Spatio-Temporal Satellite Database for Trend Analysis in Forest Ecosystem

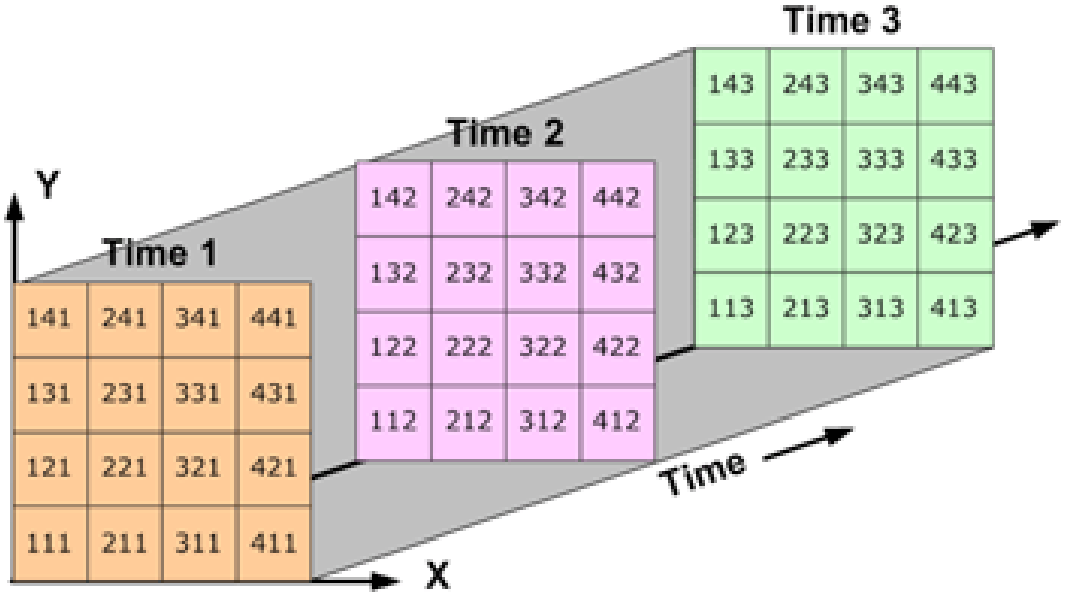
Luka Rumora[1], Damir Medak[1], Ivan Pilaš[2]

[1] University of Zagreb, Faculty of Geodesy
[2] Croatian Forest Research Institute

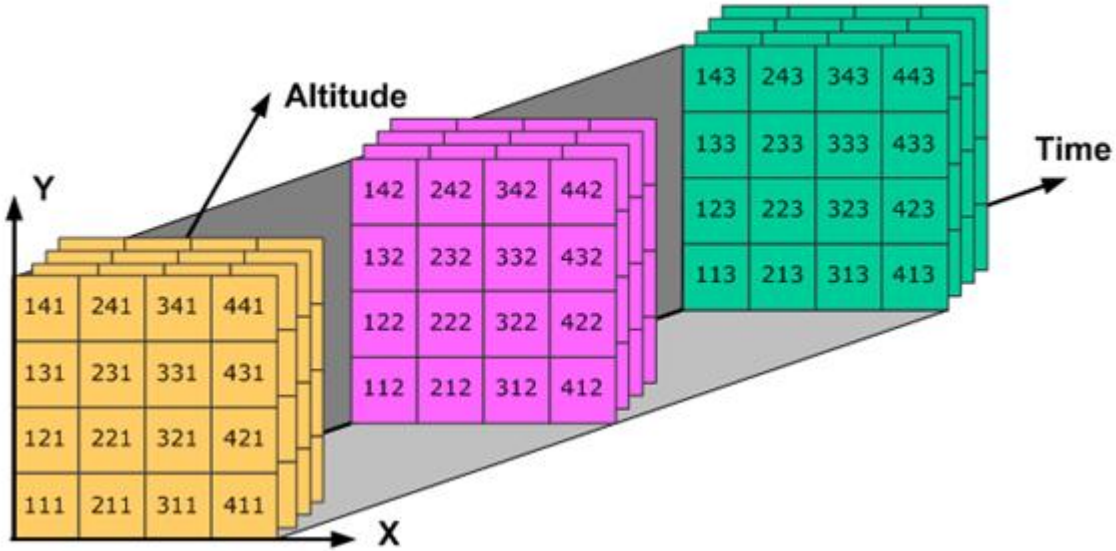
Mulidimensional raster storage – netCDF

- NetCDF (Network Common Data Form)
  - Software libraries and data formats
  - Creation, access sharing array-oriented data
  - header
- NetCDF data is:
  - Self-Describing
  - Portable
  - Direct-access
  - Appendable
  - Sharable
  - Archival
- commonly used in climatological and meteorological applications and for general scientific data exchange

NetCDF structure



Three-dimensional data: Data over an area varying with time



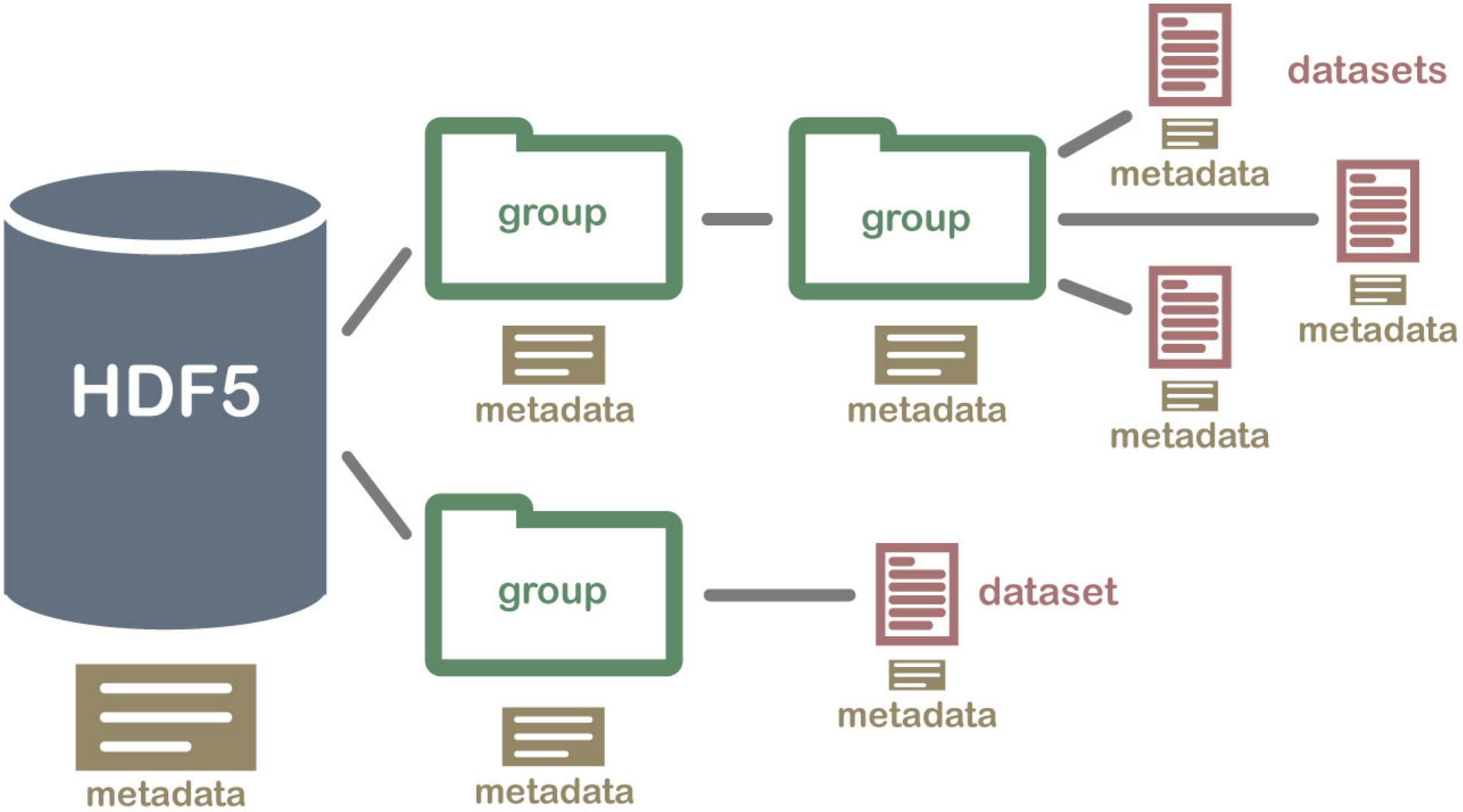
Four-dimensional data: Data over an area varying with time and altitude

http://desktop.arcgis.com/en/arcmap/10.3/manage-data/netcdf/fundamentals-of-netcdf-data-storage.htm

HDF (Hierarchical Data Format)

- physical file format for storing scientific data
- Collection of utilities and application – manipulation, viewing, analyzing
- HDF5 composed of two primary types of objects:
  - HDF5 group
  - HDF5 dataset
- Hierarchical, filesystem-like data format
- NetCDF based on HDF5
- time series data – stock price series, 3D meteorological data
- HDF5 data storage mechanism can be simpler and faster than an SQL schema

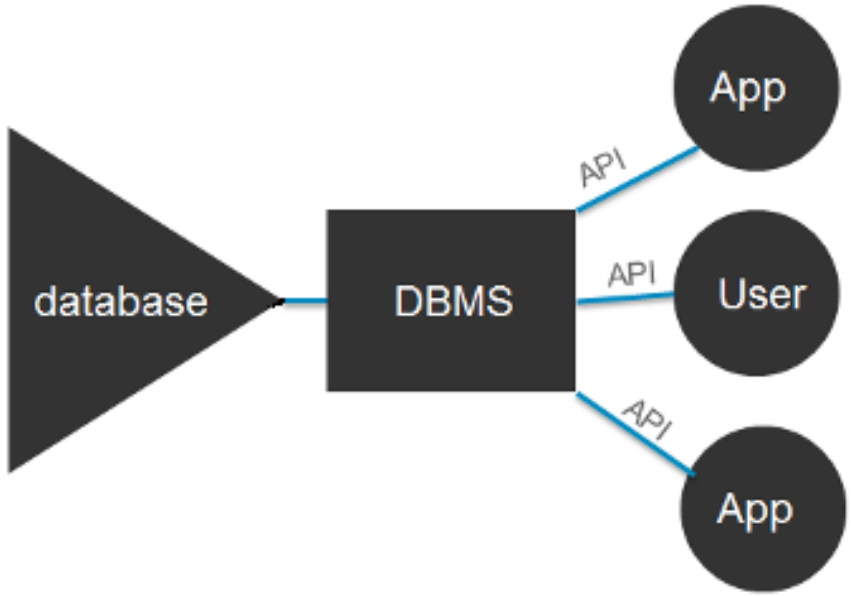
HDF structure



http://neondataskills.org/images/HDf5/hdf5_structure4.jpg

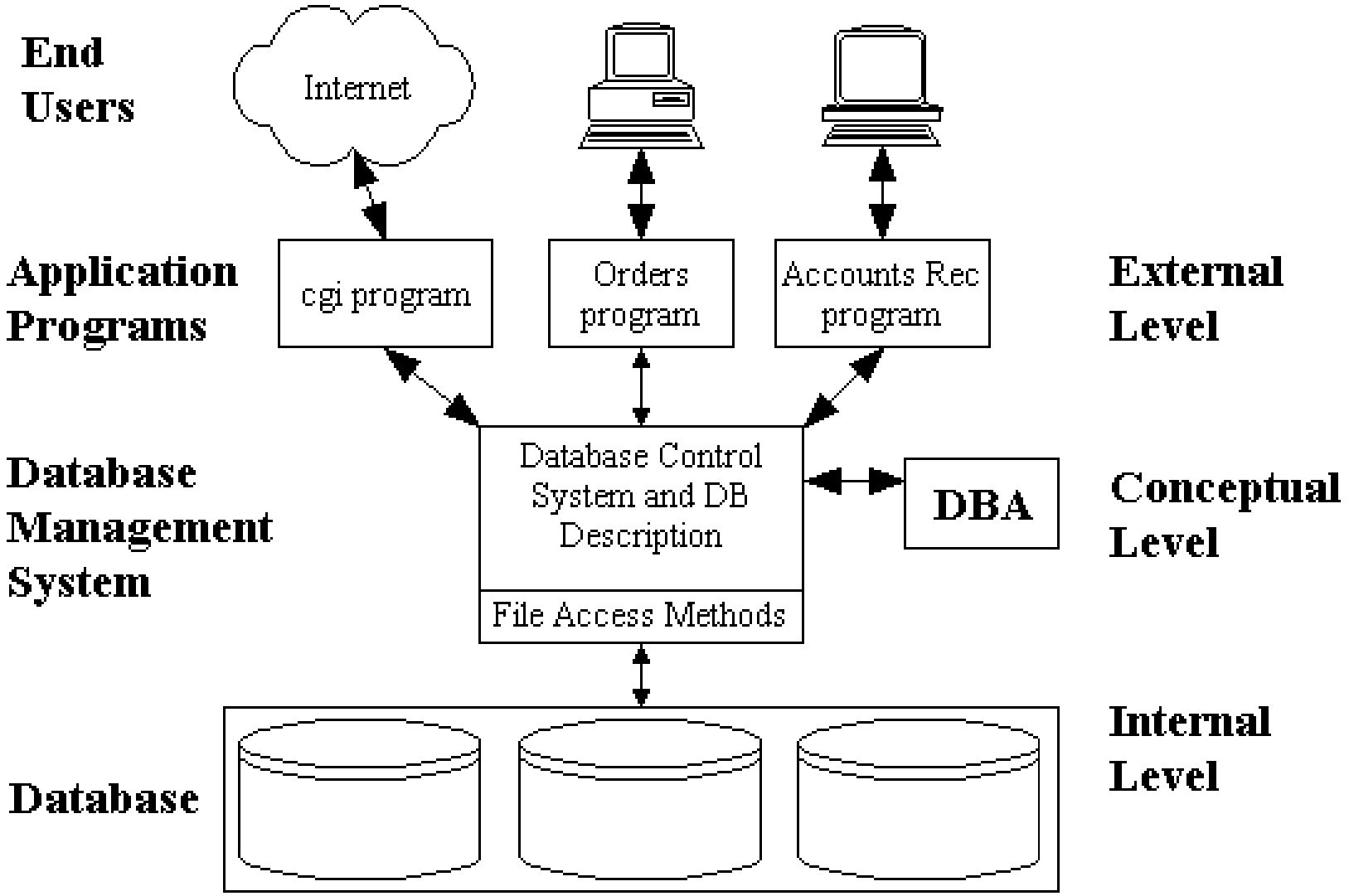
Database

- Database – collection of information
- Database management system (DBMS) – system software for creating and managing database



http://itknowledgeexchange.techtarget.com/overheard/files/2015/01/DBMS-database-management-system.png

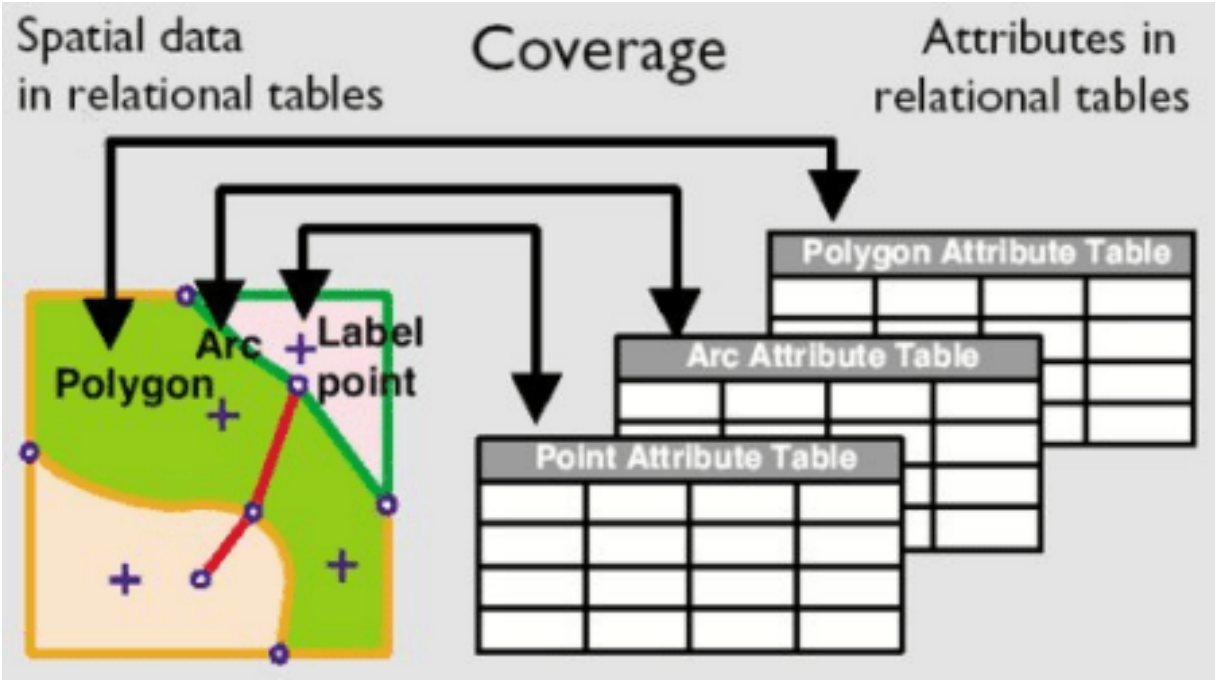
Database arhitecture



http://ironbark.xtelco.com.au/subjects/DB/2010s2/gifs/archlevs.gif

Relational database model

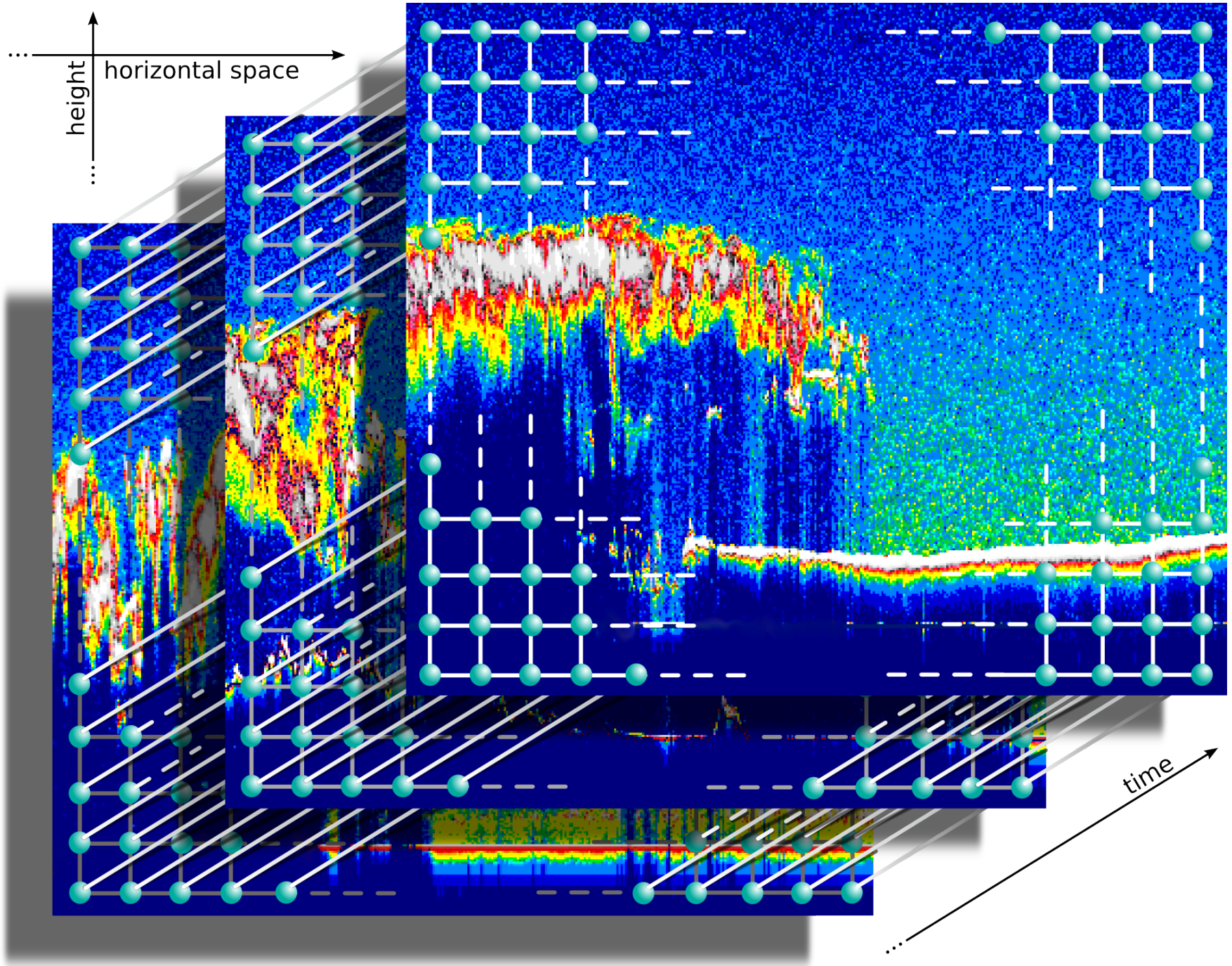
- Relational model database – Relational database
- Related records
  - linked with a "key"



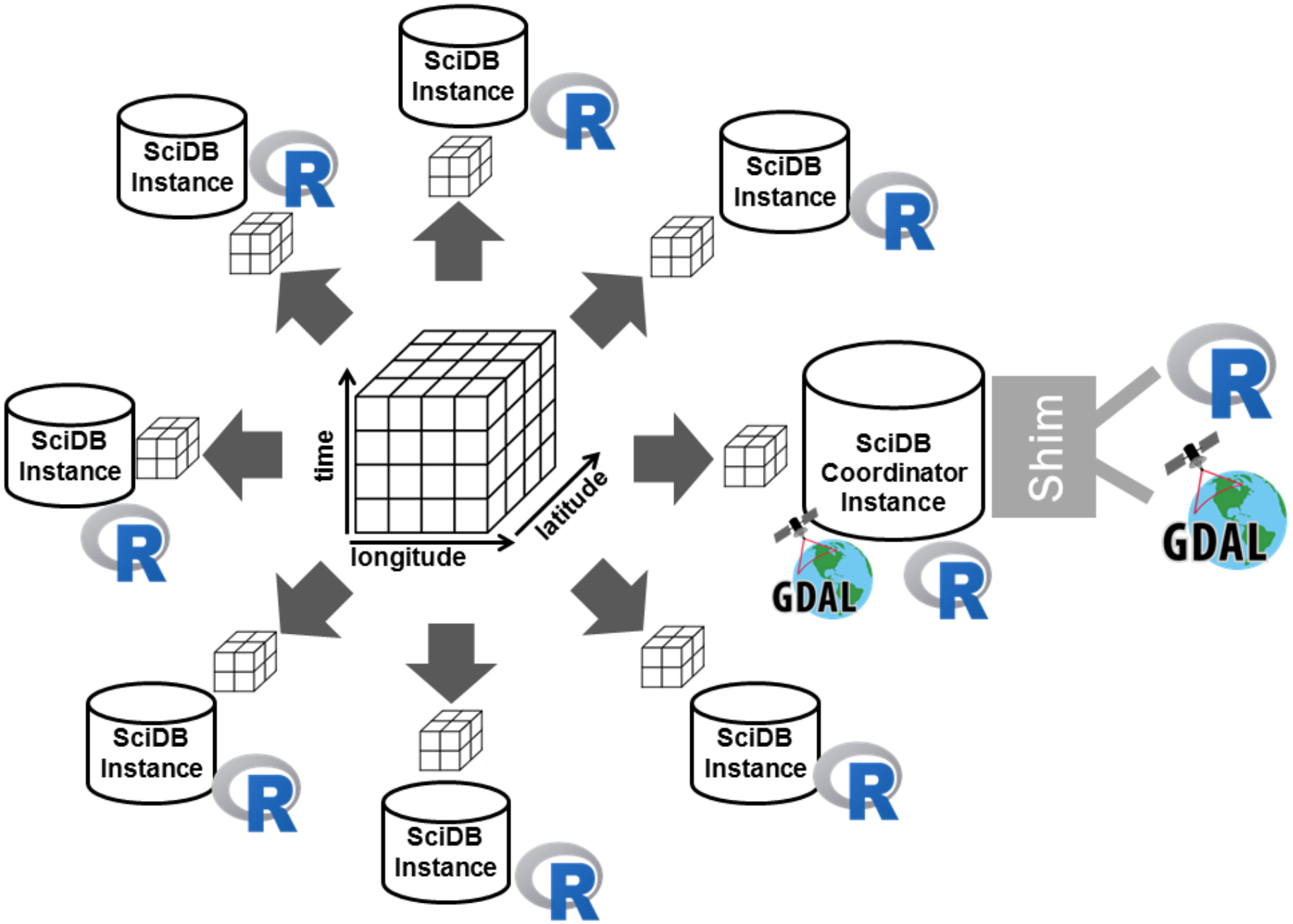
http://www.geog.ucsb.edu/~jeff/gis/data_models1_files/covdiag.gif

Array database model

- Array database management system (DBMS) – designed for arrays (raster data)
- Raster data – terabytes of data
- Different dimensions and sizes
- List of Array DBMS
  - Oracle GeoRaster
  - MonetDB/SciQL
  - PostGIS Raster
  - rasdaman
  - SciDB

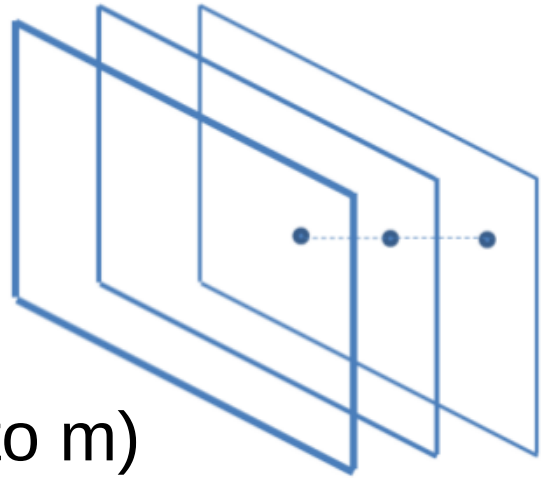


www.rasdaman.org



Lu, Meng, et al. "Spatio-temporal change detection from multidimensional arrays: Detecting deforestation from MODIS time series." ISPRS Journal of Photogrammetry and Remote Sensing 117 (2016): 227-236.

Example – relational vs array database

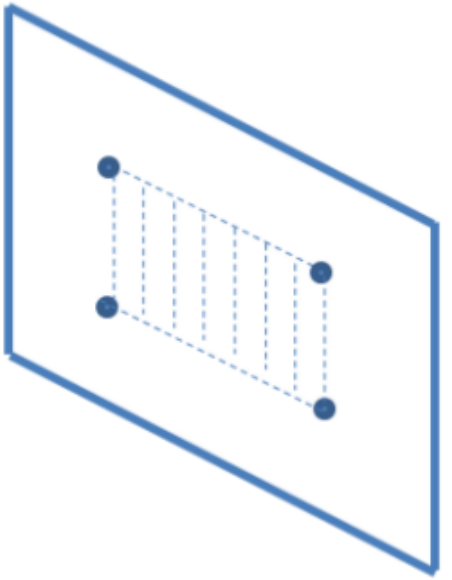


- 1. query: Fetch the spectrum-band values of the point (x=100, y=100) for images (n to m)

	SciDB (1st Model)	SciDB (2nd Model)	PostgreSQL	PostGIS Raster
(n=1 , m=10)				
Length (char)	230	84	1388	80
Performance (sec)	39	1.8	0.4	0.15
(n=1 , m=50)				
Length (char)	949	84	7193	80
Performance (sec)	1120	7.8	19	0.8

Hausen, E. (2016.) Array-database Model (SciDB) for Standardized Storing of Hyperspectral Satellite Images, dissertation

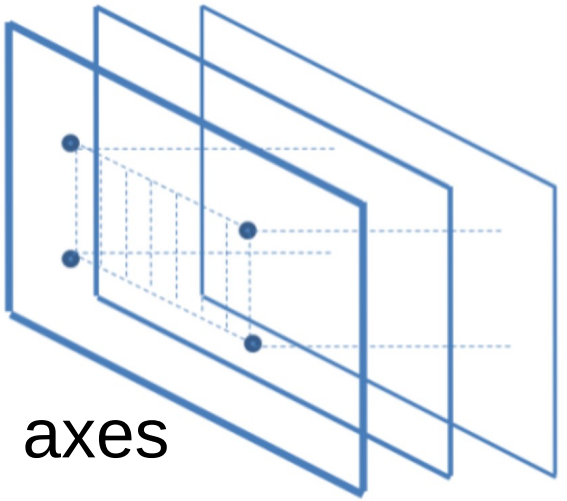
Example – relational vs array database



- 2. query: Fetch all the spectrum-band values of Image_30 for the following axes values (x1<x<x2)/(y1<y<y2)

	SciDB (1st Model)	SciDB (2nd Model)	PostgreSQL	PostGIS Raster
(x1=400 , x2=900) / (y1=1000 , y2=2000)				
Length (char)	89	92	6160	167
Performance (sec)	20	34	1.8	690
(x1=0 , x2=900) / (y1=0 , y2=2000)				
Length (char)	90	93	8224	167
Performance (sec)	67	89	2.1	1642

Example – relational vs array database



- 3. query: Fetch all the spectrum-band values of images (n to m) for the following axes values (x1<x<x2)/(y1<y<y2)

	SciDB (1st Model)	SciDB (2nd Model)	PostgreSQL	PostGIS Raster
(n=1 , m=5) , (x1=0 , x2=50) / (y1=0 , y2=50)				
Length (char)	166	103	7349	178
Performance (sec)	16	13	0.3	173
(n=1 , m=10) , (x1=0 , x2=900) / (y1=0 , y2=900)				
Length (char)	252	104	92880	181
Performance (sec)	96	112	343	1037

Conclusion

- Relational databases – past?
- Array databases
  - Time series compatible
  - possible future
- Problem – computer performance
- In-database storage or Out-database storage (metadata storage)