

---

## - Bachelor/Master Thesis -

### Efficient Storage of Network Flow Data

---

<b>CASED</b>	<p>In CASED (Center for Advanced Security Research Darmstadt) collaborate the Technische Universität Darmstadt, Fraunhofer Institute for Secure Information Technology and the University of Applied Sciences Darmstadt in the fast developing field of IT Security. In a unique cooperation, which combines different areas of expertise from these renowned institutions, progressive IT security solutions are researched, developed and implemented into industrial economy: CASED brings together computer scientists, engineers, physicists, legal experts and business economists. Read more on <a href="http://www.cased.de">www.cased.de</a>.</p>						
<b>Motivation</b>	<p>Network flow data serve as an abstract and highly aggregated description of unidirectional communication flows within IP-based networks. For a specific period of time, such flows are uniquely defined by the 5-tuple consisting of (source IP address, destination IP address, source TCP/UDP port number, destination TCP/UDP port number, layer 4 protocol number) and are enhanced with information on the intensity of the flow (e.g. number of bytes and packets per second). Flow data are exported by active network devices (e.g. routers, switches) in different formats and versions (e.g. NetFlow, IPFIX) and, among other purposes, are used for monitoring, billing, accounting and anomaly detection.</p>						
<b>Task</b>	<p>While usage of network flow data leads to highly reduced data sets in contrast to raw packet captures, in high-speed networks many flow records per second are exported and have to be stored for varying periods of time and are being retrieved using varying characteristics, depending on the area of application. In this research work, therefore, efficient and flexible strategies to network flow data storage and indexing shall be analyzed and a prototypical solution shall be developed. The prototypical solution shall be able to harmonize different flow formats and versions. Optionally flow record clustering and correlation can be discussed.</p>						
<b>Requirements</b>	<ul style="list-style-type: none"><li>• High motivation, creativity and ability to work independently</li><li>• Good communication skills</li><li>• Good programming skills (e.g. C, Perl, Ruby)</li><li>• Good knowledge of common internet protocols</li><li>• Knowledge of Linux operating system is a plus</li><li>• Very good knowledge of the German or English language</li></ul>						
<b>Start date</b>	Immediately						
<b>Contact</b>	<table><tr><td><b>Sebastian Abt</b></td><td>CASED</td></tr><tr><td><a href="mailto:sebastian.abt@h-da.de">sebastian.abt@h-da.de</a></td><td>Mornwegstraße 32</td></tr><tr><td>06151.16-8416</td><td>64293 Darmstadt</td></tr></table>	<b>Sebastian Abt</b>	CASED	<a href="mailto:sebastian.abt@h-da.de">sebastian.abt@h-da.de</a>	Mornwegstraße 32	06151.16-8416	64293 Darmstadt
<b>Sebastian Abt</b>	CASED						
<a href="mailto:sebastian.abt@h-da.de">sebastian.abt@h-da.de</a>	Mornwegstraße 32						
06151.16-8416	64293 Darmstadt						