Mail environment management and retention policies

Geert Tilborghs (KBC ICT)
Keno Torfs (BASF IT Services)
Agenda

- Introduction BASF IT Services @KBC
- Importance of Email?
- Information Lifecycle Management @KBC
- Email Management Policy
- EMP @ BASF
- EMP @ KBC
  - Concept
  - Implementation
  - Issues
  - Results
  - Next steps
BASF IT Services
Facts and Figures

- Wholly owned subsidiary of BASF founded in April 2001
- Largest IT service provider of the BASF Group
- Among the leading IT service providers for the process industry in Europe
- Sales 2009: 360 million Euros
- Around 2,300 employees (Dec. 31, 2009)
- Headquarters in Ludwigshafen/Rhein (Germany)
- Locations in 10 countries across Europe
- SAP Service Partner and SAP Special Expertise Partner
- IT Service Management according to ITIL
BASF IT Services & Collaboration

Our Focus

For everyone, from anywhere, at anytime, we offer reliable, high-performance Messaging & Collaboration Services

Service Offering

- Messaging Services Management
- Consultancy and implementation of messaging solutions, policies and mail archiving, based on market standards (Lotus Notes, MS Exchange)
- Collaboration solutions
  - Instant Messaging & Web Conferencing Systems (Sametime)
  - Social networking tools (Connections)
  - Groupware solutions (Quickr)
- Standard collaboration tool and customised specific solutions
- Incident-, Problem- and Change Management according ITIL
Cooperation between BASF IT Services & KBC ICT
- Started in 2006
- Using datacenters of KBC ICT + HW & OS services provided by KBC ICT
- Messaging application layer + SMTP/spam services provided by BASF IT Services

BASF IT Services delivering Messaging services for KBC
- 20,000 Domino / Lotus Notes users
- 8,000 MS Exchange users
- 725 Blackberry users
- Intranet running on Domino
- Productive operation of Quickr & Sametime will start in Q2 2010

Adapted toolset
- Integrated / automated user management for Domino & Exchange
Importance of email?

- Mailboxes sizes are always increasing
  - Besides important information, people get a lot of junk information in the mailbox
  - Users do not have or do not take the time to keep their mailbox in good condition

- Mailboxes contain personal data
  - Mailboxes are in most cases not organized cabinets with only relevant data
  - Important information should not be kept in the mailbox, but should be moved to the correct document repository

  ➔ result is that in most cases mass of mail could be seen as non-important, obsolete data

  ➔ It is necessary for a company to have an Information Lifecycle Management!

  ➔ Email Management Policy is needed to keep your systems in a good shape!
Information Lifecycle Management

Problem Description

- The management of all information stored in mailboxes and on file server (O-/R-/… drives) is depending on the individual employee.

- In fact, most of this information is NOT MANAGED at all:
  - From most information we do not know:
    - who is responsible
    - who is allowed to change / consult the information
    - what is the importance of the content
    - what is the kind of content
    - if it is the most recent version
    - if the content is still valid
    - when the information can and must (!!) be removed (destroyed)
Information Lifecycle Management
Consequences

- The consequences are
  - rising costs – all information is concerned by ICT as if it is critical for the company
  - lower efficiency – employees are losing time while searching for information
  - lower performance
  - rising risk to lose information – current ICT systems are reaching their limits, e.g.
    backups are running too long

- This is getting worse year by year ....
Information Lifecycle Management
Definitions

<table>
<thead>
<tr>
<th>Business Process Information</th>
<th>Structured data</th>
<th>Non structured content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution</td>
<td>Transactional system</td>
<td>Content Management solution</td>
</tr>
<tr>
<td>Implementation</td>
<td>Database, back-end, …</td>
<td>Content Management Group framework (CGF)</td>
</tr>
</tbody>
</table>

What is structured data?
- A **data structure** is a way of storing data into a computer (typically in a DBMS) so that it can be used efficiently. Can be typical easily processed by a machine.
- Structured data is managed by technology that allows for querying and reporting against predetermined **data types and understood relationships**.

What is unstructured data / rich media?
- refers to masses of electronic information which do **NOT** have a data structure **easy readable by a machine**
- which have **NO data type definition** – in textual documents a word is simply a word
- examples are audio, video, unstructured text like the body of an e-mail, a spreadsheet and other types of documents.
Content Management

Spotmap

<table>
<thead>
<tr>
<th>CAPTURE</th>
<th>IMPORT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPTURE</td>
<td>UPLOAD</td>
</tr>
<tr>
<td>Scanning of paper documents</td>
<td>Human/manual upload</td>
</tr>
<tr>
<td>Capturing from fax and e-mail channels</td>
<td>Application / automatic upload</td>
</tr>
<tr>
<td>RELEASE</td>
<td>AUTHOR</td>
</tr>
<tr>
<td>VALIDATE &amp; INDEX</td>
<td>CREATE</td>
</tr>
<tr>
<td>Manual indexing</td>
<td>Authoring in / outside tool, template support and reuse of text blocks</td>
</tr>
<tr>
<td>Generated indexing using barcode/OCR/ICR techniques</td>
<td></td>
</tr>
<tr>
<td>Automated indexing</td>
<td></td>
</tr>
<tr>
<td>MANAGE</td>
<td>TRANSLATE</td>
</tr>
<tr>
<td>CLASSIFY</td>
<td>content</td>
</tr>
<tr>
<td>Indexing of content</td>
<td>Automatic / manual translation</td>
</tr>
<tr>
<td>Categorization (based on taxonomy)</td>
<td>Translation workflow</td>
</tr>
<tr>
<td>Manual, automated allocation to business file</td>
<td></td>
</tr>
<tr>
<td>LIBRARY SERVICES</td>
<td>TEST authoring</td>
</tr>
<tr>
<td>Check-in / check-out</td>
<td>content</td>
</tr>
<tr>
<td>Version management</td>
<td>Preview content</td>
</tr>
<tr>
<td>Edit, copy, delete document and folders</td>
<td>Test in acceptance disclosure environment</td>
</tr>
<tr>
<td>Audittrail</td>
<td></td>
</tr>
<tr>
<td>WORKFLOW MANAGEMENT</td>
<td>ACCESS &amp; DELIVER</td>
</tr>
<tr>
<td>Document routing</td>
<td>SEARCH &amp; RETRIEVE</td>
</tr>
<tr>
<td>REPORT</td>
<td>Retrieval using reference metadata search, full-text search, thesaurus</td>
</tr>
<tr>
<td>STORE &amp; PRESERVE</td>
<td>TRANSFORM</td>
</tr>
<tr>
<td>STORE</td>
<td>Render to HTML or PDF</td>
</tr>
<tr>
<td>Store on file system</td>
<td></td>
</tr>
<tr>
<td>Store in database</td>
<td></td>
</tr>
<tr>
<td>Store in content repository</td>
<td></td>
</tr>
<tr>
<td>ARCHIVE</td>
<td>PUBLISH</td>
</tr>
<tr>
<td>Based on trigger, contenttype, physical location</td>
<td>For use in core / non-core workplace</td>
</tr>
<tr>
<td>Retrieve, restore, search archive</td>
<td></td>
</tr>
</tbody>
</table>
Managing Unstructured data @ KBC

Scenario 3: Starting from an offering

We made an offering of 3 different worlds:
- Content Managed by a business Application – A-ILM
- Not managed content (Basic services) – B-ILM
- Categorized content (taxonomy/context) – C-ILM

Our approach
- We believe 3 different solutions is enough
- We believe we can make/define those 3 worlds different enough to avoid ‘grey-discussions’
- Approach: define and describe the services in those 3 different worlds
- Cases choose for services and by that for one of the 3 solutions

→ Chosen scenario
Cost aspects related to mailbox size

- Mailbox size
- Clustering
- Disk category
- Backup Size
- Data not found
- User Efficiency
- Server performance-Consoldation
- Backup frequency
- Server performance-Consoldation
- BLUG

Blux Lotus User Group
Email Management Policy (EMP)

Information Lifecycle Management should
- Decide what is important data
- Define the lifecycle of information in general
- Provide the user with the correct tools to store important documents.

An Email Management Policy should
- Be aligned to the company ILM
- Give the user the chance to Export documents to Archive or other systems
- Manage all other data in the mailbox
- Remind users to keep information where it belongs
- To keep the mailing system healthy and performant
For the BASF group already in 2003 an EMP concept has been worked out.

Main concept:
- Non-important mails older than 6 months are automatically removed
- Important mails have to be ‘tagged’ and will be archived in Commonstore
  - Mandatory retention
    - Legal implications $\rightarrow$ fixed retention period per document type
  - Dedicated user group
- Worthy for Retention
  - Available to all users
  - Choice to keep 1-3 years in Archive
BASF ‘EMP’ Project

- Tag for Archiving
  - Documents not tagged, after 6 months
  - Full documents or only attachments
Mark for Archiving

BASF DELETION POLICY

Tagging of documents is performed by each individual user. To archive an email; open the message (or messages) to be archived. From the menu button, select the “archiving” button and “Set Tagging”. Enter the archiving information such as archive type and comments, and enter the amount of time you’d like to store the document. Click “OK” to confirm and archive the document.

Here is the screenshot for the settings you need to make for sametime. Just make sure use token b...
Deletion policy@ BASF

Yes, the E-mail Management procedure applies for the entire BASF Corp and does not allow any exceptions.

Q: Am I required to conform to the BASF Email Management policy?

Q: What are the main points of the E-mail Management procedure?

Q: Can I keep a copy of my e-mail messages on my own PC?

Q: What kinds of documents are subject to obligatory retention?

Q: What will happen with my calendar documents?

Q: Will my Administrative Assistant or the person covering for me during vacation have access to my moved/archived attachments?

Q: My colleague has been out sick for several weeks now and will not be back at his/her workstation for quite awhile. What can be done to make sure that they don’t lose any important e-mail?

Q: I have tagged documents for retention and the retention period has expired. I still need to retain these documents. What do I do to prevent their deletion?

Q: What happens if I delete a “Tagged” e-mail?
KBC ‘EMP’ Project

- KBC had decided to implement an EMP solution based on the same principles as the solution implemented for the BASF group.
- Main focus was to reduce the cost and increase performance (user & system).
- Would be implemented together with the R8 upgrade project (joint development / communication benefits).
- Completely Domino solution.
- Primary target:
  - 17,000 Domino users - mailboxes.
KBC ‘EMP’ Project

Mailbox

Tag for Archiving

Documents not tagged, after 6 months

Archive on Domino server

Soft deletion repository

1 month later

BLUG
Belux Lotus User Group
Mailbox versus Archive

- MailEnvironment
  - Clustering
  - Dynamic Environment

- ArchiveEnvironment
  - No Clustering
  - Static Environment
Tiering options & timeline overview

Initial scope

Documents less than 6 months old

- Default, normal user
  - All mails marked for Archiving will be moved to the Archive
  - All mails not marked for Archiving will be removed from the mailbox

Documents between 6 months and 42 months

- User with QUOTA
  - All mails not marked for archiving are kept in productive mailbox / not participating in deletion
  - All mails marked for Archiving will be moved to the Archive

- User with quo010 'Langdurig zieken'
  - All mails not marked for archiving are kept in productive mailbox / not participating in deletion
  - All mails marked for Archiving will be moved to the Archive

Documents older than 42 months

- Mails removed from Archive

Timeline

T0

T1 = T0 + 6 months

T2 = T1 + 3 years
# Project plan

## Overview

<table>
<thead>
<tr>
<th>Task</th>
<th>Jan 09</th>
<th>Feb 09</th>
<th>Mar 09</th>
<th>Apr 09</th>
<th>Mei 09</th>
<th>Jun 09</th>
<th>Jul 09</th>
<th>Aug 09</th>
<th>Sep 09</th>
<th>Oct 09</th>
<th>Nov 09</th>
<th>Dec 09</th>
<th>Jan 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOA upgrade &amp; preparations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mailtemplate development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archiving Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Production server upgrades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application &amp; iNet upgrades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Client upgrade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mailtemplate upgrade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archiving / deletion impl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Close programme</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*End R8*

*End Archiving*

*Userimpact*
**KBC Tiering**

**Enhanced scope**

<table>
<thead>
<tr>
<th>Default, normal user</th>
<th>All mails marked for Archiving will be moved to the Archive</th>
<th>Mails removed from Archive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All mails not marked for Archiving will be removed from the mailbox</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User with QUOTA</th>
<th>All mails not marked for archiving are kept in productive mailbox / not participating in deletion</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All mails marked for Archiving will be moved to the Archive</td>
<td>Mails removed from Archive</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User with quo010</th>
<th>'Langdurig zieken'</th>
<th>All mails not marked for archiving are kept in productive mailbox / not participating in deletion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All mails marked for Archiving will be moved to the Archive</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User with quo020</th>
<th>'privileged users – No Deletion'</th>
<th>All mails not marked for archiving are kept in productive mailbox / not participating in deletion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All mails marked for Archiving will be moved to the Archive</td>
</tr>
</tbody>
</table>

| User with quo030      | 'Privileged users – Archive All' | All mails marked automatically and moved to the Archive | Mails removed from Archive |

**Timeline**

- **T0**
- **T1 = T0 + 6 months**
- **T2 = T1 + 3 years**

---

Due to ‘complaints’ new options introduced for privileged users.
Mark for Archiving

- Users have to tag the required documents to mark them for Archiving

- Users can only tag 1 document at a time!
  - To force users to think twice
  - To not just move ‘garbage’ to another location

![Screen shot showing the archive function with error message: You are not allowed to archive multiple documents at a time.](image)
Overview in mailbox / Archive

Folder structure in Central Archive (is permanent, not possible to move mails!)

Please note

The folder structure in the Central Archive:
- has a permanent structure (is taken from your active mailbox)
- you cannot delete folders
- you cannot create folders
## Mail management archiving: automatic monthly actions by the system + timings

<table>
<thead>
<tr>
<th>will be deleted</th>
<th>Example:</th>
</tr>
</thead>
<tbody>
<tr>
<td>All mails (older than 6 months) in your mailbox that are NOT marked for archiving and are therefore mentioned in the status overview under 'will be deleted' – in (current) month, are deleted definitively and irrevocably.</td>
<td><img src="image1.png" alt="Example Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>will be archived</th>
<th>Example:</th>
</tr>
</thead>
<tbody>
<tr>
<td>All mails (older than 6 months) in your mailbox that are marked for archiving and are therefore mentioned in the status overview under 'will be archived' – in (current) month, are moved to your central archive.</td>
<td><img src="image2.png" alt="Example Image" /></td>
</tr>
</tbody>
</table>

If the marked mail is included in one or more folders, the same folder is created during the archiving in the central archive file and the archived message will be placed there.
Old Archives?

Where do I find my 'old' (last used) archives?

<table>
<thead>
<tr>
<th>Step</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>In the menu, select 'Actions - Get old archive location'</td>
</tr>
</tbody>
</table>

Result:

A menu appears where you will find the path and the name of your archives. The location depends on the choice you made when creating your archive folder and archive name.

Example:

```
Location of the old archive

Location of the old archive
W:\\Wanaerschot\Archief.nsf
```

3. Note down this name and click **OK**. You can also use the key combination 'Ctrl + C' (copy) instead of noting down this name.
Issues

- Users complained about upcoming deletion and forced a delay in the schedule
  - Cause: users were afraid to lose data
  - Solution: internal KBC discussion at management level and introduction of ‘Archive All (at start) user’

- User can not print / copy / forward from within Archive
  - Cause: users do not have the required role in ACL of Archive (side-effect of initially agreed setting)
  - Solution: On purpose \(\Rightarrow\) Works as designed. Helpdesk is informed about this behavior.

- Local Archiving still possible via bypass
  - Cause: Local archiving option is removed in the KBC R8 template, however it is still possible for users to access the settings via the standard Notes properties page
  - Solution: Local archiving disabled via explicit policies

- Not possible to mark documents for archiving on blackberry
  - Cause: was not in scope of project. Blackberry using propriety software
  - Solution: Solution will have to be built manually

- For 23 suspended users with QUO010 the deletion has ran in their mailbox
  - Cause: For suspended users the QUO010 status is not transferred from CRD to the NAB
  - Solution: mailboxes restored ; sync mechanism being updated
## Tiering Results

### Simulation run
- No data deletion

### Data Table

<table>
<thead>
<tr>
<th>Action</th>
<th>Pre-pilot</th>
<th>Pilot</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Pilot-Delete</th>
<th>Delete-delayed</th>
<th>Jan</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archive</td>
<td>7.874</td>
<td>72.150</td>
<td>255.750</td>
<td>406.259</td>
<td>1,071.394</td>
<td>1,929.220</td>
<td>255.750</td>
<td>406.259</td>
<td>1,071.394</td>
</tr>
<tr>
<td></td>
<td>0.79</td>
<td>7.16</td>
<td>24.45</td>
<td>51.44</td>
<td>172.48</td>
<td>1,929.220</td>
<td>1,929.220</td>
<td>1,929.220</td>
<td>1,929.220</td>
</tr>
<tr>
<td></td>
<td>43</td>
<td>743</td>
<td>3.676</td>
<td>5.008</td>
<td>5.449</td>
<td>1,929.220</td>
<td>1,929.220</td>
<td>1,929.220</td>
<td>1,929.220</td>
</tr>
</tbody>
</table>

### Deletion

<table>
<thead>
<tr>
<th>Action</th>
<th>Pre-pilot</th>
<th>Pilot</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Pilot-Delete</th>
<th>Delete-delayed</th>
<th>Jan</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archive</td>
<td>7.951</td>
<td>72.150</td>
<td>255.750</td>
<td>406.259</td>
<td>1,071.394</td>
<td>1,929.220</td>
<td>255.750</td>
<td>406.259</td>
<td>1,071.394</td>
</tr>
<tr>
<td></td>
<td>0.36</td>
<td>7.16</td>
<td>24.45</td>
<td>51.44</td>
<td>172.48</td>
<td>1,929.220</td>
<td>1,929.220</td>
<td>1,929.220</td>
<td>1,929.220</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>743</td>
<td>3.676</td>
<td>5.008</td>
<td>5.449</td>
<td>1,929.220</td>
<td>1,929.220</td>
<td>1,929.220</td>
<td>1,929.220</td>
</tr>
</tbody>
</table>

### Grand Total

- 1,317,440
- 1,349,824
- 18,618,000
- 18,618,000
- 21,737

---

### Diagram

- **Simulation run**: No data deletion
- **Run**: 7.874, 0.79, 43
- **Data**: 24.237, 2.51, 369, 47.013, 4.65, 303
- **Archive Total**: 255.750, 24.45, 3.676, 466.259, 51.44, 5.008, 1,929.220
- **Deletion Total**: 2,859,974, 276.46, 2.726, 1,066,367, 144.01, 1.659, 8,613,940, 478.92, 12,423
Mailboxes actively involved in Tiering

<table>
<thead>
<tr>
<th></th>
<th>Pilot</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Pilot-Delete</th>
<th>Delete-delayed</th>
<th>Jan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archive</td>
<td>743</td>
<td>3.676</td>
<td>5.008</td>
<td>5.449</td>
<td></td>
<td></td>
<td>6.818</td>
</tr>
<tr>
<td>Deletion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>△ Arch %</td>
<td>28%</td>
<td>23%</td>
<td>32%</td>
<td>35%</td>
<td></td>
<td></td>
<td>44%</td>
</tr>
<tr>
<td>• Del %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>79%</td>
</tr>
</tbody>
</table>

Simulation run with no data deletion
Comparison tiering versus pre-investigation (June 2007)

### Amount of Mails in mail files

- < 3 Months: 39%
- 3-6 Months: 6%
- 6-9 Months: 22%
- 9-12 Months: 17%
- 12-24 Months: 24%
- 24-36 Months: 13%
- > 36 Months: 9%

### Diskspace occupied by mails

- < 3 Months: 24%
- 3-6 Months: 19%
- 6-9 Months: 14%
- 9-12 Months: 22%
- 12-24 Months: 19%
- 24-36 Months: 31%
- > 36 Months: 4%

### Table: Run Comparison

<table>
<thead>
<tr>
<th>Action</th>
<th>Data</th>
<th>Pilot + delayed Delete</th>
<th>Jan</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion</td>
<td>#Docs</td>
<td>11.209.874</td>
<td>8.013.940</td>
<td>19.223.814</td>
</tr>
<tr>
<td></td>
<td>GBytes</td>
<td>976.50</td>
<td>978.92</td>
<td>1.955.42</td>
</tr>
<tr>
<td></td>
<td>#Db</td>
<td>11.368</td>
<td>12.400</td>
<td>23.768</td>
</tr>
</tbody>
</table>
Disk space Usage by Server

Diskspace usage by server

- **Design & Data Compression**
- **Deletion policy**
- **Users requested to clean mailbox**
Total disk space – All servers

<table>
<thead>
<tr>
<th>Date</th>
<th>Home</th>
<th>Cluster</th>
<th>Archive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/02/09</td>
<td>3.485</td>
<td>3.324</td>
<td>53</td>
</tr>
<tr>
<td>1/03/09</td>
<td>3.593</td>
<td>3.437</td>
<td>87</td>
</tr>
<tr>
<td>1/04/09</td>
<td>3.743</td>
<td>3.636</td>
<td>131</td>
</tr>
<tr>
<td>1/05/09</td>
<td>3.951</td>
<td>3.863</td>
<td>282</td>
</tr>
<tr>
<td>1/06/09</td>
<td>3.529</td>
<td>3.346</td>
<td>1.161</td>
</tr>
<tr>
<td>1/07/09</td>
<td>3.690</td>
<td>3.538</td>
<td></td>
</tr>
<tr>
<td>1/08/09</td>
<td>3.605</td>
<td>3.530</td>
<td></td>
</tr>
<tr>
<td>1/09/09</td>
<td>3.578</td>
<td>3.410</td>
<td></td>
</tr>
<tr>
<td>1/10/09</td>
<td>3.695</td>
<td>3.592</td>
<td></td>
</tr>
<tr>
<td>1/11/09</td>
<td>3.863</td>
<td>3.739</td>
<td></td>
</tr>
<tr>
<td>1/12/09</td>
<td>4.019</td>
<td>3.815</td>
<td></td>
</tr>
<tr>
<td>1/01/10</td>
<td>4.096</td>
<td>3.875</td>
<td></td>
</tr>
<tr>
<td>1/02/10</td>
<td>2.437</td>
<td>2.208</td>
<td></td>
</tr>
</tbody>
</table>

Including ‘soft’ deletion repository
Conclusions

- **Acceptance of Archiving / deletion policy by the users**
  - However deletion was initially postponed, no further delay's or problems have occurred
  - Strict Management and very thorough communication plan is needed
  - No functional user complaints have reached BASF IT Services

- **File systems**
  - Forecasting for the needed infrastructure was correct
  - Clean-up of file systems should be considered

- **Participation**
  - Archiving participation can still be considered relatively low: ~45%
  - Deletion participation is realistic with ~80%
Further disk space optimization by DAOS?

- DAOS will centralize attachments above a minimum size (standard 4KB, IBM advises 64KB)
- In case we assume ‘minimum size of object’ is 100KB
  - 10% of documents will be in scope
  - 88% of size is in scope
DAOS Estimator

Size Distribution of All Attachments Found

Diskspace Occupied by Size

Summary:
- Total DB's analyzed: 3646
- Total DB's skipped due to errors: 0
- Total Size of NSF's Examined: 1688203
- Total Attachments found: 762677
- Estimated size of DAOSified NSF's: 171.6 GB
- Estimate size of DAOS dir: 225.0 GB
- Total Disk Savings: 158.7 GB

DAOS Minimum Size versus number of NLO's and Disk Space:

- 0.0 KB will result in 1688203 .nlo files using 225.0 GB
- 4.0 KB will result in 177609 .nlo files using 224.9 GB
- 8.0 KB will result in 1443902 .nlo files using 224.5 GB
- 16.0 KB will result in 1228594 .nlo files using 223.3 GB
- 32.0 KB will result in 989142 .nlo files using 220.5 GB
- 64.0 KB will result in 708280 .nlo files using 213.4 GB
- 1.0 MB will result in 87594 .nlo files using 128.6 GB
- 2.0 MB will result in 6697 .nlo files using 34.0 GB
- 10.0 MB will result in 12 .nlo files using 4.4 GB
- 100.0 MB will result in 8 .nlo files using 1.3 GB
Long term Archiving

Due to business requests and needs, a follow up project will/should define how to treat the documents after the time span of 3 years and define possibilities to archive longer in case needed.

**Timeline**

- **T0**
- **T1** = T0 + 6 months
- **T2** = T1 + 3 years

**Default, normal user**
- All mails marked for Archiving will be moved to the Archive
- All mails not marked for Archiving will be removed from the mailbox

**User with QUOTA**
- All mails not marked for archiving are kept in productive mailbox / not participating in deletion
- All mails marked for Archiving will be moved to the Archive

**User with quo010**
- ‘Langdurig zieken’
- All mails not marked for archiving are kept in productive mailbox / not participating in deletion
- All mails marked for Archiving will be moved to the Archive

**User with quo020**
- ‘privileged users – No Deletion’
- All mails not marked for archiving are kept in productive mailbox / not participating in deletion
- All mails marked for Archiving will be moved to the Archive

**User with quo030**
- ‘Privileged users – Archive All’
- All mails marked automatically and moved to the Archive
- Mails removed from Archive
Questions and discussion
Thank you for your attention.

Contact

Geert Tilborghs
Business Analyst
KBC ICT
Bedrijvenpark 4
2800 Mechelen
Phone +32 15 35 21 85
Mobile +32 471 17 99 74
geert.tilborghs@kbc.be
www.kbc.be

Keno Torfs
Solution Architect / Project Manager
BASF IT Services
Terhulpensesteenweg 178
B-1170 Brussel
Phone +32 2 373-25.39
Fax +32 3 561-60.98
Mobile +32 477 78.79.09
keno.torfs@basf-it-services.com
www.basf-it-services.com