



PAL: An object oriented data access layer for Herschel data processing

B. Li^{1,2}, S. Guest³, M. Huang¹, P. Balm⁴, H. Siddiqui⁴, J. Bakker⁴, J. C. Segovia⁴, J. Saiz⁴, K. Edwards^{5,6}

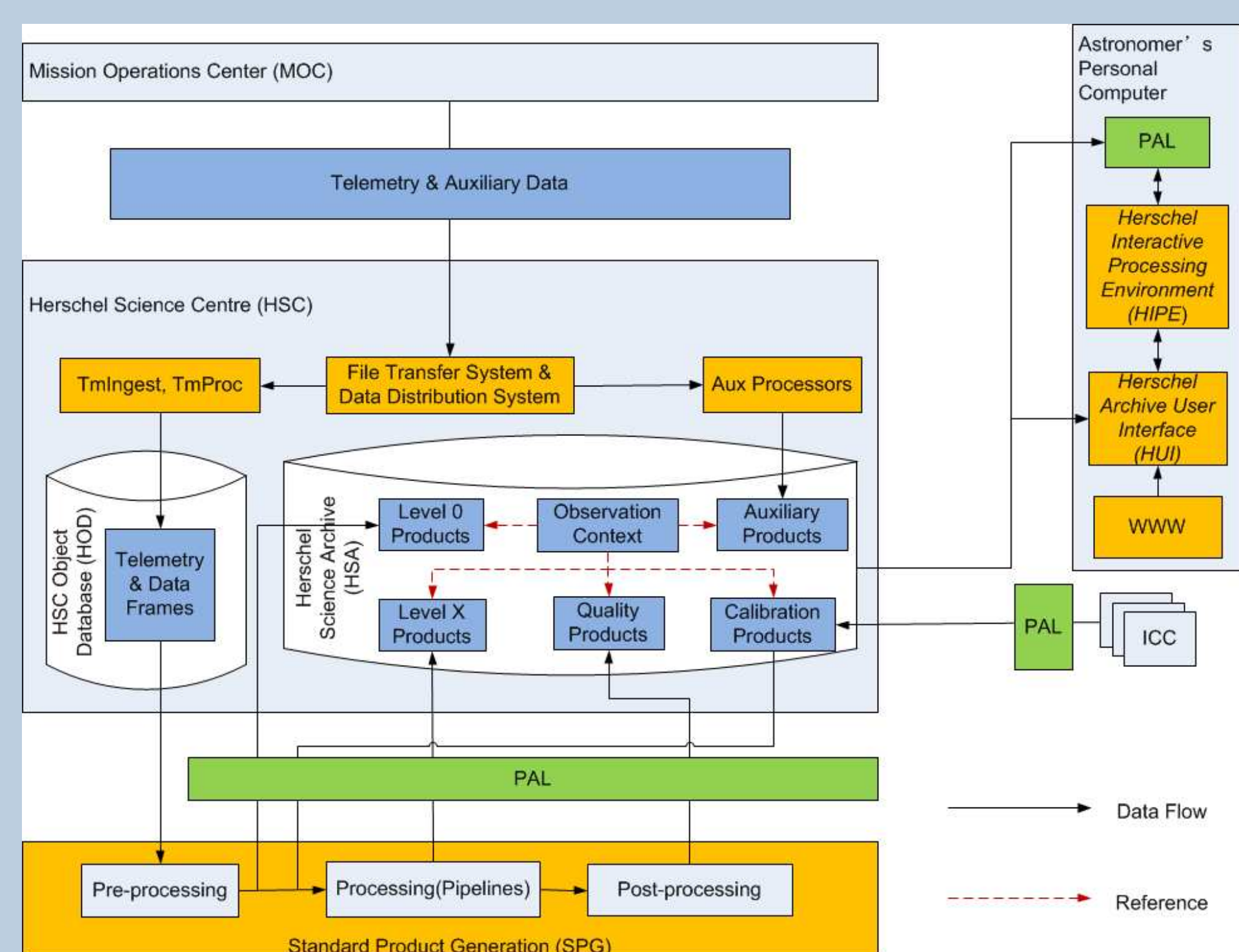
¹NAOC, China. ²GST, China. ³RAL, UK. ⁴ESAC, Spain. ⁵SRON, Netherlands. ⁶University of Waterloo, Canada.



Introduction

The Herschel Space Observatory (Pilbratt et al, 2010), launched on 14 May 2009, is an ESA Cornerstone project that will contribute to our understanding of the origin and evolution of stars and galaxies. The Herschel Space Observatory generates 9TB(Balm, 2009) of raw scientific data during its lifetime that are processed into data products of different levels and are stored at geographically distributed locations. The Herschel Ground Segment Product Access Layer (PAL), an object oriented software package that provides a consistent means for accessing these data products, isolates the complexity of underlying storage systems and makes it possible for scientific data analysis scripts to be reused among different environments and over a long period of time as software and hardware evolve. This benefits all users of Herschel data products, including astronomers, calibration scientists, instrument engineers and scientific software developers.

PAL Usage



An example of PAL usage is shown in the above figure. Raw data downlinked from the Herschel Space Observatory in the form of telemetry packets are transferred from the Mission Operation Center (MOC) in Darmstadt, Germany to the Herschel Science Center (HSC) in Villafranca del Castillo, Spain, and then ingested and stored as application telemetry and data frame objects in the HSC Operational Database (HOD) as part of the Herschel Science Archive (HSA), the principal database for storing data that are sent to or received from the spacecraft. Those data are then processed into different types of data products and stored in the HSA. Astronomers, calibration and instrument scientists in ICCs will then access these data products in the HSA through the PAL.

References

- [1] BALM, P. 2009. *Herschel Bulk Product Transfer ICD*.
- [2] GRIFFIN, M.J., ABERGEL, A., ABREU, A. ET AL. 2010, *A&A*, 518, L3
- [3] PILBRATT, G.L., RIEDINGER, J.R., PASSVOGEL, T. ET AL. 2010, *A&A*, 518, L1
- [4] S.Ott, 2010, ADASS XIX, *The Herschel Interactive Processing Environment*

Acknowledgements

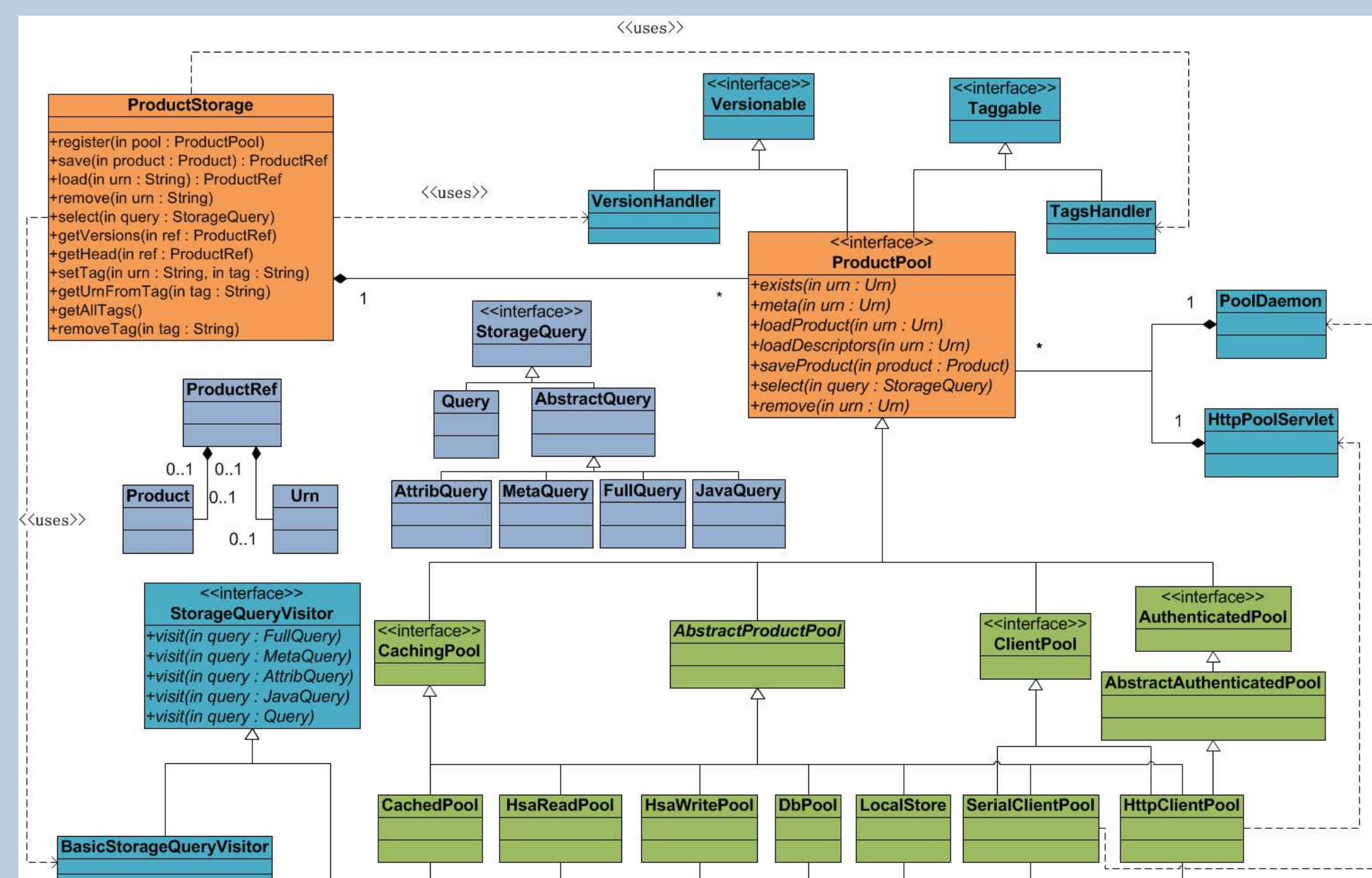
Bo Li and Maohai Huang wish to thank National Astronomical Observatories, CAS and Chinese Academy of Sciences for their funding support KJCX2-YW-T20.

PAL Features

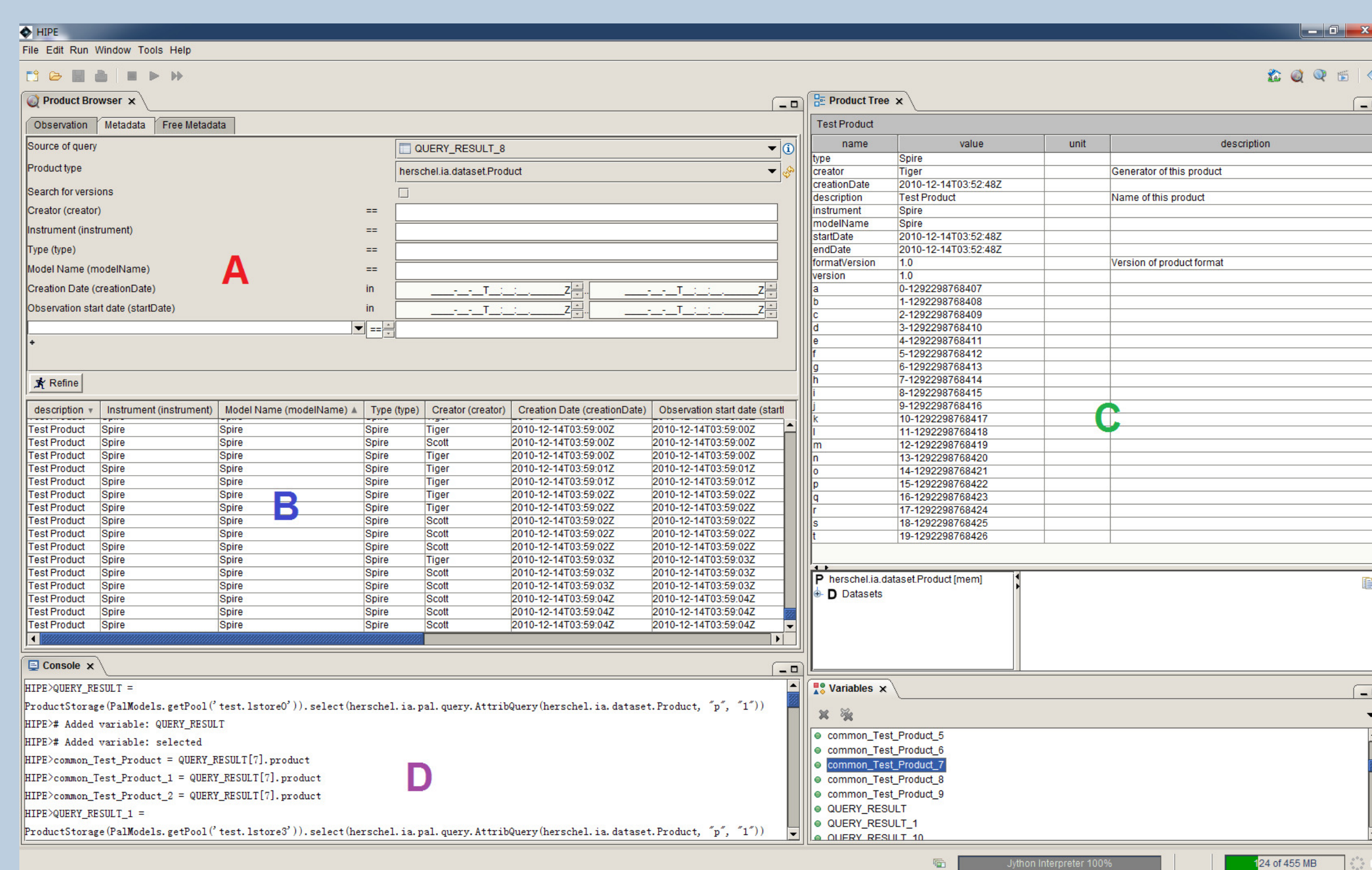
- Pluggable concrete storage system
- Default implementation uses FITS files to store data products
- Support to query on every field of data products using jython syntax or a simplified query syntax
- Light-weight access(e.g. lazy loading)
- Product Tagging
- Product Versioning
- Backward compatibility for removed or changing class definition
- Product reference counting for reliable deletion
- A common test suite to validate each product pool implementation

PAL Design

The design of PAL conforms to the general object-oriented design principles and is developed using Java and Jython. The PAL provides a consistent API for data access over any particular persistence mechanism. A general user could use this to organize and query data on their own computer, as well as from archive servers. Product pool is a PAL abstraction to data storage area. We have developed several product pools implementations to work with different data sources such as an object database, FITS files on file systems, and the Herschel Science Archive system. There are no technical limits in developing more implementations over other data sources.



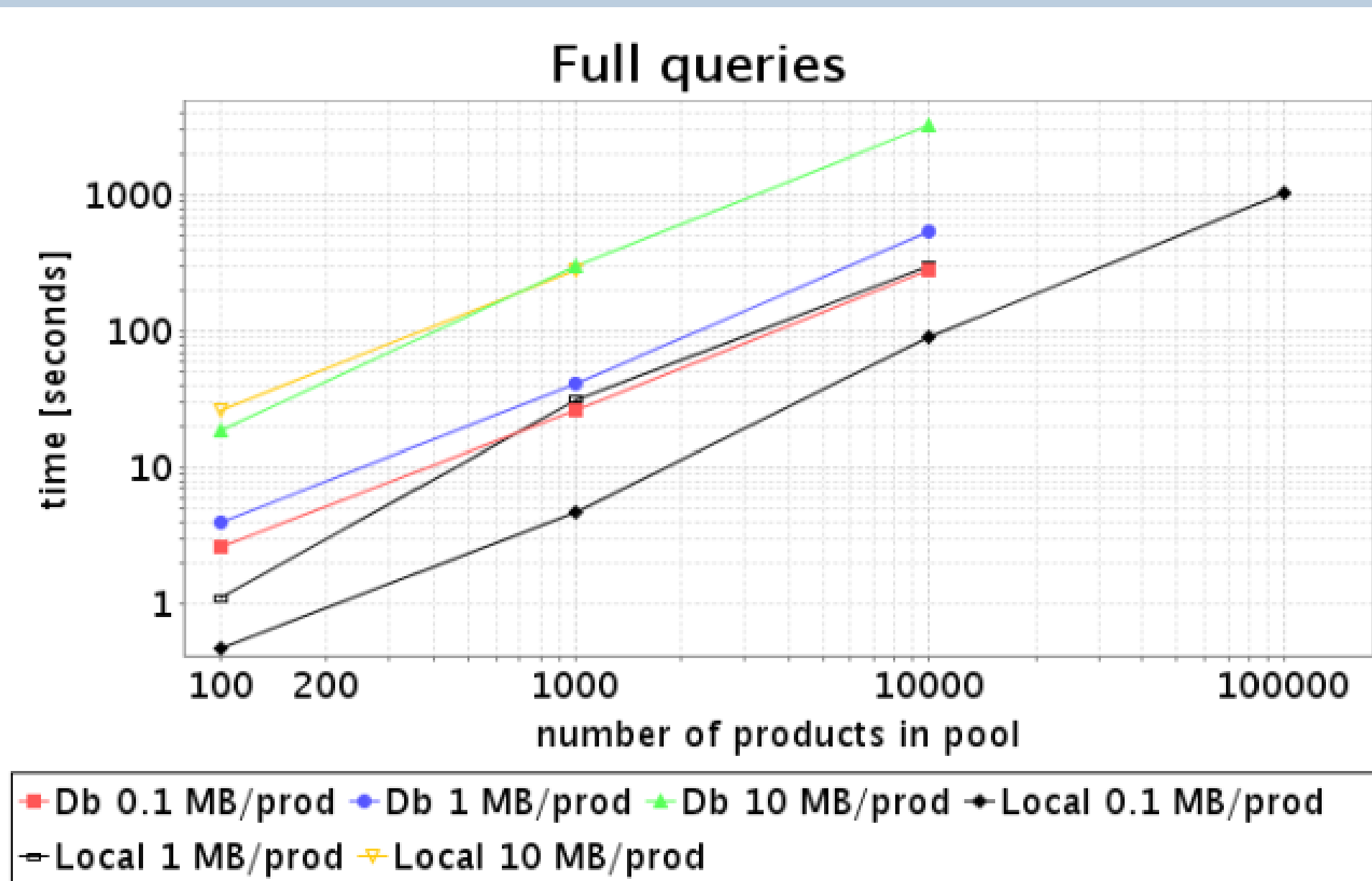
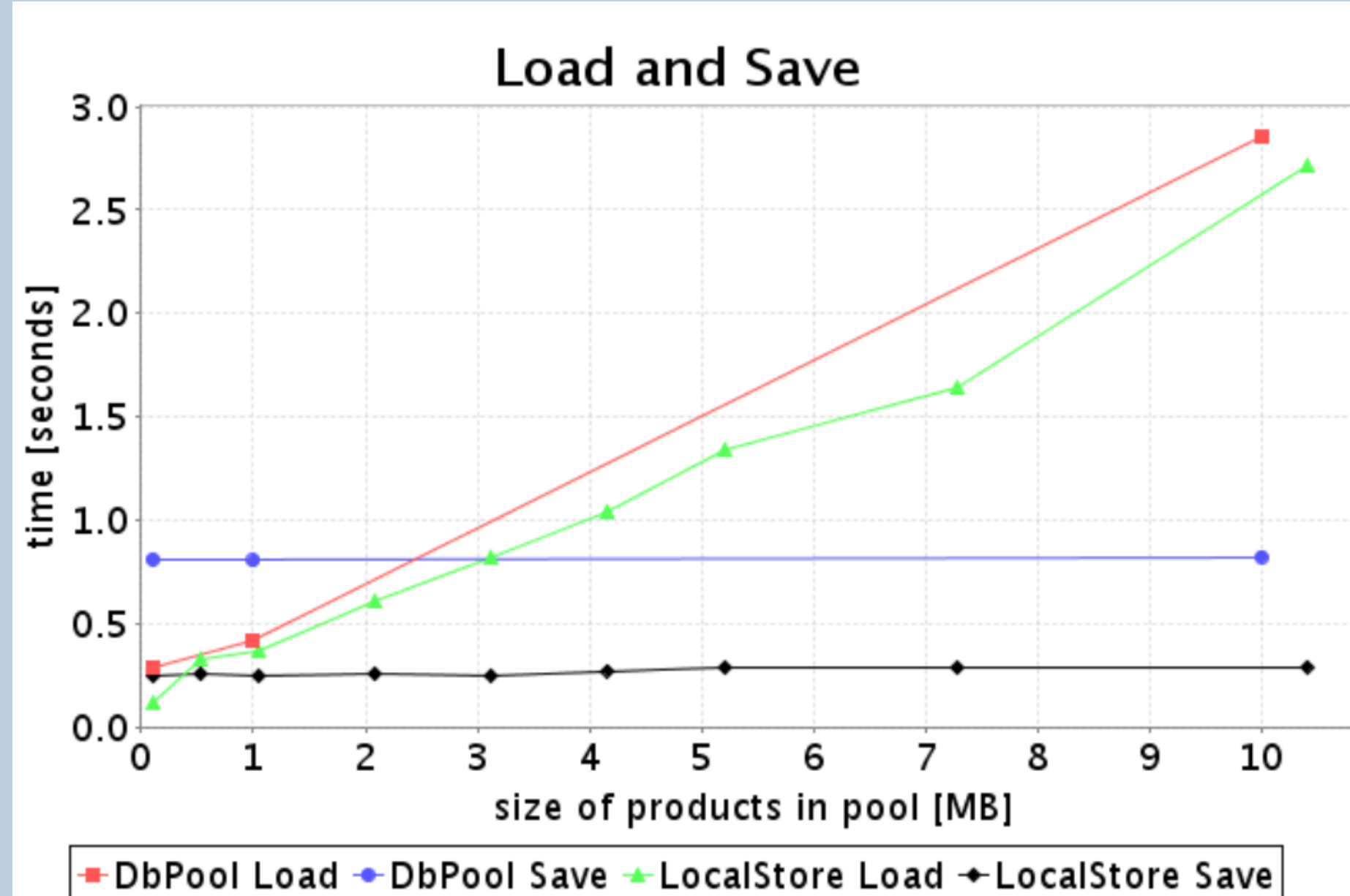
GUI



A graphical tool called Product Browser is developed with PAL to query (A), view the result (B), inspect a selected product (C) or load and initialize a product storage from the console (D).

The Product Browser is a component of the Herschel Interactive Processing Environment(HIPE, S.Ott 2010). HIPE provides a complete Herschel data processing and analyzing environment for astronomers and instrument teams.

Performance Benchmark



Contact Information

E-mail: libo@gsegment.com

Telephone: +86-10-57151257

Address: National Astronomical Observatories, Chinese Academy of Sciences.

Room B-616, 20A Datun Road, Chaoyang District, Beijing 100012, China.