



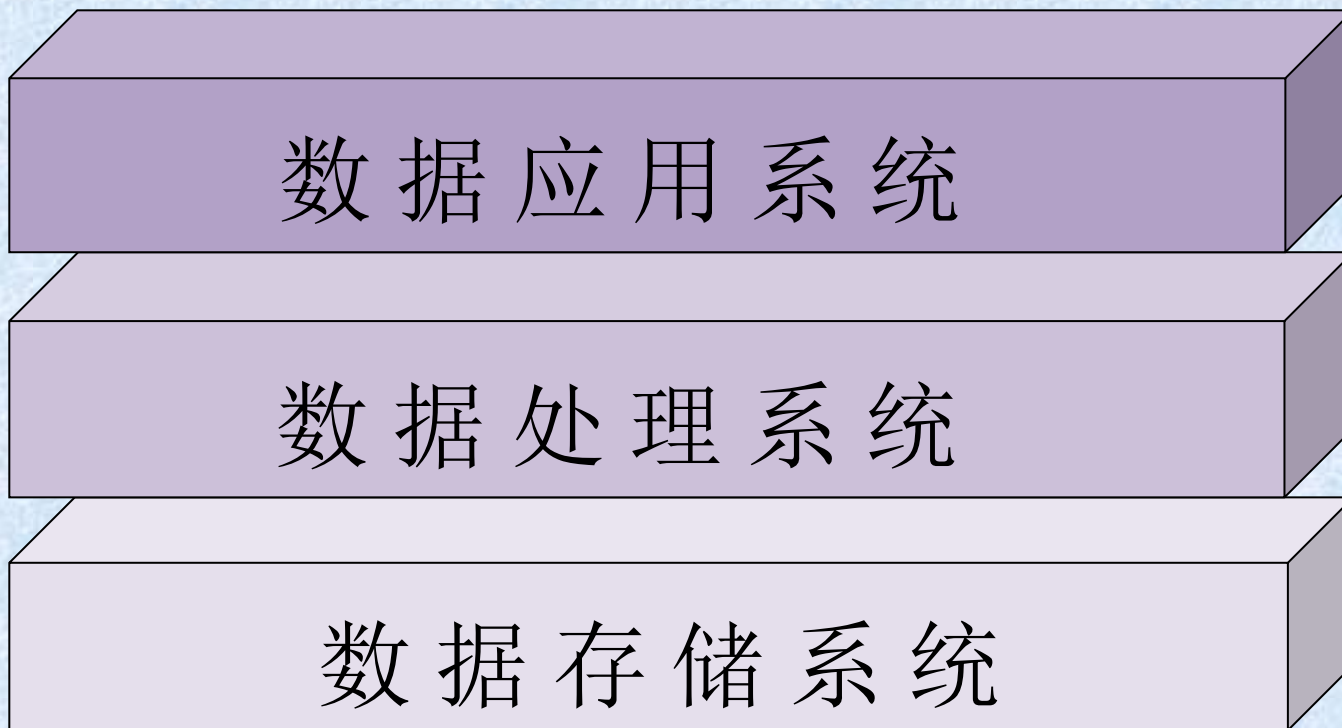
Lecture 2 大数据计算体系

- 计算总体架构
- 计算模式与平台



2.1 计算总体架构

大数据计算体系可归纳为三个基本层次：数据存储系统、数据处理系统、数据应用系统。





数据存储系统

包括数据采集层；数据清洗、抽取与建模；数据存储架构；数据统一接口等。

数据存储
系统

分布式数据库/数据仓库

分布式文件系统

数据采集与建模



数据存储系统

- 数据清洗、抽取与建模（将各种类型的结构化、非结构化、异构数据转化为标准存储格式，键值对结构，哈希表（Hash Table）检索）数据，并定义数据属性及值域）
- 数据存储架构（集中式/分布式文件系统、关系型数据库/分布式数据库、行存储数据结构/列存储数据结构
- 数据仓库与数据服务
- 统一数据接口(Unified Data Access Interface)



数据建模

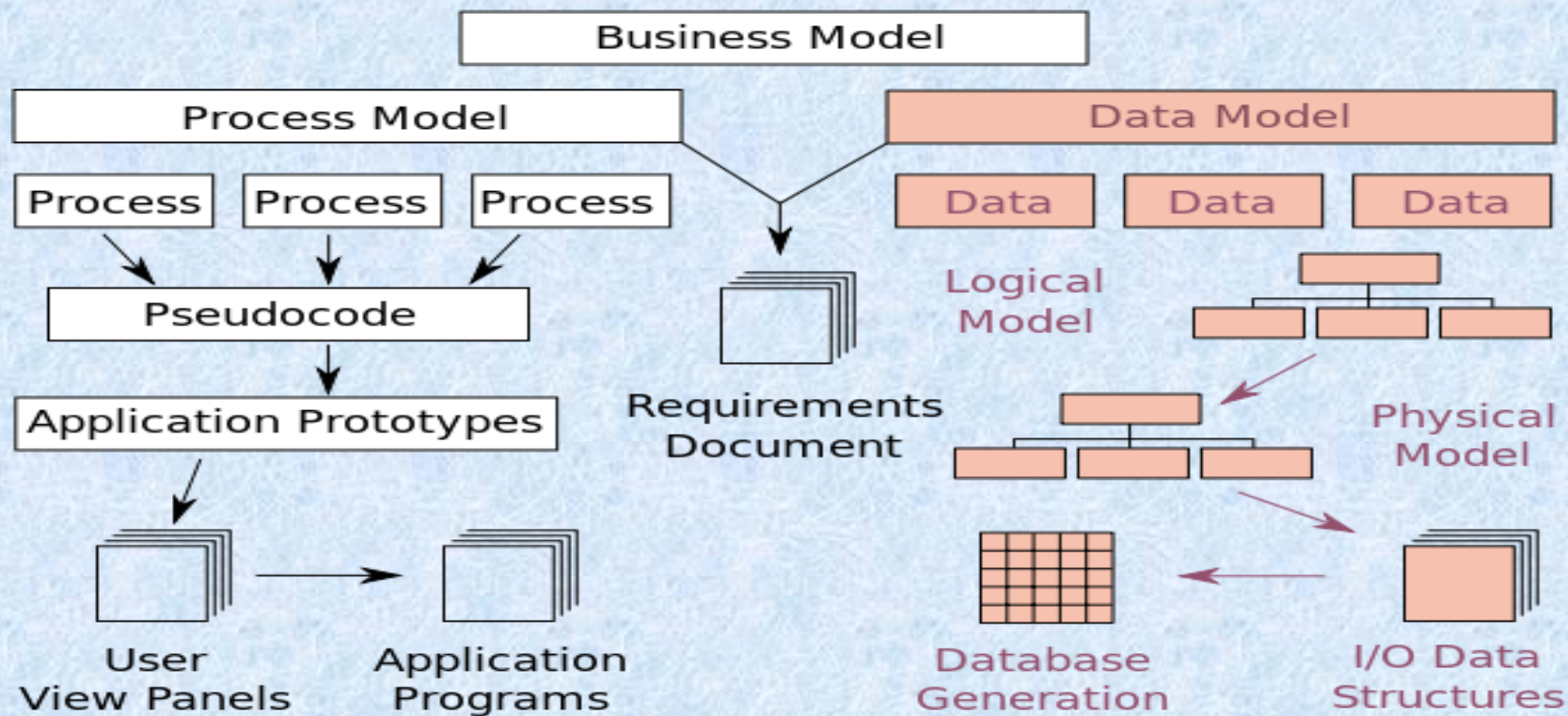
数据模型定义为三个层次：概念模型 (conceptual model)，逻辑模型 (logic model)，物理模型 (physical model)。

- 概念模型主要基于用户的数据功能需求产生，通过与客户的交流获得对客户业务要素、功能和关联关系的理解，从而定义出该业务领域内对应于上述业务要素和功能的实体类 (**entity class**)。
- 逻辑模型则给出更多的数据实体细节，包括主键、外键、属性、索引、关系、约束、甚至是视图，以数据表、数据列、值域、面向对象类(object-oriented class)、XML标签等形式来描述。
- 物理模型（有时又称为存储模型）则是考虑数据的存储实现方式，包括数据拆分(partition)、数据表空间、数据集成。



数据建模

Business Model Integration





数据处理系统

包括针对不同类型数据的计算模型，如针对海量数据的MapReduce批处理模型、针对动态数据流的流计算（Stream Computing）模型、针对结构化数据的大规模并发处理（MPP）模型、基于物理大内存的内存计算（In-memory Computing）模型；针对机器学习算法的数据流图（Data Flow Graph）模型；各类分析算法实现，及提供各种开发工具包和运行环境的计算平台，如Hadoop, Spark, Storm等。

数据处理
系统

计算引擎

计算平台

计算模型与算法



数据应用系统

基于上述计算架构和处理平台提供各行业各领域的大数据应用技术解决方案。目前，互联网、电子商务、电子政务、金融、电信、医疗卫生等行业是大数据应用最热门的领域，而制造业、教育、能源、环保、智慧交通则是大数据技术即将或已经开始拓展的行业。

数据应用
系统

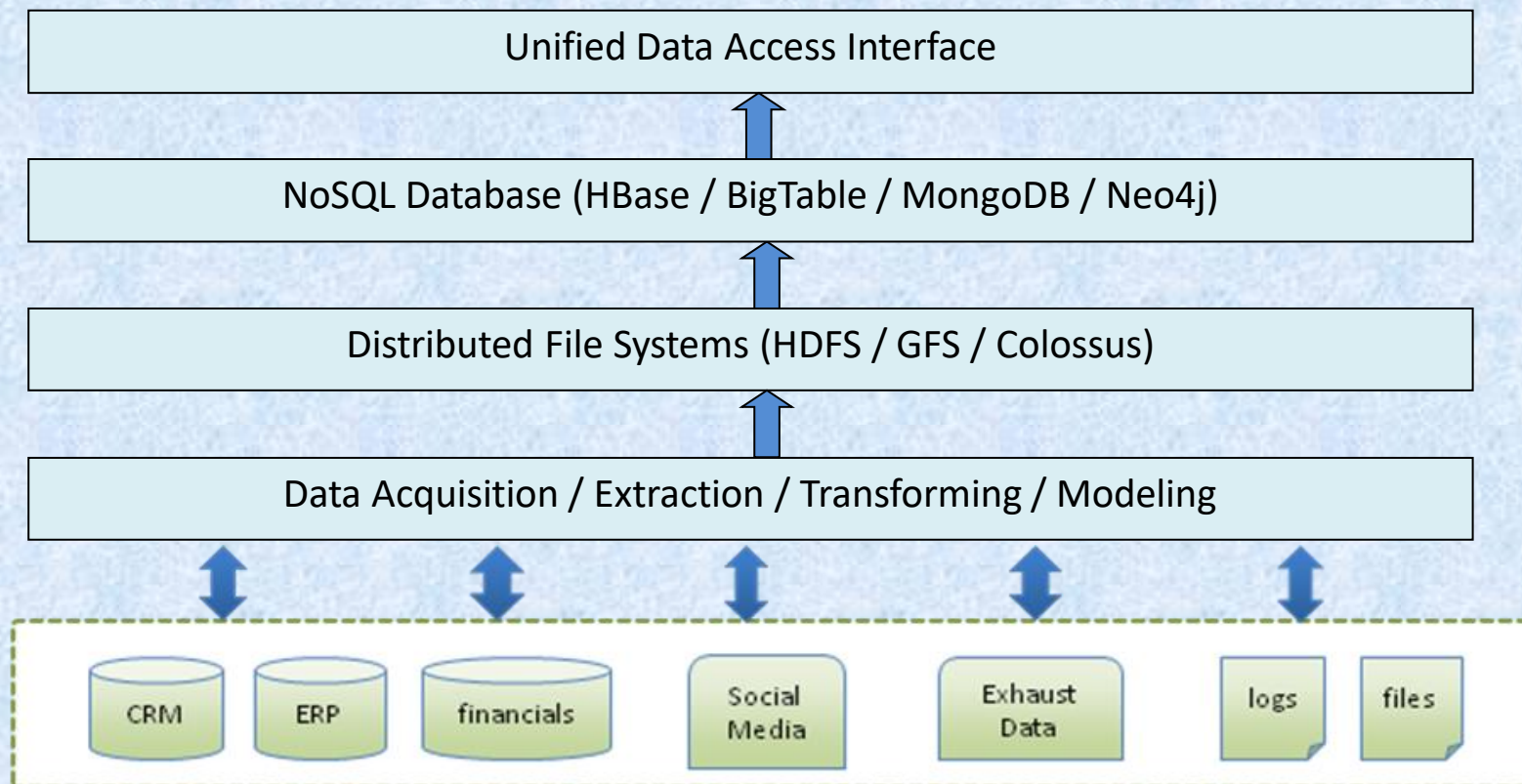
各类大数据应用

数据产品与数据服务

数据可视化

数据存储架构

在存储结构中：数据库提供了数据的逻辑存储结构；分布式文件系统提供了数据的物理存储结构。



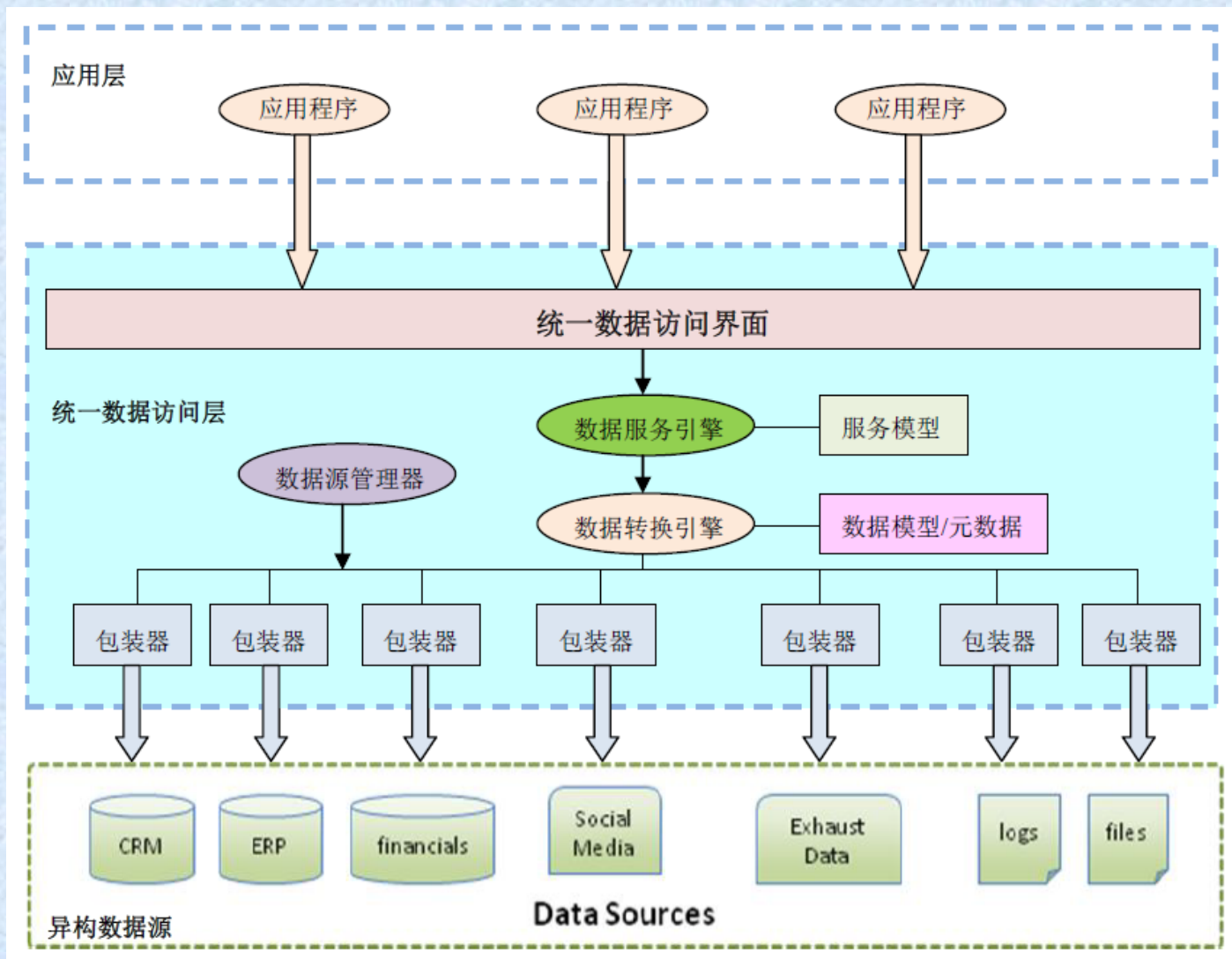


统一数据访问接口

定义：基于统一数据接口用于支持分布式环境中对跨平台异构数据库访问的数据访问层（DAL）

功能：

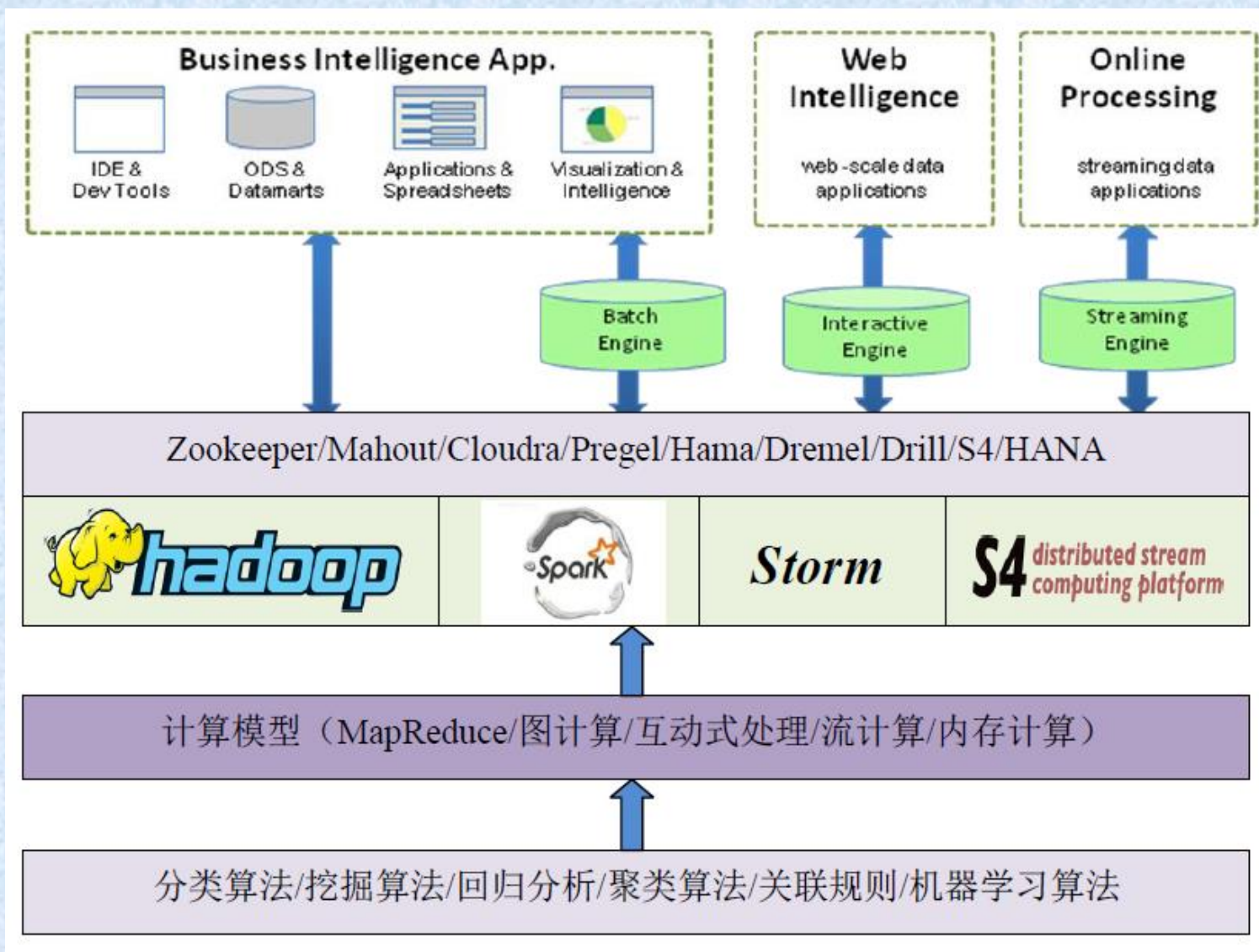
- 统一的数据展示、存储和管理
- 访问接口与实现代码分离的原则，底层数据库连接的更改不影响统一数据访问接口
- 屏蔽了数据源的差异和数据库操作细节，使得应用层专注于数据应用
- 提供一个统一的访问界面和一种统一的查询语言





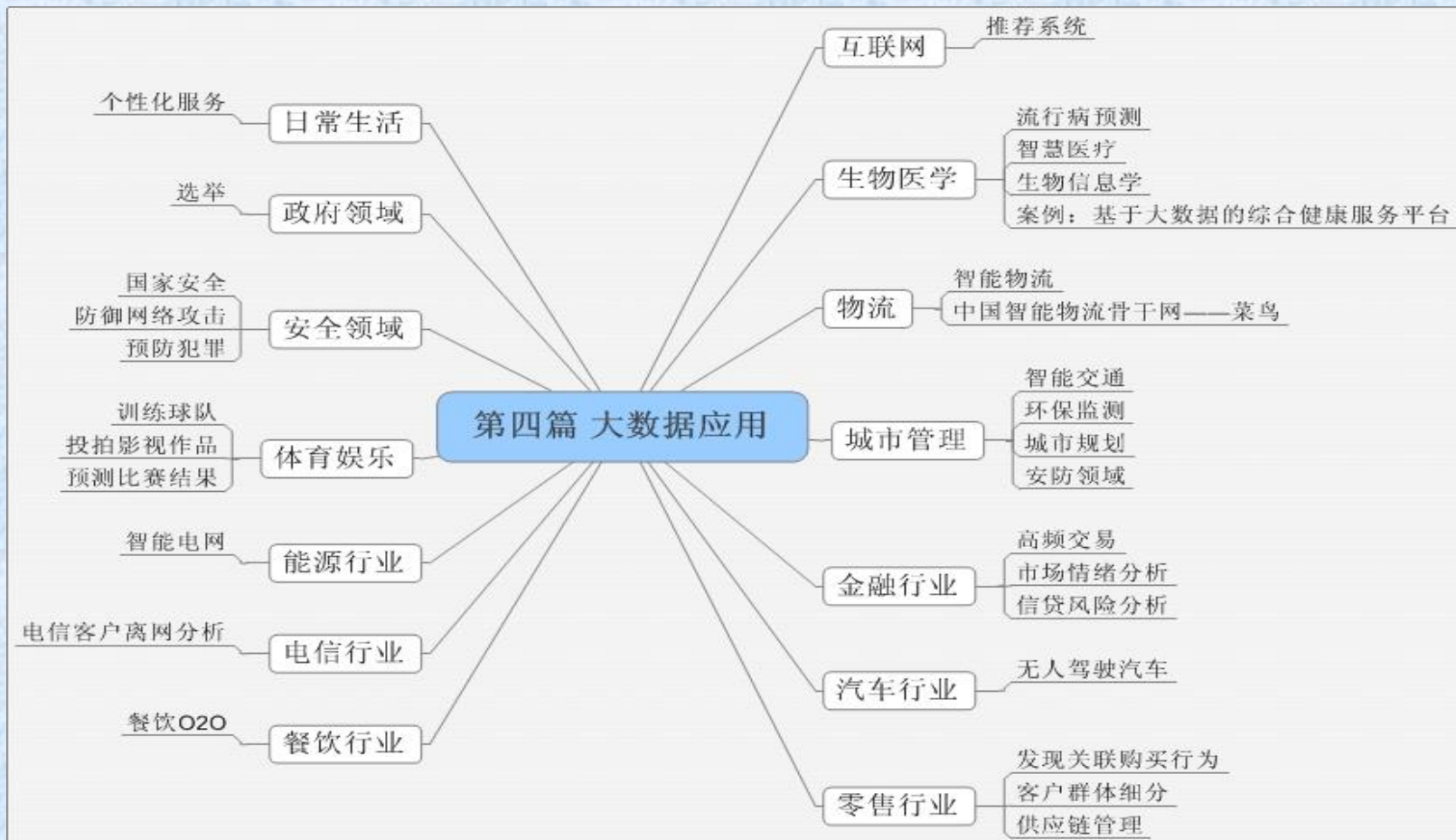
数据处理系统

- 各类算法实现（回归分析、聚合算法、关联规则算法、决策树算法、贝叶斯分析等）
- 计算模型（MapReduce批处理、流计算、交互式处理、内存计算、图并行计算、MPP模型、针对机器学习算法的数据流图（Data Flow Graph））
- 提供开发运行环境的计算平台与计算架构（Google, Hadoop, Spark, Storm, Cloudera, Impala等）
- 针对领域的计算引擎、核心框架库、智能计算系统等





数据应用系统



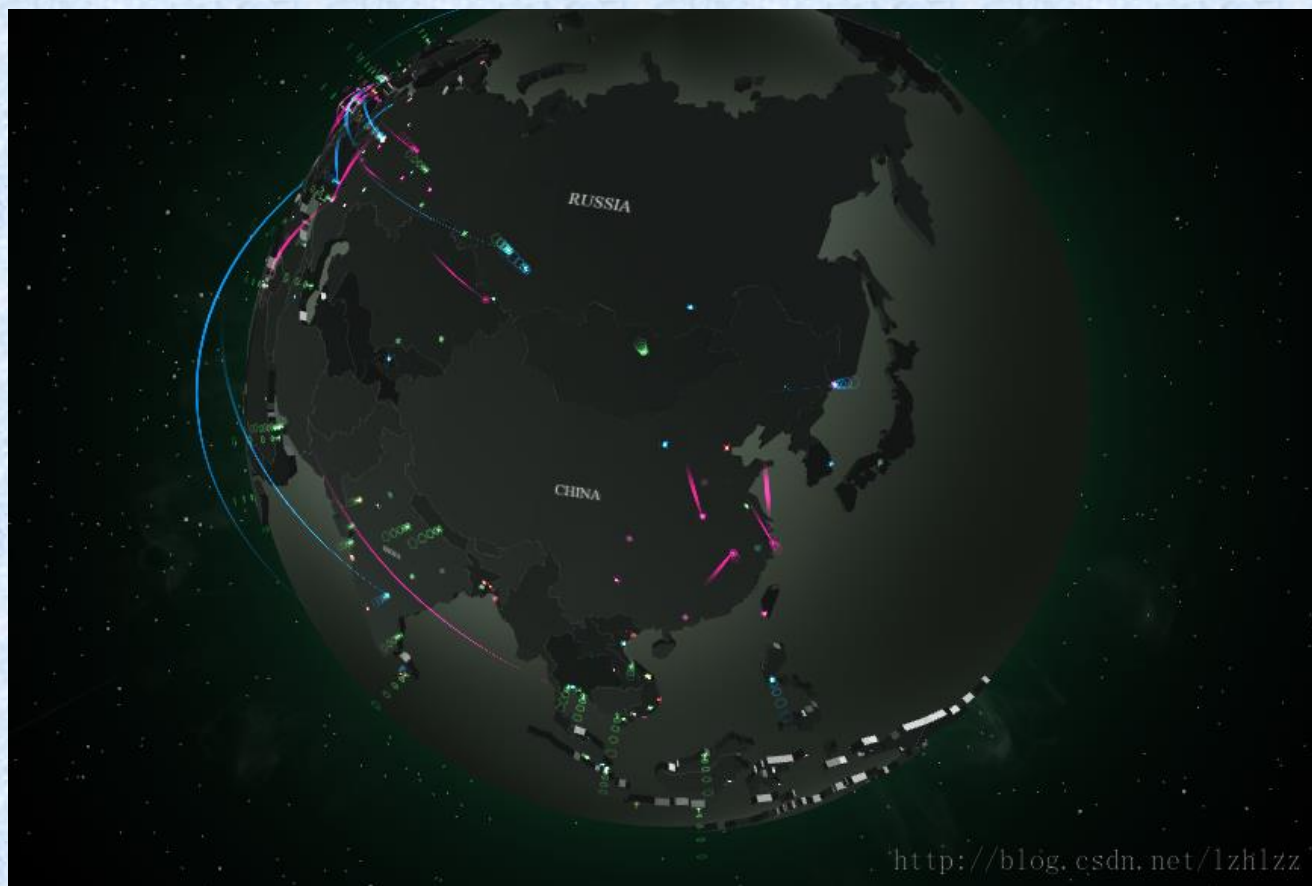


数据可视化：Facebook全球访问热点图 - 蓝色晶莹



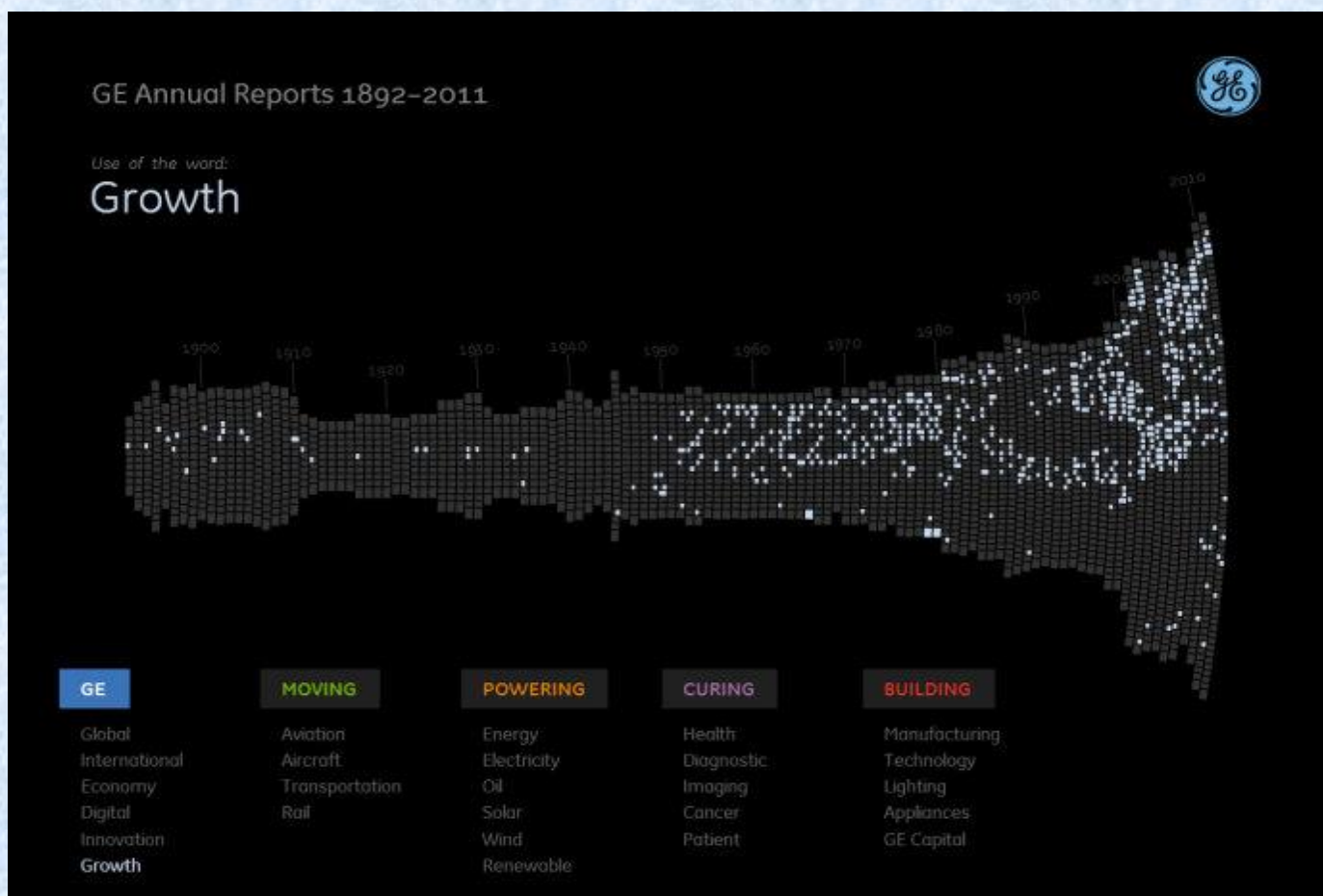


数据可视化：黑客在行动！ – 全球黑客攻击一览





数据可视化：通用电气（GE）的120年发展史浓缩





2.2 计算模式与平台

计算模式	代表产品	存储体系	计算模型	计算平台	关键技术
批处理	MapReduce	GFS, HDFS, NoSQL	MapReduce	Hadoop, Azure, InfoSphere	HDFS, Hive, Zookeeper, Mahout, Pig, Yarn
图计算	Pregel, Hama, GraphLab	GFS, HDFS, NoSQL	BSP	Google, Hadoop	Superstep, 图分 割, 数据融汇
交互式计算	Dremel, Drill, PowerDrill	GFS, HDFS, NoSQL	MapReduce + 算法	Google, Hadop	列存储结构、内存 驻存, Hash 表
流计算	Storm, S4	GFS, HDFS	流计算模型	Storm, S4	有向非循环图 (DAG), Tuple/Bolt/Topolo gy
内存计算	Spark, HANA	集中式存 储	大内存计算	Spark, HANA	列存储格式、读写 分离、内存数据库
MPP	Greenplum	多点存储 SQL	NUMA (非一致存 储访问)	Greemplum	Shared Nothing 架 构, 数据分区与并 发计算



离线批处理

在线交互式计算

大内存计算

数据规模

PB以上

TB~PB

GB~TB

时延性

离线计算(分钟~小时)

在线分析(秒~分钟)

实时计算(秒级)

计算模型

MapReduce

Dremel

MemCloud

Pregel

Drill

HANA

HAMA

PowerDrill

Spark

系统结构

分布式体系

分布式体系

集中式/分布式

采用技术

大数据迭代循环

提高数据内存存驻率

内存一次加载

硬盘读写次数多

data locality

硬件成本高

columnar data structure



大数据计算架构

- 数据存储系统
HDFS分布式文件系统、Hadoop平台、NoSQL数据库、列存储格式与检索
- 计算处理模型
MapReduce计算模型、Hama图并行处理框架、流计算、交互式计算模型、TensorFlow数据流图
- 计算关键技术
智能算法、列存储结构 (Columnar Storage Structure) 与检索、内存驻存技术 (Memory in-site)、交互式计算、数据可视化
- 技术解决方案
商业产品技术方案: Colossus/Spanner/Pregel/PowerDrill
开源技术解决方案:
HDFS/Hbase/MapReduce/Hama/Spark/TensorFlow
- 大数据应用开发



两条技术主线：商业技术 vs. 开源技术

Google	Description	Open Source
	内存计算	Spark
<u>Pregel</u>	大规模图处理系统	Hama, <u>Giraph</u>
<u>Dremel</u> <u>PowerDrill</u>	数据集交互式分析	Drill
<u>MapReduce</u>	分布式编程模型	<u>MapReduce</u>
Protocol Buffers	大数据交换协议	Avro, Thrift
<u>BigTable</u> , Spanner	分布式数据库	<u>HBase</u> , Cassandra
GFS, <u>Clossus</u>	分布式文件系统	HDFS
Google Cluster	集群架构	<u>Hadoop Cluster</u>



计算模型与计算架构

- 计算模型：抽象结构 + 计算范式 + 算法

计算模型针对领域问题提出技术解决方案的基础模型、数据结构及算法。

- 计算架构：系统架构 + 软件设计 + 实现方法

计算架构提出基于上述模型、在特定计算平台上实现的技术方案框架（系统架构、软件架构与模块、数据流与数据接口、实现原理及方法等）。



计算模式/计算模型

- ◆ MapReduce批处理
- ◆ 图并行计算
- ◆ 交互式处理
- ◆ 流计算
- ◆ 内存计算
- ◆ 数据流图模型（Tensorflow）



计算架构/计算平台

- ◆ Hadoop/HDFS/MapReduce
- ◆ 基于BSP模型的Pregel, HAMA
- ◆ Dremel/PowerDrill, Apache Drill
- ◆ Storm, Spark Stream
- ◆ Spark内存计算, MemCloud
- ◆ Tensorflow