Consortial Primo Installation at KOBV, Germany

Christian Hänger, Uni Mannheim
Andreas Sabisch, FU Berlin
Stefan Lohrum, KOBV-Zentrale

IGELU Helsinki 2009
Consortial Primo Installation at KOBV, Germany

- Partner
- KOBV hosting service
- Infrastructure
- Projekt
- Interface
- Primo consortial environment - Issues
- Primo consortial environment - Benefits
University Libraries of

- FU Berlin (32 000 FTE)
- HU Berlin, (32 000 FTE)
- TU Berlin, (25 000 FTE)
- Universität Mannheim (18 000 FTE)
- Universität Paderborn (25 000 FTE)
- Universität Düsseldorf (15 000 FTE)
- Requests from other Instituts
- Hosting by KOBV, Berlin
Hänger, Lohrum, Sabisch — Consortial Primo Installation at KOBV, Germany – IGELU 2009
Primo System Architecture
Sizing of the consortial Primo System

Requirements
- 6 institutions
- 800 simultaneous user
- 60 Mio Data sets, including 20 Mio sets from national licenses

Hardware

Test system
- 1 x Sun X4150, (4x4 CPU, 16 GB RAM),
- 2 TB SAN (Sun Array 6140)
- Virtual servers

Production system
- 7 x Sun X4150 (4x4 CPU, 16 GB RAM), 2 TB SAN
- separate Database sever (tbd.)
- 2 Front End Server (load balancing u. automatic failover, tbd.)
Infrastructure in ZIB

Magnetbandspeicher

Roboter Sun STK SL8500
Ein Roboter als Modular erweiterbares System

<table>
<thead>
<tr>
<th>Betriebsmodul + 5 Erweiterungen</th>
<th>Betriebsmodul + 1 Erweiterung</th>
<th>Insgesamt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stellplätze für Bänder</td>
<td>Laufwerke</td>
<td>13.500</td>
</tr>
<tr>
<td>16</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

Speicherkapazität
10 PB bei T9940B Bändern
40 PB bei T10000B Bändern

<table>
<thead>
<tr>
<th>Laufwerke</th>
<th>Typ</th>
<th>Kapazität eines Bandes (unkomprimiert)</th>
<th>Bandbreite</th>
</tr>
</thead>
<tbody>
<tr>
<td>bis zu 64 pro Robot</td>
<td>T9940B</td>
<td>200 GByte</td>
<td>30 MByte/s</td>
</tr>
<tr>
<td></td>
<td>T10000A</td>
<td>500 GByte</td>
<td>120 MByte/s</td>
</tr>
<tr>
<td></td>
<td>T10000B</td>
<td>1000 GByte</td>
<td>120 MByte/s</td>
</tr>
<tr>
<td></td>
<td>T10000C</td>
<td>2000 GByte</td>
<td>&gt;200 MByte/s</td>
</tr>
<tr>
<td></td>
<td>LTO und SDLT</td>
<td>möglich, aber zur Zeit nicht geplant</td>
<td></td>
</tr>
</tbody>
</table>
## KOBV - Monitoring

Filter: keywords = all

<table>
<thead>
<tr>
<th>Datum</th>
<th>Host</th>
<th>Instanz</th>
<th>Klasse</th>
<th>Agent</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.03.09, 05:50:02</td>
<td>w22</td>
<td>ZFL</td>
<td>oracle</td>
<td>ALERTLOG</td>
<td>Errors (old) found in alert_log file</td>
</tr>
<tr>
<td>17.03.09, 23:45:44</td>
<td>w13</td>
<td>ML4_Prod</td>
<td>oracle</td>
<td>ALERTLOG</td>
<td>Errors (old) found in alert_log file</td>
</tr>
<tr>
<td>17.03.09, 23:30:01</td>
<td>w22</td>
<td>ZFL</td>
<td>Z39</td>
<td>Z39</td>
<td>3 Server not OK. BARCH-facebook_kinemathek</td>
</tr>
<tr>
<td>18.03.09, 05:02:08</td>
<td>w08_kobv.de</td>
<td>Alphal8</td>
<td>filesystem</td>
<td>DISKFREE</td>
<td>1 filesystem(s) use more than 80 percent of capacity</td>
</tr>
<tr>
<td>18.03.09, 05:55:02</td>
<td>w13</td>
<td>ML4_Prod</td>
<td>system</td>
<td>LOAD</td>
<td>Average Load (last 5 min): 1.24</td>
</tr>
<tr>
<td>18.03.09, 05:55:02</td>
<td>w13</td>
<td>ML4_Prod</td>
<td>system</td>
<td>LOAD</td>
<td>Average Load (last 5 min): 0.19</td>
</tr>
<tr>
<td>18.03.09, 09:30:27</td>
<td>w07</td>
<td>ML4_Hz</td>
<td>system</td>
<td>DISKFREE</td>
<td>disk space seems to be ok</td>
</tr>
<tr>
<td>18.03.09, 05:55:02</td>
<td>w22</td>
<td>ZFL</td>
<td>system</td>
<td>LOAD</td>
<td>Average Load (last 5 min): 0.19</td>
</tr>
<tr>
<td>17.03.09, 23:30:00</td>
<td>w13</td>
<td>ML4_Prod</td>
<td>oracle</td>
<td>EXXTENTS</td>
<td>29 Tables with extents &gt; 10 % or size &gt; 50 MB found</td>
</tr>
<tr>
<td>18.03.09, 05:00:30</td>
<td>w08_kobv.de</td>
<td>Alphal8</td>
<td>oracle</td>
<td>EXXTENTS</td>
<td>157 Tables with extents &gt; 10 % or size &gt; 500 MB found</td>
</tr>
<tr>
<td>18.03.09, 05:36:00</td>
<td>w22</td>
<td>ZFL</td>
<td>process</td>
<td>dspsbbbf</td>
<td>dspsbbbf: 1 process(es) expected, 1 found</td>
</tr>
<tr>
<td>18.03.09, 02:31:00</td>
<td>w13</td>
<td>ML4_Prod</td>
<td>alphal</td>
<td>LICENSE</td>
<td>License is valid until 01.6.2013</td>
</tr>
<tr>
<td>17.03.09, 07:01:11</td>
<td>w22</td>
<td>ZFL</td>
<td>filesystem</td>
<td>DISKFREE</td>
<td>disk space seems to be ok</td>
</tr>
<tr>
<td>17.03.09, 13:25:16</td>
<td>w13</td>
<td>ML4_Prod</td>
<td>alphas</td>
<td>exprogs</td>
<td>all external programs in directory /clusters/metabiz/ml4/data/01/or_ext are valid</td>
</tr>
<tr>
<td>18.03.09, 05:45:03</td>
<td>w05_kobv.de</td>
<td>Alphal8</td>
<td>alphal</td>
<td>UE_11</td>
<td>All 1_ue_11 process(es) are alive</td>
</tr>
<tr>
<td>18.03.09, 05:45:02</td>
<td>w13</td>
<td>ZFL</td>
<td>system</td>
<td>TUNNELCTL</td>
<td>All routes are alive</td>
</tr>
<tr>
<td>18.03.09, 02:31:00</td>
<td>w13</td>
<td>ML4_Prod</td>
<td>alphas</td>
<td>SEQ</td>
<td>Count is 70633</td>
</tr>
<tr>
<td>18.03.09, 05:43:55</td>
<td>w08_kobv.de</td>
<td>Alphal8</td>
<td>alphas</td>
<td>UNION</td>
<td>b_unionp union is running, but no records waiting for update</td>
</tr>
<tr>
<td>18.03.09, 05:57:02</td>
<td>w22</td>
<td>ZFL</td>
<td>oracle</td>
<td>CONNECT</td>
<td>Connection OK</td>
</tr>
</tbody>
</table>

Hänger, Lohrum, Sabisch — Consortial Primo Installation at KOBV, Germany – IGELU 2009
Primo – Project plan

- Dez ‘08: Contract with ExLibris
- Feb/Mar ’09: Training in Primo
- Mar ’09: Start of implementation
- Mar ’09: Set up a Wiki for communication purpose
- Jun ’09: Adding Uni Mannheim
- Jul ’09: 2 day workshop with all members (and members in spe) for experience exchange
- Okt ’09: STP HU Berlin, Uni Mannheim
Hänger, Lohrum, Sabisch — Consortial Primo Installation at KOBV, Germany – IGELU 2009

Catalogue.PLUS
Including:
• OPAC
• Journals
• All Online Resources
Our Experience

- Uni Mannheim had experience on an local installation since begin of 2008
- FU, TU and HU started testing after implementation in April ‘09
- Düsseldorf and Paderborn have no experience yet
- We shared our knowledge in two workshops
- We shared the normalization rules as templates to use the local 'flavour' to change these rules
- We organize a process to share upgrades in the templates after testing and updating the local normalization rules manually
- We share collective data, for example data from our national license, in each instance, i.e. the data from one pipe is included in all views
Actual work

All:
- Training stuff
- Create and test pipes
- HU, UM, FU: Create Views, Layout
- Create and improve normalisation rules and mappings
- Check system maintenance, limits and possibilities
- Performance tests

Sources/pipes to involved:
- Aleph with MAB (German Format)
- ILS OCLC Sisis (planned)
- MetaLib as an separate tab
- SFX (in discussion)
- Institutional Repositories via OAI; different systems: Home grown, Image Catalogues, Museum catalogues, MyCoRe, OPUS…
- National Licenses Data (XX)
- Source from Publishers ()
- Sum: Pipes in consortia: Views in consortia
Pipes for sources in the consortia

- During tests created one pipe for delete and one for reload; a regular pipe for productive purpose
- One Aleph MAB normalization rule template, 4 Pipes (diff. flavours of the institutions)
- 8 OAI-Pipes; Institutional Repositories, Special and Image Catalogues
- 16 Pipes (diff. Marc Data) for National Licenses, usable for all Instances
- Additional pipes for sources in planning
  - MetaLib Data
  - SFX Data
  - Special Catalogues (i.e. Museums, Central scientific libraries, Sisis)
  - Data from Vendors (i.e. Oxford, de Gruyter, EBSCO)
- Checking the pipes/normalisation rules in El Commons
Back office issues in consortial environments: workarounds

- No instance separation in BO, no access control
  - Everybody can see and use other BO-stuff – workaround: don't touch others rules, tables..., only copy is allowed, use this own copy
  - List in Pipes, Tables, view etc. Very long, own material hard to find - workaround: indicate each name with the institute names, i.e. FUB_...

- Files for the scope with terms and notation are effective for all instances in one languages. workaround: use different 'languages' for each instance ;-) ; you can create your own language file

- No prioritization of pipes and indexing (FIFO), this blocks the other processes until it is ready – workaround: clarify via mail which pipe should start first

- Hanging threads block the system – workaround: restart the system. This is not a good option for productive systems. Threads should delete separately (like a kill in unix)

- Deploy from staging to productive system – there is no tool yet - workaround: doing by hand?

- Search scope can only accept 255 character from all scope names – not enough for all our national licence data scopes – workaround: add different licences to one scope by normalisation from pnx and reindexing for a new scope
Other issues in consortial environment without workarounds (but open SIs)

- There is only one base-URL definition for a catalog in an instance – we can not mix three different catalogue information to one 'multiple' view (Fixed now)
- PDS: one PDS-server for all instances and an authenticity provider per institution is a bottle neck. If this server stops, there is no login in Primo and Aleph and MetaLib in all institutions as well
- Some of the lists in the BO are not alphabetically ordered. It is hard to find the right thing in our long lists.
- There is only one synonym list for the instance. This list should be one per Scope/view to help for specialized views. By the way: The German synonym list is bad and was corrected by UB Mannheim. Will be available via El Commons.
- Not only consortial issues:
  - The normalisation rules don’t read 'normal' xml-files, i.e. with a regular tree structure. Will be solved in V3
  - The publishing process in Aleph is not consistent and loses data. Fixed now by a hot fix??
Benefits of the consortia

- Sharing the experience - one of the colleagues knows mostly the answer for your question or share templates for the start
- We will share our experience via El Commons as well
- Sharing normalization rules, language data etc. - in our consortia we are in a similar (German) environment with similar sources, formats etc.
- Scopes are separated by institutions, but common data (national licenses) can be shared in the search scopes
- Consortia reduce the costs for hardware and maintenance (50%)
- New instances can be added and start later on – the consortia can start with an initial group and grow up over the time with new members
Work in progress ...

Contacts:

Christian Hänger <christian.haenger@bib.uni-mannheim.de>
Stefan Lohrum <lohrum@zib.de>
Andreas Sabisch <sabisch@ub.fu-berlin.de>

Mailingliste:
<primo@ub.fu-berlin.de>