

Concepts and Objectives

The main objective of PaaSAGE is

To deliver an open and integrated platform to support both design and deployment of Cloud applications, together with an accompanying methodology that allows model-based development, configuration, optimisation, and deployment of existing and new applications independently of the existing underlying Cloud infrastructures.

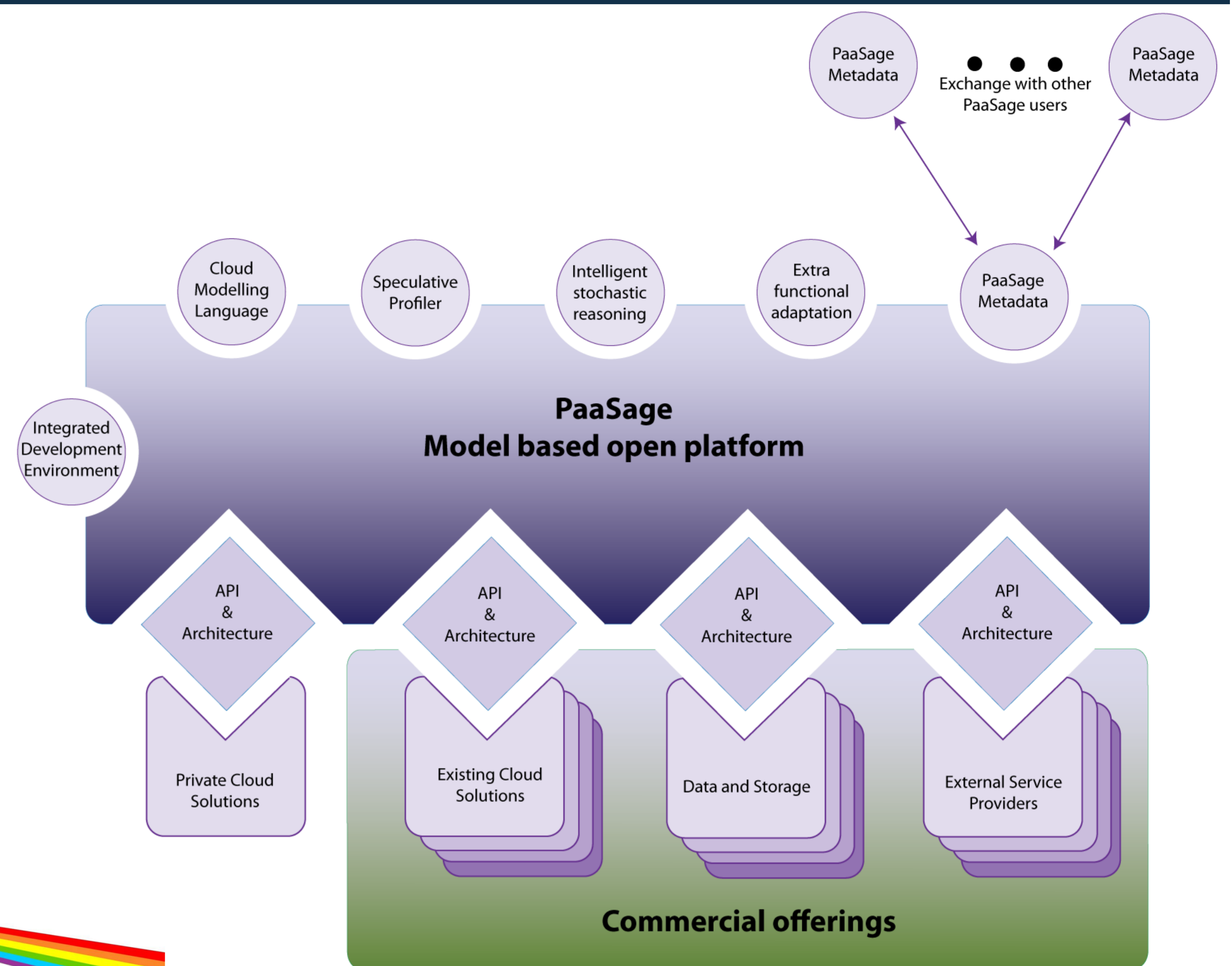
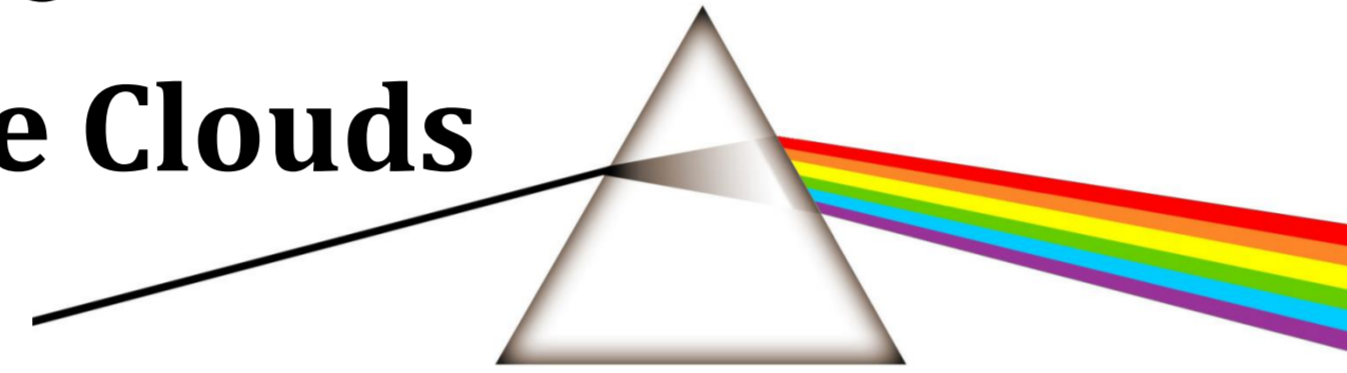
PaaSAGE architecture

- Cloud Modelling Language (CML)
- Speculative Profiler
- Intelligent Stochastic Reasoning
- Extra functional Adaptation
- Metadata
- Collaborative software development

PaaSAGE Lifecycle

- Model-driven IDE
- Deployment on Multi-Clouds
- Profiling and Adaptation
- Metadata repository
- Developer interface (social network)

Define your application once
 Deploy it at the full spectrum of the Clouds



AGH Contribution to PaaSAGE

Extended eScience Use Case

- Local domain researchers, including PL-Grid community: bioinformatics (genomics, proteomics), metals engineering (complex metallurgical processes)
- International collaborations: Virtual Physiological Human (Taverna and DataFluo workflows), multiscale applications: fusion (Kepler workflows), military mission planning support (EDA), astronomy (Pegasus workflows)

Products:

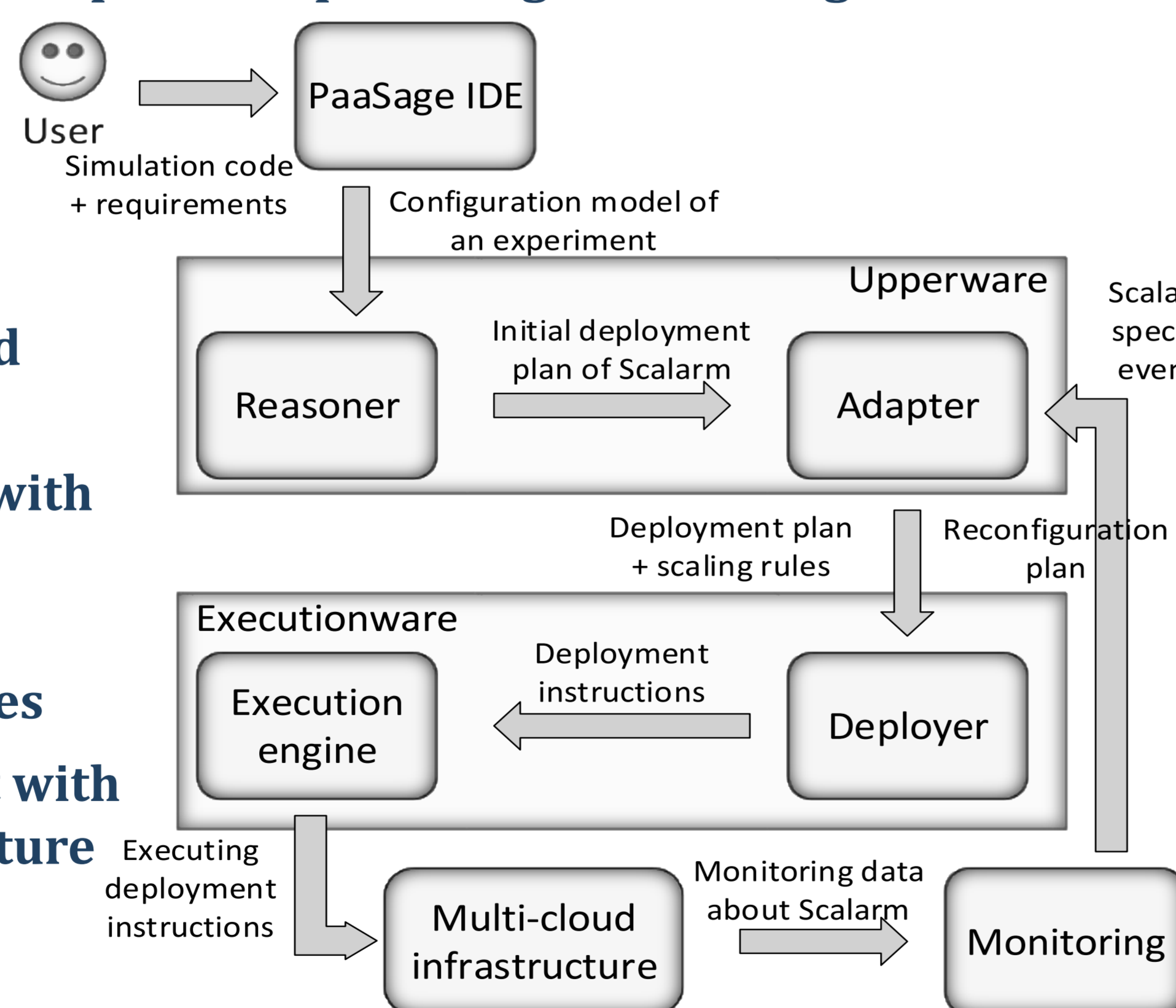
- **Hyperflow**: workflow execution engine inspired by process networks theory and hypermedia (REST) paradigm
- **Scalarm**: massively self-scalable platform for data farming workflows

Support for Data Farming Applications

Scalarm [3] is software of choice for supporting data farming in PaaSAGE

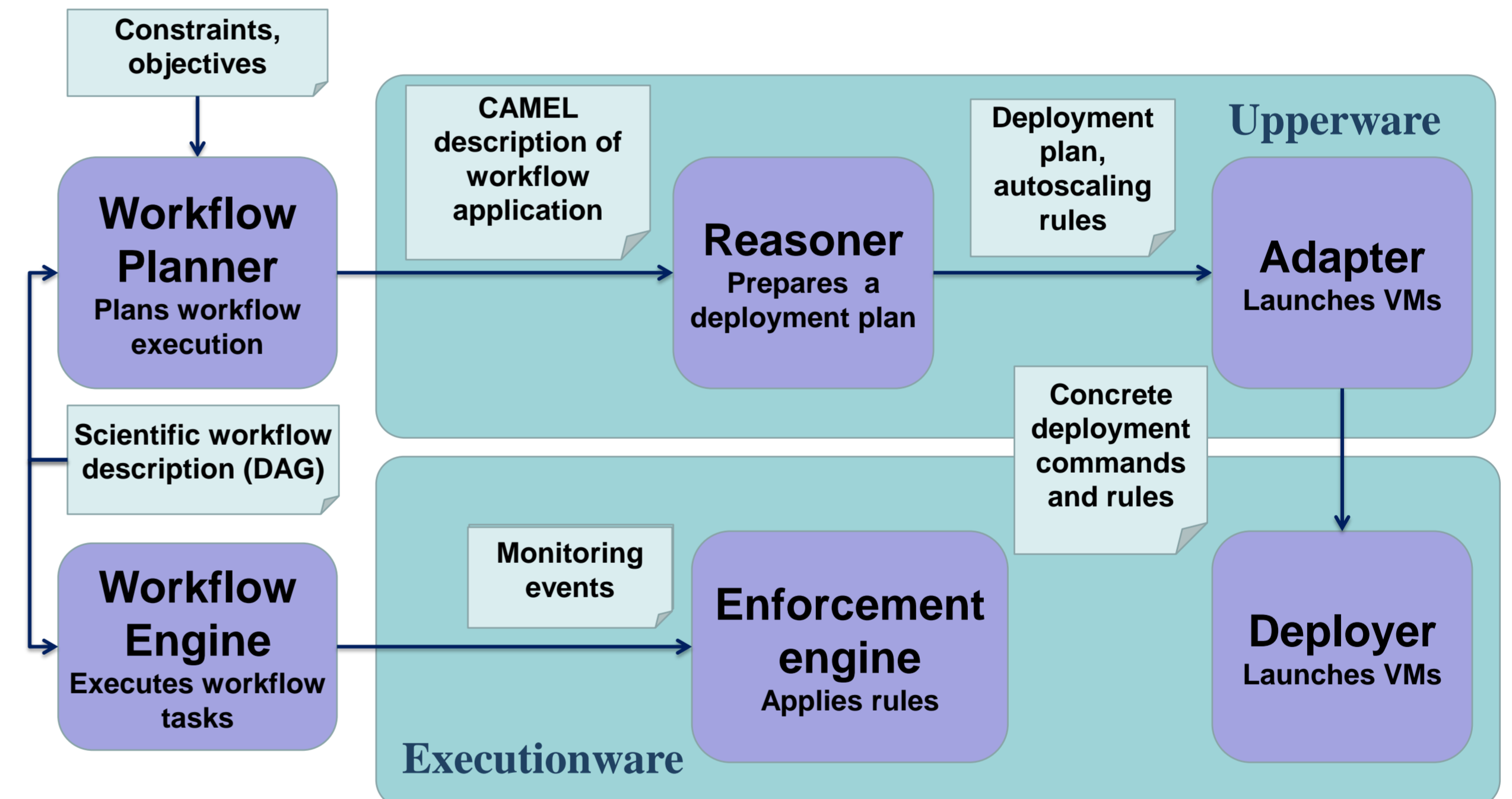
It is a complete, self-scalable platform providing with PaaSAGE:

- Design of Experiment methods
- Data exploration techniques
- Support for multi-cloud infrastructure
- Efficient computation with "pilot jobs" and "pull" mechanism
- Support for scaling rules
- Automatic deployment with Model-Driven Architecture
- QoS monitoring



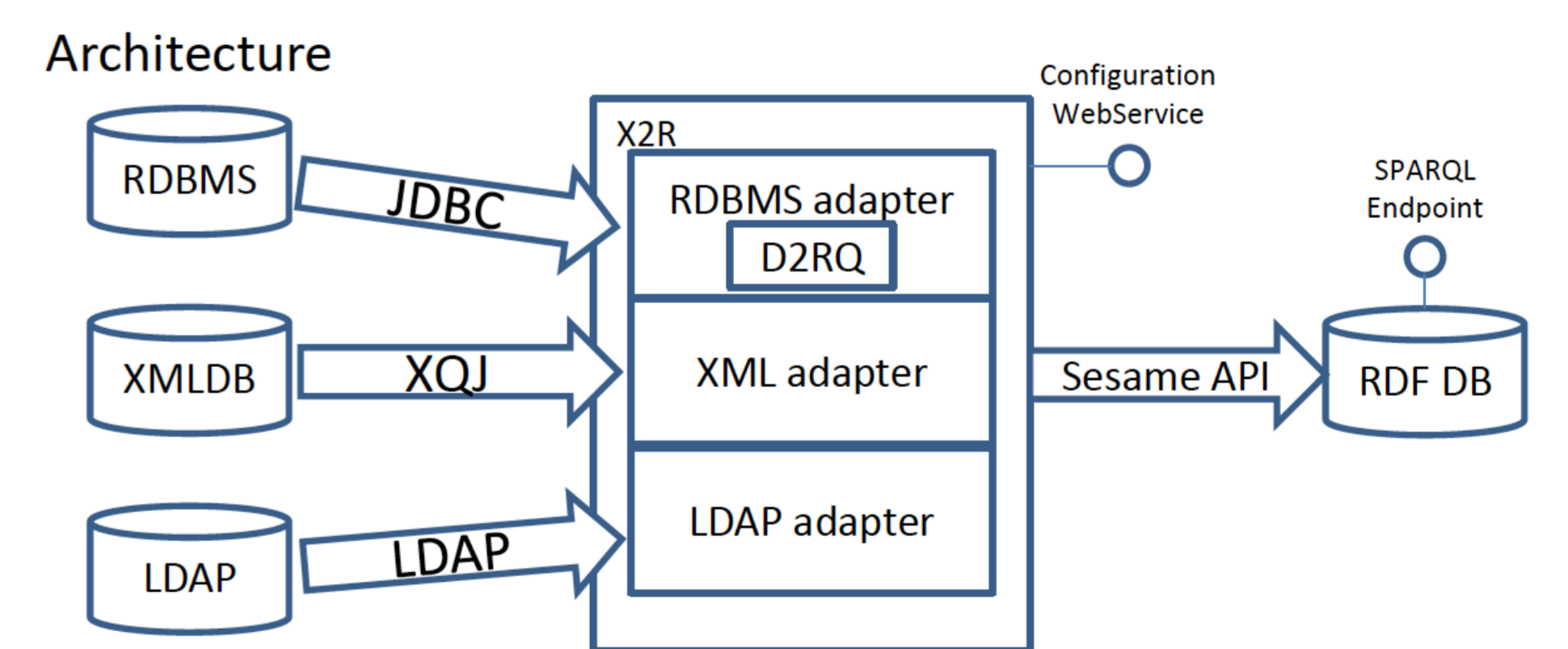
Support for Large-Scale Scientific Workflows

- Pipelines of many (100K+) resource intensive tasks
- Powered by the *Hyperflow* workflow engine
- Extension to PaaSAGE Reasoner prepares workflow deployment plan and autoscaling rules



Extended Metadata

- To extend the PaaSAGE Communication Hub with capability to integrate information from various data sources, e.g. relational databases, LDAP catalogues, or XML files, using predefined mappings and transformation methods
- The solution will be based on the X2R tool developed at AGH [4], which will semi-automatically translate data from legacy information sources to unified ontological format based on provided mappings



References

1. B. Baliś, Hypermedia workflow: a new approach to data-driven scientific workflows. In High Performance Computing, Networking, Storage and Analysis (SCC), 2012 SC Companion, pp. 100-107. IEEE, 2012.
2. M. Malawski, K. Figiela, J. Nabrzyski: Cost Minimization for Computational Applications on Hybrid Cloud Infrastructures. Future Generation Comp. Syst. 29(7): 1786-1794, 2013.
3. D. Krol, M. Wrzeszcz, B. Kryza, L. Dutka, and J. Kitowski, Massively Scalable Platform for Data Farming Supporting Heterogeneous Infrastructure, in IARIA Cloud Computing, pp. 144-149, 2013.
4. A. Mylka, A. Mylka, B. Kryza, J. Kitowski, Integration of Heterogeneous Data Sources in an Ontological Knowledge Base, Computer and Informatics. volume 31, number 1, pp. 189-223, 2012.



We thankfully acknowledge the support of the European ICT-FP7 program through the PaaSAGE (IP 317715) project