

# EAST 极向场电源的数据存储及读取方法<sup>\*</sup>

王广红, 黄连生, 高 格, 傅 鹏, 何诗英, 王泽京

(中国科学院 等离子体物理研究所, 合肥 230031)

**摘 要:** EAST (Experimental Advanced Super-conducting Tokamak) 装置的等离子体放电持续时间设计值为 1000 s, 目前放电长度已超过 100 s。极向场电源(PF)是 EAST 的核心系统, 有接近 200 路信号需要进行数据存储, 每一次放电的数据量非常大, 普通的数据库不能够满足长脉冲数据存储的要求。本文使用 HDF5 (Hierarchical Data Format) 存储长脉冲数据, HDF5 是可以存储不同类型的图像和数码数据的文件格式。本文采用 HDF5 的 C 接口, 设计复合数据类型的数据集, 为了实现实时存储与实时读取, 将一次放电时间进行分割, 并对数据进行分块存储和读取。通过可视化工具 Hdfview 查看数据集, 以及 Qt 读取 HDF5 文件并进行波形显示, 证明复合数据类型和分块数据集的正确性, 以及分块写入与分块读取的可行性, 同时对读写时间进行简单的测试, 证明分块写入与分块读取方式具有高效性。

**关键词:** EAST 极向场电源; 长脉冲; HDF5; 复合数据类型; 分块存储与分块读取

**中图分类号:** TP273 **文献标志码:** A **doi:**10.11884/HPLPB201931.180360

## Research on data storage and reading method of EAST poloidal field power supply

Wang Guanghong, Huang Liansheng, Gao Ge, Fu Peng, He Shiyong, Wang Zejing

(Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China)

**Abstract:** The plasma discharge duration of EAST (Experimental Advanced Super-conducting Tokamak) is designed to be 1000 s, the time has exceeded 100 s at present. As the core subsystem, poloidal field power supply(PF) has nearly 200 signals to be stored. For every discharge, the amount of data is very large, ordinary database can't meet the requirement of long pulse data storage. This paper uses HDF5 (Hierarchical Data Format) to store the long pulse data. HDF5 is a file format that can store different types of images and data. Using C interface of HDF5 to design a dataset with compound datatype, to realize real-time writing and reading, the primary discharge time is segmented and the data is written and read as chunks. Visualization tool Hdfview is used to view the dataset, and Qt is used to read HDF5 file and display waveform. It proves the correctness of compound datatype and chunk dataset, and the feasibility of writing and reading as chunks. Simple time test proves that writing and reading as chunks are efficient.

**Key words:** EAST poloidal field power supply; long pulse; HDF5; compound datatype; write and read as chunks

**PACS:** 55.52.Fa; 07.05.Kf

每一套 PF 电源主要包括整流器(CU)、晶闸管开关网络(SNU)、失超保护(QP)和隔离开关(DS), 以及冷却水、电压电流检测、烟雾检测等辅助系统, 12 套电源总共有超过 1000 通道的信号, 如图 1 所示<sup>[1]</sup>。需要进行数据存储的主要为整流桥、旁通的电压电流, 以及晶闸管关断电压和参考电压等, 总共接近 200 通道。假设 EAST 放电时长 100 s, 采集频率 10 kHz, 信号为浮点数, 占 4 个字节, 一次放电总数据量为 800 MB。若放电时长为 1000 s, 一次放电的数据量为 8 GB。如此庞大的数据量, 并且具有脉冲特性的信号存储, 关系数据库或者非关系数据库等不能满足需求, 而 MDSplus 和 HDF5 都是存储长脉冲大数据的数据库。目前 EAST 极向场电源采用 MDSplus 存储数据, MDSplus 在 EAST 的应用中表现出了良好的性能<sup>[2-4]</sup>, 但 HDF5 在 ITER (国

<sup>\*</sup> 收稿日期: 2018-12-03; 修订日期: 2019-02-28

作者简介: 王广红(1982—), 女, 硕士, 工程师, 从事电源控制器研究; wgh@ipp.ac.cn。

通信作者: 黄连生(1983—), 男, 博士, 大功率电源控制、控制系统设计、算法研究、控制系统电磁兼容研究; huangls@ipp.ac.cn。

际高热原子能反应堆)数据存储方面表现出了更多的优越性,HDF5 比 MDSplus 具有更优越的底层代码,以及更丰富的数据存储格式,如 HDF5 支持图片格式,同时 HDF5 具有更优越的数据读写性能<sup>[5]</sup>。HDF5 实现了一种用于存储和管理数据的模型,以分级格式存储,支持各种各样的数据类型,提供 Java,C 和 Fortran 访问库,本文使用 C 接口库,存储规模可以无限大,具有可移植性,数据文件的内部组织结构可以任意复杂<sup>[6]</sup>。为了与 ITER 接轨,同时探索未来 EAST 极向场多元化数据综合存储技术及应用可能性,所以对 HDF5 的存储与读取进行研究,本文还处于研究阶段,仅进行功能设计与测试,并未搭建完整的存储系统。

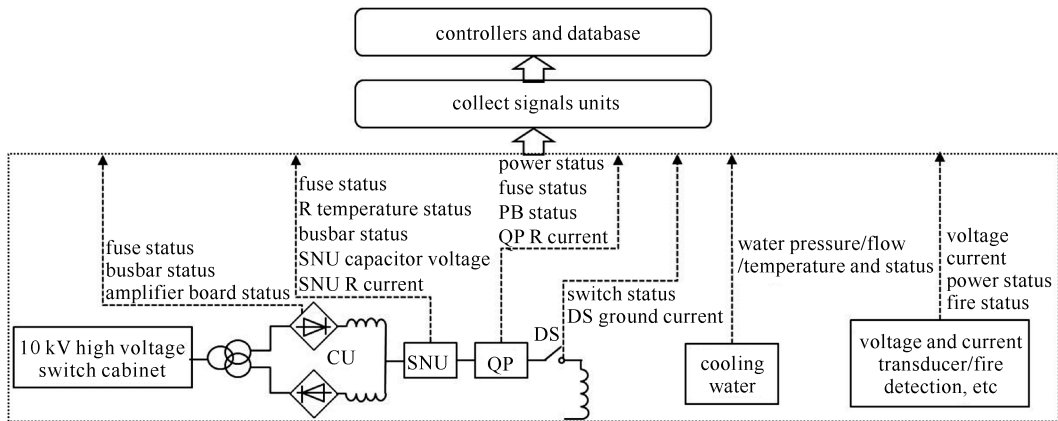


Fig. 1 Hardware and signal diagram of EAST poloidal field  
图 1 EAST 极向场电源硬件与信号简图

1 程序总体框架

图 2 为程序总体框架,接收线程、存储线程、读取线程都在一个进程中实现。以 EAST 放电时长 100 s 为例进行设计。接收线程通过 TCP 协议接收来自底层的数据。当放电开始时,存储线程检测炮号是否发生变化,如果变化则对 HDF5 数据集元数据属性进行设置,并创建新文件,接收线程将接收的数据整合到缓存中,当一块数据接收完成,标志位置位,通知存储线程写入当前块的数据。放电结束时,存储线程释放各资源。读取线程采用逐块读取的方式进行,若写完一块数据马上读取,即为实时读取。也可以在放电结束后的任何时候读取数据,此时可以逐块将所有数据读出,读取数据存放到缓存中。缓存中的数据可用于波形绘制或数据分析,此部分在本文不做说明。

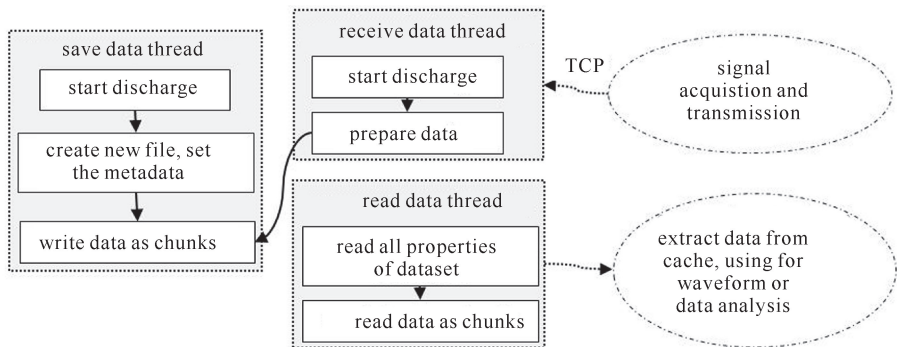


Fig. 2 Program diagram  
图 2 程序总体框架

创建新文件时的各属性设置、分块写入数据、分块读取数据等所有对 HDF5 的操作都离不开其接口函数<sup>[7]</sup>,如表 1 所示为本文使用的所有 HDF5 接口函数统计。

2 复合数据类型的分块数据集设计

HDF5 文件以分层结构组织,主要包含组和数据集,通过组的方式组织成树状结构,每个组都是树的节点,包含零个或多个组或数据集。本文没有进行分组,采用根组下面单一数据集的方式。数据集是由标题和数据数组组成的,标题在创建数据集时设定,打开数据集时指出数据集的标题,数据数组包括元数据(Metadata)和多维数组(Data),其中 Metadata 包含 Data 相关的信息,比如维度信息、数据类型、属性等,而 Data 则包含数据

表 1 HDF5 部分接口统计  
Table 1 Part of HDF5 interfaces

interface	description	functions used in this paper
H5D	datasets interface, create and manipulate dataset objects, and set and retrieve their constant or persistent properties.	H5Dcreate2, H5Dread, H5Dwrite, H5Dopen2, H5Dget _ space, H5Dset _ extent, H5Dclose, H5Dget _ type, H5Dget_create_plist.
H5A	attribute interface, create and manipulate attributes and information about attributes.	H5Acreate2, H5Aopen, H5Aget _ space, H5Aread, H5Awrite, H5Aclose.
H5S	dataspace interface, create and manipulate the dataspace in which to store the elements of a dataset.	H5Screate_simple, H5Screate, H5Sget_simple_extent _ dims, H5Sget_simple_extent_ndims, H5Sselect_hyper _ slab, H5Sclose.
H5F	file interface, designed to provide file-level access to HDF5 files.	H5Fcreate, H5Fopen, H5Fclose.
H5T	datatype interface, create and manipulate the datatype which describes elements of a dataset.	H5Tarray _ createl, H5Tcreate, H5Tinsert, H5Tget _ class, H5Tclose.
H5P	property list interface, manipulate property list objects in various ways, including reset property values.	H5Pcreate, H5Pset_chunk, H5Pset_fill_value, H5Pget _ chunk.

本身。图 3 为本文数据集设计简图，HDF5 的数据集默认是连续空间存储的，但是也可以进行分块，即将数据分割为固定大小的小块分开存储，分块存储的方式使得数据集可以改变大小，也可以进行压缩过滤，本文采用分块存储。分块太少，导致一次写入数据量过大，写文件耗时较长，同时对缓存的要求较高。分块太多，频繁写文件会增加系统负担。目前按照 1s 的时间段进行分割，后期需对不同的时间段分割进行详细的性能比较。数据集依赖于数据空间，数据空间描述数据集的维度，数据集的维度可以是固定的或者无限的，数据空间的属性包括实际维度、纬度的最大值等。本文数据集维度最大值设置为 H5P\_UNLIMITED，即数据集维度可以无限扩展。

以 100 s 放电时长进行测试，将 100 s 分割为 100 块，即 1 s 写一次数据，1 s 内每个信号有 10 000 个点，即每一块的维度为 10 000。本文需要存储的数据包括块索引号（整形，当前数据为第几次写入，数值为 0 到 99）、时间戳（长整形）、200 路信号（浮点数），设计包含上述所有数据类型的复合数据类型数据集。数据集的列分别代表块索引号、时间戳和 200 路信号，共 202 列，行代表不同时间戳下的各信号数值以及该信号所对应的块索引。写入数据时，数据集的行数不断增加，列数不变。而设计复合数据类型时，对每一列的数据类型进行设置，并逐列进行操作，对于数据集的每一列，其数据空间为一维，块维度为 10 000。所以创建数据集时，设置数据空间为一维，块纬度设置为 10 000，最大纬度不受限制，并设置复合数据类型和分块属性，如图 3 所示。随着数据集的不断写入，数据集维度不断增加，对于 100 s 放电时长，其数据集的最终维度为 1 000 000。

HDF5 数据集是数据元素的数组，按照数据存储的规范排列，HDF5 数据类型定义了单个数据元素的存储格式。一个数据集有一个单一的数据类型，用于每个数据元素。当创建数据集时，定义存储数据的类型，之后数据类型不能被修改。数据类型有两种，原子数据类型和复合数据类型。HDF5 库预先定义了一组类似于 C 类型名称的本机数据类型，例如，H5T\_NATIVE\_INT 对应于 C 的 int 类型。表 2 给出了常见的本机数据类型和相应 C 类型<sup>[8]</sup>。

复合数据类型是一个或多个原子数据类型的集合，将多个数据类型的集合表示为单个单元的数据类型。复合数据类型的成员可以是任何数据类型，包括另一个复合数据类型。首先定义复合数据类型结构体，图 4 中结构体 ArrayStruct，宏定义 SIGNO 数值为 200，注意各变量顺序，其位置决定了自身在该结构体中的偏移量。

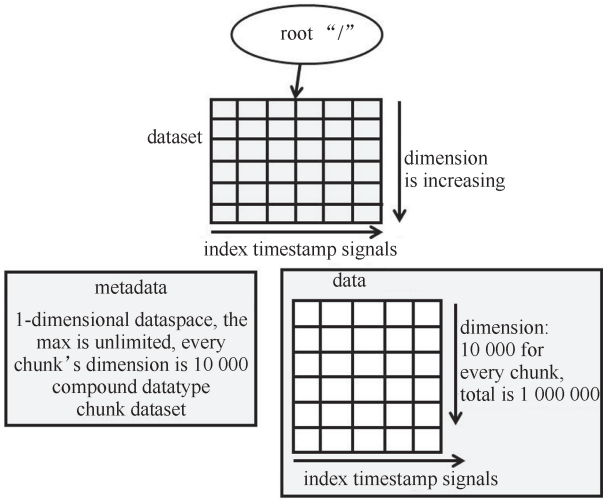


Fig. 3 Dataset design diagram  
图 3 数据集设计简图

图 4 为复合数据类型的创建框图,使用 H5Tcreate(H5T\_COMPOUND, sizeof(ArrayStruct))创建复合数据类型,返回对象 ArrayStructid,通过 H5T\_COMPOUND 指定创建复合数据类型;使用 H5Tinsert 函数将各数据类型插入 ArrayStructid 之中,比如块索引号插入程序为 H5Tinsert(ArrayStructid, "index", HOFFSET(ArrayStruct, index), H5T\_NATIVE\_INT),插入列的名字命名为 index,偏移为 ArrayStruct 结构体中的 index,元素的数据类型为 H5T\_NATIVE\_INT;然后插入时间戳和 200 路信号,注意偏移必须和结构体的定义保持一致。

表 2 HDF5 本机数据类型和对应 C 类型  
Table 2 HDF5 native type and corresponding C type

HDF5 native type	C type	HDF5 native type	C type	HDF5 native type	C type
H5T_NATIVE_CHAR	char	H5T_NATIVE_ULLONG	unsigned long	H5T_NATIVE_LONG	long
H5T_NATIVE_UCHAR	unsigned char	H5T_NATIVE_DOUBLE	double	H5T_NATIVE_LLONG	long long
H5T_NATIVE_USHORT	unsigned short	H5T_NATIVE_SCHAR	signed char	H5T_NATIVE_FLOAT	float
H5T_NATIVE_UINT	unsigned	H5T_NATIVE_SHORT	short	H5T_NATIVE_LDOUBLE	long double
H5T_NATIVE_ULONG	unsigned long	H5T_NATIVE_INT	int		

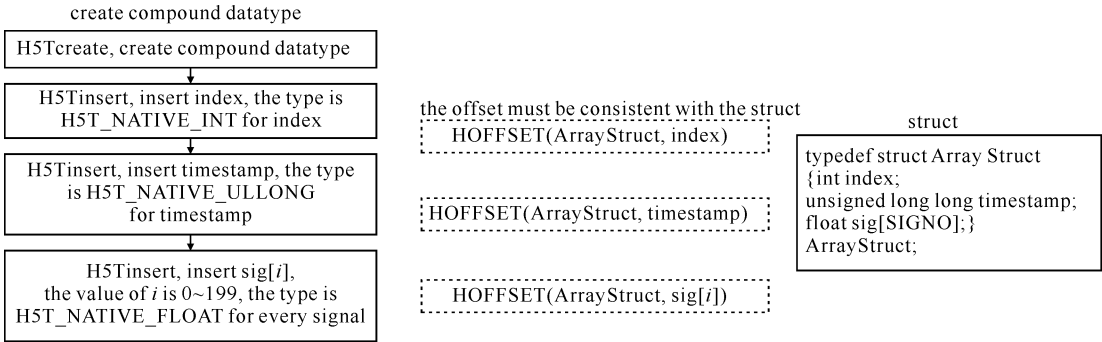


Fig. 4 Creation progress of compound datatype  
图 4 复合数据类型创建过程

3 分块存储数据

HDF5 分块的实现主要通过 H5Sselect\_hyperslab 函数对数据空间进行选择,该函数原型为 herr\_t H5Sselect\_hyperslab(hid\_t space\_id, H5S\_seloper\_t op, const hsize\_t \* start, const hsize\_t \* stride, const hsize\_t \* count, const hsize\_t \* block),start, stride, count 和 block 维数必须和数据空间的维数相同,本文将 stride 和 block 设置为 NULL,分别代表选择连续块与块大小默认为每个维度中的单个元素, start 指定起始元素的偏移量, count 指定块的元素个数,也即块维度。如图 5 所示为分块写入实现框图。当放电开始,若炮号发生变化,则创建新文件并设置相关属性。H5Pcreate 创建一个新的属性列表,返回对象 cparms。H5Pset\_chunk 为上述对象 cparms 设置用于存储分块布局数据集的块的大小,数据类型为上述新建的复合类型 ArrayStructid,块的数据空间设置为 1,块维度为 10 000。H5Dcreate2 创建数据集,数据集标题设置为宏定义 DATASETNAME,数据类型为 ArrayStructid,属性为 cparms。放电开始后,接收线程不断将接收数据存储到缓存中,当 1 s 的数据准备好后,置位标志位,存储线程开始向数据集写入一块的数据,写循环每次执行,先调用 H5Dset\_extent 函数改变数据集维度为 NX \* (index + 1),块起始位置设置为 index \* NX,块维度设置为 NX,然后使用 H5Sselect\_hyperslab 选择块,最后将缓存中已经准备好的数据写入数据集,此部分循环执行, 100 s 放电时长,共执行一百次,执行次数减 1 即为块索引号 index。放电结束后,释放所有资源,比如关闭文件、数据空间、数据集、属性等。



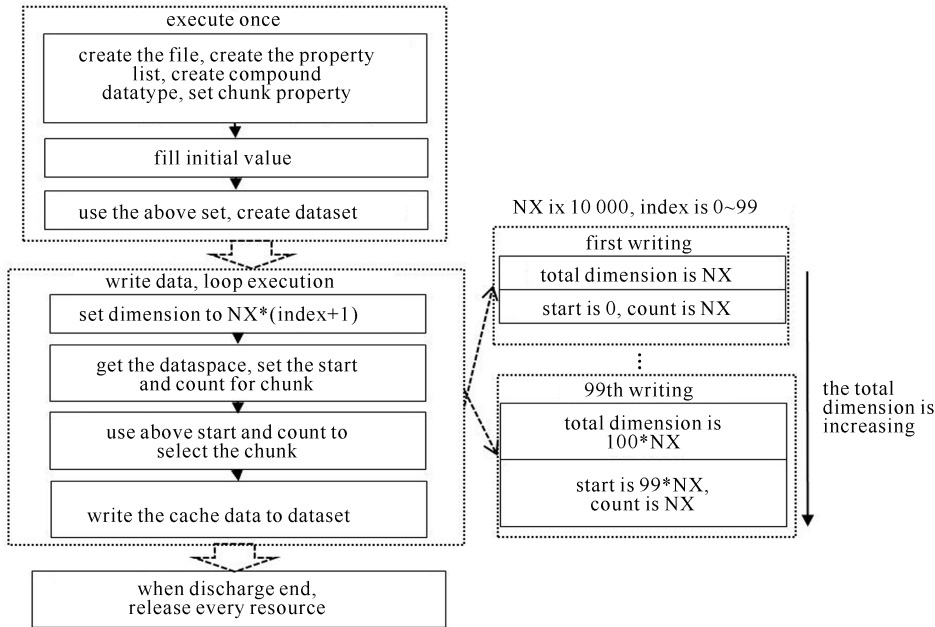


Fig. 5 Write the real-time data as chunks

图 5 实时数据分块写入

4 分块读取数据

如图 6 所示为分块读取数据框图。H5Fopen 打开文件、H5Dopen2 打开数据集、H5Dget\_space 获取数据空间、H5Sget\_simple\_extent\_dims 获取数据集维度、H5Dget\_type 获取数据类型、H5Dget\_create\_plist 获取属性列表、为缓存分配空间。如果为复合数据类型并且为分块数据集，则 H5Pget\_chunk 读取块维度，然后循环进行以下读取操作。块参数设置与写入块参数一致，即每次读取块的起始位置为  $\text{index} \times \text{NX}$ ，块维度为 NX，NX 数值为 10 000，然后使用 H5Sselect\_hyperslab 函数选择块。读取整个 HDF5 文件，则获取的数据集总维度为 1 000 000，块维度为 10 000，循环次数为 100。实时读取时，写完一块数据马上读取，块维度为 10 000，数据集总维度不断增加，存储线程写完一块数据后，将写完成标志位置位，并将当前块的索引号 index 发送给读取线程，读取线程检测到写数据完成标志位置位后，选择块的偏移设置为  $\text{index} \times \text{NX}$ ，块维度设置为 NX，然后使用 H5Dread 函数从数据集读取数据到缓存。放电结束后，退出循环，并释放各资源。

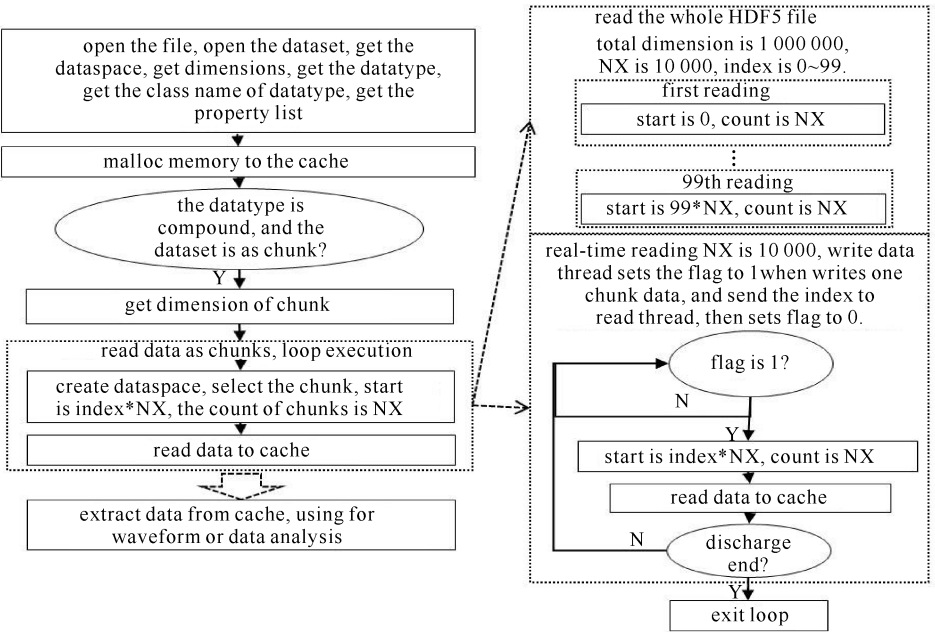


Fig. 6 Read the data as chunks

图 6 分块读取数据

5 测试结果

5.1 写入和读取数据时间测试

如图 7(a)所示为放电时长 100 s,1 s 写一次文件,每次写文件为 8 MB 时的写文件耗费时间测试。图中横轴为块索引号,纵轴为每次写文件所用的时间,写数据平均值为 9 426  $\mu$ s,最大值为 12 900  $\mu$ s。图 7(b)为读取每块数据所用时间测试,平均值为 19 933  $\mu$ s,最大值为 23 439  $\mu$ s。可见读取 HDF5 文件比较耗时,使用 Qt 进行波形显示,设计抽点读取方法,读取与显示时间共约 600 ms,读取整个 HDF5 文件用时大概 2 s。EAST 目前使用 MDSplus 进行长脉冲分块存储,其显示时间小于 1 s<sup>[9]</sup>。可见 HDF5 的分块存储与读取方式,性能可与 EAST 目前存储系统相比较,但对 HDF5 的研究还处于初步阶段,后期需要对其性能进行更为详细的测试。

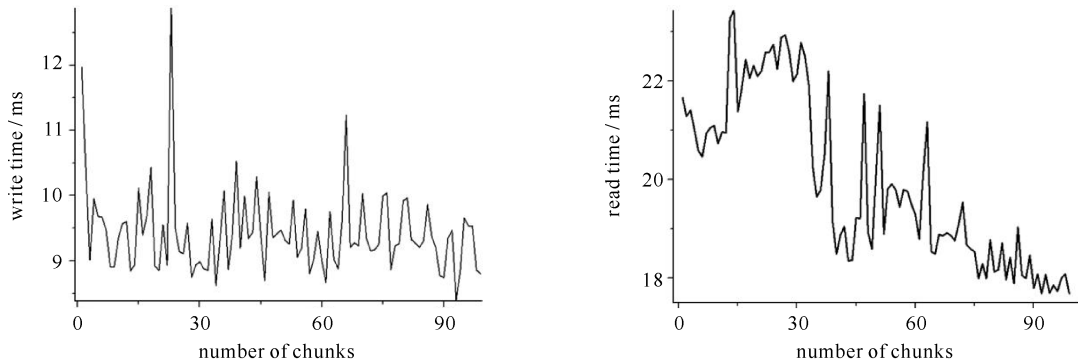


Fig. 7 Spent time for writing and reading data as chunks

图 7 分块写数据与分块读取数据占用时间测试

5.2 Hdfview 查看数据集

Hdfview 是一个用于 HDF5 文件的可视化浏览和编辑工具,可在树形结构下浏览一个文件构成,创建新文件、添加或者删除数据集,浏览和修改数据集的内容,添加、删除和修改属性等<sup>[10]</sup>。图 8 为通过 Hdfview 查看的数据集的所有数据。

Table		34, sig4 ... 0.4																				0-based	
		30	sig181	sig182	sig183	sig184	sig185	sig186	sig187	sig188	sig189	sig190	sig191	sig192	sig193	sig194	sig195	sig196	sig197	sig198	sig199	sig200	
999961	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999962	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999963	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999964	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999965	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999966	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999967	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999968	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999969	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999970	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999971	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999972	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999973	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999974	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999975	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999976	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999977	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999978	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999979	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999980	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999981	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999982	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999983	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999984	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999985	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999986	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999987	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999988	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999989	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999990	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999991	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999992	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999993	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999994	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999995	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999996	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999997	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999998	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			
999999	117.1	117.2	117.3	117.4	117.5	117.6	117.7	117.8	117.9	118.0	118.1	118.2	118.3	118.4	118.5	118.6	118.7	118.8	118.9	119.0			

Fig. 8 Check the dataset through Hdfview

图 8 Hdfview 工具查看数据集

6 结 论

本文使用 HDF5 实现了对 EAST 极向场电源 200 路信号的实时存储和读取,设计复合数据类型的分块式数据集。将 100 s 的放电时长按照 1 s 长度进行分割,写一块数据占用时间平均为 10 ms,读一块数据平均时间为 20 ms。测试中将放电时长设置为 1000 s,一次放电数据量为 8 GB,其写一块数据的平均时间也为 10 ms,读一块数据的平均时间也为 20 ms,可见写文件时间、读文件时间与总数据量无关,只与如何分块相关。通过 Qt 读取 HDF5 文件并显示波形,抽点读取与显示波形总时间约为 600 ms。通过 Hdfview 查看数据集,以

及 Qt 波形显示,证明复合数据类型和分块数据集的正确性,以及分块写入与分块读取的可行性,同时简单的时间测试证明分块写入与分块读取的方式具有高效性。HDF5 的功能繁多,库函数也非常多,本文处于研究阶段,后续需对所设计的 HDF5 存储与读取方式进行更为深入的功能与性能测试,并对 EAST 极向场电源搭建完善的实时数据存储与读取系统。

## 参考文献:

- [1] 王广红,何诗英,高格,等. EAST 极向场电源本地控制器软件设计与实现[J]. 计算机测量与控制, 2018, 26(1): 157-160. (Wang Guanghong, He Shiyong, Gao Ge, et al. Design and implementation of local controller software for EAST poloidal field power supply. Computer Measurement & Control, 2018, 26(1): 157-160)
- [2] 杨飞,肖炳甲,朱应飞,等. EAST 长脉冲放电实验实时数据系统[J]. 计算机工程, 2011, 37(4): 12-14. (Yang Fei, Xiao Bingjia, Zhu Yingfei, et al. Real-time data system in long pulse discharge experiments of EAST. Computer Engineering, 2011, 37(4): 12-14)
- [3] 张永定,秦品健,郑锐. MDSplus 在极向场电源数据管理系统中的应用[J]. 电源技术, 2011, 358(8): 966-968. (Zhang Yongding, Qin Pinjian, Zheng Rui. Application of MDSplus in data management of EAST poloidal field power supply. Chinese Journal of Power Source, 2011, 358(8): 966-968)
- [4] 刘强,肖炳甲,王华忠,等. 基于 MDSplus 的 EAST 长脉冲数据采集系统设计与实现[J]. 微计算机信息, 2012, 28(9): 59-61. (Liu Qiang, Xiao Bingjia, Wang Huazhong, et al. Design and implementation of EAST long-pulse data acquisition system based on MDSplus. Microcomputer Information, 2012, 28(9): 59-61)
- [5] Cook N. Survey of scientific data formats[R]. Cadarache: ITER Organization, 2015.
- [6] The HDF Group. HDF5 user's guide[EB/OL]. <https://support.hdfgroup.org/HDF5/doc1.6/UG/>.
- [7] 童大云,秦忠国,魏雨露,等. HDF5 在大型有限元软件数据管理中的应用[J]. 计算机应用与软件, 2014, 31(2): 58-61. (Tong Dayun, Qin Zhongguo, Wei Yulu, et al. Applying HDF5 in large-scale finite element software data management. Computer Applications and Software, 2014, 31(2): 58-61)
- [8] The HDF Group. HDF5 Datatypes[EB/OL]. [https://support.hdfgroup.org/HDF5/doc1.6/UG/11\\_Datatypes.html](https://support.hdfgroup.org/HDF5/doc1.6/UG/11_Datatypes.html).
- [9] 杨飞. EAST 实验数据系统研究[D]. 合肥: 中国科学院等离子体物理研究所, 2011. (Yang Fei. Systematic research on EAST experimental data. Hefei: Institute of Plasma Physics, Chinese Academy of Sciences, 2011)
- [10] The HDF Group. HDFVIEW[EB/OL]. <https://support.hdfgroup.org/products/java/hdfview/>.