BLAST Database Project

Beam Lattice Alignment Survey Tracker Database

Status and Plans

Dennis Box CD-CEPA-DBS

September 21 2004

Background

A Three Phase Plan was proposed to modernize Accelerator Division Magnet Properties/Location Databases and make them generally useful:

(Phase 0)

(Have a modern environment to move to)

Phase 1:

Port AD database from VAX to more modern architecture

Phase 2:

Modernize, clean up schema

Phase 3:

Integrate new data sources from other departments.

Phase 0:

Setup a suitable production environment

- Prerequisite to proceeding with 3 phase plan

- Goal is to have AD own the db server and support it.

- Currently: Sybase db is served by the warm backup machine for Control Room.

- Oustanding Issues if this is to be considered a production environment:
    - Backups
    - Web Server

- This took much longer to implement than expected.
    The remaining problems are more political than technical

Phase 1:

All of the phase 1 deliverables have been done:

1. The database environment has been set up.

2. Data from the RIM database of MTF data is now in this modern Sybase in BLASTMAN DB database. The schema has been ported to SQL, and documented by a data dictionary.

3. Validation methods demonstrate the equivalence of data from the VAX with data from the BLASTMAN DB.

4. N. Gelfand has accepted procedures for updating the data and updating the schema.

5. N. Gelfand has "signed off" on the usefulness of this database for the MTF data which formerly existed only on his RIM database. He has been using the BLASTMAN DB for his applications, and has not had to touch the VAX for data access in months. He confirms that this meets the "if the VAX got hit by a bus" criterion.

6. The main Metrology data from survey completed last fall is also loaded and validated.

7. Magnet evaluations data and sgrade codes are in the database.

This list meets the checklist for success of phase 1, presented in the original project plan.

Phase 2:

Some tasks completed

-examples: constraints, id of redundant tables

some obviously right to do

- Better automation of data update entry for the AMG data.

some that turn out to be unneeded

-complete normalization of all tables

some in the grey area and need to be discussed

-how far do we go with table refactoring to acheive maximum cost/benefit?

Phase 2:

The agreed steps now constituting phase 2 are:

1. Better automation of data update entry for the AMG data.

2. Improved descriptive attribute names (these should be based on input from AD users).

3. Validation of location information.
   -automated checks that AMG, TD, AD inputs agree

4. Cleanup of database for better long-term maintenance and easier application creation.
   a) Normalize and optimize the schema, if warranted.
   b) Assuming the schema is modified, handle in a sound manner cases of "data not available," and allow for variable numbers of harmonic coefficients and other constants per magnet.
   c) Create appropriate views and migrate applications to smooth the transition to the normalized schema and attribute names.

5. Capability for handling magnet field reshims

Phase 3:

Already done: TD mechanical inventory data

TBD: include AMG Data

- Use Cases for Data Flow exist
- Schema has been prototyped and discussed

TBD: Interface to CHEF format

- successor to MAD

Types of Data potentially useful in BLAST DB

Fact Data (location, inventory,status)

- some coming from TD Now

Alignment Data

- AMG group provides this

Field Data

- AD owns generated originally by MTF

NOT CANDIDATES FOR BLASTDB:

Lattice Descriptions

- stored in CVS by AD

FESS data

- buildings, pipelines

Settings and Readings

- kept in Control Room DB

AMG Data Expansion

Supporting information

reference sites, monuments, murphy plugs

Carefully selected extensions

Expand to other machines

Booster, MI, Accumulator....

AMG will want to use this by Oct 2005 shutdown

Standardize access interface for 'all' the machines

PPD participation is key

Lessons Learned

-Working with more than 2 divisions at a time can be hard

Example: Phase 1 complete before phase 0 (!)
due to disagreements about what a production DB server is.

Agreements between members of different divisions commonly fail when

-members lack authority/persuasiveness to implement agreement

-different organizational units may have different priorities

- unstated assumptions turn out to be wrong

In this case these issues caused delays but didnt significantly
increase total amount of work

-Accurate Effort/Schedule Estimation is Hard

Long Range Support

- Path (and AD personnel) have been identified such that
  CD will not be long term supporter of this database

- Capability and knowlege transfer will happen

- Development/Extensions (for now) being done by DBox
  In the long run (post phase III) will be transitioned to AD/PPD

Phase 2 Estimates

1. Better automation of data update entry for the AMG data.
   - -get an agreed data format (??? days)
   - -implement it (3-5 days)
2. Improved descriptive attribute names based on input from AD users.
   - -get input (??? days)
   - -implement it (3-5 days)

3. Validation of location information.
   - -automated checks that AMG, TD, AD inputs agree
   - -get an agreed data format (??? days)
   - -implement it (already done for existing data)
4. Cleanup of database for better long-term maintenance and easier application creation.
   - a) Normalize and optimize the schema, if warranted.
     - -get consensus as to extent of normalization (??days)
   - b) Assuming the schema is modified, handle in a sound manner cases of "data not available," and allow for variable numbersof harmonic coefficients (and other items) per magnet.
   - c) Create appropriate views and migrate applications to smooth the transition to the normalized schema and attribute names.
     - (2 week FTE)

5. Handling Magnet Field ReShim
   - -understand data format and user needs (??? days)
   - -implement tables and accesss application (1 week)

Phase 3 Estimates

These are being formulated.

AMG Schema has been prototyped

AMG Use cases have been proposed

AMG data extension appears to be a tractable, defined problem

CHEF interface

- Unknown difficulty
- MAD interfaces exist and were not difficult
- if we can solve this problem the same way then its ~2 week
- If we need to provide API/library then it's longer

Conclusions

Phase 1 completed successfully

- AD people use it
- Others beginning to recognize utility

Phase 2 defined, partially completed

Phase 3 dominated by AMG data for varied machines

Major progress towards Long Term Support by AD

This Project is an Important Contribution to the Lab's Mission