Project title
Data Location in Complex Information Networks

Project number: IMURA0068

Monash University supervisors: Professor David Abramson and Professor Mohan Krishnamoorthy
Monash University contact: Professor IT Faculty;
   Email: David.Abramson@infotech.monash.edu.au

Monash University contact: Professor and Associate Dean of Research (Engineering);
   Email: Mohan.Krishnamoorthy@adm.monash.edu.au

IITB supervisors: Professor Narayan Rangaraj Prof. Vishnu Narayan
IITB contact: Convenor, Industrial Engineering and Operations Research Programme
   Email: narayan.rangaraj@iitb.ac.in

Research Academy theme/s
List only the research academy theme/s that is relevant to the project
   1. Advanced computational engineering, simulation and manufacture

The research problem
With the advent of the Internet, there are a lot of data around these days and there is more being created by
the minute. Information sharing via the Web has risen to dominance in the last few years due to applications
that are data hungry. Moreover, as enterprises become more and more global, there is more data that is
generated from within large enterprise-wide software programs. Large data warehouses are routinely
accessed by companies all around the world. This access needs to be speedy and efficient in terms of both
time as well as access. As Internet-based "services" become more and more prevalent, it is important to
optimise the location of data warehouses in large and complex information networks. This will increase the
availability, reliability and accessibility of such data. In order for this to happen, latency and cost of access
needs to be minimised. Mirroring is one of the techniques to satisfy such a need. Enterprises need to locate
mirrors optimally. Moreover, depending on usage patterns from around the world, the location of these mirrors
will change over time. What is needed is a method, model and implementation framework for locating,
migrating and managing data-mirrors based on the dynamic monitoring of users’ requirements.

It is possible that this problem can be modelled as a Stochastic Hub Location Problem. Solution approaches
are expected to be both heuristic methods as well as exact solution approaches to this large and complex
problem.

Project aims
The aim of this project is to design better ways to locate data mirrors in widely distributed and routinely
accessed global data warehouses.

Expected outcomes
The project outcome is expected to be cost and latency minimisation in complex and large data warehouses
through optimally locating large data volumes.

Which of the above Theme does this project address?
Advanced computational engineering, simulation and manufacture
How will the project address the Goals of the above Themes?
Our approach will lead to optimal modelling of information stores in large and distributed global organisations. Our approach will build an abstract mathematical model and will build computational algorithms/methods for solving these models. We will test these approaches on data from industry – most likely the telecommunications industry.