



# BIOINFORMATIK-CLOUD FÜR DIE MODERNE FORSCHUNG

SVA ermöglicht mit Red Hat OpenStack, Ansible und GitLab den optimierten Datenzugriff für BIH und de.NBI.

## AUF EINEN BLICK

### AUFGABE

Aufbau, Betrieb und Weiterentwicklung einer Cloud-Umgebung für Bioinformatiker im Kontext des de.NBI-Projektes.

### SYSTEME UND SOFTWARE

- > Red Hat OpenStack Plattform
- > GitLab Premium
- > Ansible

### VORTEILE

- > Einheitliche Schnittstelle für Infrastructure as Code (IaC) durch OpenStack
- > Automatisierung/Compliance Checks mit Ansible
- > Team Workflow (Git, Issues, Pipelines) zentral in GitLab abgebildet

## BIH UND DE.NBI

Das Berlin Institute of Health (BIH) ist eine Wissenschaftseinrichtung für Translation und Präzisionsmedizin. Das BIH widmet sich neuen Ansätzen für bessere Prognosen und neuartigen Therapien bei progredienten Krankheiten (fortschreitenden Krankheiten) und ungelösten Gesundheitsproblemen, um Menschen Lebensqualität zurückzugeben oder sie zu erhalten.

Als eines von acht Instituten ist das BIH Partner im „Deutschen Netzwerk für Bioinformatik-Infrastruktur“ (de.NBI), einem nationalen, akademischen Non-Profit-Verbund, der Nutzern in den Bereichen Biowissenschaften und Biomedizin in Deutschland und Europa bioinformatische Services zur Verfügung stellt. Die Partner organisieren Schulungs-Events und Lehrgänge auf Tools, Standards und Compute Services, die vom de.NBI angeboten werden, um Forscher bei der effizienteren Untersuchung ihrer Daten zu unterstützen. Über das de.NBI sind mittlerweile aus den acht verbundenen Zentren mehr als 40 Forschungs-, Service- und Infrastruktur-Gruppen mit etwa 150 Bioinformatikern zusammengeschlossen.

## HERAUSFORDERUNG: UNEINHEITLICHER RESSOURCENZUGRIFF

Bislang gab es keinen einheitlichen Weg für Bioinformatiker, um an Ressourcen für z. B. Analysen von Genomdaten zu kommen – diese wurden sowohl aus Public Clouds verwendet als auch aus klassischen lokalen HPC Clustern an den einzelnen Standorten. Bestehende Bioinformatik-Pipelines waren aufgrund der unterschiedlichen Beschaffenheit nicht überall einsetzbar. Durch das de.NBI-Projekt sollte eine einheitliche Schnittstelle geschaffen werden, mit der Bioinformatiker über die Standorte hinweg Ressourcen anfordern und damit rechnen können. Damit sollte eine gemeinsame Basis zur schnelleren Entwicklung von Bioinformatik-Pipelines geschaffen werden. Von Beginn an war dabei ein hoher Grad an Automatisierung gefordert sowie eine zentrale Ablage des verwendeten Automatisierungscodes.



## HOHER GRAD AN AUTOMATISIERUNG

### LÖSUNG: RED HAT OPENSTACK & ANSIBLE

Für das SVA-Experten-Team lag die Lösung ganz klar bei OpenStack: Durch dessen Verwendung innerhalb des de.NBI-Konsortiums konnte eine einheitliche Schnittstelle für die Nutzung von *Infrastructure as Code (IaC)* implementiert werden. Konkret wird der Betrieb für den Standort Berlin und in Heidelberg für das DKFZ und das Bioquant (Uni Heidelberg) durch den hohen Grad an Automatisierung sowie die Skalierung vereinfacht. Die bestehenden Blueprints können einfach auf weitere Standorte ausgerollt werden. Mit der Unterstützung von GitLab und Ansible konnte ein Multi Cloud Management etabliert werden, das standardisierte Umgebungen definiert.

Um manuelle Konfigurationen und dadurch eine inhomogene Konfiguration der Systeme zu vermeiden, wurde auf das Open Source Automatisierungs-Werkzeug **Ansible** gesetzt. Mit Ansible ist es möglich, ein zentrales Inventory von allen Systemen zu pflegen, diese zu gruppieren und entsprechende Rollen darauf anzuwenden. Dabei sind die Ansible-Rollen, welche im Rahmen dieser Implementierung entwickelt wurden, modular aufgebaut, so dass ein hoher Grad an Wiederverwertbarkeit sichergestellt werden kann. Mit den Rollen werden dann auf den Systemen Konfigurationen hergestellt, wie z. B. eine einheitliche Basishärtung der Linux-Betriebssysteme oder installierte Softwarepakete. Ein weiterer Vorteil von Ansible ist die Möglichkeit, durch regelmäßiges Ausführen der sogenannten Playbooks einen Compliance Check der Systeme zu etablieren, um Abweichungen (z. B. durch ungewollte manuelle Konfiguration) schnell festzustellen.

### OPTIMIERTE TEAMARBEIT MIT GITLAB

Da das Team der „Cloud-Admins“ über verschiedene Standorte in Deutschland verteilt ist und eine enge Zusammenarbeit/Dokumentation der Tasks gefordert ist, wird **GitLab** nicht nur als Git-Repository genutzt, sondern als zentrales Tool zur Organisation des kompletten Teams. Durch die Verwendung der *Issues*-Funktion können die Aufgaben zentral organisiert und für alle Teammitglieder transparent gemacht werden. GitLab fördert hierbei durch Funktionen wie Kommentare und Labels vor allem die asynchrone Abstimmung und Nachvollziehbarkeit. Damit wird doppelte Arbeit aufgrund von mangelnder Kommunikation innerhalb des Teams vermieden. Dokumentationen von abgeschlossenen Arbeiten können mit der integrierten Wiki-Funktionalität abgedeckt werden. GitLab bietet hier also ein Tool nicht nur für Entwickler, sondern auch für das Projektmanagement sowie die Projektverantwortlichen, die sich jederzeit über den aktuellen Status erkundigen können.

Die SVA-Experten konnten das Projekt von der Erstellung des Konzeptes, über den Aufbau der Hardware und das Projektmanagement bis zur Betriebsbereitschaft erfolgreich begleiten. Stand heute sind drei Standorte mit ca. 200 physikalischen Hosts sowie diverser VMs für das Management der Umgebung verknüpft, hochgradig automatisiert und mit einer zentralen Ablage des verwendeten Automatisierungscodes. Auch weiterhin unterstützt SVA das de.NBI bei der Weiterentwicklung der Umgebung und arbeitet mit dem Kunden als ein gemeinsames „Cloud-Team“.

### KONTAKT

SVA System Vertrieb  
Alexander GmbH  
Borsigstraße 14  
65205 Wiesbaden  
Tel. +49 6122 536-0  
Fax +49 6122 536-399  
mail@sva.de  
www.sva.de

© SVA GmbH  
Alle Marken- und Produktnamen  
sind Warenzeichen und werden  
als solche anerkannt.