

SpaceGRID and Geant4

Hugh Evans

*Space Environments and Effects Analysis Section,
TOS-EMA*

What is a "GRID"

- The term "Grid" was coined to denote a distributed computing infrastructure for advanced science and engineering.
- It is principally concerned with coordinated resource (data, computing resources, experimental hardware, etc) sharing and problem solving in dynamic and multi-institutional virtual organisations.
- It is a means of sharing computational resources over a high speed network with localised control of the resources and remote user proxy authentication. This provides flexibility of control and access policies to the service provider and ease of access to the remote user.

GRIDS: wide-area computing

- Grids implement distributed task scheduling and execution.
- Grids implement distributed data:
 - Storage
 - Access
 - Replication
 - Management
- Grids provide access to remote hardware.
- Grids facilitate authentication, authorisation, and accounting across national (continental/institutional) boundaries.
- Grids give you potential access to 1000's of computers, but institutes can set their own priorities for their contribution: institutes“own” some of the resources.

What does the GRID do for the user?

- You submit your work and the GRID:
 - Finds convenient (CPU “near” data) places for it to be run;
 - Organises efficient access to your data (caching, migration, replication);
 - Deals with authentication to the different sites that you will be using.
 - Interfaces with the local site resource allocation mechanisms, policies;
 - Runs your jobs;
 - Monitors progress and recovers from problems;
 - Tells you when your work is complete.
- If the task allows, GRID can also decompose your work into convenient execution units based on the available resources and data distribution.

SpaceGRID

A European Space Agency GSP funded project with:

- a study component aimed at analysing the issues (technology, infrastructure, suitable applications) and propose a roadmap for the seamless access to, and exploitation of, distributed data, application and resources in a network of heterogeneous computing resources
- a prototype component aimed at implementation of small-scale dedicated test-beds based on existing infrastructure

SpaceGRID Domains

- Four space related disciplines have been included in the study:
 - Earth Observation (Datamat)
 - Space Environment
 - Space Weather Simulation (CS-SI)
 - Spacecraft Plasma interactions (QinetiQ, SSL)
 - Radiation Transport Simulation (QinetiQ, SSL)
 - Solar System Research (RAL, SSL)
 - Spacecraft Engineering (Alcatel)

SpaceGRID Schedule

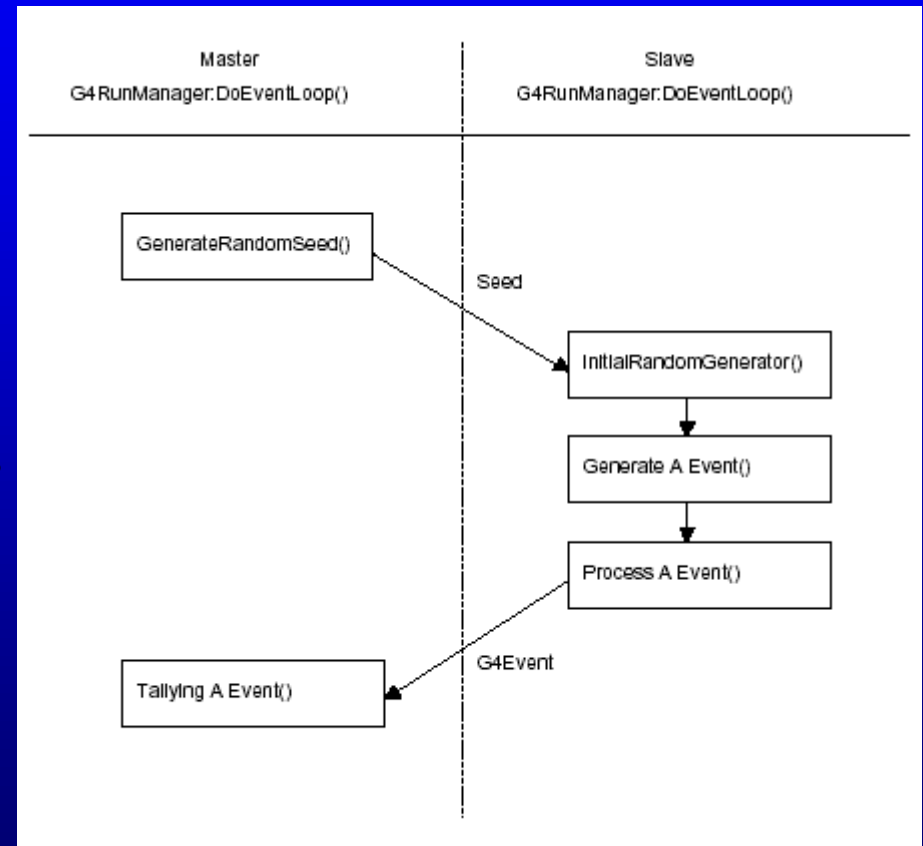
- Project duration of 18 months
- Project kicked off on Sept. 26th, 2001
- User requirements assessed and applications selected by March 2002
- Prototypes will be ready for delivery by March 2003.

Geant4 Prototype Implementation Plan

- The Mulassis application was the baseline for the prototyping a Geant 4 GRID application.
- A primary goal was to minimize the changes to the Geant-4 Kernel code.
- The prototype was to handle resource utilization and user control of the execution and GRID resources used.
- An Event level job farming model was to be used for distribution of the simulation across the GRID.

Geant4 Implementation

- Two types of nodes are used, a Master and slave.
- **The Master node:**
 - Locates available GRID computers
 - Starts the slave processes on these machines;
 - Creates a single G4RunManager which:
 - Farms out the events to the remote processors
 - Provides a random number seed
 - “harvests” the event output data and performs final data processing;
 - Handles the job control, node dropout, performance monitoring and tuning.
- **The Slave node:**
 - Receives the event information and random number seed and performs the particle transport calculation.
 - Sends the results of the particle transport back to the Master node.



SpaceGRID-Geant4 Prototype Infrastructure

- Linux based computers
- The TOP-C compiler: parallel C/C++ compiler
- Globus GRID toolkit: GRID middleware and job control/authentication/resource management software.
- AMPIC: for the inter-process communication (MPI-CH)

Further Information

- SpaceGRID web site: <http://www.spacegrid.org>
- Globus toolkit: <http://www.globus.org>
- EU-DataGRID: <http://eu-datagrid.web.cern.ch/>
- ESA GRID: <http://esagrid.esa.int/>
- Global Grid Forum: <http://www.gridforum.org/>
- TOP-C: <http://www.ccs.neu.edu/home/gene/topc.html>
- AMPIC: <http://gcl.ucsd.edu/ampic/>

Questions?

