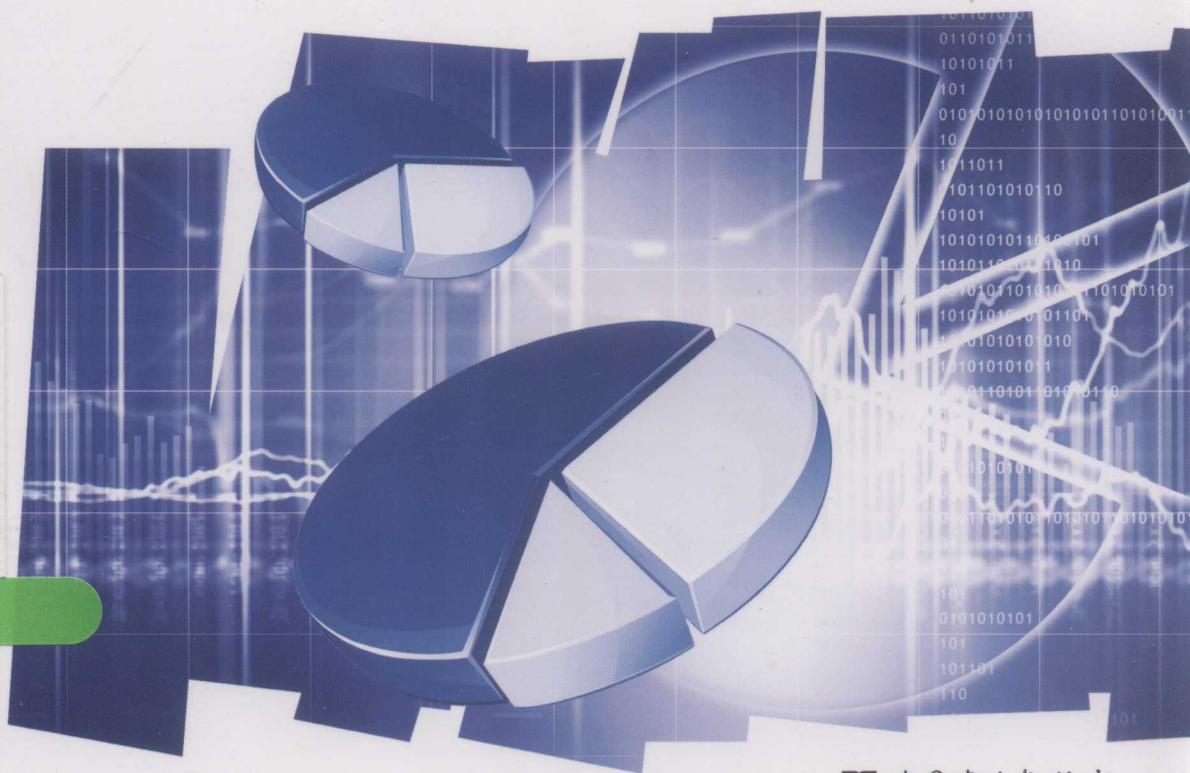


电力企业 海量实时数据中心服务能力 研究与应用

吴国诚 等 编著



中国电力出版社
CHINA ELECTRIC POWER PRESS

014032497

F407.616
12

电力企业 海量实时数据中心服务能力 研究与应用

吴国诚 等 编著



F407.616
12



北航

C1720844



中国电力出版社
CHINA ELECTRIC POWER PRESS

内 容 提 要

随着国家电网公司智能电网建设逐步深入和智能化电网信息平台建设逐步开展,各网省公司开展了实时数据中心建设,实时生产数据的应用价值在不断被挖掘与体现。实时数据平台建设贯穿数据库、数据中心、集成服务和信息展现各个层面,提升其服务能力和服务水平是实时数据平台建设中必须要考虑的问题。

实时数据中心对外提供服务功能,包括数据接入服务、数据访问服务和支撑工具服务等。本书结合网省公司海量实时/历史数据平台建设经验,对数据中心数据层优化、数据访问服务能力优化、应用性能和应用工具优化,分别展开讨论,以达到整体上提升实时数据中心服务能力目标。

全书共分六章,第一~二章介绍了数据中心定义及发展,电力企业实时数据中心,服务能力优化定义及方向等内容;第三~五章结合实际工作经验,介绍了服务能力优化中数据优化、模型与应用优化、应用工具优化等三个方面的实践总结;第六章对实时数据中心服务能力优化经验进行了总结与回顾,并对智能电网建设背景下,实时数据中心下一步的建设与服务特性进行了展望。

全书结构清晰、内容丰富,主要供电力系统管理人员和技术人员使用,也可可为其他企事业单位了解实时应用提供参考。

图书在版编目(CIP)数据

电力企业海量实时数据中心服务能力研究与应用 / 吴国诚等编著. —北京: 中国电力出版社, 2013.12

ISBN 978-7-5123-5411-1

I. ①电… II. ①吴… III. ①实时数据库—数据库系统—应用—电力工业—工业企业管理—研究 IV. ①F407.616.14

中国版本图书馆 CIP 数据核字 (2014) 第 001211 号

中国电力出版社出版、发行

(北京市东城区北京站西街 19 号 100005 <http://www.cepp.sgcc.com.cn>)

北京市同江印刷厂印刷

各地新华书店经售

*

2013 年 12 月第一版 2013 年 12 月北京第一次印刷

710 毫米×980 毫米 16 开本 8.5 印张 120 千字

印数 0001—3000 册 定价 40.00 元

敬告读者

本书封底贴有防伪标签,刮开涂层可查询真伪

本书如有印装质量问题,我社发行部负责退换

版权专有 翻印必究

副主编 董建达

李琪安 磊 陈晓杰 王勇

林明晖 蒋晓琴 吴颖嫣 鲁敏

孙晓恩 戚浩金 胡一嗔 管金胜

夏家盛 黃俊惠 錢君



前言

实时数据库在存储能力、读取速度、分析展示等性能方面的优势,近年来实时数据库行业应用快速增长。在现代工业企业中,海量实时数据库系统已经成为生产信息化的核心基础软件,在企业信息化的建设中起着至关重要的作用。目前,海量实时数据库系统广泛应用于电力、钢铁、石化、水泥、造纸等工业领域,为系统的优化运行和设备可靠性提供了坚实的数据保障和技术支撑。

海量实时数据库系统在电力企业得到应用是由电力行业生产运行的特征所决定的。电力系统的生产过程具备高度实时性,发电、输电、变电、配电、用电等电力生产环节在瞬时发生,产生了大量的实时数据。海量实时数据中心是对生产运行过程中各业务应用形成的历史/实时数据进行存储、集中、整合、共享和分析的场所,平台同时提供标准统一的访问方式,是为智能电网和 SG-ERP 各业务应用——特别是跨专业跨部门的综合业务应用在历史/实时数据层面提供技术支撑的信息基础设施。服务架构是海量实时数据中心对外提供的各种服务功能的集合,具体而言,包括数据接入服务、数据访问服务和支撑工具服务。通过这些服务,可以接入各业务应用的实时/历史数据,提供各业务应用标准、统一的访问方法,并为跨专业、跨部门的分析和辅助决策类应用提供必要的支撑。

服务能力优化是指用以提升服务能力的方法和技术。结合海量实时/历史数据库,服务能力的优化也有自身的特点,海量实时数据库建设的过程可分为两个层次,第一个层次为功能性需求,即功能架构能够满足在电力生产过程中对实时数据库的应用需求。这部分是由业务驱动的,实时数据库作为数据存储、数据共享、决策分析和统一访问的平台,满足业务的功能性需求。第二个层次

为服务能力优化，即对实时数据库本身的能力程度提出更高的要求。

本书以海量实时数据库中心服务能力改进为主线，系统阐述了数据中心建设、数据中心服务架构、服务能力优化方向等内容，涵盖了海量实时数据中心服务能力优化在数据优化、电网模型完善、组态工具升级等方面进行的有益尝试和典型经验。在编写过程中借鉴了电力行业在海量实时数据库领域研究和工程实践中取得的一些重要成果，并根据电力企业信息化和电网智能化发展的方向，对服务能力优化建设进行了拓展和展望。

本书主要供电力企业信息化建设和管理人员使用，也作为其他从事海量实时数据库系统研究及开发人员的参考书籍。在此，编者希望能与广大读者开展交流、共同学习，提高系统的建设和应用水平。

海量实时数据中心服务能力优化在日常业务需求中不断得到提升，会衍生出更多的优化方向和方法。随着国产实时数据库在智能电网建设的深入应用，必然会有大量的新需求、新技术、新应用展现，我们也需要密切跟踪海量实时数据中心的最新应用动向和研究热点，开展深层次的工作。

由于编写时间仓促，书中难免有疏漏之处，敬请批评指正。

作 者

2013.10



目 录

前言

第一章	电力企业海量实时数据中心概述	1
1.1	数据中心概述	2
1.1.1	数据中心定义	2
1.1.2	数据中心分类	3
1.1.3	数据中心发展趋势	4
1.2	电力企业数据中心建设	5
1.2.1	建设原则与目标	5
1.2.2	数据中心与服务能力	6
1.2.3	海量实时数据中心建设	7
1.2.4	实时数据中心服务架构	9

第二章	海量实时数据服务能力优化概述	12
2.1	服务能力优化定义	12
2.1.1	定义	12
2.1.2	研究步骤	13
2.2	服务能力优化研究必要性	15
2.2.1	从二次开发角度看	15
2.2.2	从服务能力的发展角度看	16
2.3	服务能力优化的方向	19
2.3.1	数据的优化方向	19

2.3.2 模型的优化方向	20
---------------	----

2.3.3 系统辅助组态工具的优化方向	21
---------------------	----

第三章 海量实时数据优化 23

3.1 “坏数据”智能过滤	23
---------------	----

3.1.1 设计思路	24
------------	----

3.1.2 技术要点	28
------------	----

3.1.3 关键算法	30
------------	----

3.2 数据测点有效性管理	35
---------------	----

3.2.1 设计思路	35
------------	----

3.2.2 技术要点	36
------------	----

3.3 数据存储压缩	40
------------	----

3.3.1 设计思路	40
------------	----

3.3.2 关键算法	41
------------	----

3.4 应用实效	43
----------	----

3.4.1 设计思路	43
------------	----

3.4.2 功能模块	45
------------	----

3.4.3 技术要点	45
------------	----

第四章 电网拓扑模型优化 63

4.1 电网全景设备模型与设备对象库	63
--------------------	----

4.1.1 设计思路	64
------------	----

4.1.2 技术要点	65
------------	----

4.1.3 关键技术	70
------------	----

4.2 图数模一体化数据接入	70
----------------	----

4.2.1 设计思路	71
------------	----

4.2.2 技术要点	73
------------	----

4.2.3 关键技术	77
------------	----

4.3	应用云端同步共享服务	78
4.3.1	总体设计思路	79
4.3.2	技术要点	80
4.3.3	关键技术	82
4.4	应用实效	82
4.4.1	设计思路	83
4.4.2	功能模块	84
4.4.3	技术要点	84

第五章

系统组态工具技术	88
5.1 应用快速开发模板	88
5.1.1 设计思路	88
5.1.2 技术要点	89
5.1.3 关键技术	92
5.2 Web 图形控件	93
5.2.1 设计思路	94
5.2.2 技术要点	96
5.2.3 关键技术	99
5.3 接口技术	99
5.3.1 设计思路	100
5.3.2 技术要点	104
5.3.3 关键技术	106
5.4 应用实效	109
5.4.1 快速开发模板应用	109
5.4.2 图形控件应用	115

第六章

结 语	120
-----	-----

缩略词	122
-----	-----

第一章

电力企业海量实时数据中心概述

企业中，数据能服务于客户、合作伙伴和员工，是企业业务活动的核心元素之一，有助于企业管理者了解业务并做出决策。同时，由于数据来源各种各样、数据随时随地产生、数据动态变化、数据之间相互关联并依存的现状，数据管理本身是一项复杂的工程。

数据集中是管理集约化、精细化的必然要求，是企业提高生产效率，提升决策水平的重要辅助工具。目前，国内外轰轰烈烈的数据大集中热潮一步一步走向深入，伴随着数据集中趋势在电子政务、企业、高校、银行、金融、保险、医药卫生、科研等领域的逐渐展开，数据中心的建设则成为数据集中趋势下的必然要求。

根据计世资讯（CCWResearch）《2010—2011 年中国企业级数据中心市场发展趋势研究报告》显示，企业级数据中心投资规模高速增长，2010 年用于数据中心投资达到 375.8 亿元，比 2009 年增长了 27.8%，企业内部海量数据尤其是非结构数据的爆炸增长，对数据的存储和处理需求日益增加，导致数据中心投资增长。数据中心作为企业 IT 的物理载体和数据平台，已经成为企业决策的核心元素，是企业管理中不可或缺的角色。由此，企业也对数据中心的可用性提出更高的要求，对其运行效率和服务能力提出新的目标。

国内最早一批数据中心是由于政府数据集中工作开始建设的，然后是金融、电信企业，进而到其他大型企业。数据中心本身作为一个复杂的 ICT 机构，其建设是一个庞大的系统工程，同时由于承担着计算、存储、应用等重要职能，



因此,它逐渐成为信息化建设的核心。当前,建设数据中心已经成为趋势,在电力、交通、保险、银行等行业的“十二五”信息化规划中,专门强调了要加强以数据中心为核心的基础设施建设。

1.1 数据中心概述

1.1.1 数据中心定义

Search Data Center.com 网站对数据中心的定义是:一个数据中心是物理或虚拟的中央数据仓储,围绕一个特定的知识团体或关于特别的商业目的来组织数据和信息,用于保存、管理、发布数据。例如美国国家气象数据中心(National Climatic Data Center, NCDC)是公开的数据中心,是世界最大的气象信息中心,一个机构可以有私有的数据中心,作为机构的专门设施进行维护和管理。

国家电网公司在实施的信息化“SG186”工程中对数据中心定义是:数据中心是企业的业务系统与数据资源进行集中、集成、共享、分析的场地、工具、流程等的有机组合。从应用层面看,包括业务系统、基于数据仓库的分析系统(数据中心的设计中不包括对业务系统的规划和建设);从数据层面看,包括操作型数据和分析型数据以及数据与数据的集成/整合流程;从基础设施层面看,包括服务器、网络、存储和整体 IT 运行维护服务。

一个完整的数据中心由支撑系统、计算设备和业务信息系统这三个逻辑部分组成。支撑系统主要包括建筑、电力设备、环境调节设备、照明设备和监控设备,这些系统是保证上层计算机设备正常、安全运转的必要条件。计算设备主要包括服务器、存储设备、网络设备、通信设备等,这些设施支撑着上层的业务信息系统。业务信息系统是为企业或公众提供特定信息服务的软件系统,信息服务的质量依赖于底层支撑系统和计算机设备的服务能力。只有整体统筹兼顾,才能保证数据中心的良好运行,为用户提供高质量、可信赖的服务。

一个数据中心的架构应该分为以下几个部分:

- (1) 应用架构：数据中心所支撑的所有应用系统以及它们之间的关系。
- (2) 数据架构：企业所有应用系统的数据构成、相互关系和存储方式，还包括数据的标准和管控手段等。
- (3) 基础架构：为上层的应用系统提供硬件支撑的平台（主要包括服务器、网络、存储等硬件设施）和一些公共的信息服务。
- (4) 安全架构：安全架构覆盖整个数据中心各个部分，包括运维、应用、数据、基础设施等。它是指提供系统软硬件方面整体安全性的所有服务和技术工具的总和。
- (5) 运维架构：运维架构用于管理执行架构和开发架构，它主要是面向企业的信息系统管理人员，为整个信息系统搭建了一个统一的管理平台，并提供了相关的管理维护工具，如系统管理平台、数据备份工具和相关的管理流程。

1.1.2 数据中心分类

各类数据中心的业务各异，其地位、规模、作用、配置和分类方法也有所不同，目前主要从以下两个维度进行分类。

一、根据数据中心服务的规模分类

数据中心按照规模划分，可以分为大、中、小型数据中心，但这也只是个相对的概念，没有严格的量化标准。在我国，从规模上来分，省、部级（或相当级别）的企业与机构所建立的数据中心一般属于大型数据中心；地市级（或相当级别）的企业与机构所建立的数据中心一般属于中型数据中心；县级（或相当级别）的企业与机构及小型企业所建立的数据中心一般属于小型数据中心。

二、根据数据中心服务的对象和范围分类

根据数据中心服务的对象和范围，常常将数据中心分为企业数据中心和互联网数据中心。

(1) 企业数据中心（Enterprise Data Center, EDC）。泛指由企业与企业所有和使用的数据中心，他们的目的是为自己的组织、合作伙伴和客户提供数据处理和数据访问的支撑。企业内部的 IT 部门或外部负责数据中心设备运行维护



的合作方。企业型数据中心是一个内部网、互联网访问、电话服务的核心。

(2) 互联网数据中心 (Internet Data Center, IDC)。由服务提供商向多个客户提供有偿的数据互联网服务 (如 Web 服务或 VPN 服务等) 的中心。互联网数据中心是一种利用电信级机房设备向用户提供专业化和标准数据存放业务及其他相关服务的中心。用户可以享受数据中心的主机托管租赁、虚拟主机等服务, 也可以租用数据中心的技术力量来搭建自己的互联网平台。国际互联网设施包括传统的电话服务商和相关的商家, 云计算数据中心属于此种类型。

1.1.3 数据中心发展趋势

随着科学技术的不断进步, 数据中心也在不断地演变和发展。从功能来看, 可以将数据中心的演变和发展分为四个阶段。数据中心经历的第一个阶段称为数据存储中心阶段。在这一阶段, 数据中心承担了数据存储和管理的功能。因此, 数据中心的主要特征仅仅是数据集中存放和管理。由于这一阶段的数据中心功能较低因此其对整体可用性需求也较低。

数据中心发展的第二阶段称为数据处理中心阶段。在这一阶段, 由于广域网和局域网技术的不断普及和应用, 数据中心已经可以承担核心计算的功能。这一阶段数据中心开始关注计算效率和运营效率。然而, 这一阶段的数据中心整体可用性仍然较低。

接着, 数据中心进入应用阶段, 需求的变化和满足成为其主要特征。随着互联网应用的广泛普及, 数据中心承担了核心计算和核心业务运营支撑功能。这一阶段的数据中心又称为“信息中心”, 人们对数据中心的可用性也有了更高的要求。

数据中心的第四阶段称为数据运营服务中心阶段。在这一阶段中, 数据承担着组织的核心运营支撑、信息资源服务、核心计算, 以及数据存储和备份等。业务运营对数据中心的要求将不仅仅是支持, 而是提供持续可靠的服务。因此, 这一阶段的数据中心必须具有高可用性。

当前数据中心处于从第三阶段向第四阶段的转型期, 传统电信企业和 IDC 企业基于数据中心进行升级, 如 AT&T、NTT 等。多家领先互联网企业则采用

新技术建设大规模的新型数据中心，一方面满足自身业务发展需要，同时也为第三方提供 IaaS、PaaS 等新型网络服务。目前多国政府已经十分重视网络数据与信息资源，将其看成影响国家科技创新和产业发展的战略性资源和核心竞争力，支持海量数据存储和处理的数据中心以及相关技术被提高到国家战略层面进行部署。

1.2 电力企业数据中心建设

1.2.1 建设原则与目标

数据中心的设计是一个系统、复杂、迭代的过程。构建数据中心需要遵守一些核心设计理念：简单、灵活、可扩展、模块化、标准化、经济性。遵守这些理念可以使得数据中心的设计清晰、高效、有条理。简单的理念要求设计容易被理解和验证；灵活的理念保证数据中心能不断适应新的需求；可扩展的理念使数据中心系统和设备易于扩展，能够随着业务的增长而扩大；模块化的理念是将复杂的工程分解为若干个小规模任务，使设计工作可控而易管理；标准化的理念要求采用先进成熟的技术和设计规范，保证能够适应信息技术的发展趋势；经济性的理念要求充分考虑数据中心的应用架构和服务能力，在提升服务能力同时，确保原有系统的扩展性，充分利用原有投资。

国家电网公司根据“十一五”信息发展规划实施的“SG186”工程，构筑了由信息网络、数据交换、数据中心、应用集成、企业门户五个部分组成的一体化企业级信息集成平台，建设/完善了财务（资金）管理、营销管理、安全生产管理、协同办公、人力资源管理、物资管理、项目管理和综合管理八大业务应用，建立健全了信息化安全防护、标准规范、管理调控、评价考核、技术研究、人才队伍六个保障体系。重点建设“一个系统、二级中心、三层应用”：一个系统就是构筑一体化企业级信息系统，实现信息纵向贯通、横向集成，支撑集团化运作；二级中心就是建设国网公司总部、网省公司两级数据中心，共享数据资源，促进集约化发展；三层应用就是部署国网公司总部、网省公司、地市（县）公司三层业务应用，优化业务流程，实现精细



化管理。

结合电力企业信息化建设现状,按照“统一领导、统一规划、统一标准和统一组织实施”指导思想,提出电力企业数据中心总体框架,根据框架需求设计数据中心的数据采集体系、数据存储体系和数据服务平台。

一、电力企业数据中心建设原则

(1) 统一采集。统一规划源头数据采集,建立采集规程,解决多头采集、层层汇总造成数据不一致、上报不及时等问题。

(2) 统一标准。建立数据存储模型、交换模型等系列标准,解决由标准不一致造成的数据交换、数据汇总分析问题。

(3) 统一管理。建立覆盖数据全生命周期及各层级的数据管理体系,解决因部门管理引起的一体化应用数据共享、数据安全问题。

(4) 统一服务。提供统一数据服务平台,解决专业人员收集整理数据困难问题,使专业人员专心关注数据分析和应用。

二、数据中心的建设目标

(1) 全面建成国网公司总部和网省公司两级数据中心,逐步实现数据及业务系统的集中。

(2) 规范化基础数据及关键数据模型,通过信息系统的规划,实现内部管理和业务语言的统一,方便信息集成。

(3) 实现数据的唯一性与共享性。

(4) 结合数据中心建设,完善数据交换体系,逐步实现数据中心的数据交换和点播。

(5) 建立企业数据仓库,提供丰富的数据分析展现功能。

(6) 实现网络、硬件、存储设备、数据的统一集中,实现业务系统和管理流程、数据交换流程的统一集中。

(7) 建立统一的安全体系,保证数据及业务系统的访问安全。

1.2.2 数据中心与服务能力

电力企业数据中心从跨应用系统的视角对数据进行组织,包括对整个数据

生命周期中数据的处理、存储、转换、整合、展现等。数据中心展现的服务能力包括但不限于以下几个方面：

(1) 数据集成。数据集成不是对某个业务系统数据的简单整理，而是跨越各个应用系统，在国网信息一体化平台的基础上，构建合理的数据存储模式。数据集成主要通过各种不同数据源之间的数据传递、转换、净化、集成等功能实现。它从用户的业务需求和实际应用出发，对现有的数据资源和处理流程进行综合分析，遵循信息资源规划，通过数据层面的整理提炼，将分散在各个“信息孤岛”中的有效信息资源进行整合存储，全面支持数据共享、管理和分析决策。

(2) 数据质量管理。对支持业务需求的数据进行全面的质量管理。依据数据在数据生命周期的各个阶段的特性，建立数据质量控制机制，及时发现数据质量问题，不断改善数据的使用质量，从而提升数据的可用性，实现数据更大的商业价值。

(3) 信息发布能力。针对具体分析应用，数据中心为最终用户提供高效的数据查询和分析服务。数据中心为最终用户提供各种标准报表和灵活的图表展现服务，为不同层次管理人员制定正确决策提供信息保障；数据中心还应该具备多维分析能力，能够允许用户把一个实体的多个重要属性定义为多个维度进行深入分析，并能对不同维度值的数据进行比较，通过展现不同模式来帮助管理人员与业务人员对业务状况和未来发展做出完整、合理、准确的分析和预测，从复杂难懂的海量数据中发掘出指引公司业务发展的正确路径。

(4) 辅助工具。数据中心提供的辅助工具能够方便快捷地实现数据访问、数据展现及数据分析等功能。

1.2.3 海量实时数据中心建设

随着国家电网公司坚强智能电网建设的逐步推进，尤其是各单位输变电设备状态监测、用电信息采集、配电自动化，以及发电集团信息化等项目的试点与推广，产生了大量的实时数据，连同调度生产大区已经形成的电网运行方式、关口电量、保护、雷电等实时数据一起，沉淀生成海量的历史/实时数据。这些



海量数据是国网公司生产运行过程中的重要财富，是实现精益化管理的重要基础。在这样的背景下，公司迫切需要建立海量历史/实时数据管理平台，以满足各业务应用对实时/历史数据进行存储、整合、共享，以及统一和标准访问的需求。

海量历史/实时数据管理平台（即实时数据中心）是指对生产运行过程中各业务应用形成的历史/实时数据进行存储、集中、整合、共享和分析的场所，平台同时提供标准统一的访问方式，是为智能电网和 SG-ERP 各业务应用——特别是跨专业跨部门的综合业务应用在历史/实时数据层面提供技术支撑的信息基础设施。

海量实时数据中心功能主要包括数据接入、数据处理、数据加工、数据存储、元数据管理、平台管理、访问服务和支撑工具八大部分，如图 1-1 所示。

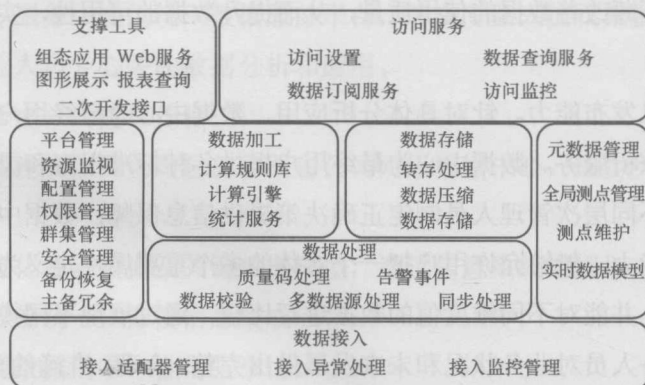


图 1-1 海量实时数据中心功能架构图

(1) 数据接入：主要负责各业务应用实时数据和历史数据的接入方式、内容、参数的定义，规约库、异常情况的处理，各接入进程的集中监控管理。具体包括接入适配器管理、接入异常处理、接入监控管理等模块。

(2) 数据处理：主要负责实时数据的整合、规范化、标准化处理，同时对数据的质量、合理性进行校验，对不合理的数据和重要的变化数据产生告警，另外，对源业务应用历史数据发生改变进行同步处理，以保证实时数据平台与业务应用源数据之间的一致。其具体包括质量码处理、数据校验、告警事件、