A Tool for Packaging and Exchanging Simulation Results

Dragan Savić
[e-mail: dragan.savic@fe.uni-lj.si]
Francesco Potortì
Francesco Furfari
Matevž Pustišek

University *of Ljubljana*
Faculty *of Electrical Engineering*

ISTITUTO DI SCIENZA E TECNOLOGIE
DELL'INFORMAZIONE "A. FAEDO"

COST 285 symposium

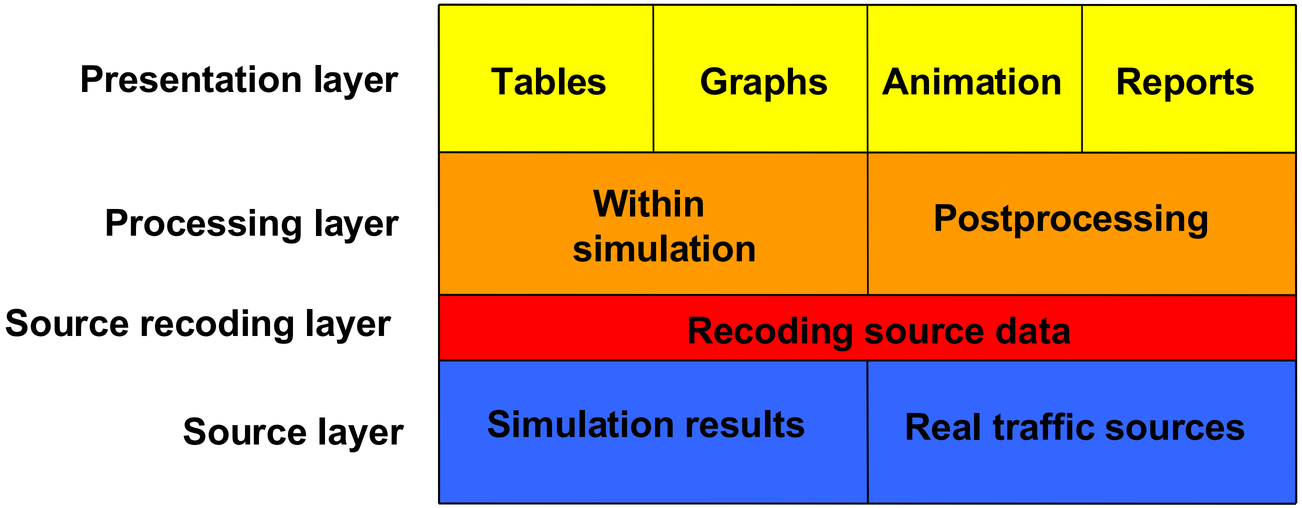
Surrey, March 29th 2007

Outline

- Reference model
- Objective
- Formats for exchanging data
- Data organization
- Metadata (XML)
- CostGlue tool architecture
- Module API
- Current status and short-term projects
- Conclusions

Reference model 1/3

- Layered decomposition of main functions of the simulation process



Presentation layer: Tables, Graphs, Animation, Reports

Processing layer: Within simulation, Postprocessing

Source recoding layer: Recoding source data

Source layer: Simulation results, Real traffic sources

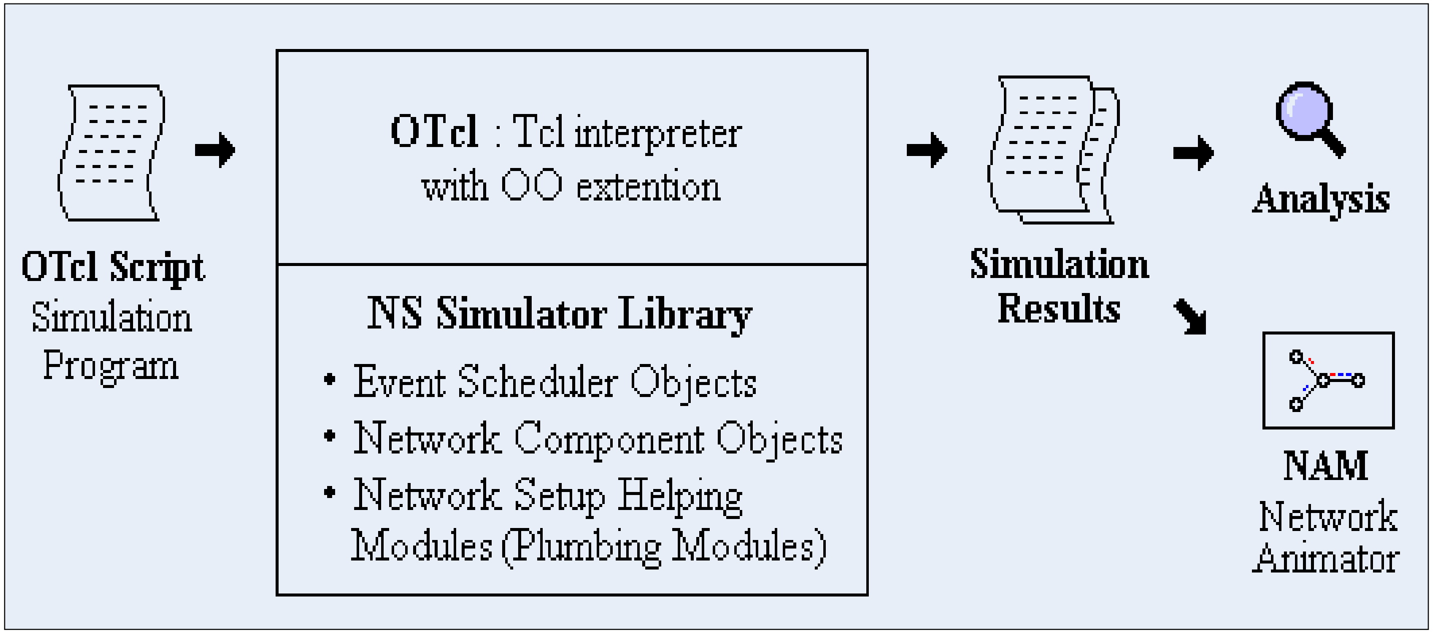
- Source layer
  - raw simulation output in form of tabular traces with ASCII or binary format
  - real network traffic traces e.g. traffic captured in a network link

Reference model 2/3

- Source recoding sublayer (optional)
  - conversion between different formats
  - data compression
  - removal of private information
- Processing layer – simulation data analysis
  - within simulation – processing carried out by simulator itself
  - postprocessing – different tools
- Presentation layer – tables, graphs, animation, reports

Reference model 3/3

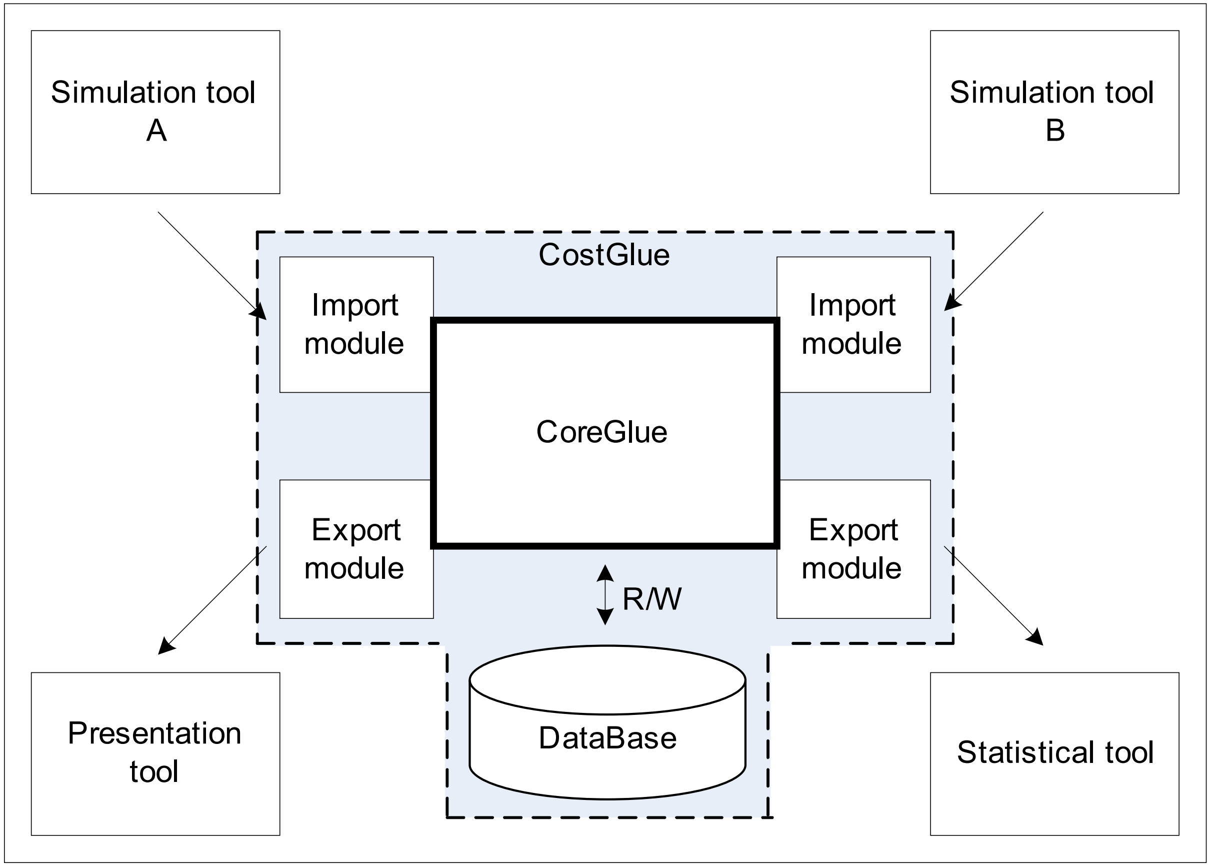
- Usage of different tools for different functional layers
  - simulation and data acquisition tools (e.g. Ns-2, OPNET, tcpdump)
  - statistical analysis tools (e.g. SPSS, awk scripts)
  - presentation tools (e.g. GNUPlot, MS Excel, nam)



OTcl Script
Simulation Program
OTcl : Tcl interpreter with OO extention
NS Simulator Library
• Event Scheduler Objects
• Network Component Objects
• Network Setup Helping Modules (Plumbing Modules)
Simulation Results
Analysis
NAM
Network Animator

Objective

- Facilitate the exchange and manipulation of simulation data
- Build a tool for efficient storage and extraction of huge quantities of data

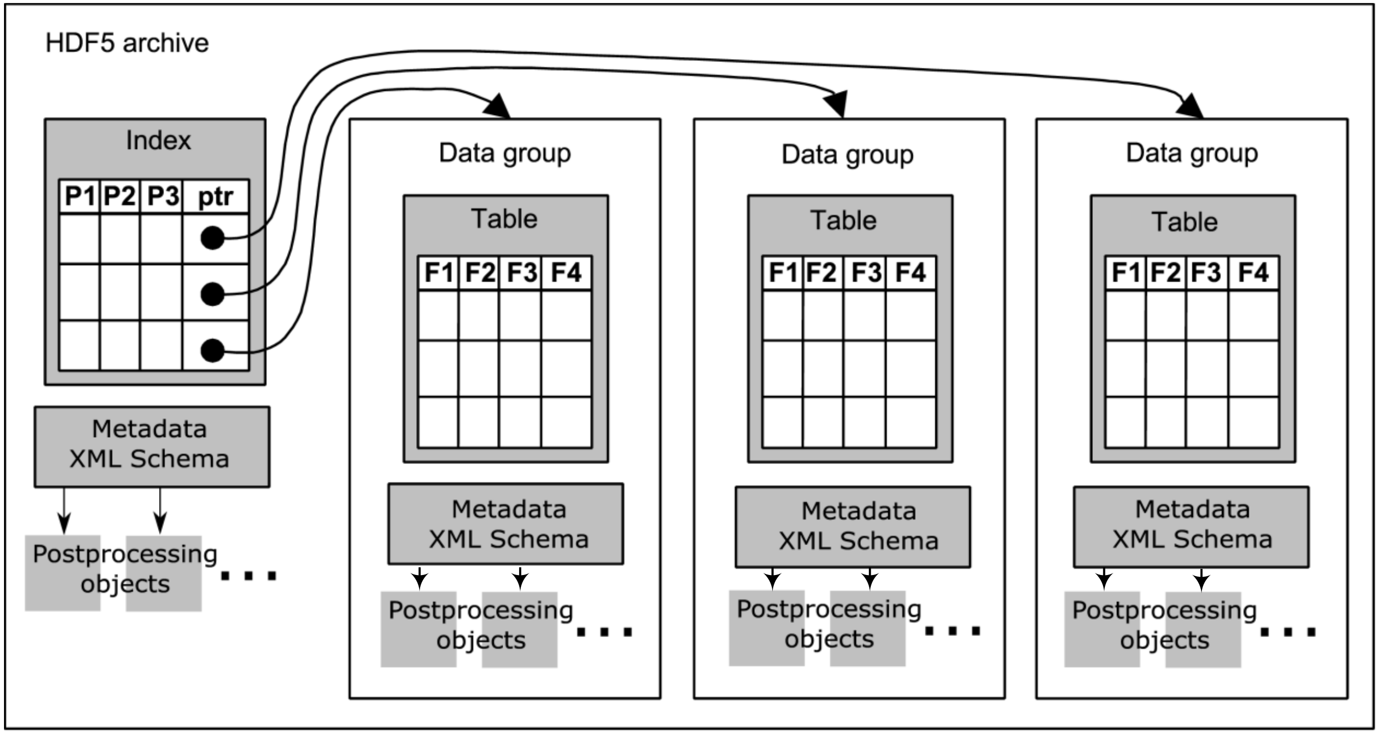


Formats for exchanging data

- The need for a common file format
- Analysis of different formats for scientific data:
  - HDF4 and HDF5 (Hierarchical Data Format)
  - netCDF (Network Common Data Format)
  - PDB (Portable Binary Database)
  - FITS (Flexible Image Transport System) and
  - OpenDX (Data Explorer Visualization Software)
- Results of analysis: HDF5 (NCSA) meets all the requirements of data organization and storage
- HDF5 format – short overview
  - two primary objects:
    - dataset – simple dataset (homogeneous data) or compound dataset (different number of data types within one dataset)
    - group – a structure containing zero or more HDF5 objects
  - support for adding attributes to HDF5 objects
  - hierarchical data organization – tree structure of HDF5 objects

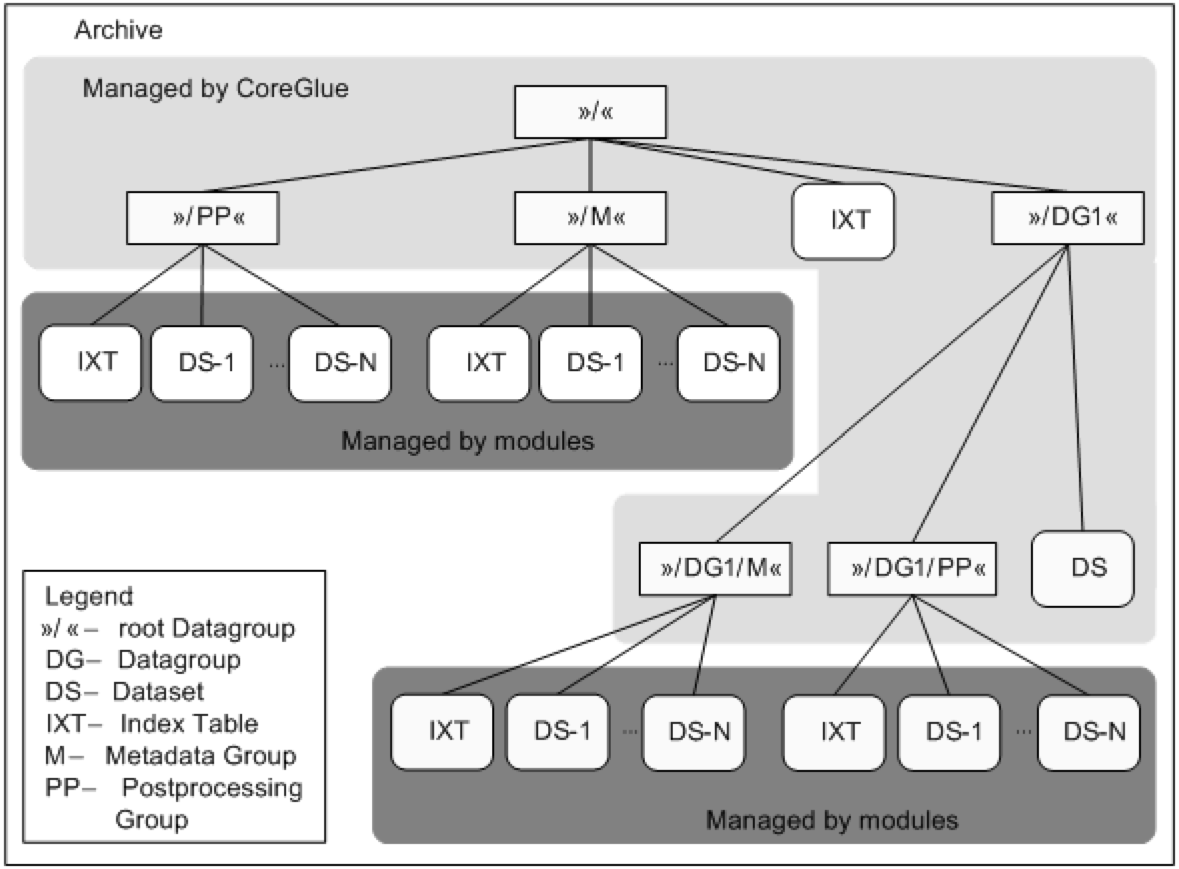
Data organization (logical view)

- Flexible representation of the stored simulation data – matrix with P (sparse) + 2 (dense) dimensions
  - indexing table – represents the logical part
    - first P indices – data group identification (arbitrary simulation run)
    - last two indices – row and column identification in data group's table
  - groups and datasets – actual raw simulation data



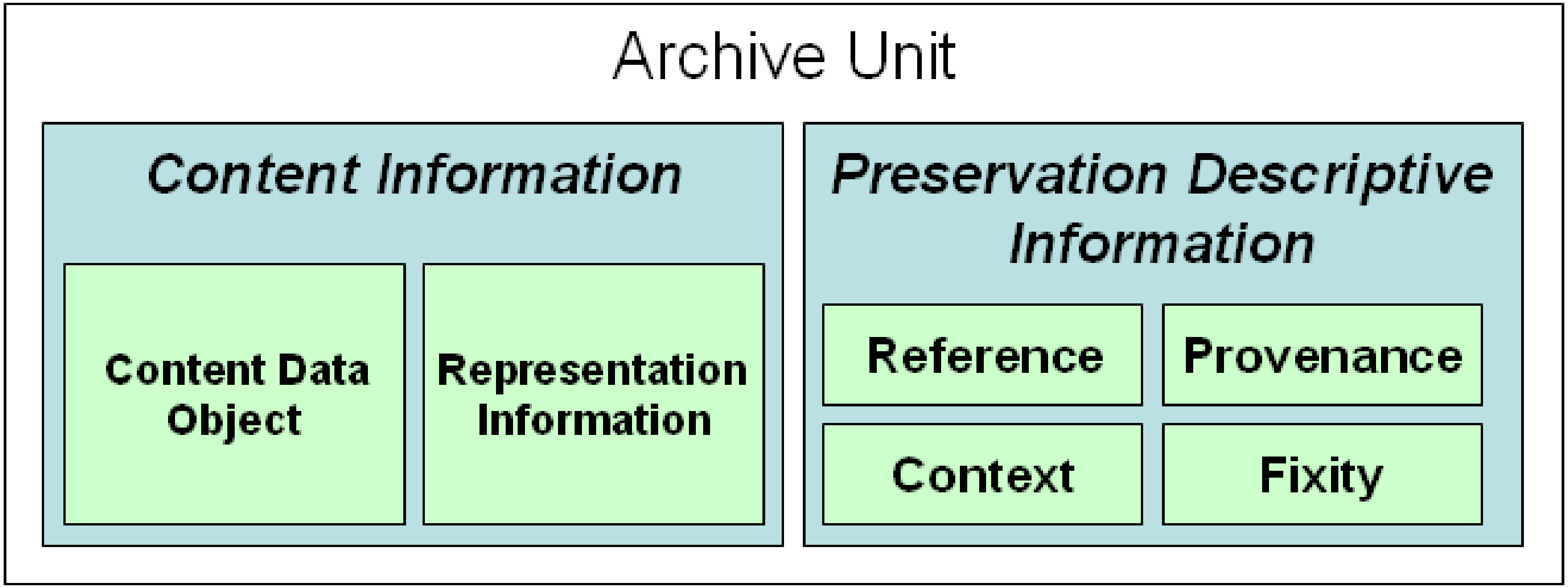
Data organization (implementation view)

- two-level tree with one "root" group
- raw simulation data – datasets under each data group
- indexing table, metadata and postprocessing data relative to the whole archive and to each simulation run



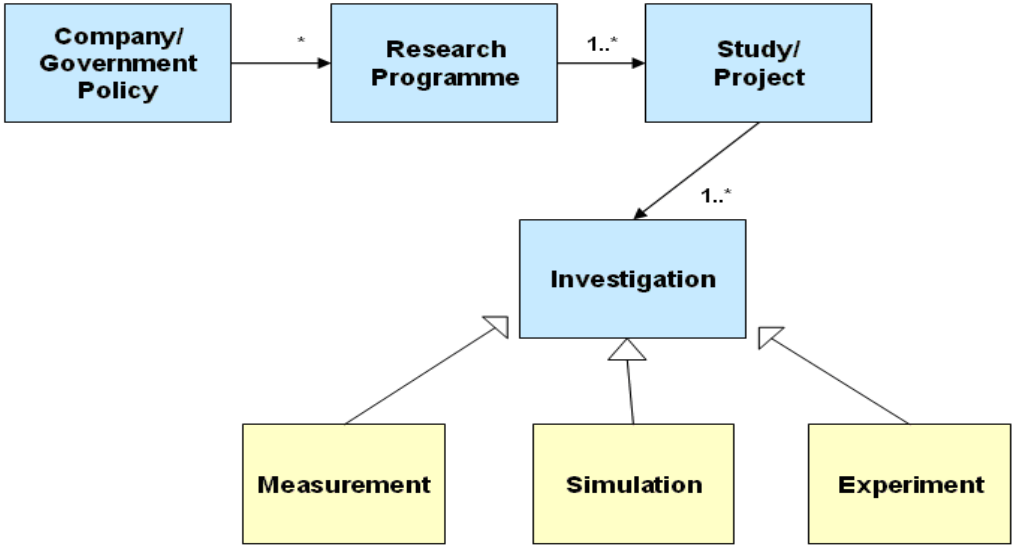
Metadata (OAIS information model)

- **Open Archival Information System** - a technical recommendation to provide permanent or indefinite long-term, preservation of digital information.



Metadata (CSMD model)

- **C**CLRC (Council for the Central Laboratory of the Research Councils) **S**cientific **M**eta**D**ata model
- Objective: to aid interoperability of scientific information systems among research organizations
- Can be used in *reference information* and in the *context information* metadata of OAIS model

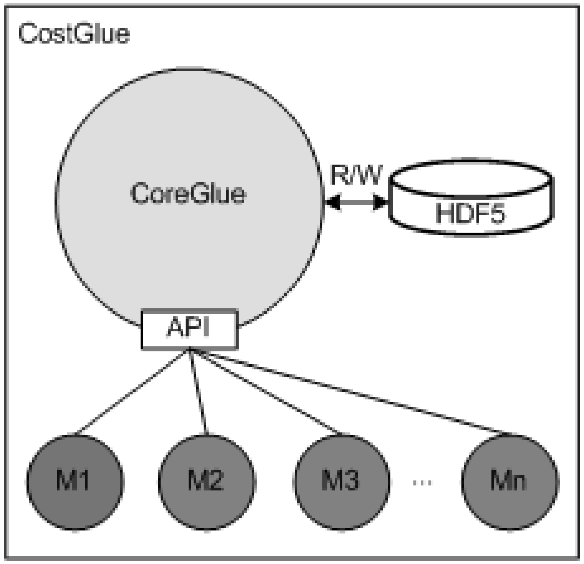


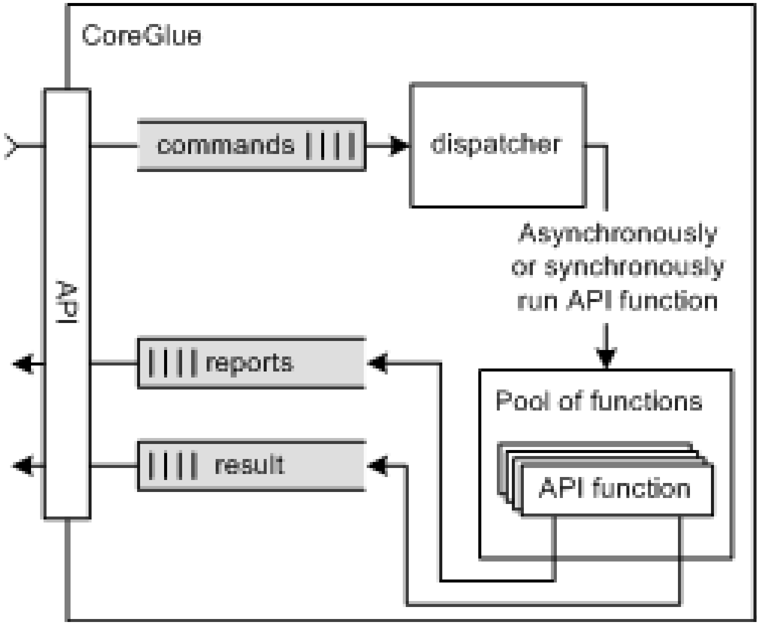
Example of a real-world simulation

- Simulating the behavior of packet switches in ns-2
- Parameters (P) in use:
  - architecture type
  - buffer size
  - number of I/O ports
  - traffic load
- Each simulation run –> new row in indexing table
- Each row uniquely identifies a simulation run and contains full path to the data group
- Datasets under data groups
  - containing: simulation outputs
- Metadata
  - containing: type of simulation scripts used, type of network topology, traffic patterns, etc.
- Post-processing data
  - containing: packet loss probability, maximum, minimum and average delay, etc.

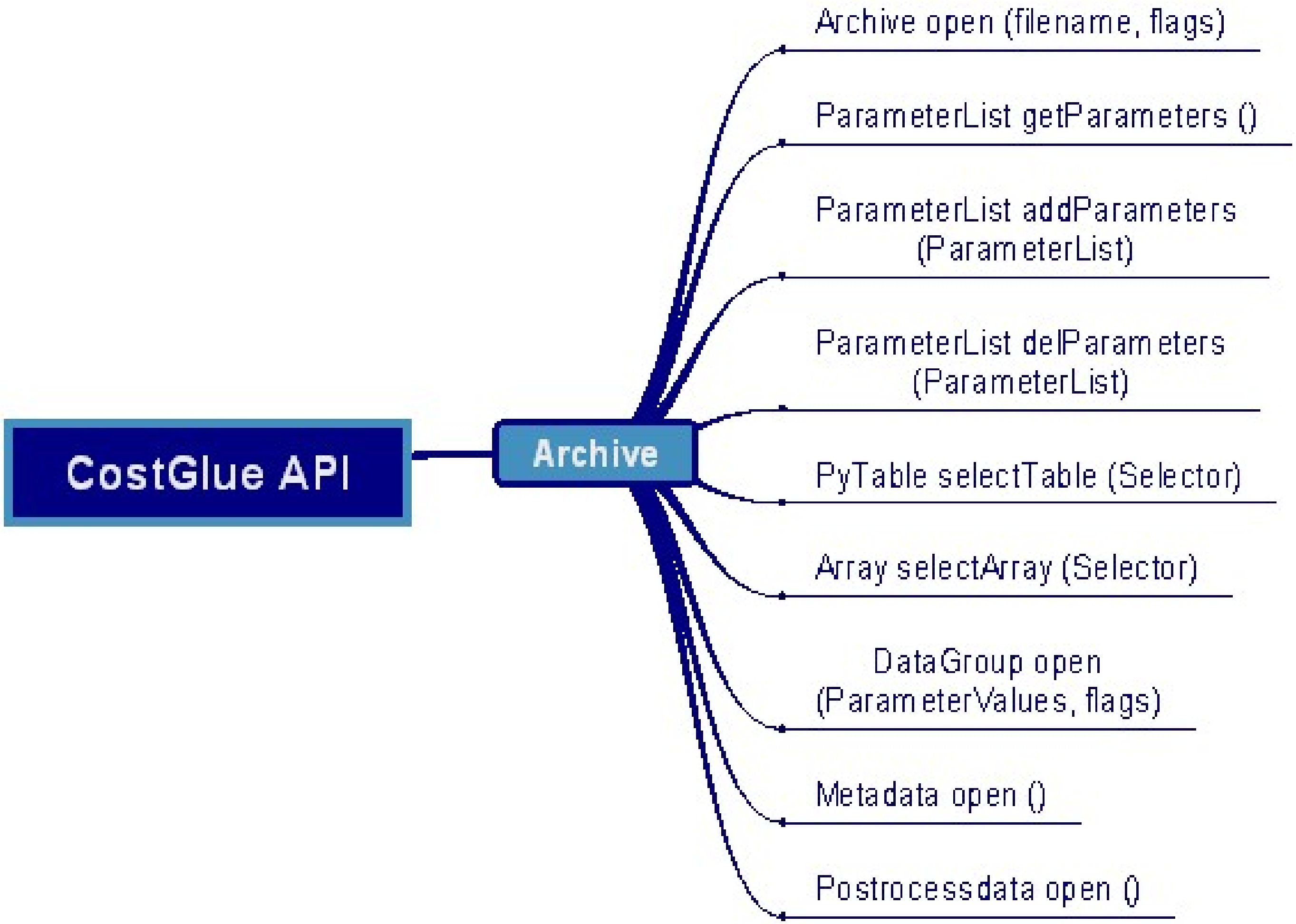
CostGlue tool architecture

- CoreGlue
  - handles HDF5
  - provides API
  - executes modules
  - passes parameters
- Modules
  - Different types for different jobs: GUI, statistical computations, export and import of data,…
  - Asyn and syn execution mode





API 1/2



CostGlue API

Archive

Archive open (filename, flags)

ParameterList getParameters ()

ParameterList addParameters (ParameterList)

ParameterList delParameters (ParameterList)

PyTable selectTable (Selector)

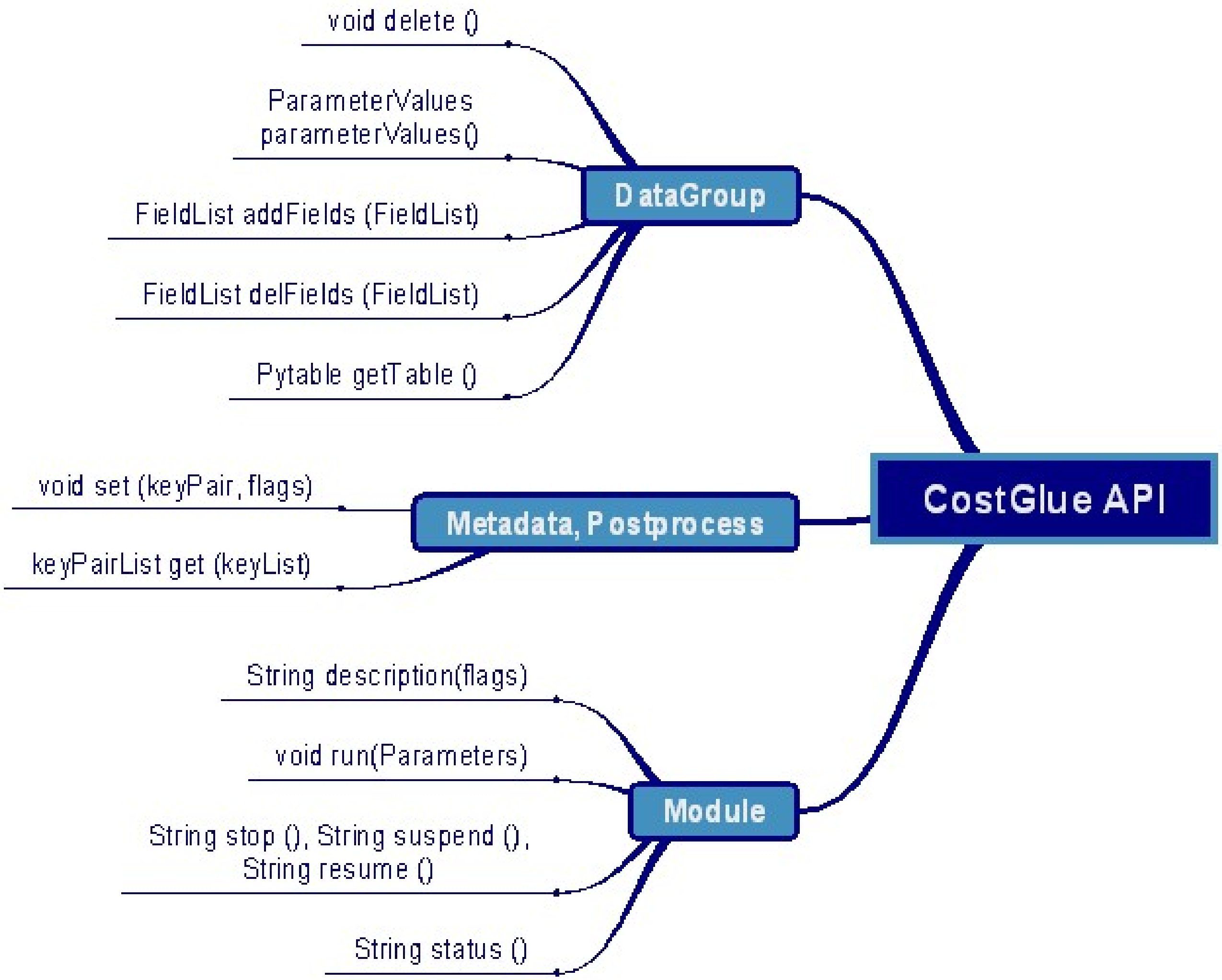
Array selectArray (Selector)

DataGroup open (ParameterValues, flags)

Metadata open ()

Postrocessdata open ()

API 2/2



CostGlue API

DataGroup

void delete ()

ParameterValues parameterValues()

FieldList addFields (FieldList)

FieldList delFields (FieldList)

Pytable getTable ()

Metadata, Postprocess

void set (keyPair, flags)

keyPairList get (keyList)

Module

String description(flags)

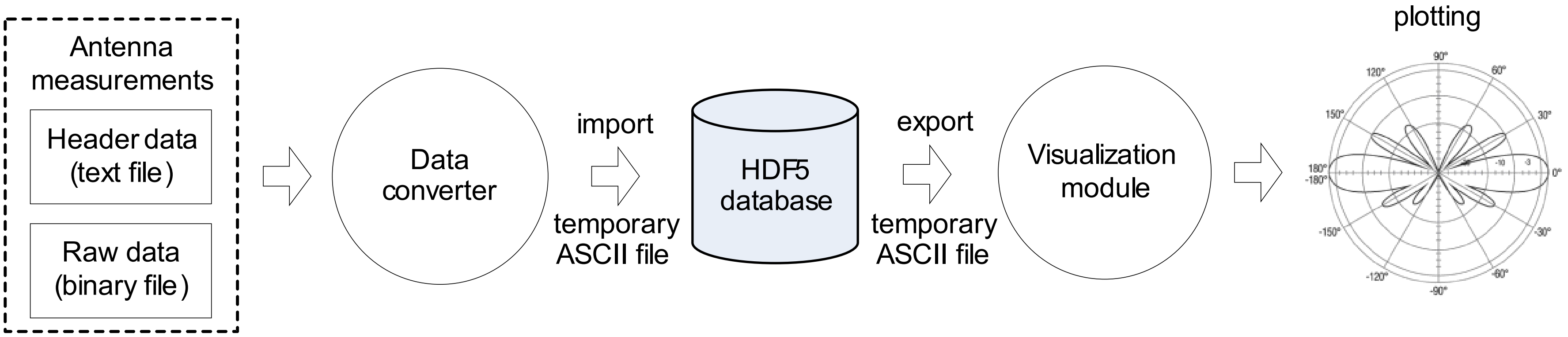
void run(Parameters)

String stop (), String suspend (), String resume ()

String status ()

Module for antenna measurements

- Design of a import module with GR-SSR group at the University of madrid
- Data is collected from antenna chamber
  - spherical
  - cylindrical or
  - plannar
- Each measurement is performed on a single frequency



Antenna measurements

Header data (text file)

Raw data (binary file)

Data converter

import

temporary ASCII file

HDF5 database

export

temporary ASCII file

Visualization module

plotting

Current status and short-term projects

- CostGlue is implemented in Python
  - Command line interface module
  - Import/export of tabular data
  - Modules for importing and exporting of antenna measurements and simulation data
- Future activity in the short term
  - Module for importing ns-2 traces
  - Module for importing tcpdump traces
  - Module for exporting to nam (network animator) format
  - Graphical user interface module

Conclusions

- Ease the task of using different simulation, measurement, data processing and visualization tools
- Software release – choosing among MIT X, GNU LGPL and GNU GPL licenses
- Future activity in the longer term
  - Generic graphical module for managing metadata and post-processing data
  - Graphical import/export module for tabular data
  - Basic statistical and plotting modules
- SVN: http://wnet.isti.cnr.it/software/costglue

Thank you for your attention!

… and enjoy data ;)