CERN Document Server on Invenio 1.1.3

CERN's open source Document Repository information system

Flavio Costa – Jean-Yves Le Meur
CERN IT Dep.

Agenda

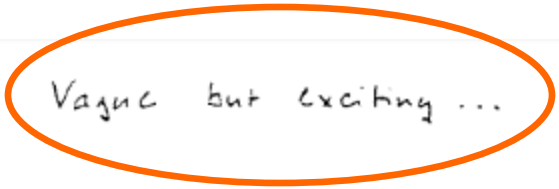
- Introduction
  - CERN "Where the Web was invented"
  - Invenio: reasons, scope, history and use at CERN
- CDS facts
  - Key facts and figures
  - Data heterogeneity
  - Communities heterogeneity
- CDS “specials”
  - Events & Performance peaks
  - E-Tendering & E-Bidding
  - Web Journal: CERN Bulletin
- Migration to v. 1.1.3
  - The process
  - The Gains
  - Next?

Agenda

- Introduction
- CERN "Where the Web was invented"
- Invenio: reasons, scope, history and use at CERN
- CDS facts
- Key facts and figures
- Data heterogeneity
- Communities heterogeneity
- CDS "specials"
- Events & Performance peaks
- E-Tendering & E-Bidding
- Web Journal: CERN Bulletin
- Migration to v. 1.1.3
- The process
- The Gains
- Next?

Presentation

- CERN: Where the web was invented



CERN DD/OC

Tim Berners-Lee, CERN/DD

Information Management: A Proposal

March 1989

Information Management: A Proposal

Abstract

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext sytstem.

Open Access:

- Background: CERN Convention (1953) contains:
  - "… the results of its experimental and theoretical work shall be published or otherwise made generally available": an early Open Access manifesto
- Open Access ≠ no more access management
  - Copyright acceptance workflows
  - Publication workflows; logins, ACLs (Access Control Lists)
- Changing the publishing model
  - Spiralling subscription costs, falling subscriptions
  - Pay to create, not download: Open Access
  - SCOAP3: http://www.scoap3.org
    - 3-way partnership: scientists, libraries, publishers

Presentation

- History:
  - 1991: SPIRES (SLAC) First database available on the web
  - 1993: First web catalogue of documents (pre-prints) at CERN, ancestor of Invenio
  - 2002: Invenio is available as Open Source software (GNU-GPL)
  - Today Invenio is being co-developed by an international collaboration comprising institutes such as CERN, DESY, EPFL, FNAL, SLAC
  - 2012: SPIRES is replaced by INSPIRE on Invenio

Agenda

- Introduction
- CERN "Where the Web was invented"
- Invenio: reasons, scope, history and use at CERN
- **CDS facts**
- Key facts and figures
- Data heterogeneity
- Communities heterogeneity
- CDS "specials"
- Events & Performance peaks
- E-Tendering & E-Bidding
- Web Journal: CERN Bulletin
- Migration to v. 1.1.3
- The process
- The Gains
- Next?

CDS key facts and figures

Institutional Repository for CERN:

- ~1’000 Digital Collections, both with digital origin or converted from a physical media
- Around 1.5M records
  - 500k with full text
  - 150k with links to external full text

CDS key facts and figures

Institutional Repository for CERN:

- Around 100 submission basic workflows
- Around 8’000 registered and 40’000 unregistered users
- 30k distinct visitors per month

CDS key facts and figures

Digital Library, definition by DL.org: "An organisation, which might be virtual, that comprehensively collects, manages and preserves for the long term rich Information Objects, and offers to its Actors specialised Functions on those Information Objects, of measurable quality, expressed by Quality Parameters, and according to codified Policies."

Digital Library <===> Institutional Repository (Wikipedia: Digital Library)

>> SINERGY <<

- 30k distinct visitors per month

CDS: Data heterogeneity

- CERN Institutional Repository
- Subject Repository for Particle Physics (=> Inspire)
- Physical collections
  - Library and Book Shop
- Preprints, “postprints”, theses
- Conference proceedings, lecture objects
  - Slides and recordings
- Experiment Support material
  - Photos, videos, animations
- Reusable information
  - Data in tables, figures , plots
  - Correlation matrices
  - Data (high-level objects)

CDS: Communities heterogeneity

- Library (curation, administration…)
- Particle Physics users (CERN Experiments, Collaborations, …)
- CERN Institutional Repository (archive, web sites documents, …)
- Document management (including administration, e-tendering/e-bidding, CERN Council and governing bodies…)
  - Multimedia Document Assets Management

Agenda

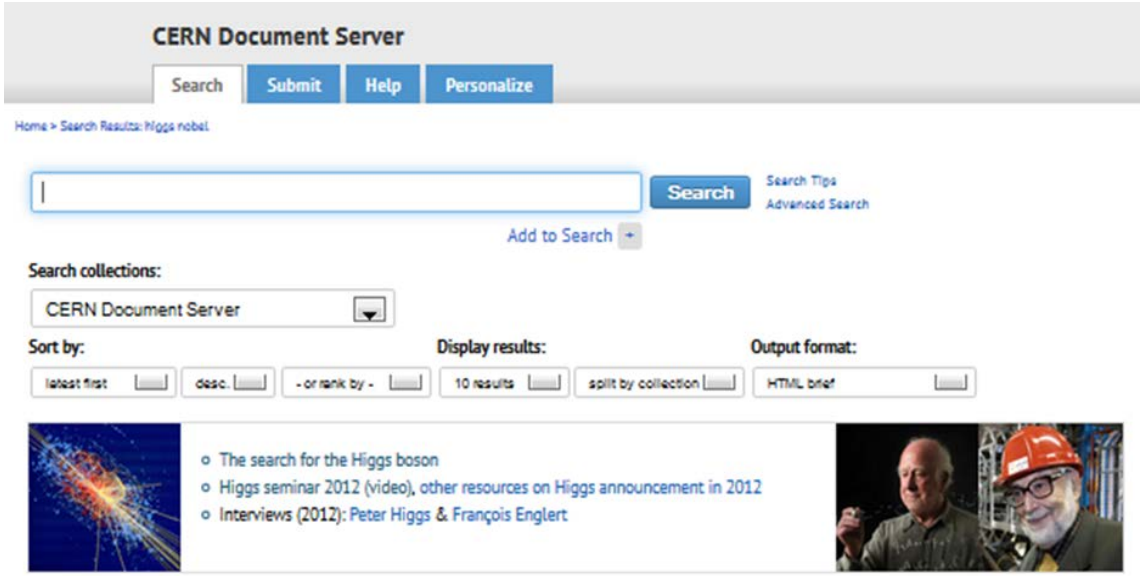
- Introduction
- CERN "Where the Web was invented"
- Invenio: reasons, scope, history and use at CERN
- CDS facts
- Key facts and figures
- Data heterogeneity
- Communities heterogeneity
- CDS "specials"
- Events & Performance peaks
- E-Tendering & E-Bidding
- Web Journal: CERN Bulletin
- Migration to v. 1.1.3
- The process
- The Gains
- Next?

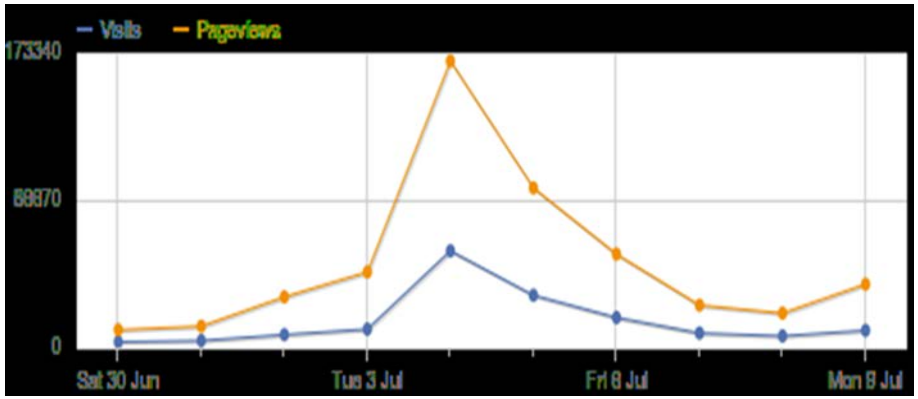
Events & Performance peaks

Peaks: ex.: Nobel Prize, Open Days, *Higgs discovery*, LHC Startup

Sustained ~230 web requests/second and rate peaks of up to 800 requests/second

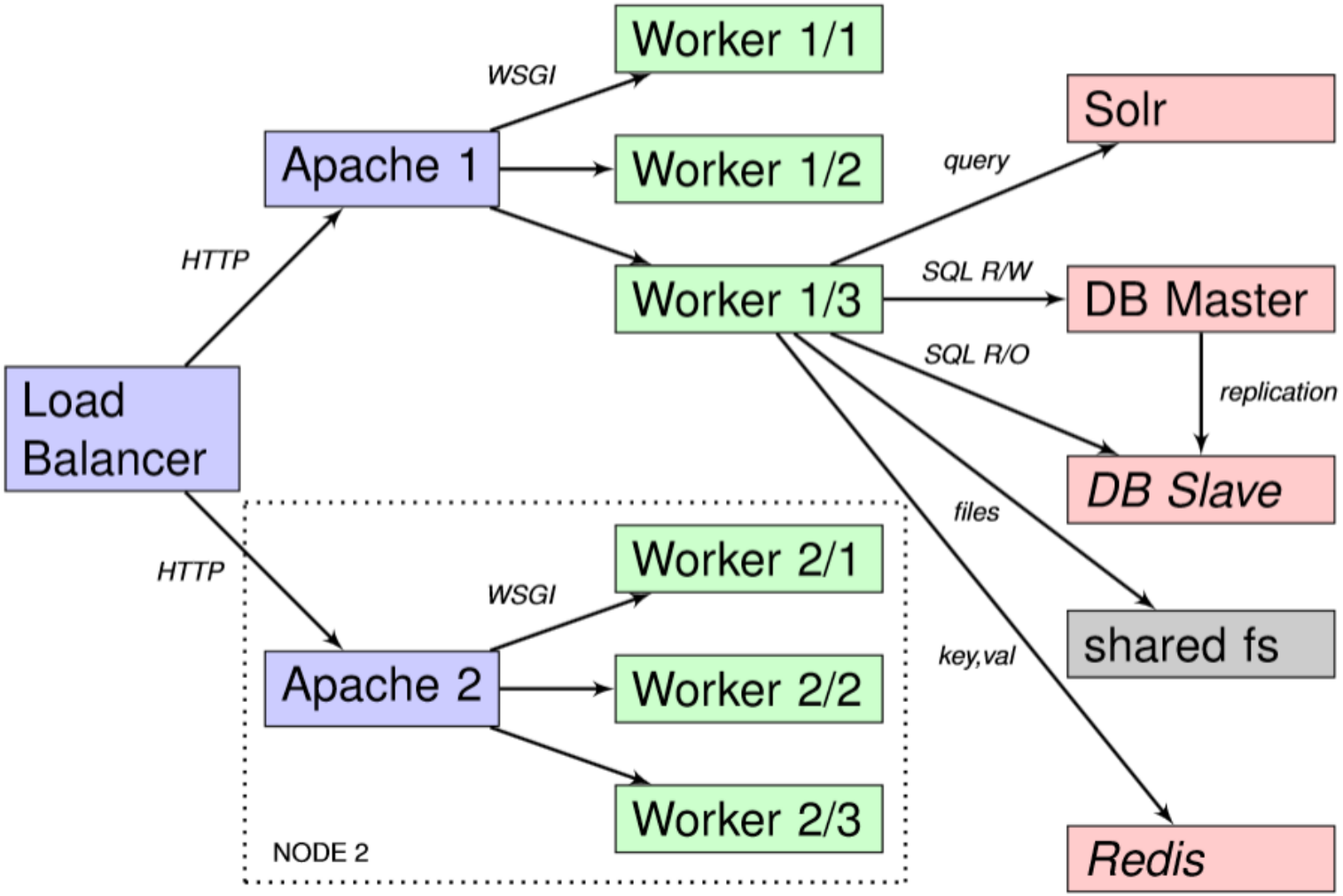
Sustained DB processing peaks of ~ 10,000 queries/second





Events & Performance peaks: architecture

- 800 hits per sec on CDS during Higgs seminar July 4th



E-Tendering & E-Bidding – B2B

CERN Document Server

CDS | Indico | Library | Bulletin | EDMS

Search | Submit | Help | Your CDS ▾ flavio.costa@cern.ch logout

Home > Supplies, Procurement & Logistics (SPL) > e-Tendering > Market Surveys > Current Market Surveys > Exhibitions at CERN Visit Points

Information | Discussion | Files

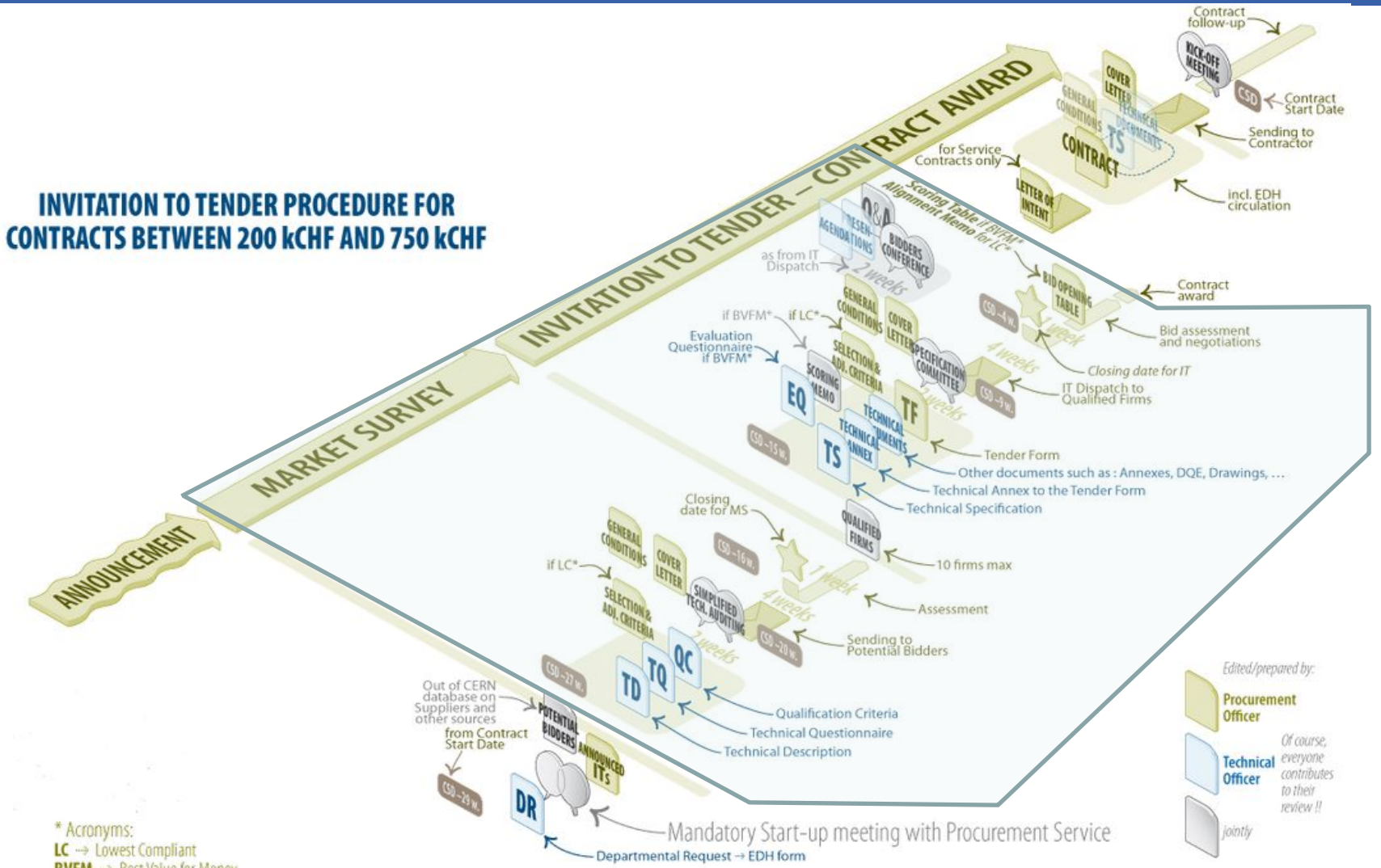
e-tendering -

Report number	MS-3797/PH
Title	**Exhibitions at CERN Visit Points**
Dispatch date (yyyy-mm-dd)	2011-07-05
Reply date (yyyy-mm-dd)	2011-08-02
Files	**Cover_Letter** : MS-3797_Cover_Letter.pdf [361.0 KB] 05 Jul 11 15:31, version 1 **List_of_Firms** [restricted access] : MS-3797_Firms.pdf [5.0 KB] 05 Jul 11 15:33, version 1 **Qualification_Criteria** : MS-3797_Qualification_Criteria.pdf [107.0 KB] 05 Jul 11 15:31, version 1 **Technical_Questionnaire** : MS-3797_Technical_Questionnaire.doc [125.0 KB] 05 Jul 11 15:33, version 1 MS-3797_Technical_Questionnaire.pdf [200.0 KB] 05 Jul 11 15:33, version 1 **Technical_Specification** : MS-3797_Technical_Desciption.pdf [961.0 KB] 05 Jul 11 15:31, version 1

[modify this record] [transer files] [new record]

Record created 2011-07-05, last modified 2011-07-05 Similar records

E-Tendering & E-Bidding – B2B



Web Journal: CERN Bulletin

CERN Bulletin

News Articles | Official News | Training | Announcements | Events | Staff Association

english | français

search

Issue No. 23-24/2013 - Monday 3 June 2013

Printable version - Subscribe:

One or many?

A shower of good news

LS1 Report: operations on the right track

Interview with Cédric Villani

Now playing in a browser near you: TEDxCERN

ACCU: a committee that addresses Users' needs

n_TOF: a new experimental area under way

FameLab Switzerland: a CERN PhD student triumphs

Don't miss the Passport to the Big Bang event this Sunday!

CERN Relay Race: sporty and colourful

Highlights from e-EPS: Physics league... for kick-ass students

The Lure of Wireless Encryption

Ombuds' corner: Virtue ethics and its applications

News from the Library: Language learners, don't forget the Library!

Subscribe by RSS

INTERVIEW WITH CÉDRIC VILLANI

On 21 May this year, CERN had the pleasure of welcoming the mathematician Cédric Villani as part of the series of lectures organised by the *Groupement des Français du CERN*. The *CERN Bulletin* took the opportunity to ask him a few questions.

Cédric Villani (left), at SM18, alongside Frédérick Bordry, head of CERN's Technology Department, is the director of the Institut Henri Poincaré and professor at the Université Claude Bernard Lyon 1. He was awarded the Fields Medal in 2010 for his work on Landau damping and the Boltzmann equation.

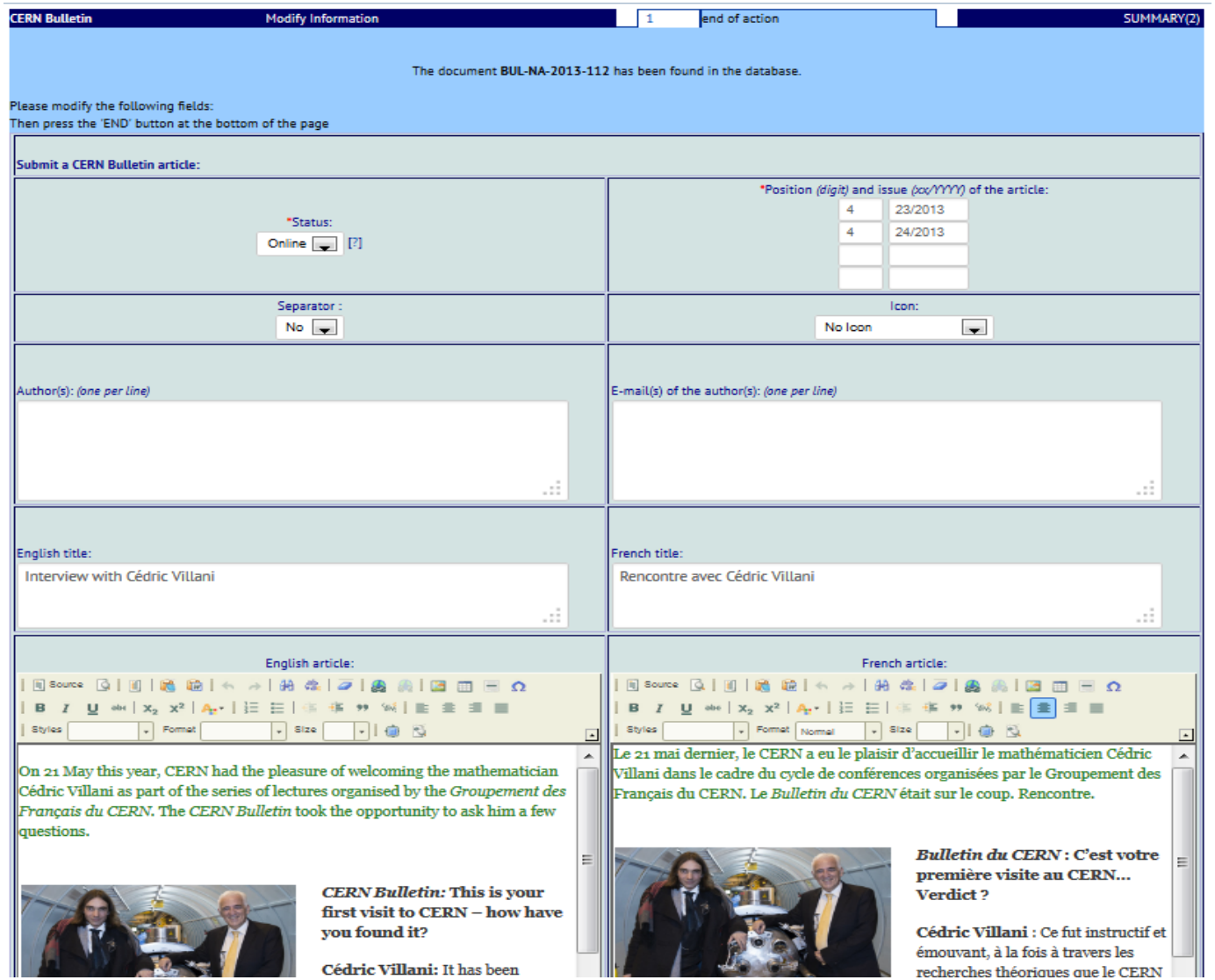
***CERN Bulletin:* This is your first visit to CERN – how have you found it?**

Cédric Villani: It has been instructive and moving, both because of the theoretical research that CERN makes possible and because of the technological feat that it represents – truly a technological work of art!

***CERN Bulletin:* You are currently working on Riemannian geometry problems. Is there any link between this work and the research carried out at CERN?**

Cédric Villani: All the work that I've done in mathematical physics builds on classical

Web Journal: CERN Bulletin



Web Journal: CERN Bulletin

Agenda

- Introduction
  - CERN "Where the Web was invented"
  - Invenio: reasons, scope, history and use at CERN
- CDS facts
  - Key facts and figures
  - Data heterogeneity
  - Communities heterogeneity
- CDS "specials"
  - Events & Performance peaks
  - E-Tendering & E-Bidding
  - Web Journal: CERN Bulletin
- Migration to v. 1.1.3
  - The process
  - The Gains
  - Next?

CDS migration to Invenio1.1.3

- From v. 1.1.1 to latest master
- On new servers managed with Puppet / SLC6
- Why ?
  - HW out of warranty!
  - SLC5 deprecated, missing libraries
  - Quattor phased out
  - Invenio v. 1.1.3 new features
    - BibSched
    - BibEdit
    - BibIndex
    - ...

CDS migration to Invenio1.1.3

- Pre-Process:
  - One new node with recent dbdump used to re-compute all indexes
  - A temporary DB node set as slave of the production DB
  - Created a cds-ops git repo to manage the 'incubating' code
  - WNs set up with invenio (cds-ops) + cds overlay
- D-Day:
  - Production service stoppage (~3:30 h)
  - Run upgrade tables on the temporary DB node (1:30 h), set as new master DB node
  - Copied new pre-created indexes to the WN(s)
  - Switch the Load balancer to use the new nodes

CDS migration to Invenio1.1.3

- Post-process
  - Some issues with the File management system
  - A few 'surprises', quickly fixed
- Lessons learnt
  - Globally smooth
  - Differences in Production and Pre-Prod environments always lead to 'surprises'
  - Downtime can hardly be avoided without losing some content, even in read mode
  - Relief to be on 1.1.3:
    - it enables deployment of many adds-on
    - it opens Next door!

CDS migration to Invenio1.1.3

- On-going prototype
  - Change in Content: focus on institute content
  - Change in Design: no 'tree navigation' by default
  - Change in Features: personalisation at all levels → collections, author pages, stats, etc.
- Based on
  - Users Feedback
  - Piwik Numbers
  - Users iterations

THANK YOU!

Flavio.Costa@cern.ch
Jean-Yves.Le.Meur@cern.ch

http://invenio-software.org/