# 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
- Certified under ISO 9001:2000 and ISO 27001:2005
- 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



### BASF IT Services @ KBC

- 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 to long
- This is getting worse year by year ....



### Information Lifecycle Management

### **Definitions**

	Busines	s Process
	Infor	mation
	Structured data	Non structured content
Solution	Transactional system	Content Management solution
Implementation	Database, back-end,	Content Management Group framework (CGF)

### 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

### **CAPTURE**

#### **CAPTURE**

Scanning of paper documents Capturing from fax and e-mail channels

### RELEASE

#### VALIDATE & INDEX

Manual indexing Generated indexing using barcode/OCR/ICR techniques Automated indexing

MANAGE RELATIONSHIPS & LINKS

### MANAGE

### CLASSIFY

Indexing of content Categorization (based on taxonomy) Manual, automated allocation to business file

#### LIBRARY SERVICES

Check-in / check-out Version management Edit, copy, delete document and folders Audit trail

### WORKFLOW MANAGEMENT

Document routing

#### REPORT

Link authoring map to disclosure view Link content to disclosure view, files, other content Link content to files / to other content

### MANAGE ROLES & PRIVILEGES

Fine-grained or low grained access control Authentication

#### COLLABORATE

Co-reading, co-writing, annotations, reader feedback, mailing list, discussion forum
Team contacts, calendaring and mailbox
Phone & video, chat
Manage and search collaboration rooms
Event notification

### STORE & PRESERVE

#### STORE

Store on file system Store in database Store in content repository

#### **ARCHIVE**

Based on trigger, content type, physical location Retrieve, restore, search archive

#### IMPORT

#### UPLOAD

Human/manual upload Application/automatic upload

### AUTHOR

#### CREATE

Authoring in / outside tool, template support and reuse of text blocks

#### TRANSLATE content

Automatic/manual translation Translation workflow

### TEST authoring content

Preview content
Test in acceptance disclosure environment

### **ACCESS & DELIVER**

#### SEARCH & RETRIEVE

Navigate through content Retrieval using reference metadata search, full-text search, thesaurus

#### TRANSFORM

Render to HTML or PDF

#### **PUBLISH**

For use in core / non-core workplace



### 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

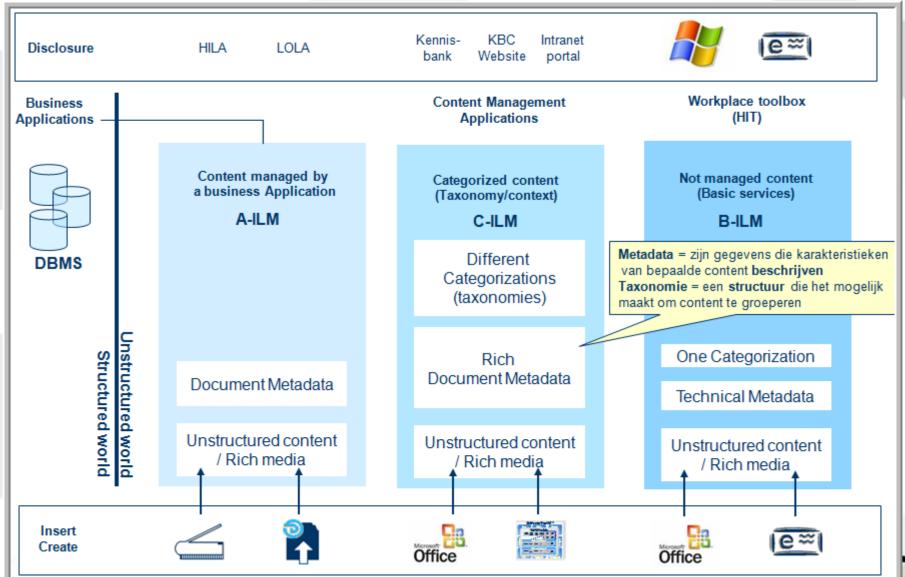
### Our approach

- 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



### Content Management

### Architecture





### Cost aspects related to mailbox size



# 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

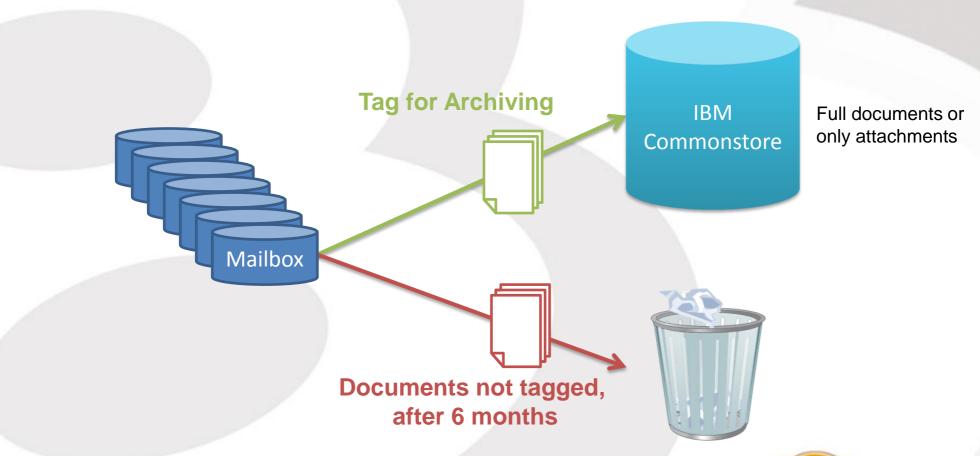


# BASF 'EMP' Project

- 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 → fixed retention period per document type
      - Dedicated user group
    - Worthy for Retention
      - Available to all users
      - Choice to keep 1-3 years in Archive

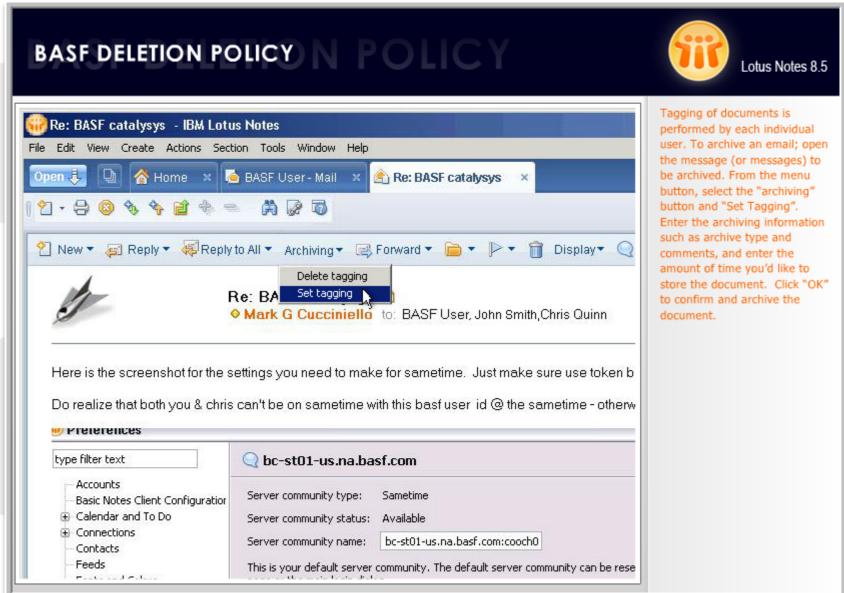


# BASF 'EMP' Project





# Mark for Archiving





# Deletion policy@ BASF

### BASF DELETION POLICY



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?

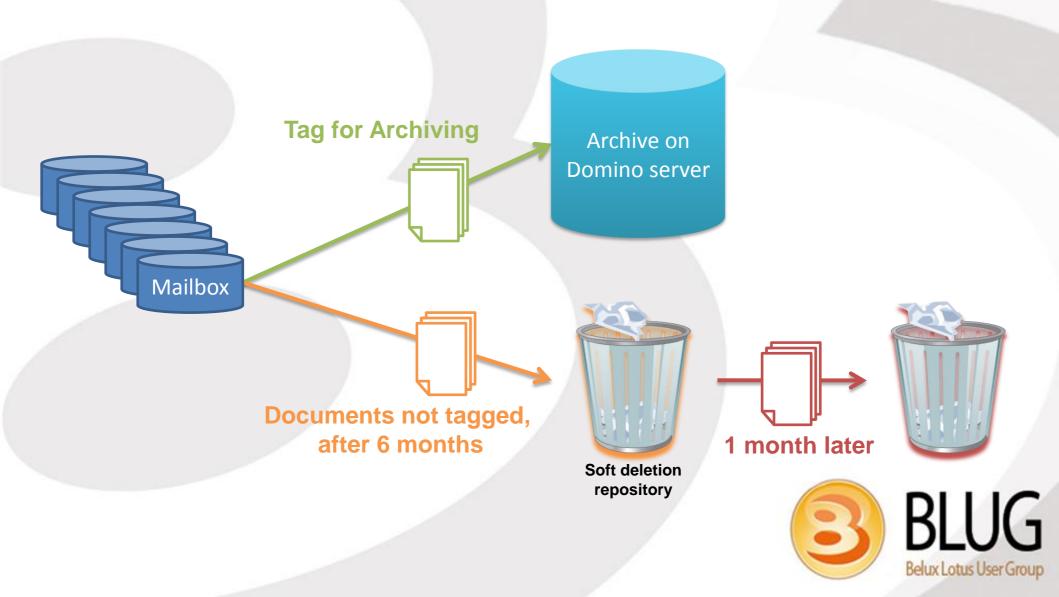


# KBC 'EMP' Project

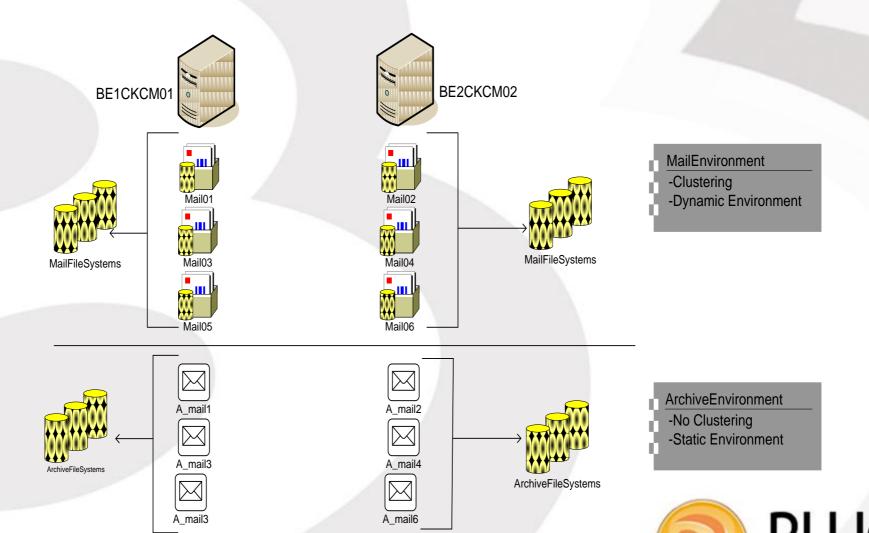
- KBC had decided to implement an EMP solution based on the same principles as the solution implemented for the BASF group
- Main focus was to reduce the cost and increase performance (user & system)
- Would be implemented together with the R8 upgrade project (joint development / communication benefits)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



# KBC 'EMP' Project



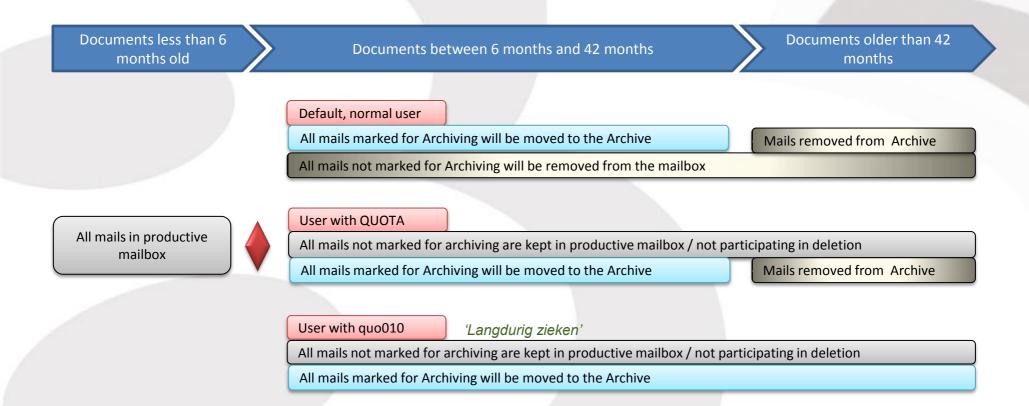
# Mailbox versus Archive



Belux Lotus User Group

# Tiering options & timeline overview

Initial scope











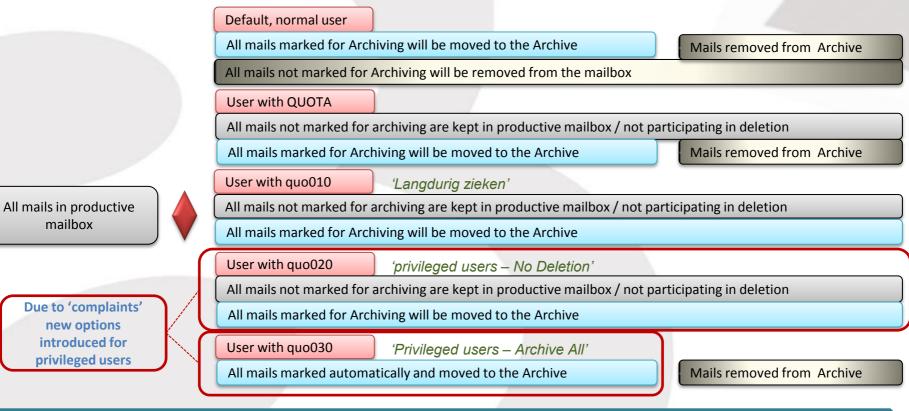
# Project plan

overview **End Archiving** End R8 apr 09 jan 10 jan 09 feb 09 mrt 09 mei 09 jun 09 jul 09 sep 09 okt 09 nov 09 dec 09 aug 09 TOA upgrade & preparations Mailtemplate development Application tests Archiving Development **Userimpact** Production server upgrades Application & iNet upgrades Client upgrade Mailtemplate upgrade Archiving / deletion impl. Close programme



# **KBC** Tiering

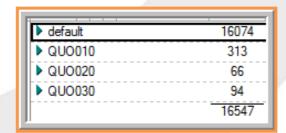
### Enhanced scope



### Timeline











# Mark for Archiving

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

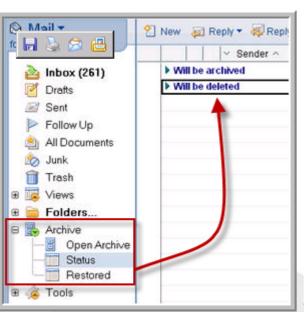


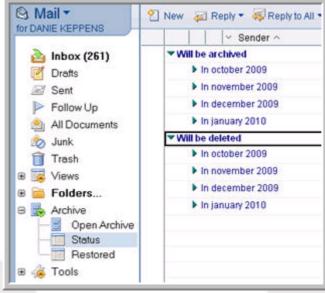
- Users can only tag 1 document at a time!
  - To force users to think twice
  - To not just move 'garbage' to another location

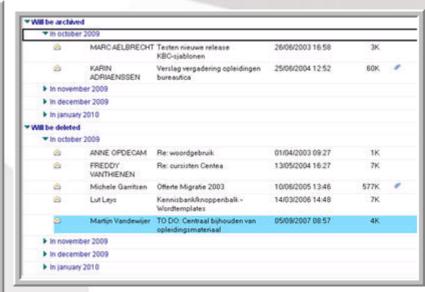




### Overview in mailbox / Archive







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



### Communication

### Kennisbank

Mail management archiving: automatic monthly actions by the system + timings

will be deleted

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.

#### Example: ▼ Will be deleted ▼ In october 2009 ANNE OPDECAM 01/04/2003 09:27 Re: woordgebruik FREDDY Re: cursisten Centea 13/05/2004 16:27 7K VANTHIENEN Michele Garritsen Offerte Migratie 2003 10/06/2005 13:46 577K 🧳 05/09/2007 08:57 Martin Vandewijer TO DO: Centraal bijhouden van opleidingsmateriaal

will be archived

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.

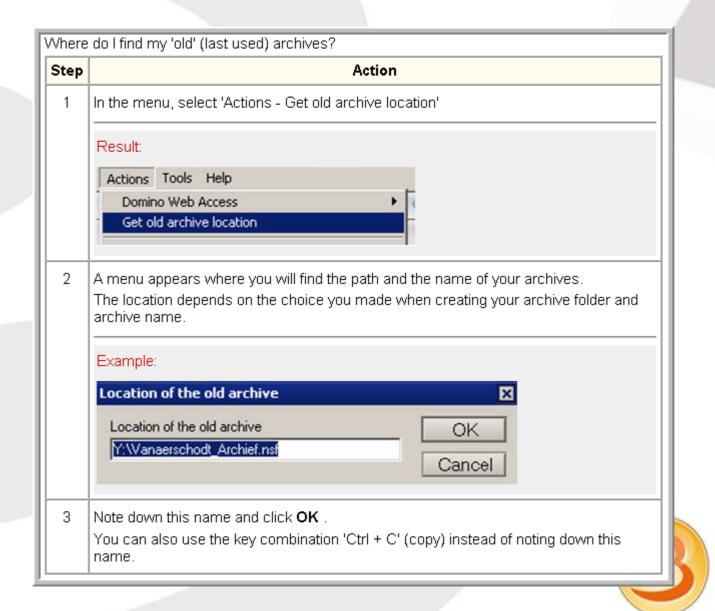


If the marked mail is included in one or more folders, the same folder is created during the archiving in the central archive file and the archived message will be placed there.



<del>02-00-2010 20.0</del>0

### Old Archives?





### Issues

- Users complained about upcoming deletion and forced a delay in the schedule
  - Cause: users were afraid to loose 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 → 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



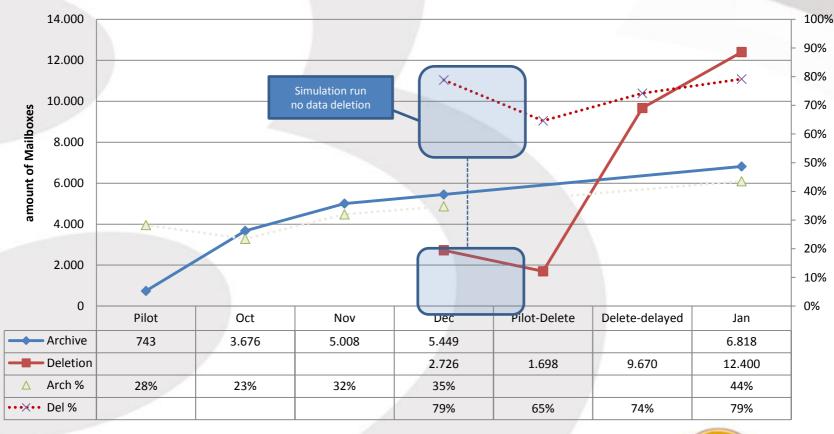
# Tiering Results

		run 🖓 [	Data																						
		Pre-pilot	Dutu	Pilot		Oct		Nov			Dec		F	Pilot-Delete			Delete-delaye	ed	J	an			Total #Docs	Total GBytes	Total #Db
Action	Server	#Docs (	GBytes #DI	b #Docs	GBytes #Db	#Docs	GBytes #Db	#Docs	GBytes	#Db	#Docs	GBytes #	Db #	Docs	GBytes #	Db i	#Docs C	GBytes #[	Ob #	Docs (	GBytes #	#Db		1	
■ Archive	BE1CKCM01/SERVERS	7.874	0,79 4	24.237	2,51 360	1.157	0,08	6 32.542	3,69	413	125.543	21,29	439							114.047	17,05	535	305.400	45,41	1.836
	BE1CKCM02/SERVERS					20.535	1,81 3	21.790		354	51.736	5,54	401							88.564	9,84	477	182.625	19,93	
	BE1CKCM03/SERVERS					47.362	3,96 34	8 29.077	3,35	390	77.251	11,18	405							76.586	11,47	492	230.276	29,96	1.635
	BE1CKCM04/SERVERS					14.649	1,26 27	0 33.532	3,53	322	109.033	21,04	344							86.005	9,38	445	243.219	35,21	1.381
	BE1CKCM05/SERVERS					16.650	1,39 32	7 21.365	2,45	358	89.063	8,37	394							58.100	8,69	495	185.178	20,91	1.574
	BE1CKCM06/SERVERS					7.050	1,22 17	2 13.702	2,77	264	50.814	10,15	298							53.282	9,93	367	124.848	24,07	1.101
	BE1CKCM07/SERVERS					16.928	1,15 20	1 36.020	5,03	371	54.256	8,06	432							104.939	17,79	535	212.143	32,04	1.539
	BE2CKCM01/SERVERS			47.913	4,65 383	2.496	0,80	8 35.289		427	77.441	16,90	454							100.651	16,28	537	263.790	43,22	1.849
	BE2CKCM02/SERVERS					27.634	2,84 37	5 37.343		373	34.955	4,54	409							75.764	10,23	527	175.696	21,25	1.684
	BE2CKCM03/SERVERS					17.826	1,40 37	7 25.591	3,52	398	77.821	13,45	437							80.783	12,75	551	202.021	31,13	1.763
	BE2CKCM04/SERVERS					20.442	1,89 34	0 25.605	3,24	355	72.878	9,16	380							66.012	10,61	495	184.937	24,89	1.570
	BE2CKCM05/SERVERS					34.349	3,00 3	4 36.295		396	108.163	18,55	417							198.447	27,09	530	377.254	53,60	1.717
	BE2CKCM06/SERVERS					16.338	2,39 20			289	62.063	10,54	320							94.102	16,46	403	198.183	32,96	1.277
	BE2CKCM07/SERVERS					12.334	-,	2 32.428	-,	298	80.377	13,71	319							120.158	20,92	429	245.297	40,25	1.228
Archive Total		7.874	0,79 4	72.150	7,16 743	255.750	24,45 3.67	6 406.259	51,44	5.008	1.071.394	172,48	5.449							1.317.440	198,49	6.818	3.130.867	454,82	21.737
■ Deletion	BE1CKCM01/SERVERS	7.951	0,36	31										862.242	70,53	877				598.441	62,68	894	1.468.634	133,57	1.802
	BE1CKCM02/SERVERS												_				768.590	62,00	767	497.807	57,62	855	1.266.397	119,62	1.622
	BE1CKCM03/SERVERS																1.016.491	77,75	812	583.877	65,59	867	1.600.368	143,34	1.679
	BE1CKCM04/SERVERS									- (							752.629	65,88	755	566.728	64,71	822	1.319.357	130,60	1.577
	BE1CKCM05/SERVERS									1							847.780	79,82	728	583.189	71,19	894	1.430.969	151,01	1.622
	BE1CKCM06/SERVERS						Simi	lation ru	ın		783.225	81,16	965				696.163	69,25	937	584.674	84,17	959	2.064.062	234,57	2.861
	BE1CKCM07/SERVERS																677.586	59,06	772	708.720	88,11	874	1.386.306	147,17	1.646
	BE2CKCM01/SERVERS						no da	ta delet	ion	L I	852.655	73,56	860	804.125	73,47	821				567.623	69,89	851	2.224.403	216,93	2.532
	BE2CKCM02/SERVERS							_									777.923	,	814	519.981	60,89	870	1.297.904	123,18	1.684
	BE2CKCM03/SERVERS																834.347	61,74	810	575.815	62,51	883	1.410.162	124,25	1.693
	BE2CKCM04/SERVERS																871.475	62,62	829	577.381	64,82	876	1.448.856	127,44	1.705
	BE2CKCM05/SERVERS										1.215.094	115,74	901				956.845	86,86	829	581.554	76,27	890	2.753.493	278,88	2.620
	BE2CKCM06/SERVERS																729.688	83,50	887	529.059	76,04	931	1.258.747	159,54	1.818
	BE2CKCM07/SERVERS	7.054									0.050.074	070.10	2 200				613.990	61,72	730	539.091	74,42	934	1.153.081	136,14	1.664
Deletion Total		7.951	0,36 3		7.40 740	055.750	04.45.0.0	0 400 050		5.000		270,46	2.726	1.666.367	144,01			832,49 9		8.013.940	,	12.400	22.082.739	2.226,24	26.525
Grand Total		15.825	1,15 7	74 72.150	7,16 743	255.750	24,45 3.67	6 406.259	51,44	5.008	3.922.368	442,95	8.1/5	1.666.367	144,01	1.698	9.543.507	832,49 9	1.670	9.331.380	1.177,41	19.218	25.213.606	2.681,06	48.262

		run	<b>\7</b>								
Action	▼ Data	F	Pre-pilot	Pilot	Oct	Nov	Dec	Pilot-Delete	Delete-delayed	Jan	<b>Grand Tota</b>
Archive	#Docs		7.874	72.150	255.750	406.259	1.071.394			1.317.440	3.130.86
	GBytes		0,79	7,16	24,45	51,44	172,48			198,49	454,8
	#Db		43	743	3.676	5.008	5.449			6.818	21.73
Deletion	#Docs		7.951					1.666.367	9.543.507	8.013.940	19.231.76
	GBytes		0,36					144,01	832,49	978,92	1.955,7
	#Db		31					1.698	9.670	12.400	23.79

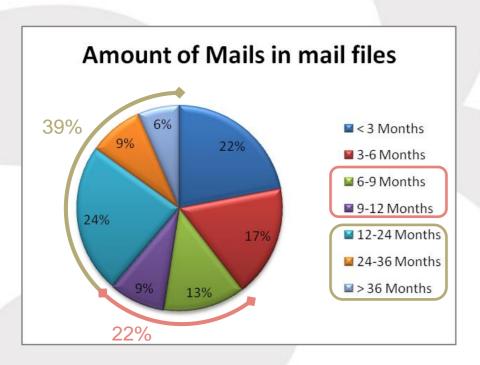
Belux Lotus User Group

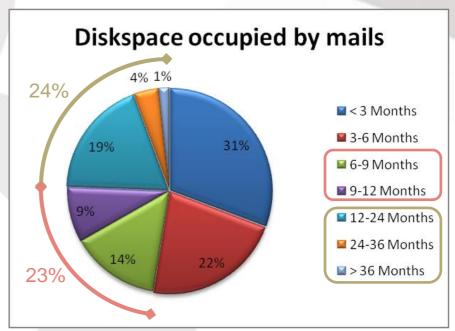
# Mailboxes actively involved in Tiering





# Comparison tiering versus pre-investigation (June 2007)

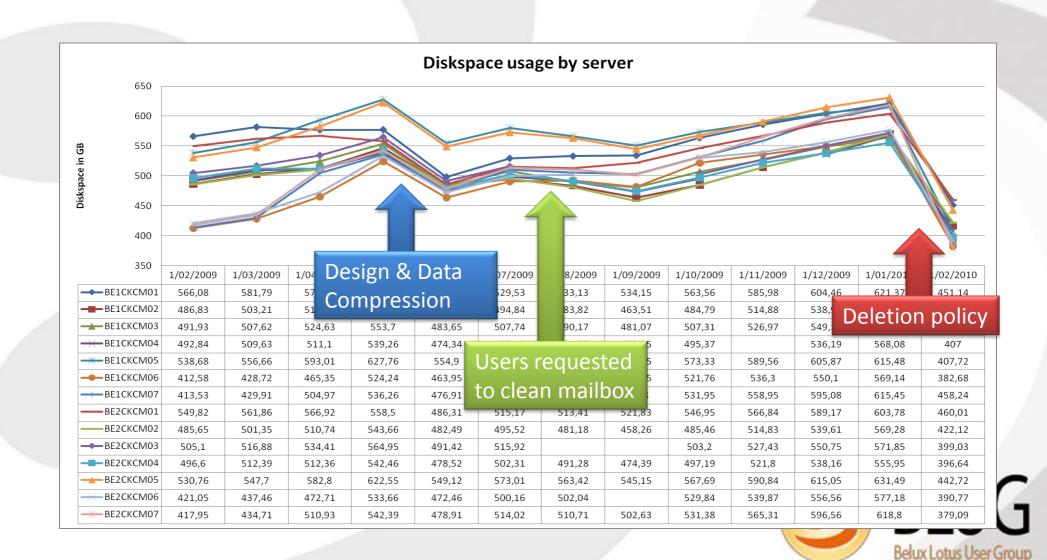




		run		
Action	Data	Pilot + delayed Delete	Jan	<b>Grand Total</b>
Deletion	#Docs	11.209.874	8.013.940	19.223.814
	GBytes	976,50	978,92	1.955,42
	#Db	11.368	12.400	23.768



# Disk space Usage by Server



# Total disk space – All servers



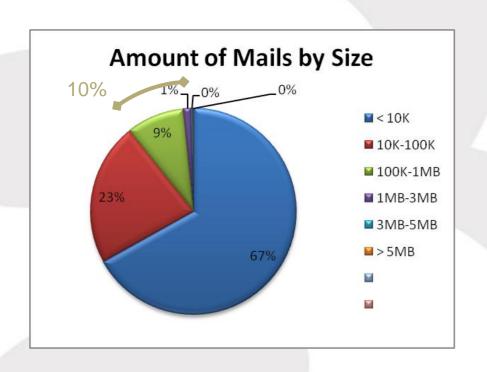
Belux Lotus User Group

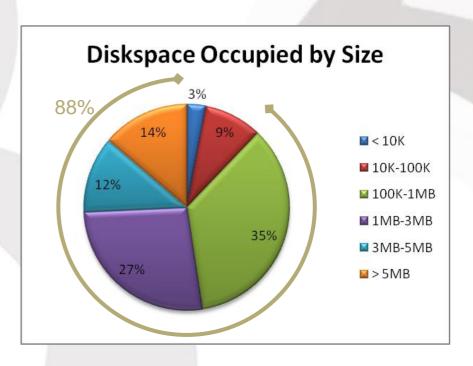
### Conclusions

- 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?



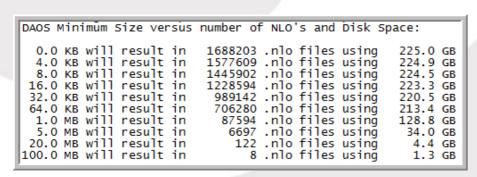


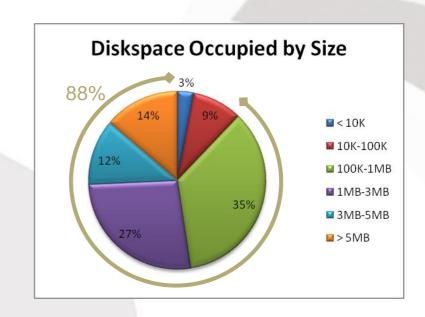
- DAOS will centralize attachments above a minimum size (standard 4KB, IBM advises 64KB)
- In case we assume 'minimum size of object' is 100KB
  - 10% of documents will be in scope
  - 88% of size is in scope



### **DAOS** Estimator

		_
Size Distribution	n of All Attachments Found	īl
		<u>-</u>
 	618686   6	<u> </u>
0.0%   0.2%   0.6%   1.4%   3.3%	39.0%  41.5%  12.6%   1.0%   0.4%   0.0%	  - 
		_
4 KB   8 KB  16 KB  32 KB  64 KB	1 MB   5 MB  20 MB  100 MB  1 GB  > 1 GB	Ц
		=





ı	Summary:	
I	Total DB's analyzed:	3646
П	Total DB's skipped due to errors:	0
ı	Total Size of NSF's Examined:	555.3 GB
ı	Total Attachments found:	1688203
ı	Total Duplicate Attachments found:	762677
П	Estimated Size of DAOSified NSF's:	171.6 GB
ı	Estimate Size of DAOS dir:	225.0 GB
ı	Total Disk Savings:	158.7 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

Default, normal user

All mails marked for Archiving will be moved to the Archive

Mails removed from 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

Mails removed from 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

### **Timeline**



All mails in productive

mailbox









# 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

### Contact

### **Keno Torfs**

Solution Architect / Project Manager

BASF IT Services Terhulpsesteenweg 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

