CDS and the Management of Multimedia Documents

Thomas Baron
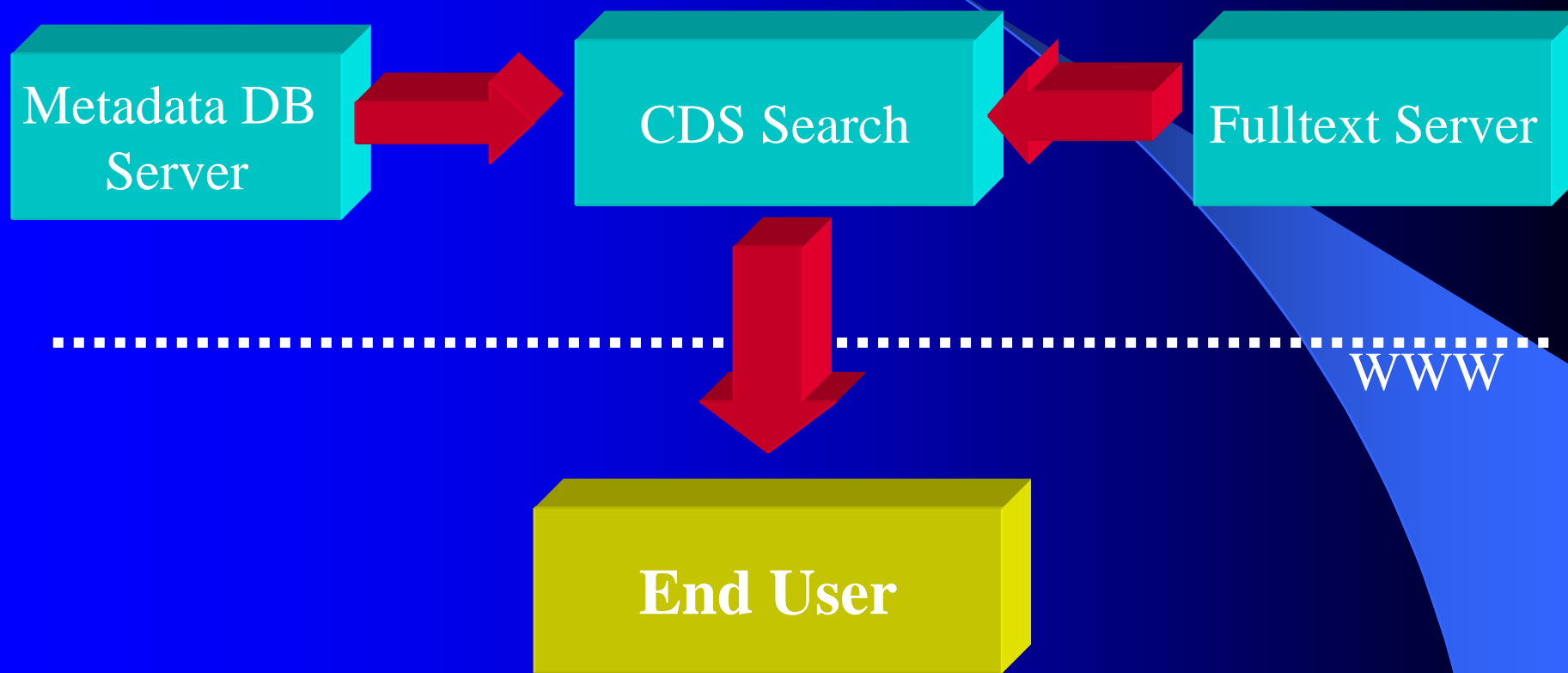
Tibor Simko

CERN ETT-DH-CDS

August 2002

Outline

1. CERN Document Server
2. Multimedia Documents in CDS
   - Description
   - Acquisition
3. CDS Agenda
4. Video Processing
5. INDICO
6. Conclusions

1 – CERN Document Server

- All particle physics literature since 1950

- And related areas documents:

  – Astrophysics, Mathematics, Life at CERN…

- 550,000 documents in more than 200 collections

- 220,000 fulltext files (preprints, photos, theses…)

- Originally developed on top of Aleph 300 in 1994

- Now free software (CDSware: http://cdsware.cern.ch) fully ALEPH independent.

1 – CERN Document Server

| Metadata DB Server | → | CDS Search | ← | Fulltext Server |

CDS Search → End User

WWW

2 - Multimedia Documents in CDS
Description

- More than <u>10 multimedia collections</u>.
  - Videoclips, pictures, objects, slides, posters…

2 - Multimedia Documents in CDS
Description

- Data formats + Services
  - Pictures:
    - Automatic thumbnail creation (gif)
    - Several sizes/resolutions/conversions proposed (original in jpeg)
    - Example
  - Videoclips
    - Manual thumbnail creation (gif)
    - Proposed format Real Media (.rm)
    - Example

- Metadata format
  - Internal representation: MARC21
  - Several possible output formats
  - Photo Example
  - Videoclip Example

2 - Multimedia Documents in CDS
Acquisition

- Web Submission form (CDS Submit)
  - Metadata collection
  - File transfer
  - Thumbnails creation
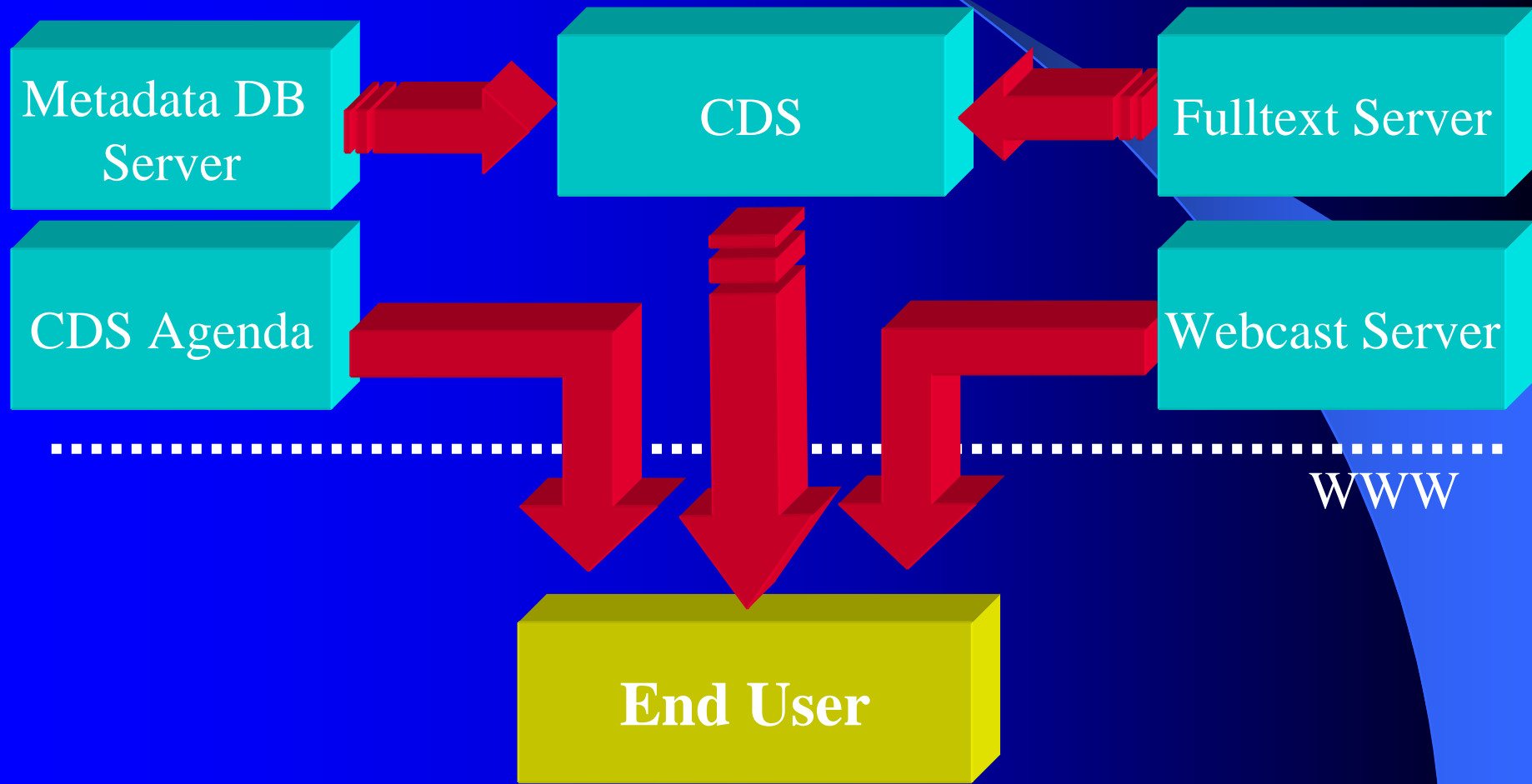 - Approval process
- Agenda Module (CDS Agenda)

3 – CDS Agenda

- Started 1999 – ATLAS Experiment
- Now widely used
- Stores meeting time-tables …
- … And associated multimedia files
- 2500 Agendas
- 15000 files
- http://agenda.cern.ch

4 – Video Processing

- Seminars and conferences are recorded and webcasted
- Specific data format needs specific file server (streaming): http://webcast.cern.ch
  - RealMedia format
- Added value:
  - Link videos and transparencies together
  - Timing acquisition (automatic when using MS Powerpoint, but manual when using mere slides)

Recap…

Metadata DB Server → CDS ← Fulltext Server

CDS Agenda

Webcast Server

WWW

End User

5 - INDICO

- INtegrated DIgital COnference
  - http://indico.sissa.it
  - EU funded project over 3 years
  - One centralized server for storing conference announcements
  - One locally installable software for managing a conference.
  - High Tech features:
    - Soundtrack indexing
    - Automatic video storyboarding

6 - Conclusions

- **CDSware**: Open source software capable of handling multimedia documents

- **CDS Agenda**: Open source Workshops and Meetings storage system

- **INDICO**: Conference Management Software

Your Questions

http://cds.cern.ch
http://cdsware.cern.ch
http://agenda.cern.ch
http://webcast.cern.ch
http://indico.sissa.it