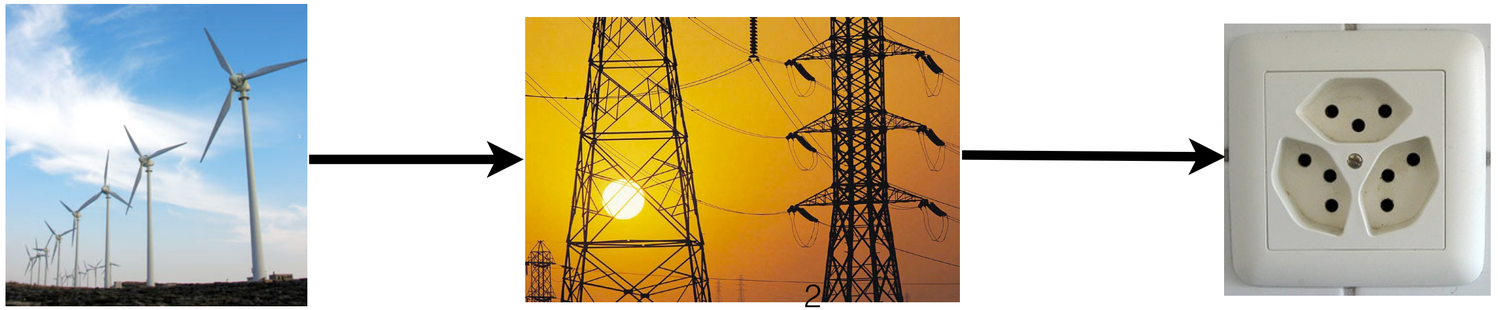
Introduction to the Grid

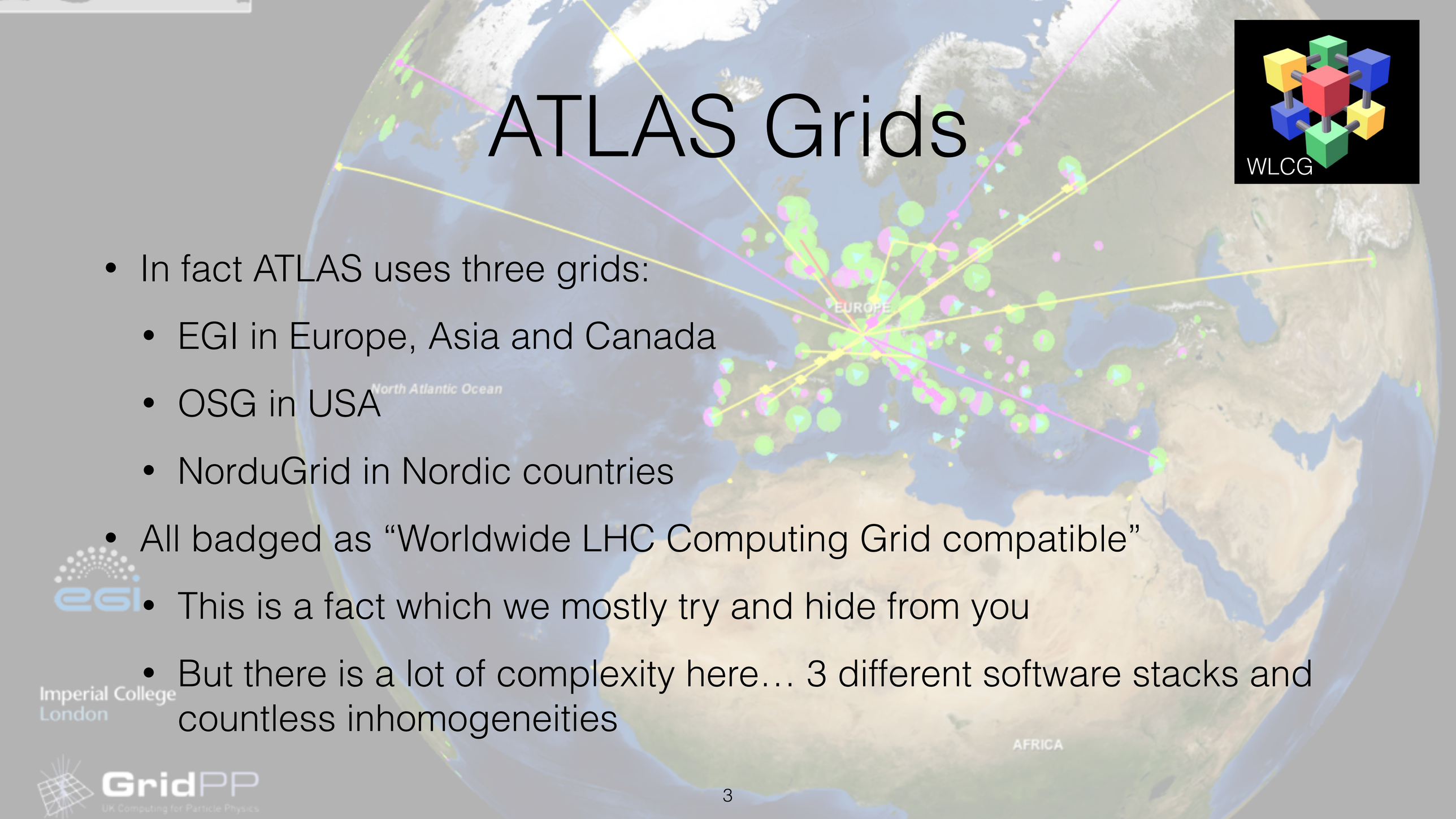
Graeme Stewart, EP-SFT

What's all this grid nonsense then?

- Recognition in the early planning of LHC computing that a single centre would not supply enough resources for the LHC experiments
  - Mainly because no one would fund a central computing facility
- Idea of a *computing grid* was current in IT circles - seamless connection of widely distributed storage and computing resources



ATLAS Grids



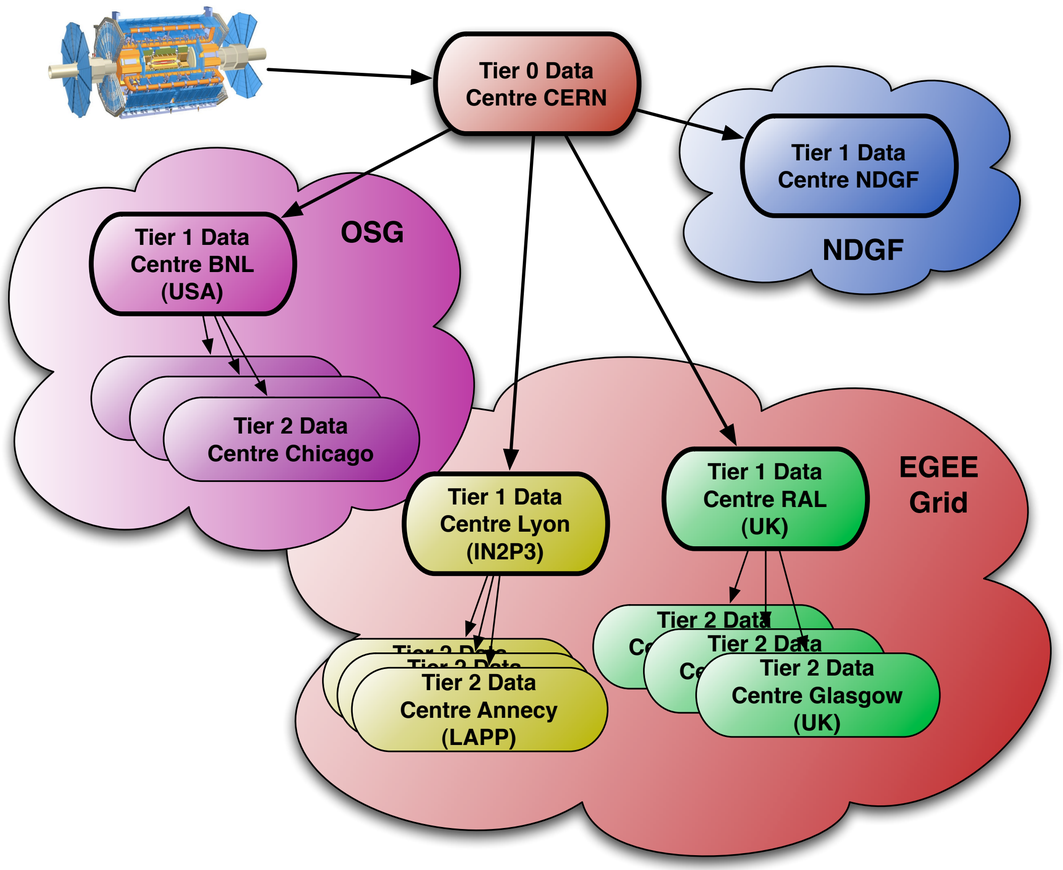
- In fact ATLAS uses three grids:
  - EGI in Europe, Asia and Canada
  - OSG in USA
  - NorduGrid in Nordic countries
- All badged as “Worldwide LHC Computing Grid compatible”
- This is a fact which we mostly try and hide from you
- But there is a lot of complexity here… 3 different software stacks and countless inhomogeneities

From Vision to Reality

- We had to scale back on many (perhaps inflated) expectations of what grids would deliver
  - Even basic services proved hard to delivery reliably in a highly distributed and heterogeneous environment
- We then built an experiment framework on top of this to deliver services we need
  - Data management
  - Book keeping
  - Production System
  - Distributed Analysis Framework
- But this has been in quite mature state and working well for years

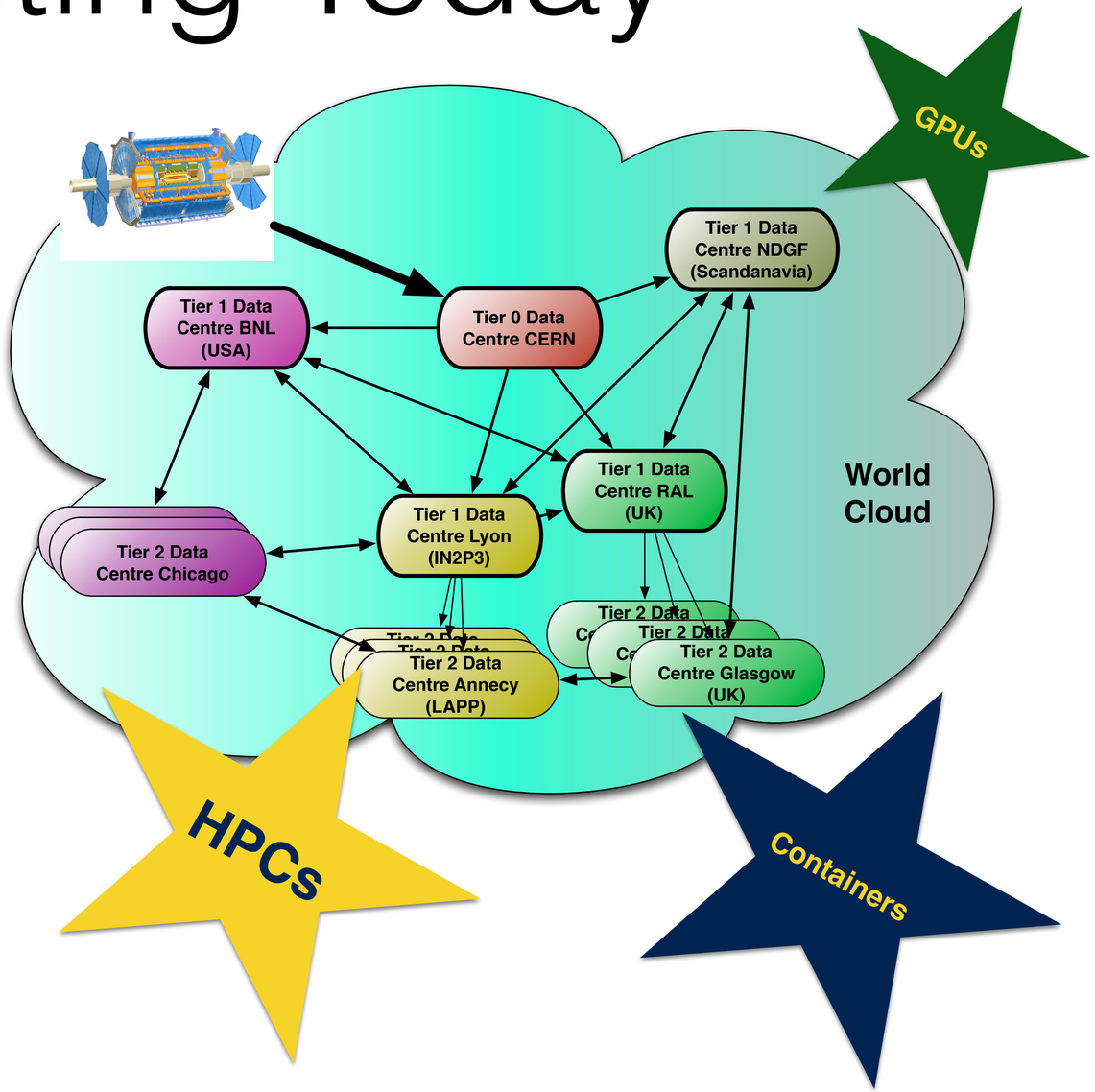
Ye Olde Computing Model Cartoon...

- 1 Tier-0: CERN
- 11 Tier-1: National Computing Centres (BNL, RAL, IN2P3, ...)
- 40 Tier-2: Regional Computing Centres (ScotGrid, Frascati, Toronto, ...)
- Composed of multiple individual sites
  - ~100 Analysis queues in PanDA
- There were quite well defined (and rather rigid) roles for each Tier



ATLAS Computing Today

- We became a lot more flexible in what we do where
  - Increases efficiency
- We removed artificial network boundaries
  - Base line mesh is ALL to ALL
- We also began to integrate new resource types
  - Today we're as happy with cloud resources as 'traditional' grid sale ones
    - In fact the whole grid evolved, adapting to new cloud technology and software
  - Plus HPCs
  - Plus Containers
  - And now with added GPUs…



ATLAS Grid Architecture

- Distributed Data Management System (DDM): Data movement and catalog system, policy engines, quotas
- Other databases and data book keeping (AMI, GRL, Coma for metadata; conditions data in distributed Oracle + frontier; Event Index; etc.)
- Production System (PanDA)
- Distributed Analysis Interface (pathena/prun/pbook)

Data Movement

- As experiment has matured so has data placement
  - AOD is mature and useful and is made widely available
    - But it's too large to run over at the single analysis level
    - We keep fighting to get its smaller (AMSG-X)
  - DAOD is made in the production system for individual analyses
    - Sometimes several rounds of reduction, depending on the use case
- Still, in general data is the harder thing to move around
  - *Jobs go to the data is still better*
- However, you can use the DDM tools to get, small samples of data to test your code or to retrieve your final outputs
- The informal rules:
  - 10GB/day - ok; 100GB/day - still ok; 1TB/day - not ok
  - Large data movements should be requested and managed via Rucio

User Disk Areas

- User output written by default to SCRATCHDISK
  - Volatile with cleanup on demand
- Most T2s have a component of grid storage called LOCALGROUPDISK
  - This is permanent DDM accessible storage
  - Just how ‘local’ this is varies, e.g., in the UK it’s UK wide
  - You can ask for data to moved here as normal
    - It’s also possible to get your analysis output subscribed auto-magically
- Physics and performance groups also have some management control over official ATLAS disk space
  - Contact your group space manager

Send in the Jobs!

- To run ATLAS jobs on the grid you need an tools which
  - Understand the ATLAS data model and data access
  - Understand the grid
  - Understand the analysis code
- You met these tools on Wednesday (Rucio and Panda)… so hopefully you made friends with them already