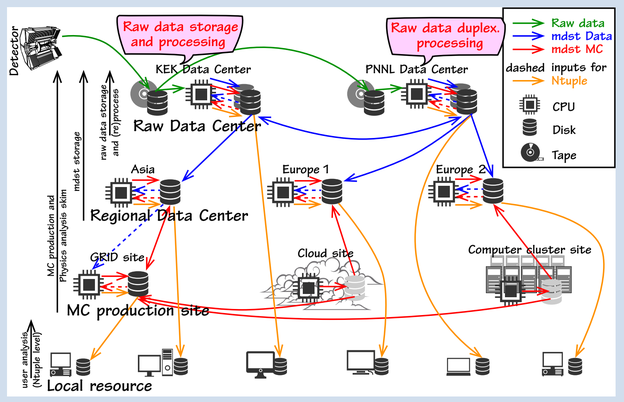
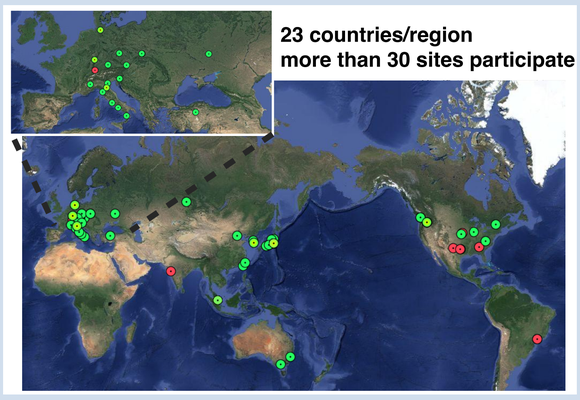
The Belle II Analysis on the Grid

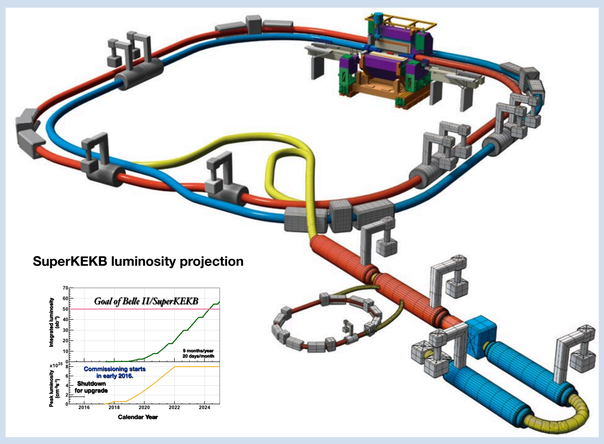
C.-L. Hsu, M. Sevior
on behalf of the Belle II distributed computing Group

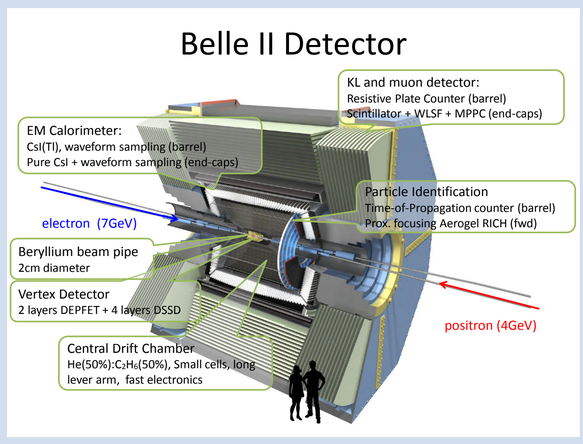
SuperKEKB and Belle II

- Next generation flavour physics experiment using the most intense beams of electrons and positrons ever produced
- **40x** higher peak luminosity than KEKB
  50x larger data size than Belle
- Aims to reveal new physics through precision studies of rare or suppressed decays
- Size of data sample is comparable to LHC experiments - to process and analyze this huge data size is a challenge for computing









Physics Goals of Belle II

The unique capabilities of Belle II include

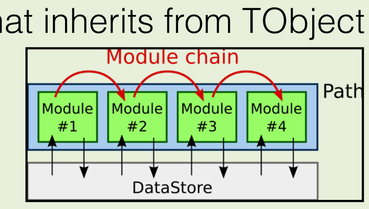
- Exactly on-resonance BB meson pairs production
- Higher flavour tag efficiency (10x LHC)
- Detection of photon, π^0, ρ^0, η^0, K_L: complete strong phase surveys
- Clean background environment

With the above features, Belle II is a perfect tool to search for the decays which are sensitive to new physics beyond the Standard Model.
Skim production helps to reduce the time consumed for user analyses.

Analysis Framework

Belle II **A**nalysi**S** **F**ramework (basf2) is a software framework developed for the use in Belle II

- All algorithms are modules
- DataStore manages all data loaded or created during event processing
- Can store any class that inherits from TObject and has a ROOT dictionary



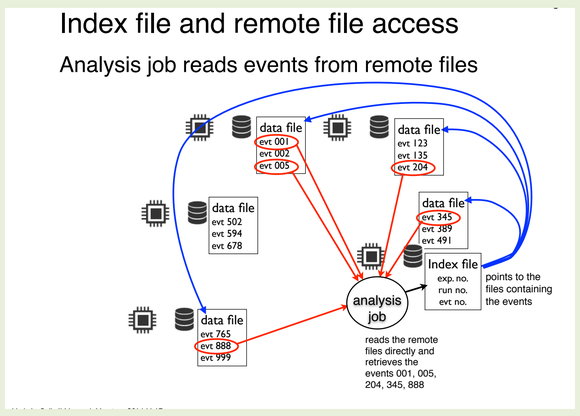
Physics Skim

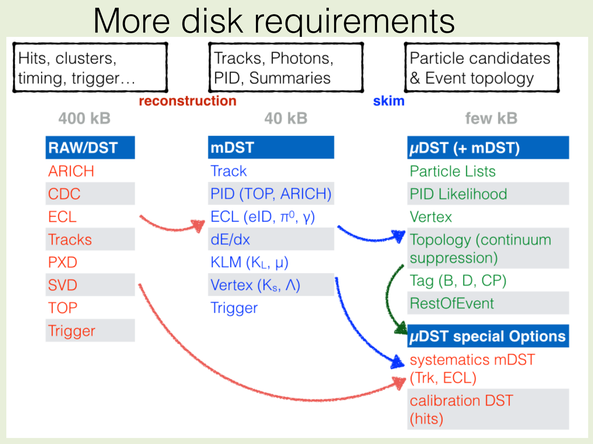
Selecting interesting events after mDST reconstruction

- Centralized reconstruction of Particle candidates
- Prepare skim group samples for different user analyses
  - Hadronic fully reconstructed B_{tag}:
    Missing energy, semileptonic decays
  - J/Ψ, D,D_s,Λ_c
    CPV in B and D systems…
- Produce multiple skim samples at once

Investigate cluster, grid and xrootd for data access

- Possible formats:
  - **index**
    Event pointers
    Small file size
    I/O intensive
  - **mDST**
  - ***μ*DST** (derived data)
    Contains persistent particle lists
    More disk requirements





Belle II Computing System

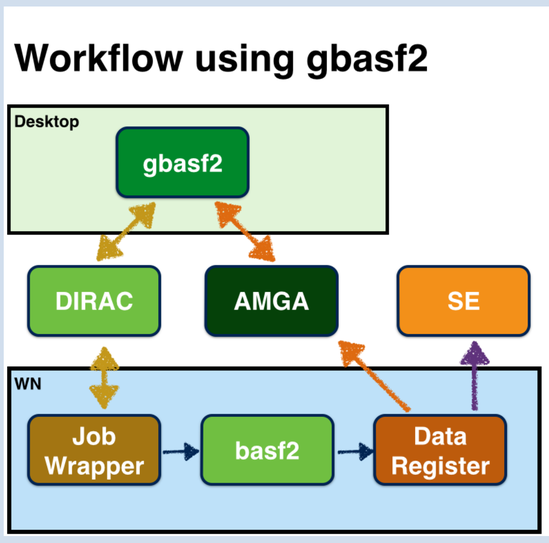
Based on existing solutions + extensions for Belle II

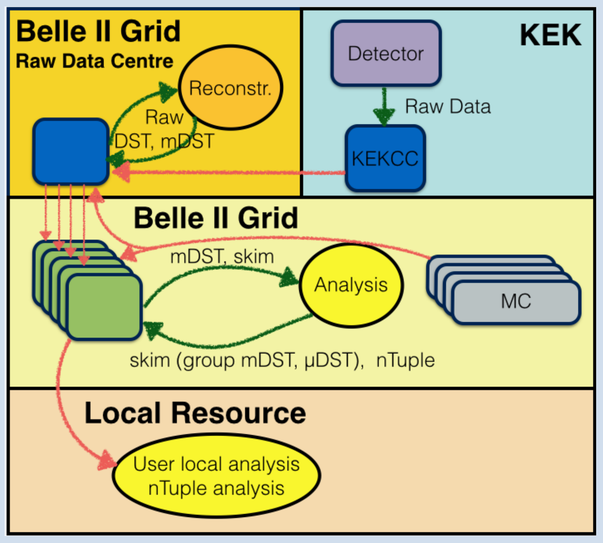
- DIRAC for job management, also provides a common UI for all sites
- AMGA for metadata
- CVMFS for software distribution
- gbasf2 for GRID user interface for basf2, provide communication with DIRAC and AMGA via API

MC production campaigns

To test Belle II software on grid and the grid design, and also to provide the official MC samples for the physics analyses

- Performed 4 MC campaigns
  - Bottlenecks and issues were recognized
  - Analysis demonstrations on100 fb-1 of Generic MC (B decays via b→c transition and e^+e^-→ qq,q = u,d,s,c) w/ & w/o beam background (MC3, MC4)
    - Test analysis tools, data flow





Analysis on the Grid

Demonstrate skim method on (g)basf2 analysis, and skim production with μDST format:

- Example analysis: radiative decay **$B^0 \to \rho^0(\pi^+\pi^-)\gamma$**
  - Analysis time is **5x** faster than post-skim analysis
  - File size increase
    - 1.3% accept fraction for hardon triggered
    - 1.3 candidate multiplicity
    - Estimate size

	Size/Event(kB)
mDST	40
Final state particle (π, γ)	0.2
Combined particle (ρ, B)	~1
Total	**41.2**

Trigger Skims	pB/ab^{-1}	Skim Group (Grid)	Retention	pB/ab^{-1}	User Physics Skims (Local)
Hadron Trigger	1	*B_{tag} (>1000)*	0.05	0.06	**V_{ub}, B→τ/μ/e ν, B→D*τ ν**
		Radiative (>50 exclusives)	0.03	0.04	**B→$X_{s,d}$ γ, B→K^*/ρ γ**
		J/ψ	0.02	0.02	**Charmonium, B CPV**
		Charm (D,D_s,Λ_c) (>100)	0.05	0.06	**D Mixing, D CPV**
		>50× Exclusive	0.20	0.22	**Many**
Total	**1**	**0.4 (0.02-0.06 PB per skim group)**			

- The goal of physics skim is to produce multiple skims in one processing pass. To achieve this, need to keep only particles used in skim candidates.