

GWDG NACHRICHTEN 07|18

Kooperation mit dem DLR
beim Hochleistungsrechnen

TSM/ISP-Updates

Erfolgreiche Umstellung
auf EN ISO 9001:2015

Academic Cloud

Digitale Langzeit-
archivierung mit koala

ZEITSCHRIFT FÜR DIE KUNDEN DER GWDG





GWDG **NACHRICHTEN**

07|18 Inhalt

.....

4 Kooperation mit dem DLR beim Hochleistungsrechnen **6 TSM/ISP-Updates: Zeitplanung und Folgen für die Nutzer**

8 Qualitätsmanagement bei der GWDG – erfolgreiche Umstellung auf EN ISO 9001:2015

10 Kurz & knapp **12 Niedersachsens Hochschulen erhalten eigene Cloud** **14 Migration von GWDG Cloud Share zu GWDG ownCloud**

15 koala – neue Plattform zur Digitalen Langzeitarchivierung für die Deutsche Nationalbibliothek **21 Personalia** **22 Kurse**

Impressum

Zeitschrift für die Kunden der GWDG

ISSN 0940-4686
41. Jahrgang
Ausgabe 7/2018

Erscheinungsweise:
monatlich

www.gwdg.de/gwdg-nr

Auflage:
550

Fotos:

© Fotolia.com: Rawpixel.com (1), pterwort (5), Maksim Kabakou (9), EtiAmos (9), Edelweiss (11), momius (13), aratehortua (15), Sashkin (20)
© MPIbpc-Medienservice (3, 21)
© DLR (4), © GWDG (2, 19, 22)
© Universität Göttingen / Christoph Mischke (5)
© <https://pxhere.com/en/photo/1246944>
//CC0 mit Bearbeitung durch die ownCloud GmbH (12)

Herausgeber:

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen
Am Faßberg 11
37077 Göttingen
Tel.: 0551 201-1510
Fax: 0551 201-2150

Redaktion:

Dr. Thomas Otto
E-Mail: thomas.otto@gwdg.de

Herstellung:

Franziska Schimek
E-Mail: franziska.schimek@gwdg.de

Druck:

Kreationszeit GmbH, Rosdorf



Prof. Dr. Ramin Yahyapour
ramin.yahyapour@gwdg.de
0551 201-1545

Liebe Kunden und Freunde der GWDG,

in dieser Ausgabe berichten wir – nach längerer Zeit einmal wieder – über unser Qualitätsmanagement. Seit 2013 ist die GWDG bereits nach ISO 9001 zertifiziert. Seitdem hatten wir bereits mehrfach erfolgreich die entsprechenden notwendigen Audits durch einen externen Prüfer absolviert. Dies führt zu einer gewissen Übung und Vertrautheit mit den Abläufen. Diesmal freut es uns dennoch besonders, dass wir hier erfolgreich waren und wieder für eine weitere Periode zertifiziert sind.

Der Prozess war diesmal aufwändiger, da bisher die Europäische Norm (EN) ISO 9001:2008 angewendet und eine Umstellung auf die neue Revision EN ISO 9001:2015 notwendig wurde. Diese beinhaltet einige größere Änderungen und erfordert diverse Anpassungen unserer Dokumentationen und Abläufe. Ein funktionierendes Qualitätsmanagementsystem ist ein wichtiger Baustein für die kontinuierliche Verbesserung von Prozessen. Nach anfänglicher Skepsis, ob sich der nicht unerhebliche Aufwand hierfür wirklich lohnt, überwiegen nun die Vorteile. Für die Mitarbeiterinnen und Mitarbeiter ist es zudem eine schöne Erfahrung, wenn ein externes Audit eine positive Bestätigung für diese Arbeit gibt. Und für unsere Kunden ist es natürlich ein wichtiges Qualitätsmerkmal und Beleg für unsere erfolgreichen Bemühungen zur ständigen Verbesserung der Servicequalität.

Ich wünsche Ihnen viel Freude beim Lesen dieser Ausgabe.

Ramin Yahyapour

GWDG – IT in der Wissenschaft



Kooperation mit dem DLR beim Hochleistungsrechnen

Text:
Gemeinsame Presseinformation der
Georg-August-Universität Göttingen,
des DLR und der GWDG vom
15.06.2018 (mit kleinen Änderungen)

Die Universität Göttingen, das Deutsche Zentrum für Luft- und Raumfahrt (DLR) und die Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) wollen beim Einsatz und Betrieb von Hochleistungsrechnern künftig enger zusammenarbeiten. Dazu unterzeichneten die Partner nun eine Absichtserklärung.

Ziel der Vereinbarung ist eine engere Kooperation in der rechen- und datenintensiven Forschung. Das Hochleistungsrechnen (High Performance Computing, HPC) gilt als Schlüsseltechnologie für wissenschaftlichen und technischen Fortschritt. Hierzu gehört der Zugang zu großen Rechnerkapazitäten, aber auch die Entwicklung von geeigneter Software. Universität und DLR wollen neben dem gemeinsamen Betrieb von Hochleistungsrechnern auch deren Anwendung und Methodenforschung vorantreiben.

Das DLR beabsichtigt, in den kommenden Jahren wesentliche Anteile seiner künftigen Investitionen in Rechnersysteme am Standort Göttingen zu bündeln. Es benötigt Hoch- und Höchstleistungsrechner unter anderem für die Entwicklung des Flugzeugs von morgen, das sowohl sicher als auch wirtschaftlich, umweltfreundlich und leise sein soll. Neben Windkanaltests und Flugversuchen wird dabei in Zeiten der Digitalisierung die hochgenaue Simulation des gesamten Fluggeräts immer wichtiger. „Mit der Universität Göttingen und der GWDG haben wir ideale Partner gefunden, um unsere Mission in Luft- und Raumfahrt durch Höchstleistungsrechnen voranzutreiben“, so die DLR-Vorstandsvorsitzende Prof. Dr. Pascale Ehrenfreund.

Die Universität Göttingen verfolgt als Teil ihrer Standortstrategie einen Ausbau der Informatik, der Datenwissenschaften und des Höchstleistungsrechnens. Hierzu gehört auch, dass Göttingen ab September 2018 Standort für den Norddeutschen Verbund für Hoch- und Höchstleistungsrechnen (HLRN) wird und gemeinsam mit dem Zuse-Institut Berlin ein großes Supercomputersystem

für die Wissenschaft und Forschung in den sieben norddeutschen Bundesländern betreibt. Außerdem baut die Universität am Standort ein neues Rechenzentrum, das 2020 seinen Betrieb aufnehmen soll. „Die Zusammenarbeit mit dem DLR ist eine ideale Ergänzung unserer Strategie im Hinblick auf die rechen- und datenintensive Forschung“, so Prof. Dr. Norbert Lossau, Vizepräsident

Cooperation between Göttingen University, DLR and GWDG in the Field of HPC

The University of Göttingen, the German Aerospace Center (DLR) and the Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) plan to cooperate more closely in the use and operation of high-performance computers in the future. The partners have now signed a Memorandum of Understanding. Purpose of the agreement is closer cooperation in computing- and data-intensive research. High Performance Computing (HPC) is regarded as a key technology for scientific and technical progress. This includes access to large computing resources, but also the development of suitable software. In addition to the joint operation of high-performance computers, Göttingen University and DLR also want to promote use of and methodological research in the field of high-performance computers.



Unterzeichneten die Absichtserklärung: GWDG-Geschäftsführer Prof. Dr. Ramin Yahyapour, DLR-Standortleiter Prof. Dr. Achim Block, Universitätsvizepräsident Prof. Dr. Norbert Lossau und DLR-Institutsleiter Prof. Dr. Andreas Dillmann (von links)



Wollen beim Einsatz und Betrieb von Hochleistungsrechnern künftig enger zusammenarbeiten: die Universität Göttingen, das Deutsche Zentrum für Luft- und Raumfahrt (DLR) und die Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)

für Infrastrukturen der Universität Göttingen.

Die GWDG betreibt eine umfangreiche IT-Infrastruktur mit HPC-Rechenclustern und zugehörigen Speichersystemen. Als gemeinsames Rechen- und IT-Kompetenzzentrum der Universität Göttingen und der Max-Planck-Gesellschaft übernimmt sie

wichtige Aufgaben in der Digitalisierungsstrategie. „Insbesondere bei der Optimierung von HPC-Anwendungen finden sich wertvolle Anknüpfungspunkte mit gemeinsamen Interessen zum DLR“, so GWDG-Geschäftsführer Prof. Dr. Ramin Yahyapour. ■



Software und Lizenzverwaltung

Der einfache Weg zur Software!

Ihre Anforderung

Sie benötigen eine Software, für die es keine von Ihnen nutzbare Rahmenvereinbarung gibt. Die Anzahl der erforderlichen Lizenzen ist nicht genau festgelegt.

Unser Angebot

Wir verfügen über eine Reihe von Rahmen- und Campusvereinbarungen mit namhaften Softwareherstellern und -lieferanten, über die Software auch in geringerer Stückzahl bezogen werden kann. Wir wickeln für Sie die Beschaffung der erforderlichen Lizenzen ab. Wir können uns bei Vertragsverhandlungen und Bedarfsanalysen engagieren. Zugriffslizenzen können auch über Lizenzserver verwaltet werden.

Ihre Vorteile

> Sie können die benötigte Software in vielen Fällen sofort nutzen.

- > Sie brauchen kein eigenes Ausschreibungs- und Beschaffungsverfahren durchzuführen.
- > Sie ersparen sich die zeitraubenden Verhandlungen mit den Softwareherstellern und -lieferanten.
- > Die Anzahl der benötigten Lizenzen wird Ihnen flexibel zur Verfügung gestellt.
- > Wir können die Nachfrage von verschiedenen Nutzern für neue Lizenzvereinbarungen bündeln.

Interessiert?

Informationen zu bestehenden Lizenzvereinbarungen sind auf der u. g. GWDG-Webseite zu finden. Falls Sie nach spezieller Software suchen, die noch nicht auf unserer Webseite erwähnt ist, kommen Sie bitte auf uns zu. Wir werden prüfen, ob wir eine Vereinbarung abschließen können und bündeln die Nachfrage mit anderen Nutzern.

>> www.gwdg.de/software

TSM/ISP-Updates: Zeitplanung und Folgen für die Nutzer

Text und Kontakt:

Björn Nachtwey
bjoern.nachtwey@gwdg.de
0551 201-2181

IBM veröffentlicht seit Version 7 im Laufe eines Jahres vier neue Minor-Releases, die jeweils neue Funktionen bringen – leider manchmal aber auch alte Fehler zu Tage bringen. Mit der Version 8 von ISP sind Sicherheitsfeatures wie SSL/TLS obligatorisch geworden, die auf Nutzerseite Anpassungen an der Konfiguration erforderlich machen. In diesem Artikel wollen wir daher die Planungen für die nächsten Updates und die Auswirkungen und notwendigen Anpassungen bei den Clients beschreiben.

UPDATESTATUS DER TSM/ISP-SERVER

Mit der Veröffentlichung des ISP-Servers in Version 7.1.4 im Dezember 2015 trat ein Fehler in der Clientsoftware zu Tage, der nur noch ein Backup, aber kein Restore von Daten ermöglichte. Es wurde zwar ein Workaround gefunden, aber seitens der GWDG haben wir uns dennoch entschieden, die Updates der Server auf neuere Versionen vorübergehend auszusetzen. Mittlerweile sind sowohl die Version 7.1.9 (März 2018) als auch die Version 8.1.5 (Juni 2018) verfügbar. Damit ist es absehbar, dass die älteren von der GWDG genutzten Versionen demnächst aus dem Support genommen werden. Außerdem lautet bei zahlreichen Problemen die erste Empfehlung sowieso, zunächst auf die jeweils neueste Version zu gehen.

Zurzeit sind die TSM/ISP-Server größtenteils noch auf der Version 7.1.3. Lediglich die Library Manager (*LM101*, *LM201*) und die Instanzen für externe Nutzer (*SM130*) und für die MPG-Generalverwaltung (*SM230*) laufen bereits auf Version 7.1.6 – sie sind damit aber auch nicht ganz auf dem neuesten Stand.

Der Server *SM130*, der von außerhalb der zentralen GÖNET-Firewall erreichbar ist, stellt zudem bereits SSL/TLS für eine verschlüsselte Verbindung auf Port 3130 zur Verfügung.

ZEITPLANUNG FÜR DIE NÄCHSTEN UPDATES

Da die nächsten Updates für TSM/ISP einige wesentliche Auswirkungen auf die Clients haben, wird nachfolgend eine grobe Zeitplanung gegeben, so dass sich die Nutzer frühzeitig auf die notwendigen Umstellungen vorbereiten können.

Zu Anfang des kommenden Jahres

Für das kommende Jahr planen wir das Upgrade auf TSM/ISP in der Hauptversion 8. Der wichtigste Unterschied besteht darin, dass die SSL/TLS-Verbindung nicht mehr optional, sondern obligatorisch ist – und entspricht damit auch den Änderungen, die ein

Update auf ISP 7.1.8 oder 7.1.9 bringen würde.

Alle Clients **müssen** sich per SSL/TLS mit einem TSM/ISP-Server in den Versionen 7.1.8 oder 8.1.2 (oder neuer) verbinden. Es muss also für jeden Client das TSM/ISP-Server-Zertifikat importiert und die Konfiguration um SSL/TLS ergänzt werden. **Unverschlüsselte Verbindungen sind dann nicht mehr möglich!**

Serverseitig funktioniert das Upgrade auf Version 8 (fast) wie ein normales Update; zahlreiche Funktionen aus den ersten TSM/ISP8-Versionen wurden auch die Version 7 als Backport integriert. Die Umstellung ist daher deutlich weniger aufwändig als die Umstellung von TSM5 auf TSM7, die wir seinerzeit bei der GWDG auch zur vollständigen Neuorganisation der TSM-Umgebung genutzt haben. Es ist lediglich mit etwas zusätzlicher Zeit für die Reorganisation der internen TSM/ISP-Datenbank zu rechnen. Insgesamt erwarten wir aber, dass die Updates/Upgrades ähnlich viel Zeit benötigen wie zuvor.

TSM/ISP Updates

The GWDG is planning to update and upgrade their TSM servers. Due to a bug in the client we suspended the updates for two years, but the actually used versions of the TSM servers will run out-of-support soon, so we have to do the updates. Starting with TSM/ISP 8.1.2 IBM changed the security policy for the client connection from optional to mandatory use of SSL/TLS. While the "restore bug" can be solved by installing an actual TSM/ISP client software, the use of SSL/TLS requires some changes in the clients option file. Therefore we do the update in two steps: first we bring all servers to ISP 7.1.7, which is the last version without mandatory SSL/TLS use and with this we configure them to allow the use of secured connections. This will be done until the end of this year. Then we will do the upgrade to ISP 8. All changes will be communicated via the TSM mailing list.

SSL-OPTION (DSM.SYS BZW. DSM.OPT)	NOTWENDIGER WERT	VOREINGESTELLTER WERT	BEDEUTUNG
SSL	Yes	No	Schaltet Verschlüsselung für die Client-Server-Kommunikation ein
SSLDISABLELEGACYtls	Yes	No	Erfordernis von TLS 1.2 oder höher
SSLFIPSMODE	Yes	No	Nutzung des „Federal Information Processing Standards (FIPS)“ für die SSL-Kommunikation
SSLREQUIRED	Yes	Default	Schaltet Verschlüsselung für die gesamte Client-Server-Kommunikation ein (Option ist unnötig; muss nicht gesetzt werden)

Tabelle 1: Optionen für die Nutzung von SSL/TLS und notwendige Werte

Eine besondere Herausforderung ist aber auch für die Server die Umstellung auf SSL/TLS, da diese auch für die Kommunikation zwischen den TSM/ISP-Servern gilt. Die nachfolgend beschriebene Bereitstellung von SSL/TLS im Vorfeld erlaubt (hoffentlich) eine schrittweise Umstellung. Sollte dies aber nicht der Fall sein, müssen alle Server (quasi) gleichzeitig auf SSL/TLS umgestellt werden, was sicherlich ein mehrtägiges Wartungsfenster erfordern würde. Wir werden darüber dann aber selbstverständlich frühzeitig informieren.

Zwischenschritt bis zum Jahresende

Zur Vorbereitung des „SSL/TLS-Zwangs“ werden wir in der Zwischenzeit die TSM/ISP-Server dahingehend konfigurieren, dass sie parallel zur unverschlüsselten Verbindung auf den Standard-Ports (Portnummer = 2000 + Server-Instanz-Nummer; z. B. SM131 => 2000 + 131 => *TCP*Port 2131) auch eine SSL/TLS-Verbindung auf Port 3xxx anbieten (SSL/TLS-Port = 3000 + Server-Instanz-Nummer; z. B. SM131 => *TCP*Port 3131). Den Nutzern bleibt somit auch die Möglichkeit, SSL/TLS zunächst auszuprobieren und gegebenenfalls die unverschlüsselte Verbindung zu nutzen.

Wahrscheinlich zum Ende des dritten Quartals werden wir für neue Knoten die Nutzung von SSL/TLS serverseitig erforderlich machen.

Parallel wollen wir aber auch die ausgesetzten Updates seit Version 7.1.3 nachholen und die Server zumindest auf den letzten SSL/TLS-„freien“ Stand in Version 7.1.7-300 bringen. Hierfür ist es aber wünschenswert, die Clients vorher auf Version 7 oder 8 zu bringen, um diesen ein reibungsloses Restore zu ermöglichen. Wir werden die Nutzer älterer Client-Versionen daher gezielt ansprechen.

Außerdem soll dieser Zwischenschritt für das Upgrade des Betriebssystems auf SLES 12 genutzt werden.

AUSWIRKUNGEN UND NOTWENDIGE ANPASSUNGEN FÜR DIE CLIENTS

Das Update auf die Version 7.1.7 hat keine direkten Auswirkungen, die Konfiguration der Clients kann unverändert beibehalten werden. Es muss zur Umgehung des „Restore-Bugs“ lediglich ein Client in der Version 7.1 oder 8.1 installiert sein.

Problematisch sind Clients, die noch ein 32-Bit-Linux nutzen, da IBM die 32-Bit-Linux-Clients mit Version 6.2 eingestellt hat, also alle 7er- und 8er-Clients nur noch als 64-Bit-Software vorliegen. Während die 6er-Clients (formal) noch für die 7er-Server zulässig sind (sie haben im April bzw. Ende September 2017 das „End-of-Support“ erreicht), fallen sie für TSM/ISP-8-Server aus der Support-Matrix vollständig heraus. Als Nutzer muss man überlegen, ob die entsprechenden Clients neu installiert werden können oder der Betrieb eines nicht mehr im Support befindlichen Clients das größere Risiko darstellt.

Die Umstellung auf SSL/TLS erfordert hingegen einige Anpassungen:

- Solange die Server noch nicht auf Version 8 umgestellt wurden, ist SSL/TLS ja optional und für verschlüsselte Verbindungen wird durch den Server ein separater Client-Port genutzt (3000er-Port-Nummer; s. o.). Es müssen also sowohl zusätzliche Optionen für die SSL-Nutzung (siehe Tabelle 1) als auch der sichere SSL-Port in der Konfiguration ergänzt werden.
- Mit der Umstellung auf TSM/ISP 8 ist lediglich die Portnummer wieder auf die alte (*TCP*Port 2xxx) zu ändern, da es dann keine extra SSL-Ports im Bereich 3130 – 3234 gibt.

Sobald einzelne Schritte umgesetzt werden, erfolgt die entsprechende Information über die Mailingliste der GWDG-TSM-Backup-Beauftragten. In der GWDG-Dokumentation finden sich unter dem URL https://info.gwdg.de/dokuwiki/doku.php?id=de:services:storage_services:backup:tsm:anleitungen:tls bereits Anleitungen zur SSL/TLS-Nutzung sowohl für Windows als auch für Linux. ■

Qualitätsmanagement bei der GWDG – erfolgreiche Umstellung auf EN ISO 9001:2015

Text und Kontakt:

Dr. Philipp Wieder
philipp.wieder@gwdg.de
0551 201-1576

Martina Brücher
martina.bruecher@gwdg.de
0551 201-2113

Seit Mai 2018 ist das Qualitätsmanagementsystem der GWDG nach der Europäischen Norm (EN) ISO 9001:2015 zertifiziert. Warum war die Umstellung notwendig und was bedeutet das für die Gesellschafter und für die Kunden der GWDG?

WARUM DIE UMSTELLUNG AUF DIE EN ISO 9001:2015?

Bislang war die GWDG nach EN ISO 9001:2008 zertifiziert (siehe die GWDG-Nachrichten 6-7/2013). Der Übergang von der EN ISO 9001:2008 auf die EN ISO 9001:2015 wurde aufgrund einer Normrevision notwendig, die erhebliche Änderungen im Vergleich zu der bisher gültigen Revision von 2008 mit sich bringt. So basiert die neue Revision unter anderem auf einem prozessorientierten Ansatz, der den Kreislauf der kontinuierlichen Verbesserung einbezieht. Sie setzt ferner einen starken Fokus auf die Interessen der Stakeholder sowie auf risikobasiertes Denken. Somit ist die EN ISO 9001:2015 weit weniger abstrakt als ihre Vorgängerin. Der anspruchsvollen Aufgabe der Umstellung hat sich die GWDG in den vergangenen zwei Jahren gewidmet und sie erfolgreich gemeistert.

WAS UNTERSCHIEDET DIE EN ISO 9001:2015 VON DER EN ISO 9001:2008?

Die Anpassung der EN ISO 9001 wurde vorgenommen, um auch im Rahmen dieser ISO-Norm der sogenannten High Level Structure (HLS) Rechnung zu tragen. Die HLS ist eine übergeordnete Gliederung, die den Aufbau neuer und überarbeiteter ISO-Managementsystem-Standards vereinheitlicht. Das Ziel der HLS ist es, die Kohärenz und Ausrichtung der entsprechenden Standards zu verbessern, indem sie Folgendes vorsieht:

- vereinheitlichende und gemeinsame allgemeine Struktur (Inhaltsverzeichnis)
- identische Kernaussagen
- identisches Begriffsverständnis

Sieben Handlungsfelder müssen dabei gemäß EN ISO 9001:2015 von Organisationen hinzugenommen oder neu betrachtet werden. Darunter fallen:

- die strategische Ausrichtung der Organisation
- die Anforderungen interessierter Parteien

- Prozessmanagement
- Verteilung der Verantwortlichkeiten
- Risikomanagement
- Wissensmanagement
- dokumentierte Informationen

WELCHE KONSEQUENZEN ERGEBEN SICH FÜR BESTEHENDE QUALITÄTSMANAGEMENT-SYSTEME NACH EN ISO 9001:2008?

Die Herangehensweise beim Aufbau eines Qualitätsmanagementsystems unterscheidet sich bei der Anwendung der neuen Revision erheblich von der vorherigen. Die EN ISO 9001:2015 normiert eine fundierte Auseinandersetzung mit den strategischen und operativen Zielen eines Unternehmens und deren Umsetzung. Der Anwender ist gewissermaßen gezwungen, das eigene Unternehmen Bereich für Bereich gemäß den oben genannten Kriterien „unter die Lupe zu nehmen“.

GWDG successful certified according to EN ISO 9001:2015

The quality management system of GWDG is now certified according to the European Standard EN ISO 9001:2015. Until May 2018, GWDG was certified according to the previous standard, EN ISO 9001:2008, and the transition has been planned over the last two years as the 2008 revision becomes invalid this year. The new revision differs considerably from the previous one. Among other things, it is based on a process-oriented approach that incorporates the cycle of continuous improvement and places a strong focus on risk-based thinking. Furthermore, it is less abstract than its predecessor. GWDG has been committed to the demanding task of improving its quality management system and has successfully mastered it.



NOCH STÄRKERE AUSRICHTUNG DER PROZESSE AN DEN ANFORDERUNGEN DER GESELLSCHAFTER, KUNDEN UND NUTZER

Die Normrevision legt besonderen Wert auf die Ausrichtung der Prozesse an den sogenannten interessierten Parteien und deren Anforderungen. Bezogen auf die GWDG sind interessierte Parteien die Gesellschafter, Kunden und Nutzer, aber auch Mitarbeiter, Behörden und Nachbarn. Die EN ISO 9001:2015 verzahnt jetzt die strategische Ausrichtung einer Organisation und ihre Prozessentwicklung enger mit den Anforderungen der interessierten Parteien.

Ausgehend von diesen Normforderungen hat die GWDG die Anforderungen der verschiedenen Kunden- und Nutzergruppen mit Blick auf existierende und sich in der Entwicklung befindliche Prozesse neu evaluiert. Die Situation der Mitarbeiter, deren Können und Engagement Garanten für ein exzellentes Service Portfolio sind, wurde ebenso einbezogen wie Lieferanten, Schulungsanbieter und weitere interessierte Parteien. Nur wenn sie in allen Bereichen gut aufgestellt ist, ist es der GWDG möglich, die an sie gestellten Anforderungen zu erfüllen.

DAS QUALITÄTSMANAGEMENTSYSTEM DER GWDG

Die GWDG nahm die anstehende Neu-Zertifizierung zum Anlass, eingehend ihre Prozesse zu prüfen. Bereits seit Einführung ihres Qualitätsmanagementsystems wendet die GWDG den prozessorientierten Ansatz an. Der prozessorientierte Ansatz ermöglicht einer Organisation, ihre Prozesse und deren Wechselwirkungen systematisch zu planen. Seit der erstmaligen Zertifizierung 2013 sind ausgewählte Prozesse der GWDG in ihrem Qualitätsmanagementsystem definiert und werden fortlaufend aktualisiert. Für jeden Prozess sind ein Prozessverantwortlicher und ein Prozessbeauftragter zuständig.

Aktuell werden weitere Prozesse für die Aufnahme in das Qualitätsmanagementsystem vorbereitet. So werden beispielsweise, um der eminenten Bedeutung des Datenschutzes bzw. der Datensicherheit für die GWDG und ihre Kunden Rechnung zu tragen, auch die Prozesse, welche entsprechend der EU-Datenschutz-Grundverordnung und der darauf basierenden Novellierung des Bundesdatenschutzgesetzes eingeführt wurden, in das Qualitätsmanagementsystem der GWDG integriert und darüber gesteuert. Hierüber wird in einer der kommenden Ausgaben berichtet.

KONTINUIERLICHE VERBESSERUNG BEI DER GWDG

Die EN ISO 9001:2015 beschreibt detailliert das Durchlaufen eines Qualitätszyklus, des sogenannten PDCA-Zyklus (Plan-Do-Check-Act), auch Deming-Kreis genannt, der unmittelbar der kontinuierlichen Verbesserung dient. Der vierstufige Zyklus, der auf den Mathematiker William Edwards Deming (1900 – 1993) zurückgeht, ist ein universelles Modell zur Qualitätsverbesserung, das für das Qualitätsmanagement von elementarer Bedeutung ist. Der PDCA-Zyklus ermöglicht einer Organisation sicherzustellen, dass angemessene Ressourcen für ihre Prozesse zur Verfügung stehen, die Prozesse gesteuert werden und dass Chancen zur Verbesserung bestimmt werden und auf diese reagiert wird.



Phase 1: Planen (Plan = P)

In dieser Phase erfolgt eine Analyse des Ist-Zustandes. Den Input hierzu können die unterschiedlichsten Quellen liefern, zum Beispiel Ergebnisse von Kundenbefragungen, Rückmeldungen von Kunden und Mitarbeitern, interne und externe Audits oder Marktbeobachtungen. Darauf basierend wird der Soll-Zustand definiert, das heißt es werden Maßnahmen zur Prozessoptimierung entwickelt sowie Verantwortlichkeiten und Termine bestimmt. Die GWDG nutzt für diese Planungen einen zentral fortgeschriebenen Maßnahmenplan, der vom Qualitätsmanagementbeauftragten und dem Qualitätsmanagement-Team gepflegt wird.

Phase 2: Durchführung (Do = D)

Die Maßnahmen einschließlich der Verantwortlichkeiten und der Termine werden an die zuständigen Arbeitsgruppen kommuniziert. Nun erfolgt die Umsetzung der beschriebenen Maßnahmen. Der Qualitätsmanagementbeauftragte und das Qualitätsmanagement-Team leisten hierbei den erforderlichen Support und achten auf die Einhaltung von Terminen.

Phase 3: Überprüfen (Check = C)

Hier erfolgt eine Überprüfung, ob der anvisierte Soll-Zustand erreicht wurde. Die Ergebnisse werden bewertet.

Phase 4: Handeln/Verbessern (Act = A)

Es erfolgt eine turnusmäßige Reflexion der Prozesse, wobei

der Turnus dem jeweiligen Handlungsbedarf angepasst wird. Bei etablierten, standardisierten Prozessen sind in der Regel größere Zeitspannen ausreichend als bei neuen Prozessen oder grundlegenden Änderungen von Parametern innerhalb etablierter Prozesse.

NEU: RISIKOBASIERTER ANSATZ

Eine große Bedeutung in der EN ISO 9001:2015 hat der sogenannte risikobasierte Ansatz. Er ermöglicht einer Organisation, diejenigen Faktoren zu bestimmen, die dazu führen könnten, dass ihre Prozesse und ihr Qualitätsmanagementsystem von den geplanten Ergebnissen abweichen. Hierzu werden spezifische Prozessrisiken und globale Risiken definiert und entsprechend der möglichen Folgen ihres Eintritts klassifiziert. Das heißt, potenzielle negative Auswirkungen werden bewertet. Darauf aufbauend werden Maßnahmen für den Fall des Eintritts festgelegt. Darüber hinaus kommt präventiven Maßnahmen ein hoher Stellenwert zu. Für jedes Risiko werden Vermeidungsstrategien ermittelt, um negative Auswirkungen zu vermeiden beziehungsweise zu minimieren.

NEU: WISSENSMANAGEMENT UND DOKUMENTIERTE INFORMATIONEN

Schon vor Inkrafttreten der EN ISO 9001:2015 betrieb die GWDG internes Wissensmanagement. Insbesondere Qualitätsdokumente im engeren Sinn werden seit Langem in einem

Dokumentenmanagementsystem abgelegt und versioniert, so dass in diesen Feldern kein Handlungsbedarf bestand.

ZUSAMMENFASSUNG

Im Ergebnis gewährleistet der Übergang von der EN ISO 9001:2008 auf die EN ISO 9001:2015, dass die GWDG ein noch besseres Qualitätsmanagementsystem als in den Vorjahren entwickelt hat, um das Spektrum und die Qualität ihrer Dienstleistungen auf bestmöglichem Niveau zu halten. Dies wurde durch ein Audit der DQS GmbH – Deutsche Gesellschaft zur Zertifizierung von Managementsystemen – am 16. und 17. Mai 2018 bestätigt. ■



Kurz & knapp

Neue Termine für drei Scientific-Computing-Kurse

Für die drei Kurse „Using the GWDG Scientific Compute Cluster – An Introduction“, „Parallelrechnerprogrammierung mit MPI“ und „Programming with CUDA – An Introduction“ gibt es im Herbst neue Termine, die kurzfristig in das GWDG-Kursprogramm aufgenommen wurden: 22.10.2018, 23./24.10.2018 und 06.11.2018. Nähere Informationen hierzu sind unter <https://www.gwdg.de/kursprogramm> zu finden.

Otto

Doppelausgabe 08-09/2018 der GWDG-Nachrichten

Die nächsten GWDG-Nachrichten erscheinen als Doppelausgabe 08-09/2018 Anfang September 2018.

Otto

Zugang zu chat.gwdg.de für UMG-Benutzer vereinfacht

Um die Kommunikation zwischen den IT-Dienstleistern am Standort zu verbessern, haben wir eine Umstellung vorgenommen, die allen Personen aus der Universitätsmedizin Göttingen (UMG) den Zugang zum Messaging-Dienst „Rocket.Chat“ über <https://chat.gwdg.de> vereinfacht. Bisher war für die Nutzung der Umweg über einen selbstregistrierten Account nötig.

Ab sofort können sich Benutzer aus der UMG bei <https://chat.gwdg.de> mit ihrem normalen UMG-Account anmelden. Hierzu werden die E-Mail-Adresse <benutzer>@med.uni-goettingen.de und das zugehörige UMG-Passwort verwendet.

Braunschweig



Servervirtualisierung

DER EINFACHE WEG ZUM SERVER!

Ihre Anforderung

Sie benötigen zur Bereitstellung eines Dienstes einen Applikations- oder Datenbankserver. Ihnen fehlen Platz, Hardware, Infrastruktur oder Manpower. Gleichzeitig soll der Server möglichst hochverfügbar und performant sein.

Unser Angebot

Wir bieten Ihnen die Möglichkeit des Hostings von virtuellen Servern für Ihre Anwendungen basierend auf VMware ESX. Sie können Ihre eigenen virtuellen Maschinen verwalten, die in unserer zuverlässigen Rechnerinfrastruktur gehostet werden, die unterschiedliche Verfügbarkeitsgrade unterstützen. Unsere Installation hält die Best-Practice-Richtlinien von VMware ESX ein. Sie bleiben Administrator Ihres eigenen virtuellen Servers, ohne sich mit der physikalischen Ausführungsumgebung beschäftigen zu müssen.

Ihre Vorteile

- > Leistungsfähiges VMware-Cluster mit zugehörigem Massenspeicher

- > Hohe Ausfallsicherheit und Verfügbarkeit durch redundante Standorte und Netzwerkverbindungen sowie USV-Absicherung
- > Bereitstellung aller gängigen Betriebssysteme zur Basisinstallation
- > Umfassender administrativer Zugang zu Ihrem Server im 24/7-Selfservice
- > Möglichkeit der automatisierten Sicherung des Servers auf unsere Backupssysteme
- > Zentrales Monitoring durch die GWDG
- > Große Flexibilität durch Virtualisierungstechnologien wie Templates, Cloning und Snapshots
- > Schutz vor Angriffen aus dem Internet durch leistungsfähige Firewallsysteme sowie ein Intrusion Prevention System

Interessiert?

Jeder Nutzer mit einem gültigen Account bei der GWDG kann das VMware-Cluster nutzen. Um einen virtuellen Server zu beantragen, nutzen Sie bitte die u. g. Webadresse.



Niedersachsens Hochschulen erhalten eigene Cloud

Text:
Gemeinsame Presseinformation der
ownCloud GmbH und der GWDG
vom 10.07.2018 (mit kleinen Ände-
rungen)

Mit dem offiziellen Start der „Academic Cloud“ steht allen Studierenden und Mitarbeitern niedersächsischer Hochschulen ab sofort eine leistungsfähige und sichere Kollaborationsplattform auf der Basis von ownCloud zur Verfügung. Sie wurde speziell für den Einsatz in Bildung und Forschung entwickelt und bietet interessante Funktionalitäten und Erweiterungsmöglichkeiten.

ownCloud und der Landesarbeitskreis Niedersachsen für Informationstechnik / Hochschulrechenzentren (LANIT) kündigen den heutigen (10. Juli 2018) offiziellen Start ihrer gemeinsamen Kollaborationsplattform „Academic Cloud“ an. Der Service bietet allen Studierenden und Mitarbeitern der niedersächsischen Hochschulen die Möglichkeit, unentgeltlich Daten im Umfang von bis zu 50 GB in einer eigenen Cloud abzulegen und mit anderen Nutzern zu teilen bzw. gemeinsam zu bearbeiten. Die Academic Cloud ist für 210.000 Nutzer konzipiert. Das Projekt wird vom Niedersächsischen Ministerium für Wissenschaft und Kultur gefördert, die technische Umsetzung erfolgt durch die Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG).

Die Academic Cloud wurde speziell für den Einsatz in Bildung und Forschung entwickelt, wo deutlich höhere Anforderungen bei den Themen Sicherheit und Datenschutz bestehen. Im Gegensatz zu den meisten kommerziellen Sync-and-Share-Diensten ermöglicht ownCloud den Betrieb auf eigenen Servern, ohne auf Kollaborationsfunktionen, wie sie zum Beispiel Dropbox bietet, verzichten zu müssen. Die Academic Cloud ist bereits jetzt an mehreren Hochschulen in Niedersachsen verfügbar, Pilotteilnehmer sind unter anderem die Technische Universität Clausthal, die Hochschule Emden/Leer, die Hochschule Hannover, die Medizinische

Hochschule Hannover, die Universität Hildesheim und die Universität Vechta. In einer zweiten Phase sollen weitere Hochschulen folgen.

Die Academic Cloud basiert auf der Technologie von ownCloud, ihre Funktionalitäten lassen sich durch die einheitliche Authentifizierungsstruktur aber um weitere zusätzliche Dienste erweitern:

Lower Saxony's Universities build up their own Cloud Infrastructure

With the official launch of the “Academic Cloud”, a powerful and secure collaboration platform based on ownCloud is now available to all students and staff of universities in Lower Saxony. It was specially developed for the use in education and research and offers interesting functionalities and expansion possibilities. The Academic Cloud is designed for 210,000 users and funded by the Ministry of Science and Culture of Lower Saxony. The technical implementation is carried out by the GWDG.

„Der Datenaustausch bildet die Basis für weitergehende Kollaborationsfunktionen, die im Jahr 2018 hinzukommen – wie zum Beispiel Office-Funktionalitäten oder andere Softwareanwendungen.“, betont Projektleiter Ralph Krimmel, der bei der GWDG für die technische Umsetzung verantwortlich ist.

Die Idee hinter der Academic Cloud ist es, Studierende und Mitarbeiter von Hochschulen in Niedersachsen mit einem wachsenden digitalen Dienstleistungsportfolio zu unterstützen. Die Academic Cloud bietet kollaborative Dienstleistungen für Studium und Forschung. Über eine zentrale Plattform können Studierende und Forschende Daten mit anderen Personen, Gruppen oder Projekten auf kontrollierte Weise teilen und kollaborative Services gemeinsam nutzen.

Open Source in Bildung und Forschung – ein Erfolgsmodell

Mit der Academic Cloud setzt Niedersachsen bereits als zweites Bundesland auf ownCloud als zentrale Filesharing-Lösung für Hochschulen. Bereits 2015 ging mit „Sciebo“ in Nordrhein-Westfalen die erste Bildungs- und Forschungs-Cloud an den Start, die sich bis heute mit über 100.000 aktiven Nutzern zu einem echten Erfolgsprojekt entwickelt hat.

„Wir freuen uns, dass immer mehr Hochschulen auf offene Cloud-Lösungen setzen. Bildung und Forschung leben vom gegenseitigen Austausch und mit unserer Software hoffen wir, dieses wissenschaftliche Grundprinzip mit Hilfe von Open Source mit Leben zu füllen“, betont Holger Dyroff, Geschäftsführer und COO von ownCloud.

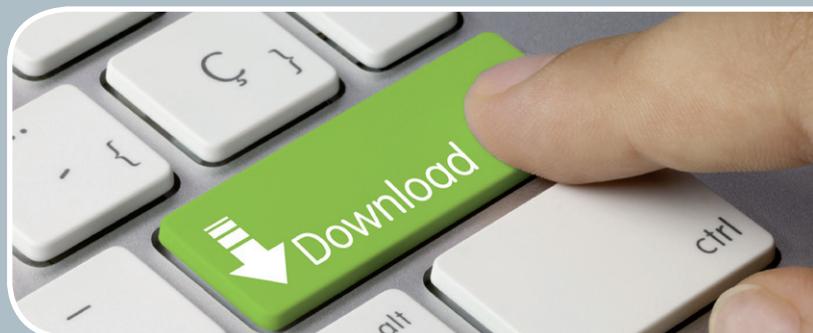
Um auf die Academic Cloud zugreifen zu können, benötigen Studierende und Mitarbeiter der Hochschulen lediglich eine gültige Benutzererkennung (Nutzername und Passwort) ihrer Hochschule. Die Anmeldung erfolgt dabei immer an der Hochschule des Teilnehmers, eine Übertragung von Passwörtern findet nicht statt, was die Sicherheit zusätzlich erhöht.

Die Plattform wird an drei Standorten des Hochschulrechenzentrums GWDG in Göttingen gehostet und ist von Anfang an für bis zu 210.000 Nutzer konzipiert. Der Gesamtspeicherplatz beträgt damit rund 1 Petabyte.

Die Academic Cloud kann entweder direkt im Browser (<https://academiccloud.de>) genutzt oder über einen installierten Client beziehungsweise eine mobile Applikation synchronisiert werden. Diese sind für alle gängigen Betriebssysteme oder Plattformen verfügbar (Android, iOS, Desktop-Apps für Windows, macOS und Linux).

GWDG ownCloud für die Universität Göttingen und die MPG

Für Nutzer der Universität Göttingen und der Max-Planck-Gesellschaft bietet die GWDG bereits seit 2014 eine vergleichbare ownCloud-Instanz an. Langfristig sollen diese Instanz und die Academic Cloud auf dasselbe technische Fundament gestellt werden. Die „GWDG ownCloud“ bietet bereits Office-Funktionalitäten zur Tabellen-, Text- und Präsentationsbearbeitung. Weitere Kollaborationstools wie ShareLaTeX, ein Chat-Dienst und Jupyter-Notebook sind für die Academic Cloud in Planung. Darüber hinaus wird an einer Archivierungsfunktion gearbeitet. ■



FTP-Server

Eine ergiebige Fundgrube!

Ihre Anforderung

Sie möchten auf das weltweite OpenSource-Softwareangebot zentral und schnell zugreifen. Sie benötigen Handbücher oder Programmbeschreibungen oder Listings aus Computerzeitschriften. Sie wollen Updates Ihrer Linux- oder FreeBSD-Installation schnell durchführen.

Unser Angebot

Die GWDG betreibt seit 1992 einen der weltweit bekanntesten FTP-Server, seit sieben Jahren mit leistungsfähigen Ressourcen für schnellen Service.

Ihre Vorteile

- > Großer Datenbestand (50 TByte), weltweit verfügbar
- > Besonders gute Anbindung im GÖNET

- > Aktuelle Software inkl. Updates der gebräuchlichsten Linux-Distributionen
- > Unter pub befindet sich eine aktuell gehaltene locatedb für schnelles Durchsuchen des Bestandes.
- > Alle gängigen Protokolle (http, https, ftp und rsync) werden unterstützt.

Interessiert?

Wenn Sie unseren FTP-Server nutzen möchten, werfen Sie bitte einen Blick auf die u. g. Webseite. Jeder Nutzer kann den FTP-Dienst nutzen. Die Nutzer im GÖNET erreichen in der Regel durch die lokale Anbindung besseren Durchsatz als externe Nutzer.

>> www.gwdg.de/ftp-server

Migration von GWDG Cloud Share zu GWDG ownCloud

Text und Kontakt:

Tobias Herbst
tobias.herbst@gwdg.de
0551 201-2135

Wie in den GWDG-Nachrichten 3/2018 bereits angekündigt, startet nun die Migration der Inhalte vom Cloud-Speicherdienst „GWDG Cloud Share“ zu „GWDG ownCloud“. Dieser Artikel gibt einen kurzen Überblick über den Migrationsprozess sowie die im Vorfeld aufgetretenen Probleme.

NEUER STARTTERMIN

Im März dieses Jahres wurde der Zeitplan für den Ausstieg aus GWDG Cloud Share bei der GWDG bekanntgegeben und eine Webseite zur Migration für den Mai angekündigt. Dabei ging der ursprüngliche Plan von einem Speicherbedarf für beide Dienste von maximal ca. 90 Terabyte aus und hätte damit das bisher gemeinsam genutzte Speichersystem praktisch allein belegt. Ein Wechsel des Speichers hätte dann Anfang 2019 erfolgen müssen. Allerdings zeigte sich diese Schätzung im April als zu konservativ und es wurde absehbar, dass der Umzug der Daten eines Speicherdienstes vor Abschluss des Migrationszeitraumes nötig sein würde. Damit war es leider nicht mehr möglich, den Mai-Termin einzuhalten: Nicht nur war der benötigte Speicherplatz zu finden, auch der Migrationsplan „internes Verschieben der Daten auf dem Speicher“ musste neu entwickelt werden.

Im Zuge der Wartung für das Update der „GWDG ownCloud“-Software auf Version 10.0.8 am 16.06.2018 wurden daher auch die Daten der GWDG ownCloud, ca. 55 Terabyte, auf einen anderen Speicher kopiert. Auf dieser Grundlage konnten nun die Webseiten und E-Mail-Vorlagen erstellt, eine Hilfe-Seite [1] verfasst und, am wichtigsten, der Teil der Migrations-Software angepasst werden, der die eigentliche Arbeit erledigt. Ab Ende Juni wurde der Vorgang intensiv getestet und optimiert.

MIGRATIONSPROZESS

Für den Migrationsprozess steht nun eine Webseite [2] in Deutsch und Englisch zur Verfügung. Der Ablauf ist, trotz der schwierigen Aufgabe, zwei nicht kompatible Systeme zu verbinden, für Nutzer sehr einfach gestaltet:

1. Anmeldung an GWDG ownCloud, um sicherzustellen, dass ein Account existiert
2. Eingabe des Namens des „GWDG Cloud Share“-Accounts in der Migrationsseite
3. Anklicken des per E-Mail zugesandten Bestätigungslinks

Im Hintergrund wird nun der Migrationsauftrag in einer Warteschlange eingetragen und die Nutzerdaten werden der Reihe nach aus GWDG Cloud Share heraus- und in GWDG ownCloud hineinkopiert. Ein besonderes Augenmerk wurde auf die Übertragung der Shares gelegt: Für jede Share in GWDG Cloud Share wird in GWDG ownCloud ein Link erzeugt und per E-Mail verschickt, mit dem Nutzer weiterhin Zugriff auf die Daten haben und der es ermöglicht, die Share in der eigenen GWDG ownCloud hinzuzufügen. Auch die Berechtigungen auf den „GWDG Cloud Share“-Shares werden dabei korrekt übernommen. Eingehende Shares, also der Zugriff auf noch in GWDG Cloud Share liegende Daten aus GWDG ownCloud heraus, können leider nicht migriert werden. Zugriff auf diese Dateien ist erst dann wieder möglich, wenn der Besitzer des Quellordners seine Daten ebenfalls zu GWDG ownCloud umzieht.

Wie bereits angekündigt wird der Dienst „GWDG Cloud Share“ im Laufe des Jahres nur noch eingeschränkt nutzbar sein. Bereits seit dem 01.07.2018 bekommen neue Nutzer keinen Speicherplatz mehr, mit dem 01.10.2018 werden alle Accounts nur noch lesend zugreifen können und mit dem Beginn des Jahres 2019 werden die „GWDG Cloud Share“-Server heruntergefahren.

LINKS

[1] https://info.gwdg.de/dokuwiki/doku.php?id=de:services:storage_services:gwdg_cloud_share:migration

[2] <https://cloudshare.gwdg.de/migration>

Migration of GWDG Cloud Share to GWDG ownCloud

As announced in the GWDG News 3/2018 the migration of GWDG Cloud Share to GWDG ownCloud has started. This article gives a short overview of the problems our team encountered and how the migration will work.

koala – neue Plattform zur Digitalen Langzeitarchivierung für die Deutsche Nationalbibliothek

Text und Kontakt:

Björn Braunschweig
bjoern.braunschweig@gwdg.de
0551 201-2133

Die GWDG ist Kooperationspartner der Deutschen Nationalbibliothek (DNB) im Bereich der digitalen Langzeitarchivierung (LZA). Hierfür wurden im Rahmen des ehemals als kopal bekannten Projektes im Jahr 2011 Hard- und Software von IBM bei der GWDG installiert. Da die Hardware mittlerweile in die Jahre gekommen ist und IBM für die verwendete Software DIAS [1] keinen Support mehr anbietet, war eine Erneuerung unausweichlich. Dieser Artikel stellt die von der GWDG entwickelte Software koala, die dafür erforderliche Infrastruktur sowie die Migration vom alten in das neue System dar.

ÜBERBLICK

Wie bereits in den GWDG-Nachrichten 9-10/2015 erwähnt, wurden Überlegungen angestellt, wie der aktuelle Soft- und Hardware-Stack ersetzt werden könnte. Hierzu gab es ein Treffen mit der DNB zur Erhebung der Anforderungen. Dabei stellte sich heraus, dass eine Inhouse-Entwicklung einer genau passenden „Ablöse-Software“ bei der GWDG aufgrund entsprechender Ressourcen möglich ist. Daraufhin wurde dieser Weg anstelle des Kaufs einer proprietären Lösung gewählt.

Hauptaspekt der Neuentwicklung ist die DIAS-Schnittstellenkompatibilität, so dass eine reibungslose Ablösung gewährleistet ist. Außerdem steht die Verwendung von bewährten Open-Source-Komponenten im Vordergrund, um eine Herstellerbindung zu vermeiden und Kosten zu senken. Die GWDG hatte zu diesem Zeitpunkt bereits einen Prototypen entwickelt. Dieser wurde als vielversprechend angesehen und im Laufe der nächsten Monate erweitert und gefestigt.



Die folgenden Abschnitte erläutern die neu entstandene Software **koala**, die dafür notwendige Infrastruktur sowie den Migrationspfad vom alten in das neue System.

INGEST-PROZESS

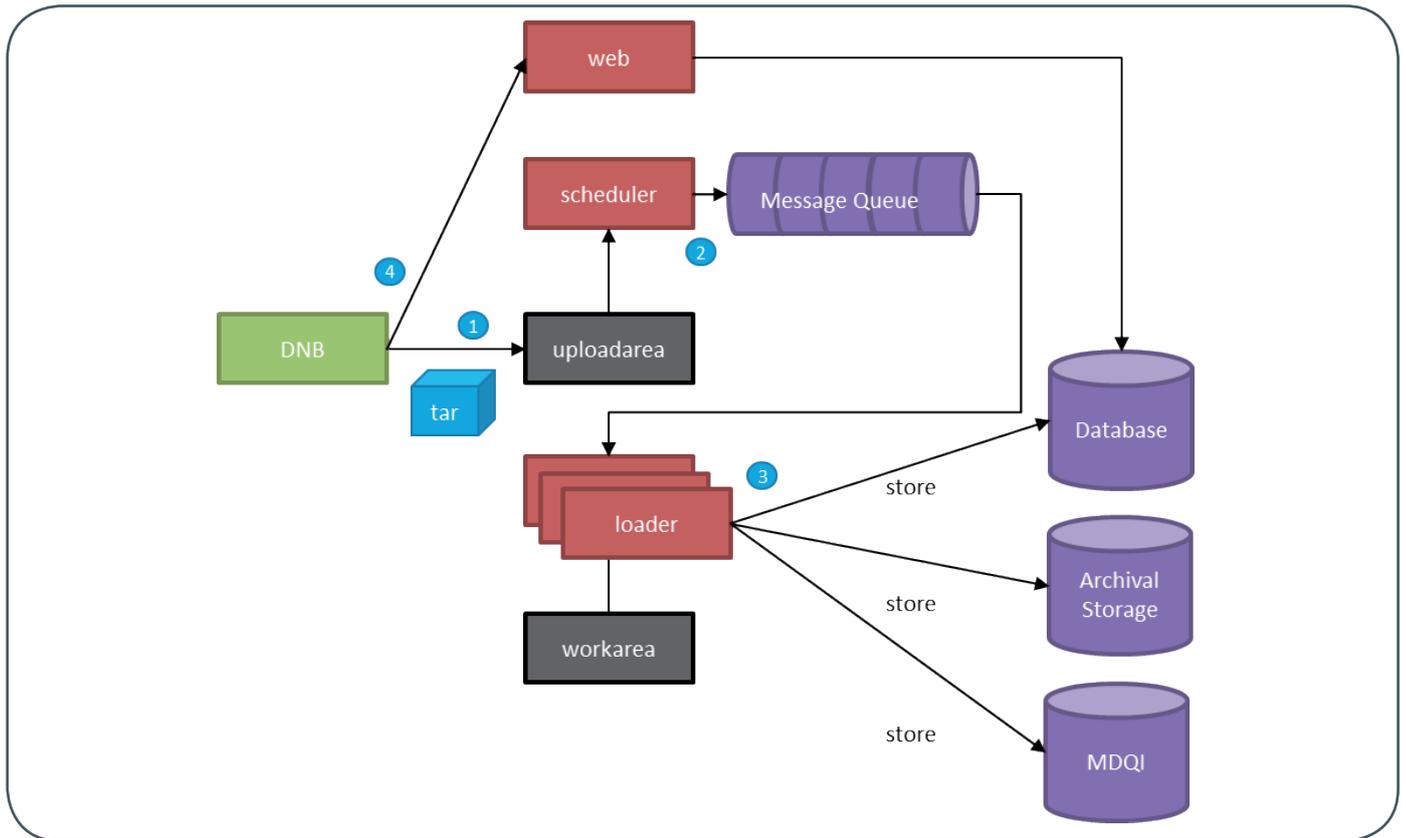
koala lehnt sich konzeptionell an die Funktionsweise der Vorgängersoftware DIAS an, ist allerdings eine komplette

Neuentwicklung und teilt daher keinerlei Quellcode. Beispielhaft sollen in Abb. 1 die wichtigsten Komponenten und Funktionsweisen anhand des Ingest-Prozesses dargestellt werden.

Von der DNB werden Archive per SFTP in in einem Hotfolder (*uploadarea* genannt) abgelegt (Schritt 1). Diese Pakete werden von der *scheduler*-Applikation markiert und notwendige Informationen in einer Message-Queue (*MQ*) abgelegt (Schritt 2). Mehrere *loader*-Applikationen sind mit der *MQ* verbunden und arbeiten die ankommenden Pakete parallel ab. Dazu werden lokale Arbeitsverzeichnisse verwendet, um die Pakete zu entpacken und die Checksummen und Metadaten zu validieren. Anschließend werden die Daten in einer Inventar-Datenbank sowie das eigentliche Archivpaket in *Archival Storage* (IBM Spectrum Protect, ehemals Tivoli Storage Manager) gespeichert. Außerdem werden die mitgelieferten Metadaten im *MDQI*, einer XML-Datenbank, hinterlegt (Schritt 3). Aus DNB-Sicht ist der Ingest-Prozess ein asynchroner Vorgang. Daher stellt koala über die Web-Applikation ein Ticket-System

Digital Long-term Archiving for the DNB with koala

The GWDG is a cooperation partner of the German National Library (DNB) in the field of digital long-term archiving. In 2011 IBM installed for this purpose hardware and software concerning the so called kopal project. Since the hardware is already getting old and IBM does not offer support for the software DIAS, a renewal was unavoidable. The following article covers the newly developed software koala, the necessary infrastructure, as well as the migration path from the old to the new system.



1_Ingest-Prozess

bereit, mit dem der Zustand eines Paketes während der gesamten Abarbeitungskette überwacht wird. Im Ticket wird vermerkt, ob der Vorgang erfolgreich war oder nicht, und der Client kann entsprechend darauf reagieren (Schritt 4), indem z. B. das Paket erneut eingespielt oder der Vorgang abgebrochen und die Administratoren benachrichtigt werden.

DASHBOARD

Im Unterschied zu DIAS wurde für koala ein umfangreiches Administrations-Dashboard entwickelt, welches einen zentralen Echtzeiteinblick auf den aktuellen Zustand von koala bietet. Es werden beispielsweise die Status der Loader und Retriever überwacht, sodass bei Problemen zeitnah eingegriffen werden kann.

Außerdem wird hier Auskunft über Anzahl, Größe und zeitlichen Verlauf der archivierten Pakete gegeben. Über die Menüs ist es möglich, diverse Einstellungen vorzunehmen oder Aktionen auszuführen, was bei DIAS nur in der Form von Config-Dateien oder API-Aufrufen möglich war. Dies erleichtert die Arbeit der Administratoren erheblich. Zum Beispiel können nun während des laufenden Betriebs die zur Metadaten-Validierung benötigten Schemata und unterstützten Dateitypen aktualisiert werden. Abb. 2 zeigt die Startseite der koala-Webanwendung.

CONTINUOUS DEPLOYMENT

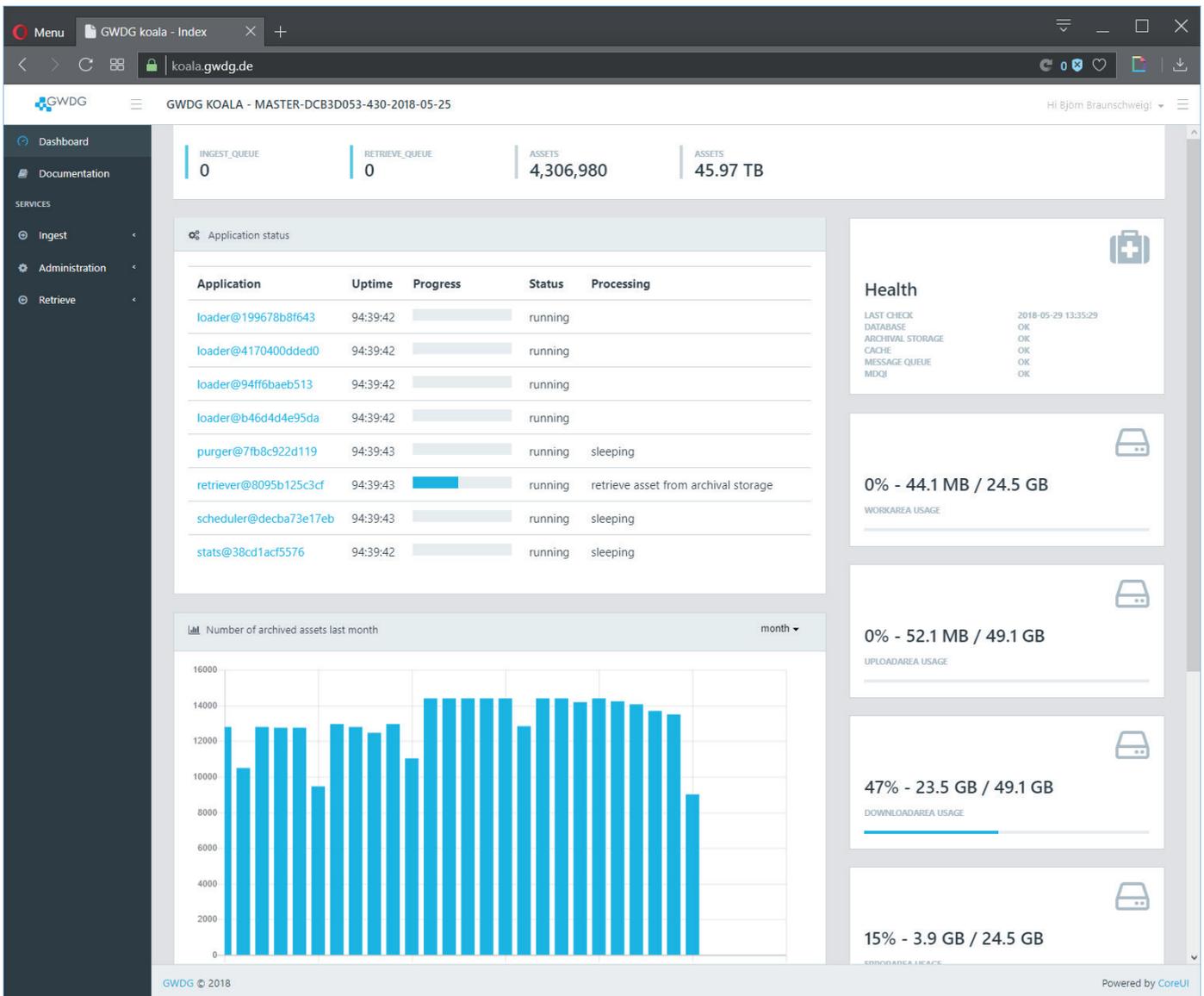
Bereits früh in der Entwicklung von koala wurde der Ingest-Prozess des Vorgängersystems DIAS erweitert, sodass eine Kopie der produktiven Daten zusätzlich in das koala-Testsystem eingespielt werden konnte. Somit war ein kontinuierlicher Test mit echten Daten möglich. Dies hatte enorm positive Rückwirkungen auf

die Entwicklung, wodurch eine Vielzahl an Fehlern frühzeitig beseitigt werden konnten. Die koala-Applikationen werden durch eine automatisierte Pipeline getestet und auf den Zielsystemen veröffentlicht. Abb. 3 zeigt die Pipeline.

Gestartet wird mit grundlegenden Unit-Tests. Anschließend werden Docker-Images auf Basis von Ubuntu erstellt und parallel mit den Tools Pylint [2] und pep8 [3] werden eine statische Codeanalyse und die Einhaltung von Coding-Standards geprüft. Nachdem die Images gebaut sind, werden lokale Docker-Container erstellt und Integrationstests anhand von API-Aufrufen und UI-Tests mit einem „headless“-Browser durchgeführt. Anschließend werden die Images in eine lokale Docker-Registry hochgeladen. Als letzter Schritt wird die Testumgebung durch diese neuen Images bzw. Container ersetzt. Ein automatisches Ausrollen auf die Produktivumgebung ist absichtlich nicht implementiert. Nach abschließender manueller Verhaltensanalyse des Testsystems kann dies allerdings per Knopfdruck nachgelagert erfolgen.

Durch das gezeigte Setup ist eine schnelle und qualitätssichernde Entwicklung möglich (siehe Abb. 4). Verwendet wird hierfür GitLab [4] zur Quellcode- und Aufgabenverwaltung, ein Jenkins [5] und eine Docker-Registry [6] zur Ablage der Images. Neben der Deployment Pipeline wird Jenkins als allgemeines Werkzeug für wiederkehrende Aufgaben eingesetzt:

- Veröffentlichung der zuletzt erfolgreich getesteten Docker-Container in das Produktivsystem
- Prüfung auf aktuellere Docker Base Images
- Prüfung auf aktuellere Pakete von denen koala abhängig ist (z. B. neuere XML-Parser-Bibliothek)
- Durchführung von täglichen TSM-Backups der Datenbanken
- Generierung der Dokumentations-Webseiten



2_Startseite der koala-Webanwendung

Verwaltet werden die Virtual Machines (VMs) durch das Systemkonfigurationswerkzeug Puppet in Verbindung mit einem Puppetserver [7]. Abb. 5 zeigt die Oberfläche einer VM im Puppetboard.

INFRASTRUKTUR

Im Unterschied zu DIAS wird koala ohne eigens angeschaffte Hardware betrieben. Genutzt wird die VMware-Umgebung der GWDG. Dies spart sowohl Strom als auch Stellplatz im Maschinenraum. Außerdem wird durch den redundant aufgebauten VMware-Cluster die Verfügbarkeit der Anwendungen erhöht.

Durch die verteilte Architektur von koala ist es möglich, die lasterzeugenden Komponenten wie Loader und Retriever auf externe Maschinen auszulagern. Bei der Bewältigung größerer Datenmengen wie der geplanten Retro-Archivierung können so weitere GWDG-Ressourcen hinzugezogen werden: GWDG Cloud Server oder der GWDG Scientific Compute Cluster. Dies hat den Vorteil, dass die eigentlichen Kernkomponenten relativ moderat dimensioniert werden können. Geplant ist, den täglichen Archivierungsbedarf durch die koala-VMs selbst zu erledigen und bei größeren Projekten die zuvor genannten Ressourcen mit einzubeziehen. Mit der

Verwendung von Docker nicht nur für die Core-Applikationen, sondern auch für die Middleware-Komponenten können eine höhere Packungsdichte pro VM und somit eine höhere Auslastung realisiert werden, was zu einer Kostenersparnis führt. Im Unterschied zur DIAS-Architektur mit neun VMs (sechs für den Produktiv- und drei für den Testbetrieb) konnte dies auf drei VMs reduziert werden. In Tabelle 1 sind die verwendeten VMs und deren Dimensionierung dargestellt.

Name	koala.gwdg.de	koala-test.gwdg.de	koala-jenkins.gwdg.de
Beschreibung	Produktion	Staging/Test	Build und Test
Hardware	4 CPUs, 6 GB RAM, 256 GB HDD, 150 GB iscsi SSD	2 CPUs, 4 GB RAM, 1 TB HDD	2 CPUs, 4 GB RAM, 96 GB HDD
Plattform	VMware Intel Cluster	VMware Bronze Cluster	VMware Bronze Cluster

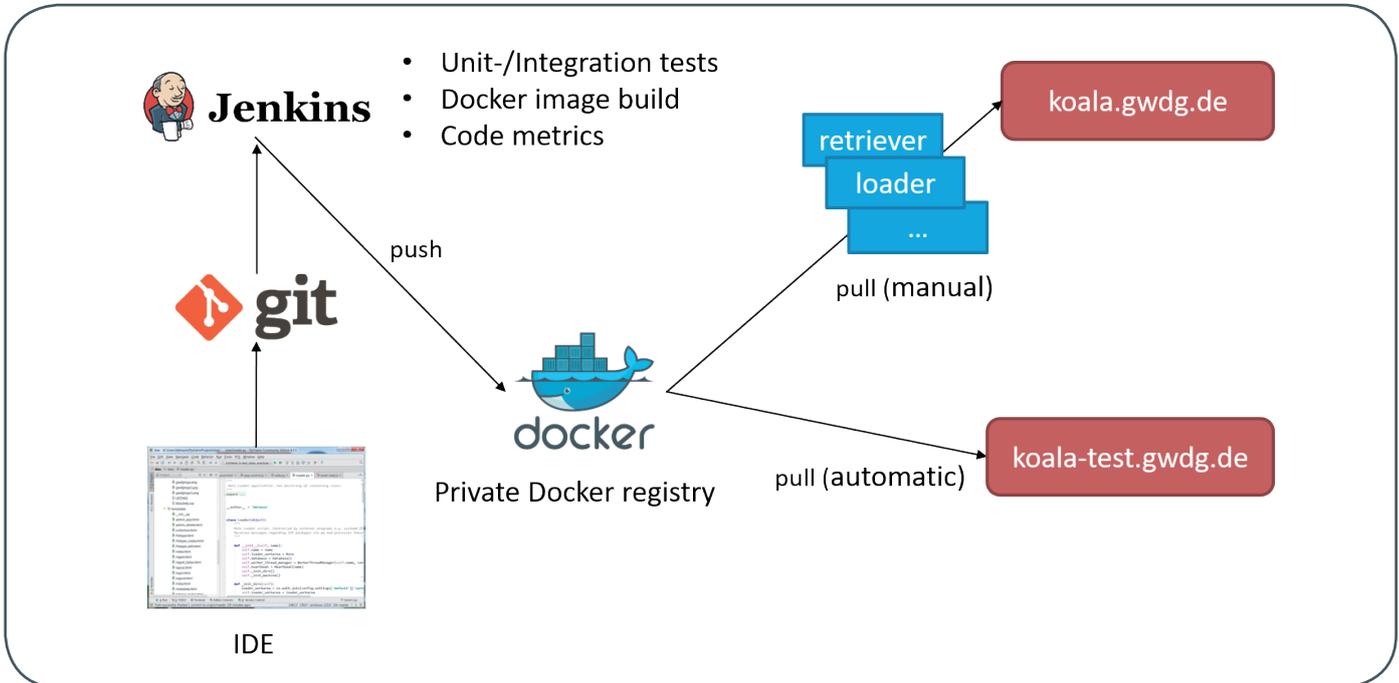
Tabelle 1: Virtuelle Maschinen

MIGRATION UND INBETRIEBNAHME

Parallel zur Entwicklung wurde ein Migrationskonzept erstellt, um die bisher in DIAS archivierten Daten verlustfrei in das

	Declarative: Checkout SCM	UnitTests	Render-documentation	Docker-build-images	Docker-integration-tests	Docker-push-to-registry	Redeploy-test	Declarative: Post Actions
Average stage times: (Average full run time: ~8min 43s)	9s	2min 15s	11s	1min 0s	2min 9s	38s	2min 0s	4s
#431 May 29 02:49 1 commits	8s	2min 30s	12s	2min 15s	2min 6s	42s	2min 1s	4s
#430 May 25 11:04 1 commits	9s	2min 17s	12s	19s	2min 5s	29s	1min 48s	4s
#429 May 24 16:01 1 commits	8s	2min 13s	11s	1min 25s	2min 15s	50s	2min 5s	2s
#428 May 17 10:15 1 commits	8s	2min 10s	10s	16s	2min 12s	29s	1min 47s	3s
#427 May 17 09:35 1 commits	13s	2min 15s	12s	17s	2min 8s	29s	1min 45s	3s
#426 May 17 08:51 2 commits	9s	2min 19s	11s	19s	2min 11s	35s	1min 58s	2s

3_koala-Pipeline



4_Development und Deployment

Folgesystem zu übernehmen. Durch Reverse Engineering von DIAS wurde ein Zugriff auf die Rohdaten im TSM möglich. Daraufhin wurde ein sukzessives Restore dieser Daten mit anschließender Aufbereitung und Ingest in das koala-System mithilfe der Umgebung des GWDG Scientific Compute Clusters implementiert. Durch die hohe Anzahl von verfügbaren Rechenknoten und die schnelle IO-Bandbreite konnte die Migrationszeit erheblich

reduziert werden. Außerdem wurde hiermit bereits nachgewiesen, dass koala hinsichtlich Skalierung auch größere Datenmengen verarbeiten kann. Nach erfolgreicher Übernahme der Daten wurden die beiden Systeme DIAS und koala bis zum endgültigen Ablösetermin parallel betrieben. In definierten Abständen wurden die in der Zwischenzeit neu hinzugekommenen Daten nach koala synchronisiert. Vor der Ablöse wurde der Datenbestand in koala

Puppetboard Overview Nodes Facts Reports Metrics Inventory Catalogs Radiator Query * 0.3.0

Details

Certname: koala.gwdg.de
 Facts: May 30 2018 - 09:04:06
 Catalog: May 30 2018 - 09:04:09
 Report: May 30 2018 - 09:04:16

Reports

Show 10 entries

End time	Status	301	0	0
May 30 2018 - 09:04:16	UNCHANGED	0	0	
May 30 2018 - 08:34:15	UNCHANGED	0	0	

Facts

Search:

Name	Value
virtual	vmware
uuid	4225B4B7-AB67-3BB9-3586-8043268DAD38
uptime_seconds	3019137
uptime_hours	838
uptime_days	34
uptime	34 days
uniqueid	4c862214
type	Other
trusted	{u'authenticated': u'remote', u'domain': u'gwdg.de', u'certname': u'koala.gwdg.de', u'extensions': {}, u'hostname': u'koala'}
timezone	CEST
swapspace_mb	2979.47
swapspace	2.91 GB

Copyright © 2013-2018 Puppet Community. Live from PuppetDB.

5_Oberfläche einer VM im Puppetboard

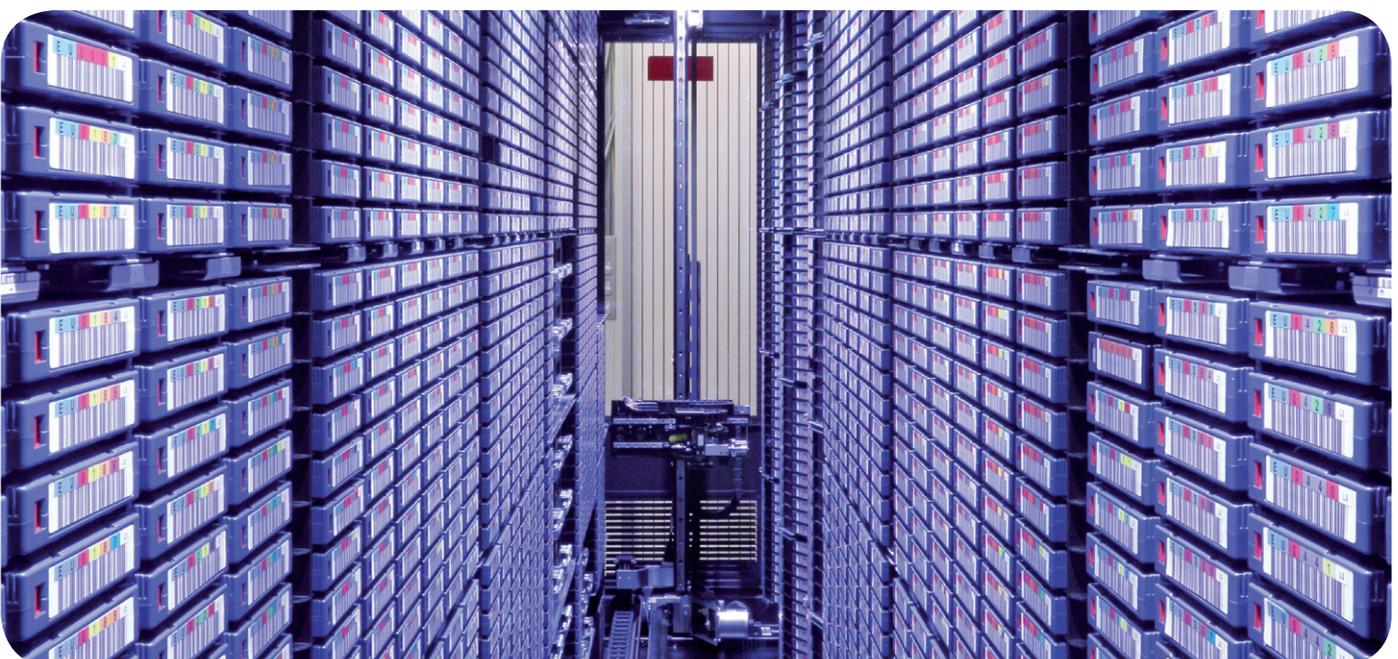
nochmals hinsichtlich Vollständigkeit und Integrität geprüft.

FAZIT UND AUSBLICK

Durch die Entwicklung einer Inhouse-Lösung wurde das Alt-System DIAS erfolgreich abgelöst. Mit der Verwendung von Open-Source-Lösungen und der GWDG-VMware-Virtualisierung konnten die Lizenz-, Hardware- und Stromkosten drastisch reduziert werden und gleichzeitig eine zukunftssichere Plattform geschaffen werden. Geplant ist, den Funktionsumfang von koala zu erweitern. So steht beispielsweise die Überlegung einer fortlaufenden Konsistenzprüfung inklusive Reporting im Raum.

FUSSNOTEN

- [1] DIAS: Digital Information Archiving System
- [2] Pylint: <https://www.pylint.org>
- [3] pep8: <http://pep8.readthedocs.io> sowie <https://www.python.org/dev/peps/pep-0008/>
- [4] GitLab: <https://gitlab.gwdg.de>
- [5] Jenkins: <https://jenkins.io>
- [6] Docker Registry: <https://docs.docker.com/registry/>
- [7] Siehe hierzu den Artikel „Neuer Dienst „Puppetserver“ im offenen Testbetrieb“ in den GWDG-Nachrichten 3/2017.





Using the Parallel Processing Power of the GWDG Scientific Compute Cluster

Upcoming Introductory and Parallel Programming Courses

GWDG operates a scientific compute cluster with currently 17,048 cores and a total compute power of 369 Teraflops ($3.69 \cdot 10^{14}$ floating point operations per second), which can be used by all scientists of the institutes of GWDG's supporting organisations, University of Göttingen and Max Planck Society.

In order to facilitate the access to and the efficient use of these computing resources, GWDG offers introductory and parallel programming courses, held at GWDG's site 'Am Faßberg' (1st and 3rd course) and 'Geiststraße' (2nd course).

The next courses in 2018 are

> October 22th, 9:30 am - 4:00 pm

Using the GWDG Scientific Compute Clusters – An Introduction

This course explains all steps for accessing GWDG's clusters, to compile and install software, and to work with the batch system for the execution of application jobs. The course is intended for new or inexperienced users of the clusters.

> October 23th - 24th, 9:15 am - 5:00 pm

Parallel Programming with MPI (Including MPI for Python)

This course introduces the message passing interface (MPI) for programming parallel applica-

tions in FORTRAN, C, and in Python. All concepts will be illustrated with hands on exercises. Examples of parallel applications will be presented and analysed.

> November 6th, 9:15 am - 5:00 pm

Programming with CUDA – An Introduction

Graphic processors (GPUs) are increasingly used as computational accelerators for highly parallel applications. This course introduces hardware and parallelization concepts for GPUs and the CUDA programming environment for C and Fortran, including the language elements for controlling the processor parallelism and for accessing the various levels of memory.

These three courses are repeated regularly. Other courses on parallel computing, dealing with more specialized topics can be arranged on demand. The possible subjects include parallel programming for shared memory systems and using extensions of C or Fortran with high level parallel constructs.

More Information about the courses held regularly or on demand at www.gwdg.de/scientific-computing-courses.

Information for registering for the courses at www.gwdg.de/courses.

If you have any further questions please contact support@gwdg.de.

NEUER MITARBEITER STEFFEN RÖRTGEN

Seit dem 1. Juni 2018 ist Herr Steffen Rörtgen als wissenschaftliche Hilfskraft in der Arbeitsgruppe „eScience“ (AG E) tätig. Herr Rörtgen bringt einen Bachelor-Abschluss in den Fächern Latein und Philosophie der Georg-August-Universität Göttingen mit. Darauf aufbauend studiert er zurzeit mit dem Ziel, seinen Master of Education in beiden Fächern zu erlangen. Herr Rörtgen wird im Forschungsprojekt „GRaCE“ arbeiten, einem Verbundprojekt am Wissenschaftsstandort Göttingen unter dem Dach des BMBF-Förderprogramms „Erforschung des Managements von Forschungsdaten in ihrem Lebenszyklus an Hochschulen und außeruniversitären Forschungseinrichtungen“. Seine Aufgaben werden im Bereich der Datenauswertung und in der organisatorischen Unterstützung des Projektes liegen. Herr Rörtgen ist per E-Mail unter steffen.roertgen@gwdg.de und telefonisch unter 0551 201-2117 erreichbar. Wieder



NEUER MITARBEITER STEFAN PFEIFFER

Seit dem 1. Juni 2018 verstärkt Herr Stefan Pfeiffer die Arbeitsgruppe „Anwendungs- und Informationssysteme“ (AG A) als studentische Hilfskraft. Er studiert zurzeit Angewandte Informatik mit Schwerpunkt Neuroinformatik an der Georg-August-Universität Göttingen. Sein Aufgabenschwerpunkt bei der GWGD wird im Bereich der Cloud-Speicherdienste liegen. Herr Pfeiffer ist per E-Mail unter stefan.pfeiffer@gwdg.de und telefonisch unter 0551 201-1821 zu erreichen. Krimmel



INFORMATIONEN:
support@gwdg.de
0551 201-1523

Juli bis
Dezember 2018

Kurse

KURS	VORTRAGENDE/R	TERMIN	ANMELDEN BIS	AE
STATISTIK MIT R FÜR TEILNEHMER MIT VOR- KENNTNISSEN – VON DER ANALYSE ZUM BERICHT	Cordes	03.07. – 04.07.2018 9:00 – 12:00 und 13:00 – 15:30 Uhr	26.06.2018	8
SHAREPOINT – EINFÜHRUNG FÜR ANWENDER	Buck, Kasper	15.08.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	08.08.2018	4
SHAREPOINT – EINFÜHRUNG FÜR SITECOLLECTION- BESITZER	Buck, Kasper	16.08.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	09.08.2018	4
GRUNDLAGEN DER BILDBEARBEITUNG MIT PHOTOSHOP	Töpfer	21.08. – 22.08.2018 9:30 – 16:00 Uhr	14.08.2018	8
SHAREPOINT – EINFÜHRUNG FÜR ANWENDER	Buck, Kasper	12.09.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	05.09.2018	4
SHAREPOINT – EINFÜHRUNG FÜR SITECOLLECTION- BESITZER	Buck, Kasper	13.09.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	06.09.2018	4
INDESIGN – GRUNDLAGEN	Töpfer	18.09. – 19.09.2018 9:30 – 16:00 Uhr	11.09.2018	8
SHAREPOINT – EINFÜHRUNG FÜR ANWENDER	Buck, Kasper	17.10.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	10.10.2018	4
SHAREPOINT – EINFÜHRUNG FÜR SITECOLLECTION- BESITZER	Buck, Kasper	18.10.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	11.10.2018	4

KURS	VORTRAGENDE/R	TERMIN	ANMELDEN BIS	AE
USING THE GWDG SCIENTIFIC COMPUTE CLUSTER – AN INTRODUCTION	Dr. Boehme, Dr. Ehlers	22.10.2018 9:30 – 16:00 Uhr	15.10.2018	4
PARALLELRECHNER-PROGRAMMIERUNG MIT MPI	Prof. Haan	23.10. – 24.10.2018 9:15 – 17:00 Uhr	16.10.2018	8
PHOTOSHOP FÜR FORTGESCHRITTENE	Töpfer	23.10. – 24.10.2018 9:30 – 16:00 Uhr	16.10.2018	8
PROGRAMMING WITH CUDA – AN INTRODUCTION	Prof. Haan	06.11.2018 9:15 – 17:00 Uhr	30.10.2018	4
EINFÜHRUNG IN DIE STATISTISCHE DATEN-ANALYSE MIT SPSS	Cordes	13.11. – 14.11.2018 9:00 – 12:00 und 13:00 – 15:30 Uhr	06.11.2018	8
ADMINISTRATION VON PCS IM ACTIVE DIRECTORY DER GWDG	Quentin	15.11.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	08.11.2018	4
SHAREPOINT – EINFÜHRUNG FÜR ANWENDER	Buck, Kasper	21.11.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	14.11.2018	4
SHAREPOINT – EINFÜHRUNG FÜR SITECOLLECTION-BESITZER	Buck, Kasper	22.11.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	15.11.2018	4
INDESIGN – AUFBAUKURS	Töpfer	27.11. – 28.11.2018 9:30 – 16:00 Uhr	20.11.2018	8
OUTLOOK – E-MAIL UND GROUPWARE	Helmvoigt	06.12.2018 9:15 – 12:00 und 13:00 – 16:00 Uhr	29.11.2018	4
ANGEWANDTE STATISTIK MIT SPSS FÜR NUTZER MIT VORKENNTNISSEN	Cordes	11.12. – 12.12.2018 9:00 – 12:00 und 13:00 – 15:30 Uhr	04.12.2018	8
SHAREPOINT – EINFÜHRUNG FÜR ANWENDER	Buck, Kasper	19.12.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	12.12.2018	4
SHAREPOINT – EINFÜHRUNG FÜR SITECOLLECTION-BESITZER	Buck, Kasper	20.12.2018 9:00 – 12:30 und 13:30 – 15:30 Uhr	13.12.2018	4

Teilnehmerkreis

Das Kursangebot der GWDG richtet sich an alle Mitarbeiterinnen und Mitarbeiter aus den Instituten der Universität Göttingen und der Max-Planck-Gesellschaft sowie aus einigen anderen wissenschaftlichen Einrichtungen.

Anmeldung

Anmeldungen können schriftlich per Brief oder per Fax unter der Nummer 0551 201-2150 an die GWDG, Postfach 2841, 37018 Göttingen oder per E-Mail an die Adresse support@gwdg.de erfolgen. Für die schriftliche Anmeldung steht unter <https://www.gwdg.de/antragsformulare> ein Formular zur Verfügung. Telefonische Anmeldungen können leider nicht angenommen werden.

Kosten bzw. Gebühren

Unsere Kurse werden wie die meisten anderen Leistungen der GWDG in Arbeitseinheiten (AE) vom jeweiligen Institutskontin-

gent abgerechnet. Für die Institute der Universität Göttingen und der Max-Planck-Gesellschaft erfolgt keine Abrechnung in EUR.

Absage

Sie können bis zu acht Tagen vor Kursbeginn per E-Mail an support@gwdg.de oder telefonisch unter 0551 201-1523 absagen. Bei späteren Absagen werden allerdings die für die Kurse berechneten AE vom jeweiligen Institutskontingent abgebucht.

Kursorte

Alle Kurse finden im Kursraum oder Vortragsraum der GWDG statt. Die Wegbeschreibung zur GWDG sowie der Lageplan sind unter <https://www.gwdg.de/lageplan> zu finden.

Kurstermine

Die genauen Kurstermine und -zeiten sowie aktuelle kurzfristige Informationen zu den Kursen, insbesondere zu freien Plätzen, sind unter <https://www.gwdg.de/kursprogramm> zu finden.



Gesellschaft für wissenschaftliche
Datenverarbeitung mbH Göttingen