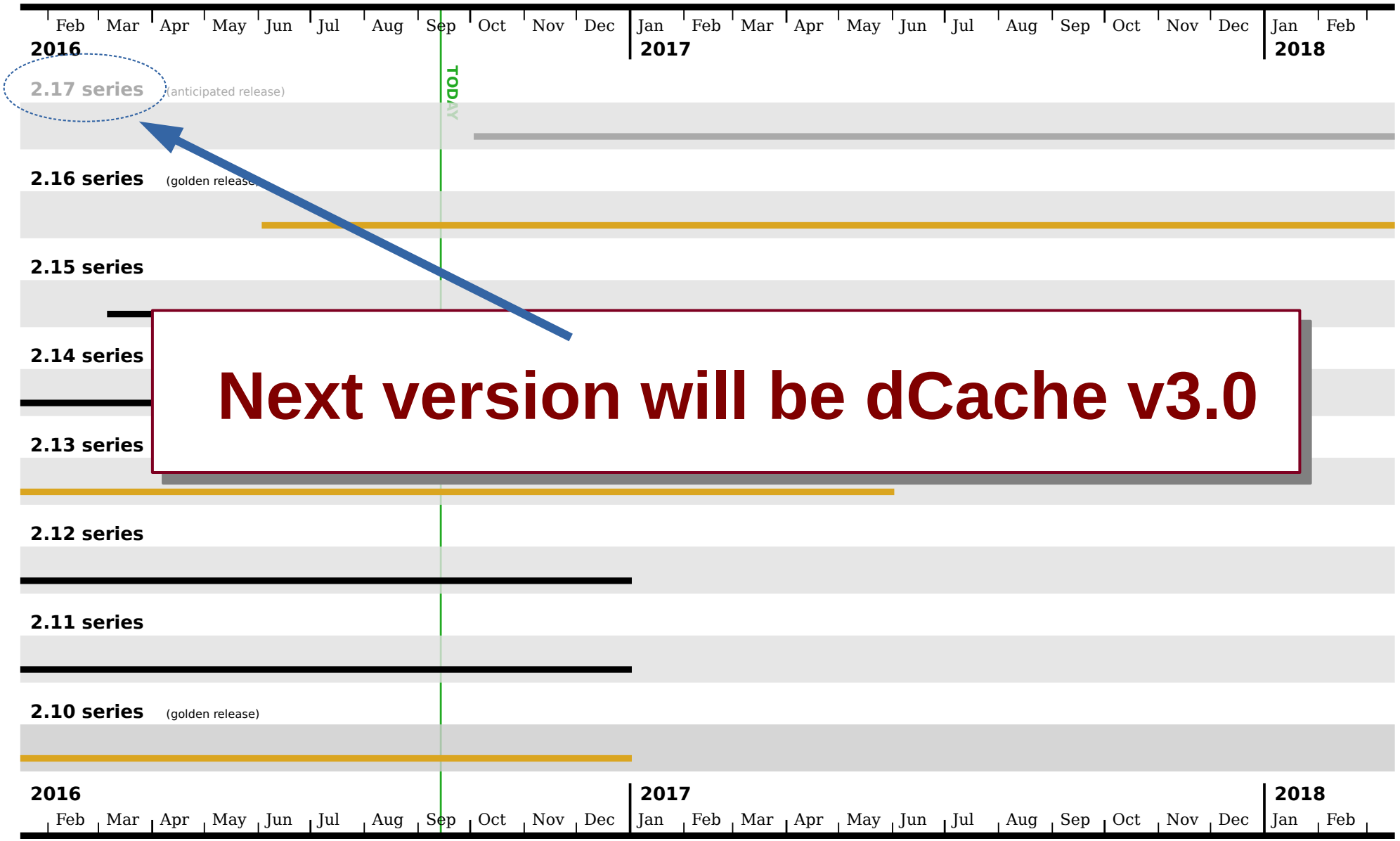
dCache

Paul Millar, on behalf of the dCache Team

pre-GDB „Data Management“ at CERN
2016-09-13
https://indico.cern.ch/event/394833/

dCache server releases

... along with the series support durations.



Why v3.0?

- Have to bump the number sooner or later.
- Better reflect backwards compatibility in mixed deployment,
- Many exciting new features,

    Optional – sites don't have to use them

- Final analysis .. just because.

New in 3.0: CEPH integration

- With dCache v3.0, dCache has **CEPH integration**:
  - Can deploy a dCache pool that provides access to a CEPH pool.
- dCache files are written as **RBD images**.

  Can be accessed directly (by PNFS-ID) outside of dCache
- All dCache features are available:

  Sites with tape integration may need to tweak their scripts
- Site driven functionality

New in 3.0: HA-dCache

- **What** is HA-dCache?
  - Multiple instances of core components can run concurrently,
  - Doors updated to support load-balancers (e.g., HAProxy).
- **Why** HA-dCache?
  - Symmetric deployment (making life easy),
  - Horizontal scaling (no CPU bottlenecks),
  - Fault tolerance (no single-point-of-failure),
  - Rolling bug-fix updates (no downtimes).

HA dCache: SRM

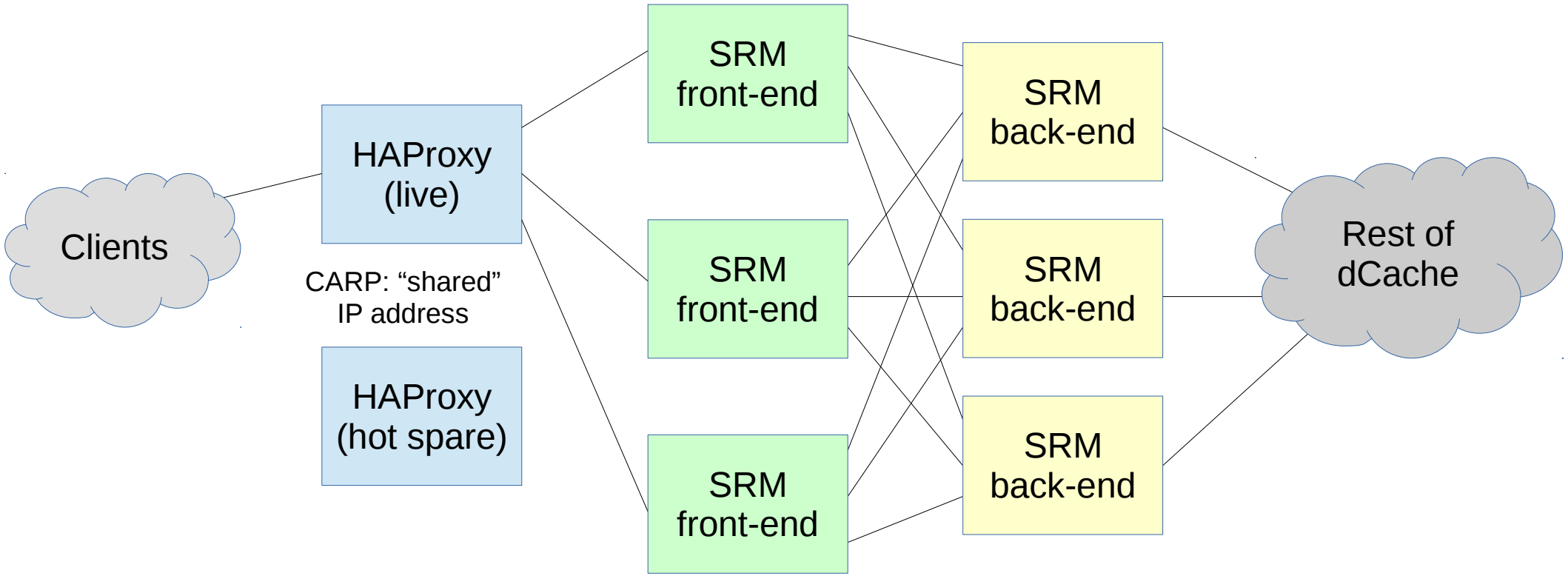
- **Split** the GSI "front-end" from "SRM engine"
- Allow **multiple front-ends**:

  horizontal scaling for encryption overhead
- Allow **multiple "SRM engines"**:

  each scheduled request is processed by the same SRM engine, load-balancing and fault-survival.
- Support for **HAProxy protocol**

  using TCP mode, rather than HTTP mode.

Pencil sketch of possible deployment



NB: works fine with just two node

HA dCache: general protocol remarks

- Should work fine for TLS-based protocols (SRM, gsiftp, webdav, gsidcap)
  - Load-balancer hostname as a Subject Alternate Name (SAN) in the X.509 certificate
- Possible to configure dCache so the SRM redirects clients to individual doors, rather than HA proxy:

  SRM already provides load-balancing.

HA dCache: FTP

- Updated to understand HAProxy protocol
- IPv4 and IPv6 supported
- Data channels connect directly to pool or door, bypassing HAProxy.

HA dCache: other protocols

- **WebDAV**: nothing major needed
- **xrootd**: updated to understand HAProxy protocol. As usual "GSI-xrootd" sucks:
  - special care needed over x.509 certificate
  - kXR_locate returns IP address; makes host name verification hard
- **dcap**: updated to understand HAProxy protocol. No other major changes.
- **NFS**: not updated to support HA.

HA-dCache: status and next steps

- Currently deployed in production at NDGF
  - Catching some bugs
- Presentations for admins at dCache workshop and "dCache Presents..." live webinar.
  - Considerable interest expressed.

Other thoughts/issues

- Deleting file with target free capacity:

  feedback loop: delete until enough space is free

- Multiple concurrent uploads of the same file:

  ATLAS – multiple FTS, CMS – hidden error recovery

  SRM mostly protects us from this (apart from "FTS srmRm bug")

  What is expected behaviour when not using SRM?

- RFC 4331 WebDAV quota support:

  Work started, anticipate being in dCache v3.0.

SRM reflections

- We (dCache.org) are NOT abandoning SRM:
  - We have invested heavily in cleaning- and speeding it up.
  - New client release, including **srmfs** an interactive SRM shell.
- It works – why replace a working system?

  By now the spec and implementations are well understood.
- Several unique features that would need to be re-implemented (e.g., see RFC-4331) – wasting effort.
- Biggest downside of SRM is NOT the protocol but the bindings; that can be fix.
- Certainly, declaring SRM dead is a self-fulfilling prophesy.

Backup slides