SOPAC–SPC merger expected to boost GIS&RS Support for Pacific Island Countries

During the first week of August Pacific Leaders endorsed the decision of the joint meeting between SPC, SOPAC and SPREP on the Regional Institutional Framework that most of SOPAC joins SPC. Prior to this decision SOPAC already carried out activities jointly with SPC in Tuvalu and Kiribati, where high-resolution satellite image data was utilised to map coconut palm, mangrove, shrubs and other land cover types at 1:10,000 scale. Long before the Pacific Leaders decision in 2008 it was already agreed between the two organisations that image data purchase and pre-processing be carried out by SOPAC's GIS&RS unit for both organisations. The demand for image data is expected to increase at SPC when the Population GIS as well as the Land Resource Division move to image backdrops for their mapping purposes. Consequently software at the GIS&RS unit is being run to fuller capacity, which necessitates the purchase new and more specialised software like GeoPDF explained in this newsletter. Currently the GIS&RS unit can provide the following services for member countries:

O   Advice on available and suitable image data and purchase.
O   Image data pre-processing and GIS backdrop production.
O   Haze removal and atmospheric correction of image data.
O   Scanning of A0 size maps and converting to GIS backdrops.
O   Scanning and converting of aerial photo negatives.
O   Establishment and editing of digital elevation models (DEM).
O   Shallow water bathymetry.

Figure 01: (Left to Right) Litia G, Wolf F, Elizabeth L, Tato M, Joy P, Vilisi T

All these tasks need special hardware and software, which can be too expensive to be purchased by SOPAC's and SPC's island member countries individually, however, utilised at one central agency cost effectively for all countries the investment pays off as the equipment and software are utilised to their fullest capacity. Before the merger of the two regional organisations SOPAC was only responsible for mapping marine non-living resources making on-land vegetation mapping 'outside' SOPAC's mandate. With the merger and the wider mandate the GIS&RS unit sees opportunities with supporters of vegetation mapping, for example GTZ (German aid) which supports it wherever it fits into adaptation to climate change. The GIS&RS tool established and honed within SOPAC will now be available to the wider Pacific Community through the merger, allowing for better utilisation and in future, more specialisation. A member of the unit has been assigned to investigate better interpretation of vegetation with the available satellite images by applying specialised image enhancement. We are also discussing if the unit will take radar data analysis and pre-processing onboard, which would not have been suitable before.

Contents

COSMO SkyMed

is an Italian satellite system having at the end 4 satellites in space flying in a sun synchronous polar orbit at an altitude of 619 km. Each satellite is capable of recording a total of 1800 images daily or 560 GB. The satellites have a radar system onboard which can penetrate the clouds and take pictures any time day or night. In **Scansar mode** the system can record an area of 100 x 100 km to 200 x 200 km with 30 x 30m to 100 x 100m pixel size. In **Stripmap mode** the system records scenes of 40 x 40 km or 30 x 30 km with 3 x 3 m to 15 x 15 m pixel resolution. In **Spotlight mode** the satellites only record scenes of 10 x 10 km but with 1 x 1 m pixel size. Due to the number of satellites already in space (3) images can be recorded from any part of the world on a daily basis. Order time can be reduced to 6 hours. This satellite system can be utilised for fast disaster response of flood mapping or coastal erosion or
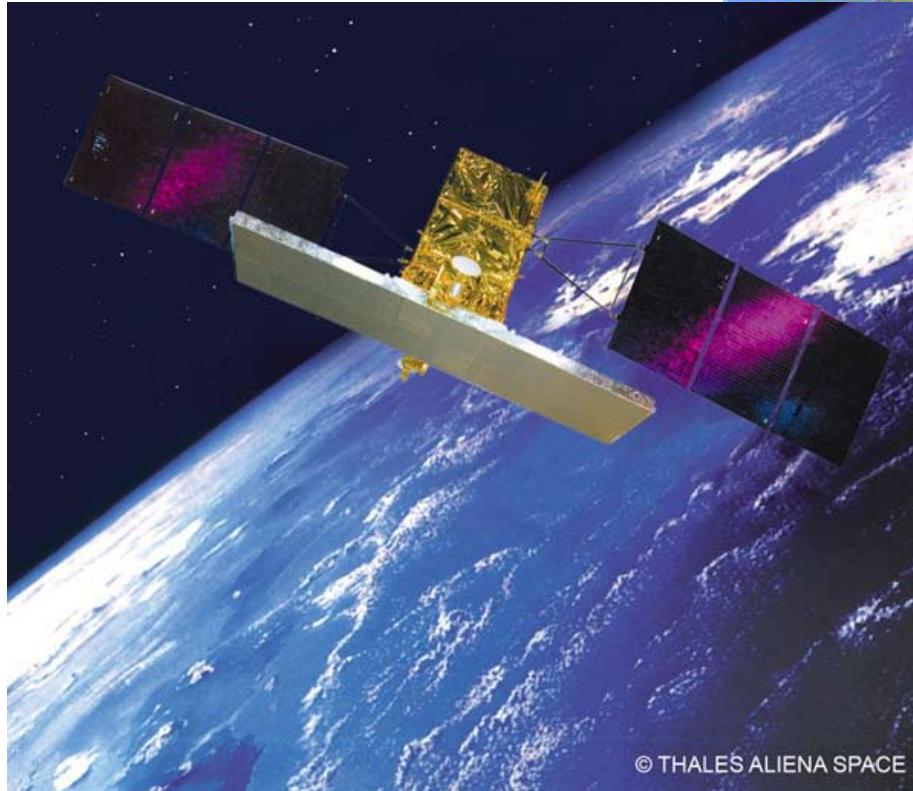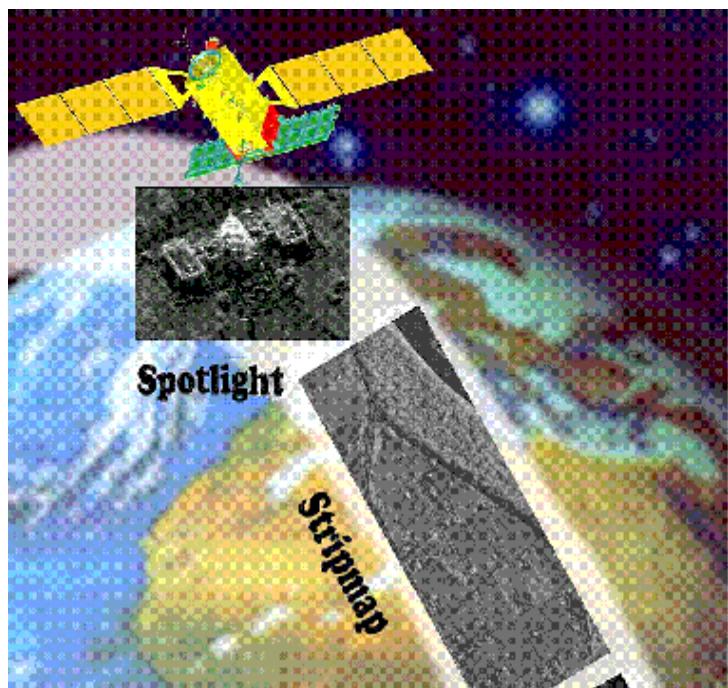

© THALES ALIENA SPACE

for shipping vessel detection to cross check the Vessel Monitoring System at FFA in Honiara. SOPAC has established a link with the company GAF AG, which distributes and pre-processes the image data.

Pacific Mapserver Status
Vilisi Tokalauvere,
GIS&RS Officer,
SOPAC

Mapservers allows users to display interactive maps of any area in the world in any projection. A user can zoom on a map, select which geographic layer to see and query any geographical object through the web.

The SOPAC EDF8/9 Project: "Reducing Vulnerabilities for Pacific ACP States" installed and helped develop mapservers in 14 countries around the Pacific during 2003-2007. The hardware for each mapserver cost the Project $US7, 000 to $US10, 000. A major reason behind it is to improve planning and better decision
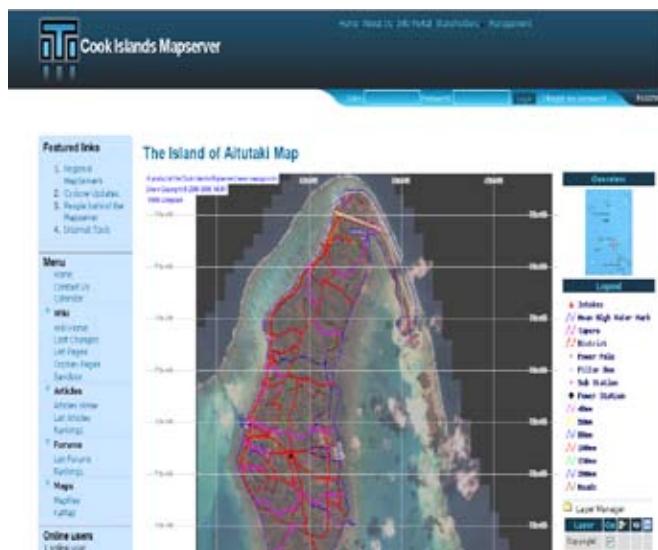
Figure 01: *The Cook Islands Mapserver*

making based on factual and accurate information by having geographical information accessible to all categories of the population through the internet.

Mapservers have a huge potential. For government departments, this can be an efficient platform to develop cartographic portals which can provide solutions to manage a range of services from environmental monitoring to assisting in disaster response, from territorial marketing to promoting tourism in the beautiful Pacific Islands. The server has benefits for the private sector as well, such as having the location of their service or sales outlets anywhere in the world. This is a key success factor which accompanies a potential customer directly "right into the virtual store". For any Pacific Islander, it can represent an easy and immediate way to locate a particular venue or the offices of public services on an interactive map.

Sadly, a number of mapservers installed in the 14

Pacific Islands Countries are currently offline. Reasons for the mapservers being offline include Internet bills being too costly, defective server UPS not replaced, mapserver managers getting a new job or no one is officially appointed to manage the mapserver. A number of the mapservers that are running have the problem of geographical data not regularly populated to the server.

A few Pacific Island countries have realized the huge potential of their mapservers. They ensure that maintenance costs of the server are made part of the budget of its host ministry.One such example is the Cook Islands Mapserver.

The Cook Islands Mapserver was installed in October 2006. Like all mapservers in the Pacific, it has also encountered its fair share of problems and challenges such as high internet costs. Despite this, the Cook Islands Mapserver rarely goes offline since

Figure 02: *Mr.* Tangiruaine

its first installment and is regularly upgraded to the latest software. How is this possible? Firstly, the Cook Island's Ministry of Infrastructure and Planning which houses the mapserver has a set budget to provide the server with Internet connection. The Ministry pays $NZ950 per month for 10GB volume at 512Kbits. Secondly, it has a full time dedicated system administrator; Mr.Timoti Tangiruaine who has both ICT and GIS skills which is rare in the Pacific. This allows him to not only manage the server; by upgrading softwares when needed, fixing hardware and software problems, but also to upload the latest GIS data and display it on the mapserver. Importantly, the managing of the mapserver is one of his official responsibilities and he reports on it to the Ministry's management and finance section. Thirdly, the server does not only have a strong support of its ministry but is also supported by other government departments that regularly provide their data to be uploaded and displayed on the Mapserver.

It is evident from the Cook Island's example, that in order to keep a country's mapserver running with the latest GIS data it will need the full commitment of the host Ministry, the national GIS Community, its system administrator and the Government. Countries that continue to have their mapservers running from its first installment are Kiribati, Nauru, Papua New Guinea, Solomon Islands,Tonga and of course the

Cook Islands.

Each country, may have their own solutions in order to keep their mapservers running. Readers are also welcome to contribute and share their ideas by sending an email to: vilisi@sopac.org.

Below is a summary of the mapservers that are offline:

| Country	Host	Reason for Offline
Federated States of Micronesia	Department of Economic Affairs	Software Installation Problem
Fiji	Mineral Resources Department	Defective UPS not replaced since 2007
Marshall Islands	Marine Resource Authority	Software Installation Problem
Niue	Internet Systems Office	Software Installation Problem
Palau	Office of the PALARIS	Defective UPS not replaced
Samoa	Samoa Meteorology Division	Internet cost too high
Tuvalu	Department of ICT	Software Installation Problem
Vanuatu	Department of Lands	No one officially apointed to manage the server

Turning Data into Information

Ken Mathers,
GIS Consultant,
Intergraph
New Zealand.

GIS has evolved over the years. GIS started out using technologies such as Computer Aided Draughting (CAD) and initially the emphasis was on data capture. It was primarily used inside an organization to improve the quality of the drawings and speed capture and editing.

The early attempts consisted of replacing a paper map with an electronically produced version. The cost justification in those early days was usually based on productivity improvements gained by no longer having to manually draught the drawings, plans and maps. Even bigger gains were actually attained by the editing ability that automatically came with the capture.
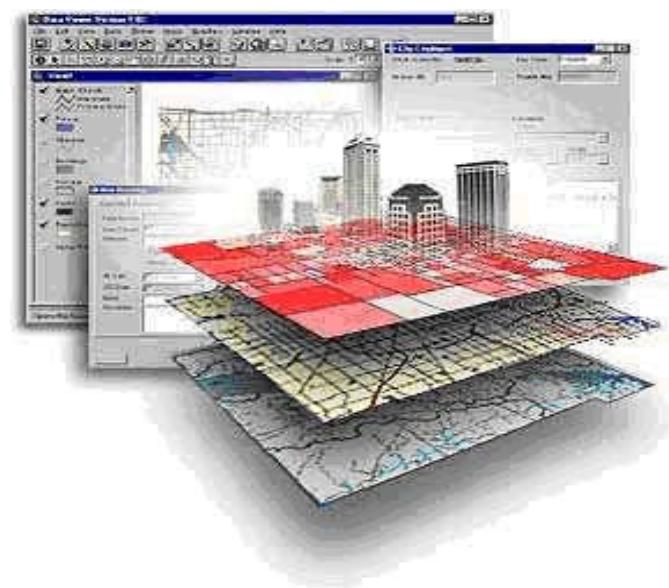
Figure 01: Turning your data into information

The next part of the evolution was all about adding additional "intelligence" to the drawings so that someone could click on a graphic and see tabular information about the graphic. This is when the terminology changed to GIS and the emphasis was now truly about "Geographic Information Systems" rather than just simple maps. But it was still all about adding some "Intelligence" to a map rather than spatially enabling Data'

With the maturity of the GIS systems came many more opportunities. People quickly discovered more and more business benefits from being able to analyse their Geographic data using a combination of the Spatial and the attribute information.

In parallel to this emergence and maturity of the GIS environment, there has been a HUGE progress over the years in the gathering and storage of "DATA" in every organization. This Data is everything from the billing information of any company to details about assets, travel plans, maintenance activities, medical treatments, emergency plans, weather patterns, etc, etc. In fact data is now available about almost any topic that you can imagine simply by using a Google Search.

All of this Data is captured to enable organizations and individuals to better achieve their daily work. It is far easier to perform almost any task if you know all of the details regarding whatever it is that you want

to do.

Now we get on to "Information". This is the differentiation between facts and figures being stored on a computer somewhere and that same "Data" being readily available to Inform someone. Once the data is readily available then it changes from being data to being Information.

This is where GIS comes to the fore. Many people, even within the GIS environment, are still struggling to understand what the term GIS actually is all about. Many people just limit their thinking to what they need to do in order to deliver what they have always delivered. And in many cases that is limited to still producing Paper Maps that may have been built using the available data and have thus achieved a small part of the goal of delivering "Information", however the potential is very limited compared to what is actually achievable.

Today, most people realize that almost every aspect of our lives and every aspect of all of that data that is stored away on computers is related to some Geographic location. We cannot do anything without first going to some location. And every physical item has some location. Some items are locked to a single fixed location and some may move, but even movable items have some precise location at any point in time.

So now, the focus has changed and a significant part of the effort of GIS systems is no longer the gathering of that Data and the production of the nice paper maps, but rather it is focused on making ALL of that Data that is stored on computers all through an organisation available as Information to all those who can benefit from the Data to make informed decisions about whatever it is that they need to do. It is all about the relationship of Data to some Spatial Location.

There are a number of technologies that have evolved and in some cases are still evolving to make this process easier. The Internet is obviously the main one of these technologies but it is the applications on top of the Internet that are truly making this possible.

With all of this power and capability comes potential for confusion. It is not sufficient to just gather data and expose it on the Internet. Exposing your Data is a relatively simple process in today's Internet and Google Map enabled world. But turning your "Data" into "Information" requires a reasonable level of care and responsibility.

The most important consideration when exposing your Data is Metadata. This is tabular Information that will allow the recipients to understand the context, accuracy, and relevance of the Data.

All GIS professionals have a responsibility to do more than just pushing data out into the world, "Just because you can…" For the data to truly be Information, it must have the context as to it's age, accuracy, method of capture, relevance to task, etc. It must also be presented in a way that will allow the Data to be used by the intended

recipient. Further, the Data must be controlled such that valuable and competitive Information is not exposed to the wrong parties.

Although this is an exciting and interesting era, it is also one that comes with responsibilities.

Some things to consider when planning to turn your data into Information include:
- Who is likely to use the data?
- What purpose will it be used for?
- What format is appropriate?
- What metadata should I include?
- What is the delivery mechanism?
- How do I keep the supplied data current?
- Does it need security?
- Will the recipient provide me with changes?
- Etc, etc

This is by no means a complete and exhaustive list of all questions and considerations, but rather a taste of the sort of things that you will need to consider when planning your project.

ArcGIS Schematics
An extension for ArcGIS desktop to manage and analyze connectivity and networks

Craig Clouet,
Account Manager,
ESRI,
Hawaii.

Maps are graphical representation of place. Maps are abstracted models used to communicate and visualize information. Most people use maps to determine where they are, their position, or where they want to go. The common reference or topographic map is used around the world. The use of abstracting reality is the only way to fit the complex and large world onto a little piece of paper or computer screen. Take for instance a highway, in the real world it may have 4, 6, or even 12 lanes, but on a map it can be represented by one thick line, with a specific color and symbol. The same is true for cities, while Mexico City, Mexico has over 20 million people, and Apia, Samoa 60 thousand people, both might be symbolized on a world map with a single star, as both are the capitols of their country.

In order to communicate and visualize information that has properties of connectivity and relationships, a further abstraction can be useful; the use of a schematic diagram can be more effective than a map. Connectivity and relationships are key information for managing and viewing networks, which are a way to use any event occurring between objects and view all the information in an organized manner that is easy

to analyze and understand. Connectivity can represent different varieties of information and can be used to generate diagrams that show how entities are related. Take for example this bus route map from London, England. It has geographic properties, such as direction; North, South, East, and West but the
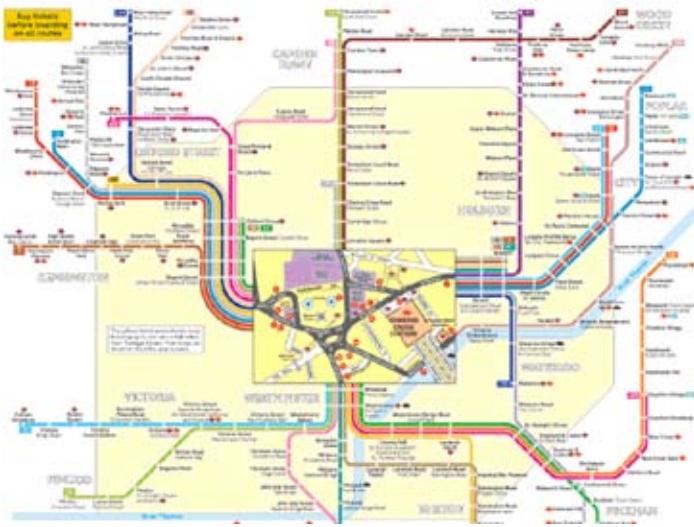
Figure 01: Bus route map, London, England

roadways are greatly simplified, removing roads that the bus
does not travel on, and also offsetting the routes that are on the same road so that all can be seen without overlapping each other.
This next example helps to demonstrate the different needs of a map. On the left is the aeronautical chart that a pilot would use for flying in the Hawaiian Islands. On the right is the same set of islands, but as a traveler would see the map, mainly just the connection from one airport to another. Transportation is

Figure 02: (Left) Aeronautical chart, Hawaii Is (Right) The Map a traveller would see during the trip

not the only use for schematics; other industries that heavily use these techniques are the utilities such as electric, telecommunications, water, and petroleum and gas.

One of the main features that the utility industries share is the need for networks. For ESRI users, networks are divided into two types, geometric networks and network datasets. The geometric network is the method of modeling flow in closed systems such as water pipes or in electric wires. Network datasets are used in transportation applications, roads for

example. Both types of networks rely on rules of connectivity. For example a power station generates electricity and sends it through the lines to the cus-

Figure 03: Non-geographic view of the geometric network

tomer. To get to the customer it may first go through high voltage transmission lines down to a substation, where smaller distribution lines bring it to a meter, and then into your home. All this is modeled in the database and has a geographic component however, for some aspects of operations, the physical location is less important than the connectivity. Electrical engineers operating the system often prefer to see a non-geographic view of the system as shown above. This allows them to focus on the electric system and not be distracted with entities such as the streets and buildings.

Further use of this technology is for modeling. System engineers can run models to see the effects of their actions. They can redirect flow, open and close switches, and even look into plant operations. This is critical for planning purposes, outage management,

Figure 04: The GeoDataBase from the Arc Catalog

and operations. The best part of using ArcSchematics is that the data is all from the same source. This means no duplication of data, and more importantly, using the real operational datasets to perform the analysis. This is done with use of the GeoDataBase. The geographic features such as the streets, buildings are stored in the database along with the GIS data that makes up the system, such as the electric grid and all the supporting assets; poles, lines, generating plants,

substations. The schematic information is built on top of the geographic data and can be view back and forth from the map view, to a schematic view. If lines get added to the base GIS, they will be reflected in the schematic view immediately. Other critical functions such as networks, domains and relationship classes are also stored in the GeoDataBase making it a robust system of data storage, plus operations. Above is an example of the database as seen from the Arc Catalog application

In conclusion, ArcSchematics is an extension to ArcGIS desktop that allows the visualization and storage of schematic diagrams built off existing GIS data. It relies on the use of the GeoDataBase that reduces data redundancy and allows for complex data operations such as the creation of functioning networks and now schematic drawings, all from the same underlying GIS data. This system is scalable from an Access database up to very large systems using Oracle or SQL Server databases. The need for this functionality is directly suited to physical networks used in the utility industries, but can also be employed in logical and social networks in health, finance, law enforcement and more.

About the author:

Craig Clouet has worked at ESRI in the Hawaii office four years; before that he was doing GIS for land management for seven years. He holds two degrees in Geography specializing in GIS & GPS.
To contact Craig, email him at cclouet@esri.com.

Digital Imagery – what is it all about?

David Napier Paul,
Bussiness Manager,
NZ Aerial Mapping Limited.

The debate still rages over whether film has or will soon be replaced by digital technology. If so, is the digital technology better and why?

The discussion about digital imaging technology for aerial applications has been going on now for some 10 years with the first major introduction of the technology at the ISPRS Conference in 2000. Since then, digital camera technology has rapidly evolved and already third generation cameras are being operated world wide.

Aerial digital aerial cameras fall into one of two categories, frame or push broom. The frame camera creates an image that resembles the traditional film format whereas the push broom camera utilises linear array scanners, acquiring data by scanning the terrain with one or more linear arrays as the aircraft moves over the terrain, in much the same way as a broom sweeps a surface. We will limit our discussion in this paper to the frame camera or "area array" cameras and specifically the Microsoft Vexcel UltraCamX.

The UltraCamX is a 136 megapixel camera (14,430 x 9,420 pixels) and is characterised by its multi lens configuration made up of 4 x Panchromatic, 1 x Red, 1 x Green, 1 x Blue and 1 x Near Infra Red lenses. This provides an immediate advantage over film cameras in allowing simultaneous Pan, Colour and Infra-red image acquisition. There is now no need to fly the same area multiple times to capture your preferred output. Also, Infra-red film is temperature sensitive and notoriously difficult to work with in tropical climates. This problem is immediately negated with this digital sensor.

Each time the UltraCam is triggered to take a photo it records data on 13 CCD sensors ("tiles") that are strategically located behind the multiply lens. The recorded image is referred to as a Level-0 image. Nine tiles are used to record the panchromatic image. One of each of the remaining tiles is used to record the red, green, blue and near infra-red images respectively.

The transition from the raw Level-0 data to the final Level-3 colour images is by post-processing using UltraCam Office Processing Centre software (OPC). The steps of processing are; Level 0 to level 2 processing stitches the 9 panchromatic tiles into a single 16bit format image, and creates a single colour image from the red, green blue, infra red sensor
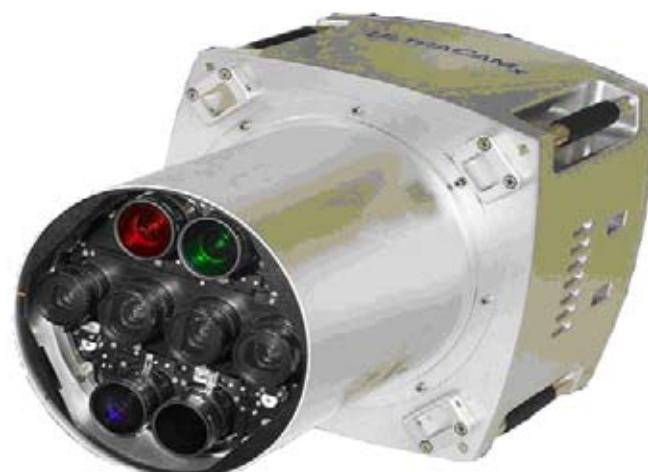
Figure 01: *The 136 megapixel UltraCamX camera*

data.

The final step is to output level 3 images – RGB colour, or InfraRed. This processing step allows the operator to adjust radiometry to get desirable colour, and to use dodging tools to correct for light fall off. Output is 8 or 16 bit format colour imagery, (or panchromatic if required). The high resolution panchromatic imagery is used to pan-sharpen the colour imagery to get the final full resolution colour images.

Once the complete Level-3 image is available, this represents the same type of file as a scanned film file would be. The advantage over film though is that you have the Pan, RGB and NIR images available. There is also a complete lack of "artefacts" on the imagery. The term artefacts relates to the dust, hair, fluff and scratches that invariably accumulate on film from the moment it comes out of the processor. These artefacts have to be laboriously "Photoshopped" out of the images to maintain a clear picture.

Furthermore, most digital cameras are paired with a Position and Orientation System (POS). POS data is used to calculate the exact position and tip, tilt, yaw and time of each image taken with the camera. The POS Inertial Measurement Unit (IMU) rotation rate and acceleration observations are combined with the POS Differential Global Positioning System (DGPS) observations in post flight processing. The combination of these two pieces of data allows a very accurate spatial location to be determined for each frame relative to mapping coordinate systems.

The most common product built from digital aerial photography is the orthophoto. Orthophoto production is the process of turning a normal perspective vertical aerial photograph into a map accurate true to scale image. Ortho images are widely used as base data layers within GIS.

The process is relatively simple yet can be a lengthy one depending upon how large the area is, what data is already available, what the terrain is like and what accuracies are required.

Typically orthophotography is specified on the resolution of the final product – GSD (Ground Sample Distance), the required accuracy of the imagery and the format and projection of the data.

The process involves taking the Level-3 images and using the POS data and an AT (Aerial Triangulation) software package and assembling a raft of images together so they are relatively connected. Once ground control is added, the absolute spatial connection to the local map coordinate system is confirmed. It should be noted here that there are a number of film systems also using POS technology.

Also noteworthy is the impact the POS technology has on the ground control requirement. It is accepted in practice that the ground control requirements can be as little 10% or less, than that used for an analogue system and DGPS alone.

Once AT has been completed, a DTM (Digital Terrain Model) must be used or created for the next phase. If existing DTM data is available, this can be imported into the processing software, checked against the stereo imagery for anomalies and corrected where necessary. If no DTM data is available, automatic processes can be used to auto correlate a DTM with human editing added to strengthen and correct the data where necessary.

With all the input datasets now on hand, the orthophoto generation process of "draping" the imagery over the DTM is run. This effectively places the pixels in their correct position on the earth so that the image is as it would appear if viewed vertically. As the original images do in fact have a slightly oblique look at their edges, sides of tall buildings and in some cases tall trees can be seen. This is accepted as normal and it must be remembered that we are interested in where the things are on the ground, rather than in space. In this case the building footprint would be correct as would the base of the tree.

Once a whole orthophoto has been built it is cut into smaller file size image tiles of a reasonable size for GIS work. The images are then copied on to delivery media. More recently this media has been a portable hard-drive (HDD) rather than multiple CDs or DVDs. HDD have the added benefit of making it quicker for the client to load their orthophotos on to their computer storage. The HDD can also be kept in a safe place to use as a back up if the operational dataset gets corrupted at all.

All this technology is available to you now. There are a number of initiatives currently being investigated to provide this in the Pacific Islands as early as this winter. The only requirement is for a definitive, confirmed project to be contracted and the rest will just happen. Be the first to have the very best in data.

Small Pacific Island Countries Makes their Submissions to the United Nations for Extended Seabed Territorities

Emily Artack,
Maritime Boundaries Project Officer,
SOPAC.

Through coordinated regional efforts, eight Pacific Island Countries lodged their submissions for about 1.8 million square kilometres of additional seabed territory before the May 13th deadline.

Over the last couple of years, SOPAC has successfully coordinated a sustained international and regional effort to ensure every SOPAC member country which has extended continental shelf (eCS) potential has managed to lodge a submission (or partial submission) which secures each eligible country's rights to claim additional seabed territory or eCS. Such claims are made under Article 76 of

Figure 01: The Pacific Islands Continential Shelf Map

the International Law of the Sea known as the United Nations Convention on Law of the Sea (UNCLOS).

The first regional workshop on the "Delineation of the Outer Limits of the Continental Shelf beyond 200 nautical miles and for Preparation of a Submission of a coastal State to the Commission on the Limits of the Continental Shelf" was held in Pacific Harbour in Fiji in February 2005. This workshop was conducted by the United Nations Division of Ocean Affairs & Law of the Sea (UNDOALOS) in collaboration with SOPAC and the Commonwealth Secretariat.

Later in February 2007 at the Pacific Islands Regional Consultation and High Level Meeting convened by

SOPAC in Brisbane, there were suggestions of the preparation of joint submissions for common areas of interest for some of the Pacific Island countries. This had been based on the practical example of the Joint Submission that was submitted by France, Ireland, Spain and the United Kingdom submitted in May 2006. Preliminary talks were initiated between Papua New Guinea (PNG), Solomon Islands and the Federated States of Micronesia (FSM) to consider the possibility of submitting a joint claim over the Ontong Java Plateau area in order to meet the 2009 deadline. This would leave unlimited time for negotiations regarding the equitable use and management of these areas.

Over the last 22 months, SOPAC has coordinated a series of 5 intensive regional training workshops. These workshops brought together country maritime boundaries teams and international and regional technical experts, where all could work together to develop joint and individual submission documents. Working from a SOPAC-commissioned Desktop Study (2006) which identified areas of potential eCS territory, each country has gradually developed their submissions towards the May 13th 2009 deadline.

The first sub-regional technical workshop was held at Geoscience Australia, Canberra in July 2007 for representatives from FSM, PNG and Solomon Islands in their efforts to prepare a joint submission for the Ontong Java Plateau, a vast area over 600 000 square kilometers.

There have since been four other technical multi-country training workshops and meetings coordinated by SOPAC and its technical partners, Geoscience Australia (GA) and the UNEP Shelf Programme. These technical workshops and SOPAC's ongoing eCS efforts have been largely funded by AusAID with a special AusAID Grant in 2008 and by the Pacific Governance Support Program (PGSP).

The Pacific Island Countries that had submitted their work to the United Nations in the past 2 months were Cook Islands, Fiji, Federated States of Micronesia, Papua New Guinea, Solomon Islands, Palau, Tonga, as well as Vanuatu and the Executive Summaries or claims of each of these countries are now publically available from the Secretariat of the United Nations Division on Law of the Sea. http://www.un.org/Depts/los/clcs_new/commission_submissions.htm

In spite of meeting the May 13th deadline, there is still much work to do since in some cases these claims are partial and must be completed over the coming months and years and all will have to be reviewed by the UN Commission on the Limits of the Continental

Shelf and each country will be required to defend their claim as these are reviewed by the Commission.

The eCS territory does not include the water column but rather these claims, if successful, give the countries the sovereign rights for exploring and exploiting mineral and other non-living resources from the seabed, subsoil and associated sedentary living organisms. This provides major poterntial benefits to countries with offshore resource potential, including oil, gas, minerals and biological resources.

The AusAID funded Pacific Islands Regional Maritime Boundaries Programme is managed by the Ocean & Islands Programme, SOPAC and has delivered its assistance through a coordinated partnership with Geoscience Australia, UNEP Shelf Programme, Commonwealth Secretariat, Geolimits, the National Oceanography Centre and GeoCap.

SOPAC and all of the technical partners would like to acknowledge the incredible dedication and hard work of all of the Pacific Island Country Maritime Boundaries teams and we look forward to continuing to work with these teams and seeing these claims through to a hopefully successful conclusion.

Using MySQL™ Master-Master Circular Replication for GIS/RS Distributed Data Applications

Sachindra Singh
ICT Systems Developer
Pacific Islands Applied Geoscience Commission
(SOPAC)

Abstract

Geographical Information Systems and Remote Sensing processes utilize a broad range of software applications, all of which involve the use of some combination of digital maps and geo-referenced data. These applications and tools are often used by geographically scattered teams of GIS/RS scientists and specialists who create, manage, analyze and display geospatial data on digital maps, which are acquired from diverse sources.

In order to track and manage various facets of the above processes, GIS/RS teams normally implement a set of customized databases. These databases are used by the teams for management of projects and tasks, and mostly handle progress (Job Tracking), time spent on particular tasks (Resource Management) and keeping track of digital sources and outputs (Asset Management) etc.

In the South Pacific region, most of these database systems are implemented in-house, and are built on top of the Microsoft Office Access platform. Microsoft Office Access is a Relational Database Management System (RDMS) which combines a database engine with a user interface and programming toolset (Visual Basic for Applications). Access enables GIS/RS scientists to quickly design and setup databases which are highly tuned to their requirements.

Problems occur when these databases are used in domains where the number of users increases, especially if they are geographically scattered. Access was designed in the early 90's, the pre-network era, and hence is mostly a "stand-alone" system. Access can be used over the Local Area Network (LAN) by file-sharing mechanisms, but since it was not designed as a multi-connection, multi-user system, data-loss and other bottlenecks are rampant in a LAN environment. Also due to the non-optimized, non-secure data transfer protocols used by Access, using it over Wide Area Network (WAN) or Internet is not practical.

Access can circumvent these issues by using an inbuilt functionality called Split Database Architecture. In this scenario the Access application is linked to another database engine (by Linked Tables), but the frontend (set of forms, queries, reports, macros, modules etc) is exposed natively in Access. Using an Access application in this manner is transparent to the user, with the backend foreign database engine seamlessly integrated with frontend forms, reports etc. Using a network-aware, enterprise-level database engine at the backend, the Access application can function smoothly over both the LAN and the internet.
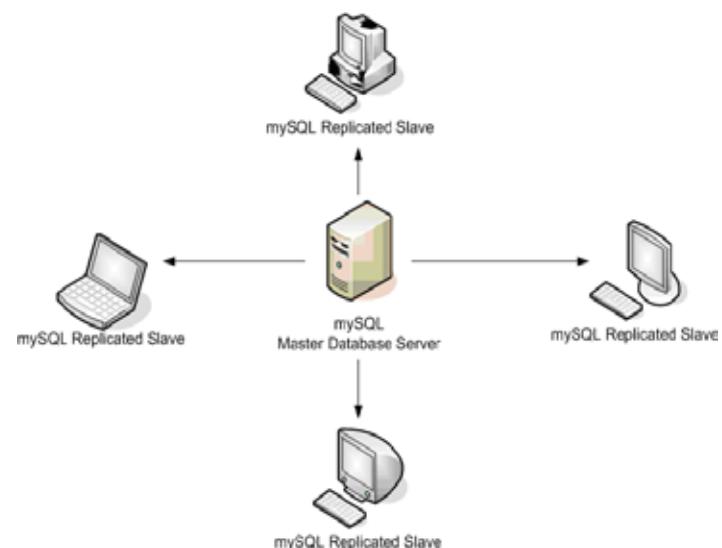
***Figure 01:** Master-Slave Replication.*

Various suitable backend database engines for integration with Access applications exist such as Microsoft SQL Server, Oracle, PostgreSQL, Firebird, mySQL etc. One of the most cost-effective, non-resource intensive, widely-deployed, highly-stable and easily maintained RDBMS is mySQL. mySQL is open-source (free to use) and is maintained by the Sun Microsystems (which has been recently acquired by another database company Oracle).

SOPAC GIS/RS team employs mySQL as the backend database engine for their Microsoft Office Access applications to enable them to be highly distributed across networks.

Replication
To enable backend database engines (such as mySQL) which are tied to Microsoft Access front-ends to communicate across networks (and over the internet for geographically scattered systems), a technique called Database Replication is employed. Replication is the process of seamlessly sharing data across different systems to enhance the data's accessibility, reliability and tolerance. Data replication is transparent to the user so the user does not need to be concerned about the source of the data, just that the data available is timely, current and shared with other users on the same networked system.

Typically, replication is done in the Master-Slave manner, where a mySQL database server (called the master) to be replicated to one or more mySQL database servers (slaves). Replication is asynchronous - your replication slaves do not need to be connected permanently to receive updates from the master, which means that updates can occur over long-distance connections and even temporary solutions such as a

dial-up service. This is ideal for databases that are distributed in the Pacific Island Countries which has irregular Internet services. Upon availability of Internet connection, all data is transparently synchronized between the database servers.

As seen above, Master-Slave Replication has significant disadvantages. Only data inserted or updated in the Master server gets replicated to Slave servers, not vice versa. That is, if any data is written to a Slave database server (node), it does not get updated or reflected in the primary Master database. In this scenario, the Master-Slave setup is nothing more than a glorified real-time backup system.

What is actually required, is a setup where regardless of which node (Database server) the data is written to, it gets automatically and safely replicated on all other connected nodes, without overwriting existing rows (data). This entails that all nodes are configured as both masters (dispersing updates) and slaves (accepting and updating updates). This particular setup is called Master-Master Circular Replication. In simple terms, multi-master replication is a method of database replication which allows data to be stored by a group of computers, and updated by any member of the group.

While Master-Master Circular Replication can be implemented in mySQL or PostgreSQL (Microsoft SQL Server does not allow this functionality), this process is neither officially documented nor supported by their respective vendors. Third-party plug-in developers have released software which adds Master-Master replication for the above database systems, but again, these are not officially sanctioned.

Rationale
This article aims to implement a Master-Master Circular Replication tied to standalone Microsoft front-ends that can function seamlessly over local area networks and the internet to facilitate real-time data reproduction and availability across geographical boundaries.

Prerequisites
Desktop-Class Computers running recent Microsoft Windows (or GNU/Linux) operating systems connected to the either a Local Area Network and/or the Internet.

- mySQL Database Server 5.1 or above
(http://dev.mysql.com/downloads/mysql/5.1.html)

- mySQL GUI Administration Tools
(http://dev.mysql.com/downloads/gui-tools/5.0.html)
- mySQL ODBC Connector for your platform
(http://dev.mysql.com/downloads/connector/odbc/5.1.html)

Figure 02: *Master-Master Circular Replication*

Microsoft Access 2003 or above
• mySQL 5.1 Documentation
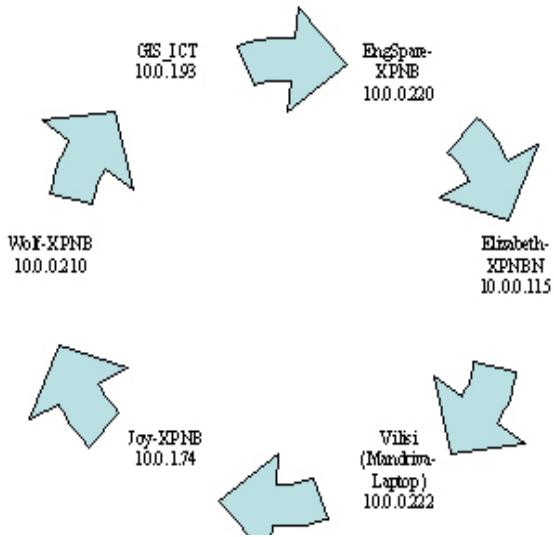
(http://dev.mysql.com/doc/refman/5.1/en/index.html)
Working knowledge of RDBMS systems, basic SQL statements etc.

Configuration.
Initial setups are planning out your deployment scenario, and then migrating data from Access applications to mySQL database engine.

1. Identify the individual systems and their network connectivity. Logically label each system that would be on the Master-Master setup with an ID, and note down their IP address and system name. eg: 1 – 192.168.0.10 (Computer1) , 2 – 192.168.0.12 (Computer2) etc.

2. Install mySQL Server on all target machines; ensure that remote root access is allowed during setup configuration. To simplify circular replication, ensure all root passwords across the target installations are the same.

3. Migrate data from your existing Microsoft Access applications to mySQL engines. For each Access application, create a corresponding database in mySQL, then migrate the data from Access to the created databases.

• Login to mySQL. On the command prompt type mysql –u root –p, press enter, and type in your root password.

• At the mySQL command prompt, enter create database <access_db_name> ;eg: create database gis_rs_jobs; and press enter. Create all databases that are required.

• Use mySQL Migration Toolkit (installed with mySQL GUI Administration Tools) to migrate data from source Access application to target mySQL database created in the steps above. Please ensure you use mySQL target hostname as localhost, and also use the correct root password you specified during setup. The toolkit wizard is quite user-friendly but should you need help, please refer to the in-built help documentation. Migrate all the Access databases to the databases created in mySQL.

• Repeat steps i to iv for all the target machines. You can use mySQL Administration tools and mySQL Migration toolkit (included in the mySQL GUI Administration setup) to connect to remote target machines via their noted IP address to create databases and migrate data from Access. Please ensure that in the initial steps, the data migrated to each individual target machines are exactly same. Only one (preferably the latest) version of the Access database should be used to migrate data to the target mySQL databases.

4. Once data from the front-end Access applications have been transferred to the backend mySQL database engines for each target machine, it is time to link back the tables in mySQL to Access fronted. This would enable your Access forms, modules, reports ,queries etc to function as with the tables/data in the mySQL backend.

• Install mySQL ODBC Connector on target machine.

• Make a backup of your original Access database application and then open your Access database.

• Under tables, select all tables and delete them.

• Under File menu in Access, select Get External Data, and Link Tables.

• In the Link File Dialog, select the File of Type as "ODBC Database" (scroll all the way down to the bottom of the drop-down list)

• Select "Machine Database Source" tab, click New button, Select "System Data Source" and click next.

• From the list of database drivers, select "mySQL ODBC 5.1 Driver", and click Finish.

• In the mySQL Connector dialog, enter your datasource name as anything, server as "localhost", username as "root", password as

your root password and select your migrated database from the drop-down list. Click Ok.

- From the list of tables, select all tables and link them to your Access application. After finishing, your forms, reports, queries etc should function smoothly using linked tables from mySQL Database engine.

Implementation.
After mySQL databases have been created on all target machines, data migrated from Microsoft Access to mySQL database engine and the data tables in Access replaced with linked tables from mySQL, it is now time to configure the actual replication of the target machines mySQL engine in a multi-master circular manner.

For each target machine, carry out the following steps:

1. Login to mySQL and stop slave threads.
mysql –u root –p (press enter, and then type in your password and press enter again)
mysql> stop slave thread;

2. Stop MySQL server after stopping the slave threads.
C:/>net stop mysql
(on Linux, /etc/init.d/mysql stop)

3. On each target installation, edit my.ini configuration file (found in the installation directory of MySQL Server, eg :C:\Program Files\MySQL\MySQL Server 5.1 in Windows or /etc/mysql/ in Linux) and specify the following under [mysqld] header :

server-id = ID
log-bin= mysql-bin
log-slave-updates
replicate-same-server-id   = 0
auto_increment_increment = 10
auto_increment_offset = 1
master-host = NEXT_IP
master-user = root
master-password = password
report-host = HOST

ID is the incremental number for target machines specified under Configuration Part 1 (see above). NEXT_IP is the IP address of the target machine this particular machine is connecting too, eg: For server Id 1, the NEXT_IP would be the IP Address for server whose Id is 2, and so on. The HOST is the name of the current target machine, eg: Computer 1.

Example:
For the setup above, the my.ini configuration files for each target machine would be something like the following.

GIS_ICT	ENGSPARE-XPNB
server-id = 1 log-bin = mysql-bin log-slave-updates replicate-same-server-id = 0 auto_increment_increment = 10 auto_increment_offset = 1 master-host = 10.0.0.220 master-user = root master-password = password report-host = gis_ict	server-id = 2 log-bin = mysql-bin log-slave-updates replicate-same-server-id = 0 auto_increment_increment = 10 auto_increment_offset = 2 master-host = 10.0.0.115 master-user = root master-password = password report-host = engspare-xpnb
ELIZABETH-XPNBN	**VILISI (Mandriva Laptop)**
server-id = 3 log-bin = mysql-bin log-slave-updates replicate-same-server-id = 0 auto_increment_increment = 10 auto_increment_offset = 3 master-host = 1.0.0.222 master-use = root master-password = password report-host= elizabeth-xpnb	server-id = 4 log-bin = mysql-bin log-slave-updates replicate-same-server-id = 0 auto_increment_increment = 10 auto_increment_offset = 4 master-host = 10.0.1.74 master-user = root master-password= password report-host= vilisi.pc.com

JOY-XPNB	WOLF-XPNB
server-id = 5 log-bin = mysql-bin log-slave-updates replicate-same-server-id = 0 auto_increment_incre-ment = 10 auto_increment_offset = 5 master-host = 1.0.0.210 master-user = root master-password = password report-host = joy-xpnb	server-id = 6 log-bin = mysql-bin log-slave-updates replicate-same-serv-er-id = 0 auto_increment_incre-ment = 10 auto_increment_offset = 6 master-host = 10.0.1.93 master-user= root master-password = password report-host = wolf-xpnb

4. Start MySQL server after stopping the slave threads.

```
C:/>net start mysql
(on Linux, /etc/init.d/mysql start)
```

5. Login to MySQL command-line console and execute start slave :

```
C:/mysql –u root –p
<enter root password>
mysql > start slave;
```

6. Master-Master Circular Replication should now be enabled. On any target machine, use your linked table Access application to insert or update an entry. On all other target machines, the Access application should reflect the changes.

Troubleshooting
In case replication fails on any target machine; follow the steps below until replication is working again.

1. Check network connectivity between target machines. Ping all machines from target machines to ensure that network connectivity is available.
2. Reset Slave threads on the target machines if network connectivity is available, but replication is still non-functional.
   ```
   mysql> reset slave;
   ```
3. To debug slave and master thread status, use the following commands respectively in mySQL Command prompt:

   ```
   mysql> show slave status;
   mysql> show master status;
   ```

Ensure that the status do not show any errors. If errors are found, refer to mySQL documentation on steps to resolve them. mySQL errors are beyond the scope of this article.

Summary
Database Management strategies such as Master-Master Circular Replication and Split Database Architecture can easily enable legacy Microsoft Access Applications implemented by GIS/RS teams to be functional over the network/internet. This is turn enables team members to better collaborate electronically and transparently regardless of their geographical location. The end-goals for any Information System is to enhance productivity and optimize resource utilization (whether human or systems), and the methodologies described in this article will hopefully make achieving them easier for GIS/RS teams in the Pacific.

Geographical Information Systems from an Open Source Perspective

Geographic Resources Analysis Support System (GRASS) The Complete GIS

Part One – Background

Sachindra Singh
ICT Systems Developer
Pacific Islands Applied Geoscience Commission

In what is, hopefully, the first part in a series of articles, I hope to explore the proficiency of Open Source applications in Geographical Information and Remote Sensing systems utilized around the globe; and how OSS GIS/RS systems can play a vital role in and around the South Pacific. This article explores the motivations and politics surrounding the birth of such an Open Source GIS system a little under three decades ago.

On January 20th of 1981, at 69 years of age, Ronald Reagan was inaugurated as the oldest president ever, till date, of United States. Age though, didn't hamper him from implementing policies and strategies that drove scientific innovation and research to even newer frontiers since John F. Kennedy's challenge to NASA, that eventually led to the Apollo missions landing on the moon over 10 years prior.

Most rationalize Reagan's keen interest in science, especially space science, to the fact that he was

not only a Hollywood actor in his hind years, but also a keen science fiction aficionado. Especially since some of his initiatives bordered more on the fantasy side than actual science, in comparison to the technology available then. Even though, compared with the controversies surrounding our latest science frontiers, such as the deterring of Stem Cell research by the last few U.S. administrations, Reagan is now seen as a visionary. This is also well due to the fact that he filled his Advisory Councils with astronauts, engineers, and, more prominently, science fiction writers, including Jerry Pournelle, Poul Anderson, Greg Bear, Robert Heinlein, Greg Benford, Dean Ing, Steve Barnes, and Jim Baen.

One of the boldest initiative brought about by Reagan was Strategic Defence Initiative (commonly referred to as the Star Wars programme) to use ground and space-based systems to protect the United States from attack by strategic nuclear ballistic missiles. Though never officially completed (and deemed a failure by many), SDI saw an unprecedented sum of funding redirected to satellite-based research and development in public, private and academic sectors. Most of the billions invested for SDI eventually found its way into non-military applications, such as, environmental impact assessment, disaster management, land resource planning, climate change observations, geographical surveys and countless others.

As the number of Satellite deployments increased, corporate and academic sectors saw the benefits of having access to the endless wealth of information that the "eyes in the skies" can provide regarding our planet. But when it came to analysing the huge amounts of data to extract requisite core information, bottle necks were encountered. While computer processing power and storage were exponentially growing in that ear, software applications were solely missing.

Prior to 1982, only one publically available GIS system existed, called Map Overlay and Statistical System (MOSS). A pioneer vector-based GIS system, it was developed by then U.S. Department of Interior. Nonetheless it was a very important milestone in the evolution of GIS systems. But since MOSS ran only on older, propriety main-frame computers and was a pain to easily and quickly program, there was a huge gap for a flexible, modern and powerful GIS system in the early 1980s.

Enter Star Wars. Reagan's Strategic Defence Initiative (SDI) was overseen by U.S. Department of Defence. Foreseeing the need for an all encompassing GIS software system, they commissioned their support programme, United States Army Corps of Engineers

(USACE) to design and develop the required system. USACE, in turn, assigned development to their Construction Engineering Research Laboratory (CERL) in Champaign, Illinois.

For the next 13 years, CERL-USACE kept on developing and releasing improved iterations of the GIS system, called GRASS (Geographic Resources Analysis Support System). GRASS initially focused on land management and environmental planning, but quickly evolved with functionalities for other areas of scientific and military research.

When under the Bush administration scientific research funds were significantly slashed, development halted at version 4.1, around 1995. After a brief hiatus, an academia group took over the source code and migrated the application from the obsolete Unix/VaX platform to modern POSIX compliant operating systems. The group continued development and added newer functionalities such as modern user interface and inter-operability with a wide variety of data formats, until version 5.0.

Version 5.0, in 1999, saw GRASS re-licensed as an open source system under the GNU General Public Licence. The decades of continued development has seen GRASS labelled as the complete GIS, and is widely recognized as the most flexible and powerful GIS processing application available. GRASS version 6.x is actively utilized by major governmental agencies, corporates and universities across the globe.

After 27 years, GRASS continues to see incremental development, and with nearly a hundred developers in the core team of Grass 6.x, it will remain a formidable force in the world of GIS applications, whilst also being totally free.

This article ends with a quote from the visionary world leader, who, however indirectly, brought about GIS software applications research and development into the mainstream:

"You on the cutting edge of technology have already made
yesterday's impossibilities, the commonplace realities of today" Ronald Reagan, 40th President of United States, in 1985
Part Two of this series will look into the installation, configuration and basic feature sets of GRASS version 6.4.

What a GeoPDF Does....
Vilisi Tokalauvere,
GIS&RS Officer,
SOPAC

"Just about anyone reading this is a GIS Pro so remember GeoPDFs are maps for masses, maps for non-mappers, maps for people who do not have big honking GIS/Mapping software on their "tricked" out PCs. It's like having a miniature GIS system inside of Adobe Reader." Ray Caputo, Geographer, US Army.

GeoPDFs are normal PDF documents, but with interactive maps - map layers can be switched on or off, users can zoom in or out, measure distances, view coordinates and attributes and change projections. They allow a fuller presentation of your GIS dataset to your audience. The end users do not need to have elaborate or expensive tools to view GIS datasets; they can simply use Adobe Reader.

To produce GeoPDF files, software called TerraGo Publisher for ArcGIS (formerly known as Map2PDF) from TerraGo Technologies is needed. The TerraGo software prices range from about $US3, 000 to $US36, 000 depending on the product(s). Acrobat Reader is needed to view the PDF, and in order to view the interactive elements in a GeoPDF, one must first install a free Acrobat Reader 1.9 MB plugin (GeoPDF Toolbar).

The GIS Section in SOPAC is planning to install TerraGo Publisher on one of its Desktop. GIS Users around the Pacific are welcome to use the software to convert their data in GeoPDF format.
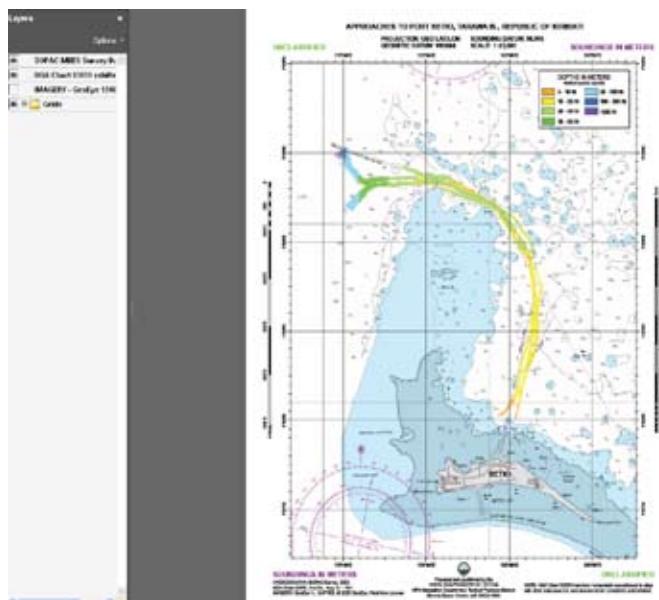

Figure 01: *Several Map layers in GeoPDF format*

The 2009 Pacific GIS&RS Conference

Dates: 1st-4thDecember, 2009
Venue: Marine Studies Complex, University of the South Pacific, Suva, Fiji
Theme: GIS&RS: The Pacific Way Forward

This year's conference is promised to be a bigger and better event for Pacific GIS Users. Unlike previous conferences, two days will be allocated for workshops during the week long event.

Workshop topics include "Softwares for Surveyors", "Handheld DGPS Mapping", and "GIS for Disaster Manage". Discussions and Presentations are scheduled for the remaining three days which will focus on GIS&RS applications and method development in the Pacific.

Presentation abstracts are open until 20th November, 2009.
Register for free on;
http://www.picisoc.org/PacGISRS2009.

For more information please contact:
Elizabeth Whippy
SOPAC Secretariat,
Private Mail Bag, GPO
Suva,
Fiji Islands
Phone: +679 3381377 ext 243
Fax: +679 3370040
Email: elizabeth@sopac.org