D1 INITIATIVE TITLE

The *EarthByte* software and database system

Initiative summary

Earth processes over geological timescales cannot be understood outside of a plate tectonic context. However, no standard tool exists to explore the causes and effects of lithosphere-mantle interaction in accordance with past plate configurations. Our aim is to develop a Palaeo-Geographic Information System called EarthByte that will connect the open source and architecture-independent GPlates and GMT software, and implement XML-based service interfaces and databases. EarthByte will create the foundation for an e-geoscience framework for grid-based data access and Earth process modelling by linking geological and geophysical observations to palaeogeographic models for constraining mantle convection and lithospheric deformation.

D2 AIMS, SIGNIFICANCE AND BACKGROUND

Background

Earth Science has a history of being data rich and information poor, a balance which is worsening year by year, especially due to a flood of remotely-sensed data. The problem we face now is how to amalgamate data and to connect them to analysis and process modelling tools. Mineral and petroleum resources typically form over time periods of hundreds of millions of years. In order to make digital Earth data sets suitable for resource exploration, one needs to be able to restore the geographic positions of all data back through geological time, and link them to process models. A "Plate Tectonic GIS" in which all data are attached to moving tectonic plates is a fundamental goal toward successful resource exploration and understanding Earth processes through geological time.

Even if all relevant data sets can be retrieved and merged, it is virtually impossible to analyse them in a plate tectonic context. However, both hydrocarbon and mineral resources have a clear association with the proximity to different plate boundaries through time, to hot mantle upwellings (plumes) or to downwellings associated with subduction and particular plate tectonic processes. In order to associate the likelihood of resource formation with a particular basin or geological terrain, we must be able to trace all relevant data through geological time.

Ultimately, it is not the data themselves that lead to success in understanding fundamental Earth processes and exploration, but rather the well-informed, far-sighted and inspired interpretation of the data in a plate tectonic context, which ultimately leads to the generation of new ideas and more accurate predictive models. Yet, no standardised, platform-independent, web-extensible software and data base system exists to achieve this aim. At the moment, we cannot take full advantage of the unprecedented amount of digital geodata now available through emerging electronic networks and the rapidly growing availability of open-source software. This severely limits the capacity of scientists, government agencies and companies to understand and predict the Earth processes that govern the sustainability and management of our mineral and energy resources.

The *EarthByte* initiative project is designed to fill this fundamental technological gap. *EarthByte* will start exploiting the enormous potential of the World Wide Web and grid computing to organise the research community and provide access to 4D computing and data resources to allow users to turn observational and computational Earth and ocean data into knowledge. *EarthByte* will facilitate collaborative multi-disciplinary research through 4D data synthesis through space and time that will help determine when and where tectonic plates break up, diverge, converge, collide and deform. *EarthByte* will aid in the search for natural resources in deeper parts of the Earth's crust in deeper fault systems and deeper basins. With an open-source platform, the universities, government and industry will—with their own computer connected to the internet—be able to harness the power of *EarthByte* to work on their geological problems.

In order for e-research infrastructure development to succeed in creating useful and long-lasting technologies with market success, on a rather limited budget, two things are essential:

(1) We need to tap strategically into ongoing software and database development that is showing signs of early adoption by the community, or has already been adopted, and

(2) We need to collaborate with key individuals and e-research centres that are driving this ongoing development.

Successful e-research infrastructure development is not a solitary activity, as the old-style "hero-code" development was, where an individual person spends years in a dark, isolated room to develop a brilliant piece of software. It rather depends on constant communication with key players across the globe to ensure that the infrastructures we create are compatible with standards developed by major e-research groups, especially in Asia, North America and Europe.

This pilot project is designed to build upon an international network of e-research centres and individuals to leverage common technology, databases and software, including the USA (SOEST (Hawaii), CHRONOS and Geoframework projects), the UK (NERC Data Grid), France (IFREMER), Germany (Alfred Wegener Institute for Polar Research), Norway (Norwegian Geological Survey), Japan (University of Tokyo/Ocean Research Institute) and China (Univ. of Geosciences).

Aims

Our aim is to develop a plate tectonic Geographic Information System (GIS) called the *The EarthByte software and database system*, which would revolutionise frontier resource exploration and Earth process modelling by creating an infrastructure for linking geological/geophysical observations and models to each other via palaeogeographic/plate kinematic models. Users would be able to seamlessly connect observations in distributed databases to palaeo-plate configurations and to visualisation/multimedia tools.

Specifically, the *EarthByte software and database system* aims to connect the open source and architecture-independent *GPlates* (Müller et al., 2005) and *GMT* software (Wessel and Smith, 1991) into a unified software system embedded in the APAC Geoscience grid infrastructure, and implement GML-based service interfaces and databases. *GPlates* is currently under development through an international consortium headed by the University of Sydney (www.gplates.org) and will represent a 4D database/software system, able to interactively reconstruct geodata spatially and temporally through geological history. *GMT*, the Generic Mapping Tools (http://gmt.soest.hawaii.edu), is a software package for the processing and visualisation of data, used by more than 10,000 users world-wide.

The *EarthByte system* will provide the enabling technology for a new resource exploration and Earth process modelling tool unlike anything that exists today. It will be a catalyst for novel research that was previously impossible, and the outcomes of this pilot project will demonstrate this through a number of test-bed projects. Our aim is to develop an *EarthByte Data and Computing Grid* at a later stage.

Significance

In the next 25 years, Australia faces the challenge to adapt to global environmental change, shifting economic landscapes, and diminishing fossil-energy and mineral resources. Despite the difficulties lying ahead, enormous opportunities for Australia lie with the resources contained in its continental crust, its vast unexplored ocean territories, and its continental shelves. To harvest these resources, exploration companies are increasingly forced to explore for both ore deposits and hydrocarbons in more remote onshore and offshore areas, and deeper in the Earth's crust where data are difficult and expensive to obtain. Soaring prices of non-renewable resources, especially oil and gas are calling for the development of smarter exploration tools.

EarthByte will deliver the basis for an integrated understanding of Earth processes from the heart of the continent to the continental shelves and the deep marine environment. It will aid in the search for natural resources in deeper parts of the Earth’s crust, in deeper fault systems and deeper basins. 4D data synthesis through space and time will help determine when and where mineral deposits and petroleum accumulations develop and under what tectonic settings. With an open-

source platform, anyone from universities, government or industry will be able to harness the power of *EarthByte* — with their own computer connected to the internet — to work on their geological problems, and contribute to the growth of its databases.

The *EarthByte software and database system* will be the foundation for the next generation e-science framework in the geosciences. We envision that *EarthByte* will play a significant role in resource exploration and our understanding of planetary-scale processes similar to that played by bio-informatics tools in DNA analysis and the development of new drugs.

D3 OUTLINE OF THE PROPOSED INITIATIVE

EarthByte Pilot Project initiative outline

A standard Geographic Information System (GIS), which stores data in present-day geo-referenced coordinates is sufficient for many environmental applications. However, most Earth processes over geological timescales cannot be understood outside of a plate tectonic context for which a "fixed coordinate" GIS is inadequate. In order to make digital Earth datasets suitable for resource exploration purposes, one needs to be able to restore the geographic positions of all data back through geological time, and link them to process models. This requires a system which consists of the following components: **(1) Processor** (linking tectonic plate-coded data to a global plate rotation model), **(2) 4D Visualisation, (3) GML/GPML-based input/output, (4) Database design, (5) Populating the database with quality controlled data, (6) GMT-based Mapmaking, (7) a web portal and (8) ancillary software linking *EarthByte* data to geodynamic modelling software** (Figure 1). The comprehensive linking of the *EarthByte* system to a variety of geodynamic modelling tools (8) is not part of the APAC Geoscience project or this proposal, with the exception of initial test beds.

Building on APAC Geosciences infrastructure

This project is complementary to the APAC Geoscience initiative focused on grid-infrastructure and services interfaces. In particular, *EarthByte* will benefit from the basic GPML information model currently being developed within the APAC Geoscience project (3), PostGIS database structure for GPML-formatted data (4), and web portal design (7) to allow users to submit queries and extract data. *EarthByte* will add several key software and database ingredients to create a Plate Tectonic GIS of immediate utility for a large community of end users both on the web (or APAC Grid) as well as on their individual computers without depending on web services (e.g, in the field, on a research cruise, or on a plane). For example, *EarthByte* will cover the inclusion of gridded data into GPML, such as digital elevation data models and potential field images.

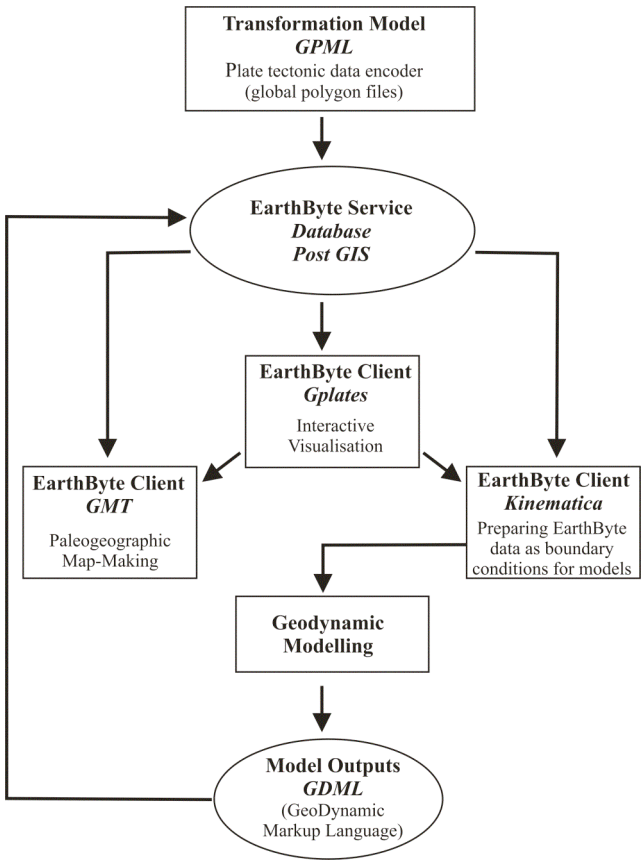

Figure 1: *Flowchart illustrating the components of the Earthbyte system and how each component is linked to the Earthbyte Database.*

Detailed outline of methodology

Processor and Visualisation

The *EarthByte* processor (1) is already completed, and visualisation (2) is partly completed for simple point and line-type data through software development by *GPlates* project members. We plan to develop visualisation of raster data within this project.

GML/GPML-based Input/Output

We propose to develop a comprehensive version of the GPML (GPlates Markup language) by applying the OpenGIS Consortium's (www.opengis.org) GML (Geography Markup Language) (www.opengis.org/docs/02-023r4.pdf) to the particularities of geophysical and geological data coded by tectonic plate and geological time, especially including gridded geophysical data. XML (eXtensible Markup Language) is the de-facto standard language framework for the serialisation of complex data for transmission between computer systems. The combination of a serialisation system and freely available domain-specific conceptual model (also known as the domain's "ontology") is the basis of all interoperability solutions.

We will utilise a concept-centric modelling approach instead of geometry-centric modelling as the next step in the evolution of database systems: it enables unprecedented flexibility in the processing, transformation and visualisation of data. This flexibility is essential if GPML is to form the transmission layer of the *The EarthByte software and database system* with GPlates at its centre. This effort will depend on a close collaboration between the APAC Geoscience project members, the GPlates consortium, the CHRONOS project, and the NERC data grid group.

Database

We will start compiling a quality-controlled GPML-based database. This necessitates the construction of a global set of plate boundaries that can be connected to form time-dependent closed polygons. The database will include a library of coastlines, faults, terrane boundaries, volcanic edifices, and hotspots, broken up by tectonic plate, and enabling the immediate derivation of plate reconstructions. The database will facilitate the generation and customisation of palaeographic reconstructions on all research/exploration scales. Key contributors to this database are the members of the GPlates consortium. Müller and Cervato will be co-chairing a US NSF-funded CHRONOS palaeogeography workshop just before the European Geosciences Union (EGU) Assembly in Vienna in April this year. This workshop will bring together key players in providing open data towards a CHRONOS/EarthByte database (http://www.chronos.org/community/paleogeography.html). This workshop will also help to compile an extended feature list of all data types to be included in a more fully developed GPML information model.

GMT-based mapmaking

Critical elements of the GMT software need to be turned from stand-alone programs into Application Program Interfaces (API's) so that the GMT mapmaking capability can be seamlessly connected to *EarthByte*. This "unbundling" of GMT is not a trivial undertaking, as one needs to carefully consider what the *EarthByte* community will use GMT component interfaces for, and to ensure that the interfaces are stable, as people will be building tools that use the interfaces. The interfaces become as important as the GMT file formats. In particular, the interface design needs to: (1) minimise the number of arguments to a function call (e.g., through combining them in structures or getting them from an environment), (2) ensure that it is self-checking as far as possible, and (3) easily allow for future, backwards-compatible expansion.

Prof. Paul Wessel, who will be spending a sabbatical at the University of Sydney from June 2005-June 2006, independently funded by a US NSF grant, will be driving this development, but the involvement of a local software developer and APAC's Geoscience Project professional programmers with their considerable experience in interface design are essential to achieve this aim. Prof. Wessel's interaction with the APAC Grid community during the e-research pilot phase represents a golden opportunity to tap into two decades of GMT-based expertise of open geoscience software development.

Gridded data and XML

We need to include gridded data in the widely used cross-platform "netCDF" format in our GPML-based data model. This requires a carefully designed GPML schema. It also requires overcoming another problem. When the GMT grid default format was originally designed, it was chosen to write grids as one long 1-D array, since GMT does its own 2-D internal bookkeeping. However, an unintended consequence of that choice is that other netCDF-aware tools (i.e. ferret, ncbrowse) will not recognise GMT grdfiles as 2-D grids as the 1-D encoding eliminates any reasonable possibility of sub-setting or striding on the data as it is pulled from netCDF. Users who must bridge two worlds must do time-consuming reformatting. There are two ways to address this shortcoming: (a) rewrite the subroutine responsible for writing netCDF grids so that future grid files are truly 2-D while ensuring that the read routine can handle both kinds of grids, or (b) offer 2-D netCDF grid input/output as an optional file format along with the dozen or so existing formats. Resolving this issue will involve close collaboration with A. Woolf from the NERC data grid, who is working on related problems (e.g. Climate System Markup Language – CSML).

EarthByte *and geodynamic modelling*

GPML forms the backbone of the ongoing work in coupling the GPlates plate reconstruction software and the CitComS mantle convection simulation software. CitComS is a component of the Geoframework (www.geoframework.org), a simulation architecture that enables the coupling of, and data sharing between, specific simulations. The coupling of GPlates and CitComS is a prototypical instance of the successful linkage of a software framework to a 4D Data Portal. Linkage requires the platform-independence and expressability of GPML and is enhanced by the web-friendly nature of mark-up languages. The expressability of GPML allowed the segregation of data according to the plate on which it is located, and also the capture of surface velocity grids over time, which are then used to constrain the CitComS geodynamic models. In this project we intend to explore test-beds for linking GPML-based data to numerical modelling software such as CitcomS and Snark/Underworld (developed by APAC/VPAC), focussed on subduction modelling.

Innovation and Significance

GIS-based data within a GIS-like framework—with the added capacity to reconstruct point, line, and raster (gridded) data in accordance with past plate configurations—is fundamental toward understanding the causes and effects of lithosphere-mantle interaction, associated with the practical aim of successful resource exploration. The *EarthByte software and database system* will provide novel IT protocols for storing, accessing and simulating data in a palaeogeographic framework, thus allowing the user to access distributed palaeogeographic databases on the fly, and link these data to process models. It will provide the enabling technology for a new generation of process models by allowing users to track observations through geological time and visualise reconstructed data on the screen and on printed maps, providing geodynamic model boundary conditions and/or to validate process simulations. *EarthByte* will lay the foundations for a virtual Earth exploratorium that will enable the user to follow any observations attached to geographic coordinates through geological time, from supercontinental assembly and breakup to the present-day Earth with widely dispersed land masses. Without the control imposed by the top surface —the Earth's plates—mantle convection models can never be used as a predictive tool to decipher the interaction of features such as plumes and superswells with processes such as continental uplift/subsidence or changes in relative plate motion. *EarthByte* will integrate all types of geodata in distributed databases with computational models within a single visualisation tool.

Addressing the Priority Area *Environmentally Sustainable Australia: Developing Deep Earth Resources*

Enormous opportunities for Australia lie with the resources contained in its continental crust, vast unexplored ocean territories, and continental shelves, extending from the tropics to Antarctica. Common to exploration and mining is the need to integrate disparate data sets at multiple scales and test a hypothesis through the creation of static geological models, forward simulations of process

models, inversion of simulations against tested models and measurement of uncertainty or confidence in the prediction. Examples of the data to be integrated include geophysical, geochemical and chronological data, drill hole data, rock core, and point sample analyses of the properties or rocks, fluids, and gases.

A successful search for natural resources, in deeper parts of the Earth's crust in deeper fault systems and deeper basins depends on understanding how tectonic convergence and break-up are accommodated and how, and under what conditions mineral deposits and petroleum accumulations develop (National Strategic Plan for the Geosciences, 2003). The exploration industry has realised that the search for deep Earth resources must be based on understanding how the Earth's mantle has evolved through time, stirred by subducting plates and mantle plumes. This is expressed in the National Research Priority area "Developing Deep Earth resources" and in multidisciplinary conferences such as the Exploration Geodynamics AGU Chapman Conference, held in Dunsborough, WA in 2001 and hosted by Müller, Moresi and Hobbs, where the rising importance of e-geoscience as an exploration tool was highlighted.

Australia derived 47% of it merchandise exports from its mineral and energy resources in 2001 constituting a significant component of the Australian economy. For future mineral exploration, it is critical to have ready access to a wide range of data, linked to process model outputs, in digital form and developed as national data sets. This project will directly address these goals by providing the following outcomes: 1) Exploring the links between tectonic activity and the transfer of material and energy in the Earth through time; 2) Improved understanding of the nature and rate of change in the Earth system, and the links between tectonics and the hydrosphere and atmosphere; 3) Development of geoinformatics and e-science required to model plate boundary processes; 4) Correlation of tectonic events mapped in crustal rocks with mantle processes through time. Break-throughs in constructing e-geoscience frameworks provides new opportunities for systemic understanding of Earth processes, linked to frontier exploration. The benefits will include reduced costs, and improved efficiency for mineral exploration, e.g. for hydrothermal and orogenic ore deposits related to subduction.

D4 COLLABORATION

Collaborating Partners

EarthByte will bring together research strengths from academia and government organisations across a wide range of national and international Earth science and information technology-based disciplines (Figure 2). The international *GPlates* Steering Committee (www.gplates.org) has jointly developed a global model over the past 15 years that defines the movement of the major plates on the Earth. The APAC (Australian Partnership for Advanced Computing) Geosciences project members are uniquely placed to develop the next generation e-research framework for the Geosciences in Australia. Both GPlates and APAC Geoscience efforts are coordinated by A/Prof. D. Müller. The additional research groups involved include the CHRONOS project (managed by Dr. C. Cervato, and closely linked to the US GEONGRID), CSIRO Exploration and Mining, the Tectonic Special Research Centre (TSRC) (headed by Prof. P. Cawood), the UK-based NERC Data Grid (represented by Dr. A. Woolf), IFREMER (Brest, France, represented by Dr. W. Roest, Director of its Marine Geosciences Department), the Alfred Wegener Institute for Polar Research, Germany (represented by Dr. K. Gohl), the Integrated Ocean Drilling Program (IODP, represented by M. Coffin) and the China University of Geosciences, Wuhan, (Prof. X. Xie), whose Geosphere Evolvement and Mineral Resource Laboratory is a National Key Project in China. By contributing to the development of an architecture of loosely coupled distributed databases and soft infrastructure, efforts can focus on problem solving utilizing 4D databases and some of the most advanced Earth process simulation and visualisation groups such as the Australian Computational Earth Science Simulator (ACcESS), and the GeoFramework (www.geoframework.org) based at the California Institute of Technology.

A proper "palaeo-GIS" requires scientifically sound plate reconstruction model parameters. The members of the *GPlates* Consortium are internationally renowned for their cutting-edge expertise in reconstructing ocean basins, reconstructions based on palaeomagnetic data, hotspot/mantle modelling, and palaeogeography. They have demonstrated their abilities to develop computer codes for plate tectonic reconstruction, modelling mantle convection, and deforming the lithosphere by contraction or extension. We intend to utilise this joint expertise in software development and associated research to turn *EarthByte* into an international, widely used standard, based on already accepted technologies/standards such as GML, XMML (developed in Australia by CSIRO) and the GMT map-making software by Prof. Paul Wessel and Dr. Walter Smith. Paul Wessel will be joining this effort through a US-National Science Foundation funded 1-year sabbatical at the University of Sydney from June 2005 to June 2006.

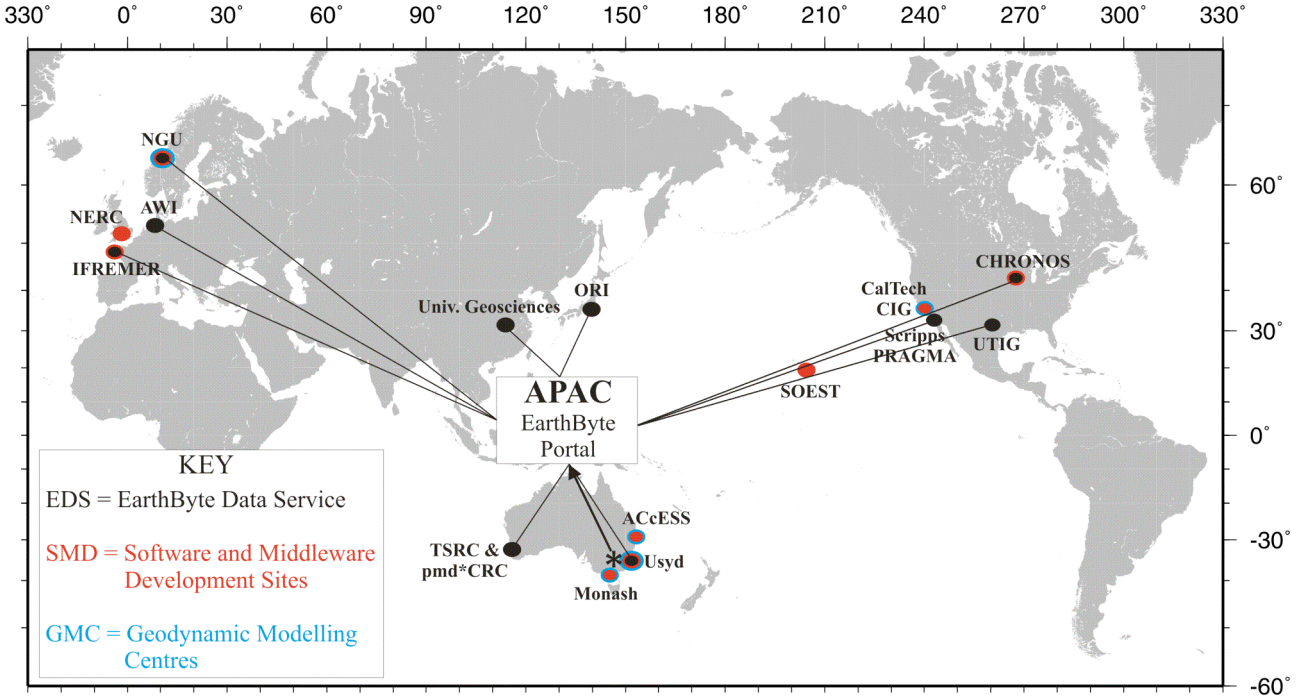

Figure 2: *The partner organisations and the type of service that each organisation will provide. ACcESS=Australian Computational Earth Science Simulator; AWI=Alfred Wegener Institute for Polar Research; CalTech CIG=California Institute of Technology, Computational Infrastructure for Geodynamics; CHRONOS=Interactive Network of Data and Tools for Earth System History; IFREMER=French Institute of Research and Exploitation of the Sea; Monash=Monash University; NERC=Natural Env. Research Council Data Grid; NGU=Norwegian Geol. Survey; ORI=Ocean Research Institute; pmd*CRC=Predictive Mineral Discovery Cooperative Research Centre; Scripps PRAGMA=Scripps, Pacific Rim Applications and Grid Middleware Assembly; SOEST=School of Ocean and Earth Science and Technology, University of Hawaii; TSRC=Tectonics Special Research Centre; Univ. Geosciences=Univ. of Geosciences, Wuhan, China; USyd=University of Sydney; UTIG=Institute of Geophysics, Univ. of Texas at Austin.*

Matching this initiative to the strategic plan of the University of Sydney

The initiative fits well into the University of Sydney's strategic plan, as it has promoted innovation in spatial science in a number of ways. The recent founding of the Spatial Science Innovation Unit (SSIU) best expresses this strength. SSIU is a GIS hub for spatial data and is aimed at building upon the University's intellectual capacity to combine spatial data with theoretical knowledge to get more information than contained in the data alone. SSIU has close links with VISLAB, and the University's Grid Computing group, and the University of Sydney Institute of Marine Science (USIMS), which has a major focus on developing cutting edge e-science technologies for simulating the history of ocean basins and building palaeogeographic databases.

Long-term alliances

All *EarthByte* project participants are connected by their interest in building research capacity in e-research in the geosciences. They have either assembled substantial databases, that need to be turned into palaeogeographic databases through *EarthByte* for use in either resource exploration or

fundamental research, or they have developed numerical codes for Earth process simulation which require *EarthByte* development so that models can be more easily ground-truthed through a diversity of geological and geophysical data connected to a plate tectonic model. The main two pillars of the proposed *EarthByte* pilot project are (1) The GPlates Consortium and (2) the APAC Geosciences Project. This pilot project will result in further strengthening the connection between both projects and the University of Sydney through the added infrastructure development support it would provide. Most *EarthByte* project members have a long history of collaborating, and this proposal reflects the collective goals and momentum of this international group of researchers. *EarthByte* will not only consolidate these links, but also bring the expertise of this group to large emerging e-research infrastructure projects in Australia, North America and Europe, to collectively build an e-geoscience network that is long-lived because it is built on open standards and open software, based on the consensus of the major players in this field globally.

D5 INFRASTRUCTURE

Even though an increasing number of regional and global digital Earth data sets are becoming available every year, planetary processes can only be unravelled by using these observations as direct constraints for computer simulations via a (palaeo-) geographic information system linked to a distributed software system for iterative inverse modelling of Earth processes. *EarthByte* will provide the enabling technology for this new generation of process models. We plan to create a Spatial Data Infrastructure (based on ISO and Open GIS interfaces and data encodings, developed in part through SEE Grid (www.seegrid.csiro.au) to link data sources into an open software computational framework (based on EScript and Pyre) developed by the Australian Computational Earth Systems Simulator (www.access.edu.au) and the GeoFramework (www.geoframework.org), respectively. This one-year pilot project is aimed at creating the core of an *EarthByte* software and database system that would allow users to seamlessly connect geological and geophysical observations to simulation software and visualisation tools in a palaeogeographic context.

This project builds upon the APAC Grid infrastructure, which draws on significant amounts of background intellectual property and related R&D occurring in the Australian geoscience research community. Significant portions of this community have been working together through the Solid Earth and Environmental Sciences Grid (SEE Grid), pmd*CRC and ACCESS MNRF. SEE Grid is of particular relevance since it represents a concerted effort by the community to form a geoscience Community of Practice for the development of open standards and interfaces for the provision of Grid and Web enabled services and applications. The SEE Grid community website (www.seegrid.csiro.au) hosts the community collaboration, standards, architectures, patterns and the ongoing development. The Geosciences APAC Grid project seeks to ensure that services useful to the Geoscience community are provided via the APAC Grid network and partners by ensuring they support the appropriate open standards and interfaces. Several APAC partners and proposed APAC Grid projects include components that are future building blocks or plug-ins for the *EarthByte software and database system* including: (1) Numerical codes designed for HPC and Grid computing – Snark, eScript, Finley, SPModel, Snac, (2) libraries for mesh generation (3) management of large scale jobs and workflow engines (4) federated data access, (5) large scale data storage, (6) visualisation, (7) grid middleware services – single sign on, security, remote access, and (8) cross-domain interoperability – Connecting some geoscience information with other Earth sciences (eg. marine) via service chaining functionality that converts compatible data on the fly.

D6 EXPECTED OUTCOMES

EarthByte will facilitate the development, implementation and advancement of currently lacking IT-infrastructure to extract a wealth of information from existing spatial and time series geodata. *EarthByte* is designed to enhance the Research and Development capabilities of all partners involved. A primary outcome of *EarthByte* will be a vastly improved capability to transform data into knowledge by coordinating protocols of linking geo-referenced information to the plate tectonic history of the Earth. *EarthByte* will foster data and software infrastructure development in

Australia that is in sync with sister projects especially in the USA and the UK (ie the CHRONOS, Geoframework and NERC Data Grid projects). This pilot project will contribute to the development of the APAC Grid to support *EarthByte* (ie an application driver). Wide-spread access to *EarthByte* and contributions to its development by researchers around the world, results in an enormous amount of leverage for the APAC Grid effort.

A test-bed outcome of this pilot project will be the integration of time-dependent meta-data sets with geodynamic modelling to allow a quantitative analysis of observations from the world's subduction zones. Global oceanic palaeo-age grids, a new global plate model based on moving hotspots as reference frame, palaeo-plate velocity grids, seismic tomography and geological data together with *CitcomS* geodynamic modelling software will be used to create a preliminary digital library for subduction containing multidimensional data and 4D model outputs. A second test-bed outcome will be the testing of the *NIMROD*/*OI* software for interactive inversion. Here a lithospheric extension model using *Snark* and/or *Ellipsis* software will be fine-tuned by a visual inspection of model results after each batch of model runs, in which the total parameter space is explored by random perturbations via genetic algorithms. After completion of the *EarthByte* pilot-project and test bed demonstrators, we intend to expand our software and database system towards an *EarthByte Grid* that will provide services to industry partners, government, University researchers and educators not only in Australia but internationally. The *EarthByte Grid* will open new markets for geodata, software and services. The benefits will include reduced costs, improved research and exploration efficiency, a broadened range of users, enhanced merging of data from different sources, and improved decision-making by industry and government.

D7 DESCRIPTION OF PERSONNEL

A/Prof. Müller will oversee the planning and accountability of the project, coordinating national and international collaborative infrastructure and data integration efforts. Müller, jointly with Moresi and Gurnis, will also help implement testbed projects for linking data to dynamic models.
Prof. P. Cawood (Director, TSRC) will coordinate the plate tectonic data integration effort with the Tectonic Special Research Centre (TSRC), especially with regard to palaeomagnetic data.
A/Prof. L. Moresi, Director of the Monash Cluster Computing (MC2) Facility will play a key role in test-beds for mantle convection and lithospheric deformation models constrained by *EarthByte*.
Prof. H. Mühlhaus, Director of the Comp. Geodynamics Group of the Australian Comp. Earth Science Simulator (AcCESS) MNRF (UQ), will coordinate *EarthByte* activities with AcCESS.
Dr. P. Rey (University of Sydney, Geosciences) will be testing the connection between *Ellipsis*-based numerical lithospheric extension and compression models and *EarthByte* data.
Prof. A. Zomaya, CISCO Systems Chair of Internetworking at the University of Sydney, will provide expertise in parallel and distributed computing, parallel algorithms, and cluster computing.
Dr. C. Cervato (Manager, CHRONOS project) will form the liaison between the *EarthByte* and CHRONOS projects, ensuring compatible and mutually complementary database development.
A/Prof. G. Clarke (HOS, Geosciences, Univ. of Sydney) will provide key crustal data from strike-slip and subduction plate boundaries, associated with 2D numerical test models (Ellipsis, Snark).
Prof. M. Coffin (Univ. of Tokyo Ocean Research Institute) will represent a liaison with the Integrated Ocean Drilling Program (IODP) and contribute a database for Large Igneous Provinces.
Dr. S. Cox (CSIRO EM) is a recognised international expert in information modelling and service interface design for the geosciences and will oversee the development of the GPML data format.
Dr. K. Gohl (Alfred-Wegener Institute for Polar Research – AWI) will form a liaison to link the extensive Pangaea sediment database as well as circum-Antarctic tectonic data into *EarthByte*.
Prof M. Gurnis is the leader of the GeoFramework project which aims to develop a suite of tools to model multi-scale deformation for Earth Science problems and will run test beds in this area.
Dr W. Roest (Director of Marine Geosciences Department at IFREMER) will contribute large scale oceanic data compilations and key software API's for grid rotations to *EarthByte*.
Prof. T. Torsvik (Nowegian Geol. Survey – NGU) is already number one (if you believe the inscription on his favourite T-shirt). Even though therefore he should not have to work any harder,

he will contribute travel funds, a global database of tectonic data and plate rotations, 20 years of software and database development expertise and the madness and good humour that is actually required to get this project done in the time frame given.
Prof. P. Wessel's (Univ. of Hawaii) open GMT software system is used globally by tens of thousands of geoscientists. He will spend a US-NSF-funded sabbatical at the University of Sydney from June 2005 to June 2006 to help integrate relevant parts of GMT into *EarthByte*.
Prof. X. Xie (China University of Geosciences), who recently spent a sabbatical at the University of Sydney, will contribute relevant data from Chinese sedimentary basins to *EarthByte*.
Dr. R. Woodcock (CSIRO EM) will provide the primary liaison between the *EarthByte* project and the SEE Grid community, ensuring the implementation of compatible open standards and interfaces development, especially information standards compliant with SEE Grid XMML.
Dr. A. Woolf (UK NERC Data Grid) will form a liaison to the NERC Data Grid community and ensure the implementation of a compatible, open and extensible infrastructure that allows data access services to be chained via plug-in data translation tools, especially for gridded data.

Technical staff and other Initiative personnel

Mr. Rob Atkinson (USYD) will provide expert capability to help design the detailed information model and structure of the *EarthByte* database in close collaboration with key partners such as the APAC, the GeoFramework and ACcESS MNRF.
Mr. James Boyden (USYD) will be further developing the maths back-end of *EarthByte*, ie the connection of a plate rotation model to the *EarthByte* database, including user interaction.
Prof. I. Dalziel (Univ. Texas Inst. Geophysics – UTIG) will form a liaison to the Plates project and help coordinate the merging of open databases and plate rotation models for the Paleozoic.
Dr. Maria Sdrolias (research fellow, USYD), Mr. Christian Heine (pgrad student, USYD), Mr. Scott Dyksterhuis (pgrad student, USYD), Mr. Stuart Clark (pgrad student, USYD). Lydia Taylor (pgrad student, USYD), and Mr. David May (Monash) will be involved in test-bed projects.
Dr. M. Smethurst and Dr. T. Redfield (NGU, Norway) will be involved in database population and the design of a global plate polygon data set.
Dr. H. Staudigel (Scripps Inst. of Oceanography), as Head of the Geochemcial Earth Reference Model (GERM) initiative and PRAGMA Grid member, will contribute a variety of GERM data.
Dr. S. Pisarevski (TSRC, UWA) will provide an updated TSRC Global Palaeomagnetic Database, global rotation parameters and will help resolve database architecture compatibility problems.
Dr. M. Takatsuka, as Director of VISLAB at the University of Sydney, will provide key expertise in Component-Oriented Software Engineering and Data Visualization for the *EarthByte* software.
Mr. Mark Turner (Caltech) will work on interfacing GPlates with geodynamic models such as CitcomS and SNAC *EarthByte*.
A Senior Scientific Programmer (USYD) will help develop the core functionality of the *EarthByte* system, including GPML schema development and embedding GMT-based API's into the *EarthByte* system for map making and visualisation.

D8 ADMINISTERING ORGANISATION CAPACITY

The University of Sydney has an extensive network of research infrastructure, high quality research capacity and a cohort of leading edge research teams providing a strong and innovative research capacity across a diverse range of e-research disciplines. Spatial science, tied with geophysics, advanced computing and visualisation, are internationally competitive areas of institutional strength at the University of Sydney. Related research activities include the ARCO Geophysical Imaging Laboratory/CRC Mining, and simulation and visualization, lead by VISLAB. Earth Science and spatial science is primarily coordinated by the School of Geosciences the Spatial Science Innovation Unit (SSIU) and the University of Sydney Institute of Marine Science (USIMS). A total of over 30 full-time academic staff are involved in spatial science, engineering and modelling/visualisation at the University of Sydney, which collectively have strong links to an enormous range of universities, government institutions and industry with international research

strengths. The University's long-term commitment to supporting Earth and spatial science, as well as simulation and visualisation, puts it in an excellent position for supporting e-research.

REFERENCES

Müller, R.D. et al., 2005. GPlates: Open source software and data base for plate reconstructions, European Geosciences Union General Assembly, Vienna, pp. Session SSP23.

Wessel, P. & Smith, W., 1991. Free software helps map and display data. EOS Trans. AGU, 72: 445-446.

For additional references see one-page CV's of participants.