

一、 企业数据仓库

1、企业数据仓库

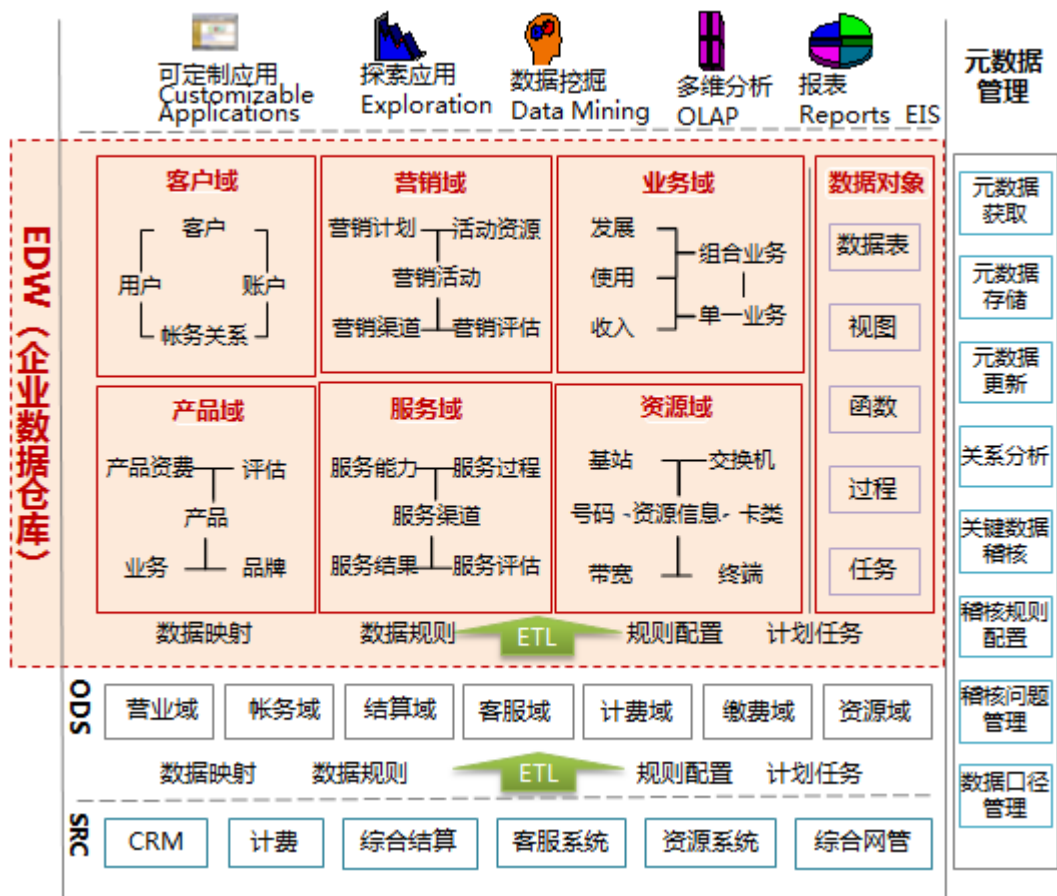


概述

企业数据平台是商业智能体系的根底，通过建立层次化的数据仓库体系结构、统一的数据模型，整合企业生产系统的数据，从而为数据分析平台和 CRM 等上层应用提供数据支撑。企业数据平台沉淀了企业多年来的生产经营数据，数据按照存储粒度分为细节级、轻度综合级、汇总级，其数据量通常非常庞大，少那么几 TB，多那么上百 TB，而且每天都有大量的生产数据被抽取、整合进入企业数据平台，并传送到数据分析平台和其他系统，以满足其数据需要。如此庞大的数据量、吞吐量和复杂的业务逻辑，对企业数据平台的存取效率、数据质量、易管理性，易维护性等均提出了很高的要求。公司在多年的数据仓库建设过程中研发出一套适用于电信行业的统一数据模型（该模型中的核心局部已经被中国联通作为数据仓库模型的标准标准）。基于统一数据模型、结合数据仓库技术、元数据技术，并对 ETL 过程进行优化处理，公司实现了电信行业海量数据仓库的标准化构建和高效运转。

企业数据仓库（EDW）是企业分析型数据的统一存储，它和运营数据存储一起构成企业的统一数据平台。企业数据仓库在企业内部建立统一的数据架构和数据模型，沉淀来自运营数据存储、生产系统的数据，采用多维分析和数据挖掘手段，细分市场和客户，为企业内各个部门、各个层面的使用者的信息查询获取、经营分析、决策支持等需求提供全方位的数据支撑效劳。企业数据仓库从根本上解决了沉淀数据分散重复、共享困难和信息孤岛问题，充分发挥数据资源价值，提高了企业在信息化建设方面的投资回报率。

系统架构



系统功能

1. 数据域划分

EDW 的数据存储按域可以分为：客户域、产品域、市场营销域、效劳域、业务域、竞争域。

客户域以三户模型为根底，客户域包含客户、客户对产品的订购等相关信息，如客户的根本属性、

衍生属性，客户的消费、欠费、缴费、行为、结算、效劳、产品订购和业务订购等。数据加工主要是以

ODS 中的营业域、计费域、帐务域、结算域数据进行整合，按照分层次、分类别、多维度进行数据汇总。

产品域以产品模型为根底，包括产品的定价、产品根底信息、产品目录、产品生命周期、产品定位、

产品营销等信息。数据加工以 ODS 的营业域、计费域、帐务域、结算域等数据进行整合，按照分层次、

分类别、多维度进行数据汇总。

市场营销域包含营销活动、营销能力、营销过程、营销结果、渠道等信息。数据来源是 CRM 系统的营销活动、营销方案、营销执行、渠道等内容，数据加工在 ODS 的营业域、帐务域、结算域等数据根底之上按照分层次、分类别、多维度数据汇总加工。

效劳域主要包含对主动效劳和被动效劳的信息的分析，包括效劳能力、效劳过程和效劳结果等。数据加工在 ODS 的客服域根底上进行主动效劳和被动效劳的数据汇总。

业务域从单一业务和组合业务的角度分析企业的经营信息，包含了各业务的开展、收入、使用等信息。数据加工在 ODS 的客户域、营业域、计费域、帐务域、结算域等数据根底上，按照业务为主线进行分层次多维度数据汇总。

竞争域包含竞争对手及其产品、营销、开展、话务等信息。数据加工在 ODS 的客户域、营业域、计费域根底上，按照对端号码归属电信运营商进行多维度、分层次数据汇总加工。

2. 数据加工（ETL）

ETL 是英文 Extract-Transform-Load 的缩写，提供数据的抽取、转换和加载功能。从源系统中抽取目标数据，按照一定的规那么进行清洗、集成、转换、汇总等处理工作，并将处理之后的数据加载到数据仓库的目标层中。为了利用目标数据仓库的并行处理能力，一般先抽取、加载，再做转换处理。

3. 数据模型支撑

企业数据仓库（EDW）是为了便于多维分析和多角度展现而将数据按特定的模式进行存储所建立起来的关系型数据库，它的数据基于 OLTP 源系统。数据仓库中的数据是细节的、集成的、面向主题的，以 OLAP 系统的分析需求为目的。星型模型聚合快、分析效率高，雪花型模型结构明确，便于与其他数据交互，所以 EDW 一般采用星型、雪花型综合架构采用面向主题方式设计模型，采用自顶向下逐层分解方法

设计企业数据仓库模型。主题域是模型针对的业务某一关注领域或关注点，同一个域内的实体具有高内聚性，不同域的实体之间具有低耦合性。主题域的引入有助于建立模型框架的整体视图。

4. 元数据管理

元数据（Metadata）是关于数据、操纵数据的进程，以及应用程序的结构、意义的描述信息，其主要目标是使数据、程序具有自描述性，使数据、程序更易于管理和维护。在企业数据平台中提供统一的元数据管理手段，对数据库仓库中的元数据进行采集、存储、挖掘，支持查询、分析、维护等操作，使技术人员和业务人员可以统一地对数据仓库中的元数据进行管理和分析，从而更有效地理解和控制企业数据平台的运行。在元数据信息的根底上，通过数据稽核及稽核问题管理，实现对数据质量的管控。

1) 元数据采集、同时

指对表、视图、过程、方法等数据库对象进行扫描，获取各数据库对象所包含的元数据，将其存储于系统中。系统自动检查数据仓库中的对象，保持元数据同时，并建立元数据的版本管理。

2) 元数据关系分析

通过分析元数据信息，获取各数据对象之间的关系。元数据关系分析重点关注元数据的血统分析和元数据之间的影响分析。通过元数据的关系分析，提供冗余发现、流程优化、故障诊断等功能。

3) 数据稽核过程

包括稽核规则配置、稽核流程调度、质量报告异常统计等。支持对各种类型、各种粒度的数据对象采用多种稽核规则方法进行质量稽核。

4) 稽核问题管理

包括问题发现管理、问题发布管理、问题处理与跟踪过程管理等，支持对数据质量的流程化管理。

5) 稽核质量查询

系统按数据仓库的层次、主题域等自动生成数据质量报告，并支持以多种方式进行查询分析，为系统管理、优化、故障诊断提供依据。

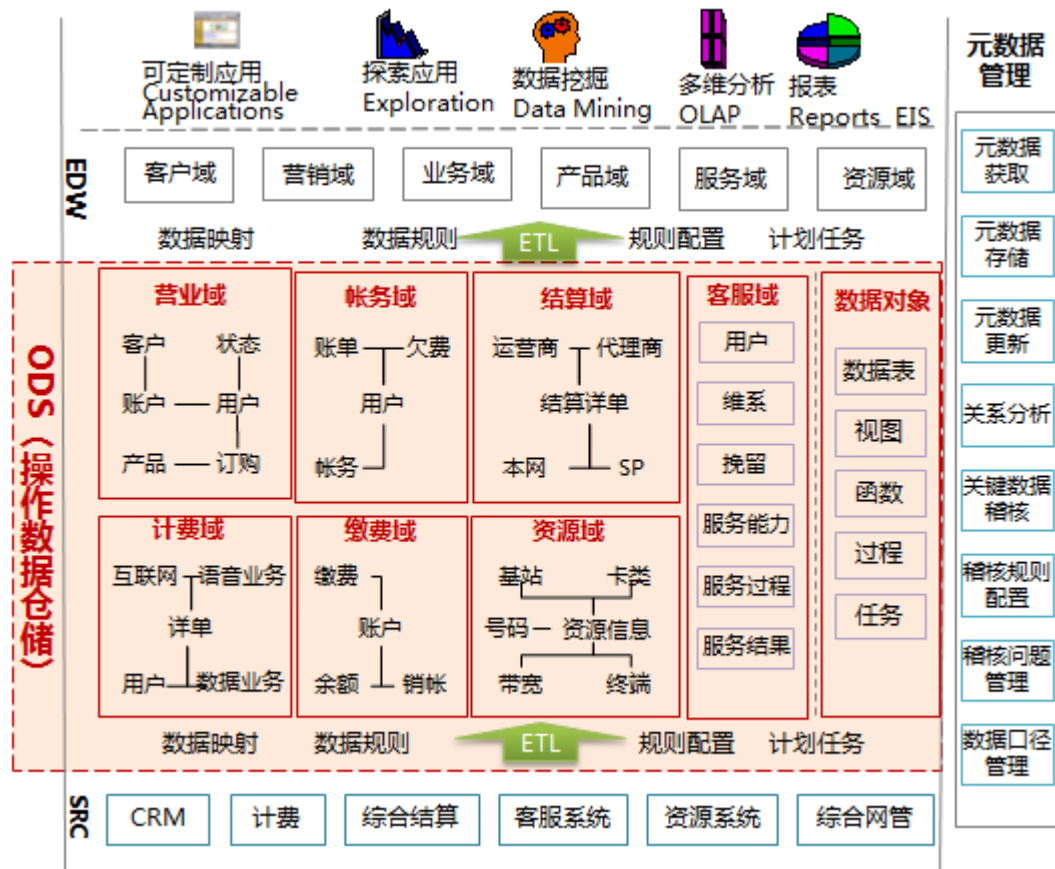
2、操作型数据仓储



概述

操作数据仓储（ODS）是企业数据平台的重要组成部分，整合企业生产系统的数据，是各生产系统的统一接口平台，为企业各个系统提供接口数据效劳。运营数据存储在企业运营数据层面起到数据总线的作用，消除了运营数据的全局冗余，降低了系统间数据交互的复杂度，降低了上层应用的开发本钱，并使数据质量得到保障。

系统架构



系统功能

1.数据域划分

ODS 的数据存储按域可以分为：营业域、计费域、帐务域、缴费域、结算域、资源域、客户域。

营业域数据来源是营业厅营业系统、CRM 系统、网上营业厅系统、电子渠道等系统。营业域包含客户资料信息和用户受理、开通等业务相关的内容，是其他数据域的根底。由于数据来源于不同厂家的各种业务系统，数据加工重点做数据的完整性、合法性、一致性校验。实体主要包含客户信息、账户信息、订购资料、状态变更流水、资费变更流水以及用户协议信息等。

计费域数据来源是不同业务的计费系统。模型核心为用户行为模型。计费域包含与用户行为相关信息，是其他数据域的根底。数据加工主要做用户通话行为的分业务、分类别、分层次汇总加工。实体主要包含用户通话行为、用户增值行为、用户上网行为等。

帐务域数据来源是不同业务帐务系统、计费帐务系统，以帐目帐务模型为根底。模型核心为分层次分类别帐目模型。帐务域描述客户的账户、帐目、账单、帐务关系、欠费等信息，是根底数据域模型。数据加工主要按照账户的帐目做分业务、分类别、分层次汇总加工。

缴费域数据来源是各种缴费系统，包括营业系统、CRM 系统、电子渠道、网上营业厅、空中充值等系统。缴费域描述帐目费用与帐户的对应关系，帐本与帐单数据的销帐关系，帐本金额的使用记录及帐单、发票的数据和格式信息。数据加工主要按照用户缴费类型、缴费方式等分层次、分类别、多维度汇总加工。

结算域数据来源是各种结算系统，包括网间结算系统、网内结算系统、合作伙伴系统、渠道管理系统等。结算域描述结算清单、结算明细账单、合作效劳方等信息。数据加工主要是对结算明细数据进行清理和按照结算类型、结算去向等分类别、分层次、多维度数据汇总加工。

资源域数据来源是网管系统、终端系统、资源管理系统等，资源可以分为网络资源、基站资源、交换机资源、带宽资源、号码资源、卡类资源、终端资源等信息。数据加工主要是对资源明细数据进行清洗、抽取、转换和按照资源域相关维度进行分层次、分类别、多维度的数据汇总加工。

客服域数据来源是客服系统、维系挽留系统、俱乐部管理、积分等系统。模型核心为客服模型。描述客服、维系挽留、营销、渠道、积分、俱乐部等效劳质量相关内容。数据加工主要对明细数据做数据一致性、完整性、合法性校验，以及对明细数据按照客服业务进行数据汇总。

2.数据整合

ODS 具备完整的数据整合流程。由 ODS 承载企业数据模型（EDM），促进企业各系统数据逻辑模型的统一。通过数据模型的统一，减少系统间复杂的转换，提高各系统的效率。ODS 中建立标准的数据编码目录，源系统数据依据标准的数据编码目录，经过整合后进入 ODS 中存储，实现企业运营数据的标准化与统一存储。数据整合主要包括数据的抽取、清洗、转换三种。

3.数据质量管控

ODS 对数据存储、整合流程等提供全周期的数据质量管控。ODS 系统对抽取到的各生产系统的数据进行属性合法性检查、数据表关联性检查、数据比对稽核等以发现数据质量问题，反响给源生产系统，源生产系统进行数据修正后再提交给 ODS 系统，形成数据质量管理的闭环流程，提升运营数据质量。

4.数据模型支撑

ODS 具有统一、稳定的模型支撑。运营数据存储（ODS）是一个面向主题的、集成的、可变的、当前的细节数据集合，用于支持企业对于即时性的、操作性的、集成的全体信息的需求。采用面向主题的模式设计思想设计统一的企业数据模型，来集成来源于多个生产系统的数据，ODS 数据又提供应生产系统和企业数据仓库使用。ETL 的主要工作是按照主数据思想集成、清洗来源于多个系统数据为统一的数据。主题域是模型针对的业务某一关注领域或关注点，同一个域内的实体具有高内聚性，不同域的实体之间具有低耦合性。主题域的引入有助于建立模型框架的整体视图。

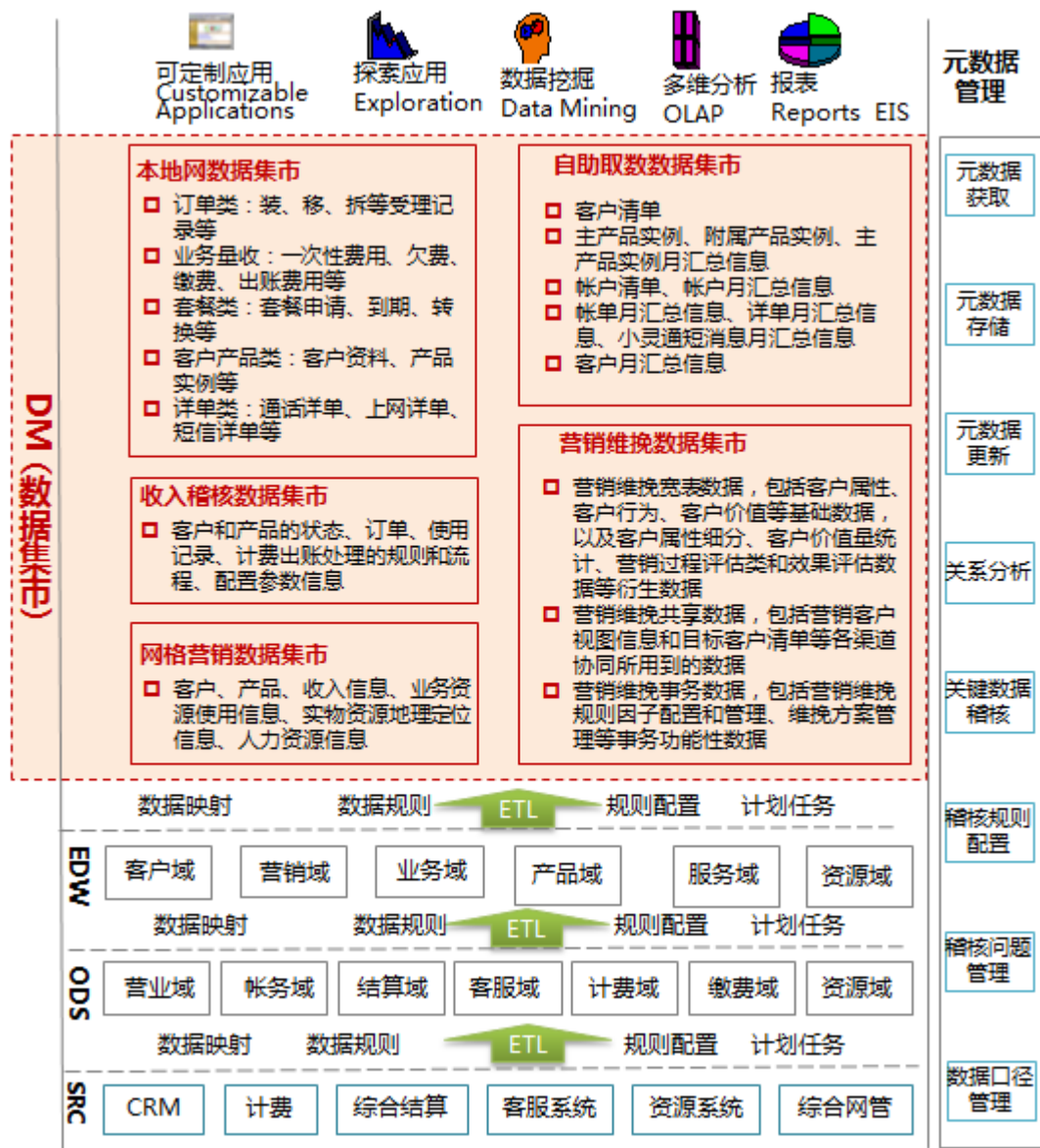
3、数据集市



概述

数据集市是将数据仓库中的数据按照不同角度进行组织和存储的过程，数据集市作为 EDW、ODS 的数据子集，主要目的是满足特定用户的应用需求，根据不同的应用需求，建立相应的业务主题。特点是数据粒度较粗，数据采取星型或雪花型结构，较少保存历史数据，便于访问分析和快速查询。数据集市具有面向特定用户群所需、合理的查询响应时间、便于扩展的特点。在实现方式上，数据集市原那么上采用逻辑部署的模式，考虑到系统性能和应用效果可以采用物理部署的模式，但是数据必须由 ODS、EDW 统一提供。

系统架构



数据集市分类

数据集市的数据按不同的应用需求，可以分为本地网数据集市、自助取数数据集市、营销维挽数据

集市、收入稽核数据集市、网络营销数据集市等。

1.本地网数据集市

本地网数据集市的数据包括两类：

1) 整合层、汇总层数据

包含整合的客户根本信息、客户行为类信息、客户消费类信息、客户接触类信息等，具体数据有：

每日提供经过整合的用于考核的各种业务量〔如装、拆、在网等〕根底数据等，每月提供业务量考核统计根底数据；每月提供经过整合的分产品、分行业的月计费根底数据等；每日提供经过整合的网间分运营商、分类型的结算数据等。

2) 接口层数据

来源于 ODS、EDW 接口层，为本地网数据集市提供生产系统原始数据。每日提供昨日 CRM 的客户资料、产品资料、订单资料、营销活动资料等增量数据，每月提供 CRM 全量数据一份；每日提供昨日计费账务系统的话单增量数据、日欠费全量数据，每月提供全量月计费帐单数据等；每日提供结算系统的分运营商、分结算类型的各类结算增量数据；每日提供资源管理系统、效劳开通系统中用户资源与业务开通等增量根底数据，每月提供对应全量数据。

2.自助取数数据集市

自助取数主要满足对客户、产品、帐户三户信息及详单级的取数，数据内容涵盖客户信息、产品、附属产品根本信息、账单、详单汇总信息等。

按月汇总表：包括帐单月汇总信息、详单月汇总信息、小灵通短消息月汇总信息、主产品实例月汇总信息、帐户月汇总信息和客户月汇总信息等。

按日更新表：主产品、附属产品实例、客户的根本信息等。

3.营销维挽数据集市

营销维挽数据集市包含营销维挽宽表数据、营销维挽共享数据和营销维挽事务数据三局部。

营销维挽宽表数据是经过一定的 ETL 处理流程，通过一定粒度汇总的，面向主题的数据集，包括客户属性、客户拥有、客户行为、客户价值、客户效劳类数据、客户营销接触类数据、客户离网、过网行为、移动增值业务、客户历史营销信息等根底类数据，以及基于根底数据生成的客户属性细分、客户拥有量统计、客户行为量统计、客户价值量统计、移动增值业务量统计、增值业务量统计、营销过程评估类和效果评估数据等衍生类数据。

营销维挽共享数据包括营销客户视图信息和目标客户清单等各渠道协同所用到的数据。其中客户营销视图信息数据主要包括：客户属性信息、客户拥有信息、客户行为信息、客户价值类信息、客户交互类信息、客户离网类信息、移动业务专有类信息、移动业务增值信息等多种客户相关信息；客户名单主要包括对应维挽方案的目标客户清单，包含客户编码，客户名称等信息；营销维挽方案包括方案类型、渠道标识、执行人等信息；客户接触信息主要包括接触时间、接触渠道、接触人员等信息。

营销维挽事务数据那么包括营销维挽规那么因子配置和管理、维挽方案管理等事务功能性数据。其中维挽规那么因子信息，包括规那么因子的类型、用途、维挽规那么类型、维挽规那么用途、紧急程度、可见范围、生失效时间等；营销维挽方案包括方案类型、营销维挽方案编码、营销维挽方案类型、渠道标识、执行人等信息。

4.收入稽核数据集市

收入稽核数据集市以支撑收入保障为目的，围绕收入稽核点，抽取某个时间点所有涉及收入稽核范围的相关系统的数据，并开展一致性比对和分析，从而发现收入流失问题。

收入稽核数据集市的数据内容包括客户和产品的状态、订单、使用记录、计费出账处理的规那么和流程、配置参数信息等。数据频度根本以月为单位。

5. 网络营销数据集市

网络营销数据集市的目的是支撑以网格为最小单元进行市场营销、指标分配、绩效考核等工作任务的执行，网络营销数据集市整合了客户、产品、收入信息、业务资源使用信息、实物资源地理定位信息、人力资源信息。

二、 数据分析平台

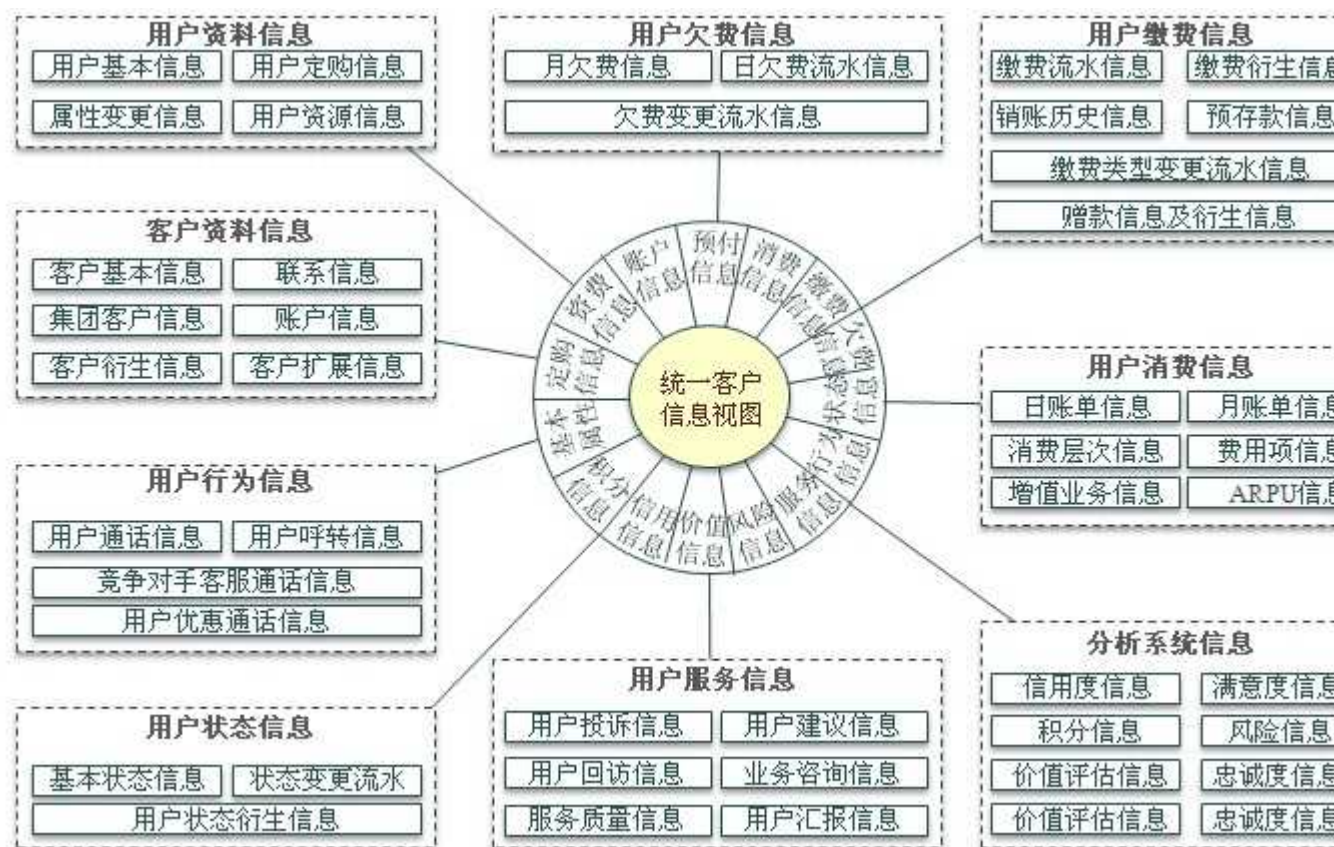
1、 客户统一视图



概述

客户统一视图是一个企业对客户的了解，是一个完整的 360 度的视图，而不是财务部门看到的仅仅是财务、销售部门看到的仅仅是销售时机、市场部门看到的是客户活跃度等等，客户统一视图中可以看到以上的所有角度，这就需要把企业数据整合成以客户为中心的数据集合。

系统架构图



功能描述

以客户为中心，聚集各个业务系统中的各类客户信息，如根本资料、帐务信息、历史信息、效劳信息等。经过分解、合并、审核、修正等环节的处理，实现客户信息的统一视图。主要包括公众客户视图和集团客户视图。

1、公众客户视图提供公众客户（个人客户等）360度综合的信息展示，本视图提供营销中心所关注的公众客户的根本信息和状态指标、消费特征、支付特征、接触记录、投诉咨询、以往联系方式等综合信息的展示；包含本公众客户的"固话"、"互联网"、"无线市话"等月消费指标。

(1)、客户主视图根本信息：包含"在网时长"、"客户状态"、"当前产品"、"客户地址"、"ARPU"、拥有电话、无线市话和增值效劳等。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/918106132063006051>