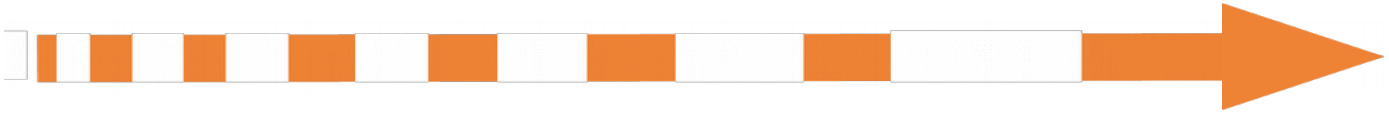
Data Lakes and DOMA



eXtreme DataCloud

Data Management for extreme scale computing

Paul Millar

paul.millar@desy.de

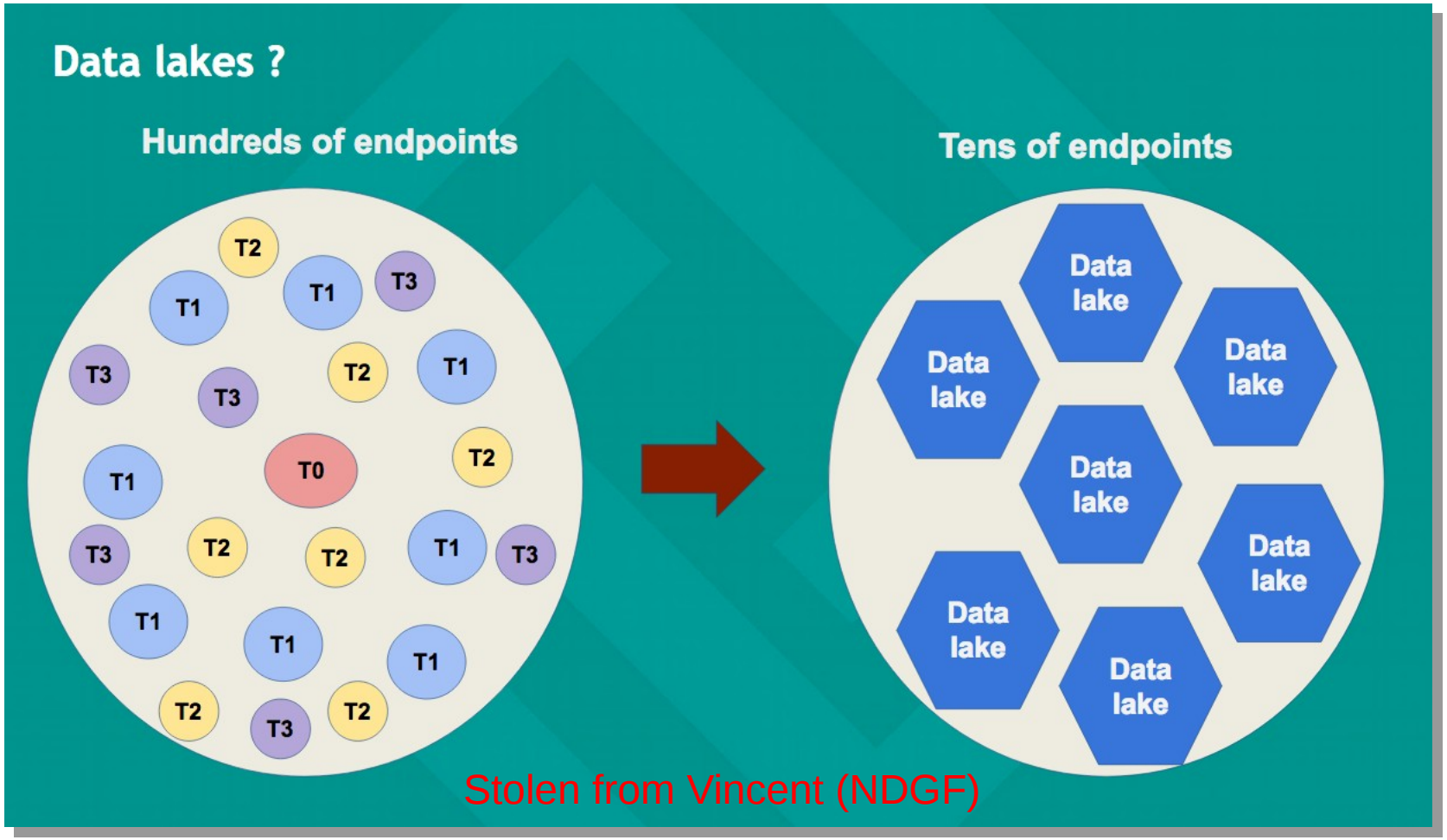
Perspektiven HEP 2018

27-28 September 2018, Wuppertal

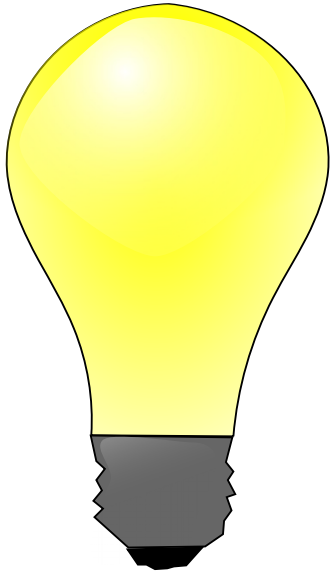
eXtreme DataCloud is co-funded by the Horizon2020 Framework Program – Grant Agreement 777367

Data Lake

Data lakes: an overview



Data Lake → DOMA



Data Lake
An idea

Data Organisation Management Access (DOMA)
A WLCG working group

DOMA: why are we doing this?

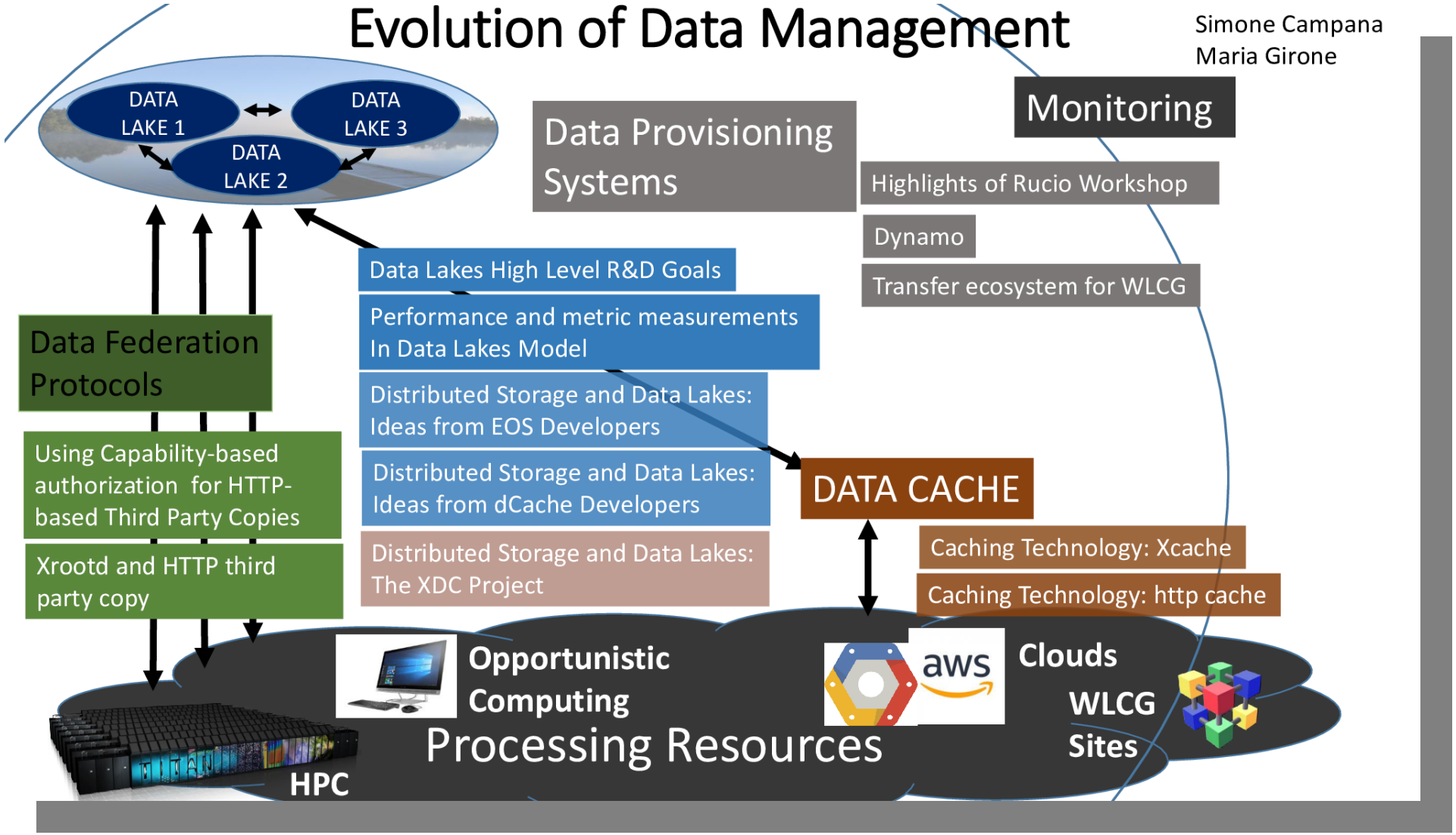
- Reduce cost for storage
  - Global (WLCG level) and local (Site level)
  - Hardware
  - Operations
- Scale out
  - Does the current model really have a scale-out problem ?
  - Which architecture would solve that ?
- Shared Infrastructure
  - E-Infrastructure
  - Research Infrastructure
- Resource Usage Optimization (Summary of above)
- Or simply : Evolution forced by ’external’ technologies or methodologies. (Best example is the ‘cloud’)

Slide from P. Fuhrmann

DOMA – a forum for discussion

- DOMA is a working group, covering storage activity.
  - Two co-chairs: **Simone Campana** and **Maria Girone**.
  - Co-exists with existing storage WGs (e.g., Archive WG)
- Meetings:
  - 2018-03-26 Napoli Joint WLCG & HSF
  - 2018-05-30 Storage Interop. meeting
  - 2018-06-04 Kick-off meeting
  - 2018-07-12 CHEP 2018 DOMA BoF
  - 2018-07-26 Video meeting
  - 2018-09-12 Video meeting
- Relatively low volume egroup `wlcg-doma` for email discussion.

2018-03-26 Napoli joint meeting

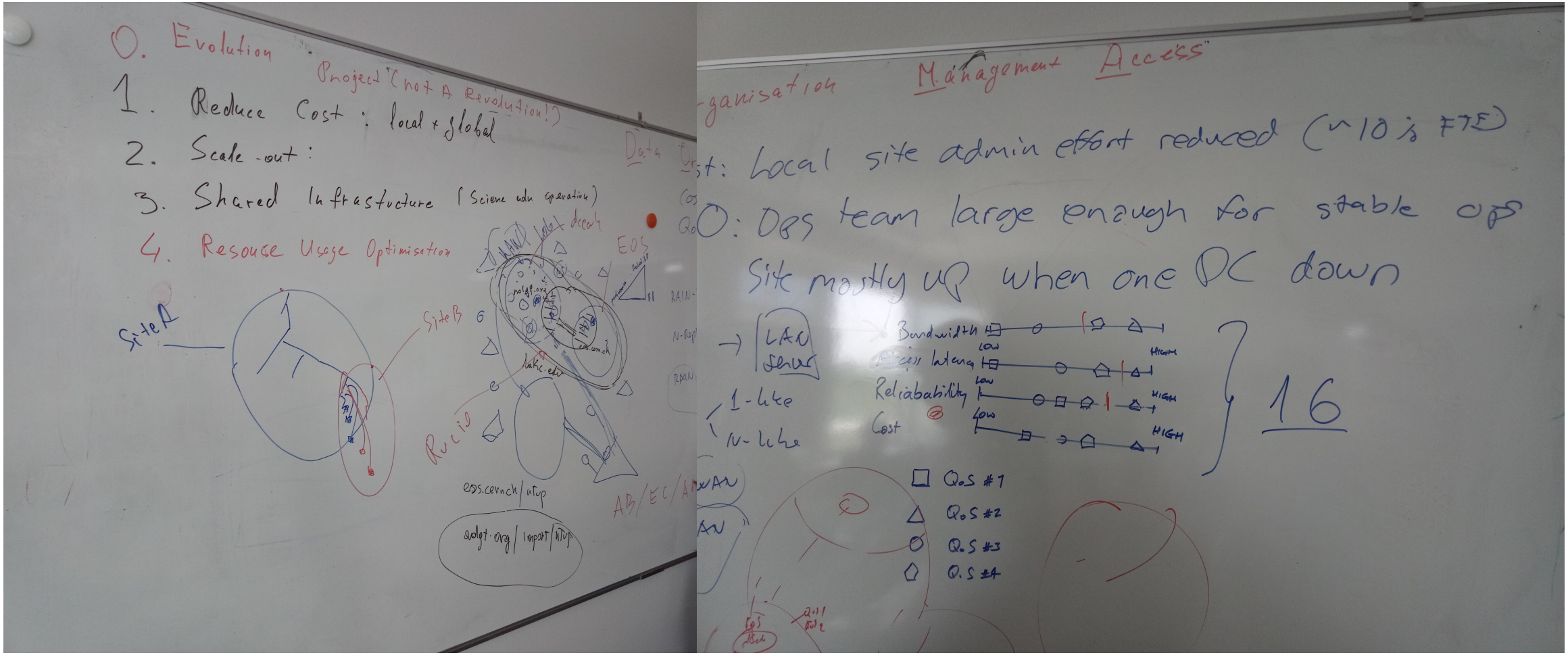


2018-05-30 Storage Interop. meeting

- Meeting at **DESY**, just prior to the official DOMA kick-off meeting.
- In attendence...
  - From the dCache team …
    - Paul, Tigran, Patrick (DESY)
    - Al, Dmitry (Fermilab)
    - Vincent (NDGF)
  - From the EOS team …
    - Andreas, Xavier (CERN)
  - The chairs of WLCG/DOMA …
    - Maria, Simone (CERN)
  - Site with considerable DataLake experience: NDGF …
    - Mattias Wadenstein (NEIC)

Content stolen from Patrick

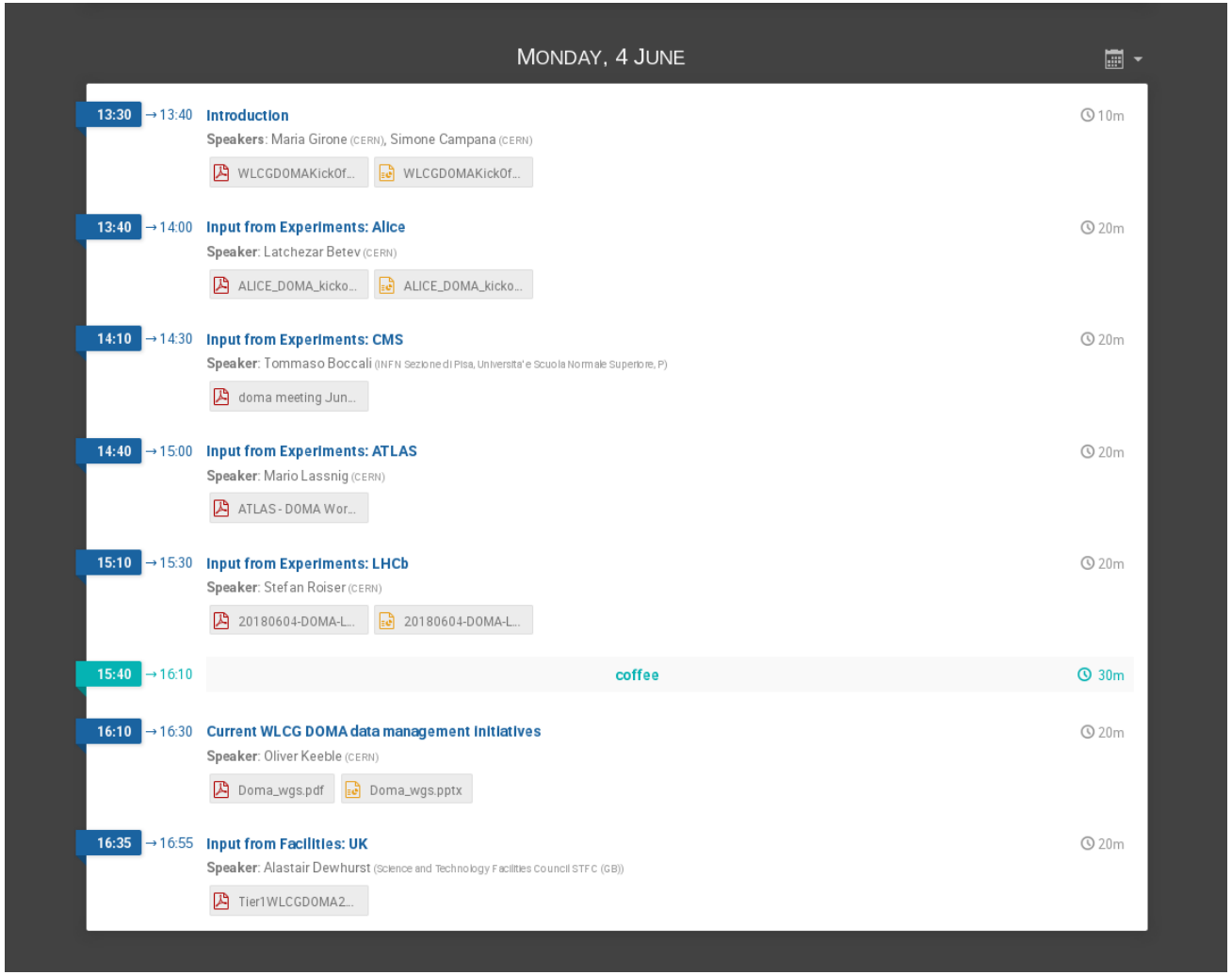
2018-05-30 Storage Interop. meeting



Slide from P. Fuhrmann

2018-05-30 Storage Interop. meeting

2018-06-04 Kick-off meeting



MONDAY, 4 JUNE

13:30 → 13:40 **Introduction** 10m
Speakers: Maria Girone (CERN), Simone Campana (CERN)
WLCGDOMAKick0f... WLCGDOMAKick0f...

13:40 → 14:00 **Input from Experiments: Alice** 20m
Speaker: Latchezar Betev (CERN)
ALICE_DOMA_kicko... ALICE_DOMA_kicko...

14:10 → 14:30 **Input from Experiments: CMS** 20m
Speaker: Tommaso Boccali (INFN Sezione di Pisa, Universita' e Scuola Normale Superiore, P)
doma meeting Jun...

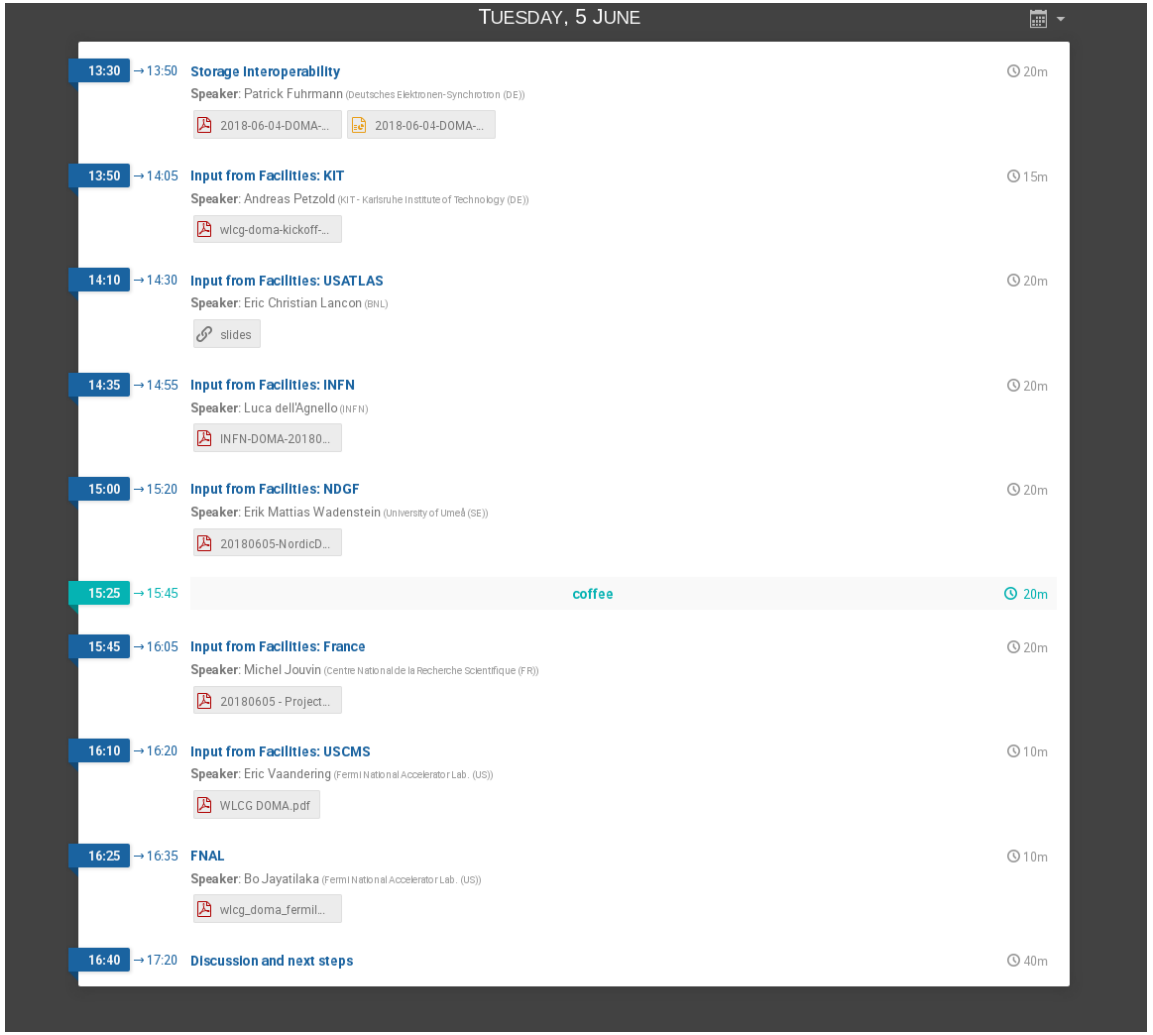
14:40 → 15:00 **Input from Experiments: ATLAS** 20m
Speaker: Mario Lassnig (CERN)
ATLAS - DOMA Wor...

15:10 → 15:30 **Input from Experiments: LHCb** 20m
Speaker: Stefan Roiser (CERN)
20180604-DOMA-L... 20180604-DOMA-L...

15:40 → 16:10 coffee 30m

16:10 → 16:30 **Current WLCG DOMA data management initiatives** 20m
Speaker: Oliver Keeble (CERN)
Doma_wgs.pdf Doma_wgs.pptx

16:35 → 16:55 **Input from Facilities: UK** 20m
Speaker: Alastair Dewhurst (Science and Technology Facilities Council STFC (GB))
Tier1WLCGDOMA2...



TUESDAY, 5 JUNE

13:30 → 13:50 **Storage Interoperability** 20m
Speaker: Patrick Fuhrmann (Deutsches Elektronen-Synchrotron (DE))
2018-06-04-DOMA-... 2018-06-04-DOMA-...

13:50 → 14:05 **Input from Facilities: KIT** 15m
Speaker: Andreas Petzold (KIT - Karlsruhe Institute of Technology (DE))
wlcg-doma-kickoff-...

14:10 → 14:30 **Input from Facilities: USATLAS** 20m
Speaker: Eric Christian Lancon (BNL)
slides

14:35 → 14:55 **Input from Facilities: INFN** 20m
Speaker: Luca dell'Agnello (INFN)
INFN-DOMA-20180...

15:00 → 15:20 **Input from Facilities: NDGF** 20m
Speaker: Erik Mattias Wadenstein (University of Umeå (SE))
20180605-NordicD...

15:25 → 15:45 coffee 20m

15:45 → 16:05 **Input from Facilities: France** 20m
Speaker: Michel Jouvin (Centre National de la Recherche Scientifique (FR))
20180605 - Project...

16:10 → 16:20 **Input from Facilities: USCMS** 10m
Speaker: Eric Vaandering (Fermi National Accelerator Lab. (US))
WLCG DOMA.pdf

16:25 → 16:35 **FNAL** 10m
Speaker: Bo Jayatilaka (Fermi National Accelerator Lab. (US))
wlcg_doma_fermil...

16:40 → 17:20 **Discussion and next steps** 40m

2018-07-12 CHEP 2018 DOMA BoF

Photos courtesy of Brian Bockelman

DOMA – a framework for activities

DOMA – a framework for activities

- **Official duties** [from Maria & Simone's DOMA kick-off intro.]
  - Keep track of developments and advances in all DOMA areas,
  - Provide a forum to discuss ideas,
  - Foster interoperability of solutions,
  - An umbrella for stakeholders, national initiatives, EU projects, existing working groups.
- **Spins off "activities"** for interested people to focus on a particular topic:
  - Three direct activities (so far): TPC, Access, QoS
  - Existing WLCG activities as "place-holder activities": AAI, AuthZ, Networks, SRR.
  - Activities operate independently from DOMA and report back.

DOMA Third-Party Copy

- **Main task**: move away from GridFTP
- Typical questions:
  - How do I initiate a transfer between two sites?
  - What protocol should be used between these sites?
  - How should these transfers be authorized?
- Three phases:

2018-12-31 Complete a survey of available replacements,

2019-06-30 All sites pledging >3 PiB provide at least one solution,

2019-12-31 All sites provide a non-GridFTP endpoint.

- Two solutions available: xrootd TPC and HTTP TPC

DOMA Content delivery & caching

- **Main task**: make “compute-only” sites work.
  - “Improve data access performance and costs by addressing latency, bandwidth management, and data structures and access patterns”
- Typical questions:
  - What is known about the application IO patterns?
  - What is known/planned regarding “data reuses”?
  - How do we best handle “compute-only” sites?
- Lots of activity in this area (14 projects collected)
- Milestones:
  - HEPix 2018: face-to-face (maybe)
  - WLCG+HSF Workshop (early 2019): presentation
  - TDR (2021+): R&D projects advanced sufficiently to actively contribute.

DOMA Quality of Service

- **Main task**: reduce cost of storage / better use of available storage
- Typical questions:
  - Can we trade performance/reliability for additional capacity?
  - How do sites describe their different storage possibilities?
  - How does an experiment adopt new technologies as they become available?
- Currently WLCG experiments have two words to describe storage: DISK and TAPE
  - Can we introduce a more precise (experiment-focused) and technology agnostic new vocabulary?
- Initial investigation in H2020 project **INDIGO-DataCloud** and continues in the follow-on project **XDC**.
- High-level working group in Research Data Alliance (RDA) **Storage Service Definition WG** on how to describe storage QoS.

~~DOMA~~ Networks

- **Main task**: adopt advances in networking.
- Typical questions:
  - How to best handle busts of network activity?
  - How to best use site-local network resources?
  - How to best for sites to provision sufficient bandwidth?
- Transfer Optimisation Projects (~7 being tracked)
  - Tracking technology: Software Defined Networks (SDN), dynamic circuits, …
  - Dynamic provisioning: using FTS to drive bandwidth provisioning.
  - REN / ISP: balancing provisioning.

(DOMA) Storage Resource Reporting

- **Main task**: accurately understand what storage is available
- Typical questions:
  - How do sites publish information about themselves?
  - What information should they publish?
  - In which format should this information be presented?
  - How up-to-date should this information be?
- Storage topology description: SRR
  - Topology: protocols, storage shares (“quotas”)
  - Accounting: non-SRM storage accounting
- Mechanisms for getting this information:
  - Protocol to provide “live” information: xrootd, GridFTP, WebDAV (SRM)
  - Write a file into the storage system.
- Integration into CRIC / AGIS

WLCG AAI group

- **Main task**: move users away from X.509
- Typical questions:
  - How do we know who are our users (really)?
  - How problems are reported (in both directions)?
- Benefits:
  - Easier for users to get started
  - Avoid bad X.509-specific toolage
  - Better web-driven experience.

WLCG AuthZ group

- **Main task**: investigate centralised authorization
- Typical questions:
  - How to encapsulate what a user/agent is allowed to do?
  - How to handle traceability?
- Benefits:
  - Caches can honour authz decisions
  - Delegation could allow caches to acquire data
  - Separates services from authentication: switching away from X.509 becomes simpler.
- Challanges:
  - Traceability – how much does the storage service need to know

DOMA summary

- DOMA is a mechanism for building the next-generation storage infrastructure for WLCG
  - This is a multi-facited approach, touching all aspects of storage.
  - Split into distinct “activities”, where interested people can come up with common solutions.
- The expected limited (“flat”) budget is the main motivation for DOMA activity.
- Timescale is to have projects deliver prototypes for Run 3, to have solutions ready for Run 4/HighLumi.
- dCache is strongly involved in the DOMA process, and (through the XDC project) is leading the QoS activity.

Thanks for listening!

As always, there's certain to be the odd surprise along the way ...