



Introduction to WP9



IBM-VM 360, CERNVM, 1988

Portable Analysis Environment Using Virtualization Technology

<http://cern.ch/cernvm>

Predrag Buncic (CERN/PH-SFT)

- Software @ LHC Experiment(s)
 - Millions of lines of code
 - Complicated software installation/update/configuration procedure, different from experiment to experiment
 - Only tiny portion of it is really used at runtime in most of the cases
 - Often incompatible or lagging behind OS version on desktop/laptop
- Multi core CPUs with hardware support for virtualization
 - Making laptop/desktop ever more powerful and underutilised
- Using virtualization and extra cores to get extra comfort
 - Zero effort to install, maintain and keep up to date the experiment software
 - Reduce the cost of software development by reducing the number of compiler-platform combinations
 - Decouple application lifecycle from evolution of system infrastructure

How do we want to do that?

- Build a “thin” Virtual Software Appliance for use by LHC experiments
- This appliance should
 - provide a complete, portable and easy to configure user environment for developing and running LHC data analysis locally and on the Grid
 - be independent of physical software and hardware platform (Linux, Windows, MacOS)
- This should minimize the number of platforms (compiler-OS combinations) on which experiment software needs to be supported and tested thus reducing the overall cost of LHC software maintenance.
- All this to be done
 - in collaboration with LHC experiments, OpenLab
 - reusing existing solutions where possible
- Resources
 - 7.5 FTE (staff) + 7.5 FTE (fellow) over 4 years

Initial Project Goals

1. Evaluation of the available virtualization technologies
 - Understand and validate technologies by checking their performance in HEP environment
2. Evaluation of the tools to build and manage Virtual Appliances
3. Collect User Requirements from experiments
 - Confront them with available technologies
 - Suggest an optimal choice for a given use case
4. Development and deployment of a read-only distributed Network File System for software distribution
 - Essential to keep the basic appliance small in size (<100MB)
 - Validate performance, scalability and usability of such approach
5. Provide prototypes for at least two LHC experiments
 - Assist experiments in adapting their software practices to this platform
6. Setup a service and support infrastructure

About virtualization, appliances and some recent trends in industry

- The industry is gradually redefining meaning of the grid computing
 - Shift from glorious ideas of a large public infrastructure and common middleware towards an end-to-end custom solutions and private corporate grids
 - Grid & Grow (IBM)
- Emerging new platforms:
 - Amazon Elastic Computing Cloud
 - Everything is for rent (CPU, Storage, Network, Accounting)
 - Blue Cloud (IBM) is coming
 - Software as a Service (SaaS)
 - Google App Engine
- Key enabling technology: Virtualization
 - CPU, OS, network, file virtualization and storage virtualization
 - Rich tools built on top of virtualization environments



virtualization

vs

grid computing

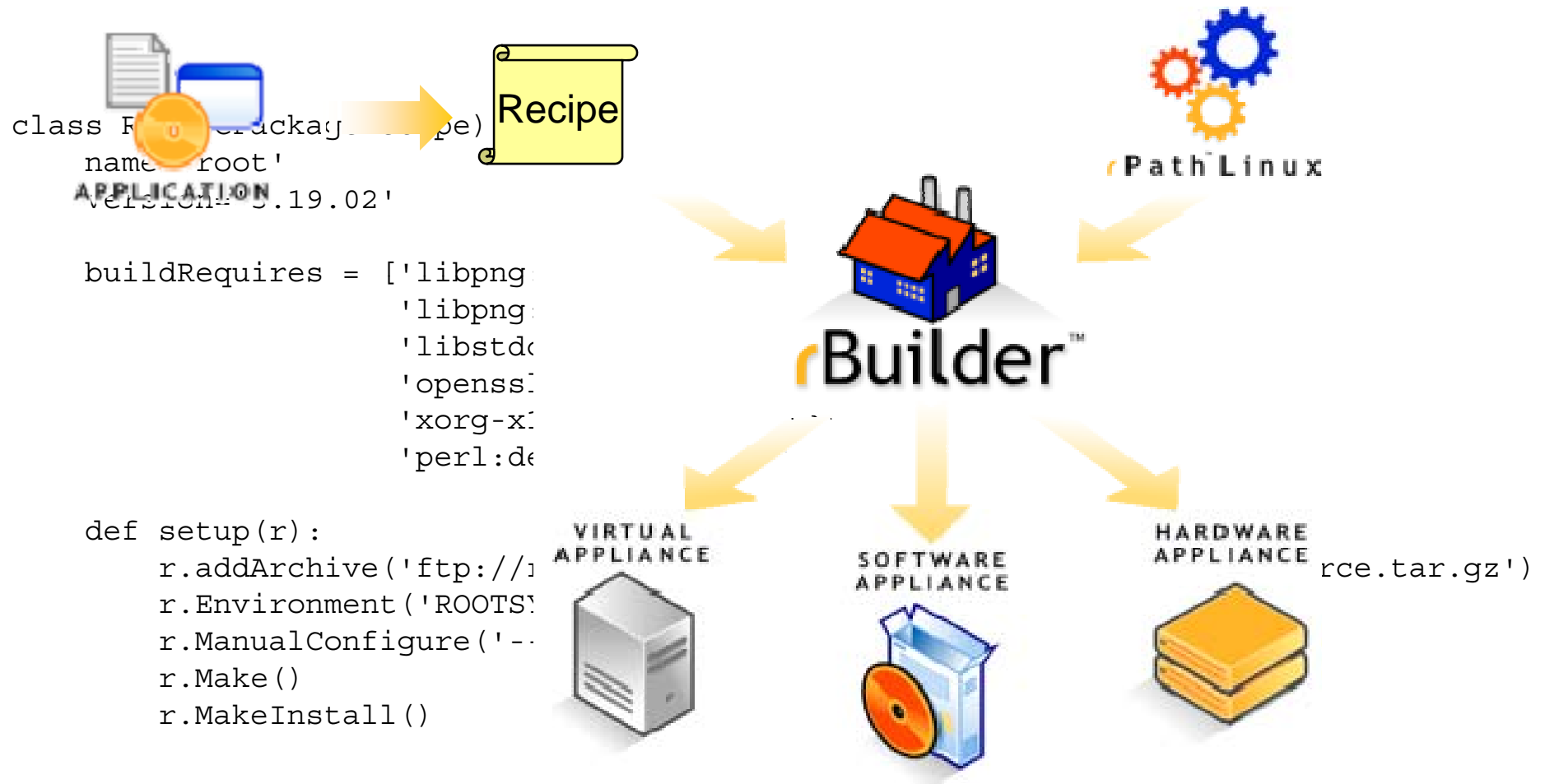
- Virtual Software Appliance is a lightweight Virtual Machine image that provides easy configuration and maintenance and combines
 - minimal, “just enough” operating environment (JeOS)
 - specialized application functionality
 - These appliances are designed to run under one or more of the various virtualization technologies, such as
 - VMware , Xen, Parallels, Microsoft Virtual PC, QEMU, User mode Linux, CoLinux, Virtual Iron...
- 👉 For evaluation of some of these technologies in HEP environment, see talk of [J. Blomer](#)

Evaluation of the tools to build and manage Virtual Appliances

AliEn Appliance (2006)

QuickTime™ and a
decompressor
are needed to see this picture.

Easy packaging: rBuilder



Ready for Amazon EC2...

The screenshot shows the EC2 Firefox UI interface. The main window displays a list of Available AMIs (Amazon Machine Images) with columns for AMI ID, Manifest, Owner, and Visibility. A 'Launch Permissions' window is open on the right, showing the Amazon Web Services logo. A 'Your Instances' table is visible at the bottom left, and a file manager window is open in the foreground, showing a directory structure with a file named 'pbuncic'.

AMI ID	Manifest	Owner	Visibility
ami-a37792ca	rbuilder-online/bioinformatics-0.0.1-x86_15351.img.manifest.xml	099...	public
ami-e630d58f	rbuilder-online/login-appliance-0.21-x86_14113.img.manifest.xml	099...	public
ami-5681643f	rbuilder-online/alien-0.4.0-x86_9430.img.manifest.xml	099...	public
ami-e0cb2e89	rbuilder-online/opermix-2.0-x86_13261.img.manifest.xml	099...	public
ami-f2fa1f9b	rbuilder-online/intaio-2.2-x86_12595.img.manifest.xml	099...	public
ami-936a8ffa	rbuilder-online/nclcds-base-0.0.1-x86_15299.img.manifest.xml	099...	public
ami-018e6c68	rbuilder-online/phonehome-1.5.6-x86_10132.img.manifest.xml	099...	public
ami-9d1ffa4	rbuilder-online/conary-proxy-2C080310-x86_15138.img.manifest.xml	099...	public
ami-5695703f	rbuilder-online/erberus-0.3.0-x86_10516.img.manifest.xml	099...	public
ami-0683666f	rbuilder-online/fedoracore6-1.0-x86_9677.img.manifest.xml	099...	public
ami-1be30572	rbuilder-online/freeswitch-0.0.1-x86_11952.img.manifest.xml	099...	public
ami-5481643d	rbuilder-online/alien-0.4.0-x86_9427.img.manifest.xml	099...	public
ami-fb10f592	rbuilder-online/conary-proxy-2C080310-x86_14943.img.manifest.xml	099...	public
ami-ec13f18f	rbuilder-online/...	099...	public

Reservation ID	Owner	Instance ID	AMI	State	DNS

File Name	File Size(KB)	Upload Time
pbuncic	0	04/03/2007 10 07 PM

QuickTime™ and a
decompressor
are needed to see this picture.



Groups & Builds

Build types

- Installable CD/DVD
- Stub Image
- Raw Filesystem Image
- Netboot Image
- Compressed Tar File
- Demo CD/DVD (Live CD/DVD)
- Raw Hard Disk Image
- VMware® Virtual Appliance
- VMware® ESX Server Virtual Appliance
- Microsoft® VHD Virtual Appliance
- Xen Enterprise Virtual Appliance
- Virtual Iron Virtual Appliance
- Parallels Virtual Appliance
- Amazon Machine Image
- Update CD/DVD
- Appliance Installable ISO

- Decision to use rPath's rBuilder was easy one
 - There is no other comparable product offering so much flexibility
 - It is possible to use rBuilder Online to setup test project and exercise the tools
 - Fully functional
 - Free
 - Occasionally overloaded
 - We bought a license for one instance of rBuilder
 - <http://rbuilder.cern.ch>
 - Now hosts development of CernVM Appliance
- 👉 If you are interested to learn more about rBuilder, please come to tutorial on Wednesday

Tools to Manage VMs

- There are many prototypes and a few really working ones
- We tried in house solution (Vgrid) and Open QRM
- Finally settled for Enomalism 2.0
 - Simple easy to use web based user interface
 - VM Agnostic, migrate to and from various virtual environments including KVM/Qemu, Amazon EC2 and Xen. (OpenVZ, VirtualBox and VMware coming soon)
 - Fine grained user permissions and access privileges
 - One click deployment
 - Can start/stop/pause/save VM instances
 - Free Open source License



👉 More about Enomalism later today in talk of Carlos Aguado Sanchez

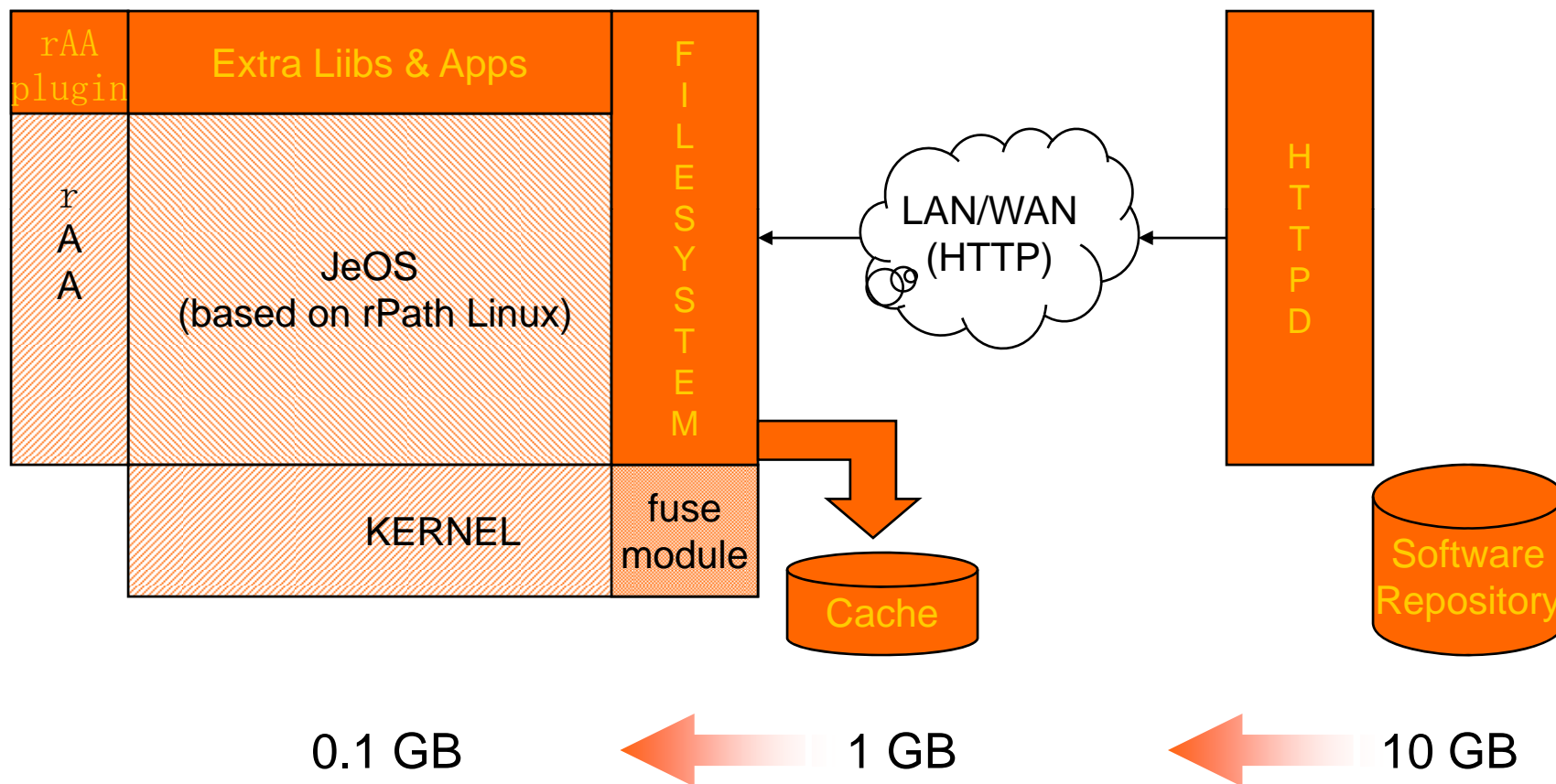
User Requirements

User Requirements

- This should normally happen before any work starts
- The initial screening of opinions happened in direct discussion with people from experiments who got interested in Virtualization project (end of 2007, beginning of 2008)
 - This feedback was taken into account when initial plan was created
- Expecting more feedback during discussion session on Wednesday
 - No enthusiastic response to calls for help in writing URD
 - <http://cernvm.web.cern.ch/cernvm/index.cgi?page=UserRequirements>
- We have to decide if User Requirements document is needed and if yes, we will need at least one person from participating experiment who will feel responsible for specific requirements

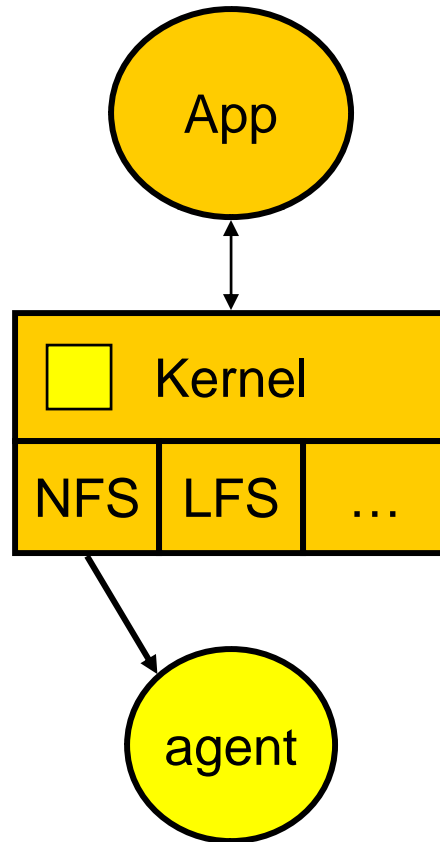
Development and deployment of a read-only distributed Network File System for software distribution

“Thin” Software Appliance

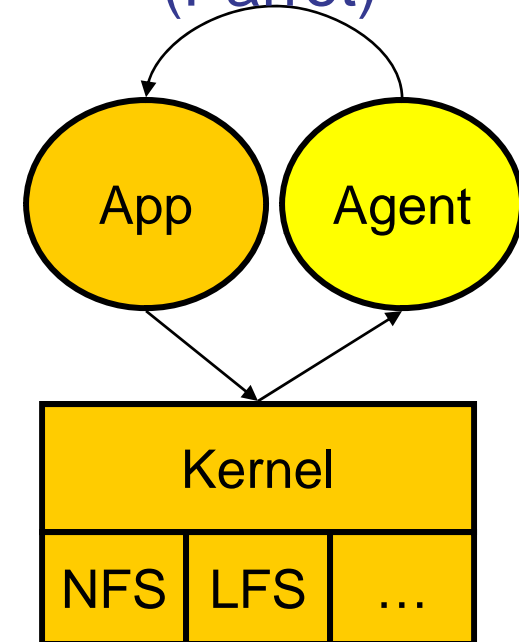


Network File System Techniques

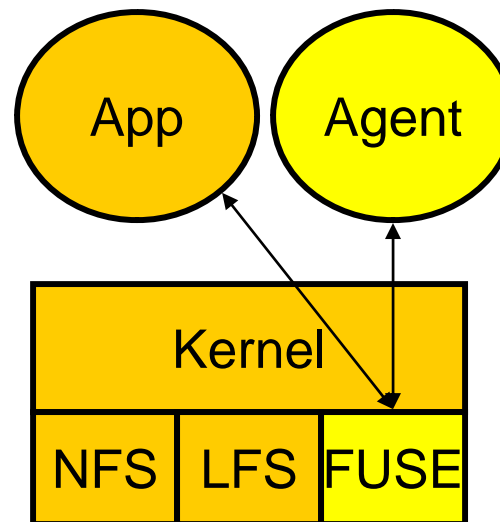
Kernel File System
(NFS)



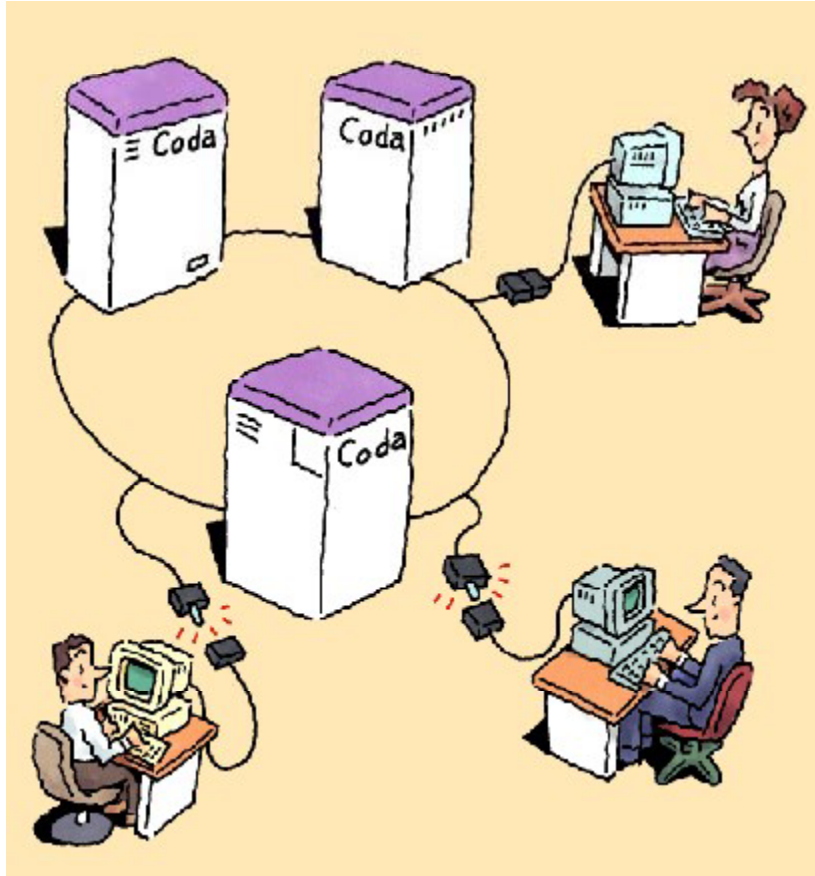
Debugger Trap
(Parrot)



Kernel Callout
(Fuse)

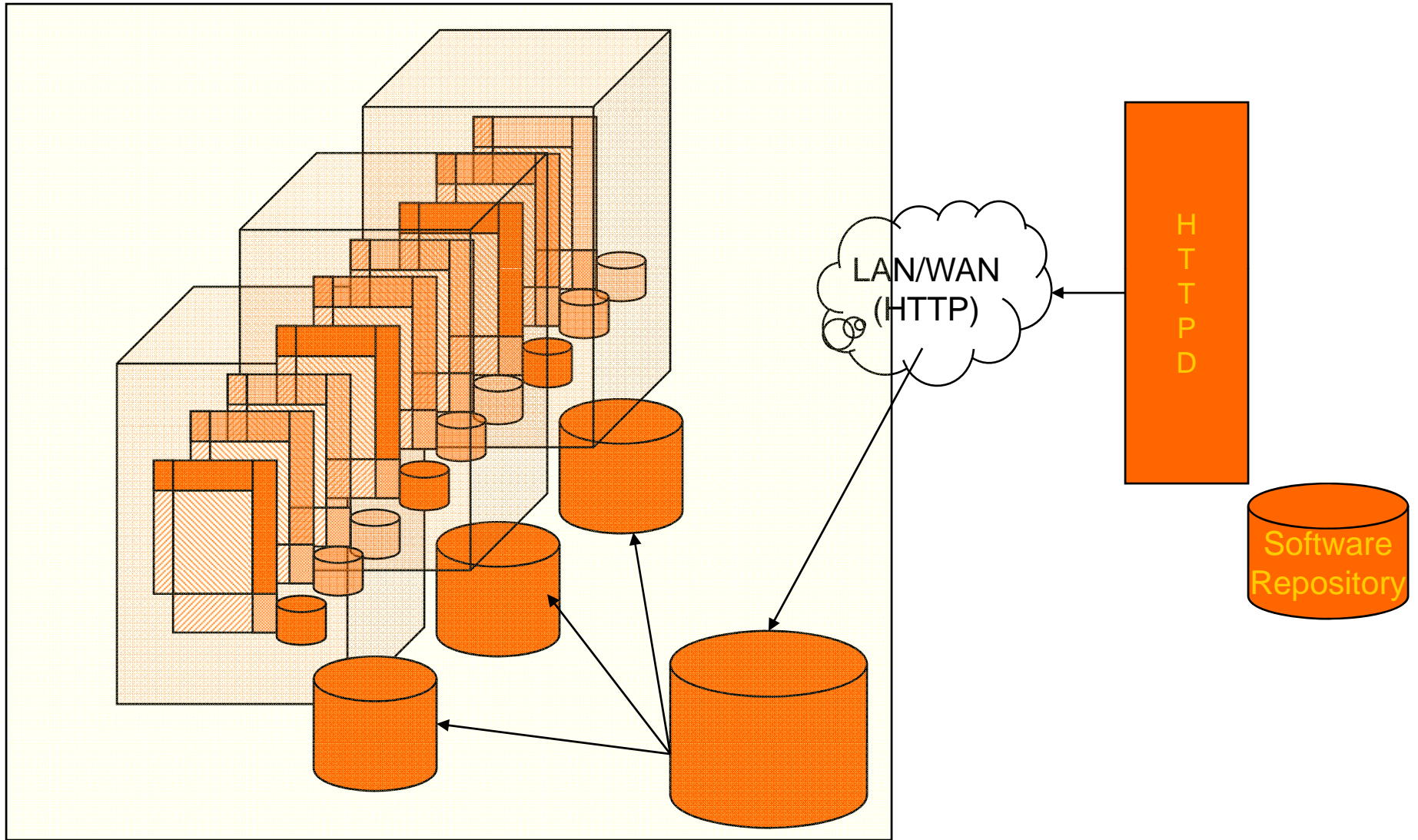


Experience with AliEn

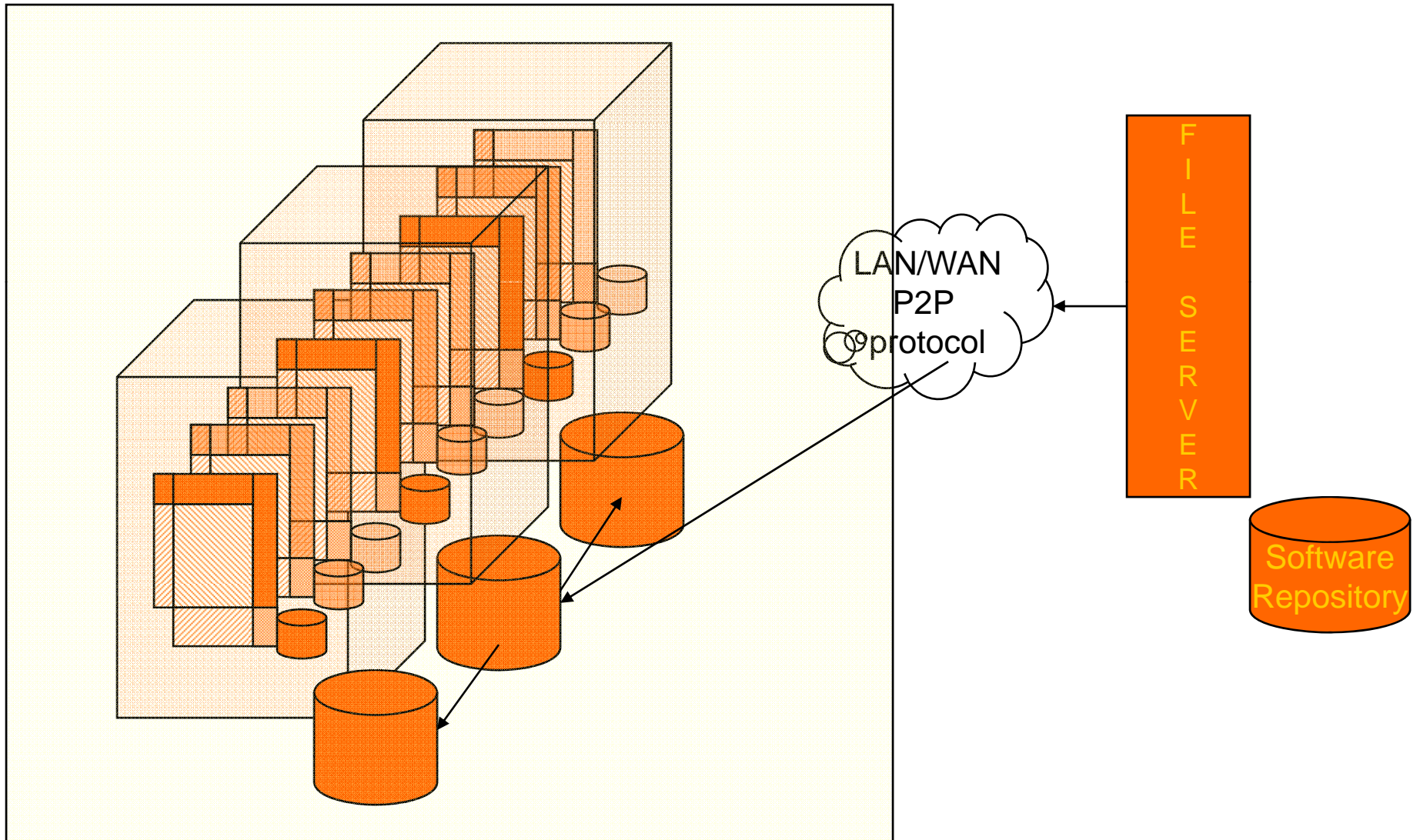


- Based CODA, distributed file system with its origin in AFS2
 - **disconnected** operation for mobile computing
 - client side persistent caching and **server replication**
 - **security** model for authentication, encryption and access control
 - **continued operation during partial network failures** in server network
 - network **bandwidth adaptation**
- **Problems**
 - Latency on WAN is killing the response time
 - We do not really need its write and synchronization capability

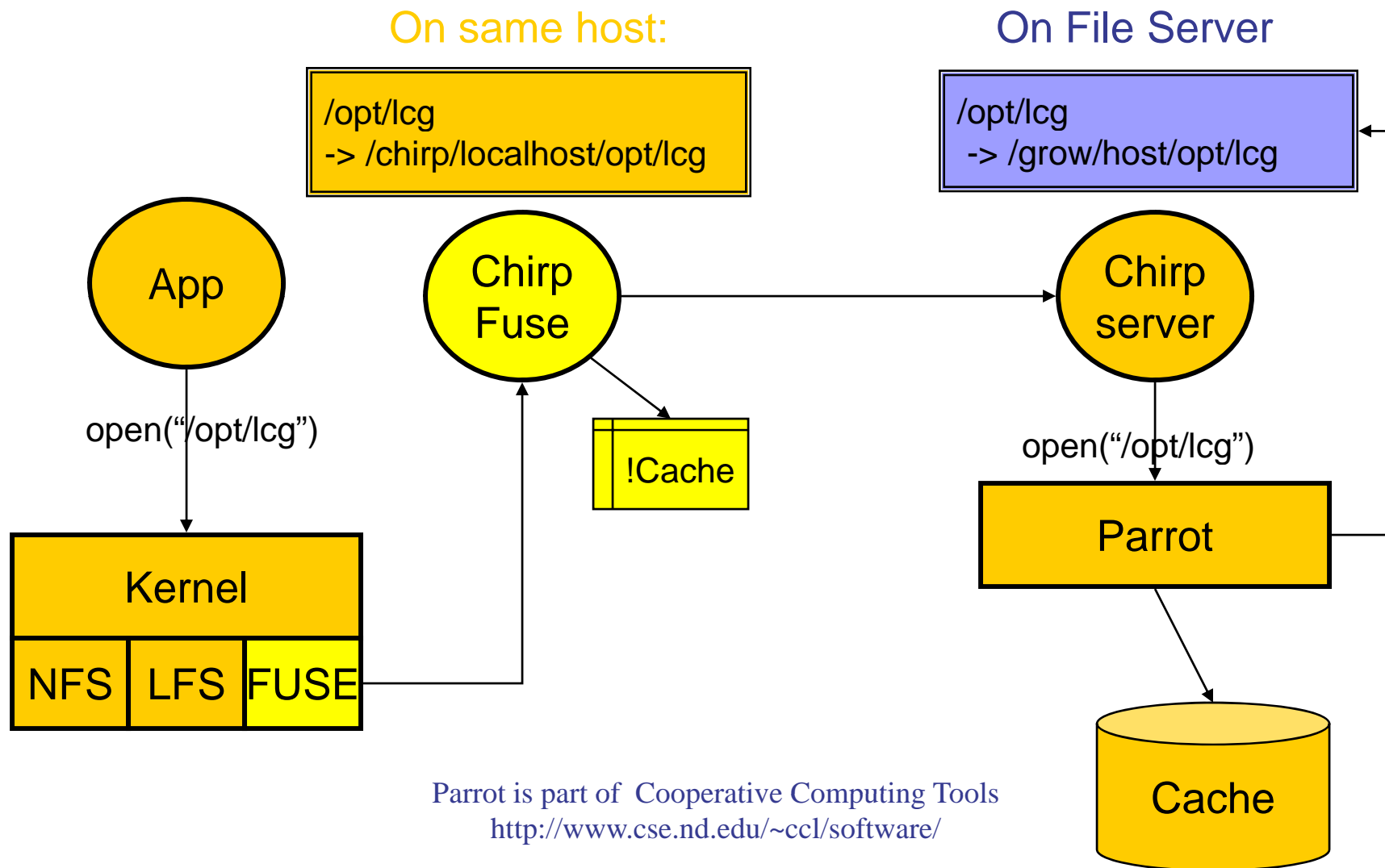
Scalability: Cache Hierarchy



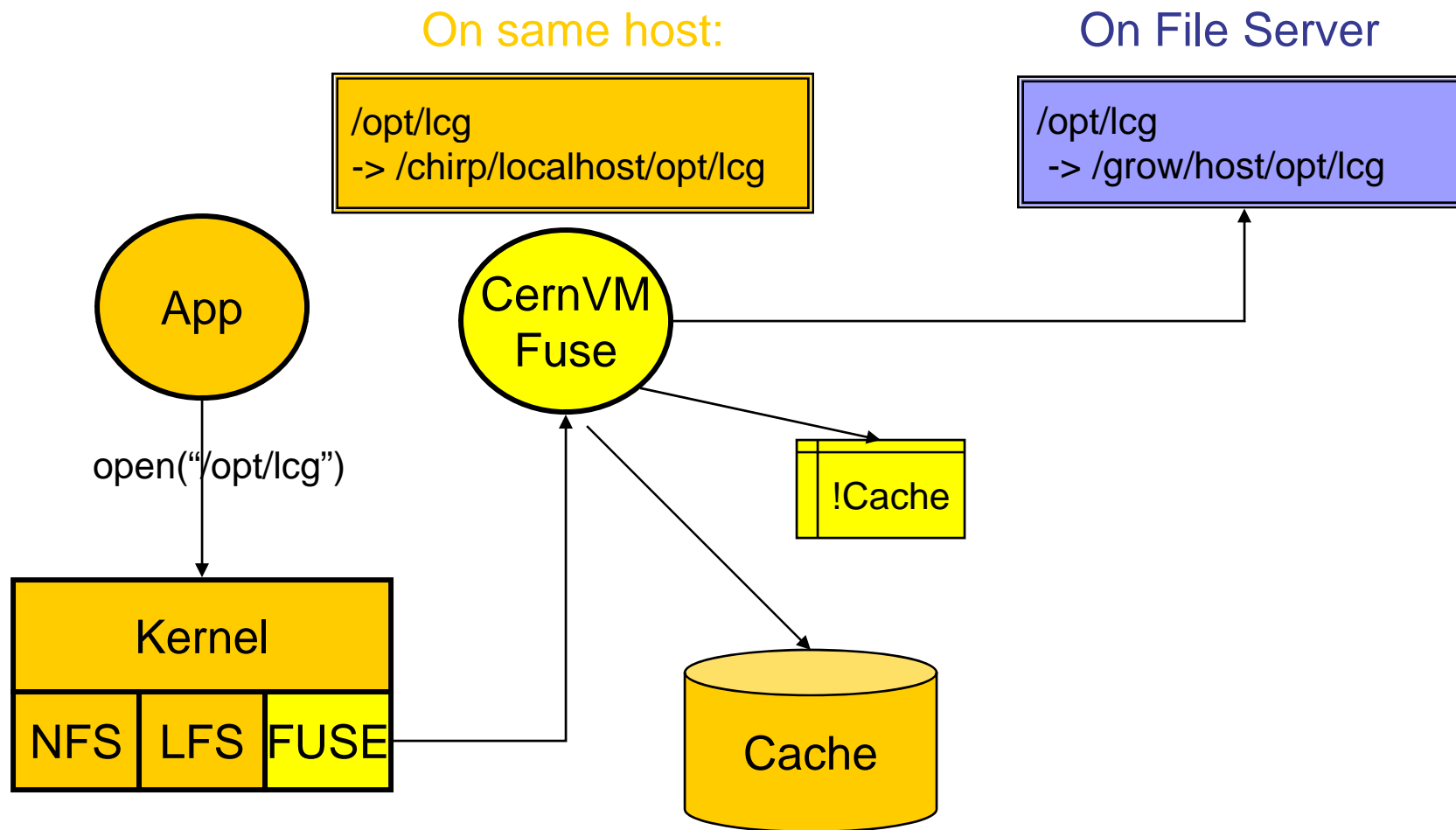
Scalability: P2P



Current CernVM setup



Foreseen development



Provide prototypes for at least two
LHC experiments

Alpha release 0.3

QuickTime™ and a
decompressor
are needed to see this picture.

- Available for download from
 - <http://rbuilder.cern.ch/project/cernvm-devel/releases>
- Can be run on
 - Linux (Xen,VMware Player)
 - Windows(WMware Player)
 - Mac (Fusion, Parallels)
- Minimal HowTO at
- <http://cernvm.web.cern.ch/cernvm/?page=HowTo>
- Limited testing to verify that appliance can be configured and used with Alice and LHCb software frameworks
- ATLAS is next in line
- Waiting for input from CMS

Demo time...

- The work on the project has started according to initial planning and is progressing well
 - Initial discussions with experiment trying to identify interest for collaboration
 - Contacts have been established with goal to carry out further research and look into alternative solutions
 - We have evaluated performance of several virtualization technologies in HEP context
 - We purchased and installed rBilder instance at CERN
 - Work on setting up scalable service infrastructure is in progress
 - Prototype of file system for s/w distribution is implemented
 - Alpha release is available and is ready for 2 experiments
 - Work on Atlas will start next
 - Waiting for expression of interest from CMS
 - We are expecting input from experiments
 - User Requirements document
 - Plan of work