



# NorduGrid 2018 conference summary

---

GDB JULY 2018 / OXANA SMIRNOVA

# NorduGrid conferences – an overview

---

Corridors of LRZ



Annual events, started as conferences in 2007

- Traditionally free of charge
- Various locations around Europe, following NorduGrid members and users
- Was held in Garching by Munich this year
- The next is planned in Lund in June 2019

Consist of a plenary part and a technical ARC workshop

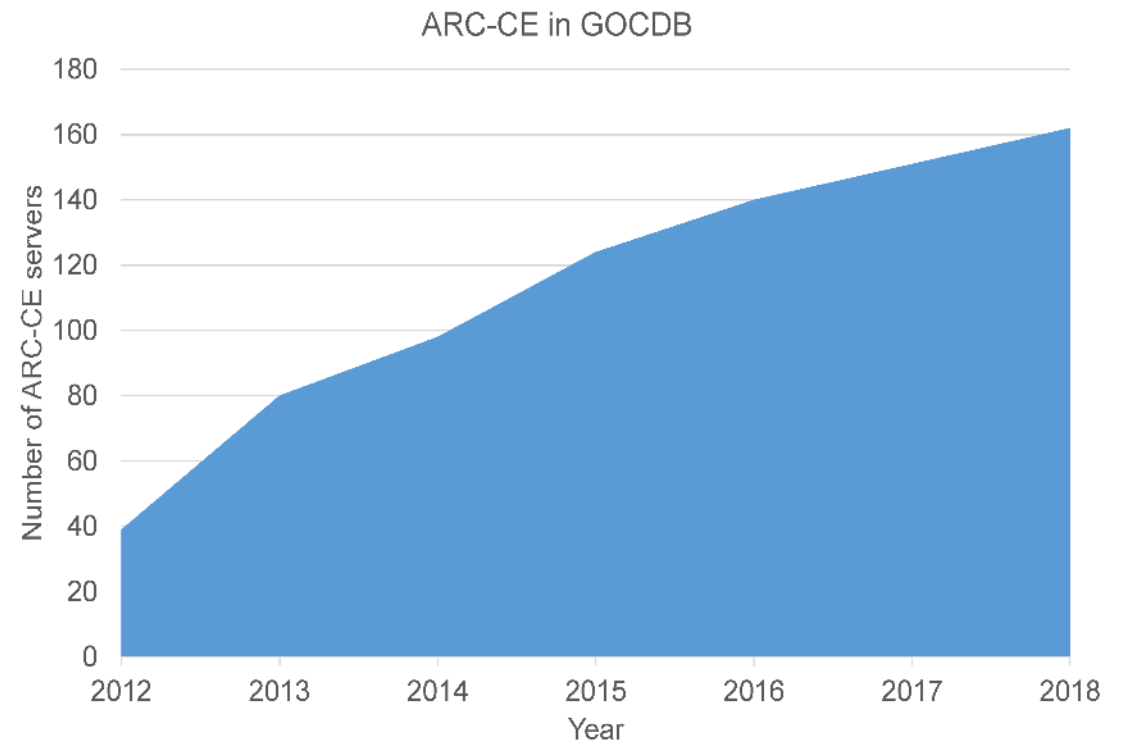
- Plenary talks: overview reports, success stories, technology trends
- Technical workshop: informal round-table discussions with flexible agenda
- A football game is always on the agenda

# ARC status

---

ARC5 is the current stable release

- First released in 2015, latest update in December 2017
- Is distributed via UMD, standard Linux repositories and NorduGrid's own repo
- Components:
  - **ARC-CE** including cache, data handling, accounting etc
  - CLI, infosystem
- Deployments steadily increase
- Main difficulty: challenging configuration and parameter tuning



# Science cases

LHC experiments are the primary users, and LHC use cases drive ARC development

- Current workloads are well supported, but HL-LHC will be a serious challenge, *we all may need to do things very differently*

Non-WLCG science cases presented:

- Belle II
- EISCAT\_3D
- Astroparticle experiments (GERDA, LEGEND, CRESST, MAGIC, CTA)

Commonalities: previously unseen data volumes vs limited availability of suitable hardware and human resources

- Differ from LHC ones by smaller s/w stacks and fewer hands, need turn-key solutions

Possible direction: non-traditional Grid resources, especially HPC facilities, since they are likely to see lots of investment in future

- ARC itself was developed as a front-end for HPC systems comprising the Nordic Tier1



Belle II in 3D

# ARC6 – next major release

Main change: new configuration, **manual upgrade required**

- No interface changes

Teaser: new features

- REST interface
- Internal submission mode: off-line submission, no external interface, tailored for HPC
- Event-driven job submission: significant speed-up
- DDS: Data Delivery Service: incl. file copying between caches, can be a standalone service and a lightweight SE
- ARCHERY: DNS-record-based service discovery
- `arcctl` – a new CE management script for sysadmins (handles configuration, RTEs, jobs)
- Concept of defaults for turn-key deployment
- And, ARC is now in GitLab (via NeIC CodeRefinery)

Coding mostly finished, tests ongoing

- A **production** site in Oslo runs ARC6, successful ATLAS jobs
- Another configuration re-organisation undertaken after the conference
- Release expected in Fall 2018

EVENTSERVICE (1)	1 (3918)
ATLASRELEASE (2)	Atlas-21.0.15 (3918) Atlas-20.7.5 (1)
PRODUSERNAME (2)	dsouth (3918) gangarbt (1)
JOBSTATUS (8)	closed (320) activated (23) merging (753) running (208) failed (1)
COMPUTINGSITE (1)	UIO_MCORE_LOPRI (3919)
JEDITASKID (1)	14525796 (3918)

*Bigpanda screenshot of the test site*

# Ongoing ARC6 tests overview

---

- **Job management:** GridFTP, EMI-ES, **REST and internal submission** interfaces
- **Authorisation:** gridmapfile, gridmapless (VOMS attributes only), authgroups (incl. on queue level) etc
- **Data management:** cache, data staging (in, out – gridftp, xroot, https)
- **Batch back-ends:** SLURM (**new** Python back-end), HTCondor, fork – all with the new configuration
- **RTEs:** ENV/PROXY (**new RTE handling**)
  - To be tested: Singularity RTE, container support in HTCondor
- **Infosys:** nordugrid-arc, GLUE2; VO info
- **Startup scripts:** a-rex, gridftp, infosys, **data delivery service** (new script names)
- **Accounting:** JURA with APEL and SGAS; manual re-publishing
- Server logfiles
- Configuration validator
- **Clients:** compatibility with ARC5 client, CondorG client, DIRAC client; consistent interpretation of job description attributes, esp. units of memory, disk, time etc



# Other things discussed

---

CondorG can only submit via GridFTP interface, does not support EMI-ES standard – perhaps REST interface will be more welcomed

GitLab is not open for everybody – considering mirroring ARC code repo on Github

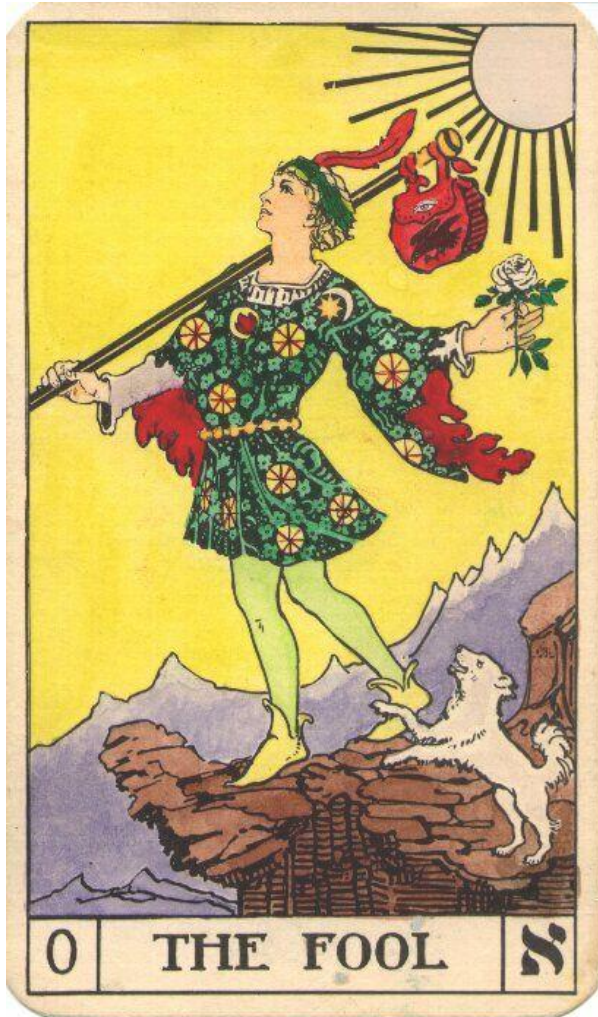
Training is needed – co-locate a **training event with ISGC 2019 in Taipei**

- Another event is being considered in November 2018, likely co-located with the NDGF all-hands in Umeå



Ume river

# What the future may hold



Data lakes – ARC is known to be a fitting CE for this environment, as e.g. in the Nordic Tier1; Rucio is a well-integrated dedicated tool

Moving away from GridFTP for disk-only sites is a reality, brings us closer to an X509-free world

Federated identities: solutions for multiple federations and trust models are being developed

Lightweight sites: ARC services, esp. cache, can be a generic solution

Heavyweight sites: ARC-CE + HTCondor + virtualisation/containerisation can be attractive for HPC sites (including PRACE)

Neural networks are back, on a more suitable hardware – machine learning might challenge traditional software



# Conclusion: a wishlist

Simplicity, stability, robustness, usability, magic and wizardry – ARC-CE should just work, no matter the sysadmin

This magic simplicity should hide the heterogeneity of resources, architectures, data sources, applications and workflows

- ARC6 should be a step in the right direction

Evolve from a classic Grid service to a thing more agile – while still supporting old architectures and old applications code

