Longterm Development of SuperB Offline

David Brown

SuperB Offline Development

- Current Focus: facilitate TDR studies
- Must transition focus soon to long-term issues
  - computing TDR
- Development research must be tightly focused
  - comparisons between a few options
  - realistic evaluations based on BaBar use cases
- New manpower is required
  - CSG should organize well-defined research projects for new groups joining SuperB
  - code cleanup should be farmed out (undergrads?)

Major longterm comp. issues

- Multi-core processors (parallelism)
- Leveraging the BaBar code base
  - fix quality, performance, + maintenance issues
  - refine, reverse-engineer, or rewrite?
- Integration of external developments (LHC)
- Data storage
  - access model (grid, cloud, hierarchical tiers, ...)
  - Long-term data access (+ outreach)
- Reduced manpower (compared to BaBar)
  - simplification + consolidation
  - integration of validation and performance testing

Possible Guidelines

- Open-source tools
- Industry-standard solutions wherever possible
- gcc
- Linux+1 (Mac OS X?)
- Make major migrations early
  - Framework, persistence, analysis interface
- Try to avoid a ‘rewrite’ (CM2)
  - Make choices now looking to operation in 2020
- treat major development projects as branches
  - Integrate as necessary

Simulation

- Generators
  - EvtGen replacement
    - Code is unmaintainable, unsupported
    - switch to HepMC?
  - Common output format
    - LTDA requirement to separate from Det. Sim.
    - FastSim example
- Digitization code
  - Common framework for subsystem contributions
- Luminosity bookkeeping in datastream

Framework

- Use existing LHC frameworks?
  - ties to LHC persistence, build system
  - multi-core compatibility?
  - Coordination/conflict with LHC development
    - Branch off now or later?
- Dynamic or static load?
- Scripting language
  - Python
- Re-evaluate state diagram
  - Make beginJob a real state?

Analysis Interface

- Protect users from memory management
  - Separate ‘sandbox’ functionality from event data
- Replace candidate implementation
  - separate bases for composites, det. based, MC, ...
  - explicit MC truth, matching interface
  - flexible ‘overlaps’ function
- Replace vertexing interface
  - clear separation of operator and data objects
  - replace inefficient 6X6 low-level implementation
- Replace quals with direct reco data access
- User code migration

Performance Optimization

- Parallelize where possible
  - multicore compatibility
- Rewrite/replace critical algorithms
  - POCA
  - calibration
  - matrix manipulation
  - geometry functions
- replace string comparisons with hashes (PDT)
- Performance monitoring
- Rewrite high-level reconstruction
  - DIRC, IFR, EMC, track-finding, ...

Persistence

- underlying technology
  - root, mdf5, ...
  - compatibility with LTDA, framework
- Data model
  - hierarchical structure as in BaBar?
  - integration of ‘analysis’ tuples?
- Compatibility with data distribution model
  - peer-to-peer or centralized?

QA

- Code cleanup
  - fix all compile warnings
  - fix all memory leaks
- Install unit tests as part of every package
  - flexible definition of success
- Nightly build of the trunk
  - find errors early
  - run unit tests, standard validation
- Release validation

Release Infrastructure

- Release system + scripts
- Build system
  - dependency management and monitoring
  - separate library granularity from packages
- Packaging
  - sub-directories for source, include, test, script, ...
- Amalgamation of repositories
  - framework, detector code, release, tools
- GIT for repository?
  - peer-to-peer vs client-server model
  - SVN interface for backwards-compatibility

Adaptive algorithms

- Useful where combinatorics are large and linear algorithms fail to find global optimum
  - pat. rec.
- Can be integrated with multi-core
  - many solutions explored in parallel
- Promise real improvement in physics output
  - track finding
  - clustering
  - BReco

Conclusions

- Lots of work is needed for SuperB offline
- clearly-defined projects will help attract new manpower