

XRootD

LSST Data Management
Database Architecture Review
SLAC National Accelerator Laboratory
August 6-7, 2013

Andrew Hanushevsky, SLAC
<http://xrootd.org>

A Bit Of **XRootD** History

- # 1997 – Objectivity, Inc. collaboration
 - Design & Development to scale Objectivity/DB
 - First attempt to use commercial DB for Physics data
 - Successful but problematical
- # 2001 – BaBar decides to use root framework vs Objectivity
 - Collaboration with INFN, Padova & SLAC created
 - Design & develop high performance data access system
 - Work based on what we learned with Objectivity
- # 2003 – First deployment of **XRootD** system at SLAC
- # 2013 – Wide deployment in LHC & Astrophysics
 - ALICE, ATLAS, CMS, EXO, Fermi/GLAST, LSST; among others
 - Protocol also available in dCache, DPM, and EOS

What Is XRootD?

- # A system for scalable cluster data access



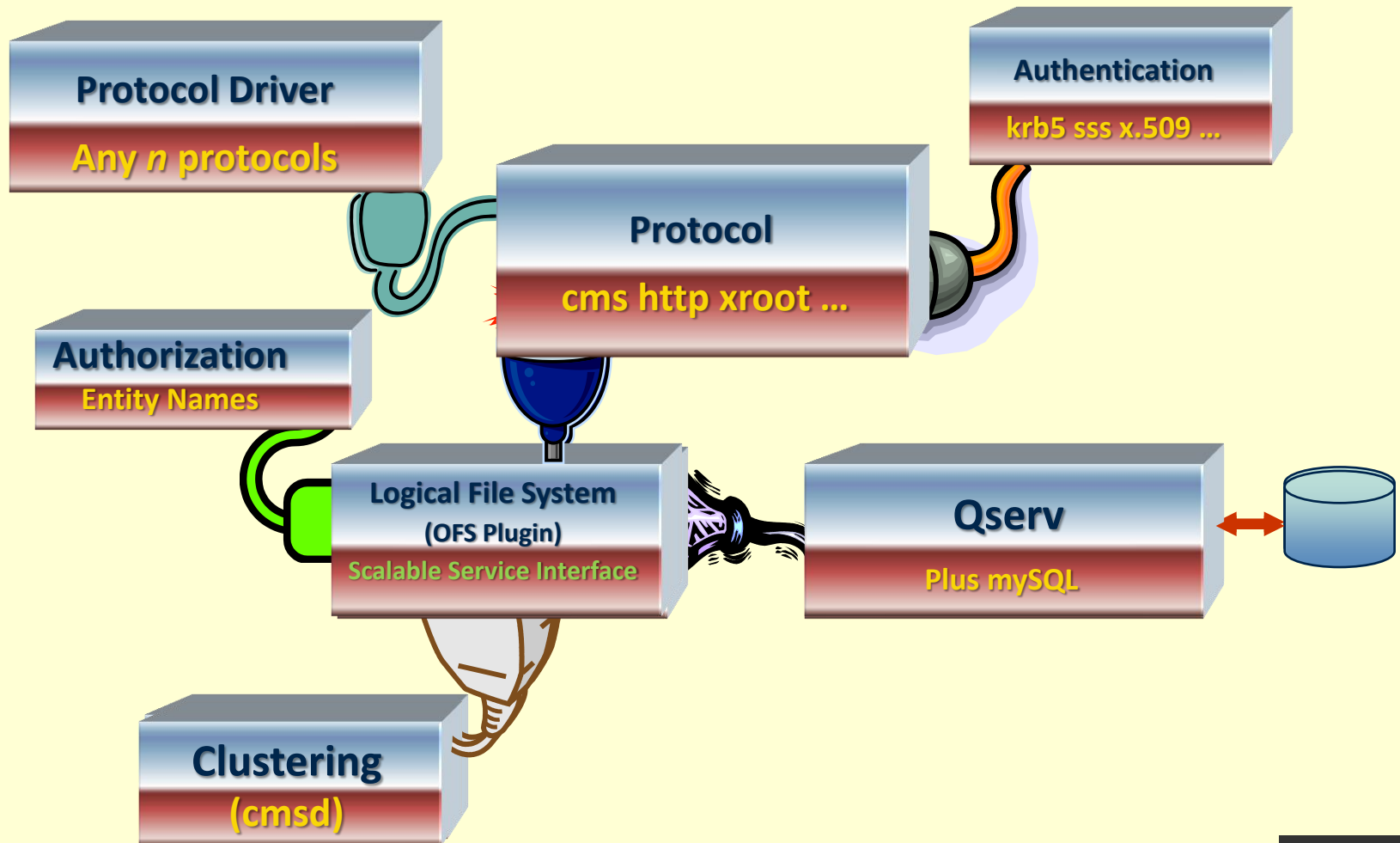
Data Access



Data Clustering

- # Not a file system & not *just* for file systems
- # If you can write a plug-in you can cluster it
 - The essential key for Qserv

XRootD Plug-In Architecture



The Original Challenge

- # The physics analysis regime “problem”
 - Write once read many times access mode
 - Small block sparse random I/O
 - Hundreds of servers of uneven reliability
 - Thousands parallel batch jobs
 - 100’s of thousands file sessions

The LSST Challenge

XRootD Synergistic Solution

Minimize latency

Parallelizable protocol, file sessions, lockless I/O, sticky threads

Minimize hardware requirements

Short code paths, compact objects

Cache aware members

Minimal data movement

No cross-thread data sharing

Minimize human cost

Single configuration file

Cookie cutter installation

No database needed

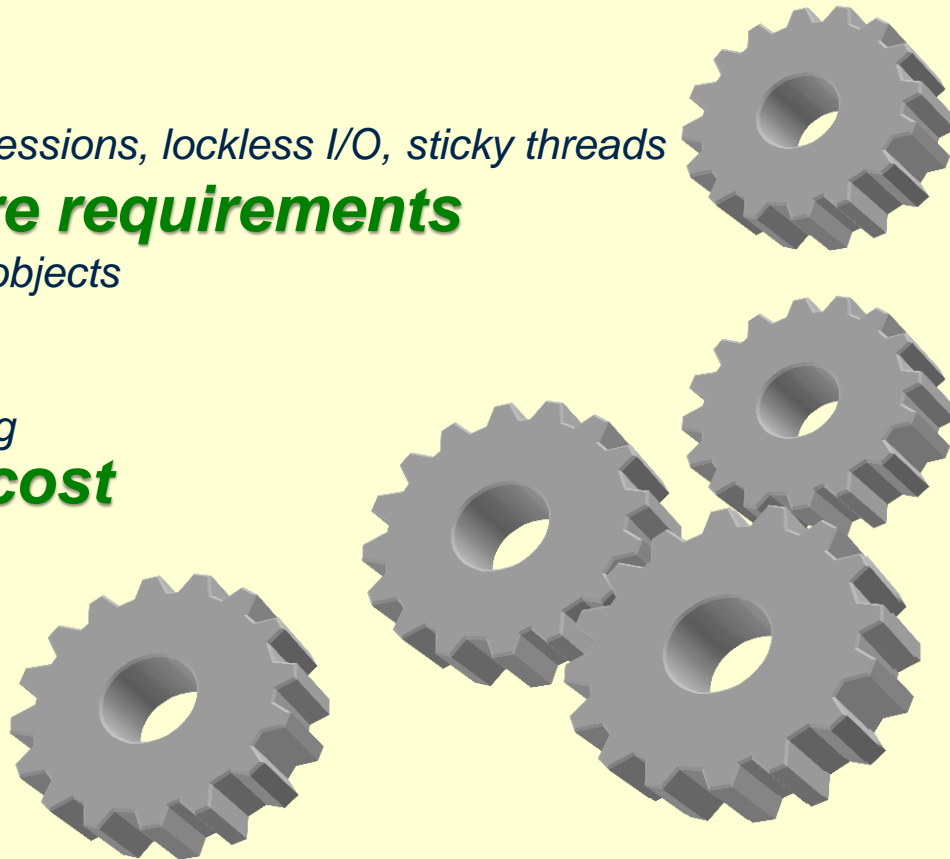
Real time node change,

Native FS admin tools

Maximize scaling

Result

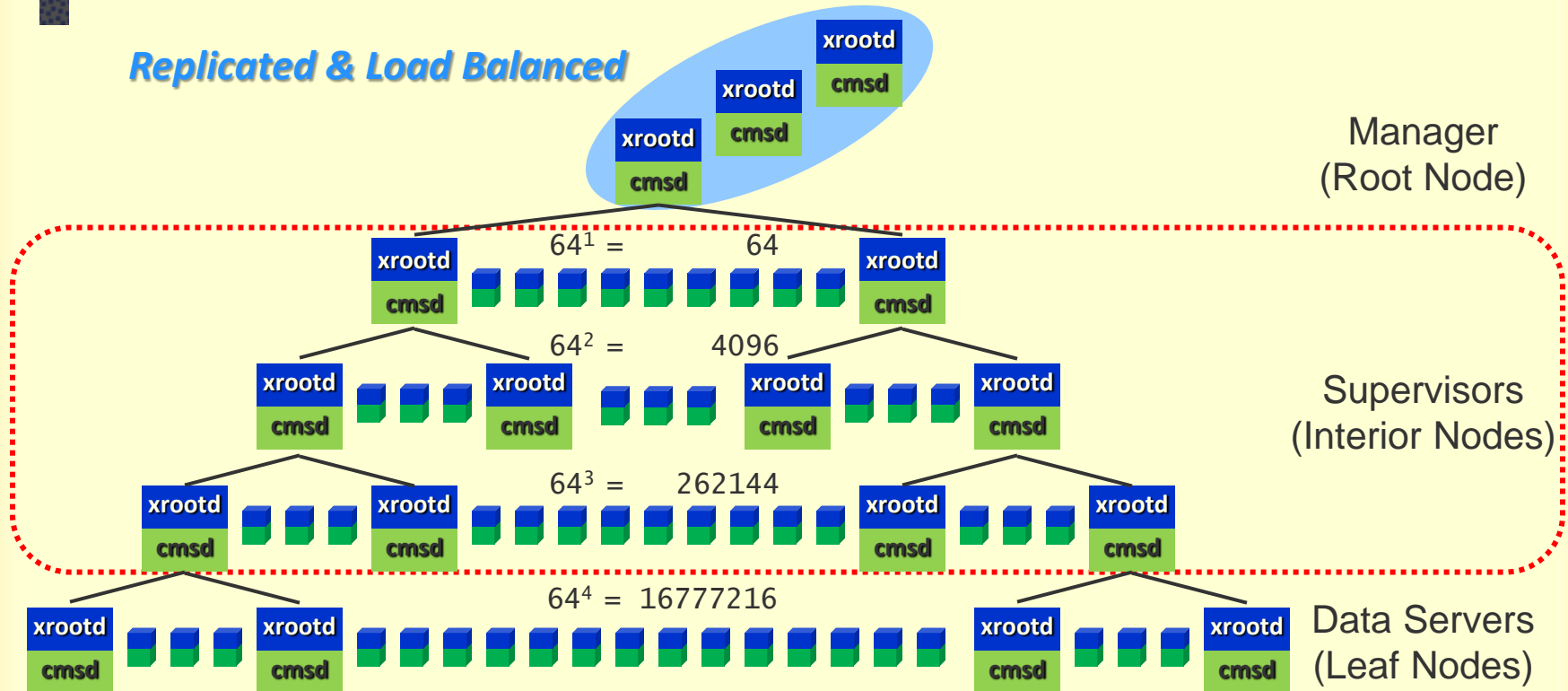
<7 us overhead & <100 MB footprint & unlimited servers





XRootD B⁶⁴ Scaling

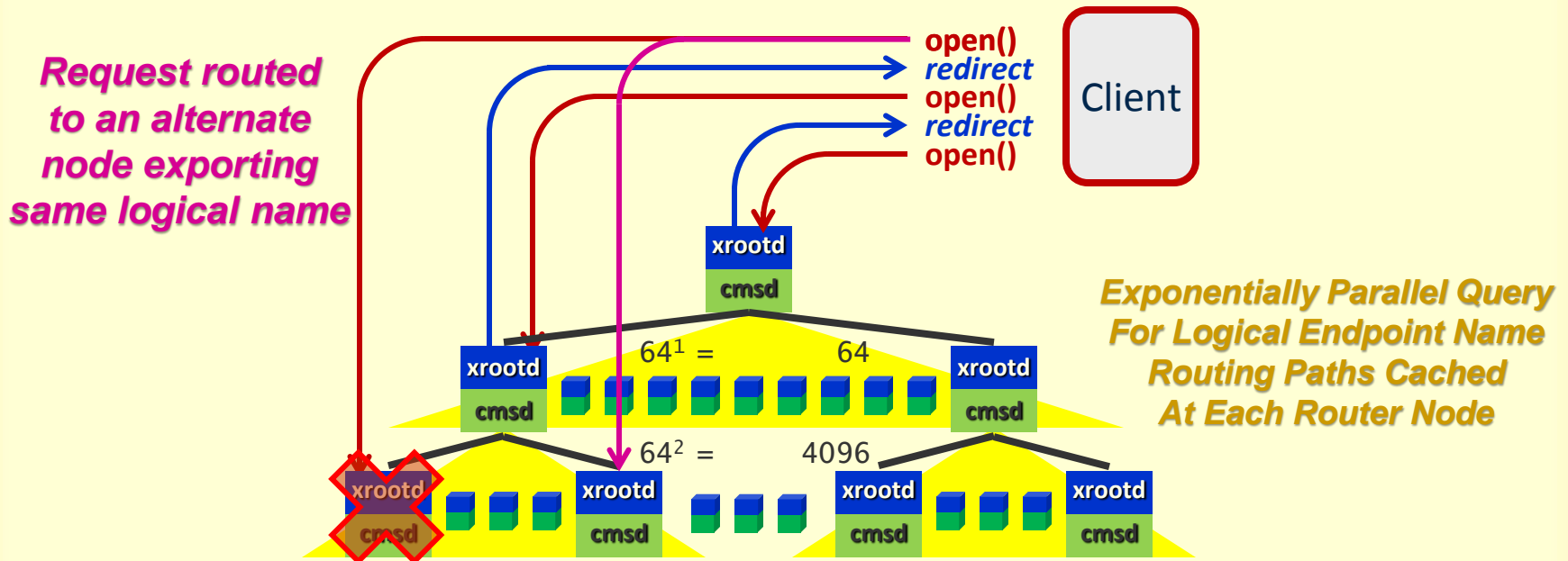
Replicated & Load Balanced



Qserv worker nodes reside *only* at the leaves
Each identified by a dynamic set of arbitrary names



WYSIWYG Scalable Access



Request routing is very different from traditional data management models
This implements a structured network of request routers
Capable of automatically recovering from adverse conditions
Much like internet routing

Conclusion

- # A facile, flexible, and sound system
 - Applicable to a wide variety of problems
- # LGPL open-source
- # Managed by the **XRootD** collaboration
 - SLAC, CERN, Duke, JINR, UCSD, & UNL (fall)

The logo for SLAC (Stanford Linear Accelerator Center), featuring the letters "SLAC" in a bold, red, sans-serif font.The logo for UNL (University of Nebraska-Lincoln), featuring a blue square with the letters "UNL" in white.

More at <http://xrootd.org/>

Acknowledgements

Current Contributors

- ATLAS: Doug Benjamin, Patrick McGuigan, Ilija Vukotic
- CERN: Lukasz Janyst, Andreas Peters, Justin Salmon
- Fermi: Tony Johnson
- JINR: Danila Oleynik, Artem Petrosyan
- Root: Gerri Ganis, Bertrand Bellenet, Fons Rademakers
- SLAC: Andrew Hanushevsky, Wilko Kroeger, Daniel Wang, Wei Yang
- UCSD: Matevz Tadel, Alja Mrak-Tadel
- UNL: Brian Bockelman
- WLCG: Fabrizio Furano, David Smith

US Department of Energy

- Contract DE-AC02-76SF00515 with Stanford University