

**How-to Guide
SAP NetWeaver 7.0**



How To... Tune the Performance of Knowledge Management

**Version 1.40 – January 2009
Public**

**Applicable Releases:
SAP NetWeaver 7.0
(Portal, Knowledge Management)**

© Copyright 2007 SAP AG. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP AG. The information contained herein may be changed without prior notice.

Some software products marketed by SAP AG and its distributors contain proprietary software components of other software vendors.

Microsoft, Windows, Outlook, and PowerPoint are registered trademarks of Microsoft Corporation.

IBM, DB2, DB2 Universal Database, OS/2, Parallel Sysplex, MVS/ESA, AIX, S/390, AS/400, OS/390, OS/400, iSeries, pSeries, xSeries, zSeries, z/OS, AFP, Intelligent Miner, WebSphere, Netfinity, Tivoli, and Informix are trademarks or registered trademarks of IBM Corporation in the United States and/or other countries.

Oracle is a registered trademark of Oracle Corporation.

UNIX, X/Open, OSF/1, and Motif are registered trademarks of the Open Group.

Citrix, ICA, Program Neighborhood, MetaFrame, WinFrame, VideoFrame, and MultiWin are trademarks or registered trademarks of Citrix Systems, Inc.

HTML, XML, XHTML and W3C are trademarks or registered trademarks of W3C[®], World Wide Web Consortium, Massachusetts Institute of Technology.

Java is a registered trademark of Sun Microsystems, Inc.

JavaScript is a registered trademark of Sun Microsystems, Inc., used under license for technology invented and implemented by Netscape.

MaxDB is a trademark of MySQL AB, Sweden.

SAP, R/3, mySAP, mySAP.com, xApps, xApp, SAP NetWeaver, and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP AG in Germany and in several other countries all over the world. All other product and service names mentioned are the trademarks of their respective companies. Data

contained in this document serves informational purposes only. National product specifications may vary.

These materials are subject to change without notice. These materials are provided by SAP AG and its affiliated companies ("SAP Group") for informational purposes only, without representation or warranty of any kind, and SAP Group shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP Group products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

These materials are provided "as is" without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. SAP shall not be liable for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials.

SAP does not warrant the accuracy or completeness of the information, text, graphics, links or other items contained within these materials. SAP has no control over the information that you may access through the use of hot links contained in these materials and does not endorse your use of third party web pages nor provide any warranty whatsoever relating to third party web pages.

SAP NetWeaver How-To Guides are intended to simplify the product implementation. While specific product features and procedures typically are explained in a practical business context, it is not implied that those features and procedures are the only approach in solving a specific business problem using SAP NetWeaver. Should you wish to receive additional information, clarification or support, please refer to SAP Consulting.

Any software coding and/or code lines or strings ("Code") included in this documentation are only examples and are not intended to be used in a productive system environment. The Code is only intended to better explain and visualize the syntax and phrasing rules of certain coding. SAP does not warrant the correctness and completeness of the Code given herein, and SAP shall not be liable for errors or damages caused by the usage of the Code, except if such damages were caused by SAP intentionally or grossly negligent.

Document History

Version	Date	Change
Version 1.10	July 2006	Initial version
Version 1.20	January 2007	<p>Additions and changes to the following sections:</p> <p>New section: Before You Start: Important Information</p> <p>Section 5 extended:</p> <ul style="list-style-type: none"> • Using ACL caches in CM repositories • Shorter character strings in the URL generator service • Using wrappings to accelerate search queries <p>Section 6 extended:</p> <ul style="list-style-type: none"> • Recommendation for the persistence mode DBFS and reference to the corresponding backup documentation in the administration guide added. <p>Section 6.2 added:</p> <ul style="list-style-type: none"> • Distance to the LDAP system <p>Section 6.3 extended:</p> <ul style="list-style-type: none"> • Using the <i>SimpleExplorer</i> layout set for faster navigation • Deactivating content snippets and "Did you mean" terms <p>Section 6.5.5 added:</p> <ul style="list-style-type: none"> • Using the modified XML forms resource renderer
Version 1.30	May 2008	<p>Additions and changes to the following sections:</p> <p>Section 5 extended:</p> <ul style="list-style-type: none"> • Caching Static Resources in the Browser Cache of the Client PC subsection added.
Version 1.40	January 2009	<p>Section 1 extended:</p> <ul style="list-style-type: none"> • SAP enhancement package 1 for SAP NetWeaver 7.0 added as an applicable release. <p>Section 2 changed and extended:</p> <ul style="list-style-type: none"> • The storage location of Sizing SAP NetWeaver Portal changed. • Reference to the <i>Installation Guide - SAP NetWeaver 7.0 including EHP 1</i> added. • Reference to SAP Note 900571 added. <p>Section 5 extended:</p> <ul style="list-style-type: none"> • <i>Caching Static Resources in the Browser Cache Forever</i> subsection added. <p>Section 6.1 extended and restructured:</p> <ul style="list-style-type: none"> • Information about the persistence modes of the CM repository manager enhanced and structured into three subsections per mode

		providing a general description, information about the possible applications and special features, and information about the performance impact of each mode.
--	--	---

Before You Start: Important Information

These performance adjustments specified in the performance how-to-guide are part of the following releases:

- SAP enhancement package 1 for SAP NetWeaver 7.0
- SAP NetWeaver 7.0 SPS 10 and higher
- SAP NetWeaver 2004 SPS 19 and higher (with minor changes)

Therefore, we recommend the following strategy to customers:

1. If your project planning permits, upgrade to SAP NetWeaver 7.0 in the near future. SAP NetWeaver 7.0 provides new functions that are not available in SAP NetWeaver 2004.
2. If you plan to continue using SAP NetWeaver 2004 in the near future, we recommend updating to SPS 22. However, in the medium term you should schedule the upgrade to SAP NetWeaver 7.0.

1 Introduction

This guide provides recommendations that allow you to improve the performance of the *Knowledge Management Platform* in SAP NetWeaver 7.0



The collective SAP Note 739829 contains corrections to this document and new recommendations. Read this SAP Note before reading this document.

At the beginning of this guide is a list of documents that are either prerequisites for this guide or provide you with related information. The going-live checks offered by SAP are also listed. You should carry out these checks before starting to use the portal for production operations. The *Best Practices* section contains tips that help you to avoid common errors. This section also contains many recommendations on how to configure your system minimally and to optimally modify it to your needs.

2 Prerequisites

This guide applies to a portal system with an installed *Knowledge Management* of version SAP NetWeaver 7.0 SPS 0 and higher, or SAP NetWeaver 2004 SPS19 and higher. SAP Search and Classification (TREX) should be installed.

Before you implement the performance recommendations in this guide, make sure that you have read the following documents and taken note of the described procedures and recommendations.

Document	Content	Version	Storage Location/Note Number
Sizing SAP NetWeaver Portal	Recommendations on CPU and memory sizing	2.0	http://service.sap.com/sizing → <i>SAP NetWeaver</i> → <i>SAP NetWeaver in Detail</i> → <i>Solution Life-Cycle Management</i> → <i>Hardware Sizing</i> → <i>Guidelines</i> → <i>Solutions & Platform Sizing Guidelines</i> → <i>Solutions & Platform</i>
Installation Documentation – SAP NetWeaver 7.0 SR3	Installation of SAP NetWeaver 7.0	Depending on OS	http://service.sap.com/installnw70
Installation Guide - SAP NetWeaver 7.0 including EHP 1	Installation of SAP enhancement package 1 for SAP NetWeaver	Depending on OS	http://service.sap.com/installnw70
Java VM settings for J2EE 6.40/7.0	Configuration of the Java Virtual Machine	24	723909
Performance Tuning Windows 2000/2003 (Advanced) Server	Contains a procedure to improve performance when using TREX to index larger systems	4	500408
KM administration guide	Describes the administration and configuration of Knowledge Management	SPS Specific	http://help.sap.com/nw70 <u>KM Administration Guide</u>
CM Repository Manager: Additional Information	Contains additional information about the persistence modes of the CM Repository manager.	15	900571

3 Related Documents

The following table contains information on related topics:

Document	Content	Version	Storage Location/Note Number
Technical Operations Manual for TREX	Entry-level document for administrators containing procedures for optimum operation and use of TREX	SPS-specific	SAP Library http://help.sap.com/nw70 Search and Classification (TREX) → Tasks On Demand → Optimizing Performance
Distributed TREX Systems (Multiple Host Installation)	Explains how to plan and implement a distributed system.	SPS-specific	SAP Library http://help.sap.com/nw70 Distributed TREX Systems (Multiple Host Installation)
Collaboration Room Meta Attribute Cleanup in PCD	Describes how you can improve the performance of virtual rooms by deleting PCD attributes.	5	991480

4 Going-Live Checks

We recommend performing the *SAP Going-Live Checks* for your production system. Depending on the phase of your project, these checks, offered by SAP, analyze the hardware sizing, check the configuration of the server and components, find performance bottlenecks, and verify the capacity of CPU and memory. The results of the checks are issued in a report.

The following contains a list of the recommended checks:

Check	Description
EP Going-Live Check Analysis Session (GA)	The GA session checks whether the available hardware is sufficient for the planned workload. Recommendations for the component settings are made.
EP Going-Live Check Optimization Session (GO)	The GO session simulates workload on the system and identifies performance bottlenecks and potential risks.
EP Going-Live Check Verification Session (GV)	The GV session checks the behavior of the production system after going live and provides appropriate recommendations.

You can find more information on the *SAP Service Marketplace* at service.sap.com/safeguarding.

5 Performance-Optimized KM Configuration

Network Drive Letters in Lowercase

You should always enter network drive letters in lowercase in the configuration. This allows you to avoid time-consuming system calls.

Setting Write-Protection Mode for the Folder `/etc/public` (valid for SAP NetWeaver 7.0 SPS 10 and SAP NetWeaver 2004 SPS19 only)

Set all items in the folder `/etc/public` and its subfolders to write-protected through the operating system so that these items are stored temporarily in the browser. All `.gif` files are stored in this folder. For this reason, the cache process significantly reduces the loading time on the client side. The load is also reduced in the backend system because fewer operations and resources are saved.



You must undo the write-protection in the `/etc` repository before a system update. Otherwise, the update fails. It is possible that the write-protection will be undone automatically during future updates.

In the case of a portal that runs on Windows, proceed as follows: Open the Windows Explorer and navigate to the folder `.../usr/sap/<instance number>/SYS/global/config/cm/etc`. Click the `public` folder with the right-hand mouse button and choose *Properties* in the context menu. Activate the *Read-Only* property and choose *Apply*. In the dialog box, choose *Apply changes to folders, subfolders, and files*. Then choose *OK*.

To make the above changes take effect, increase the timeout value (parameter `Read-Only Content Expiry Delay`) for the repository `/etc` to several days, for example, five days.

To launch the configuration of the `/etc` repository, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Repository Managers* → *File System Repository* in the portal.



This setting causes client-side caching, that is, caching in the browser on the client side. When choosing the cache duration, possible changes in the relevant folders of the repository `/etc` should be taken into account.

You can set the write-protection mode and increase the timeout value not only for the file system repository `/etc`, but also for all customer-specific CM repositories that are operated in the persistence modes DB or DBFS.

To launch the configuration of CM repositories, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Repository Managers* → *CM Repository Manager* in the portal.



Set write-protection mode for resources **only** in customer-specific CM repositories. Do not do this in CM repositories delivered by SAP because this can lead to function restrictions and problems during updates.

Caching Static Resources in the Browser Cache of the Client PC (as of SAP NetWeaver 7.0 SPS 11 and SAP NetWeaver 2004 SPS20)

Use the client cache service to cache static resources, such as graphics or stylesheets in the browser cache of the client PC. The cache process significantly reduces the loading time on the client side. The load is also reduced in the backend system because client-side caching reduces the amount of requests sent to the portal server. Thus, fewer operations are performed on the server side and

resources are saved. The client cache service is activated by default, and it is independent of the repository used to store the resources.

The client cache service analyzes client cache patterns that you created in KM Configuration. In a client cache pattern, you specify the KM folder and/or MIME types for which resources are to be cached in the browser cache of the client PC for the defined time of call.



If you modify or delete a resource in KM during its lifetime, the cached version still appears on the client of the user until the configured lifetime has expired, unless you manually delete the browser cache on the client side.

To create client cache patterns for the folders whose resources are to be cached on the client, in the portal, navigate to *System Administration->System Configuration -> Knowledge Management -> Content Management -> Global Services (-> Show Advanced Options) ->Related Areas -> Client Cache Patterns*.



When specifying the lifetime of resources, take into account possible changes to the resources.

Caching Static Resources in the Browser Cache Forever (as of SAP Enhancement Package 1 for SAP NetWeaver 7.0 SPS0)

As of SAP enhancement package 1 for SAP NetWeaver 7.0 SPS0, you can configure certain MIME resources to be cached forever in the Web browser of the client PC without losing the flexibility to change the resources on the server side. To optimize performance, change the `Cache Timeout` value for the SAP standard configuration `etcPattern` to `-1`. As a result, all MIME objects coming from the configured path and matching the configured MIME types are cached forever in the Web browser. The browser retrieves the objects from the server only if the image file is modified. However, the browser does not retrieve the modified image immediately because of the small cache that stores the resource URLs for one minute. Thus, it may take up to one minute for the browser to display the modified image.



Set the `Cache Timeout` value to `-1` for KM resources, only. It is required that the code of the application that generates the HTML containing links to the affected MIME resources uses the KM service `URL Generator Service`. Thus, the feature is supported for KM resources only.

Configuration in a Cluster Environment

If you are operating KM in a cluster environment, the `Cluster Installation` parameter must be activated in the system landscape service (see the KM installation guide). Check the `Cluster Caching Timeout` parameter as well. To launch the configuration of the service, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Global Services* → *System Landscape Service* in the portal.



The cluster caching timeout determines the period for which local caches remain valid, regardless of whether another cluster node has invalidated the cache in the meantime. Changes to a cluster node are only distributed to other nodes in the cluster after this time period. For example, a new ACL entry only takes effect after this time. This leads to *dirty reads* and, in the case of write operations, to *lost updates*, therefore this value should be kept as small as possible (30 seconds for a scenario with many read operations and fewer write actions, 10 seconds for a scenario with many write operations).

Using ACL Caches in CM Repositories

If you are using ACLs in CM repositories, you should activate ACL caching to reduce response times. The caches `ca_cm_rep_acl` and `ca_rsrc_acl` are used for this caching. To launch the configuration of the caches, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Utilities* → *Caches* → *Memory Cache* in the portal. Open the entries for the `ca_cm_rep_acl` and `ca_rsrc_acl` caches and set the following values:

```
Restart lifetime on access = activated
Capacity = 10000
Default Time-to-Live = 1800
Singleton = activated
Assumed Entry Size = 0
Max Cache Size = 0
Max Entry Size = 0
```

Once you have configured the caches, you must enter them in the configuration for all repositories that you are using ACLs in. To launch the configuration of the CM repositories, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Repository Managers* → *CM Repository Manager* in the portal. Open the entries that use ACLs and set the following values:

```
Security Manager = AclSecurityManager
ACL Manager Cache = ca_rsrc_acl
```

Shorter Character Strings in the URL Generator Service

Shorter character strings for some parameters in the configuration of the URL generator service can improve performance. In addition, they avoid problems with URLs that are too long in certain WebDAV clients. This conversion also suppresses the basic authentication messages when opening MS Office documents.

To call up the configuration of the URL generator service, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Global Services* → *URL Generator Service* in the portal. Make the following changes:

1. Replace the prefix

`/irj/servlet/prt/portal/prtroot/com.sap.km.cm.docs` in the parameters *Image Path*, *Viewer*, *XML Forms CSS URL*, and *Content Access Path* with `/irj/go/km/docs`

The entry in the *Image Path* parameter must look as follows:

`/irj/go/km/docs/etc/public/mimes/images`

2. Replace the prefix

`/irj/servlet/prt/portal/prtroot/com.sap.km.cm.uidetails` in the parameters *Resource Properties Page* and *New Resource Properties Page* with `/irj/go/km/details`

3. Replace the prefix

`/irj/servlet/prt/portal/prtroot/com.sap.km.cm.navigation` in the parameters *Explorer Servlet* and *Navigation Servlet* with `/irj/go/km/navigation`



If you are using a release earlier than NW '04 SPS 14, you cannot change the following two parameter entries.

4. Replace

`/irj/servlet/prt/portal/prtroot/com.sap.km.cm.highlightedcontent` in the parameters *Highlighted Content* with `/irj/go/km/highlightedcontent`

5. Replace

`/irj/servlet/prt/portal/prtroot/com.sap.km.cm.basicsearch` in the parameter *Basic Search Servlet* with `/irj/go/km/basicsearch`

Using Wrappings to Accelerate Search Queries

In the index management service, you can specify repository managers, for which wrappings are generated for search result objects when searching takes place. The wrapping wraps all required properties to the returned search result objects rather than all available properties. This increases performance for search queries.

In the portal, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Global Services* → *Index Management Service*. Enter the repository managers in the *Repository Manager for Which Search Results Are Wrapped* parameter.

You may only specify repository managers of the type CM repository manager or file system repository manager. If you want to use wrapping for repository managers you developed yourself, you must carry out tests first to make sure that searching takes place without errors if you enter your own repository managers here. If you have developed your own repository managers and specified them here, remove them again if problems occur during searching.

6 Best Practices

6.1 Things to Consider

You should make the following decisions before starting the specific configuration of KM:

- Which persistence mode (DB, DBFS, or FSDB) for the CM repository is most suitable for you?
 - **DB mode** means that all data, including hierarchy information and content (documents, folders, and metadata) is stored in the database.

Possible applications and special features of the DB mode:

If there is a large number of write requests in your CM usage scenario, set up the CM repository in database mode. Since all documents are stored in the database, this avoids unintentional external manipulation of the data.

The procedure for data backup and restore for CM repositories in DB mode is easy since only the database needs to be backed up.

Performance impact:

Since the content is stored in the database the streaming speed is lower than in the other persistence modes. For CM repositories set up in DB mode the streaming speed depends on the performance of the database. (In a non-standard scenario, many large documents (>1MB) are often called. The threshold between standard scenarios and non-standard scenarios depends on the database used: The calculated value of the average file size multiplied by the average number of accesses must not be larger than the database-specific throughput threshold.) The streaming speed might also depend on the JDBC driver version used.

- **DBFS mode** means that the metadata and the folders as well as the hierarchy information are stored in the database, but the content is stored in the file system.

Possible applications and special features of the DBFS mode:

The DBFS mode minimizes connections to the database. Streaming is carried out by the file system, not by the database. This mode is suitable for managing larger amounts of data that a database can no longer handle. This mode is also recommended if you do not want the database to become too large. (The content requires considerably more space in the database than the hierarchy information.) The DBFS mode is also suitable for CM usage scenarios with larger amount of read requests.

This mode allows you to control the size of the database more easily since documents are stored in the file system.

The documents and metadata are stored in different places, so you have to take into account both the database and the file system and synchronize them both when backing up and restoring data. Therefore, the backup and restore mechanism is more complicated than in pure DB mode.

Performance impact:

This mode is faster than the database mode if you have large documents since there is no database data streaming

- **FSDB mode** means that folders and documents are stored in the file system, but metadata is stored in the database.

Possible applications and special features of the FSDB mode

In FSDB persistence mode the file system is predominant. The FSDB mode is used to mount a file system with an existing structure. This mode is recommended for selective outsourcing of large files, for example, media files, from the database to prevent the database overloading or crashing in the event of high number of accesses.

For a scenario with write accesses, this mode cannot ensure consistent data backup and restore. Consistent data backup and restore can therefore be ensured only if read accesses alone are taking place.



Note that metadata can be lost if you edit or manipulate documents or folders directly in the file system.

Performance impact

As far as streaming is concerned, this mode can sometimes be better than the DB mode.

Since file systems are not transactional, this mode has restrictions and affects performance. If read and write operations take place for one document in the file system at the same time, these operations have to be coordinated by the repository manager. This happens by recording both write accesses and read accesses in the database. This affects the performance.

From a performance perspective, we recommend the DBFS mode. However, you should note that the backup and restore process is more complicated than for the pure DB mode.

For more information about the persistence modes and about which points you need to note during a backup, see the KM administration guide, *CM Repository Manager* section, and SAP Note 900571. You can find the link to the KM administration guide in section 2 of this how-to guide.

- Plan which filters and services you want to use for your repositories. Note that each additional filter or service used costs additional performance.
- You should store resources that are relevant for rendering (for example, graphics, gif files) in a separate KM repository. Make sure that the repository is modified appropriately: You must deactivate repository services that are not required, such as *accessstatistic*, because it is time-consuming when they access the repository. This can significantly affect performance. Do not use the repository configurations in the standard system because they are not optimized to suit your project requirements. For more information, see the Repository Services section of the documentation.
- Create a quantity structure for expected data sets. When creating your data or folder structures, note the restrictions on the number of documents that can be stored in a folder. Special folders that are regularly supplied with documents (such as a news archive) are critical here. If necessary, folders with more than 1000 documents should be distributed into subfolders by date or another characteristic, or you should make sure that the oldest documents are deleted regularly.

6.2 Connections To Other Systems

The LDAP you are using should be located close to the portal to guarantee response time of 15 ms for access checks.

6.3 Minimal Configuration

Some functions can negatively affect the performance of your system. You should therefore only implement these functions if they are really necessary.

Note the following recommendations for a minimal configuration:

- Deactivate global services that you are not implementing (such as the quick poll service).
- In the configuration of repository managers: Deactivate the repository services that you do not need.

You can find the configuration of the repository managers in the portal at *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Repository Managers*.

- Deactivate commands that you do not require (for example, *Move* or *Delete*).
Check the configuration of commands for your resource renderers. Reduce the number of commands to a minimum. Reason: Rendering commands, for example, in hover mode, uses a relatively high amount of CPU time. Each additional command increases the CPU time required for rendering.
You can find the configuration of the commands and command groups in the portal at *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *User Interface* → *Commands*.
- For fast navigation, provide a layout set that does not contain a context menu. For example, you can use the *SimpleExplorer* layout set. Set this as the default layout or inform your users that they are to use this layout set in their personalized iView settings.
- Reduce the level of detail for the KM application log to ERROR. For example, in the case of a correctly configured production system, debug output is not useful and reduces system throughput. For more information about the KM log, see the KM administration guide.
- Check the configuration of the crawler parameters that you use to crawl for Web repositories. If the Maximum Depth parameter is empty or contains 0, this stands for unrestricted depth. You should definitely restrict the depth. Each additional layer of depth causes the number of visited links to rise.
- By default, TREX queries content snippets or "Did you mean" terms every time you execute a search. If you do not display content snippets or "Did you mean" terms in your search iView, you can deactivate this query to improve performance during searches.

To do this, add the following parameters to the existing entry in the *query* iView property:

```
disContSnippets=true
```

```
disDidYouMean=true
```

6.4 Optimum System Configuration

6.4.1 General Tips

To achieve optimum system utilization, take the following points into account:

- Distribute scheduler tasks among different portal nodes and if possible, set these up as dedicated nodes to which the load balancer does not send user requests.
- Define new scheduler times if the default times fall in main periods of operation and thus in periods of high system load in your company.
- Use the cache monitor to check the load and size of the active caches. If necessary, correct their settings. For more information, see the KM administration guide.

6.5 Implementation Tips

6.5.1 General Tips

- If you want to display content in a folder that contains a large number of documents, make available a modified KM search iView for users instead of a KM navigation iView. We recommend this in news or taxonomy scenarios, for example.
Reason: When a user uses the KM navigation iView to call up a folder that contains several hundred documents, the system has to take all documents into account when preparing for display. However, only a limited number of documents are displayed on the first page of the folder. All other documents are not taken into account at this point. Performance is significantly better if you use a KM search iView instead. You can use a query (for example, *.doc) to call up the documents in this folder directly from the search index and structure the display according to a prescribed sort order.
- For news scenarios, you can use the persistent caching of the portal runtime (PRT). Set the cache timeout to ten minutes.
Reason: The data is read completely as an HTML stream from the database.
- Assign only a small number of iViews to a page (less than six).
Reason: Each iView in *isolated mode* generates a request to the server. Internet Explorer allows a maximum of two parallel server requests; it serializes requests, which results in slow screen output. This is not relevant for iViews in *embedded mode*.
- Place a maximum of two KM navigation iViews that display folders containing more than 200 documents on a page. Refer also to the *Tips for the Homepage* section.
- Avoid using the tree view in end user layout sets for the navigation iView. The tree view of the root directory is particularly complex.
- Do not use any “full” iViews, that is, iViews that display only fixed HTML.
Reason: A “full” iView that displays only static HTML, for example, to generate a title or padding between two iViews, generates at least one request to the server. It makes more sense to integrate the title in the iView in question to prevent the additional request to the server.
- If you want to use only search and not classification, you should use only pure search indexes. These are faster than combined search and classification indexes.

6.5.2 Tips for the Homepage

- Avoid assigning a KM navigation iView containing more than 20 documents to the portal homepage.
Reason: Displaying a folder with a large number of documents requires a lot of CPU time, for example, for sorting. As a consequence, the response time for logging on to the portal increases.

6.5.3 Tips for Web Repositories

- Make sure that backend systems with fast response times are used. Note the following tips for this:
 - Avoid switches, routers, and so on between the portal server and Web servers.
 - Make sure that there is a high bandwidth between the portal server and the Web servers.
 - Use fast storage mediums, such as hard disks instead of tape drives.
- Values for the configuration of a Web repository cache:
 - `Capacity = (20 x number of providers) + 10`
 - `Default Time to Live = 7200` (2 hours)
- Values for the configuration of a Web repository manager:
 - `Cache Content Persistently = switched off` (the parameter is relevant for the Legacy Web Repository manager only)
 - `Cache Timeout =` Choose a value that causes the cache content to be retained until the next update of the index.
 - `HTTP Timeout = 10000` (10 seconds). If timeouts occur, longer timeouts cause longer index periods.

6.5.4 Tips for TREX and Related Processes

6.5.4.1 Crawling and Indexing (General)

- If possible, crawl and index repository content in packages that do not exceed 20-30 GB or 1-2 million documents each. This procedure optimizes the crawling process and indexing performance as follows:
 - Memory shortages in the AS Java when crawling too many objects are avoided.
 - New initiations of crawling and indexing processes can be carried out incrementally.
 - You avoid overloading the TREX preprocessor.

However, this recommendation requires a suitable repository structure.

- If there are regular and/or significant content changes in KM or in connected, indexed repositories, we recommend that you scale the TREX preprocessor by placing it on one or more separate servers.
- Do not create one large index or an extremely large number of small indexes for indexed document sets of 10 to 100 GB. Both extremes would negatively affect search operations. Depending on the quantity of data, a value of 2-20 indexes is recommended.
- Increase the size of the pool in the `Trexqueueserver.ini` file (number of threads in `Trexpreprocessor + 2`)
- If you are using index replication, increase the number of threads in the `Trexpreprocessor.ini` file (3 threads for each CPU).

6.5.5 Modified XML Forms Resource Renderer (as of SAP NetWeaver '04 SPS 19, SAP NetWeaver 7.0 SPS 10)

The default resource renderer for XML forms uses a memory cache with a LRU displacement strategy (**Least Recently Used**). In news scenarios, in which most users display only the latest news items, you can generally achieve a good hitrate with this caching mechanism.

In other scenarios, in which access to resources is distributed more evenly, for example, bulletin boards, the hitrate can be lower. In this case, it can make sense to use the XML persistence resource renderer. If you use this renderer, the system persists the HTML fragment in a property of the XML resource. The caching mechanism then has two stages:

1. Memory cache as for the default renderer
2. Property of the XML resource

The first time that the resource is accessed, the system writes the HTML fragment to the document property. However, the XSL transformation is avoided for each subsequent access, which can lead to improved performance in this scenario.

If you want to use the XML persistence resource renderer, proceed as follows:

1. Create a set of resource renderer settings.
 - a) Choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *User Interface* → *Settings* → *Resource Renderer Settings* → *Resource Renderer Settings* in the portal.
 - b) Duplicate one of the following renderers:
 - NewsResourceDisplayRenderer
 - NewsResourceRenderer
 - c) In the *Resource Renderer* field, choose the *NewsPersRenderer* entry.
2. Create a layout set.
 - a) Choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *User Interface* → *Settings* → *Layout Set* in the portal.
 - b) Duplicate one of the following layout sets:
 - NewsBrowser
 - NewsExplorer
 - c) Select the resource renderer settings that you created in step one.
3. Create an iView with this layout set.

6.5.5.1 Crawling Web Repositories

- To crawl Web repositories, use crawlers with 4 to 6 provider threads and a similar number of retriever threads. In the portal, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Global Services* → *Crawler Parameters* and enter the required number of provider and retriever threads.



If the crawlers are too fast, the remote servers and networks may not be able to keep up. If necessary, reduce the number of retrievers and providers. Otherwise you may get unintentional *Denial of Service Attacks*.

If you want to crawl external Web sites, use the repository access delay function to avoid too high a load (*Denial of Service Attacks*) on the external system. In the portal, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Global Services* → *Crawler Parameters* and configure the required delay in the *Repository Access Delay* parameter.

6.5.5.2 Crawling CM Repositories

- To crawl CM repositories, use crawlers with 4 to 6 provider threads and one retriever thread.



Note that using *more* provider and retriever threads increases the crawler performance, but also increases the system load when performing other tasks.

- Provided that the crawling of CM repositories is not time-critical, you do not need a dedicated machine or multiple provider and retriever threads.

6.5.5.3 Prompt Indexing

Carry out the following to make document indexing as prompt as possible:

- Configure the crawlers so that they run on a dedicated machine. This machine should not be connected to the load balancer, because the high load caused by the crawlers leads to high response times for user interaction.

To configure a dedicated machine, choose *System Administration* → *System Configuration* → *Knowledge Management* → *Content Management* → *Global Services* → *Scheduler Tasks* → *Index Management Task Queue Reader* in the portal. Enter the required machine in the CM Systems parameter.

6.5.5.4 Search Operations

- Make sure that the TREX server has enough physical main memory to keep indexes in the main memory when the search is accessed. An index that has the size HD on the server hard disk uses the following amount of main memory: HD/10 to HD/30 (a tenth to a thirtieth of the index size on the server hard disk).
- Using only one TREX HTTP server simplifies tracing when troubleshooting. However, using multiple HTTP servers in distributed TREX scenarios increases the number of requests that can be processed in a period of time.

6.5.5.5 Automatic Classification

- Instead of creating an index and taxonomies, defining queries, and then starting the indexing run for the documents, proceed as follows. This procedure is faster than the one just described.

Initial classification of all document for a query-based taxonomy:

- a. Create an index.
 - b. Index all documents. Use the TREX monitor to check the status of indexing. All indexed documents should have the status OK.
 - c. Create a taxonomy and the required categories.
 - d. Define queries for all categories.
 - e. Perform an update of the query-based taxonomy.
- The amount of time required for indexing a document depends on the number of taxonomies used in the index in question.

7 FAQs

How can I optimize performance when scrolling in repositories?

You can define "read-only access" for files and folders that do not have to be changed through the portal. The system can then store these in the local browser cache. You configure this setting in the *Details* dialog box. Choose *Settings* → *Properties*. On the *System* tab page, activate the *Read-Only* property.

This setting means that users can no longer update the content using the CM user interface, but they can continue to update properties. If read-only access is defined for a folder, users cannot create any more subfolders, but can delete, move, or change existing documents.

Does it matter whether we use the *AdminExplorer* or the *ConsumerExplorer* layout set?

For the most part, the performance of the layout sets depends on the complexity of the layout set, the number of resources that are displayed on one page, and the number and type of commands that are assigned to the resources. For example, the *AdminExplorer* layout set displays a longer document list than normal and a folder list in the left-hand screen area. In addition, it contains a comprehensive range of complex commands that are required for administrative tasks. These additional functions mean that the server load is three times higher than that for the *ConsumerExplorer* layout set. *SimpleExplorer* is an even faster layout set, in which the number of commands has been reduced to an absolute minimum.

I want to make several individual documents available to various employees. Is it a good idea to save all documents in one folder and use the ACLs to control access?

This would cause unnecessary server load because the system would have to check the access permissions for all documents each time that the content of the folder is displayed. Particularly in the case of a large number of documents, of which only a few are accessible, these access checks would be very time-consuming. A better solution is to store the documents in different folders and create specific ACLs for each folder.

I want to use a document display with a tree structure like the one in *AdminExplorer*. Which layout set should I use?

The *ConsumerTreeList* layout set is the most suitable for this purpose. This layout set is approximately twice as fast as the *AdminExplorer* layout set. However, it does not support mass commands and provides few commands.

Performance is poor when displaying a folder. How can I improve the performance?

- Reduce the number of documents in the folder to a maximum of 1000.
- Choose the layout set that are delivered as standard in this folder:
 - SimpleExplorer* (provides the smallest number of functions but the best performance)
 - ConsumerExplorer*, *ConsumerTreeList*, *AdminExplorer* (provides the largest number of functions but are time-consuming)
- Reduce the number of configured commands.

8 Constraints

- There are known issues with opening and saving documents through the portal (especially with Office documents and PDF files). SAP Note 700177 describes the problems and suggests solutions.
- In some cases, problems can occur when Office documents are opened more than once through the portal. SAP Note 855500 describes the problems and suggests solutions.
- Note that the performance of the portal with regard to throughput and response time is reduced by about 10%, if you use HTTPS in place of HTTP.
- Note that a maximum of 1000 documents should be stored in any one folder

www.sdn.sap.com/irj/sdn/howtoguides

