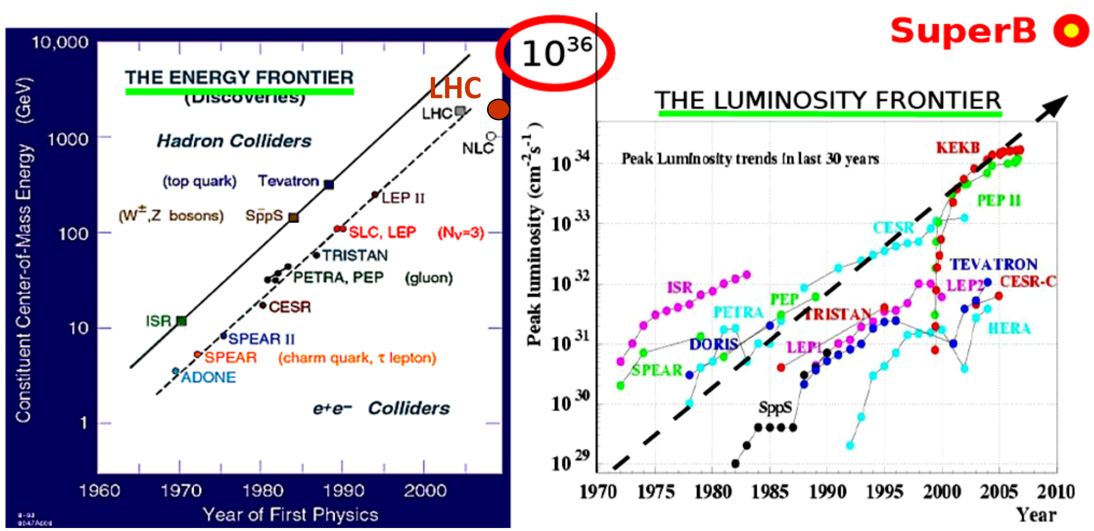
The SuperB Real (and Virtual) Organization

Steffen Luitz, Armando Fella

OSG Council 07/10/12

What is SuperB?

- Next-generation "Flavor Factory" to be built near Frascati (Rome) in Italy
  - 1st-generation B-Factories (BaBar and Belle) have collected ~1.5ab-1 together
    - Many physics results
  - Goal: collect 50-100ab-1 in 5 years
- Search for New Physics - complementary approaches
  - Relativistic approach – increase the energy and look for the production of new particles. ("Energy Frontier")
  - Quantum approach – increase the luminosity (and number of collisions collected) and look for effects of physics beyond the standard model in loop diagrams. ("Intensity Frontier")



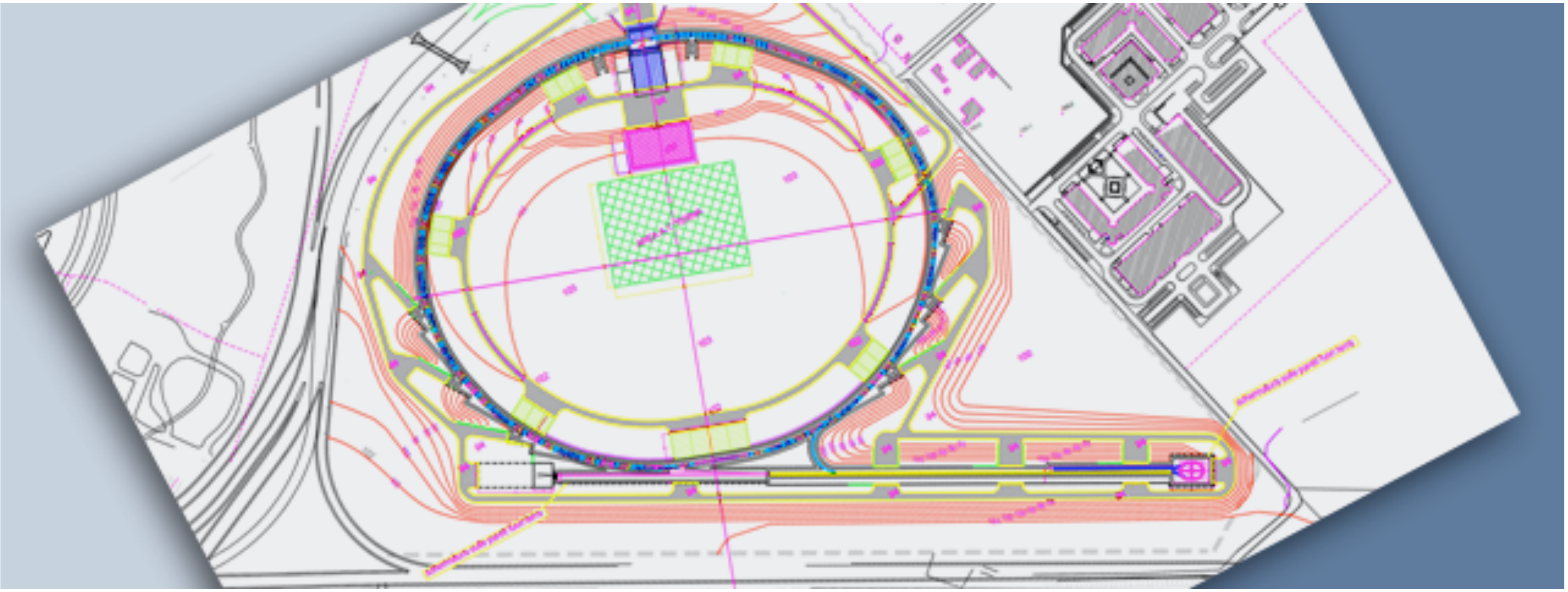
SuperB

SLAC NATIONAL ACCELERATOR LABORATORY

SuperB Physics Goals

- Some of the goals
  - Increase precision of BaBar and Belle by a factor of ~10
  - Challenge the CKM at the 1% level
  - Improve sensitivity for Lepton Flavor Violation in tau decays by 10-100
  - Explore T-violation in tau
  - CPV in Charm sector
  - Spectroscopy
- This rich menu can be effectively mined with 75 ab^{-1} in 5 years at Υ(4S) and a few months at Charm threshold with peak luminosity of 1×10^{36} cm^2 s^{-1}.

The Nicola Cabibbo Lab



- Created in October 2011 as joint venture of INFN and University or Rome
- Located on a green field site of Tor Vergata University (near Frascati / LNF and Rome)
- Ring circumference ~1200m

http://www.cabibbolab.it

The SuperB Machine

- Fundamental improvements over PEP-II:
  - Low emittance (small focus in the interaction point)
  - Crab waist technique
  - Low beam currents

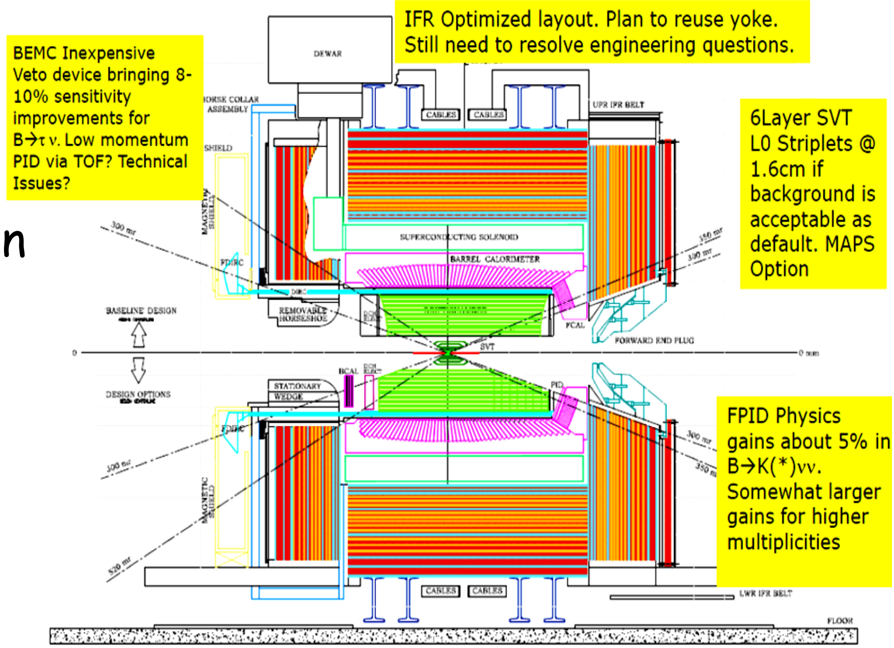
The Detector

Baseline reusing BaBar:

- Fused silica bars (DIRC)
- Barrel EMC CsI(Ti)
- Superconducting coil and flux return

Elements to be upgraded / replaced:

- Beam pipe, interaction region
- Silicon vertex detector
- Drift chamber
- Photon detectors for DIRC
- Forward EMC
- IFR scintillator
- Electronics, Trigger, DAQ



The Computing Model

- A lot of experience from BaBar computing
  - We can make a good guess of the requirements and resource needs of SuperB
  - Data Flow, (Re-)Processing, Skims, analysis, etc.
  - In fact, BaBar has provided the code base to SuperB
- It's a starting point for the SuperB Computing Model, but we expect major challenges and evolution over the next few years
  - Framework & Code
    - Multi/ManyCores + GPUs
    - Adopt existing frameworks?
  - Storage and Data Access
    - Parallel / cluster file systems, Hadoop-like FS? …?
    - Databases?
  - Distributed Computing
    - Grid computing? It's the baseline.
    - Cloud technologies applicable on SuperB timescales?
    - Funding will most likely drive us to having ~5-10 mid-size data centers, mostly in Europe
      - Flat hierarchy. No or a distributed "Tier-0"? Peer-to-peer topology of Tier-1 centers?
      - Group some smaller regional centers (in Italy) into "Virtual Tier-1" (common management)
- SuperB Computing TDR in 2013

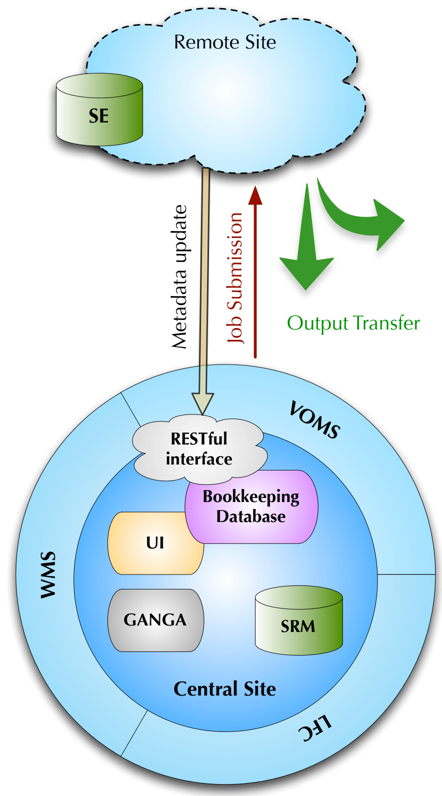
SLAC NATIONAL ACCELERATOR LABORATORY

The Data

- Computing requirements (full lumi, steady-state, **storage includes replication**):
  - Raw data
    - 25kHz x 200kByte = 5 Gbyte/s
    - ~160 PByte/year
  - Disk storage growth ~ 10-20PByte/year
  - Tape storage growth ~ 200 Pbyte/year
  - CPU growth ~ 2 MHEPSpec06 / year
    - 1 reprocessing cycle per year

Distributed Computing System (1)

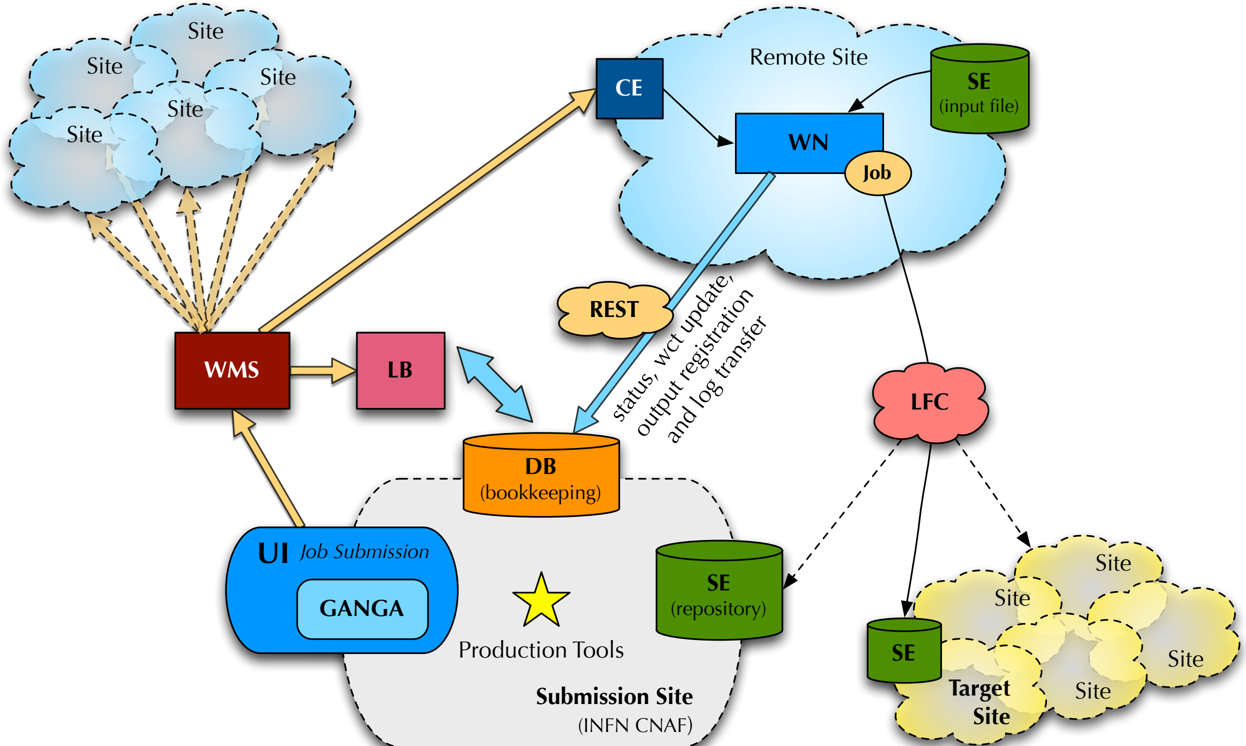
- GANGA system performs job submission from CNAF UI to sites
- Direct submission via WMS to site CEs
- Job run time tasks
  - Access DB for initialization and status update via REST
  - Retrieve/access input files from local Storage Element
  - Transfer the output to SE at CNAF or other target site
- Prototype has provided SuperB with about 170 CPU-years in 2010 to produce ~10^{10} FastSim events



	Sept. '09	Feb. '10	Jul. '10
Analysis stream	2	5	6
job done, failure rate	5K, 10%	20K, 8%	160K, 10%
Number of event	2.25 x 10^8	1.6 x 10^9	8.6 x 10^9
Involved site	1	9	15
WallClockTime	6 years	19 years	150 years
Disk occupancy (TB)	0.5	5	25
Peak job running	500	2500	7000

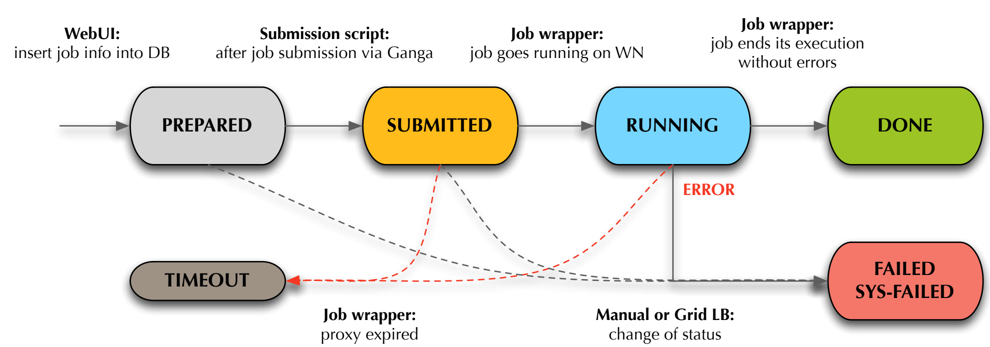
NATIONAL ACCELERATOR LABORATORY

Distributed Computing System (2)



INFN CNAF in Bologna is our central site:

- Job submission
- Bookkeeping DB
- File catalog



Sites

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, P. Franchini
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	L. Fano'
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	S. Lo, F. Porter, P. Ongmongkolkul
Fnal*	50	400	1	dCache	OSG	M. Slyz
OhioSC*	?	?	?	dCache	OSG	R. Andreassen, D. Johnson
Victoria	50	100	5	dCache	EGI	A. Agarwal
McGill*	100	200	1	StoRM	EGI	S. Robertson, S.K. Nderitu
Cyfronet	100	500	10	DPM	EGI	L. Flis, T. Szepienie, J. Chwastowski
Total	3570	11510	440			

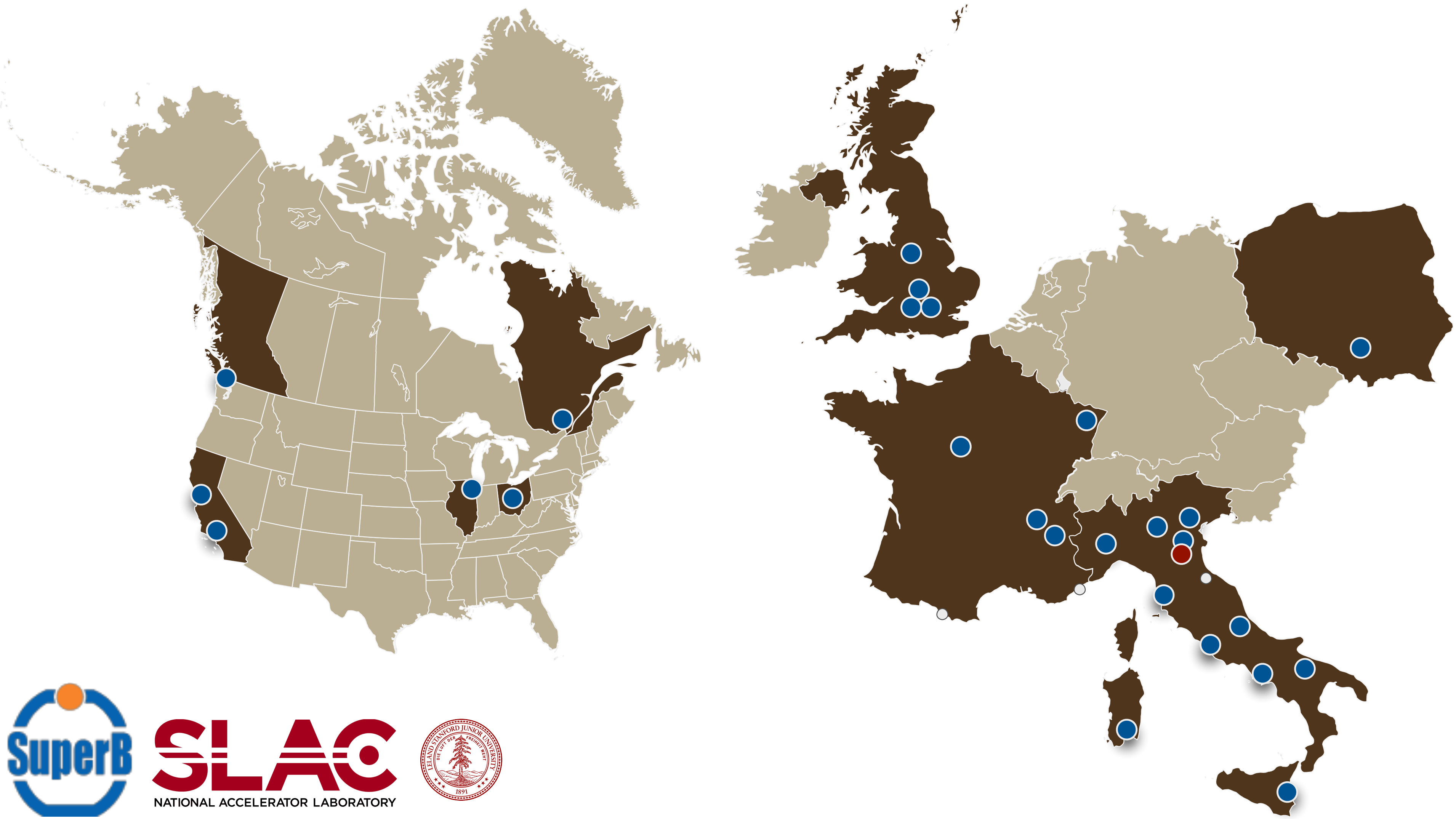
* VO enabling procedure in progress

~3000 - ~10000 cores!
- In a mix of temporary and permanent allocations

Predominantly EGI

Current OSG sites with SuperB resources:
- SLAC
- Caltech
- OSC (in progress)
- FNAL (in progress)

Geographical Locations of our Sites



SuperB

SLAC NATIONAL ACCELERATOR LABORATORY

Current Computing Activities

- Main activities
  - Support tools and computing to study beam and detector for the upcoming Technical Design Report and beyond
    - Full Sim (GEANT4) and FastSim
    - All Grid-Based (see sites on next slide)
    - We are interested in using OSG opportunistic cycles for our production peaks.
  - Computing R&D for the future
    - What do we need to do to stay on Moore's law?
      - Multi-/Many-cores, GPU
    - Storage (internet-scale technologies)
    - Distributed computing (Grid / Cloud / … ?)
- SuperB has inherited the BaBar software
  - Major rewrites expected per outcome of the R&D
    - Framework, distributed computing, etc., etc.

SuperB SLAC NATIONAL ACCELERATOR LABORATORY

SuperB and OSG

- We are an "OSG VO" (superbvo.org)
- Resource allocations at SLAC and Caltech
- Collaboration with OSG support group has been excellent
  - SuperB requirements have been mapped on OSG general services
  - VO has been enabled for simulation production at SLAC, FNAL and Caltech (WIP at OSC)
- We will run a simulation production campaign in September

Future Needs

- SuperB will be an exciting experiment but a relatively small collaboration
- Use existing tools (middleware, data distribution, etc.) as much as possible
- Very interested in lightweight tools
  - Not so much in "big solutions"
- Very interested in technical collaborations and sharing of experience in our R&D areas
  - Frameworks, storage, distributed computing
- On the practical side
  - If we get really serious about OSG use, we would probably want a VOMS replica in OSG-land.