ETSI GS INS 009 V1.1.1 (2012-09)

Group Specification

Identity and access management for Networks and Services (INS); Security and privacy requirements for collaborative cross domain network monitoring

Disclaimer

This document has been produced and approved by the Identity and access management for Networks and Services ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference

DGS/INS-009

Keywords

access control, data sharing, multi-party computation, network monitoring, policies, policy management, privacy, security

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Copyright Notification

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://ipr.etsi.org).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Group Specification (GS) has been produced by ETSI Industry Specification (ISG) Identity and access management for Networks and Services (INS).

Introduction

The threat landscape of today's Internet is characterized by highly distributed attacks (e.g. botnets) which leverage and target multiple domains. Attackers do not care about boundaries between networks and jurisdictions; malware and cyber-attacks are criminal activities focused in having the best tools for the intended purpose (e.g. information theft, extortion), while remaining difficult to detect and defend against. Spreading malware as widely as possible across multiple networks is an effective tactic to this end, in large part because current detection and mitigation measures are taken from a the point of view of a single administrative domain; each operator or enterprise fights the threats locally, with limited or non-existent collaboration with other peers.

Collaborative cross-domain network monitoring and mitigation seems to be a natural approach to fight these threats more efficiently. This approach includes technical solutions delivering distributed processing and computation, protocols, and data exchange allowing network operators to cooperate in network security monitoring efforts in a dynamic and efficient manner. This collaboration allows the scaling the defensive measures to the same level as those applied by the attackers. A keystone in this cross-domain collaboration is the communication and sharing of monitoring information among network operators, but this is limited by many factors:

- Network operators are not eager to share information about their operations, especially sensitive data such as event logs from intrusion detection systems for a variety of reasons, including privacy, regulatory compliance, and protection of information of commercial value.
- Current collaborative procedures are slow human-driven communications, as there are limited solutions available in the market for collaborative network defence.
- Research in the field is ongoing; there is little work to date on providing a holistic view of the problem and its potential solution (the state of the art in applicable research is explored in clause 7 of the present document).

The purpose of the present document is:

1) to assess the context of any potential data exchange among different network operators, from different points of view (technical, regulatory, business) and extract security and privacy requirements to govern those exchanges;

2) to identify existing technologies, protocols and specifications helping to fulfil the defined requirement; and

3) to identify the gaps and to propose new technologies and protocols to fill them.

1 Scope

The present document places some general requirements applying on any security monitoring data exchange aiming at cross-domain detection and mitigation.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the reference document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are necessary for the application of the present document.

Not applicable.

2.2 Informative references

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] SEPIA library.

NOTE: Available at http://sepia.ee.ethz.ch/.

[i.2] CoRR abs/1101.5509 (2011): "Reduce to the max: A simple approach for massive-scale privacy-preserving collaborative network measurements (extended version)", F. Ricciato and M. Burkhart.

[i.3] IETF RFC 5070 (December 2007): "The Incident Object Description Exchange Format", R. Danyliw, J. Meijer, and Y. Demchenko.

[i.4] IETF RFC 6545 (April 2012): "Real-time Inter-network Defense (RID)", K. Moriarty.

[i.5] IETF RFC 6546 (April 2012): "Transport of Real-time Inter-network Defense (RID) Messages over HTTP/TLS", B. Trammell.

[i.6] draft-ietf-mile-sci-02.txt: "IODEF-extension to support structured cybersecurity information", February 2012.

[i.7] IETF RFC 6235 (May 2011): "IP Flow Anonymization Support", E. Boschi, B. Trammell.

[i.8] IETF RFC 5101 (January 2008): "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information", B. Claise (ed.).

[i.9] IETF RFC 5655 (October 2009): "Specification of the IP Flow Information Export (IPFIX) File Format", K B. Trammell, E. Boschi, L. Mark, T. Zseby and A. Wagner.

[i.10] In 19th USENIX Security Symposium (August 2010): "SEPIA: Privacy-Preserving Aggregation of Multi Domain Network Events and Statistics", M. Burkhart, M. Strasser, D. Many and X. Dimitropoulous.

[i.11] Computer Law & Security Report 24(6), 508-520 (2008): "The EU Data Protection Directive: An engine of a global regime", Birnhack, M..

[i.12] Directive 2006/24/EC, European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC, Official Journal of the European Communities, No. L 105, April 2006, pp. 54-63. 2006/24/EC.

[i.13] Directive 2002/58/EC, European Parliament and of the Council concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), Official Journal of the European Communities, No. L 201, pp. 37-47, July 2002.

[i.14] Commission Decision 2000/520/EC of 26 July 2000 pursuant to Directive 95/46/EC of the European Parliament and of the Council (Safe harbor principle).

[i.15] "Access control: Policies, models, and mechanisms": in FOSAD 2000: Foundations of Security Analysis and Design, Vol. 2171 of Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2001, pp. 137-196, P. Samarati, S. D. C. di Vimercati.

[i.16] IEEE Computer 29, 2 (February 1996), pp.38-47: "Role-based access control models", R.S. Sandhu, E.J. Coyne, H.L. Feinstein, C.E. Youman.

[i.17] OASIS: "eXtensible Access Control Markup Language (XACML) 2.0", February 2005.

NOTE: Available at http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml.

[i.18] "Leveraging Access Control for Privacy Protection: A Survey": in Privacy Protection Measures and Technologies in Business Organizations: Aspects and Standards, , (Ed.), pp. 65-94, 2012, IGI Global, ISBN: 978-161-3505-014, G. Yee, A. Antonakopoulou, G. V. Lioudakis, F. Gogoulos, D. I. Kaklamani, I. S. Venieris.

[i.19] "Hippocratic databases": in: VLDB '02: Proceedings of the 28th international conference on Very Large Data Bases, VLDB Endowment, 2002, pp. 143-154, R. Agrawal, J. Kiernan, R. Srikant, Y. Xu.

[i.20] "Purpose based access control for privacy protection in relational database systems": the VLDB Journal 17 (4) (2008) 603-61 9, J.-W. Byun, N. Li.

[i.21] "Limiting Disclosure in Hippocratic databases": in Proceedings of the 30th International Conference on Very Large Databases (VLDB' 2004), pp.108-119, Toronto, Canada, August 31-September 3, 2004, K. LeFevre, R. Agrawal, V. Ercegovac, R. Ramakrishnan, Y. Xu, D. DeWitt.

[i.22] "Exploiting cryptography for privacy-enhanced access control: A result of the prime project": Journal of Computer Security 18 (1) (2010) 123-160, C. A. Ardagna, J. Camenisch, M. Kohlweiss, R. Leenes, G. Neven, B. Priem, P. Samarati, D. Sommer, M. Verdicchio.

[i.23] "PuRBAC: Purpose-aware role-based access control": in R. Meersman, Z. Tari (Eds.), On the Move to Meaningful Internet Systems: OTM 2008, Vol. 5332 of Lecture Notes in Computer Science, Springer, 2008, pp. 1104-1121, A. Masoumzadeh, J. Joshi.

[i.24] "Privacy-aware role-based access control": ACM Transactions on Information and System Security 13 (3) (2010) 1-31, Q. Ni, E. Bertino, J. Lobo, C. Brodie, C.-M. Karat, J. Karat, A. Trombetta.

[i.25] "Organization Based Access Control": in: 4th IEEE International Workshop on Policies for Distributed Systems and Networks (Policy"03), 2003, pp. 120-131, lake Come, Italy, A. Abou-El-Kalam, R. E. Baida, P. Balbiani, S. Benferhat, F. Cuppens, Y. Deswarte, A. Miège, C. Saurel, G. Trouessin.

[i.26] "Modeling Contextual Security Policies": International Journal of Information Security 7 (4) (2008) 285-305, F. Cuppens, N. Cuppens-Boulahia.

[i.27] "Dynamic deployment of context-aware access control policies for constrained security devices": Journal of Systems and Software 84 (2011) pp. 1144-1159, S. Preda, F. Cuppens, N. Cuppens-Boulahia, J. Garcia-Alfaro, L. Toutain.

[i.28] "Contextual privacy management in extended role based access control model": in J. Garcia-Alfaro, G. Navarro-Arribas, N. Cuppens-Boulahia, Y. Roudier (Eds.), Data Privacy Management and Autonomous Spontaneous Security, Vol. 5939 of Lecture Notes in Computer Science, Springer, 2010, pp. 121-135, N. Ajam, N. Cuppens-Boulahia, F. Cuppens.

[i.29] "Privacy-aware access control and authorization in passive network monitoring infrastructures": in CIT 2010: Proceedings of the 10th IEEE International Conference on Computer and Information Technology, 2010, F. Gogoulos, A. Antonakopoulou, G. V. Lioudakis, A. S. Mousas, D. I. Kaklamani, I. S. Venieris.

[i.30] "Legislation-aware privacy protection in passive network monitoring": in I. M. Portela, M. M. Cruz-Cunha (Eds.), Information Communication Technology Law, Protection and Access Rights: Global Approaches and Issues, IGI Global, 2010, Ch. 22, pp. 363-383, G. V. Lioudakis, F. Gaudino, E. Boschi, G. Bianchi, D. I. Kaklamani, I. S. Venieris.

[i.31] "A Contextual Privacy-Aware Access Control Model for Network Monitoring Workflows: Work in Progress": in Proceedings of the 4th MITACS Workshop on Foundations & Practice of Security (FPS), Paris, France, May 12 - 13, 2011, LNCS, Vol. 6888, Springer-Verlag, E. I. Papagiannakopoulou, M. N. Koukovini, G. V. Lioudakis, J. Garcia-Alfaro, D. I. Kaklamani, I. S. Venieris.

[i.32] "How to share a secret": Communications of the ACM vol. 22 (1979), pp. 612-613, A. Shamir.

[i.33] "Completeness theorems for non-cryptographic fault-tolerant distributed computation": in Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing, STOC '88, pages 1-10. ACM, 1988, M. Ben-Or, S. Goldwasser and A. Wigderson.

[i.34] "How to generate and exchange secrets": in Foundations of Computer Science, 1986, 27th Annual Symposium on, pages 162-167, 1986, A. C. Yao.

[i.35] "How to play any mental game": in Proceedings of the nineteenth annual ACM Symposium on Theory of Computation, STOC '87, pp. 218-229. ACM, 1987, O. Goldreich, S. M. Micali and A. Wigderson.

[i.36] "Fully homomorphic encryption using ideal lattices": in Proceedings of the 41st annual ACM Symposium on Theory of Computing, STOC '09, pages 169-178. ACM, 2009, C. Gentry.

[i.37] "Asynchronous multiparty computation: Theory and implementation": in PKC, pp. 160-179, 2009, I. Damgard, M. Geisler, M. Krigaard and J: B. Nielsen.

[i.38] "Fairplaymp: A system for secure multi-party computation": in Proceedings of the 15th ACM Conference on Computer and Communications Security, CCS '08, pages 257-266. ACM, 2008, A. Ben-David, N. Nisan and B. Pinkas.

[i.39] "P4P: Practical Large-Scale Privacy-Preserving Distributed Computation Robust against Malicious Users": in The 19th USENIX Security Symposium, August 11-13, 2010, Washington, D.C., Y. Duan, J. Canny and J. Zhan

[i.40] "Assisting Server for Secure Multi-Party Computation": 6th Workshop in Information Security Theory and Practice, June 19-22 2012, Egham, UK, J.-M. Bohli, W. Li, J. Seedorf.

[i.41] "Enabling conditional cross-domain data sharing via a cryptographic approach": IEEE 5th International Conference on Internet Multimedia Systems Architecture and Application (IMSAA), 12-13 Dec. 2011, G. Bianchi, H. Rajabi and M. Sgorlon.

[i.42] "Large-scale collection and sanitization of network security data: risks and challenges": in Proc. 2006 workshop on New security paradigms. NY, USA: ACM, 2007, pp. 57-64, P. Porras and V. Shmatikov.

[i.43] "Attribute-based encryption for fine-grained access control of encrypted data": in Proceedings of the 13th ACM conference on Computer and communications security (CCS '06), 2006, V. Goyal, O. Pandey, A. Sahai and B. Waters.

[i.44] "Ciphertext-Policy Attribute-Based Encryption": in Proceedings of the IEEE Symposium on Security and Privacy 2007: 321-334, J. Bethencourt, A. Sahai and B. Waters.

[i.45] "Decentralizing Attribute-Based Encryption": EUROCRYPT 2011: 568-588, A. Lewko and B. Waters.

3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

DAC	Discretionary Access Control
DDoS	Distributed Denial of Service
DNS	Domain Name System
DMZ	Demilitarised Zone
GCR	Globally-Constrained Randomization
GUI	Graphical User Interface
IDS	Intrusion Detection System
IESG	Internet Engineering Steering Group
IODEF	Incident Object Description Exchange Format
IP	Internet Protocol
IPFIX	IP Flow Information Export
ISP	Internet Service Provider
IT	Information Technology
MAC	Mandatory Access Control
QoS	Quality of Service
RBAC	Role-Based Access Control
RID	Real-time Inter-network Defense
SEPIA	Security through Private Information Aggregation
SIEM	Security Information and Event Management
SMC	Secure Multiparty Computation
SOAP	Simple Object Access Protocol
TLS	Transport Layer Security
TTP	Trusted Third Party
XML	eXtensible Markup Language

4 Scenario description and basic concepts

The goal of a collaborative cross-domain network monitoring service is to permit multiple (two or more) operators to mutually exchange information so as to permit each peer to achieve a wider perspective of on-going attacks or anomalies extending beyond a domain perimeter, hence achieving a more efficient and effective detection and mitigation. Such a cross-domain network monitoring service is not meant to be defined as a specific solution, but it is rather meant to specify a general framework, characterized by the following crucial and structural properties:

1) Involved entities (domains) should not be bound to agree on any specific or common choice of internal monitoring platforms or systems; rather, the collaborative monitoring service shall be specified in terms of interfaces between domains, leaving every domain in charge of deploying or running its own preferred internal monitoring system.

2) The information to be exchanged across domain is not necessarily limited to the sharing of actual monitoring data, meta-data, or alarms/reports (although especially this latter is the most obvious use case), but may extend (or conversely, restrict) to the exchange of pre-processed and/or suitably encrypted information which can be used to perform a joint (i.e. cooperative) detection of events or anomalies.

3) The collaborative cross-domain monitoring service should not specify a set of pre-established monitoring tasks or specific types of data to be shared, but should rather provide interfaces and control primitives to permit involved domains (or a subset of) to agree on, and correspondingly deploy a cooperative monitoring task or use-case. Thus, the focus is on the framework, rather than on the specific application to a target precise monitoring goal (such as Botnet detection, DDoS mitigation, fraud detection, anomaly detection, etc).

Unless otherwise specified, in what follows for simplicity we consider the scenario depicted in figure Figure 1, involving just two domains (owned by two different ISPs/Operators). In this reduced scenario we focus on the particularities of the proposed communication. Any reasoning for this scenario may be extended to situations with more participants like the one presented above, but we use this one as a starting point for the sake of clarity and simplicity.

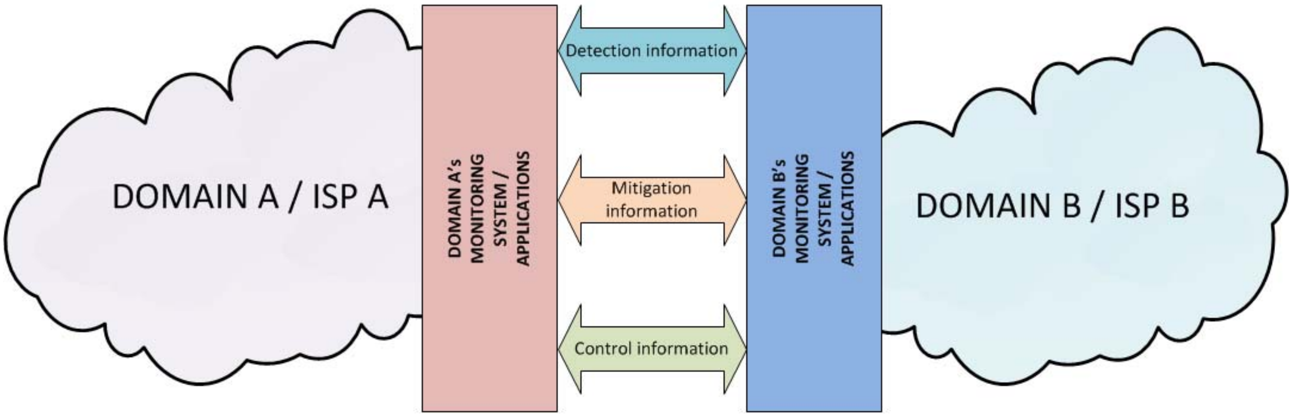

Figure 1 Communication between two domains

Each domain may have its own monitoring systems composed by an array of tools and solutions (network probes, SIEMs, IDS, etc.) Details of those systems are out of the scope of the present document, but we assume that they are providing alarms, security events, etc. comprising the **detection information**. This information, which is today customarily used for internal consumption only, now may be shared with the other peer through a dedicated cross-domain interface. Exchange of information further comprises, when applicable, mitigation information, as well as control information needed to support more advanced cooperative and sharing schemes (e.g. cryptography based).

Specific variants of this scenario are elaborated below.

4.1 Cooperative incident handling by network operators

"Operator A," a country-wide operator, with a large data network, is concerned about their infrastructure. Therefore they have signed an agreement with other carriers to share relevant security-related incident information in order to improve the detection of distributed threats.

Operator A is providing security events detected locally in its network by its security systems, composed of a wide array of monitoring probes, IDS, collectors, etc. In exchange, Operator A receives similar information coming from the other operators in the consortium. Operator A is using these data to feed its already existing correlation systems, getting a broader view with data that originate from other domains and improving the detection rate and/or the characterization of certain attacks with this information. So the information is spread "globally" and used locally by each member of the consortium according to their own preference.

Eventually, Operator A is detecting some traffic volume in the order of several Gigabits per secondconverging on some few nodes of its infrastructure, which is having some impact on the performance of some services offered to customers. Beyond taking the appropriate countermeasures to mitigate that attack on those nodes, Operator A is not able to characterize the whole attack since it is highly distributed and really complex to track down, but decides to share this incident with other operators. The other operators receive the incident information and correlate it with their current traffic activity in their networks, resulting in the association of the sources and the target of the attack characterizing a DDoS attack. Although they are not the target of the attack, hosting the elements of a botnet has a severe impact on their public reliability and reputation, so they also take measures to cut that traffic, causing the end of the attack or at list some noticeable relief on Operator A's attacked infrastructure.

Thus, both the first immediate victim (Operator A failing to keep the QoS) and the secondary ones (other operators hosting attacker and losing reputation) obtain some benefit from this collaborative model.

4.2 Cooperative incident handling among enterprises

"Bank B," a bank, has a large IT network hosting many critical services. Bank B is really worried about the security of its services, and, as part of its Security policy, enforces traffic management policies (particularly outbound traffic) in order to detect data extrusion, malicious users, malware or network traffic that may pose a threat to the security of neighbouring systems. The final goal here is protecting the service and their users (customers) more than the infrastructure by anticipating the attack at the origin. Thus, analysing the data received at and sent by the service is paramount. Since data exfiltration techniques are normally the same in many incidents, they have joined some related partners (including other banks) in a cooperative initiative focused on sharing incident information. Bank B is using incident info from other partners to better identify similar events, and the same in the other way round: they are sharing their own detected incidents to help the others. Moreover, an extrusion detection system is more generally aimed at identifying successful and unsuccessful attempts to use the resources of a computer system to compromise other systems so it tries to prevent attacks from being launched in the first place. Collaboration among as many as actors as possible makes a lot of sense to detect more attacks at the origin and incident sharing let the targeted systems protect from the attack.

Bank B monitors all its critical services. Eventually, a suspicious activity in one of the services is detected; part of the outbound traffic seems to be malicious. Immediately, it reports the incident the other enterprises (including relevant information but also applying some data transformation to comply with privacy restrictions and regulations). Other partners receive the information and set up security means to protect from the potential attack. The symmetric situation is also possible: a partner reports a suspicious activity generated in its network and Bank B takes some mitigation actions to protect from the threat.

4.3 Cooperative anomaly and misuse detection

The two previous variants consider final results at some domain that are shared to others. This last variant considers the case of intermediate results that are shared to get an overall final result. Let us consider Operator A again; the company has decided not only to share their own detection results but also to set up real collaborative cross-domain application with other Telcos. This means that they can create applications that provide larger coverage since they operate with input data from different domains. This is noticeably different from the first scenario; now we deal with intermediate data instead of final results. We do not have applications at each operator that are sharing results, but just raw intermediate data that are analysed by one distributed application.

As a good initial application they decide to deploy a DNS misuse application. Each participant provides data related to DNS traffic in its network; these data are neither relevant nor conclusive itself but when aggregated with data provided by the other partners lead to some more relevant results. Each partner computes the results respecting the privacy and business constraints of the external data and all the partners get finally the same output. Similar collaborative applications can be developed using any kind of input data that operators are able to share.

5 Sharing schemes used in cooperative incident handling

In the previous clause we have presented a general scenario and some variations motivating the existence of some means to exchange information about security incidents between domains. Now we present three different schemas for that sharing (and many others that we may consider). These schemas are general approaches fitting any general cross-domain monitoring situations. They apply in one way or another to the described scenarios.

When considering network security monitoring in a multi-domain environment, it becomes necessary to define additional requirements for the sharing of intermediate or final results across administrative domains or jurisdictional boundaries. There are three general approaches to cross-domain cooperation in network monitoring. The simplest of these is the sharing of desensitized final results, where the publisher of the data trusts the recipient to use the results properly, and not to forward the results to untrusted parties. That is, while there may be transformations done to the data to reduce their sensitivity while maintaining their utility (e.g. anonymization of IP addresses) and annotations on the data which describe the data and the expectations the publisher has as to what the recipient may do with the data, there are not any technical protections on the data. We call these annotated sharing schemes.

Several applications require the aggregation, union, or intersection of final results from multiple publishers into a single shared result, generally to be distributed back to all participating publishers, or more publicly. These applications share the property that while the information from each publisher is still sensitive, the aggregate or intersection/union, without attribution, is less so. There are two ways to arrange such applications. First are trusted-third-party (TTP) sharing schemes, where each publisher hands the sensitive data to a third party trusted by all publishers. This third party then performs the calculations, and returns or further publishes the final results, destroying the per-publisher data. Nevertheless, such approaches are in principle unacceptable, because not only sensitive data are disclosed, but also the third party becomes a warehouse of sensitive information, with the dangerous potential of combining and correlating this information and, eventually, extracting sensitive meta-information and generating even more sensitive information.

Many of the implications with TTP sharing schemes can be remedied through the application of cryptographic technology, specifically secret sharing schemes from secure multiparty computation. Here, the TTP is replaced by a collection of privacy peers operated by the publishers, such that any one entity does not have access to any published data from any other publisher, or to any intermediate results. Schemes for secure multiparty computation specifically applicable to cooperative network monitoring include SEPIA [i.1] and GCR [i.2].

The subsequent clauses explore and define requirements for each type of scheme.

5.1 Annotated Sharing Schemes

Annotated sharing schemes refer to any network traffic data sharing schemes between or among organizations in a, whether manual or automated, wherein the data is annotated to indicate its contents and the preferences which the publisher has with respect to the use and further publication of the data.
These annotations can take virtually any form: they may be appended explicitly inline, attached externally to the data in a metadata file, made available via a web service, implicit to an entire collection of data, or even contained in some contract between the publisher and the recipient.

There are several examples of such schemes. IODEF [i.3], the Incident Object Description Exchange Format, is an XML format designed for the description and interchange of data about network security incidents. These incident reports are derived from network traffic measurements as well as information about the involved entities and limited information about remediation (e.g. "denial of service attack from IP address A, against host B, which is a webserver; suggest rate-limiting the source address"). RID [i.4] [i.5] adds a protocol atop IODEF for on-line exchange of incident reports to support cooperative mitigation of network security events.

IODEF provides limited annotations for data sharing restrictions, through its restriction enumeration, which limits information to "public", "need-to-know", or "private" dissemination, with an additional value "default", which references an external agreement between the publisher and recipient. Each element of an IODEF document may have a different restriction, allowing public, redacted versions of IODEF documents to be derived from those containing; note, however, that the only method of redaction involves the removal of elements. The charter of the IETF MILE working group includes an internet-draft [i.6] to add additional elements to IODEF for a richer description of data sharing restrictions; this work is ongoing.

Annotations can also include information about how a particular data set was handled, e.g. [i.7], which adds metadata to IPFIX exports [i.8] or IPFIX files [i.9] for noting which fields of the exported records are anonymized, and how. This specifically does not add data handling markers, but would work together with external annotations to provide a complete picture of the contents of a data set, and what can be done with it.

First, annotations are advisory only, and do not provide any technical protection for the contents of the annotated data. They should only be used in situations where the publisher explicitly trusts the recipient to follow the data handling annotations, and that the recipient is competent to protect the data as well as the publisher itself. It is recommended that such schemes only operate within the framework of an explicit agreement between the publisher and recipient, and that annotated data only be published to the general public when it contains no sensitive information.

Given the advisory nature of data sharing annotations, the security of data in transit shall be guaranteed though cryptographic means. When sharing data using protocols which specify a binding to a transport-layer security scheme (e.g. TLS), as in the case of the RID and IPFIX examples above, the transport-layer security shall be used in order to guarantee integrity and confidentiality of the shared data. It does no good to agree with a recipient to follow a set of data sharing markers when the data can be intercepted and republished by a third party not party to the agreement.

Network traffic data shared with an annotated sharing scheme should be reduced or anonymized to remove personally-identifiable information about the network's end users. However, as there has been much work in the past years on the limitations of anonymization of network traffic data [i.10], such schemes are most applicable to prevent accidental disclosure of anonymized data. Anonymization should only be applied with a full understanding of its limitations, and anonymization alone should not be relied upon in the absence of trust arrangements between the publisher and recipient.

5.2 Trusted-Third-Party Sharing Schemes

For the subclass of problems in network monitoring where the source data of a given computation is more sensitive than the result (e.g. traffic aggregates, heavy-hitter aggregates, or set unions/intersections of watch lists), the simplest approach is an evolution of the annotated sharing scheme. A group of publishers in a consortium agree to trust a third party, publish their intermediate result to the third party, which then performs a calculation and returns the (less sensitive) results of the calculation to the publishers and/or to some other designated recipient.

As this is an elaboration of the annotated sharing scheme (the publishers can be said to use a form of annotated sharing when transmitting the initial data to the TTP), the requirements as for annotated sharing apply as well: transport-layer security shall be applied to the communication between the publishers and the TTP, and anonymization of source data shall not be relied upon as a complete solution for data protection.

There are additional requirements both for the TTP and the publishers in such arrangements. First, the TTP should only provide coordination and calculation services. The TTP should not retain any source data or intermediate data generated by the aggregation operation; otherwise, it would represent a potential single point of security failure.

For aggregation and set operations, each publisher has access to its own source data, and can in effect subtract this from the aggregated data in order to know what information came from all the other publishers taken together. There is no way to keep each publisher from having this knowledge; however, it publishers collude outside the scheme, they could further compare their results in order to break the privacy protections afforded by the scheme. Therefore, no communication among publishers outside the TTP scheme with respect to the shared data should be permitted. Additionally, each application of a TTP scheme should ensure an adequate number and diversity of participants so that no one participant, with its own data as well as the aggregate results, can deduce the inputs of the other participants. The degenerate case of this is the obvious insight that a TTP aggregation scheme with two participants is useless; but note that in the general case, these requirements are difficult to enforce through technical means.

TTP schemes are further complicated in practice in that the trust each organization places in the third party shall be absolute; since the TTP sits at the centre of the scheme, it can modify both the input or output data, or not calculate the aggregate in the way agreed with each publisher. The prohibition on side channels among the publishers makes verification difficult or impossible

5.3 Secure Sharing Schemes

Secure sharing schemes are a further elaboration of TTP schemes, which eliminate the TTP (and therefore the problems presented thereby) and replace it with a cryptographic scheme. In the general arrangement, each publisher generates private shares from their own source data, from which the source data itself cannot be derived, but from which aggregation is possible, and publishes these to a set of privacy peers, each operated by one of the publishers. The peer processes then calculate the aggregate and return it to the publishers. Examples of such systems include SEPIA [i.1] and [i.10].

Since these schemes replace a TTP with a federation of privacy peers, all of the requirements which apply to TTP schemes but not to the TTP itself apply as well to secure sharing schemes. Specifically, no communication among publishers outside the secure sharing scheme with respect to the shared data should be permitted. Additionally, each application of a secure sharing scheme should ensure an adequate number and diversity of participants. The use of cryptographic primitives to generate the private shares reduces the need use transport layer security when sending them securely to the privacy peers. But any communication of the data to be shared before it is converted to private shares shall be protected by transport-layer security.

With consideration for the attacker model supported for the given scheme, note that secure sharing schemes remove the need for absolute trust in a trusted third party, and remove the requirement to prevent data retention on the privacy peers, as the source data cannot be derived from private shares.

6 Requirements

Note that the requirements in this clause refer only to privacy and security aspects of distributed network monitoring.

6.1 Business requirements

Business requirements refer to those requirements posed by the processes and business environment in which any distributed network monitoring system will operate. These include requirements to keep business information confidential, as well as to minimize the impact of distributed network monitoring on other operations.

6.1.1 Confidentiality of business-sensitive data

Any incident information shall not include any piece of data revealing sensitive business information.

Access to data shall be controlled. The definition of access control policies shall take into account all applicable parameters, such as the role of the entity being the data recipient, the semantics of the data themselves, the underlying purpose of foreseen processing and the data confidentiality level.

Access control policies shall be defined in a fine-grained manner, following the granularity of the data, as well as of the recipient (internal role model vs. external organisation).

There shall be some policy (or policies) defining the level of confidentiality of the data (this data is exportable, this one is restricted to certain partners, etc.).

The format of the policy shall be machine-readable.

There shall be a mechanism to inspect the data and enforce the defined policies regarding confidentiality prior to the sharing of data.

It may be required to transform the data before being shared to meet the confidentiality policies (e.g. data anonymization).

6.1.2 Impact on network operations

6.1.2.1 Roles

Roles and attached rights should be defined for the system to clearly separate different activities (development, operation, management).

6.1.2.2 Same domain operations

There shall be some application deployment procedures available: e.g. deploy, update and remove application.

There shall be some runtime procedures for the operational exploitation of the monitoring system: e.g. start, stop and restart application.

There should be a GUI making available all the relevant information and operations in the domain regarding the monitoring activities. This includes statistics and indicators for any running applications.

6.1.2.3 Cross-domain operations

There shall be a way to configure different inter-operator exchange strategies activated per operator's peer process, with an application level granularity. Each telecom operator may adopt a different cooperation level according to the considered peer process and application.

A dedicated role should be made available in order to handle the specific constraints associated to inter-domain collaboration (e.g. business reasons, legal constraints, strategic relationships).

There shall be an explicit authorisation mechanism for inter-domain requests. For example, even if an inter-domain mitigation could be automatically defined when an alert is raised, an operator should validate it.

6.1.2.4 Legacy systems

The system shall envision the integration with legacy monitoring systems. Two types of legacy systems shall be considered: the ones that provide information (legacy inputs) and the ones that consume information (legacy outputs).

It may be required to develop adapters to facilitate the integration of such legacy elements.

6.2 Regulatory requirements

Regulatory requirements refer to those requirements posed by the legal and regulatory environment in which any distributed network monitoring system will operate. While these requirements may vary widely from jurisdiction to jurisdiction, this clause focuses on the legislation of the European Union, since it comprises the most representative, influential and mature approach worldwide, that seems to pull a general framework and has been characterised as an "engine of a global regime" [i.11]. The requirements are as follows.

6.2.1 Lawfulness of data processing

A monitoring system should be able to evaluate the lawfulness of each request for personal data, in accordance to applicable laws and regulations. In practice, the system should be structured in a way that enables assessing the legitimacy of a request of access to data submitted by the different components of the system.

6.2.2 Purposes for which data are processed

A monitoring system should provide the means for identifying the purpose of each request, which shall be lawful and made explicit to the data subject, and should function so that it allows the collection and processing of personal data only when said activities are carried out for specified, explicit and legitimate purposes. In addition, the system should prohibit that personal data collected for some specific and legitimate purposes are used for other purposes, incompatible with those for which the data have been originally collected.

6.2.3 Necessity, adequacy and proportionality of the data processed

A monitoring system should be able to guarantee that only the data that are functional, necessary, relevant, proportionate and not excessive with regard to the sought processing purpose are processed.

6.2.4 Quality of the data processed

A monitoring system should ensure that the data processed are correct, exact and updated. Moreover, the system should be able to perform corrective actions, as well as audits, in order to delete or correct inaccurate data and to delete or update data that are outdated or redundant.

6.2.5 Minimal use of personal identification data

A monitoring system should minimise to the extent possible the use of identification and personal data only when this is a prerequisite to the specific monitoring function that is to be performed.

6.2.6 Storage of personal data

A monitoring system should keep personal data in an identifiable form only for the time that it is strictly necessary to the specific monitoring function that is carried out. Personal data that are redundant or no longer needed should be deleted or anonymised.

6.2.7 Data retention

A monitoring system should comply with the requirements set forth by applicable data retention regulations, such as, in the EU area, the Directive 2006/24/EC [i.12] and the corresponding national laws. This implies that the system should store the specific data that are subject to the data retention regulations for the time periods specified under the applicable regulatory framework.

6.2.8 Access limitation

A monitoring system should authenticate all users of the system, should provide different levels of access to the stored data and should provide for the logging of all access to the stored data in order to detect attempted or successful unauthorised access.

6.2.9 Information to and rights of the data subject

A monitoring system should be able to provide that data subject is duly informed on the purposes and features of the relevant data processing according to peculiarities of applicable data protection legislation. As to the privacy rights of the data subjects, the data subject should be provided with the possibility to access his/her personal data, to ask for specific information about the processing of his/her personal data, to ask for his/her personal data to be integrated, updated, rectified, deleted or transformed in an anonymous form. The data subject should also be enabled to block the processing of his/her personal data in case of breach of applicable laws and to object the processing of his/her personal data for legitimate reasons.

6.2.10 Consent of the data subject

A monitoring system should guarantee that, when required by applicable data protection legislation, the data subject's consent to the data processing is requested and obtained and that the data processing is further performed according to the preferences expressed by the data subject. The data subject should be enabled to revoke at any time the consent previously granted (even temporarily in case of location and traffic data processed for the performance of value added communications services).

6.2.11 Data security measures

A monitoring system should adopt appropriate technical and organisational measures with the purpose of protecting the personal data that are collected and processed against the risks of accidental or unlawful destruction, accidental loss, alteration, unauthorised disclosure or access, in particular where the processing involves the transmission of data over a network, as well as against any other unlawful possible data processing operation or set of operations.

6.2.12 Special categories of data

A monitoring system should guarantee that the processing of special categories of data (for example, but not limited to, traffic or other location data) is performed in compliance with the specific requirements that the applicable data protection legislation sets forth for said categories of data. For instance, in the European Union, the Directive 2002/58/EC [i.13] defines specific requirements, security and others, regarding the processing of communication data.

6.2.13 Coordination with competent Data Protection Authority

A monitoring system should monitor compliance with the notification requirement and with the provisions of the authorisations of the competent Data Protection Authorities, as ruled under applicable data protection legislation. Moreover, the system should allow communications between the system and the competent Data Protection Authorities in order to validate and verify that the notification and/or authorisation requirements have been duly complied with.

6.2.14 Supervision and sanctions

A monitoring system should provide the competent Data Protection Authorities with the means for supervising and controlling all actions of personal data collection and processing. This function is very important, as it often happens that the competent Data Protection Authorities encounter difficulties in auditing the processing of personal data carried out through technical means and over the Internet; this is due to the peculiar nature of the technical means deployed, that allow the hiding of the data processing activities performed.

6.2.15 Communications confidentiality

A monitoring system should be structured consistent with the protection of the confidentiality of communications over the monitored networks. Indeed, the European Union legislation prohibits the listening, tapping, storage or other kinds of interception or surveillance of communications and the related traffic data, unless the user has given consent and such surveillance is technically necessary to provide the data subject with the requested communication service.

6.2.16 Dissemination of data to third parties

When components of the service logic are outsourced to third party providers and, in this context, personal data are disseminated for being processed, the monitoring system should be able to provide certain guarantees that the consequent processing of information complies with the underlying fair data practices and the contract with the data subject.

6.2.17 Transfer of data to third countries

The dissemination requirement above applies especially when data are transferred to third countries, possibly with essentially different legislation regarding personal data collection and processing. The system should be able to provide for compliance with the specific provisions ruling on transfer of data. For instance, consider the Safe Harbour Principles regulating data transfer between the European Union and the USA [i.14].

6.2.18 Flexibility and adaptability of legal compliance provisions

Given the complexity of the legal environment in which a monitoring system operates, the different legal requirements across different jurisdictions and the nature of the law to change from time to time, the system's design should to the extent possible be flexible and adaptable with respect to all the provisions described be the requirements outlined above.

6.3 Technical requirements

Technical requirements are derived from the business and regulatory requirements declared above as well as from the implementation and deployment environment in which a distributed network monitoring system will operate. This includes additional privacy and security requirements which are listed in the following clauses.

6.3.1 Privacy requirements

This clause draws mostly from the Regulatory Requirements (clause 6.2), translating the provisions summarised there into technical requirements for the preservation of privacy. Nevertheless, not all the regulatory requirements are within the scope of technical implementation. For instance, the provisions related with the interaction with the data subject, e.g. regarding consent, are typically dealt with contractual provisions. Moreover, the bottom-line for privacy protection and the enforcement of several associated requirements is security; the corresponding requirements are covered in clause 6.3.2.

6.3.1.1 Purpose specification and binding

A very fundamental concept for privacy is the purpose for which data are processed, which appears often in the requirements, being a core part of lawfulness. The system should provide the means for identifying the purpose of each request, in order to allow the collection and processing of data only when said activities are carried out for specified, explicit and legitimate purposes. Moreover, the purpose principle prescribes mechanisms for specifying the compatibility between processing purposes, while also providing for checking whether the processing purposes are consistent with these for which data have been collected. In this respect, the system should also provide the means for defining prevention rules regarding incompatible purposes. In order to achieve the above, the formal description of purposes is deemed necessary.

6.3.1.2 Necessity, adequacy and proportionality

The principles of necessity, adequacy and proportionality are also tightly connected to purpose. In fact, the system should provide the means for necessity specification, as well as be able to examine whether the collection or processing of specific data is necessary for the provision of the service in question, i.e. the fulfilment of a purpose. This implies a relation between data, processing activities, roles and purposes that shall enable the definition of necessity and proportionality constraints, as well as the determination of what constitutes minimal use. Furthermore, it should enable the definition of different levels of data granularity, so that the accuracy of disclosed data can be adjusted depending on the purpose and the subject requesting access to the said data, among others. It should be noted that the requirement for different levels of data accuracy and granularity, along with the existence of different categories of data, prescribe their semantic description and discrimination; similar needs for semantic taxonomies apply to the associated concepts, such the roles and the processing functions.

6.3.1.3 Cooperation with third parties

Especially in the context of inter-domain cooperative scenarios, data collected or generated in a domain, which in this case constitutes the Data Controller, may be further subject to processing within another domain, notably being the Data Processor. This feature makes of particular importance the adoption of protection mechanisms for data that are exchanged across domains, i.e. disseminated to third parties. Similarly, for effective protection of privacy as well as business information confidentiality, the system should enable flexibility and adaptability of provisions. To this end, the system should provide technical features that allow interoperation, as well as negotiation between domains, during which the system should be able to verify that the Data Processor has complied with the aforementioned requirement towards the Data Controller. Even better, the system should provide the means for preventing the Data Processor from misusing data; therefore, the third parties should not in principle receive data in identifiable form, in line with the associated minimisation principle. In general, as the system shall support communication of data between different domains, the type of transferred data and the circumstances under which they are being transferred shall be checked. To this end, the system shall additionally support the negotiation of policies in the presence of regulatory discrepancies.

6.3.1.4 Complementary actions

In some cases the access to data involves complementary actions, which will precede or follow the access. Thus, the system should impose this kind of actions, expressed as pre-/post-actions that need to be fulfilled; this consists, for instance, in the anonymisation of data on the fly for adapting with the proportionality principle.

In this context, the associated functionalities implementing the complementary actions should be provided, along with the means (e.g. policies) for specifying and enforcing their incorporation to the monitoring operational chain.

In addition, logging functionalities shall be provided; apart from being an often required complementary action itself, logging enables the traceability and, therefore, the accountability of actions performed.

6.3.1.5 Data storage and retention

Regarding data storage, the system should provide the means for automatic update or deletion of data, when the latter are outdated or incorrect. This prescribes the development of machine-specific code deriving from associated rules for the automatic administration of the organisation's databases without the need for a human administrator's intervention. The automatic administration of databases is of great importance also in the case of data retention periods enforcement. In general, the system shall be able to control as dictated by the regulations the storage of the data for retention purposes; therefore, the use of policies for control realisation is implied.

6.3.1.6 Data protection mechanisms

The monitoring system should be empowered with effective mechanisms for the active protection of data, mostly leveraging cryptographic means. Thus, the appropriate mechanisms for anonymisation, pseudonymisation, selective disclosure, secure multiparty computation, etc., should be available.

6.3.1.7 Access control

Any privacy violation certainly includes illicit access to personal data and, intuitively, access control constitutes a fundamental aspect of privacy protection. The enforcement of effective access control is highlighted not only regarding access to the data, but also to other system resources. The definition of access control policies should take into account all applicable parameters, such as the role of the entity being the data recipient, the semantics of the data themselves, the underlying purpose of foreseen processing and the data confidentiality level. Moreover, access control policies should be defined in a fine-grained manner, following the granularity of the data, as well as of the recipient (internal role model vs. external organisation).

6.3.1.8 Semantics

As highlighted by the regulatory requirements, the particular type characterising each personal data item constitutes a parameter of significant importance for the determination of the procedures subject of which the data will be and, consequently, of the underlying monitoring operations. The semantics of the operation itself, the underlying collection and processing purposes, as well as the entities involved (such as their roles), are of equal importance. Therefore, a monitoring system should be enabled to take into account the semantics of the associated concepts.

6.3.2 Security requirements

We enumerate here the most relevant security needs to be considered in any eventual implementation and deployment of a cross-domain security information sharing system. These requirements are intended to be implementation-agnostic (we do not assume any specific implementation choice). Security requirements should be extended and further detailed when considering specific implementations. The very general assumption is that we have some degree of component distribution.

6.3.2.1 Confidentiality

Authentication mechanisms are required at any node of the system to allow the access just to the authorized users or modules.

6.3.2.2 Integrity

Communication between nodes shall keep the integrity. Communication shall be encrypted to avoid eavesdropping and data tampering.

6.3.2.3 Availability

The system shall have high availability at any time since it delivers critical security functionality.

In case of having centralized control elements, they shall be redundant to avoid single points of failure.

Any eventual monitoring network used to connect the different components of the system shall not compete for resources with the protected network.

The distributed monitoring system shall have its own resources (communication network, storage, etc.).

6.3.2.4 Backup

A backup mechanism shall be in place storing configuration data, executable components and auxiliary data of every node to allow quick reaction to any failure or disaster.

6.3.2.5 Access audit logs

Some audit functionality is recommended to empower forensics in the event of any security incident requiring further investigation, let identifying operational errors, security breaches, internal attacks, etc. This requirement is often compulsory in some cases due to legal reasons.

6.3.2.6 Network Segmentation

The system shall contemplate a segmented design to isolate components exposed to external networks from the internal network (DMZ model).

Any component requiring access to external networks (e.g. internet) shall be divided in two parts each: one part shall be placed in the internal domain and is in charge of the core operations. The other part is deployed in a DMZ and just takes care of the external communications.

7 Available solutions and gaps

7.1 Incident information sharing (IETF INCH and MILE)

The **Extended Incident Handling (INCH)** Working Group was part of the Security Area of the Internet Engineering Task Force (IETF). It was chartered to create an exchange format for computer security incident data used by Computer Security Incident Response Teams (CSIRTs).

The working group authored five documents. The requirements for INCH were documented in the Format for Incident Reporting (FINE). A data model, the Incident Object Description Exchange Format (IODEF), for exchanging this incident information was specified. The Real-time Inter-Network Defense (RID) protocol provided a messaging format for IODEF which with an associated binding to SOAP was provided. An initial extension of the core IODEF data model describing phishing information was also authored.

The INCH WG was closed in October 2006. Prior to this closure, the requirements and IODEF data model were submitted to the IESG as WG documents. Ultimately, the requirements document was not sponsored for publication. The remaining documents were resubmitted as individual drafts on the standards track.

The **Managed Incident Lightweight Exchange (MILE)** working group has been recently created within IETF with the goal of developing standards and extensions for the purpose of improving incident information sharing and handling capabilities based on the work developed in the IETF INCH working group, mainly IODEF and RID.

The MILE working group aims at creating extensions and guidance to assist with the daily operations of CSIRTs at an organization, service provider, law enforcement, and at the country level. The application of IODEF and RID to inter-domain incident information cooperative exchange and sharing has recently expanded and the need for extensions has become more important. Efforts continue to deploy IODEF and RID, as well as to extend them to support specific use cases covering reporting and mitigation of current threats such as anti-phishing extensions.

General introduction to IODEF and RID are provided in the following sub-clauses.

7.1.1 The Incident Object Description Exchange Format - IODEF

The Incident Object Description Exchange Format (IODEF) is an Internet Engineering Task Force (IETF) Standard which defines a data representation that provides a framework for sharing information commonly exchanged by Computer Security Incident Response Teams (CSIRTs) about computer security incidents. The specification (RFC 5070 [i.3]) describes the information model for the IODEF and provides an associated data model specified with XML Schema.

Some general assumptions of IODEF:

- Incidents are not IDS alarms; "Incidents are composed of events".
- Agnostic to specific incident taxonomies; "Your definition/threshold of an incident may be different than mine".
- Incidents are numbered and there is state kept about them; "Organizations assign incident IDs and have ticketing/handling/correlation systems that process them".
- Merely a wire format; "Sharing is different than storage and archiving".

- Incomplete information; "You may require more complete information than I need, can get, or have right now".

7.1.2 Real-time Inter-network Defense – RID

Real-time Inter-network Defense (RID) is an ongoing IETF standardization activity (RFC 6545 [i.4]) which outlines a proactive inter-network communication method to facilitate sharing of incident handling data while integrating existing detection, tracing, source identification, and mitigation mechanisms for a complete incident handling solution. Combining these capabilities in a communication system provides a way to achieve higher security levels on networks. Policy guidelines for handling incidents are recommended and can be agreed upon by a consortium using the security recommendations and considerations.

RID's goal is enable the exchange of incident information by:

- Facilitating secure communication of incident information between providers, entities, regions, or countries
- Enabling tracking of incidents as investigations evolve
- Tracing incidents to the source
- Stopping or mitigating the effects of an attack
- Integrating with existing and future infrastructure components

RID has an important focus on Security and Privacy:

- Session and stored encryption: XML digital signatures and encryption, TLS used in transport
- Authentication for single and multi-hop scenarios
- Consortiums to establish trust relationships
- Regional and international security and language barriers addressed via IETF Internationalization
- Privacy: Data restriction markings, ability to optionally provide, possibility to anonymize or encrypt data

Both items are closely related: RID is intended to be a communication method to share IODEF formatted incident reports (RFC 6546 [i.5]). RID provides security, privacy and policy support.

7.1.3 Applicability to collaborative cross-domain network monitoring

Attending to the aforementioned characteristics of both IODEF and RID, they emerge as very appropriate solutions for inter-domain communication of security incidents, which is a key part of the scenarios described in the present document. IODEF seems to be a good choice to use as preferred format when exchanging information between peers. Having a standard format is something more than desirable; it makes much simpler the adoption of a collaborative cross-domain monitoring by new partners (new enterprises, operators, etc.).

RID is a natural choice to use as protocol to transport IODEF. However, it is not so clear that it is the best option for all the cases. Some further work is required to clarify the applicability of RID.

7.2 Access Control

Access control constitutes a fundamental aspect in network monitoring as far as security and privacy are concerned; it can be interpreted in a variety of ways, including access to monitored data, monitoring devices and processing operations. Moreover, access policies may reflect operational aspects, mostly related to security, such as the behaviour of a firewall or the routing table of a router redirecting malicious traffic to a honeypot. In addition, access control can provide the means for privacy preservation, controlling the access to sensitive personal data.

A variety of models are now applied in real systems, starting from traditional ones, like *Discretionary Access Control* (DAC) and *Mandatory Access Control* (MAC) [i.15], as well as the family of *Role-Based Access Control* (RBAC) [i.16]. An important standard in the area is the eXtensible Access Control Markup Language (XACML) [i.17], an OASIS initiative that provides the means for expressing and interchanging access control policies, offering the functionalities of most security policy languages and enabling the definition of new resource types and functions.

Nevertheless, these models fail to meet the requirements stemming from the fundamental privacy principles, as described in clause 6.3.1. In this context, the trend of *privacy-aware access control* has emerged, typically concerning the enhancement of RBAC in order to incorporate privacy-related criteria in access control decisions [i.18]. A common, baseline characteristic of all models falling into this category is that a central role is held by the concept of *purpose* for which *Personal Identifiable Information* (PII) is collected and/or being processed. The first influential approach has been the so-called *Hippocratic Databases* [i.19], that leveraged purpose as a core parameter for providing access to information, resulting in purpose-aware transformations of queries requesting data from relational databases, an approach that has been followed by other models later (e.g. [i.20] and [i.21]). Research has been also driven by the idea of integrating access control with privacy policies, in the sense that the former should ensure the enforcement of the latter; representative models of this family are the ones described in [i.22], [i.23] and [i.24]. These models introduce additional aspects in line with the technical requirements of clause 6.3.1, such as automation of retention periods' enforcement, automatic update or deletion of data, pre- and post- obligations and contextual constraints.

However, these approaches have not been designed for meeting the particular requirements of collaborative, cross-domain network monitoring. While approaches relying on query transformation (e.g. [i.19] and [i.20]) cannot be suitable for real-time monitoring systems dealing with very high data rates, the main drawbacks of all previously mentioned approaches are that their models fail at conceptualising the corresponding functionalities and infrastructures and that they do not generally provide support for collaborative and dynamic cross-domain environments. Moreover, in environments with dynamic operations, contextual information becomes very important. Legacy approaches either do not support context-awareness or they only support straightforward contexts; therefore they are not suitable for highly dynamic and distributed environments and for automating security and privacy-awareness. In fact, when access control has to be coordinated across multiple components that interoperate for the fulfilment of complex goals, including both detection and mitigation operations, the automation of the enforcement of security and privacy requirements becomes a critical issue, while also controlling the access to monitoring infrastructures proves to be more complex.

In this direction, the *Organization-Based Access Control* (OrBAC) [i.25] and [i.26] approach, one of the most important and cited access control models proposed, incorporates a variety of features mitigating such drawbacks. Moreover, it has been subject to numerous extensions that cover a miscellany of needs in a range of domains, e.g. for dynamically deploying security policies [i.27] or enhancing privacy [i.28]. The specificity of the OrBAC approach, apart from the semantically rich formal representation of contextual parameters and their management, is that traditional triples of subjects, actions and objects are abstracted at the organisational level. The concept of organisation itself holds a prominent position, allowing for definition of access control rules beyond the boundaries of a single domain.

Recently, some models specifically devised for network monitoring have been proposed, such as PRISM [i.29], which aimed at meeting the legal and regulatory requirements related with privacy [i.30]. The PRISM approach specifies a semantic policy model capturing a variety of concepts related to network monitoring, as well as an architecture for the enforcement of the associated policies. It anticipates a transformation mechanism for efficiently tuning the granularity of data to which access is granted, while another important feature is its mechanisms for avoiding resource-demanding real-time reasoning. Its basic limitation is that it only applies to single-probe environments; thus, it misses concepts such as organisations and it is not suitable for distributed network monitoring.

This gap is filled by the FP7 ICT project DEMONS (http://fp7-demons.eu/), in the frame of which an innovative access control model is being developed [i.31], espousing the advantages of OrBAC and PRISM and taking into consideration the requirements of clause 6.3.1. Similar to PRISM, it is conceived on the basis of data protection legislation, while it is designed for meeting the requirements of distributed network monitoring. Moreover, it fosters the realisation of the *Privacy by Design* vision, by driving the automatic verification of network monitoring workflows' compliance with the privacy principles and their enhancement with privacy features already at design-time. It relies on a rich in semantics information model that captures all concepts related with collaborative network monitoring and upon which the access rules are specified. The rules, along with the detailed interrelations of the monitoring concepts, enable the evaluation of the workflows - from both a security and a privacy point of view - and their appropriate modification at design-time, taking into consideration a variety of parameters, including access rights, purpose, Separation of Duty (SoD) and Binding of Duty (BoD) constraints, dependencies between actions and contextual variations.

7.3 Secure multiparty computation and other cryptographic approaches

Secure multiparty computation is a cryptographic concept for a group of parties to compute a function on their private input data, without giving access to the private input data to the other parties. It implements a secure sharing scheme as described in clause 5.3 of the present document.

For a flexible scenario, generic schemes that can realise any functions are preferable. The foundations of such secure multiparty computation schemes are now three decades old [i.32] and [i.33]. The basis for secure multiparty computation is Shamir's secret sharing scheme, which uses polynomials to represent secrets and evaluation points as shares.

Other secure multiparty computation schemes are garbled circuits which were introduced by Yao as well as Goldreich et al. [i.34] and [i.35]. The function to be computed is first transformed to an encrypted circuit; then an evaluator obtains obliviously the keys for each bit of the input from the circuit creator and evaluates the circuit through partial decryption. The main concern is the performance, as public-key primitives are needed to retrieve every bit of the input. Another way for secure computation is the use of homomorphic encryption. Fully homomorphic schemes were recently proposed [i.36], however, the performance of these fully homomorphic schemes is currently far from being practical.

Several implementations of generic secure multiparty computations frameworks exist [i.10], [i.37], [i.38] and [i.39]. Among them, the SEPIA system has recently been applied to network monitoring problems. One of the issues with applying secure multiparty computation to network monitoring is performance: for the most part, these schemes have evolved to deal with relatively small volumes of data, while network monitoring typically deals with larger data sets derived from packet or flow analysis. The SEPIA library implements Shamir's scheme in Java, and is targeted toward network monitoring applications, providing a set of appropriate basic primitive protocols (addition, multiplication, and comparison) with a specific focus on performance through parallelization and reduction of synchronization among peers. A further speed-up, if necessary, could be achieved by carefully chosen security/efficiency tradeoffs [i.40].

Secure multiparty computation on its own is not, however, a panacea for cooperative network monitoring and defence. First, most operations in secure multiparty monitoring (distributed top-N, attribution-free blacklisting, etc.) can be compromised by the fact that each participating domain knows its own input, and may be able in infer information about the inputs of other domains from the resulting output, especially if the number of participating organizations is low or there is limited diversity among them. Second, even with recent developments improving the scalability of SMC techniques, care shall be taken in the size of the input data sets and number or peers; therefore, such techniques are primarily applicable to the computation of aggregates from pre-aggregated or otherwise pre-processed data, and as such should be considered for deployment in parallel with other techniques outlined in the present document, depending on the application.

Finally, other cryptographic approaches appear to be promising for application in a cross-domain, multi-authority, network monitoring scenario. Scalable techniques to share encrypted data, meanwhile conditioning the decryption to the occurrence of a same monitoring event across a threshold number of different domains have for instance been addressed in [i.41], thus providing a preliminary response to a specific challenge raised in [i.42]: "design efficient distributed protocols and similarity metrics for network security data to ensure that each contributor only reveals the data if a threshold number of other participants are ready to reveal similar data". Another area of research which may have important impact in the cross-domain sharing of security data is that of attribute-based encryption, pioneered in [i.43] and [i.44] and designed as a fully decentralized scheme (hence more appropriate to a multi-operator scenario) in [i.45]. However, so far, such techniques have not yet challenged with large volumes of data.

8 Conclusion

In the present document we have identified some of the major security and privacy requirements for information sharing in collaborative cross-domain network monitoring. Clause 4 presented scenarios showing the potential of data sharing between different domains. Clause 5 presented some existing sharing solutions. The identified requirements for sharing monitoring data across domains are identified in clause 6. Finally, clause 7 points out gaps in existing solutions and proposes first ideas in research for extensions in order to fulfil the requirements.

Annex A (informative): Authors & contributors

The following people have contributed to the present document:

Rapporteur:
Jens-Matthias Bohli, NEC Europe Ltd.

Other contributors:
Giuseppe Bianchi, CNIT

Alvaro Armenteros, Telefónica I+D

Brian Trammell, ETH Zurich

Georgios V. Lioudakis, ICCS

Eugenia I. Papagiannakopoulou, ICCS

Maria N. Koukovini, ICCS

History

Document history		
V1.1.1	September 2012	Publication