An Overview of the HDF5 Technology Suite and its Applications

Mike Folk, Gerd Heber, Quincey Koziol,
Elena Pourmal, Dana Robinson

The HDF Group

{mfolk,gheber,koziol,epourmal,derobins}@hdfgroup.org

Outline

Brief History

The HDF5 technology suite

Applications

Feature supply and demand

Brief History

1987: Graphics task force at NCSA began work on architecture-independent format and library, **HDF**.

1990: NSF provided funding to improve documentation, testing, and user support.

1994: NASA selected HDF as standard format for Earth Observing System.

1996–1998: DOE tri-labs and NCSA, with additional support from NASA, developed **HDF5**, initially called "BigHDF".

2005: NASA funded development of netCDF-4, a new version of netCDF that uses the HDF5 file format.

2006: **The HDF Group**, a non-profit corporation, spun off from NCSA and the University of Illinois.

The HDF5 Technology Suite

HDF5 **abstract data model**

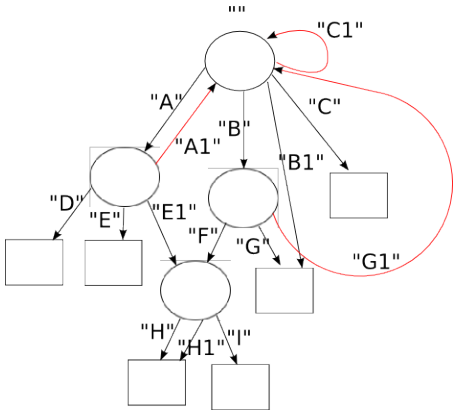
The “building blocks” for data organization and specification

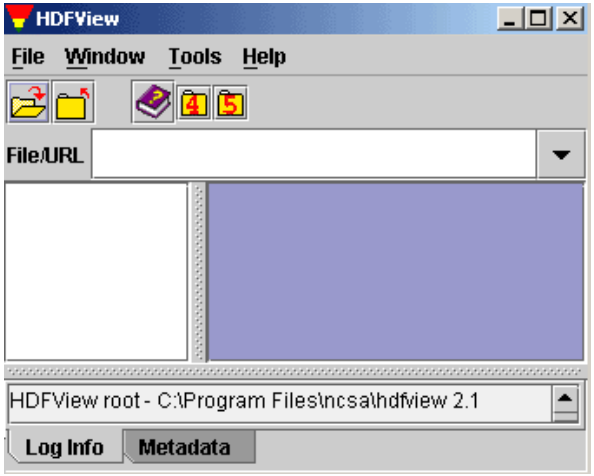
HDF5 **software**

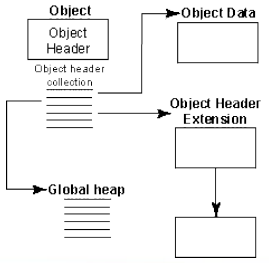
Library, language interfaces, tools

HDF5 **file format**

Bit-level organization of HDF5 file







HDF5 Abstract Data Model

An HDF5 information set is a container for annotated associations of array variables and types.

Container (HDF5 File)

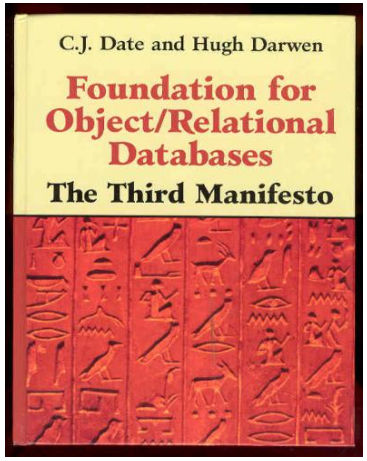
Annotation (HDF5 Attribute)

Association (HDF5 Group, Link)

Array variables (HDF5 Dataset)

Types (HDF5 Datatype)

We love and admire The Third Manifesto but decided to go into a different direction.



~~*"All information in the database at any given time must be cast explicitly in terms of values in relations and in no other way"*~~
(Codd's Information Principle, Date/Darwen: *Third Manifesto*)

A Few Analogies

File system (in a file)

Binary XML file

PDF for numerical data

Database (container for array variables)

Click to edit Master text style
Second level
- Third level
  - Fourth level
    - Fifth level

Some analogies are less flattering than others…

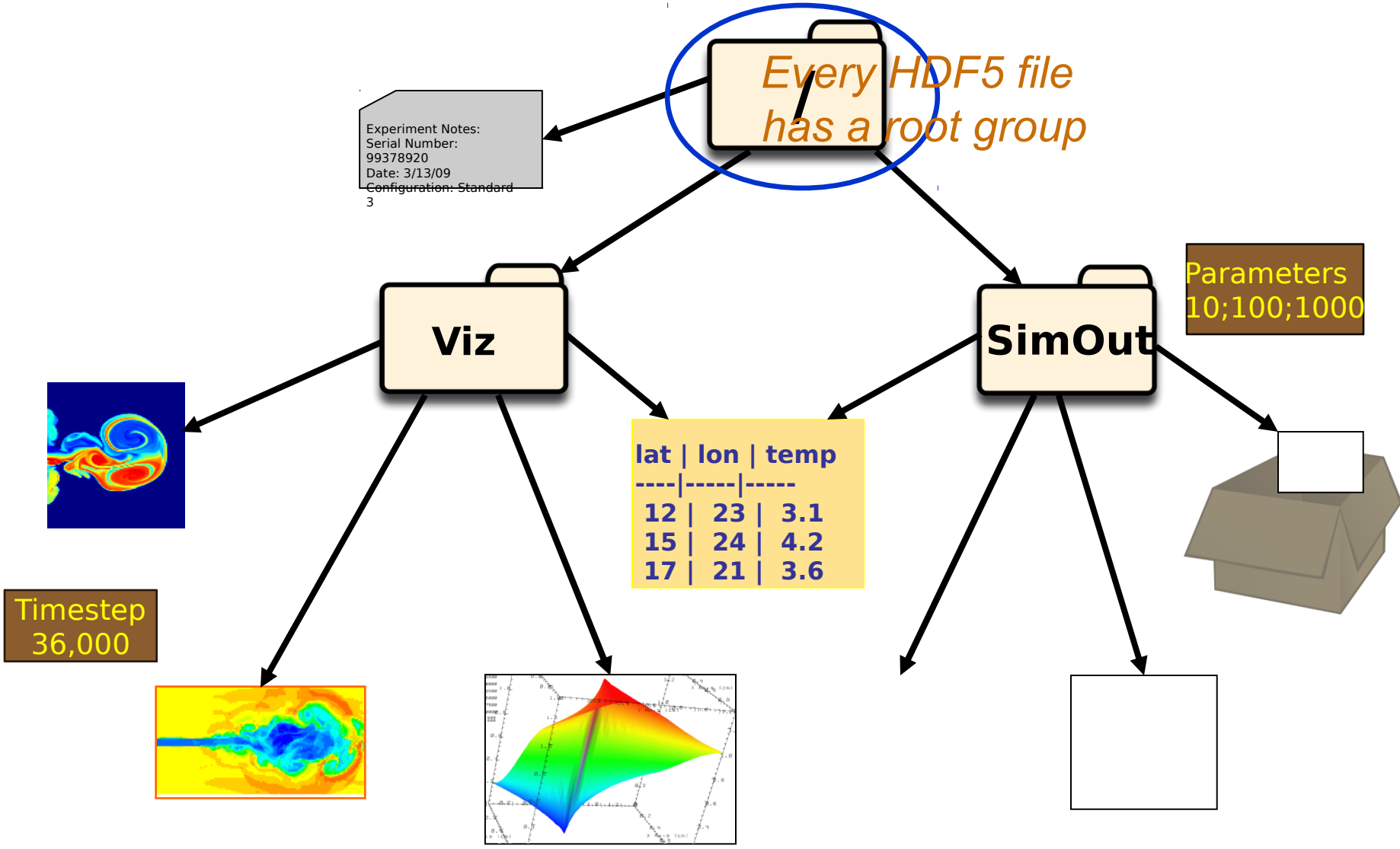
Analogy [ad. L. analogia, a. Gr. ἀναλογία equality of ratios, proportion (orig. a term of mathematics, but already with transf. sense in Plato), f. ἀνάλογ-ος adj.: see analogon. Cf. mod.Fr. analogie.]

...

3. Equivalency or likeness of relations; 'resemblance of things with regard to some circumstances or effects' (J.); 'resemblance of relations' (Whately); a name for the fact, that, the relation borne to any object by some attribute or circumstance, corresponds to the relation existing between another object and some attribute or circumstance pertaining to it. Const. to, with, between.

This is an extension of the general idea of proportion from quantity to relation generally, and is often expressed proportionally, as when we say 'Knowledge is to the mind, what light is to the eye.' The general recognition of this analogy makes light, or enlightenment, or illumination, an analogical word for knowledge. [**Source:** OED 2nd Edition]

HDF5 Information Association



HDF5 Datasets (Array Variables)

Metadata

Dataspace

Rank 3

Dimensions Dim_1 = 4, Dim_2 = 5, Dim_3 = 7

Datatype

IEEE 32-bit float

Storage info

Chunked

Compressed

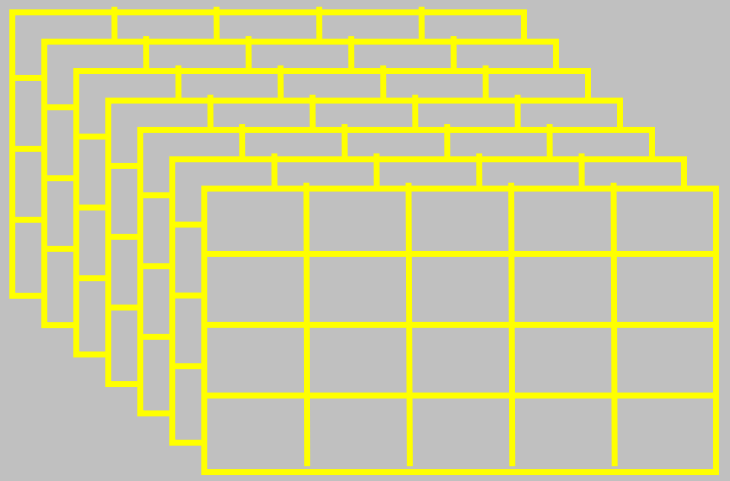
Attributes

Time = 32.4

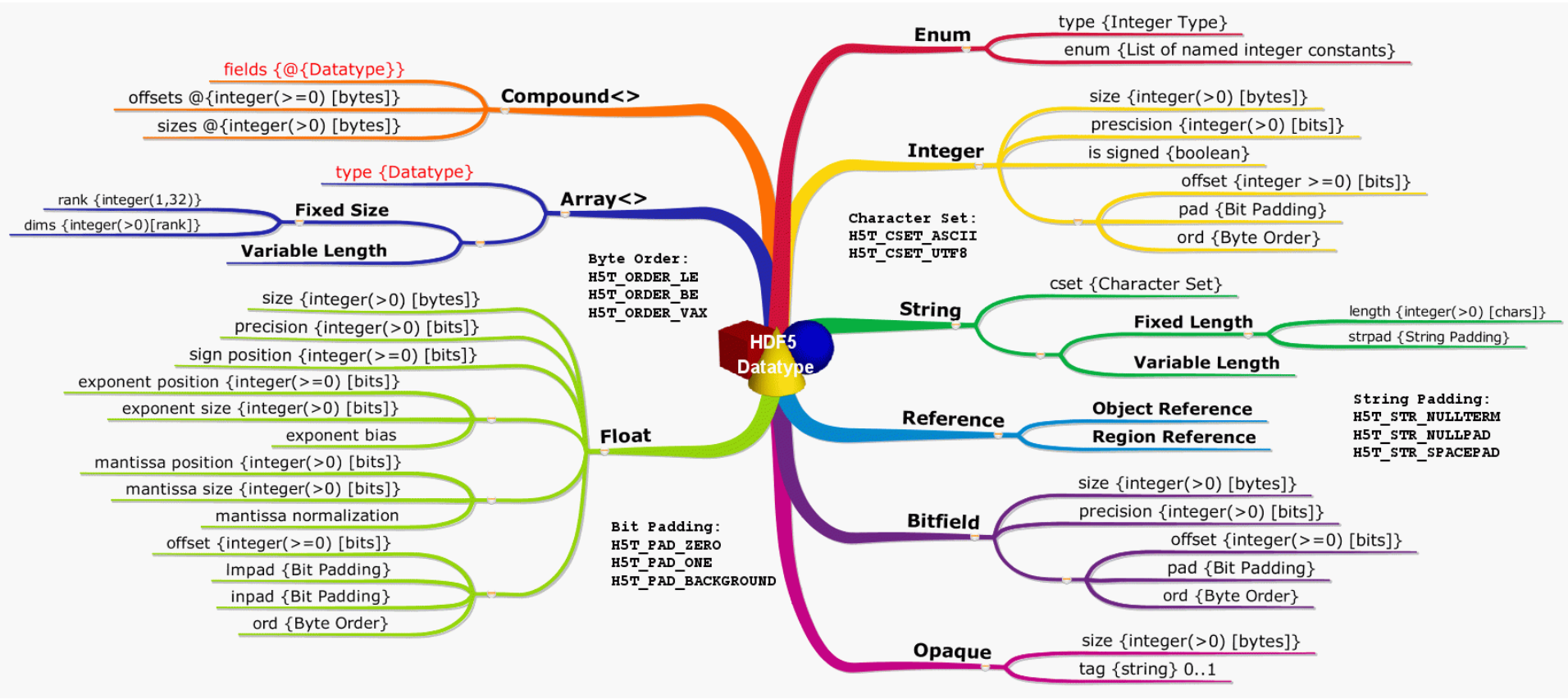
Pressure = 987

Temp = 56

Dataset data



HDF5 Datatypes

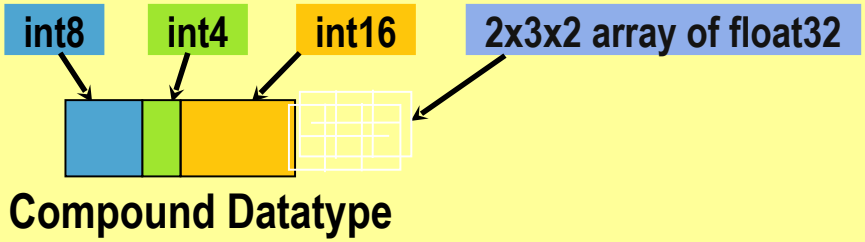


- Unlimited mix and match
- DIY integers, floats etc.
- Embarrassment of riches
- All-you-can-handle buffet (**Careful!**)

Example: Compound Datatype

Committed Datatype

PurpleGreenRedBlue



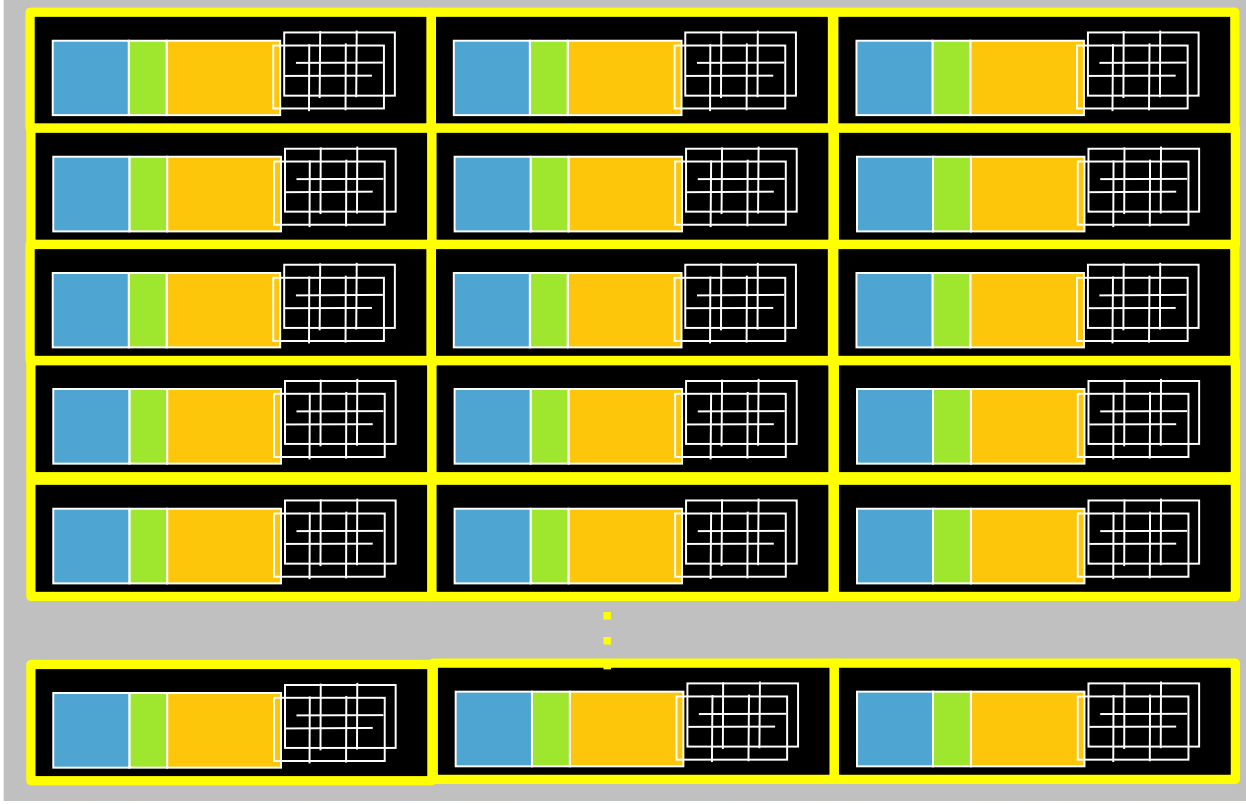
HDF5 Dataset

HDF5 Datatype

PurpleGreenRedBlue

HDF5 Dataspace

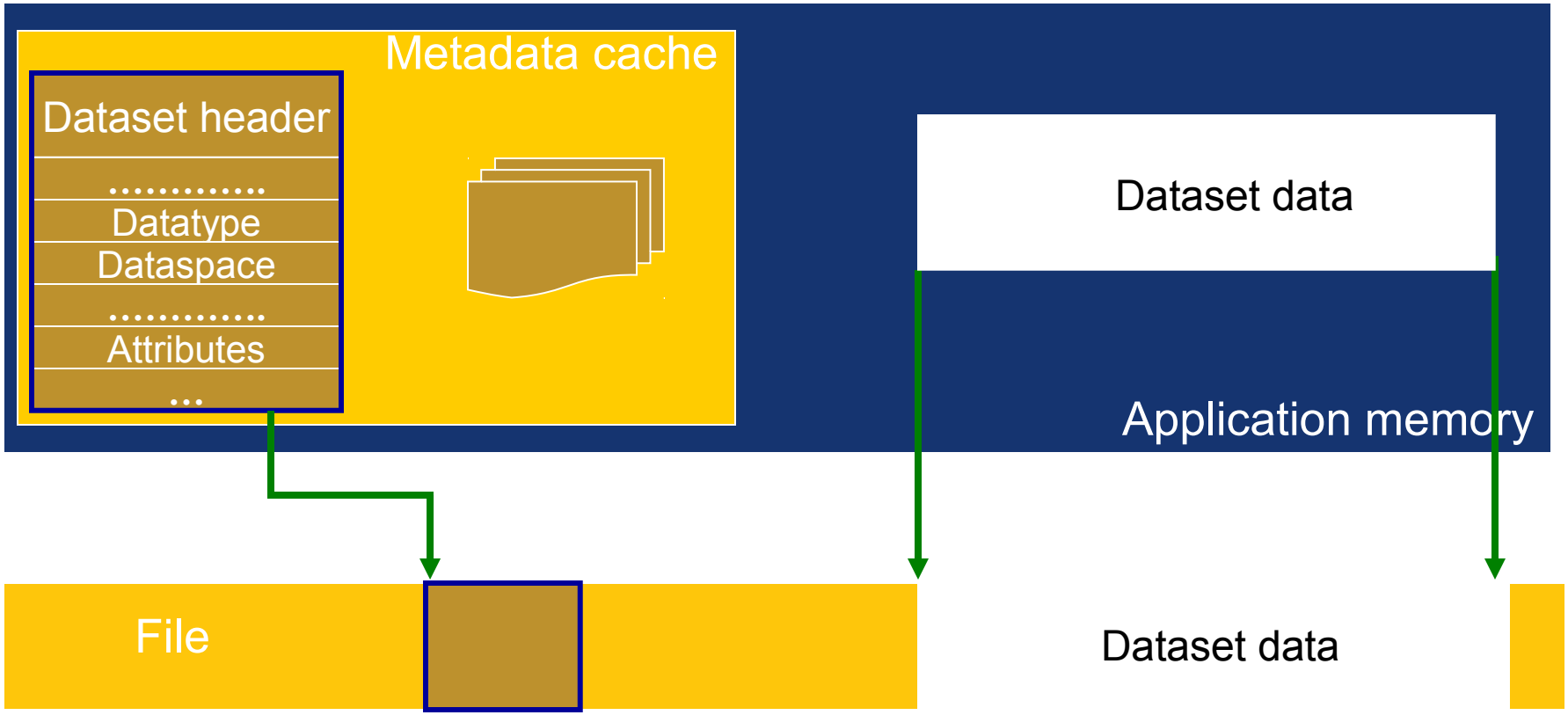
Rank	DimSizes	Max DimSizes
2	5	unlimited
	3	3



Contiguous Layout

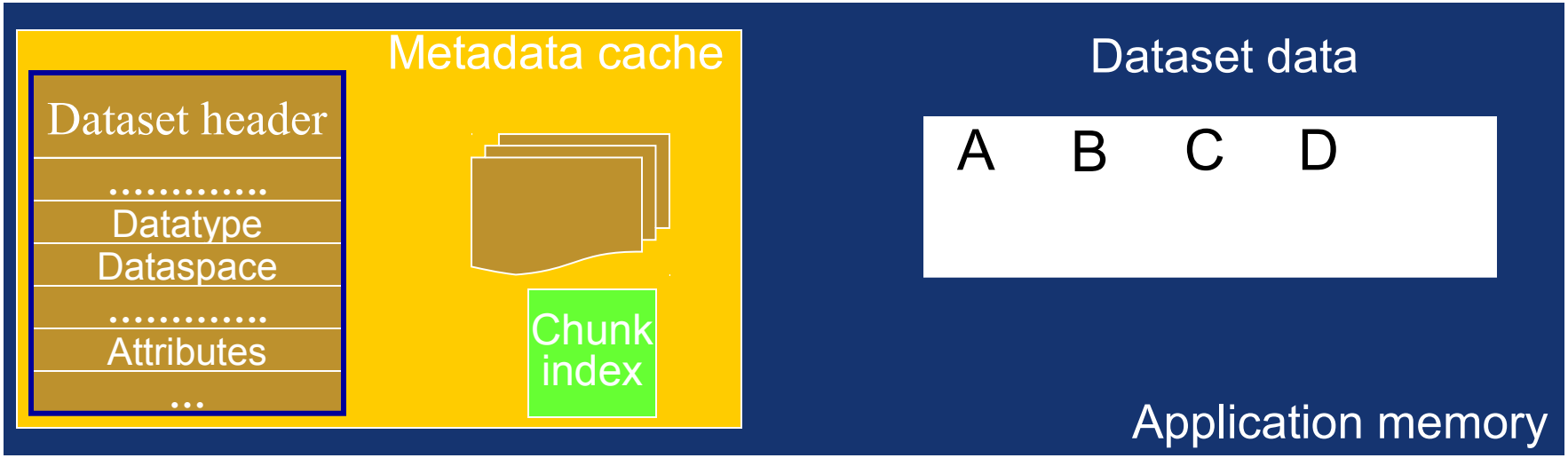
Metadata header separate from dataset data

Data stored in one contiguous block in HDF5 file

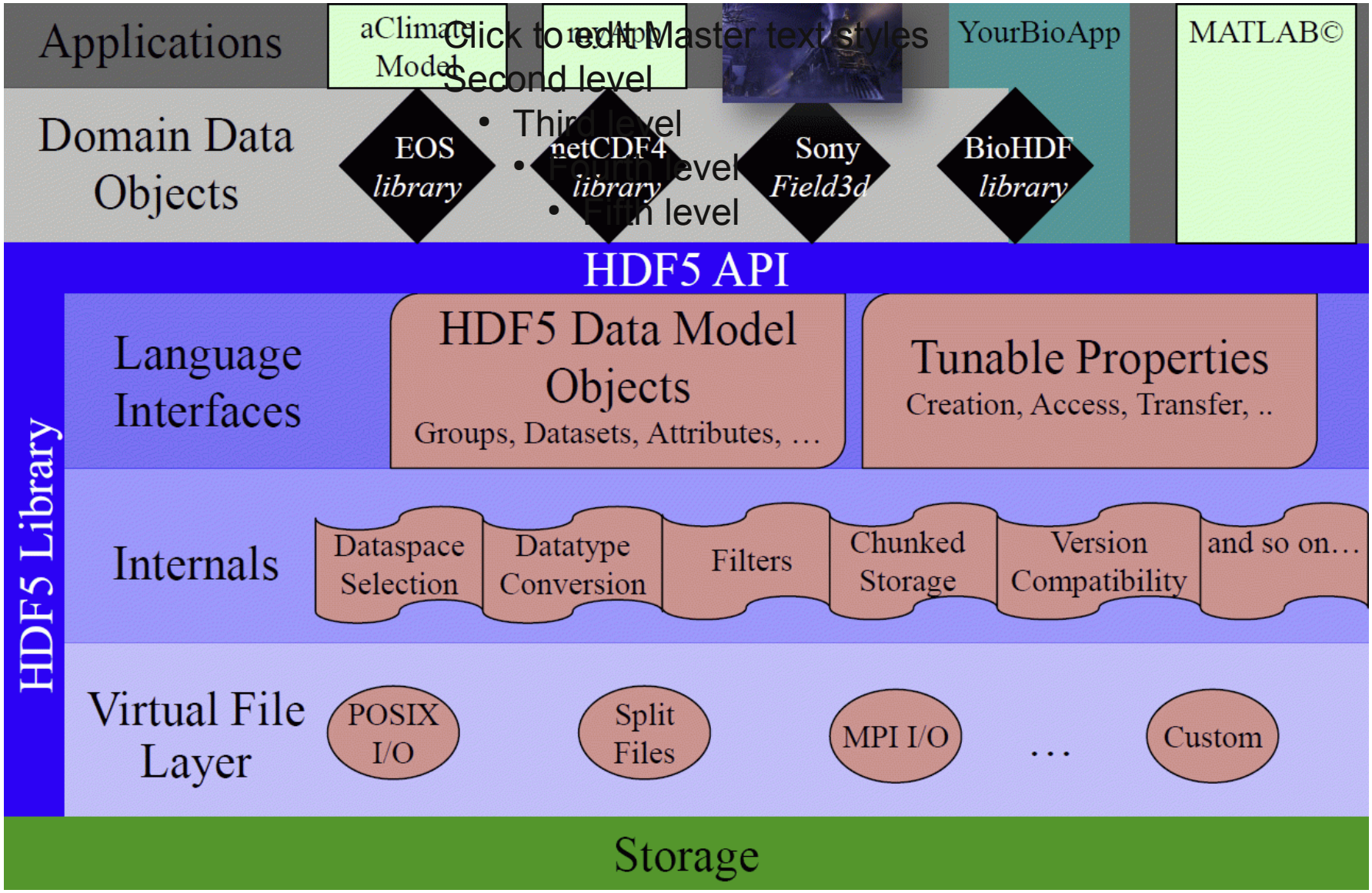


Chunked Layout

Dataset data is divided into equally sized blocks (chunks). Each chunk is stored separately as a contiguous block in HDF5 file.



HDF5 Software Architecture



HDF5 Data Pipeline

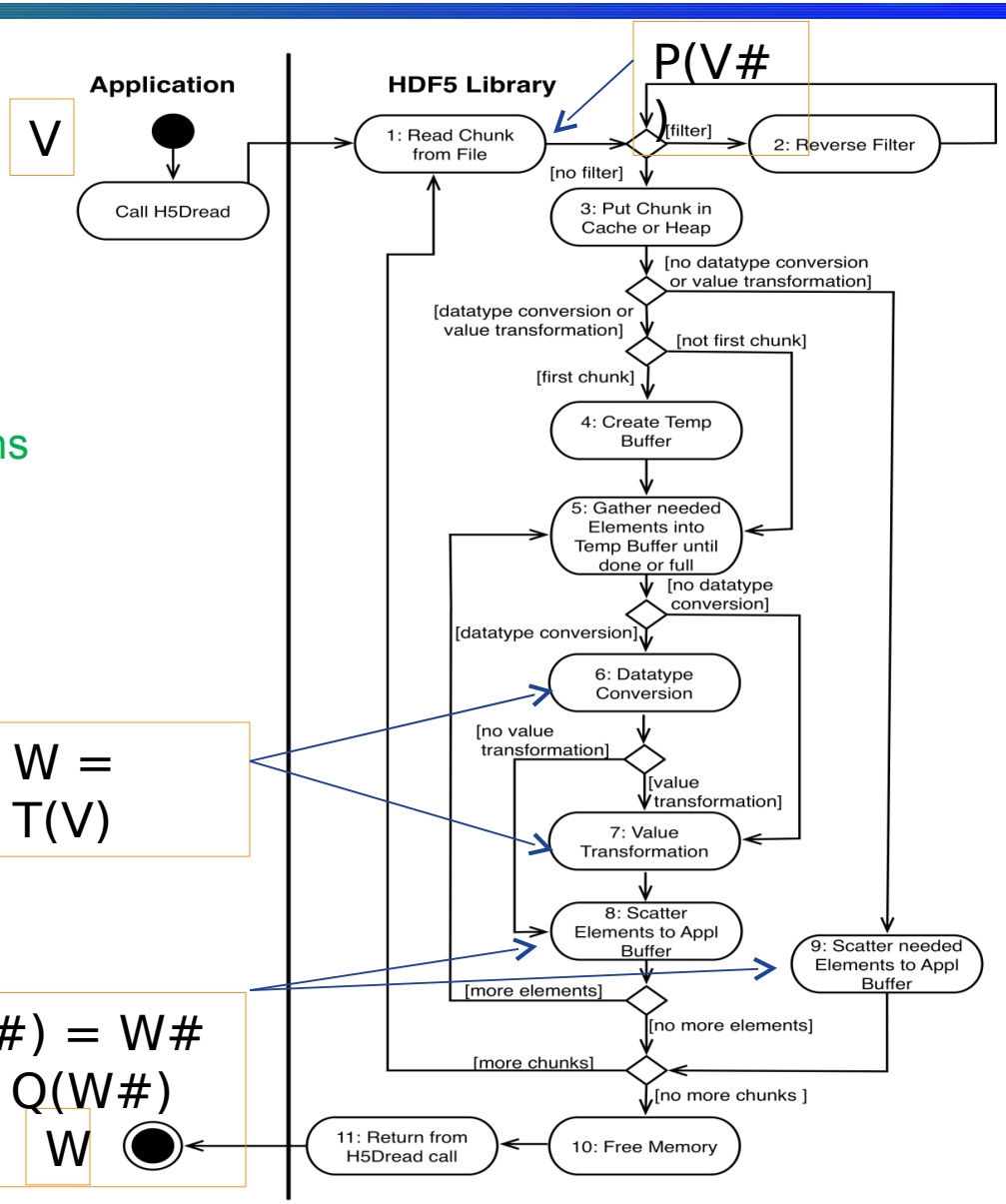
T - Data Transformation or Filter

UPDATE W SET W = **T(V)**

FROM V JOIN W ON S(V#) = W#

WHERE **P(V#)** AND **Q(W#)**

P, **Q** - Hyperslab and/or point selections



It is the "marriage of a format and an access library" that "allows one to access the data without knowing anything about the actual representation of the data layout of the file."

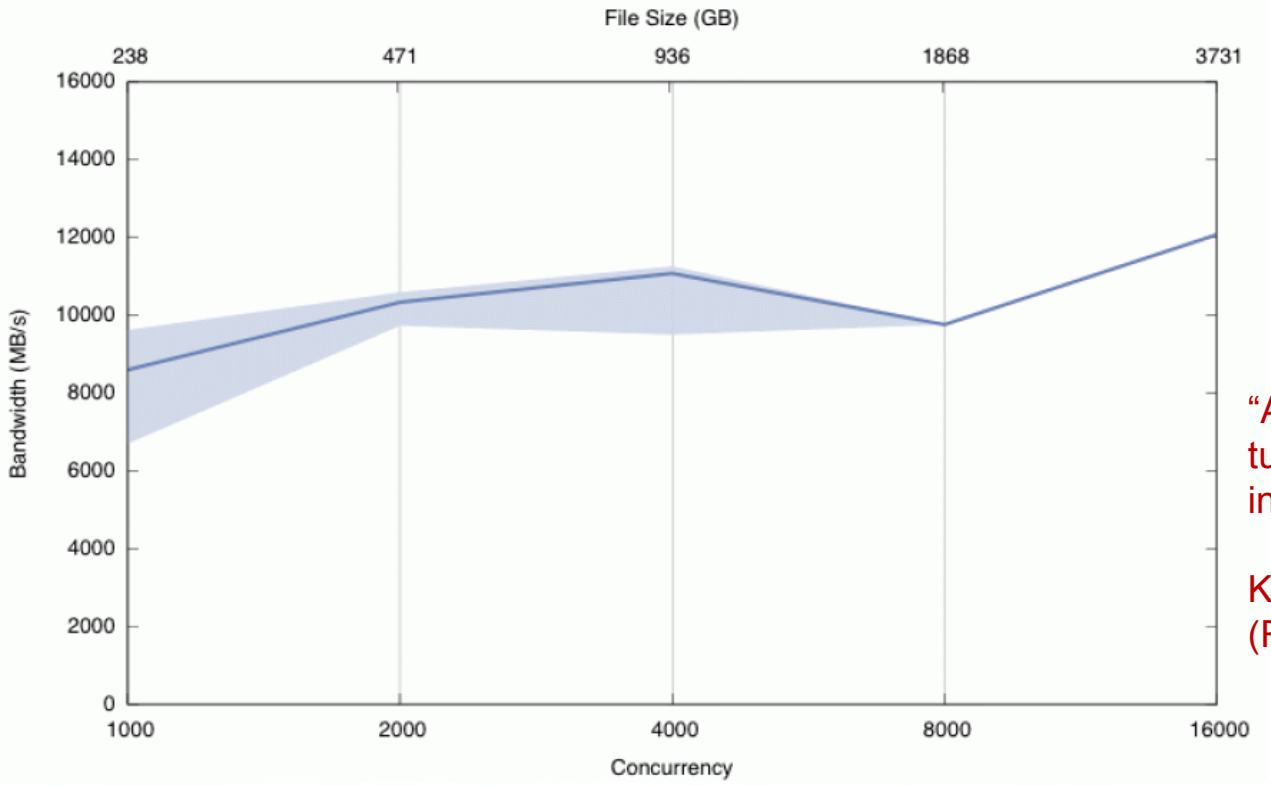
(Jeff Kuehn, netCDF)

Parallel I/O in HDF5

Joint project w/ NERSC

Goal: same performance w/ HDF5 as w/ MPI-IO

Up to 12 GB/s to shared file (out of 15 GB/s) on NERSC's Franklin system (Cray XT4)



"A supercomputer is a device for turning compute-bound problems into I/O-bound problems."

Ken Batcher
(Ray Paden, GPFS Best Practices)

HDF5 & XML

XML representation(s) for HDF5

Leverage XML machinery

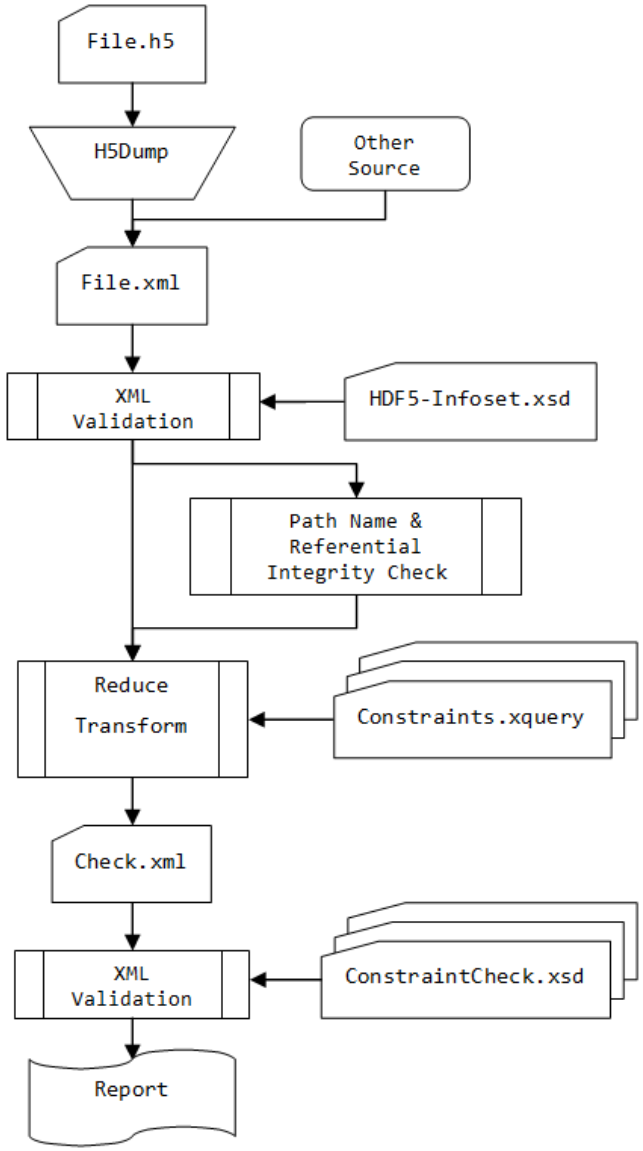
HDF5 profiles

Constraints and validation

Limitations of XML schema

XQuery to the rescue

Automation



HDF5 as a DBMS Backend

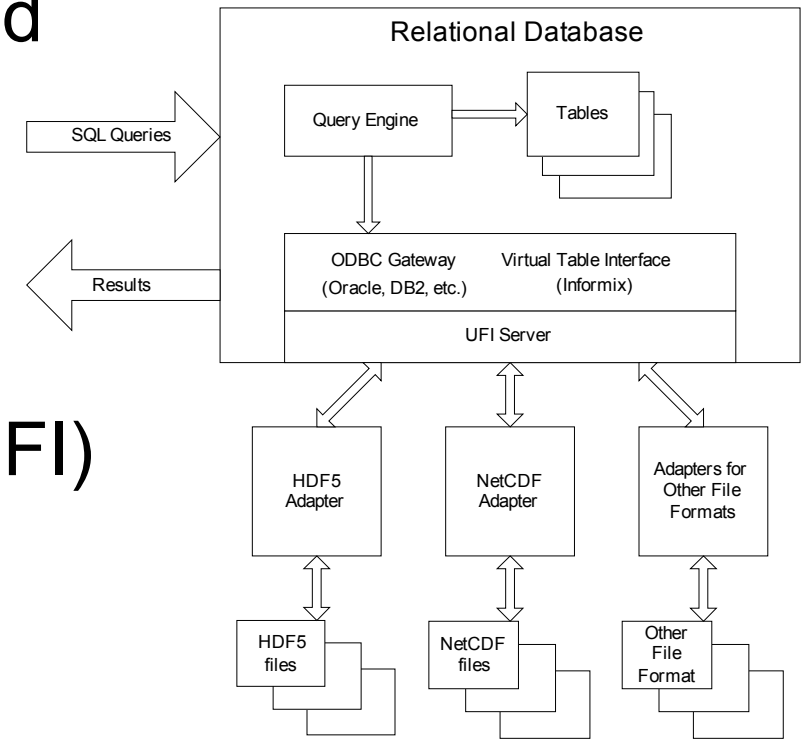
Old (pipe?)dream: RDBMS and non-relational data

External data management SQL/MED (SQL:2003)

BCS Universal File Interface (UFI)

UFI w/ HDF5

1. Define virtual table
2. Associate table w/ one or more HDF5 files



Demo and Documentation at
http://www.barrodale.com/bcs/universal-file-interface-ufi

Server-side vs. client-side indexing

Applications of HDF5

We are aware of 100+ distinct applications

Some users prefer to remain anonymous†

All sectors (enterprise, government, research)

HDF-Forum: http://hdf-forum.184993.n3.nabble.com/

Third-party tool support

Two examples:

LOFAR (LOw-Frequency ARray, radio astronomy)

- One slide by courtesy of Anastasia Alexov, Astronomical Institute Anton Pannekoek, Amsterdam

BioHDF (Bioinformatics/Genomics)

- Four slides by courtesy of Todd Smith, Geospiza, Inc.

† Not because they are ashamed of using HDF5…

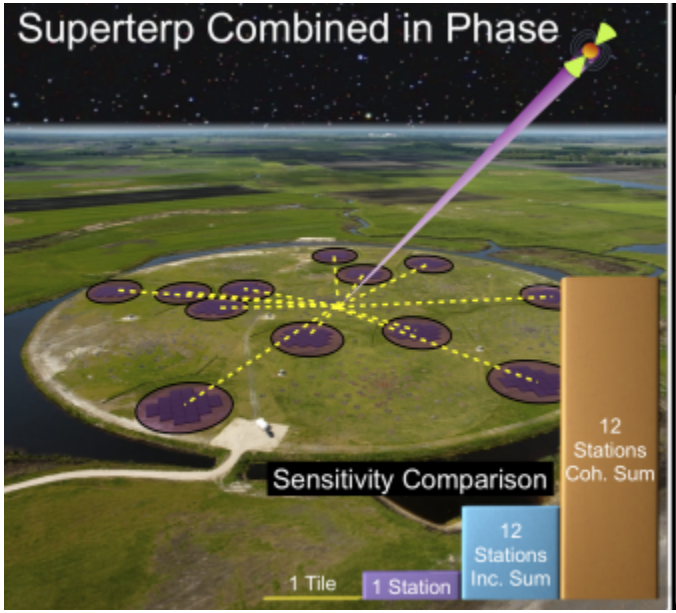
LOFAR

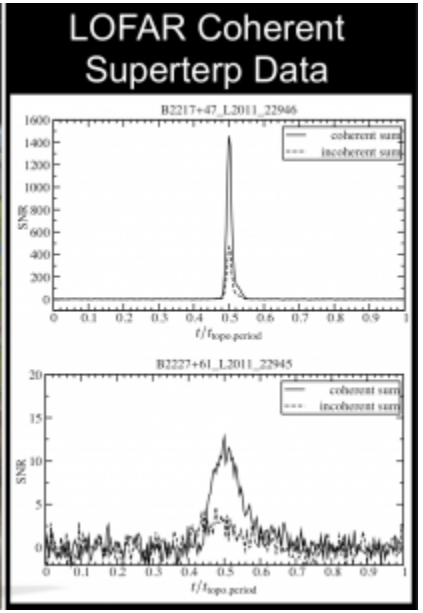
Real-time multiple sensor array

Exponential increase in file size over last decades

Most of LOFAR's standard data products stored in HDF5

Epoch	Nominal File Data Volume
1970	210 bytes
1980	220 bytes
1990	230 bytes
2000	240 bytes





[LOFAR] http://www.lofar.org/

[LOFAR&HDF5] Anderson et al. *Toward a New Radio Data Standard*, PoS(ISKAF2010)062, arXiv:1012.2266v1

[Alexov2010] Alexov, A. *LOFAR: Data Challenges and HDF5*, ADASS XX, November 2010

LOFAR Sample Dataset Sizes

Exposure Time	Number of Subbands	Number of Stations	File Size Known Mode	File Size Search Mode
1 min	248	5	11.2 GB	56 GB
1 min	248	20	11.2 GB	244 GB
10 min	248	5	112 GB	560 GB
10 min	248	10	112 GB	1.1 TB
10 min	248	20	112 GB	2.2 TB
10 min	248	30	112 GB	3.3 TB
20 min	248	5	224 GB	1.1 TB
30 min	248	5	336 GB	1.7 TB
1 hr	248	5	672 GB	3.4 TB
1 hr	248	10	672 GB	6.7 TB
1 hr	248	20	672 GB	13.4 TB
1 hr	248	30	672 GB	26.8 TB
2 hr	248	5	1.3 TB	6.7 TB
12 hr	248	5	8.0 TB	40.3 TB
12 hr	248	15	24.0 TB	120.1 TB

Why LOFAR chose (yet) another data format: Hierarchical Data Format, version 5 (HDF5)

Question: Can only ONE of the astronomical formats (like FITS or CASA) do ALL these things?

- HDF5 is a data model, library, and file format for storing and managing large and complex scientific data (images, N-D arrays, tables, metadata).
- It supports an unlimited variety of datatypes, and is designed for flexible and efficient I/O and for high volume and complex data.
- Self-describing and portable to a diversity of computational environments
- No inherent file size limitations; no header Attributes/keys character length limits
- C, C++, Java, Fortran 90 interfaces
- Can be run on single node or massively parallel/distributed systems (~600 cores)
- Built-in compression (GNU zlib, but can be replaced with others)
- Parallel reading and writing (via MPI-I/O)
- Partial I/O: “Chunked” (tiled) data for faster access
- Free and in use for 20+ years by NASA and other projects
- Inspection and visualization tools exist (HDFView + command line tools, Visit + pluggin, PyTables, h5py, MATLAB)

BioHDF Project

NIH STTR

Geospiza, Seattle WA

The HDF Group, Urbana/Champaign IL

Goal: Move bioinformatics problems from organizing and structuring data to asking questions and visualizing data

Develop data models and tools to work with NGS data in HDF5

Create HDF5 domain-specific extensions and library modules to support the unique aspects of NGS data => BioHDF

Integrate BioHDF technologies into Geospiza products

Deliver core BioHDF technologies to the community as open-source software

Next Generation DNA Sequencing

"Transforms today's biology"

"Democratizing genomics"

NGS is Powerful

"Changing the landscape" "Genome center in a mail room"

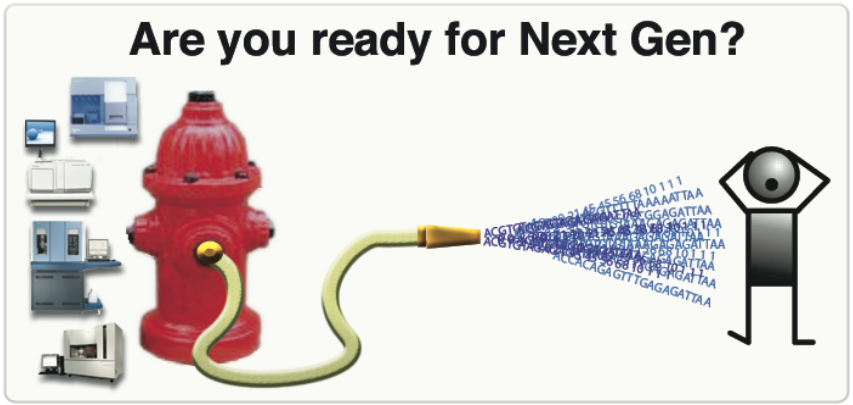
"The beginning of the end for microarrays"

… And Daunting

"Prepare for the deluge"

"Byte-ing off more than you can chew"

"These sequencers are going to totally screw you"

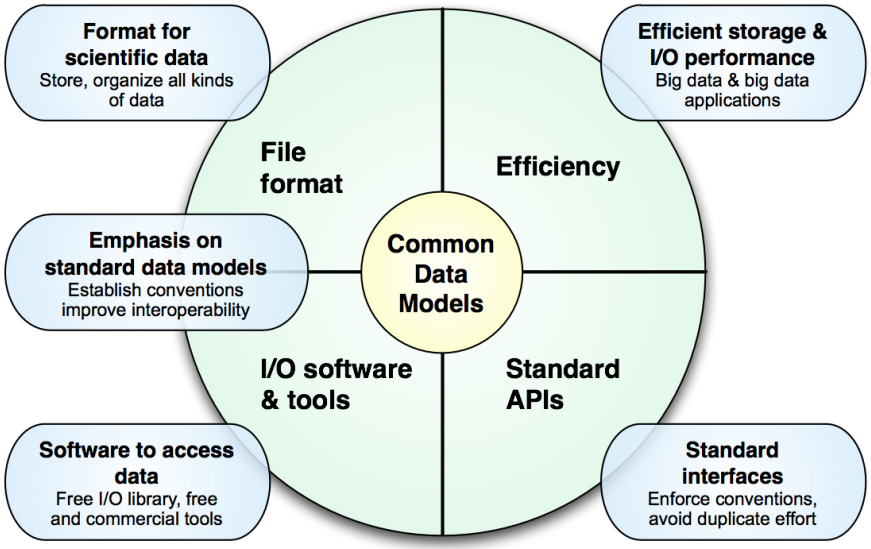


"If the data problem is not addressed, ABI's SOLiD, 454's GS FLX, Illumina's GAII or any of the other deep sequencing platforms will be destined to sit in their air-conditioned rooms like a Stradivarius without a bow."

Why HDF?

HDF5: 20 Years in Physical Sciences

HDF - Hierarchical Data Format



- Arrays, rich data types, groups accommodate every kind of data
- Store any combination of data objects in one container.
- Performance: fast random access **and** efficient, scalable storage
- Portability, data sharing: platform independent, self describing, common data models
- Tools for viewing, analysis: HDFview, MATLAB, others
- Widespread: used in academia, govt, industry - MATLAB, IDL, NASA-Earth Observing System

A platform for creating software to work with many kinds of *scientific* data

geospiza

HDF5 Philosophy

A platform of many use(r)s

Open is better than secretive.

Communal is better than lonesome.

“Pauca sed matura.”† (C.F. Gauss)

One abstract data model

Visible is better than obscure.

Types are better than stereotypes.

Pragmatic is better than dogmatic.

One software library

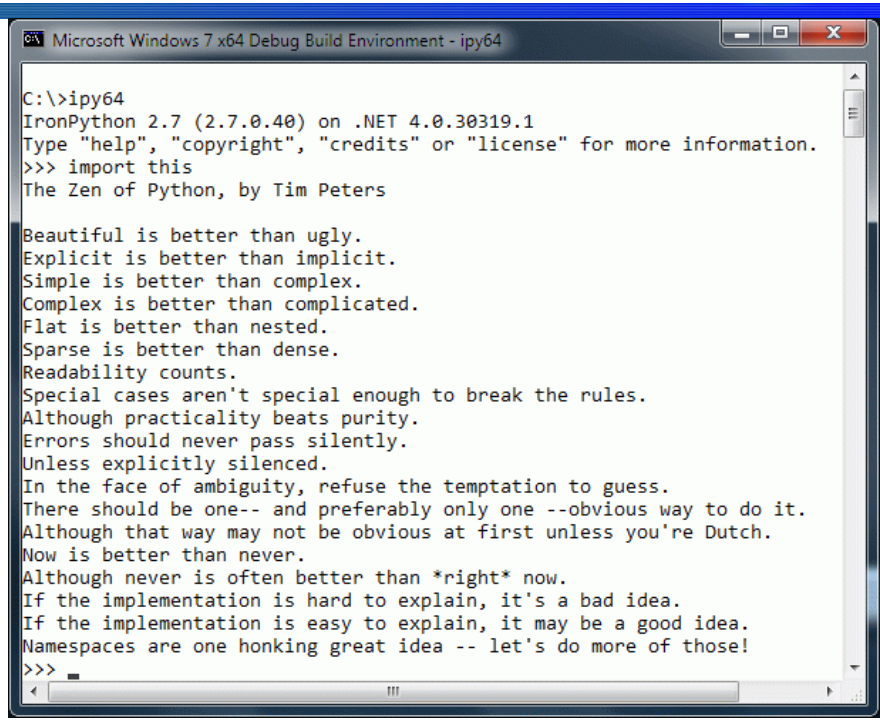
Adaptation embraces change.

In-between is better than at the extreme.

One general file format

Self-describing is better than documented.

Speed and storage efficiency matter.



† Few, but ripe.

What Users Get…

A capability that:

Lets them **organize** large and/or complex collections of data

Gives them efficient and **scalable** data storage and access

Lets them integrate a wide **variety** of types of data and data sources

Takes advantage of *rapidly* **evolving** data and storage technologies, but through *slowly* maturing interfaces

Guarantees **long-term** data integrity and preservation

What Users Want…

Fancy tools

Safety net

Performance

Standardized indexing

Query engines

Engage the ArrayDB community!

Thank You!

Questions & Comments?