Ian Bird

LHCC Referees;

CERN, 2nd June 2015

Project Status Report

RRB summary

- Reviewed 2016 requests (slides)
  - ➔ viewed as reasonable, only minor adjustments
  - Significant increases in tape requests cf earlier expectations – previously mentioned here

ATLAS

		2015 ATLAS	2015 CRSG	2016 ATLAS	2016 CRSG
CPU (kHS06)	T0	205	205	257	257
	T1	462	450	520	520
	T2	530	520	566	566
Disk (PB)	T0	14	14	17	17
	T1	39	37	47	47
	T2	55	52	72	72
Tape (PB)	T0	33	33	42	42
	T1	65	65	116	116

- Revision to requirements for 2015 and beyond

CMS

		2015 CMS	2015 CRSG	2016 CMS	2016 CRSG
CPU (kHS06)	T0	271	271	317	317
	T1	300	300	400	400
	T2	500	500	700	700
Disk (PB)	T0	15	15	16	16
	T1	27	26	35	33
	T2	31	29	40	38
Tape (PB)	T0	35	35	44	44
	T1	74	74	100	100

- Requests unchanged since October 2014
- Will take advantage of doubled HLT capacity in Run 2

LHCb

		2015 LHCb	2015 CRSG	2016 LHCb	2016 CRSG
CPU (kHS06)	T0	36	36	51	51
	T1	118	118	156	156
	T2	66	66	88	88
	HLT+Yandex	20	20	20	20
Disk (PB)	T0	5.5	5.5	7.6	7.6
	T1	11.7	11.7	13.5	13.5
	T2	1.9	1.9	4.0	4.0
Tape (PB)	T0	11.2	11.2	20.6	20.6
	T1	23.7	23.7	42.1	42.1

- Requests unchanged since October 2014

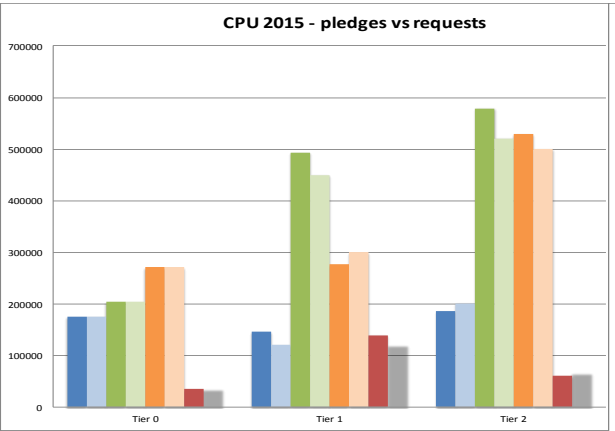
ALICE

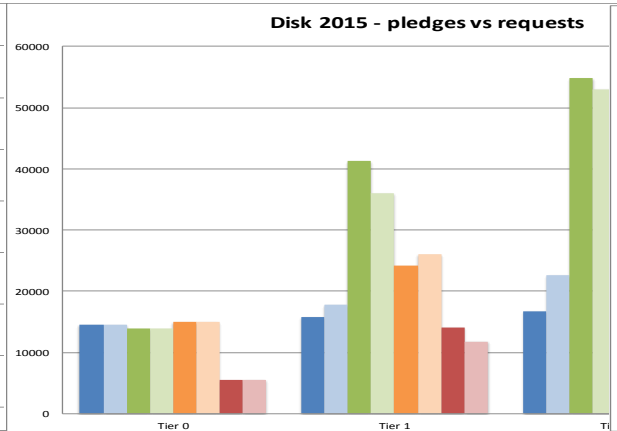
		2015 ALICE	2015 CRSG	2016 ALICE	2016 CRSG
CPU (kHS06)	T0	175	175	215	215
	T1	120	120	160	157
	T2	200	200	240	237
Disk (PB)	T0	14.5	14.5	16.7	16.7
	T1	17.8	17.8	21.0	21.0
	T2	22.7	22.7	26.1	26.1
Tape (PB)	T0	16.2	16.2	21.6	21.6
	T1	10.2	10.2	15.6	15.6

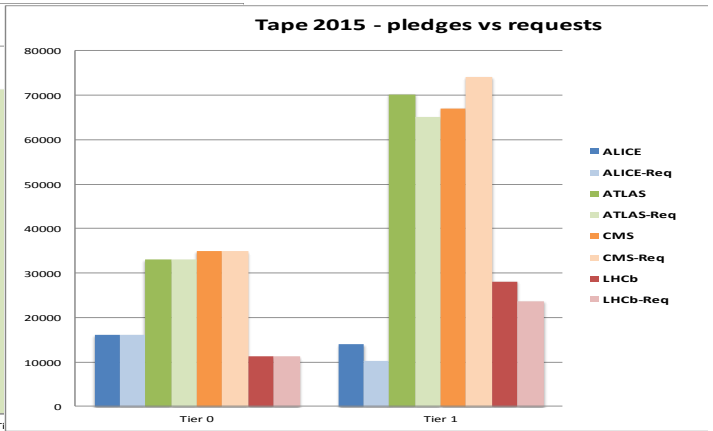
- Requests unchanged since October 2014

Run 2 readiness

- WLCG sites prepared for Run 2
- Some delays in deploying new resources, but no major issues reported at Tier 1s
  - Tier 0 delayed due to firmware issue on a large batch of hardware
  - All experiments now at 90% pledges, rest coming online asap
- Major procurement exercise has essentially doubled existing installations
  - Major procurement activities – that took significant effort and time







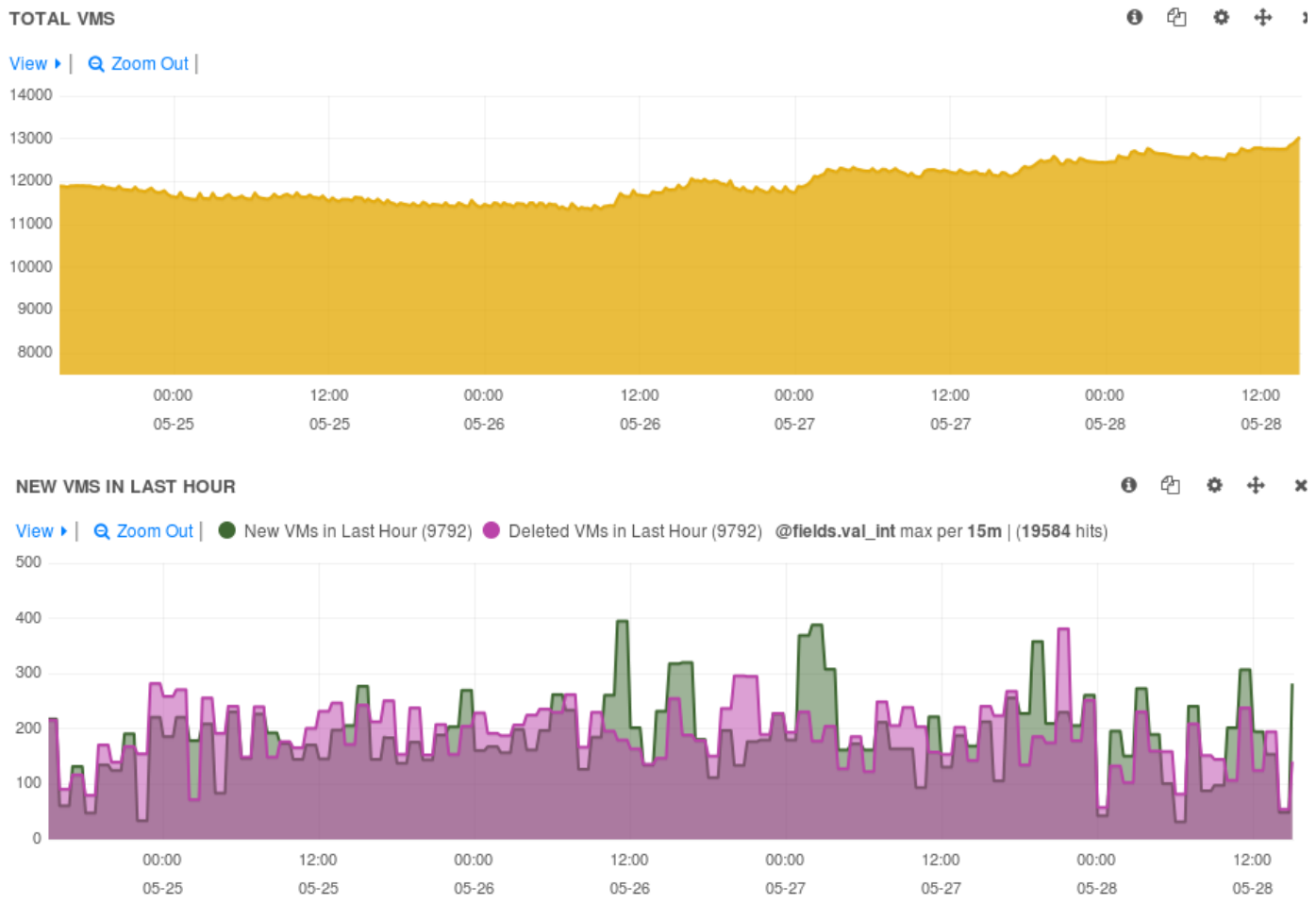
Status of CERN 2015 capacity

- Disk and tape fully installed at the beginning of 2015
- Current CPU status:

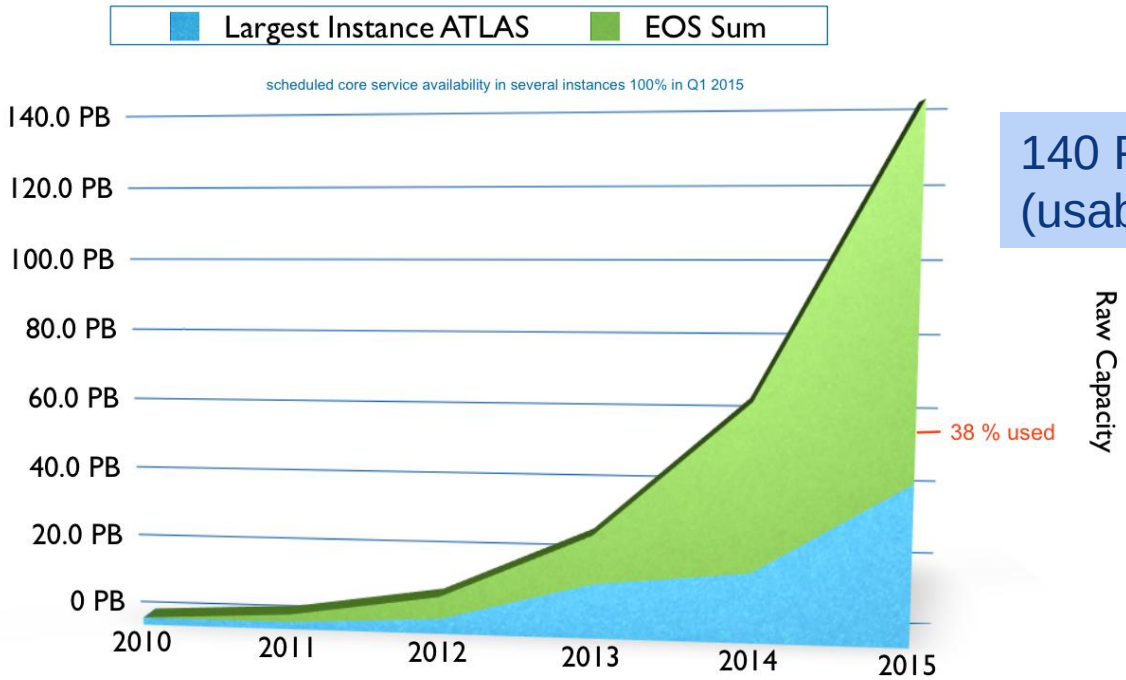
VO	Installed capacity vs. pledge
ALICE	75%
ATLAS	100%
CMS	100%
LHCb	100%

- Late fulfilment because of
  - Late delivery of hardware
  - Firmware issue detected after burn-in on 850 machines
  - KVM security issue
- Prospects for completion
  - Capacity being added right now
  - Full level: second half of June 2015

CERN Cloud Service in numbers

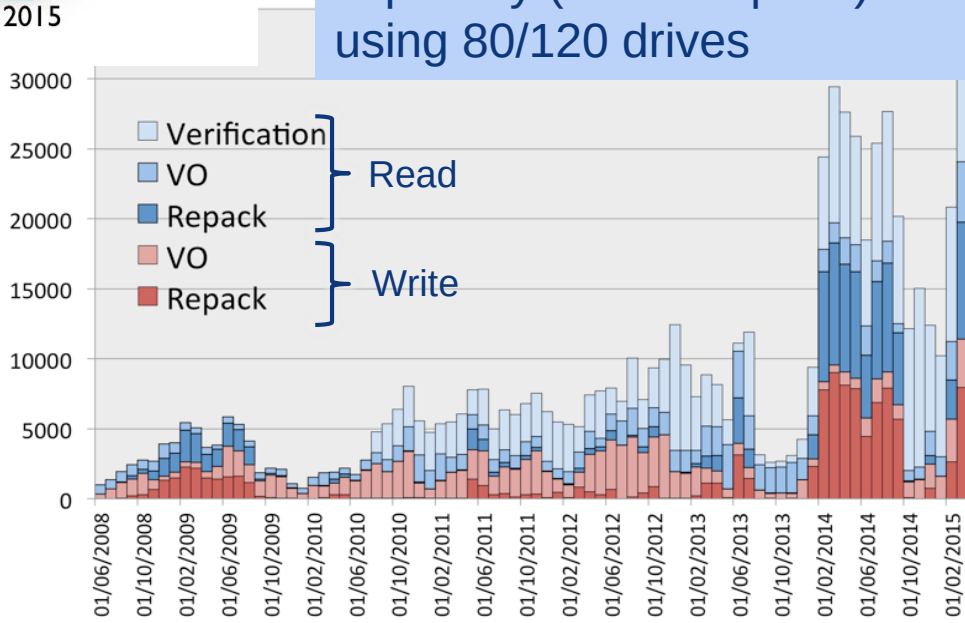


Tier 0 storage



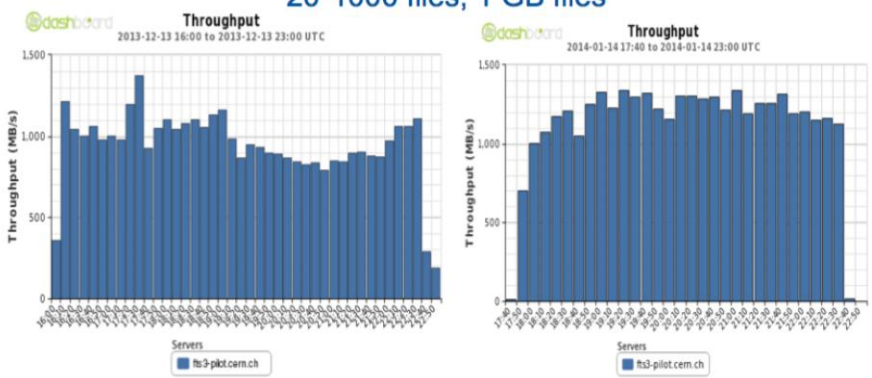
140 PB raw space in EOS
(usable space is half of this)

Tape system shown to have very high performance capability (18 GB/s peak) using 80/120 drives



New FTS in production

20*1000 files, 1 GB files



Fixed config (60 files, 20 streams)

Auto-config optimiser

HSF Next steps

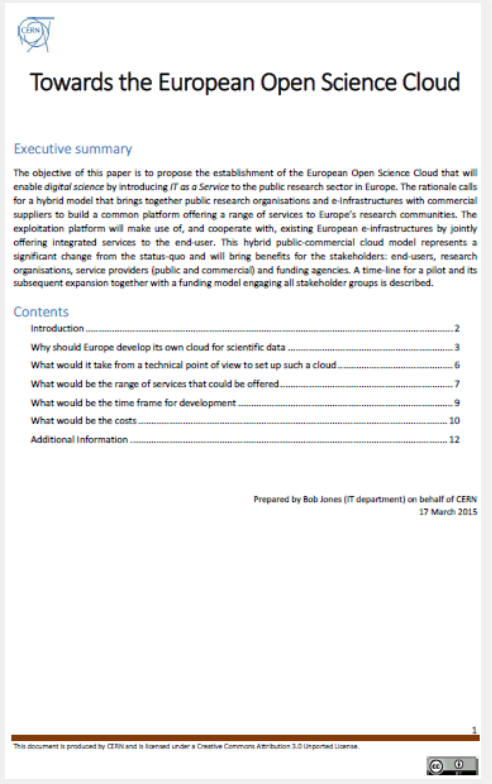
- Technical forum
  - Place for technology discussion and dissemination of experiences
  - Publish technical notes
  - Help build expertise in the community
  - Concurrency forum – continue as prototype
  - Interest group – optimising reconstruction software
- Training
  - Consensus that is important initially
  - Several suggestions and volunteers to work with existing schools etc
  - Learn from experience of the UK SSI
  - Working group set up
- Set up SW Knowledge Base
  - Prototype exists
  - Initially try and gather/catalogue software in use and available – provide ability to comment and cross-ref usage
  - Important that community contributes to this
- Build/test/integration infrastructure
  - Mentioned by several groups; examples exist in labs and projects
  - Under consideration
- Under consideration:
  - Licensing issues – must be open source - recommendations needed?
  - Consultancy/SWAT teams – ready to start some activities here – to be better defined and scoped

Pre-Commercial Procurement proposal

- A Pre-Commercial Procurement proposal was submitted to the H2020 ICT 8a call in April 2015 for innovative IaaS cloud services
- The proposal is currently under evaluation by the EC
- Official feedback expected in September
- These slides give an overview of the proposal

Towards the European Open Science Cloud

- **Hybrid** – link public research organisations, e-Infrastructures & commercial cloud services
  - Use GEANT network to link Research Infrastructures, repositories (EUDAT, OpenAIRE), EGI, PRACE etc. to commodity commercial cloud services (multiple providers)
  - A cornerstone of the Open Science Commons*
- **Trust** - Researchers keep control of the cloud and their data
  - Guarantee a copy of all the data is kept on public resources
  - Ensure long-term preservation of the data
  - Insulate users from changes of service supplier and technology
- **Economy** - Must be cheaper than the '*build our own*' approach
  - Avoid separate 'silos' for each Research Infrastructure/Community
  - Profit from the economies of scale in commercial data centres



Towards the European Open Science Cloud

Executive summary

The objective of this paper is to propose the establishment of the European Open Science Cloud that will enable *digital science* by introducing *IT as a Service* to the public research sector in Europe. The rationale calls for a hybrid model that brings together public research organisations and e-Infrastructures with commercial suppliers to build a common platform offering a range of services to Europe's research communities. The exploitation platform will make use of, and cooperate with, existing European e-infrastructures by jointly offering integrated services to the end-user. This hybrid public-commercial cloud model represents a significant change from the status-quo and will bring benefits for the stakeholders: end-users, research organisations, service providers (public and commercial) and funding agencies. A time-line for a pilot and its subsequent expansion together with a funding model engaging all stakeholder groups is described.

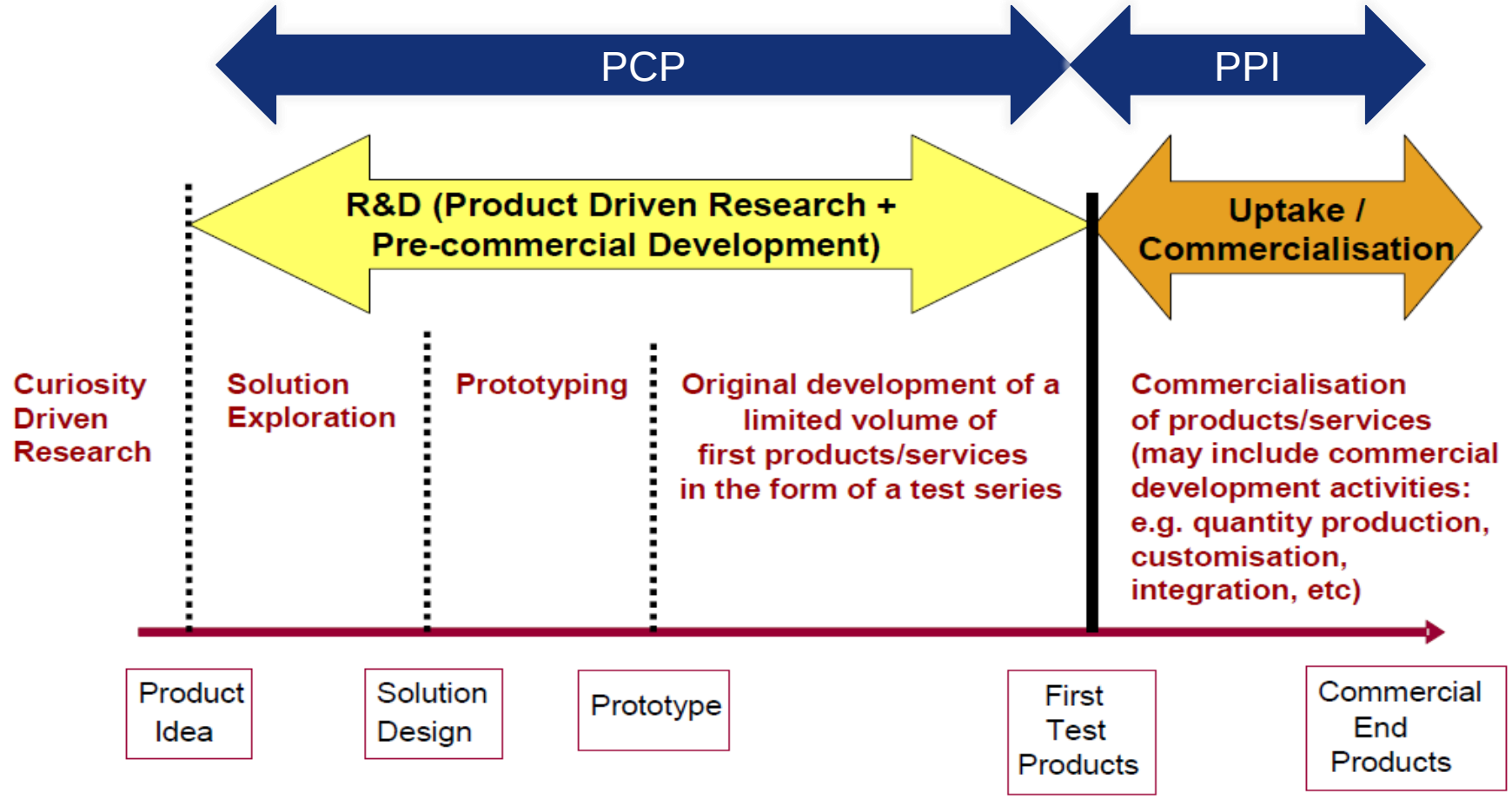
Contents

Prepared by Bob Jones (IT department) on behalf of CERN

17 March 2015

* http://go.egi.eu/osc

http://dx.doi.org/10.5281/zenodo.16140



Typical Product Innovation Life Cycle

Why PCP?
Commercial IaaS exists but not certified, integrated with public e-infrastructures, offering std interfaces with suitable SLA and contractual terms & conditions.

PPI
Potential follow-on project if this PCP project is successful

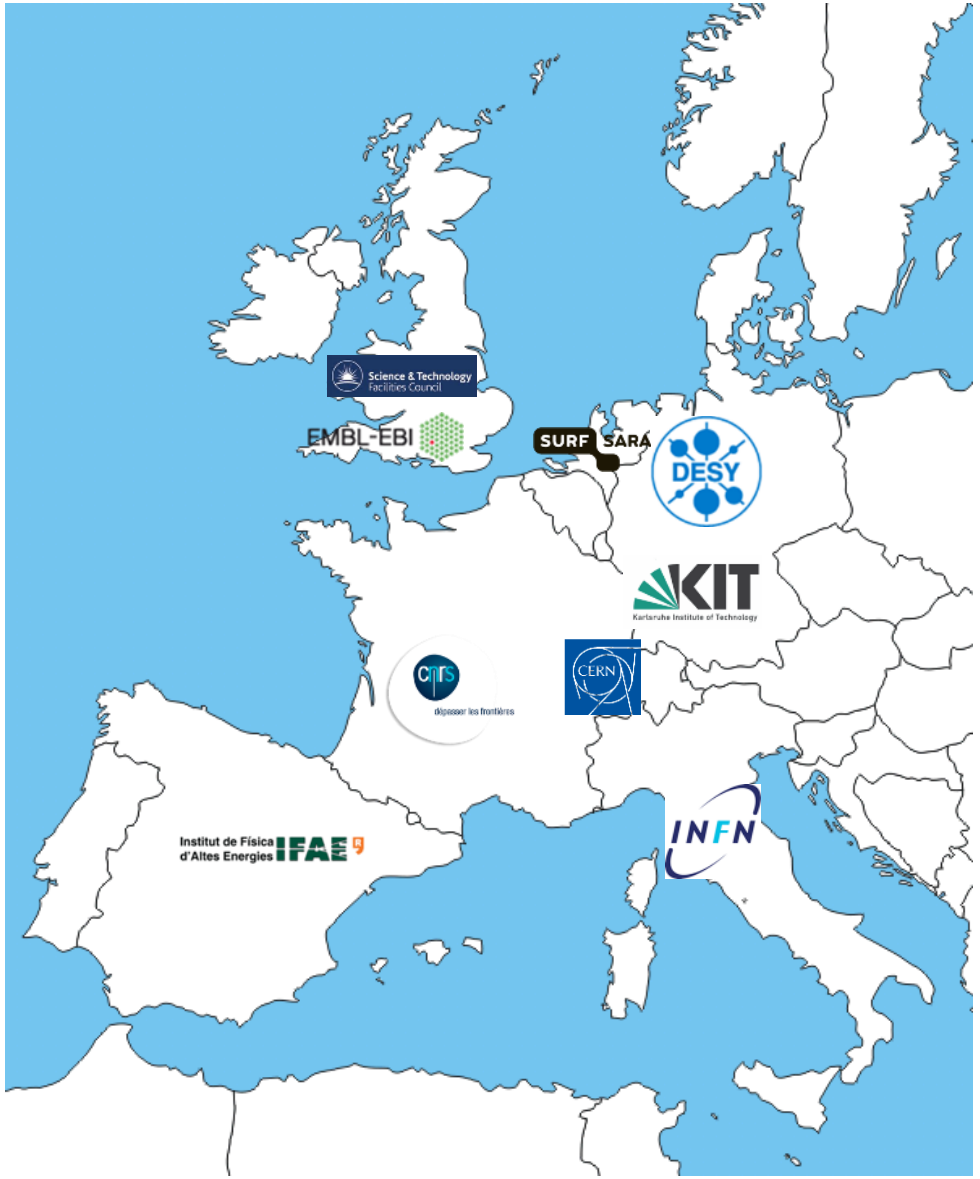
Proposed Joint Pre-Commercial Procurement

The group of buyers have committed
- >1.6M€ of funds
- Manpower
- Applications & Data
- In-house IT resources

To procure innovative cloud services

Integrated into a hybrid cloud model:
- Commercial cloud services
- e-Infrastructures
  - GEANT network
  - eduGAIN Fed. Id mgmt.
  - EGI Fed Cloud
- Potentially host data services from EUDAT, INDIGO-Datacloud, etc.
- In-house IT resources

Made available to end-users including BBMI, DARIAH, ELIXIR, EISCAT_3D, EPOS, INSTRUCT, LifeWatch, LHC, etc.



User Groups

The cloud resources procured will be made available to user groups during the pilot phase

- LHC experiments via WLCG
  - Procured resources will count against the buyers' pledges (du e)
  - CERN will provide the interface via Tier-0 (OpenStack in tend
- ELIXIR
  - Managed by EMBL-EBI via the ELIXIR Compute platform
- Other research communities via EGI Fed Cloud
  - Request OCCI interface in tender spec.
- Local users at each buyers site
  - Each buyer is responsible for integration
  - Request web GUI interface in tender spec.
    - Sites can also use OpenStack or OCCI interface as well

BBMRI
DARIAH
EISCAT_3D
EPOS
INSTRUCT
LifeWatch
Long-tail

Each buyer decides what fraction of their procured resources is made available to each user group but cannot assign only to their local users

Collectively the users will form a user group with a role in the project to define requirements and provide feedback on pilot deployments

Science cloud – key points

- Essential to understand how to procure commercial services
  - Highly likely to be an important component of LHC(HEP) computing in future
  - Costs are becoming very interesting
- Today still more cost effective to operate our own facilities, but this situation is expected to change
- Hybrid model gives us flexibility
  - Does not save staff effort as we still need to operate services there, as well as maintaining in-house services
- Important to understand long-term sustainability of WLCG services and infrastructures
- Under assumption that long term computing funding remains flat, essential that we understand how best to make use of funding