



**NORDUGRID**

*Grid Solution For Wide Area  
Computing and Data Handling*

# **ARC7 features and release plan**

*July 13, 2022 GDB*

*Balázs Kónya, Lund University  
Nordugrid Technical Coordinator*

# Previous ARC major release: ARC6

- ARC 6, May 2019: **lots of backward incompatibility**

- **ARC.CONF**
- Event-driven internal job loop
- **Re-engineered RTE framework**
- ARCCTL utility
- **Accounting system changes**
- **ARCHERY**

- *Postponed to ARC7*

1. *Completion of the tech preview solutions: REST Interface, Internal Interface, Python LRMS, Candypond service*
- ~~2. Any major work on the client-side~~
3. *Embracing non-x509 security trends*
4. *Controldir restructuring*
5. *Phasing out gridftp and glue-1*
- ~~6. Better error messages~~



# Since then on the ARC6 code line... several releases came out

- Switching from ARC5 to ARC6
  - was painfull but not that bad
- ARC5 reached its end-of-life in 2020
  - although we still provide some „minimal operation site support”
- „Only” ... regular bug fix and enhancement releases on the ARC6 code line:
  - 6.1 June 2019, . . . 6.15.1 March 2021
  - 6.4: introduction of a new accounting subsystem
  - *6.6: OIDC tech preview release*
  - **6.8: removal of the pre 6.4 accounting subsystem**
  - 6.8.1 express release (fix for a too strict validator option)
  - *6.9.0: REST tech preview re-implemented*
  - 6.10-DO-NOT-USE RELEASE
  - 6.15.1 python3 related fix

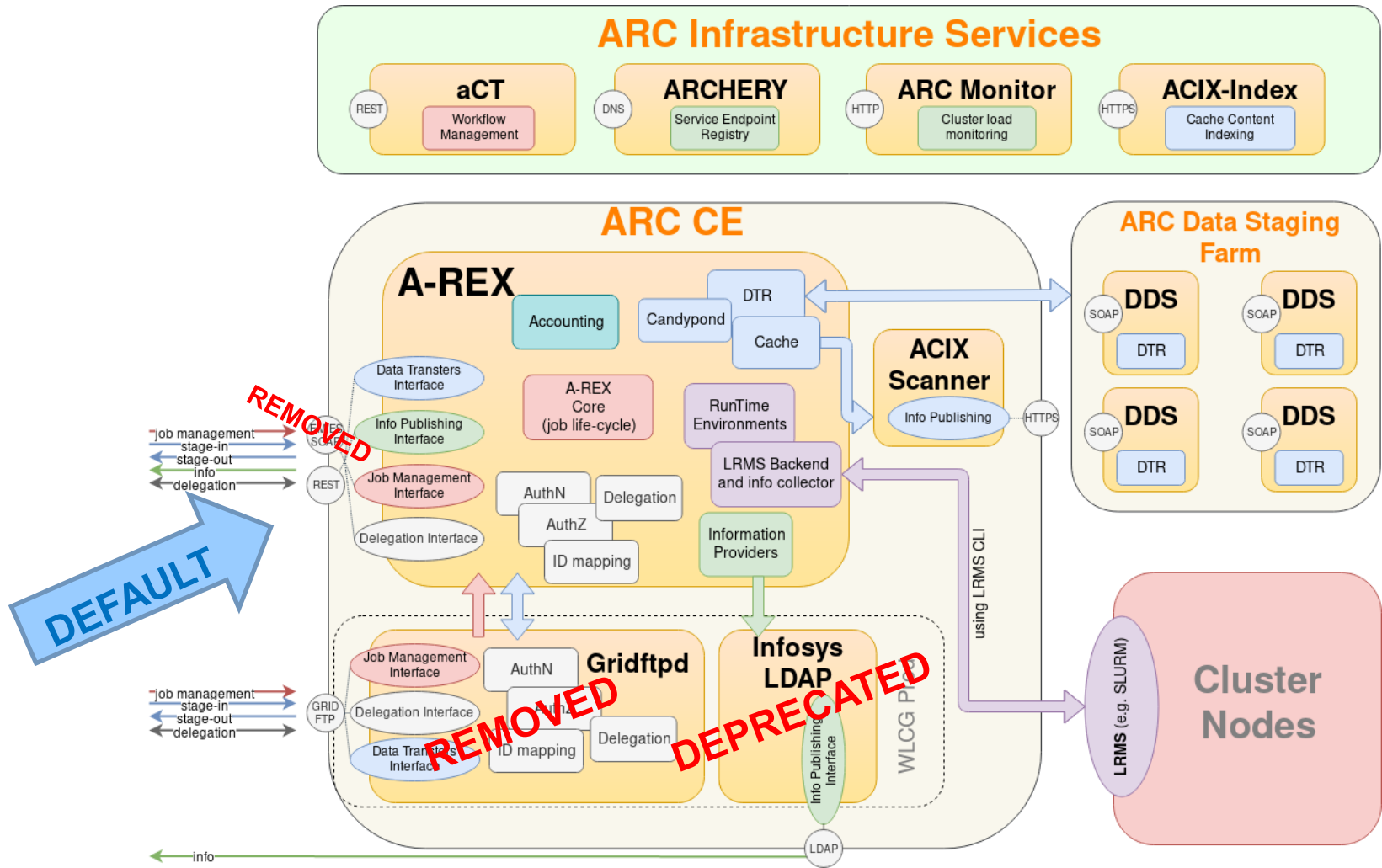
# Development targeted for ARC7

- ARC 7 as the next **major** ARC Release
  - Backward incompatibilities are allowed
  - One can remove not-used components, obsolete interfaces
  
- 1. „Final”, production version for the REST interface
- 2. Production OIDC support as to satisfy the WLCG experiments... (?)
- 3. **Major internal ARC CE restructuring (controldir, scalability)**
- 4. Lots of discussion to drop-not to drop gridftp jobsubmission and/or LDAP infosys from ARC7

# ARC7: REST CE interface

- REST specification
  - [www.nordugrid.org/arc/arc6/tech/rest/rest.html](http://www.nordugrid.org/arc/arc6/tech/rest/rest.html)
  - API endpoint: <https://arc-ce.example.org:443/arex/rest/1.0/<functionality>>
  - Supported functionalities: **info, delegation, jobs**
    - **post** <base>/jobs?action={new,info,status,kill,clean,restart}
    - **get** <base>/jobs/<job id>/session/<file>, <base>/jobs/<job id>/diagnose
  - V1.0 spec is complete and „closed”, already since ARC 6.9 release
    - All defined functionality got implemented as part of ARC CE (server-side)
    - The code is released and enabled as a technology preview within 6.9
    - The ARC client SDK and the „arcsb” CLI supports REST
  - REST is used in production
    - ATLAS sites enabling REST on ARC-CE: <http://novastore.farm.particle.cz/ce/arc-rest/>
- REST in ARC7:
  - Intensive testing (stress tests): ACT (25–30 Hz bulk submission), DIRAC, Condor teams
  - V1.1 specification extension: queue parameter, better support for bulk jobs, and token-delegation may need additional changes
  - ARC7 will support both v1.0 and 1.1
  - *Switching from gridftp/emies to REST as the default and only ARC CE job management interface!*

# ARC CE interfaces in ARC7



# Tokens in ARC7 (support for OIDC)

- Not a new feature in ARC7
  - technology preview of the handling of OIDC tokens, ARC 6.6, April 2020
- Need a third-party tool to obtain a token and store it in „BEARER\_TOKEN” variable
- Tokens can be used as a replacement of x509 for authenticating against ARC-CE
  - But: data staging still requires x509 proxies
- Tokens-based server-side authorization and user identity mapping on the ARC-CE
  - Tokens are processed by ARC-CE by enabling [authtokens] config block
  - „Token information” can be used in server-side authorization and identity mappings: **subject, issuer, audience, scope, group**
- ARC-CE and HTCondor token hackathon, Amsterdam
  - Any necessary code change will be included to ARC7
- TESTING:
  - Done: tokens conforming to the WLCG profile, Atlas and IAM (Petr Vokac)
  - Soon: EGI profile with EGI Check-in as a token issuer

# Controldir restructuring in ARC7

- A-REX stores information about jobs in files in the control directory
- *Previously all files were kept in one single flat directory: scalability issuse with 100k jobs*
- ARC7 introduces a hierarchy of subdirectories:  
**jobs/SUBID/SUBID/SUBID/SUBID**
  - SUBID are arbitrary ASCII string which if put together form ID of the job.
  - each SUBID consists of 3 lowercase+ hex symbols. But that may change.
- Internal change, should not have any effect on 3rd party software
  - All ARC internal components were ported to use the new structure (e.g. LRMS scripts, accounting, etc..)
  - **BUT: some site admins may have developed scripts against controldir files.... Those may be broken with ARC7**



# REMOVED or OBSOLETE in ARC7

- gridftp interface on ARC-CE: both for job submission and as a storage element, data staging
  - ARC7 will be „globus free”
- EMIES as a job interface (older SOAP WS interface)
- LDAP infosys is kept but to be marked as DEPRECATED
- ACIX component
- LRMS-ssh backend, Python LRMS scripts
- Support for python2
- Use C++11
- Support for every interface (both job and infosys) in the ARC7 client except the REST

# REMOVED or OBSOLETE in ARC7

Smaller features, options to be removed from ARC7:

- Glue1 schema
- Provisioning of the „Fake site-bdii” instance
- Support for Argus
- Nordugridmap-like gridmapfile generation
- Resubmit option in xrls should be retired (related to the accounting and diag file discussion)

# ARC7 compatibility statements

- **ARC7 packages and distribution** will be very similar to ARC6, no major changes expected
  - ARC7 packages in all distributions will be named arc7 in their package names
- **ARC7 configuration:** no major changes, ARC6 config file will be compatible with ARC7 CE
  - modulo discontinued blocks/functionality
- **Controldir restructuring** is the major change in ARC7
  - it is a transparent internal change within ARC7
- **ARC7 will deliver a REST-centric ARC CE** with production-level token support.
- **Server-client compatibility, based on REST** (ARC 6 and ARC 7):
  - ARC6 clients will work with ARC7 CE-s (only via REST)
  - ARC7 clients will work only with REST-based ARC6 CE-s.
  - ARC7 server will be compatible with ARC7 and ARC 6 (REST) client, i.e. accept both REST 1.0 and 1.1
  - **Neither ARC7 server nor the client will support gridftp and LDAP**

# ARC7 Release timeline

- Documentation branching for ARC7: **done**
- Technical description of the restructured controldir is available: **pending**
- Propose a working process for the branch(es) for the ARC7 release work: **done**
- Controldir new structure is available /up-to-date/merged into the ARC7 branch: **done**
- Bug review day, decide what to fix for ARC7: **done**
- Testing campaign for both controldir structure and REST interface: **started**
- Very first alpha for ARC7 is announced for developers: **after summer**
- REST 1.1 defined and implemented: **ongoing**
- All obsoleted components, code, library removed from ARC7 branch: **after the summer**
- Upgrade procedure in place: **early September**
  - No draining needed, the conversion happens by ARC automatically
- First release candidate **late September**
- ARC7 is announced **5th November 2022**

# Not for ARC7: ARC8 candidate content

- LRMS Backend rationalizing, cleanup
- Json as a supported job description language
- REST-based webmonitor
- Some replacement for ganglia
- Containerized version of the ARC-CE REST only server
- ARC storage element (drop-in replacement for gridftp-storage server)
- Intelligent Garbage collector for controldir leftovers
- LRMS error propagation
- possibility to keep the \*.errors file of jobs for sysadmin
- Find some alternative to swig
- Integrated own minimalistic internal task scheduler, an “ARC batch system”,  
[lrms/slurm] [lrms/condor] etc configuration subblocks
- Generalise the job concept so that ARC e.g. could handle data transfer jobs
- Complete, powerful rest-based new generation client replacing the “old cli”
- Support for “modern” openpbs/pro
- A mechanism to blacklist DNs dynamically

# Support statements for ARC6

- **No new feature development** is planned or going on for ARC6 and no bug-fixing development will happen on ARC6 code base in the future except for security issues.
- **Security fixes for ARC6** will be provided till the end of <release-date+1 yr>.
- Production sites already running ARC 6 will be able to get deployment and configuration troubleshooting help via GGUS till <release-date + 2 yrs>. This we call "**operational site support**".
- ARC 6 will be available in EPEL 7– 9, in the future EPEL10, only ARC7 will be available. ARC 7 will be available in EPEL 7–9 with a different package name as arc7
  - EPEL 7: arc, arc6, arc7
  - EPEL 8: arc (arc 6), arc7
  - EPEL 9: arc (arc 6), arc7 to-do rename to arc6
  - EPEL 10: arc7

# ARC FACTSHEET

Latest version: ARC 6.15.1  
ARC7 pre-view: early autumn...

- Code:
  - <https://source.coderefinery.org/nordugrid/arc>
- Installable packages:
  - Global Linux repositories (CentOS, Debian, EPEL)
  - Upstream: <http://download.nordugrid.org/repos.html>
- Documentation:
  - ARC6: complete documentation online at <http://www.nordugrid.org/arc/arc6>
    - Including tech specifications and internals such as the OIDC tokens and REST interface
  - ARC7: will have its own documentation area, work started here: <http://www.nordugrid.org/arc/arc7>
- Support:
  - Dedicated skype support channel: <https://join.skype.com/dyf3A6Uutjy2>
  - Email lists: [nordugrid-discuss@nordugrid.org](mailto:nordugrid-discuss@nordugrid.org) – generic, CERN e-group [wlcg-arc-ce-discuss@cern.ch](mailto:wlcg-arc-ce-discuss@cern.ch) – WLCG-specific
  - Own bugzilla and EGI GGUS: <https://bugzilla.nordugrid.org>; “ARC” unit in GGUS

# Summary

- ARC 7: is going to be a **SOFT major release**
  - Almost transparent for sysadmins!
- New developments to be released with ARC7:
  - either internal restructuring, improvements
  - or already available on the ARC6-line as tech preview
- ARC7 backward incompatibility is mostly due to
  - **OBSOLETED** interfaces, components
- ARC-CE from ARC7 will be a very clean single interface CE
  - **REST-only**
- Lots of cleanup, consolidation, internal improvements