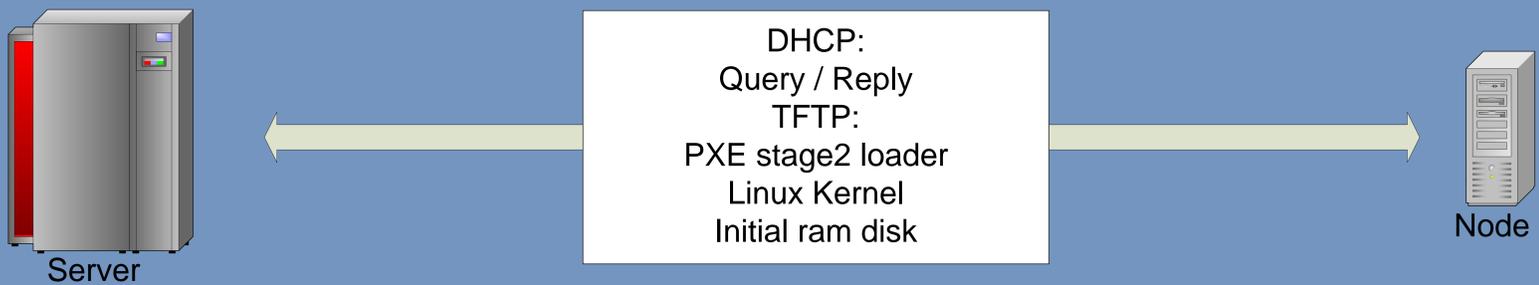


Fabric management with diskless servers and Quattor in LHCb

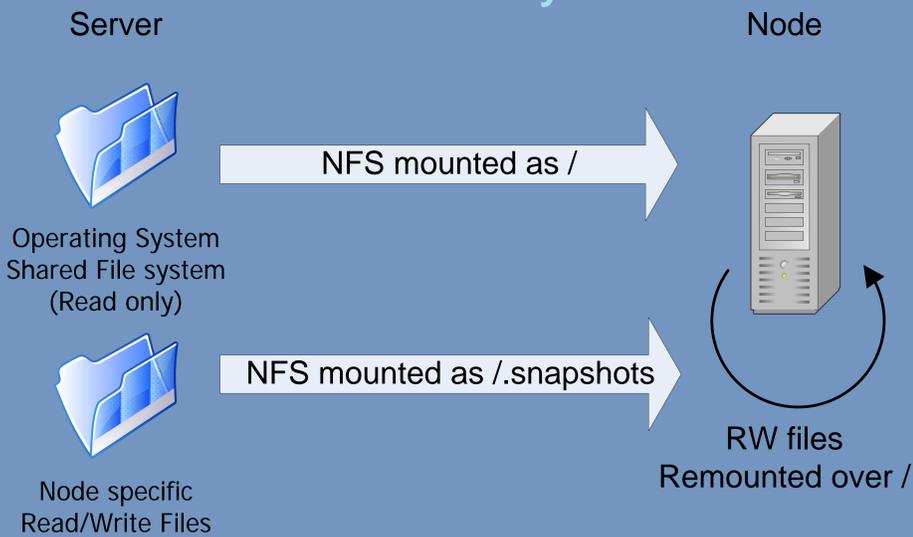
P. Schweitzer*, E. Bonaccorsi, L. Brarda, N. Neufeld, CERN, Geneva, Switzerland

LHCb is using many computers running Linux without disks : ~ 1500 computing nodes in the event filtering nodes and ~ 500 Credit Card computers (CCPCs) embedded in frontend electronic boards. As all LHCb Linux computers, these diskless computers are configured using the Quattor Toolkit. This poster presents how we evolved from an unmaintainable Quattor component using deprecated RedHat tools to a modern and configurable component using file system union to provide the Operating System to the diskless nodes.

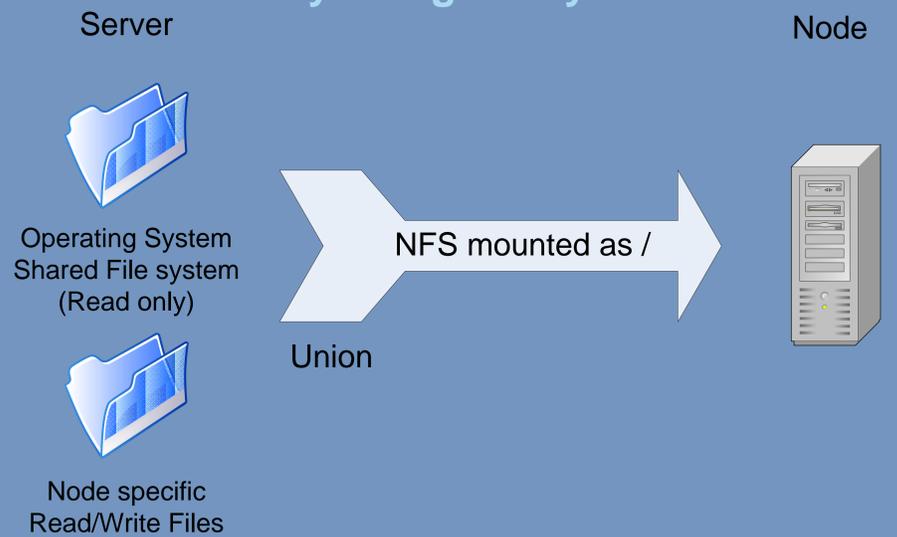
Diskless boot : common part



Diskless boot RedHat way



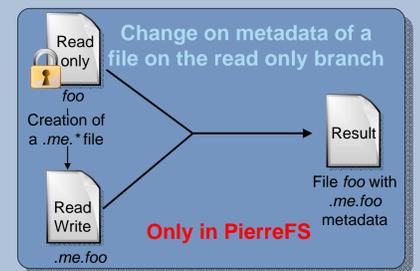
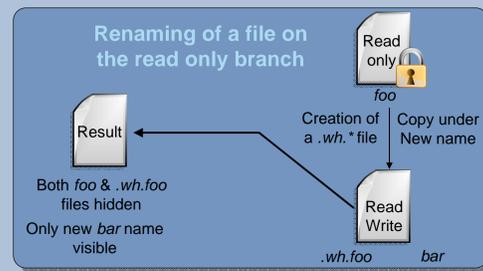
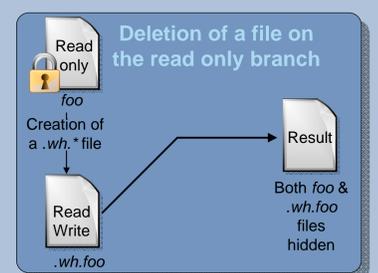
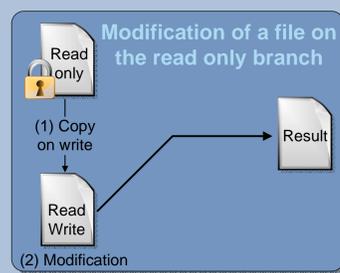
Diskless boot New way using file system union



Comparison between PierreFS and other union FS

Fusion Mode	Solution	UnionFS	aufs	funionfs	unionfs-fuse	PierreFS	Legend
		Kernel	Kernel	FUSE	FUSE	FUSE	
Cow		☑	☑	☑	☑	☑	☑ = supported
Whiteout		☑	☑	☑	☑	☑	☑ = Documented
NFS Support		☑	☑	☑	☑	☑	☑ = As supported but is not
NFS Export		☑	☑	☑*	☑*	☑*	☑* = supported under conditions
Metadata separation		☑	☑	☑	☑	☑	
Copyup deletion		☑	☑	☑	☑	☑	
Kernel rebuild		☑	☑	☑	☑	☑	☑ = unsupported

Union FS specific cases



PierreFS unique features

- After evaluation of the current union FS implementations, it was decided to implement our own, to gain in performance and memory footprint.
- Copy only metadata on read/write branch when metadata is modified
 - Remove files on read/write branch whenever possible
 - when there is a .me.* file and it's metadata is now the same as the file on the read only branch
 - when the file have been removed (there is a .wh. File) and the same file is written back
 - When the file have been modified and the original file is written back
- PierreFS is currently implemented using FUSE but a project to implement it as a real Linux file system will start soon.

New Quattor Diskless component

The LHCb experiment is using the Quattor toolkit (www.quattor.org) to manage the configuration of all its Linux servers and nodes. A Quattor component had been made several years ago to handle the configuration of the servers for diskless nodes. This component have been completely rewritten.

Previous component weaknesses

- Based on Red Hat's system-config-netboot package
 - some scripts were buggy and were failing without any explanations
 - Have been entirely removed from RHEL6/SLC6
- Unmaintainable
- Not supporting Union File Systems

New component features

- Supports several Operating System versions and architecture
- Creation 'from scratch' of the shared root file system(s)
- Nicer generated dhcpd.conf, using groups
- Easy support of new Operating Systems, using plug-in
- Easier to maintain

Diskless nodes configuration

- The configuration of diskless nodes is coming from two sets of templates:
- Shared root file system templates : configures all parameters common to nodes that uses that shared root file system : package installation, running services, nfs config, ...
 - Node templates : everything specific to that node : mainly the network config