Betriebs- und Sicherheitskonzept

Nutzung der Large Scale Data Facility (LSDF) im BioQuant (BQ) / Uni Heidelberg durch das DKFZ

Stand: 28.7.2010
Version: 2010-07
Autoren: Holger Haas, Jürgen Eils, Marc Hemberger

Abzeichnung:
(ok) Holger Haas, DKFZ, ITCF
(ok) Roland Eils, DKFZ, TBI
(ok) Marc Hemberger, BQ
(ok) Rolf Kabbe, DKFZ, TBI
(ok) Jürgen Eils, DKFZ, TBI
(ok) Tobias Reber, DKFZ, ITCF
(ok) Nils Kückenhoff, DKFZ, ITCF
(ok) Hartmuth Heldt, Uni HD, URZ

Management Overview

Das vorliegende Konzept dient zur Beschreibung der Sicherheitsrisiken, die sich durch die Einrichtung der 10-Gbit-Standleitung zwischen DKFZ und Bioquant ergeben und ist vor allem an die verantwortlichen Sysadmins gerichtet. Das Dokument ist von den IT-Hauptverantwortlichen Holger Haas und Marc Hemberger sowie dem ICGC-Projektverantwortlichen Datenmanagement Juergen Eils konzipiert worden.

Es stehen momentan nur wenige zentrale Sicherheitsmechanismen wie die Einschränkung der Dienste und der beteiligten Rechner auf dem Link zur Verfügung.

- Es gibt keine zentralen Virenscanner auf der LSDF.
- Eine Verkehrskontrolle auf dem Link ist technisch nicht möglich und
- durch das eingesetzte Protokoll NFS V3 ist eine Rechte-Verwaltung auf User-Ebene nicht möglich.

Das Risikopotential ist bezüglich

- Verteilung von Windows-Viren,
- Störungen auf dem 10GBit/s-Link zwischen 2 Institutionen und Seiteneffekte auf das jeweilige LAN,
- Probleme bei der Zuordnung von Daten und Usern

nicht unerheblich. Deshalb müssen alle beteiligten Sysadmins die dezentralen Maßnahmen und Vorgaben verantwortlich umsetzen.
Die Zeit, in der NFS V4 nicht zur Verfügung steht (angekündigt für Q1 2011 für LSDF im BQ) ist ausdrücklich als Übergangszeit deklariert mit eingeschränktem Nutzungsspektrum (Gensequenzierungsdaten aus Projekten ICGC, Abtlg. TBI).

Maßnahmen und Vorgaben

1. Der direkte Zugriff auf die LSDF im BQ (bq-lsdf) erfolgt vom DKFZ aus ausschließlich über den Spezial-Link „DKFZ-URZ-Link“ zwischen DKFZ und URZ Heidelberg. Dieser Link hat eine Kapazität von 10 GBit/s und ist als IP-Link ausgestaltet.

2. Auf dem Link werden ausschließlich die notwendigen Dienste (nfs über Nicht-Standard-Ports) und die benannten Rechner (DKFZ-Abtlg. TBI, LSDF BQ) für Produktion und Test (SRVnet, DKFZBBnet, Netzwerkmanagement URZ) freigeschaltet. Es können keine Adressen nach RFC 1918 verwendet werden.

3. Solange bq-lsdf selbst keine Viren scannt, müssen alle Rechner, die darauf zugreifen, eigene Virenscanner betreiben und dafür Sorge tragen, dass nur Daten auf bq-lsdf geschrieben werden, die keine Viren enthalten. Dies betrifft auch die Windows-Rechner, die über einen temporären Zwischenspeicher (z. B. Illumina Pipelineserver im BQ) Daten in der LSDF indirekt ablegen.

4. Die nfs-gemounteten Fileshares dürfen nicht weiterexportiert werden.

5. **Zugriff vom DKFZ aus**: Solange nfs v4 nicht zur Verfügung steht, ist keine Rechte-Verwaltung auf User-Ebene möglich. Der nfs-mount-point auf den benannten Rechnern ist so angelegt, dass nur Mitglieder der Gruppe B080-ICGC vom DKFZ darauf zugreifen können (zur Zeit 16 Personen). Die Sequenzier-Messdaten und die von B080 generierten Analyse-Daten werden auf der LSDF unter zwei verschiedenen User-IDs und in getrennten Dateistrukturen abgelegt. So sollte weitestgehend vermieden werden, dass die Sequenzier-Messdaten versehentlich überschrieben werden können.

6. **Verantwortlichkeiten:** Die Netzwerkeinstellungen auf den Netzwerk-Komponenten ebenso wie die Netzwerk-Management-Einstellungen verantworten
   - für das DKFZ die ITCF/AGNetz
   - für die Uni HD das URZ/Team Infrastruktur

   Die verantwortlichen Sysadmins für Rechner sind
   - im BQ Marc Hemberger
   - im DKFZ/ITCF Tobias Reber
   - im DKFZ/TBI Rolf Kabbe

7. **Zusatz: LSDF-Zugriff der TBI vom BQ**: Die Sequenzier-Messdaten werden auf die LSDF nicht nur über die Standleitung vom DKFZ aus geschrieben. Daher wird dieser Passus hier als Ergänzung aufgeführt, obwohl er nicht direkt die Standleitung betrifft. Das Speichern der Sequenzier-Messdaten wird von externen Kooperations-Partnern auf einem FTP-Server erfolgen. Die übertragenen Daten werden in Zukunft (geplant Juli/Augst 2010) lokal zwischen gespeichert, auf Viren gecheckt und in einem zusätzlichen Prozess auf die LSDF kopiert werden. Auf diesem Rechner haben nur die Personen Zugang, die für das Daten-Management des ICGC-Projekts verantwortlich sind.