

NORDUGRID

*Grid Solution For Wide Area
Computing and Data Handling*

Distributed Computing in Norway – *Software and middleware*

Jon Kerr Nilsen, Oslo, NO, 2 Oct 2015

RECFA open session

Norwegian Grid Team

- Farid Ould-Saada, *UiO/phys.dept.* – NorduGrid coord, ATLAS ICB, Nordic LCG
- David Cameron, *UiO/phys.dept.* – ATLAS DC lead, aCT, ATLAS@HOME, ARC developer
- Vincent Garonne, *UiO/phys.dept.* – DDM, Rucio developer and coordinator (ATLAS-funded)
- Cedric Serfon, *UiO/phys.dept.* – DDM, Rucio developer (Atlas-funded)
- Silje Raddum, *UiO/phys.dept.* – PhD student, ARC developer
- Dmytro Karpenko, *UiO/NeIC* – Nordic T1 operator, ARC debugger
- Jon Kerr Nilsen, *UiO/NeIC* – aCT, ARC release manager, ARC developer
- Håvard Helstrup, *HiB* – Research, Nordic LCG
- Boris Wagner, *UiB* – Nordic T1 operations, research
- Saerda Halifu, *UiB* – Nordic T1 operations

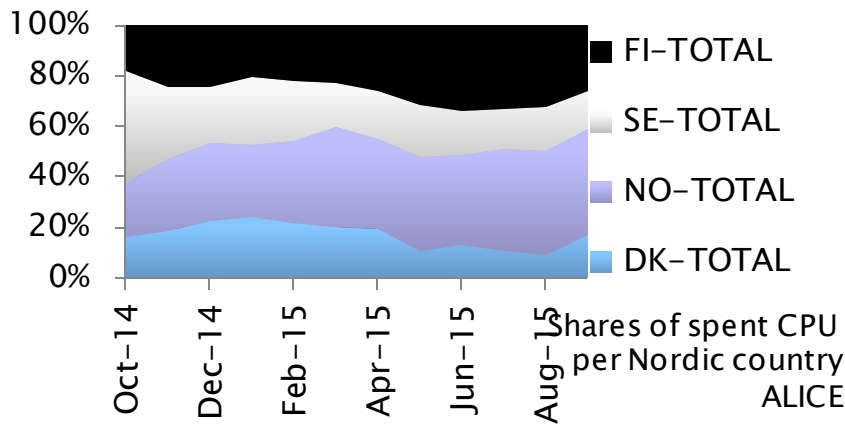
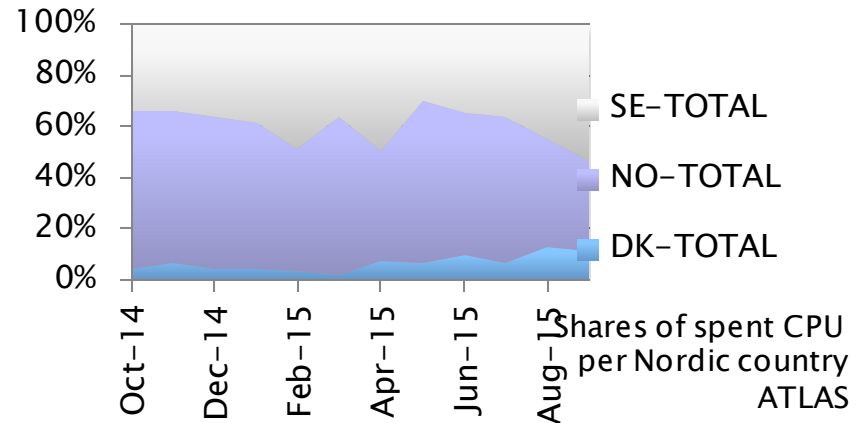
- *et al.*

The Full Chain

- ATLAS Nordic T1 operations
- ALICE Nordic T1 operations
- Advanced Resource Connector (ARC)
Distributed Computing middleware
- Rucio – Distributed Data Management
- ARC Control Tower (aCT) – connecting
ATLAS with ARC
- ATLAS software

Computing infrastructure

- Two sites
 - ALICE: fimm.bccs.uib.no in Bergen
 - ATLAS: grid.uio.no in Oslo
- Nordic T1 pledges
 - 9% of total ALICE resources, 6% of total ATLAS resources
 - Of which Norway should contribute ~50% for ALICE, ~25% for ATLAS

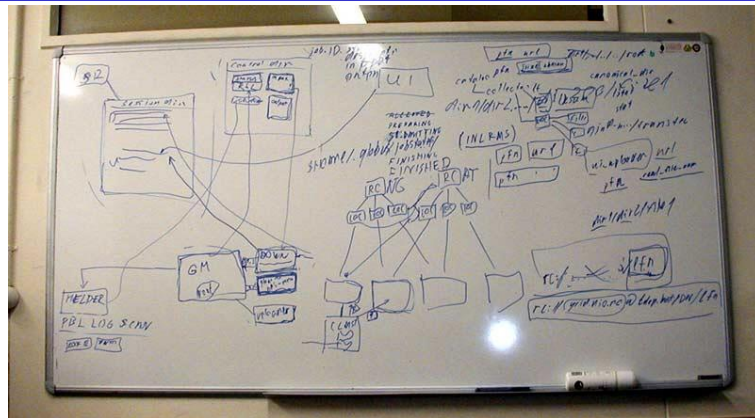


Norwegian ALICE Tier-1

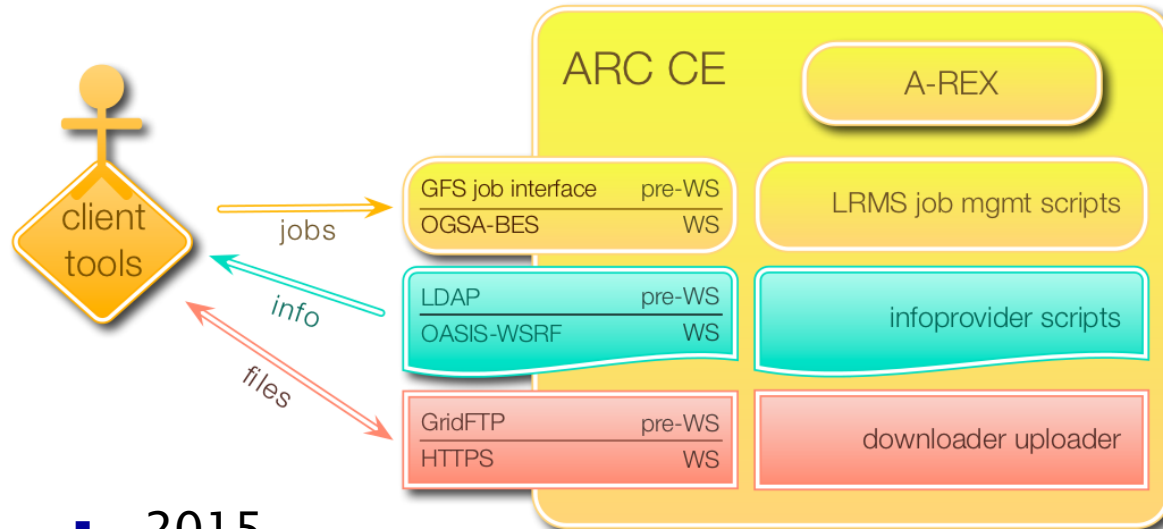
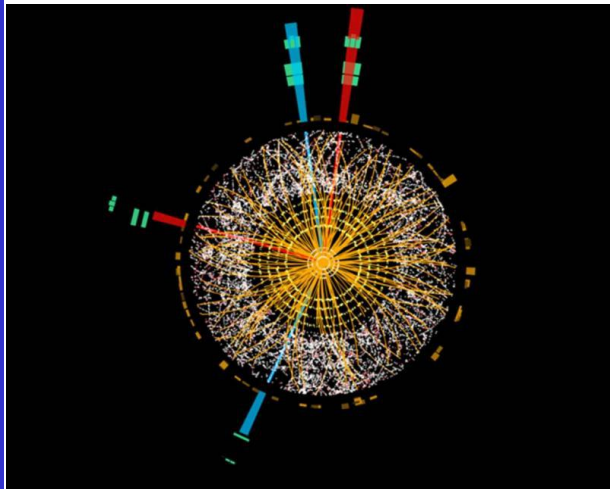
- The ALICE Tier-1 is part of the IT department of the University of Bergen (UiB).
- Resources provided:
CPU: 704 cores
Disk: 1 PB
Tape: 200 TB
- The storage resources are provided via Nordic T1 with the dCache middleware.
- Grid jobs come directly from CERN through the ALICE Grid middleware (ALiEn)



ARC Achievements and Highlights



- 2001 – 2002
 - “Transverse momentum distribution” of pre-LHC data simulated on HPC resources ...
- 2012 – 2013
 - ☞ Higgs discovery, Nobel prize

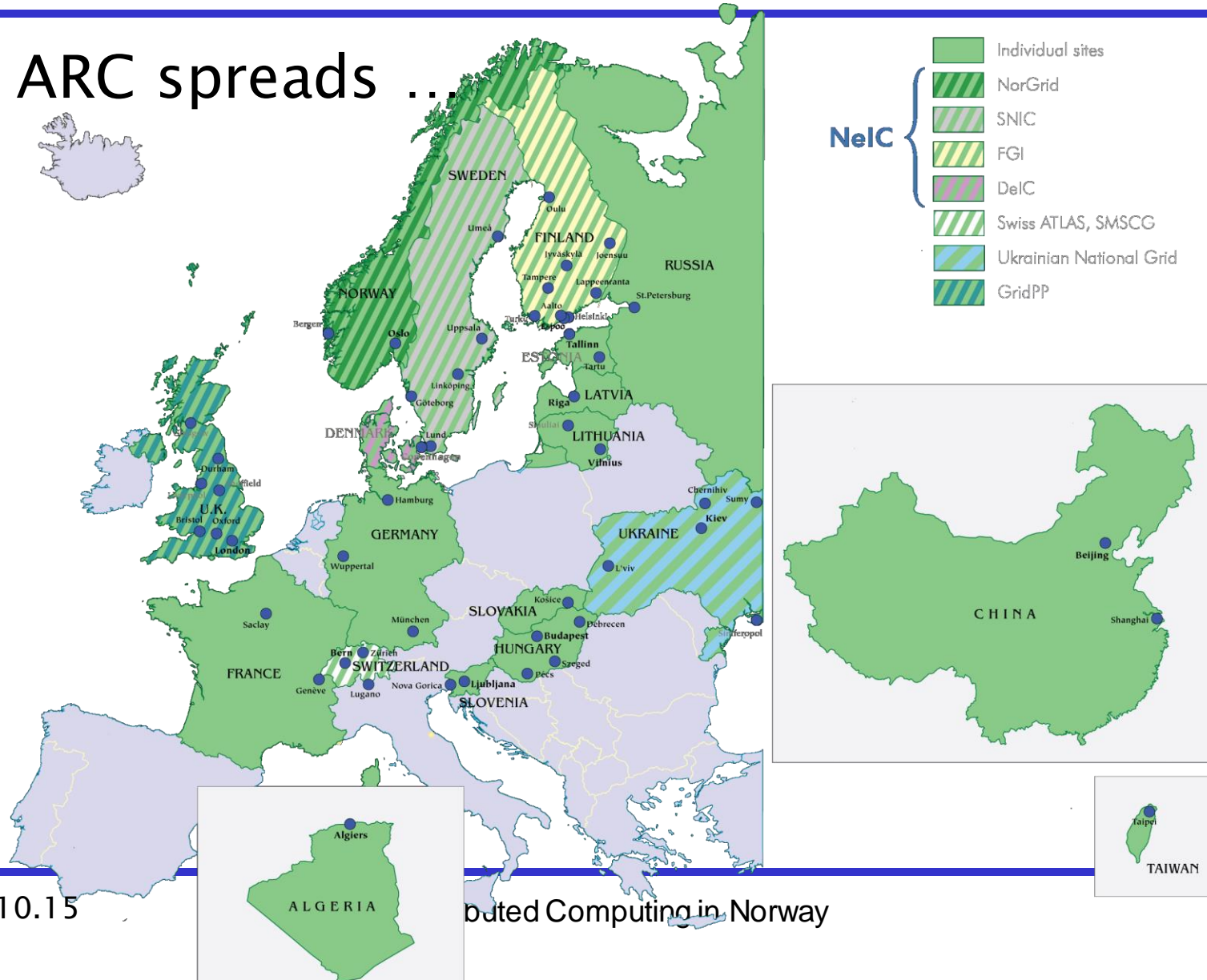


- 2015 – ...
 - ☞ 13 TeV collisions, Dark Matter?

ARC Deployment Map

ARC spreads ...

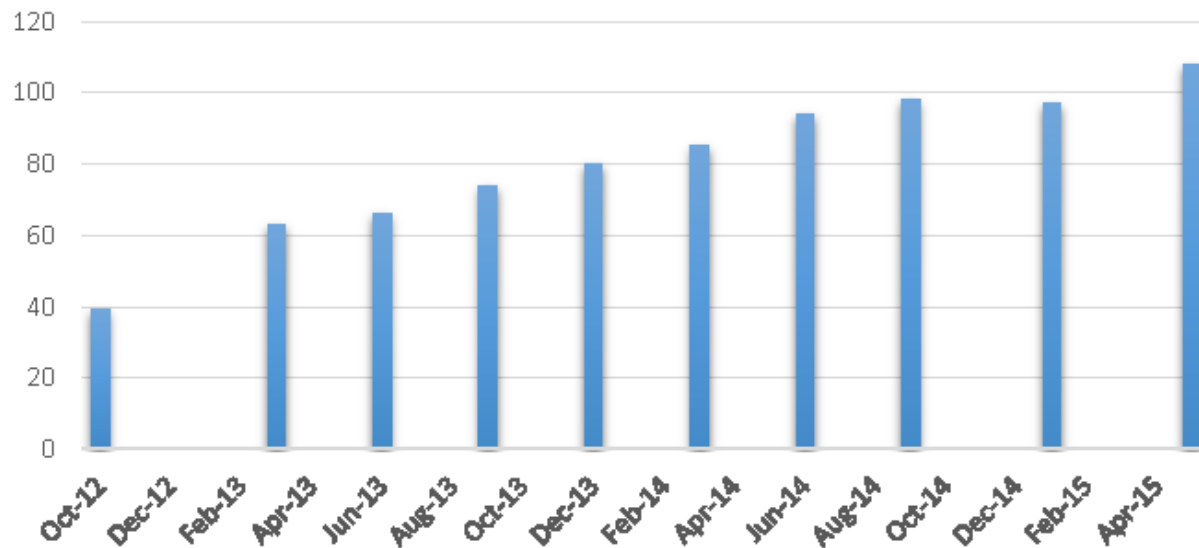
- NeIC**
- Individual sites
 - NorGrid
 - SNIC
 - FGI
 - DeIC
 - Swiss ATLAS, SMSCG
 - Ukrainian National Grid
 - GridPP



ARC CE in EGI

- Some trend ...

ARC CE services in EGI database - GOCDDB



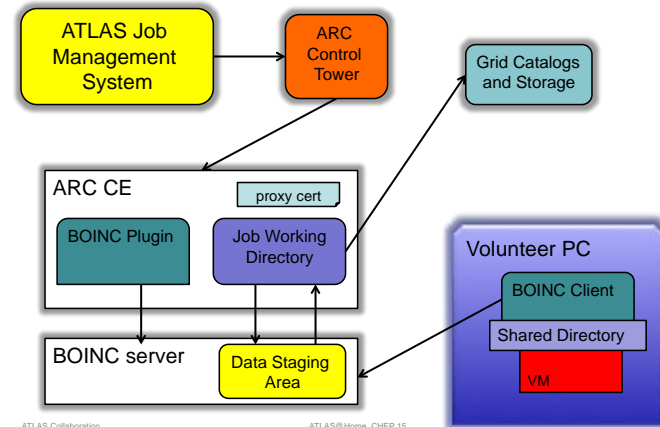
- ARC CE at CERN
 - ☞ 2 ARC-CEs and HTCondor attached to it
 - ☞ Production jobs expected soon
 - 2 ARC CEs for BOINC

ARC and Volunteer resources

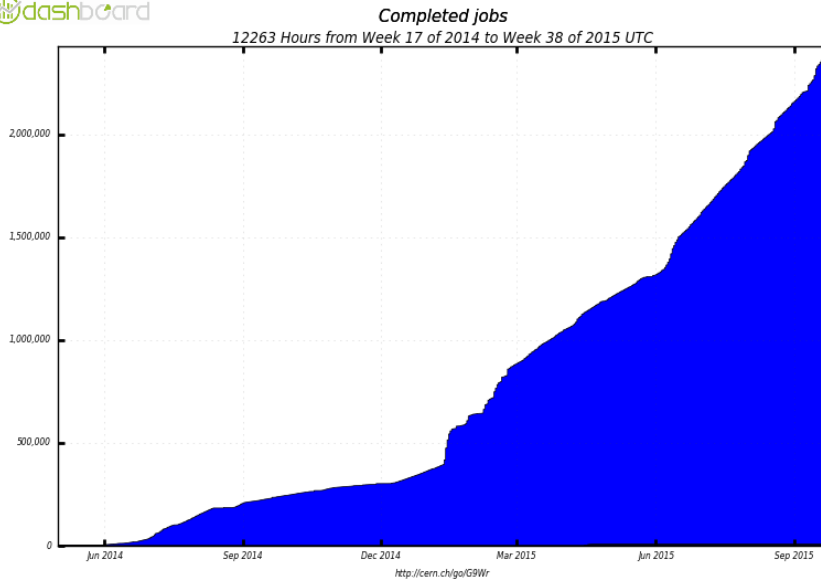
- [ATLAS@home](#)
- ARC and ARC Control Tower (aCT) as gateway between ATLAS job management system and BOINC – software used to distribute jobs to volunteers

UiO : Department of Physics
University of Oslo

Basic ATLAS@Home Architecture



dashboard



- Equivalent to a large Tier-2 centre
- 2.5M completed jobs so far

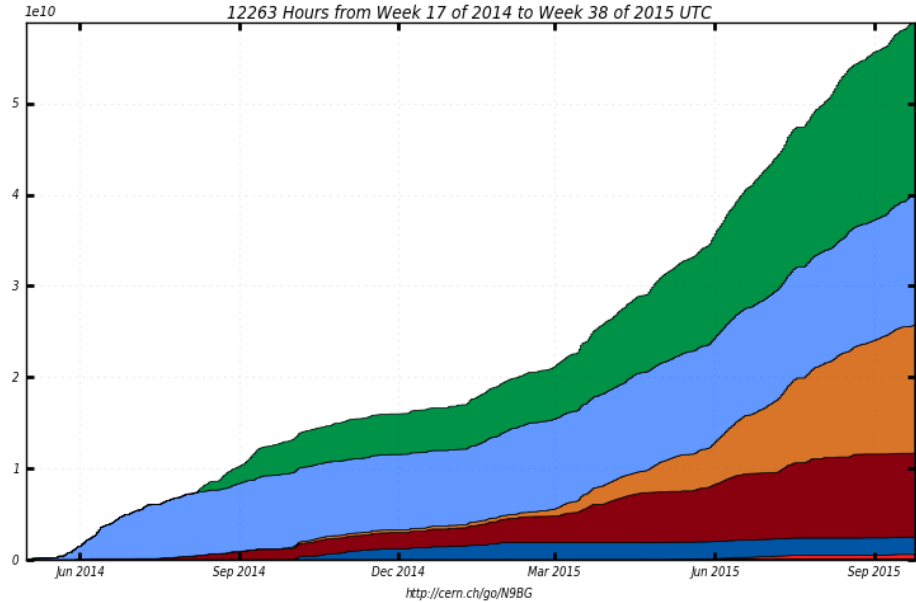
ARC back to HPCs

LRZ-LMU-C2PAP MC, DE
 MPPMU-HYDRA MC, DE
 LRZ-LMU-C2PAP, DE
 LRZ-LMU MC, DE
 CSCS-TODI, CH
 BEIJING-PI, Shanghai China
 BEIJING-ERA, Chinese HPC "grid"

CPU consumption Good Jobs in seconds (Sum: 58,875,929,637)

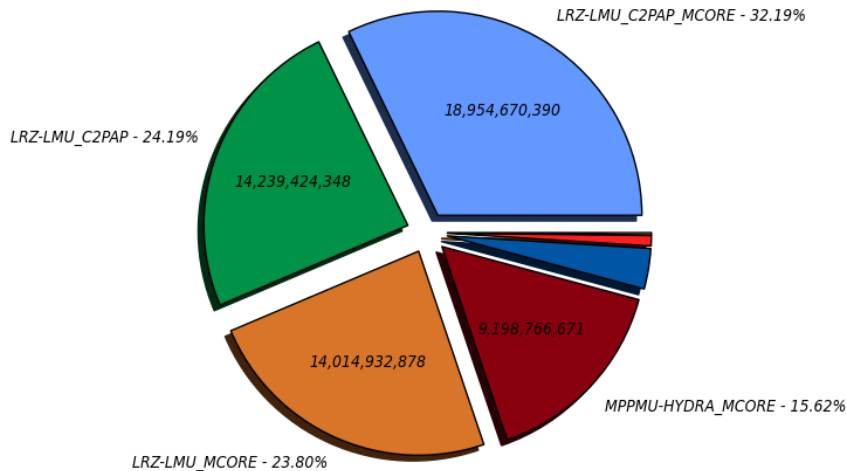


CPU consumption Good Jobs in seconds
 12263 Hours from Week 17 of 2014 to Week 38 of 2015 UTC



■ LRZ-LMU_C2PAP_MCORE (18,954,670,390)	■ LRZ-LMU_C2PAP (14,239,424,348)	■ LRZ-LMU_MCORE (14,014,932,878)
■ MPPMU-HYDRA_MCORE (9,198,766,671)	■ CSCS-TODI (1,885,430,511)	■ BEIJING-ERA_MCORE (480,542,621)
■ BEIJING-PI_MCORE (102,162,218)		

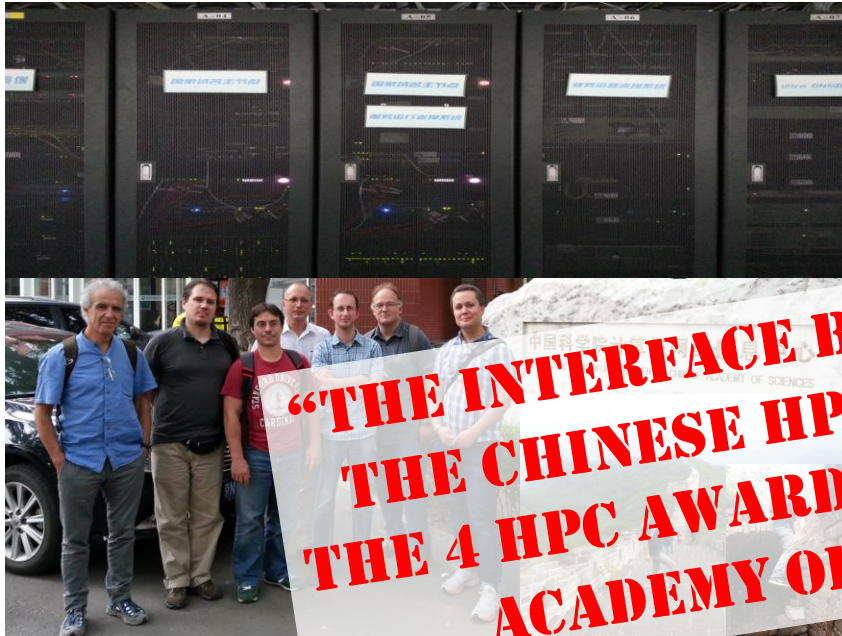
Total: 58,875,929,637 , Average Rate: 1.333 /s



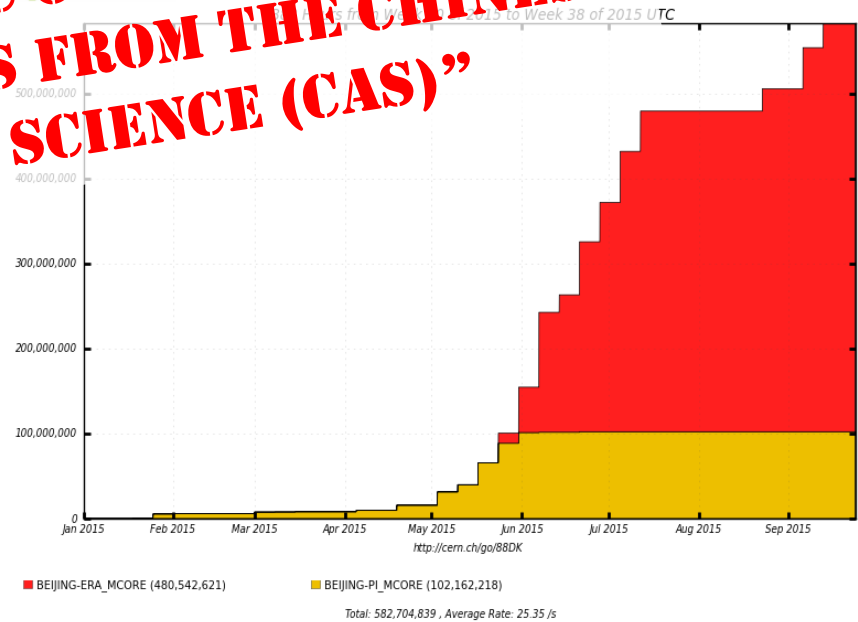
<http://cern.ch/go/N9BG>

■ LRZ-LMU_C2PAP_MCORE - 32.19% (18,954,670,390)	■ LRZ-LMU_C2PAP - 24.19% (14,239,424,348)
■ LRZ-LMU_MCORE - 23.80% (14,014,932,878)	■ MPPMU-HYDRA_MCORE - 15.62% (9,198,766,671)
■ CSCS-TODI - 3.20% (1,885,430,511)	■ BEIJING-ERA_MCORE - 0.82% (480,542,621)
■ BEIJING-PI_MCORE - 0.17% (102,162,218)	

ATLAS+ARC in China



- Remote ARC CE
- New Python backend
- David Cameron, Andrej Filipcic, Jon Kerr Nilsen, Silje Raddum (as ATLAS qual. team), Chao Qiang, Hao Haili



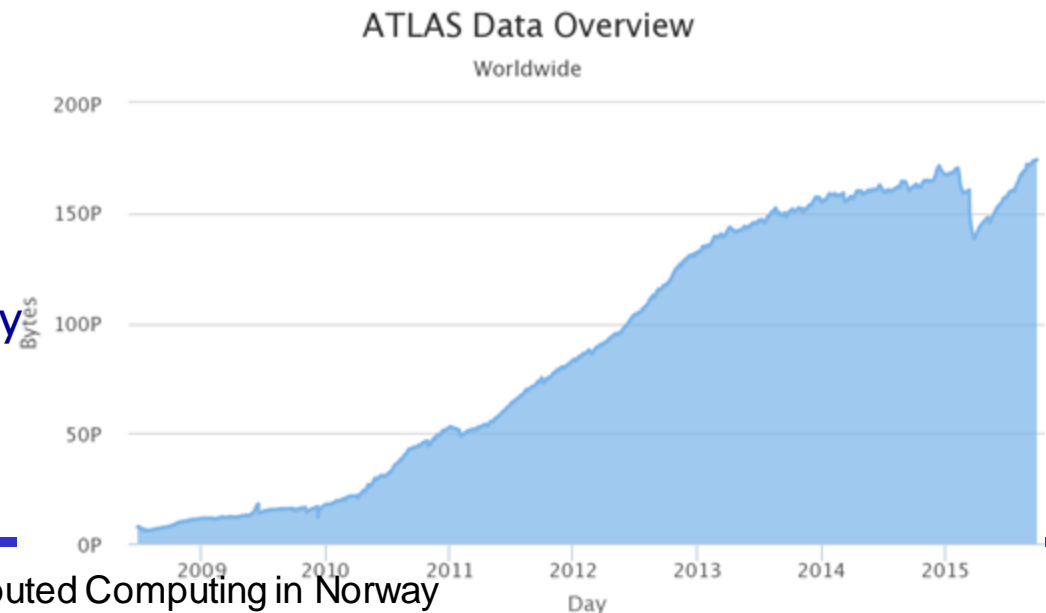
Two systems:

- Pi – CE in Beijing, jobs through ssh to Shanghai
- ERA – CE in Beijing connects to Chinese HPC Grid

The ATLAS Distributed Data Management

- The Distributed Data Management project manages ATLAS data on the grid
- The Rucio Software project
 - new major version to ensure
 - system scalability
 - reduce operational overhead
 - support new ATLAS use cases for LHC
- Rucio has demonstrated very large scale data management
 - 130 Storage sites
 - 2000 users
 - 170 PB, 700 M files
 - 25 M files transferred/day

- Rucio is an open-source project
- 24 contributors/authors from 8 institutes
- People involved: V.Garonne (Project lead, Core developer), C.Serfon (DDM operations lead, Core developer), D.Cameron (ATLAS Distributed Computing operations lead)



- J. Phys.: Conf. Ser., 2015 ☺
 - Dynamic Resource Allocation with the arcControlTower, A. Filipčič et al.;
 - ARC Control Tower, A Flexible Generic Distributed Job Management Framework, J. K. Nilsen et al.
 - The ATLAS ARC backend to HPC, S. Haug et al., [ATL-SOFT-PROC-2015-039](#)
 - ATLAS computing on CSCS HPC, M. Hostettler et al., [ATL-SOFT-PROC-2015-010](#)
 - Bringing ATLAS production to HPC resources – A use case with the Hydra supercomputer of the Max Planck Society, J. Kennedy et al., [ATL-SOFT-PROC-2015-055](#)
 - ATLAS@Home: Harnessing Volunteer Computing for HEP, D. Cameron et al., [ATL-SOFT-PROC-2015-012](#)
 - 2 more not submitted ...
 - We need more ... Still referring to old ARC publications
- Publications in refereed journals ☹
 - Task force necessary – work starting

Achievements

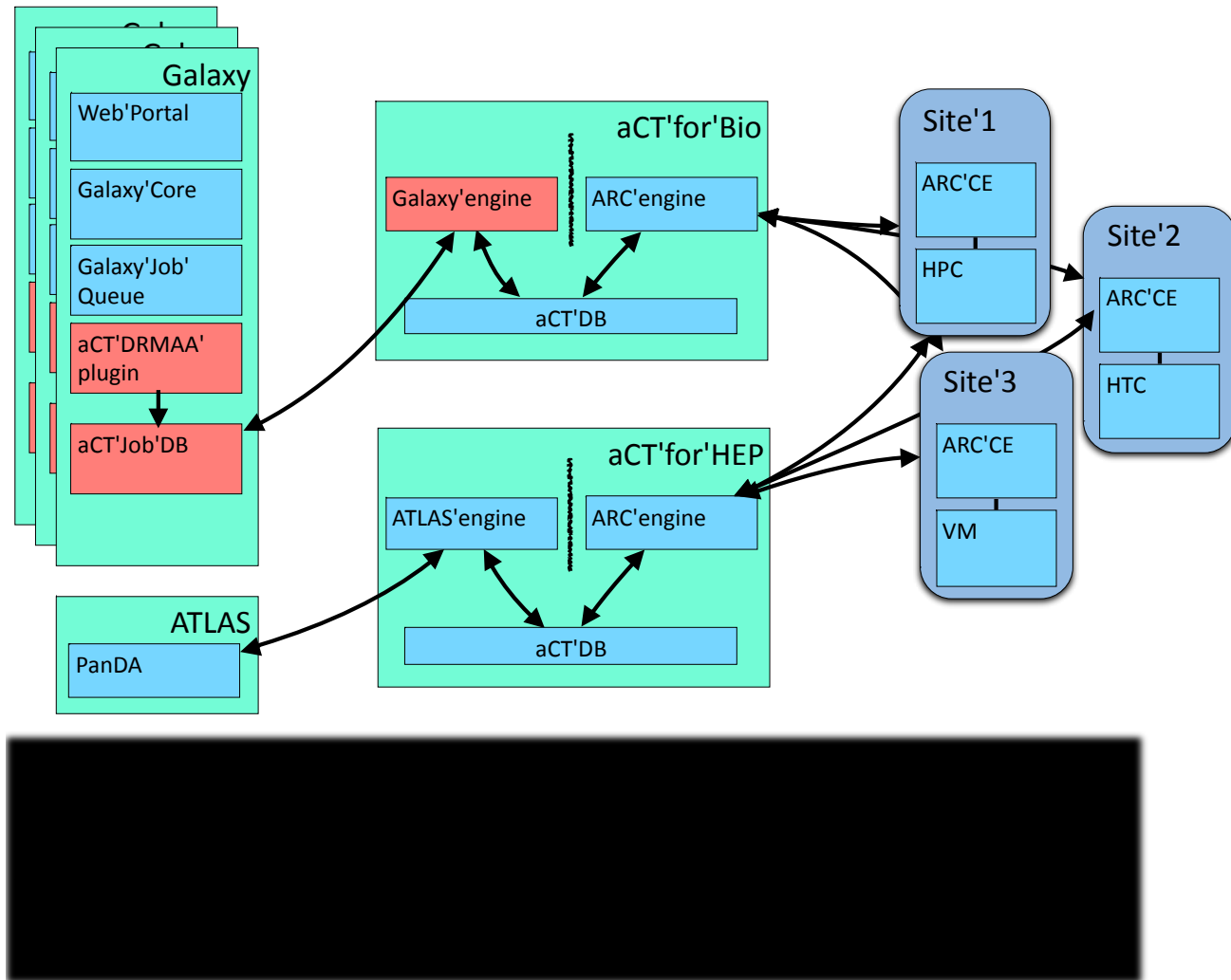
- NorduGrid and ARC still – Year 14
- ARC CE recommended European CE at least for ATLAS
- ARC CE deployment at CERN on-going
- Continuous steady increase of ARC CE usage in WLCG bringing new countries to the ARC community
- ARC successfully integrated with volunteer computing (via Boinc) AND HPCs
- The ARC Control Tower has been used efficiently and reliably to manage WLCG production over ARC resources
- The aCT is being investigated/used by non-HEP communities
 - (Galaxy and other open possibilities)
- In 2015 delivered a major release, the NorduGrid ARC 15.03 release
 - Major release represents outcome of an important consolidation, cleanup process.
- HPC award from the Chinese Academy of Science for interfacing ARC and ERA
- We have plans and dreams and/but need funding!

The need for New Projects

- Maintenance & support
 - ARC for infrastructures
 - NeIC supervision
- Long-term R&D project
 - EXERCISE
 - Opening HPCs to data intensive science at exascale
 - Submitted to H2020 call for Future and Emerging Technology
 - failed for lack of IMPACT!
- Shorter-term R&D project
 - ARChestrates
 - Sensitive & Intensive data
 - Beyond HEP
 - Submitted to Research Council Norway
- HEP
 - Synergy with research, especially ATLAS
 - Big customers, requiring extra resources and efforts
 - HEP is main source of sustainability fuel
- ARC4eInfrastructure
 - NorduGrid's maintenance of the Advanced Resource Connector (ARC) software over a period of three (3) years.
 - 2015-2017
 - Overall goal
 - Provide sustainable support for the ARC technology platform that enables NeIC infrastructure and thus provides a gateway to international e-infrastructures

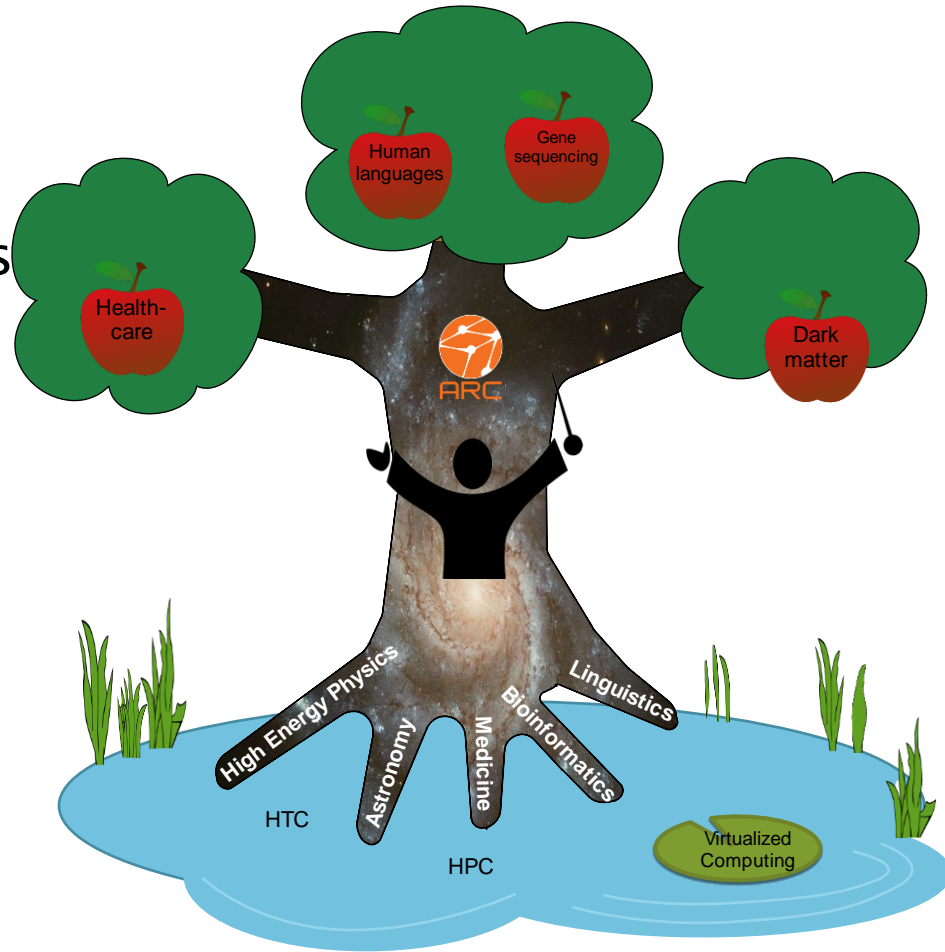
ARChestrator – Advanced Resource Connector for Secure Data–Services everywhere

- University of Oslo, NorduGrid and USIT
- Integrate portals for science and data analysis into the distributed computing environment to be orchestrated by aCT.
- Resulting framework capable of intelligent adaptive data-aware job brokering that will maximise job efficiency, optimise data I/O and have inherent resiliency.

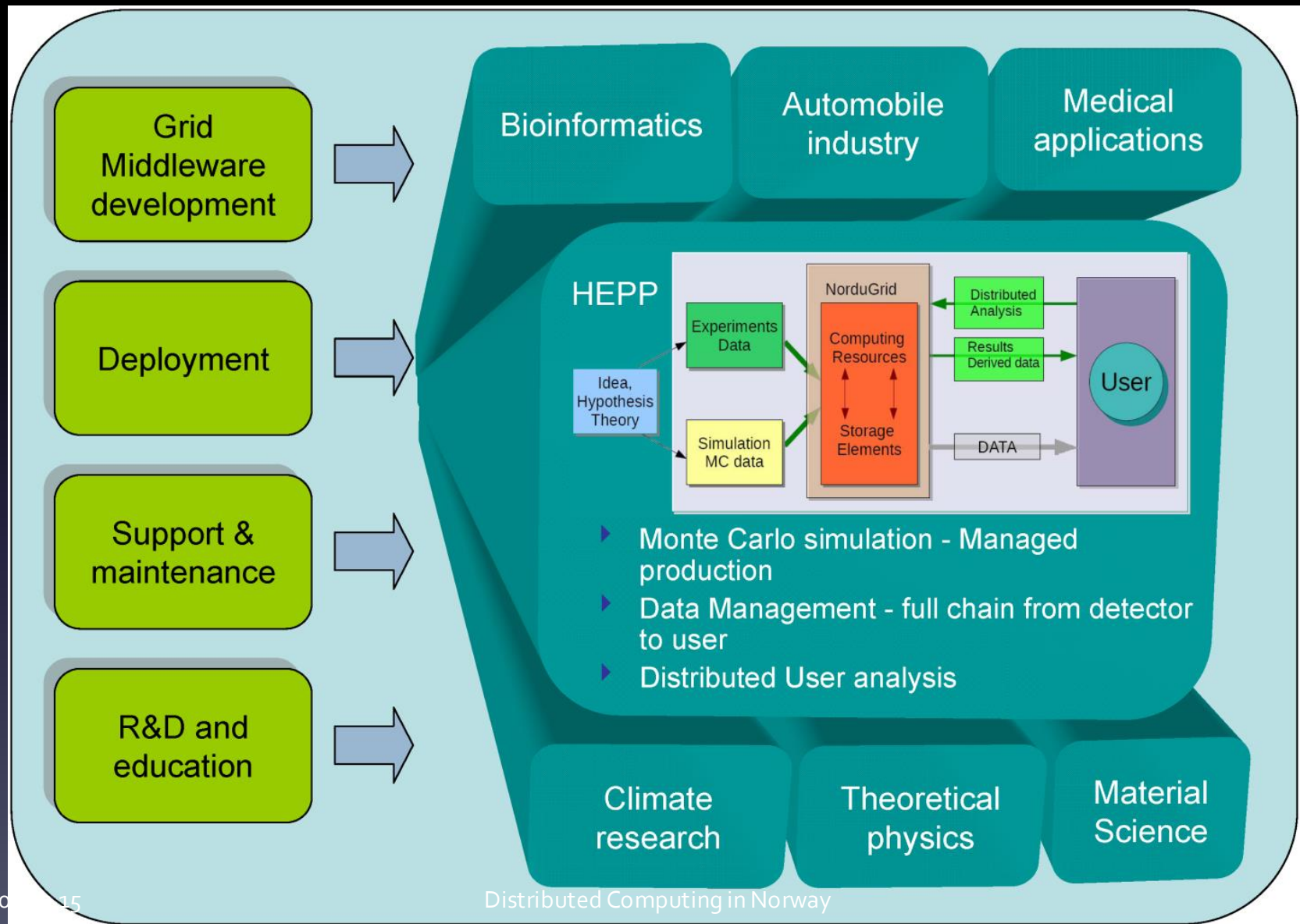


ARChestrator – Advanced Resource Connector for Secure Data–Services Everywhere

- IMPACT of ARChestrator
- Access to and efficient utilisation of diverse, complementary, heterogeneous resources, ranging from powerful supercomputers to personal computers, whether dedicated, shared or voluntary
- Satisfy computational needs of diverse user communities through a shared Grid of heterogeneous resources



Is a complete, simple, efficient “Distributed” Computing environment for research on HPCs ++ possible?



Conclusion

- Grid activities in Norway centered in Oslo and ATLAS
 - Heavily involved in the full ATLAS computing chain
 - Focus on ARC as the Grid middleware
- ARC usage is growing
 - ARC is the preferred choice for new CEs in ATLAS
- Basic funding just enough to maintain ARC – project funding needed to push development further
 - Will try Exercise again
 - ARChestrated application pending
 - ARC4eInfrastructures kick-off in October
- We are connecting data intensive science to advanced resources
 - HTC, HPC and more

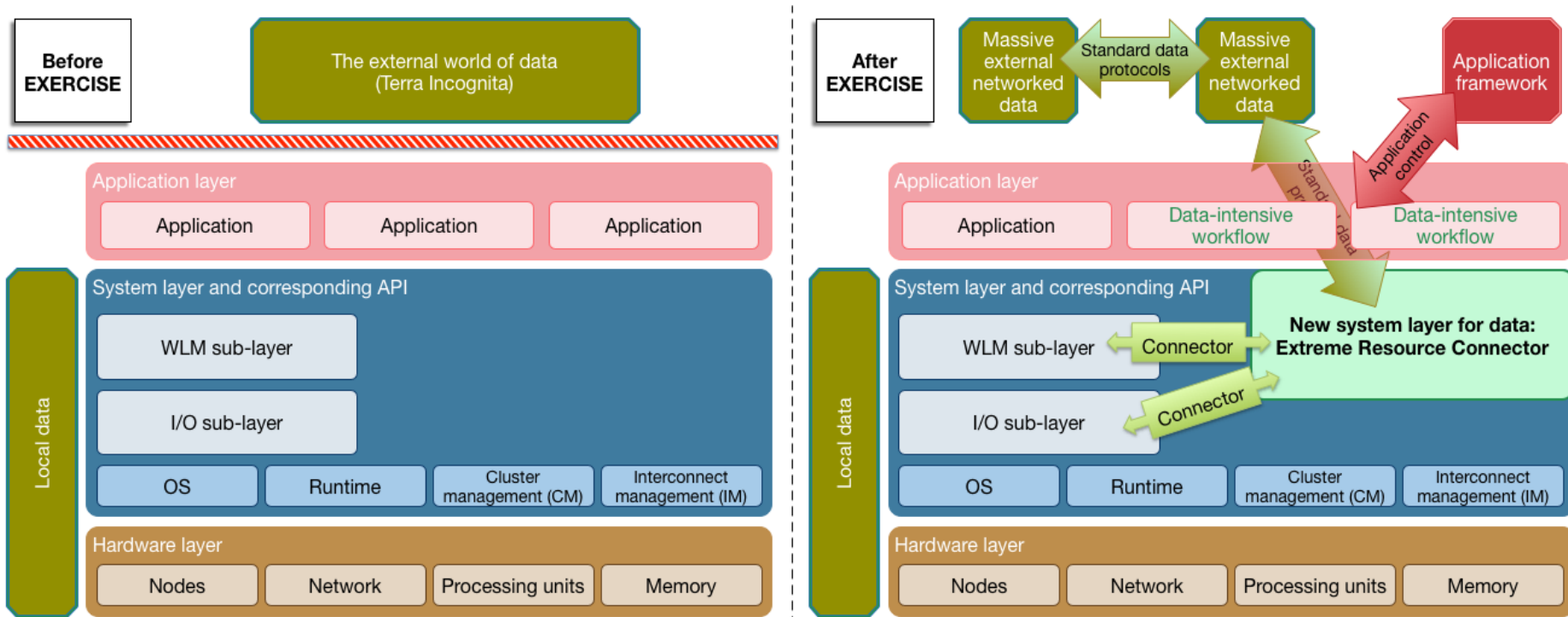
THANK YOU!

BACKUP SLIDES

Exercise

EXTREME RESOURCE CONNECTOR FOR DATA INTENSIVE SCIENCE AT EXASCALE

- NorduGrid, HPC centres and experts, Researchers in data intensive science, technology providers ...
- Extreme Resource Connector for Data Intensive Science at Exascale
- From “in-house” HPC computing to external data-intensive EXERCISEs



EXERCISE Objectives

1. *Design and develop a new layer in the HPC system software stack for efficient automated processing of extreme scientific data on future Exascale HPCs – **Extreme Resource Connector (ERC) will open up HPC centres for massive networked data processing.***
2. *Provide extensions to the foundational software layer of the HPC systems to address the Exascale I/O needs of data-intensive scientific workflows.*
3. *Demonstrate scalability of our novel approach to Exascale requirements through deployment and evaluation of the new system software developed by the project on some existing Top-500 HPC systems, future vendor prototypes, and by simulation.*
4. ***Provide recommendations for future HPC systems design and usage, addressing data-intensive processing use cases.***
5. *Exploit system software developed project through the HPC centres beyond the immediate project participants and disseminate results to relevant standards bodies, HPC service providers, vendors, technology providers and researchers.*