CDF Databases: Summing it Up

Rick St. Denis
Glasgow University
August 30, 2001

"If your end users are surgeons or nuclear physicists, they may have more urgent requirements"[1]

Outline

- Organization
- Achievements
- Problems
- Plans
- Manpower
- Ruth's 2000 Quiz

[1] *Oracle 24x7* V.S. Devraj, Osborne/McGraw-Hill, 2000, page 186

Organization: Three Tiers of Management

Oracle DBA/Applications

CDF Oracles

C++ Application code writer

C++ Application code user: Physicists

$\Longrightarrow$ Three Tiers of Code

Organization: Roles of the Tiers

- Oracle DBA/Applications isolate most users from SQL and the vagaries of keeping Oracle running (emergencies, backups, installation)
  They speak SQL$\longrightarrow$ 5 people

- CDF Oracles: link DBA and Physicist
  They speak SQL and C++ $\longrightarrow$ 7 people.

- Application writers interested in USING calibration, just want containers stored/filled
  Speak C++, order breakfast in SQL $\longrightarrow$20 people.

- Physicists just want to know a module is getting the right calibration
  They speak $E = mc^2$ and make a dog's dinner of C++ $\longrightarrow$ 500 people.

Organization: Who's Who

Schema	Physicist(s)	dBa
Hardware	Bill Badgett	Nelly/Julie
RunDB	Bill Badgett	Nelly/Julie
Trigger	Alexei Varganov	Nelly/Julie
	Kirsten Tollefson	Nelly/Julie
Calib	Jack Cranshaw	Nelly/Julie
FileCat	Dmitri Litvintsev	Nelly/Julie
MCS	(Jack Cranshaw)	M. Vittone
	(Larry Kirsch)	D. Box
	Steve Hahn	Nelly/Julie
(Module)		S.Jones
(Matrix)	Sidet	D.Ritchie

- GoodRun needs Czar (in RunCond)
- Randy Herber has excellent understanding of entire database.
- Loosing Kirsten soon, Jack next year.

Summary of the Year's Achievements

- Extensive operational experience gained during Commissioning run.
- Machines in Final Configuration (but disk)
- $> 99.9\%$ uptime (offline)
- Oracle Software Version Cuts allowed
- Calibrations routine (consumers)
- Applications programmers producing tables, calibration and analysis.
- Schema evolution routine
- Databases operational: Run Conditions, Hardware, Data File Catalog, Calibration, Trigger

- MCS developed and close to operational
- Database Browser an Essential tool widely used
- Replication functioning since Nov 2000 (RC, HwDB, Calib, Trig)
- Database Simulation to 4 mo. and database sizing estimated.
- OEM fuctioning (but Kerberos hinderences)
- API Code being informally reviewed and revised
- Export reviewed but progress stopped.
- Online behind a firewall.
- Very Few Emergencies/Pages: Root Cause analysis great

Achievements: Machines

- Got ultra5 for offline int/dev on Nov 12 2000
- Warehouse decommissioning started 2/19/01, done 5/1/01
- onprd had problems dec 29-jan 24: bad mid-board. Sun replaced cpu and midplane.

Achievements: Oracle upgrades

- 8.1.6.0 tested upgrade during commissioning run. Ran on integration
- Procedural mods to account for two instances on one machine
- 8.1.6.2 upgrade on 30-Nov-2000 offline
- 8.1.6.2 upgrade on 16-Dec-2000 online
- 8.1.7.0 terminal version 15-Feb-2001
- 8.1.7.1 this time for sure (not fully tested in integration run) 18-July-2001

Integration testing hurt by run numbers: cure worse than disease.

Achievements: Oracle upgrades

- Replication complicates this
- Need to drop create/public synonyms
- Need to drop create views
- Need to compile procedures
- Requires some number of Oracles.

Achievements: Consumers

- C++ templates developed and used widely
- Many small packages: responsibility devolved
- CalibDBTables now split into xxTables for code generation
- Lack a Calibration Coordinator
- DBANA supported (but thin) and used widely.

Achievements: Replication

- Takes 2 min typically
- Both on demand and with cron
- Full refresh: hours? With larger database?
- Went to having only an online db except dfc
- offline is a copy (except dfc)
- Monitoring needed:
  - - need to know when cron will fire, signal error if failed
  - - need to keep track of the time it takes - signal 5sigma effects.
- Need to keep track of time for a full refresh.

Achievements: Simulation and Sizing

- Calibdb
  - 75 tables
  - 4 months running
  - size is about 80G data: some large tables missing
  - learned we needed to vectorize
  - Need to study times and performance for various sized databases

- DFC
  - 175 datasets
  - 209853 filesets
  - 1.9 million files
  - 4.6 million run sections
  - time to update one run section 20-40 s and flat as a function of the number of

files and linear as number of run sections to update.

- 268 for 10. (so for 10^6 is one year?)
- This is a failure
- Designer gives correct estimate of sizes: 100% error on input though

Achievements: API

- Common misnomer: CODEGEN
- Spaghetti code /code bloat.
- 10% knowledge from each of 3 authors. Reverse documented/engineered by Dennis. Took time!
- Manager still has cut/past method for some cases.
- Proposed changes 1 year ago but still sitting (we did not even try to schedule them, but suffer)
- Lots of requests
- OCI gone, OTL used

- User perspective: works fine
- Support perspective: cut and paste extrodannaire
- What happened? CODEGEN was for Calibration schema only, but trying to extend requirements.
- Failed to convince collaboration to give us help/support: Fell on Dennis.
- Solution? Break up/clean up DBObjects, CalibDB, DBManager packages with clear lines of responsibility (ie. librarians). Meet next week with Paulo.

Larger question of C++ API access to all the database.

Achievements: Export

- Review recommended home brew solution: afraid of oracle admin @ small institute
- need design review
- need code review
- Implementation/testing stalled by UCL and hindered by Kerberos: Mark on leave
- Dave Waters and Richard Hughes evaluating code and situation

Status and Problems

Systems:

- Biggest problem: Online much better, Offline improving Online–Offline comminication better but only recently in offline

Status and Problems

Personnel:

- The CDF Oracles are Critical, Few, Diminishing and hard to train: need for Simulation, Growth.
- Trigger: Serious condition: reliant on two people, one fading.
  - L1/2:One talented but strong-willed 2nd year grad student
  - L3: Kirsten on the way out
- Data File Catalog: Even Dmitri is feeling powerless.
- The CDF Applications code (nb. calibration consumers, used sets): Leader: Online Cali-

bration Coordinator is STILL Jack and STILL he does the calibration schema.

- SVX, COT, muon, CLC, TOF are superb
- Strip chambers much improved
- Silicon struggling to existence

- Woefully short on Common Code writers Expertise! (nb test code)

- Slow Control only now beginning to come together (3rd leader)

- Alignment needs some advanced tools

- Online: Usage at the (b)leading edge, but *has pulled* it off. Reliant on one person: The BILL problem.

- Used and Valid Sets: Rob Snihur doing well, Jean Slaughter pioneering, it is coming together but shows cracks in a non-unified schema and lack of farms information *in Oracle*. Nb: What version was production run with and which constants were used?

- Export to Institutes did not happen. Now being put together.

Plans:Structural

- Code and Database Review and package breakup: DBObjects, CalibDB, DBManager (Jim/Dennis)
- Database Validation (Connected to responsibility and review)
- Database Simulation (Igor/Yuyi)
- Review 24/7 CDF notes, failover (ongoing) Nelly/Anil
- Backups: Situation still not satisfactory with respect to tape.

Plans: BigList Items

- Calib DB Schema (Jack/Dennis/Jim)
- Export (Dave Waters/Richard Hughes, Jim, Jack)
- “Books” for Data File Catalog (Dmitri)
- Sizing (Igor/Yuyi)
- File Catalog lacks functionality: 1 yr to update Lumi
- Bugs
  - Fixing taking 30-40% of Dennis’ time
  - New usage patterns reveal new rewrites and create another plague

We still need more than 2 Dennis’ and 2 Jim K’s.

Manpower

- **Database:**
  - Need 1.5 dBa FTE for design reviews, cuts, operations.
  - Also have 0.5 of Badgett as Online DBA.

  The combination of Nellie, Anil, John, Julie and Diana remains effective and are at the right level of support. BUT need more of Julie plus Randy to get the unification of the schema to work.

- **API**
  - Dennis Box (80%) and Yuyi Guo (80%)
  - Jim Kowalkowski (40% since May 2001)

  As use Cases develop, this is woefully inadequate.

Ruth's Quiz

1. **Crest of Wave**
   - Each phase
   - Run Startup
   - Development support
   - Usage patterns, Unification patterns

2. **Experience**
   - simulate (needs development and DBA manpower)
   - consultancy may help (Buy experience)
   - Real problem is now API.

3. **Backup** Commissioning, validation, MDBC

4. **Turnover** Schema Team, DBA Team

5. **Offsite** Replication helps 24/7 (90/10)
Pacific Rim, Europe, America
Freeware

6. **Sizes**

HdwDb	10 GByte	ok
RunDb		
Calib	20-2200 Mrow	?
	(.06-6.6 Gbyte)	160 G
Trigger	10 Gbyte	ok
MCS	60 GByte	5 G
	(Unfiltered)	(filtered)
File Catalog	10 GByte	35 G: books!

Kevin Mc Farland "We know the order of magnitude of the log of the silicon needs"
Still True: 10 G table, need 10 G indexes, need 80G temp space: 100 G!!!

7. **Transaction Rates**

- Online: 1Mbyte in calib run, also Begin Run. Should measure

- Offline: MDC2 results, need to add user load, need more use! 1000 connections/hour

8. **Simultaneous Downtime** Like backup question Schedule offline Thursday, Online during shutdowns.

9. **Online Share** See Umesh Bill

10. **Schema and Responsibility** See slide

11. **MOU** Essential to have Commissioning Run to Determine Have draft, By librarian

Summary/Conclusions

- It seems to work.
- Much Progress: Need to unify schema (DFC to source machine) and address CDF Oracle problem.
- We need to keep on top of use patterns and identify warehousing needs vs replication.
- Need to tune, understand performance.
- Need export.

Continuing Deep Thanks and Compliments

to ODSDBA for their help!