Harnessing Big Data

High-speed electronics and large-scale arrays of sensors can produce data sets that are too big to analyse in one straightforward step. Frazer-Nash Consultancy has comprehensive experience of delivering analyses in this challenging environment.

SOURCES OF BIG DATA

Cutting-edge research projects can require complex sensor arrays or detectors with a large number of readout channels. These detectors can produce very large data sets, requiring petabytes (millions of gigabytes) of storage space. As more and more devices are added to the Internet of Things, the amount of data associated with a given array of devices will also produce large data sets that cannot be processed in a straightforward manner.

DATA STORAGE

Big data problems often imply that the associated information normally cannot be stored on one hard disc. Depending on the size of the data set and the application, cloud or distributed RAID storage solutions will be required. We can provide advice on how to best stream data to an optimal storage solution, using either commercial or bespoke storage implementations.

To enable rapid analysis of large data sets, it is often necessary to use a mixture of file-based relational storage and true relational databases. To perform analyses on very large data sets, we have written bespoke data serialisation tools and deployed applications with Oracle, MySQL and PostgreSQL databases.

DISTRIBUTED ANALYSIS

When analysing very large data sets, the data access speed associated with a single hard disc can limit analysis processing time. We have an in depth understanding of how to mitigate this and create high-speed storage solutions.

It isn't normally possible to execute complex data analysis programmes directly on disc servers. But we have extensive experience of operating in this environment, and have written a range of distributed data analysis algorithms.

It is often too expensive to rerun complex analyses on large data sets. To avoid this, we have written data selection algorithms that provide data summary formats. We can provide advice on how to create optimised data summary formats, reducing the need to process the complete data set.

ANALYSIS TOOLS

Fast, flexible analysis tools are a vital ingredient in performing efficient data analyses. We have written software frameworks that provide the fastest possible processing time together with a flexible implementation that can be easily expanded.

To allow standard queries and high-level analyses, it can be necessary to provide a web application or a graphical user interface (GUI). We have extensive experience of developing web applications, using Microsoft, Oracle and open-source software stacks. We have also implemented platform-independent GUIs using Java and Python

CLOUDS & DATA GRIDS

When local storage and data processing facilities are not sufficient, cloud computing can offer an agile alternative. We can propose solutions with an optimal mixture of local and cloud computing for a given application or operational outlook.

Data grids are used for particle physics and bioinformatics, enabling analysis algorithms to be run on several different data centres. We have a comprehensive understanding of data grids, allowing us to provide complete solution proposals.

For the most complex projects, it can be necessary to replicate large amounts of data between more than one data centre. We have designed and implemented simulation tools that allow us to predict the best balance of computing, storage and network bandwidth for a given set of user requirements.

OUR CAPABILITIES

- Sixteen years of working at the forefront of data analysis.
- Detailed understanding of hardware, middleware and software infrastructure solutions.
- Proven track record of producing distributed analysis frameworks and designing associated algorithms.
- Wide range of experience of designing, operating and using distributed computing facilities.

For more information about Frazer-Nash please visit our website.

www.fnc.co.uk

www.fncaustralia.com.au

Offices throughout the UK and Australia

FRAZER-NASH CONSULTANCY