# Hierarchical Data Acquisition (HDAQ) and Task Control (HTC)

## Definitions

### Hierarchical Task Control

1. A *Study* consists of multiple *Root* *Tasks* that may be carried out independently from each other, in an arbitrary order and at arbitrary time points, in a user-defined way.
2. A Root Task such as acquiring a stack of images may contain an arbitrary number of Subtasks grouped hierarchically at arbitrary nesting levels that form a *Task Tree* structure. The Root Task is the topmost Task node within the Task Tree.
3. Task Trees function independent of each other.
4. Parent Task nodes can have multiple child Task nodes or Subtasks. A child Task node may have only one parent Task node.
5. Adding a Child Task node to a Parent Task node makes the Child Task node share the same Root Task with the Parent node.
6. Once started, a parent Task is complete when all child Tasks are complete.
7. Tasks are carried out by *Task Controllers* (TCs) which model the actions of a software or hardware module during the execution of the Root Task.
8. TC names within a Task Tree are unique. TC names between Task Trees may be the same.
9. Only one TC is allowed per module/device and the TC may participate in only one Task Tree at a time during the execution of the Root Task. This promotes the reuse of TCs within different Root Tasks and decreases their total number required to carry out the Study.
10. The execution of a Task by a Task Controller is understood as an iteration of the Task.
11. Depending on the number of iterations, there are 3 kinds of Task Controllers: 1) null iterations – Null Task Controller (NTC), 2) finite iterations – Finite Task Controller (FTC) and 3) continuous iterations – Continuous Task Controller (CTC).
12. An NTC participates in the execution of the Task Tree but does not execute its iteration function. Instead it relays messages from its parent TC to its children TCs.
13. A FTC once started executes its iterations a finite number of times after which the task completes or the iterations are manually stopped.
14. A CTC has an arbitrary number of iterations. For a CTC to complete it must be stopped manually or a certain stopping condition must be met after an arbitrary number of iterations.
15. The execution of a Task Tree can be started and stopped only from its Root Task Controller which relays these commands to its child TCs. The child TCs in turn relay commands further down to their child TCs, to the deepest nesting level of the Task Tree.
16. The execution state of child TCs is relayed back to their parent TC and thus affects the execution state of their parent.

### Task Control Iterators

1. Each TC, including the Root Task TC, may have several TC *Iterators* (TCIs) that form a TC *Iterator Set* (TCIS) of the form {TCI1, TCI2,.., TCIN}. The elements in a TCIS are dimensions in an N-dimensional iterator space and the convention taken here is that elements in this set starting from left to right increase the dimensionality of this space.
2. A single TCI from a TCIS may have a name and keeps track of the total number of iterations, the current iteration index, and a TCI parameter set (TCIPS). This is summarized as TCI = ("iterator name"= name (optional if there are more than one TCIPs), "current iteration index/total number of iterations" = i/N, "parameter set" = {TCIP1, TCIP2, .., TCIPN}).
3. TCI names from a TCIS are unique within the set.
4. The total number of TC iterations is the product of the individual TCI total number of iterations.
5. A single TCIP from a TCIPS contains the name of the parameter that is iterated (such as frame number or z-stage position), the current “value” of the parameter which is a pointer (numeric, string, images, etc.) and the unit of measure if any. This can be summarized as TCIP = ("name" = name, "data type" = dtype, "value" = val, "unit" = unit).
6. The name of a TCI may be the same as the name of a TCIP.
7. Just as there are 3 different types of TCs: NTCs, FTCs and CTCs, there are 3 different kinds of TCIS elements: 1) null number of iterations – null task controller iterator (NTCI), 2) finite number of iterations – finite task controller iterator (FTCI), 3) indefinite number of iterations – continuous task controller iterator (CTCI).
8. A FTC finishes when all iteration indices from the TCIS are equal to the total number of required iterations for each index.

### Virtual Channels

1. Data exchange between the TCs occurs through *Virtual Channels* (VCs).
2. VCs of a TC have unique names and each TC manages its own list of VCs.
3. Depending on the data flow direction, there are two types of VCs: *Sink* VCs and *Source* VCs.
4. Data exchange between TCs is done by connecting Sink and Source VCs.
5. A Source VC can be connected to multiple Sinks, however, a Sink may have only one Source connected to it.
6. Sink and Source VCs from the same Task Tree can be connected if their data types are compatible. VCs belonging to different Task Trees cannot be connected.

## Hierarchical data format (HDF5)

### Introduction

Since data is acquired and indexed in a hierarchical way by the TCIS, it is convenient to also save it in a similar manner. To this end, using the time-tested industry standard HDF5 file format is the most logical approach:

http://www.hdfgroup.org

http://www.hdfgroup.org/HDF5/doc/UG/index.html

"*HDF5 is an open-source technology suite for managing diverse, complex, high-volume data in heterogeneous computing and storage environments. HDF5 includes: (1) a versatile self-describing data model that can represent very complex data objects and relationships, and a wide variety of metadata; (2) a completely portable binary file format with no limit on the number or size of data objects; (3) a software library with time and space optimization features for reading and writing HDF5 data; and (4) tools for managing, manipulating, viewing, and analyzing data in HDF5 files.*"

" *An HDF5 file is a data container, similar to a file system. Within it, user communities or software applications define their organization of data objects. The basic HDF5 data model is simple, yet extremely versatile in terms of the scope of data that it can store. It contains two primary objects: groups, which provide the organizing structures, and datasets, which are the basic storage structures. HDF5 groups and datasets may also have attributes attached, a third type of data object consisting of small textual or numeric metadata defined by user applications.*

*An HDF5 dataset is a uniform multidimensional array of elements. The elements might be common data types (for example, integers, floating-point numbers, text strings), n-dimensional memory chunks, or user-defined compound data structures consisting of floating-point vectors or an arbitrary bit-length encoding (for example, 97-bit floating-point number). An HDF5 group is similar to a directory, or folder, in a computer file system. An HDF5 group contains links to groups or datasets, together with supporting metadata. The organization of an HDF5 file is a directed graph structure in which groups and datasets are nodes, and links are edges. Although the term HDF implies a hierarchical structuring, its topology allows for other arrangements such as meshes or rings.*

*HDF5 is a completely portable file format with no limit on the number or size of data objects in the collection. During I/O operations, HDF5 automatically takes care of data-type differences, such as byte ordering and data-type size. Its software library runs on Linux, Windows, Mac, and most other operating systems and architectures, from laptops to massively parallel systems. HDF5 implements a high-level API with C, C++, Fortran 90, Python, and Java interfaces. It includes many tools for manipulating and viewing HDF5 data, and a wide variety of third-party applications and tools are available.*

*The design of the HDF5 software provides a rich set of integrated performance features that allow for access-time and storage-space optimizations. For example, it supports efficient extraction of subsets of data, multiscale representation of images, generic dimensionality of datasets, parallel I/O, tiling (2D), bricking (3D), chunking (nD), regional compression, and the flexible management of user metadata that is interoperable with XML. HDF5 transparently manages byte ordering in its detection of hardware. Its software extensibility allows users to insert custom software "filters" between secondary storage and memory; such filters allow for encryption, compression, or image processing. The HDF5 data model, file format, API, library, and tools are open source and distributed without charge.*"

### HDAQ mapping to HDF5 data storage model

## Case studies

### Imaging: Finite frame scan movie

### Imaging: Continuous frame scan movie

Until stopped by the user.

### Imaging: Z-stack acquisition

### Imaging: Mosaic z-stack acquisition

### Imaging: ROI scan

### Ephys: Continuous measurement and plotting

### Ephys: Finite measurement

### Ephys: Stimulation and measurement

### Imaging & Ephys: In parallel and decoupled measurement of Ephys and continuous frame scan movie

### Imaging & Ephys: Synchronized finite frame scan movie and recording of Ephys

### Imaging & Ephys: Synchronized point recording of fluorescence and Ephys

### Imaging & Ephys: Two-photon glutamate uncaging and simultaneous point recording of fluorescence and Ephys

### Imaging & Ephys: Two-photon uncaging, light induced dendritic inhibition and recording of Ephys