Computing Status

F. Bianchi
Torino

III SuperB Collaboration Meeting
Frascati, Mar 20, 2012

UNIVERSITÀ
DEGLI STUDI
DI TORINO
ALMA UNIVERSITAS
TAURINENSIS

Outline

- Cmake Build System.
- High Availability Cluster for Collaborative Tools.
- FastSim & Physics Tools. -> Physics Tools session.
- FullSim. -> FullSim & Background session.
- Distributed Computing -> Distributed Computing session.
- Stratus of PON ReCaS.
- R&D -> R&D + Overflow sessions.

Cmake build system

- CMake was modified to adapt it to the current FullSim developments and get rid of some SRT dependencies
  - namely .mk files which are used to build dependencies among packages
- We discovered that the CMake feature used for partial build led to target problems.
  - Fixed forcing the build of libraries right before that of binaries.
  - Not the most elegant solution, but working. . .

High Availability Cluster

- All the Collaborative Tools will be migrated to an High Availability cluster.
  - To have more stable hardware.
  - To have better response time (avoid timeout with some Java clients)

High Availability Cluster

- EqualLogic storage for the High Availability Cluster is arrived and is online at CNAF.

- Schedule:
  - Storage : initial setup already completed
  - Platform migration compatibility tests : in progress
  - Cluster VMs setup: by the end of April
  - Migration of CAS/Portal/database systems: starting in May (with software upgrades performed at the same time)
  - Migration of other services: tentatively starting in the second half of May

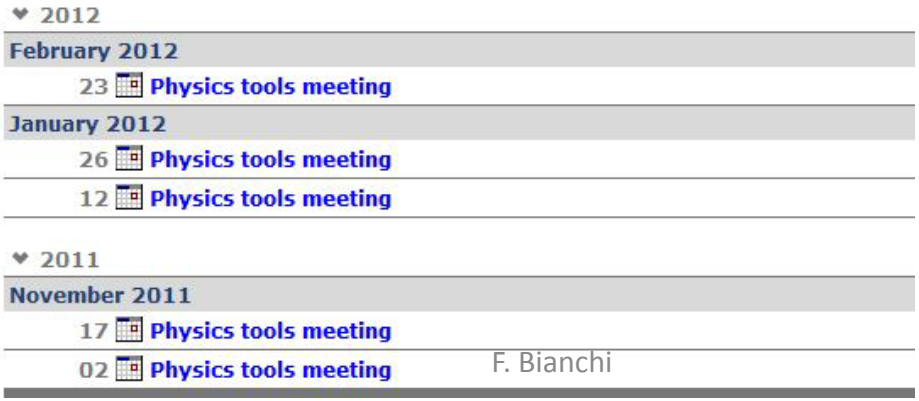
Activities of the Physics Tools Group

since the Frascati meeting, Dec 2011

- FastSim developments
  - improved sub detectors response
  - quality monitoring tools
  - skims
  - improved documentation

Physics tools meetings

- Every other Thursday at 17:30 CET. Agenda in Indico. http://agenda.infn.it/categoryDisplay.py?categId=491
- Mailing list: superb-fastsimu@lists.infn.it



Improvements in detectors simulation

Improved simulation of detector response of EMC and IFR

Example:

smarter EMC clustering to deal with high background conditions

before

CPU time per event (ms/event) consumed by modules (running 10 B0B0bar events)

Module	no bkg	1x	2x	3x	4x
PmcReconstruct	205	340	572	917	1191
PmcSimulate	60	176	290	522	625
BtaLoadMcCandidates	0	303	1671	8987	16147
PacCaloSplitMerge	1	190	3086	18178	54315
PmcRadBhabhaNeutronBkgInput	0	177	227	336	319
RacTestInput	6	80	158	317	366

background level w.r.t. nominal level

CPU time per module (ms/event)

~1 min/event with old algorithm

new algorithm + other fixes

CPU time per event (ms/event) consumed by modules (running 10 B0B0bar events)

Module	no bkg	1x	2x	3x	4x
PmcReconstruct	279	355	705	1035	1442
PmcSimulate	82	192	376	550	696
BtaLoadMcCandidates	0	34	130	243	450
PacEmcReclustering	1	4	11	15	21
PacCaloSplitMerge	0	1	1	4	11
PmcRadBhabhaNeutronBkgInput	0	171	260	291	299
RacTestInput	9	73	190	299	393

<0.5 s/event with new algorithm

Quality Monitoring tools

- Tools to monitor and validate the output of the FastSim MC production
- Needed to validate the release during pre-production and to monitor the data quality during production
- Two steps:
  - Each subsystem developed a module that stores quality-sensitive variables into output ROOT files
  - Each subsystem wrote a macro to produce quality histograms from the ROOT files of the previous step
- Development basically completed for all detectors

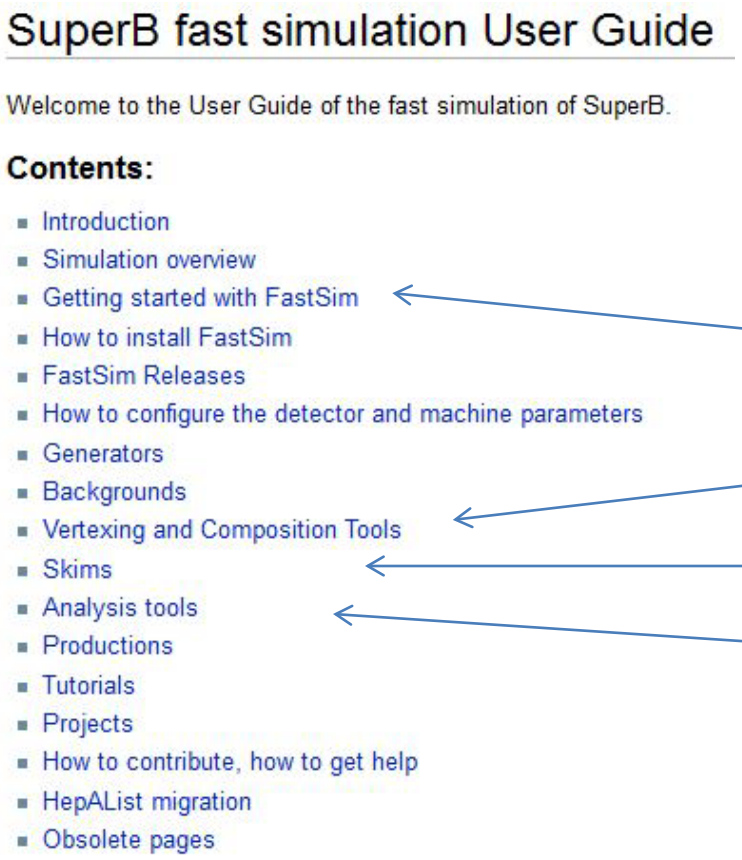
Skims in FastSim

- A *skim* is a subset of data defined by a given set of selection criteria
  - example: hadronic or semileptonic tag skims, tau skims, charm skims etc..
- Skims are necessary to make physics analysis at BaBar
  - they allow to run the selection of a given analysis over a small fraction of all events, with a huge save of time and CPU resources
  - for the same reason they're needed also at SuperB
- In FastSim the simulated events are not *persistent* (there are no *collections of events* saved on disk)
  - how can a skim be defined?
- Idea: define a *skim* as a collection of seeds
  - It requires that two FastSim jobs with same configuration and seed produces exactly the same output
  - Some tricky changes to the FastSim code needed

Code ready and committed. Validation in progress.

Improved documentation

http://mailman.fe.infn.it/superbwiki/index.php/SuperB_fast_simulation_User_Guide


SuperB fast simulation User Guide

Welcome to the User Guide of the fast simulation of SuperB.

Contents:

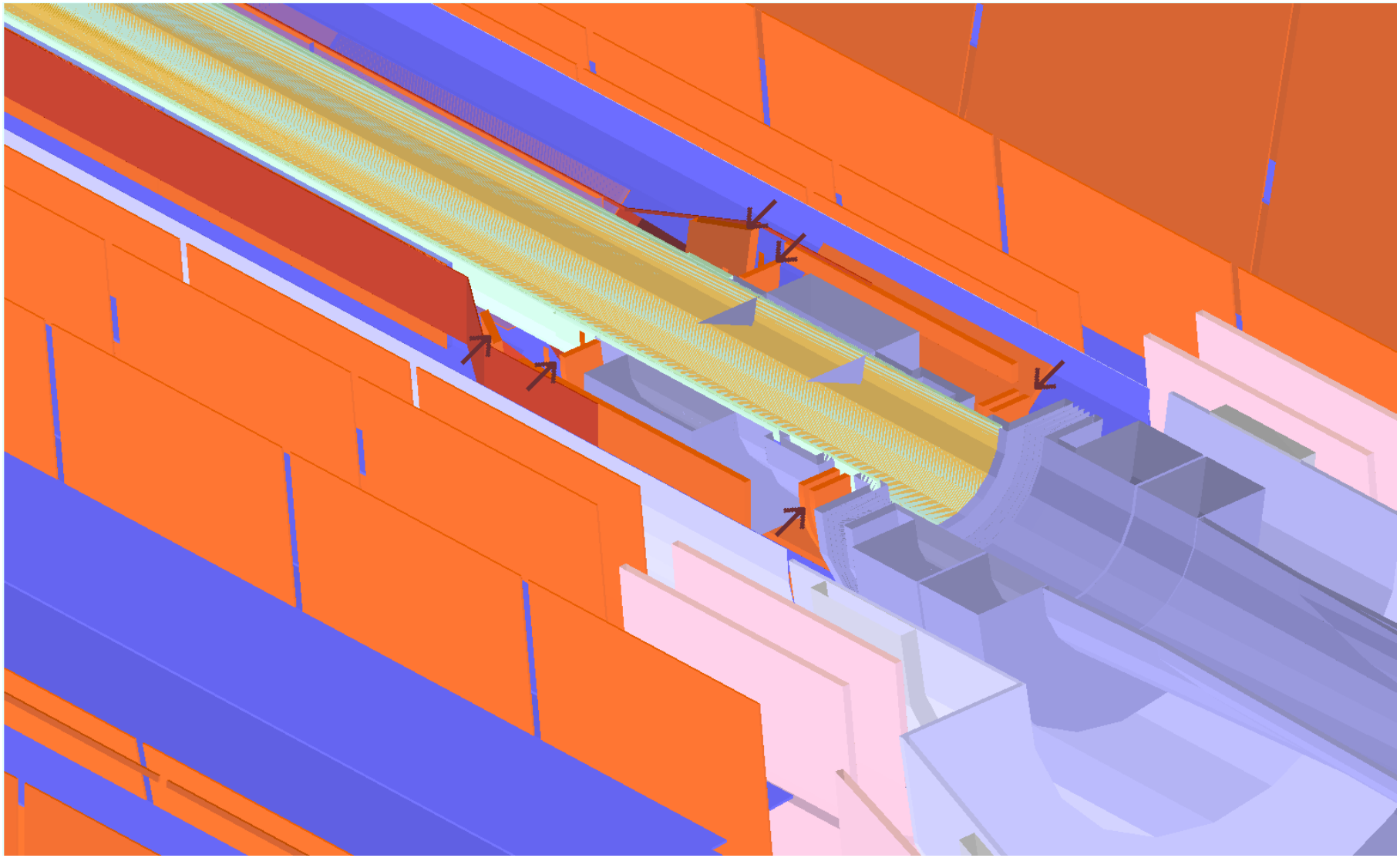
- Introduction
- Simulation overview
- Getting started with FastSim
- How to install FastSim
- FastSim Releases
- How to configure the detector and machine parameters
- Generators
- Backgrounds
- Vertexing and Composition Tools
- Skims
- Analysis tools
- Productions
- Tutorials
- Projects
- How to contribute, how to get help
- HepAList migration
- Obsolete pages

- FastSim documentation is improving
  Examples:
  - up to date section for beginners and new section for developers
  - Vertexing and Composition tools guides translated from BaBar
  - Section on skims
  - Analysis tools section is expanding
  - and more…
- Further work and manpower needed

FullSim: Overview

- Implemented and available for the next round of production
  - More precise tuning of the bgframes.
  - More generator info in metadata.
  - Radiation Monitor: added scoring volumes to assess doses.
  - Optical photons :
    - Already done for DIRC
    - Work in progress for FTOF
- In addition, there are a number of more general, medium-term, longstanding issues we need to start dealing with
  - Event display: Root display ok for the time being.
  - Event structure: move to a more modular event structure
  - Runtime configuration
  - Handling many geometries
- There are MANY MORE issues we should be working on, but one must be realistic, and ***with present manpower they'll have to wait***

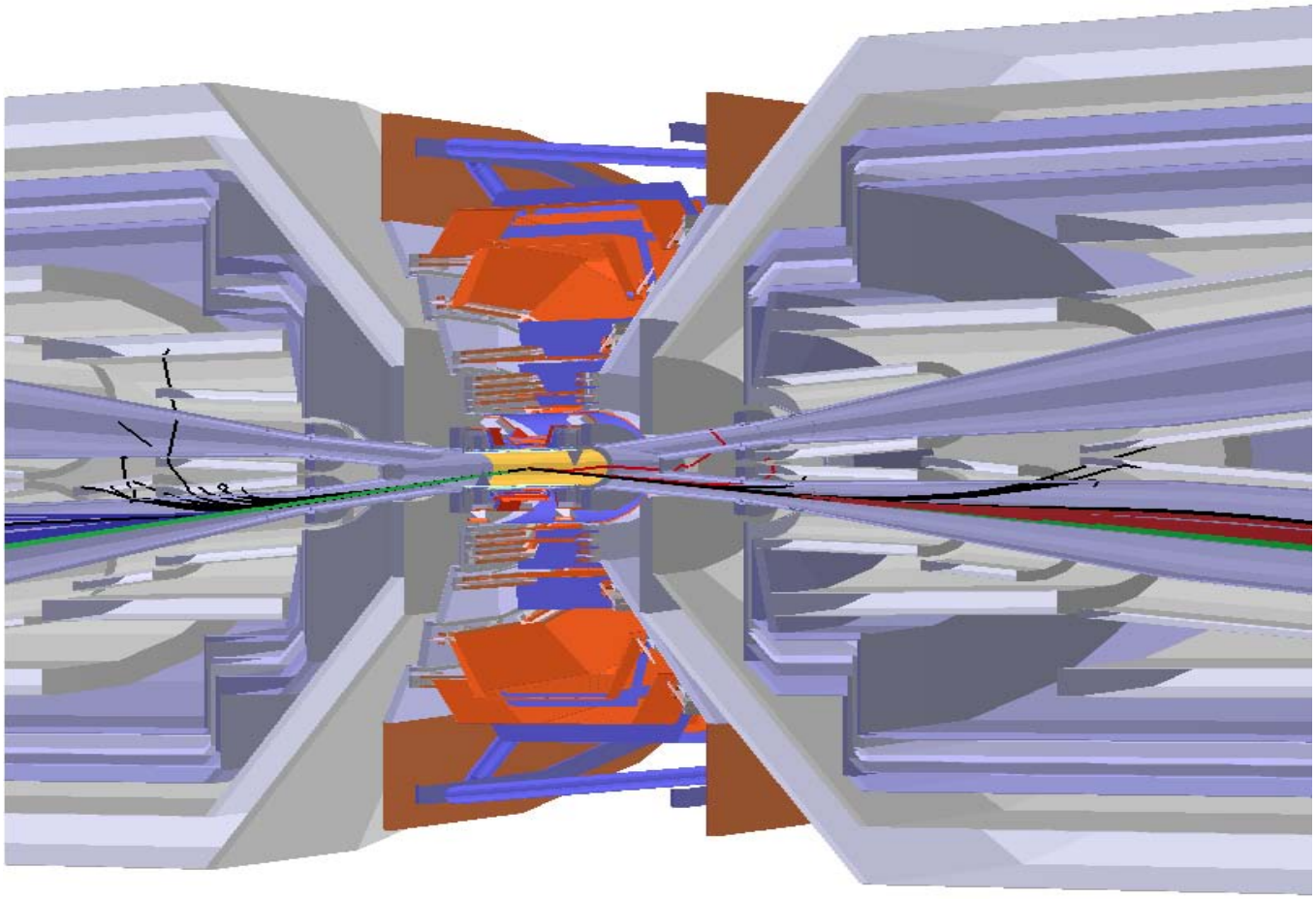
Rad Mon



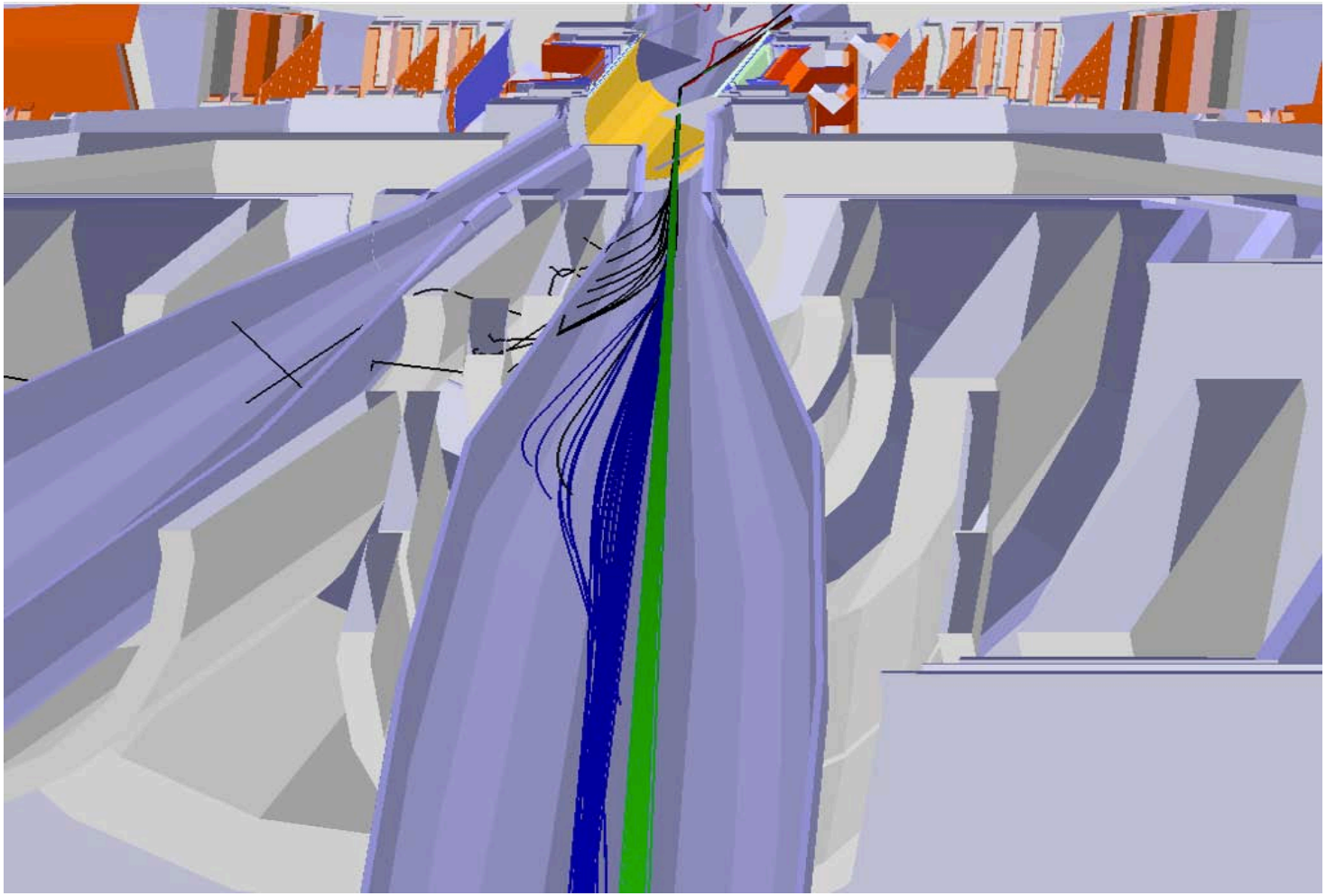
Event display

- All the truth informations (including trajectories) were persisted using native ROOT classes
  - One of the reasons was to allow easy drawing within root
  - Never really exploited until the Vienna meeting, where Eugenio produced some astounding ROOT displays
    - See next slides

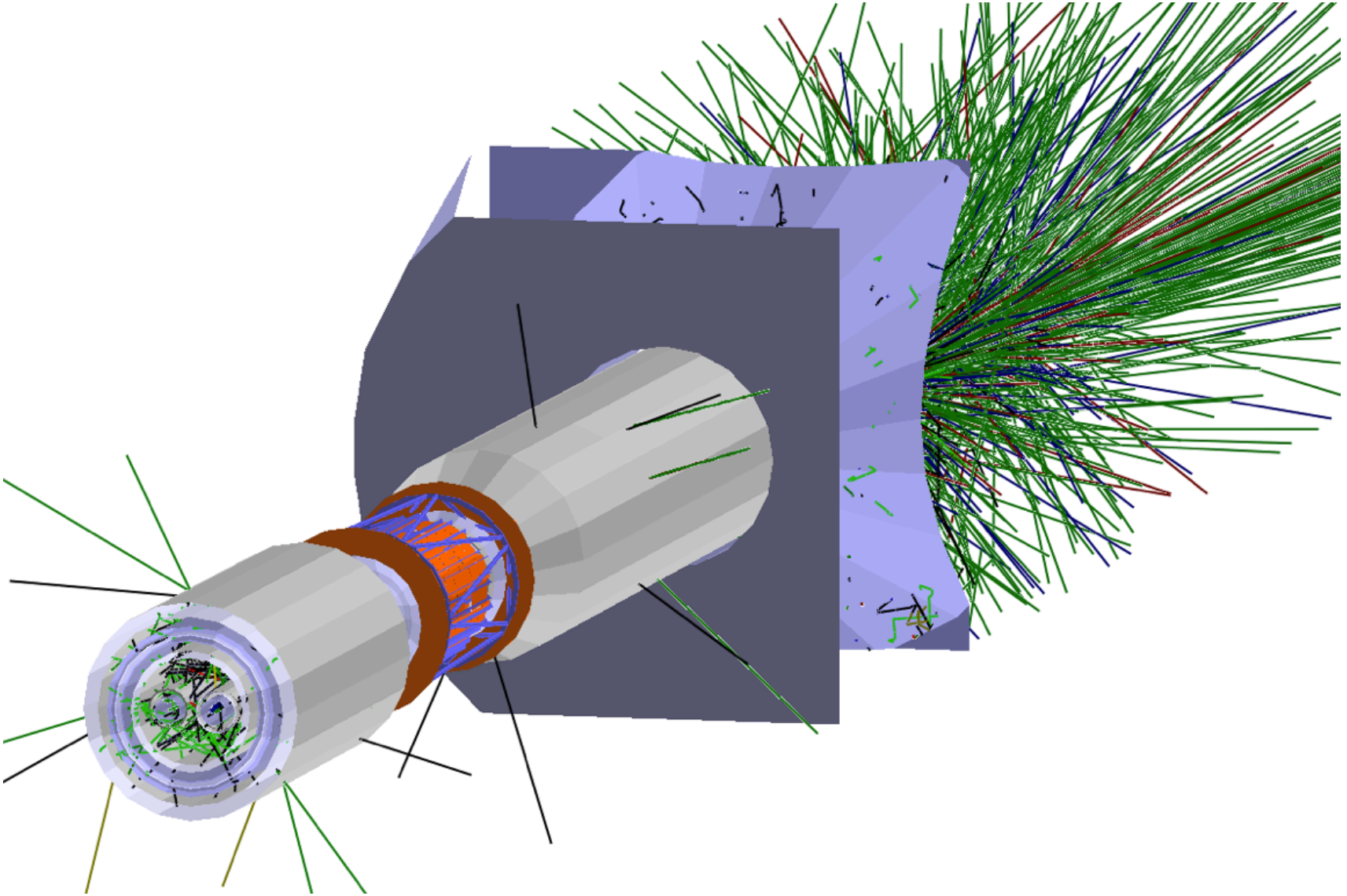
Event display



Event display



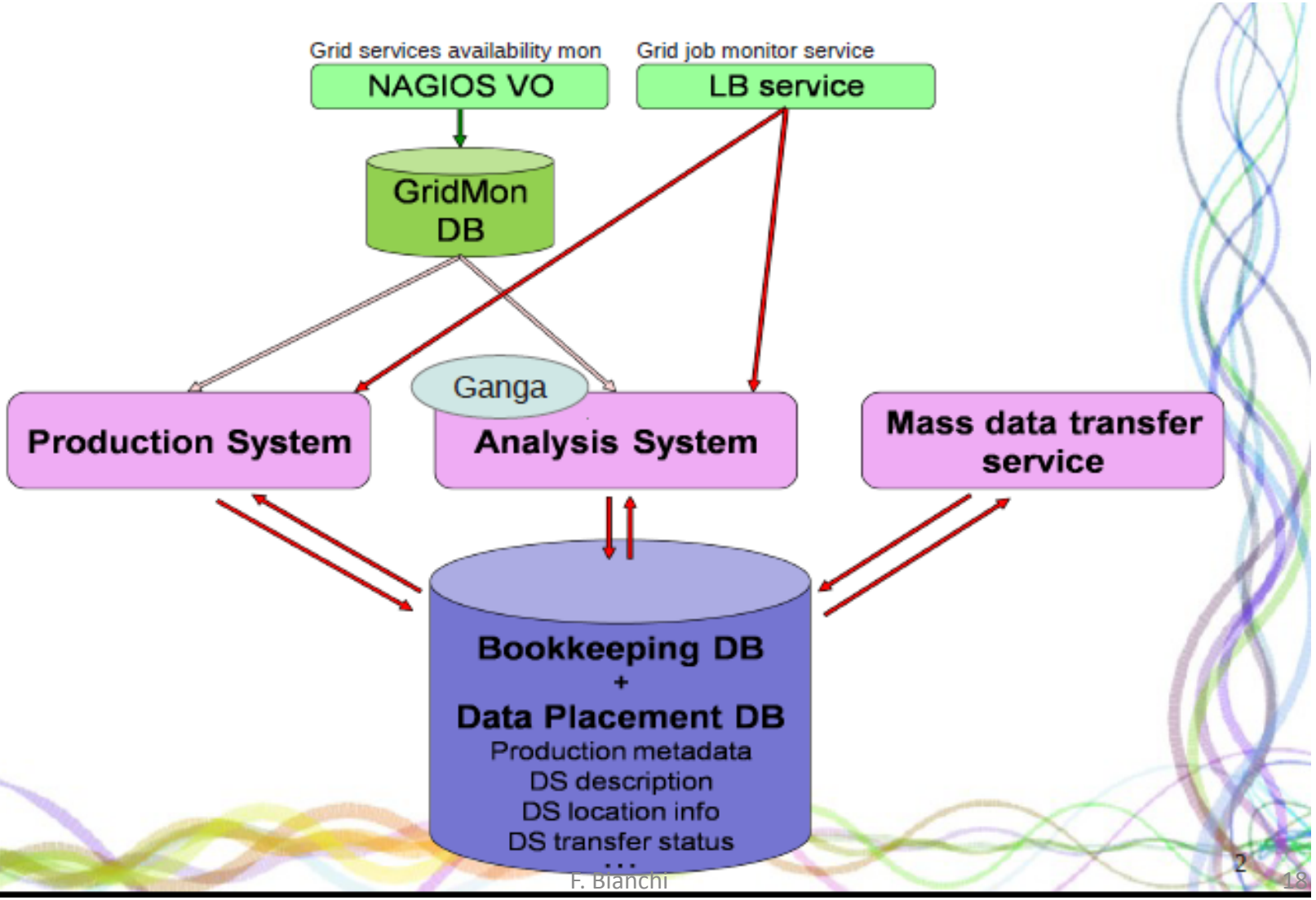
Event display



Handling Many Geometries

- Present schema is that a geometry is a folder within BrnRunTime
  - Each modification to the geometry (or a new geometry altogether) requires a new release of the whole Bruno code
    - Burden for people doing releases
    - Delay for people doing detector/background studies
- Work ongoing to decouple geometry releases from Bruno releases
  - First implementation will use svn to tag geometries, and some dedicated scripts to couple a given geometry to a Bruno release

Distributed Computing Tools Overview



Ganga SuperB Layer

- Implemented use cases and features:
  - Analysis: official and personal FastSim and FullSim dataset analysis
  - Personal simulation production
    - Automatic job data preparation
    - Automatic resources selection
    - Job resubmission
    - On line job monitoring
    - Integration with bookkeeping DB
- Report on procedures, use case design and framework capabilities:
  - Thursday 22nd 15:00 at " Computing + Physics:Production Plan"
- A technical report on development status and work in progress:
  - Wednesday 21st 11:30 at "Distributed computing"

Nagios Monitoring (CNAF)

- Site status is monitored with NAGIOS:
  - https://sb-serv01.cr.cnaf.infn.it/nagios/
  - The EGI Standard checks are executed
  - Specific SuperbVO checks will be added
  - What is monitored is the service, not the host
- The site status is stored in an external DB.
- Every time a service status changes, a nagios plugin updates this DB.
- The informations present in this DB are used to prevent the production job submission to problematic sites.
- Host monitoring tool under development in Napoli
  - Primary target is monitoring the resources in the PON sites

Book-keeping Database

- Porting of BK DB from MySQL (5.1) to PostgreSQL (9.1) decided after the 2nd SuperB Collaboration Meeting
  - PostgreSQL is more SQL compliant
  - Exploiting its hstore data-type, allows to solve some major architectural issues concerning the dataset management
- As first step, the MySQL *sbk4* book-keeping database has been reproduce under PostgreSQL
  - Compatibility between the two *sbk4* versions has been successfully tested
- A new book-keeping database (*sbk5*) has been developed from sbk4 under PostgreSQL 9.1
- Extensive tests to check PostgrSQL and HTTP interface system robustness have been done.
  - Result were good, the system is capable to sustain a peak of 10000 DB connections and transactions in ~100s.

Distributed Resources

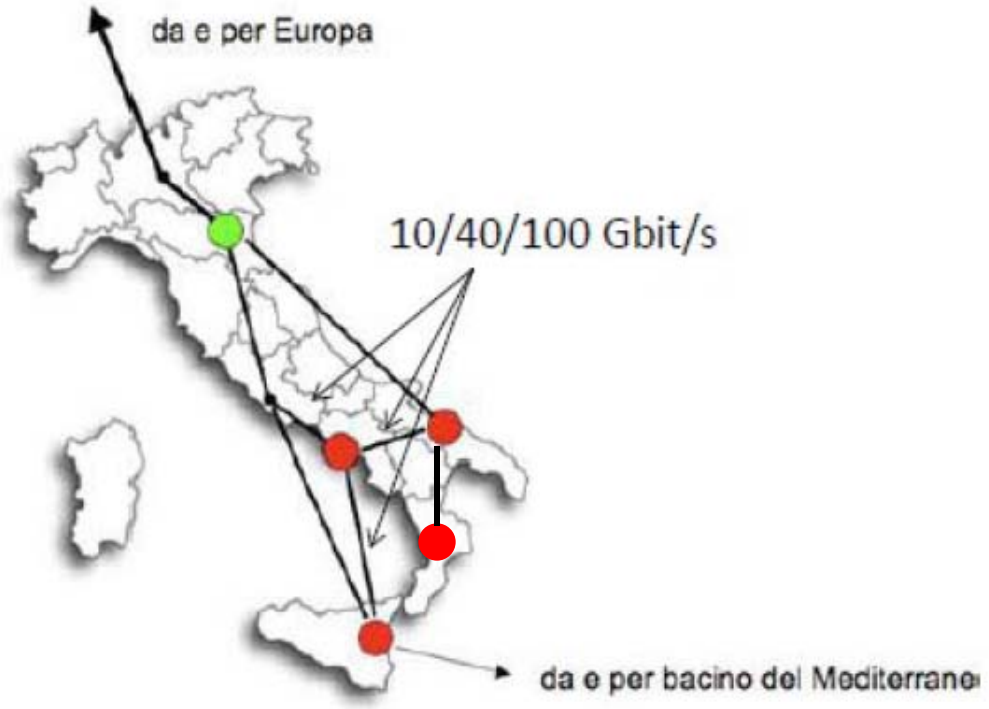
- 26 sites are available to the SuperB VO.
  - From: Canada, France, Italy, Poland, UK and USA

Site	Min (cores)	Max (cores)	Disk (TB)	SRM layer	Grid Org.	Site contacts
RAL(T1)	200	1000	25	Castor	EGI	F. Wilson, C. Brew
Ralpp	50	500	5	dCache	EGI	F. Wilson, C. Brew
Queen Mary	300	2000	150	StoRM	EGI	A. Martin, C. Walker
Oxford Univ.	50	200	1	DPM	EGI	K. Mohammad, E. MacMahon
IN2P3-CC(T1)	500	1000	16	dCache	EGI	N. Arnaud, O. Dadoun
Grif	50	300	2	DPM	EGI	N. Arnaud, O. Dadoun
in2p3-lpsc	50	100	2	DPM	EGI	J.S. Real
in2p3-ires	50	100	2	DPM	EGI	Y. Patois
CNAF(T1)	500	1000	180	StoRM	EGI	A. Fella
Pisa	50	500	0.5	StoRM	EGI	A. Ciampa, E. Mazzoni, D. Fabiani
Legnaro	50	100	1	StoRM	EGI	G. Maron, A. Crescente, S. Fantinel
Napoli	500	2000	15	DPM	EGI	S. Pardi, A. Doria
Bari	160	260	0.5	StoRM/Lustre	EGI	G. Donvito, V. Spinoso
Ferrara	10	50	0.5	StoRM	EGI	L. Tomassetti, A. Donati
Cagliari	10	50	1	StoRM	EGI	D. Mura
Perugia	10	50	1	StoRM	EGI	R. Cefala'
Torino	50	100	2	DPM	EGI	S. Bagnasco, R. Brunetti
Frascati	30	100	2	DPM	EGI	E. Vilucchi, G. Fortugno, A. Martini
Milano	50	100	2	StoRM	EGI	N. Neri, L. Vaccarossa, D. Rebatto
Catania	?	?	?	StoRM	EGI	G. Platania
Slac	400	400	10	NFS	OSG	S. Luiz, W. Yang
Caltech	200	400	4.5	NFS	OSG	F. Porter, P. Ongmongkolkul, S. Lo
OhioSC	?	?	?	dCache	OSG	R. Andreassen, D. Johnson
Victoria	50	100	5	dCache	EGI	A. Agarwal
McGill	100	200	1	?	EGI	S. Robertson (in progress)
Cyfronet	100	500	10	DPM	EGI	L. Flis, t.Szepienie,J. Chwastowski
Total	3520	11110	439			

Computing Infrastructure

- In Italy:
  - CNAF
  - 4 new centers in Bari, Catania, Cosenza, Napoli

- + centers in other participating countries



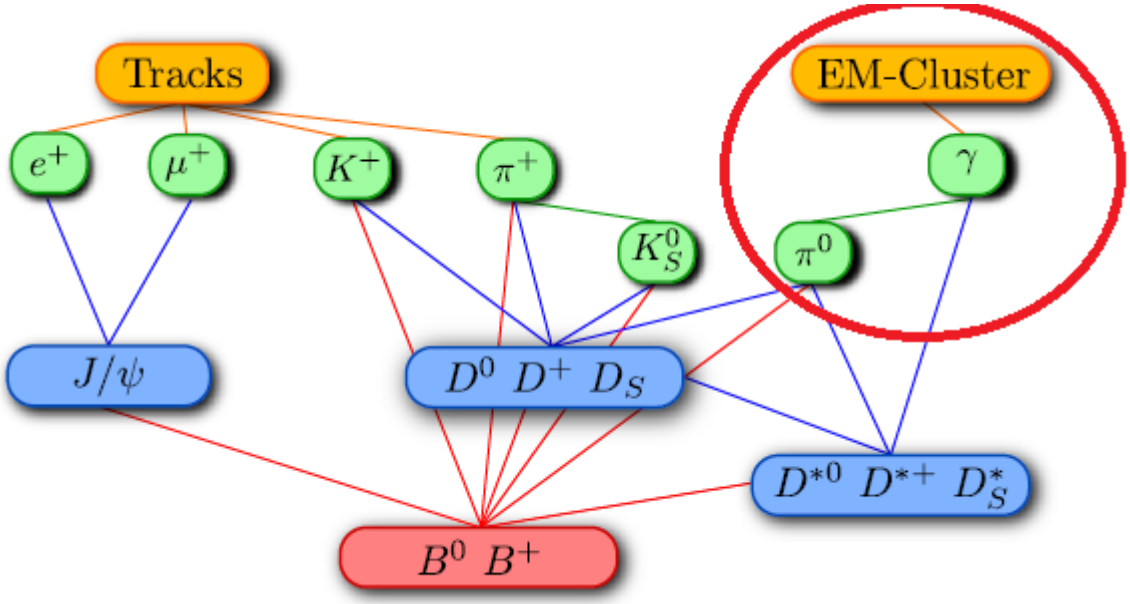
Pon ReCaS

- Approved and funded with 13.7 ME.
- Contract with MIUR signed on 2011.11.30.
- Spending chapters approved on 2012.02.16.
- Meeting of internal Committee on 2012.02.10, 2012.02.24, 2012.03.13.
- Personnel: procedure to hire 1 person in Na, Ba, Ct started.
- Bari: civil engineer work design started.
- Napoli & Catania: power distribution design started.
- Napoli: procedure for limited hardware acquisition started.

R&D: Parallel Framework (Padova & Cnaf)

- Evaluation of Art, the framework developed by Fnal from a fork of CMSSW:
  - How much work to migrate SuperB Framework modules to Art ones?
  - How to parallelize Art modules using Intel Threading Building Blocks?
  - Is the Art configuration language a good candidate to substitute tcl files?
- Exploit concurrent programming paradigms:
  - Parallelize execution of modules inside an event by means of generic graphs
  - In particular evaluation of Tbb graph objects.
- Make the Framework able to delegate computational load to an external accelerator (e.g. Intel® MIC®)
  - An exercise in this direction is being done with EvtBtoXsgammaKagan::computeHadronicMass() function

R&D: GPU (Napoli)



4 stages process to reconstruct B meson decays

stage	particelle
1	tracks, K_S, γ, π^0
2	$D^{\pm}_{(s)}$, D^0, e J/ψ
3	$D^{*\pm}_{(s)}$ e D^{*0}
4	$B^{\pm}$ e B^0

Studied performance of π^0 reconstruction as function of the number of γ on GPUs
From 1000 γ parallel algorithm is more performing

SEQUENZIALE	PARALLELO
0,013 ms	0,351 ms
0,159 ms	0,457 ms
15,211 ms	0,621 ms
1422,742 ms	19,699 ms
10637,217 ms	170,23 ms
18613,341 ms	301,756 ms
23266,988 ms	380,989 ms

Conclusions

- The computing group is supporting the collaboration by providing:
  - Collaborative Tools
  - Physics Tools: FastSim, etc.
  - FullSim
  - Production Tools
  - Bookkeeping Tools
- There is an active R&D program aimed at the design of the computing model.
- The activities funded under the Pon ReCaS are an important step forward into building the computing infrastructure.
- A severe lack of manpower is affecting us.
- Come and join the fun !