dCache.org

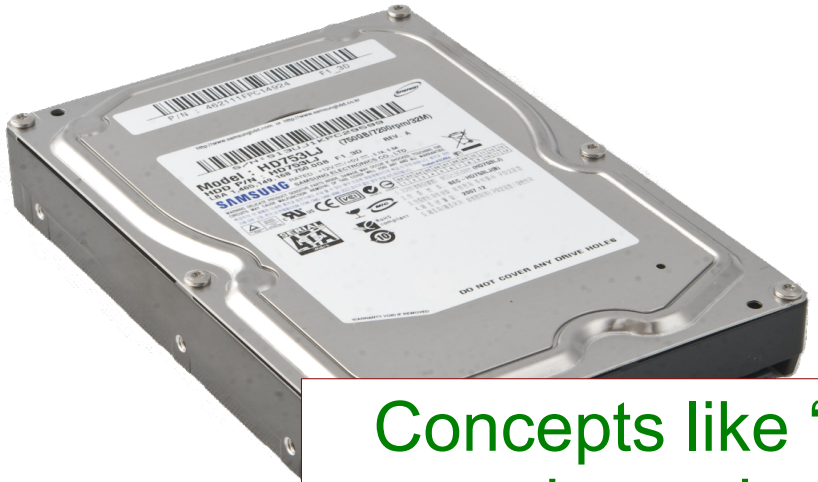
GLUE, StAR and dCache: a random walk in storage accounting

Paul Millar

Nordic Accounting Workshop
Uppsala, Sweden

Fermilab NDGF Nordic DataGrid Facility DESY HELMHOLTZ ASSOCIATION LSDMA

HDD: a short story with a moral



- LBA introduced in 1994.
- HDD simply has n-blocks (of fixed size)

Concepts like “Total Space” “Free space” depend on your storage model.

- Block stored “somewhere”
- Some capacity is inaccessible

  Used when block is difficult (but not impossible) to read

Who wants to know about accounting?

- The users
- The admins
- Local management
- Different collaborations

… and why?

Users:

Does the system agree with the amount of data I think should be there?

Can I upload more data?

Admins:

When do I need to buy more disks?

Local Management:

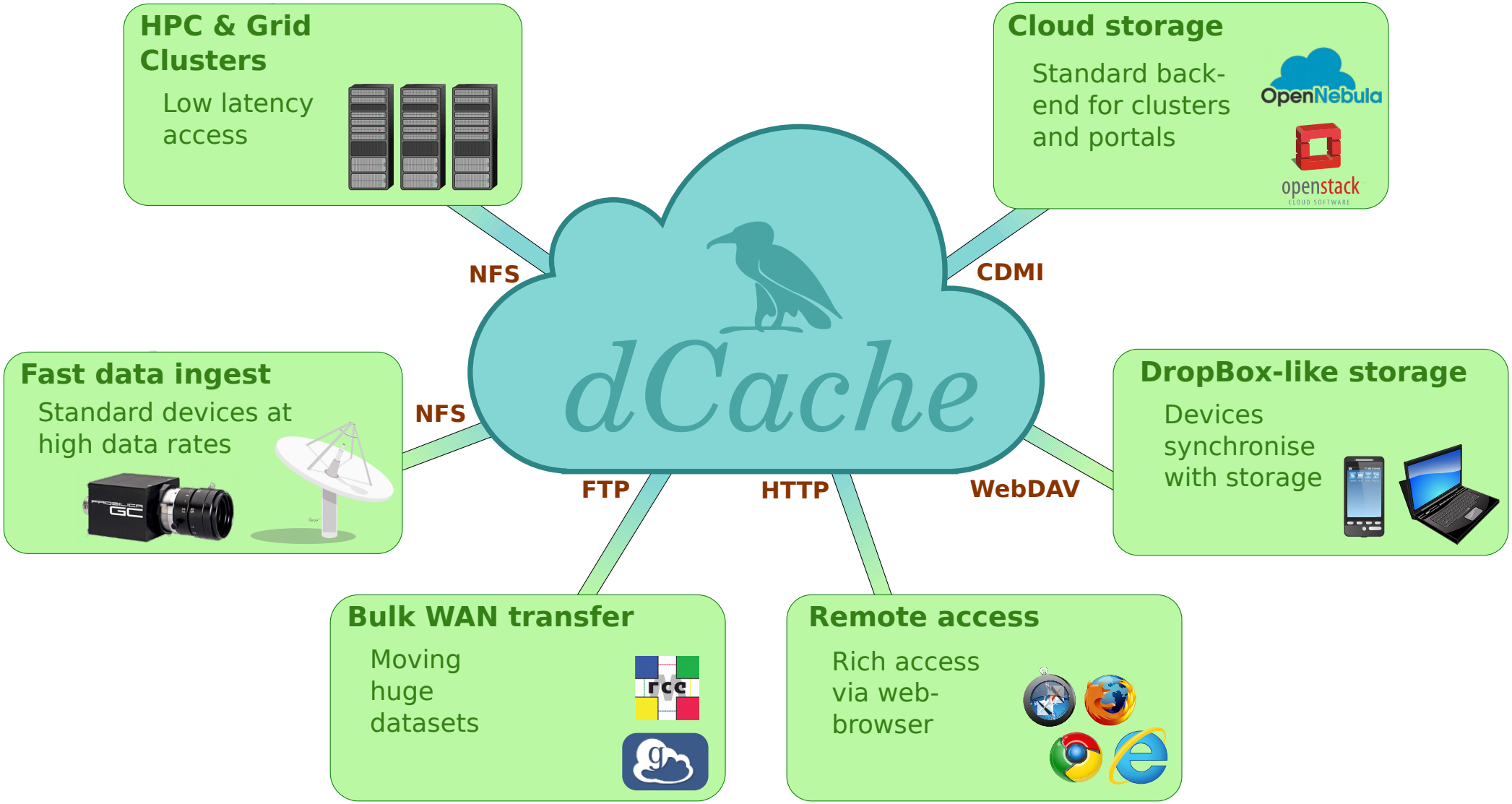
Is the storage being used efficiently?

How much to charge users (if pay-as-you-go model)?

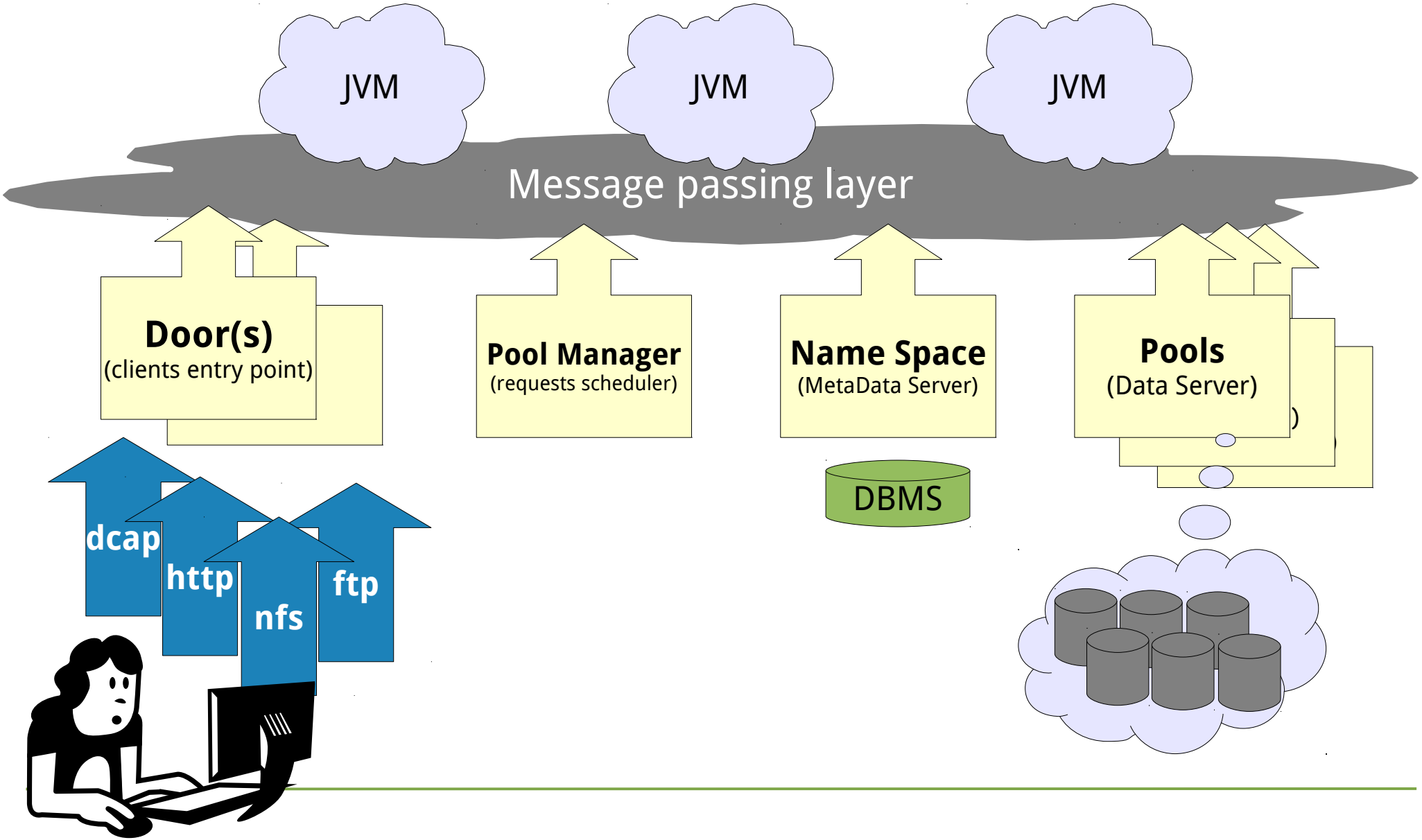
Collaboration:

Is the site providing the capacity they promised?

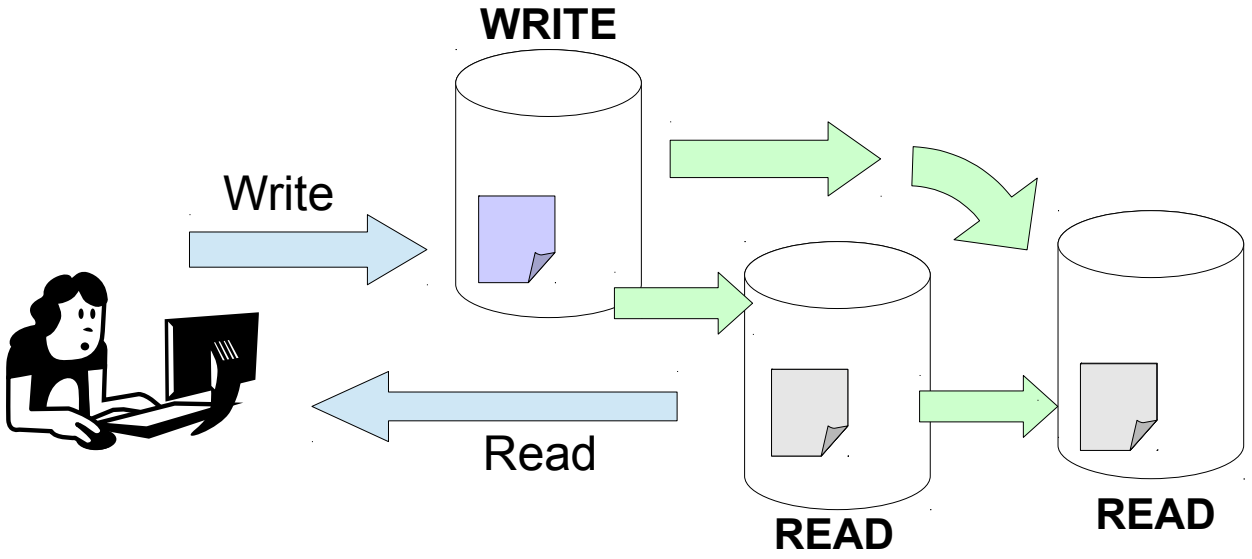
dCache: the scientific cloud

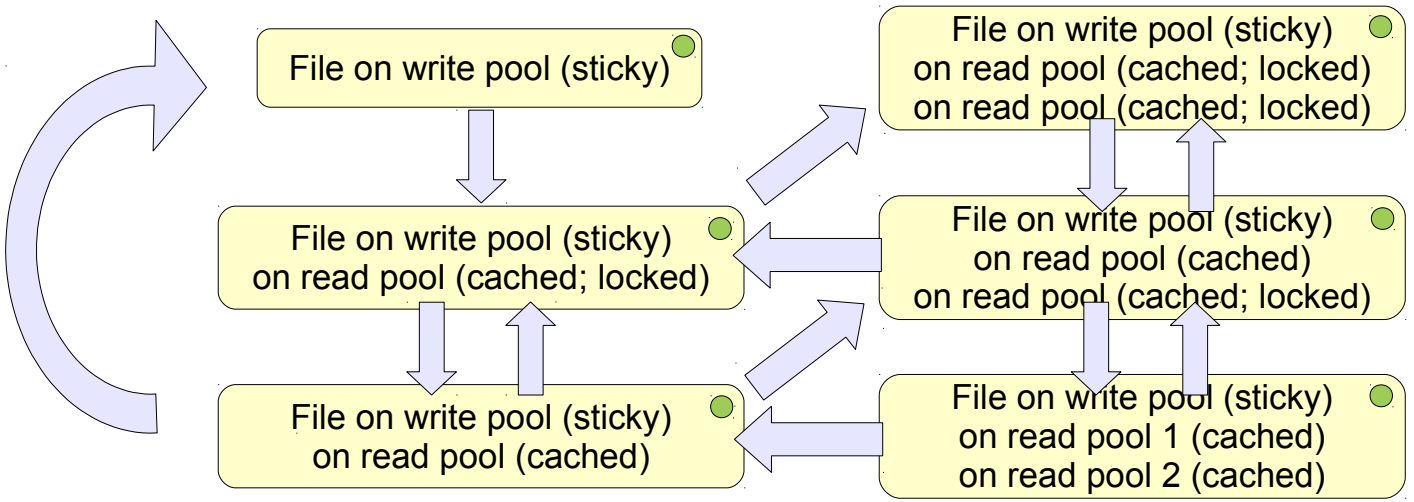


dCache – under the hood



How much space is used?





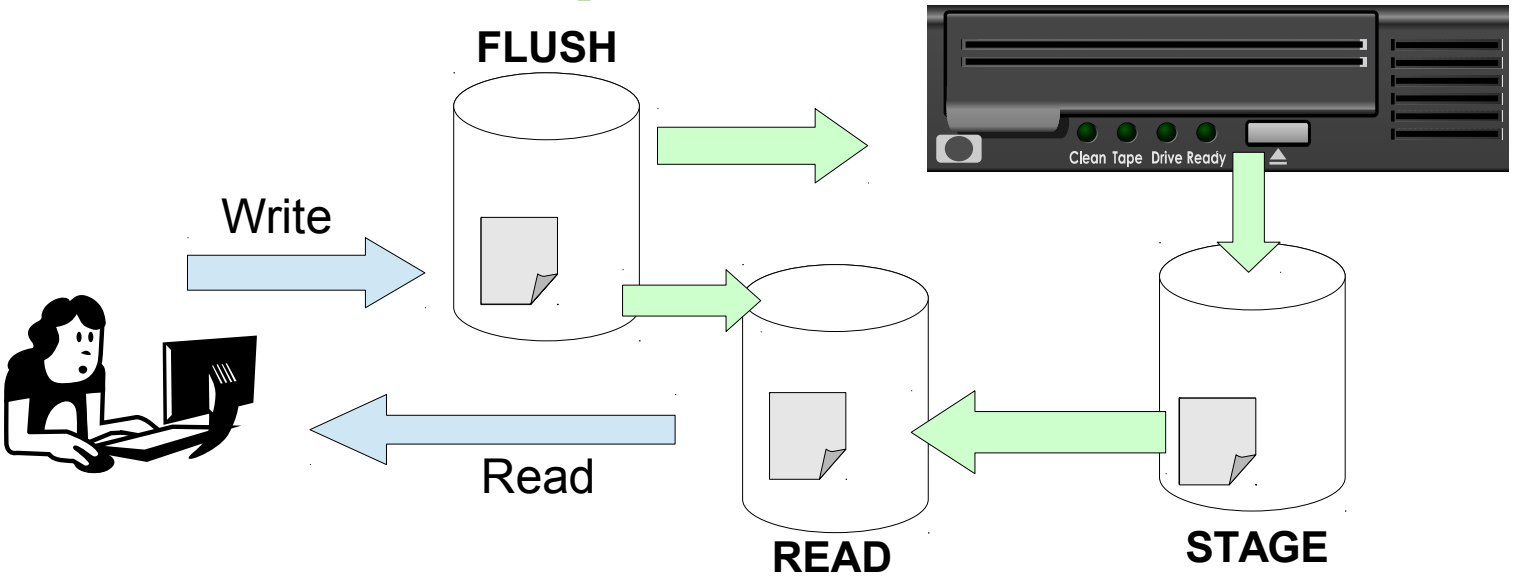
Now... tape

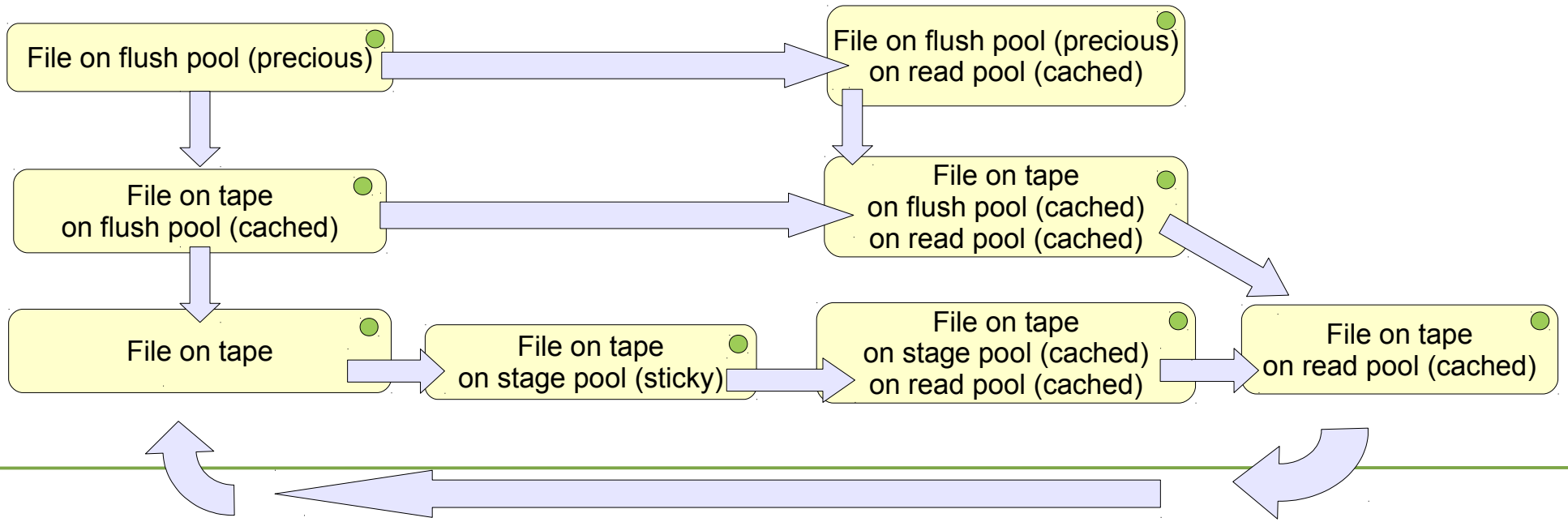
Not this

... or this

More this

How much space is used?





Some pitfalls in accounting

- **QoS problem**:
  - If QoS ensures there are two internal replicas of each file, what does Total, Used, Free mean?
  - If disk/tape stores files compressed or dedup, what is the used space?
  - If there are two QoS classes (1-replica & 2-replicas) what should Total, Used, Free mean?
- **Free space fallacy**:

  People like seeing a "free space" – they want to know if uploading their dataset will work.
- **The shared capacity**:

  Communities want to see how much space is available to them. What if they share resources: how much free space is there?

More than just capacity...

- Largely limited talk to **capacity accounting**: accounting at number of bytes.
- Storage Accounting can include **other aspects**:
  - Use of networks (WAN, LAN; read, write),
  - Object or file count,
  - Data churn (tape, SSD),
  - Cost of migrating between different QoS (e.g., tape → disk),
  - Power usage,

  …

Accounting: SRM and spaces

- SRM is a **standard protocol** for managing storage systems:

  Includes concept of "space reservations", a promises to store an amount of data with certain QoS.

- SRM is most widely used in **WLCG**:

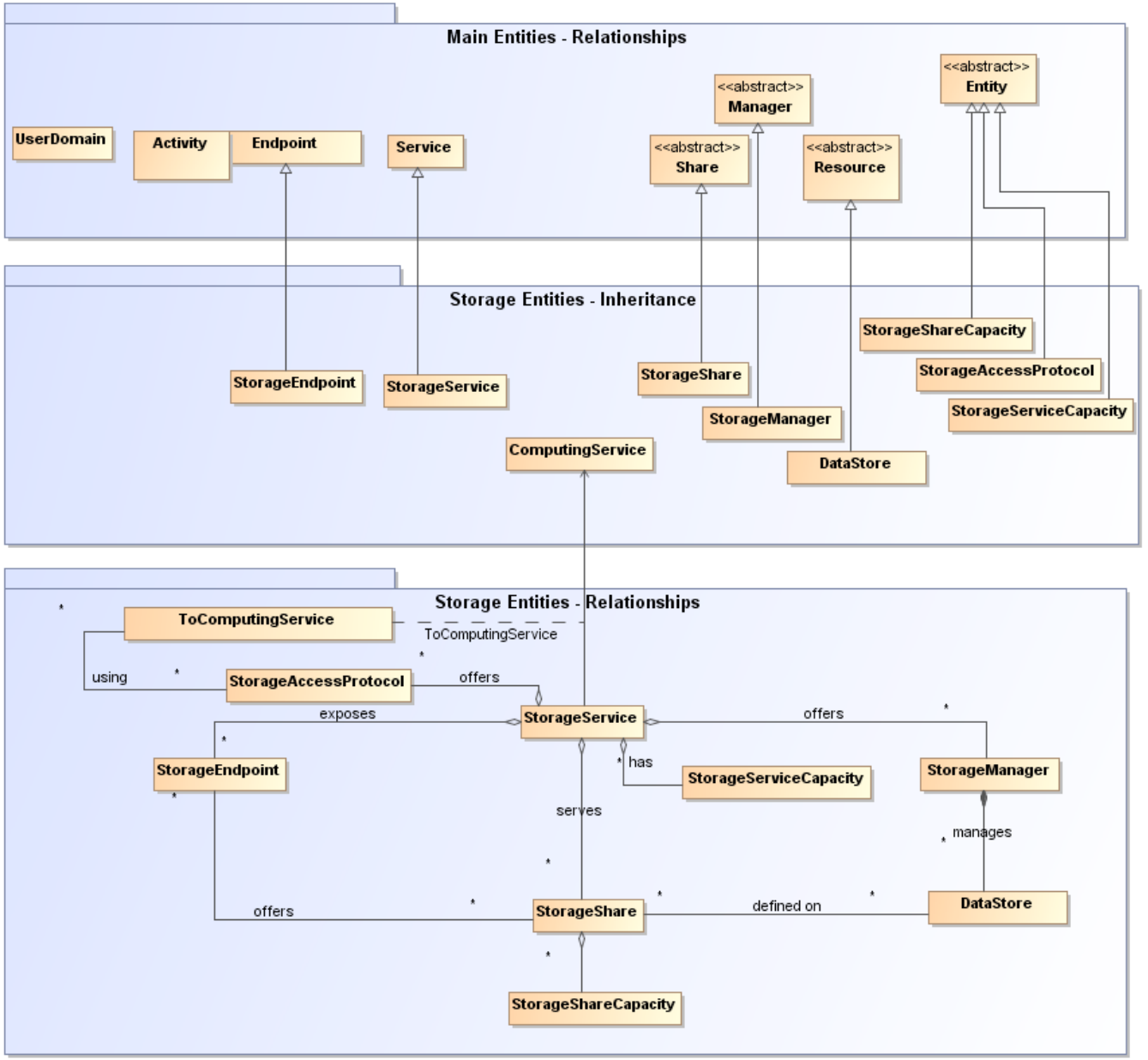
  Communities mostly use space reservations for accounting.

  Metrics are: **Total Size**, **Guaranteed Size**, and **Unused Size**

- WLCG MoU describes what these mean:

  Disk reservations: Unused Size = Total Size – the sticky replica.
  Tape reservations: Unused Size = available space on flush pools.

Accounting: GLUE



Accounting: StAR

- Standard format for describing capacity usage.
- Simplest mode: records describe current status, cut periodically

  More sophisticate modes (e.g., dynamic rates) are possible
- Record describes describes **only *used* capacity**:
  - Two reporting metrics: **LogicalCapacity** (opt) and **ResourceCapacity** (req).
  - Four profiling metrics: **StorageSystem** (req), **StorageShare** (opt), **StorageMedia** (opt), **StorageClass** (opt).
  - dCache publishes:

    StorageShare based on gid-ownership of files, StorageMedia as disk or tape, StorageClass based on tape-specific information.

    ResourceCapacity based on total number of replicas, LogicalCapacity based on bytes uploaded.

Accounting: CDMI

- CDMI is a standard **cloud storage protocol**.
- Includes an "Administrative Domain" concept:
  - Provides 16 different accounting metrics
  - It's cloud, so only used capacity is available
- Accounting information is provided for **specific periods**:
  - For each {object-count, bytes-used, energy-usage} concept, minimum, maximum, average and hour-product metrics are provided.
  - Also includes some counting metrics (e.g., files created)
- dCache currently doesn't publish any accounting information through CDMI.

Current status

- SRM: widely used in WLCG, but never gained much traction outside
  - WLCG considering moving away from SRM.
- GLUE: mostly used in WLCG with some other communities:
  - EGI FedCloud being an example of another community.
  - Accounting information available but not widely used.
- StAR: initial push within WLCG, but effort currently stalled.
  - Likely see increased usage in the future.
- CDMI: nice (but ambitious) protocol
  - remains open to what extent we'll see adoption.

Not covered in this talk

- Non-capacity accounting:
  - Network, power, data-churn, altering QoS, …
- Transport, infrastructure:
  - Data has to get somewhere to be useful.
- Visualisation, Analysis:
  - Turning data into information.
- Automated activity:
  - Agents reacting to accounting information.
- Privacy (legal) issues:
  - Who can know what, and at what granularity?
- Identity issues:
  - With a distributed community, do you know your users?

Summary

- Accounting is more difficult that people think:
  - Partly the problem is vocabulary/model: what does 'Free' really mean?
- Abstractions, models or profiles are necessary to compare different systems:
  - most times people don't realise they're describing a model.
  - must be worded ***very*** carefully to avoid ambiguity.
- No one existing solution for accounting:
  - StAR is a good start.
- Metrics only become reliable if someone cares.

Thanks for listening … any questions?