Von Einheit und Mannigfaltigkeit

Erfahrungen bei der sicheren Fernanbindung der Fakultäten

Till Hoke
till.hoke@cms.hu-berlin.de

Seit nunmehr über drei Jahren sind Daten der zentralen Verwaltungsdatenbanken der Humboldt-Universität der direkten Bearbeitung durch Sachbearbeiterinnen in den Fakultäten zugänglich. Da es inzwischen kaum eine Fakultät oder Zentraleinrichtung ohne einen derartigen Zugang zum Verwaltungsnetz gibt, scheint es angebracht, die Bedingungen für den dezentralen Zugriff sowie die sich ergebenden Schwierigkeiten zu erläutern.

Zentral versus dezentral

Das Kernproblem ist kurz umrissen. Auf der einen Seite gibt es eine zentrale Verwaltung der Universität mit einem durch eine Firewall geschützten Netz, in dem sich unter anderem die zentralen Datenbanken für die Studenten-, Personal- und Haushaltsdaten befinden. Auf der anderen Seite gibt es Verwaltungen in den Fakultäten, in der Regel mit fakultätsinterner EDV-Abteilung und eigener Netzwerkorganisation. Und es ergeben sich inhaltliche Zwänge aus den Arbeitsprozessen in den Verwaltungen, die einen direkten Zugriff auf zentrale Datenbestände erfordern. Es sind momentan zwei Anwendungen, über die dieser Zugriff erfolgen soll: HISPOS für die Prüfungsverwaltung und HISFSV für die Arbeit mit Haushaltsdaten. Aus technischer Sicht gleichen sich die Vorgänge beim Start beider Anwendungen. In diesem Artikel sollen allerdings die Probleme rund um die Prüfungsverwaltung (ssoftware) näher beleuchtet werden.

Mit der Einführung modularisierter Bachelor- und Masterstudiengänge ergeben sich neue Anforderungen an die Verwaltung der anfallenden Daten. Zum einen macht die gewachsene Komplexität der Prüfungsverwaltung (messbar an der zunehmenden Anzahl der Prüfungsfälle, aber auch an der Qualität der Prüfungsordnungen mit Studienkonten, Bonuspunkten etc.) eine elektronische Datenverarbeitung unerlässlich, zum anderen erfordern die Verflechtung der Studiengänge über Grenzen der Fachbereiche hinweg und die angestrebte internationale Vergleichbarkeit von Abschlüssen eine Vereinheitlichung der Arbeitsabläufe.

Als Beispiel sei hier nur die sich abzeichnende Notwendigkeit, Prüfungsaktivitäten über Strukturgrenzen hinweg zu koordinieren, genannt. Dieser Bedarf an Vereinheitlichung ist zunächst ein inhaltlicher, und es ist nicht Sache des Rechenzentrums, Standards hierfür zu formulieren. Der geäußerte Vorwurf, die Lehre müsse sich nun den Anforderungen der EDV unterordnen, erscheint vor diesem Hintergrund mehr als merkwürdig. Es ist der Strukturwandel in der Lehre selbst, welcher ein Nachdenken über die Organisation der Prüfungsverwaltung an der Universität mit sich bringt. Warum gibt es beispielsweise eine strikte Trennung von Studien- und Prüfungsverwaltung? Wem unterstehen eigentlich die Prüfungsämter? Dem Prüfungsausschuss, dem Verwaltungsleiter? Diese Fragen führen aus der rein verwaltungstechnisch relevanten inhaltlichen Dimension des Begriffspaares zentral/dezentral hinaus. Hier gibt es zum Beispiel datenschutzrechtliche Implikationen. Während eine Mitarbeiterin in der Studienverwaltung (zentral) sämtliche Daten aller Studenten einsehen und bearbeiten darf, kann eine Mitarbeiterin des Prüfungsamtes (dezentral) einem Studenten keine Modulabschlussbescheinigung für ein an ihrem Fachbereich bestandenes Modul ausstellen, wenn diese im Kopf private Daten des Studenten enthält und der Student das Modul im Rahmen eines Studienganges aus einem anderen Fachbereich besucht hat. Eine Prüfungsamtsmitarbeiterin kann einem solchen Studenten eine Leistung verbuchen, hat es dabei aber ungleich schwerer, einen Fehler zu korrigieren, wenn der Student außerhalb ihrer Studiengangskompetenzen

liegt. Ohne eingehende SQL-Kenntnisse kann eine Mitarbeiterin im Prüfungsamt der Juristen nur unvollständige Daten über eine Prüfung aus einem vergangenen Semester einholen, wenn unter den Prüflingen Studenten sind, die inzwischen den Fachbereich gewechselt haben. Diese Aufzählung ließe sich problemlos verlängern und wird durch die Praxis weiter wachsen. Rechte auf die Daten werden zwar durch das Rechenzentrum in der Datenbank gesetzt, aber nach Vorgaben der Eigentümer dieser Daten und des Datenschutzes. Und momentan enden die Möglichkeiten der Prüfungsämter an den Grenzen der Fachbereiche.

Dieser kurze Abriss inhaltlicher Probleme ist nicht der eigentliche Gegenstand dieses Artikels. Er wurde nur eingefügt, um aufzuzeigen, dass die Schwierigkeiten, die mit dem Zugang zur Datenbank verbunden sind und die eher mit der dezentralen Struktur der DV zu tun haben, ihre Entsprechung in der inhaltlich-organisatorischen Ebene haben. Im folgenden Abschnitt wird dieser schwierige Zugang einmal so beschrieben, wie eine Mitarbeiterin ihn erleben muss, allerdings mit einer kurzen Erläuterung der dahinter liegenden Prozesse.

Einfach versus vielfach

Aus Nutzersicht geht es einfach darum, durch das Anklicken eines Icons eine Anwendung zu starten. Aus technischer Sicht handelt es sich um eine Lawine von Prozessen, ausgelöst von dem einen Doppelklick auf das Icon. Jene Prozessflut sollte eigentlich für den Benutzer durchsichtig, nicht wahrnehmbar sein. In der Praxis fällt das Brodeln in der Tiefe der Systeme der Nutzerin[1] allerdings durch eine Vielzahl von Logins auf – im günstigsten Falle, denn technische Probleme verringern die Transparenz zusätzlich. Werfen wir einen kurzen Blick auf diese Mannigfaltigkeit von Hürden, die sich vor dem Start einer einfachen Anwendung auftut.

1. Die Nutzerin startet einen PC (im Folgenden kurz der *VPN-PC* genannt), der für den Zugriff auf die Verwaltungsdatenbanken besonders konfiguriert wurde. Es handelt sich um ein Windows 2000- bzw. Windows XP Professional System mit einem Terminalservice-Client und einer Software (im Folgenden *VPN-Client*), welche die Absicherung des Netzverkehrs von und zu dem VPN-PC besorgt. An diesem PC erfolgt ein erstes Login.
2. Die Nutzerin startet per Doppelklick auf ein Icon eine Anwendung. Diese Anwendung ist aber noch nicht die HISPOS- oder FSV-Software, sondern der erwähnte Terminalservice-Client, das Gegenstück zu einem Serverdienst, dem Terminalservice. Der Terminalservice läuft auf Servern, die Benutzern Anwendungen und Ressourcen wie Speicherplatz und Prozessorleistung zur Verfügung stellen. Der Terminalservice-Client, der durch die Nutzerin gestartet wurde, versucht nun – anhand besonderer Konfigurationsinformationen – einen Terminalservice zu erreichen, einen Terminalservice, der die in Rede stehenden HIS-Programme anbietet.
3. Der Start des Terminalservice-Client löst auf dem VPN-PC weitere Prozesse aus. Denn die Suche des Terminalservice-Client nach einem passenden Terminalservice geht über das Netzwerk. Der Zugang zum Netz aber wird kontrolliert durch den VPN-Client. Dieser verfügt über seine eigene Konfiguration, ein Regelwerk, in dem aufgelistet ist, welche Art Netzwerkverkehr von und zum VPN-PC gestattet ist. Er verhindert etwa das Surfen im Internet oder den Versand von E-Mails. Sein Regelwerk hält aber auch für die Konversation des Terminalservice-Client mit seinem Terminalserver gewisse Spielregeln bereit. Und diese verlangen zunächst einmal Identifikation und Authentifikation (d. h. Identitätsnachweis). Hier nun erfährt die Nutzerin von dem Treiben im Hintergrund. Sie wird nämlich aufgefordert, eine Smartcard einzulegen und Ihre PIN anzugeben. Dies ist das zweite Login.
4. Während der VPN-Client die Authentifizierungsinformationen der Nutzerin prüft bzw. mit ihrer Gegenstelle, dem VPN-Gateway, die Bedingungen für die Absicherung der Kommunikation zwischen beiden Rechnern aushandelt, ist der Terminalservice-Client zur Untätigkeit verurteilt. Er muss einfach warten. Dauern ihm die Verhandlungen durch die *Sicherheitsbehörden* zu lange, legt er sich wieder schlafen, nicht ohne die Nutzerin mit einer Fehlermeldung darüber zu informieren. Die Nutzerin muss in diesem Falle den Terminalservice-Client erneut per Doppelklick wecken.
5. Waren die Verhandlungen erfolgreich, entsteht zwischen dem VPN-PC und dem VPN-Gateway eine sichere Verbindung, ein so genannter Tunnel. Durch diesen Tunnel läuft in der Folge der gesamte Verkehr zwischen dem Terminalservice-Client und dem Terminalserver. Dabei werden die Daten signiert (mit einer Art Fingerabdruck des Senders versehen) und verschlüsselt. Durch diesen Tunnel und mit der Legitimation der VPN-Client findet der Terminalservice-Client einen passenden Terminalserver. An dem Server erfolgt ein drittes Login durch die Nutzerin. Durch dieses Login wird sie an der Windows-Domäne des Verwaltungsnetzes angemeldet, sie erhält ein Homeverzeichnis, Netzlaufwerke und Konfigurationseinstellungen zugewiesen.
6. Bisher standen ausschließlich der VPN-Client und der Terminalservice-Client im Lichte der Betrachtung. Was aber ist mit den HIS-Programmen, welche zu Starten der Sinn hinter dem ganzen bisherigen Treiben war? Nun, bei erfolgreichem Login auf dem Terminalserver werden diese dort automatisch gestartet. Für die Nutzerin ergibt sich ein viertes Mal die Gelegenheit, Geduld und Gedächtnis bei einem Login zu beweisen. Die HIS-Anwendungen greifen nämlich auf die Verwaltungsdatenbanken zu. Dabei erfolgt die vierte Anmeldung.

1 Da in den Prüfungsämtern ausschließlich Frauen beschäftigt sind, wird im Artikel die weibliche Form gebraucht.

Einheit versus Vielheit

Während das vielfache Login den einfachen Programmaufruf begleitet, entsteht auf dem Netz zwischen den beteiligten Rechnern eine neue Einheit, ein VPN.

VPN

VPN steht für Virtual Private Network. Ein Netz ist privat, wenn dessen Übertragungswege ausschließlich von einer Gruppe benutzt werden, die Übertragungskanäle und sämtliche Netzwerkknoten dieser Gruppe gehören bzw. von ihr kontrolliert werden und damit ein legitimer Zugang zu den übertragenen Daten nur aus der Gruppe heraus erfolgen kann. In Zeiten zunehmender universeller Vernetzung gibt es so etwas nur in Hochsicherheitsbereichen. In der Regel müssen für die Kommunikation Wege und Vermittlungseinrichtungen benutzt werden, welche auch von – gutwilligen oder böswilligen – Fremden beschritten werden. Erst in der Wechselwirkung mit dem Öffentlichen konstituiert sich überhaupt auch das Private, beide Begriffe sind nur im Bezug aufeinander zu denken. Und so entwickelt sich neben der universellen Vernetzung die Abgrenzung des Privaten durch Firewalls, Datenschutzbestimmungen etc. Wenn sich aber das Private erst mit dem Öffentlichen konstituiert, was bedeutet dann die Rede von einer »virtuellen« Privatheit neben einer »eigentlichen« oder »echten« Privatheit? Nun, das Private wurde gerade über die Zugehörigkeit des Materials zu einer Gruppe beschrieben. Dieser *exklusive Zugang* zum Übertragungsmedium machte die Informationen zu privaten – nicht die *Beschaffenheit* der übertragenen Daten selbst. Eine virtuelle, eine Quasi-Privatheit kommt nun dadurch zu Stande, dass die Daten vor der Übertragung auf besondere Weise behandelt werden, so dass Dritte, wenn ihnen diese Datenpakete in der Öffentlichkeit begegnen, den Verkehr zwar sehen (und auch stören) können, aber mit dem Inhalt der Pakete – den privaten Informationen – nichts anzufangen wissen, da der Inhalt sich ihnen (auf noch geheimnisvolle Weise) verschließt. Damit legt sich über die Vielheit der Subnetze an der Universität eine Einheit, und der exklusive Zugang zu dieser privaten Einheit ist an den Besitz gewisser geheimer Informationen gebunden. Wenn nun ein VPN-Rechner zwecks Arbeit an den Datenbanken dieser Einheit beitritt, muss er den Besitz dieser besonderen Informationen nachgewiesen haben.

IKE

Der Prozess dieses Nachweisens, jenes verzögernde Moment beim Starten des Terminalservice-Client, nennt man IKE (Internet Key Exchange).

IKE ist ein Standardverfahren, nach dem sich zwei Parteien über den Aufbau einer sicheren Verbindung verständigen. Dazu gehören vor allem drei Merkmale:

- die gegenseitige Authentifizierung,
- der Austausch von Parametern, nach denen die Absicherung des weiteren Verkehres erfolgen soll,
- die Berechnung eines gemeinsamen geheimen Schlüssels für die Signatur und die Verschlüsselung von Daten.

Die Authentifizierung soll wechselseitig die Identität der Kommunikationspartner bestätigen. Das ist deshalb wichtig, weil sich im Verkehr zwischen Rechnern recht viel fälschen lässt. So erfolgt der Identitätsnachweis mit Mitteln, die sich sehr schwer (d. h. mit unverhältnismäßig hohem Einsatz an Ressourcen) fälschen lassen. Wir arbeiten beispielsweise mit sogenannten RSA-Signaturen. Jeder Rechner bzw. jede Nutzerin verfügt über ein Schlüsselpaar bestehend aus einem öffentlichen (Public Key) und einem privaten Schlüssel (Private Key). Beide stehen in einem mathematischen Verhältnis zueinander dergestalt, dass der Public Key nach einer bekannten Formel aus dem Private Key abgeleitet wurde, die Berechnung des Private Key aus dem Public Key, d. h. die Umkehrung der Formel, aber eine wesentlich höhere Komplexität besitzt. Die Zuordnung beider Schlüssel zueinander und damit die Sicherheit, d. h. die (Un)Möglichkeit, aus dem öffentlichen Schlüssel den privaten abzuleiten, beruht allein auf der Schwierigkeit, bestimmte mathematische Probleme mit einem Algorithmus zu lösen, dessen Aufwand wesentlich geringer ist als das Durchprobieren aller möglichen Konstellationen. Wobei diese Umkehrung möglich ist. Solche asymmetrischen Schlüssel müssen also entsprechend lang sein, um den Berechnungsaufwand in astronomischen Dimensionen zu halten. Und mit der Steigerung und Verbilligung von Rechenleistung muss auch die Schlüssellänge wachsen. Die beiden Partner unterschreiben (d. h. verschlüsseln) nun jeder ein Stück Information und legen dieses sowie ihren Public Key der Gegenseite vor. Diese prüft mit dem Public Key die signierte Information. Lässt sich eine zugesandte Information mit dem Public Key des Senders verifizieren, handelt es sich bei dem Sender um den Besitzer des zugehörigen Private Key. Zertifikate bestätigen die Zugehörigkeit eines bestimmten Schlüsselpaares zu einer bestimmten Identität, einer Person oder einem Rechner. Sie werden von Zertifizierungsstellen ausgestellt, die sich vor dem Ausstellen des Zertifikates von der Identität des Antragstellers überzeugen müssen. Auf diese Weise lässt sich also der Besitz eines bestimmten Schlüsselpaares einer Identität zuordnen.

Nun nützt ein Schlüssel jedem, der ihn besitzt und die Tür kennt, die er öffnet. Deshalb darf man diesen Schlüssel nicht aus der Hand geben. Was nun unsere PCs anbetrifft, so ist der Schlüssel, der auf dem PC liegt, nur so sicher wie der PC selbst. Die größte Gefahr ist das Kopieren durch unbefugte Dritte außerhalb der normalen Arbeit des Betriebssystems. Darum sollte man den PC vor unbefugtem Zugriff und den Schlüssel mit einem Passwort schützen, besser noch, mit Schlüsseln arbeiten, die nicht auf der Maschine liegen, sondern die Benutzern gehören und sich z. B. auf Smartcards befinden. Dieses höhere Maß an Sicherheit erfordert allerdings zusätzlichen Verwaltungsaufwand. Der Administrator muss erstens einen PC-Benutzer einrichten, er muss diesem Benutzer zweitens eine VPN-Identität verschaffen, er richtet drittens einen Domänenbenutzer ein und erstellt viertens einen Datenbankbenutzer. Die Frage, ob alle diese Benutzeraccounts nicht vielleicht durch ein und dieselbe Identität (oder vielleicht auch zwei) repräsentiert

werden können, soll im nächsten Abschnitt angerissen werden. Halten wir fest, dass die Identität, um die es beim IKE-Vorgang geht, kein bloßer Name, sondern die (beglaubigte) Einheit aus einem eindeutigen Namen und einem kryptographischen Geheimnis ist. Das ganze Sicherheitsverfahren mit IKE als Auftakt heißt IPsec. IPsec verfügt unter anderem über folgende Sicherheitsmerkmale:

- Verschlüsselung der Daten (mit je aktuellen kryptographischen Algorithmen und Schlüssellängen),
- Authentizität des Datenursprunges, d. h. alle Datenpakete tragen die gleiche Signatur, entstammen derselben Quelle und sind während der Übertragung nicht manipuliert worden (Integrität),
- Authentizität des Senders, d. h. Signatur- und Verschlüsselungsschlüssel stammen von einer ausgewiesenen, vertrauten Datenquelle.

IKE (oder um auf den ersten Abschnitt zurückzukommen *das zweite Login*) ist ein kritischer Vorgang, eine Art Nadelöhr beim dezentralen Zugriff auf die Verwaltungsdatenbanken. Nicht nur, weil während dieser Phase der Rechenaufwand auf dem VPN-Gateway besonders hoch ist. Der reibungslose IKE-Verlauf beruht auf der hohen Verfügbarkeit einer ganzen Reihe allgemeiner Dienste des Universitätsnetzes, des Time-Service, des DNS und selbstverständlich der Dienste der Zertifizierungsstelle.

Nun wird man fragen: Wozu dieser ganze Aufwand: separate PCs, auf denen kein normaler Zugang zum Netz möglich ist, Zertifikate und Smartcards, Terminalservice, vierfaches Login ...? Weil es sich darum handelt, vertrauliche Daten über ein öffentliches, eben dezentral verwaltetes Netz zu transportieren. Dabei werden nicht einfach Daten verschlüsselt und authentifiziert von A nach B bewegt. Genauso muss sichergestellt werden, dass die Daten – kaum in B angekommen, wo sie ja dann unverschlüsselt zur Bearbeitung vorliegen – nicht wieder durch eine bösartige Software in die weite Welt verschickt werden. Wer wäre erfreut, seine Personaldaten auf irgendeinem Web-Server wiederzufinden.

Fazit versus Ausblick

Wie wir sehen, zieht sich der Widerspruch bzw. die Einheit von zentral/dezentral, einfach/vielfach, Einheit/Vielheit wie ein roter Faden von der inhaltlichen Bestimmung der Arbeitsprozesse bis zu deren maschineller Begleitung. Am erfolgreichen Start der HISPOS-, HISFSV-Software auf den Terminalservern über eine VPN-Verbindung sind zum Beispiel folgende Servergruppen unmittelbar beteiligt:

- der lokale PC,
- die DNS-Server der HU,
- die Server der HU-CA,
- das VPN-Gateway,
- die Domänencontroller des Verwaltungsnetzes,
- die Fileserver des Verwaltungsnetzes,
- die Terminalserverfarm des Verwaltungsnetzes,
- die Firewall des Verwaltungsnetzes und
- die Datenbankserver des Verwaltungsnetzes.

Des Weiteren müssen folgende Komponenten zumindest hin und wieder erreichbar sein:

- der VPN-Server und
- der Timeservice der HU.

Endlich sollen auch die beteiligten Netzwerkkomponenten wie Router nicht vergessen werden.

Neben der Vielzahl der Komponenten spielt die Konfiguration, die Abstimmung der einzelnen Softwareschichten eine große Rolle. So besitzen z. B. sowohl der VPN-Client als auch der Terminalservice unabhängig voneinander Timeouts bzw. Keep-Alive-Mechanismen, welche den Auf- bzw. Abbau von Verbindungen steuern. Eine optimale Abstimmung der Komponenten aufeinander ist sicher noch nicht gegeben.

Ausblick

Die Anzahl der Logins für die Smartcard-Nutzerinnen ließe sich reduzieren, indem die Rechner komplett in die Verwaltungsdomäne aufgenommen würden. Die Smartcard würde nicht mehr für die IKE-Authentifizierung eingesetzt, sondern zur Anmeldung an der Domäne. Die Zahl der Anmeldevorgänge reduzierte sich auf drei, wobei Rechner-/Domänenlogin bzw. Terminalserverlogin über die Smartcard laufen würden. Es gäbe statt vier Nutzeridentitäten nur noch zwei. Ein Abgleich der Passwörter zwischen Windows-Domäne und Datenbank ist dabei noch nicht vorgesehen. Eine Schranke bildet dafür die HIS-Software selbst, die offenbar eine (nicht dokumentierte) Begrenzung in der Zeichenlänge des Strings [DSN, Servername, Nutzername, Passwort] besitzt. Nicht zuletzt würde sich das Nutzer- und Rechnermanagement spürbar vereinfachen. Die nötigen Voraussetzungen müssen allerdings noch geschaffen werden.

Sollte diese Vereinfachung Wirklichkeit werden, wäre das immerhin auch ein Schritt in Richtung eines einheitlichen Identitätsmanagements an der HU, erfordert aber auch eine hundertprozentige Verfügbarkeit der CA. Schon heute bedeuten ungültige CRLs[2] den Totalausfall des VPN.

Komplexität durch Verquickung vieler Dienste macht Fehlersuche nicht einfacher, lässt sich aber gerade dann nicht vermeiden, wenn Nutzer möglichst bequem und durchweg mit persönlichem Profil unterschiedliche hochverfügbare Netzdienste in Anspruch nehmen wollen. Für den Nutzer soll die Mannigfaltigkeit der Dienste möglichst transparent sein, aus der Perspektive seiner Arbeitsprozesse zu einer neuen Einheit verfließen. Das mag Zukunftsmusik sein; ein VPN, wie wir es in der Universitätsverwaltung für dezentrale Datenbankzugriffe betreiben, ist aber schon heute eine hochgradig personalisierte, komplexe kleine Netzwerkwelt. Für uns Administratoren ist es lebenswichtig, diese Komplexität mit vertretbarem Zeitaufwand zu beherrschen.

2 CRL – von einer CA ausgestellte Rückruflisten für Zertifikate. Die Zertifikate werden damit ungültig.