An Ontological Scheme for Specifying Time in HDF5

Ana Elena Buleu
Electrical Engineering Department
Center for Computation and Technology
Louisiana State University
Faculty Advisor: Dr. Gabrielle Allen
Research Advisors: Dr. Werner Benger,
Ms. Shalini Venkataraman
Hurricane Katrina

Heavy social and economic effect on the Gulf Coast...

At Louisiana State University a concentrated effort was directed towards visualization research. Why?

Scientific Visualization can be used to provide the capability to better communicate forecast models for such events.
Raw datasets…

Cloud data -->

Atmospheric data -->

Surge data -->
Combined datasets

Cloud data + Atmospheric data + Surge data

- each kind of data has a different origin, which is why it has time notions that are different from the other two

- for them to be combined, we need a common time standard --> we need an API to help handle time issues
Time notions

**Time semantics** -> What does this mean?

- 7 calendars currently used worldwide. The Gregorian calendar is the international standard
- Leap days, weeks, months (intercalation)
- Time zones
- Atomic Time, Universal Time (UT), Coordinated UT, Dynamical Time (relativistic and non-relativistic), Geocentric Coordinate Time, Barycentric Coordinate Time, Sidereal Time
- Calendar and non-calendar time
1. Permit the user to specify both points in time as well as intervals (in order to avoid round-off errors when subtracting two time stamps, each of which might have a different level of accuracy).

2. Allow the usage of arbitrary units and scales, and conversions between different time semantics.

3. Be as much as possible architecture (both software and hardware) independent (should not use calls, or structures specific to certain operating systems i.e. struct timeval – Unix specific data type).

4. Maintain precision of the original time specification (i.e. no loss in accuracy due to time unit conversions or data type conversions).

5. Permit indexing of datasets within HDF5 files.

6. Have a reasonable range of specification for a given time unit (i.e. avoid issues like ‘the year 2038 problem’, when the 32 bits used to define one memory location on most computer systems will not be enough to store time any more on Unix systems, which count time in seconds starting since January 1, 1970).
What is HDF5?

- I/O library and a file format designed for storing scientific data
- Hierarchical structure
What is the problem?

- ‘The time datatype, H5T_TIME, has not been fully implemented and is not supported. If H5T_TIME is used, the resulting data will be readable and modifiable only on the originating computing platform; it will not be portable to other platforms.’
  (http://www.hdfgroup.org/HDF5/doc_1.8pre/doc/H5.user/Datatypes.html)

- Other file formats designed for scientific data storage do not satisfy our needs either…
What is F5?

It is a C API built on top of HDF5.

Why? What is HDF5 missing?

HDF5 provides the syntax for building data models, while F5 adds semantics to it.

‘What does a number mean?’ is a question that we are trying to answer.
typedef enum {
    UNSPECIFIED,
    UNITLESS,
    NANOSECONDS,
    MICROSECONDS,
    MILLISECONDS,
    SECONDS,
    MINUTES,
    HOURS,
    DAYS,
    YEARS,
    MEGAYEARS,
    ELECTRONVOLTS,
    METRES
} F5_TimeUnits;
Value ? Offset ? Datetime ?

```c
struct __F5_TimeParameter{
    double  value;
    time_t  offset;
    time_t  datetime;
    F5_TimeUnits TimeUnits;
    char   *comment;
};
```
What functionality have we added so far?

- Set the time unit globally for the given file.
- Read the time unit information which is specified in the file.
Future work

- Conversions between different time semantics.
- Capability for being very specific about time (e.g. time zones, different calendars, leap seconds, various timescales).
I would like to thank the Center for Computation and Technology for providing me with the funding to work on this project, as well as write the research paper and present it at this conference. Moreover, I would like to thank the people who have made it a success because of their scientific expertise and continuous effort and patience in guiding me through my research: Dr. Werner Benger, Ms. Shalini Venkataraman, Dr. Gabrielle Allen.
The End…

Questions ?