

ARC FOR WLCG PILOTS

Gianfranco Sciacca

AEC - Laboratory for High Energy Physics, University of Bern, Switzerland

HTCondor & ARC workshop - ISGC 2019, Taipei - 1 Apr 2019



A gateway from Grid to batch systems

- ▶ In use since 2004
- ▶ Widely deployed on European Grid sites, and beyond
- ▶ Modular, well-suited to heterogeneous environments:
 - ➔ from standard grid sites to restrictive HPCs
- ▶ Can pass complex job requirement and extra env variables to the LRMS
- ▶ **SDK for client applications (C++ and python)**
- ▶ CLI for job and data management (*arcsub, arcstat, arccp, arcls, ...*)
- ▶ Integrated information system and accounting tool

▶ Pilot gateway

- Suited to the standard pilot model used by the WLCG experiments

▶ NorduGrid

- Suited to heterogeneous, distributed and HPC sites (*e.g. no middleware on the wn's*)

▶ Truepilot

- Offers advanced job-to-site resource matching

▶ Pilot gateway

- Suited to the standard pilot model used by the WLCG experiments

▶ NorduGrid

- Suited to heterogeneous, distributed and HPC sites (*e.g. no middleware on the wn's*)

▶ Truepilot

- Offers advanced job-to-site resource matching

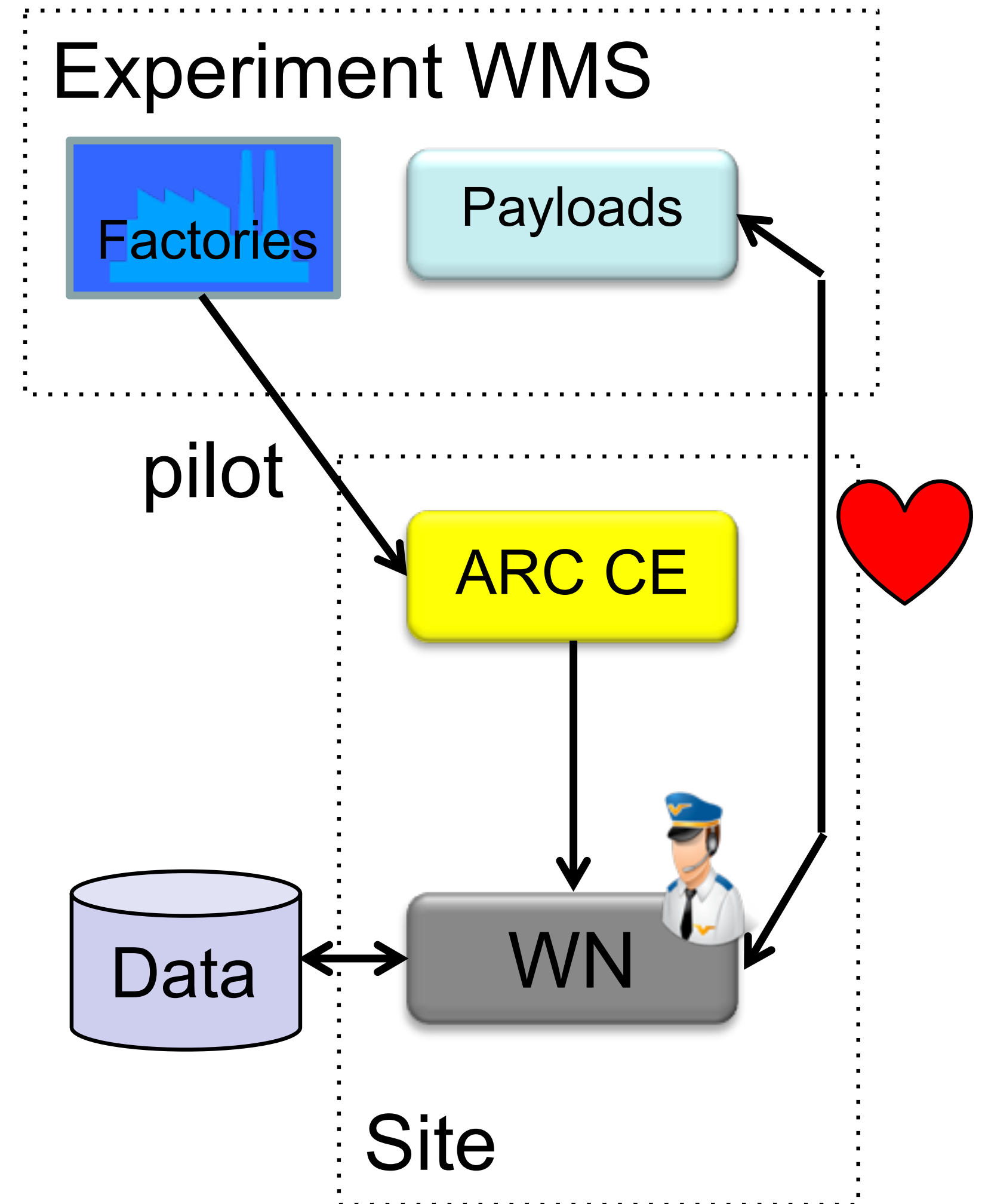
➔ **The ARC operational mode is transparent to the site:**

no impact on ARC installation and operation, or site architecture

➔ **The experiment frameworks have since long integrated ARC with their WMSs**

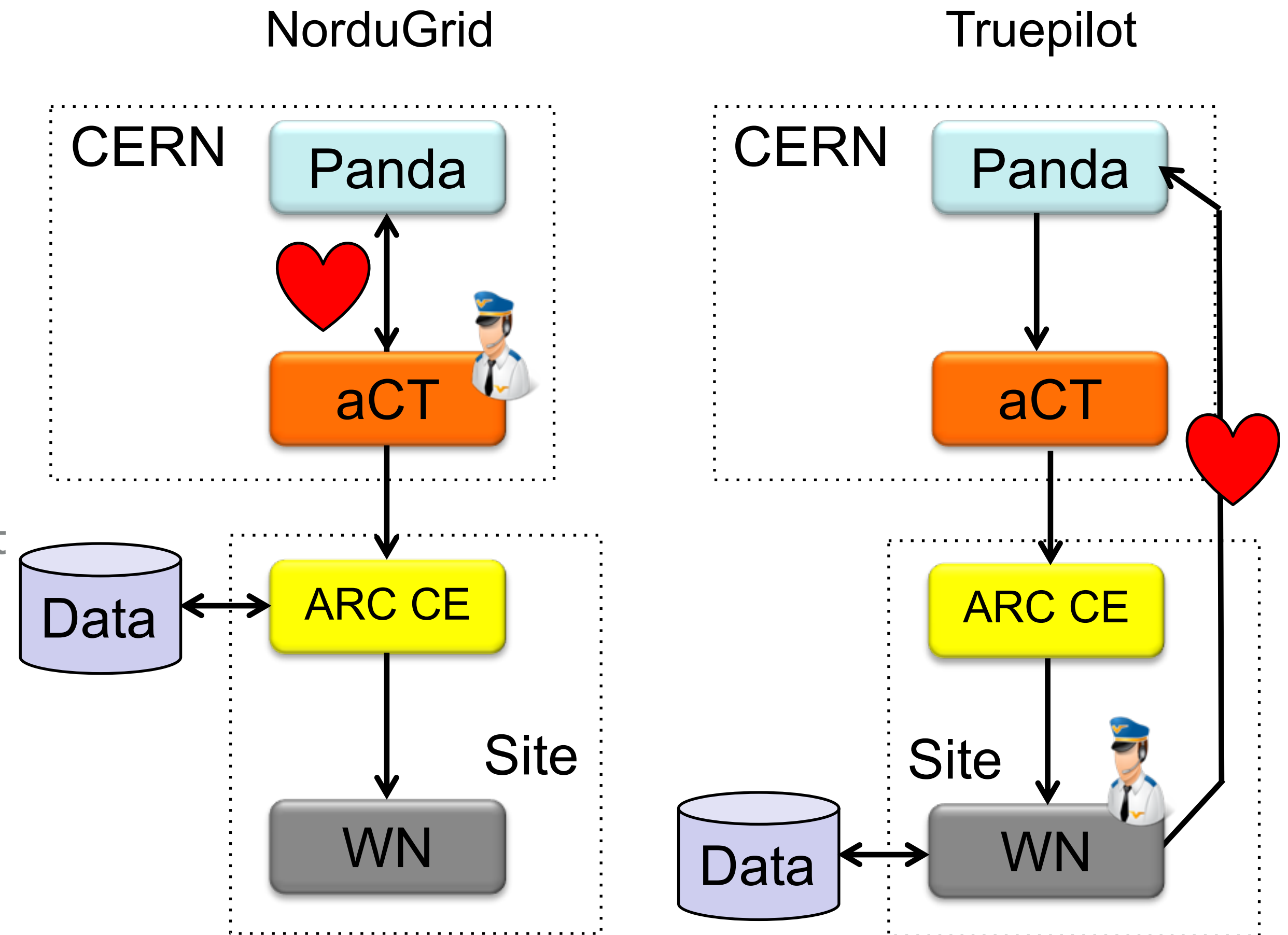
Used by ALICE, ATLAS, CMS, LHCb

- Pilot factory submits wrappers / agents to the sites
- Pilot starts on worker node and pulls real payload(s)
- Data staged-in and out of WN
- Pilot reports back regularly to the WMS



Used by ATLAS

- aCT/harvester pulls the payload from Panda and pushes to ARC CE with the correct requirements
- ▶ **NorduGrid**
 - ARC CE performs the staging of input/output
 - aCT/harvester handles the communication with Panda, acting as pilot
- ▶ **Truepilot**
 - Pilot on WN uses pre-placed payload
 - Pilot takes care of data staging and Panda communication just like pilot pull



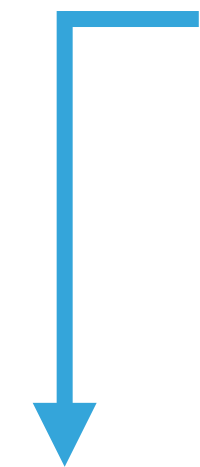
- ▶ **ALICE:** 16 ARC CEs at 10 sites in 8 countries
- ▶ **ATLAS:** 133 ARC CEs at 46 sites in 20 countries
(about half of them in pilot gateway mode)
- ▶ **CMS:** 38 ARC CEs at 15 sites in 5 countries
- ▶ **LHCb:** 54 ARC CEs at 13 sites in 3 countries

Swiss ATLAS sites have been using ARC CEs since the early days

- ▶ 2 ARC CEs at the CSCS Tier 2
- ▶ 2 ARC CEs at the Bern Tier 2
- ▶ 1 ARC CE at the Geneva Tier 3

Swiss ATLAS sites have been using ARC CEs since the early days

- ▶ 2 ARC CEs at the CSCS Tier 2
- ▶ 2 ARC CEs at the Bern Tier 2
- ▶ 1 ARC CE at the Geneva Tier 3



CSCS-LCG2: WLCG Tier 2 supporting ATLAS, CMS, LHCb

Running *ARC and CREAM* CEs since the beginning of WLCG operations

- ▶ ARC for ATLAS
- ▶ CREAM for ATLAS, CMS, LHCb (*CMS and LHCb not ready for ARC until ~late 2014*)

Motivation for migration to ARC only

- ▶ Hefty hardware requirements for CREAM
- ▶ Modest VM for ARC services
- ▶ ARC ease of setup, configuration and maintenance vs. CREAM (yaim)
- ▶ CREAM operation not too smooth

Decommissioned CREAM CEs in 2015

- ▶ Subject to readiness of the CMS Glide-in WMS and LHCb DIRAC
- ▶ Transparent to experiments
- ▶ Allowed a transparent transition from PBS to SLURM
- ▶ Job accounting to EGI out of the box using the existing APEL client initially
- ▶ Moved later to ARC native Jura publisher

Admins overall very happy with ARC

- ▶ Lightweight operation, maintenance, upgrades
- ▶ Some modifications to the information system needed to cater for LHCb needs
- ▶ Deployed in 2014 to access one of CSCS HPC systems remotely
- ▶ Catalised the HPC integration for ATLAS, later on for CMS and LHCb
- ▶ Also operated in this mode for fronting cloud resources
- ▶ As of today operating up to 6 ARC CEs and 2 ARC data stagers, in front of grid and HPC resources
- ▶ In Truepilot and NorduGrid mode (ATLAS), in WLCG pilot mode (CMS, LHCb)
- ▶ Fully transparent to all experiments supported

▶ Some sites need containers

• Non compliant OS on WNs

➔ Add suitable job option arguments:

- in the ARC LRMS backend script →
(e.g. *shifter* at CSCS, *singularity* at HPC2N)
- in a RTE script →
(e.g. *singularity* at SiGNET)

• External requirement

➔ Singularity for CMS

- singularity called by the job payload
nothing to do at the site

In `arc.conf`:

```
run_thru_shifter           = true
shifter_vo_whitelist       = 'cms' # comma separated values
shifter_image              = 'cscs/wlwg_wn:20180710'
shifter_module             = 'shifter-ng'
shifter_job_store_path     = '/scratch/lhc/arc_shifter_job_store'
shifter_job_store_keep     = false
shifter_job_store_keep_path = '/var/spool/nordugrid/shifter_slurm_mods'
shifter_options            = '--mount=type=bind,source=/scratch/lhc/apps/
cle60.up04/bin,destination=/usr/local/bin/ --mount=type=bind,source=/
scratch/lhc,destination=/scratch/lhc --mount=type=bind,source=$
{HOME},destination=${HOME}'
features_base_path        = '/scratch/lhc/features_per_job' #
location where machinefeatures/jobfeatures will be generated
```

```
pikolit ~ # cat /ceph/grid/runtime6/ENV/CONTAINER

# description: executes the job inside singularity container
# param:SINGULARITY_IMAGE:string:TO BE SET:singularity image or tree
# param:SINGULARITY_OPTIONS:string: :singularity options

SINGULARITY_IMAGE="${SINGULARITY_IMAGE:-}"
SINGULARITY_OPTIONS="${SINGULARITY_OPTIONS:-}"

IMAGE="${2:-SINGULARITY_IMAGE}"

if [ "x$1" = "x0" ]; then
    joboption_args="/usr/bin/singularity exec $SINGULARITY_OPTIONS
$IMAGE $joboption_args"
```

ALICE

- ▶ ARC modules for AliEn were implemented using the features of ARC in 2014 and successfully worked until 2017. It was not too difficult to implement these modules
- ▶ This makes the ARC CE deployment at sites fully transparent
- ▶ Infosys content need fixing at the sites

ATLAS

- ▶ The use of the ARC CE is widespread in ATLAS, both for the classic pilot pull mode and the advanced push modes
- ▶ It is a versatile product that allows the VO to make transparent use of diversified resources: from classic grid sites to HPCs, voluntary, opportunistic and distributed computing resources

CMS

- ▶ Rely on ARC CEs quite a bit in the CMS distributed computing grid and are happy with their overall performance in conjunction with the Glide-in WMS
- ▶ There is consensus that the ARC setup is very easy

LHCb

- ▶ LHCb/DIRAC currently uses a fairly simple setup with resources expected to be approximately static at a given site. These requirements are met by ARC
- ▶ Given DIRACs plugin nature and ARC modular structure, it is very easy / seamless to deploy ARC CEs for any experiment using DIRAC
- ▶ LHCb and DIRAC based VOs benefit directly from advanced features like job-to-site resource matching, specially at shared sites

- ▶ **The ARC CE has been operated with the WLCG pilot model of ALICE, ATLAS, CMS and LHCb for years**
 - ▶ Also operated in parallel with the CREAM CE at sites like CSCS, and in a mix of modes
- ▶ **For the experiments, it has been proved to be a drop-in replacement for CREAM**
 - ▶ It integrates seamlessly with the experiment WMSs
 - ▶ The SDK allows to support the experiment specific frameworks evolving needs
- ▶ **Site admins find it easy to deploy and operate**
 - ▶ Minor adaptations at sites might be needed, mostly depending on the local job accounting scheme for EGI, or for the infosys
 - ▶ Helps managing efficiently resources at sites, specially useful at shared sites