



agorum core **optimal betreiben.**

Holen Sie das Beste aus Ihrem
Dokumentenmanagement-System heraus

Plattensystem - Prozessor - RAM - Datenbanken - Betriebssystem -
DMS-Laufwerk (SMB/CIFS) - Datensicherung - Monitoring

Hard- und Software Infrastruktur

Abhängig davon, was Sie mit Ihrem Dokumentenmanagement-System erreichen möchten, benötigen Sie eine dafür ausgelegte Hard- und Software-Infrastruktur. Mit diesem Dokument geben wir Ihnen eine Hilfe an die Hand, mit der Sie den optimalen Nutzen aus Ihrem DMS ziehen.

Bauen Sie sich eine wirtschaftlich sinnvolle Infrastruktur auf, die es Ihnen erlaubt, bestmöglich mit *agorum core* zu arbeiten und Ihre Datenbank und Dokumente zuverlässig zu sichern.



Zunächst einmal ist es wichtig, zu wissen, was Sie mit *agorum core* erreichen möchten. Anhand Ihrer Anforderungen können Sie in diesem Dokument nachlesen, mit welcher Konfiguration Sie am besten fahren.

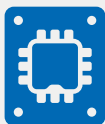
Bevor wir ins Detail gehen, sollten Sie sich daher selbst ein paar Fragen beantworten:

- **Wie viele Nutzer sollen mit dem DMS arbeiten?**
- **Soll das DMS als der zentrale Arbeitsplatz für die Mitarbeiter genutzt werden, die den Tag über nach Dokumenten suchen und diese öffnen, bearbeiten und wieder speichern?**
- **Oder nutzen Sie das DMS hauptsächlich zur Archivierung von Belegen?**
- **Welche Geschäftsprozesse sollen automatisiert werden?**
- **Möchten Sie Dokumente automatisiert erfassen? Wenn ja, wieviele Dokumente müssen pro Tag per OCR erfasst werden und wie zeitkritisch ist die Erfassung? Müssen zum Beispiel in Spitzenzeiten viele Dokumente schnell erfasst werden oder können die Dokumente über den Tag verteilt im Hintergrund abgearbeitet werden?**

Kurzübersicht Hardwareinfrastruktur*

*Alle Werte beziehen sich auf die durchschnittliche Nutzung eines DMS, wie sie im Unternehmensalltag üblich ist.

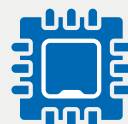
Plattensystem



Für den performanten Betrieb einer datenbankgestützten Software ist es unerlässlich, ein schnelles Festplattensystem zu haben.

Daher empfehlen wir, die agorum core-Datenbank und den Suchindex auf **SSD-Platten** zu betreiben. Die im Betrieb wichtigen **IOPS (Input/Output Operations Per Second)** sind auf einer SSD um ein Vielfaches höher als auf HDD-Platten. Den **Content können Sie ohne Weiteres auf ein Storage mit drehende Platten auslagern** und haben immer noch ein hochperformantes System ohne Einschränkungen. Es spricht aber nichts dagegen, Ihren Server **komplett mit SSD-Platten** auszustatten.

Prozessoren



Die benötigte Prozessorleistung hängt von der **Anzahl der Personen** ab, die mit **agorum core** arbeiten und **wie viele Dokumente Sie per OCR** ins DMS einlesen möchten. Als Faustformel gilt, dass Sie für ein performantes System vier Kerne mit hoher Taktung als Basiskonfiguration einkalkulieren sollten und pro 1000 Anwender weitere vier Kerne. Dazu kommen weitere Kerne für die OCR. Hier sollten Sie berechnen, wie viele Dokumente Sie in Spitzenzeiten abarbeiten müssen, pro Seite benötigt die OCR im Schnitt ca. 10 Sek. (abhängig von der Taktung) – mit jedem weiteren Kern vervielfachen Sie die Leistung entsprechend.

Arbeitsspeicher



Als Basis sollten für jeden agorum core Server mindestens **18 GB RAM** zu Verfügung stehen. 8 GB für die Javamaschine + 4 GB für die Solr-Suchmaschine + 1 GB für den Zookeeper + 2GB für die Datenbank. Der Rest ist für das Betriebssystem und dient als Cache für Solr und das Filesystem. Je mehr Personen gleichzeitig mit dem System arbeiten und je größer die Datenmenge wird, sollte auch der Speicher entsprechend erhöht werden. Weitere Details finden Sie auf der Seite 7.

Datenbank



Als Datenbank empfehlen wir aus langjähriger Erfahrung eine **MySQL-Datenbank**. Sie hat sich im Livebetrieb als vergleichsweise wartungsarm, zuverlässig und performant gezeigt. Eine Beschleunigung ist sehr einfach über die Erhöhung des Hauptspeichers möglich. **Insbesondere bei großen Datenmengen empfehlen wir den Einsatz von MySQL.**

Sonstiges

Detaillierte Infos zu den oben genannten Punkten und was Sie in Ihrer Infrastruktur noch beachten sollten, haben wir Ihnen auf den folgenden Seiten ausführlicher zusammen gestellt. Wir informieren Sie über **Serverbetriebsysteme, die Spezialitäten des DMS-Laufwerks, das Monitoring und die optimale Datensicherung**. Damit Ihrem perfekten DMS-Betrieb nichts im Wege steht!

Plattensystem

Eine Investition in ein schnelles Festplattensystem macht sich sehr schnell bezahlt – Ihren Mitarbeitern stehen alle Informationen innerhalb von Sekunden zur Verfügung, sie vergeuden keine Zeit mit Warten und sie arbeiten gerne mit dem DMS. Ein nicht zu unterschätzender Vorteil, der Ihnen viel Zeit und Geld spart!



Verhältnis SSD + HDD

agorum core ist eine datenbankgestützte Software. Daher hängt die Gesamtperformance des Systems stark davon ab, wie schnell die Festplattensysteme sind, auf denen die Datenbank läuft und die Dokumente gespeichert werden. Wir empfehlen, zumindest die *agorum core*-Datenbank und den Suchindex auf schnellen SSD-Platten zu betreiben. Der Grund sind die im laufenden Betrieb wichtigen IOPS (Input/Output operations Per Second) die auf einer SSD um ein Vielfaches höher sind als auf herkömmlichen, drehenden HDD-Platten.

Das DMS erzeugt pro Einzelanwender, der auf *agorum core* zugreift, viele kleine Anfragen, die das DMS sequentiell abarbeiten muss. Dieses Lesen und Schreiben vieler kleiner Blöcke hintereinander erzeugt viele IOPS, die von SSD-Platten schneller verarbeitet werden können als von HDD-Platten, auch wenn diese im RAID-Verbund laufen.

HDD	RPM	IOPS
SATA	5400	50
SATA	7200	80
SAS	10k	140
SAS	15k	175
SSD	-	4.000
Lesen im HDD-RAID (SAS 15K) - RAID 10: 10 x 175 IOPS		1750
Schreiben im HDD-RAID (SAS 15K) - RAID 10: (10 x 175 IOPS) / 2		875
Lesen im SSD-RAID - RAID 10: 10 x 4.000 IOPS		40.000
Schreiben im SSD-RAID - RAID 10: (10 x 4000 IOPS) / 2		20.000

Die Tabelle zeigt, wieviel schneller SSD-Platten sind.

Selbst im RAID-Verbund, in dem auf mehreren Platten parallel gelesen und geschrieben werden kann, erreichen drehende Platten nur einen Bruchteil dessen, was SSD-Platten zu leisten vermögen.

Maximale Geschwindigkeit mit SSD-Platten

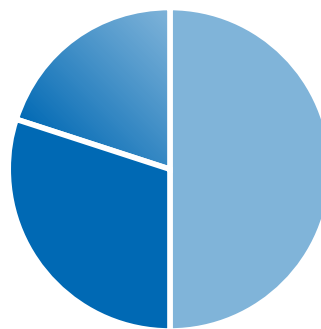
Wie aus der Tabelle zu ersehen, erreichen Sie die maximale Leistung mit einem mit SSD-Platten ausgerüsteten System. Das ist unsere Empfehlung, wenn viele Anwender intensiv mit dem DMS arbeiten. Also Arbeitsumgebungen, die *agorum core* voll ausnutzen, indem viele Mitarbeiter im Arbeitsalltag Dokumente suchen, öffnen, speichern und über automatische Prozesse verarbeiten. SSD-Platten tragen hier einen Großteil dazu bei, die Performance hoch zu halten, schnell und flüssig zu arbeiten und die Zufriedenheit der Mitarbeiter zu steigern.

Hochperformantes System im Mischbetrieb mit SSD und HDD

Ein wirtschaftliches, hochperformantes System können Sie betreiben, indem Sie Ihren Content aus der *agorum core* Datenbank auf drehende Platten auslagern. Die Datenbank läuft in diesem Fall auf SSD-Platten und der Content wird mit dem *agorum core storage* Modul auf ein externes Storage mit herkömmlichen HDD-Platten geschrieben. Damit haben Sie ein hochperformantes System, das im Preis-Leistungsverhältnis aktuell etwas günstiger ist als eine Komplettausstattung mit SSD, aber der Leistung praktisch in nichts nachsteht. Der Platzbedarf kann, je nach Nutzung des DMS, variieren.

Typische Verteilung der Daten bei 1 TB Datenvolumen (gesamt)

Je kleiner die einzelnen Dateien sind, die im DMS gespeichert werden, desto höher ist im Verhältnis der Platzbedarf für Datenbank und Index.



Datenbank + Index SSD
(ca. 300 - 500 GB)

Content HDD
(ca. 500 - 700 GB)

Speicherbedarf bei agorum core

Ein DMS braucht naturgemäß mehr Speicherplatz als ein gewöhnlicher Fileserver. Das liegt daran, dass *agorum core* zum Beispiel Versionen von Dokumenten anlegt, gelöschte Objekte für eine gewisse Zeit im Serverpapierkorb hält, Metadaten mit abspeichert oder Vorschaubilder erzeugt, um schnell darauf zugreifen zu können. **Daher können Sie für den Platzbedarf ungefähr das 1,5 - 2-fache rechnen wie für einen Fileserver, der lediglich die Dokumente vorhält.** (Achtung, bei Nutzung von Previews kann dieser Faktor weiter steigen)

Für die Archivierung von E-Mails rechnen Sie ungefähr das doppelte Datenvolumen eines herkömmlichen Mailserver. Das ist der Tatsache geschuldet, dass wir den Fokus darauf legen, den Anwendern Informationen so schnell wie möglich zur Verfügung zu stellen. Daher legen wir E-Mails einmal komplett ab und zusätzlich die Anhänge und E-Mail-Bodys nochmal als Einzeldokumente, um deren Anzeige und die Suche danach zu beschleunigen.

Einfluß auf den Speicherverbrauch können Sie nehmen, indem Sie Previews nur für einen gewissen Zeitraum vorhalten. Wie lange, ist individuell einstellbar, danach werden die Previews gelöscht und erst wieder erzeugt, wenn das Preview wieder angefordert wird. Auch die Größe des Serverpapierkorbs hängt davon ab, wie lange Sie gelöschte Dokumente im Papierkorb zwischenlagern möchten, die Zeitspanne ist auch hierfür frei definierbar. Metadaten machen sich nur bemerkbar, wenn sehr viele an einem Dokument hängen. Vererbte Metadaten belasten den Speicher kaum und sind zu vernachlässigen.

Tip: Überwachen Sie die Kapazität Ihrer Festplattensysteme und reagieren Sie frühzeitig, wenn der Platz zur Neige geht. Vollgelaufene Platten sind ein häufiger Grund für Störungen.

Prozessoren

Je mehr Nutzer mit Ihrem Dokumentenmanagement-System arbeiten und je mehr Dokumente Sie per OCR erfassen müssen, um so mehr Prozessorkerne sollten Sie Ihrem Server gönnen. Dabei kommt es nicht nur auf die Anzahl der Kerne an, sondern auch auf die Taktrate eines einzelnen Kerns. Viele sequentielle Einzelaktionen profitieren von höheren Frequenzen, so dass die Performance stark ansteigt. Für ein performantes System empfehlen wir daher ein gutes Verhältnis zwischen Anzahl der Kerne und der Taktfrequenzen. Die dargestellten Berechnungen sind beispielhaft für ein durchschnittlich genutztes System. Die Anforderungen können je nach Nutzung variieren.



Prozessorleistung für Basissystem und Nutzer

Je mehr Personen gleichzeitig mit dem Dokumentenmanagement-System arbeiten, desto mehr Kerne sollten *agorum core* zugeteilt werden. Für das Basissystem rechnen Sie vier Kerne ein - pro 1000 User weitere vier, was acht Kernen im Minimum entspricht. Mit 2,4 GHz getaktete Prozessoren sind dabei die Untergrenze für ein gut laufendes System, eine optimale Performance erhalten Sie mit Prozessoren, die mit über 3,4 GHz getaktet sind. Vor allem bei der Massenverarbeitung von Dokumenten, bei der viele Aktionen nacheinander ablaufen, wie zum Beispiel beim Import von Dokumenten, macht sich die Investition in ein performantes System sehr schnell bezahlt.

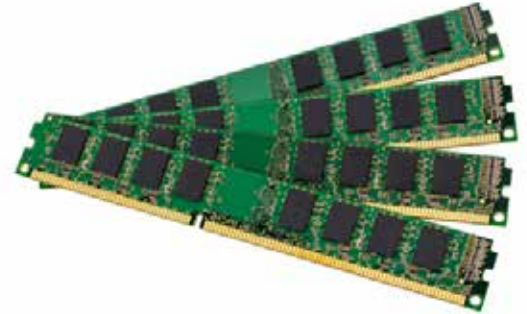
Berechnung des Bedarfs für die automatische Texterfassung

Für die Erfassung von gescannten und undurchsuchbaren Dokumenten per OCR sollten weitere Kerne einkalkuliert werden. Wieviele und mit welcher Taktung ist davon abhängig, wie viele Dokumente Sie erfassen müssen und in welcher Zeit das stattfinden soll. Möchten Sie beispielsweise die Eingangspost innerhalb kürzester Zeit erfassen und verteilen, benötigen Sie mehr und höhergetaktete Prozessorkerne als wenn Sie Dokumente zeitunkritisch zum Beispiel über Nacht verarbeiten lassen. Zur Berechnung der Kapazität können Sie einen Durchschnittswert von ca. 10 Sekunden pro Seite annehmen, die tatsächliche Geschwindigkeit hängt von der Taktung der Prozessoren ab. Als Basis sollten Sie mindestens einen Kern für den OCR-Prozess und einen Kern für die Verarbeitung durch *agorum core docform* vorsehen. Je mehr Kerne Sie dem OCR-Prozess zur Verfügung stellen, um so mehr Seiten können parallel verarbeitet werden – je höher die Taktung ist, um so schneller werden die einzelnen Seiten verarbeitet.

Faustformel für die Berechnung der Kerne für das Gesamtsystem:
8 Kerne als Basis + (Anzahl User/250) Kerne + Kerne für die OCR. D.h. bei
250 Benutzern + 1 OCR Kern wären es 10 Kerne.

Arbeitsspeicher

Auch der benötigte Arbeitsspeicher ist natürlich mit davon abhängig, wie viele Personen mit *agorum core* arbeiten. Aber auch die Anzahl der Objekte im System beeinflussen den Speicherbedarf: Je mehr Speicher Sie zur Verfügung stellen, um so mehr kann der Suchindex belegen, was die Performance des Systems deutlich verbessert. Die dargestellten Berechnung sind beispielhaft für ein durchschnittlich genutztes System. Die Anforderungen können je nach Nutzung variieren.



Berechnung des Speicherbedarfs

Der Speicherbedarf von *agorum core* hängt davon ab, wie viele Nutzer gleichzeitig mit dem DMS arbeiten und wie viele Dokumente im System gespeichert sind. Als Basis sollten für jeden *agorum core* Server mindestens 18 GB RAM zu Verfügung stehen. 8 GB für die Javamaschine + 4 GB für die Suchmaschine + 1 GB für den Zookeeper + 2 GB für die Datenbank (MySQL), der Rest ist für das Betriebssystem und dient als Cache für Solr und das Filesystem. Für jede weitere Person, die gleichzeitig mit *agorum core* arbeitet, können Sie mit ca. 10 - 20MB kalkulieren (je nach Verwendung)

Einstellung für sehr große Datenmengen (viele Millionen Objekte)

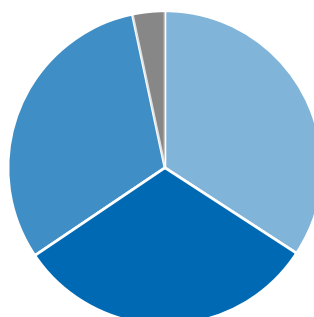
Hierbei muss die Speichereinstellung der Suchmaschine erhöht werden auf ca. 16 - 20GB. Zudem sollte die Datenbank (MySQL) auf ca. 16 - 20GB erhöht werden (InnoDBRamBufferSize).





Die Suchmaschine Solr benutzt frei verfügbaren Speicher als Cache. Je mehr RAM frei zur Verfügung steht, umso schneller arbeitet die Suchmaschine und muss nicht laufend vom Plattensystem nachladen. Durch den Einsatz von SSDs kann die Nachladezeit stark minimiert und die Performance erhöht werden.

Ansonsten gilt stets: je mehr RAM verfügbar ist, desto besser.

Bei einem Linux-System sollte zudem der Linux-Kernel-Parameter `vm.swappiness` auf 0 stehen (deaktiviert), um zu vermeiden, dass das System unnötigerweise RAM-Bestandteile auf Platte auslagert. Für die Einstellung dieses Parameters gibt es Anleitungen im Internet.

Speicherbedarf bei typischer Nutzung (bis 1.000 Nutzer mit ca. 1TB Datenvolumen)



-  Betriebssystem 2 GB
-  Index 20 GB
-  Datenbank 20 GB
-  Basissystem 24 GB

Datenbank

Für den Betrieb von *agorum core* wird eine SQL-Datenbank benötigt. Unterstützt werden aktuell MySQL, PostgreSQL, MS SQL und Oracle Database. Um Ihnen die Auswahl zu erleichtern, haben wir Ihnen im Folgenden unsere Erfahrungswerte mit den unterschiedlichen Datenbanken kurz zusammen gestellt. Die Datenbank kann auf dem gleichen Server wie *agorum core* installiert sein oder auf einem eigenen Rechner. Bei einem externen Rechner ist auf eine möglichst geringe Netzwerklatenz zu achten!



Unsere Erfahrungen mit den unterstützten Datenbanken

Alle von uns unterstützten Datenbanken sind in unterschiedlichen Kundenprojekten im Einsatz. Grundsätzlich wird die Datenbank immer von unseren Kunden bereitgestellt und gewartet. Wenn Sie einen Spezialisten für eine der gelisteten Datenbanken zur Verfügung haben, sollte es im Betrieb mit keiner Datenbank Probleme geben. Trotzdem haben wir mit den verschiedenen Datenbanken unterschiedliche Erfahrungen gemacht, die wir gerne mit Ihnen teilen möchten.

MySQL

Die MySQL-Datenbank ist die von uns empfohlene Datenbank für *agorum core*, wir haben sie selbst bei uns im Einsatz und bisher damit die besten Erfahrungen gemacht. Sie benötigt wenig Wartung und kommt mit sehr großen Datenmengen sehr gut zurecht.

PostgreSQL

Die PostgreSQL ist aus unserer Erfahrung etwas langsamer als MySQL, Sie ist aber ebenso wartungsfreundlich, das Tuning ist hier fast ausschließlich über schnellere Platten möglich.

MS SQL

Die MS SQL-Datenbanken benötigen relativ viel Pflege und Wartung, die Indexe der DB müssen häufig defragmentiert werden und für die Datensicherung muss ein schneller Speicher für das Transactionlog bereitgestellt werden. In der Standard-Edition kann während des Betriebs nicht defragmentiert werden, so dass man dafür ein Wartungsfenster braucht. Zudem gibt es Einschränkungen beim Arbeitsspeicher. In der Enterprise-Edition ist die Defragmentierung im laufenden Betrieb möglich und es kann mehr Arbeitsspeicher zugeteilt werden. Zudem ist diese Datenbank sehr Speicher und CPU intensiv.

ORACLE

Auch die Oracle-Datenbank benötigt aus unserer Sicht laufende Wartung und häufig eine Optimierung der Indexe. Damit ist sie nicht so wartungsfreundlich wie beispielsweise die MySQL-Datenbank.

Serverbetriebssystem

agorum core basiert auf Java und läuft sowohl auf Linux- als auch auf Windows-Servern. Je nachdem, welchen Server Sie verwenden, sind einige Dinge zu beachten. Welche das sind, möchten wir Ihnen kurz erläutern.



Windows- und Linuxserver

Welches Betriebssystem Sie verwenden möchten, hängt davon ab, womit Sie sich auskennen und welche Ansprüche Sie an den Betrieb stellen. Sie können beide Betriebssysteme auch im Mischbetrieb miteinander kombinieren. In ausführlichen Tests konnten wir keine Unterschiede in der Performance beim Betrieb von *agorum core* feststellen.

Linux

Wenn Sie *agorum core pro* unter Linux installieren, müssen Sie lediglich beachten, dass für die *agorum core ocr-engine* ein zusätzlicher Windows Server (ab 2008, auch virtualisiert) benötigt wird.

Tipp: für einen durchgängigen, stabilen Betrieb empfehlen wir für große Systeme den Betrieb auf einem Linux-System.

Windows Server ab 2008

Der Windows Server hat den Vorteil, dass zusätzliche Software verfügbar ist, die sich als sehr nützlich erweisen kann. Das wäre zum Beispiel:

- **Serverseitiges MS Office**
Mit einem serverseitig installierten MS-Office konvertieren Sie zum Beispiel alle Officeformate in PDF, ersetzen Feldern innerhalb von Office-Vorlagen (z.B. können Sie in eine Vorlage automatisch die Adresse des Empfängers einsetzen lassen).
- **OCR-Engine**
Aktuell ist für die als Zusatzmodul verfügbare, in *agorum core* integrierte OCR-Engine ein Windows-Betriebssystem notwendig. Sie können mit einigen Einschränkungen auch eine eigene OCR-Software betreiben.

Für den reibungslosen Betrieb von *agorum core* unter Windows Server wird ein regelmäßiger Reboot des Betriebssystems empfohlen (z.B. einmal wöchentlich).

DMS-Laufwerk

Es sieht aus wie ein herkömmliches Netzlaufwerk, funktioniert exakt wie ein Netzlaufwerk, ist aber keins. Das *agorum core* DMS-Laufwerk mounten Sie genau wie einen Fileserver, aber es steckt viel mehr dahinter, als Sie auf den ersten Blick sehen können. Denn das DMS-Laufwerk ist eine von vielen Möglichkeiten, auf die DMS-Datenbank von *agorum core* zuzugreifen.



Das *agorum core* DMS-Laufwerk

Mit dem *agorum core* DMS-Laufwerk arbeiten Sie mit Ihrem Dokumentenmanagement-System wie mit einem Filesystem. Damit ist das DMS-Laufwerk die am einfachsten einsetzbare Schnittstelle, um Daten in *agorum core* zu speichern und zu öffnen. Alle Benutzer und jede Software, die mit einem Netzlaufwerk arbeiten können, sind auch in der Lage, genau so mit dem DMS-Laufwerk zu arbeiten. Das DMS-Laufwerk ist aber ein rein virtuelles Laufwerk, die Daten liegen nicht in einem Filesystem, sondern in *agorum core*.

Im direkten Performance-Vergleich mit einem herkömmlichen Netzlaufwerk scheint die Geschwindigkeit des DMS-Laufwerks langsamer zu sein. Scheinbar deshalb, weil hier Äpfel mit Birnen verglichen werden. Beim Schreiben von Daten über das DMS-Laufwerk geschieht im Hintergrund wesentlich mehr als bei einem reinen Filesystem. Das sind Prozesse, von denen Sie stark profitieren und die es mehr als wert sind, ein paar Millisekunden länger zu warten. Welche Prozesse das sind, finden Sie in der folgenden Auflistung:

Ablaufende Prozesse:

- Das integrierte Mitteilungssystem informiert die Benutzer über Änderungen in den von ihnen überwachten Bereichen (zur Beschleunigung kann dieser Service abgeschaltet werden).
- Metadaten werden erzeugt, vererbt und beim Ablegen indiziert, um sofort auffindbar zu sein. Damit können automatisierte und regelbasierte Prozesse nahezu in Echtzeit ausgeführt werden.
- Metadaten werden automatisch über Ordnerstrukturen vererbt, um zu vermeiden, dass Anwender diese von Hand eingeben müssen.
- Für jedes abgelegte Dokument untersuchen Tasks, ob weitere Metadaten ausgelesen oder vordefinierte Regeln ausgeführt werden müssen, sogenannte Contenttasks. Damit können z.B. Metadaten aus Word oder PDF-Dateien ausgelesen werden.
(Contenttasks lassen sich abschalten, wenn Sie auf diese automatisierten Prozesse verzichten möchten).
- In einer Audittabelle werden alle Änderungen eines Dokumentes protokolliert.

- Die Zugriffe auf Dokumente werden protokolliert und können ausgewertet werden.
(Dieser Service kann abgeschaltet werden).
- Es werden Prüfungen ausgeführt, um die Konsistenz des DMS zu gewährleisten.
- Abgelegt wird in Transaktionen. Das bedeutet, sollte beim Schreiben etwas schief gehen, bleiben keinerlei Fragmente zurück, sondern die komplette Transaktion wird rückgängig gemacht.
- Der Inhalt des Dokumentes wird im Volltext indiziert.
(Dieser Service kann abgeschaltet werden, dann ist aber das Dokument nicht mehr durchsuchbar.)
- Previews werden erstellt, die es ermöglichen, die Vorschaubilder eines Dokuments sofort anzeigen zu lassen.
(Dieser Service kann abgeschaltet werden, dann müssen Anwender bei ersten Aufruf eines Dokuments kurz warten, bis die Vorschau im Hintergrund erstellt wurde.)
- Wenn ein Dokument überschrieben/neu gespeichert wird, wird automatisch eine Historie des Dokumentes angelegt.
- Das Ablegen des Dokumentes in einer Datenbank ist langsamer als das Speichern auf einer Platte, weil das Speichern in der Datenbank transaktionsgesteuert ist und daher im Hintergrund deutlich mehr passiert.
Das kann aber durch das *agorum core storage*-Modul beschleunigt werden.
- Zusätzlich zum *storage*-Modul kann noch das *dedup*-Modul zur Deduplizierung installiert werden. Dieses sorgt dafür, dass gleicher Content auf den selben Speicher verweist. Das spart Speicherplatz und ist vor allem bei der Erzeugung von Versionen wesentlich schneller, weil immer nur der Teil des Dokuments neu geschrieben werden muss, der sich verändert hat.

Einige Vorteile des DMS-Laufwerks:

- Volltext- und gezielte Metadatensuche - Millionen von Dokumenten innerhalb weniger Sekunden durchsuchen und schneller das richtige Dokument finden
- Dokumente miteinander verknüpfen - z.B. Lieferschein und Rechnung
- Dokumente revisionssicher ablegen
- Dokumente können einfach in weitere Prozesse mit ein gebracht werden, indem sie in einen aktiven Ordner gespeichert werden. Sobald ein Dokument dort hineingespeichert wird, können im Hintergrund Aktionen ausgelöst werden.
- Der Lebenszyklus eines Dokumentes wird vollständig protokolliert.
- Prozesse für eine externe Freigabe können einfach implementiert werden, Dokumente können beispielsweise direkt aus dem DMS geteilt und zum Download zur Verfügung gestellt werden.
- Automatisiertes Ablegen - Dokumente werden anhand ihrer Inhalte und/oder Metadaten erkannt und automatisch abgelegt.
- Vergabe beliebiger Metadaten an Dokumente und Verzeichnisse

Datensicherung

Daten zu verlieren, ist immer ärgerlich. Egal, ob es sich dabei nur um einzelne Dokumente handelt, oder um einen Datenverlust im größeren Stil. Im schlimmsten Fall kann der Wegfall von Daten sogar die Existenz eines Unternehmens bedrohen. Wie Sie Ihr DMS sichern, kommt dabei darauf an, wie Ihr System genutzt wird. Lläuft es zum Beispiel 24/7 durch oder gibt es in der Nacht Pausen, in denen das System heruntergefahren werden kann?



Daten sichern und wiederherstellen

Einzelne versehentlich gelöschte oder überschriebene Dokumente stellen Sie in *agorum core* über die Dokumentenhistorie oder den Serverpapierkorb schnell und einfach wieder her. (Es sei denn, diese sind revisionssicher markiert, dann sind sie ohnehin nicht löscht- oder änderbar). Für den Disasterfall müssen Sie etwas mehr vorsorgen und ein regelmäßiges Backup fahren. Dafür gibt es unterschiedliche Möglichkeiten, die wir hier kurz anreißen.

Wie eingangs erwähnt, stellt *agorum core* Ihnen für das kleine Malheur im Alltag mit der Dokumentenhistorie und dem Papierkorb zwei Funktionen zur Verfügung, mit denen Sie Ihre Daten schnell zurückbekommen. Die Historie stellt Ihnen vorherige Versionen eines Dokuments unkompliziert zur Verfügung, so dass Sie schnell darauf zurückgreifen können, wenn Sie zum Beispiel eine Datei versehentlich überschrieben haben.

Gelöschte Dokumente liegen erst einmal für einen definierten Zeitraum im Serverpapierkorb von *agorum core*, bevor sie komplett gelöscht werden. Bis dahin können irrtümlich gelöschte Dokumente oder auch ganze Verzeichnisstrukturen mit einem Klick vom Benutzer wiederhergestellt werden.

Problematisch wird es erst, wenn durch einen Hardware- oder Datenbankdefekt oder Schlimmeres der Datenbestand in größerem Umfang geschädigt wird. Dann brauchen Sie eine effiziente Datensicherung, die es Ihnen ermöglicht, Ihre Daten sicher und schnell wieder zur Verfügung zu stellen.

Sicherung in virtuellen Umgebungen - Von uns empfohlene Variante

In der Variante, die wir selbst einsetzen, sind Sie in der Lage, im laufenden Betrieb regelmäßig Snapshots zu machen, auf die sie innerhalb von Minuten zurückgreifen können. Mit virtualisierten Systemen und einer Hochverfügbarkeits- und Backup-Lösung werden die Snapshots stündlich erstellt, inkrementell gesichert und danach wieder gelöscht. Der laufende Betrieb wird dabei praktisch nicht gestört und das System ist 24/7 verfügbar, die Tabellen in der Datenbank werden für maximal 1-2 Sekunden gesperrt, gesichert und wieder freigegeben. Selbst wenn Nutzer genau zu diesem Zeitpunkt ein Dokument öffnen oder speichern sollten, wird ihnen diese minimale Verzögerung nicht auffallen. Für eine ressourcenschonende Ablage dedupliziert und komprimiert die Software die anfallenden Backupdaten.

Innerhalb kürzester Zeit sind Sie in der Lage, Ihr System auf einen beliebigen Zeitpunkt zurückzufahren oder wieder-

herzustellen. Ein weiterer Vorteil: Sie können aus einem Backup heraus innerhalb von Minuten ein Testsystem etablieren, in dem Sie neue Konfigurationen oder Updates bequem testen können, bevor Sie sie in Ihr produktives System einspielen. Bitte beachten: Machen Sie trotzdem regelmäßig einen Dump der Datenbank, um festzustellen, ob diese in sich konsistent ist.

Die Vorteile im Überblick:

- Regelmäßige, inkrementelle Sicherung im laufenden Betrieb
- Hochverfügbarkeit
- Testsysteme einfach aus einem Backup starten
- Ressourcenschonend durch Komprimierung und Deduplizierung

Virtualisierung des Backup-Systems

Ähnlich in der Vorgehensweise ist die Sicherung über die Virtualisierung eines Backup-Systems. Dazu brauchen Sie ein Filesystem, mit dem Sie Snapshots machen können, beispielsweise VSS unter Windows oder XFS unter Linux, außerdem Datenbank-Replikation und eine virtuelle Umgebung für das Backup, mit der Sie temporäre Snapshots starten können. Richtig eingerichtet, können Sie auch so recht einfach auf einen älteren Stand des Systems zurückspringen, um an gelöschte Daten zu kommen oder schnell ein aktuelles Testsystem einrichten.

Sicherung der Datenbank - nur für kleinere Systeme empfohlen

Die einfachste Möglichkeit, *agorum core* zu sichern, ist ein Datenbankdump. Dazu liefert *agorum core* ein Backup-Skript mit, mit dem der Vorgang gesteuert wird. Dazu wird der Server gestoppt und ein Datenbankdump generiert, der alle Systemdaten enthält. Nach Beendigung des Backups wird *agorum core* wieder gestartet. Je nach Größe der Datenbank und Performance der Hardware kann dieser Vorgang relativ lange dauern, daher wird dies üblicherweise über Nacht gemacht, wenn das DMS nicht genutzt wird. Zudem ist eine Wiederherstellung relativ zeitaufwändig. Vorteil: Das Backup ist betriebssystemneutral und unabhängig von der Datenbankversion. Beachten Sie bitte, dass dabei allerdings kein Backup von Storage Daten erfolgt, sofern Sie *agorum core storage* verwenden. Zudem wird der Index für *agorum core smart search* dabei nicht mitgesichert - dieser ließe sich jedoch im Notfall wieder aus den Daten erzeugen.

Hochverfügbarkeit mit *agorum core sync*

agorum core sync verteilt die Daten auf mehrere Server oder innerhalb eines Serverclusters, so dass die Benutzer beim Ausfall eines Servers ohne lange Ausfallzeiten mit den Daten von anderen Servern weiterarbeiten können. Das funktioniert auch dann, wenn einer der synchronisierten Server außer Haus, z.B. in einer Cloud, liegt. Obwohl die Daten in beiden Fällen mehrfach vorhanden sind, ersetzt die verteilte Speicherung keine vollwertige Datensicherung, die das gesamte System inklusive Workflows und MetaDB sichert.

Hinweis: Machen Sie keinen Datenbank-Dump im laufenden Betrieb, wenn es dafür nicht ein explizit dafür vorgesehenes Verfahren gibt!

Monitoring/Testing

Wichtig für den sicheren Betrieb eines Dokumentenmanagement-Systems ist ein regelmäßiges Monitoring Ihrer Hard- und Software und die Bereitstellung eines Testsystems, in dem Sie neue Konfigurationen und Updates testen können. Damit gewährleisten Sie einen reibungslosen Arbeitsalltag für Administratoren und Mitarbeiter.



Regelmäßiges Monitoring

Wie im Abschnitt Datenbank bereits erläutert, benötigen Datenbanken mehr oder weniger viel Aufmerksamkeit. Bei einigen Datenbanken ist die Pflege sehr wichtig. Schauen Sie daher im laufenden Betrieb darauf, dass die Datenbank immer in einem guten Zustand ist. Prüfen Sie regelmäßig, ob die Datensicherung korrekt läuft und komplett ist.

Ein häufig vorkommendes Ärgernis sind vollgelaufene Festplatten. Prüfen Sie auch hier regelmäßig, ob noch ausreichend Speicherplatz vorhanden ist, und sorgen Sie frühzeitig für Nachschub, wenn Sie merken, dass es eng wird. Nutzen Sie ein Plattensystem, das es Ihnen ermöglicht, den Speicherplatz jederzeit nachträglich erhöhen zu können.

Testsysteme bereitstellen

Installieren Sie sich unbedingt ein Testsystem, das Sie ständig auf dem Stand Ihres Produktivsystems halten. Damit können Sie schnell und einfach neue Konfigurationen und Updates testen, bevor Sie damit in den Livebetrieb gehen. Mit gut getesteten Änderungen oder Erweiterungen des Systems vermeiden Sie im Ernstfall Datenverlust und Ausfallzeiten, zudem ersparen sich viel Ärger mit Ihren Anwendern und Führungskräften.

Zusammenfassung/Empfehlung

In folgender Auflistung erhalten Sie unsere Empfehlungen für den bestmöglichen Betrieb von *agorum core*, basierend auf langjähriger Erfahrung mit unterschiedlichsten Konfigurationen und Installationsgrößen. Für große Systeme empfehlen wir:

- **Server virtualisiert:** Empfohlene Virtualisierungslösungen VMWare ESXi, Proxmox, OpenStack.
- **SSD-Plattensystem:** RAID Verbund, direkt im Host. Von Storage-Systemen mit drehenden Platten + SSD-Cache raten wir ab. Zudem sollten Sie darauf achten, das Plattensystem vergrößern zu können, falls Ihre Datenmenge wächst.
- **CPU:** mind. 2 CPUs mit jeweils mindestens 6 Kernen + HT, moderne Intel CPUs mit mindestens 2.4 GHz Taktung
- **RAM:** je mehr, je besser, mindestens 32GB
- **Betriebssystem:** Linux (debian, CentOS, RedHat) für *agorum core*, Datenbank und Suchmaschine, Windows Server für OCR Engine.
- **Datenbank:** MySQL mit InnoDB Engine
- **Dateisystem:** unter Linux: Ext4, virtualisiert. Unter Windows: NTFS
- **Sicherung:** Snapshots des virtuellen Filesystems, inkrementelle Sicherung der Snapshots. Snapshots, die einen Konsistenten Stand der Datenbank, des *agorum core storage*, sowie des *agorum core* Suchindexes beinhalten.
- **Virens Scanner:** auf dem Server sollte kein Virens Scanner installiert sein, da dieser massiv den Betrieb stören kann. Falls dies dennoch notwendig ist, müssen Ausnahmeregeln für alle Java Prozesse, sowie alle von *agorum core* benutzten Verzeichnisse definiert werden.



open source **dokumentenmanagement**

Vereinbaren Sie eine persönliche Online-Demo, um *agorum core* unverbindlich kennenzulernen:

www.agorum.com/demo

agorum Software GmbH

**Vogelsangstr. 22
73760 Ostfildern**

Tel.: +49 (0)711 - 358 718 0

Fax.: +49 (0)711 - 346 106 3

E-Mail: info2@agorum.com

Web: www.agorum.com