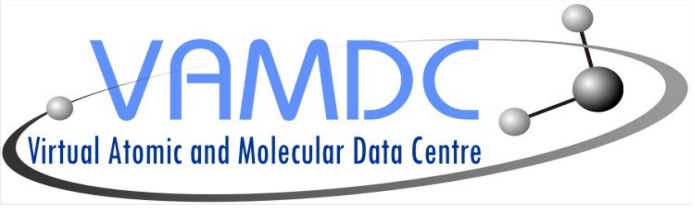
Integrating BASECOL DB into VAMDC infrastructure

Misha Doronin

LPMAA, CNRS, Paris

Joined project in January



BASECOL: stored data

- Collision properties:
  - Rate coefficients vs temperature
- Chemical elements: atoms, molecules, ions, particles, around 60 elements
- Energy tables
- References and description of work

Data access for users

- Traditional web interface
  - Well-polished, user-friendly, fast
- MOLPOP service
  - Service accessible using command-line tool
- Astrogrid DSA/Catalog service for direct tables access
  - Provides standard TAP service
  - Custom, user needs to know database structure
- Prototype XSAMS service
  - Provides output in unified XSAMS format
  - Access and query protocols are not standardised yet

Access: Traditional web interface

- Query:
  - Choose by element
  - Search collisions filtering by elements, process type, author and year of publication
- Output:
  - Set of web pages for verbose data (description, references)
  - Plain text tables, HTML tables, VOTable for tabular data (energy tables, rate coefficients)

Access: Traditional web interface

- Pro's:
  - Well-polished, convenient interface
  - Fast data access
  - Multiple output formats
- Con's:
  - Specific interface for each database

Access: DSA/Catalog

- DSA/Catalog:
  - Direct access to database tables
  - Access to predefined views
- Query:
  - SQL query (ADQL dialect)
- Output:
  - VOTable

Access: DSA/Catalog

- Pro's:
  - Virtual Observatory TAP standard implementation
  - client/registry/server software is already written and is working almost out-of-the-box
- Con's:
  - Queries are database-specific
  - Not very well suited for representing hierarchical data
- Needs:
  - Standard set of table views to ensure uniform access and interoperability

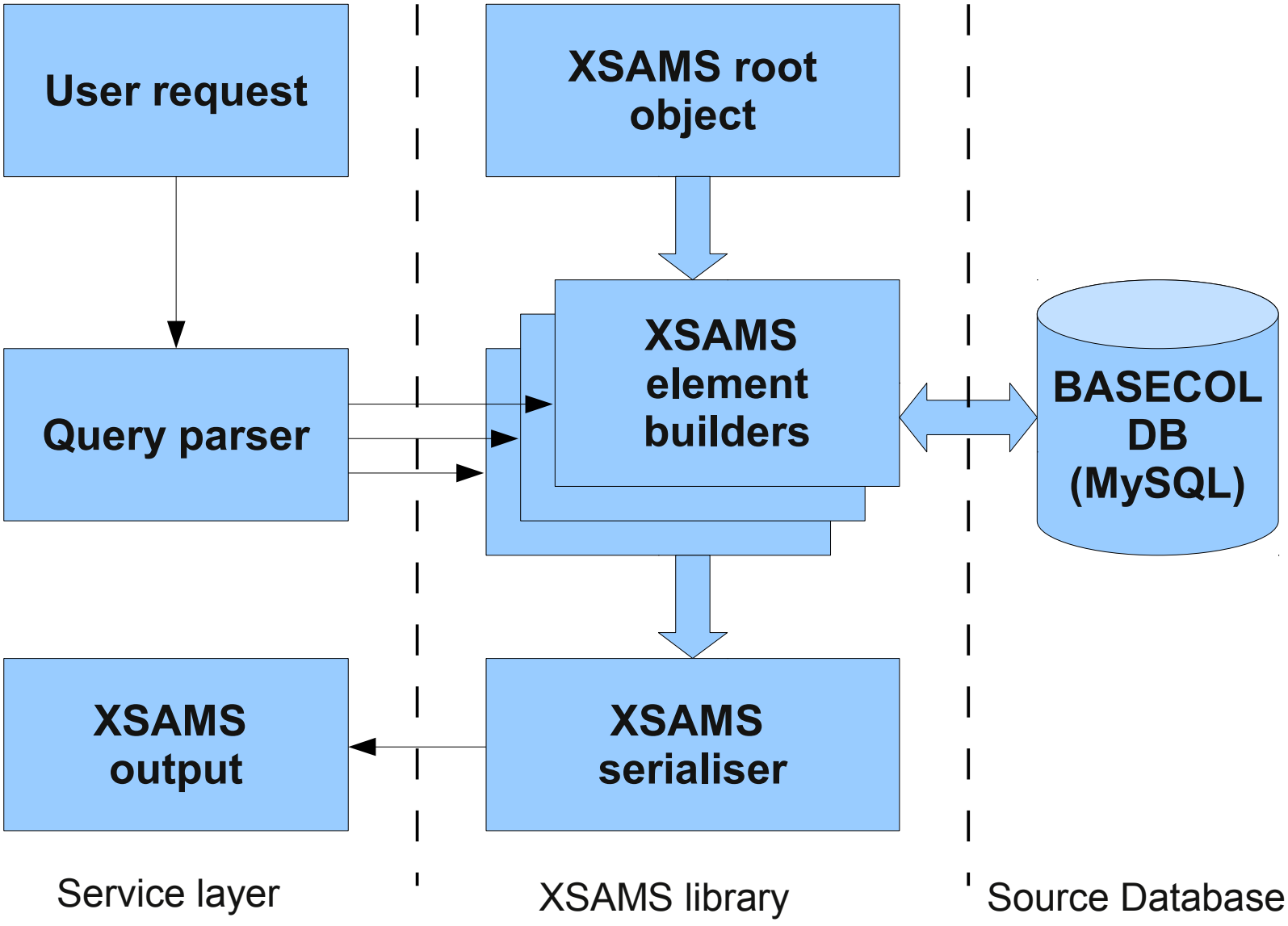
XSAMS service: access and query

- Command-line tool/library for local access
  - Provides routines for generating elements of XSAMS
- Java servlet
  - accessible using HTTP GET method.
  - SLAP/TAP like service needs to be standartised and implemented
- Query
  - self-defined non-standard query language
  - Standard, common query language needs to be defined and implemented

XSAMS service: output

- Full basecol output for single collision: 100K-2M uncompressed XSAMS, exaustive information in single file.
- XSAMS molecular dialect: flattened full schema
- Compressed output: 2 to 40 KB
- Retrieval time for full requests: 0.5 - 1.5 s
- Retrieval time for simple requests: 10-50 ms

XSAMS service: architecture



Review: current state

- Prototype: VAMDC next generation service, providing output in XSAMS format
- Prototype: client software for reading XSAMS
- Deployed-testing: Virtual Observatory architecture services
- Deployed-mature: Web interface, MOLPOP service

Review: pending tasks

- Standard access protocol(s) for XSAMS service
- Decide how will XSAMS-0.2 look like
- Standard query language for XSAMS service
- Standard table layout for VO service
- Standard chemical elements identification