Input summary from the R&C Sector

1st RCS-ICT Technical Committee
9th December 2022

Xavier Espinal and Pere Mató

Introduction and a disclaimer

- Highly valuable input received from all of you - **many thanks!**

- We made an attempt to **digest** the input and summarize it

    - Not an exhaustive summary - if a service/tool is not mentioned does not mean is not used

- We made an attempt to identify **areas of common interest**, to be addressed in subsequent technical discussions and in potential cross-collaboration projects

- All received input **will be made available** to the IT Technical Delivery groups

    - Some items are specific and could be addressed via direct dialogues

    - Some items have wider scope and needs further discussion and agreement - **this is where we want to put the accent**

Summary of received inputs

Storage and Data Management

- EOS [1:4,6,9,C,D], CernBox [3,6:8]

- CTA [1:4,9,D], TSM centralized backup [6]

- AFS [7,9]

- Ceph (Block storage, FS, S3) [C] and S3 Object Stores [6]

- CVMFS [B,D] and software releases (by SFT)

- FTS ensure long-term support and intensified development support [4,9,D]

- XrootD: IT Involvement critical to improve the performance and integration with HEP [3,D]

- Rucio: further engagement in the Rucio project and operational support for the experiments [3,4,D]

- Grid File Access Library (GFAL) [D]

Computing Infrastructure (1/2)

- LXPLUS and LXBATCH (HtCondor): data processing, simulation, interactive, etc. [∀]

- Cloud infrastructure: OpenStack [∀]
  - Magnum [C]
  - Physical Machines (ironic) [6]

- Common infrastructure for software development (nightly builds, CI, GPUs, performance, etc) [1,2,4,6,7]

- Heterogeneous computing (non-x86 architectures): PowerPC, ARM and accelerators/FPGAs integrated with the common infrastructure [1:4,6]
  - *Involvement of IT in the entire life cycle: procurement, installation, software setup and operation*

- Exploitation of opportunistic resources (HPCs and cloud facilities) [1:4,7,9]

Computing Infrastructure (2/2)

- Hosting externally owned servers in IT premises [6,7]

- HPC cluster [B], dedicated high-end machines [B]

- GPU resources [1:4,B,C]

- Support plans/collaboration for the future linux distributions [1:4,6,7,B,D]

  - Long-term support releases [9], convergence on a well supported RHEL clone [3,6]

- Service Availability Monitoring (SAM) and Experiment Testing Framework (ETF) [3,D]

- HammerCloud maintenance and operation [3,D]

- Large Clusters Management knowledge exchange [2]

Communication Systems

- Refurbishment and extension of the network in experimental areas (LS3) [3,4]

- Collaboration on HW procurement [4]

- Areas for improvement
  - Network monitoring: issues/bottlenecks, traffic visibility, xrootd [3,4]
  - Network orchestration, accounting and packet marking, optimization of flows (QoS) [3,4]
  - Landb: programmatic interface, VLAN support, etc (LS3 focus) [4]
  - Campus Network connectivity (P1 testbed) [4]
  - End-user friendly interfaces for Network Connection/LANdb (when registering/updating equipment) [8]

- CERNPhone (inc. integration with other tools, improvements) [3,8]

Databases

- Online Oracle Database, critical for data taking [2:4]

- DBoD (inc. service level and closer support) [5:7,C]

- Support for alternative database technologies: MySQL [9], Postgres [4,C]

- ElasticSearch/OpenSearch [C]

Applications and Platforms

- Container services
  - Harbour, Openshift, unpacked [3,4,6]

- MONIT/Grafana (inc. modernization possibilities and individual support) [1:4,7,D]

- Authentication & Authorisation: VOMS to token-based authentication migration and long-term support for IAM [3,4,D]

- Web hosting
  - Drupal migrations (inc. eventual need for professional support) [B]
  - Openshift/OKD, Static/Jekyll [5,6,7]
  - Reverse proxy support [6]

- Grappa / CERN e-groups [1,8]

Collaborative and Computational services

- GitLab [5,6]
  - CI [9,B]
  - Improved and facilitated access for external (non-CERN) collaborators [4]
  - Browsable artifact HTML pages [4]

- Issue Tracker: JIRA [5,6,9]

- Strategies for documentation and collaborative tools (codimd, Twiki, Sharepoint, Google docs, Indico, CDS, mattermost, Discourse) [2,4,6]

- KubeFlow (ML training) [C]

- Centrally supported e-Log [9]

- Desktop/notebooks/SWAN [5,9,D]
  - Training, tutorials [6]

- Support for Data and Analysis preservation, Open Data and Open Science [3,4,9]

Licenced Software

- Software & Engineering tools:
  - Mathematica, Maple [7,B],
  - CAD, FEA, LabView [7,5]
  - Altium, Ansys, Modelsim, Igor, Sigasi, Vivado,... [7]

Additional input

- Service Catalogue, status of critical services and support, BC/DR strategies [1:3,8,C]

- PCC: updated roadmap is needed to anticipate and verify the viability of possible usages [2]

- Energy efficient computing [1]

- Collaboration on procurement [1,7]

- Application development support: Burotel, Labotel [8]

- Enhance communication and foster participation, e.g. IT tools/services pilots [8]

- Bi-directional personnel exchange practice (IT-Experiments) to foster knowledge transfer [2]

- Partnership with other experiments and sciences leveraging EU initiatives [D]

- Perception support provided by IT is often targeting computing scientists with sound expertise, closer/individual support from IT would be appreciated [7]

Potential initiatives and projects

Potential (not prioritized) initiatives and projects (1/3)

- Common infrastructure for software development, including discussion on:
  - Access and support for non-x86 arch: PowerPC, ARM and Accelerators (GPU, FPGA)
  - Access and support for GPU and Mac nodes (build)

- Scientific and Distributed Computing tools and services, including discussion on:
  - Rucio: engagement and experiment support
  - FTS: support and development
  - XrootD, SAM/ETF, HammerCloud
  - Token-based authentication (IAM)

- Collaborative tools strategy, including discussion on:
  - Atlassian products future (JIRA, Confluence). Prospects with: GitLab (CI,artifacts), GitHub runners, Discourse
  - Web: Drupal (migrations) support, static webs/Jekyll
  - Mattermost, Indico, Twiki, codimd, CDS and Google tools

Additional ideas on initiatives and projects for next TCs are more than welcome!

Potential ^(not prioritized) initiatives and projects (2/3)

- File Systems and Object Stores evolution, support model and operations. Including discussion on:
  - Filesystems, AFS, Home directories, Ceph, CVMFS

- Fabric and infrastructure, including discussion on:
  - Prevessin Computing Center update
  - Collaboration on hardware procurement and "hosting"

- Linux distributions and other OS, including discussion on:
  - LTS releases, convergence on RHEL clone and support,
  - OS/Windows and MacOS

- Networks, including discussion on:
  - LS3 network-related works in the experimental areas
  - Suggested improvements on: Monitoring, Landb, QoS, etc.
  - R&D activities for Run4

Additional ideas on initiatives and projects for next TCs are more than welcome!

Potential (not prioritized) initiatives and projects (3/3)

- Service Catalogue, Critical Services, Business Continuity and Disaster Recovery

- Current usage of licensed software and future outlooks
  - Mathematic, Maple, Cadence, CAD, FEA, LabView and others

- Container services (registry) strategy, including discussion on:
  - CERN common repository for images
  - Best practices

- Monitoring, including discussion on:
  - Specific network monitoring
  - Monitoring for the experiments: MONIT

Additional ideas on initiatives and projects for next TCs are more than welcome!

Possible cross-collaboration ideas and projects (1/2)

- Software/hardware optimisation
  - Software developers and service engineers together. Understand how to optimise the code with a given hardware and vice versa

- Prototyping of an Analysis Facility
  - Involving IT groups, experiments and EP-SFT

- Modernization of the experiment software and common libraries
  - Joint IT, EP-SFT and experiments to facilitate the software development process.

- IT infrastructure for microelectronics design (ASIC)

Additional ideas on cross-collaboration ideas and projects are more than welcome!

Possible cross-collaboration ideas and projects (2/2)

- Co-design of energy efficient computing
  - Fabric (hardware and cooling) but also software, algorithms and workload management

- Study of the CMS premixing workflow
  - optimization of the access pattern of remote data and code for xrootd, ROOT or CMSSW

- Specific network R&D for Run4 on RoCE (RDMA over Ethernet)

- Implement advanced monitoring of LHCb-DIRAC agents and services in MONIT

AOB

- How we proceed in defining thematic(s) for the next TC(s)?

- Agenda privacy?

- TC frequency ?

- Next tentative RCS-ICT TC in **February**. Please let us know in case you have preferred week slots.