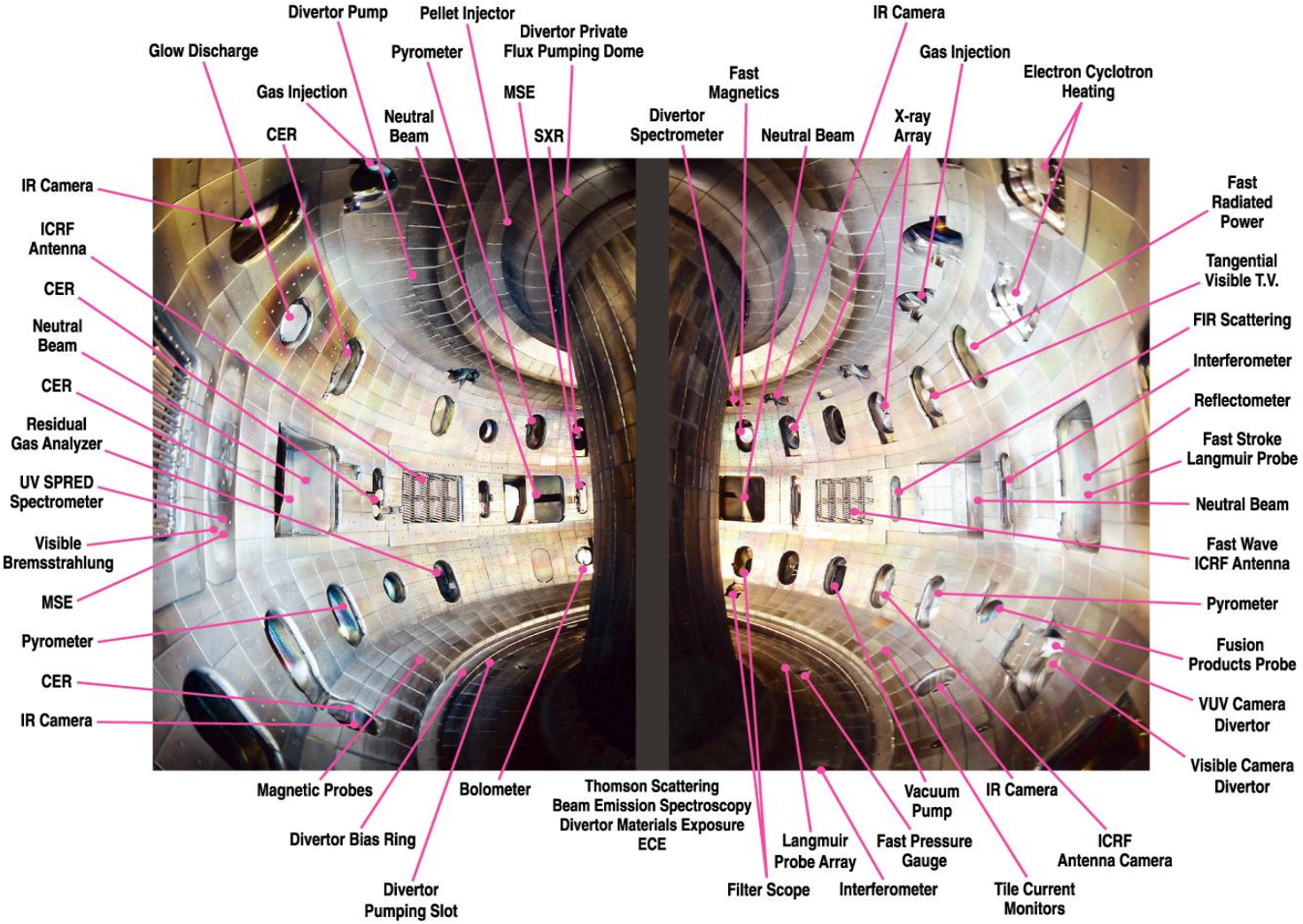
Storage Data Analysis with Nirvana SRB

Presented for 2014 IEEE MSST Conference
Santa Clara, CA
June 2-6 2014

Nirvana

Explosion of Data Growth from Experiments, Computer Simulations, and Advanced Sensors



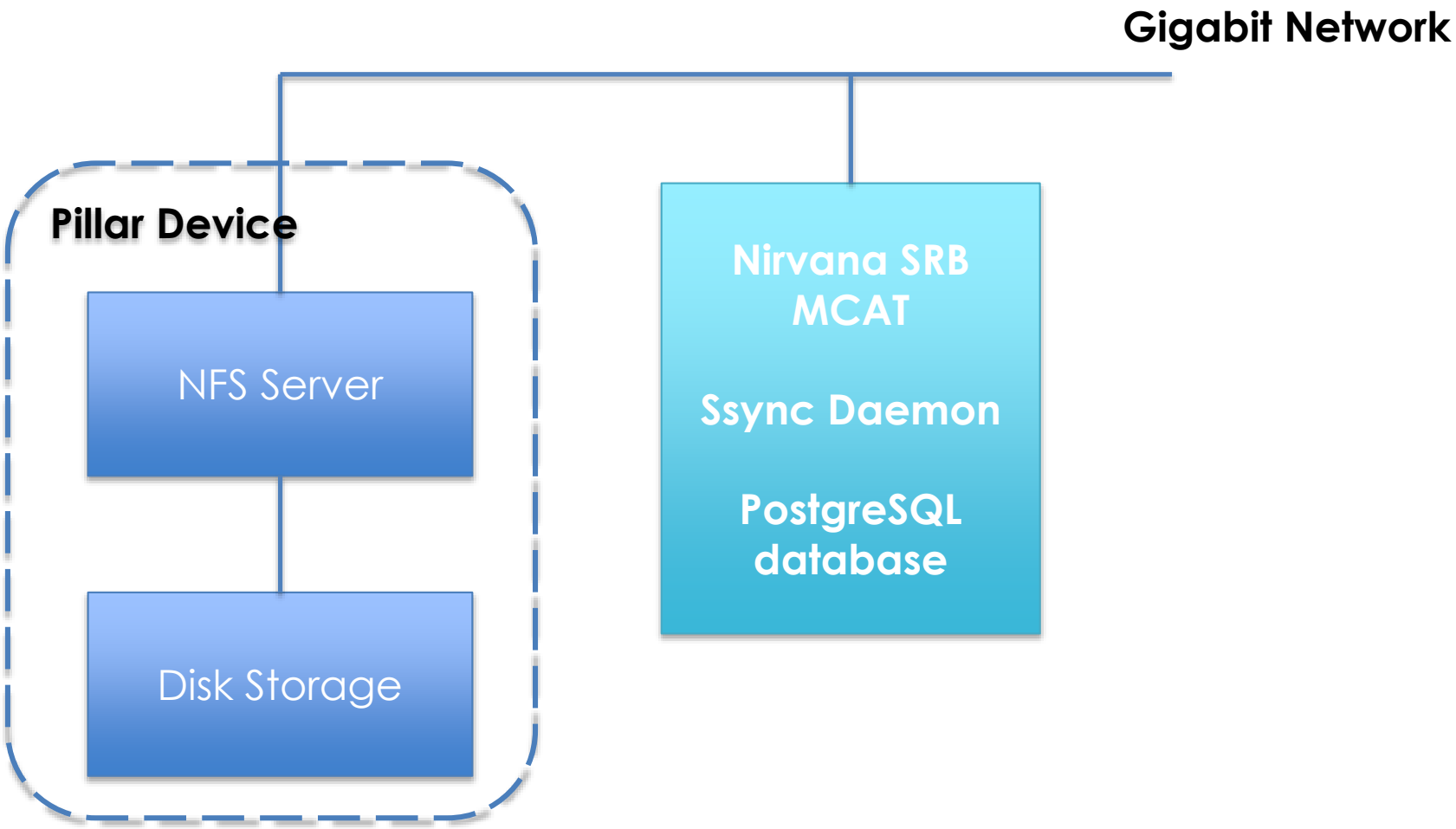
Challenge: User Data Rests On Expensive High Performance Storage System That Affects Tokamak Experiments

- **Pillar Axiom Storage (now Oracle) was bought in 2008**
- **Three major partitions exported via NFS v3**
  - /u (user home directory) 23 TB used
  - /data (work / scratch space) 21 TB used
  - /vidata (visualization) 17 TB used
- **Limited visibility to what is stored in Pillar, forcing SysAdmin to back up everything**
- **Compute jobs that produces lots of small files slow down compute resource crucial to experiments on D-3D Fusion Reactor**
- **We call this type of unmanaged storage Data Junkyard**

Using Nirvana SRB To Inventory And Analyze The Junkyard

- **Build Nirvana SRB application server**
- **Exports 3 NFS filesystems to SRB application server**
- **Use Nirvana SRB fast multithreaded synchronization daemon to register user file metadata**
- **Analysis queries and reports against Nirvana SRB Metadata Catalog**
- **Analysis reports basis for:**
  - Remediation of duplicates and other conditions with individual users
  - Creating Information Lifecycle Management policy proposals for the next quarter, year, etc.
    - Migrate accessed files to less expensive data-store technology
    - Bulk data protection and DR capacity planning

System Block Diagram



Analysis: Summary File Inventory

Location	Total TB	Count	Avg Size
/u	23	70,664,216	332 KB
/data	20	66,421,040	316 KB
/vidata	17	215,288	82.9 MB
Total	**60**	**137,300,544**	

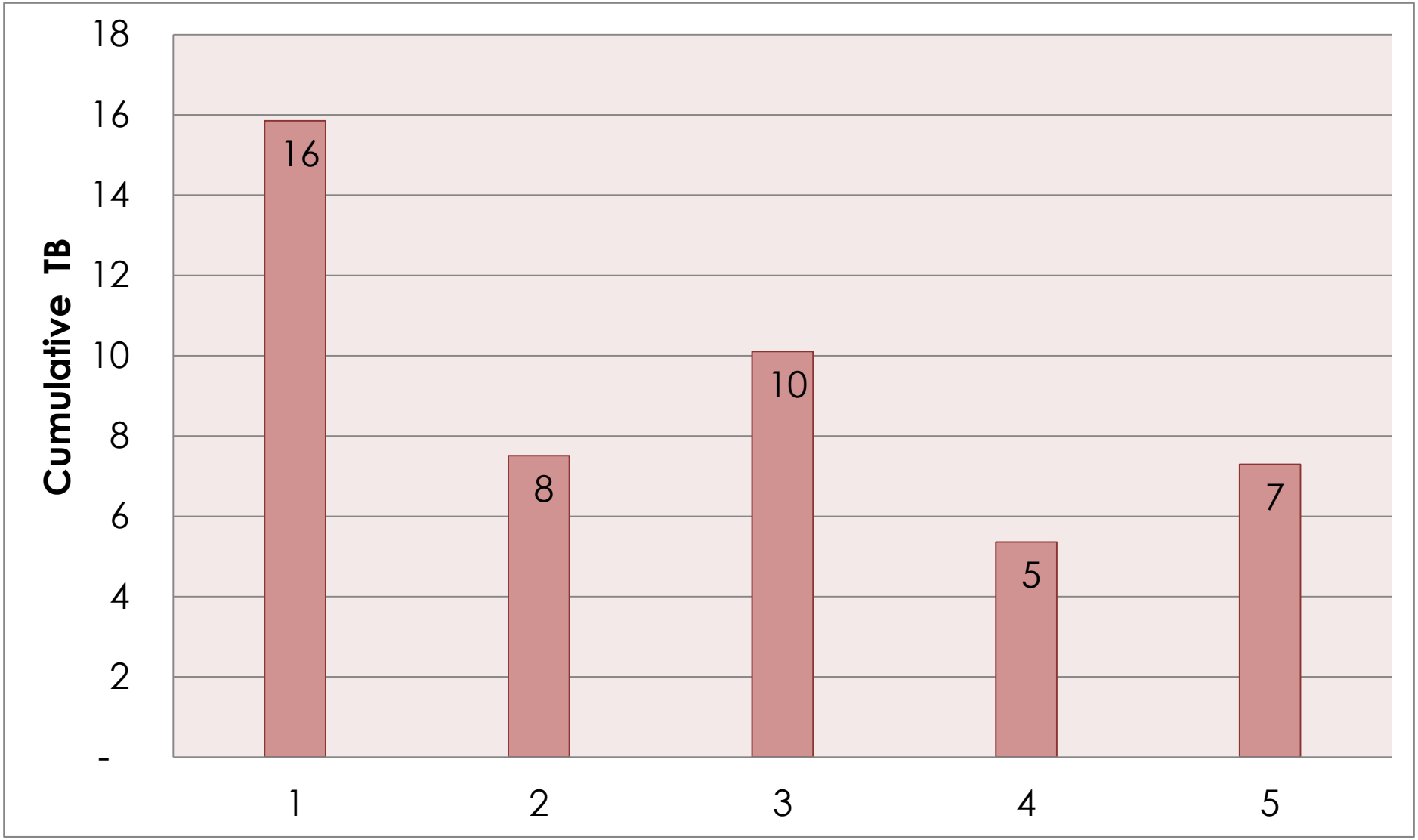
Analysis: Top 10 Users By File Count

User	TB Total Use	Count	Avg Size
1	3.64	39,376,868	92 KB
2	0.13	4,145,415	31 KB
3	1.30	3,282,744	396 KB
4	1.81	1,883,409	961 KB
5	1.66	1,605,967	1 MB
6	0.80	1,516,296	520 KB
7	0.36	1,092,901	329 KB
8	2.21	902,696	2.4 MB
9	0.12	901,450	133 KB
10	0.06	767,108	78 KB

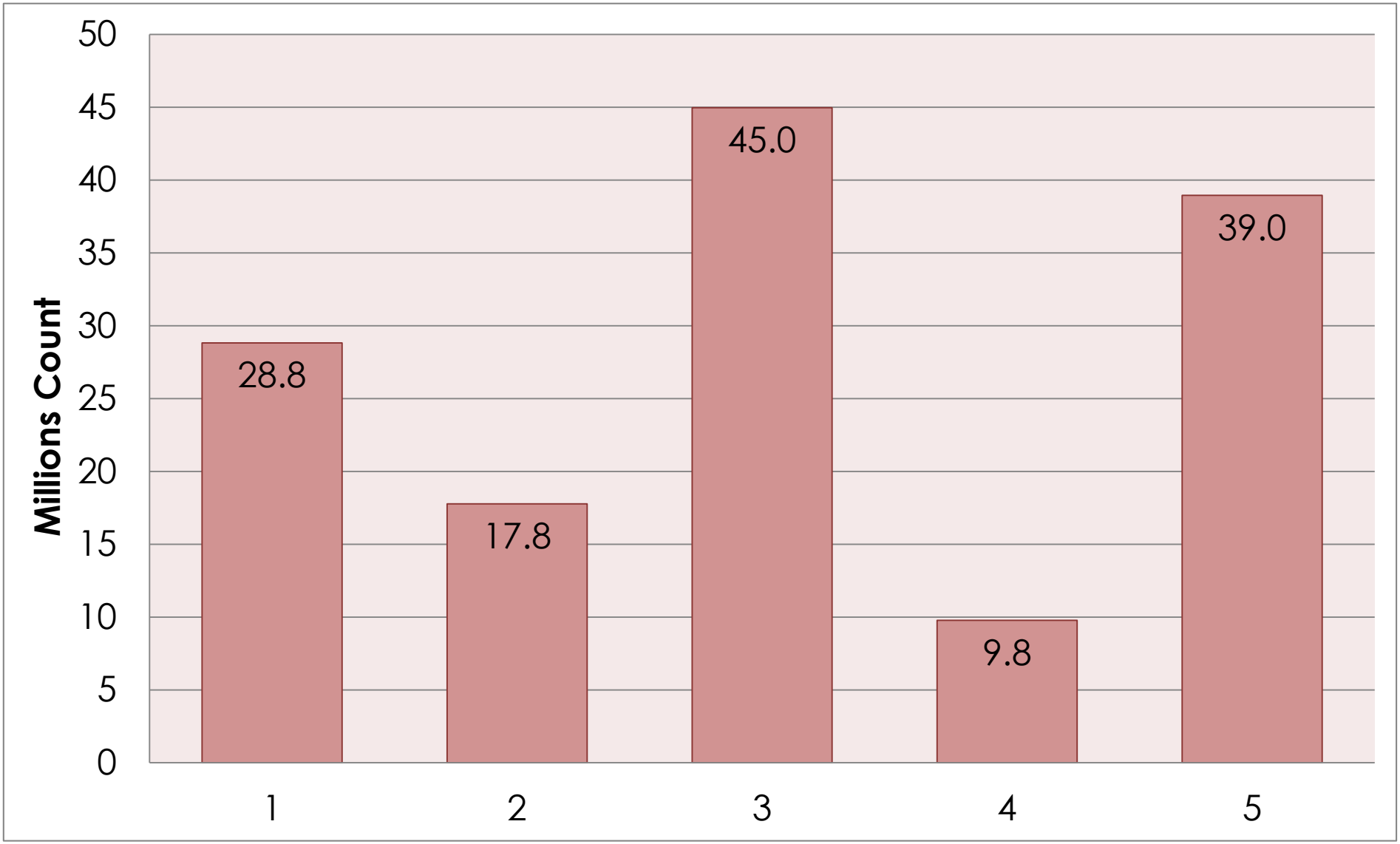
Analysis: File Type Inventory

Type Designator	Total TB	Count	Avg Size
EFIT	2.88	$36x10^6$	80 KB
GAProfile	1.53	$24x10^6$	64 KB
CINE	10.68	5,400	197 MB
CIN	3.65	3,200	1.1 GB

Analysis: Cumulative TB By Last Access Age In Years



Analysis: Cumulative File Count By Last Access Age In Years



Analysis Capability Enables…

- **Nirvana SRB gives us ability to identify top 10 users of codes that generate thousands of small files per run**
  - Affects overall system responsiveness
- **Nirvana SRB gives us ability to locate the knee of the curve to set archival policy based on last access age**
- **Nirvana SRB gives us**
  - The business evidence to justify less expensive second tier storage for WORM data (video files CIN, CINE)
  - The mechanism to migrate the data from high $$/TB Pillar disk using Nirvana SRB ILM automation
- **Nirvana SRB gives us hard data evidence for**
  - Negotiating a quota policy and/or data-store hardware funding with the user communities
  - Designing next generation data store

Lessons Learned

- **Tune Linux for SSD storage**
  - Deadline scheduler, fifo_batch, no barrier, no atime, swappiness, PGSQL work_mem
- **Nirvana SRB Ssync registered an average of 600 files / second**
  - Full scan took 2.6 days
- **Nirvana SRB application server HW was under-powered**
  - Nirvana SRB known to be capable of 2 X or better than this perf
  - NFS mount did not help
  - Better: 2 sockets with at least 64 GB DRAM for Nirvana Metadata Catalog service + Nirvana SRB filesystem scanner
  - Better: 2 sockets with at least 128GB DRAM for RDBMS
- **Over all CINE files, SRB found 940 duplicates files occupying 1.23 TB total**
- **Over all CIN files, SRB found 521 duplicate files occupying 298 GB total**

Thank You!