Задача хранения контента SAP-систем и ее решение с помощью SAP Archiving by OpenText

Олег Хрипко, SAP SE Декабрь 2014







Содержание

- 1. Типовые задачи хранения контента в SAP
- 2. Инструменты и механизмы SAP по работе с контентом
- 3. База данных SAP как место хранения контента
- 4. Cpabhehue SAP Content Server u SAP Archiving by OpenText
- 5. Миграция контента во внешнее хранилище
- 6. Архивация данных в SAP



Типовые задачи хранения контента в SAP







Когда стоит задуматься о хранении контента?

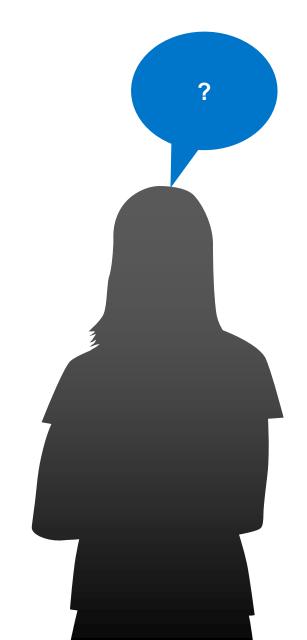
- При внедрении
- После внедрения (чем чаще, тем лучше)
- При миграции на SAP HANA





Ключевые проблемы бизнес-пользователей

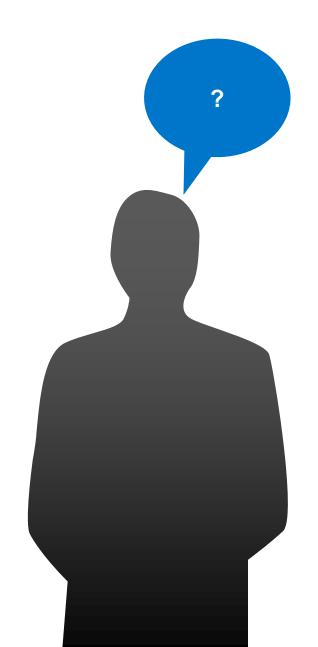
- Производительность
- Доступность
- Риск потери и повреждения документа





Ключевые проблемы IT-отдела

- Затраты на администрирование
- Затраты на оборудование для физического хранения
- Миграция на SAP HANA



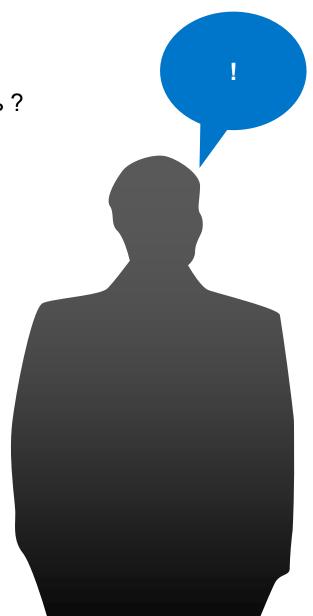


Что дальше?

Как не допустить появление проблем ?

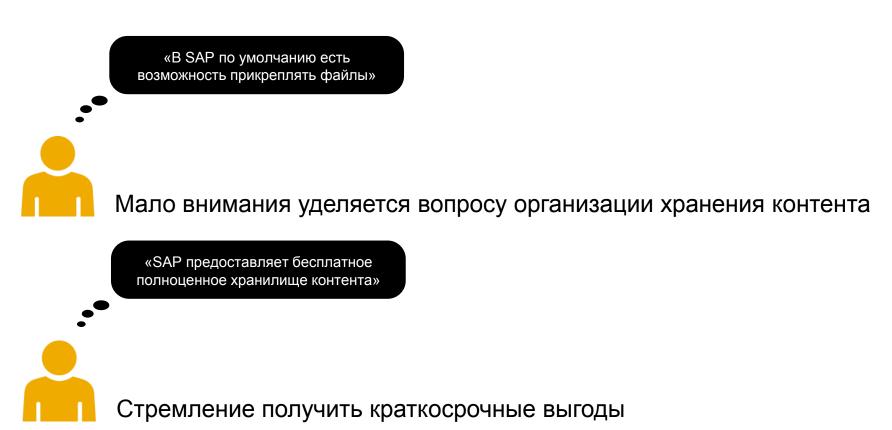
• Что делать для их устранения, если они уже есть?

 Какая при этом роль решения SAP Archiving and Document Access by OpenText





Распространенные заблуждения



«У меня уже есть решение для похожей задачи»



Узкий (локальный) взгляд на задачу



Типовые задачи управления контентом

Бизнес- процессы Компании	Производство	ТОРО	Управление проектами	Закупки	Сбыт	Финансы	Управление персоналом	Организационно- распорядительные процессы	
Документационное обеспечение бизнеспроцессов	Управление технической и эксплуатационной документацией		Управление проектно- сметной документа- цией	Управление договорами			Управление командировками	Управление приказами и распоряжениями	
				Управление тендерной документацией					1
				Управлени	ие претензионно работой	-исковой	коминдировкими	Управление входящими и исходящими	
Архивное обеспечение бизнес- процессов	Эл. архив технической и эксплуатационной документации		Эл. архив проектных документов	Эл. архив договоров			Эл. архив	Эл. Архив	
				Эл. архив д-тов поставщика	Эл. архив д-тов клиента	Эл. архив финансов ых д-тов	кадровых документов	организационно- распорядительных документов	2
									3
	Ведение бумажного архива								
	Сервисный слой по работе с хранилищем								
	Централизованное хранилище документов								

- Группа процессов по управлению документооборотом. Для данного класса задач требуется регистрационная карточка документа (РКК), функционал управления документами (версионность, связи между документами, поиск), автоматизация процесса согласования (Workflow), работа с шаблонами, отчеты, контроль исполнения.
- Группа процессов ведения электронного архива. Для данного класса задач требуется функционал ввода сканобразов со сканера, в том числе с потокового сканера, связь с учетным документом (бизнес-объектом), возникающим на основе бумажного документа. В то же время, для скан-образов отсутствует версионность.
- Процесс ведения бумажного архива. Для данной задачи требуется ведение номенклатуры дел со справочником статей хранения, архивная учетная карточка, справочники мест хранения и коробов, процедура выдачи оригиналов во временное пользование, инвентаризация, процедура уничтожения документов.



Содержание

- 1. Типовые задачи хранения контента в SAP
- 2. Инструменты и механизмы SAP по работе с контентом
- 3. База данных SAP как место хранения контента
- 4. Cpabhehue SAP Content Server u SAP Archiving by OpenText
- 5. Миграция контента во внешнее хранилище
- 6. Архивация данных в SAP

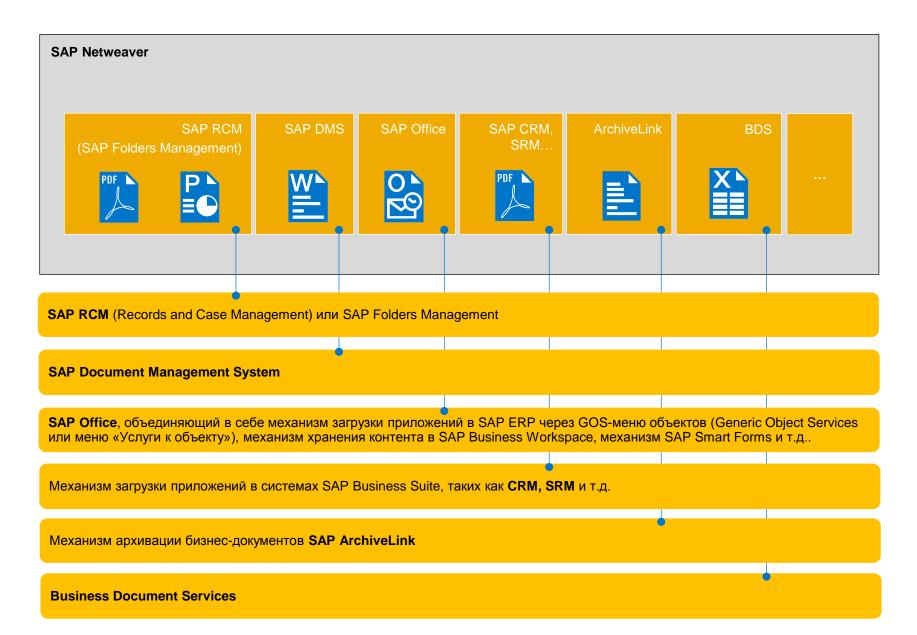


Инструменты и механизмы SAP по работе с контентом



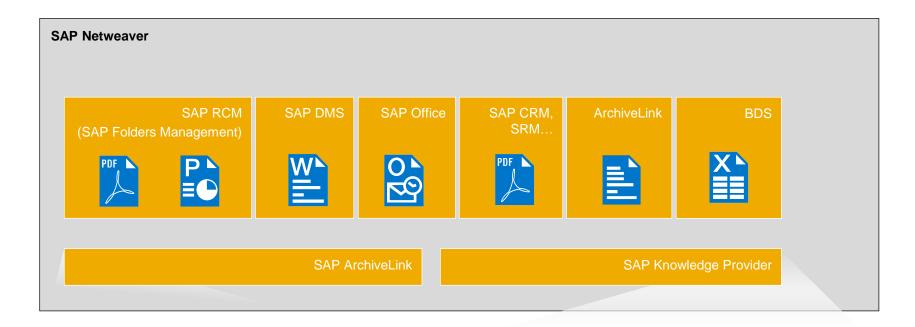








Механизмы хранения SAP KPRO и SAP ArchiveLink



SAP ArchiveLink - механизм хранения контента внутри платформы SAP NetWeaver, обеспечивающий связь заархивированных документов с бизнес-объектами SAP. Примечание: иногда термином «ArchiveLink» обозначают также унифицированный протокол взаимодействия с внешними системами хранения контента — «HTTP Content Server». Это не вполне корректно, так как, начиная с версии 4.5, в протоколе была реализована поддержка HTTP-соединения и механизма KPro, и он был переименован на «HTTP Content Server».

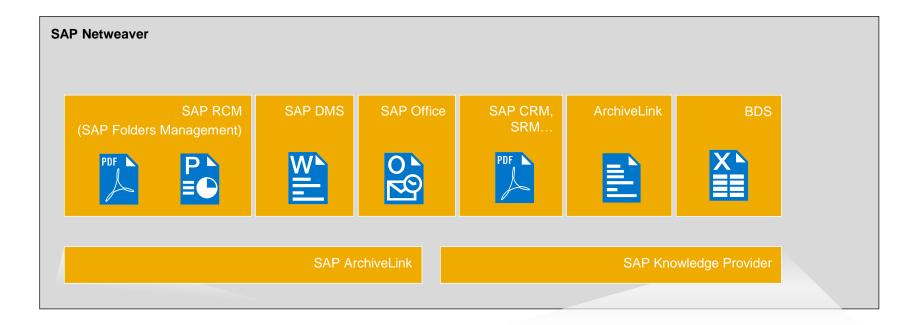
Knowledge Provider (KPro) - механизм, технологическая инфраструктура хранения контента внутри платформы SAP NetWeaver. Принципы работы механизма универсальны для всех модулей SAP и не зависят от типов устройств и способов хранения. КРго полностью реализует функционал взаимодействия с внешними системами хранения, включая функции администрирования.

Для хранения технически могут использоваться как внутренние репозитории, т.е. таблицы собственной БД SAP (не рекомендуется!), так и внешние репозитории, т.е. репозитории внешних систем хранения контента.

Исключение составляет механизм **SAP Office**, который в некоторых случаях может хранить контент в отдельных таблицах внутренней БД SAP, таблица SOC3



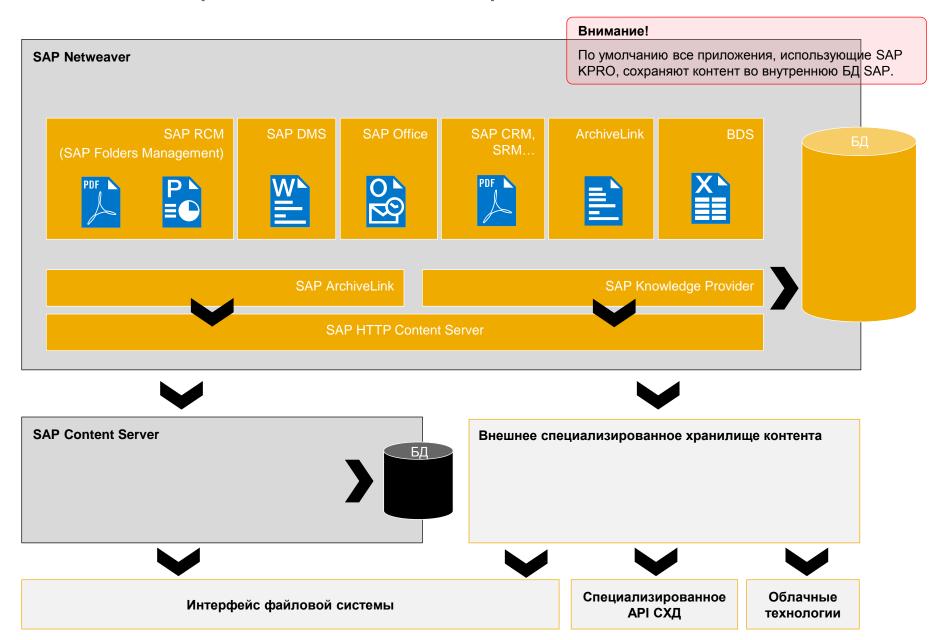
Механизмы хранения SAP KPRO и SAP ArchiveLink



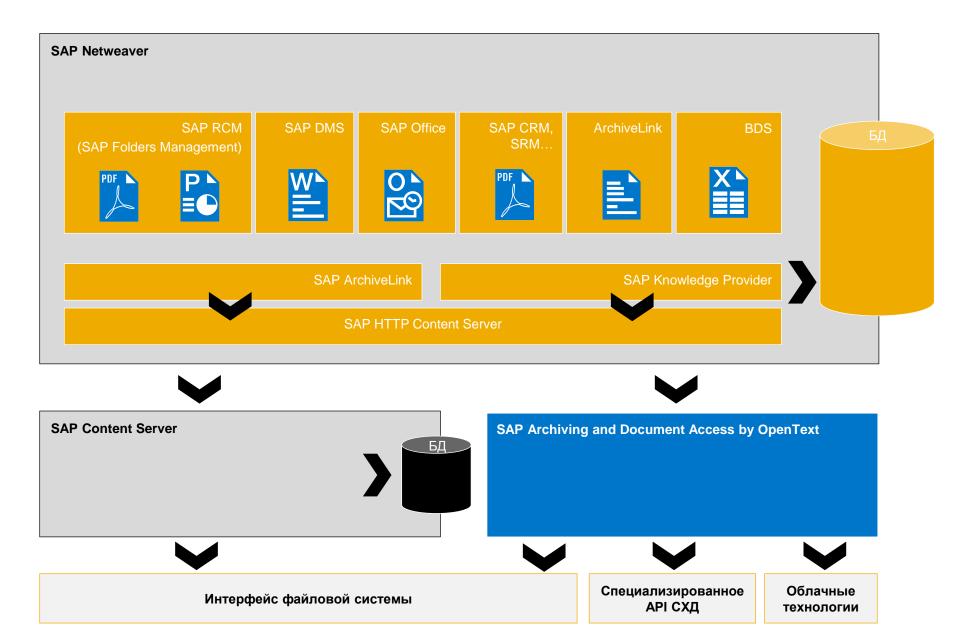
- Возможности полнотекстового поиска в стандартном функционале SAP - нет. Однако, такая возможность может быть добавлена с помощью сторонних приложений (например, SAP Archiving and Document Access by OpenText).
- Содержит интерфейс для ввода документа со сканера, однако в стандартном функционале SAP собственного решения по сканированию – нет.
- Для хранения информации о связях между объектами SAP и контентом используются таблицы TOA0x.

- Поддерживается возможность индексации контента и полнотекстового поиска по содержимому текстовых документов. Поддерживаемые форматы: DOC, XLS, PPT, PDF, TXT, HTML.
- Отсутствует возможность ввода документа со сканера.
- Для хранения информации о связях между объектами SAP и контентом не используются таблицы TOA0x.

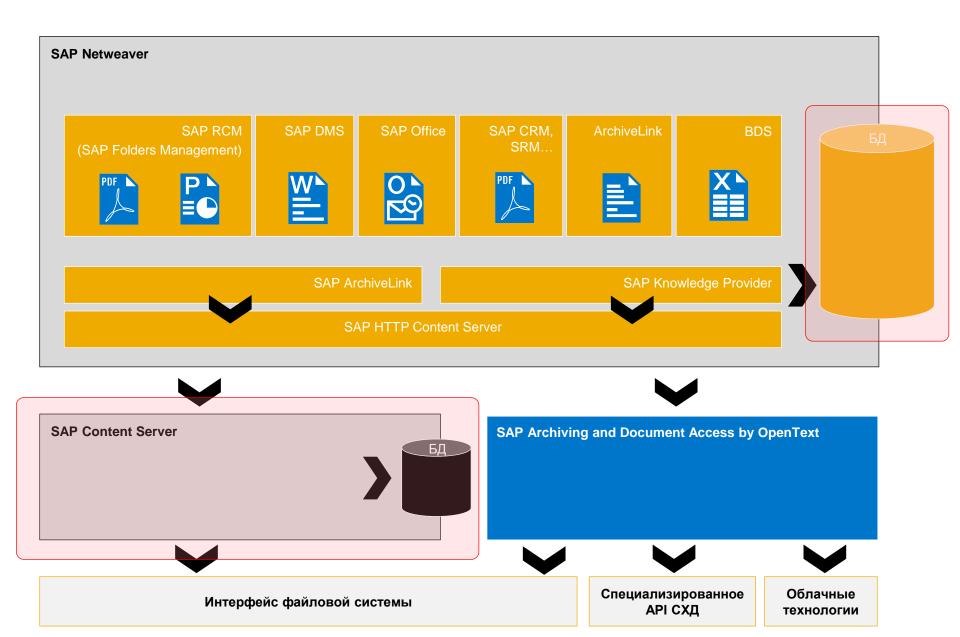






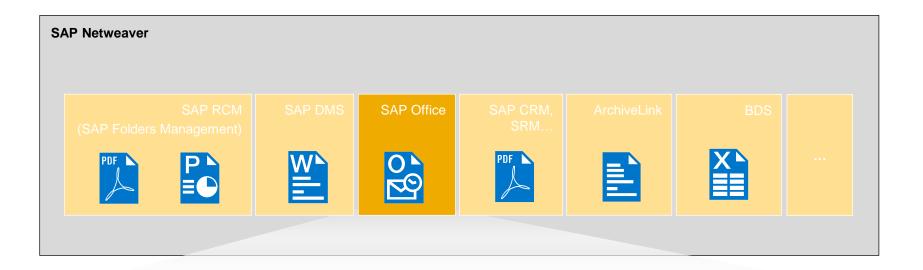








SAP Office

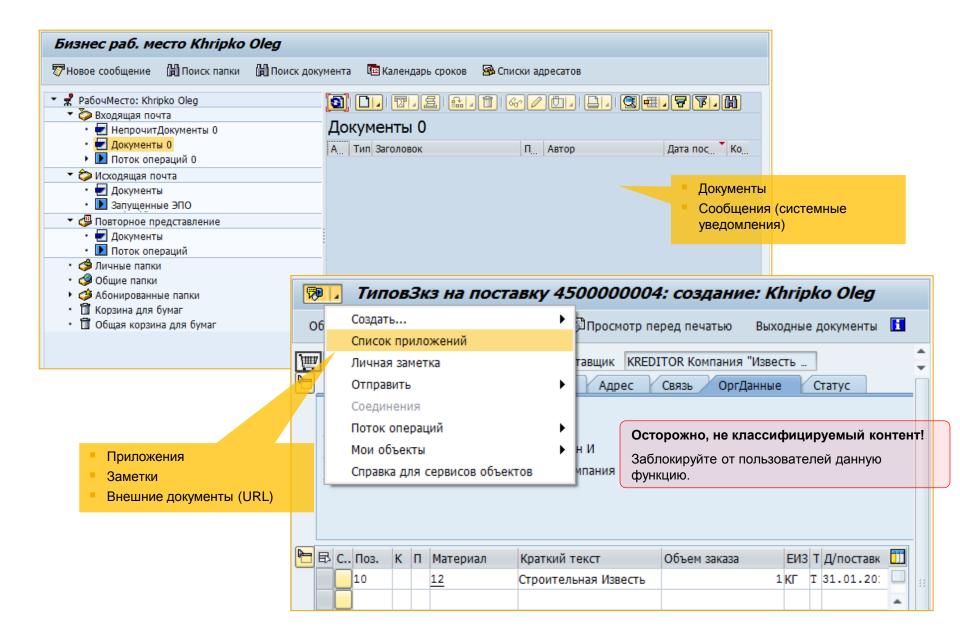


Механизм SAPOffice используется для хранения документов, сохраняемых в системе через общие папки и прикрепляемых к бизнес-объектам через GOS-меню. При этом, SAPOffice поддерживает следующие способы хранения документов:

- С использованием механизма KPro. Такие записи идентифицируются значением "K" в поле EXTCT таблицы SOOD. Применимо для бинарных файлов (формата, отличного от TXT). В случае использования механизма KPro, по умолчанию документы SAPOffice сохраняются в репозиторий SOFFDB, настроенный на хранение во внутренней таблице SAP SOFFCONT1.
- С использованием механизма ArchiveLink. Такие записи идентифицируются значением "A" в поле EXTCT таблицы SOOD. Применимо для документов, заархивированных через интерфейс Business Workspace либо для принтлистов (при условии, что для них настроена архивация).
- Напрямую в таблице SOC3. Такие записи идентифицируются значением " " в поле EXTCT таблицы SOOD. Применимо для примечаний, ссылок и текстовых файлов (формата TXT), сохраняемых в общих папках либо в качестве приложений к бизнес-объектам.

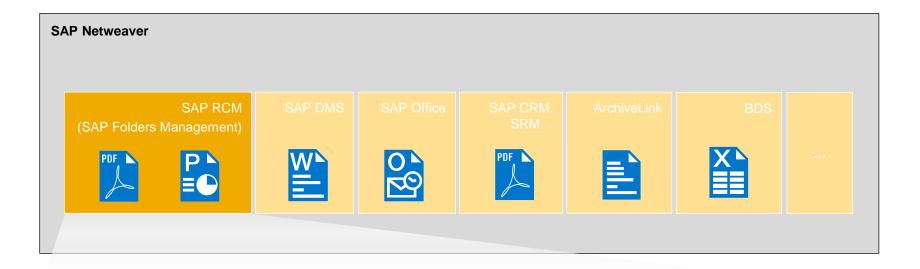


SAP Office





SAP RCM (SAP Folders Management)

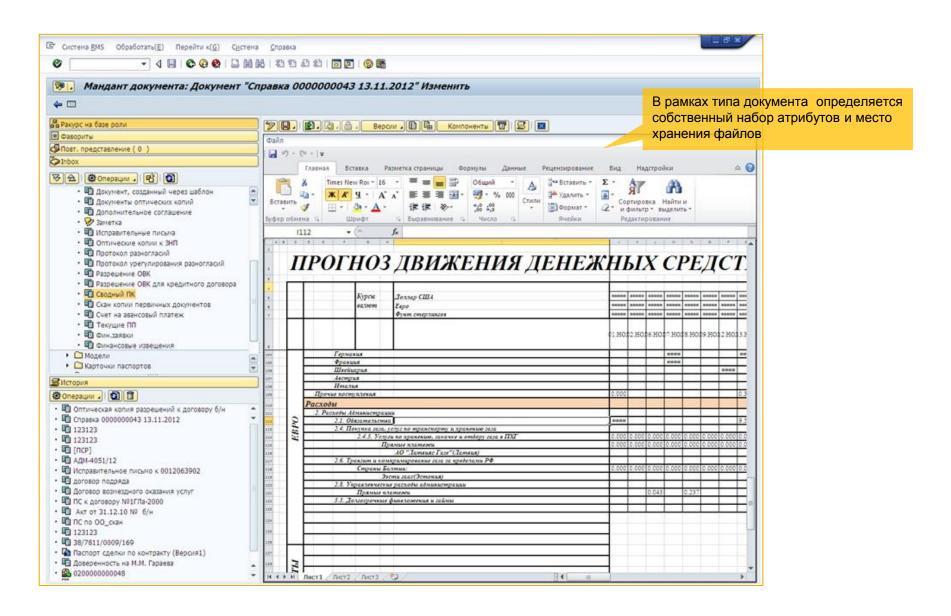


Механизм RCM предполагает работу с контентом (файлами) посредством карточки объекта и позволяет прикреплять к карточке документы различных типов, в зависимости от настроек типа карточки (тип карточки имеет название «модель документа»).

В рамках механизма RCM, каждый файл, сохраняемый в систему, представляет собой отдельный объект типа "DOCUMENT". Соответствующий объект (включая свойства и методы) доступен для анализа из транзакции SWO1. Для создания экземпляра объекта необходимо указать составной ключ документа, состоящий из физического класса и уникального идентификатора (GUID)

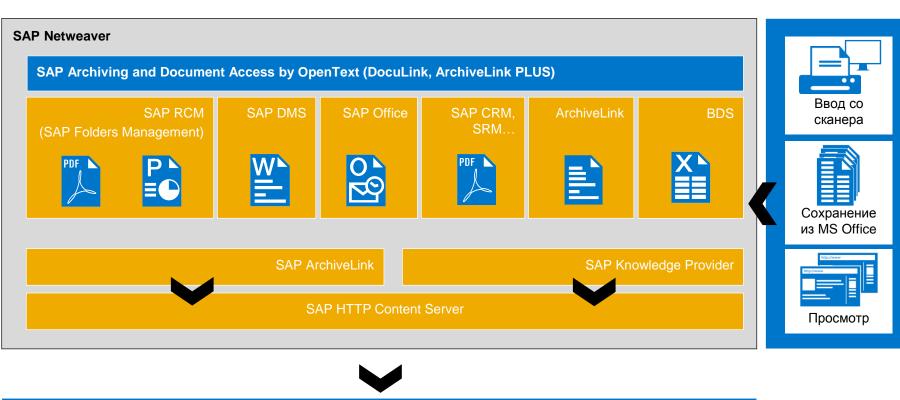


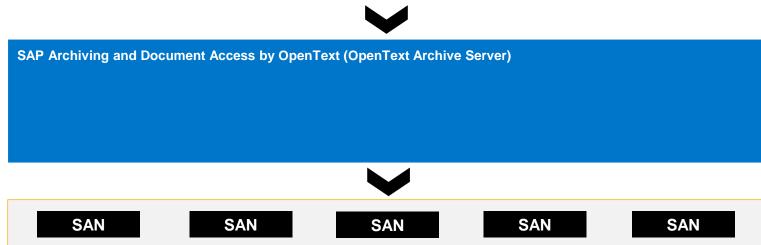
SAP RCM (SAP Folders Management)





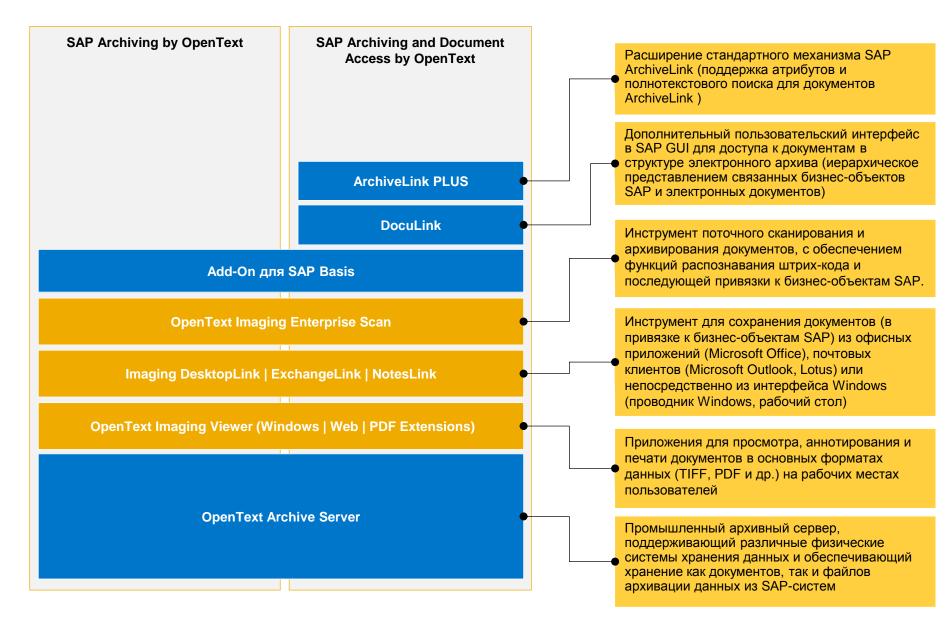
Pасширение стандартного функционала средствами SAP Archiving and Document Access by OpenText





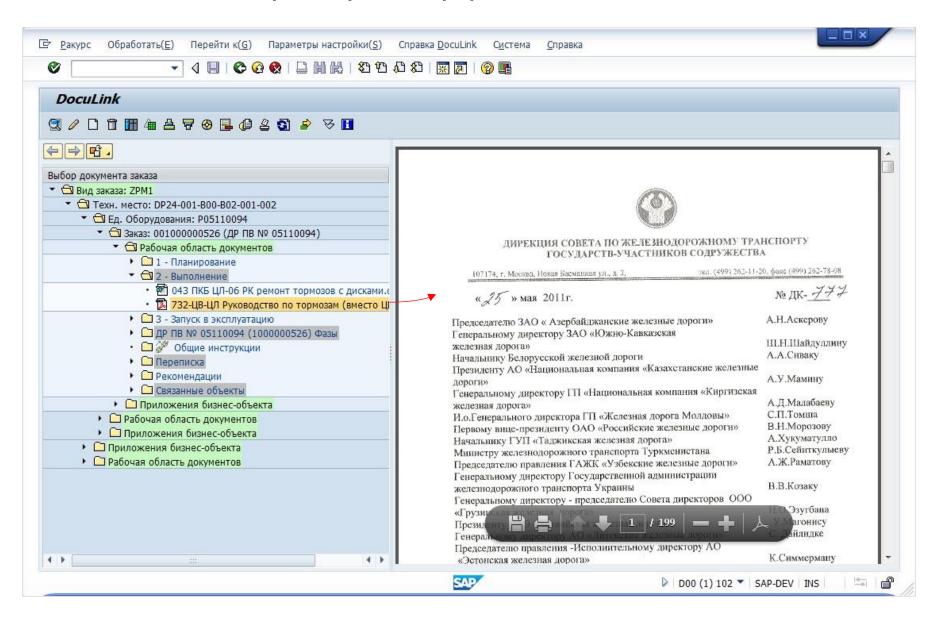


Компоненты SAP Archiving and Document Access by OpenText



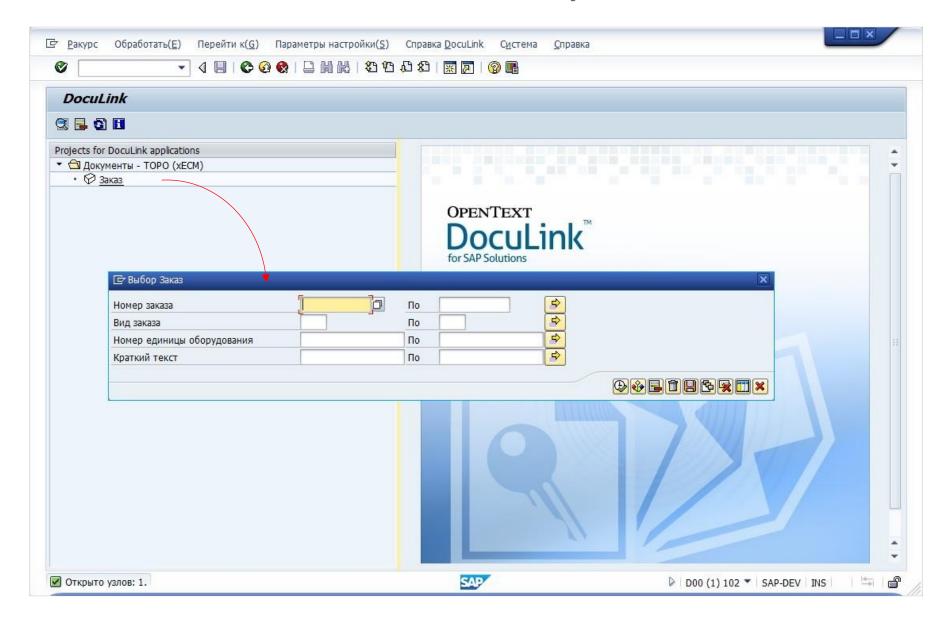


SAP DocuLink. Пример интерфейса



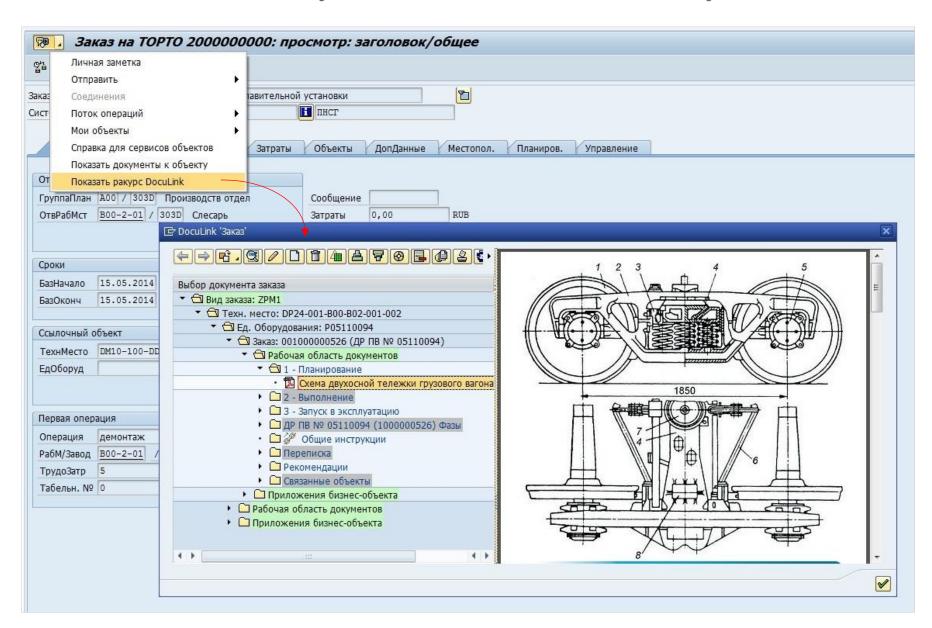


SAP DocuLink. Поиск связанных документов



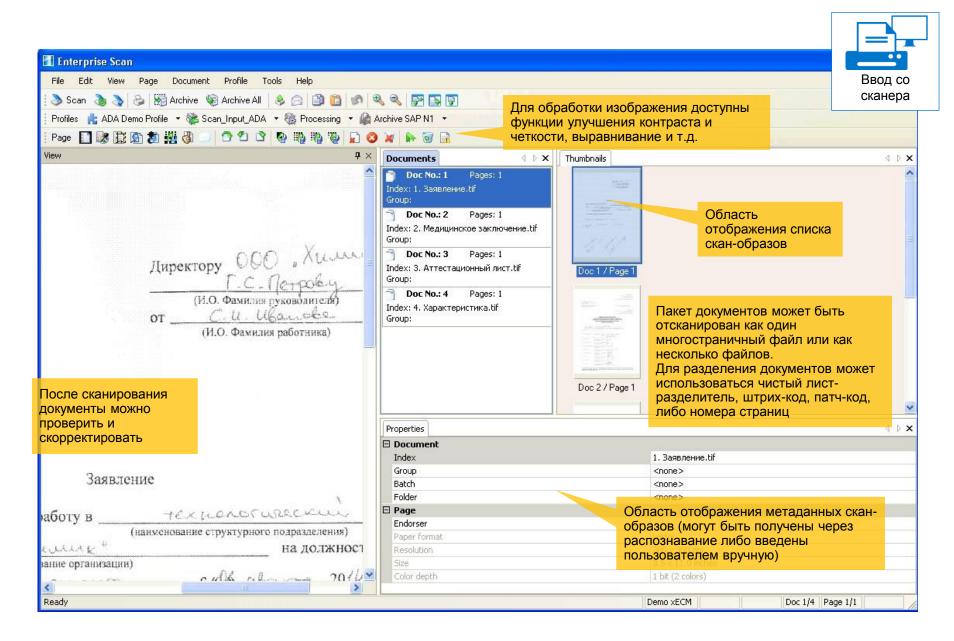


SAP DocuLink. Доступ ко всем связанным документам



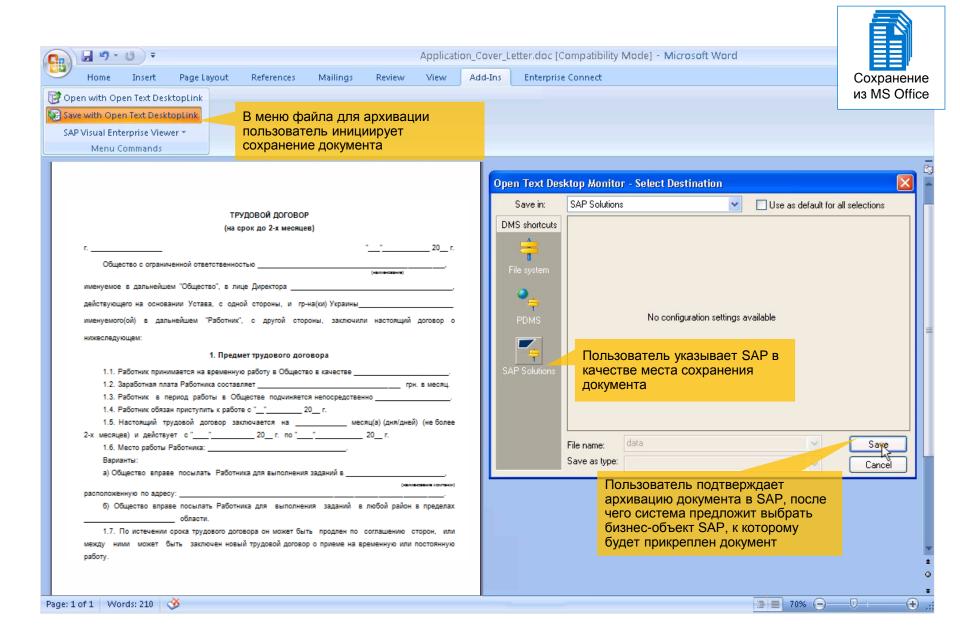


EnterpriseScan. Ввод документа со сканера



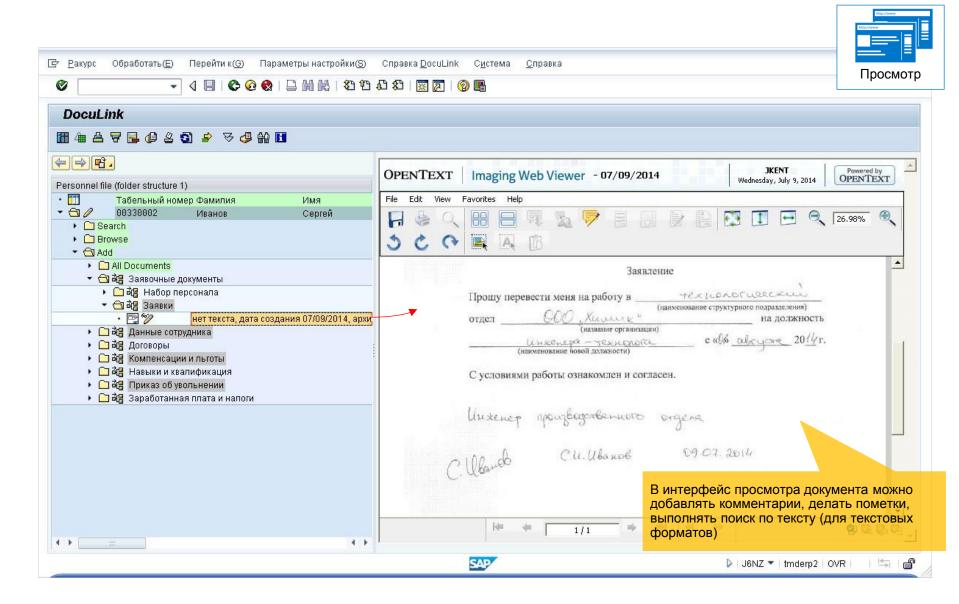


DesktopLink. Сохранение документа из MS Office





Web Viewer. Интерфейс просмотр документа





Содержание

- 1. Типовые задачи хранения контента в SAP
- 2. Инструменты и механизмы SAP по работе с контентом
- 3. База данных SAP как место хранения контента
- 4. Cpabhehue SAP Content Server u SAP Archiving by OpenText
- 5. Миграция контента во внешнее хранилище
- 6. Архивация данных в SAP



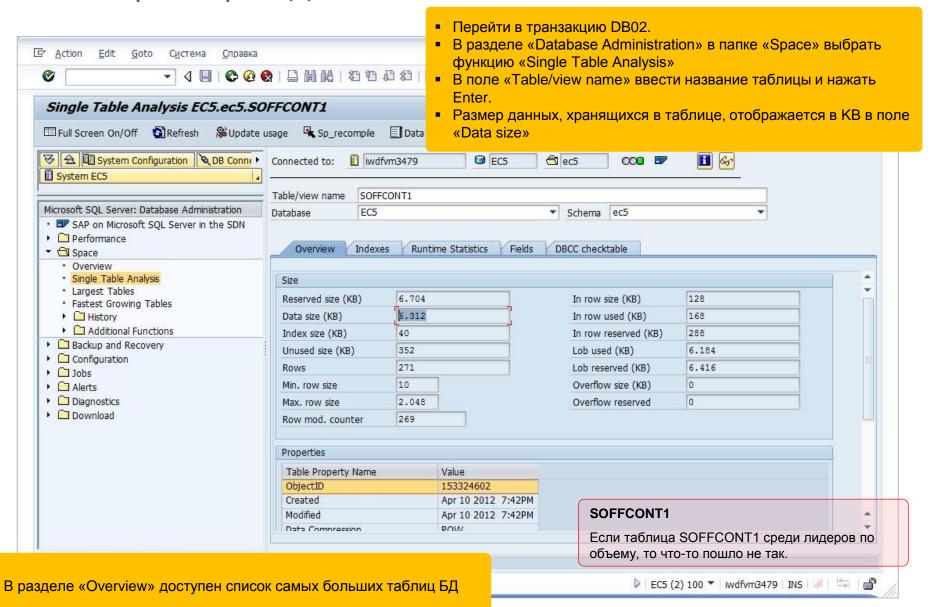
База данных SAP как место хранения контента





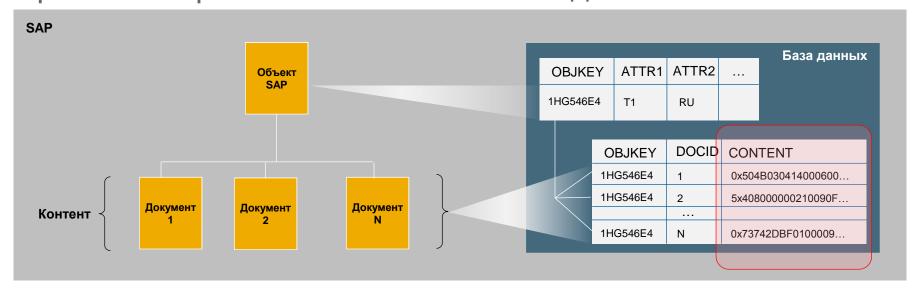


Анализ размера БД





Решение задачи хранения контента в SAP. Проблемы хранения контента в базе данных

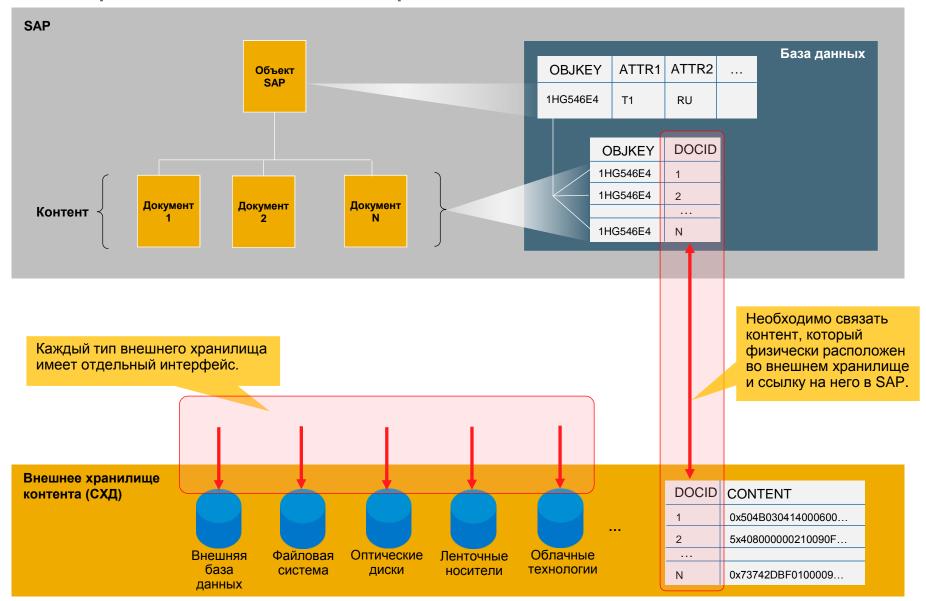


Проблемы, связанные с хранением контента в БД

- Снижение производительности работы продуктивной системы. С ростом объема таблиц в БД увеличивается время на выполнение запроса к этим таблицам.
- Ограничение на максимальный размер прикрепляемого файла. Максимальный размер бинарного файла, который можно сохранить в БД, в основном 2 ГБ (зависит от СУБД).
- Сложность администрирования базы данных. С ростом объема БД возникают, к примеру такие проблемы, как нехватка свободного места на дисках, где хранятся файлы БД и индексы, или увеличение времени и сложности выполнения ежедневного бэкапа.

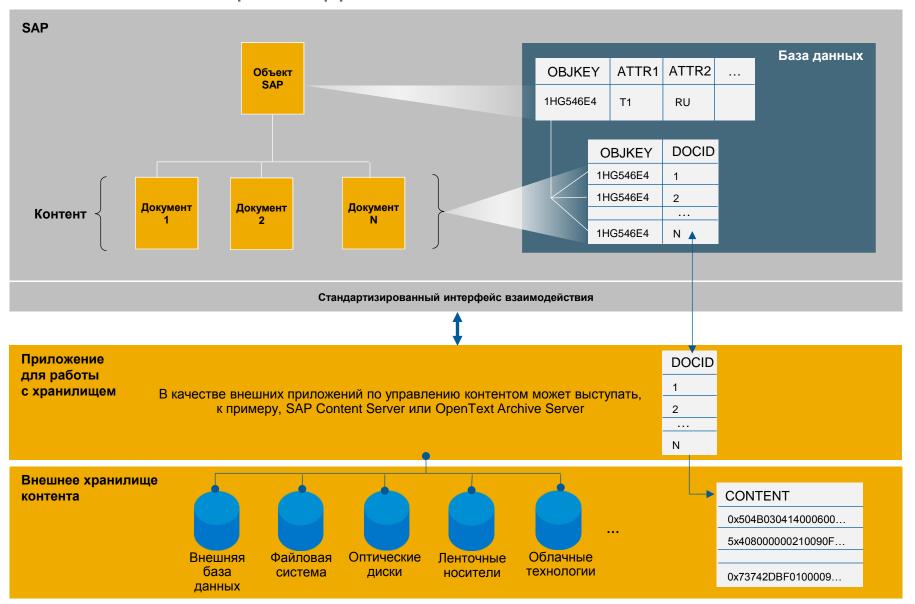


Решение задачи хранения контента в SAP. Альтернативный способ хранения контента





Решение задачи хранения контента в SAP. Решение в общем виде





Содержание

- 1. Типовые задачи хранения контента в SAP
- 2. Инструменты и механизмы SAP по работе с контентом
- 3. База данных SAP как место хранения контента
- 4. Сравнение SAP Content Server и SAP Archiving by OpenText
- 5. Миграция контента во внешнее хранилище
- 6. Архивация данных в SAP



Сравнение SAP Content Server и **SAP Archiving by OpenText**





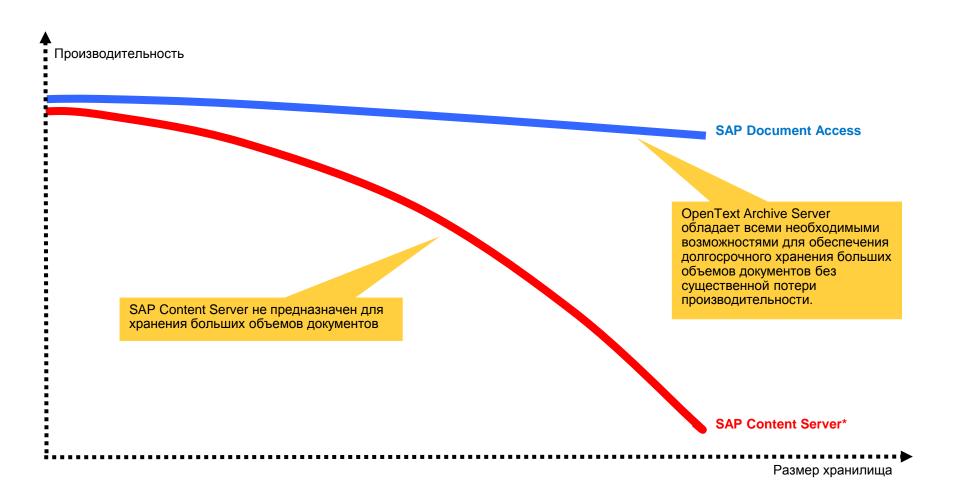


Проблемы, связанные с использованием SAP Content Server

- Снижение производительности хранилища при росте объемов и типов сканируемых документов может напрямую повлиять на скорость и эффективность работы пользователей
- Ограничения SAP Content Server на общий объем архива документов 64 Тб и на размер отдельного документа - 2 Гб (при хранении в БД).
- Отсутствие ряда функций, обеспечивающих безопасное и надежное хранение документов (нет встроенных средств резервирования, отсутствует механизмы репликации, нет механизма шифрования документов и поддержки жизненного цикла документа).
- Не подходит для реализации долговременного хранения документов (ограниченное количество поддерживаемых систем хранения данных, нет компрессии данных, нет поддержки преобразования в форматы длительного хранения документов).
- Отсутствие готового инструмента сканирования, отсутствие удобного пользовательского интерфейса для работы со скан-образами документов в SAP, отсутствие стандартных возможностей поиска документов ArchiveLink, отсутствие функции создания структуры хранилища скан-образов.



Сравнение SAP Content Server и SAP Archiving by OpenText

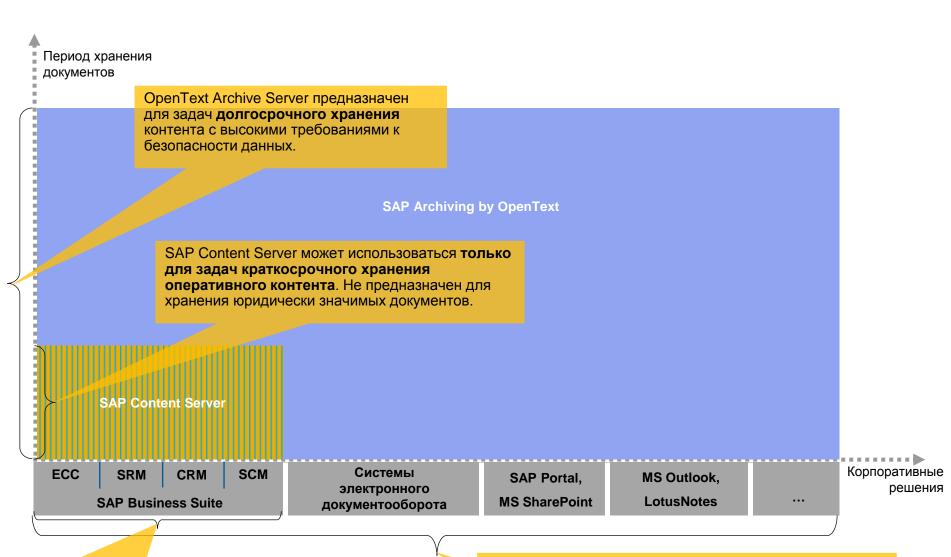


Детальное техническое сравнение приведено в Приложении

*Более подробная информация о продукте - см. SAP Note 719971



Сравнение SAP Content Server и SAP Archiving by OpenText



Использование SAP Content Server ограничено бизнес-приложениями SAP

OpenText Archive Server имеет **стандартные** коннекторы для хранения контента и предоставляет API для реализации хранения из **любых** корпоративных приложений и в любых бизнес-сценариях.



Содержание

- 1. Типовые задачи хранения контента в SAP
- 2. Инструменты и механизмы SAP по работе с контентом
- 3. База данных SAP как место хранения контента
- 4. Cpabhehue SAP Content Server u SAP Archiving by OpenText
- 5. Миграция контента во внешнее хранилище
- 6. Архивация данных в SAP



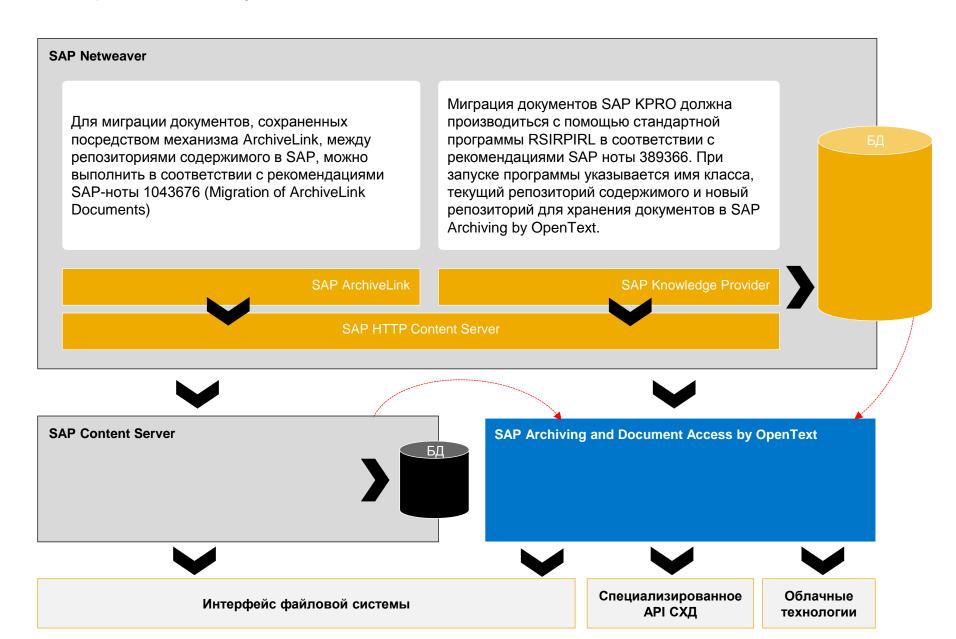
Миграция контента во внешнее хранилище







Миграция документов в SAP





Содержание

- 1. Типовые задачи хранения контента в SAP
- 2. Инструменты и механизмы SAP по работе с контентом
- 3. База данных SAP как место хранения контента
- 4. Cpabhehue SAP Content Server u SAP Archiving by OpenText
- 5. Миграция контента во внешнее хранилище
- 6. Архивация данных в SAP



Архивация данных







Цели и задачи архивации данных

Основная цель архивации данных – снизить затраты, связанные с большим размером БД SAP

Доступность системы

Из-за большого объема данных увеличивается время обновления и восстановления системы.

Потребление ресурсов

Речь не только о размере места на дисках, но также и о человеческих ресурсах на администрирование и т.д.

Производительность

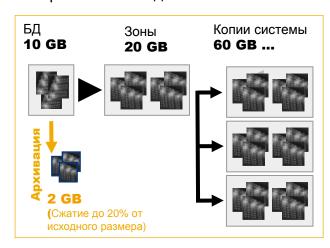
Время отклика системы снижается при большом объеме данных в БД.



Для монетизации необходимо оценить стоимость 1 часа простоя компании.

Для монетизации необходимо оценить совокупную стоимость хранения 1 ТБ данных.

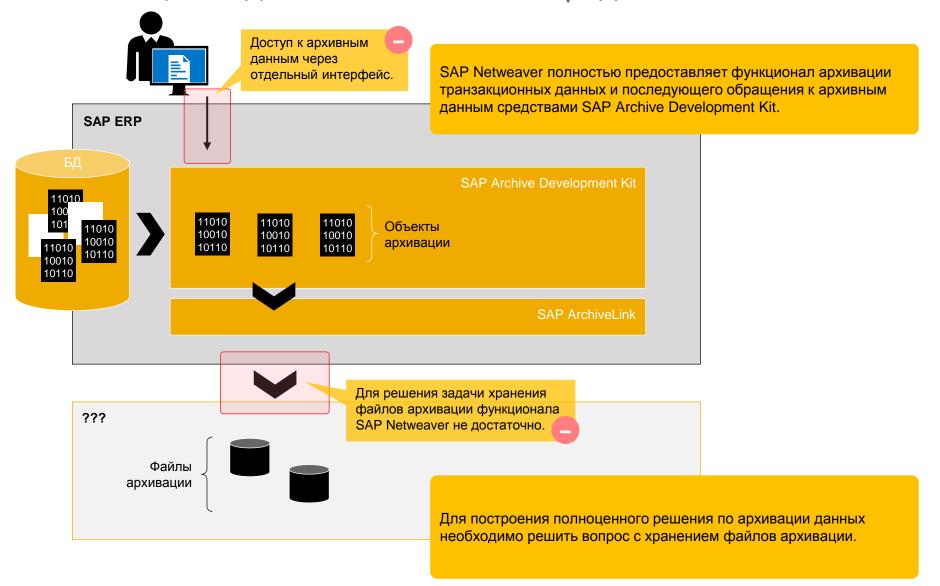
Сложно монетизировать, отрицательный эффект – недовольство пользователей.



Задача архивации – сделать естественный процесс роста объемов БД SAP управляемым. Для реализации данной задачи в должна быть разработана и внедрена стратегия архивации данных.

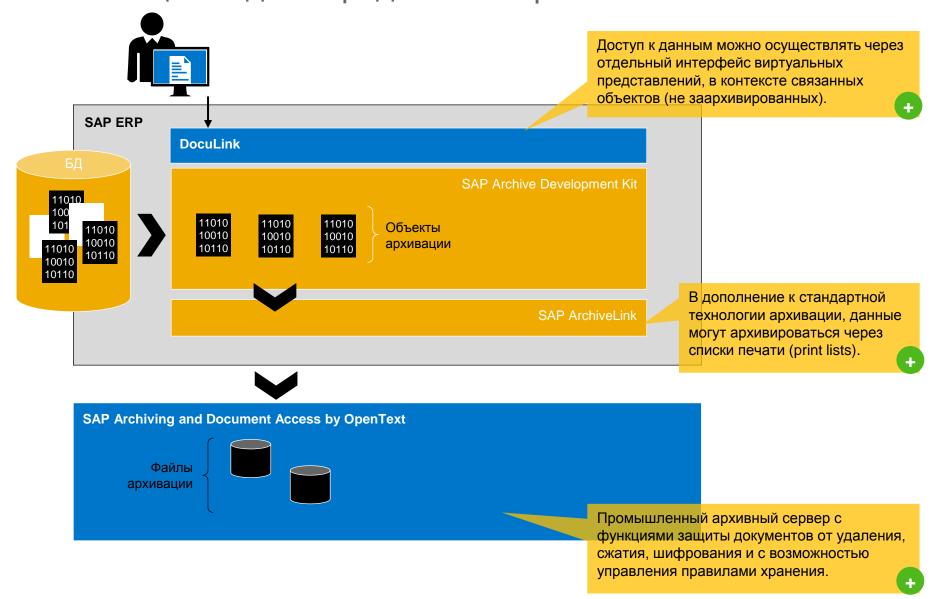


Выгоды от использования SAP Archiving by OpenText. Реализация задачи собственными средствами SAP



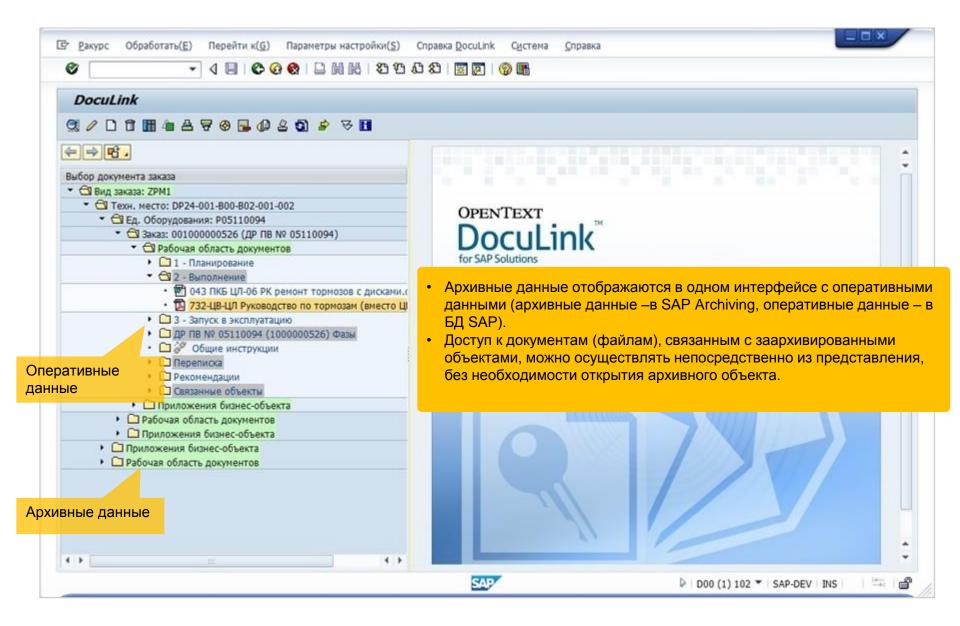


Выгоды от использования SAP Archiving by OpenText. Реализация задачи средствами OpenText



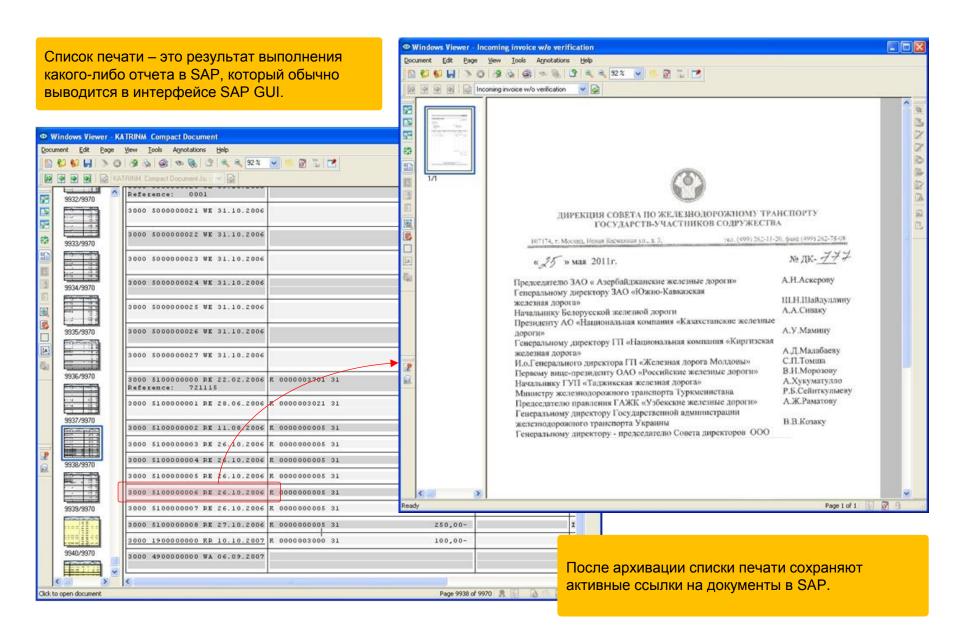


SAP DocuLink. Вариант доступа к архивным данным





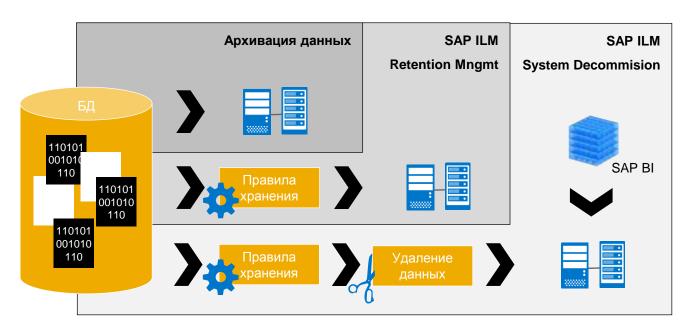
Архивация данных через списки печати





Поддержка SAP NetWeaver Information Lifecycle (ILM)

SAP ILM – результат эволюции подходов к хранению, управление жизненным циклом



- SAP ILM не имеет собственного хранилища для хранения архивных данных.
- SAP ILM не имеет собственного хранилища для хранения документов.
- SAP Archiving and Document Access полностью поддерживает решение SAP NetWeaver ILM
- Архивация данных производится посредством интерфейса ILM WebDAV



- SAP Archiving and Document Access by OpenText

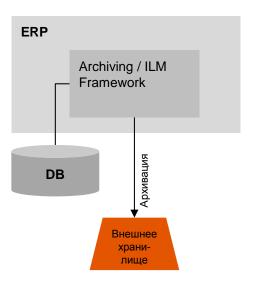


Проблема архивации аналитических данных

Для архивного хранения разных типов данных требуются различные подходы:

Транзакционные данные

Данные могут храниться в традиционных системах хранения.

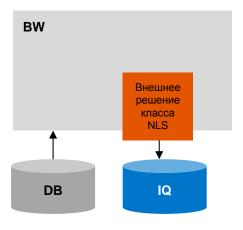


SAP Archiving and Document Access by OpenText

Аналитические данные

Для хранения требуется решения класса «Near-Line Storage».

Задача: доступ к архивным данным для анализа



SAP IQ Near-Line Storage



Проект архивации данных. Пример подхода

Анализ и проектирование

5-40 дней

Анализ внутренней БД SAP.

Подготовка проектного решения по архивации.

Анализ и определение для каждого объекта архивации условий архивации.

Типовая ошибка

заключается в том, что недостаточно внимания на данном этапе уделяется требованиям к хранению данных, которые могут включать законодательные нормы, требования к доступности данных в оперативном режиме, к срокам хранения, и способам доступа.

Настройки инфраструктуры архивации

10-40 дней

На данном этапе должна быть полностью реализована инфраструктура архивации (состоит из сервера архивного хранения контента, платформы физического хранения, настроек механизма SAP ArchiveLink).

Настройки для объектов архивации.

Тестирование архивации

Продуктивная архивация

10-40 дней

Выполняется архивация и мониторинг архивации.

Создаются необходимые для работы с архивами структуры.

Фиксируются результаты архивации (число с архивированных документов, объём удалений из базы данных и тд).

Формируется перечень рекомендованных базисных работ для уменьшения базы данных (перестроения индексов, реорганизации таблиц).

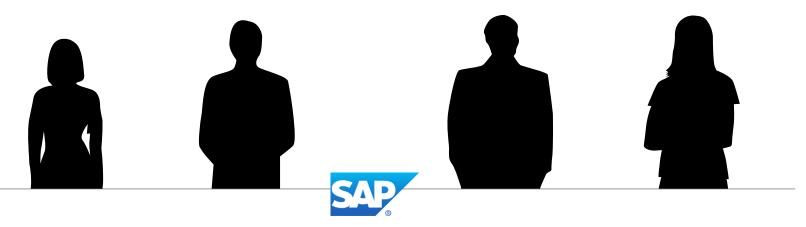
Передача знаний, обучение

5-10 дней

Обучение подразумевает передачу знаний по администрированию сервера промышленного хранения архивных файлов и настройки связи с SAP, в достаточном объеме для поддержки настроенной инфраструктуры архивации и для ее развития, т.е. для добавления новых сценариев архивации



Мы с радостью ответим на Ваши вопросы



Команда SAP Consulting по управлению контентом и по решениям OpenText



Максим Лысенко

Руководитель группы maxim.lysenko@sap.com +7 906 710 3524



Олег Хрипко

Ведущий консультантoleg.khripko@sap.com+7 903 136 7675



Приложение





Сравнение SAP Content Server и OpenText Archive Server. Критерии сравнения

Вводная информация

SAP Content Server – стандартный компонент SAP для хранения документов. Согласно официальной информации производителя, не предназначен для долгосрочного хранения.

OpenText Archive Server – система управления хранением для безопасного, долгосрочного, централизованного хранения документов. Входит в состав продуктов SAP Archiving & Document Access и SAP xECM by OpenText.

Группы критериев сравнения

- **Требования к стороннему ПО**. Данная группа критериев сравнения описывает поддерживаемое внешнее ПО, необходимое для работы каждого из репозиториев.
- Реализация функции хранения документов. Данная группа критериев позволяет сравнить репозитории с точки зрения реализации основной функции хранения контента.
- Обеспечение долгосрочного хранения документов. Данная группа критериев позволяет сравнить репозитории с точки зрения возможностей долгосрочного хранения контента.
- Производительность. Критерий позволяет сравнить репозитории с точки зрения обеспечения производительности с ростом размера хранилища.
- Реализация масштабирования хранилища. Данная группа критериев позволяет сравнить возможности репозиториев по функционированию в распределенной среде и под большими нагрузками.
- Реализация отказоустойчивости и предотвращение потери данных. Данная группа критериев позволяет сравнить репозитории надежности работы.
- Реализация безопасного хранения документов и аудита. Данная группа критериев позволяет сравнить репозитории с точки зрения контроля доступа к документам и обеспечения безопасного долгосрочного хранения.
- Реализация возможностей администрирования. Данная группа критериев позволяет сравнить репозитории с точки зрения удобства и полноты средств администрирования.



Группа критериев сравнения



Требования к стороннему ПО

ответствие критериям		
Критерий	SAP Content Server	OpenText Archive Server
Поддерживаемые типы ОС	MS Windows	
	Solaris	
	HP-UX	
	AIX	
	RedHat	
Поддерживаемые СУБД		Oracle
		MS SQL Server
	MaxDB	SAP HANA
Необходимый сервер приложений / web- сервер		Oracle WebLogic
	Apache Tomcat	
	IIS	



Группа критериев сравнения



Реализация функции хранения документов

ответствие критериям		
Критерий	SAP Content Server	OpenText Archive Server
Хранение файлов больших размеров	Существующее ограничение на максимальный размер файлов - 2 GB	Подтверждена тестированием возможность загрузки файлов размером 100 GB
Логические разделы хранилища	Хранилище состоит из логических разделов для хранения разных типов документов с индивидуальными настройками правил хранения. Организация хранения внутри логического раздела скрыта для SAP.	Хранилище состоит из логических разделов для хранения разных типов документов и настроек правил хранения. Для документа может быть определен ярус хранения (конкретный диск внутри логического раздела)
Компрессия (сжатие контента)	Уровень сжатия отличается в зависимости от формата файла. В среднем, процент сжатия – 20-25%.	Уровень сжатия имеет величину «gzip level 6», что соответствует сжатию, в среднем в 2 раза.
Дедупликация (предотвращение дублирования)	Данная функция отсутствует.	Каждый файл физически сохраняется только один раз (Single Instance Archiving). При повторном сохранении используется ссылка на уже существующий файл.
		Оптические устройства хранения, дисковые библиотеки
		Облачные системы хранения
Поддерживаемые способы хранения		Специализированные системы хранения Oracle, EMC, IBM, NetApp
	Файловая система	
	База данных	



Группа критериев сравнения



Обеспечение долгосрочного хранения документов

Соответствие критериям		
Критерий	SAP Content Server	OpenText Archive Server
Поддержка технологии WORM	Не поддерживается.	Для хранения контента могут использоваться оптическое носители с однократной записью и многократным чтением, без возможности удаления информации.
Защита документа от удаления на протяжении определенного периода хранения	Данная функция отсутствует.	Все документы в хранилище могут быть защищены от удаления пользователями (включая администраторов) на протяжении всего срока хранения (retention period), в соответствии с требованиями бизнеса.
Миграция контента на более дешевые носители (по принципу HSM)	Данная функция отсутствует.	По истечению периода оперативной работы с документом, документ средствами OpenText Archive Server может быть перемещен на более дешевые носители для долгосрочного хранения (например, оптические устройства и ленты).



Группа критериев сравнения



Производительность

Соответствие критериям

Производительность OpenText **Archive Server** Результаты нагрузочного тестирования подтверждают возможности OpenText Archive Server по хранению больших объемов документов без существенных потерь производительности. По информации от производителя, SAP Content Server не предназначен для хранения больших объемов документов. **SAP Content** Server



Группа критериев сравнения



Реализация масштабирования хранилища

Соответствие критериям		
Критерий	SAP Content Server	OpenText Archive Server
Возможность хранения больших объемов документов	При хранении контента в БД (данный способ хранения рекомендован SAP) ограничение на размер хранилища - 64 ТВ .	Ограничения по объему документов отсутствуют.
Масштабирование нагрузки по количеству пользователей	Обеспечивается за счет горизонтального масштабирования средствами load balancer.	
Масштабирование нагрузки по количеству документов	Обеспечивается путем установки дополнительного инстанса БД.	Реализуется путем установки дополнительного инстанса Archive Server и объединения всех инсталляций в логический централизованный Archive Server.
Территориальное масштабирование	Возможность полноценной работы с хранилищем для территориально удаленных пользователей обеспечивается за счет установки Cache-сервера.	



Группа критериев сравнения



Реализация отказоустойчивости и предотвращение потери данных

Соответствие критериям		
Критерий	SAP Content Server	OpenText Archive Server
Поддержка кластера Active- Passive	Для возможности работы с документами в случае выхода из строя какого-либо из компонентов или всей системы, установка производится в режиме кластера с высокой доступностью (High Availability Cluster). Поддерживается аппаратный кластер соответствующей операционной системы либо кластер на основе Solaris Zones.	
Средства контроля целостности данных	Данная функция отсутствует.	Поскольку атрибуты документов внутри Archive Server хранятся в БД, а сам контент – на внешних СХД, то при восстановлении из горячего бэкапа возможны расхождения между данными в БД и в хранилище. В этом случае все расхождения можно исправить с помощью встроенной функции контроля целостности данных.
Поддержка удаленной репликации	Данная функция отсутствует.	Возможна синхронная репликация данных на отдельную удаленную инсталляцию Archive Server – технология Remote Standby. В случае отказа основного инстанса Archive Server, пользователи смогут продолжать работу с документами через Remote Standby.



Группа критериев сравнения



Реализация безопасного хранения документов и аудита

оответствие критериям		
Критерий	SAP Content Server	OpenText Archive Server
Поддержка безопасных каналов доступа	Поддержка SSL.	
Подпись запросов к хранилищу (SecKey)	Все запросы к хранилищу из SAP подписываются его закрытым ключом, и содержат срок действия подписи. Для проверки подписи на стороне хранилища используется сертификат SAP (открытый ключ). Благодаря данному механизму, доступ к документам возможен только из SAP, все запросы «напрямую» исключены.	
Шифрование документов в хранилище	Данная функция отсутствует.	Шифрование может быть включено на уровне отдельного логического хранилища. Расшифровка контента выполняется сервером в момент чтения документа.
Временные метки (timestamps)	Данная функция отсутствует.	Для данной функции может использоваться любой внешний Timestamp Service, который для хэша документа генерирует временную метку (timestamp) и подписывает их объединение своим ключом.
Защита документов от удаления	Отсутствует.	Да.
Аудит доступа к данным	Только при использовании БД в качестве хранилища.	Да.