BLINK : Billion Lines INdexing in a clicK

Nicolas KAMENNOFF
ACSEL, France

Sebastien REYBIER
SoaMI, France

Sebastien FOUCAUD
NTNU-ES, Taïwan

MENG-FENG Tsai
NCU-CSIE, Taïwan

CHENG-HSIEN Tang
NCU-CSIE, Taïwan

Project

BLINK is an open source collaborative project which aims at creating a computational framework. The goals are :

- Take advantage of available resources
- Optimized for concurrent multi-users
- Fast and accurate cross-matching
- ease of access to a complete tool chain

Indexing the Sky

One of the major steps shown by the functional design of BLINK is the Index Database. It answers two needs :

- Finding a basic way to organize data

- Match objects from various inputs

The main, maybe the only, way to determine an unique identifier for objects in space is by their position. We are currently implementing the Hierarchical Triangular Mesh [Szalay and al.]. HTM describes a Quad-Tree system able to locate and identify objects on a sphere.

Cross-matching

Astronomy, today, faces a data avalanche. Sky surveys generate terabytes of images and catalogs. In order to keep track of information about an object one has to identify the object on many inputs. As we are identifying object by their coordinates we have to face another problem. Precision and noise on inputs lead to errors. The HTM data structure does not handle this. We are actually working on new, more accurate solution which will be the first of the BLINK's tool.

Middleware

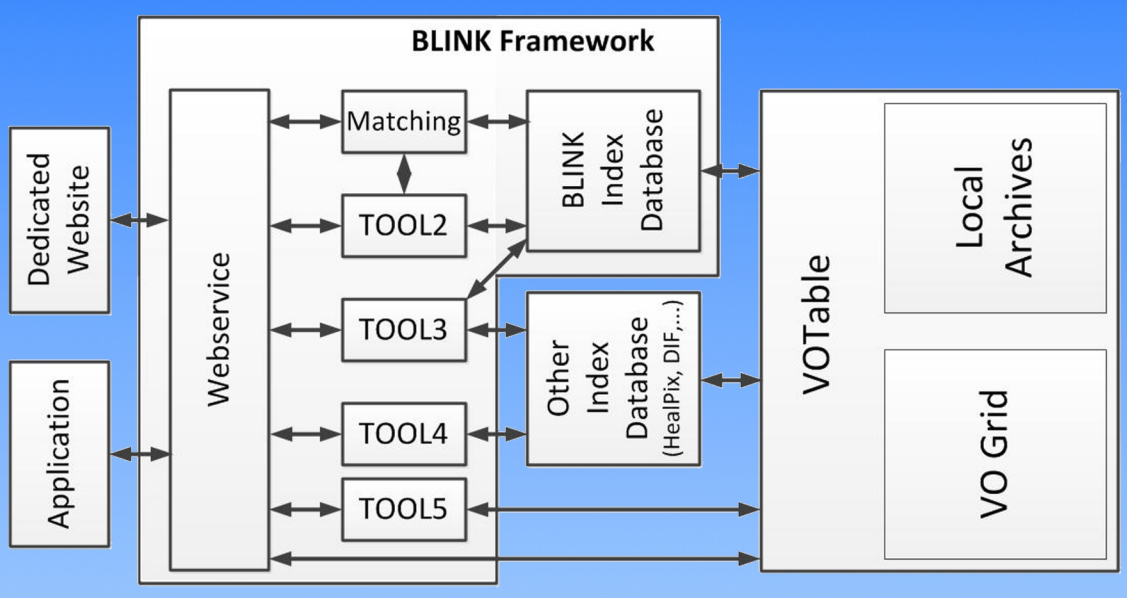
BLINK is designed as a middleware : It stands between user interface and data. User interface are external tools : through a web or software application. Needed data are gathered from local or on-line archives using International Virtual Observatory Alliance ; DAL (Data Access Layer) and VOTable format from the International Virtual Observatory Alliance [IVOA].

Even if BLINK as its own index database, we are convinced that it cannot fulfill all of the user's needs. Therefore BLINK will also be able to access other indexation systems like HEALPix [HealPix].

Accessibility

BLINK frontend is managed as a web service. A dedicated web application will allow users to send their requests and gather results. Results can be directly displayed as logs, plots or they can be downloaded in different file format.

This web service will also provide an interface for thick client applications.



Tool box

One of the major added-values of the BLINK framework is to provide computation tools. As shown on the functional design, tools get data from the BLINK index database, from external index systems, directly from the VOTable archives or from the output of other tools.

Like IVOA specifications describe a query language for astronomical data, BLINK uses a script to describe data requests as well as computation using those tools.

Multi-user

BLINK is intended to run on parallel IT systems. The software architecture is designed and optimized for concurrent multi-user access. Taking advantage of a various kind of hardware : multi computer systems (grid, cluster, cloud), GPGPU, AMD's APU and more to come...

[Szalay and al.] A.S. Szalay, J. Gray, G. Fekete, P. Kunszt, P. Kukol, A. Thakar, Indexing the sphere with the hierarchical triangular mesh, Microsoft Research Technical Report, 2005
[HEALPix] K. Gorsky, E. Hivon, A.J. Banday, HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, APJ, 2005.
[IVOA] http://www.ivoa.net/