PDS Repository Survey Summary

M. Martin

10/15/2007

The Planetary Data System is a distributed federation consisting of discipline nodes, subnodes, data nodes and functions. There are five science discipline nodes which provide expertise, tools and data for the atmospheres (ATM), geosciences (GEO), imaging (IMG), planetary plasma interactions (PPI), rings and small bodies (SBN) disciplines. The navigation node (NAIF) provides expertise, tools and data to support multiple science disciplines. There is also a radio science (RS) function that supports the collection and archiving of radio science data from all missions. Several of the discipline nodes have subnodes which focus on special areas within the discipline. There are also data nodes which prepare and provide access to specific data sets, generally for new mission data (IMG-THEMIS, IMG-HIRISE, IMG-LROC). The nodes and subnodes essentially operate as individual systems, each with their own hardware, software, system administration and maintenance. The Data Repository Survey was carried out in 2007 to determine the characteristics of the current node storage systems and plans for future growth. Responses were received from nine nodes and subnodes, three data nodes and radio science. Due to some confusion in interpreting the data node system costs (partially borne by the projects) and some potential duplication in accounting for future year expansion they have not been included in Tables 3 and 4. Statistics for THEMIS and HIRISE are presented in Appendix 2. LROC is not operational yet.

The key finding of the survey include:

- Nearly all planetary data sets are accessible online at the discipline nodes.
- The nodes primarily utilize LINUX or Solaris operating systems, Apache servers and RAID arrays for storage.
- A variety of applications and data base systems are used to develop node data access systems.
- The nodes have 84 TB of storage capacity with 38 TB (45%) used.
- Current total system cost (hardware, software, maintenance) is about $560K or $6.7K per TB.
- Yearly labor cost to operate the systems is an additional $200K.
- Only a few of the nodes have industrial strength backup systems in place and there are no mirror sites.
- Most nodes would seem to require substantial effort to rebuild their Primary Data Repository from backups.
- Data deliveries to the deep archive at National Space Science Data Center (NSSDC) are spotty.
- The GEO and IMG nodes account for 85 percent of current data volume and system cost and over 96 percent of anticipated data volume and cost over the next three years.

Data Repository Architecture

For this discussion the terms Primary Data Repository (PDR) and Secondary Data Repository (SDR) are used. The Primary Data Repository is the main storage site for a node's data collection. The Secondary Data Repository provides an alternate copy of the data repository for local backup purposes. PDS policy states that there should also be a copy of all archival data sets in the deep archive and the National Space Science Data Center. Table 1 identifies the PDR and SDR characteristics at each site.

NODE	Percent PDR Online	Percent PDR offline and storage type	SDR contains all of PDR?	SDR storage type	Percent archived at NSSDC	Tape system
RS	0%	100% CD&DVD			TWO-THIRDS	
SBN (Maryland)	100% RAID		Yes	CD&DVD	MOST	
SBN (PSI)	100% RAID				UNK	
NAIF	100% RAID				PARTIAL	
ATMOS	85% RAID	15% TAPE&CD			UNK	TAPE
PPI	100% RAID		Yes	CD&DVD&TAPE	PARTIAL	LTO TAPE
RINGS	100% RAID		Yes	DATA BRICK	PARTIAL	
GEO	85% RAID	15% TAPE&CD			UNK	LTO TAPE
IMG (USGS)	100% RAID		No	CD&DVD	MOST	
IMG (JPL)	95% RAID	5% OTHER	No	CD&DVD	PARTIAL	TAPE
IMG (THEMIS)	100% RAID				NONE	TAPE
IMG (HIRISE)	100% RAID				N/A	LTO TAPE J/B
IMG (LROC)	100% RAID		Yes	ONLINE	N/A	TBD

Table 1 PDR and SDR Characteristics

All the discipline nodes and several subnodes and data nodes provide an on-line PDR. The radio science "function" does not provide any on-line data access but some radio science volumes are on-line at other nodes. Most of the sites have 100 percent of their PDR collection online. ATMOS has 15% offline, mostly radio science volumes. GEO also has about 15% off line, but should have most data online by the end of 2007. IMG-JPL has 5% offline but also expects to have all data on-line in the near future. Five of 12 sites provide an SDR. For three of these sites all PDR data is represented in the SDR. For two sites IMG – USGS and IMG – JPL there is data in the PDR that is not represented in the SDR. Most of these are offline collections of tapes, CDs and DVDs stored on shelves or cabinets at the site. The engineering node has for many years maintained an online collection of the entire collection of PDS CD-ROM volumes but this was not specified as a secondary repository by any of the nodes.

The responses regarding delivery of data to the deep archive make it difficult to assess the situation. It would appear that most of the data sets that have been delivered to NSSDC are older data sets that were produced on CD-ROM or DVD-ROM media for widespread distribution to the science community. Most newer data sets that have been received electronically or on recordable media probably have not been delivered, pending a new electronic delivery system that is currently being tested. Six out of 12 sites report having tape backup systems available, with at least three using newer LTO systems (Linear Tape – Open).

Data Repository System Characteristics

Table 2 summarizes the hardware and software systems at all the nodes. Primary servers are Sun (3 sites), Dell (4 sites) and unspecified rack mount servers. Every site uses some variant of UNIX including: Solaris (3 sites), RedHat (5 sites) and SuSE (2 sites) Linux. One site uses Windows XP for its main web site but also has a Linux server. Nearly every site uses some version of the Apache server. One site uses standalone Apache Tomcat and one uses Microsoft Internet Information Server

(IIS). Other software includes the object oriented data technology profile and data servers (4 sites), java, javascript, VisualStudio, proftpd, IAS JPIP Server, PHP, Perl, Flash, mod_perl, mason. Database software specified includes Oracle, MySQL, Sybase and PostgresSQL. Most of the sites use some sort of firewall including iptables (2 sites), Firestarter, UNIX security, hardware firewalls, netscreen, Cisco FWSMs and facility firewall. Communication links to the outside are either 1 Gbps (5 sites), and .045 Gbps (3 sites). One site has an internal 10 Gbps network.

NODE	SYSTEM TYPE	OPERATING SYSTEM	SERVER AND DATABASE	FIREWALL	COMMUNICATION RATE
SBN (Maryland)	SUN+DELL	SOLARIS +RH LINUX	Apache 2.0.55, Apache 2.2.3, Oracle 8	Iptables	1 Gbps
SBN (PSI)	HP	RH LINUX	Tomcat	firestarter firewall	UNK
NAIF	DELL	LINUX	Apache	UNK	.045 Gbps (T3)
ATMOS	SUN+DELL	SOLARIS	Apache/tomcat/java/OODT	unix security (ssh)	UNK
PPI	RACKMT SERVERS	RH LINUX	Tomcat	Cisco, iptables	1 Gbps
RINGS	UNK	UNK	UNK	UNK	UNK
GEO	DELL	RH LINUX + WIN	IIS, VisualStudio, JavaScript	Hardware firewalls	.045 Gbps (Shared T3) lines; gigabit internal between servers
IMG (USGS)	DELL+ EINUX+ G/W	SUSE LINUX	Apache / Proftpd / OODT	NetScreen Firewall	.045 Gbps (DS-3) link from Geonet3
IMG (JPL)	RACKMT SERVERS	RH LINUX	Apache / Tomcat / OODT / MySQL, Sybase	UNK	1 Gbps Ethernet internal network, .100Gbps external I/F
IMG (THEMIS)	RACKMT SERVERS	RH LINUX	Apache 1.3 with mod_perl/mason/php	facility firewall	1 Gbps gigabit Ethernet to ASU core
IMG (HIRISE)	RACKMT SERVERS	SOLARIS	Apache, PDS-D, Tomcat, java, IAS JPIP Server	Cisco FWSMs	1 Gbps Gigabit Ethernet via dual-failover connection
IMG (LROC)	RACKMT SERVERS	SuSE Linux 10.x, OpenBSD v4.x	Apache v2.0, PHP, Perl, Flash	Iptables, pf	10 Gbps campus link, 1 Gbps offsite

Table 2 Node System Characteristics

Online storage is RAID (7 sites), Storage Area Network (SAN) RAID (2 sites), Network Attached Storage (NAS) RAID (2 sites). We presume that most of the raid systems have some fault tolerance. Six sites mention offline storage, including SAN attached LTO-3 (Linear Tape Open 3 - 400 GB), Sun/Storage Tek SL500 LTO-3 library, tape backup systems, facility tape backups, databrick, hard drive and DVD.

Primary Data Repository Capacity and Cost

The storage capacity at existing sites ranges from 100 GB to 1 petabyte (LROC). For this discussion we will leave out the three data nodes and just focus on the PDS nodes and subnodes. Table 3 summarizes current capacity and cost at the nodes.

The total storage capacity is 83 TB, of which 38 TB (45%) is used. The actual data storage is estimated to be slightly less, at 33 TB. Two nodes, Geosciences and Imaging comprise 72 TB (87%) of the total capacity and 32 TB (84%) of the storage used. The PDR costs vary markedly from node to node and are difficult to interpret. One would expect the nodes with smaller data collections to have a higher cost per terabyte and this is generally the case. However, the cost figures submitted by Geosciences node are more than all the other nodes combined, and that is even after reducing the raw numbers to attempt to reflect current hardware costs. Overall we see a hardware cost of $6K per TB. The software costs specified were negligible for all the nodes except Geosciences, reflecting the use of primarily open source software. The maintenance costs are also negligible, probably reflecting that maintenance is done by the node staff when something fails, not via a maintenance contract. The yearly labor costs for operation have been converted from FTE's to dollar figures using $80K per FTE. They amount to 40% of the total hardware cost.

	SBN	SBN-PSI	NAIF	ATMOS	PPI	RINGS	GEO*	IMG–USGS	IMG–JPL	Total
Online Storage Capacity (TB)	3.5	.4	.1	1.5	4.0	2.0	18.0	20.7	33.0	83.2
Online Storage Used (TB)	2.1	.2	.1	1.0	1.6	1.0	14.5	6.6	10.9	37.9
PDR Size (TB)	1.0	.1	.1	.5	2.0	1.0	12.0	7.0	9.0	32.6
Hardware Cost $K	$12.0	$22.0	$16.0	$.0	$4.0	$.0	$300.0	$113.0	$45.0	$512.0
Software Cost $K					$1.0		$30.0		$1.0	$32.0
Maintenance Cost $K		$.0			$1.0		$12.0	$1.0	$3.0	$17.0
Total Cost $K	$12.0	$22.0	$16.0	$.0	$6.0	$.0	$342.0	$114.0	$49.0	$561.0
Total Cost per TB $K	$5.7	$137.5	$320.0	$.0	$3.8	$.0	$23.6	$17.2	$4.5	$14.8
Yearly Labor Cost $K**			$16.0	$20.0	$2.0		$120.0	$3.2	$40.0	$201.2

* GEO HW and SW costs reduced by 40% to reflect current cost

** Labor cost estimated using $80K per person year

Table 3 Discipline Node Capacity and Cost

Primary Data Repository Expansion Plans

The GEO and IMG nodes account for 98 percent of anticipated new data over the next several years and 96 percent of system expansion. Most of the nodes plan continuous upgrades. IMG-USGS plans substantial upgrades ($73K for 11.2 TB) this year. These numbers have been included in Table 3 and 4. They also project expenditures for absorbing MRO and Lunar data sets on the order of $500K per year in 2008, 2009 and 2010. IMG-JPL plans major replenishments in 2008 ($32K) and 2009 ($46K). Three nodes plan upgrades in 2010. The data volumes that go along with these expenditures are 150 TB in 2008 (five times the current PDR size), 205 TB in 2009 and 235 TB in 2010. Thus PDS is planned to expand to 18 times its current storage in three years. The cost per TB works out to $3.4K, $3.0K and $2.7K for these years, not including software cost, which would add about 10 percent.

	SBN	SBN-PSI	NAIF	ATMOS	PPI	RINGS	GEO*	IMG–USGS	IMG–JPL	Total
Online Storage Used (TB)	2.1	.2	.1	1.0	1.6	1.0	14.5	6.6	10.9	37.9
Online Storage Capacity (TB)	3.5	.4	.1	1.5	4.0	2.0	18.0	20.7	33.0	83.2
New data – 2008 TB	1.3		.0	.5	1.0		24.0	93.0	30.0	149.8
New data – 2009 TB	1.0		.0	.5	2.0		42.0	100.0	60.0	205.5
New data – 2010 TB	1.0		.0	.5	2.0		42.0	100.0	90.0	235.5
Current Hardware Cost $K	$12.0	$22.0	$16.0	$.0	$4.0	$.0	$300.0	$113.0	$45.0	$512.0
Hardware Cost 2008 $K				$10.0	$.0		$40.0	$424.0	$30.0	$504.0
Hardware Cost 2009 $K				$10.0	$4.0		$100.0	$450.0	$45.0	$609.0
Hardware Cost 2010 $K			$15.0	$10.0	$15.0		$140.0	$450.0	$3.0	$633.0
Current Software Cost $K					$1.0		$30.0		$1.0	$32.0
Software Cost 2008 $K				$.0	$.0		$4.0	$50.0	$2.0	$56.0
Software Cost 2009 $K				$.0	$.0		$20.0	$50.0	$1.0	$71.0
Software Cost 2010 $K				$.0	$.0		$20.0	$50.0	$1.0	$71.0
Current H/W Cost per TB $K	$3.4	$55.0	$160.0	$.0	$1.0	$.0	$16.7	$5.5	$1.4	$6.2
2008 H/W Cost per TB $K	$.0		$.0	$20.0	$.0		$1.7	$4.6	$1.0	$3.4
2009 H/W Cost per TB $K	$.0		$.0	$20.0	$2.0		$2.4	$4.5	$.8	$3.0
2010 H/W Cost per TB $K	$.0		$750.0	$20.0	$7.5		$3.3	$4.5	$.0	$2.7
* GEO HW and SW costs reduced by 40% to reflect current cost										

Table 4 PDR Expansion Plans

Legacy Archival Data Holdings

Several questions were included to determine the amount of older media that might have to be copied to new media. Since there are no large sets of offline archival data holdings the questions did not yield much useful information. Radio Science has 100 volumes from before 1990. Imaging has some Science Digital Data Preservation Task (SDDPT) volumes that are older than 10 years but is working to put these volumes online. There are some Magellan F-BIDR tapes and CD's at GEO and IMG-USGS that have not been identified anywhere, but GEO hopes to have these volumes online this year. There are 1858 Magellan DODR volumes on CD-R media that nobody in PDS seems to own.

Comments Submitted by Survey Respondents

The following sections present comments provided by the respondents. The names of the respondents are listed in Appendix 1. Among the issues raised in the comments: One would like to preserve data sets on hard media and requests continuing research of higher density hard media formats; several indicate a desire for a more robust system architecture; one indicates that the data node architecture seems to be working very well and one requests "clear policies on administration, availability and preservation..."

1. Specific data repository future plans

[EE] "would like to produce hard media of the HIRISE archive". [CI] "upgrade to gigabit speeds. Continue migration to clustered storage solution." [MM] "Buttress server infrastructure, investigate fail-over mechanism for repository access." *NOTE: MM should talk to GEO who is already doing this.* [KM] "online Secondary Repository being established. Contacts and schedules need to be negotiated if THEMIS offline archives are desired." [TK] "continue placing all data on RAID. Backups made at regular intervals on different technology (eg tape). " [TS] "Working on security and data management plan." [BS] "Establish mirror at ESTEC or ???. Complete submissions to NSSDC."

2. PDS-wide recommendations

[BS] Better visibility of SPICE data via central catalog. [EE] Having the HiRISE PDS data node co-located at HiROC has (from our perspective) really eased data management and distribution, and the present plan of relocating the associated PDS hardware to the Imaging Node at end of mission is workable, but for long-term (multi-decade) storage, higher capacity approved optical storage media should be viable and preferred for long term storage (lower yearly management costs and high reliability, assuming a mistake is not made in investing in a dead-end technology). Historically the optical archival of PDS data has been a great boon. Even now the Planetary Image Research Laboratory at the University of Arizona and the Tucson branch of the NASA Space Imagery center, also at the University of Arizona, depend on archive optical media (CD-ROMs) for earlier planetary datasets, though we also keep copies of these datasets on-line on host attached disk arrays - our CDROM jukebox was retired years ago.

3. Additional comments.

[EE] - We also plan complete backups on tape media (LTO-3), but we would not like to consider this as a long-term archive format. We plan only to use the tape backups for disaster recovery. [KM] Data Subnode; this protocol has worked extremely well for this mission. As we have the instrument expertise to generate and provide this data, we are therefore, ideally suited to administrate the data products. With the ever increasing volumes of data returned by imaging instruments, it is the experience of the THEMIS mission that online storage, with redundant backups, is more straightforward to maintain than other media types. [TK] The usefulness of surveys such as this is unclear. There is currently no need to redesign the basic architecture of the PDS. Clear policies on administration, availability and preservation would be more useful. Nodes would then implement the policies within their respective environments.

Appendix 1. Survey Respondents

[RS] Richard Simpson (RS)
[AR] Ann Raugh (SBN)
[BS] Chuck Acton / Boris Semenov (NAIF)
[LH] Lyle Huber (ATMOS)
[TS] Thomas Stein (GEO)
[CI] Chris Isbell (IMAGING-Flagstaff)
[EE] Eric Eliason (IMAGING-HiRise)
[MM] Myche McAuley (IMAGING-JPL)
[EB] Ernest Bowman (IMAGING-LROC)
[KM] Kimberly Murray (IMAGING-THEMIS)
[CN] Carol Neese (SBN-PSI)
[TK] Todd King (PPI-UCLA)
[MG] Mitch Gordon (RINGS)

Appendix 2. Data Node Tables

	THEMIS	HIRISE
Online Storage Capacity (TB)	25.5	6.0
Online Storage Used (TB)	10.7	4.3
PDR Size (TB)	4.9	8.6
Hardware Cost $K	$50.0	$23.0
Software Cost $K		
Maintenance Cost $K	$5.0	$7.0
Total Cost $K	$55.0	$30.0
Total Cost per TB $K	$5.1	$6.9
Yearly Labor Cost $K**	$2.0	$20.0

	THEMIS	HIRISE
Online Storage Used (TB)	10.7	4.3
Online Storage Capacity (TB)	25.5	6.0
New data – 2008 TB	13.0	40.0
New data – 2009 TB	15.0	.0
New data – 2010 TB	18.0	.0
Current Hardware Cost $K	$50.0	$23.0
Hardware Cost 2008 $K	$8.0	$68.0
Hardware Cost 2009 $K	$40.0	$.0
Hardware Cost 2010 $K	$8.0	$.0
Current Software Cost $K	$5.0	$7.0
Software Cost 2008 $K	$.0	$5.0
Software Cost 2009 $K	$.0	$.0
Software Cost 2010 $K	$.0	$.0
Current H/W Cost per TB $K	$2.0	$3.8
2008 H/W Cost per TB $K	$.6	$1.7
2009 H/W Cost per TB $K	$2.7	$.0
2010 H/W Cost per TB $K	$.4	$.0