CERN**IT**
Department

IT/DM (Data Management) Group meeting

Alberto Pace

Group photo

- Now !

- Electronic Engineer, age 43, various jobs at CERN … in 20+ years

- Worked in many areas

  - AA & LPI vacuum system, PS Workshop, CAD (euclid) suport, PS accelerator operation, Isolde and Booster Control Systems, Novell deployment, NICE95/NT, Web Services, Mail Services, Windows desktop management

  - Many sections, many groups

    - PS/ML/VAC, PS/ML/MI, PS/ML/ES, PS/OP, PS/CO, CN/DCI/NCS, IT/DIS, IT/IA/WS, IT/IS, IT/DM

    - 8 different group leaders in the first 15 years …

Who is Alberto Pace

- ## Current responsibilities
  - Leading the Internet Services group since 2002
  - (and few other things …)

- ## Knowledge / Skills
  - Networking, Web / Mail services, Desktop management, Databases, Storage
  - Services, Operation, Support
  - Programming / Architecture
  - (Accelerator operations, Control systems, …)

DM group mandate

- Ensure a coherent development process for CERN Physics Data management activities, strongly driven by operational and user needs

- CERN Physics Data management activities includes Castor, DPM, FTS, LFC and associated grid tools, Physic database operation and development of related applications and tools

**Lack of data management could
trigger the failure of the whole chain**

- Merge from 4 groups
- From IT/PSS
  - Radovan Chytracek; Romain Basset; Dirk Duellmann; Giacomo Govi; Zsolt Molnar; Andrea Valassi; Miguel Anjo; Luca Canali; Eva Dafonte Perez; Dirk Geppert; Maria Girone; Ioannis Papadopoulos; Gianni Pucciani; Jacek Wojcieszuk; Dawid Wojcik
- From IT/FIO
  - German Cancio Melia; Rosa Maria Garcia Rioja; Giuseppe Lo Presti; Steven Murray; Sebastien Ponce; Giulia Taurelli; Dennis Waldron
- From IT/GD
  - Lana Abadie; Jean-Philippe Baud; Akos Frohner; Robert Harakaly; Remi Mollon; Krzysztof Nienartowicz; Andreas Joachim Peters; David Smith; Paolo Tedesco
- From IT/IS
  - Alberto Pace

- **2 sections**
  - The IT-PSS-DP section moves unchanged
    - Section leader: Dirk Duellmann
  - Developments in Grid data management and Castor areas will merge
    - Section leader: German Cancio
  - Dirk Duellmann will also act as deputy group leader
  - Jean-Philippe Baud will be directly attached to me and act as a "technical architect"
- **Adjustments may happen following the POW meeting of next week**

People in the DM group

- The two sections
- From IT/PSS
  - Radovan Chytracek; Romain Basset; Dirk Duellmann; Giacomo Govi; Zsolt Molnar; Andrea Valassi; Miguel Anjo; Luca Canali; Eva Dafonte Perez; Dirk Geppert; Maria Girone; Ioannis Papadopoulos; Gianni Pucciani; Jacek Wojcieszuk; Dawid Wojcik
- From IT/FIO
  - German Cancio Melia; Rosa Maria Garcia Rioja; Giuseppe Lo Presti; Steven Murray; Sebastien Ponce; Giulia Taurelli; Dennis Waldron
- From IT/GD
  - Lana Abadie; Jean-Philippe Baud; Akos Frohner; Robert Harakaly; Remi Mollon; Krzysztof Nienartowicz; Andreas Joachim Peters; David Smith; Paolo Tedesco
- From IT/IS
  - Alberto Pace

- Train newcomers and reallocate tasks to a smaller team

- Various software consolidation projects
  - SEAL replacement, CORAL server, maintain POOL, COOL scalability test

- Operate physics database
  - migration to the new hardware, transition to 64bit, Oracle 11g
  - Prepare / optimize operation in view of the LHC startup

- Integrate experiment funded DBAs for setting-up and support their online RACs

- LCG 3D Tier 1 coordination and consultancy, including license needs and move to multi-core / 64-bit

Data replication and Managed Storage Development

- Build a common team with people coming from FIO and GD
- Prepare / optimize / finish in view of the LHC-start-up
  - FTS 2, Consistency check tools, monitoring, quotas
- Integration/enhancements of common DPM/CASTOR modules
  - Nameserver, RFIO interfaces, security / Strong authentication / VOMS integration, Data checksums
- Integration/enhancements in tape area, improve resource optimization
  - handling of small files, policies for recalls and migrations
- Consolidate / improve CASTOR operations
- Improve test/certification process
  - Efficient low-latency data access in particular for analysis ("xrootd")
- Ensure continued good communication with stakeholders (users and operations in/outside CERN)

- I have / will be interviewing all group members

- Issues raised
  - Improve communication / information
  - Improve training / coaching / mentoring of junior group members
  - Avoid duplication of efforts
    - Operational aspects
    - Development
  - Discussions on technical architectures
  - Maintain/improve the link with operational teams

- **Weekly technical meetings**
  - These meetings will take place on Tuesdays, 14:00 to 16:00 in room 513/1-024
  - Series of tutorials / presentations / proposals
  - An opportunity for
    - Explain to other group members the inner architecture of DM software and services
    - Report the work done to other group members. Raise awareness of what is available
    - Submit technical proposals
    - Raise group wide issues
    - Improve communication in general
  - Every group member is expected to give *at least* one presentation every year
  - Maria Girone has accepted to coordinate and chair this activity

Prepare the future

- Can we simplify our architecture ? for maintainability ? Can we reduce the number of DBs and simplify the schema ?

- Do we need formal software development processes ? testbeds for automated tests ?

- Are we at the right balance between generic solution versus optimized solutions ?

- Do we need disk servers for all applications (DAQ for example) ? several layers of disk ?

- What should be the role of tapes?

- Could / should we mix archive and active data on same spindle? how do we handle small files?

- Test different file systems ? other platforms for disk servers ? new access protocols to replace RFIO/Xroot ?

- Can we reduce the number of daemons ?

- Can we reduce external dependency on licensed commercial applications ?

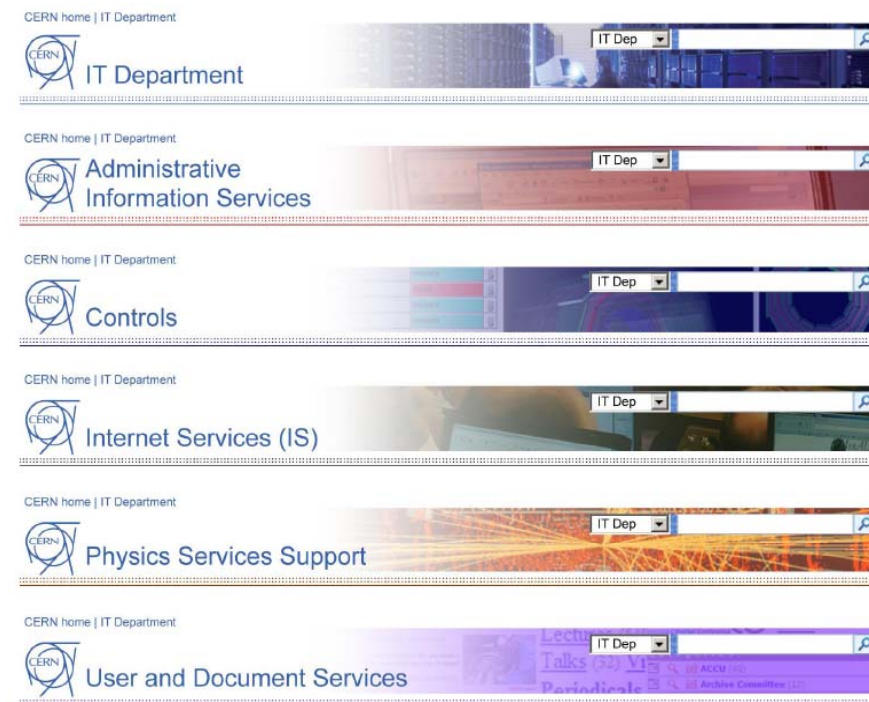It will take years to answer all these questions …

- A small team to discuss the architecture, the vision for the future, the strategy, the priorities in developments
  - Details not yet finalized. Discussions animated by Jean-Philippe Baud and myself.
  - Experts will be called in as necessary …
- All existing cross group meetings must continue.
  - Operational coordination meetings with FIO
  - Collaborations and discussions with DES in database service areas

- **Following IT recommendations**
  - See group leader meeting from October 17
    http://indico.cern.ch/conferenceDisplay.py?confId=22651

Group image

- ## Powerpoint template
  - Available from http://it-comm-team.web.cern.ch/IT-Comm-Team/repository/ITgroups/

- ## Posters

- ## Group site
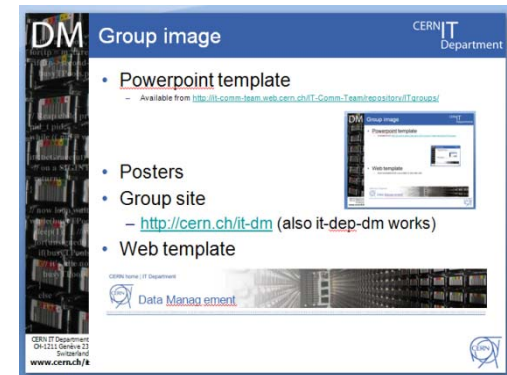 - http://cern.ch/it-dm (also it-dep-dm works)

- ## Web template

CERN home | IT Department

Data Management

Questions ? / Discussion ?