Recent and Future Activities in HPC and Scientific Data Management

Siegfried Benkner

Research Group Scientific Computing

Faculty of Computer Science

University of Vienna

AUSTRIA

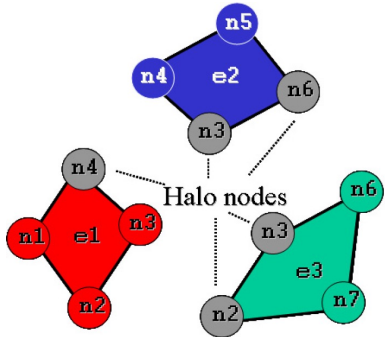
http://www.par.univie.ac.at

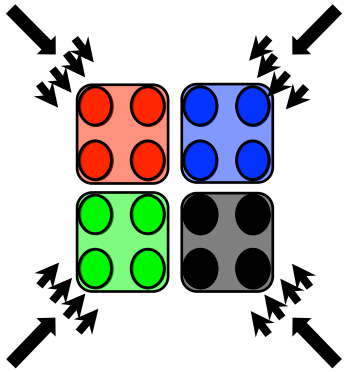
Research Group Scientific Computing

One of twelve research groups at the Faculty of Computer Science, University of Vienna.

Parallel Computing / HPC

- Programming Models and Languages
- Compiler and Runtime Technologies
- Programming Environments and Tools

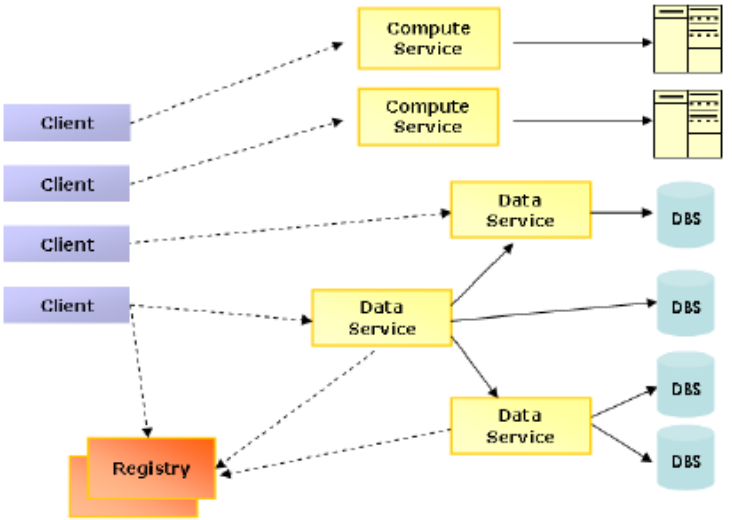




Vienna Fortran, HPF+, Hybrid Programming, Multicore, GPU, Phi…

Grid/Cloud/Big Data

- Compute and Data Services
- Semantic Data Integration
- Big Data and HPC



Vienna Cloud Environment, VPH-Share, …

Recent Research Projects

- **PEPPHER**, Performance Portability and Programmability for Heterogeneous Many-core Architectures
  European Commission, FP7, 2010-2014, **Coordinator**

- **VPH-SHARE,** Virtual Physiological Human - Sharing for Healthcare,
  European Commission, FP7, 2011-2015

- **AutoTune -** Automatic Online Tuning,
  European Commission, FP7, 2011-2015

- **RETIDA** - Real-Time Data Analytics for the Mobility Domain,
  FFG, 2014-2017

EU Project PEPPHER

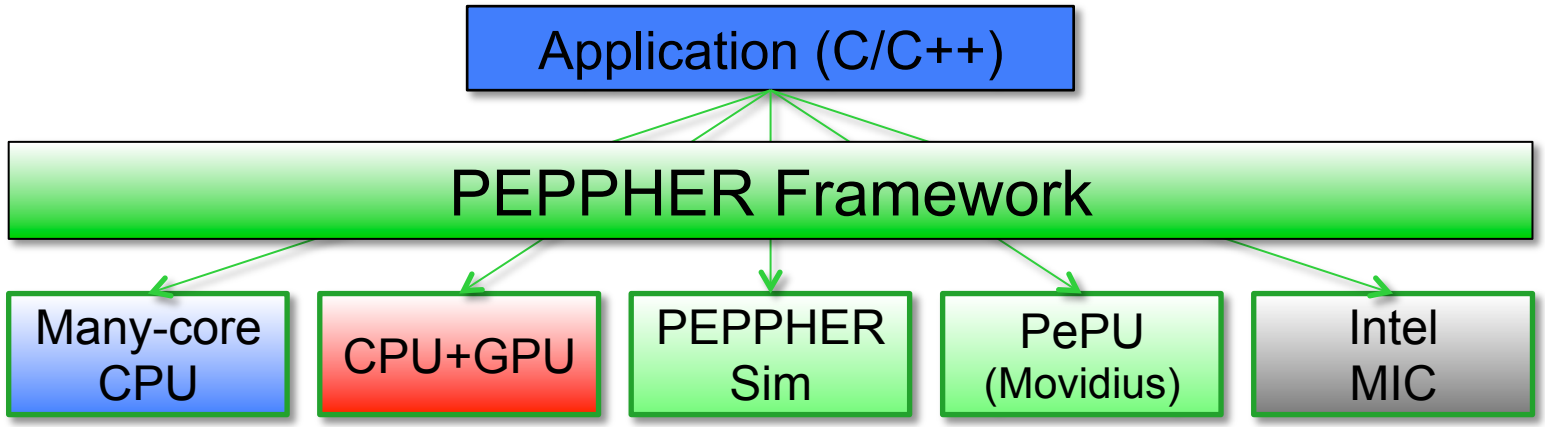
http://www.peppher.eu

Performance Portability & Programmability for Heterogeneous Many-Core Architectures

- FP7 ICT, Computing Systems; 2010-2014
- Partners: UNIVIE, INRIA, LIU, Intel, Movidus, Codeplay, KIT, Chalmers, TUW

Methodology & framework for development of performance portable code

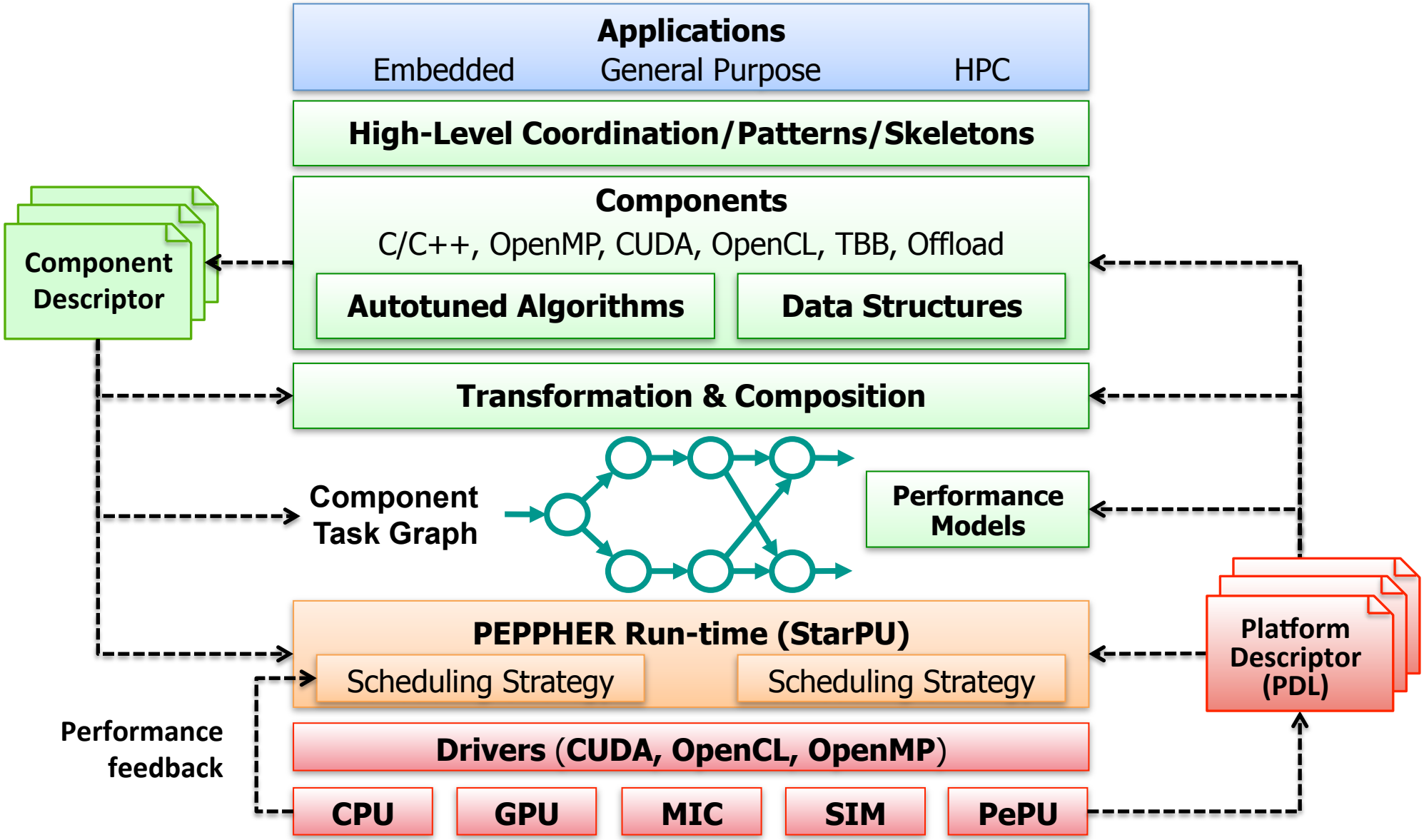
- Execute same application efficiently on different heterogeneous architectures
- Support real hybrid execution to exploit all available computing units



Focus: Single-node heterogeneous architectures

PEPPHER Framework

http://www.peppher.eu



The AutoTune Project

http://www.autotune-project.eu

- FP7 ICT, Computing Systems; 42 months, 2011-2015
- Partners: TUM, UAB, CAPS, LRZ, ICHEC, UVIE

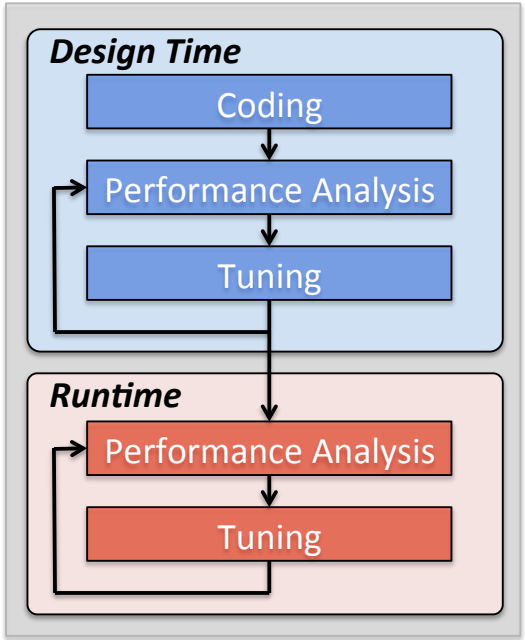
Combine performance analysis and tuning into single framework.

☞ **Periscope Tuning Framework (PTF)**

Extend Periscope with automatic tuning plugins for **performance** and **energy efficiency** tuning.

Execute whole tuning process **online** (performance analysis and tuning during single application run).

Use **expert knowledge** to guide search for performance properties and tuned versions.



Periscope Tuning Framework (PTF)

Parallel architectures

- Multicore servers
- Supercomputers (SuperMUC)
- Accelerated systems (GPU, Xeon Phi)

Programming paradigms

- MPI, MPI/OpenMP
- OpenCL/CUDA
- Parallel Patterns (PEPPHER)

Tuning (at) Different Layers of SW Stack

- High-level language (directives/annotations)
- Compilers / Transformation systems
- Runtime systems and libraries
- Operating system

Developed Tuning Plugins

- MPI Parameter Tuning
- Pipeline Patterns for CPU/GPU
- DVFS Plugin
- Compiler Flags Selection
- MPI Master/Worker Tuning
- OpenCL Worksize Tuning

PTF available as open source:

- http://periscope.in.tum.de/releases/latest/tar/PTF-latest.tar.bz2

EU Project VPH-Share

http://www.vph-share.eu

Virtual Physiological Human: Sharing for Healthcare

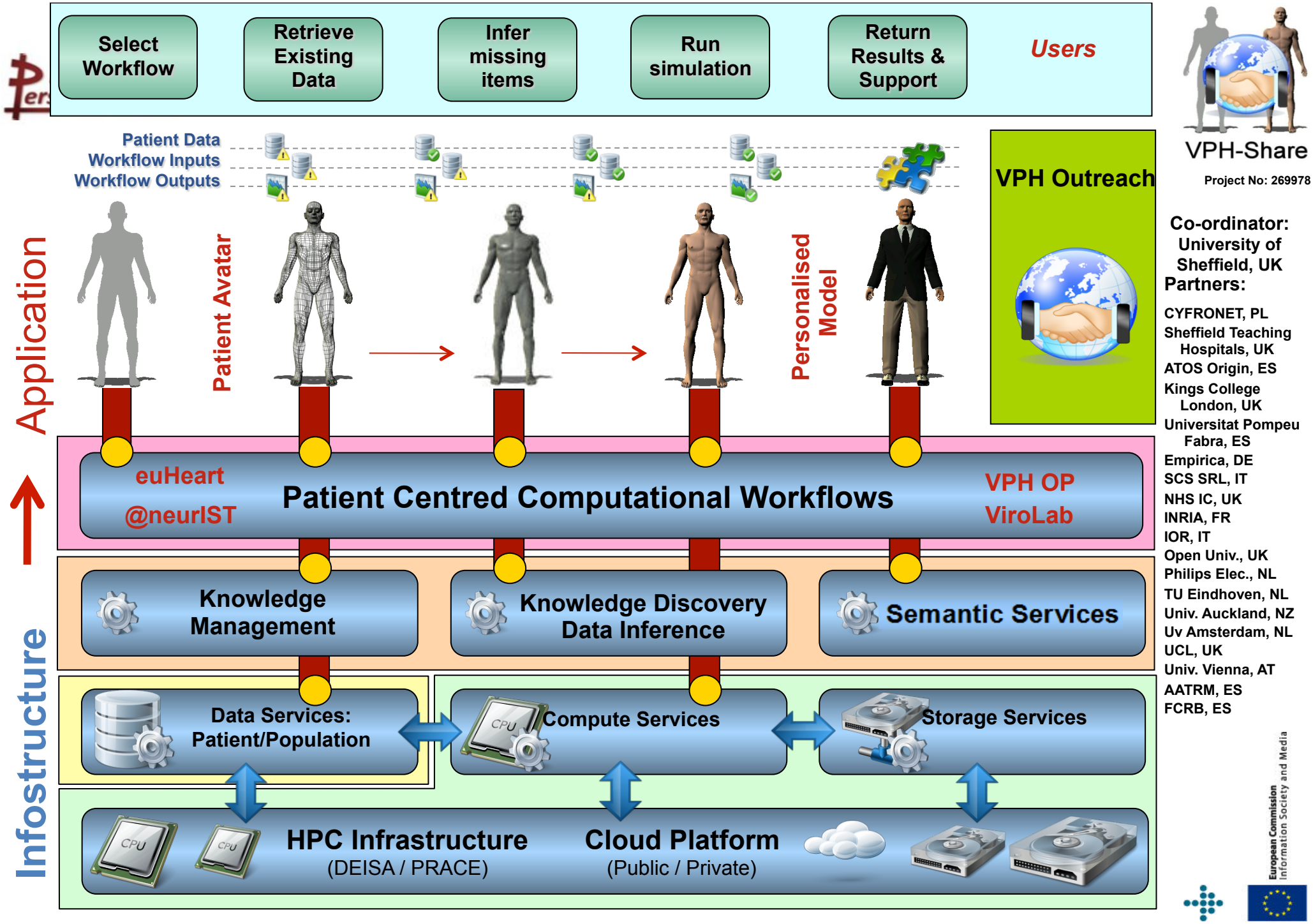
(European Commission, FP7, 2011-2015, Coordinator: The University of Sheffield)

VPH-Share developed a **Cloud *infostructure*** to facilitate integration of/ access to...

- **Patient data** across different systems, hospitals, countries ...
- **Information/models** related to various parts and processes of human body
- **Knowledge** (guidelines, standards, protocols in research and clinical practice)

Specific requirements

- **Complex, distributed, heterogeneous data** (multi-scale, multi-modal)
- **Sensitive data** (security, privacy, legal issues, data quality)
- **Specialized analytics** to integrate bioinformatics and systems biology information with clinical observations



VPH-Share

Project No: 269978

Co-ordinator:
University of Sheffield, UK

Partners:

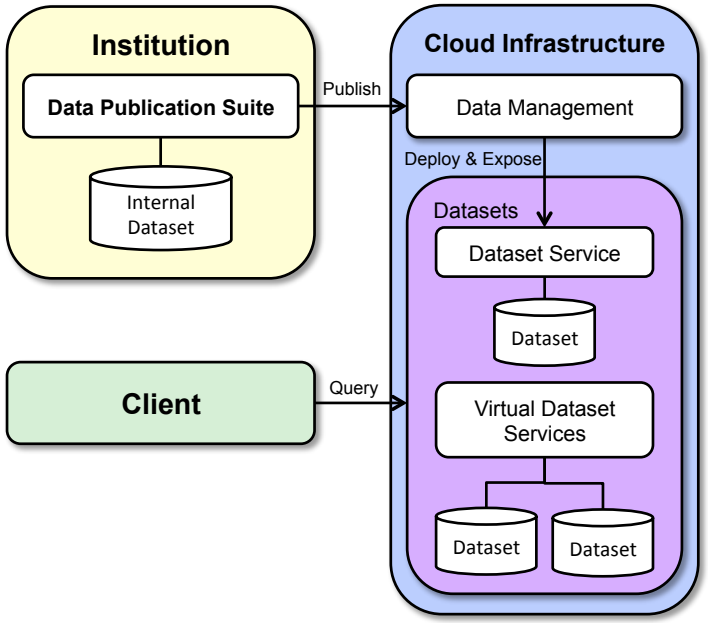
CYFRONET, PL
Sheffield Teaching Hospitals, UK
ATOS Origin, ES
Kings College London, UK
Universitat Pompeu Fabra, ES
Empirica, DE
SCS SRL, IT
NHS IC, UK
INRIA, FR
IOR, IT
Open Univ., UK
Philips Elec., NL
TU Eindhoven, NL
Univ. Auckland, NZ
Uv Amsterdam, NL
UCL, UK
Univ. Vienna, AT
AATRM, ES
FCRB, ES

European Commission
Information Society and Media

VPH-Share Data Management

Data Publication Suite

- select & annotate data with semantic concepts,
- de-identify selected data items
- expose data as Cloud services

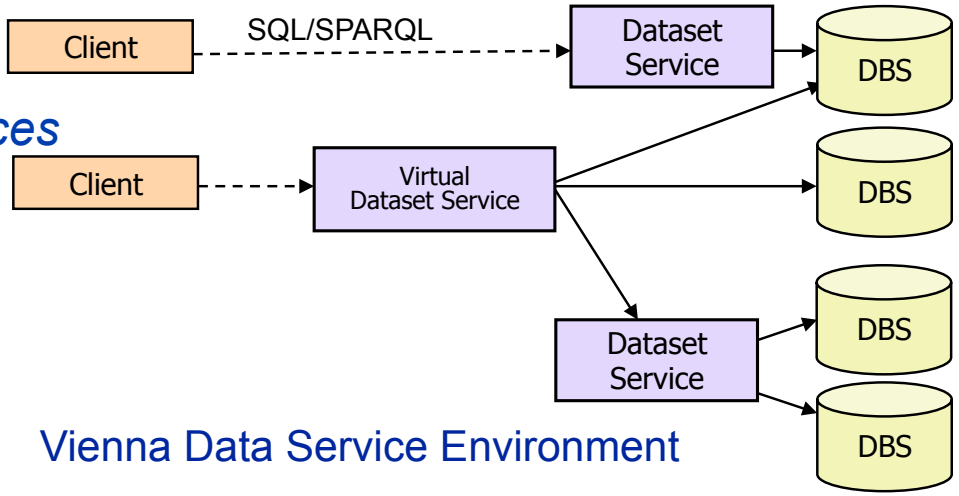


Data Service Environment (VDSE)

- based on Vienna Cloud Environment (VCE)
- provision of data sets (RDBs) as Cloud services
- access via SQL or SPARQL
- provide on-demand *customized views*
- federation across multiple data sets
- preserve *autonomy of underlying data sources*

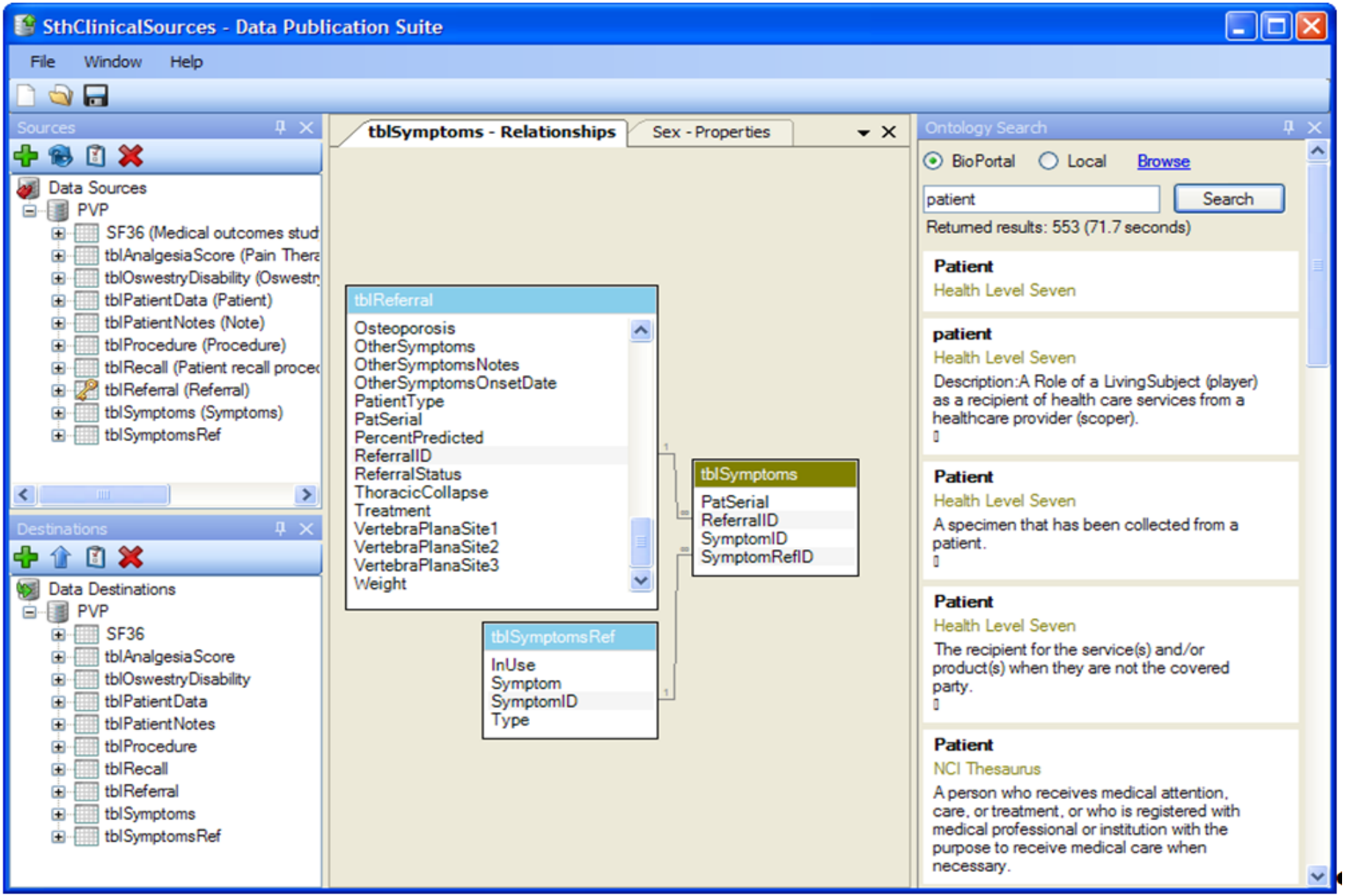
Semantic mechanisms

- to discover, link and search data



Vienna Data Service Environment

VPH-Share Data Publication Suite

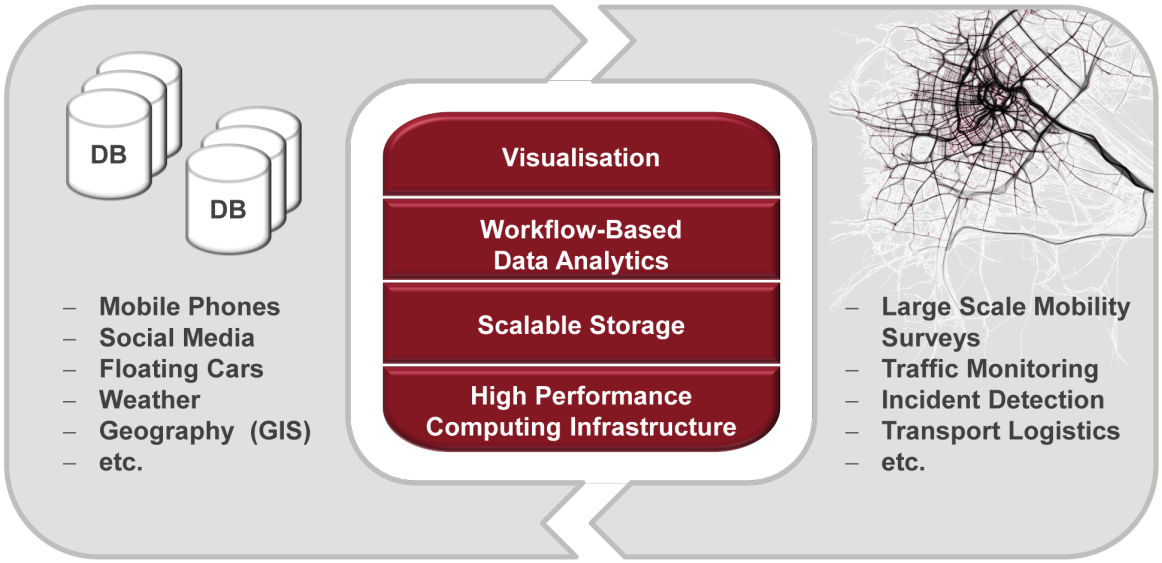


RETIDA

https://dts.ait.ac.at/projects/retida/

Real-time data analytics for the mobility domain

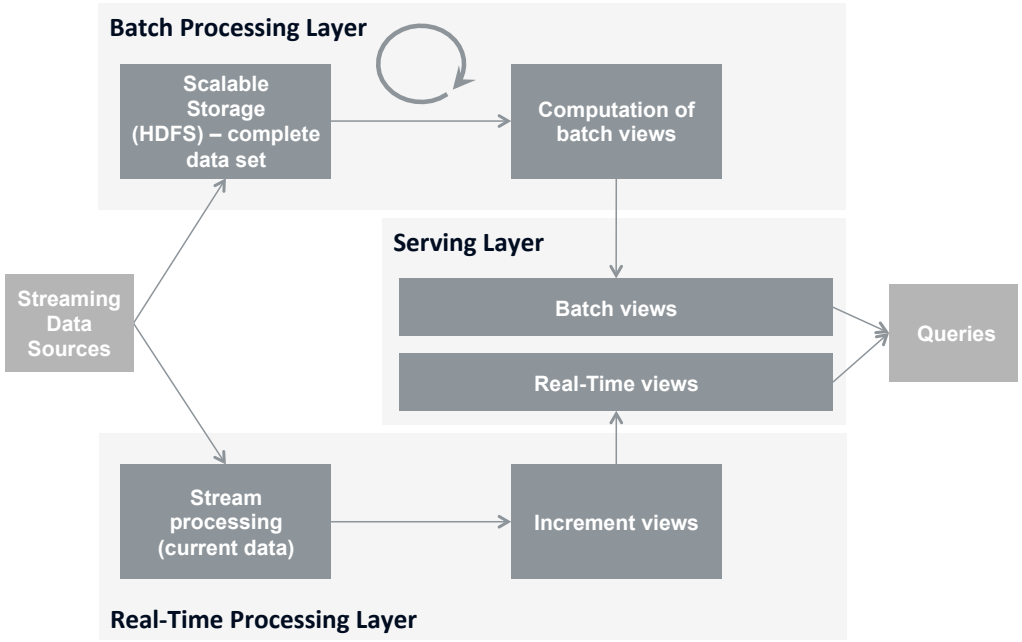
Austrian FFG, ICT of the Future, 2014-2016

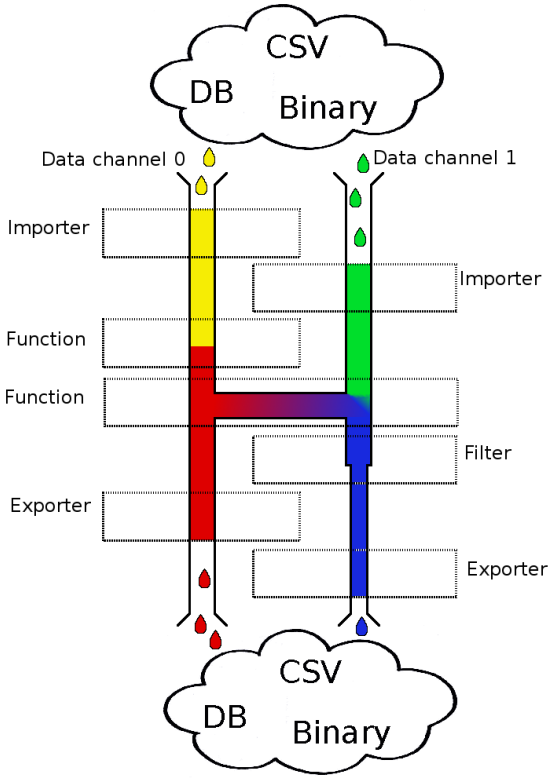


- Real-time integration and analytics of large-scale heterogeneous data sources
  mobile phone data, floating car data, GIS, weather, social media, …
- Massively parallel, adaptive execution of generic data analytics workflows
  support for heterogeneous architectures (GPUs, Xeon Phi, …)
- Application-specific visualizations

RETIDA

https://dts.ait.ac.at/projects/retida/





Hadoop-based lambda architecture

- scalable and fault-tolerant processing
- real-time vs. batch

High performance data pipeline

- C++ re-configurable framework
- real-time processing capabilities

Future Research

- Programming Support for Big Data Applications
- Taking Advantage of Heterogeneous Architectures
- Streaming and Real-Time Support
- Runtime systems for Big Data Applications
- Cloud-based Scientific Data Management (VDSE)