Science & Technology
Facilities Council

Service Evolution Ideas

Andrew Lahiff, Ian Collier

GridPP 37, Ambleside

1st September 2016

Introduction

- In July there was an explicit question, and some follow up discussion, on TB-SUPPORT about whether larger site(s) could run services for smaller sites
- This talk begins exploring some scenarios arising form this idea
- The aim is to seed a discussion – not particularly to advocate for any of the scenarios

Introduction

- There are many common services at sites (Grid + other)
  - CE, SE, BDII, Argus
  - monitoring, logging, accounting
- Large sites run these services anyway
  - And have larger teams dedicated to them
  - Running additional instances doesn't require much effort
- Could large sites run some of these services on behalf of smaller sites?
  - resulting in less duplication of effort across multiple sites
  - freeing up time of admins at the smaller sites
  - instead of maintaining grid middleware, people can spend more time improving computing for GridPP and science in general

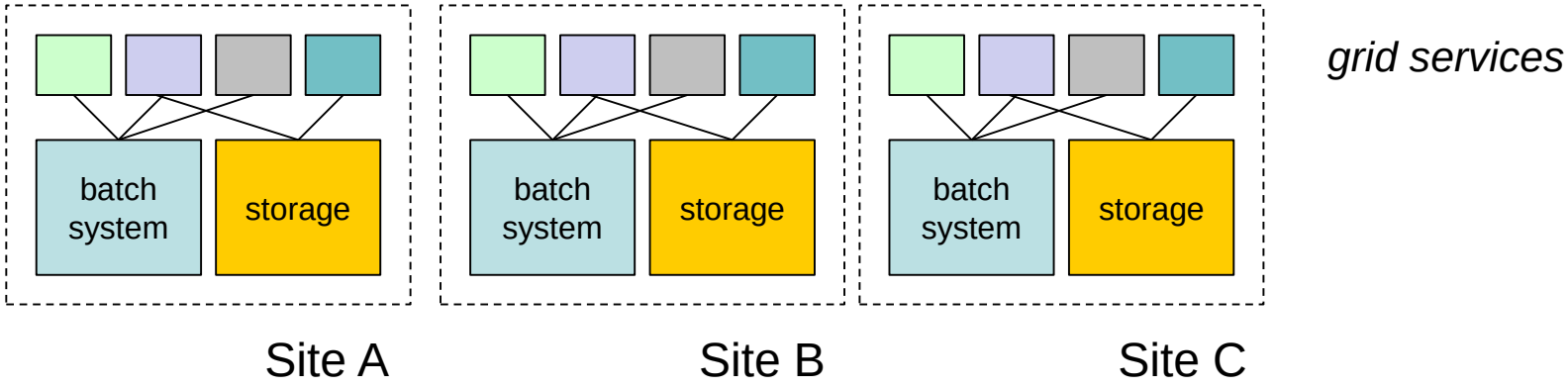
Introduction

- The idea of a distributed site is of course not new, e.g.
  - NDGF Tier-1
  - ATLAS Midwest Tier 2
  - GRIF

  but remotely running Grid services is not commonplace
- Why not use Vac?
  - Not everyone wants to run jobs in virtual machines
  - It's not appropriate at all sites
    - e.g. if local batch is an important (or dominant) use case

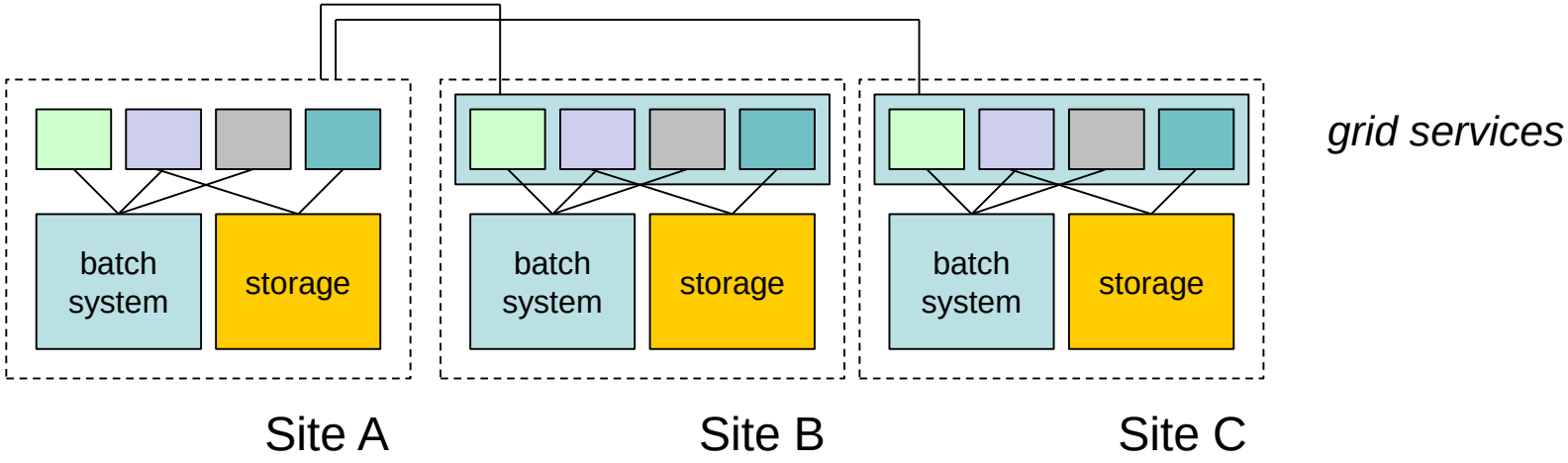
The situation today

- Similar grid services deployed at multiple sites
  - effort required to manage these services at each site
  - duplication of effort



Possible alternative #1

- A larger site remotely manages Grid services ***at*** the smaller sites
  - 2 variations of this proposed in talks at the May GDB (OSG)



Possible alternative #1

Pros

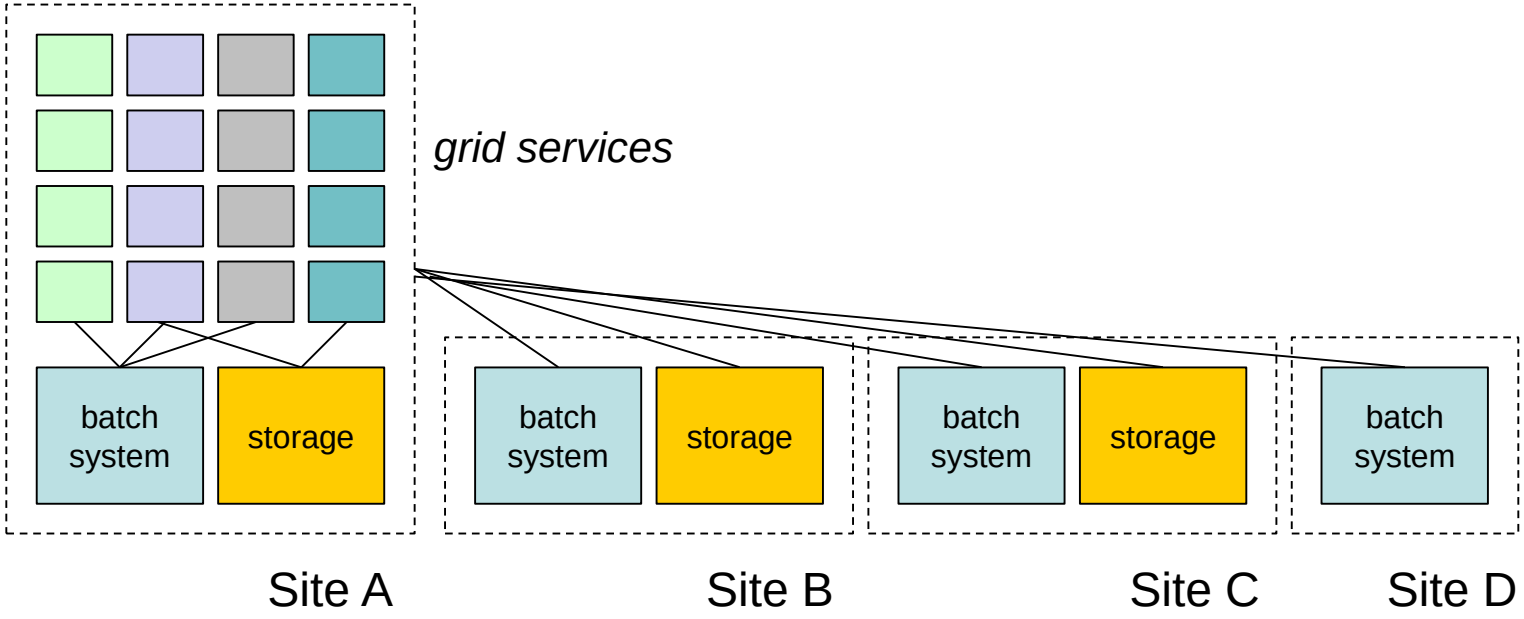
- Smaller site
  - less effort required
  - still a traditional grid site

Cons

- Differences in configuration management systems in use
- Sites unlikely to have existing cloud infrastructure which could simplify deployment of services

Possible alternative #2

- A larger site provides multiple instances of grid services
  - Smaller sites just run a batch system (& storage)



Possible alternative #2

Pros

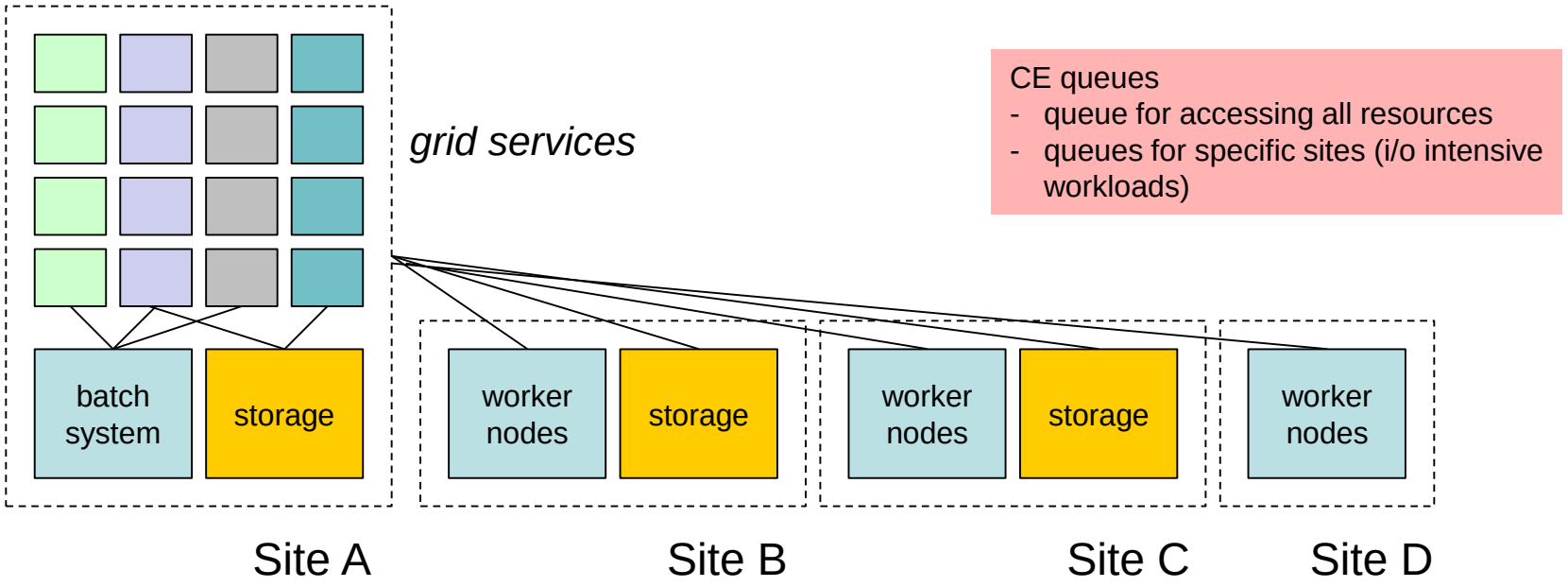
- Smaller site
  - less effort required
  - runs only more standard services (batch system, storage)
  - still a traditional grid site
- Larger site
  - already runs (many instances of) Grid services: additional instances don't require a huge amount of effort

Cons

- Need to think carefully about security
- A large site becomes a single point of failure for multiple smaller sites
- Possible network latency issues
- More scope for networking problems affecting operation

Possible alternative #3

- Smaller sites run worker nodes, squids (& storage)
  - distributed batch system with centralized CEs



Possible alternative #3

Pros

- Smaller site
  - much less effort required
  - still a traditional grid site
- Larger site
  - already runs Grid services: additional instances don't require a huge amount of effort

Cons

- Smaller site
  - Job submission for local users more complex
  - Needs to ask the larger site to make fairshare changes

Examples

- Services that could be run at a larger site
  - PhEDEx (CMS)
    - this has been done in the UK for many years
  - Site BDII
    - smaller site would need to ensure site firewall holes for resource BDIIs for services still at the site
  - (ARC) CE, Argus, batch system central managers
    - main complication would be pool accounts: need consistency with worker nodes at the remote site
  - Monitoring & logging
    - Metrics (e.g. InfluxDB, Grafana) straightforward
    - ELK: some thought about security required

Questions/Comments?