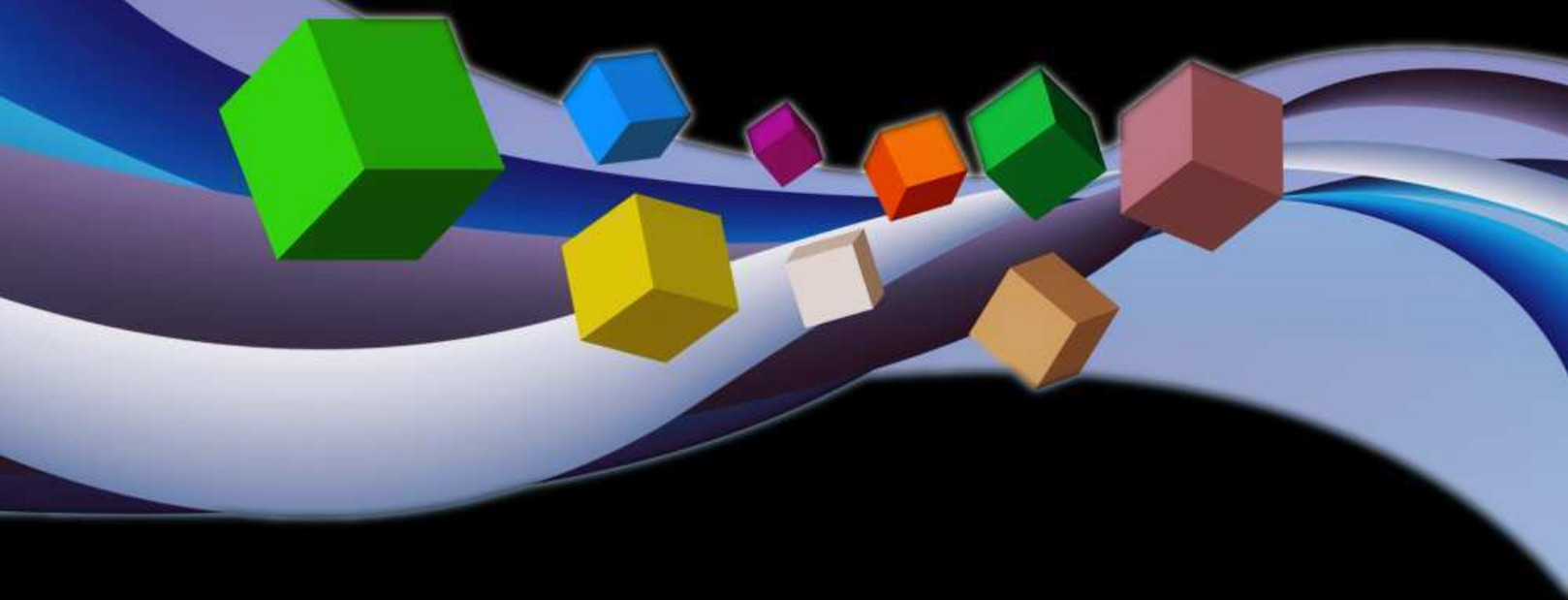
Grid Computing Cluster

The Development and Integration of Grid Services and Applications

Grid@USM
2008/09 Report



Edited by Bahari Belaton and Lim Lian Tze
Platform for Information & Communication Technology Research
Universiti Sains Malaysia

CONTENTS

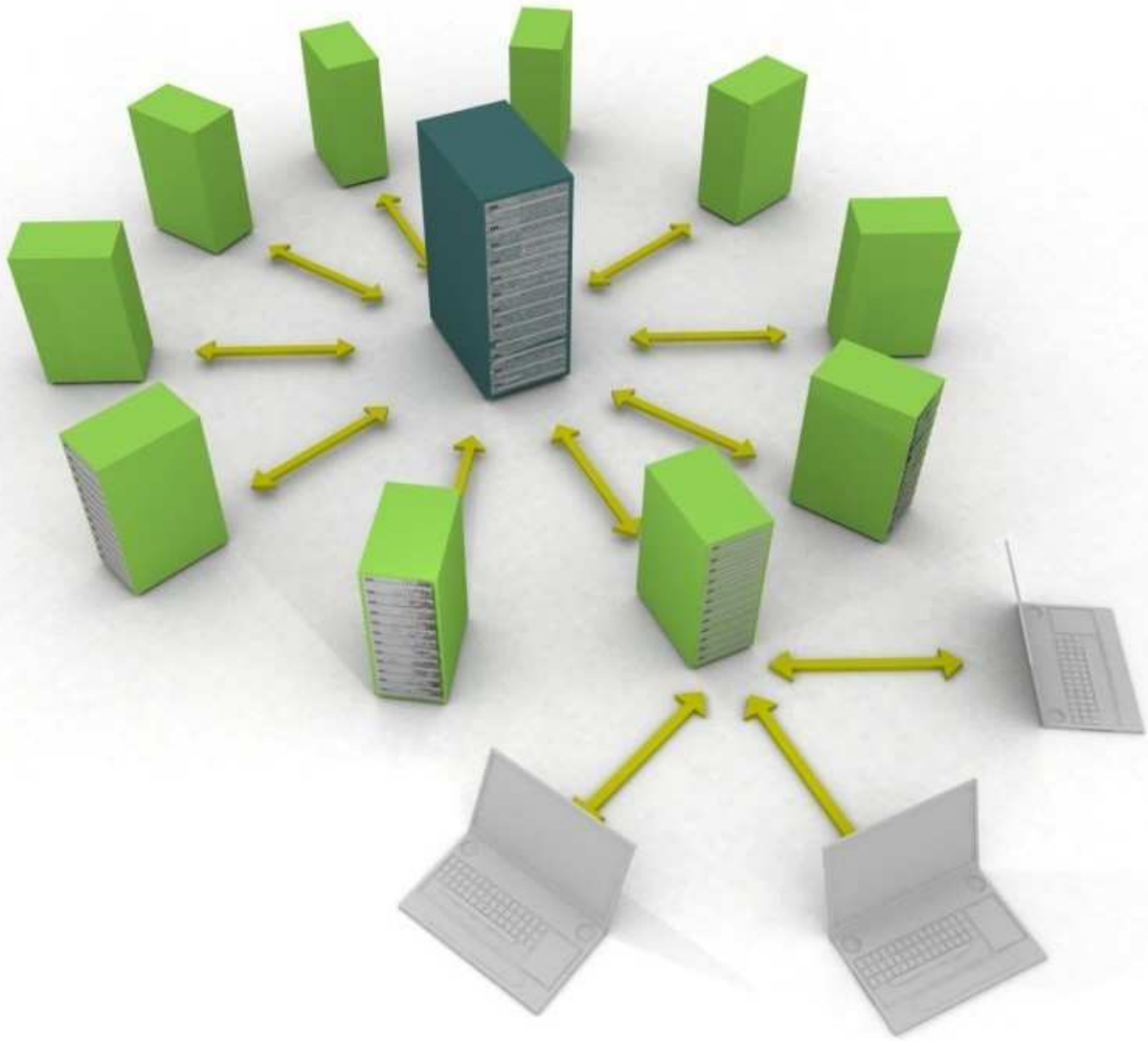
Project Overview

Sub-Projects

Project Activities

Note: The abbreviations Prof. and AP are used for *Professor* and *Associate Professor* respectively throughout this book.

Image: A simple network. Illustration by ©Svilen Mushkatov (http://www.sxc.hu/profile/svilen001).



Grid@USM: DEVELOPING & INTEGRATING GRID APPLICATIONS & SERVICES

USM'S R&D SUSTAINABILITY

To attain the level of excellence as a research university (RU), USM in recent years has seriously engaged herself in various efforts to formulate and chart her future research and development (R&D) directions. The pinnacle of these efforts was translated into a 2 year strategic plan "USM Research-Intensive University 2007–2009".

One of the crucial input elements for attaining and subsequently sustaining RU status is the creation of highly conducive research environment and infrastructure. Grid computing is an example of such infrastructure, aiming at providing excellent research facilities by ensuring good computational horsepower to support high impact domains in the bio-sciences, physical sciences, information and communication technology (ICT), Environment sciences, Education, Arts and others.

Grid computing is by its nature highly distributed geographically, consists of highly specialised equipment (storage, grid engines and management tools), as well as expensive. It is as such an ideal example of common, core computational resources to be created and pooled among aspiring researchers who require high performance resources.

Equally important, USM needs a focused, holistic and dedicated effort

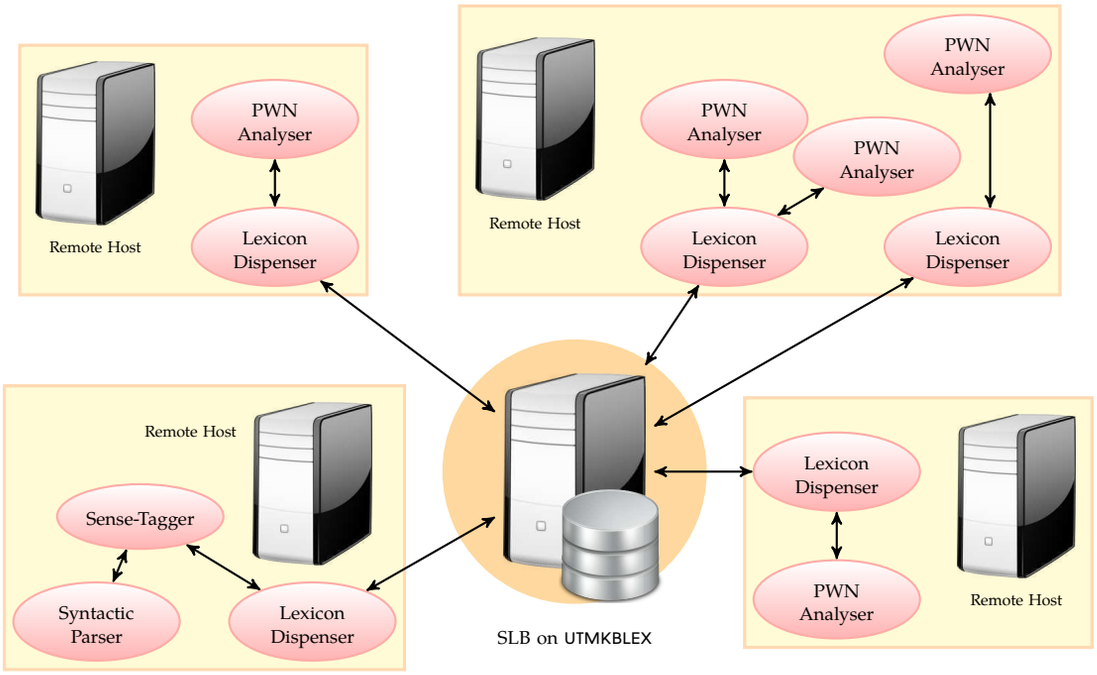


Figure 3: *Blexisma2* agents are deployed on USM Campus Grid nodes using the Globus job submission toolkit.

Type an English sentence: Spica is a star in the Virgo constellation. Disambiguate!

Spica$_{n.1}$ is a star$_{n.1}$ in the Virgo$_{n.2}$ constellation$_{n.2}$.

Spica	n.1	the brightest star in Virgo
star	n.1	(astronomy) a celestial body of hot gases that radiates energy derived from thermonuclear reactions in the interior
Virgo	n.2	a large zodiacal constellation on the equator; between Leo and Libra
constellation	n.2	a configuration of stars as seen from the earth

Type an English sentence: she is the star actress in our play. Disambiguate!

she is the star$_{n.5}$ actress$_{n.1}$ in our play$_{n.1}$.

star	n.5	an actor who plays a principal role
actress	n.1	a female actor
play	n.1	a dramatic work intended for performance by actors on a stage

(a) Different meanings of *star* in context

Type an English sentence: a school of fish swam past. Disambiguate!

a school$_{n.7}$ of fish$_{n.1}$ swam$_{v.2}$ past$_{r.1}$.

school	n.7	a large group of fish
fish	n.1	any of various mostly cold-blooded aquatic vertebrates usually having scales and breathing through gills
swim	v.2	be afloat; stay on a liquid surface; not sink
past	r.1	so as to pass a given point

Type an English sentence: the school's classrooms are spacious. Disambiguate!

the school's$_{n.1}$ classrooms$_{n.1}$ are spacious$_{a.1}$.

school	n.1	an educational institution
classroom	n.1	a room in a school where lessons take place
spacious	a.1	very large in expanse or scope

(b) Different meanings of *school* in context

Figure 4: SLB data generated by *Blexisma2* agents used to determine most probable meanings of ambiguous lexical item. A Web interface to the Sense-Tagger is available at `http://utmkblex.usmgrid.myren.net.my:8080/Blexisma2Servlets/CVSenseTagger`.

interested to improve its design to a multilingual setting, where more complicated cross-lingual phenomena will have to be considered. We also hope to improve the agent communication mechanisms to reduce latency. Apart from creating more agents that implement different learning algorithms and heuristics, we are also planning agents responsible for other NLP tasks such as deep parsing, detecting named entities, etc. In other words, we hope to have agents of various responsibilities so that they can be 'mixed-and-matched' to construct new NLP applications, such as those mentioned in the introduction.

On the performance issues, we may attempt several solutions for the SLB bottleneck problem mentioned earlier:

- hardware (RAM) upgrades,
- further optimisation of PostgreSQL server settings,
- database connection pooling,
- load balancing,
- parallelising queries.

PROJECT PUBLICATIONS

Lim, L. T. & Schwab, D. (2008). Limits of Lexical Semantic Relatedness with Ontology-based Conceptual Vectors. In *Proceedings of the 5th International Workshop on Natural Language Processing and Cognitive Science (NLPCS'08)*. Barcelona, Spain; pp. 153–158.

Schwab, D. & Lim, L. T. (2008). Blexisma2: a Distributed Agent Framework for Constructing a Semantic Lexical Database based on Conceptual Vectors. In *Proceedings of the International Conference on Distributed Frameworks & Applications 2008 (DFmA 2008)*. Penang, Malaysia; pp. 102–110.