

Record Format for Storage Accounting

Henrik Thostrup Jensen <htj@ndgf.org>
Nordic Data Grid Facility
OGF 30, Brussels

Disclaimer

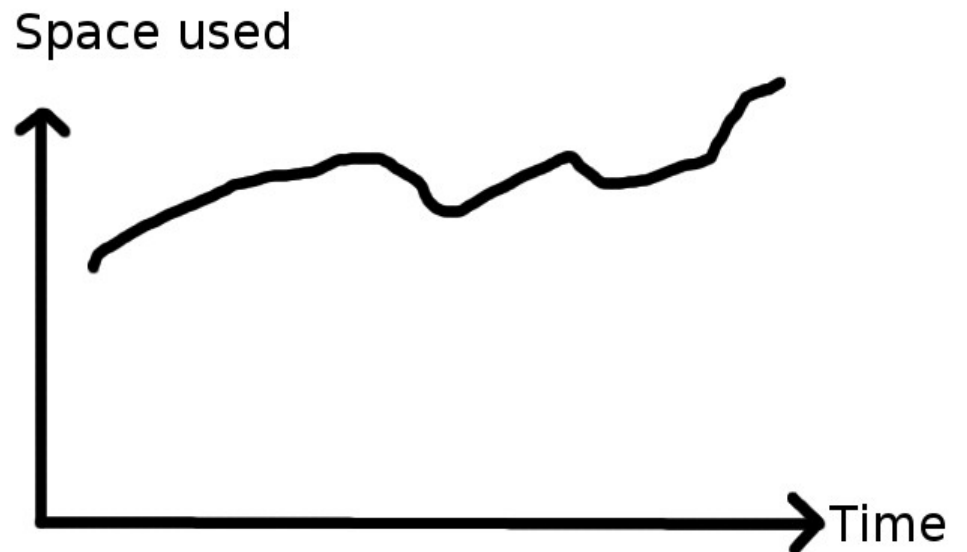
- Everything is tentative
- Original draft by Jon Kerr Nilsen and me
- I've included some changes after feedback from the mailing list

What to measure?

- Used space – yes!
- Reserved – probably
- Available space – perhaps
- Transfer statistics – probably
 - Resolution is the tricky part
 - Per site, storage element, partition, file?

Measuring Space

- Use is continuous
- Records are discrete
- Split into start + end + used bytes
 - Not unlike integration
 - Whatever time resolution you like



Measuring Transfers

- Bytes read
- Bytes written
- Bytes deleted
- I think we should have these
 - Just not per file

Concepts

- Where, what, who

Where (is data stored)

- Site
- Storage Element
- Storage Partition (part of SE, think pools)
- Storage Type (disk, tape, ssd)
- Partition and Type are optional
- Site and Element are perhaps optional

Structuring the where

- Is hierarchical
- Site
 - Storage Element
 - Storage Partition
 - Storage Type
- Some are optional
 - This becomes a bloody mess
 - Have flat structure, split into seperate records

What (is stored)

- Storage Class (replica, precious, pinned, etc.)
- File Count
- Directory Path
- Filenames?
- NDGF has 11M files
 - Medium WLCG site



Who (stores the data)

- Global
 - User Identity
 - VO data (vo name, group, role, issuer)
 - Names could be better (please suggest)
- Local
 - Local user name
 - Local group
- Can some of these elements be repeated?

UR2 / Seperate format

- EMI needs a record format soon (this year)
- UR2 will take time
 - No draft for service record (AFAIK)
- Implementation benefits of unified record structure is rather low
- Not against the idea, just the timespan
- Sharing lingo – yes please

UR2 – Meeting it halfway

- Make a standard for the identity block
 - (and use it in SAR)
- Create new UR 1.1 with the identity block
 - Missing VO information is biggest problem in UR
 - Of course a lot of people would like to add something

Complete Record Format

- StorageRecord
 - recordId + createTime (similar to UR fields)
 - StartTime, EndTime
 - SpaceUsed, SpaceReserved, SpaceAvailable
 - BytesRead, BytesWritten, BytesDeleted
 - Site, StorageElement, StoragePartition, StorageType
 - StorageClass, FileCount, DirectoryPath, **FileNames**
 - UserIdentity, VOInformation, LocalUserName, LocalUserGroup