Leitfaden zur Nutzung von HPC- und GPU-Ressourcen

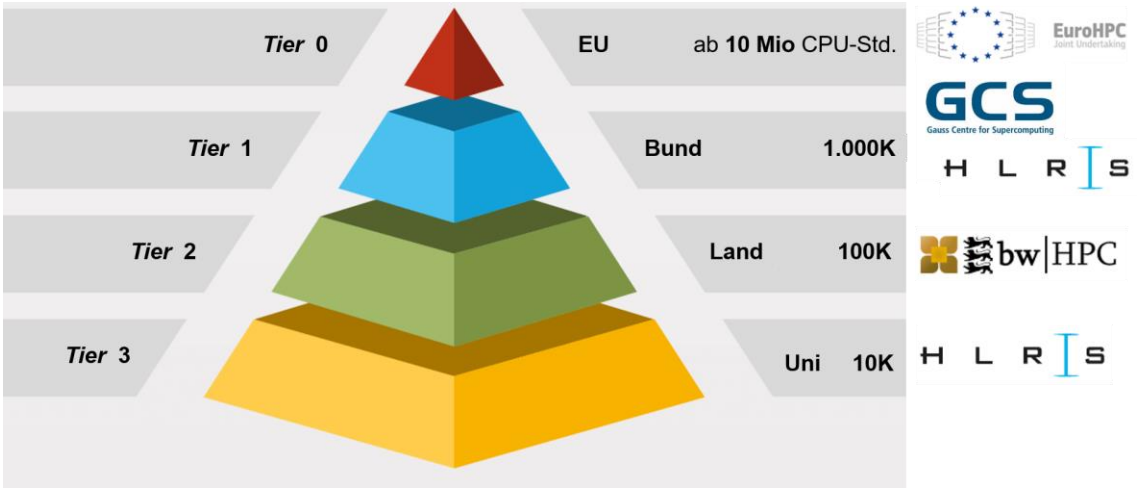
Inhaltsverzeichnis

- Einleitung
- Überblick der Tiers
  - Uni-Stuttgart-Systeme (Tier 3)
  - BW-Landessysteme (Tier 2/3)
  - Bundessysteme (Tier 1/2)
  - EU-Systeme (Tier 0)
- Zugangsformalitäten und finanzielle Aspekte
  - Uni-Stuttgart-Systeme (Tier 3)
  - BW-Landessysteme (Tier 2/3)
  - Bundessysteme (Tier 1/2)
  - EU-Systeme (Tier 0)

Einleitung

Hoch- und Höchstleistungsrechnen oder **High Performance Computing (HPC)** bezeichnet eine Form des Rechnens, die eine Computer-Infrastruktur mit sehr hoher Rechenleistung und/oder Speicherkapazität (zusammen mit dazugehörigen Software- und Anwendungs-Technologien) benötigt. Die Infrastruktur besteht in der Regel aus sehr vielen parallel geschalteten Rechenknoten (bestehend aus sog. Central Processing Units, kurz CPUs, welche wiederum aus einzelnen Kernen bestehen). Damit kann eine enorme Gesamtleistung erreicht werden, die in FLOPS – der Anzahl der erledigten Fließkommaoperationen pro Sekunde – inzwischen in der Peta-Dimension gemessen wird. Dank dieser HPC-Infrastruktur lassen sich viele Aufgabenstellungen in Bereichen wie bspw. Simulation, Data Intensive Computing (DIC) und Künstliche Intelligenz (KI) lösen, die aufgrund ihrer numerischen Komplexität lange nicht als lösbar galten. Ein großer Meilenstein in der Entwicklung des Hoch- und Höchstleistungsrechnens war die Einführung von **Graphics Processing Units (GPUs)** zu Computing-Zwecken, die Rechenoperationen parallel ausführen können, was für viele Anwendungen eine erhebliche Leistungssteigerung bedeutete.

Den Wissenschaftler*innen der Universität Stuttgart stehen an der Universität selbst, im Land Baden-Württemberg, in Deutschland und in Europa zahlreiche Systeme oder Systemverbünde für HPC zur Verfügung. Gemäß einer Leistungs- und Organisationspyramide sind diese Systeme in vier Ebenen (sog. Tiers) unterteilt. Die HPC-Rechner-Verbünde auf europäischer Ebene werden organisatorisch in einem „Partnership for Advanced Computing in Europe (PRACE)“ gebündelt und der Leistungsklasse Tier 0 zugerechnet. Auf nationaler Ebene befinden sich drei HPC-Rechner der Leistungsklasse Tier 1 – einer davon steht an der Universität Stuttgart beim Höchstleistungsrechenzentrum (HLRS). Die drei Tier-1-Rechner zählen organisatorisch zum Gauss Centre for Supercomputing (GCS). Ergänzend gibt es im Land Baden-Württemberg weitere Rechencluster, sowohl auf Tier-2- wie auf Tier-3-Ebene, siehe z. B. https://www.zki.de/fileadmin/user_upload/Downloads/HPC-Positionspapier_ZKI.pdf.



Welches System sich für welche Aufgabenstellung am besten eignet, hängt u.a. von der benötigten Rechenkapazität ab. In der obigen Skizze sind die für die in Baden-Württemberg Forschenden erreichbaren Leistungsklassen Tier 0 bis Tier 3 grob den Laufzeiten typischer Jobs zugeordnet. Auf jeder Ebene werden sowohl CPU- als auch GPU-Ressourcen angeboten.

Eine gute Übersicht über weitere regionale und überregionale HPC-Infrastrukturen findet man auch auf der deutschsprachigen Seite für Forschungsdatenmanagement https://forschungsdaten.info/themen/speichern-und-rechnen/hochleistungsrechnen/.

Dieser Leitfaden richtet sich an die Forschenden der Universität Stuttgart und beschreibt, welche Schritte im Einzelnen notwendig sind, um die Rechenmöglichkeiten der Ebenen Tier 0 bis Tier 3 zu nutzen. Meistens sind dazu mehrere Schritte notwendig, beginnend mit der Einholung verschiedener Zugangsformalitäten, die aus rechtlicher und technischer Sicht notwendig sind, um die wertvollen Rechenressourcen sorgfältig zu verwalten. Er enthält außerdem Hinweise, welches System für welche Aufgabenstellungen geeignet sein könnte.

Überblick der Tiers

Uni-Stuttgart-Systeme (Tier 3)

Diese Ebene wird für die Forschenden der Universität Stuttgart hauptsächlich durch die Ressourcen des Höchstleistungsrechenzentrums Stuttgart (HLRS), einer zentralen Einrichtung der Universität, abgedeckt. Das HLRS ist zwar auf ingenieur- und naturwissenschaftliche Anwendungen spezialisiert, wie z. B. Strömungsmechanik, Molekulardynamik und Chemie, es hat aber ebenso Expertise in anderen Anwendungsbereichen des Scientific Computing wie bspw. in Künstlicher Intelligenz, Datenanalytik, usw.

Aktuell bietet das HLRS folgendes Cluster-System auf Tier 3 an:

Cluster	Hersteller	Disziplinen	CPU-Knoten	GPU-Knoten	Dokumentation
Vulcan	NEC	Alle Disziplinen, mit Schwerpunkt auf Ingenieurwissenschaften	ja	ja	Link

BW-Landessysteme (Tier 2/3)

Das Land Baden-Württemberg fördert mehrere HPC-Systeme, die von allen Forschenden in Baden-Württemberg und somit auch von den Forschenden der Uni Stuttgart genutzt werden können. Dazu zählen in erster Linie die HPC-Ressourcen der bwHPC-Initiative:

Cluster	Standort/Betreiber	Disziplinen	CPU-Knoten	GPU-Knoten	Dokumentation
bwUniCluster 2.0	KIT, Karlsruhe	Alle Disziplinen	ja	ja	Link
bwForCluster Justus2	Uni Ulm	Chemie, Physik, Quantenphysik	ja	ja	Link
bwForCluster Helix	Uni Heidelberg	Biomedizin, Biologie	ja	ja	Link
bwForCluster NEMO	Uni Freiburg	Neuro-, Materialwissenschaften	ja	ja	Link
bwForCluster BinAC	Uni Tübingen	Bioinformatik, Astrophysik, Geowissenschaften	ja	ja	Link

Eine hilfreiche Übersicht und weitere Informationen findet man auch im bwHPC-Wiki (Link).

Bundessysteme (Tier 1/2)

Auch auf Bundesebene werden größere Systeme zum Durchführen besonders leistungshungriger HPC-Anwendungen angeboten. Am KIT und an der Universität Stuttgart gibt es dazu die folgenden Systeme:

Cluster	Tier	Standort/Betreiber	Disziplinen	CPU-Knoten	GPU-Knoten	Dokumentation
HoreKa	2	KIT, Karlsruhe	Alle	ja	ja	Link
Hawk	1	HLRS, Stuttgart	Alle	ja	ja	Link
Hunter	1	HLRS, Stuttgart	Alle	ja	ja	Link

EU-Systeme (Tier 0)

Auf der EU-Ebene werden mehrere Systeme der Höchstleistungsklasse (mit Kapazitäten im dreistelligen PetaFLOPS-Bereich) für die Anwendungen mit höchsten Anforderungen an Rechenleistung angeboten. Mehr Informationen zu EuroHPC-Systemen findet man unter diesem Link.

Zugangsformalitäten und finanzielle Aspekte

Uni-Stuttgart-Systeme (Tier 3)

Der erste Schritt zur Nutzung des Tier-3-Systems am HLRS (Vulcan siehe oben) ist die Benennung einer Ansprechperson ("Kontaktperson") an Ihrem Institut bzw. Ihrer zentralen Einrichtung, welche später die Aufgabe hat, die Zugänge für die Mitarbeiter*innen Ihres Instituts zu beantragen und zu verwalten sowie sog. Rechenprojekte (mit kurzer Beschreibung, Fachgebiet, Projektleitung, etc.) zu erstellen und zu administrieren. Das Formular zur Benennung der Kontaktperson

und eines Vertreters / einer Vertreterin findet man unter https://www.hlrs.de/fileadmin/for_users/AnschreibenWebInterface_Uni_v193.pdf.

Unter https://www.hlrs.de/de/rechnen-am-hlrs sind verschiedene Möglichkeiten beschrieben, um die Ressourcen des HLRS zu nutzen. Mitglieder und Angehörige der Universität Stuttgart können diese Ressourcen gegen ein im Vergleich zu den Preisen für die Industrie-Anwender*innen reduziertes Entgelt nutzen. Hier finden Sie die aktuelle Entgeltordnung (die im Jahresrhythmus angepasst wird). Falls in Ihrem Institutsetat nicht ausreichend Mittel zur Verfügung stehen, kann Ihr Institut einen Antrag auf Kostenübernahme bei der zentralen Verwaltung stellen. Den "Antrag auf Rechenzeit", der früher immer zum Jahresende für das Folgejahr gestellt wurde, gibt es jedoch nicht mehr. Für diese Anträge gibt es das Rundschreiben Nr. 42/2022 (Fonds für sonstige Maßnahmen), in dem ein Formular verlinkt ist, sowie das Rundschreiben Nr. 31/2023 (Verteilung der Haushaltsmittel 2024).

BW-Landessysteme (Tier 2/3)

Die BW-Landessysteme stehen Mitgliedern und Angehörigen aller Landesuniversitäten zur Verfügung, ohne dass dafür momentan ein Entgelt erhoben wird. Der Zugang zu den sog. bw-Clustern basiert auf dem persönlichen Benutzer*innenkonto der Universität (für Stuttgart ac- oder st-Account), das vom HLRS auf Antrag für die Nutzung eines bw-Clusters freigeschaltet wird.

Registrierung für bwUniCluster:

Es ist keine Beschreibung des Projekts erforderlich. Die Registrierung erfolgt durch die Nutzenden selbst in drei Schritten.

1. **bwUniCluster Entitlement** (Antragsberechtigung) der eigenen Hochschule einholen:
    - Prüfen, ob man die Antragsberechtigung bereits hat: Dazu meldet man sich auf der Webseite https://login.bwidm.de/user/index.xhtml mit einem Benutzer*innenkonto der Universität Stuttgart an. Im Tab "Shibboleth" werden dann die Antragsberechtigungen angezeigt, Wenn in der Liste unter `urn:oid:1.3.6.1.4.1.5923.1.1.1.7` das folgende enthalten ist, hat man die Antragsberechtigung bereits und kann zum Schritt 2. gehen: http://bwidm.de/entitlement/bwUniCluster (dies ist kein Link, sondern die Bezeichnung einer Berechtigung).

    - Antragsberechtigung/Freischaltung der Universität Stuttgart beim HLRS einholen (siehe auch: bwUniCluster-Zugang (hlrs.de)):

      Die Freischaltung des eigenen Benutzer*innenkontos für bwUniCluster wird mit einer E-Mail an bwunicluster@hlrs.de mit folgenden Angaben beantragt:

        - vollständiger Name und akademischer Titel
        - Geschlecht
        - eindeutiges Benutzer*innenkonto (ac-Account oder st-Account)
        - Kontakt-E-Mail-Adresse
        - Betreuer*in (Name & Institut) (nur für Studierende erforderlich)
        - Projekttitel
        - Kurze Zusammenfassung des Projekts

      Das HLRS wird den Antrag bearbeiten, das Benutzer*innenkonto für die Nutzung des bwUniClusters freischalten und eine Benachrichtigung an die angegebene E-Mail-Adresse senden.

2. Für den **bwUniCluster-Service registrieren:** Nach der Freischaltung des Benutzer*innenkontos durch das HLRS kann man sich für die Nutzung des bwUniClusters am KIT über die Webseite https://bwidm.scc.kit.edu registrieren.

3. **bwUniCluster-Fragebogen innerhalb von 14 Tagen ausfüllen** (sonst verliert man den Zugang): Es wird nach dem wissenschaftlichen Fachgebiet, den verwendeten numerischen Methoden, dem Tätigkeitsbereich und der Art der Tätigkeit gefragt.

Registrierung für bwForCluster: Die Registrierung für ein bwForCluster erfolgt in drei Schritten, wobei die Schritte 1 und 2 parallel durchgeführt werden können:

1. **bwForCluster Entitlement** (Antragsberechtigung) der eigenen Hochschule einholen:
    - Prüfen, ob man die Antragsberechtigung bereits hat: Dazu meldet man sich auf der Webseite https://login.bwidm.de/user/index.xhtml mit einem Benutzer*innenkonto der Universität Stuttgart an. Im Tab "Shibboleth" werden dann die Antragsberechtigungen angezeigt, Wenn in der Liste unter `urn:oid:1.3.6.1.4.1.5923.1.1.1.7` das folgende enthalten ist, hat man die Antragsberechtigung bereits und kann zum Schritt 2. gehen: http://bwidm.de/entitlement/bwForCluster (dies ist kein Link, sondern die Bezeichnung einer Berechtigung).
    - Antragsberechtigung/Freischaltung der Universität Stuttgart beim HLRS einholen (siehe auch: bwForCluster-Zugang (hlrs.de)):

      Die Freischaltung des eigenen Benutzer*innenkontos für bwForCluster wird mit einer E-Mail an bwunicluster@hlrs.de mit folgenden Angaben beantragt:
        - vollständiger Name und akademischer Titel
        - Geschlecht
        - eindeutiges Benutzer*innenkonto (ac-Account oder st-Account)
        - Kontakt-E-Mail-Adresse
        - Betreuer*in (Name & Institut) (nur für Studierende erforderlich)
        - Projekttitel
        - Kurze Zusammenfassung des Projekts

      Das HLRS wird den Antrag bearbeiten, das Benutzer*innenkonto für die Nutzung von bwForCluster freischalten und eine Benachrichtigung an die angegebene E-Mail-Adresse senden.
2. **Registrierung bei der Zentralen Anmeldestelle (ZAS)**, um die Genehmigung des Cluster Assignment Teams (CAT) und die Zuordnung zu einem bwForCluster zu erhalten: Man kann ein neues Rechenvorhaben anmelden oder an einem vorhandenen Rechenvorhaben mitarbeiten (falls es bereits ein Rechenvorhaben der eigenen Arbeitsgruppe gibt). Ein Rechenvorhaben umfasst die geplanten Rechenaktivitäten einer Gruppe von Wissenschaftler*innen auf einem der bwForCluster.
3. **Web-Registrierung auf der bwForCluster-Seite:**
   Nach der Registrierung bei der Zentralen Anmeldestelle erhält man eine E-Mail mit einem Link zu einer Webseite des zugewiesenen Clusters, auf der man sich dann einen Account einrichten kann.

Informationen für registrierte Benutzer*innen:

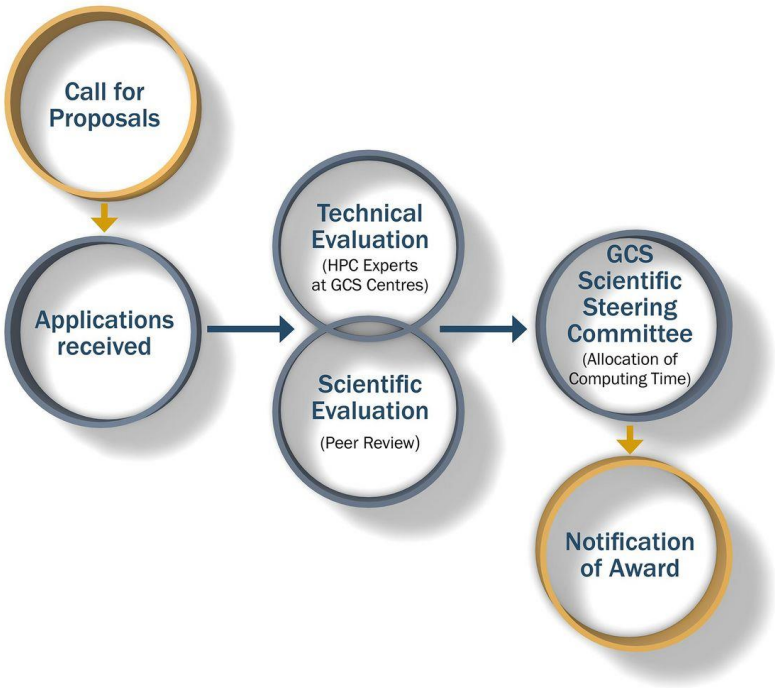
- Anleitung für den Login auf einem bwForCluster

Bundessysteme (Tier 1/2)

Das Höchstleistungszentrum Stuttgart (HLRS)

Da das HLRS Teil des Gauss Centre for Supercomputing ist, können über die entsprechende Governance-Struktur Rechenprojekte beantragt werden. Dabei wird zwischen "GCS Large-Scale

Projects" (für Details siehe HPC Access: Gauss Centre for Supercomputing e.V. (Link)) und kleineren Projekten (siehe Link) unterschieden.



Bei einer entsprechenden Bewilligung durch das Gauss Centre for Supercomputing fallen keine Kosten an. Bei Ablehnung kann die Nutzung gegen die vollen Gebühren erfolgen, es gibt dann keine Reduzierung für die Universität Stuttgart.

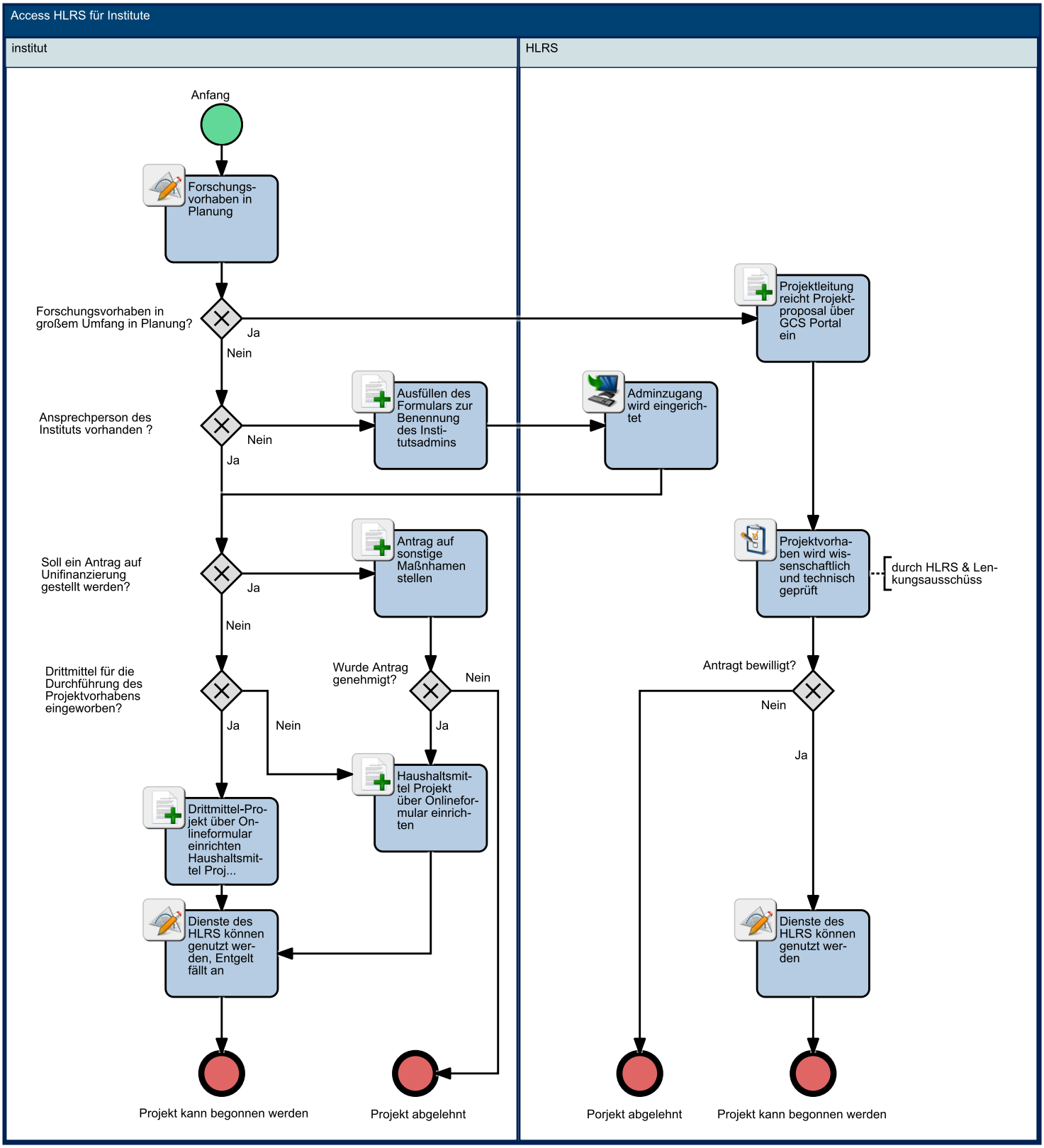
Hochleistungsrechner Karlsruhe (HoreKa)

Auch hier können Forschende der Universität Stuttgart jederzeit Anträge auf Rechenzeit stellen. Dazu genügt eine kurze Projektbeschreibung. Bei Bewilligung fallen keine Kosten für die Nutzung an.

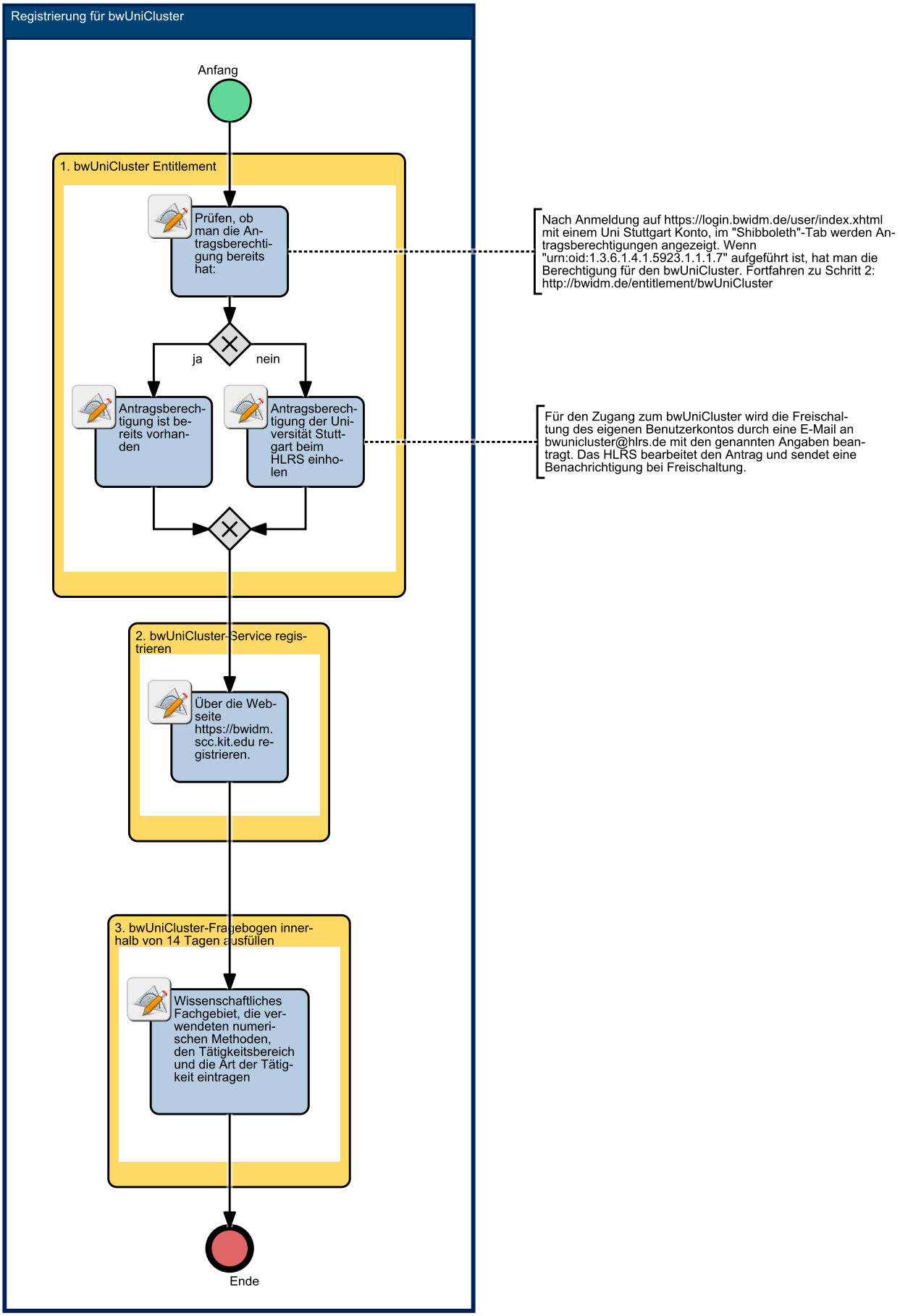
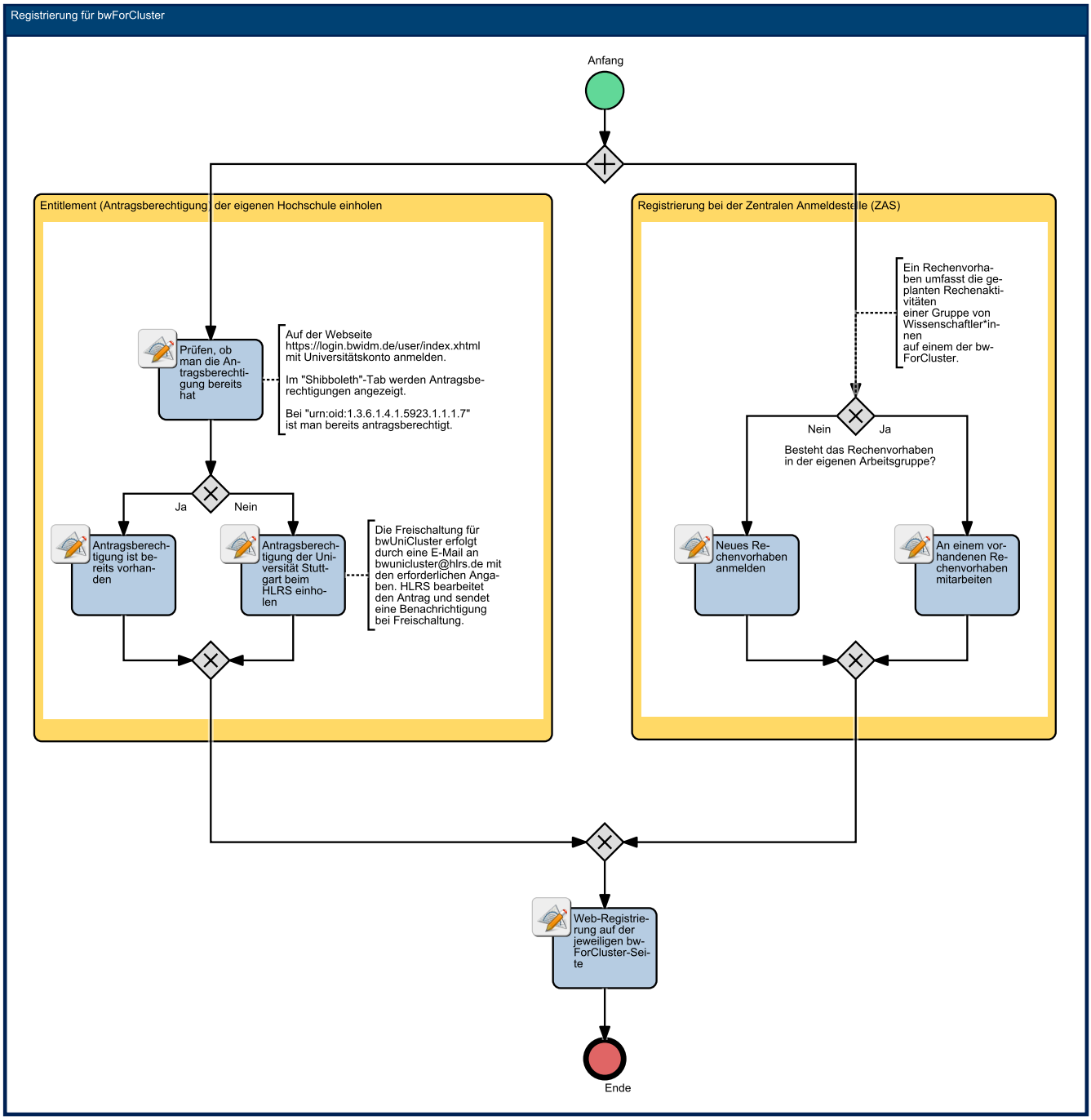
Es stehen CPU-Ressourcen ("HoreKa Blue, 60.000 CPU-Kerne) und GPU-Ressourcen ("HoreKa Green", 668 NVIDIA A100-GPUs) zur Verfügung. Weitere Informationen findet man auf der Seite von HoreKa (Link). Eine genaue Beschreibung der Zugangsmodalitäten findet sich in der HoreKa-Systemdokumentation (Link).

EU-Systeme (Tier 0)

Auf EU-Ebene (Tier 0) bietet die Seite "The European High Performance Computing Joint Undertaking (EuroHPC JU)" (Link) einen allgemeinen Überblick über die vorhandenen Systeme. Hier findet man auch die Bedingungen und Wege zur Nutzung (Antragsverfahren) der entsprechenden Systeme, siehe EuroHPC Access Calls (europa.eu).


Access HLRS für Institute
institut
HLRS
Anfang
Forschungs-vorhaben in Planung
Forschungsvorhaben in großem Umfang in Planung?
Ja
Nein
Projektleitung reicht Projekt-proposal über GCS Portal ein
Ansprechperson des Instituts vorhanden ?
Nein
Ja
Ausfüllen des Formulars zur Benennung des Insti-tutsadmins
Adminzugang wird eingerich-tet
Soll ein Antrag auf Unifinanzierung gestellt werden?
Ja
Nein
Antrag auf sonstige Maßnhamen stellen
Projektvorha-ben wird wis-senschaftlich und technisch geprüft
durch HLRS & Len-kungsausschüss
Drittmittel für die Durchführung des Projektvorhabens eingeworben?
Ja
Nein
Wurde Antrag genehmigt?
Nein
Ja
Antragt bewilligt?
Nein
Ja
Haushaltsmit-tel Projekt über Onlinefor-mular einrich-ten
Drittmittel-Pro-jekt über On-lineformular einrichten Haushaltsmit-tel Proj...
Dienste des HLRS können genutzt wer-den, Entgelt fällt an
Dienste des HLRS können genutzt wer-den
Projekt kann begonnen werden
Projekt abgelehnt
Porjekt abgelehnt
Projekt kann begonnen werden

Registrierung für das bwUniCluster



Registrierung für das bwForCluster