Geant4 developments in multi-threading, reproducibility, and physics

J. Apostolakis*, M. Asai+, G. Cosmo*, A. Dotti+,
G. Folger*, M. Gayer*, V. Ivantchenko¥*, M. Kelsey+,
T. Koi+, T. Nikitina*, W. Pokorski*, A. Ribon*,
V. Uzinskiy#, M. Verderi§ and D.H. Wright+

Institutions: * CERN, ¥Ecoanalytica, # JINR, § LLNR, + SLAC

Overview

- Recent Geant4 releases
- Latest release 10.0 – why ‘major’ ?
  - Multi-Threading
  - Strong reproducibility of events
- Improvements in Hadronic Physics
  - Improvement of cross sections and physics models
- Performance improvement
- New in Geometry – the first *USolids* library

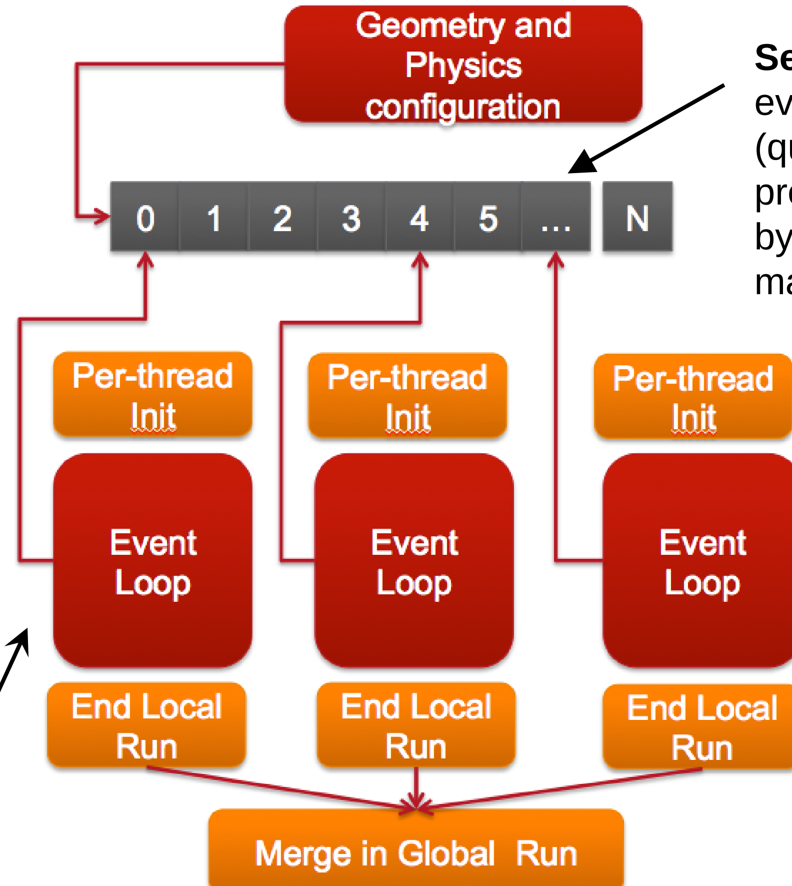
- Note: Highlights of latest developments

Recent G4 releases

- Geant4 9.4 (Dec 2010) – baseline for Run 1
  - Used in production: ATLAS(most), CMS(all) of **LHC Run 1**
- Geant4 9.6 (Dec 2012) is '**long-term support**' version
  - ATLAS production Run2; patches for LHC exp. to 2015+
- Geant4 10.0 (Dec 2013) was first **major** release in 6.5 years (since rel. 9.0 in June 2007)
  - Added big feature (MT)
  - Removed some features & physics models, ones which are obsolete, unnecessary or hard to maintain
  - CMS preparing to use it (10.0.p02) for production in Run2
- Recent development release: 10.1-beta (June 2014)
  - For the upcoming minor release 10.1 (Dec 2014)

MultiThreading - Intro

- Geant4 10.0: First release with Multi-Threading
  - Events are **independent**: each event can be simulated separately
  - Chosen event-level parallelism as current target
- Key goals: more cores within a memory ‘budget’
  - Scaling to use more CPU cores
  - Reduce extra amount of memory per worker thread
  - Use caches (L2/L3) better by sharing ‘constants’/tables
  - Small effort to port an application (compared to the effort to develop it.)
- First experiences
- Recent improvements

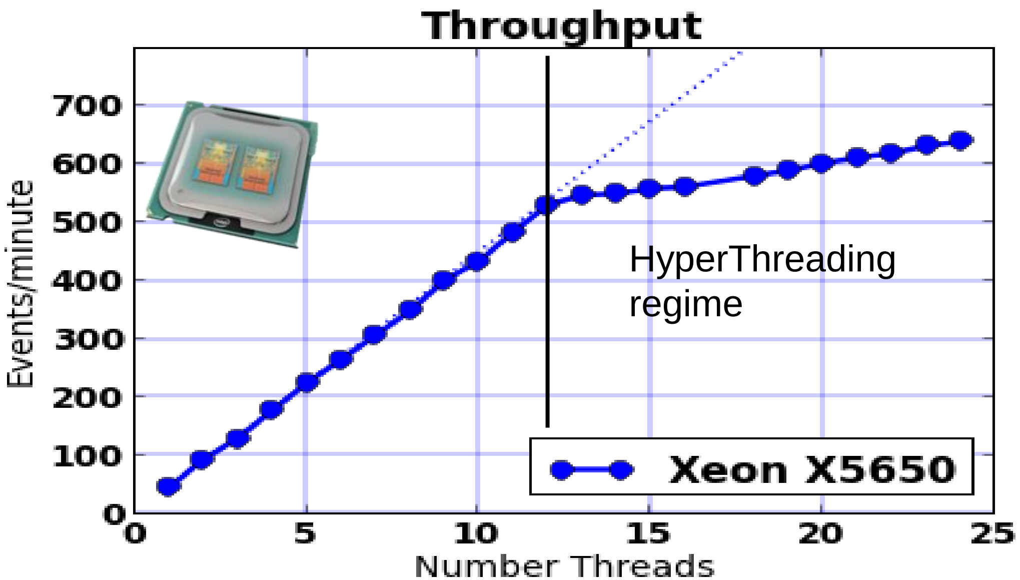


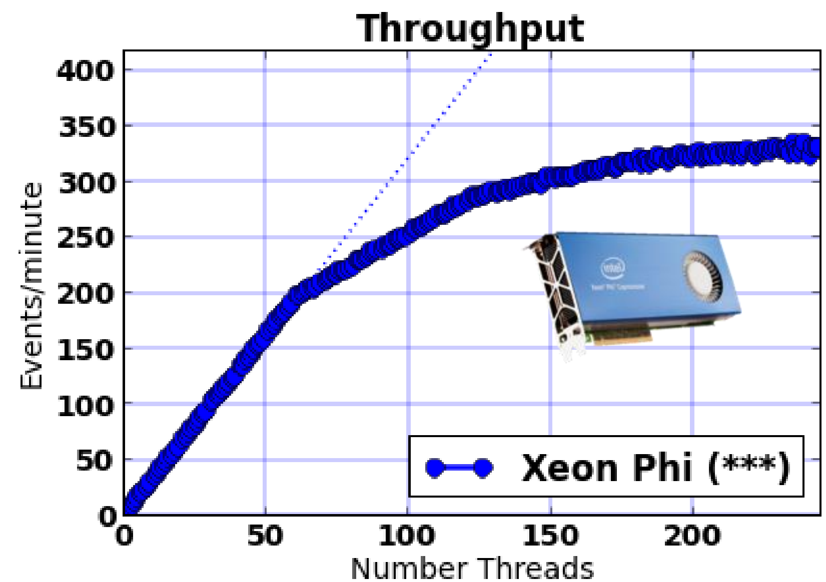
Seeds for events (queue prepared by the master)

Each thread obtains a set of events, simulates it and comes back for more. (**Dynamic** scheduling)

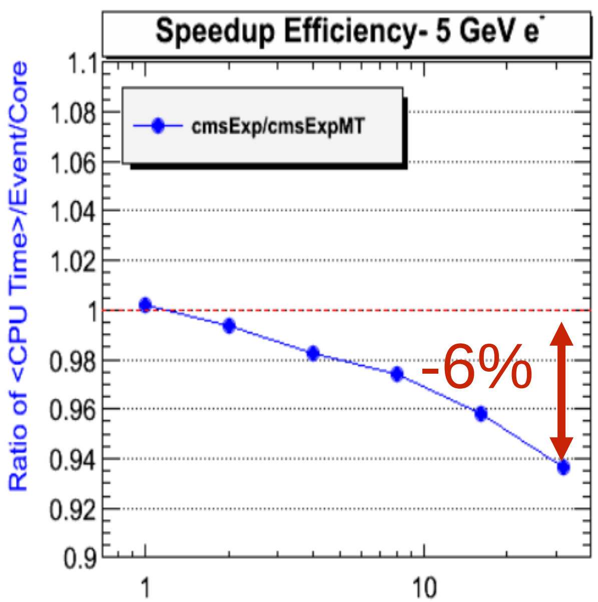
Global **output** for G4 scorers & histograms.
Per event hits ‘just work’.

CPU / Memory performance

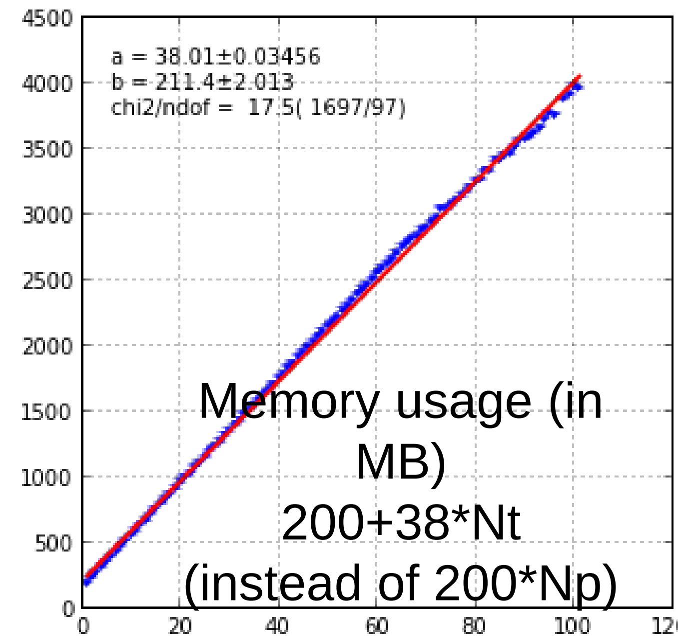




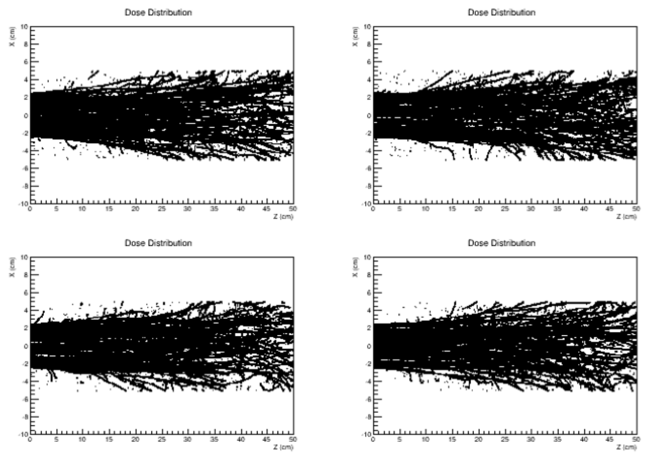
Geant4 10.0p02 on 12-core Xeon X5650



Obtained with simple application but realistic HEP geometry (CMS gdml 2008) - no hit output.



Multi-threading: recent/upcoming



- Control of per-thread memory improved
  - Ensuring more ‘constants’ (tables) are shared;
  - reduce memory reported ‘lost’ at the end of a job
- Extended to other parallelism ‘frameworks’
  - MPI: having 2+ MPI ‘jobs’, each 2+ threads
  - TBB: Changing workers to be task-based
- Tens of standalone applications ported
  - Limited number of changes, esp. to user action classes
- Big experiment applications limited by complex frameworks
  - must be thread-aware or thread-safe; CMS leads (task/TBB)
- Improve control of thread-local memory objects
  - create ‘workspace’ that owns thread-local objects, allowing it to be passed between threads (an inactive and a newly activated one.)
  - Needed to reduce memory use on task-based frameworks (e.g. TBB)
- Goal: Reduce thread memory to fit more threads on ~8-16 GB RAM (of accelerators) – free room for per-event data (hits)
  - Already identified several areas for improvements
- Goal: Enable threads to join/leave ‘work pool’.

What is Reproducibility?

- ‘Strong’ reproducibility
  - same result for an event, independent of its ‘location’ in a job (first, middle or last) or whichever the series of previous events.
- Required for HEP experiment applications
  - To enable easier debugging of problems/issues
  - Enables repeatable results in MT mode – where the order of events depends on number of workers (and load.)
- Testing: run a job; then rerun each event in separately
  - Use same RNG seed/state, input for the event
  - Compare number of RNG calls (weaker), observables (stronger)

Fixes and benefits

- First deployment of tests found several ***discrepancies*** (events with repetition errors)
  - Careful, long **investigation** needed to find source of each discrepancy
  - Typical error: ***caching a value*** of a cross-section using a ***different input*** (Energy within a bin, or one isotope of an element)
- Expanded, regular testing introduced
  - Found both new and other existing errors.
  - *The wait to **start debugging** problems in a long ATLAS jobs is not 16 hours, but 45 minutes!*

Hadronic Physics

Goal: improve physics performance, in particular for LHC calorimeters

- Energy response
- Energy resolution
- longitudinal shower 'size' (moment)
- lateral shower 'size' (moment)

Starting point:

- QGSP_BERT, the production physics list for HEP experiments (2005-2012)

Some key interaction properties which correspond to observables:

- Nuclear breakup, EM fraction
- Variance of EM frac., p/n inter.
- X-sections, products of quasi-elastic interactions
- neutron production, transport

Means:

- **replace** 'weak' physics models
- plug **gaps** between the stronger 'microscopic' models
- improve models, extend E range

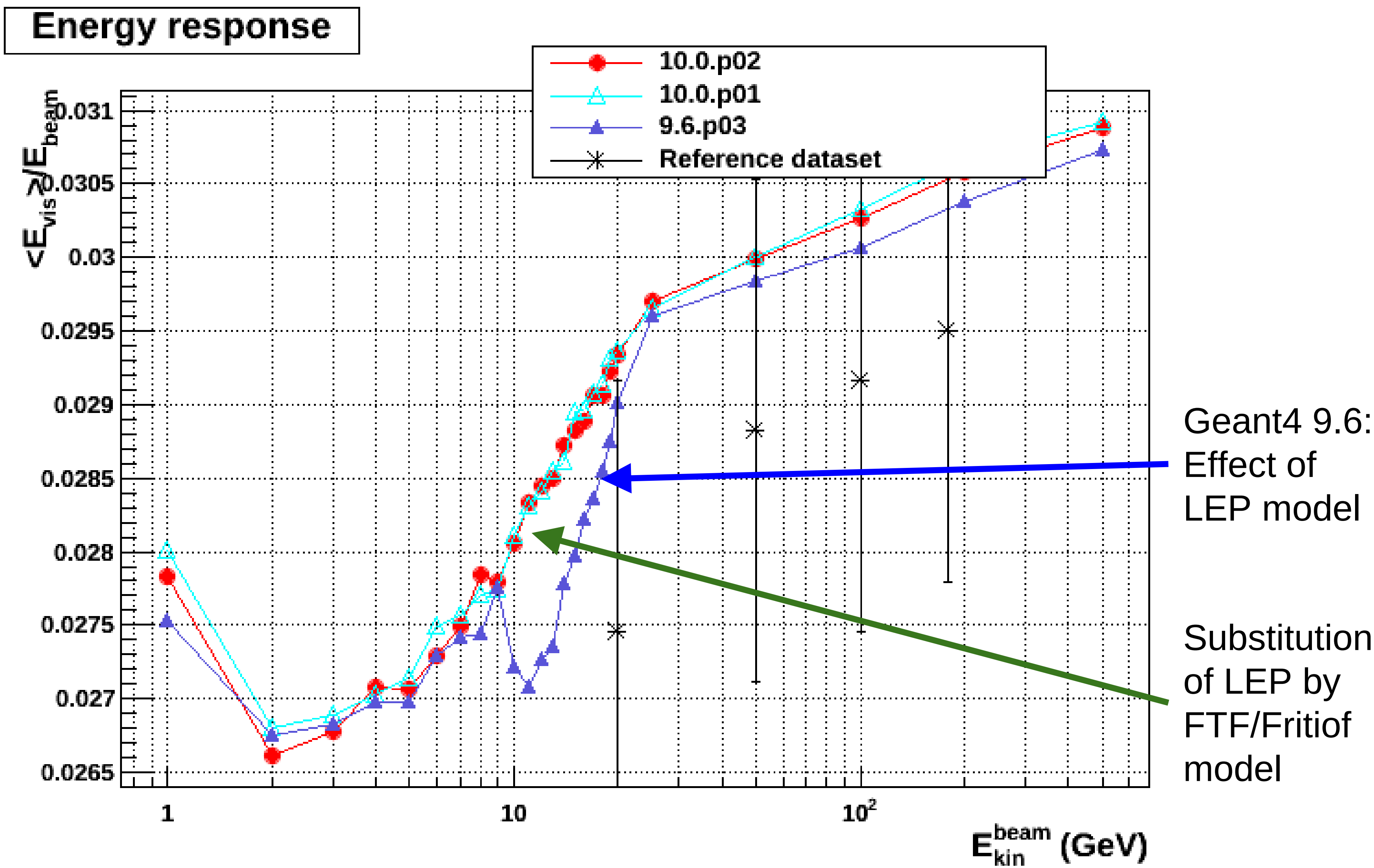
Starting point: Physics lists 9.4

Type of Physics	**QGSP_BERT (9.4)**	**FTFP_BERT (9.4)**
Electromagnetic Physics	Bremstrahlung	Bremstrahlung
	Urban93 tune	Urban93 tune
Hadron elastic	CHIPS (H), coherent (other)	CHIPS (H), coherent (other)
Hadron Inelastic (p,n,π, K)	Bertini cascade: 0-9.9 GeV	Bertini cascade: 0-5 GeV
	LEP Parameterised**:**9.5-25 GeV	**Fritiof FTF model:** 4GeV - 1TeV
	QGS model: 12 GeV - 1TeV	
Anti-baryons	***LEP***	CHIPS (p-bar) / LEP (other)
De-excitation	Precompound, evaporation	Precompound, evaporation
Quasi-elastic	CHIPS	FTF
Stopping particles	CHIPS	CHIPS
γ-nuclear, e-nuclear	CHIPS	CHIPS
n-capture	CHIPS	CHIPS

p/n
K/π

QGSP_BERT **RED**: Need(ed) to replace/improve **FTFP_BERT**

$E_{response} / E_{beam}$ in QGSP_BERT



π^- QGSP_BERT in slabs Fe/Sci

sizes and materials of ATLAS 'Tilecal'

Hadronic Improvement (2009-11)

- Developed & tuned **Fritiof/FTF** model
  - to plug gap below 5 GeV (and up to 12-15 GeV)
  - provide alternative model at high E
- Re-engineered **Bertini** cascade
  - Enabled use of alternative cross-sections
- **HP Neutron** option greatly improved (v 9.5)
  - ENDF/B VII.0 data files + alternatives
  - Results closer to MCNP with same data
- Many improvements (2009-13) in cascades, string models (later)

Hadronic Improvement (2012-13)

- Improved cross-sections (elastic, inelastic)
- ‘XS’: better cross-section, model of **neutron capture**
  - x-sec derived from HP, accuracy close to HP, without HP’s CPU penalty
- Improved model of **stopping** (negative particles)
- Dropped LEP/HEP parameterised models (<GHEISHA)
  - at core of LHEP, and used also in QGSP_BERT
- **Consistent** handling of **nuclear levels** in decay, de-excitation
  - revision of databases for levels with lifetime down to 1 ns
- Create **isomers** in pre-compound, evaporation, decay

Determines lateral hadronic shower size.

Energy-response, track multiplicity

necessary to enable simulation of activation

Fritiof (FTF) improvement

Impacts multiplicity, energy/type of tracks => energy response and resolution

- Add/tune Reggeon cascade de-excitation (G4 v9.3)
- Light anti-ion interactions (9.5) - $\bar{p}$, $\bar{n}$, $\bar{d}$, $\bar{t}$, $\bar{\alpha}$, $\overline{He^3}$
- Extension to nucleus-nucleus (10.0)
  - Interfaced to Binary Cascade, Pre-Compound
  - Support energy from 3-4 A*GeV to RHIC

Enabled analysis of anti-nucleons and light anti-ions production in ALICE

- Recent / ongoing:
  - Correction in string decay (10.1-β)
  - Retuning of parameters to improve description of h+p & h+A interactions - with NA49 data
  - more validation of nucleus-nucleus interaction

Impacts on shower size (long. & lateral), energy-response.

Cascades

- Treat projectiles 0.1-3 GeV (extended to 5-10 GeV)
  - Bertini-like: shells, (E<10GeV) wide use “BERT”
  - Binary QMD-like: static nucleons, pot. V, E< 3 (π: 1.5)
  - INCL++ Liege cascade: mature alternative (systematics)
- Components
  - Nucleus = 1-6 shells (density) or individual nucleons
  - Collisions create ‘holes’, secondaries => tracked
- Remnant de-excited by pre-compound, evaporation, fission, ..

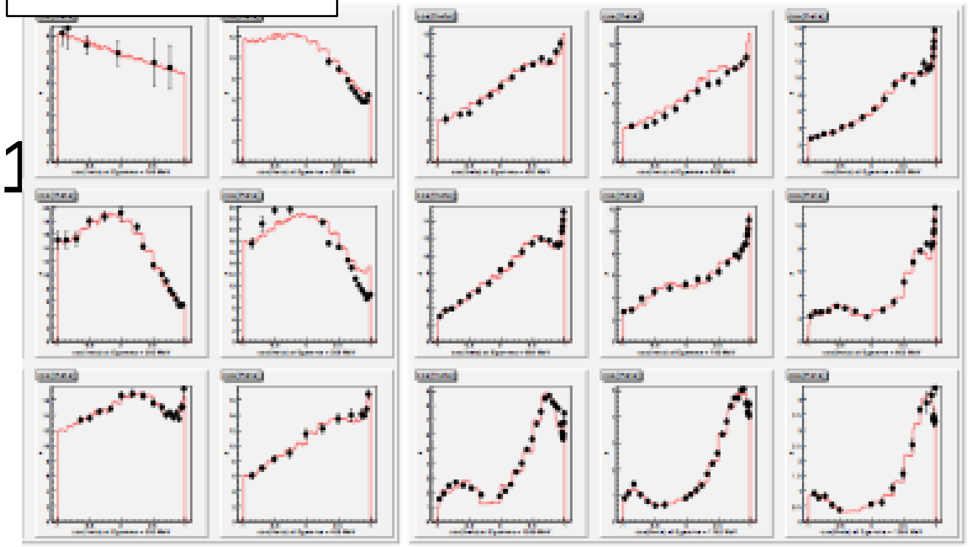
Cascades: Bertini

- New angular distributions for two-body final states
  - γ p, N-N elastic + pi-N (10.1
- Added γ- and electro-nuclear interactions (9.6)
  - replace CHIPS model (10.0**?**)
- Improved multi-body final state (N>5)
  - Kopylov N-body phase space generation (optional)
- Leptons as projectiles
  - Stopping/capture of muons generate cascade
- Improvement of selection of first interaction

Verification plots

(Comparison to data used to generate these fits.)

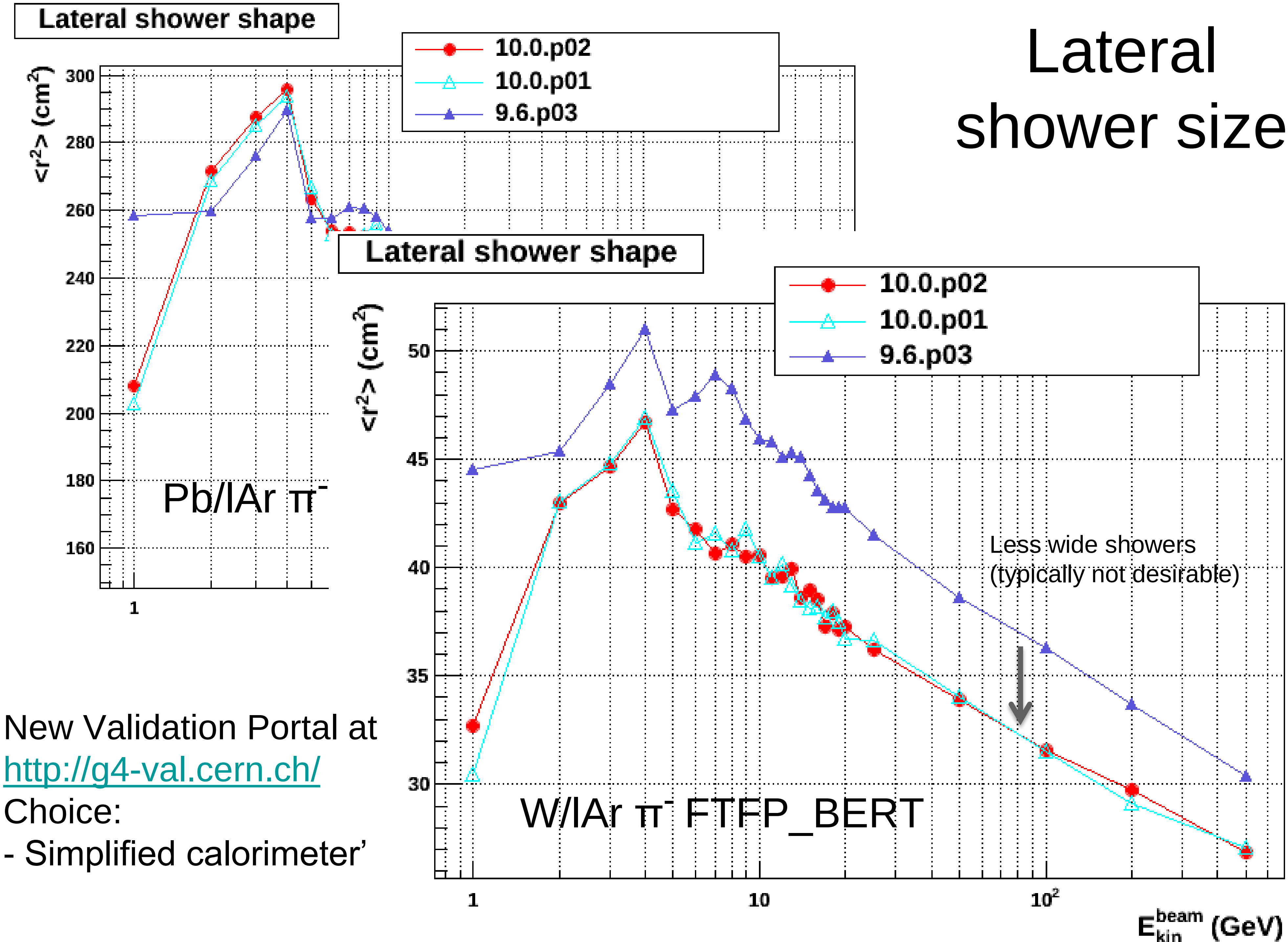
450 MeV



400 MeV

1.27 GeV

Lateral shower size



New Validation Portal at
http://g4-val.cern.ch/
Choice:
- Simplified calorimeter’

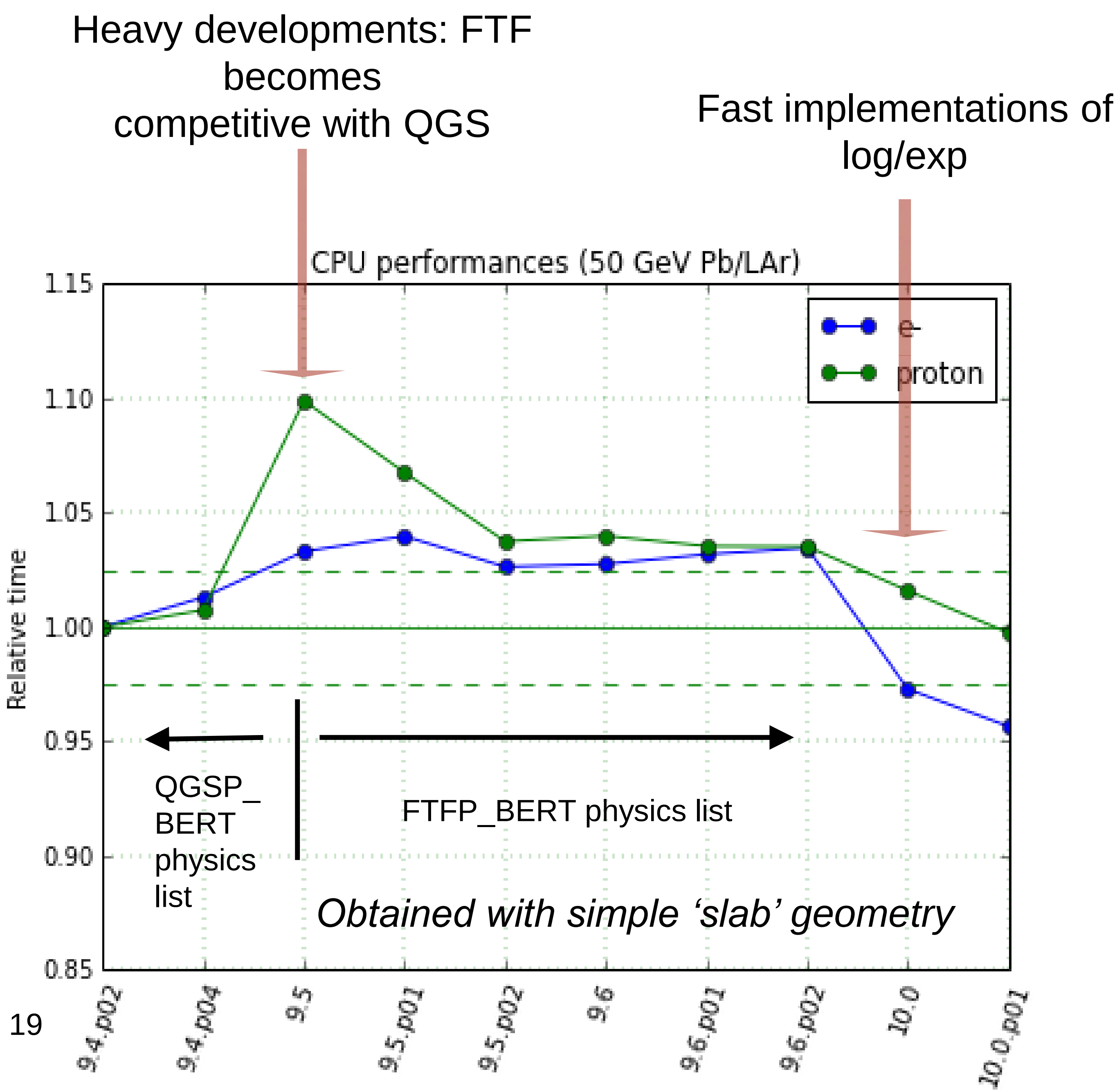
Current state: Physics lists in 10.0

Type of Physics	**QGSP_FTFP_BERT (CMS)**	**FTFP_BERT (ATLAS)**
Electromagnetic Physics	Improved Brem, γ-conversion	Improved Brem, γ-conversion
	Urban96 tune + Wentzel MSc	Urban96 tune + Wentzel MSc
Hadron elastic	CHIPS (H), coherent (other)	CHIPS (H), coherent (other)
Hadron Inelastic (p,n,π, K)	Bertini cascade: 0-5 GeV	Bertini cascade: 0-5 GeV
	Fritiof FTF **:**4-25 GeV	**Fritiof FTF model:** 4GeV - 1TeV • Reggeon cascade (in FTF)
	QGS model: 12 GeV - 1TeV	
Anti-baryons	**FTF**	**FTF**
De-excitation	Better Precompound (9.5)	& better Evaporation (9.5)
Quasi-elastic	a-la CHIPS	FTF
Stopping particles	**Bertini**	**Bertini**
γ-nuclear, e-nuclear	**Bertini**	**Bertini**
n-capture	**Bertini**	**Bertini**

p/n
K/π

QGSP_BERT **Green**: New replace/improved **FTFP_BERT**

Absolute throughput (sequential)



Improvements for MT brought CPU benefits also to sequential

Improvements (3-5% each):

- Refactor **eA, γA** X-sections (per isotope to per element)
- Fast log/exp (from VDT)

Smaller improvements (1-2%)

- Fluctuations
- UrbanMSc

Continuous monitoring of CPU time of benchmarks

- Regular (monthly+) runs
- Slabs and complex geom.

Challenge of diminishing returns (few parts >2%)

Geometry: USolids

- **Usolids**: started as unique solids library for Geant4 and Root.
- First goal: fast, robust implementations for all solids
- Available in 10.0 (first version) & improved in 10.1-β

- New fast ***tesselated solid*** (voxelized)
- New type: ***multi-union*** (10.1β)
  - voxels make it faster > 4+ unions
- Better implementations (< VecGeom)
  - sections of solids (e.g. Tube)
- Seek ***feedback*** on existing, upcoming versions

- Strong new focus of Usolids ('VecGeom'): performant vector (and sequential) implementation, using the same source code (also for GPU)
- USolids will become the *default* in Geant4, Root and Geant-V
  - Single-track vectorization for a few key solids: polycone, trd

Conclusions

Multi-threading

- MT saves ~150 MB/thread
  - Effort to reduce the 38MB of G4 – leave room for app hits
- CPU Performance per core maintained, seq. & MT
  - Sequential perf. ~ +5%
  - MT: scaling >90%*$N_{threads}$ for 1 thread/core, (for 60 avail.)
  - HT gave +30%
  - Easy Porting of simple app
- Achieved repeatability
  - A must for MT & debugging

Physics, Geometry

- Replaced obsolete LEP models with FTF
  - ➢ Corrects E-response QGSP_BERT in 10-20 GeV
- Improved, tuned FTF
  - Enabled extension to anti-nucleon, anti-ion
- Refactoring, retune of Bertini
  - ➢ Improved lateral shower shape
- New USolids library
  - Available for testing now – will become the default
  - Common with Root & GeantV

Backup slides

Shared Vs Thread-local

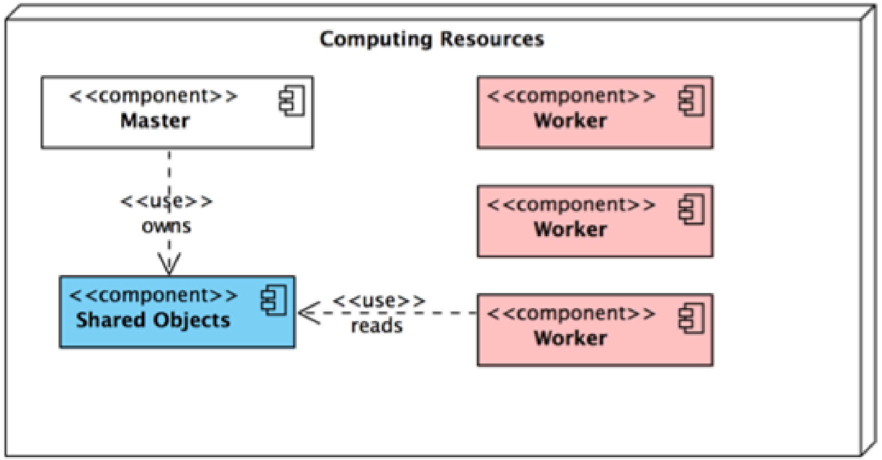
- To reduce memory footprint, threads must share part of their memory
- General rule in G4: ***threads are allowed to share whatever is invariant during the event loop*** (e.g. threads do not change those objects while processing events - they are “read-only”)

The ‘constant’ parts which we identified are:

- ○Geometry definition
- ○Electromagnetic physics tables

A few other processes have large tables

➢ a few are now shared, others tbd



- Memory cost for each additional thread requires 20-40MB for Geant4 data
  - ○ the user hits, histograms and other output are in addition to this
  - ○ saves ~150MB per thread for applications O(100) materials

Physics Lists: Alternatives in 10.0

Evolution of old recommended LHC production physics list: **QGSP_BERT** (ver **10.0**, Dec2013)

EM

- Urban(96) & Wentzel Multiple Scattering
- Improved Bremstrahlung
- Improved γ-conversion (E>80 GeV)

Hadronic

- Elastic
  - CHIPS (H), Coherent (other)
- Inelastic (p,n,π)
  - Bertini cascade: 0 - 9.9 GeV
  - String Fritiof/FTF: 9.5 - 25 GeV
  - String QGS model : 12GeV - 1TeV
  - Precompound & Evaporation
  - Anti-baryons: **FTF**
- Quasi-elastic model: FTF
- Stopping particles - FTF
- γ-nuclear, e-nuclear: Bertini

New recommended LHC production physics list: **FTFP_BERT** (Geant4 ver 10.0, Dec 2013)

EM

- Urban(96) & Wentzel Multiple Scattering
- Improved Bremstrahlung
- Improved γ-conversion (E>80 GeV)

Hadronic

- Elastic
  - CHIPS (H), Coherent (other)
- Inelastic (p,n,π)
  - Bertini cascade: 0 - 5 GeV
  - String Fritiof/FTF model 4GeV - 1TeV
    - Reggeon cascade (in FTF)
  - Improved Precompound (G4 v9.5)
  - Anti-baryons: **FTF**
- Quasi-elastic model: FTF
- Stopping particles: Bertini
- γ-nuclear, e-nuclear: Bertini

Cascades

- Treat projectiles 0.1-3 GeV (extended to 5-10 GeV)
- Nucleus = 1-6 shells (density) or individual nucleons
- Collisions create ‘holes’, secondaries => tracked
  - Bertini-like: shells, (E<10GeV) wide use “BERT”
  - Binary QMD-like: static nucleons, pot. V, E< 3 (π: 1.5)
  - INCL++ Liege cascade: mature alternative (systematics)
- Remnant de-excited by pre-compound, evaporation, fission, ..

Cascades: INCL++

- Mature ‘first-principles’ cascade
- Provides alternative for **systematics** studies
  - to Bertini for π/p/n**,** and to Binary, QMD for ions
- INCL++ = Liege cascade INCL recast in C++
  - Full implementation of models from INCL 4.6
- Extended **validity** up to 10-15 GeV (G4 10.1-β)
- Improved cross-sections for **small clusters** (d, t, α)
- Restored independent ABLA module for de-excitation

Isomers

- **Isomer levels** existed for all nuclei already
  - before G4 10.0 they were decayed promptly
- In G4 10.0 isomers with $\tau_{1/2}$ > 1 μsec are created/tracked
  - Option: use ENSDF State file with more data
- Need G4RadioactiveDecayPhysics to decay isomers
  - this must be added to physics list (it is not default)
- Relevant processes create/handle isomers
  - Cascades and decay create excited residuals
  - Capture, absorption, quasi-elastic can excite nucleus
- Improvement of radioactive decay / database

Improved Radioactive decay

- Fix: entire decay chains occur with their correct timing
  - in the past if a short-lived nuclide was reached, all subsequent decays would happen instantaneously
- Database synchronization
  - New radioactive decay database (4.1)
  - New photon evaporation database (3.1)
  - Synchronized to have identical energy levels and lifetimes down to a lifetime of 1 ns
- Model adapted to new isomer creation scheme (previous slide)

Event biasing

- New design for event biasing
  - Physics biasing
  - Angle biasing
  - Geometry/important biasing
- Flexible coupling of different types of biasing

Starting point: Physics lists 9.4

QGSP_BERT (Geant4 9.4, Dec 2010)

EM Standard ('hard-wired')

- Urban93 Multiple Scattering
- 'Standard' Bremstrahlung

Hadronic final state models

- Elastic
  - CHIPS (H), Coherent (other)
- Inelastic (p,n,π, K)
  - Bertini cascade: 0 - ***9.9*** GeV
  - ***Parameterised* (LEP)**: 9.5 - 25 GeV
  - **QGS model:** 12 GeV - 1TeV
  - Precompound, evaporation, fission
  - Anti-baryons: LEP
- Quasi-elastic: CHIPS
- Stopping particles - CHIPS
- γ-nuclear, e-nuclear: CHIPS
- n-capture: LEP

FTFP_BERT (Geant4 9.4, Dec 2010)

EM Standard ('hard-wired')

- Urban93 Multiple Scattering
- 'Standard' Bremstrahlung

Hadronic final state models

- Elastic
  - CHIPS (H), Coherent (other)
- Inelastic (p,n,π, K)
  - Bertini cascade: 0 - 5 GeV (revised)
  - **Fritiof FTF model:** 4GeV - 1TeV
  - Precompound, evaporation, fission
  - Anti-baryons: LEP
- Quasi-elastic: FTF
- Stopping particles: CHIPS
- γ-nuclear, e-nuclear: CHIPS
- n-capture: LEP

p/n
K/π