# 《超大规模关系型数据管理关键技术及系统》调研报告

本报告以天文学中对关系型数据库管理的需求为研究动机，系统全面地介绍了天文学领域产生数据的过程和特点，详细列出了有价值的研究问题，并给出了任务分析和面临的挑战。最后，总结了超大规模关系型数据管理的关键技术及针对天文学设计的数据库原型系统AstroDB。

# 引言

天文学是信息爆炸的起源，是最早迎接大数据挑战的领域。进入21世纪，随着天文观测技术的发展，天文学已经进入了一个信息丰富的大数据时代，天文数据正在以TB量级甚至PB量级的速度快速增长。2000年斯隆数字巡天（Sloan Digital Sky Survey）项目启动的时候，位于新墨西哥州的望远镜在短短几周内收集到的数据，已经比天文学历史上总共收集的数据还要多。到了2010年，信息档案已经高达1.4×242字节。不过，预计2019年在智利投入使用的大型视场全景巡天望远镜（Large SynopticSurvey Telescope，LSST）能在五天之内就获得同样多的信息。目前国际上已有多个国家进行了大规模的巡天项目，除了SDSS外，还有PanSTARRS（The Panoramic Survey Telescope and Rapid Response System）、 WISE (Widefield Infrared Survey Explorer)、 2MASS (Two Micron All Sky Survey)、 Gaia of the European Space Agency (ESA)、 UKIDSS（UKIRT Infrared Deep Sky Survey）、NVSS（The NRAO VLA Sky Survey）、 FIRST (Faint Images of the Radio Sky at Twenty-cm)、2df (Two-degree-Field Galaxy Redshift Survey)、我国的 LAMOST（The Large Sky Area Multi-Object Fiber Spectroscopic Telescope –郭守敬望远镜）和GWAC（The Ground-based Wide-Angle Camera array –地基广角相机阵）等，这些巡天项目每天都在产生着海量的天文数据。

# 背景知识

我国兴建中的地基广角相机阵( Ground Wide Angle Camera，GWAC) ， 由36台口径为 18cm的广角望远镜组成， 每台望远镜配备 4 k × 4 k 的电荷耦合器件(charge coupled device, CCD)探测器。整个相机阵的天区覆盖5000平方度，时间采样率为15秒。每个观测夜对固定天区目标的持继观测长达4至5小时。从观测视场的大小和观测时间的采样频度上，地面广角相机阵在时域天文观测中具有特殊的优势。巨大的数据量和高时间采样率，对数据的管理和处理提出极大的挑战。

地面广角相机阵的星表数据指标是: 星表数据每幅图像大约有1. 5 × 105 条记录， 整个相机阵在15秒内共产生 5. 4 × 106 条记录， 每晚约有2400 × 36 = 86400 幅图， 大约有2 TB的存储开销。对数据库管理系统的要求: ( 1) 快速的大数据入库能力，所有相机阵 15秒内产生的观测星表入库时间控制在 15秒以内， 每个观测夜的2TB 星表最晚完成入库时间保证在下一个观测夜开始前; ( 2) 在数据高速采集下能够完成实时分析， 面对持续不断的高密度海量星表的快速关联计算能力，即每个 CCD 每 15秒产生的星表数据与参考星表相关联形成光变曲线；(3) 以10年为设计周期，GWAC总计将产生3PB~6PB量级的超大规模星表，这就要求存放星表数据的数据库管理系统具有极强的海量数据管理能力。

海量天文数据对数据存储和管理提出了巨大的挑战。因此，研究海量天文数据的快速处理和存储机制是至关重要的。GWAC天文数据库是实现一个光变曲线实时生成并提供查询服务的数据库，设定的科学目标需满足以下两点：

1. 快速的数据入库能力。

所有相机阵15秒内产生的观测星表入库时间要控制在15秒内，每个观测夜的2TB星表最晚入库时间要保证在下一个观测夜开始观测前。

1. 实时分析高速采集的数据(包括高效的检测和动态识别天文目标)。

GWAC的主要科学目标是实时寻找所观测天区里的光学瞬变源并定位，之后制定星表目录。GWAC的巡天节奏是15秒，面对如此短时标下的持续不断观测密集星场的海量星表，数据处理系统必须有快速识别天体目标的关联计算能力及数据处理算法，即每个CCD每15秒内产生的星表数据与参考星表相关联形成光变曲线。目标是开发一个集成在数据库内部的点源证认算法，并使之具有高扩展性。

# 问题定义与基础知识

## 3.1 问题定义

定义1. 点源提取. 给定由36个广角望远镜(T1,T2,…T36)构成的相机阵Ω，每个广角望远镜Ti(i=1,2,…36)配备4 k × 4 k的CCD探测器。点源提取的过程即通过CCD探测器，把光学影像转化为数字信号，形成星表数据。

星表中的每行数据为一个星的记录。每个星的属性信息如表1所示。星表数据每幅图像大约有 1. 5×105条记录，整个相机阵在15 s 内共产生 5. 4×106条记录，每晚约有 2400×36 = 86400幅图，需占用2 TB的存储空间。

定义2. 交叉认证(cross-match). 动态生成的星表数据与模板星表进行对比认证的过程称为交叉认证。

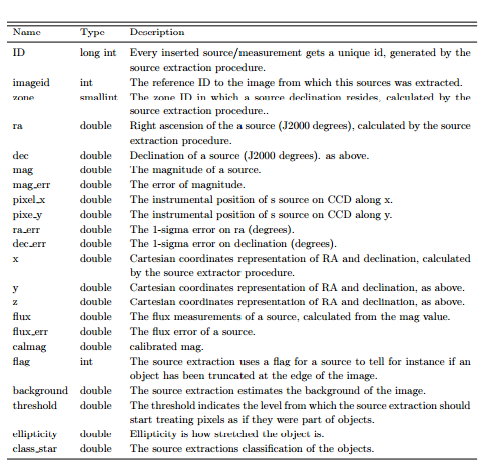
通过观测目标星表与模板星表的交叉认证，如果能与模板星表匹配的则进入时序测光通道进行光变曲线的处理并进行管理。 如果从模板星表无法找到匹配星， 则认为是瞬变源(候选体) 。交叉认证是地面广角相机阵巡天中搜索瞬变源和光变曲线生成的关键算法。交叉认证问题必须依靠有效的分区策略，通过将天区按赤纬进行条带分区，将天区划分成一个个水平条带，每个源有一个属于自己的条带属性，交叉认证前首先比较条带属性可以大大降低比较次数。条带可集成到数据库内部，以减少数据输入输出的时间损耗 。

## 相机阵数据处理流程

图1所示，地基广角相机阵数据处理流程图。通过对原始图像的基本预处理后，进行点源提取和星表天测定标，然后通过与模板星表的交叉认证，完成相对流量定标、实时动态识别天文目标和光变曲线挖掘等任务。

.

表1. 星表属性

.

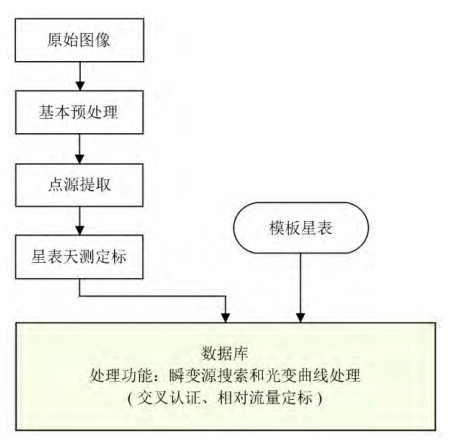


图1. 地基广角相机阵数据处理流程

设计AstroDB的总体目标是：（1）快速的大数据入库能力，所有相机阵15秒内产生的观测星表入库时间控制在15秒以内，每个观测夜的2 TB星表最晚完成入库时间保证在下一个观测夜开始前；( 2) 在数据高速采集下能够完成实时分析，面对持续不断的高密度海量星表的快速关联计算能力，即每个 CCD 每15 秒产生的星表数据与参考星表相关联形成光变曲线 。

# 功能分析

设计的数据库原型系统AstroDB需要实现的核心功能分析如下：

## 4.1 实时入库

可将所有相机阵15秒内产生的观测星表快速入库。每个观测夜的2TB星表以增量入库的思想，保证数据的实时入库。并在数据入库的低峰期（非观测夜的其余时间）进行增量数据的每日合并，以提升数据的入库能力，降低入库延迟。

## 交叉认证

建立高效的索引机制，优化数据库的连接操作，提升星表关联和交叉认证的效率。

## 高扩展性

随着每个广角望远镜采集数据的不断增加，难以支撑各用一台服务器存储和分析不同望远镜的数据，我们需设计高可靠性的分布式集群架构。针对每个广角望远镜应有不用的子集群用以存储星表数据，以确保星表数据的一致性，并实现整个系统处理能力的线性增长，实现高吞吐率和低延迟。

## 实时分析

针对观测时间小于1小时的短时标数据，实时监控和分析窗口内部的流数据，用于动态识别天文目标。

## 光变曲线挖掘

针对观测时间大于1小时的长时标数据，利用全量的历史数据预测光变曲线的波形，并判断曲线的变化周期，以便于更好地分析星体的变化特性。

# 挑战与展望

设计AstroDB面临的挑战性问题可归纳为：

## 5.1 数据库选型

针对天文数据的特点和功能需求，本报告需测试多种不同数据库并进行原型系统的选型工作。目前，拟对比的数据库系统和其相关特点如表2所示。

表2. 数据库系统对比

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Features | SciDB | OceanBase | MonetDB | MongoDB | Spark |
| 存储引擎 | Array Data Model | 内存事务存储 | Binary Association Table | 集合存储 | RDD |
| 优点 | Shared-Nothing设计；  SciDB-R | 支持ACID事务；  自动容错与负载均衡； 分块存储 | 自动建立索引；索引不占用额外的存储空间 | 查询语言强大；支持动态查询；  支持完全索引 | 便于横向扩展集群规模；迭代计算速度快 |

## 定制化天文数据库AstroDB

针对天文数据的特点，定制化天文数据库AstroDB。其中，重点考虑和设计的算子包括：

1) 增量入库算子DeltaInsert，保证数据的实时入库。

2) 范围连接算子RangeJoin，实现快速地交叉认证。

3) 引入数据挖掘算子库，实现关联分析、分类、聚类、异常点检测等功能。

此外，设计同时进行批量处理和流式处理的查询计划适配器， 通过统一的查询系统接口 ， 不仅可以实时获取当前数据流的查询结果， 而且可以从离线的历史数据中获取全量(full-history)的查询结果．

## 5.3 数据存储

长期存储(Long-term Storage)原始的图像数据，可印证后期分析工作的正确性，且对深度的图像分析工作提供了原始数据的保证。

# 结论

本报告调研了在天文学领域对关系型数据库的需求，介绍了天文学领域的相关背景知识，指出了需要解决的核心问题，并阐述了AstroDB需要实现的核心功能，以及面临的挑战性问题。

# 参考文献

[1] W. Meng. Column store for gwac: A high cadence high density large-scale astronomical light curve pipeline and distributed shared-nothing database. Publications of the Astronomical Society of the Pacific, 2016.

[2] 万萌，吴潮 等. GWAC 海量星表数据处理的数据库系统选型研究. 天文研究与技术, 2016.

注: 1. 天文台给的真实数据在wamdm249上，具体路径是：/home/wamdm/wangchunkai/AstroDB/catalog

  2. 模拟数据的生成器：<https://github.com/wan-meng/gwac_pipeline>