Changes from ARC5 to ARC6

Florido Paganelli, Lund University

ISGC 2019, Academia Sinica in Taipei, Taiwan 31 March ~ 5 April 2019

Why?

- Needs for major release:

    - internal **scalability** issues

    - regain **manageability**
      (arc5 configuration had grown
      out of hand)

INSURANCE CLAIM WILL BE INTERESTNG

shareitsfunny.com

Why?

- Needs for major release:
  - interface **consolidation**
  - allow retirement of **unused components**

Goals

- Simplification
  - For users:
    deployment, reference documents
  - For developers:
    - code cleanup
    - "How to contribute" guide is coming

- Cloud-readiness:
  - Automatic configuration
  - Unattended deployment
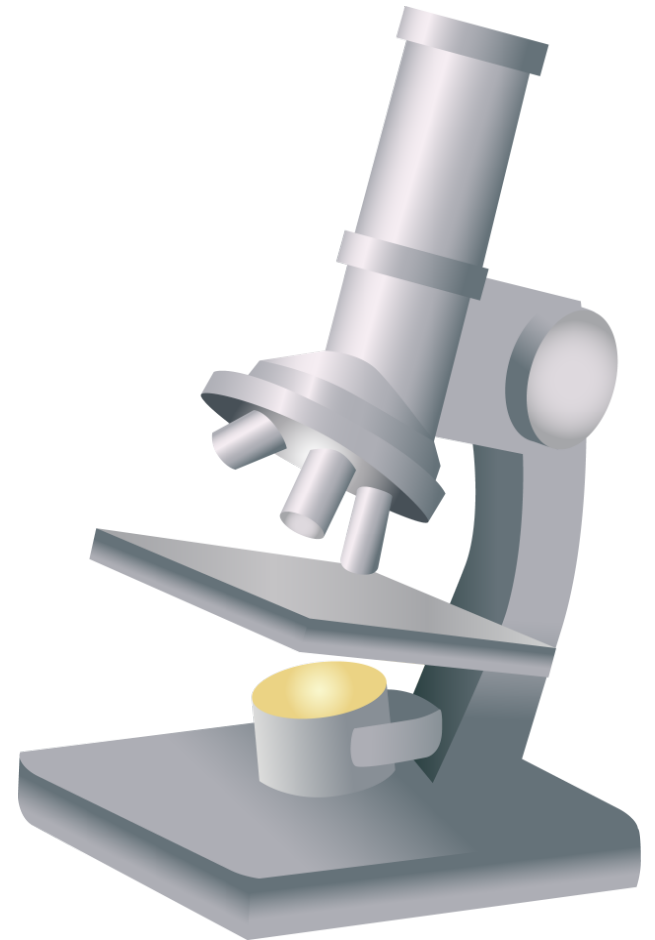
- Improved control
  - Swiss-knife management tools

Focus on:

- Better Configuration

- Stability

- Performance

- Code Cleanup

- Consolidation

- New Features

- Experimental Features

Highlights

- Totally revised and rewritten **configuration**

- Old, unused code removed,**reducing** points of failure

  - Some (never used) submission and info interfaces **removed**

  - API changed but Client-Server backward compatibility preserved

  - **WallTime** unit of measure changed:
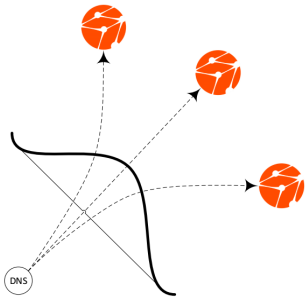   `<Wallclock> * <num. of cores>`

- New development cycle based on **Gitlab**, hoping it will help us improve external contributions

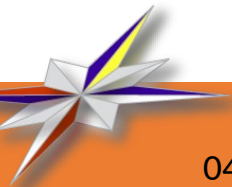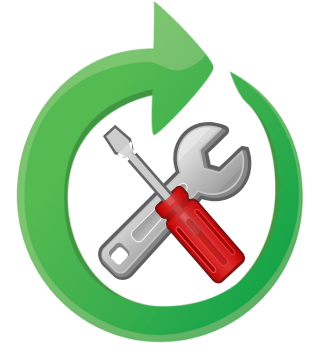- Gitlab **CI** now provides build, deploy and test platform

Highlights cont'd

- **Quick installation** that can be done by configuration managers (ansible etc.)

- Administration tools like **arcctl**

- New simplified index **ARCHERY**

- Took almost 3 years of development, maybe also influenced by lower interest of funding agencies into GRID development.

Better Configuration

- Total **revision** of the configuration **options**

- Rewritten, streamlined *Cloud Ready* configuration: make it easier for tools like Ansible, Puppet etc to customize the options

- Configuration **validation** is automatically run at startup, informing about issues.

- New **authorization** and **authentication** configuration

  - gridmap-file now legacy and not recommended

- Proper JURA **accounting** section

- Some real time changes with **arcctl**

- Many other requested tweaks (maxcputime, maxwalltime...)

Stability
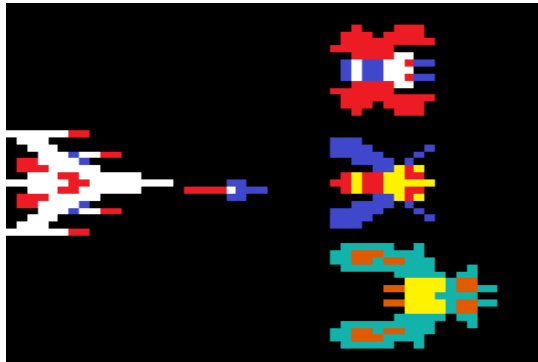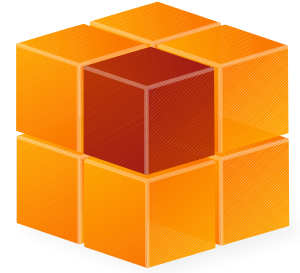
- Hunt for memory leaks, both in job handling and data staging

◼ Better logs

 – Messages try
 to point in the right direction

Performance

- Event-driven engine for better real-time handling of jobs.
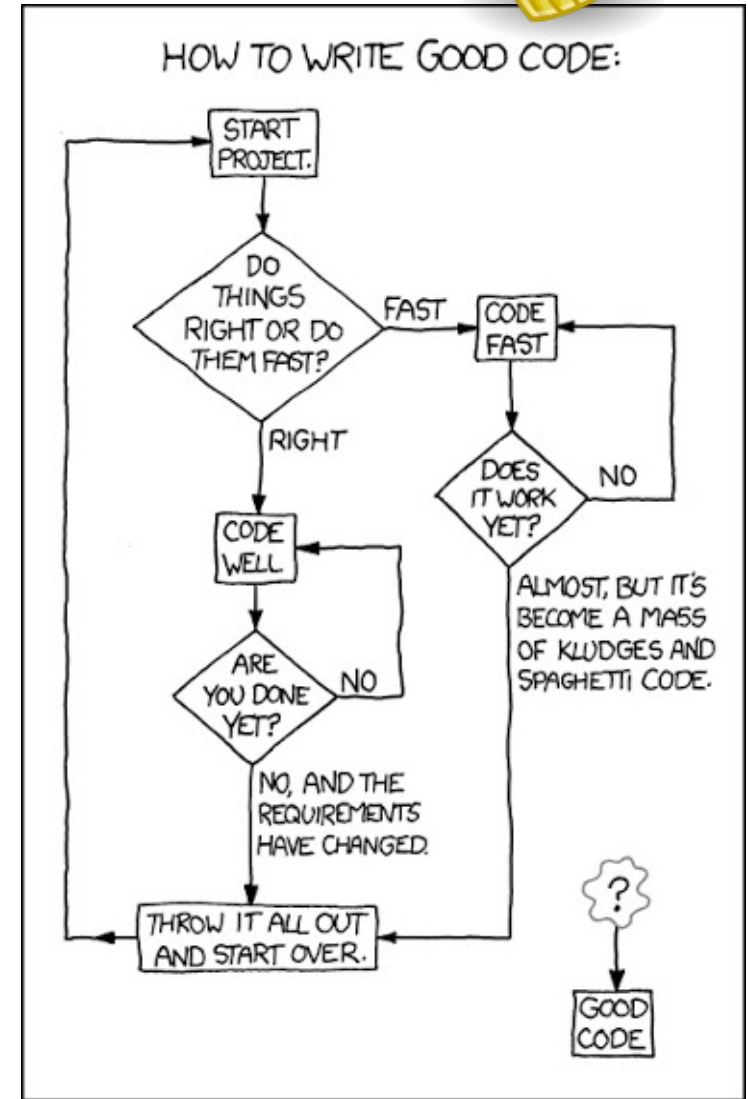 Target: 100000 jobs managed by A-REX

Programmer (noun.)

A person who fixed a problem that
you don't know you have,
in a way you don't understand.

Code Cleanup

- Rewrite/Review of obscure code

- Removal of features that were never used. Retired components:

    - EGIIS, the ldap-based grid index

    - EMIR, a WS grid index

    - Windows port

    - JAVA Bindings

Consolidation

- **Commonly used configuration schemes** and services have been **simplified** and their internal workflow revised (relationships between a-rex, gridftp, accounting, infocollection, lrms submission)

- **Unified defaults** for all subsystems

- **Obsoleted** components that will be removed on the next release (ARC7 ?)

  – Glue 1 support

New Features

- **arcctl** command line tool to manage A-Rex

- Better RTE handling, default RTEs **distributed with ARC** (e.g. ENV/PROXY)

```
[root ~]# arcctl rte list
<output omitted>
APPS/HEP/ATLAS-20.8.0-X86_64-SLC6-GCC48-OPT (user, enabled)
APPS/HEP/ATLAS-20.8.1-X86_64-SLC6-GCC48-OPT (user, enabled)
APPS/HEP/ATLAS-20.8.2-X86_64-SLC6-GCC49-OPT (user, enabled)
<output omitted>
ENV/LRMS-SCRATCH                (system, default)
ENV/PROXY                       (system, masked, disabled)
ENV/PROXY                       (user, enabled)
ENV/RTE                         (system, disabled)
ENV/RUNTIME/ALIEN-2.17          (user, enabled)
VO-biomed-CVMFS                 (dummy, enabled)
```

See dedicated talks!

Experimental Features
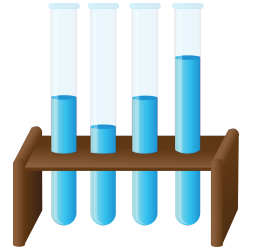
- INTERNAL plugin to allow minimal intrusive setup on a closed system. Very good for restrictive HPC's. (see talk about aCT!)

- Proof of concept of the REST interface
  - Usable but can be made better, may be in one of the minor releases.

Good to know:

- **No automatic configuration conversion tool.**
  ARC team offers help converting ARC5 configuration to ARC6 and gives support for configuration issues.

- Due to privacy regulations **no user information** should be **published** by ARC.
  Therefore gridmap-file based authorization where user DNs where stored is **deprecated**.
  ARC6 still supports it but gives more ways of expressing mapping and authorization rules.

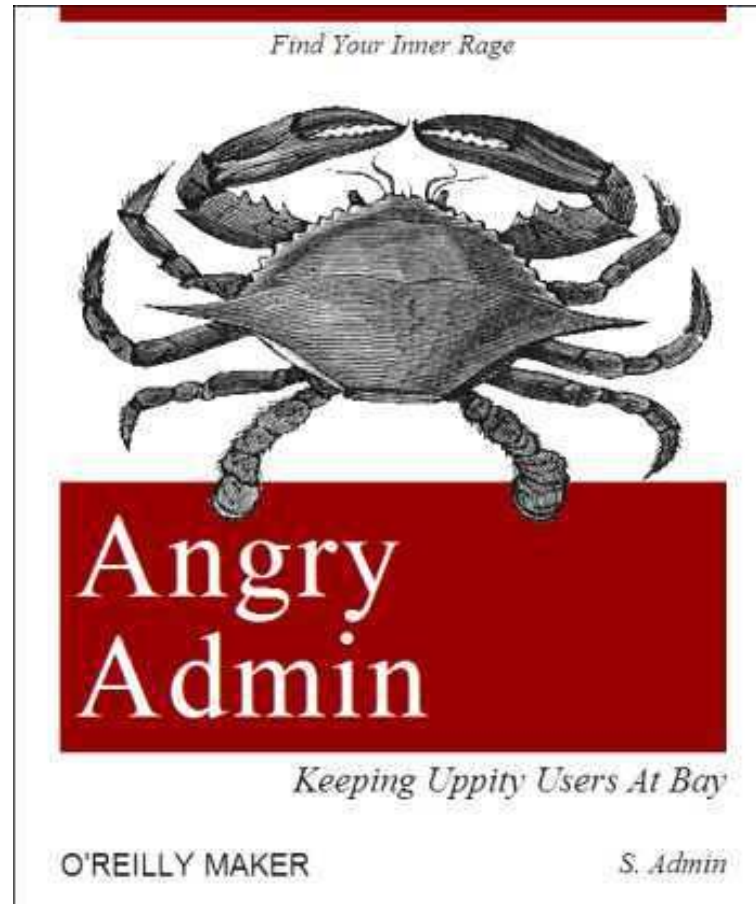 – Authorization can be fine tuned for each submission interface.

Good to know:

- ARC is trying to **get away from gridftpd + ldap**, so our default interface is https Web Services (WS) and supports the EMI-ES protocol. This is completely transparent when using ARC clients.

- WLCG decided that **GLUE2 is the current official information system schema**. Therefore ARC will by default publish GLUE2 information in the WS interface and no LDAP.

- As a consequence of the above, the classic WLCG configuration
  LDAP NorduGRID/Glue 1 + Gridftpd
  MUST be configured **manually** by WLCG sites.
  We are preparing dedicated documentation.

Questions

Picture references

- LibreOffice Gallery

- https://www.openclipart.org/

- https://xkcd.com/844/

- https://i.pinimg.com/736x/5b/eb/35/5beb35543f67789276dfad6d75805bab--computer-books-programming-humor.jpg

- http://www.xkcd.com

- https://www.quora.com/What-are-the-best-programmer-jokes-you-want-to-share

- https://shareitsfunny.com

- http://wheeltowheel.blogspot.com/2016/09/sept-9-heavy-load.html

- https://orig00.deviantart.net/42db/f/2010/024/d/4/galaga_by_keredlo.png

- https://www.recyclart.org/2016/01/beautiful-way-reusing-old-wood-logs-ceramics-stools/

References

- ARC6 presentation @ Umeå, Sweden, 2018

  http://indico.lucas.lu.se/event/1020/material/slides/9.pptx