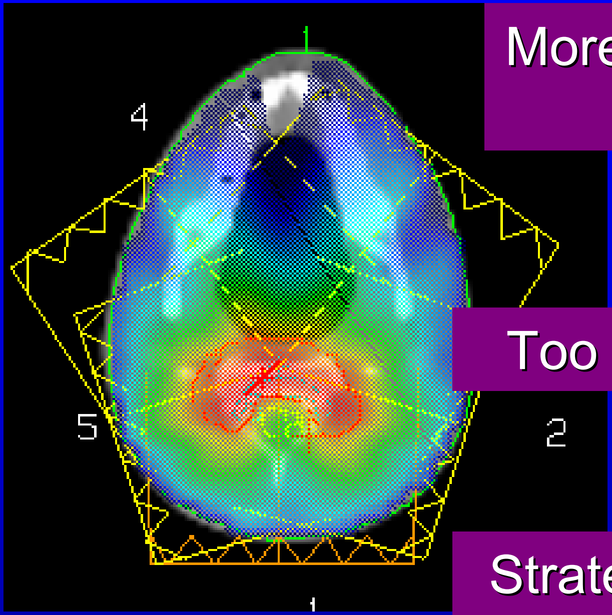
IEEE NSS 2007
Honolulu, HI, USA
27 October – 3 November 2007

Benchmarks of medical dosimetry simulation on the grid

S. Chauvie[1,6], A. Lechner[4], P. Mendez Lorenzo[5], J. Moscicki[5], M.G. Pia[6],G.A.P. Cirrone[2], G. Cuttone[2], F. Di Rosa[2], F. Foppiano[3], M. Piergentili[6], G. Russo[2]

[1] Santa Croce and Carle Hospital, Cuneo, Italy , [2]INFN Laboratori Nazionali del Sud, Italy, [3]IST, Genova, Italy, [4]Technical University Vienna, Austria, [5]CERN, Geneva, Switzerland, [6]INFN Genova, Italy

Monte Carlo methods in radiation oncology



More precise than approximated analytical methods used in commercial treatment planning systems

BUT

Too slow to be realistically usable in the clinical practice

HENCE?

Strategies to reduce the overall simulation execution time

- Variance reduction techniques *(event biasing)*
- Inverse Monte Carlo methods
- Analytical transport methods
- "Fast simulation" techniques
- Parallelisation

The GRID?

This project

- Let’s take a few representative dosimetry simulation examples
- Let’s run them on the grid
- Let’s evaluate (quantitatively) gains and drawbacks in realistic situations

- Three Geant4 Advanced examples of dosimetry simulation covering major techniques in radiotherapy
  - Brachytherapy
  - Hadrontherapy
  - IMRT (Intensity Modulated RadioTherapy)

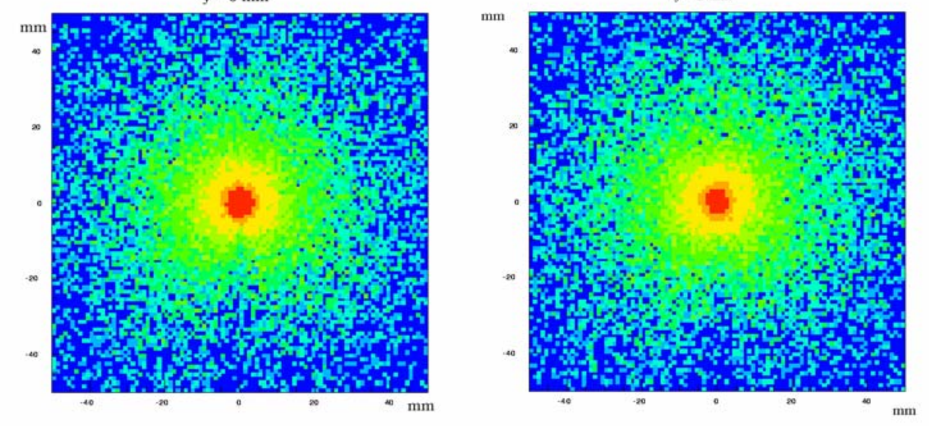
- LCG infrastructure, Geant4 Virtual Organization
  - The exercise will be repeated one step further as Mr. Nobody

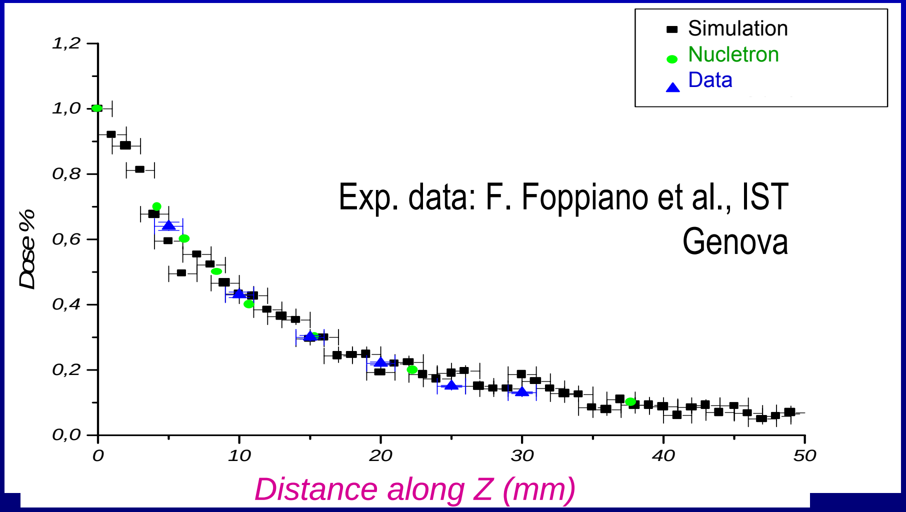
brachytherapy

Italian National
Institute for
Cancer Research

Geant4 brachytherapy Advanced Example

MicroSelectron-HDR source

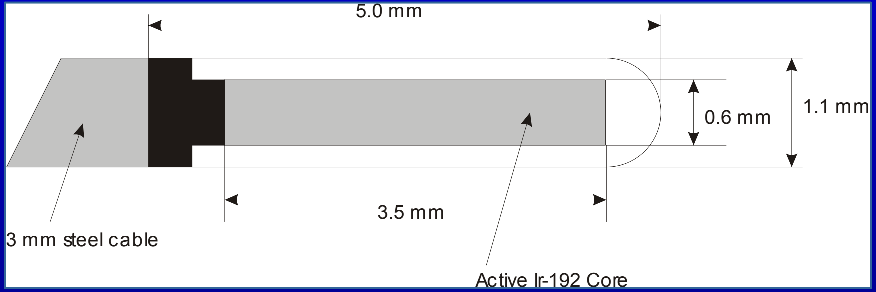




Radiation Source: hundreds keV photons

Exp set-up: simple geometry

Target: homogenous “tissue and tumor”



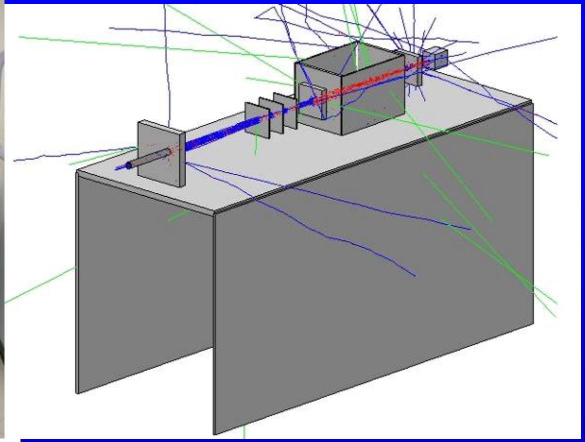
Relatively “fast” simulation

~ 7 CPU hours on an “average” PC to produce meaningful statistics for clinical studies

hadrontherapy

Geant4 hadrontherapy Advanced Example

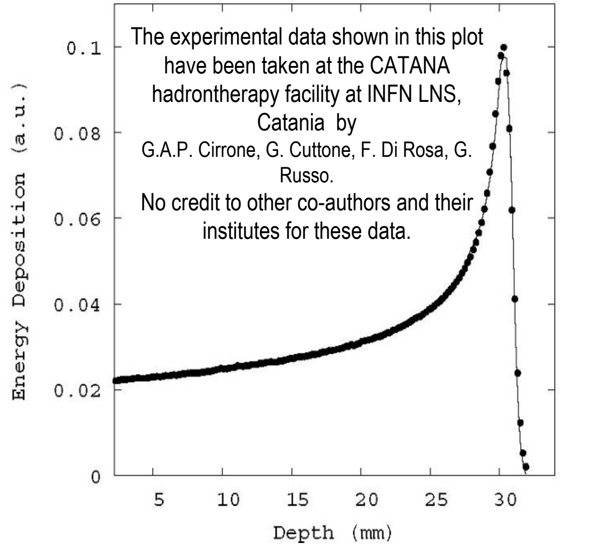
INFN LNS,
Catania, Sicily,
Italy



Radiation Source: tens of MeV protons

Exp set-up: accurate reproduction of the CATANA hadrontherapy facility

Target: homogenous "tissue and tumor"



G. A. P. Cirrone et al., "Implementation of a New Monte Carlo GEANT4 Simulation Tool for the Development of a Proton Therapy Beam Line and Verification of the Related Dose Distributions", *IEEE Trans. Nucl. Sci.*, vol. 52, no. 1, pp. 262-265, Feb. 2005.

Production aimed at validating physics modelling of the simulation

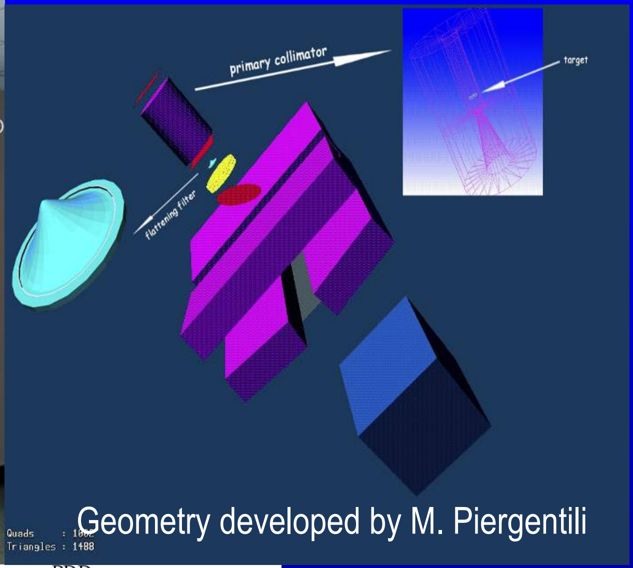
~ 150 CPU hours on an "average" PC to produce meaningful physics results

~ 16 hours for calibration results

medical LINAC for IMRT

Italian National Institute for Cancer Research

Geant4 medical_linac Advanced Example

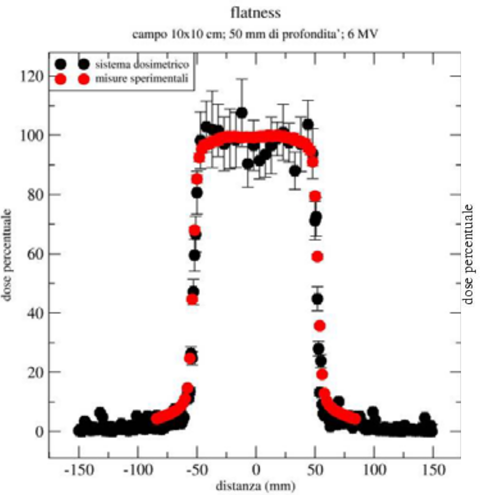


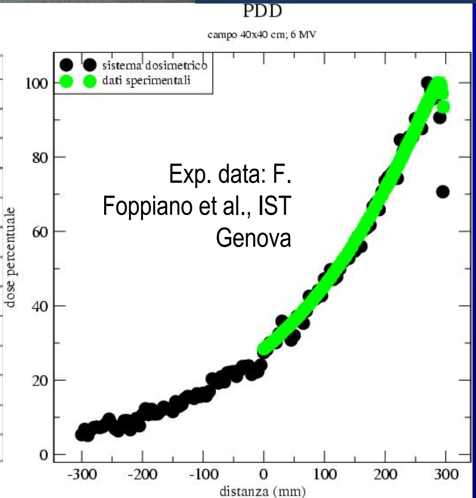
Geometry developed by M. Piergentili

<u>Radiation Source</u>: order of MeV electrons

<u>Exp set-up</u>: accurate reproduction of accelerator's head

<u>Target</u>: non homogenous soft tissue, air, bones and tumor





High demand of CPU resources for meaningful statistics

(e.g. for treatment planning verification)

≈ tens of CPU-days

Requirements

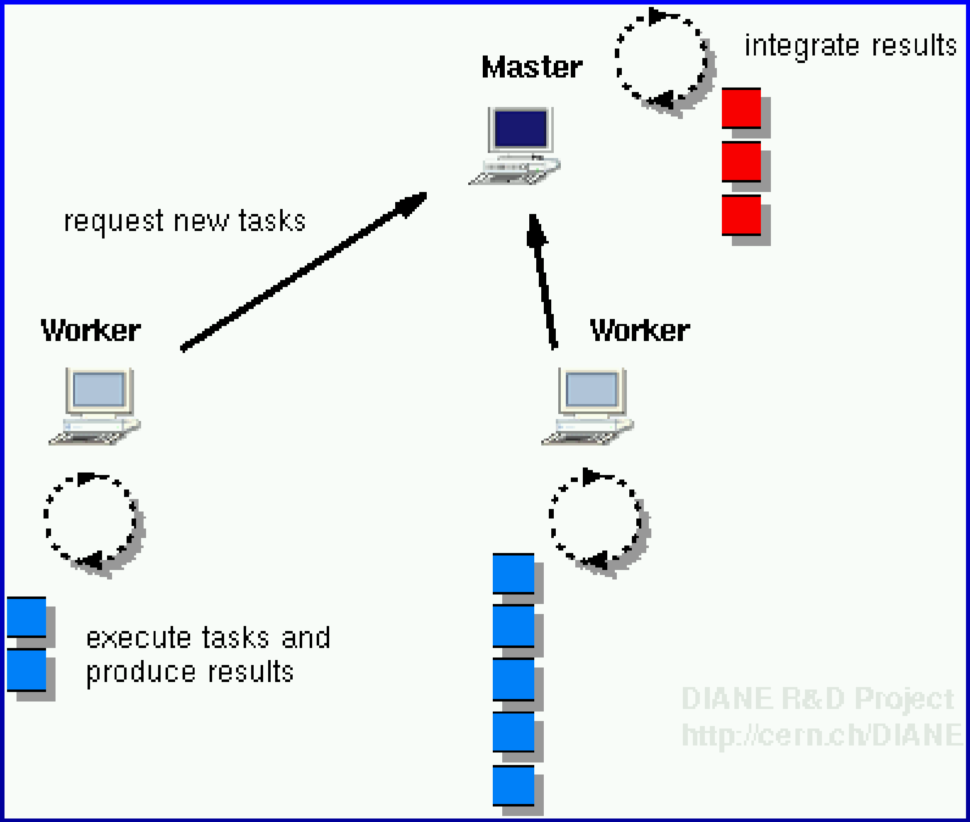
- Submit jobs to a Grid system
- Plug-in Geant4 application
- Split into tasks (without biasing the results)
- Schedule tasks intelligently

Geant4 + AIDA iAIDA + Ganga

+ DIANE + LCG

- Hide the underlying complexity to the end user

DIANE…



a workflow manager for distributed Master-Worker applications

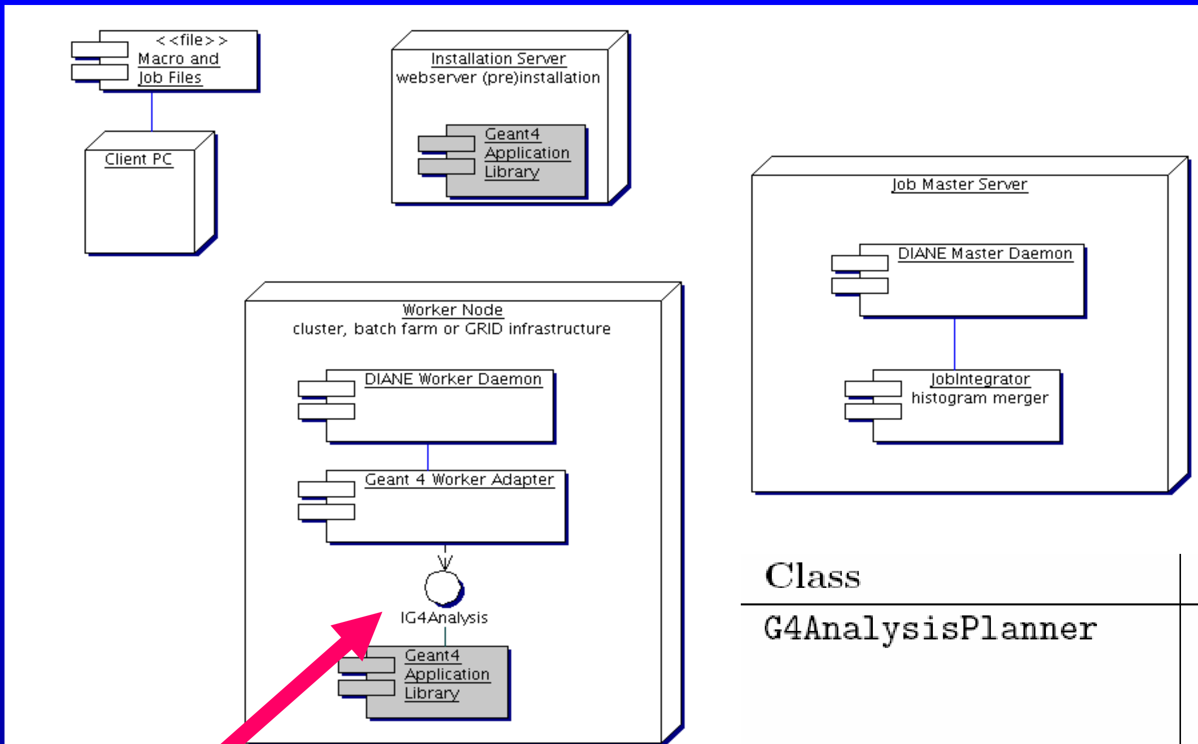
Master-worker model
- client starts a job
- workers perform tasks and produce results
- master integrates results

Distributed Processing

task = N events
job = M tasks

tasks may be executed in parallel
tasks produce results
task output is automatically combined

… DIANE with geant4



DIANE Geant4 Interface

Interface class which binds together Geant4 application and DIANE framework

Class	Tasks
`G4AnalysisPlanner`	splits total simulation into tasks according to the specifications in the DIANE job file; assigns a random seed to each task to be executed
`G4AnalysisWorker`	initialises the application according to the specifications in the job file, loads the application module runs the application, and executes the termination procedure (all performed on the worker node)
`G4AnalysisIntegrator`	integrates simulation output and merges the results if required

Computing environment

- Identification of resources
  - The site must support the Virtual Organization (Geant4)
  - Shared file system accessible through VO_GEANT4_SW_DIR environment variable
  - Operating system and architecture compatible with Geant4

- Software installation
  - Geant4, CLHEP, AIDA, gcc
  - Problems encountered: missing write permission, user authentication etc.
  - Requires familiarity with grid middleware

Requires experienced user

One must do one's own installation on the grid

Interface to grid middleware *(e.g. lxplus)* must be set-up, otherwise additional (not trivial) installation needed *(e.g. gLite)*

Problem set-up

- Determination of the total # of events to be produced
  – Defined for each application on the basis of the desired statistical precision of dosimetric variables relevant to each use case
- Splitting into tasks ← **Critical issue**
  – Task = smaller group of events to be produced
- Perform tasks sequentially on a worker node
- Use a range of worker nodes in parallel in a grid environment

Task splitting

Task size (execution time) ⟷ (balance) # workers involved

- Important for applications with short execution time
  - Time elapsed between submission of grid jobs and beginning execution on a worker node comparable to total duration
- High granularity → more effective use of CPU resources
  - Drawbacks: in general, larger output size to deal with

Splitting optimization

“One size DOES NOT fit all”

“Fast” use case brachytherapy

- Goal: fastest simulation results
- Many small tasks
  - 20000 events, ~ 25 s each
- 40 worker nodes
- Higher # nodes would not be more effective
  - Time for last registered workers to become active might be higher than total duration

“Intensive” use case hadrontherapy

- Goal: sensitive simulation study
  - beam energy calibration
- 1 task = 1 worker
  - Not the fastest option
- Compromise with large disk space requirements for output
- Simplify distribution of random number seeds passed to tasks

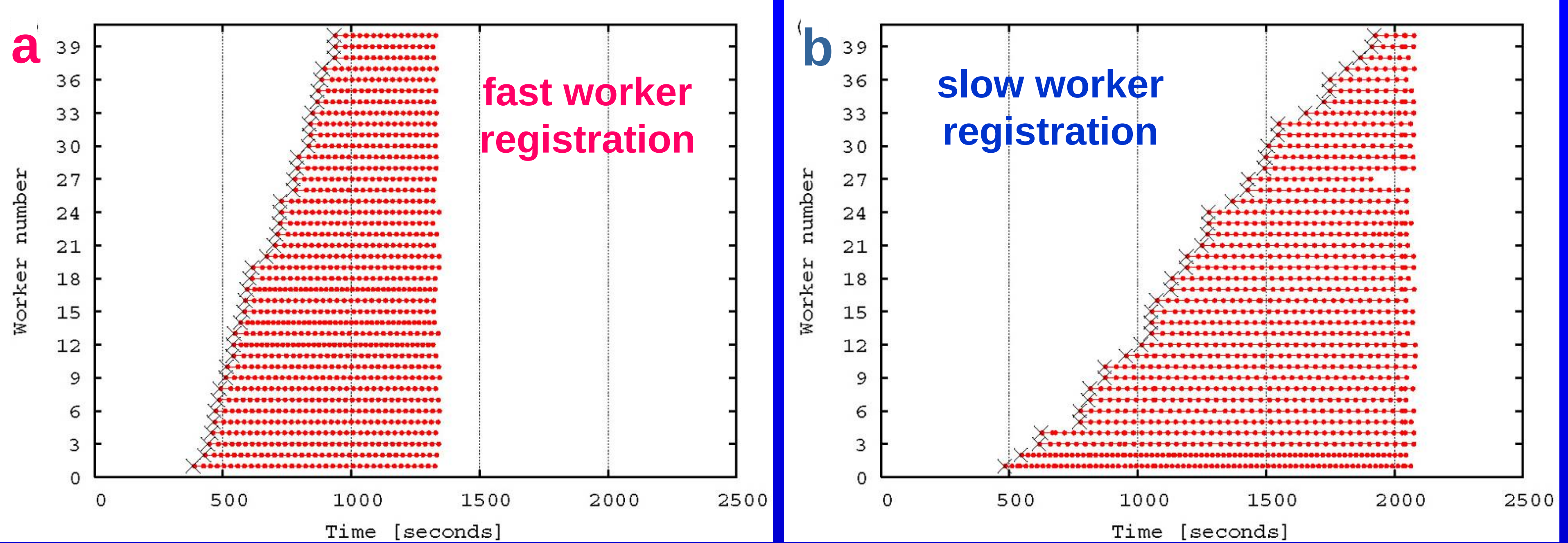
GRID Baby-sitting

- Selection of resources
  – Availability, priority
- Set-up a list of Computing Elements
  – Passed to Ganga through DIANE
- Criteria for selection
  – State of Computing Element (production, draining mode)
  – Number of CPU allowed to VO at a given site
  – Estimated waiting time in batch queue
  – Estimated additional workload on given nodes
  – CPU speed

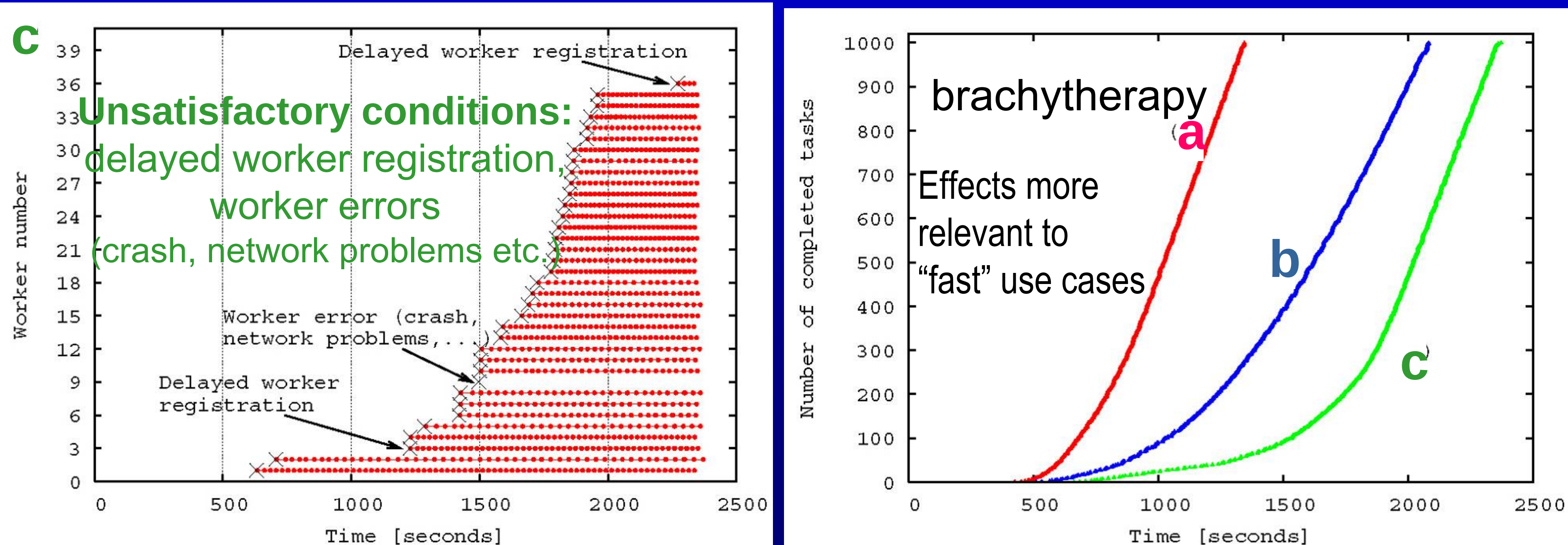
Must be done on a regular basis

Automated tools not always effective

- Relatively easy retrieval of information
  – Used for selection criteria



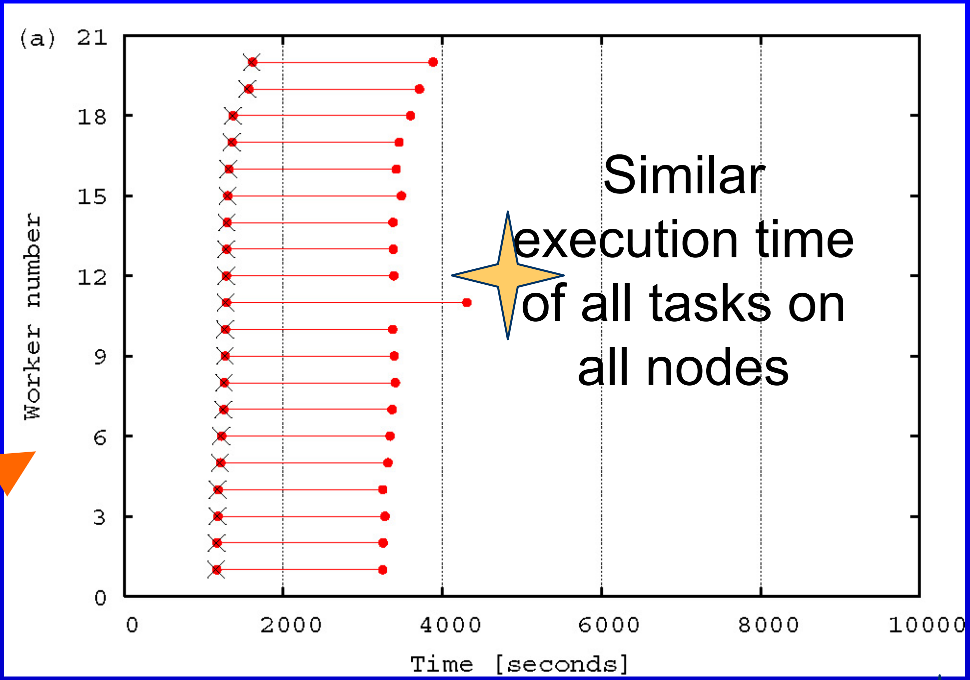
Effects of delayed worker registration

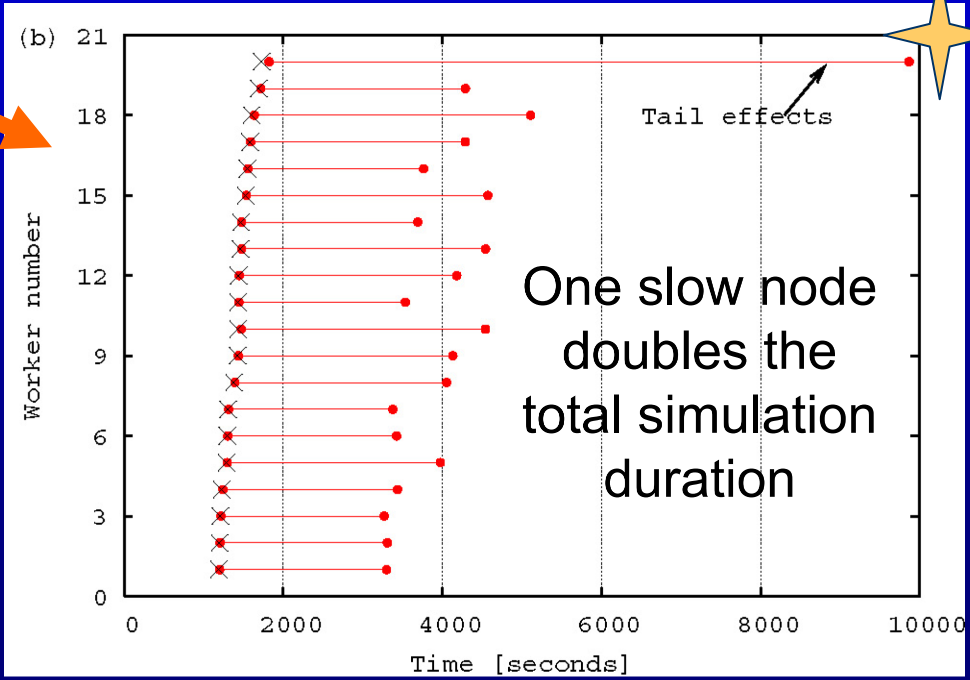


Effects of slow worker nodes

More visible when few worker nodes

Both simulations executed in the same environment in identical Computing Elements





Possible explanation: variable workload on the same nodes

These are the overall simulation time!

"Fast" use case: brachytherapy

Period of testing	3weeks
Number of runs performed	50
Number of events simulated (per run)	$2 \cdot 10^7$
Number of DIANE workers applied (per run)	40
Number of tasks (per run)	10^3
Number of events per task	$2 \cdot 10^4$

40 DIANE workers

Sequential

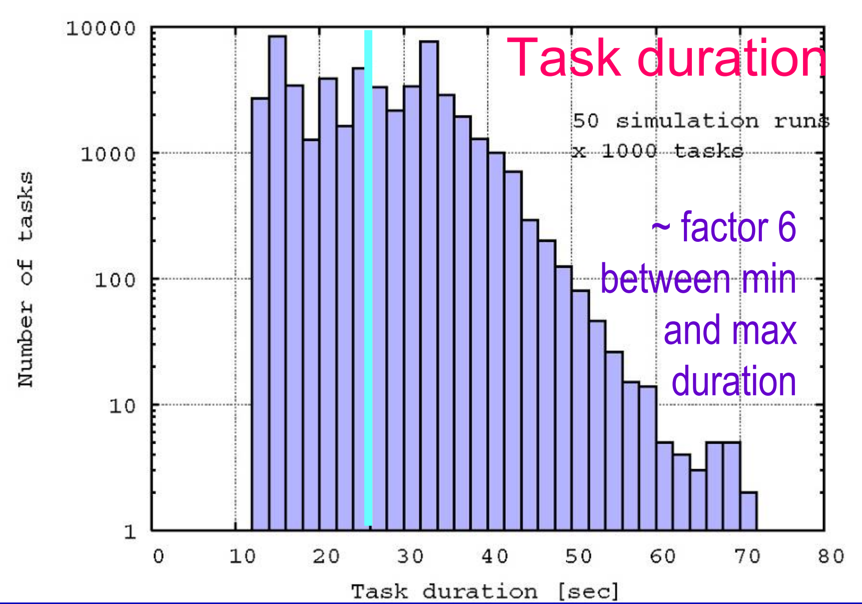
1 task = 25 ± 0.5 CPU s

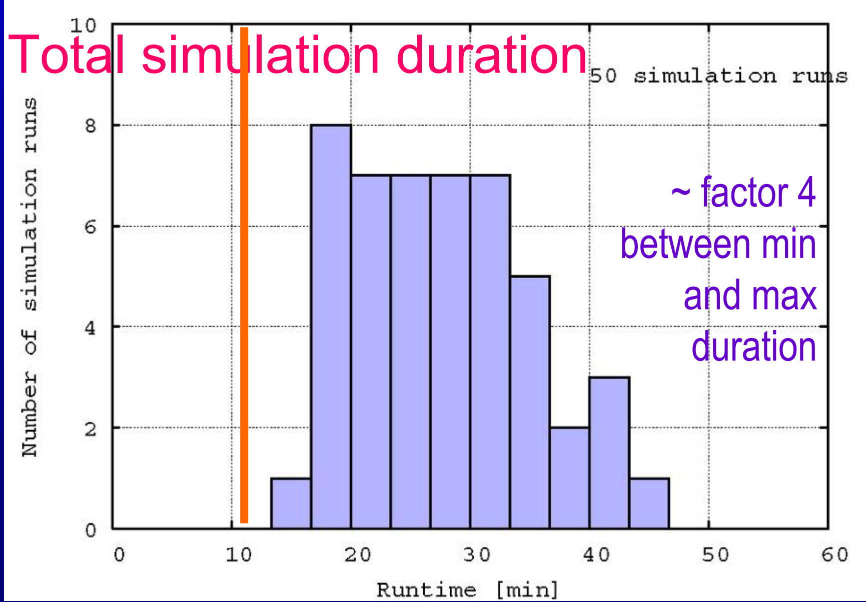
Total simulation = 417 ± 8 min on lcgui003

Ideal expectation with 40 workers:
~10 min for the whole simulation

On the grid

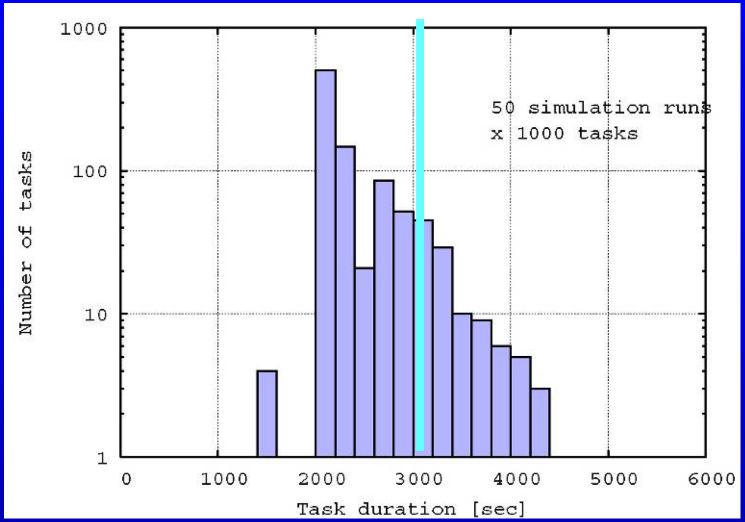
64% runs terminated < 30 min
96% runs terminated < 40 min





“intensive” use case: hadrontherapy

Period of testing	5weeks (with breaks)
Number of runs performed	50
Number of events simulated (per run)	10^5
Number of DIANE workers applied (per run)	20
Number of tasks (per run)	20
Number of events per task	$5 \cdot 10^3$



20 DIANE workers

Sequential

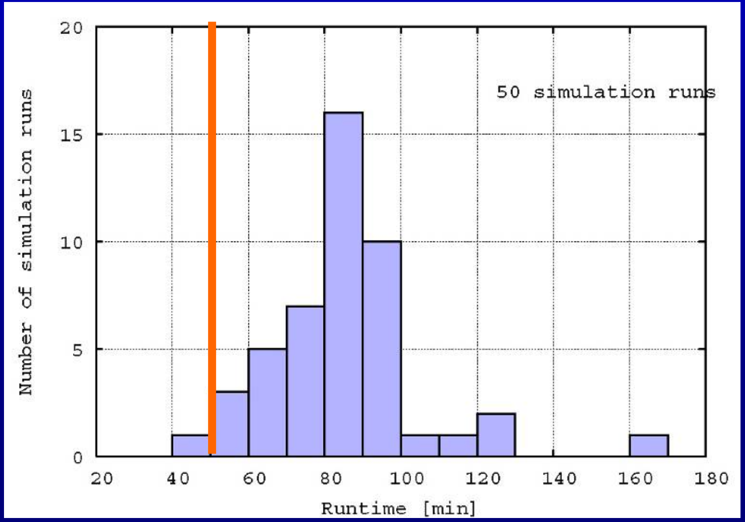
1 task = 50’ ± 50 s CPU time

Total simulation = 16.7 h ± 17 min

Ideal expectation with 20 workers:
~50 min for the whole simulation

On the grid

84% runs terminated < 100 min



Large scale medical simulations

- Test case: hadrontherapy
  - Increase statistics by 2 orders of magnitude
  - 10M events for high precision physics studies
  - Expected duration of whole production: 2-3 days
- 3 large statistics runs
  - 2 did not complete (80% done)
  - 1 terminated successfully
- Medical linac
  - Even more demanding than high statistics hadrontherapy
  - “Low” statistics demonstration OK
  - Realistic scale stable production does not look practical yet

Conclusions

Politically correct	➡	Grid technology offers unprecedented opportunities to the medical physics community for fast, high precision Monte Carlo simulation
Politically incorrect	➡	Grid technology is not mature yet for stable operation in a clinical environment by non-expert users
Scientific	➡	Both statements above reflect reality Quantitative tests evaluate the extent of success and contribute to identify problems and bottlenecks Measurements in real-life use cases are the path to solve problems

Anton Lechner

Austrian Doctoral Student
CERN IT/PSS

Major contribution to the results presented!

Alfonso Mantero
Graduate Student at
INFN Genova
1st Geant4-DIANE project

A BIG thank you!

Jürgen Knobloch Patricia Mendez Lorenzo Kuba Moscicki
CERN IT/PSS

and

Hurng-Chung Lee

Alberto Ribon