**HDF5（**Hierarchical Data Format 5**）调研**

**摘要：**本文主要针对HDF5的特性和特点进行阐述，第一节简单介绍一下HDF5的应用和优势，第二节介绍HDF5的文件存储格式，第三节介绍HDF5文件的数据模型，第四节介绍HFD5软件，其中包含的库、语言接口和工具等，第五节介绍HDF5中的并行方法，第六节SWMR的并发控制机制，第七节VDS虚拟数据集机制。

**1 HDF5优势及其应用**

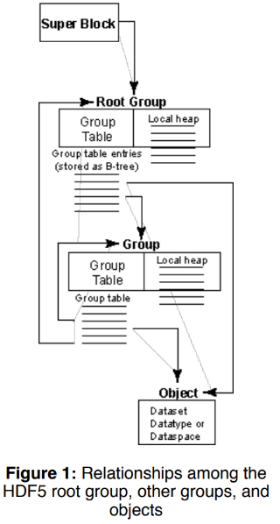
HDF5是一种独特的技术套件，可以管理极其庞大和复杂的数据集。HDF5包含用于存储HDF5数据的文件格式，用于从应用程序逻辑组织和访问HDF5数据的数据模型，以及用于处理此格式的库、语言接口和工具。其广泛应用于金融服务、天然油气、航天、汽车、医药生物、芯片制造、电子器件、政府、国防安全等部门和领域。HDF5有以下优势：

* + 可移植的文件格式，对集合中数据对象的数量或大小没有限制。
  + 多功能数据模型，可以表示非常复杂的数据对象和各种元数据。
  + 丰富的软件库，其中包含的库、工具和语言接口，用于管理，操作，查看和分析集合中数据。
  + 高效灵活的数据压缩和并行处理的机制，可实现访问时间和存储空间优化。

**2 HDF5的文件存储格式**

在最低级别，因为信息实际写入在磁盘上，HDF5文件由以下内容组成对象：

* 超级块（Superblock）
* B树节点(B-tree nodes)
* 堆积块(Heap blocks)
* 对象头部(Object headers)
* 对象数据(Object data)
* 可用空间(Free space)

  
图1.文件存储格式

HDF5使用低级对象表示高级对象，其中高水平的对象通过API为用户提供接口。如图1所示，一个组（group）由一个对象头部(object headers)和B树节点（B-tree nodes）组成。其中对象头部包含一个局部堆栈，用于存储指向组中对象链接，B树节点用于索引该链接。一个数据集（dataset）是一个对象头部（object headers）,其包含数据类型（datatype），数据空间(dataspace)，布局(layout)，过滤器(filters)，外部文件(external files)，填充值(fill value),指向其他原始数据块或B树的外部信息。

**3 HDF5逻辑组织和数据模型**

HDF5数据模型，也称为HDF5抽象数据模型或逻辑数据模型，由HDF5中的数据组织和规范构建块组成。HDF5文件可以被看作是容纳各种异构数据对象的容器。数据集可以是图像、表格、图形甚至文档，例如PDF或Excel。HDF5数据模型包含group、datasets、datatype、dataspaces、properties和attributes等对象，其中group、datasets是基础对象。如图2所示，“/”表示root group，“/A”、“/B”、“/C”、表示root group包含的group，其中group“/A”包含“/A/temp”和“/A/k”两个dataset. 下面我们详细介绍group、datasets、datatype、dataspaces、properties和attributes等对象。

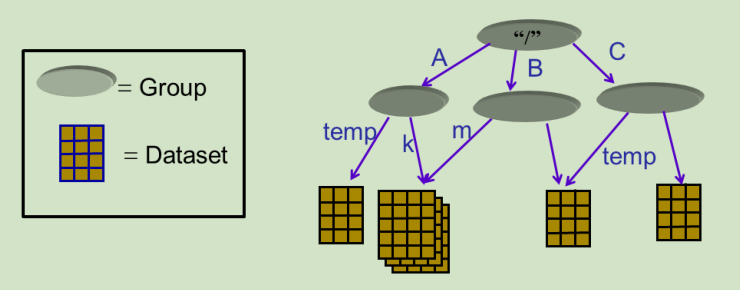


图2. HDF5数据模型整体结构

**3.1 Group**

group或link组织数据对象。每个HDF5文件都包含一个root group，该group通过包含子group或link来包含其他文件中的对象。

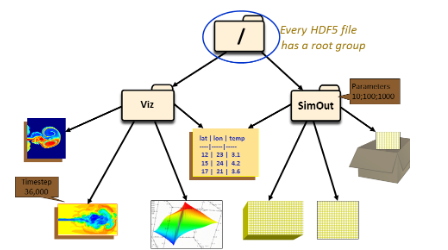


图3.group

如图3所示，HDF5文件中有两个group：Viz和SimOut。在Viz下有各种图像和一个与SimOut共享的表格。SimOut包含一个三维数组，一个二维数组和另一个HDF5文件中二维数组的link。

**3.2 dataset**

HDF5的dataset由元数据和数据本身组成，其中元数据为描述数据的数据。 如图4所示，一个数据集包含元数据和数据本身，其中元数据由Datatypes、dataspaces、properties and attributes描述。

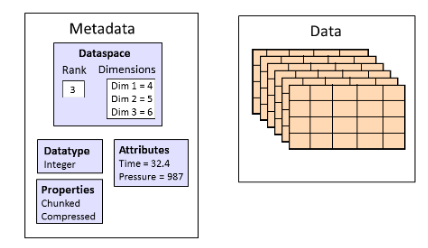


图4.dataset

**3.3 datatype**

Datatype用于描述dataset中个体的数据元素的信息。如图5中，每个元素的数据类型为32位整数。Datatype包含两种数据类型：预定义数据类型和派生数据类型。图5表示预定义的数据类型，这些数据类型由HDF5创建，例如5T\_NATIVE\_INT表示int。图6表示派生数据类型，这些是从预定义数据类型创建或派生的数据类型。常用派生数据类型的示例是包含多个字符的字符串。复合数据类型也是派生类型。复合数据类型可用于创建简单表，也可以嵌套，其中包含一个或多个其他复合数据类型。

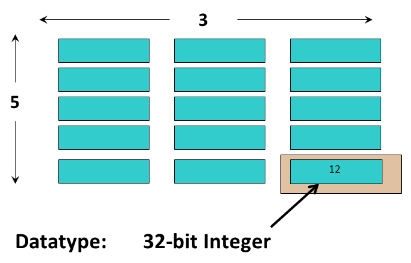


图5.预定义数据类型

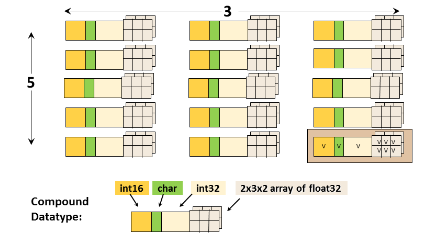


图6.派生数据类型

**3.4 dataspace**

Dataspace描述dataset中数据的元素的分布，其可由空元素、单个元素、简单数组组成。如图7所示，该dataspace是一个尺寸为5 x 3且等级为2的数组。

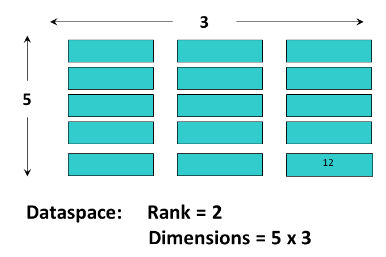


图7.dataspace

**3.5 Properties**

Properties是HDF5对象的特征。其中默认Properties可以处理最常见的需求。可以使用HDF5属性列表API修改这些默认Properties。例如图8，默认情况下，数据集的数据存储布局属性是连续的。为了获得更好的性能，我们可以将布局修改为分块或分块压缩。

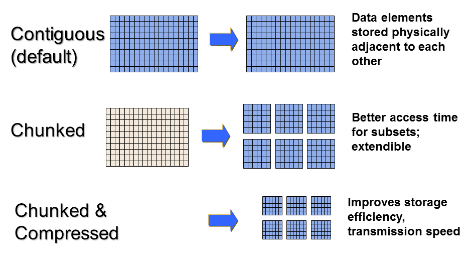


图8. Properties

**3.6 Attributes**

Attributes与HDF5对象相关联。它们有两部分：名称和值。通过打开它们所附加的对象来访问属性。通常属性的大小很小，并且包含有关其附加到的对象的用户元数据。属性看起来类似于HDF5数据集，因为它们具有数据类型和数据空间。但是，它们不支持部分I/O操作，并且无法压缩或扩展它们。

**4 HFD5软件**

**4.1 HDF5 API和库**

HDF5软件包含三种类型对象的API和高级API。其中Attribute、Dataset、File的接口如下：

* H5A—Attribute接口
* H5D—Dataset接口
* H5F—File接口

HDF5高级API简化了创建和访问对象所需的许多步骤，并提供了存储对象的模板。以下是高级API列表：

* HDF5 Lite（H5LT）--简化了创建数据集和属性的步骤
* HDF5 Image（H5IM）--定义了以HDF5格式存储图像的标准
* HDF5 Table（H5TB）--浓缩创建表所需的步骤
* HDF5 Dimension Scales（H5DS）--为尺寸标度存储提供标准
* HDF5 Packet Table（H5PT）--提供存储数据包数据的标准

同时HDF5包含了一些简单的工具：

* h5dump--转储或显示HDF5文件内容的实用程序
* h5cc，h5c ++，h5fc--用于编译应用程序的Unix脚本
* HDFView--用于查看HDF（HDF4和HDF5）文件的Java浏览器

**5 PHDF5(Parallel HDF5)**

HDF5中的并行基于MPI并行方式。MPI中包含两个重要的术语：MPI Communicator和Collective。MPI Communicator允许一组处理器之间进行交流，而Collective保证Communicator中的所有处理器以一个正确的顺序参与任务。PHDF5打开带有Communicator的并行文件，返回一个文件句柄，并用于将来访问该文件。所有进程都需要参与Parallel HDF5 API的Collective。可以使用不同的Communicator打开不同的文件。如下为使用PHDF5创建和访问文件程序：

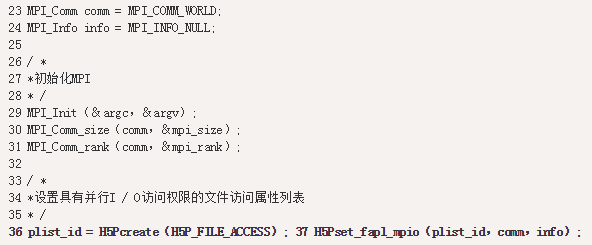


图9. PHDF5创建和访问文件

**6 SWMR (Single-Writer / Multiple-Reader)**

Single-Writer / Multiple-Reader（SWMR）功能允许多个进程在写入时（通过单个进程）读取HDF5文件，而无需使用锁或进程之间的通信。这项工作的基本工程挑战是确保HDF5文件的读者总能看到连贯的HDF5文件。SWMR解决方案实现了何时可以将元数据刷新到文件的依赖性。这确保了元数据缓存刷新操作以正确的顺序发生，因此物理文件中永远不会有指向无效（未刷新）文件地址的内部文件指针。使用SWMR访问的有益副作用是更好的容错。使用SWMR时更难破坏文件。

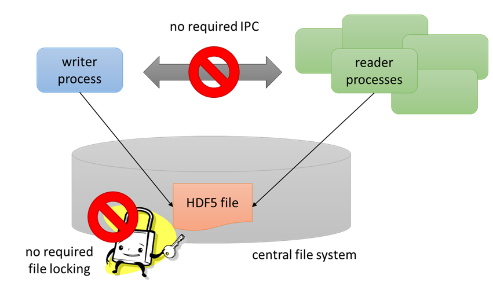


图10.SWMR

**7 VDS (Virtual Dataset)**

VDS的作用是将多个文件映射到一个虚拟的数据集中，并通过该虚拟的数据集来访问数据，访问方式和其它一般的数据集一样。如图11所示，我们创建了一个4x6大小的VDS,并将右边的a.h5、b.h5、c.h5三个文件的数据集映射到了左边的VDS中。

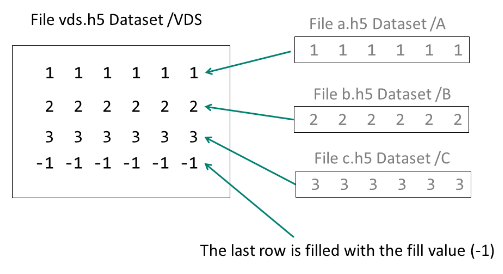


图11.VDS

**7 总结**

HDF5适合于庞大及其复杂的数据集，应用领域广泛。HDF5底层采用树结构进行数据的组织和存储，其上层抽象出group和dataset等概念，并提供了相应的API。此外HDF5为加速文件存储和访问、并发控制和多源数据集映射提供了并行化接口、SWMR以及VDS机制。

**UDA(Unified Data Access)调研**

**摘要：**本文简单介绍Iter用户手册中涉及到的UDA部分，主要将UDA分为脉冲数据库、UDA索引数据库、UDA索引器、UDA服务、UDA客户端工具以及API接口等六个方面。首先我么概括UDA的整体架构，接着分别介绍每个小模块，最后总结并概括UDA。

**1 整体结构**

UDA主要包括两个方面分别为UDA的索引服务和UDA服务。UDA索引服务包括索引数据库和索引代理组件，其中索引数据库用于存放数据的索引数据，而索引代理组件用于代理索引和数据的连接。UDA服务包括标准的IDAM服务、标准的MDSplus服务、索引服务和UDA插件。本文主要介绍标准的MDSplus服务、索引服务。

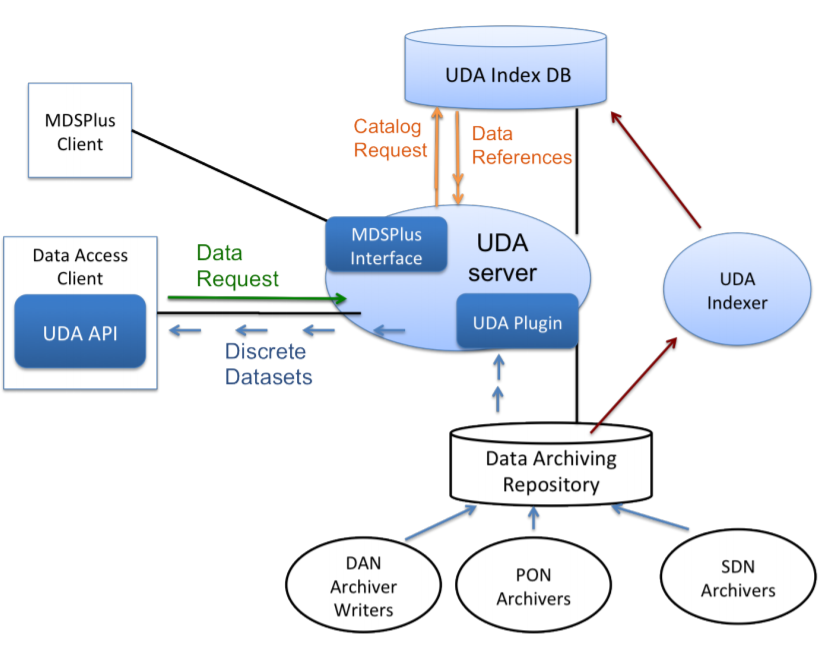


图1.UDA整体架构

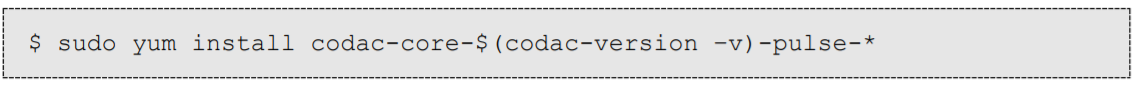
从图1中可以看出，UDA服务处于整个架构的核心位置，向上连接索引数据库，向下连接脉冲数据库，并为数据请求客户端和MDSplus客户端提供相应的接口。此外索引器连接索引数据库和脉冲数据库，用于建立索引和数据之间的连接。

**2 UDA子模块**

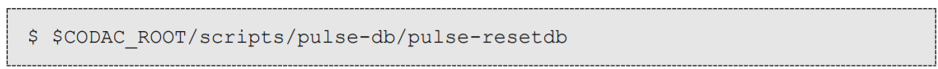
**2.1 脉冲数据库**

脉冲数据库主要包括脉冲数据库原型版本和脉冲数据库客户端。脉冲数据库主要用于存放原始数据。下面介绍脉冲数据库的安装和管理。

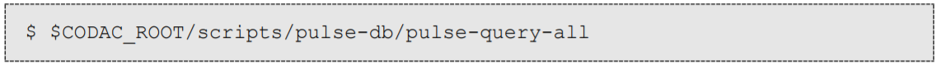
* 安装



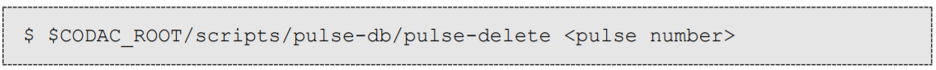
* 数据重置



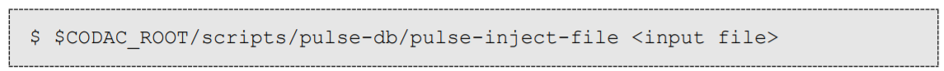
* 数据查询



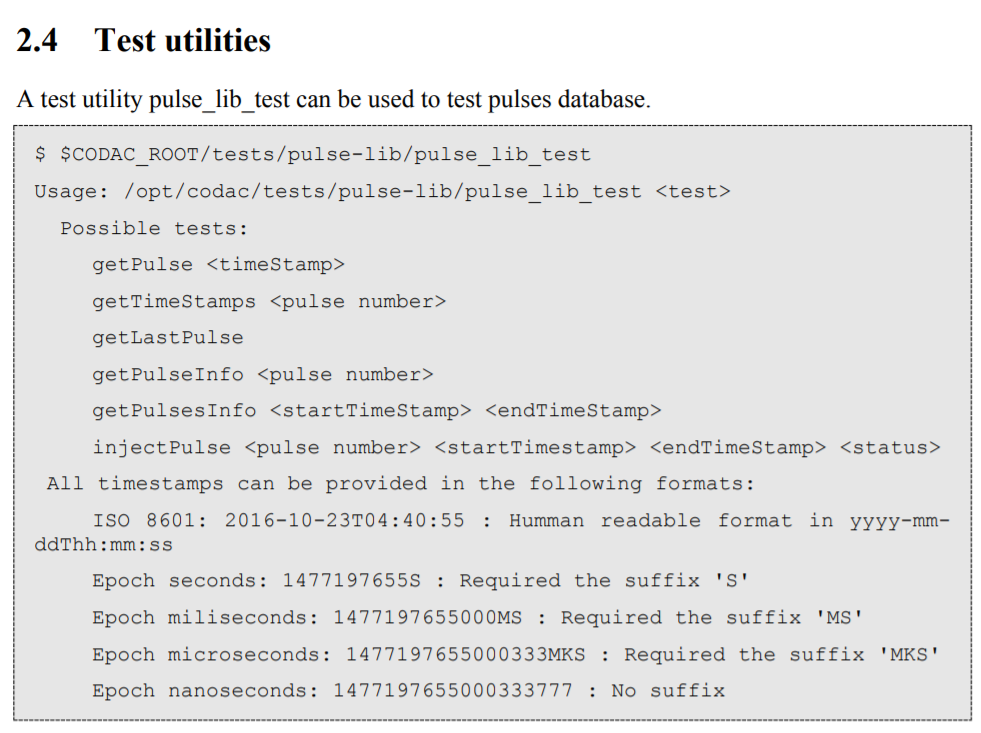
* 数据删除



* 数据插入



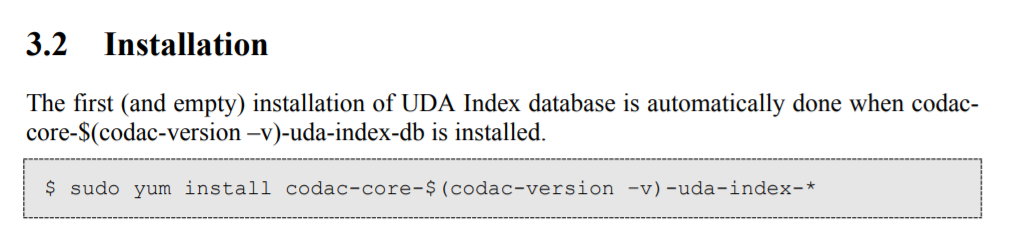
* 数据测试



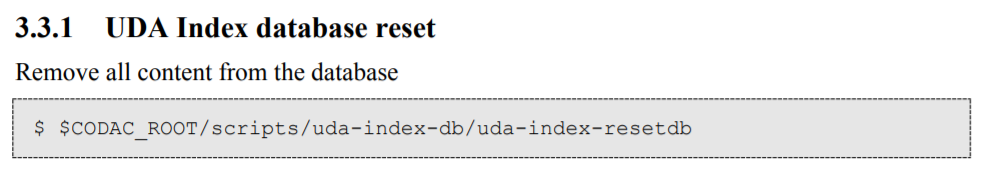
**2.2 索引数据库**

UDA索引数据库包含UDA可访问的所有存档变量的信息。 UDA索引还保存有关不同时间间隔的存档数据所在资源的信息。UDA索引数据库数据主要有两个来源，一个是在线的UDA索引器，其包含实时监测新文件、更新文件或移除文件等检测数据，另一个来源是离线UDA索引器实用程序，用于索引明确目录下的所有文件。下面介绍索引数据库的安装和管理。

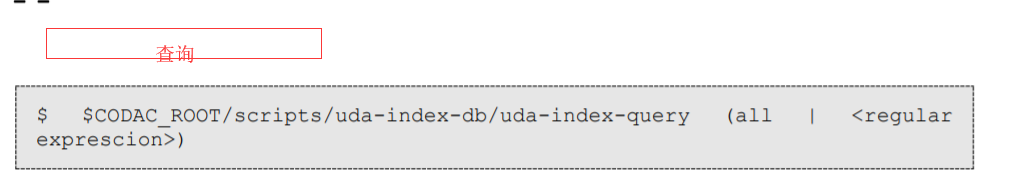
* 安装



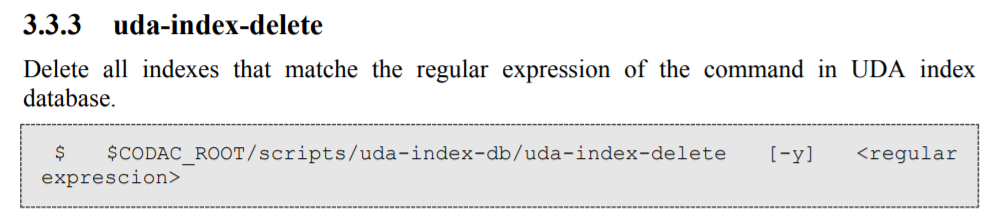
* 数据重置



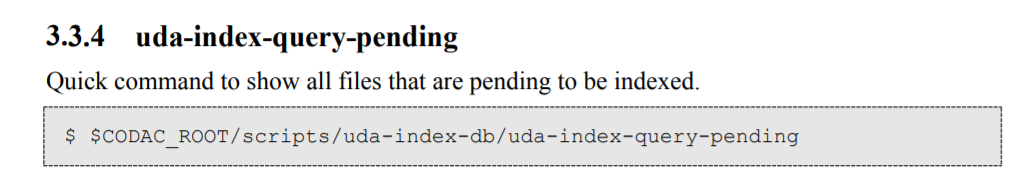
* 数据查询



* 数据插入



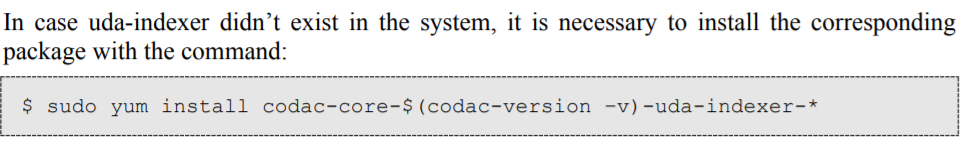
* 显示要所有挂起的索引



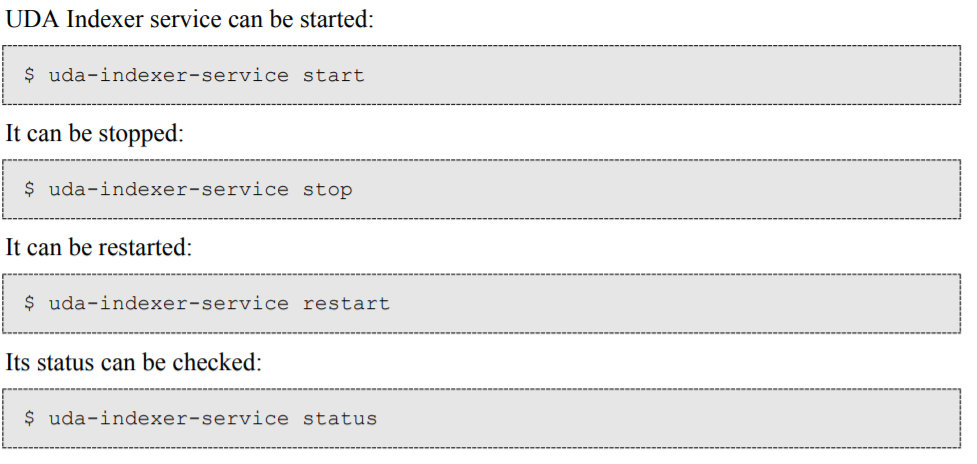
**2.3 UDA索引器**

UDA索引器是一个索引器代理，它根据在文件系统归档目录中创建和填充的HDF5文件来实时更新UDA索引数据库。UDA索引器监视已配置目录中的文件系统事件并更新UDA索引信息。下面我们介绍UDA索引器的安装、启动和停止。

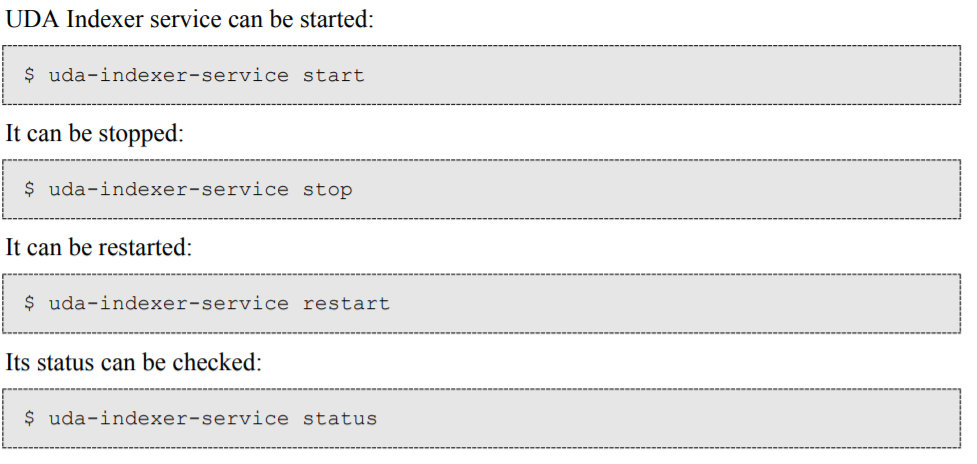
* 安装



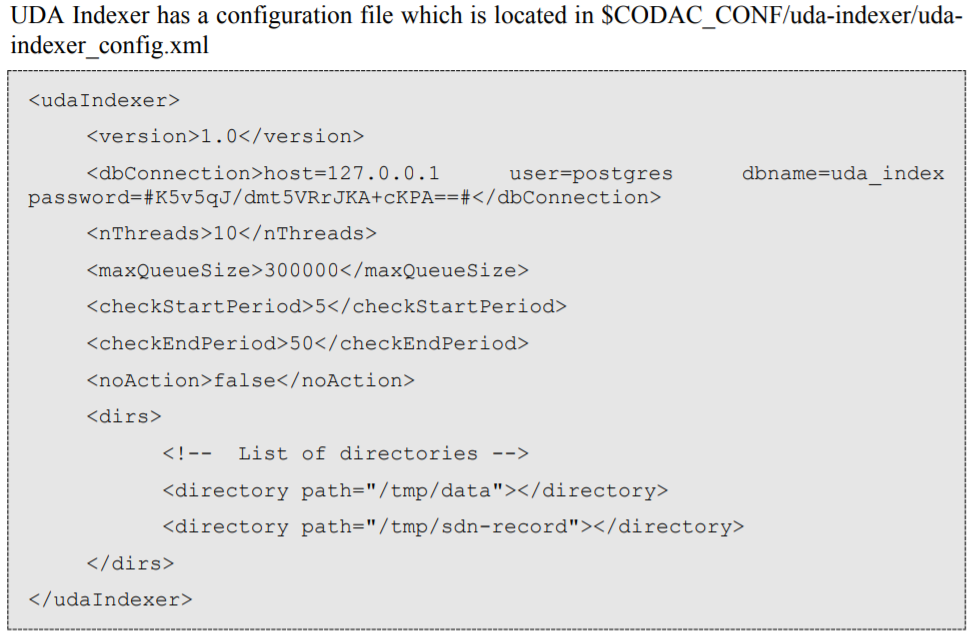
* 启动



* 停止



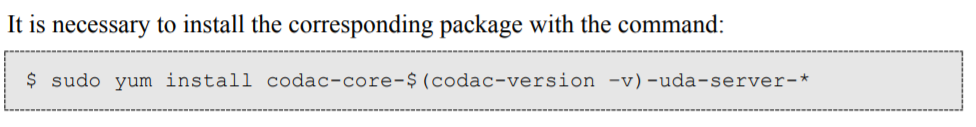
* 配置



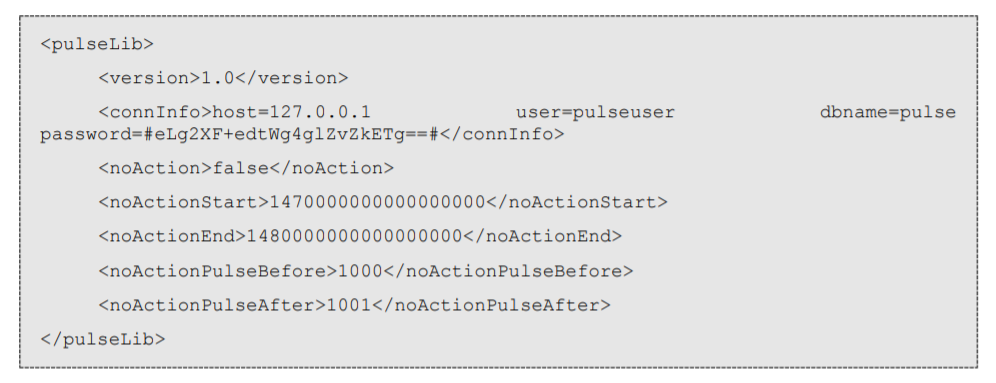
**2.4 UDA服务**

UDA服务是整个架构的核心，向下主要用于访问脉冲数据库，向上用于访问UDA索引数据库。下面介绍UDA服务的安装和配置。

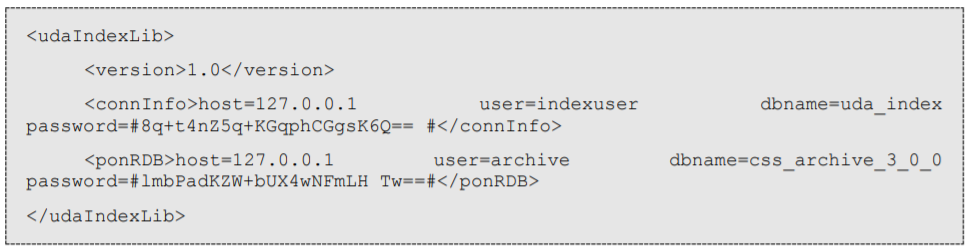
* 安装



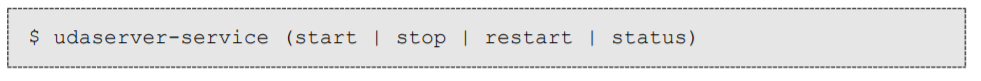
* 脉冲数据库访问配置



* UDA索引数据库访问配置

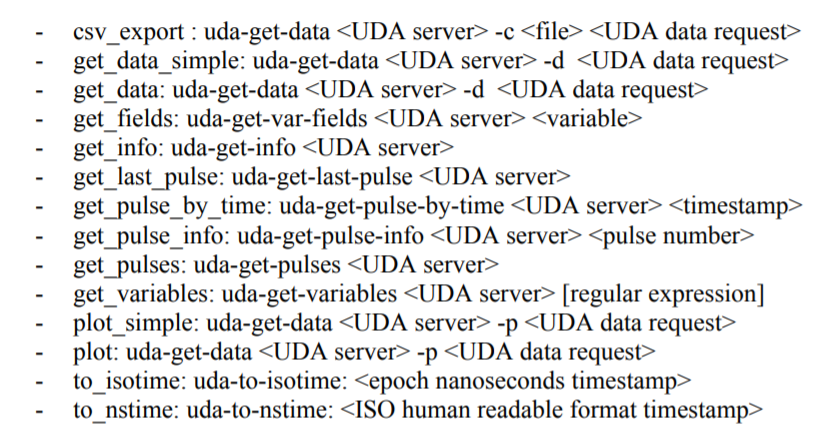


* 运行UDA服务

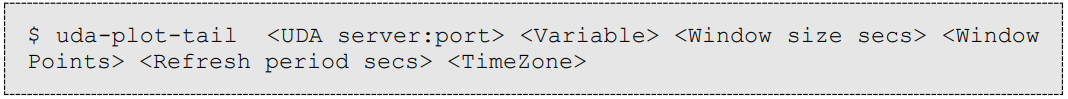


**2.5 UDA客户端工具**

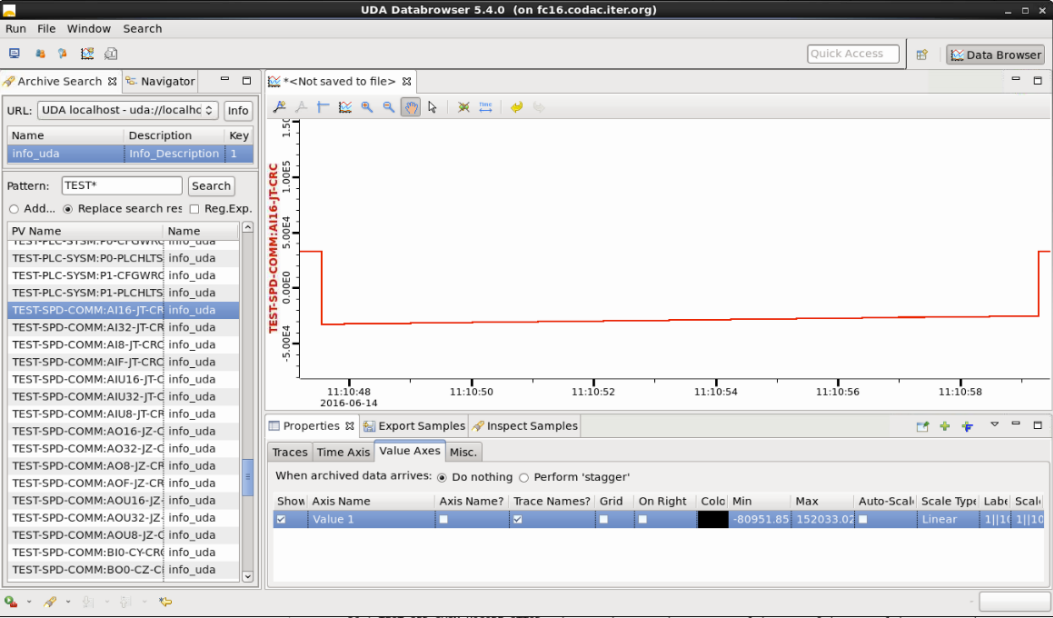
UDA有丰富的客户端工具，包括uda-itool、uda-plot-tail、UDA数据浏览器、JScope等。uda-itool是一种命令式交互工具，它包含帮助命令、获取数据命令、画图命令等。



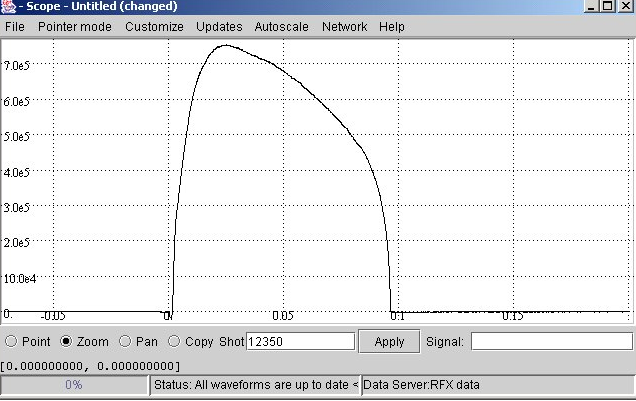
uda-plot-tail是UDA客户端包，包含一个绘图工具。它实现了基于UDA服务器轮询的存档数据绘制。



UDA Databrowser是EPICS世界的主要可视化工具之一。 通常，这种可视化工具能够在许多其他特征中绘制时间演变数据EPICS变量，也称为PV（过程变量）。

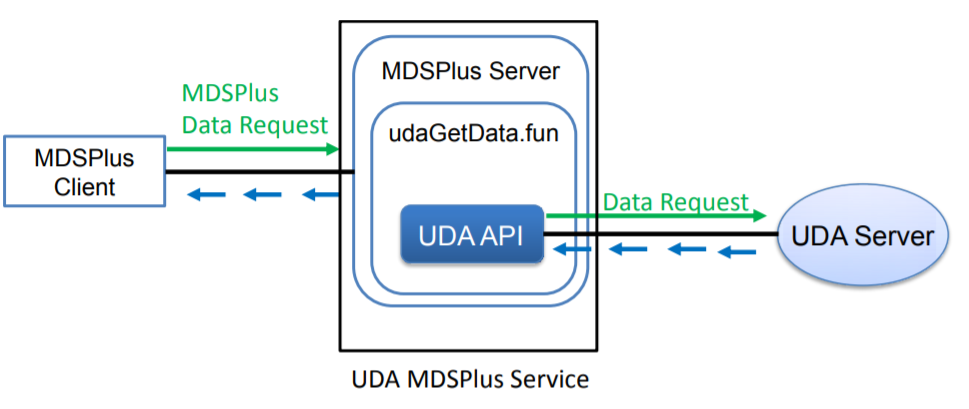


JScope是MDSPlus发行版中包含的绘图仪实用程序。



**2.6 API接口**

UDA提供了丰富的高级访问API的接口和MDSplus接口，其中高级访问API包括C++、java、Python等接口。高级访问API的基本功能在C++中实现，并由SWIG实用程序处理，以便在C++，Java，Python等中获得相同的功能。除此之外，调用用户定义的回调函数的迭代器类是单独实现的。MDSplus接口是通过MDSPlus服务实现的，该服务接收MDSPlus请求并使用idamget.fun TDI功能将请求传递给相应的IDAM服务器。当idamget.fun函数收到响应数据时，它会构建一个由MDSPlus服务器传递给MDSPlus客户端的响应。 此通信通过端口8000完成。



**3 总结**

UDA具有灵活的结构，底层存储不同格式的数据，可以开发各种插件用于查询不同类型数据，还可以通过索引数据库对数据进行更加快速的查询，且通过索引器及时同步脉冲数据库和索引数据库，提供MDSPLUS接口访问MDSPLUS里的数据。UDA为远程用户提供Web API和Web GUI界面，同时也提供了C++、PYTHON等封装好的接口进行访问。

**参考资料**

1.UDA: <http://202.127.205.42/doc/v6.0/pdf/UDA_User_Manual.pdf>

2. HDF5: <https://portal.hdfgroup.org/display/HDF5/HDF5>