Das Datenschutzkonzept für das föderierte Frühwarnsystem im D-Grid und seine technische Umsetzung

Nils gentschen Felde[1], Wolfgang Hommel[1], Jan Kohlrausch[2],
Helmut Reiser[1], Christian Szongott[3], Felix von Eye[1]
[1]Leibniz-Rechenzentrum, [2]DFN-CERT, [3]Regionales Rechenzentrum Niedersachsen
felde@nm.ifi.lmu.de, wolfgang.hommel@lrz.de, kohlrausch@dfn-cert.de,
helmut.reiser@lrz.de, szongott@rrzn.uni-hannover.de, felix.voneye@lrz.de

Abstract: Im Projekt GIDS wird ein föderiertes Intrusion Detection System für das D-Grid konzipiert, implementiert, evaluiert und produktiv geführt. Dabei müssen zur D-Grid-weiten Erkennung von Angriffen auf Grid-Ressourcen lokale Alarmmeldungen der einzelnen Ressourcenanbieter organisationsübergreifend ausgetauscht, korreliert und u. a. zur Anzeige in einem zentralen Benutzerportal aufbereitet werden. Die dafür technisch vorhandenen Möglichkeiten werden praktisch sowohl durch restriktive *Information-Sharing-Policies* der beteiligten Organisationen als auch durch rechtliche und datenschutzrechtliche Randbedingungen eingeschränkt. In diesem Artikel werden das GIDS-Datenschutzkonzept und seine technische Umsetzung vorgestellt.

1 Ausgangssitutation und Herausforderungen

In einem organisationsübergreifenden Verbund muss die Informationssicherheit nicht nur jeweils organisationsintern, sondern auch für den Verbund als Ganzes betrachtet werden. Zu diesem Zweck müssen präventive, detektierende und reaktive Maßnahmen konzipiert und implementiert werden, die explizite Schnittstellen zwischen den beteiligten Organisationen vorsehen. In [gJMT06, HKK+11] wurde gezeigt, dass durch den Einsatz föderierter bzw. kooperativer Sicherheitsmechanismen ein verbessertes Sicherheitsniveau großer organisationsübergreifender IT-Infrastrukturen erzielt werden kann. Zentrale Herausforderungen für solche Sicherheitssysteme sind die potentiellen Kollisionen mit Datenschutzvorschriften und zum Teil sehr restriktiven organisations- und fachspezifischen Leitlinien zur Heraus- bzw. Weitergabe von IT-sicherheitsspezifischen Informationen an Dritte.

Die Idee eines föderierten, D-Grid-weiten Angriffserkennungssystems, die im Projekt GIDS (ein Grid-basiertes, föderiertes Intrusion Detection System zur Sicherung der D-Grid Infrastruktur) umgesetzt wird, ist in [Hgv+10c] beschrieben. Die Soll-Erkennungsleistung und die daraus resultierenden technischen Anforderungen wurden in [Rgv+10] analysiert. Auf dieser Basis wurde eine verteilte Architektur erarbeitet, die eine lose Kopplung der am D-Grid bzw. an GIDS beteiligten Ressourcenanbieter vorsieht, um die im Grid-Umfeld gewünschte organisatorische, administrative und technische Unabhängigkeit und Autonomie zu berücksichtigen.

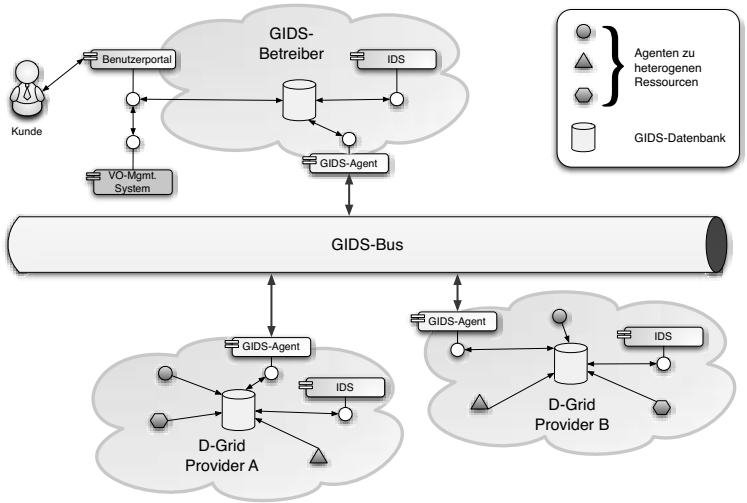

Abbildung 1: Architekturidee von GIDS nach [Hgv+10c] und [Hgv+10a]

Abbildung 1 gibt einen Überblick über den Architekturentwurf von GIDS. Die (lokalen) IDS-Instanzen basieren auf *Prelude* (`http://git.lrz.de/prelude/`), das nativ zu einer Vielzahl anderer IDS kompatibel ist und zudem die ausgereifte Bibliothek *libprelude* bereitstellt. Um die Idee der Föderation bestehender sicherheitsrelevanter Komponenten zu realisieren, bedarf es einer Kommunikationsinfrastruktur, die durch einen Nachrichten-Bus (*GIDS-Bus*) realisiert wird. Er basiert auf einer organisationsübergreifenden und Multicast-fähigen VPN-Infrastruktur. Als VPN-Lösung kommt die freie Software *OpenVPN* (`http://www.openvpn.net`) zum Einsatz und stellt einen SSL-geschützten Tunnel zwischen allen Ressourcenanbietern bereit. Über den GIDS-Bus können alle an GIDS beteiligten Ressourcenanbieter unter Wahrung der Integrität, Authentizität und Vertraulichkeit der übermittelten Informationen Daten über erkannte Angriffe austauschen. Es ist jedoch zu beachten, dass jede über den Bus gesendete Nachricht von allen an den Bus angeschlossenen Parteien gelesen werden kann.

Da solche Nachrichten personenbezogene Daten wie beispielsweise Benutzerkonten oder IP-Adressen enthalten, sind die Vorgaben des Bundesdatenschutzgesetzes (BDSG) zu berücksichtigen, das die Weitergabe stark einschränkt. Allerdings sind diese Daten oftmals für eine Angriffserkennung und die Einleitung von Gegenmaßnahmen notwendig. Um im GIDS einen effektiven Betrieb realisieren zu können, wurde ein Datenschutzkonzept entwickelt, das technische und juristische Mittel anwendet, um eine föderierte Erkennung von Angriffen zu ermöglichen ohne die rechtlichen Rahmenbedingungen zu verletzen.

Um auf Seiten der Ressourcenanbieter erhobene, für GIDS relevante Daten (dies können z. B. Log-Daten von Firewalls, Berichte von lokalen IDS-Instanzen o. ä. sein) GIDS-weit austauschen und automatisiert verarbeiten zu können, bedarf es gemeinsamer Datenformate und Protokolle. Die Übersetzung der vorliegenden Daten und die Verbreitung der

Informationen über den GIDS-Bus nehmen Agenten vor, die von den jeweiligen Ressourcenanbietern selbst betrieben und konfiguriert werden, so dass die volle Kontrolle und Unabhängigkeit der Ressourcenanbieter gewährleistet bleibt.

Als Grundlage für das in GIDS verwendete Datenaustauschformat dient das XML-basierte *Intrusion Detection Message Exchange Format* (IDMEF [DCF07]). Neben `Heartbeat`-Nachrichten sind `Alert`-Nachrichten der entscheidende Teil zur Übermittlung von Alarminformationen. Hier können neben Informationen über die Quelle und das Ziel eines Angriffs vielfältige weitere Daten aufgenommen werden. Insbesondere sind eine Klassifizierung und eigene Erweiterungen im IDMEF-Standard vorgesehen.

Dieser Artikel beschreibt ausgewählte Aspekte des Datenschutzes beim Einsatz eines föderierten Frühwarnsystems im D-Grid und die praktische Umsetzung. Abschnitt 2 skizziert das GIDS zugrundeliegende Datenschutzkonzept. Abschnitt 3 widmet sich der technischen Umsetzung des Datenschutzkonzepts. Abschnitt 4 fasst die Ergebnisse zusammen und gibt einen Ausblick auf anstehende Weiterentwicklungen.

2 Ein Datenschutzkonzept zur Verwendung innerhalb von GIDS

Im Rahmen von GIDS werden Daten über Angriffe auf Grid-Systeme erhoben und verarbeitet. Die Daten werden von verschiedenen Sensoren aufgezeichnet und betreffen detaillierte Informationen über Angriffe, die z. B. Quelle und Ziel des Angriffs beinhalten. Da ein Teil dieser Daten den Bezug zu Personen zulässt – dies können z. B. kompromittierte Benutzerkonten sein – fallen diese mindestens unter den Schutz des Bundesdatenschutzgesetzes (BDSG), das die Erhebung, Verwendung und Weitergabe dieser Daten stark einschränkt. Des Weiteren können diese Daten kritische Sicherheitslücken auf der Seite der beteiligten Partner offenlegen. Deshalb sind Sicherheitsanforderungen der beteiligten Partner zu berücksichtigen, die die Verarbeitung, Verbreitung und den Zugriff auf diese Daten betreffen.

Das GIDS-Datenschutzkonzept muss diese Anforderungen in sich vereinen: Neben allgemein gültigen rechtlichen Regelungen des BDSG müssen die technischen Spezifika und die Anforderungen der beteiligten Organisationen berücksichtigt werden. Wir fassen im Folgenden kurz die Grundlagen zusammen, die zu einem Datenschutzkonzept [Hgv+10b] geführt haben, und gehen anschließend auf praktische Aspekte seiner Umsetzung ein. Zwar wird die Erhebung, Verwendung und Weitergabe personenbezogener Daten durch das BDSG stark eingeschränkt, jedoch sieht der Gesetzgeber auch Mittel vor, die eine Weitergabe unter bestimmten Bedingungen ermöglichen. So erlaubt beispielsweise das Telekommunikationsgesetz die Nutzung personenbezogener Daten zur Beseitigung von Störungen in IT-Anlagen.

Grundaspekte des Datenschutzkonzepts Für den nachhaltigen Betrieb von GIDS sind verschiedene Aspekte zu berücksichtigen, die den Schutz und die Verwendung der Angriffs- und Betriebsdaten betreffen:

Technische Anforderungen. Im Datenschutzkonzept muss berücksichtigt werden, dass der Zweck erhobener und verarbeiteter Daten die Absicherung von Grid-Systemen ist. Angriffe müssen sowohl lokal als auch im Grid als Ganzes effektiv erkannt werden. Zudem muss rasch und möglichst weitgehend automatisiert reagiert werden können, um auch die von einem Angriff noch nicht betroffenen Systeme besser schützen zu können.

Sicherheitsaspekte bei Erhebung, Transport und Speicherung der Daten. Da im Rahmen von GIDS sicherheitskritische Informationen verarbeitet werden, müssen die grundlegenden Sicherheitsaspekte Vertraulichkeit, Integrität, Authentizität und Verfügbarkeit gewährleistet sein. Dies betrifft insbesondere Daten, die auf der Seite der Betreiber von Grid-Ressourcen erhoben werden und deren Kenntnis sich möglicherweise für Angriffe ausnutzen lassen würden. Weiterhin kann potentieller Missbrauch durch Anwendung des „*least privilege*"-Prinzips minimiert werden.

Sicherheitsanforderungen der beteiligten Organisationen. Neben Angaben zu Angriffen sind bei verteilten IDS häufig weitere Informationen über Interna der beteiligten Organisationen zu verarbeiten. Dies umfasst einerseits Informationen der Betreiber über die Position ihrer Sensoren und die interne Struktur ihrer Netze. Andererseits liefern die Sensordaten indirekt auch Informationen über die Verwundbarkeit und den Zustand der überwachten Netze. Damit diese Daten nicht an Unbefugte weitergegeben werden, ist eine genaue Festlegung der Verwendungszwecke der Daten notwendig, der alle beteiligten Organisationen zustimmen müssen.

Rechtliche Regelungen. Der Betrieb eines verteilten, in Deutschland eingesetzten IDS unterliegt rechtlichen Regelungen, beispielsweise dem Datenschutzgesetz und dem Telekommunikationsgesetz; die entsprechenden Vorgaben bezüglich der Erhebung, Verarbeitung und Speicherung personenbezogener Daten müssen zwingend eingehalten werden.

2.1 Verwandte Arbeiten

Die für das GIDS-Datenschutzkonzept relevanten Vorarbeiten lassen sich in mehrere Bereiche aufteilen. Zuerst gibt es verwandte Arbeiten im Grid-Umfeld, die analog zu GIDS die rechtlichen Rahmenbedingungen im produktiven Betrieb berücksichtigen. So wurde vom D-Grid-Projekt „PneumoGrid" ein eigenes Datenschutzkonzept erstellt und veröffentlicht [KC09]. Zwar fokussiert dieses Konzept auf die datenschutz- und arzneimittelrechtlichen Vorgaben, die durch Pseudonymisierung personenbezogener Daten erfüllt werden, jedoch wird auch auf den Schutz der Daten im Benutzerportal eingegangen.

Des Weiteren sind viele Ansätze vorgeschlagen worden, um IDS-Daten rechtlich konform zu erheben und zu verarbeiten, wie beispielsweise in [BF00]. Der Schwerpunkt dieser Arbeiten liegt aber auf einem Konzept für die Pseudonymisierung der IDS-Daten.

Einen Schritt weiter gehen Flegel et al. in [FHM10] mit dem Ziel, eine Kooperation in einem zentralen Frühwarnsystem zu etablieren. Hier werden neben den gesetzlichen Bestim-

mungen zusätzlich die Anforderungen der beteiligten Partner an die Vertraulichkeit der übertragenden Daten berücksichtigt. Jedoch bildet auch dieser Ansatz die Anforderungen an das Datenschutzkonzept des GIDS nur teilweise ab. So müssen im GIDS die Anforderungen der föderierten, autonom agierenden Partner berücksichtigt werden. Dies beinhaltet zum Beispiel, dass die Partner entscheiden können, welche Daten weitergegeben werden können und welche herausgefiltert werden. Im Gegensatz zu [FHM10] basiert GIDS auf einer dezentralen Struktur und sieht auch die Verarbeitung von Daten vor, die nicht direkt im Zusammenhang mit Angriffen stehen, sondern Betriebsdaten der Grids sind. Ein weiterer Unterschied ist die Datenquelle, die beim GIDS beliebige IDS-Sensoren sind. Neben der Anwendung des Datenschutzkonzepts auf IDS-Daten sind also zusätzliche Anforderungen zu berücksichtigen, die die Erstellung eines eigenen GIDS-Datenschutzkonzepts notwendig machen.

2.2 Ansätze zur rechtlich-konformen Umsetzung des Datenschutzkonzepts

Das in [Hgv+10b] vorgelegte GIDS-Datenschutzkonzept dient zunächst als rechtlich fundierte Grundlage für die Verarbeitung von Angriffsdaten in GIDS. Seine praktische Umsetzung, also die technische Implementierung der organisatorischen und rechtlichen Vorgaben, ist ein wichtiger Schritt, der von allen an GIDS teilnehmenden Organisationen durchlaufen werden muss.

Der Gesetzgeber bietet die Möglichkeit durch das Telekommunikationsgesetz (TKG), Verkehrsdaten im Allgemeinen oder IDS-Angriffsdaten im Speziellen für die Erkennung und Beseitigung von Störungen zu erheben. Allerdings ist die Erhebung an diesen Anlass gebunden und es gelten strikte Einschränkungen für die Verwendung und Weitergabe. Grundlegend ist die Verwendung auf die Beseitigung der Störung beschränkt (Zweckbindung) und die Weitergabe an Dritte untersagt. Die Weitergabe der Daten im Rahmen des GIDS kann aber durch eine geeignete vertragliche Bindung zwischen den Nutzern und Betreibern des GIDS ermöglicht werden.

Der gesetzliche Schutz bezieht sich auf personenbezogene Daten. Fehlt dieser Bezug, entfallen Einschränkungen bezüglich der Erhebung und Weitergabe. Deshalb ist es eine wirkungsvolle technische Maßnahme, personenbezogene Daten vor der Weitergabe zu anonymisieren oder zu pseudonymisieren. Ein Vorteil der Pseudonymisierung ist, dass der Bezug zu den ursprünglich erhobenen Daten rekonstruiert werden kann und die Korrelation von Daten ohne Einschränkungen möglich ist. Da allerdings eine Zuordnung zu einer Person zumindest noch in der Theorie existiert, ist nicht unumstritten, in welchem Umfang die Einschränkungen für personenbezogene Daten entfallen. Im Zweifelsfall ist eine Anonymisierung der Daten zu bevorzugen, durch die der Bezug vollständig vermieden wird.

Als Beispiel eines Datums mit Personenbezug gilt bereits eine IPv4-Adresse. Hier reicht es im Regelfall beispielsweise, das letzte Byte einer IPv4-Adresse zu löschen, um einen Personenbezug im juristischen Sinne aufzuheben. Im Gegenzug müssen aufgrund des damit verbundenen Informationsverlusts möglicherweise Einschränkungen bei der Erkennungs- und Korrelationsleistung in Kauf genommen werden.

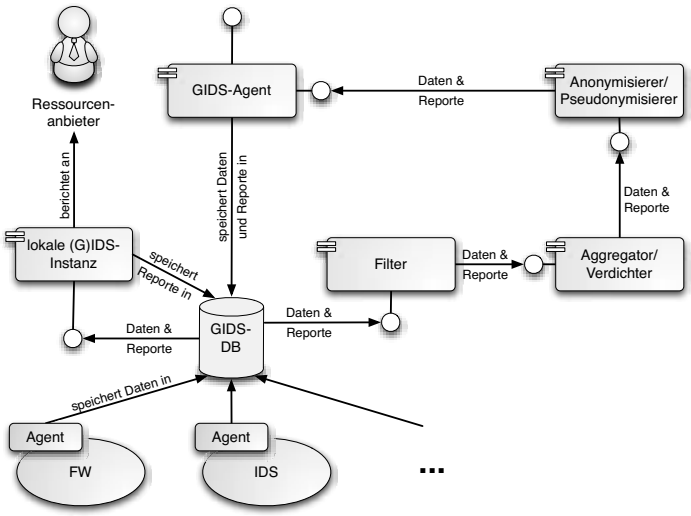

Abbildung 2: GIDS-Architektur auf Seite einer teilnehmenden Organisation nach [Hgv+10a]

3 Technische Umsetzung des Datenschutzkonzepts

Um das Datenschutzkonzept vor dem Übergang der Informationen auf den GIDS-Bus technisch realisieren zu können, wurden in einem ersten Schritt alle Attribute und XML-Knoten von IDMEF auf ihre datenschutzrechtliche Bedeutung geprüft. Da die Semantik der Attribute und Knoten durch den Standard [DCF07] festgelegt ist, ist eine für die Ressourcenanbieter zuverlässige Zuordnung der datenschutzrechtlich und sicherheitspolitisch relevanten Teile einer IDMEF-Meldungen möglich.

Abbildung 2 stellt den dreistufigen Vorverarbeitungsprozess dar, den eine IDMEF-Nachricht durchlaufen muss, bevor sie über den GIDS-Bus verteilt wird. Fällt an einer Informationsquelle (hier beispielhaft eine Firewall oder eine lokale IDS-Instanz) eine Nachricht an, so wird diese von einem für das jeweilige System entwickelten Agenten ins IDMEF-Format übersetzt und in eine lokale Datenbank geschrieben. Alle neuen Nachrichten in dieser Datenbank durchlaufen zur Veröffentlichung anschließend zuerst eine Filterung, die die Umsetzung lokaler Datenschutzrichtlinien unterstützt. Danach findet eine Korrelation/Aggregation der Daten statt, die die Skalierbarkeit des Systems sicherstellt, bevor vor der Weitergabe der IDMEF-Nachrichten an Dritte schlussendlich noch eine Anonymisierung oder Pseudonymisierung personenbezogener Daten vorgenommen wird, um den rechtlichen Anforderungen genüge zu tun.

Die Reihenfolge der Schritte ist so gewählt, damit Datensätze, die die Domänengrenzen nicht überschreiten dürfen, als erstes aussortiert werden, danach eine Aggregation auf den komplett vorhandenen Datensätzen durchgeführt werden kann und erst im letzten Schritt eine Anonymisierung der (aggregierten) Datensätze stattfindet. Damit erreicht man eine

maximale Erkennungsleistung bei gleichzeitiger Umsetzung des Datenschutzkonzeptes. Die Umsetzung der drei letztgenannten Komponenten und ihr Bezug zum gesamten Datenschutzkonzept ist im Folgenden beschrieben.

Nachrichtenfilterung Zunächst werden Meldungen, die nicht für die Weitergabe an die anderen Teilnehmer des GIDS bestimmt sind, ausgefiltert. Dies hat den praktischen Grund, dass Sicherheitsmechanismen, die bei den beteiligten Ressourcenanbietern installiert sind, in der Regel mehr als nur die am D-Grid angeschlossenen Ressourcen zu überwachen haben. Somit kommt es laufend vor, dass Alarmmeldungen nur organisationsintern verarbeitet werden, aber nicht an die anderen GIDS-Teilnehmer versendet werden sollen. Das GIDS-System bietet die Möglichkeit, solche Nachrichten komplett auszufiltern. Somit dient diese Maßnahme nicht primär den rechtlichen Anforderungen, sondern in erster Linie der Erfüllung lokaler Sicherheitsrichtlinien. Dabei ist es möglich, die Nachrichten, die ausgefiltert werden sollen, über reguläre Ausdrücke zu identifizieren. Ein praktischer Nebeneffekt der Filterung ist das frühzeitige Entfernen ungewünschter bzw. unnötiger Daten, was direkt zur Ressourcenschonung und somit Skalierbarkeit des Gesamtsystems positiv beiträgt.

Datenverdichtung durch Korrelation und Aggregation Bei der Datenverdichtung geht es in erster Linie darum, das über den GIDS-Bus zu übermittelnde Datenvolumen zur Sicherstellung der Skalierbarkeit zu minimieren, indem mehrere Meldungen Site-lokal zu einer Meldung zusammengefasst werden, so dass das Datenschutzkonzept hier nur am Rande von Bedeutung ist. Eine nähere Betrachtung im Rahmen dieser Arbeit bleibt außen vor.

Anonymisierung und/oder Pseudonymisierung Im letzten Schritt kann jede Nachricht anonymisiert oder pseudonymisiert werden. In diesem Fall kann man einzelne Attribute oder XML-Knoten, die datenschutzrechtlich relevant sind, auswählen, damit sie anonymisiert oder pseudonymisiert werden. Dies ist der zentrale Schritt, in dem die Einhaltung der rechtlichen Randbedingungen, speziell der Datenschutzgesetze, realisiert wird.

Im praktischen Einsatz ergeben sich jedoch eine Reihe von Herausforderungen. Zum einen ist der Umfang einer IDMEF-Nachricht im Wesentlichen eine Kompromisslösung, um verschiedenartige Alarmmeldungen u. a. von Logfile-Analyzern, Firewalls oder Intrusion Detection Systemen abzubilden. Somit werden viele, für eine tiefere Analyse wichtige Daten, in das Feld `AdditionalData` ausgelagert, dessen genaue Verwendung im IDMEF-RFC bewusst nicht spezifiziert wurde, um die Möglichkeiten zur Erweiterung des XML-Schemas nicht einzuschränken. Da jedoch die Semantik der darin enthaltenen Daten nicht a priori festgelegt ist, sind diese Felder im Sinne einer strengen Auslegung des Datenschutzgesetzes vollständig zu löschen. Ansonsten kann beispielsweise nicht ausgeschlossen werden, dass personenbezogene Daten (ggf. in binärer Form) enthalten sind. An dieser Stelle wäre eine standardisierte Erweiterung von IDMEF hilfreich, die mehr Felder bereit stellen könnte, deren Semantik a priori bekannt ist.

3.1 Erweiterungen des IDMEF-Formates mit datenschutzrechtlichem Bezug

In Abbildung 1 ist als Schnittstelle zwischen den Kunden und dem GIDS ein Benutzerportal dargestellt. Dieses dient der Information bzw. Alarmierung von Grid-Benutzern und Site-Administratoren. Benutzer müssen sich authentifizieren und werden Rollen zugeordnet, um nur die für sie relevanten Meldungen einsehen zu können.

Bei der Löschung von z. B. IP-Adressen aus Alarmmeldungen oder einer zu starken Anonymisierung wird jedoch auch die eindeutige Zuordnung eines Alarms zu einer Ressource erschwert oder gänzlich unmöglich gemacht. Dabei ergeben sich Probleme, um beispielsweise die Zugriffsregelungen im Portal eindeutig durchsetzen zu können.

Dieses Problem lässt sich jedoch mit Mitteln lösen, die im D-Grid bereits vorhandenen sind. Im *Grid Ressource Registry Service* (GRRS) existiert für jede Ressource eine Zuordnung zu einem Ressourcenanbieter, dem eine *Short-ID* zugeordnet ist, also ein Kurztitel. Jeder IDMEF-Nachricht wird daher ein zusätzlicher XML-Knoten vom Typ `AdditionalData` hinzugefügt, in dem die Short-ID des GIDS-Teilnehmers, der die Nachricht erstellt hat, vermerkt ist. Damit ist auch nach einer vollständigen Anonymisierung der Nachrichten eine eindeutige Zuordnung zu den Ressourcen und den dazugehörigen VO-Nutzern möglich. Dies eröffnet zwar potentiell neue Datenschutzrisiken, vor allem wenn es sich bei den Short-IDs um Grid-Ressourcen bestehend aus einer einzelnen Maschine handelt, auf der immer nur ein Nutzer Grid-Jobs rechnet. Es unterliegt jedoch den Ressourcenanbietern, Alarmmeldungen, die einen solchen eindeutigen Nutzerbezug haben, mit den oben genannten Anonymisierungs- oder Filterfunktionen datenschutzkonform vorzubereiten.

Ein ähnlicher Weg wird für die Durchsetzung lokaler Datenschutzrichtlinien bei von anderen Domänen empfangenen Alarmmeldungen gegangen. Da es unwahrscheinlich ist, dass bei allen Ressourcenanbietern die anfallenden Daten nach der gleichen Anzahl von Tagen gelöscht werden, ist es nötig, ein System zu entwickeln, das garantiert, dass die Daten, die die Domänengrenzen überschreiten, trotzdem nach den Richtlinien der ursprünglichen Domäne behandelt werden. Auch hier wird wieder ein weiterer XML-Knoten vom Typ `AdditionalData` hinzugefügt, dessen Inhalt ein Ablaufdatum für die Nachricht ist. Sobald das hier definierte Datum erreicht ist, muss die Nachricht auch bei den anderen GIDS-Teilnehmern komplett gelöscht werden. Für die Löschung „abgelaufener“ Daten sind regelmäßig auszuführende Automatismen verantwortlich, zu denen sich alle an GIDS beteiligten Ressourcenanbieter a priori verpflichten.

3.2 Erfahrungen im Praxiseinsatz

Im Projekt werden derzeit verschiedene Anonymisierungs- und Pseudonymisierungsstufen evaluiert. In einem ersten Schritt wurde die strengste Auslegung des Datenschutzes erprobt. Diese führt dazu, dass alle Attribute bzw. XML-Knoten, die personenbezogene Daten enthalten können, komplett aus der Nachricht gelöscht werden. Damit wird ein Personenbezug vollständig ausgeschlossen. Dies geht jedoch zu Lasten der theoretisch

möglichen und praktisch erzielten Erkennungsleistung: In diesem Fall ist keine ordentliche Korrelation von Meldungen über Domänengrenzen hinweg mehr möglich. Daher wird schrittweise empirisch ermittelt, welche und wie viele Teilinformationen beispielsweise über die Ziel- und Quelladressen eines Angriffs nötig sind, um angriffsspezifische Korrelationen durchzuführen, die die Grid-globale Erkennungsleistung verbessern.

Ein parallel verfolgter Ansatz nutzt Pseudonymisierungswerkzeuge wie z. B. Crypto-PAn. Dabei werden „kritische“ Teile einer Nachricht kryptographisch verschlüsselt und somit für Außenstehende nicht mehr erkennbar gemacht. Bisher wurde in der D-Grid-Praxis von dieser Möglichkeit noch kein Gebrauch gemacht. Dies hat zwei Gründe: Zum einen ist es juristisch noch umstritten, ob eine Pseudonymisierung den Auflagen der Datenschutzgesetze genügt, da ein Personenbezug, wenn auch mit Aufwand, immer noch herstellbar ist. Zum anderen ist für ein solches Verfahren ein wirkungsvolles Krypto-Schlüsselmanagement zu integrieren, da der verwendete Schlüssel regelmäßig erneuert werden muss, um die Vertraulichkeit der verschlüsselten Daten langfristig gewährleisten zu können.

4 Zusammenfassung und Ausblick

Föderierte Frühwarnsysteme, wie sie im Rahmen von GIDS für das D-Grid aufgebaut werden, unterliegen Einschränkungen, die organisationsinterne Intrusion Detection Systeme nicht haben: Zum einen ist die uneingeschränkte Weitergabe von Informationen über Angriffe an externe Dritte oftmals unerwünscht und zum anderen müssen gesetzliche Auflagen, insbesondere mit Bezug auf den Datenschutz, umgesetzt werden.

In diesem Artikel wurden das GIDS-Datenschutzkonzept und seine technische Umsetzung vorgestellt: Von einer Organisation intern erzeugte Alarmmeldungen werden zunächst gefiltert, verdichtet und anonymisiert bzw. pseudonymisiert, bevor sie über eine VPN-geschützte Multicast-Bus-Infrastruktur an die anderen beteiligten Organisationen weitergegeben und zur Anzeige im GIDS-Webportal aufbereitet werden. Als Transportformat wird das weit verbreitete IDMEF eingesetzt, für das präzise definiert wurde, welche Attribute und XML-Knoten potenziell kritische bzw. personenbezogene Daten enthalten können. Eine aktuell laufende Evaluation analysiert die Korrelations- und Erkennungsleistung des Grid-weiten Frühwarnsystems in Abhängigkeit vom Informationsverlust, der durch Anonymisierung eintritt.

Neben der Produktivführung des erarbeiteten Systems, die sich an die Evaluation anschließen wird, ist geplant, das Konzept der so genannten *sticky policies* in GIDS zu integrieren. Momentan werden Angriffsmeldungen, die an andere an GIDS beteiligte Sites weitergegeben werden, nur mit einem Soll-Ablaufdatum versehen. Die Empfänger dieser Meldungen verpflichten sich organisatorisch bzw. vertraglich dazu, diese Ablaufdaten zu honorieren und die entsprechenden Meldungen rechtzeitig wieder bei sich zu löschen. *Sticky policies* verknüpfen die ausgetauschten Daten hingegen untrennbar mit fein granulierten Regelungen, die neben der maximalen Aufbewahrungsdauer beispielsweise auch eine genaue Zweckbindung und weitere Auflagen umfassen können. Sie sind eine wichtige Voraussetzung für die Akzeptanz des GIDS aufgrund der sehr heterogenen Information-Sharing-

Policies der D-Grid-Ressourcenanbieter.

In GIDS wurden auch einige Erweiterungen des IDMEF-Formats vorgenommen. Diese nutzen zwar die bereits vorhandenen Schnittstellen, stellen zunächst jedoch eine D-Grid-proprietäre Entwicklung dar. Um die Übertragbarkeit der Ergebnisse auf andere Grids und unternehmensübergreifende Verbünde sicherzustellen, soll mit den IDMEF-RFC-Autoren diskutiert werden, wie eine Standardisierung erreicht werden kann.

Literatur

[BF00] Joachim Biskup und Ulrich Flegel. On Pseudonymization of Audit Data for Intrusion Detection. In Hannes Federrath, Herausgeber, *Designing Privacy Enhancing Technologies*, number 2009 in Lecture Notes in Computer Science, Berkeley, California, USA, Juli 2000. Springer, Heidelberg.

[DCF07] Herve Debar, David A. Curry und Benjamin S. Feinstein. The Intrusion Detection Message Exchange Format (IDMEF), 2007. IETF RFC 4765.

[FHM10] Ulrich Flegel, Johannes Hoffmann und Michael Meier. Cooperation enablement for centralistic early warning systems. In Sung Y. Shin und Sascha Ossowski, Herausgeber, *Proceedings of the 25th International ACM Symposium on Applied Computing (SAC 2010)*, Sierre, Schweiz, März 2010. ACM Press.

[gJMT06] Nils gentschen Felde, Marko Jahnke, Peter Martini und Jens Tölle. Impact of Sanitized Message Flows in a Cooperative Intrusion Warning System. In *Proceedings of the 25th Military Communications Conference (MILCOM 2006)*, Washington, DC, USA, Oktober 2006.

[Hgv+10a] Wolfgang Hommel, Nils gentschen Felde, Felix von Eye, Jan Kohlrausch und Christian Szongott. Architekturkonzept für ein Grid-basiertes IDS. Technical report, D-Grid, Oktober 2010.

[Hgv+10b] Wolfgang Hommel, Nils gentschen Felde, Felix von Eye, Jan Kohlrausch und Christian Szongott. Datenschutzmodell für ein Grid-basiertes IDS. Technical report, D-Grid, Juli 2010.

[Hgv+10c] Wolfgang Hommel, Nils gentschen Felde, Felix von Eye, Jan Kohlrausch und Christian Szongott. Grobskizze einer Architektur. Technical report, D-Grid, April 2010.

[HKK+11] Wolfgang Hommel, Jan Kohlrausch, Jan Köcher, Christian Szongott, Nils gentschen Felde und Felix von Eye. Ein föderiertes Intrusion Detection System für das D–Grid. In *18. DFN Workshop „Sicherheit in vernetzten Systemen“*, Hamburg, Deutschland, Februar 2011. DFN-CERT GmbH.

[KC09] Dagmar Krefting und Sebastian Canisius. Anforderungen in PneumoGrid an Sicherheit und Nutzerfreundlichkeit. Grid Security Workshop in Göttingen, Oktober 2009.

[Rgv+10] Helmut Reiser, Nils gentschen Felde, Felix von Eye, Jan Kohlrausch und Christian Szongott. Anforderungs- und Kritierienkatalog (MS 6). Technical report, D-Grid, Januar 2010.