



数据流计算模型及其在大数 据处理中的应用

汇报人：

汇报时间：2024-01-23

目录



- 引言
- 数据流计算模型基本原理
- 大数据处理中数据流计算模型应用
- 数据流计算模型关键技术

目录



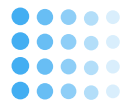
- 数据流计算模型在大数据处理中优势与挑战
- 未来发展趋势及前景展望



01

引言





背景与意义

01

大数据时代的到来

随着互联网、物联网等技术的快速发展，数据量呈现爆炸式增长，传统数据处理方法已无法满足实时性、高效性等需求。

02

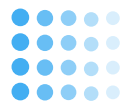
数据流计算模型的兴起

针对大数据处理的需求，数据流计算模型应运而生，为实时数据处理提供了有效的解决方案。

03

数据流计算模型的意义

数据流计算模型能够实时处理大量数据，提取有价值的信息，为决策支持、业务分析等提供实时数据支撑。



数据流计算模型概述



数据流计算模型的定义

数据流计算模型是一种基于数据流的计算模型，它将连续不断的数据流作为处理对象，通过实时计算和分析，提取出有价值的信息。



实时性

数据流计算模型能够实时处理数据流，保证数据的时效性和准确性。



连续性

数据流是连续不断的，数据流计算模型需要能够持续处理数据流，保证数据的完整性和一致性。



可扩展性

数据流计算模型需要能够处理大规模的数据流，因此需要具备良好的可扩展性，以适应不断增长的数据量。



容错性

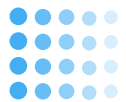
在数据流处理过程中，可能会出现各种异常情况，数据流计算模型需要具备一定的容错能力，保证系统的稳定性和可靠性。



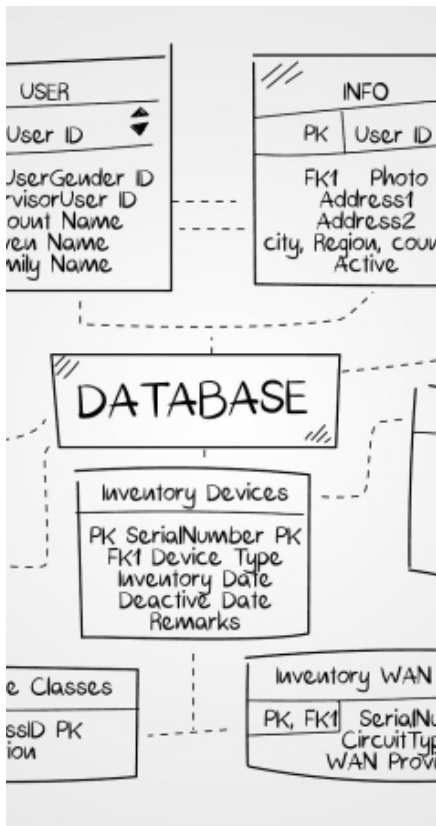
02

● 数据流计算模型基本原理 ●





数据流模型定义



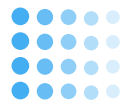
01

数据流模型是一种计算模型，用于处理连续、无界的数据流。



02

在数据流模型中，数据以流的形式持续不断地到达，计算任务需要在有限的时间内对数据流进行实时处理。



数据流模型特点

01

连续性

数据流中的数据是连续不断到达的，形成一个持续的数据序列。

02

无界性

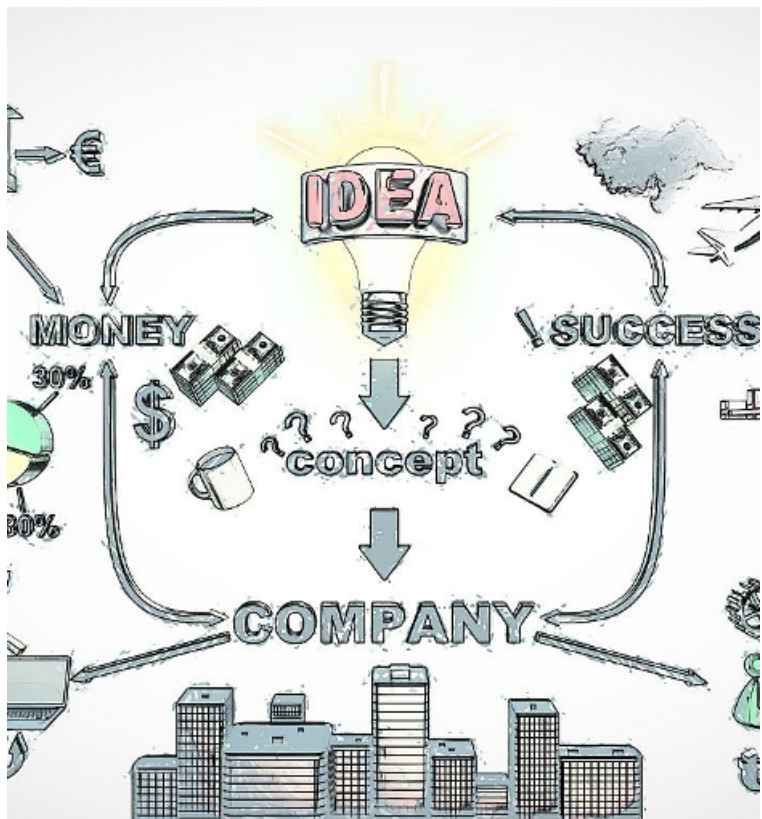
数据流中的数据量是无界的，无法预知数据的总量和结束时间。

03

实时性

数据流计算需要实时地处理到达的数据，以满足实时分析和响应的需求。

数据流计算过程



数据接收

接收来自数据源的数据流，并将其转换为内部数据格式。



数据处理

对接收到的数据进行实时处理，包括数据清洗、转换、聚合等操作。



结果输出

将处理结果输出到指定的目标系统或应用程序中，以供后续分析和使用。



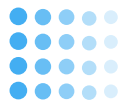
资源管理

对数据流计算过程中的资源进行管理和优化，以提高计算效率和资源利用率。



03

● **大数据处理中数据流计算** ●
模型应用



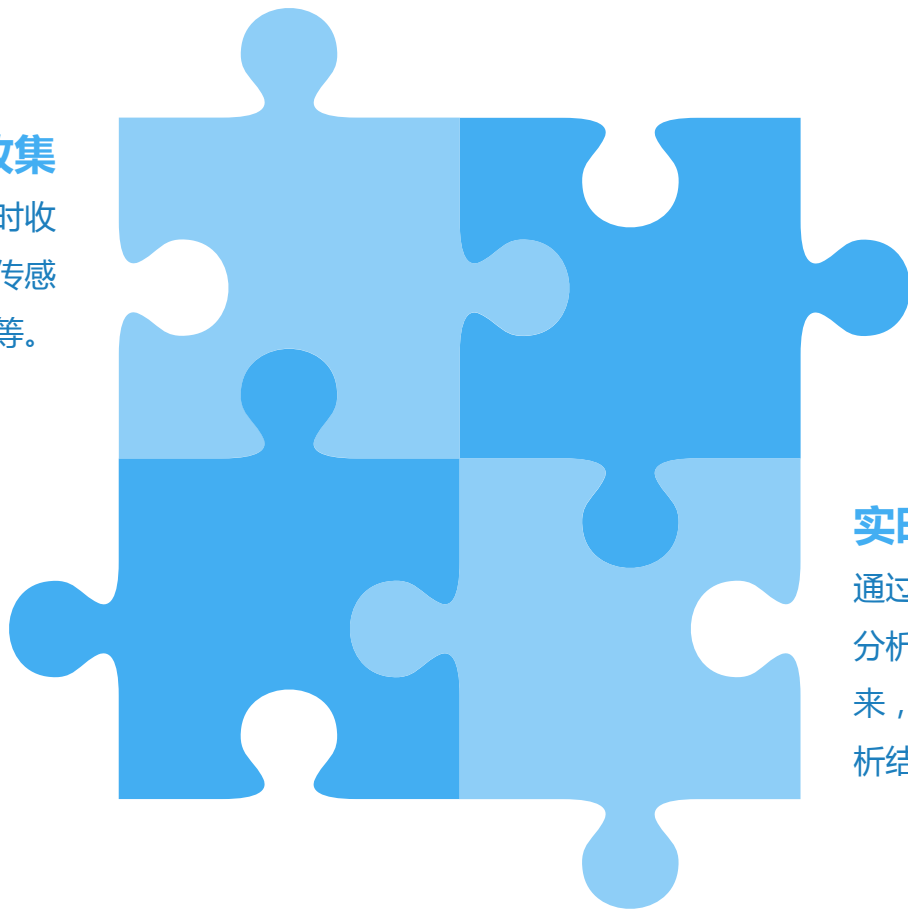
实时数据分析与处理

实时数据收集

通过数据流计算模型，可以实时收集来自各种数据源的数据，如传感器、日志文件、社交媒体等。

实时数据清洗

在数据流入处理系统之前，可以使用数据流计算模型进行实时数据清洗，去除噪声、异常值和重复数据。



实时数据分析

利用数据流计算模型的实时处理能力，可以对收集到的数据进行实时分析，提取有价值的信息和洞察。

实时数据可视化

通过数据流计算模型，可以将实时分析的结果以可视化的形式展现出来，帮助用户更好地理解数据和分析结果。

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/827025054016006122>