Update on CERN Search: SharePoint 2013

E. Alvarez, S. Fernandez, A. Lossent, I. Posada, A. Wagner [CERN]

CERN's enterprise Search solution 'CERN Search' provides a central search solution for both users and CERN service providers. Public and protected documents from a wide range of documents are indexed and available for retrieval.

- High Availability
- 2000 daily queries
- 29 servers
- <5 seconds switch.

- Improved design
- 33 million documents
- 1100 daily unique visitors
- 3 server farms
- 12000 daily page views

- High Reliability
- 33 million documents
- 1100 daily unique visitors

Content Sources

- Push model.
- Access control for each source.

- Pull model.
- Dynamic URI list.

- Stores data in protected space.
- Extracts metadata from binary files.
- Sets ACL for each document.

- Microsoft SharePoint 2013 search engine.

- Redundancy of key components.

- Staging farm as high availability failover.

Search Engine

- User interface upgraded.
- Improved user experience.
- Thesaurus for acronyms.
- Promoted CERN most relevant sites.
- Map integration for locations.

- Improved integration as search backend.

Prod

Stag

Frontend

REST

Prod

Stag

Frontend

REST

Crawler

Pre-processing

Indexing

Push web services

Push model.
Access control for each source.

Pull model.
Dynamic URI list.

- Analyze and split documents in several stages.
- Custom processing.
  - Different processing by content source.
  - Dynamic property mapper.
  - Property mapping by configuration.
  - Auto load-balanced CEWS call.

http://cern.ch/IT © CERN CC-BY-SA 4.0

https://search.cern.ch

https://home.cern