. The easiest to get WRONG because we don't concentrate on the other aspects of architecture FIRST!
- > What do we have and need to support the other 3 Views without limitation?
- > Choosing an ETL tool before you have defined an architecture is "SUB OPTIMAL"

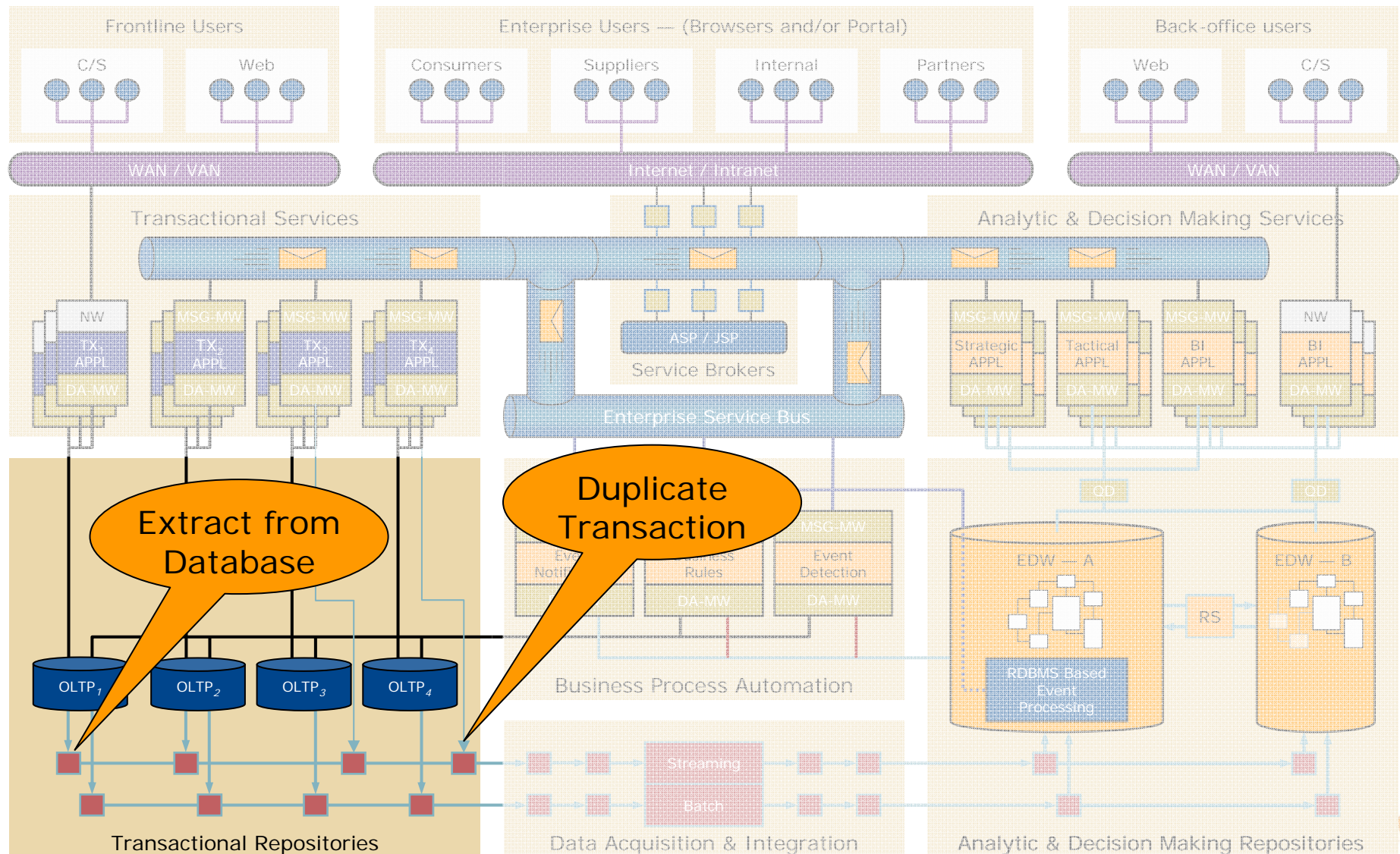
- A separation of Business, Data, Process and Technology as appropriate



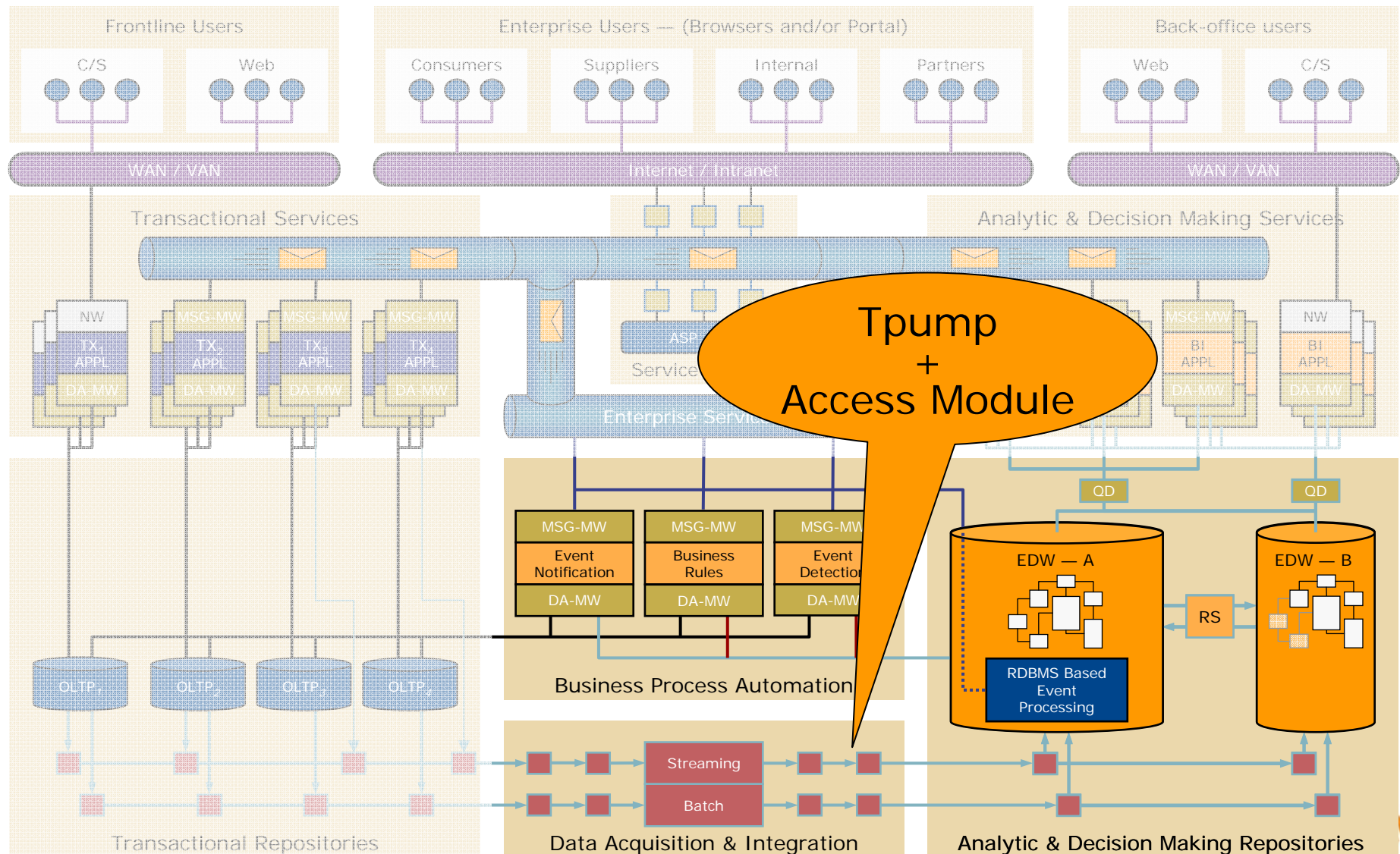
EDW 应用逻辑架构



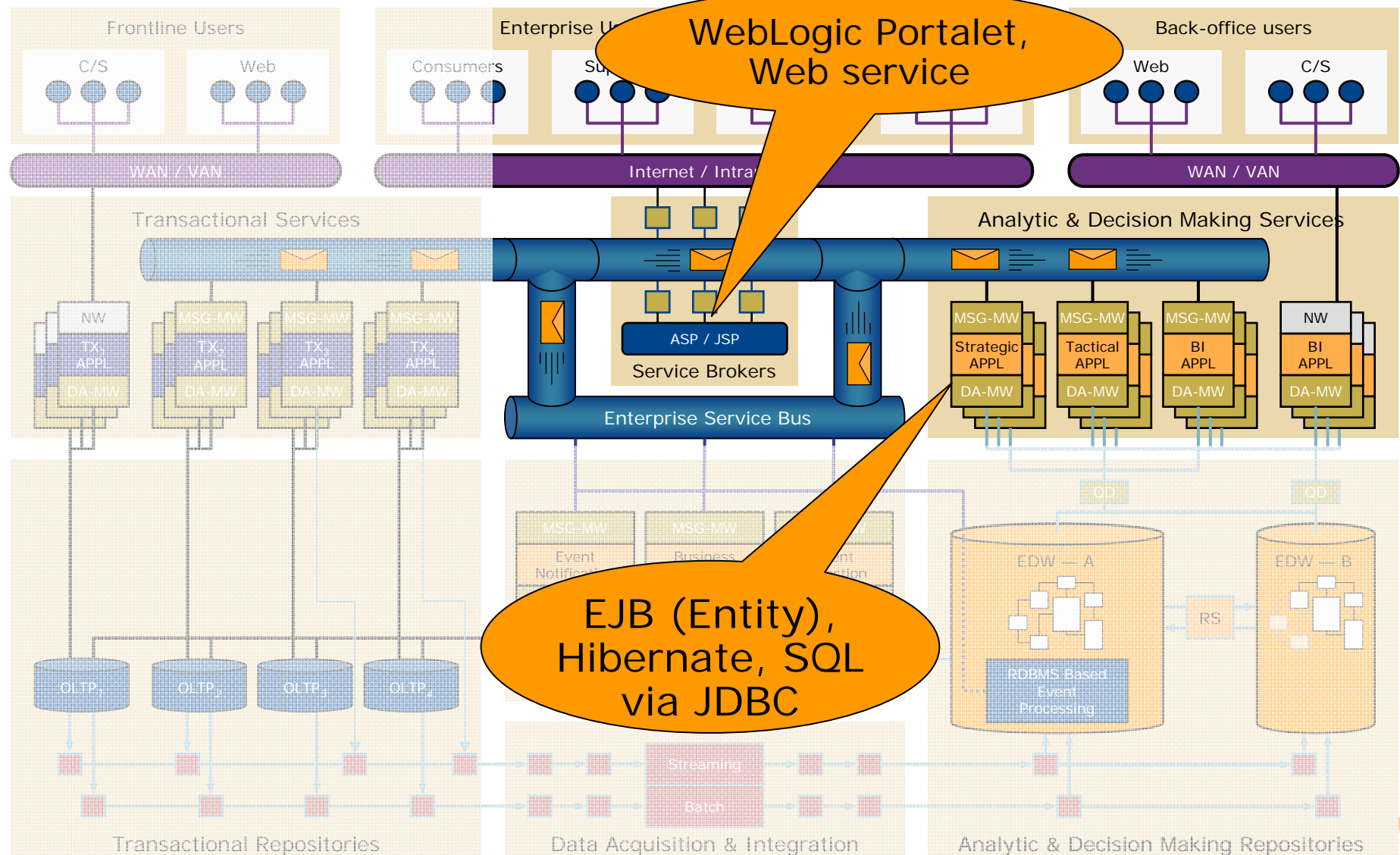
Reference Architecture - Source View



Reference Architecture – Data loading View



Reference Architecture - Application View

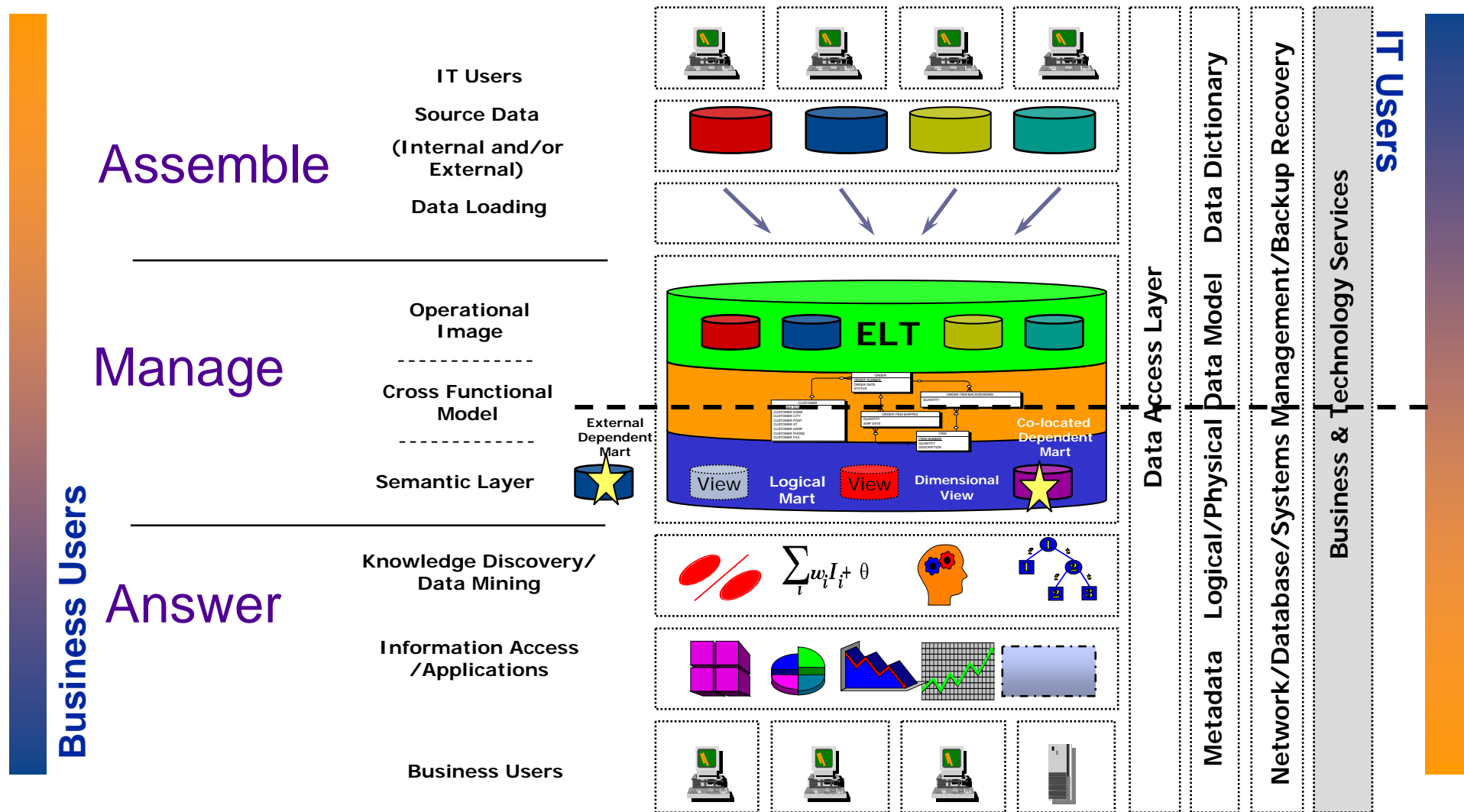


Agenda

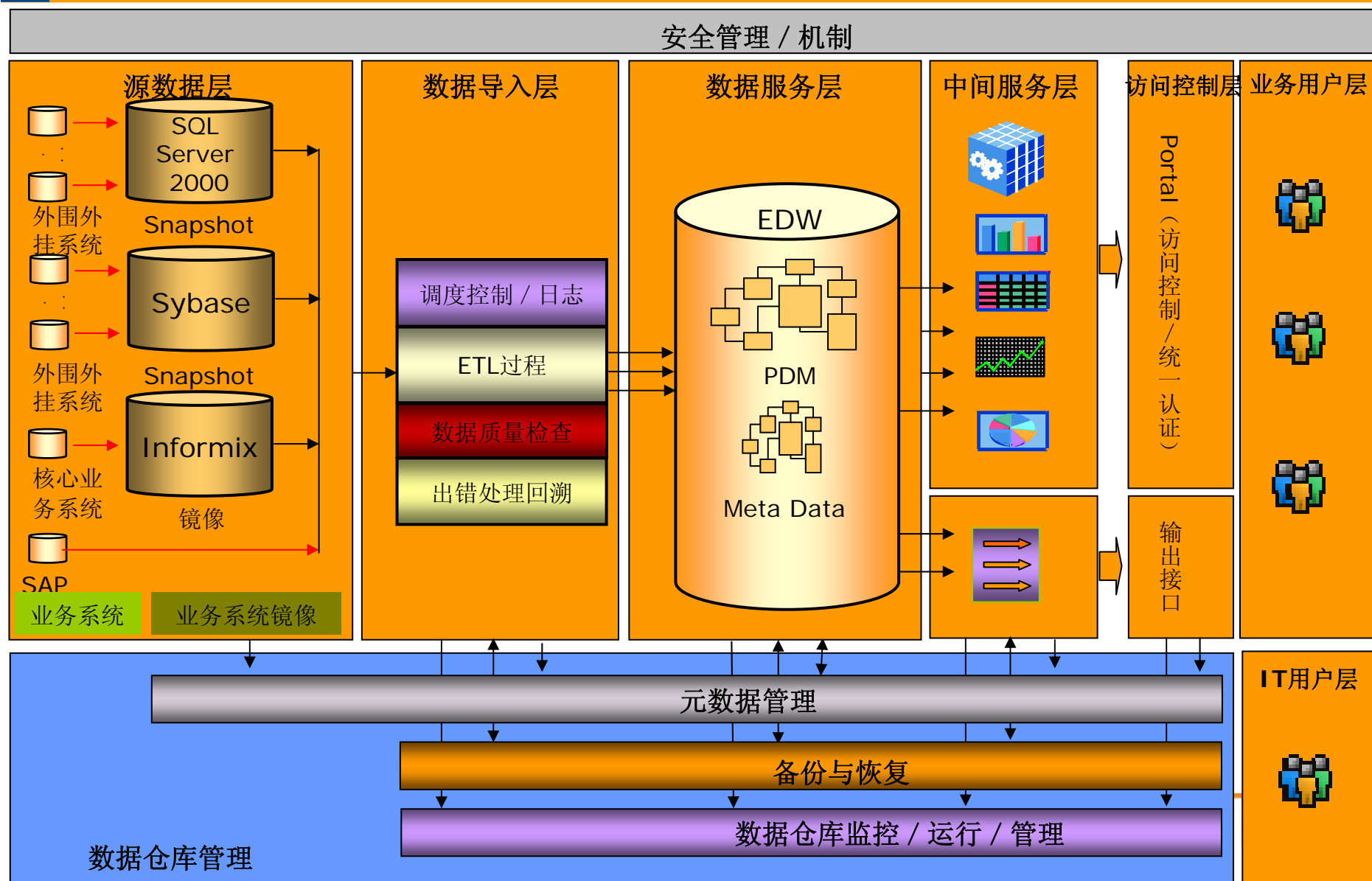
- **Teradata**简介
- 架构设计原理
- 整体架构说明
- **ETL**架构说明



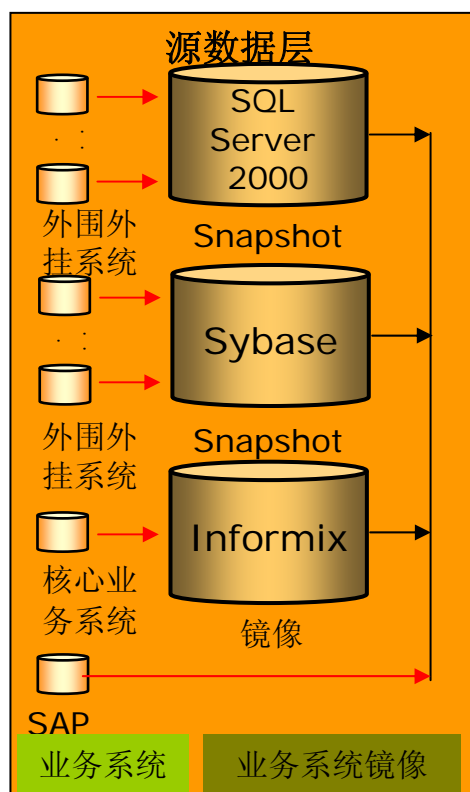
Teradata Warehouse Framework – Since forever...



技术体系架构设计



源数据层



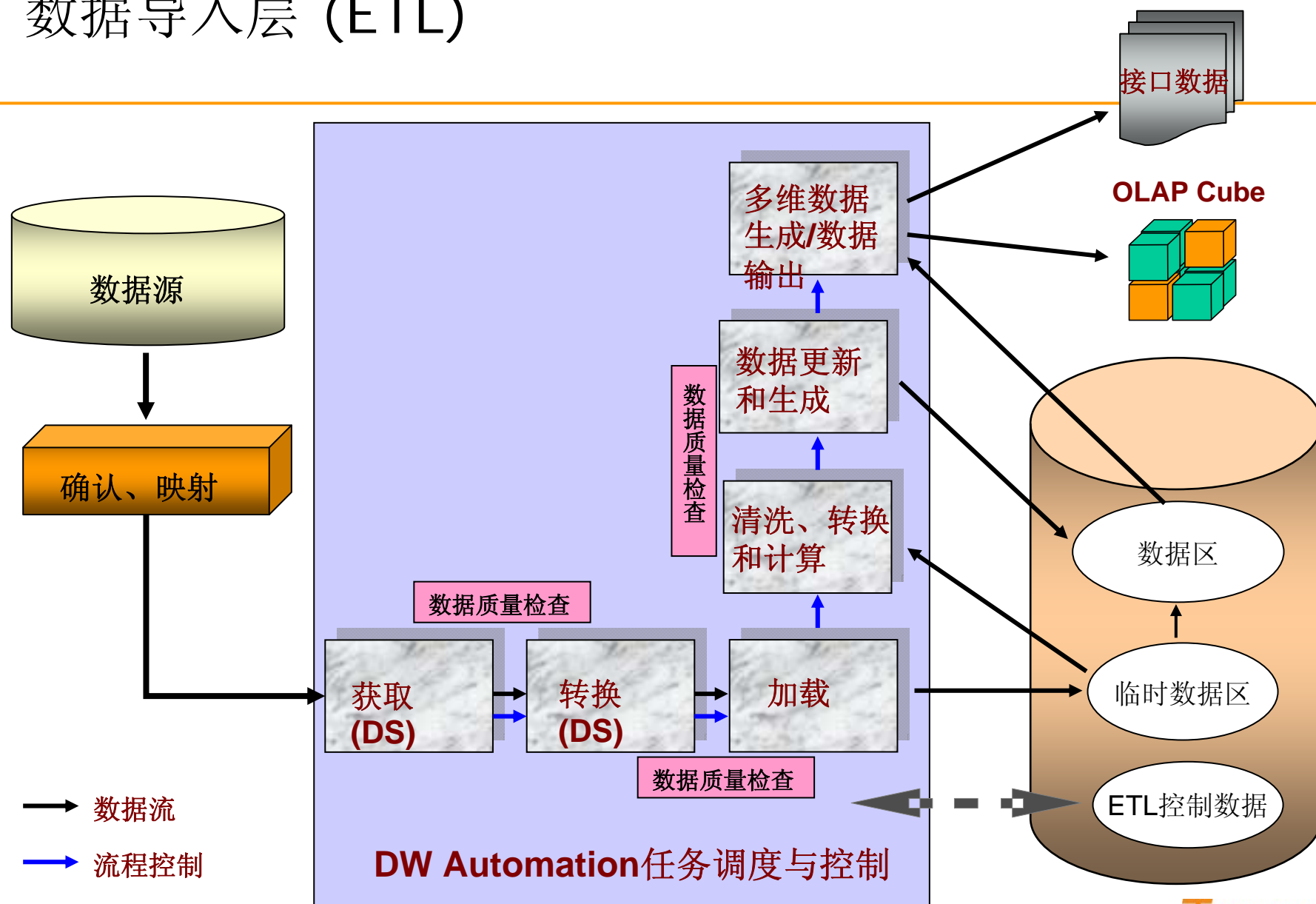
- 【组成部分】
 - > 核心业务系统的镜像系统 (**Informix**平台)
 - > 数据汇集平台1 (**SQL Server 2000**平台)
 - > 数据汇集平台2 (**Sybase**平台)
 - > **SAP**系统 (**SQL Server 2000**平台)
- 【存在形式】
 - > 关系数据库**RDBMS**及其对象如**View** (在线数据)
 - > 磁带备份 (历史数据)
- 核心业务系统通过数据库镜像提供对核心业务数据的访问
- 数据汇集平台通过数据库**Snapshot**来获取对相关源系统业务数据的访问
- **SAP**系统直接提供对业务数据的访问

数据导入层 (ETL)

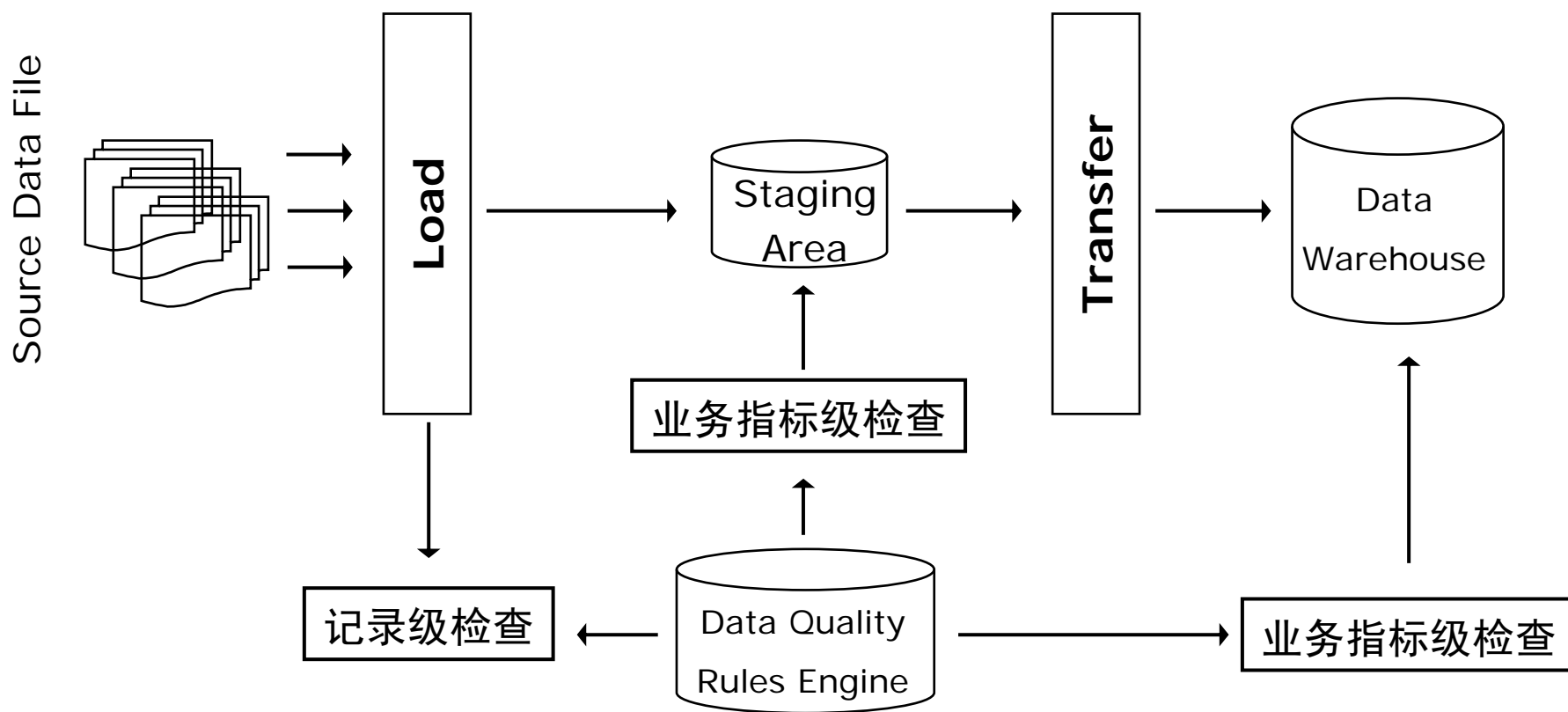


- 【功能与作用】
 - > 从源数据层的数据抽取采用**DataStage**
 - > 向数据仓库中央数据库的数据加载采用**Teradata**的数据加载工具或**DS**插件
 - > 整个**ETL**流程管理采用**Teradata**的**DW Automation**工具。
 - > **ETL**系统的具体要求需要与其他系统的**ETL**需求综合考虑。
- 【组成部分】
 - > **ETL**调度与日志跟踪管理模块
 - > **ETL**运行脚本模块
 - > 数据质量检查模块
 - > 出错处理与回溯模块

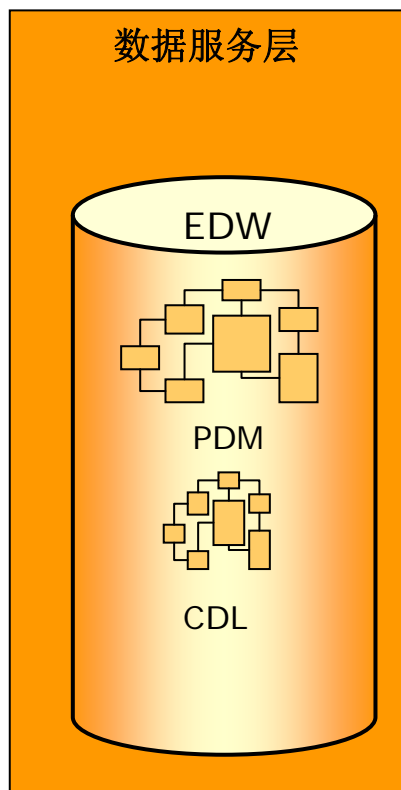
数据导入层 (ETL)



数据质量检查实施框架

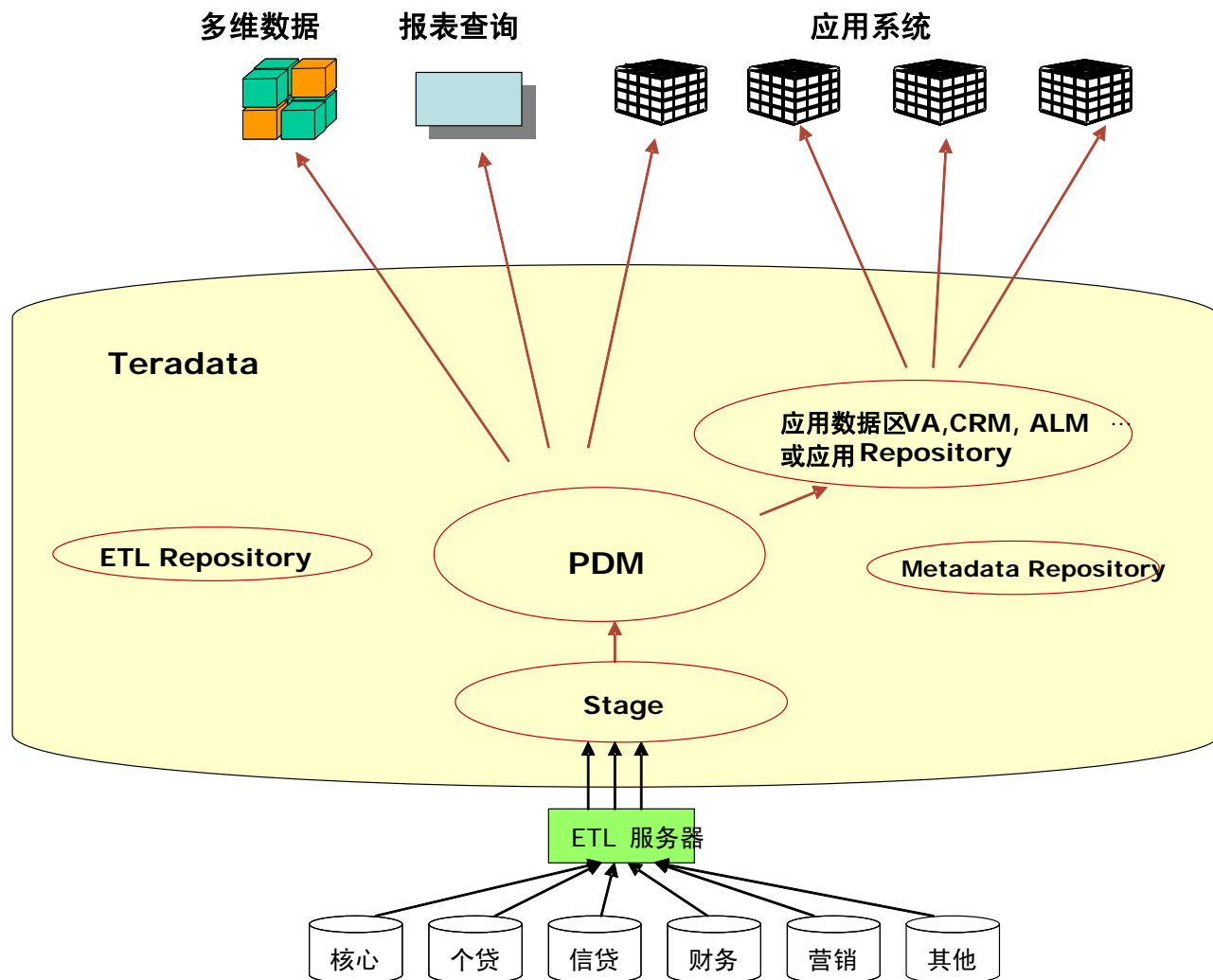


数据服务层 (中央数据库)

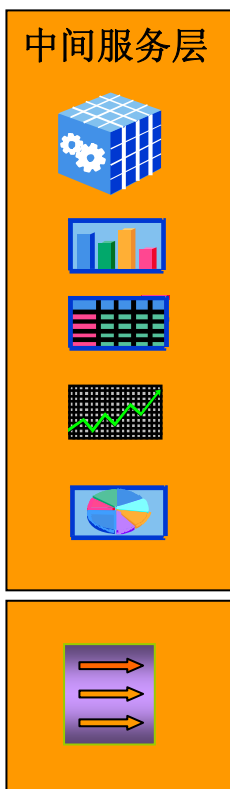


- 【功能与作用】
 - > 作为本系统的数据核心部分，负责存储和管理来自各种源数据系统的数据，并为访问用户提供数据服务。
 - > 数据按照逻辑数据模型分主题存放
 - > 采用Teradata的Teradata关系型数据库。
- 【组成部分】
 - > 设计(Design)：逻辑数据模型(LDM)和物理数据模型(PDM)，同时也包括维护和管理数据模型各个版本的过程与方法。工具是ERWin。LDM采用Teradata的FS-LDM作为设计的基础。
 - > 存储(Store)：数据存储体系是专业数据仓库引擎Teradata以及Teradata 5400 MPP硬件平台组成。
- 数据存储内容包括两部分组成：
 - > 数据仓库基础物理数据模型：用于保存各类业务的基础数据；
 - > 数据仓库汇总数据模型：用于保存各类业务的汇总数据；

数据服务层与前端应用的关系

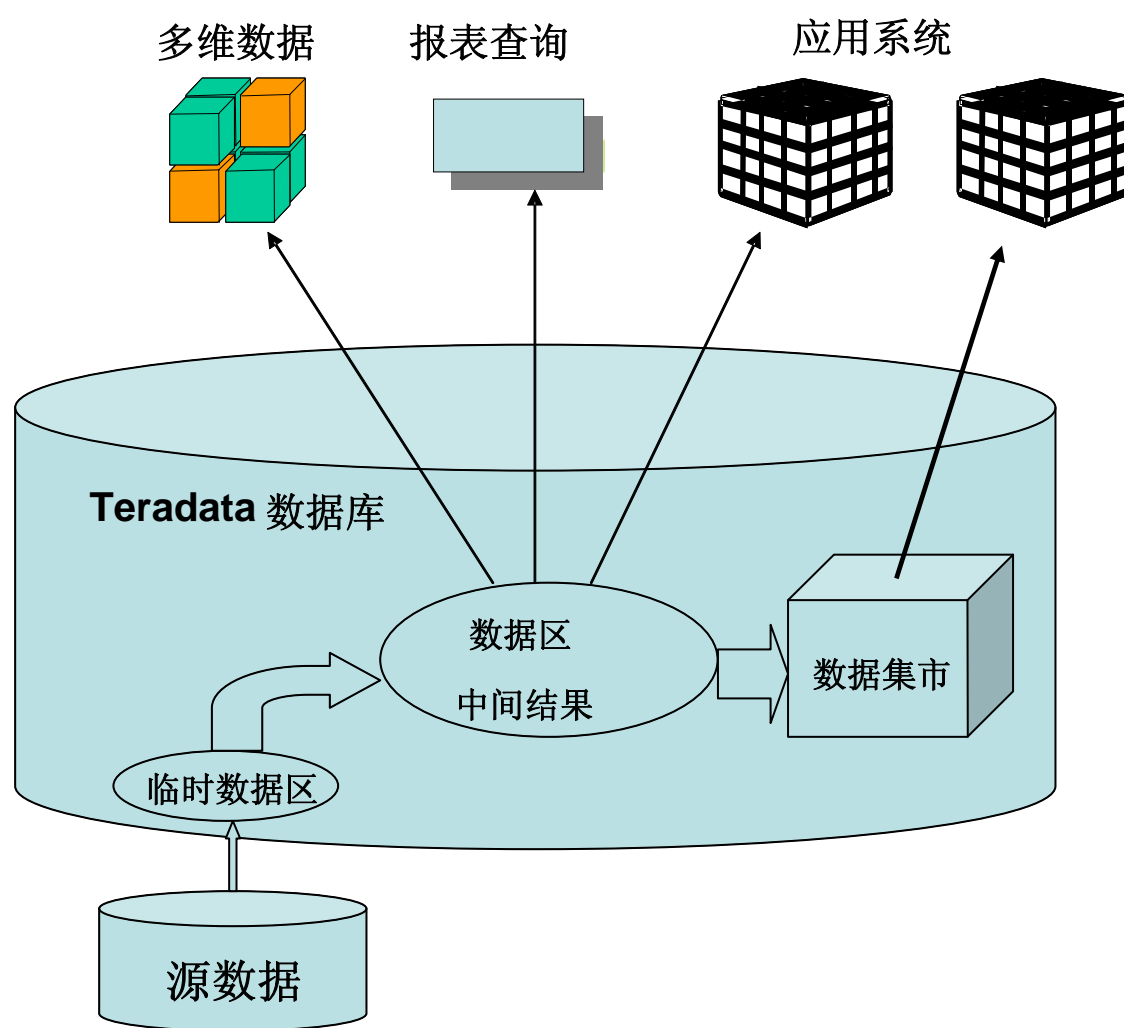


中间服务层 (业务应用)



- 【功能与作用】
 - > 向用户提供包括**OLAP**服务、报表服务、查询服务、数据挖掘服务和中间件服务等多种服务应用。
 - > 为用户对中央数据的访问提供各种方式的服务，从而实现访问方式的多样化和信息存取的透明化。
- 【组成部分】
 - > 中间数据服务
 - > 数据访问

中间数据服务



- 【功能与作用】
 - > 建立从属数据集市，或 **OLAP Cube**
 - > 满足用户对数据访问的多样化、灵活化。

数据访问

- 通过各种应用的定制与展现，用户对数据进行访问，包括：
 - > 报表与动态查询服务
 - 包括**Hyperion Performance Suite**的访问接口与工具，用于定制与发布报表和动态查询应用；
 - > 元数据服务
 - 包括实现一组实现元数据应用报表和分析功能的应用接口；
 - > 数据仓库管理服务
 - 提供数据仓库管理的信息查看、报警等服务。
 - > 其他服务
 - 包括实现与其他应用系统集成的组件。
 - > 此外，还包括用于对外进行数据输出的服务
 - 这些服务是由一些数据导出组件组成的，为其他部门或机构提供数据仓库的对外数据输出，例如为**SAP**等其他业务部门提供业务数据的输出。

访问控制层

访问控制层

Portal (访问控制/统一认证)

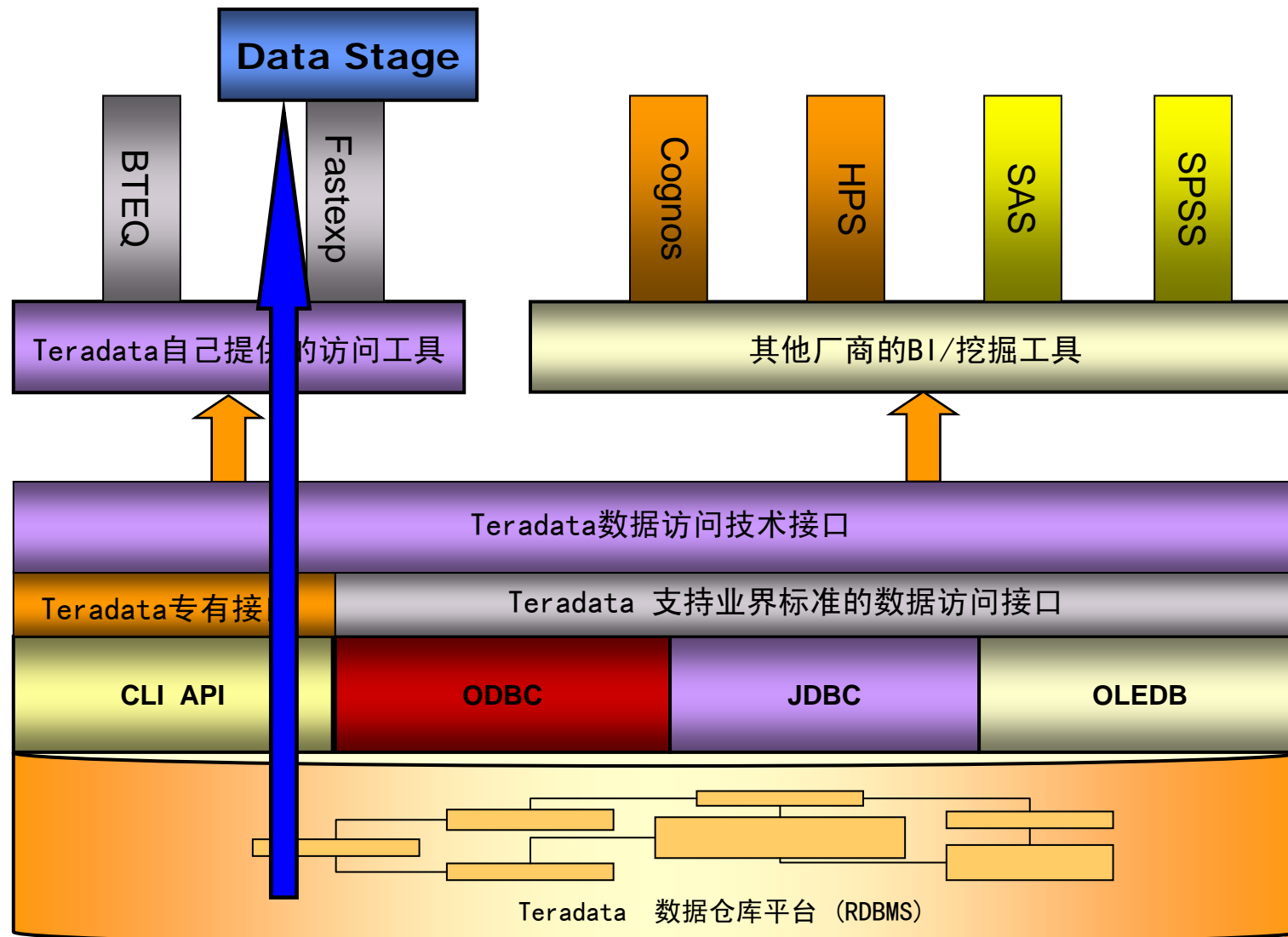
- 【功能与作用】

- > 主要包括**WEB**、认证、安全、门户四方面的服务。提供**HTTP Web**服务、门户的用户登录、用户认证、门户**Web**应用、并提交用户层请求到中间服务层，对用户实施安全策略，为用户管理报表、查询文档，提供个性化定制等。

- 【组成部分】

- > **Web Server**采用**WebLogic**。
 - > 从产品的集成性角度出发，建议在数据仓库中采用**HPS**自带的**Portal**，日后等浦发银行有了相应的技术标准之后，再来讨论具体的对应策略。
 - > 经过与浦发银行的讨论，建议采用**2**次登录的方式实现用户的安全认证，而这也是目前浦发银行在其他项目中所采用的登录方式。
 - > 由于数据仓库的用户来源于**OA**网段和生产网段，因此，需要在生产网段设置**Portal**服务器的同时，在**OA**网段设置**Portal**服务器的代理服务器。所有通过**OA**网段对数据仓库的访问，都将通过该代理服务器转到生产网段的**Portal**服务器上。

数据输出接口



用户安全管理及系统安全管理

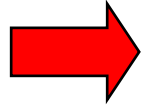
- 用户分类
 - > 业务用户
 - 决策管理层用户（决策用户）
 - 部门专业用户（一般用户）
 - 信息分析型用户（超级用户）
 - > 技术用户
 - > 操作用户
- 用户管理
 - > 用户分组
 - > 用户级别
- 用户安全管理
 - > 操作系统的安全控制
 - > 数据权限控制
 - > 数据库权限控制
 - > 前端功能的用户管理和访问权限控制
- 用户日志及审计
- 数据转移安全
 - > 网络安全与协议
 - > 操作系统安全
 - > 软件工具安全
- 数据存储安全
 - > **MPP**架构
 - > 物理安全、用户登录与数据存取
 - > 数据保护 — 磁盘阵列**RAID**技术、**Teradata**数据库级数据保护
 - > 备份与恢复
- 数据访问安全
 - > 数据库级访问 — 登录管理、存取权限、视图对象、存取日志

Agenda

- Teradata简介
- 架构设计原理
- 整体架构说明
- **ETL架构说明**



ETL 架构说明

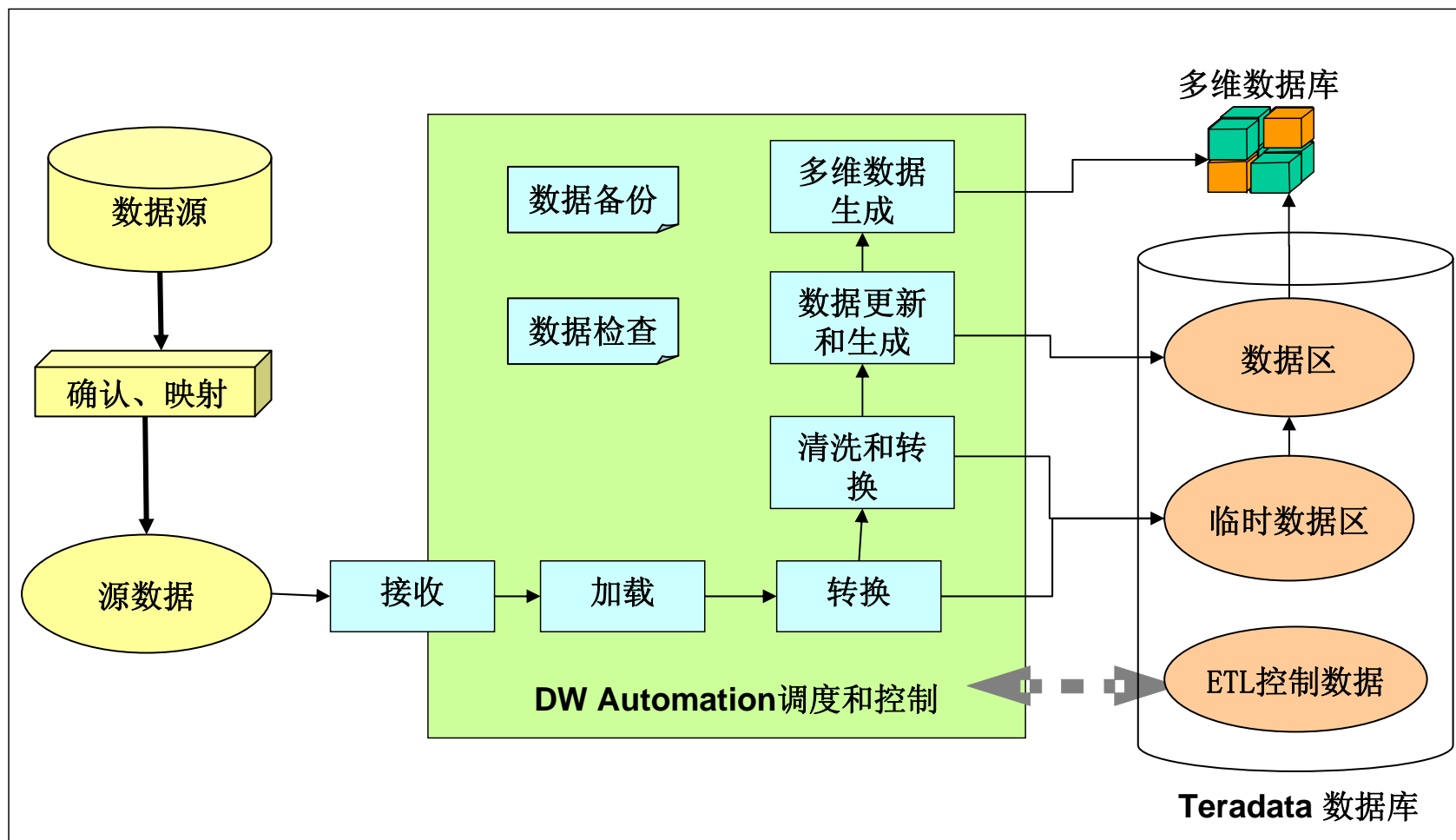


- 体系架构
- ETL Automation 简介



ETL架构

ETL典型流程



ETL架构

完整的ETL加载系统由三部分组成

1.ETL Automation 运行环境

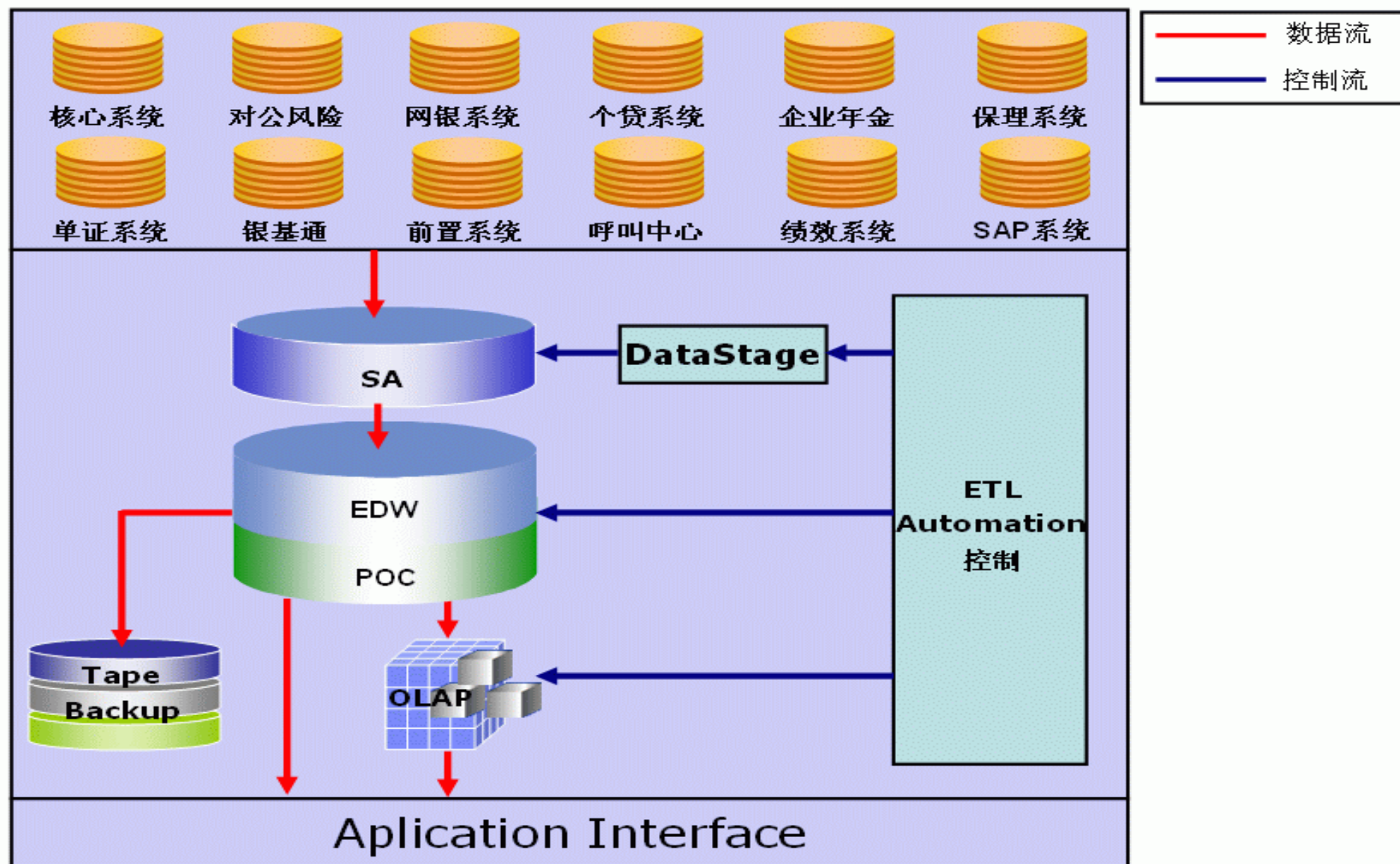
- > 标志文件到达，则启动作业运行流程
- > 判断作业的执行条件
- > 运行作业的加载脚本
- > 作业的日志管理

2.客户化的数据加载流程

- > 实现和业务系统的数据接口
- > 确保数据加载的同步和一致
- > 优化的作业加载流程
- > 合理安排加载时间窗口，满足客户的要求

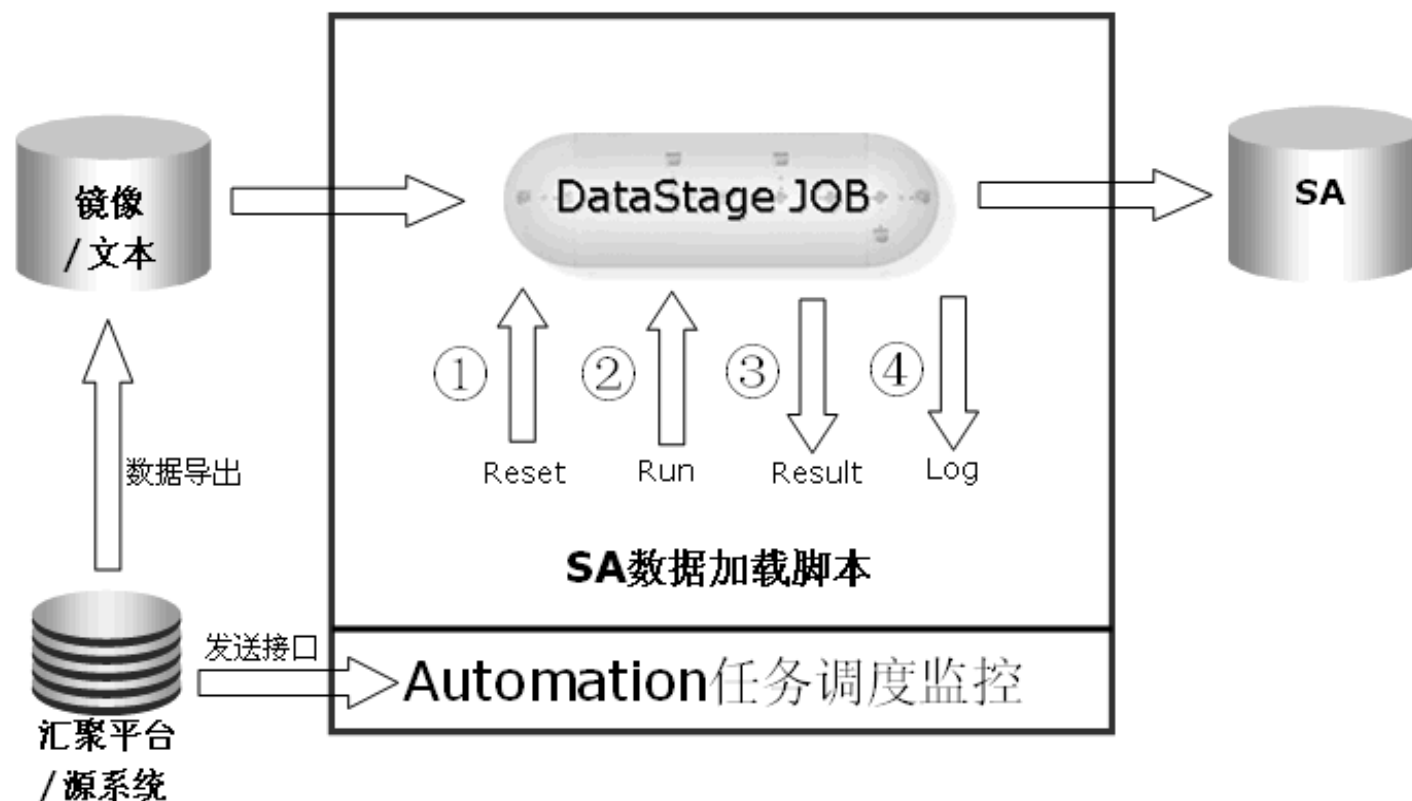
3.加载脚本

SPDB ETL 架构

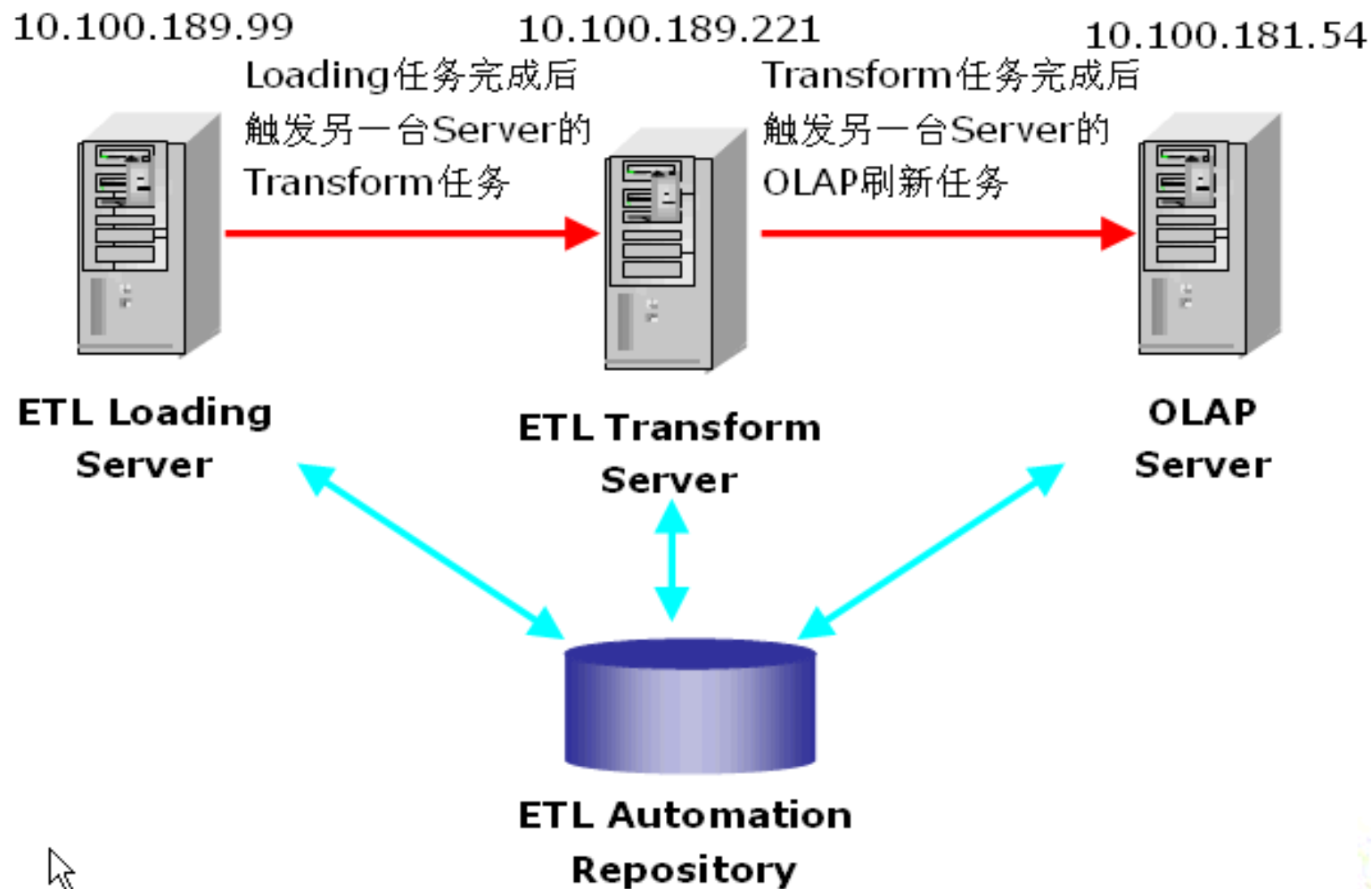


DataStage 集成

- **DataStage**工具将完成由镜像数据库或文件到**Teradata**数据库**SA**的数据抽取和加载过程
- **ETL Automation**来完成整个数据仓库**ETL**的统一管理和流程控制

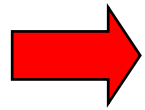


SPDB ETL 环境部署



ETL 加载

- 体系架构



- ETL Automation 简介



ETL Automation概况

- **Teradata** 自主开发
- 运行环境: **Windows 2000/XP/2003**
- **Perl** 封装
- 实现加载流程自动化管理
- 数据仓库自动化管理
- 支持多服务器并行
- 图形化的管理和监控

ETL Repository

- **ETL Automation** 知识库(etl repository)建立在Teradata 数据库中
- **Repository Tables**可分为两类：配置表（用户维护）和日志表（系统维护）
- **ETL Automation**的参数设置依赖于**Repository Tables**中的配置表
- **ETL Automation**在运行过程中不断地向**Repository Tables**读写记录

Repository Tables

表名	用途	类别
ETL_Sys	系统配置表	用户维护（常用）
ETL_Job	作业配置表	用户维护（常用）
ETL_Job_Source	源作业配置表	用户维护（常用）
ETL_Job_Stream	作业触发关系配置表	用户维护（常用）
ETL_Job_Dependency	作业依赖关系配置表	用户维护（常用）
ETL_Job_Group	作业组配置表	用户维护（常用）
ETL_Job_GroupChild	作业组成员配置表	用户维护（常用）
ETL_Job_TimeWindow	作业时间窗口配置表	用户维护（常用）
ETL_Server	DW Automation客户端配置表	用户维护（常用）
DataCalendar	日历表，用于检查作业日期	用户维护（可选）
DataCalendarYear	日历年表	用户维护（可选）
ETL_RelatedJob	相关作业配置表	用户维护（可选）
ETL_Notification	自动发邮件、短信配置表	用户维护（可选）
ETL_User	自动发邮件、短信目标用户表	用户维护（可选）
ETL_UserGroup	自动发邮件、短信目标用户组表	用户维护（可选）
ETL_GroupMember	自动发邮件、短信目标用户组成员表	用户维护（可选）

Repository Tables(Cont.)

表名	用途	类别
ETL_TimeTrigger	定点触发配置表	用户维护（可选）
ETL_TimeTrigger_Calendar	定制定时作业定义表	用户维护（可选）
ETL_TimeTrigger_Monthly	月定时作业定义表	用户维护（可选）
ETL_TimeTrigger_Weekly	周定时作业定义表	用户维护（可选）
ETL_Received_File	数据文件信息表，用于load作业	系统维护
ETL_Record_Log	加载记录信息表，用于load作业	系统维护
ETL_Job_Status	作业状态表，用于所有作业	系统维护
ETL_Job_Log	作业日志表，用于所有作业	系统维护
ETL_Job_Queue	作业队列表，用于定时作业和异地作业	系统维护
ETL_Job_Trace	作业跟踪表，用于定时作业	系统维护
ETL_Event	事件表，用于监控所有作业的运行状态	系统维护

ETL Automation的功能

- 加载作业的图形化设置、维护和修改
- 以数据驱动的作业自动执行
- 协调作业执行的相互关系
- 作业组管理
- 作业日志管理
- 数据定时清理
- 支持多加载服务器
- 作业并行处理
- 加载作业的实时监控
- 加载作业的远程报警-**Email**
- 开放的加载作业数据接口

系统驻留程序

- 七个系统驻留程序完成数据加载控制

Etlagent

Etlclean

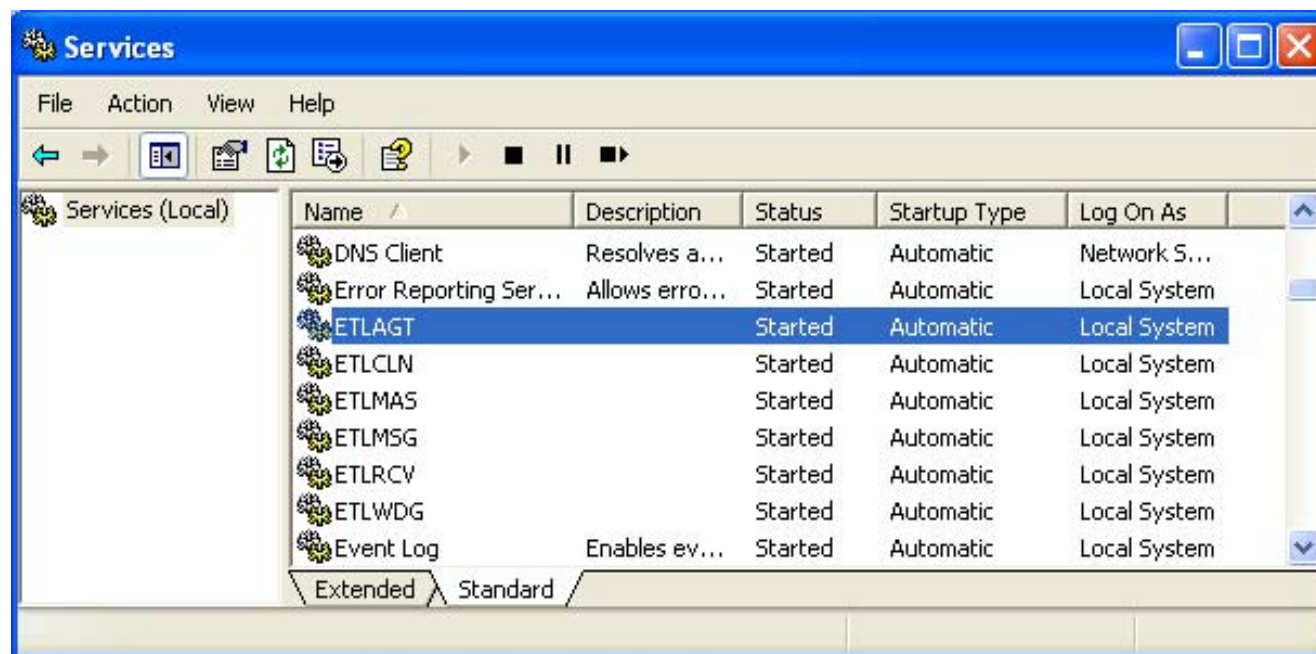
Etlmaster

Etlmsg

Etlrcv

Etlwdog

Etlschedule



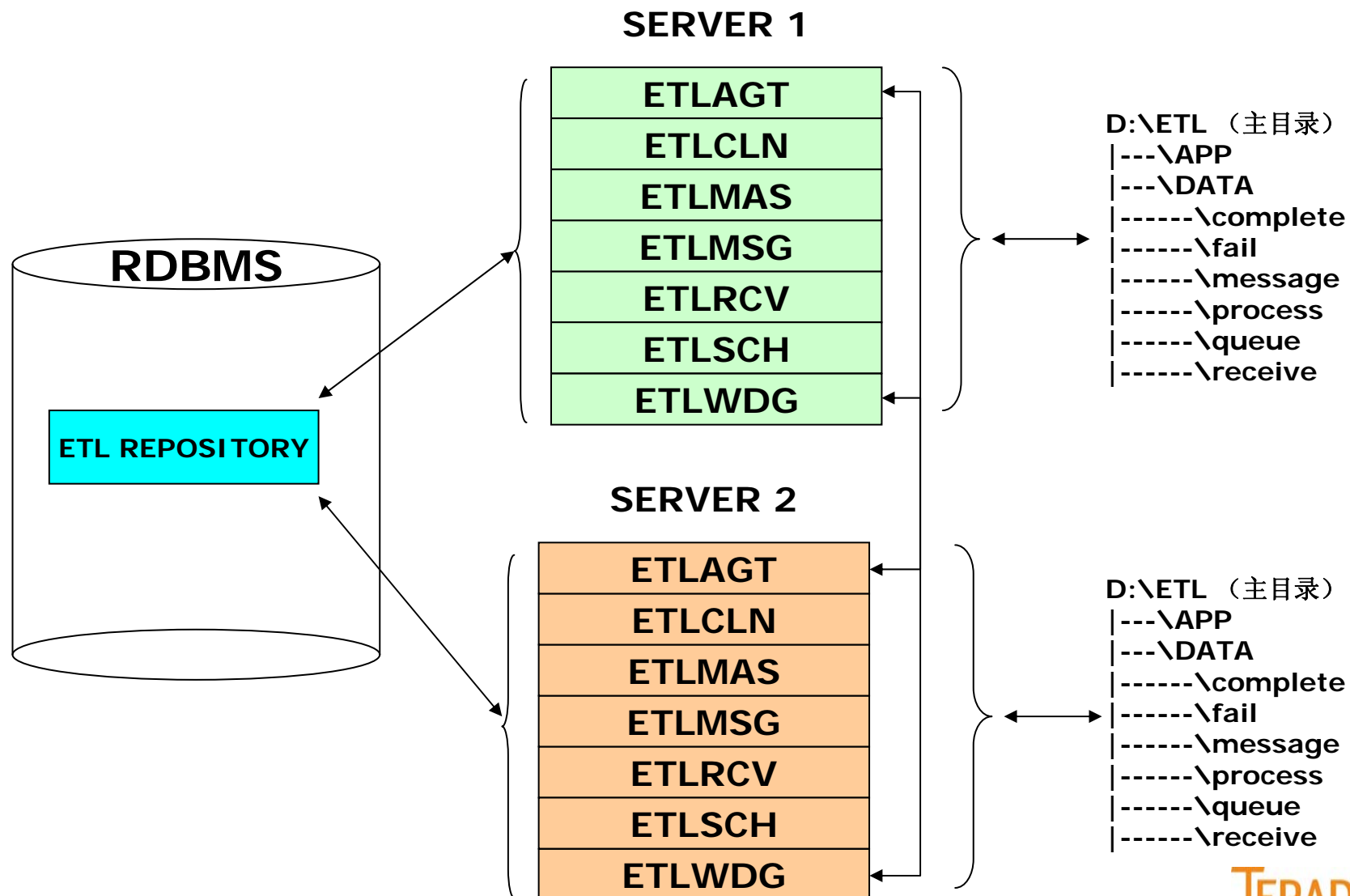
系统驻留程序

程序名称	说明
etlagent.exe	负责执行远程操作的动作
etlclean.exe	负责执行数据清除 (House Keeping) 的动作以及在每天晚上 12 点一过时会将所需要执行的作业新增至作业的追踪记录中, 以便后续的作业执行记录查询
etlrcv.exe	负责处理从各来源系统所传送来的档案. 当档案通过检查后, 会将档案交由 etlmaster.pl 来处理
etlmaster.exe	负责处理由 etlrcv.pl 所交付的档案. etlmaster.pl 会找出档案所对应的作业并在作业的执行条件满足时呼叫 etlslave_nt.pl 或 etlslave_unix.pl 来执行作业.
etlmsg.exe	是负责发送讯息通知 (Message Notification) 的动作.
etlwdog.exe	在主要服务器(Primary server) 上是负责接收各次要服务器 (Secondary Server)所传送来的存活封包(Heartbeat Packet) 以判断各服务的状态并更新记录. 而在发现有次要服务器未能在一段时间内送来存活封包的话, 则会认定此次要服务器为停止状态, 而会将当时正在此部次要服务器上执行的作业的状态设成失败 (Failed). 而在次要服务器(Secondary Server)上则是定时发送存活封包 (Heartbeat Packet) 至主要服务器 (Primary Server)上以通知目前次要服务器的状态
etlschedule.exe	定时任务调度程序

并行多ETL 服务器配置

- **ETL Automation**支持多**ETL**加载服务器
- **Primary Server**
- **Secondary Server**
- 服务器独立完成各自的作业
- 各服务器有各自的加载作业脚本
- 服务器的状态由**Primary Server**统一监控
- 统一的流程控制
- 服务器内作业并行
- 服务器之间作业并行

并行多ETL 服务器配置

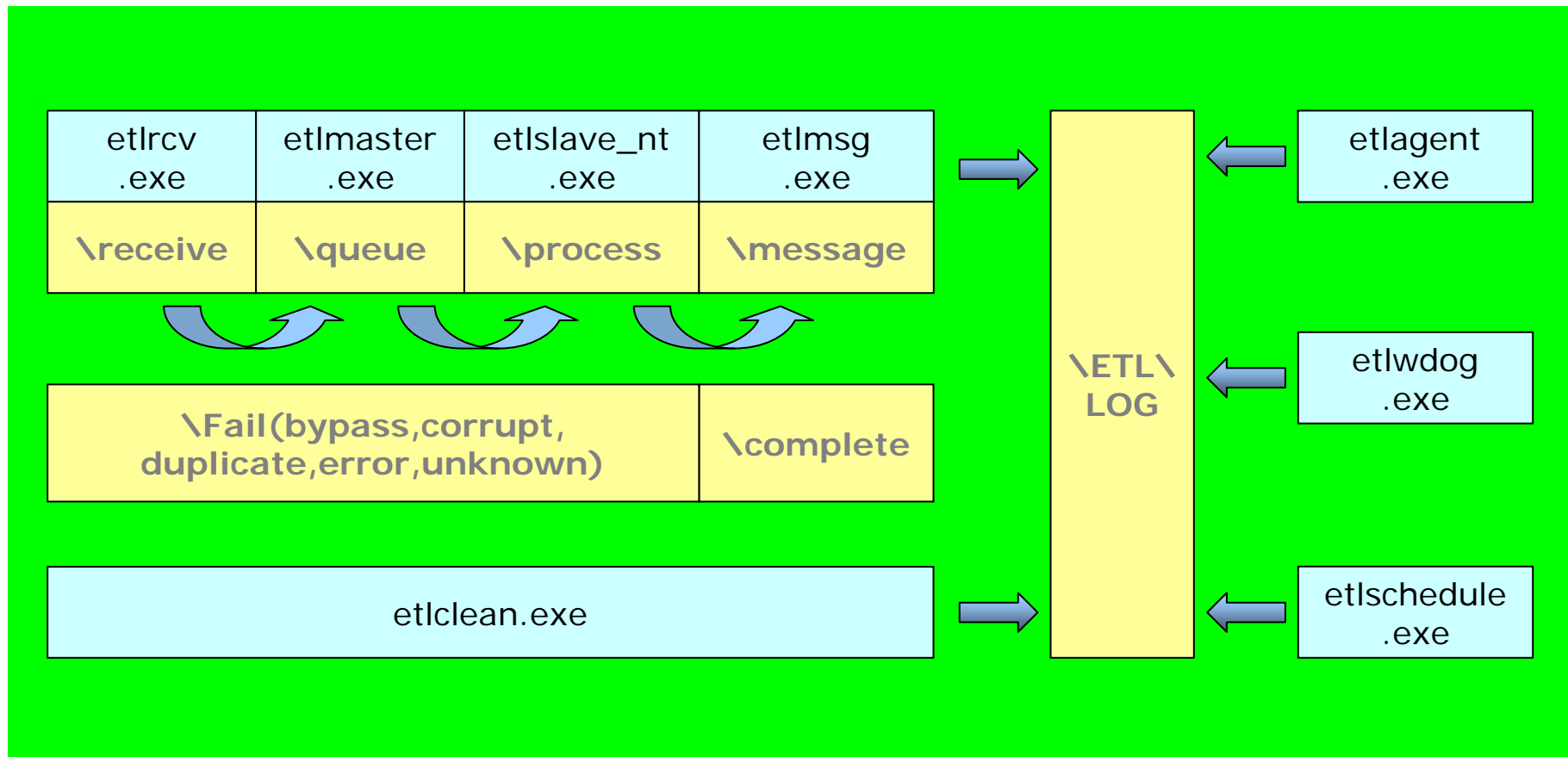


ETL Automation目录结构

D:\ETL (主目录)

---\APP	存放作业的作业脚本
---\DATA	
-----\complete	存放已成功执行过的数据. 以系统名称以及日期来建立子目录
-----\fail	
-----\bypass	存放不须执行的档案. 以系统名称以及日期来建立子目录
-----\corrupt	存放大小不符合的档案. 以系统名称以及日期来建立子目录
-----\duplicate	存放重复接收的档案. 以系统名称以及日期来建立子目录
-----\error	存放执行作业过程中产生错误的档案. 以系统名称以及日期来建立子目录
-----\unknown	存放未定义在 ETL Automation 机制中的档案. 以日期来建立子目录
-----\message	存放要发送讯息通知的控制文件
-----\process	存放正在执行中的作业所使用的数据文件及控制档
-----\queue	存放准备要执行的作业所使用的数据文件及控制档
-----\receive	用来接收各来源系统所传送来的数据文件及控制文件
---\LOG	存放 ETL Automation 系统程序及各作业在执行时所产生的记录档案
---\bin	存放 ETL Automation 系统程序的执行文件
---\etc	存放 ETL Automation 机制的一些设定档案
---\lock	存放 ETL Automation 系统程序及各作业在执行时所产生的 lock 档案
---\tmp	

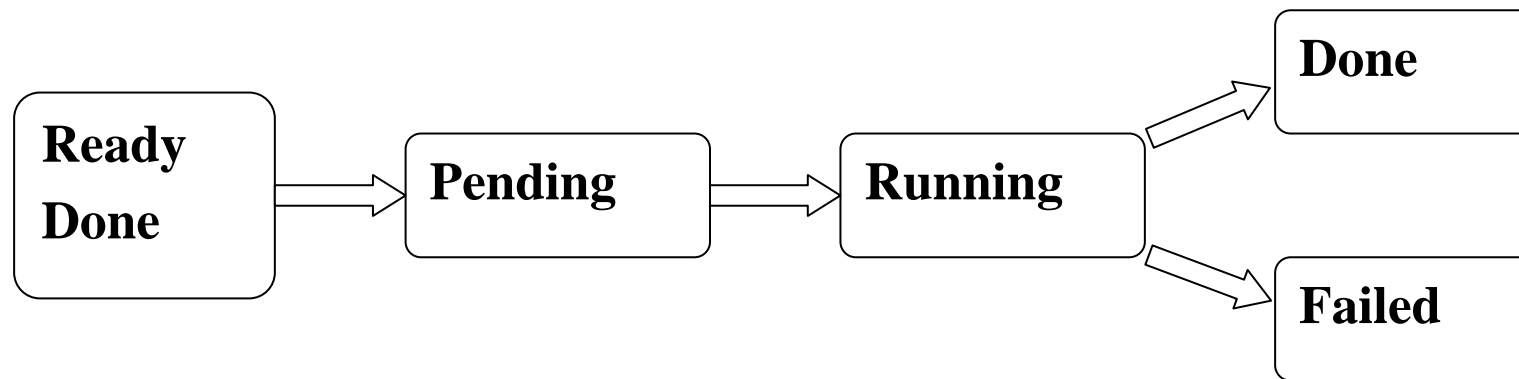
ETL Automation处理流程



ETL Automation 的特定概念

- 数据驱动
- 控制文件 - 控制作业处理的文件
 - dir.job_source_nameYYYYMMDD**
由**etlrcv** 监控并处理
 - System_Name_Convert_Job_Name_YYYYMMDD.dir**
由**etlmaster**监控并处理
- 数据文件 - 作业处理的数据文件，随控制文件一起在处理时移动
 - 如果控制文件是空文件，则表明该作业无数据文件
 - 如果控制文件是非空文件，则其中必须包含数据文件的名称、大小

ETL Automation作业的状态



ETL Automation 作业触发机制

作业由控制文件驱动触发，当满足以下一个条件时被调用：

情况1) 由 **ETLRCV** 服务监控并调用

监控 **receive** 目录下

dir.sysname_jobname_yyyymmdd

作业处于READY/Done状态时

情况2) 由 **ETLMAS** 服务监控并调用

监控 **queue** 目录下

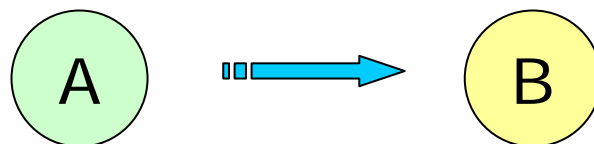
System_Name_Convert_Job_Name_YYYYMMDD.dir

作业处于PENDING状态时

作业被触发的几种方式

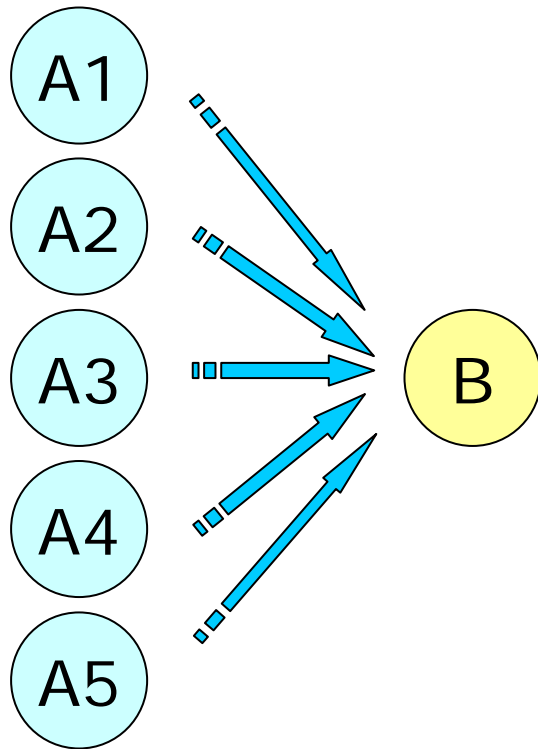
- 通过作业的触发关系 (**job stream**) 自动触发
控制文件由 **DW Automation** 产生
不能处理需要数据文件的作业
- 通过作业的群组关系 (**job group**) 自动触发
控制文件由 **DW Automation** 产生
不能处理需要数据文件的作业
- 通过脚本生成控制文件自动触发
控制文件由脚本自动产生
可以处理需要数据文件的作业
- 通过图形操作界面手工触发
控制文件由 **DW Automation** 产生
不能处理需要数据文件的作业

触发关系（Job Stream）



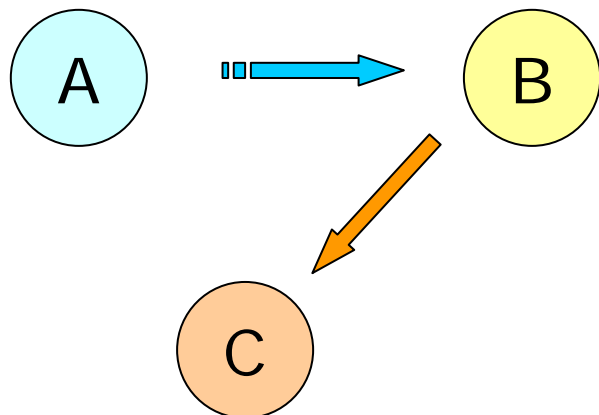
- **Job Stream**用来定义两个作业间的执行关系
- **Job A**执行成功后，自动触发执行**Job B**
- **A**为**UpStream Job**，**B**为**DownStream Job**
- 一个**UpStream Job**可以同时触发多个**DownStream Job**

群组关系 (Job Group)



- **Job Group**用于定义多个作业与一个作业间的执行关系
- **Job A1, A2, ..., A5**全部执行成功后, 自动触发执行**Job B**
- **A1, A2, ..., A5**为**Child Job**, 属于同一个**Job Group**。 **B** 为该**Job Group**的**Head Job**
- 一个**Job Group**只能定义一个**Head Job**, 但可以包含多个**Child Job**
- 同一个**Child Job**可以属于多个**Job Group**

依赖关系（Job Dependency）



- **Job Dependency**用于定义作业在执行之前，检查依赖条件是否满足
- **Job A**与**Job B**之间为触发关系，同时**Job B**依赖于**Job C**
- **Job A**执行成功后，**Job B**虽被触发但并不立即执行。如果检查发现**Job C**执行成功，**Job B**随即执行；否则，**Job B**一直等待**Job C**而不会执行
- 一个作业可以被定义为依赖于多个作业

ETL的加载脚本

- 加载脚本以**perl** 封装
- 同一个作业可以设置多个脚本（理论上可以有几百个脚本）
- 同一个作业的脚本共同完成一个作业
- 脚本执行按顺序号先后顺序
- 用户密码加密存放
- 加载脚本自动生成或采用统一模板
- 脚本按主题分系统组织存放
- 脚本名 = 转换后作业名 + 4位顺序号 + .pl
 - > 注意：顺序号'0100'专指**fastload/multiload**脚本

Email 和短信通知功能

- **Email**和短信用户和用户组的设定
- 作业可以设定通知条件：
 - > 作业执行成功
 - > 作业执行失败
 - > 作业有错误纪录
 - > 作业源未在指定时间到达
 - > 文件接受错误

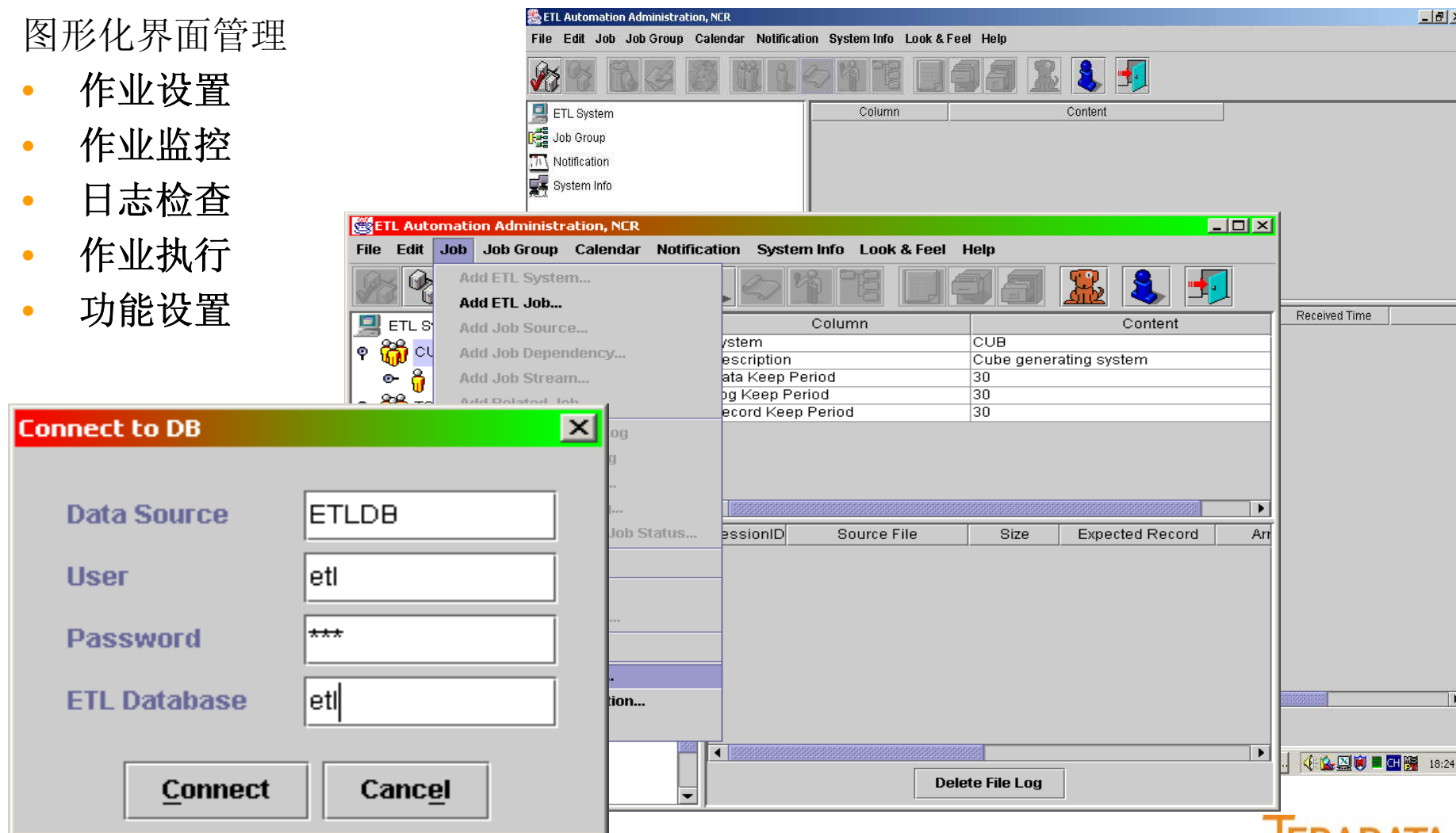


短讯 (Short Message) 通知方式必须在使用者提供短讯服务的接口时才能够整合至讯息通知的功能中, 如果使用者无法提供适当的短讯服务接口则行动短讯通知方式则无法使用.

图形化操作界面

图形化界面管理

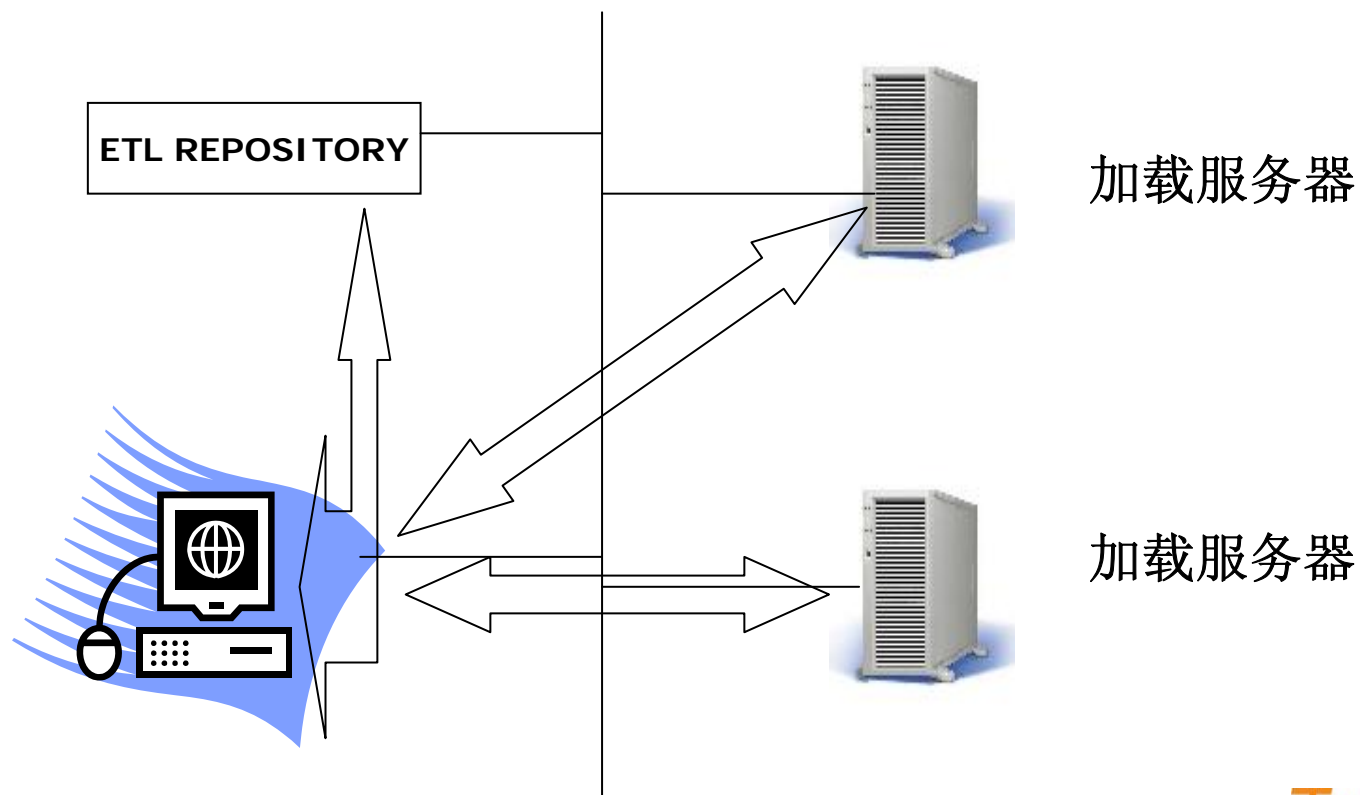
- 作业设置
- 作业监控
- 日志检查
- 作业执行
- 功能设置



图形化操作界面

作业的管理和监控

- **ETLAdmin** - 作业管理、控制、监视
- **ETLMonitor** – 作业运行监视



Questions.....



Thank You ✓