



REAL-TIME ACCURACY

ELIZABETH WILKINSON FROM SNOWFLAKE SOFTWARE, SHARES HER OPINIONS AND SUGGESTIONS FOR A SMOOTH RUNNING OF THE 2012 OLYMPICS

The 2012 Olympics for the United Kingdom will be as much about providing safe and reliable facilities as it will about proving our sporting prowess. With the world focussed on the country's network infrastructure, the ability to respond quickly and accurately, even under potential adverse circumstances, will be paramount. With a failing IT infrastructure, there could be far more than the medal table under scrutiny for the British.

With the Olympic venues scattered across the country, the importance of place will play a major role in coordinating both the smooth running of the Olympics and any potential security efforts. But how do you manage to have one unified view of all locations when each location has numerous data sources, different applications and different data models? One concept would be to demand that each venue and each application stores its data – at source – in exactly the same manner. However, this would be unreasonable and not practical. The alternative method of collating the data into a unified view is via the concept of data harvesting.

What is data harvesting and why is important

The concept of data harvesting can be quite daunting to most users. It provides yet another theory to learn, more industry terms (round-tripping) to follow and, undoubtedly, there will be more acronyms to study. But the concept itself is really very simple. In effect, data harvesting is the gathering of data from numerous disparate databases (and different data models) into a single database from which it can be re-published in a unified manner.

The organisations involved in the Olympics will be under pressure to make the information they hold in their databases available to centralised sources or “hub databases”. The “hub” in turn collates the

data from multiple suppliers and presents it in a unified manner – easily understandable to the outside world – often against a geographic background. Whereas this used to be achieved through time-consuming paper trails and documents, the continued adoption of the internet as a source of information means that every one now expects up-to-date intelligence reporting all the time at the click of a button.

To ensure that the Hub Database is constantly up-to-date, the suppliers need a quick and easy mode of making their data available to the hub. Data harvesting has gone digital.

Local and Hub Database Relationships

To understand the concept of data harvesting it is important to reiterate the relationship between the Local Databases and the Hub Database.

The Local Database holds the information as collected from the field and input according to the business process of a particular organisation (organisations store similar data in different databases and often use different database models). The data then needs to be fed to the hub. The hub is responsible for collecting – or harvesting – the data from multiple Local Databases, loading that data into a centralised database and then hosting the data and making it available to external users.

In order to harvest data to a unified hub, a common model for the exchange of data needs to be agreed so that the hub receives data in consistent form from all the local databases. With an agreed common model, each Local Database then needs to be mapped to this model in order for it to be harvested by the hub.

Defining a common model is ideally suited to XML Schema, which in turn makes XML / GML the ideal format for harvesting geospatial

CORE TECHNOLOGY

data. Utilising OGC standards' technology also provides defined parameters around the model and schema definitions, ensuring interoperability and openness. The organisations, then, agree to what is now known as a "community schema" – an agreed model against which each can map their data and consequently feed it to the hub.

Since the community schema is utilising OGC standards, the organisations can also make use of all the standard validation tools that come with GML and XML to test that the data provided to the hub is valid according to the community schema. This is provided without additional software licensing costs. However, it is true to say that the accuracy of the source data itself is still a trust issue and must be manually checked if in question.

The Duty of the Local and the Hub Database

Once the mapping to community schema / common model has occurred, the easiest way to respond to harvesting demands from the Hub Database is to make the data available to the internet (in some cases, intranet) via a Web Feature Server.

As such, the supply is achieved via these three easy steps:

- 1) Understanding of the common model / community schema
- 2) Mapping to the common model / community schema
- 3) Publishing the data via a Web Feature Server

Steps 1 and 2 take place locally within the respective organisation. An additional

advantage of this is that these are one-off actions. Once the mapping has taken place and a connection has been established via the Web Feature Server, the local database need only be updated and saved. The hub "harvests" all the suppliers periodically to gather any updates.

Per above, the first duty of the Hub Database is to periodically harvest the data from all the Local Databases, collate the data and load it into the Hub Database. Since mapping to the common model has already taken place at the Local Database, no further action need take place here. Once hosted, the prime responsibility of the Hub Database is to make the collated data available to others. It can do this via either (or both) of the following options:

- 1) The harvested data can be collated and integrated with GIS applications to display the information in an interactive and functional environment. Often, this application geospatially represents and displays the data. In a typical environment, this would then be hosted as a service for multiple end user access via the internet.
- 2) The Hub Database may be used to supply the collated data to various third party applications. The data has already been provided according to the model required by the hub, but it may have to be re-published according to one or more alternative models to meet the needs of different applications.

Alternatively, the Hub Database might be required to supply the collated data to another Hub Database. And so the cycle continues: the Hub Database becomes, in itself, a Local

Database in a hierarchy of databases. A regional hub, for example, could supply data to a national hub.

A Unified View

By harvesting data into a central hub, end users of the data are provided with a unified view of the data even though it has been created in a disparate group of local databases. The local databases are responsible for providing data to the hub, but only the hub is responsible for providing data to the end users. This means that only the hub needs the scalability and high availability that may be required to support those users. Moreover, by translating the data into a common model for exchange, the local databases can continue to operate using their own local model. This allows existing databases to be brought into the network without modification and therefore without disruption to existing business activity or extra cost.

Writing now, four years before the games are to start, it is easy to imagine that 2012 might well be Britain's finest hour or, indeed, its worst nightmare. Whereas it can not take responsibility for the much of the infrastructure and certainly has no control over the medal table, the geospatial industry does have a strong responsibility to ensure that its side of the IT infrastructure runs smoothly, on time and under budget. By putting best practice into place now, this can certainly be achieved.

Elizabeth Wilkinson, Snowflake Software,
elizabeth.wilkinson@snowflakesoftware.com



Create a Safer World®

→ Uniquely, Intergraph provides real-time operational systems for public safety and security organisations around the world including some of the largest airports and metro systems, major events from the Winter Olympics to the Pan-American Games, border security and emergency services organisations from Toronto to Mumbai.

INTERGRAPH

For more information email:
sgi-uk@intergraph.com
or visit
www.intergraph.co.uk/sgi