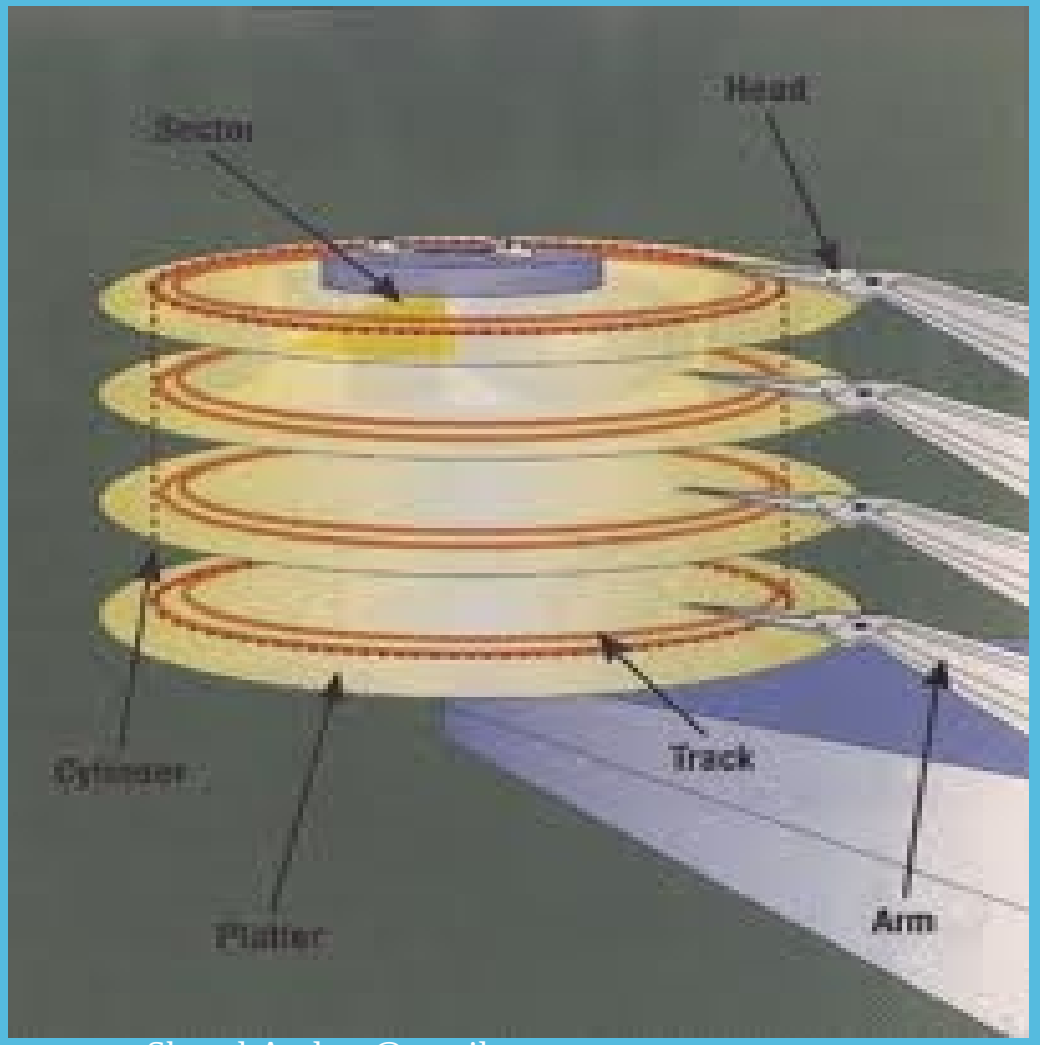
Lecture 3: Data store and transfer

https://sites.google.com/site/clustergateorg/

- Where to store the data
- Disk drives
- Organisation of the data store on disk drives (RAID)
- Data transfer in LAN
- Distributed file systems
- Data transfer on large distance (intercity, between countries)

Shevel.Andrey@gmail.com

Disk drive with magnetic method write/read

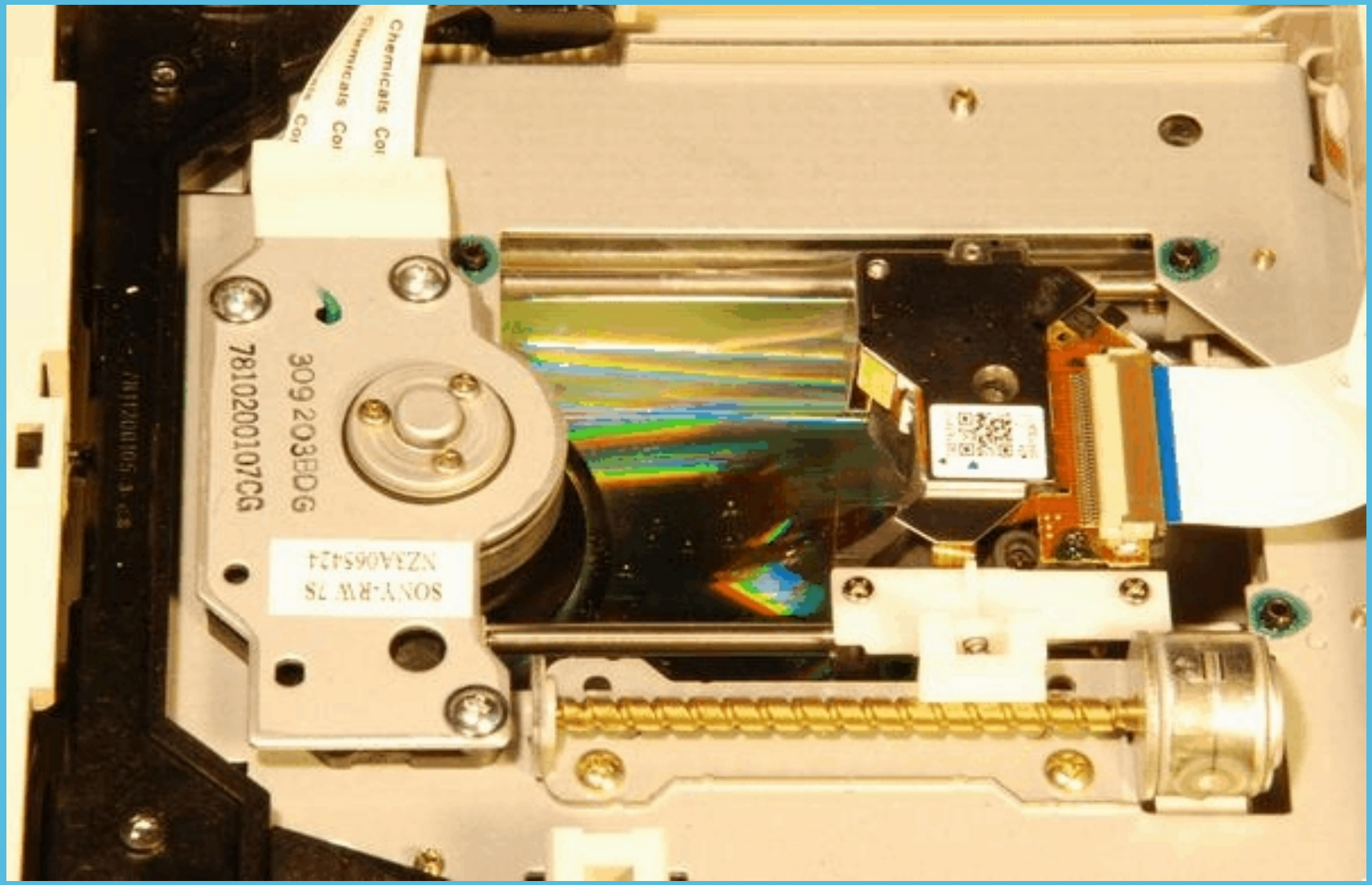


Shevel.Andrey@gmail.com

Disk drive

Shevel.Andrey@gmail.com

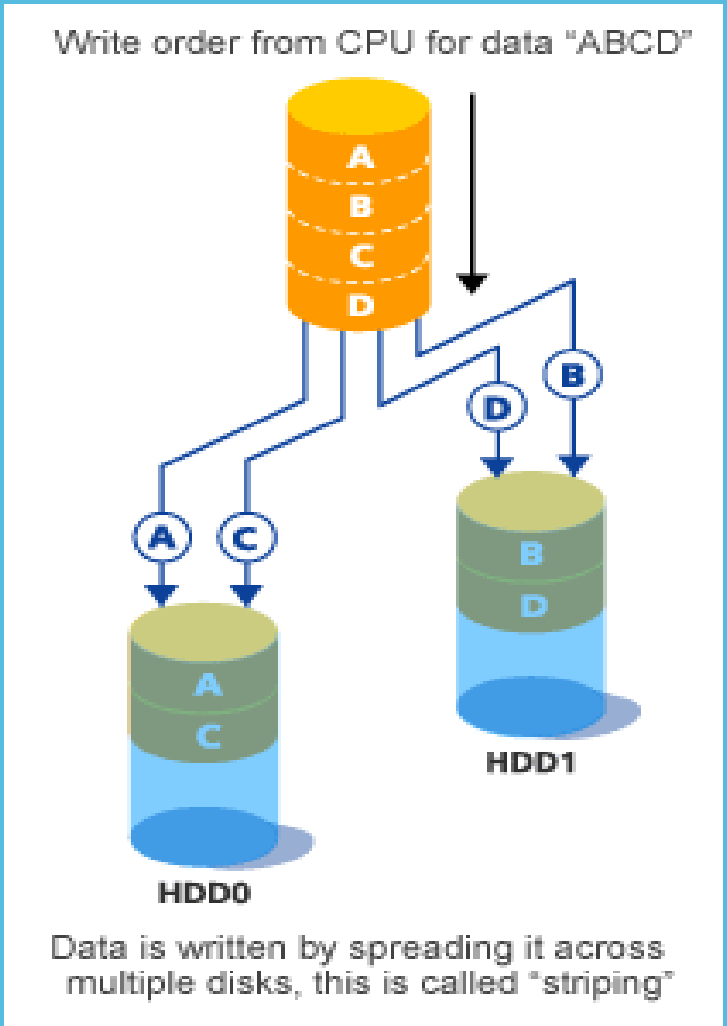
Optical disk drive

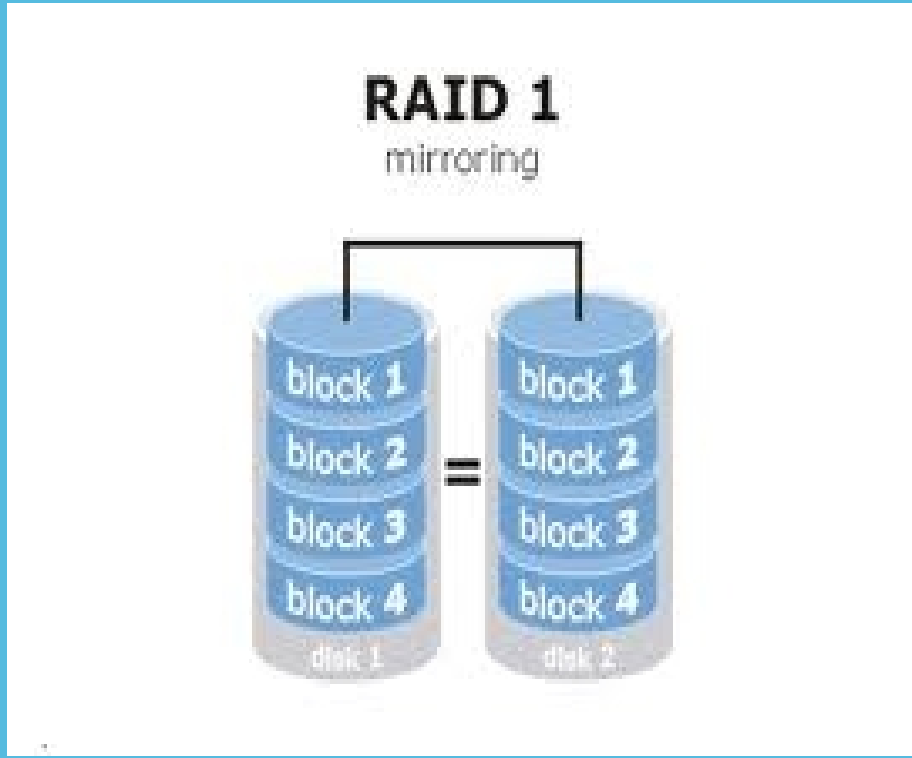


Shevel.Andrey@gmail.com

redundant array of independent *(inexpensive)* disks

RAID0

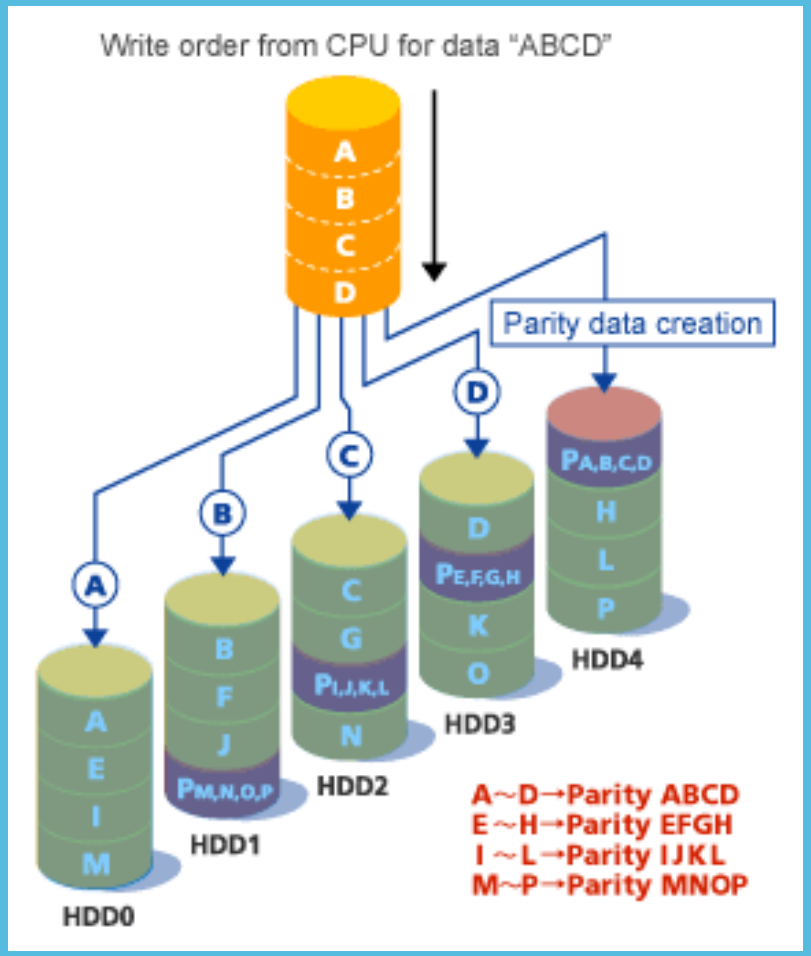




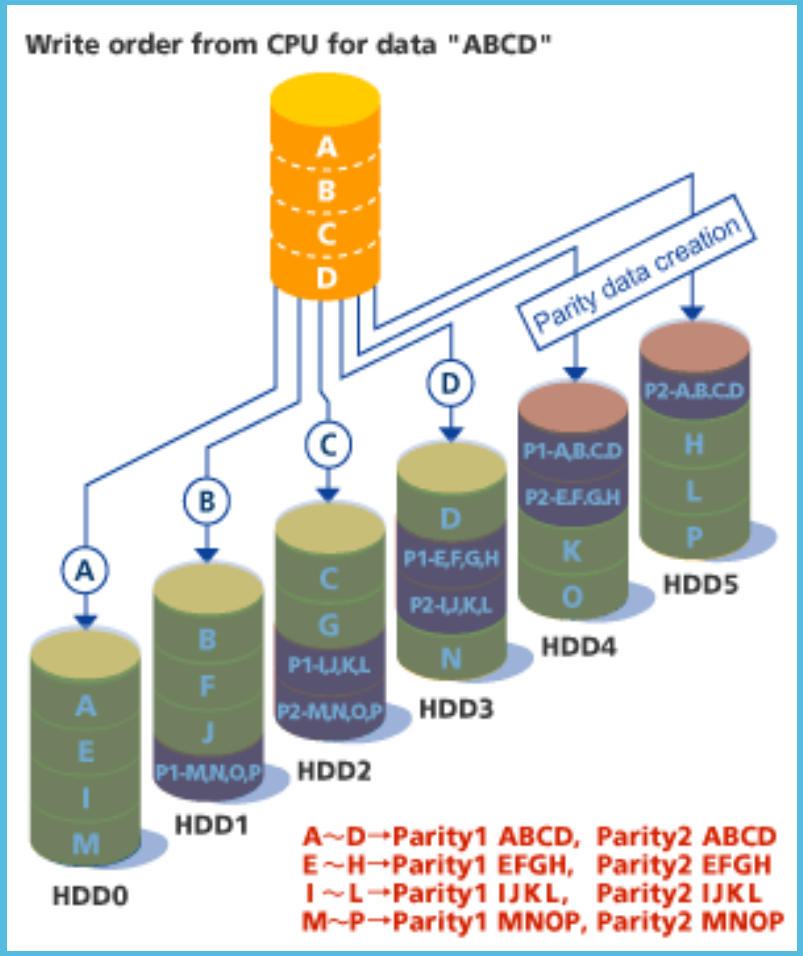
Shevel.Andrey@gmail.com

RAID5 и RAID6

RAID5



RAID6



Shevel.Andrey@gmail.com

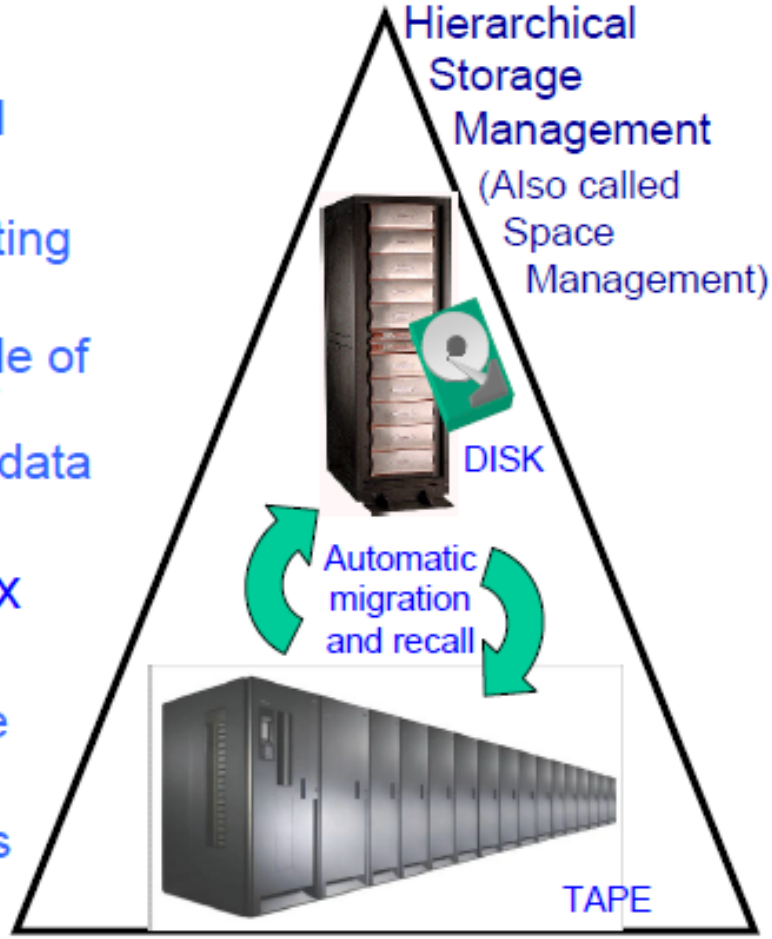
Robotic storage

Shevel.Andrey@gmail.com

High Performance Storage System

HPSS

- Disk and tape file repository
  - Hierarchical storage management (HSM) with automatic migration and recall
  - Highly scalable for high-end computing and storage customers
  - A single instance of HPSS is capable of concurrently accessing hundreds of tapes for extremely high aggregate data transfers.
- User sees HPSS as a single Unix file system
  - "Classic" HPSS presents its own file system
  - New HPSS for GPFS extends IBM's most scalable file system to tape



Data storage in large systems

- **High Performance Storage System (HPSS) http://www.hpss-collaboration.org/**
  - HPSS (High Performance Storage System) is a storage management system especially designed for moving large files and large amounts of data around a network that may consist of parallel processing computers, supercomputers, and clusters of high-end workstations.
- **Who uses large volume storages http://www.hpss-collaboration.org/learn_who_petabyte_data.shtml**
  -
  - NCSA - 380 PB http://www.hpcwire.com/2013/05/30/blue_waters_seals_off_with_tape/
  - **NSA** - http://nsa.gov1.info/utah-data-center/

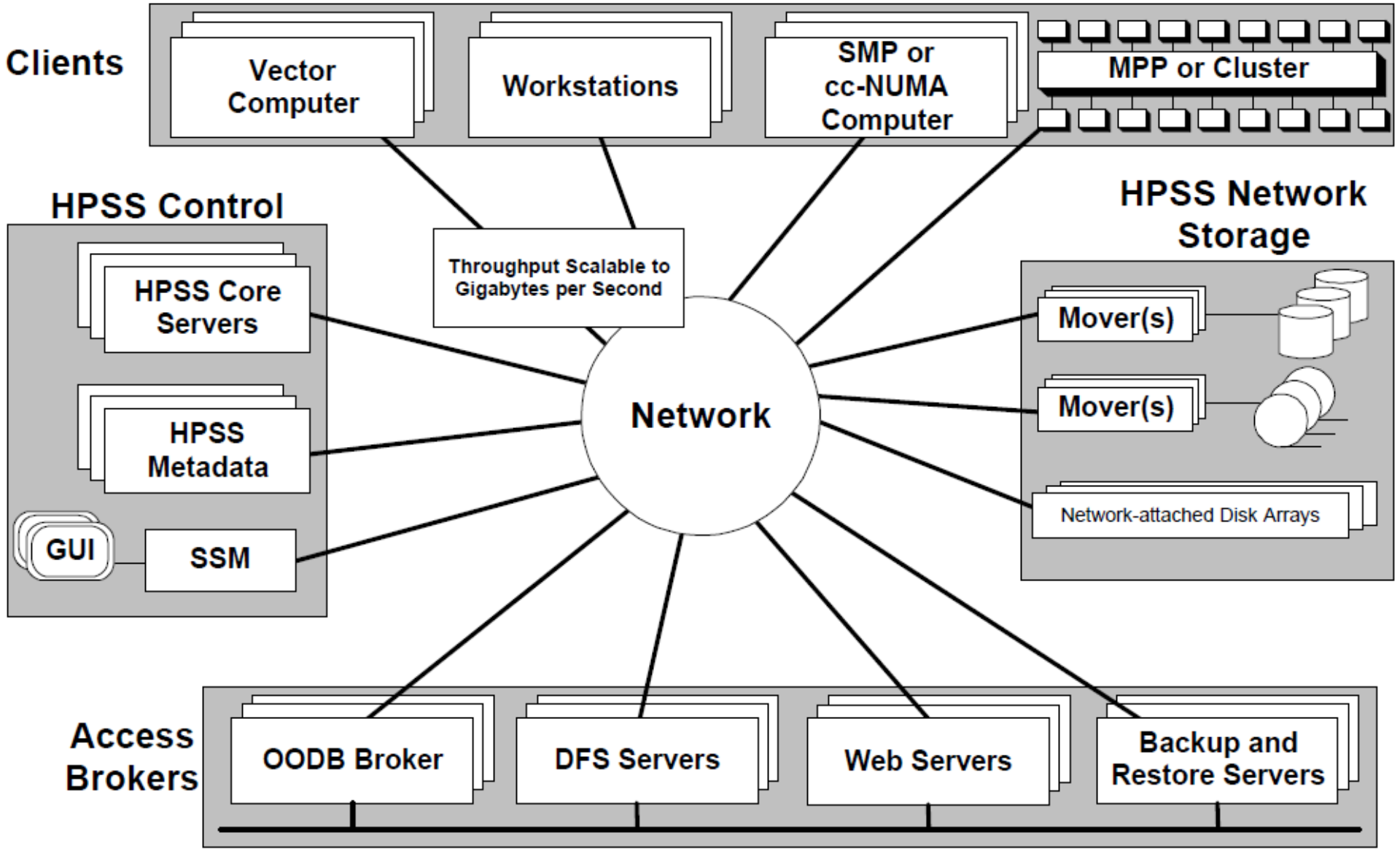
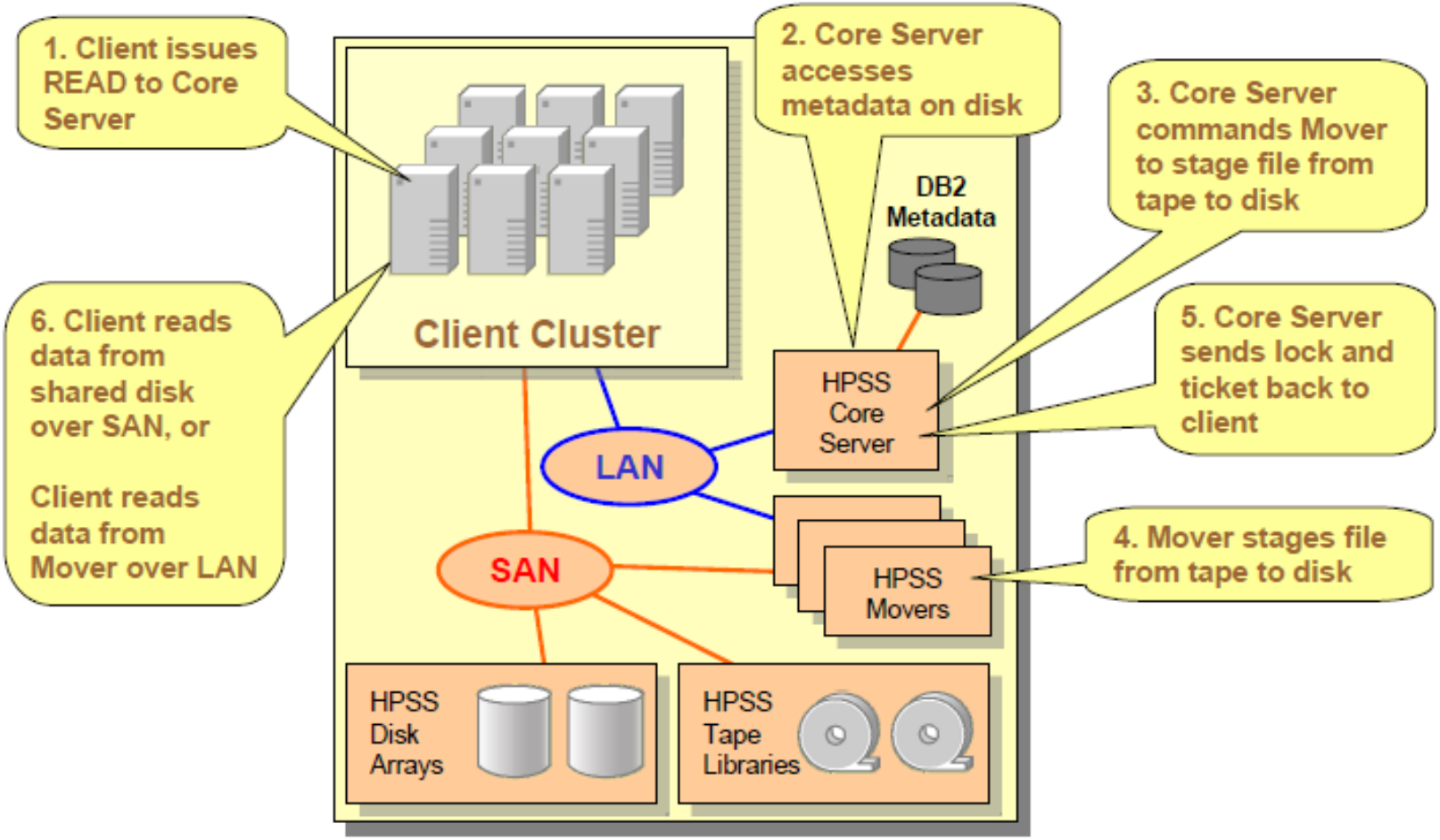


Figure 1 - HPSS Network Centered Architecture

How HPSS works

Example of an hpss_read

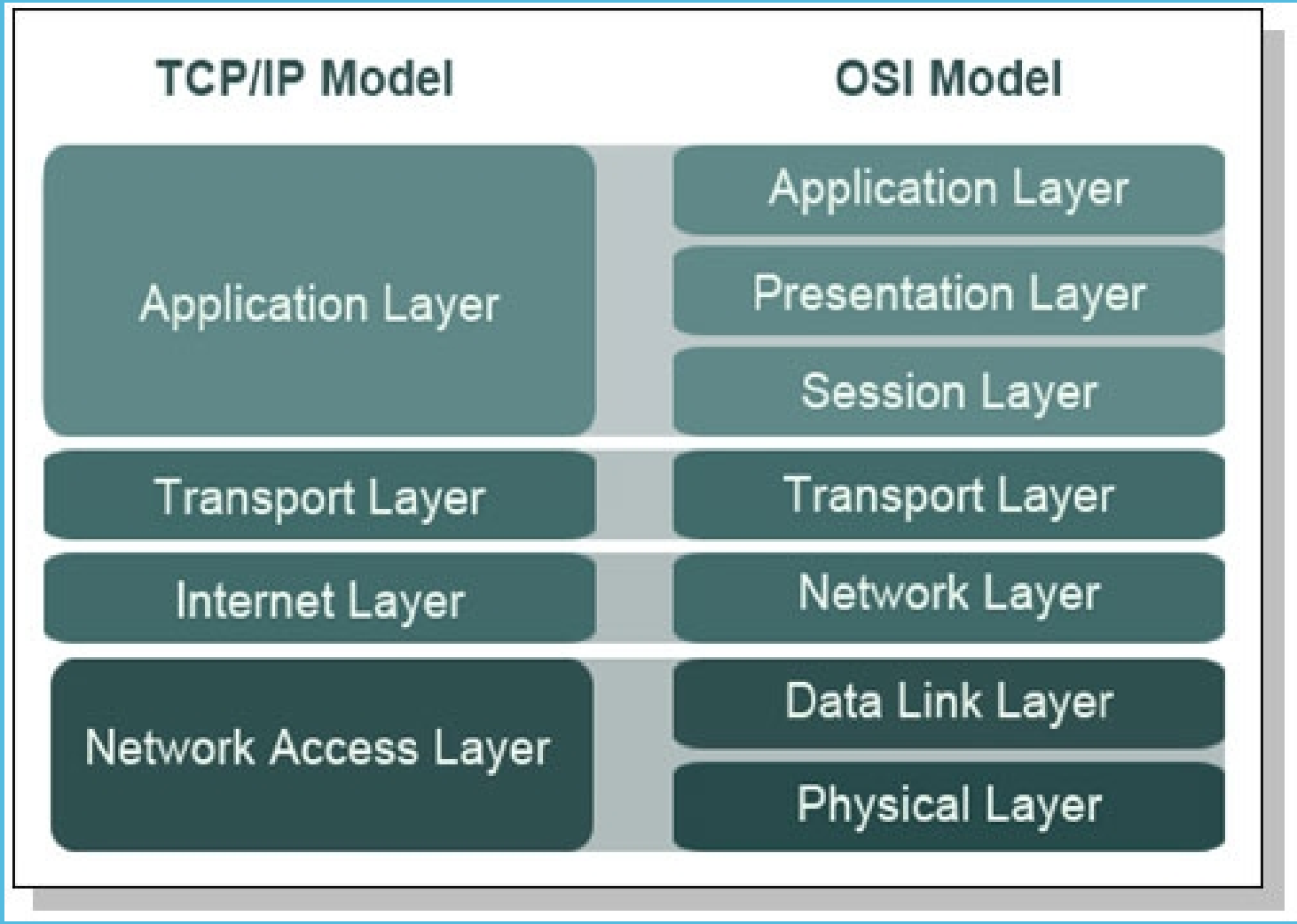


Shevel.Andrey@gmail.com

Data transfer in LAN

- **In LAN it is still used the stack of TCP/IP**
  - Initial and most longest used protocols for data transfer is ftp and its successor sftp;
  - Later on a lot of protocols/utilities for data transfer have been appeared
  - http://en.wikipedia.org/wiki/List_of_file_transfer_protocols

Data transfer models



Network filesystems

- ***Distributed filesystem*** - AFS
- *Global filesystem (in RedHat GFS2)*
- ***Symmetric filesystem*** – clients perform also manager codes for metadata.
- ***Asymmetric filesystem*** – there are several managers for metadata, which supprt filesystem. Examples: Panasas ActiveScale, Lustre. Traditional client/server filesystems like NFS and CIFS are also asymmetric.
- ***Cluster filesystem*** – distributed filesystem, which is not one server, but cluster, mainly for data storing. For clients such the cluster is just "filesystem".
- ***Parallel filesystem*** – filesystem to support parallel computing, all nodes might use same files. Data in the file is distributed by strips among many servers in order to increase the performance.

Type of the access to the disk storage

- **By File, e.g. NFS**
- **By Block, e.g. SAN**
  - In SAN might be used SCSI, iSCSI, Fibre Channel, Network Block Device, Infiniband
- **By Objects**

CAP theorem

- **Not possible to meet all of requirements:**
  - Consistency
  - Availability
  - Partitioning

Cluster filesystem

- **http://en.wikipedia.org/wiki/Clustered_file_system**

Data Transfer Utilities

- **The list of the protocols (quite often they are also utilities)**
  - http://en.wikipedia.org/wiki/List_of_file_transfer_protocols

Long distance Data Transfer

- **Long distance: in between cities, countries, continents, planets.**
- **Tasks:**
  - Reliable transfer;
  - Time of the transfer;
  - Volume of the transfer;
  - Interruption and restart the transfer;
  - Forecast when data transfer is accomplished;
  - API, Statistics.

Data Transfer systems

- Physics Experiment Data Export (**PhEDEx**) http://iopscience.iop.org/1742-6596/219/6/062010/pdf/1742-6596_219_6_062010.pdf
- FTS3 - https://svnweb.cern.ch/trac/fts3/wiki/UserGuide
- “Bittorrent”, http://www.bittorrent.com
- “GnuTella”, http://www.gnutella.com

End of Lecture

Shevel.Andrey@gmail.com