Large Scale Production Using Project Maps

Youssef BENNANI

Nov 13, 2016
This presentation will show how SAP has integrated the **DITA Open Toolkit** to create a large-scale production infrastructure able to build 60 000+ outputs daily for SAP product documentation. This covers the end-to-end process from DITA CMS to delivery channels, using **project maps** as a key ingredient.

- **Introduction : SAP environment**
  - DITA Adoption and growth
  - DITA Process and Landscape overview

- **DITA Production process @SAP**
  - Project(*) production life cycle based on **project map**
  - **DITA Open Toolkit** integration into the production landscape
  - Peer linking across outputs as example of advanced feature
  - Infrastructure optimization to scale

(*) Project : documentation set for a product shipment
Introduction: SAP environment
## DITA implementation @SAP

<table>
<thead>
<tr>
<th>General decision to use DITA</th>
<th>CMS evaluation and selection</th>
<th>Initial implementation and migration</th>
<th>Continuous improvement and incremental growth</th>
<th>Evaluation and decision to broaden</th>
<th>Migration and scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>150 authors</td>
<td></td>
<td>300 authors</td>
<td>700+ authors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First products</td>
<td></td>
<td>100 products</td>
<td>300 products</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First 12 languages</td>
<td></td>
<td>21 languages</td>
<td>47 languages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First outputs based on DITA Open Toolkit</td>
<td></td>
<td>12k+ outputs (daily)</td>
<td></td>
<td>60k+ outputs (daily)</td>
<td></td>
</tr>
</tbody>
</table>

© 2016 SAP SE or an SAP affiliate company. All rights reserved.
Some statistics

1000+ CMS users (registered, not necessarily all active)

3+ Million DITA files x2 increase in 2Y

50% Migrated content x1.4 increase Y2Y

3200+ Projects x2 increase Y2Y

CMS Content(# k-DITA files)
Day-to-day process overview

- Content validation, technical checking, and translation of documentation begins early and is performed continuously
- Final preparation before publication is reduced to minimum

Reviewers/contributors
- Access latest doc content
- Fix/add information directly in the right place

Translators
- View source content
- Preview translation

Authors
- Co-create with developers
- Review with stakeholders

Internal Preview server

DITA builds

Automated pushes through daily builds

Translation (SOLITT)

Every 1h

Daily

Nexus

Software builds

2016 SAP SE or an SAP affiliate company. All rights reserved.
SAP DITA Landscape overview

- ~40 servers in central data center, operated by IT
- Central admin team
- RDP access for users around the world
SAP DITA Landscape detailed view

**Web Author (Browser)**

**CMS Client / RDP environment**
- Production plugins: Production center, Build dashboard, Project Map editor...
- Oxygen Customizations
- IXIASOFT extensions
- Schematron
- Acrolinx
- Oxygen editor
- IXIASOFT DITA Eclipse

**RDP Infrastructure**
- (DC, Storage, Servers, Remote access)

**Build environment**
- SAP DITA Build Framework
- DITA Output toolset
  - DITA OT customization
  - DITA Open Toolkit
  - Antenna house
- DITA Build system
  - & scheduling
- Build Infrastructure
  - (DC, Build servers, Logistic Servers)

**CMS back-end environment**
- Web Author Server
- Output Generator Server
- Cloning Scheduler
- PNG creation
- Delta fetch
- Translation exchange
- Replication
- TEXTML Authoring
  - Connection, Replication, Document Base, Configuration, DTD ...
- TEXTML Localization
  - Translation, Build, Back-up ...
- Back-end Infrastructure
  - (DC, Storage, cluster, Servers, failover ...)

**Translation system**
- Translation exchange
  - (Hourly)
- Structure (Daily)
- Content (Daily, On demand)

**Delivery channels**
- Internal Previewe Server
- Nexus...
- ...

© 2016 SAP SE or an SAP affiliate company. All rights reserved.
DITA Production process
Project production lifecycle

- The **project map** defines the configuration of project’s outputs and publishing
  - The consistency and enablement of the project production lifecycle is defined in one place: **project map**
  - It’s a ticket free life-cycle: moving across the cycle is either self-service or automated
DITA build landscape

- Build lifecycle uses highly performant production framework with **parallel compilation**
- Project configuration information defined in project maps is:
  - transformed into properties passed to DITA Open Toolkit for producing branded SAP outputs (with SAP customizations)
  - used for automated publishing to delivery channels (e.g. Nexus, UA Content Platform…)

- 2016 SAP SE or an SAP affiliate company. All rights reserved.
- Continuous improvements in **DITA Output toolset**
  - Output formats (20+) and their customizable properties (100+) are **exposed** to Doc Leads through Project Map Editor
  - Integration of Open source components where applicable for better user experience
  - Toolset changes are **pushed** automatically to daily builds after passing acceptance tests → Quality assurance to +60k daily outputs

**DITA Output toolset**

- **Output formats & their customizations**
  - Customizable properties ...

  - SAP HP HTML Help
  - SAP DHTML Help
  - SAP HTML5
  - SAP PDF
  - UACP Product page
  - SAP Help Portal
  - SAP Eclipse help

**SAP DITA Open Toolkit Framework**

- Advanced rendering features e.g. datatable, interactive image maps
- Conditional processing: profiling, filtering & flagging
- Customized linking behaviors
- Language support, including right to left

**Open source components (JavaScript)**

- Jquery
- DataTable
- Pretify
- Lunr

**DITA Open Toolkit**

- DITA Open Toolkit 1.6 (migration to v2.3 in progress)

**Exposed to**

**Pushed to**

**Average daily #outputs**

<table>
<thead>
<tr>
<th>12</th>
<th>10</th>
<th>8</th>
<th>6</th>
<th>4</th>
<th>2</th>
<th>0</th>
</tr>
</thead>
</table>

**Project Map Editor**
Output-to-output (O2O) linking within documentation set

- Enables linking between outputs of the same collection (output format e.g. html5) within the same project or in the project dependency tree
- Implemented in daily Build (compile step)
Recent scaling activities (to accommodate growth)

- New architecture to support **project dependencies** in the CMS (project map) and in the build process
- New Document base split architecture (one document base for authoring, another for localization and production) with **replication process** to insulate authoring work from the performance impact of translation and production
- New **Delta Fetch** mechanism for faster build process
- New optimized build process with **Delta compilation** (only impacted outputs are re-built and published)

---

*Recent scaling activities (to accommodate growth)*

- 2016 SAP SE or an SAP affiliate company. All rights reserved.
Key takeaways and learnings

• Rely on leading industry solutions and Open Source technologies where applicable

• Large volume requires continuous optimization of the landscape

• It’s key to build an internal community of stakeholders and experts to foster innovation and internal support

• It’s important to have a regular upgrade strategy (e.g. DITA Open Toolkit) and engage with external community to take advantage of updates and recent enhancements
Thank you

Contact information:

Youssef Bennani
Head of Engineering Services France