DATA WAREHOUSE ON THE WEB FOR ACCELERATOR FABRICATION AND MAINTENANCE

A. Chan, G. Crane, I. MacGregor and S. Meyer

Stanford Linear Accelerator Center*, Stanford, CA 94309 U.S.A.

Abstract

A data warehouse grew out of the needs for a view of accelerator information from a lab-wide or project-wide standpoint (often needing off-site data access for the multi-lab PEP-II collaborators). A World Wide Web interface is used to link legacy database systems of the various labs and departments related to the PEP-II Accelerator. In this paper, we describe how links are made via the 'Formal Device Name' field(s) in the disparate databases. We also describe the functionality of a data warehouse in an accelerator environment. One can pick devices from the PEP-II Component List and find the actual components filling the functional slots, any calibration measurements, fabrication history, associated cables and modules, and operational maintenance records for the components. Information on inventory, drawings, publications, and purchasing history are also part of the PEP-II Database [1]. A strategy of relying on a small team, and of linking existing databases rather than rebuilding systems is outlined.

1. DATA WAREHOUSE—DO WE NEED IT?

In an accelerator laboratory we are inundated with data related to the fabrication and maintenance of the machine, such as hardware, personnel, finance [2]—yet it is more often than not quite difficult to obtain such related information in a useful format. To resolve a maintenance problem, one may need to log onto diverse computer systems and run different database software, frequently devoting much time to trying out obscure search strings.

Such compartmentalization of databases of the various accelerator departments evolved because different software was used to suit particular problems and situations. Although there is growing recognition to achieve database implementation under the same database software, we will still be faced to some degree with these differences. [3]

Another reason for this compartmentalization is that databases developed by the departments, of necessity, focus on their area of business functions. This short-term focus can be mitigated by recognizing the need to make linkages to the overall accelerator during the design of the database. However, the implementation of a data warehouse is required for a lab-wide view of decision support information for the accelerator.

2. DATA WAREHOUSE—FOR AN ACCELERATOR ENVIRONMENT

A data warehouse can be simply defined as [4] "a single, complete, and consistent store of data obtained from a variety of sources and made available to end users in a way they can understand and use in a business context."

There are 3 main architectures for a data warehouse (see Fig. 1), although in reality what is implemented is most likely a mixture of these different types. The data layers that Fig. 1 refers to are conceptual, rather than physical.

For the single-layer architecture, its strength is that data is only stored once, which avoids the need to synchronize multiple copies of data. Its weakness is that contention can occur between the online transaction processing systems and the decision support systems.

For the two-layer architecture, its strengths lie in solving the contention between the online transaction processing systems and the decision support systems. It addresses the fact that end-user needs for information are different from what is easily available from real-time data.

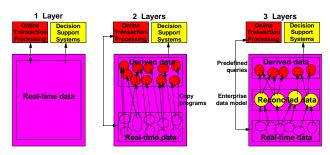


Figure 1 -- Types of implementation of data warehouse

^{*}Supported by the U.S. Department of Energy under contract DE-AC03-76SF00515.

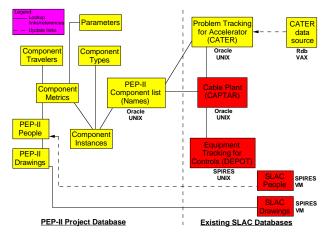


Figure 2 -- Linked SLAC accelerator databases

However, there is a high level of data duplication in the two-layer approach (with a tendency to become 'spaghetti code').

For the three-layer architecture, its strengths lie in the reconciled layer which is based on enterprise data modeling (i.e., a normalized database). This reconciled layer can support new, unanticipated end-user needs. The derived layer can be used to fill most end user needs, being the equivalent of predefined queries. The enterprise data modeling is a much more committed effort, and needs to be done incrementally.

At SLAC, we have mostly implemented the single layer and two-layer approaches.

The World Wide Web (WWW) technology has made it possible to access and link disparate databases, allowing the data warehouse to be a reality for us, and at a fraction of the programming effort in today's resourcescarce environment.

3. LINKING THE LAB DATABASES

Figure 2 is an overview of the lab's major accelerator-related databases that we have linked for the accelerator data warehouse. These systems are:

- the PEP-II Component system, containing the formal device names, component fabrication history and calibrations. This is integrated into the PEP-II project-wide database which is in Oracle, most of it developed using Oracle*CASE tools
- the problem tracking database for the accelerator (CATER), which is in VAX Rdb and replicated into Oracle tables
- the cable plant database (CAPTAR) in Oracle
- the equipment tracking database for Controls (DEPOT) in SPIRES

Via WWW Common Gateway Interface (CGI) scripts, these databases are linked together using common data elements (i.e., variations of formal device names shown in Table 1). The WWW interface gives excellent 'drill-down' capabilities. Since linkages of the

	0		
PEP-II Component	Primary	Micro	Unit
List	QUAD	PR02	6072
Problem Tracking for	Area	Micro	Primary
Accelerator (CATER)			Unit
	HER	PR02	(not a
	LER		required
			field)
Cable Plant	System F	unction	
(CAPTAR)	QUAD:F	PR02,60	72
	Micro	Crate	Model
	LI09	CR02	233-002-00
Equipment Tracking	Location		Model
	1		
for Controls (DEPOT)	LI09/CR	02	233-002

Table 1.	Linkage	of data	elements.
rable r.	Linkage	or unit	cientento.

/vhat's I	Newl		What's Cooll	Handbook	Ne	et Search	Net E	irectory	Soft	ware					
			PEPII HER F	ormal Device	Nar	nes With F Yields 18	~		r "1" a	nd Cell Nu	mber ''12'				
Region No.	Cell No.		Component String	Primary Micro	imary Micro Unit 1		Serial No.		Len. [m]	IP Dist. [m]	PEP Station No. [Ft.]	Note			
1	12	12	12	VA	AC. PUMP	VACP-PR02-6071				Α		2060.855	3552	LIP-60 liters/sec.	
1	12	QD		QUAD-PR02-6072		PEP000398	124	Α	.56	2061.638	3555				
1	12	BI	PMV BPMS-PR02-6072					A		2062.058	3556				
1	12		PRIMARY = (RIMARY = OUAD			OUAD								
1 12 B			•												
1	12	D	MICRO = P				PR02 6072								
		Ļ	UNIT = 6									_			
			CATER CAPTA CAPTA Raft Ph Travele Polynoi	MI: AR - D oto - D er-Data -D er-Images - D	CRO: isplay isplay Barco isplay	de 398, Seria	ormatio graph of pertain al Numb ages of	n pertaini QUAD: ng to me er 124. travelers	ing to QI PR02:607 asurement for Barc	JAD:PR02:6 2. nt, manufactu	072. are and/or re				

Figure 3 -- WWW screens of PEP-II Components List, and menu choices for each formal device identified by 'Primary, Micro, Unit'

different databases are done by the CGI scripts, huge savings in programming time are realized as we do not have to create physical tables to join the databases.

Users can query the PEP-II Component List on WWW, and the search results will contain hypertext references that request information from the other database systems. In Fig. 3, among the search results returned for components in Region 1 Cell 12 is the Primary/Micro/Unit QUAD PR02 6072 (the three fields comprising the formal device name), which has a hypertext link that produces a menu with additional hypertext links to:

⇒ the CATER database summary and detail screens for Area HER or LER, Micro PR02 (Fig. 4).

						. 1		. 1]			
What's N	lew! What's Co		Hand	book	NetSei	arch	Net Dire	sctory	Software			
Line M	Aode Proble	m R	epo	rt Sy	nopsis							
Search return	o ed 3 resolts											
view soluti	ions											
Num	Rep.Date	T P	S	By	Area	Mic	r Sys	Un	it	Fac	Description	
47395 48491 48653	04/04/97	H L H S H I	C	CJB MIL SSM	HER HER HER	PRO PRO PRO	2 CAN	A CR	AT 3 AT 5 DU 5		Crates 3 and 4 ar PEP camac crate 5 Missing fiducials	in PRO2 fails verify
Detail	Problem R	еро	rt I	nfor	matio	n						
Number		48	1491		T	тре			Hardwar	e	Solutions	1
Urgency	rgency Scheduled			St	atus			Closed		Ready-to-Close	No	
Report	by	М	ILLE	R	P/	UL M	ILLER				Report Date	04-APR-1997 19:16
Modify 1	by	J	₩H		Ru	sty Hu	mphrey				Modify Date	05-APR-1997 08:41
Assign t	0	R	JJ		RA	ALPH.	JACOB	IR.			Finish Date	
Closed t	by	R	JJ		RA	ALPH.	JACOB.	IR.			Close Date	18-APR-1997 07:34
Fix hour	r s	0			Di	isposi	tion				Class	
Beam Lo	ost	0			D	ocume	ents		Unknow	n		
Cow/Cal	lf/₩S				A	rea			HER		Shop Main	CTL
Facility					M	icro			PR02		Shop Alt	
Display					Pr	im. &	Unit		CRAT 5			
Reprodu	uce				St	ıbsyst	em		CAMAC			
Track m	sg											
PEP cama	Description ac crate 5 in PR02 f	ails ve	erify'.	.fails co	umand lin	ue tests	e -					
New Sear	rch]											
Solution	1 Information											
Solved t	b y :					RJJ			RALPH J	ACOBJR		
Hours to	o Solve:					0			Date Sol	ved:		18-APR-1997 07:34
Solution	Туре:					Repla	ace		Problem	Number		48491
	ze Module:					SCC			Colorado -	Number		1

Figure 4 -- Summary and detail WWW screens of accelerator problem report database (CATER) linked from 'Area HER, Micro PR02' in figure 3

- ⇒ the CAPTAR database of cables linked to QUAD PR02 6072. From the Micro field, one can obtain crate profiles which are database reports in WWW tables (Fig. 5)
- ⇒ the DEPOT database with summary screen of the history of modules in this crate location, and detail screens of the maintenance records of each module is linked from the CAPTAR crate profile. In Fig. 5, this link is shown for Micro LI09 Crate CR02 Model 233-002.
- ⇒ the scanned paper fabrication and alignment travelers, online calibration measurements, drawings, photos, polynomials files, etc.

		Location: Micro: Assemble	KF09-2/ LI09-CI		IAC CRATE		
Slot	Status	SLAC Number	:	Module Na	me	Date Entered	
1	0	123-589-00	CRATE V	ERIFIER		27-SEP-90	
2	0	123-624-00	PIOP	PIOP			
3	0	123-624-00	PIOP	PIOP			
		100 204 00	DIAD			APD OD	
20	0	123-976-00	BEAM PO	DSITION MONI	TOR	27-SEP-90	
21	0	233-001-00	STB II			10-SEP-90	
22	0	233-002-00	PDU II			27-SEP-90	
		137-037-00		CONTROLLER			

SPIRES DEPOT Database Search results

Printed: 05/07/97											
ID	Nickname	Make	Model	Rev							
18005113	PDUII-F	SLAC	233-002	F							
18003950	PDUII-E	SLAC	233-002	Е							
18003562	PDUII-F	SLAC	233-002	F							
16030282	PDUII-F	SLAC	233-002	F							
16001539	PDUII-F	SLAC	233-002	F							
16000704	PDUII-D	SLAC	233-002	D							
16000698	PDUII-F	SLAC	233-002	F							
16000667	PDUII-F	SLAC	233-002	F							
16000672	PDUII-F	SLAC	233-002	F							

Figure 5 -- WWW screen of crate LI09/CR02 -- crate profile from cable plant database (CAPTAR), and linked history of modules in the Equipement Tracking for Controls database (DEPOT) which leads to maintenance history screens

4. LESSONS AND ISSUES

This accelerator-wide view of the information delivered through the WWW interface has been popular with users, enabling more efficient work.

There are a few key issues that have enabled us to get to this stage for an accelerator data warehouse

- We 'web-ified' both new or existing main disparate databases at the lab. This provided immediate benefits to the departments, and they were very willing to work with us.
- We approached each system with a lab-wide view of the information-not a departmental view.
- We focused on linking the existing departmental databases-not on rebuilding systems. Re-engineering is a major effort that our scarce programming resources cannot easily undertake, so we have left that effort to the departments-though, hopefully, there is coordination amongst related database work.
- We built incrementally, for each increment giving quick turnaround and visible benefits to the users and the departments.
- We relied on experience database programmer(s) for data modeling.

With the ability now to link databases via a few common data fields, the integrity of the data stored in these fields need to be enhanced by cleansing the data and/or applying better database constraints. But this effort becomes easier when the benefits are now apparent to the users (even if it may not benefit their own department immediately).

In the future, we would like to rely more on commercial WWW-database tools (such as Oracle Designer2000, etc.), and less on CGI scripts which are time consuming and less secure.

The URL for the PEP-II Project Database is: http://www.slac.stanford.edu/accel/pepii/db.htm

5. ACKNOWLEDGMENTS

We are grateful for the continued support of the PEP-II management. We thank the people in PEP-II and the various Technical Division departments who manage the different databases, and who helped us analyze the linkages of these different systems. We also thank the student programmers who contributed to the WWW interface. We have drawn many lessons from discussions with members of the International Accelerator Database Group (IADBG).

REFERENCES

- [1] A. Chan, G. Crane, I. MacGregor, S. Meyer, "The PEP-II/BABAR Project-Wide Database Using World Wide Web and Oracle*CASE", Proceedings of 1995 International Accelerator Database Group Workshop, Argonne National Lab, November 6-8, 1995.
- [2] J. Poole, "Databases for Accelerator Control–An Operations Viewpoint", Proceedings of 1995 Particle Accelerator Conference, Dallas, May 1–5, 1995, p. 2157.
- [3] J. Poole, "Summary and Conclusions of the IADBG Workshop on Databases for Accelerators", Proceedings of 1995 International Accelerator Database Group Workshop, Argonne National Lab, November 6-8, 1995.
- [4] Barry Devlin, "Data Warehouse: from Architecture to Implementation", Addison-Wesley, 1997.