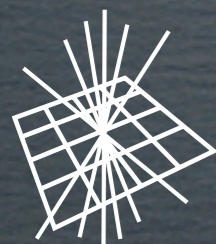# ATLAS Requirements and Plans for SRM2

Graeme Stewart
University of Glasgow

GridPP
UK Computing for Particle Physics

UNIVERSITY
of
GLASGOW

Overview

◆ The ATLAS Computing Model

◆ Storage Essentials

◆ SRM v2.2

  ◆ Current Tests

  ◆ New Features

  ◆ Information System and Client Tools

◆ CCRC and FDR

◆ Conclusions

SRM v2.2 Workshop, Edinburgh

# ATLAS Computing Model

- Tier0
  - Copy RAW data to Castor tape for archival
  - Copy RAW data to Tier-1s for storage and reprocessing
  - Run first-pass calibration/alignment (within 24 hrs)
  - Run first-pass reconstruction (within 48 hrs)
  - Distribute reconstruction output (ESDs, AODs & TAGS) to Tier-1s
- Tier-1s:
  - Store and take care of a fraction of RAW data
  - Run "slow" calibration/alignment procedures
  - Rerun reconstruction with better calib/align and/or algorithms
  - Distribute reconstruction output to Tier-2s
  - Keep current versions of ESDs and AODs on disk for analysis
- Tier-2s:
  - Run simulation
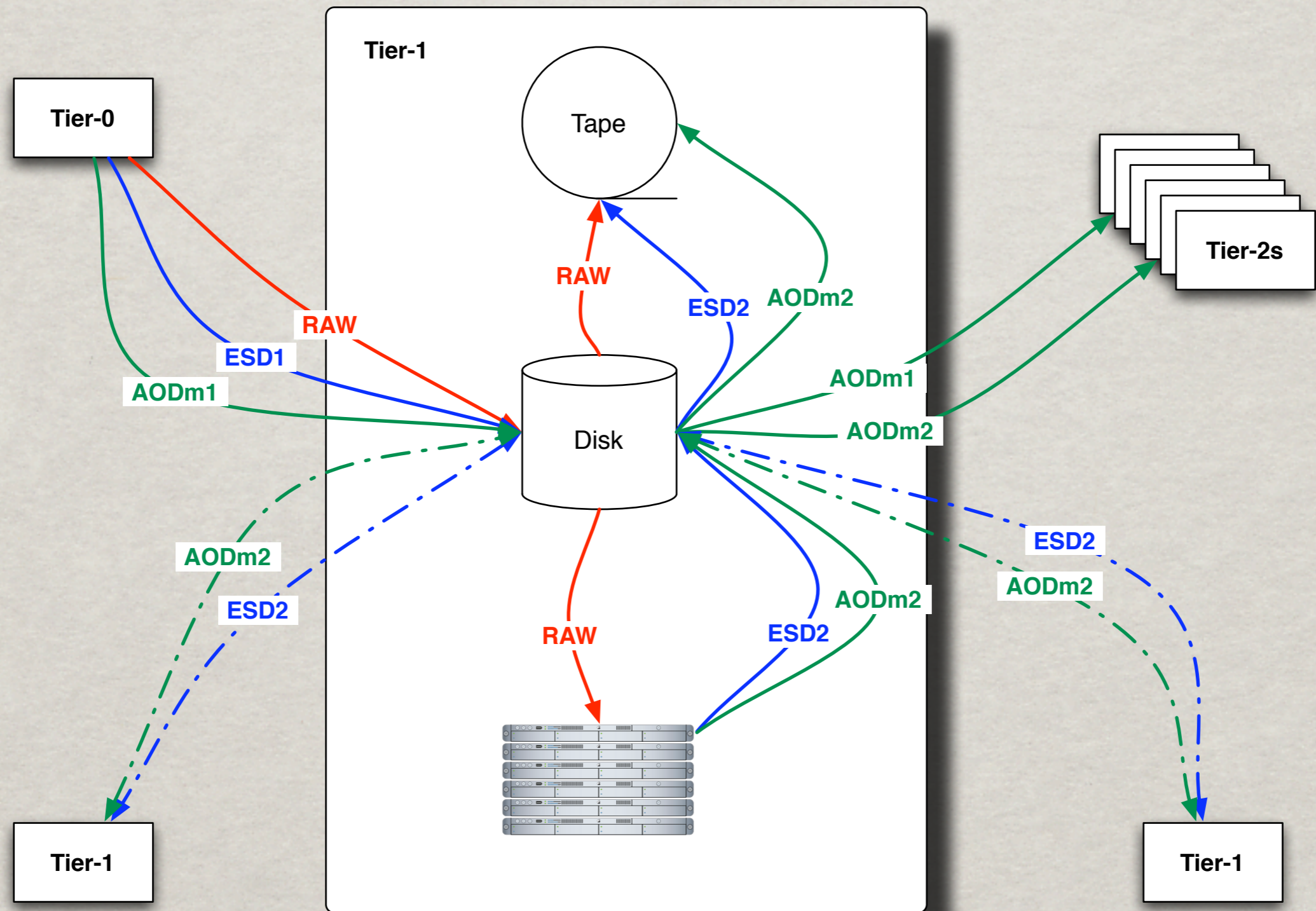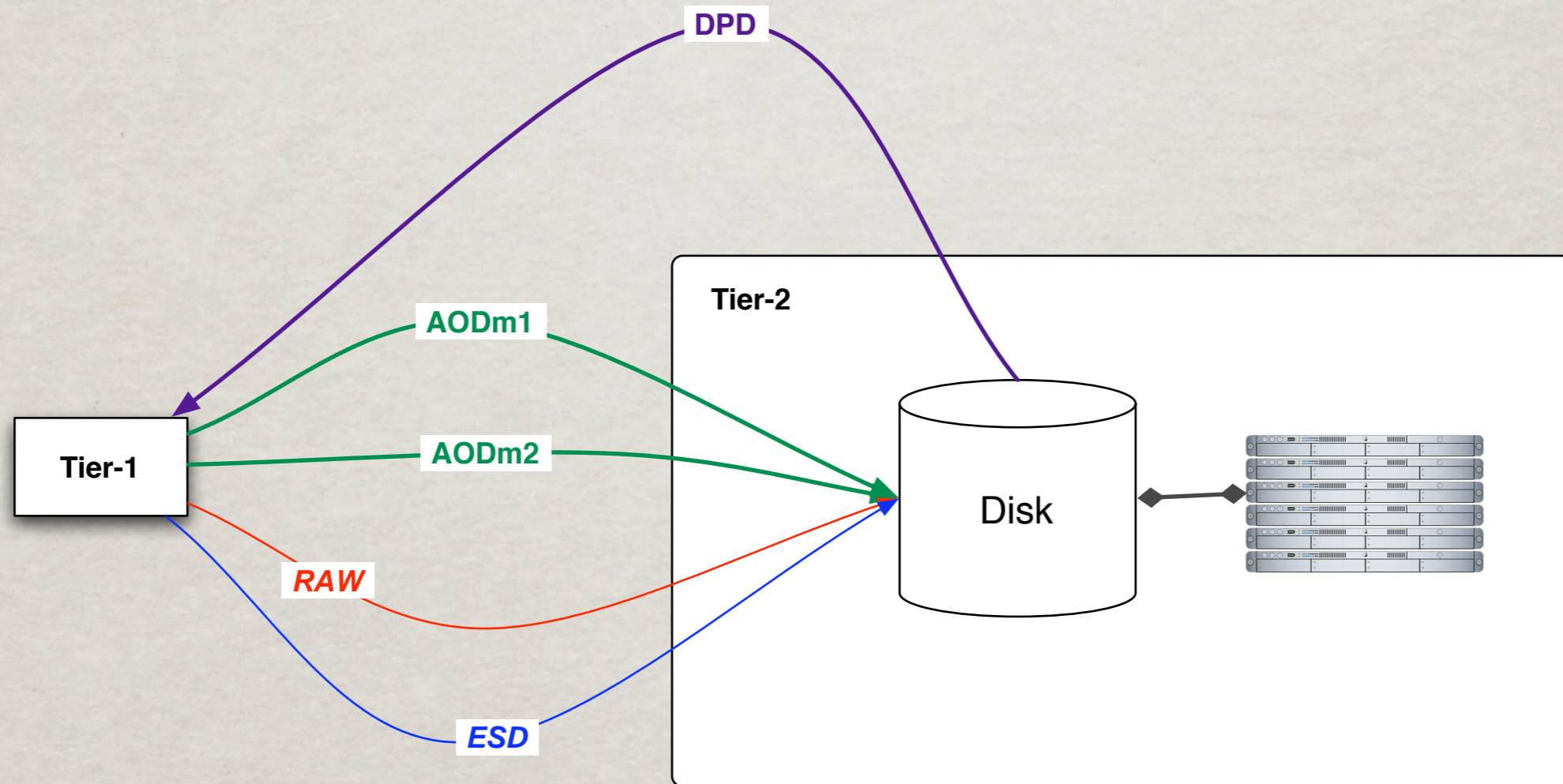 - Keep current versions of AODs on disk for analysis
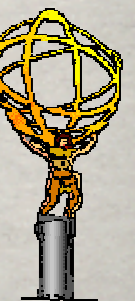
Overview of Real Data Flows
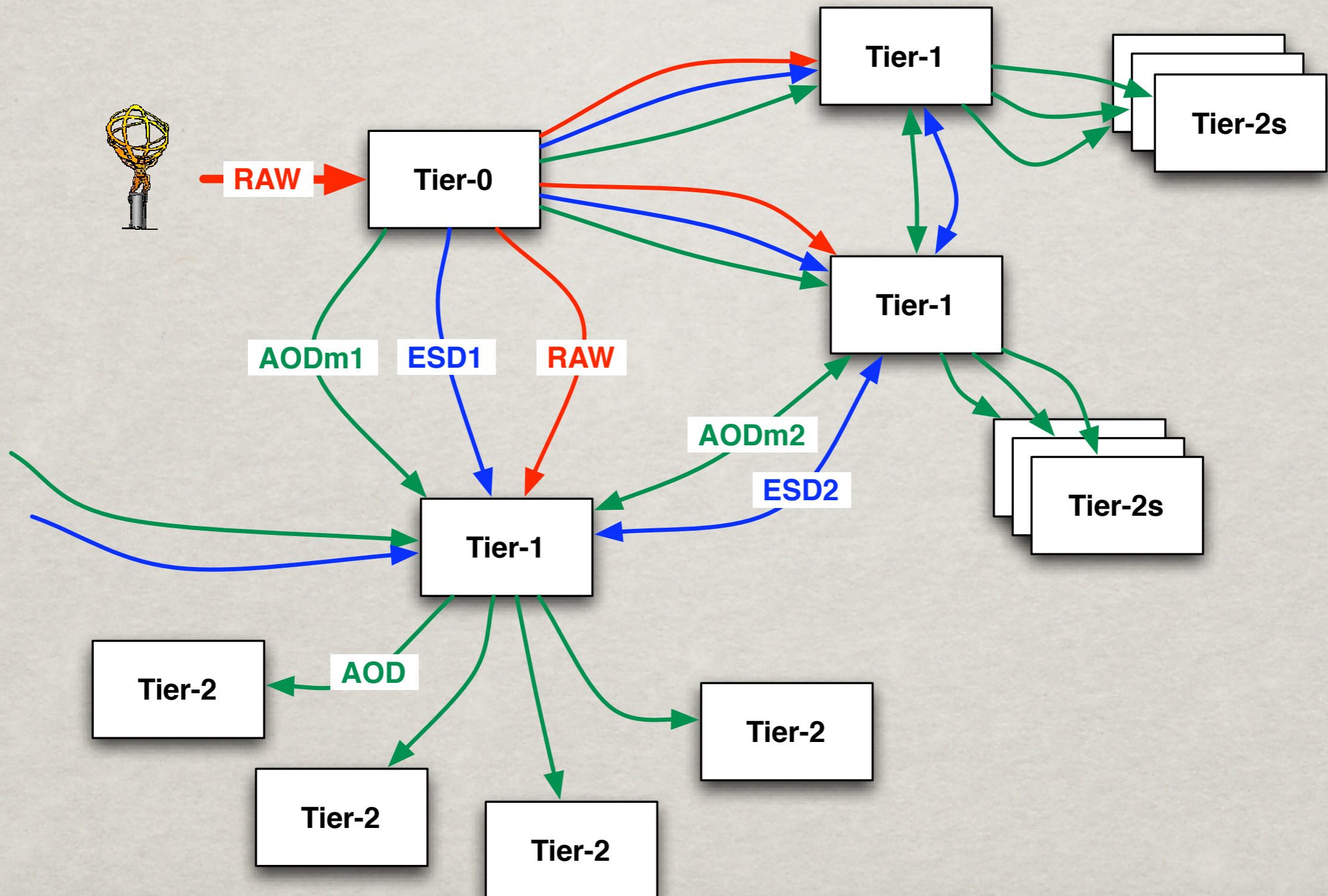
At The Tier-0

At The Tier-1

At The Tier-2

SRM v2.2 Workshop, Edinburgh
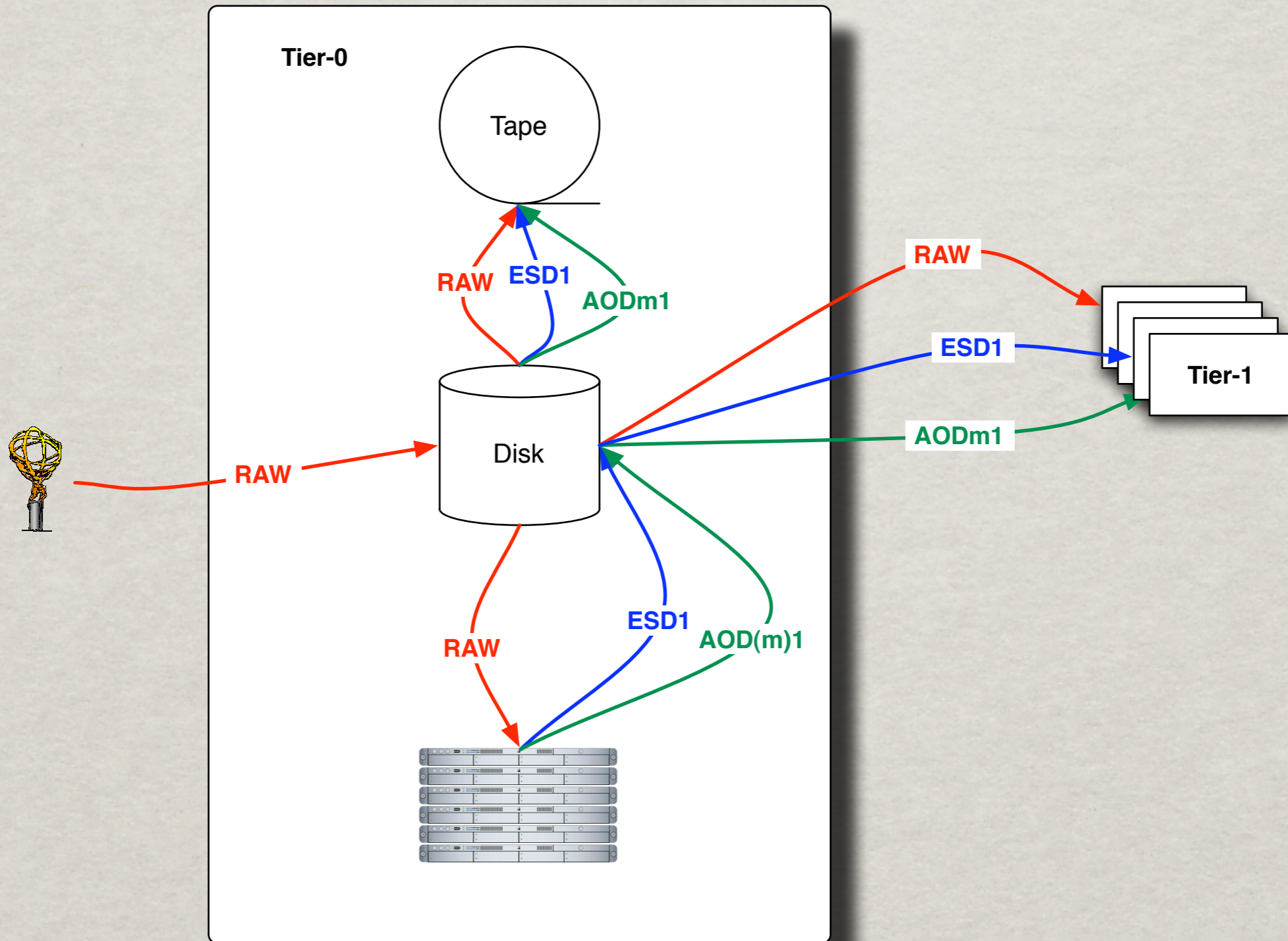
MonteCarlo Production

- ◆ ATLAS MonteCarlo production constantly simulates events, digitises, reconstructs, etc. at T1 and T2

- ◆ This provides flows of simulated RAW, ESD and AOD from T2s to their T1

  - ◆ (There will be attempts to zip, e.g., hits files on the T2, but the T1 will still need to run merge jobs)

- ◆ Extensive use of DDM to stage input data to T2s and recover outputs from T2 to T1

Some Notes

- All data flow is mediated by ATLAS Distributed Data Management system (DDM)

- Actual transfers in and out of SEs done by FTS

- A very important use case is *reprocessing* - recover RAW from tape and reprocess to generate new ESD and AOD

  - We also have to exercise expected failure modes (disk array on fire!)

Essential Features of Storage for ATLAS

1. Storage systems should be reliable and robust.

2. If failures do occur error, then error messages should be clear and consistent.

3. We need to know what's on tape and what's on disk.

(The brief answer to the title of this session are the 3 points above.)

Report Card: SRM v1.1

1. Reliable and robust?

   - *No.*

2. Clear error messages?

   - *No.*

3. What's on tape and what's on disk?

   - *Not really.*

SRM 1.1's only advantage is a long history in production already - it's the devil we know.

SRM v2.2: New Kid on the Block

- ◆ ATLAS have been testing SRM v2.2

  - ◆ However these tests are still an an early stage

- ◆ Moving fully to SRM v2 will imply a lot of changes to ATLAS Distributed Data Management

  - ◆ So we want to get it right

  - ◆ And we plan to incrementally use SRM v2.2 features

SRM v2.2 Tests: dCache

SRM v2.2 Tests: StorM

SRM v2.2 Tests: Conclusions So Far

◆ SRM v2 works well as a replacement for SRM v1

◆ For dCache and STORM at least, reliability seems good

◆ Unfortunately problems with the DDM configuration meant the test endpoints got muddled with production

◆ So we need new rounds of testing

　◆ We haven't tried any new features!

SRM v2.2 Workshop, Edinburgh

SRM v2.2 Report Card

◆ Reliable and robust?

◆ Dedicated tests for dCache, STORM went well (errors were from CASTOR source)

Very hopeful on this front - we really do need improvements in this area.

GSSD Stress tests are very valuable; however, real production scale tests of SRM v2.2 have not yet happened.

SRM v2.2 Report Card

- Good error messages?

  - Significant improvements expected here

    - FTS 2.0 already categorises errors as belonging to source or destination

    - We need this to be exposed to ATLAS tools so DDM shifters can take appropriate actions

Looks as if we can expect good news here too

Space Tokens and New Functionality

- ATLAS requirement at T0 and T1 is to know what's on tape and what's on disk (online and nearline, if you prefer)

  - Note we don't want too much complexity or abstraction from our point of view

- At the moment we have not tested space tokens, and they are not integrated yet into DDM

  - Aim will be to start simple, so we shall ask, in the first instance

    - ATLAS_DISK

    - ATLAS_TAPE

  - At T2s we will probably not declare space tokens at all initially, i.e., we will be using the "default" (whatever that is!)

Space Tokens in the Future

◆ Dependent on the success of basic use of space tokens we can envisage extending to

    ◆ A T1D1 token for cloud produced AOD and ESD

    ◆ Extending the use of space tokens at T2s to protect production areas, e.g.,

        ◆ ATLAS_PROD

        ◆ ATLAS_USER

◆ But nothing is settled yet (sorry!)

◆ Operation and management of space tokens is almost at important as functionality

◆ Dynamic space reservation *will not be used* by users, but could be explored for high level user tools, e.g., ganga

SRMLs

◆ Is a *good thing* and will help in data placement (recovery from aborted or partial transfers)

◆ It will also help greatly in validating the contents of the LFC

　　◆ (This is envisaged as a constant background task)

Staging Files: srmBringOnline

- Reprocessing requires recall of RAW data from tape

  - This can be seriously painful

- srmBringOnline would provide a good way of doing this across T0/1 implementations

  - But it needs to be tested

    - (There are always nasty ways to recall files from tape!)

Glue Information System

- ◆ ATLAS has partly isolated itself from problems with the grid information systems

  - ◆ It's unacceptable to drop important sites because of BDII problems

- ◆ However, information is gathered from here to populate more dynamic systems, like pilot scheduler systems

  - ◆ So accurate publishing of information through GLUE will help us a lot

Client Tools

- We'd like command line tools to issue, e.g., srm-ls and srm-stage

  - These should be scriptable

  - Should also not *need* to use BDII

- Python bindings for SRM access will also be very useful

  - Thread safe, no BDII, catalog plugins

SRM v2.2 Workshop, Edinburgh

THE BIG TESTS
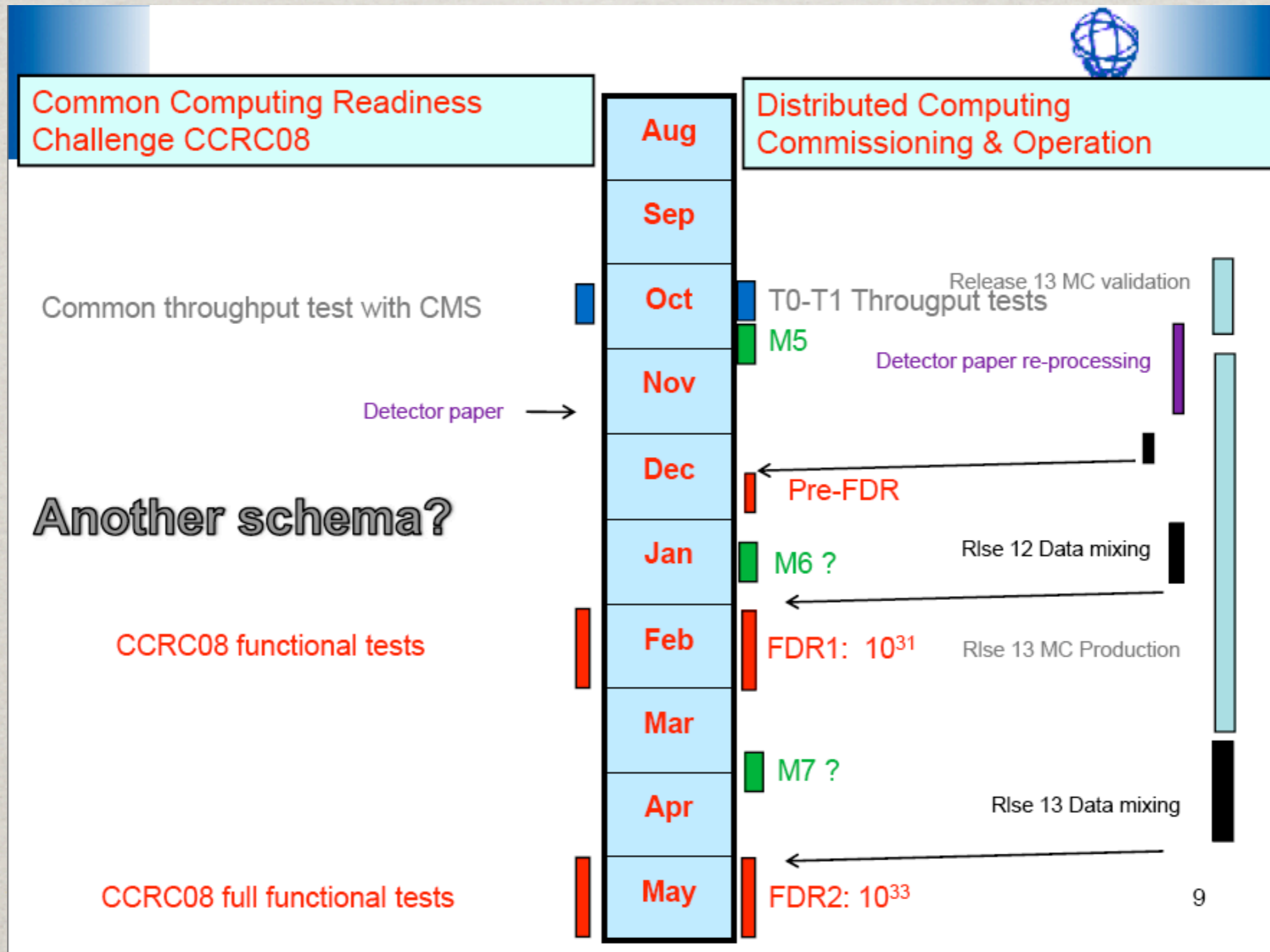
Testing Schedules

- Two weeks of dedicated SRM v2.2 testing in November (i.e., now)

- After this, we, anticipate moving to using SRM v2.2 endpoints as these become available at sites

- In FDR 1 we must exercise reprocessing - plan is to use SRM v2.2 functionality for this

- Conversion of DDM to use space tokens properly will happen in time for FDR2

SRM v2.2 Workshop, Edinburgh

Conclusions

- GSSD group have done a good job (well done!) of guiding SRM v2 development and deployment

- ATLAS certainly hope to be able to move to SRM v2 smoothly in advance of data taking

- New features will be tested at scale and incorporated more widely into the system if successful

- At the moment there isn't really a recipe :-(

- It will be a process of evolution and we'll have to work together to ensure success

SRM v2.2 Workshop, Edinburgh