Max-Planck-Institut
für Informatik

-- FOR PUBLIC RELEASE --

Confidential information has been redacted in this version of the document.

Report on the Security State of Networks of Max-Planck Institutes

Findings and Recommendations

Author:
Dr.-Ing. Tobias Fiebig, MPI-INF
(tfiebig@mpi-inf.mpg.de)

October 6, 2023
https://doi.org/10.17617/2.3532055

Disclaimer

The evaluation presented in this document has been requested by the CISO of the Max-Planck Society, Guntram Rupp, and authorized by VP Prof. Dr. Klaus Blaum. Findings in this document may already have been addressed between the time of scanning and the publication of this report. For any questions or comments you can contact the CISO of the MPG as indicated in the contact details below.

Any opinions, conclusions, and recommendations expressed in this document are those of the author, and do not necessarily reflect those of the Max-Planck Society, any of its institutes, or personnel involved in requesting this report. The analysis in this document is based on active measurements of digital infrastructure associated with the Max-Planck Society[redacted].

Confidential information in this document has been redacted and it was cleared for public release.

Copyright

Metadata

Publication Date:
October 6, 2023

DOI:
https://doi.org/10.17617/2.3532055

Contact

CISO:
Guntram Rupp,
Informationssicherheitsbeauftragter der MPG

Max-Planck-Gesellschaft
Hofgartenstrasse 8
80084 München

Email: guntram.rupp@gv.mpg.de
Phone: +49 89 2108-1317

Report Author:
Dr.-Ing. Tobias Fiebig,
Senior Researcher

Forschungsgruppe Internet Architecture (INET)
Max-Planck-Institut für Informatik
Campus E14
66123 Saarbrücken

Email: tfiebig@mpi-inf.mpg.de
Phone: +49 681 9325-3527

Table of Contents

Executive Summary

This report documents and analyzes the findings of a continuous port scan of networks attributed to the MPG between 2023-04-18T01:00+00:00 and 2023-05-01T04:12+00:00. It was requested by the CISO of the MPG in the beginning of 2023 in order to identify potential vulnerabilities in the MPG's networked infrastructure and to identify structural challenges for IT security in institutes of the MPG in response to a report received between the end of 2022 and the beginning of 2023 by the president of the MPG, drawing an alarming picture of the MPG's IT security state, presenting findings in excess of 10000 events.

Objectives: The main objectives of this survey were:

- Identify security issues that require immediate or immanent mitigation due to a high likelihood of causing events or incidents.
- Assess the overall security posture of the MPG, focusing on structural and organizational shortcomings.
- Identify action items and strategies to sustainably improve the security posture of the organization.
- Evaluate findings of this internal investigation against the information that was provided by the [redacted] project.

Results: In our scans, we find a plethora of issues of varying severity. Among 3541 addresses found to expose services, 1997 (56.40%) hosts present findings that require mitigation or further investigation:

1514 (75.81%/1997, 42.76%/3541) *Low:* Issues that need verification, or require structural activity but are unlikely to cause harm.

378 (18.93%/1997, 10.67%/3541) *Medium:* Violations of best practices, hardening, and issues unlikely yet plausible to cause incidents.

83 (4.16%/1997, 2.34%/3541) *High:* Misconfigurations, exposure of data, etc., carrying a possibility of causing incidents or lateral movement.

22 (1.10%/1997, 0.62%/3541) *Critical:* Issues that require immediate attention and/or verification as they may lead to incidents.

Clustering our findings, we identify organizational root-causes in the nature of the digital infrastructure in the MPG as a research organization, proliferating the emergence of these issue-clusters:

- Unique requirements of research infrastructure and researchers as users.
- Employee churn among scientific employees.
- Use of enterprise toolchains and infrastructure that is not tailored to the organization's requirements.
- Challenges in capability retention in digital infrastructure operations.

Especially the characteristics as a research network create a unique environment that is distinct from standard enterprise networks, restricting the applicability of standard enterprise practices to improve IT security. In fact, we find two distinct events that may be related to a recent activity in terms of top-down security management, i.e., security issues related to 2FA implementations. While we do not establish a causal relationship, such a relationship would be a viable explanation following prior work on security

misconfigurations. However, we also find that the high degree of decentralization in terms of operation is a major asset for the organization's security status, as it makes lateral movement after the compromise of an individual institute less likely.

Recommendations: Based on our analysis, we suggest to improve IT security within the MPG by addressing root-causes, mostly focusing on bottom-up governance[1] and organizational improvements. A corner stone of these improvements is leveraging and expanding existing distributed expertise, improving network and service segmentation, the introduction of an operational excellence framework, and the formation of a team that actively addresses operational requirements of research infrastructure, ensuring high levels of operational excellence.

Furthermore, we explicitly caution against the introduction of top-down management of digital infrastructure. Our analysis suggests that such an approach–especially when leveraging centralization and security-by-control as common in enterprise scenarios–would create high-impact high-likelihood risks for the society's security posture, by increasing issues of orphaned systems and Shadow IT, contributing to personnel churn and capability loss, and breaking the effective decentralization. The issue of breaking decentralization is crucial, as it improves attackers' abilities for lateral movement, i.e., a full compromise, while proliferating security issues.

Comparison to the [redacted] report: Finally, we compare our results to the report by the [redacted] project received in early 2023. We find that the [redacted] report neither documents the used methodology nor does it present an analysis of the provided raw data. Based on the shared data, we infer that the report utilizes an automated methodology, focusing on pre-determined web related vulnerabilities, leaving, e.g., misconfiguration and individual instances of issues out of scope. Web related issues are scored based on version matching, prompting high CVSS scores for issues without reported manual verification of severity, while underestimating security issues outside this narrow perspective. Hence, the used methodology is unsuited to assess the security posture in non-standardized infrastructures, as found in research networks like those of the MPG.

Comparing our results, we find that the [redacted] report failed to identify the bulk of our critical and high-severity findings in its 'action items'. We also find that the report includes port-scans that would have allowed the identification of the critical/high severity issues we identified, if they had been analyzed. Furthermore, despite lacking ground-truth on the utilized scope, we find indications in the provided data that the [redacted] report included foreign entities in its vulnerability assessment.

Finally, in conjunction with our observations on automated evaluation and a missing analysis, we note that the presentation and structure of the report carry the risk of harming the security posture of an organization, mainly due to socio-organizational effects in the context of operator visibility, inner-organizational trust, 'just culture', and notification fatigue.

We conclude that, while likely well intentioned, the report holds risks for causing harm (out-of-scope scans, organizational effects), takes a narrow and automated perspective that is methodologically unsuited for investigating the security posture of research infrastructure (web-focus, automated analysis). The report requires extensive follow-up work to assess the severity of findings, which have to be manually assessed and validated, without providing guidance on that process.

[1] Governance here does not mean a strict top-down approach as it is often understood. Instead, we refer to bottom-up governance, i.e., the process of nurturing emergent (self) organization and enabling self-improvement in an organization, see also D2.1 of CS4E [16].

1 Introduction

This report documents and analyzes the findings of a continuous port scan of networks attributed to the MPG between 2023-04-18T01:00+00:00 and 2023-05-01T04:12+00:00. It was requested by the CISO of the MPG in the beginning of 2023 in order to identify potential vulnerabilities in the MPG's networked infrastructure and to identify structural challenges for IT security in institutes of the MPG.

1.1 Initial Situation

The MPG is a German NGO directly founded by the federal government and several state governments consisting of 85 different research centers. With over 24,000 employees and an annual budget in excess of 2.5b Euros it is one of Germany's largest independent research organizations. The organizational structure of the MPG relies on a relatively high autonomy of independent research centers, with the standard organizational form being that of an institute.

Institutes regularly have dedicated digital infrastructure and their own mostly independent IT staff with a direct chain of command below the directorate of the corresponding institute. Various services offered across institutes, for example, a central documentation and knowledge database system [2], are handled via a federated OpenID/SAML based authentication system, with authentication and user management still remaining within individual institutes. This leads to the challenge that independent institutes form independent infrastructure, for which they are responsible, while security incidents will regularly be attributed to the society as a whole.

In between the end of 2022 and beginning of 2023, the president of the MPG received a report from the [redacted] project[3], which drew an alarming picture of the MPG's IT security state, presenting findings in excess of 10000 events. However, a discussion of provided data within the distributed organizational structure of the MPG and individual spot-checks open several questions with regard to the reliability of the report. Furthermore, several departments found the presentation of the report to inhibit actionability.

1.2 Goals

Based on the initial situation, the objective of the study was to:

- Identify security issues that require immediate or immanent mitigation due to a high likelihood of causing events or incidents.
- Assess the overall security posture of the MPG, focusing on structural and organizational shortcomings.
- Identify action items and strategies to sustainably improve the security posture of the organization.
- Evaluate findings of this internal investigation against the information that was provided by the [redacted] project.

[2] https://max.mpg.de/ or services offered by MPG's library.

[3] [redacted]

Table 1: Netblocks included in scans.

ASN	Prefix	Institute/Comment	ASN	Prefix	Institute/Comment
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

1.3 Scope

The scope of this survey was limited to active measurements of networks attributed to the MPG and independent subsidiaries of the MPG, or in which the MPG holds a major stake. This attribution could take place via reverse DNS entries, existing thread intelligence feed indicating a high number of MPG related systems in a netblock, and RIR/WHOIS data. Furthermore, all networks announced by ASes owned by the MPG were within scope. Systems operated by the MPG but not hosted in an MPG related netblock, including systems run under MPG related domain names, for example, Software-as-a-Serivce offerings were not within scope.

Please note that the list has been compiled from available public information including RIR databases, names under various MPG domains, and internal documentation[4]. Additionally, blocks were reduced to sub-blocks for which active systems could be observed. As a result, while generally tight, the scope partially covered some adjacent networks of other research organizations[5], or did not cover the full network allocated to a sub-organization of the MPG[6]. See Table 1 for a list of networks within scope.

[4] [redacted]

[5] See, for example, [redacted].

[6] See, for example, [redacted] in [redacted].

Rules of engagement for bulk activity were limited to non intrusive techniques, specifically:

- ICMP echo requests.
- SYN requests.
- Establishment of connections with open TCP and UDP ports to obtain banners and default unauthenticated data supplied.
- Retrieval of TLS certificates for TLS enabled endpoints under all known names for the associated IP address in SNI in addition to `localhost` and no explicit SNI indicator being set.
- Retrieval of HTTP index pages for ports exposing an HTTP related protocol, using both HTTP and HTTPS per port and all known names for the IP address, in addition to `localhost` and a HTTP host being set.

The explicit identification of software versions, even though data on this may have partially been collected in the above process was not in scope.

In case of events that may indicate incidents, events were followed up manually to ascertain the seriousness of the incident by verifying exploitability, e.g., by testing RCE with non-intrusive commands[7]. Please see Section 2 for a detailed description of the methodology, also providing the reasoning behind restrictions on the evaluation of software versions and how limitations in the dataset and methodology should be considered in the interpretation of the results. Critical vulnerabilities have been communicated to the CISO prior to delivery of this report, please see Section 2.8 for details.

Document Structure

The remainder of this document is structured as follows: We introduce the methodology used in the measurements this report is based on in Section 2, where we also discuss limitations and document the reasoning behind specific scope limitations and methodological choices. In Section 3, we present an overview of our findings and derive clusters of comparable/similar vulnerabilities. Next, we interpret our results in Section 4, where we also distinguish the situation in the MPG as a research institution from that of classical enterprise environments. Our analysis then allows us to derive recommendations for the MPG, presented in Section 5. Thereafter, we compare this study in terms of scope, methodology, findings, recommendations, and analysis to the [redacted] report the MPG received earlier this year in Section 6. Finally, we briefly summarize the core findings of this report and key action-items in Section 7.

[7] Non-intrusive commands would, e.g., be `ls` or the creation of an empty file in a non-application-critical repository.

2 Methodology

Here, we document the methodology used to execute the scans of networks held by MPG affiliated organizations. This includes the used infrastructure, the software stack used for the measurements, the composition of the final scan toolchain, and the scanning schedule we executed. Furthermore, we will discuss limitations and document interactions with institutes due to either abuse complaints, or if incidents occurred that necessitated an early notification and mitigation.

2.1 Used Infrastructure

Scans originated from [redacted] (IPv4) and [redacted] (IPv6) within [redacted], i.e., fully external to any infrastructure hosted by the MPG or any of the Germany research networks, most notably DFN with AS680. This, in contrast to using, e.g., a network of the MPG, was done to ensure potentially existing ACLs do not influence the results of the evaluation. [redacted]

For the allocated networks, corresponding IRR/whois objects were created, documenting the purpose of the ongoing measurements, referring to an abuse address, contact information of [redacted], and a reference to a website providing further information. See the listing below for the IRR entry of [redacted]; The setup for [redacted] mirrored this.

```
% This is the RIPE Database query service.
% The objects are in RPSL format.
%
% The RIPE Database is subject to Terms and Conditions.
% See http://www.ripe.net/db/support/db-terms-conditions.pdf

% Note: this output has been filtered.
%       To receive output for a database update, use the ”-B” flag.

% Information related to ’XXX.XXX.XXX.XXX - XXX.XXX.XXX.XXX’

% Abuse contact for ’XXX.XXX.XXX.XXX - XXX.XXX.XXX.XXX’ is ’x@example.com’

inetnum:        XXX.XXX.XXX.XXX - XXX.XXX.XXX.XXX
netname:        FOR-SCANNING-AT-THE-MAX-PLANCK-SOCIETY
country:        XX
descr:          You can contact the responsible researcher at: r@example.com
descr:          We are scanning to assess the MPG’s security state.
descr:          Find further information at https://example.com/
admin-c:        XXXXXX-RIPE
tech-c:         XXXXXX-RIPE
abuse-c:        XXXXXX-RIPE
status:         ASSIGNED PA
mnt-by:         XXXXXXX-MNT
created:        2023-XX-XXTXX:XX:XXZ
last-modified:  2023-XX-XXTXX:XX:XXZ
source:         RIPE

role:           XXXXXX-RIPE
address:        XXXXXXXXXXXXXXX, XXXXXXXXXXXXXXx, XXXXXXXXX, XXXXXX
phone:          +XXXXXXXXXXXXXXXX
nic-hdl:        XXXXXX-RIPE
mnt-by:         XXXXXXX-MNT
created:        XXXX-XX-XXTXX:XX:XXZ
last-modified:  XXXX-XX-XXTXX:XX:XXZ
source:         RIPE # Filtered

% Information related to ’XXX.XXX.XXX.XXX/XXASXXXXXX’

route:          XXX.XXX.XXX.XXX/XX
origin:         ASXXXXXX
mnt-by:         XXXXXXX-MNT
created:        XXXX-XX-XXTXX:XX:XXZ
last-modified:  XXXX-XX-XXTXX:XX:XXZ
source:         RIPE

% This query was served by the RIPE Database Query Service version 1.106.1 (ABERDEEN)
```

Within that network segment, 13 virtual machines were created, see Table 2, 12 to execute these scans, with an additional host in place to orchestrate scanning as per the planned schedule, see Section 2.4. Each virtual machine received matching reverse and forward DNS under the domain [redacted], prefixed with an identifier, e.g., [redacted]. On all machines a webserver was installed, using [redacted], redirecting requests to the base domain ([redacted]). On that site, a webpage provided information and additional contact details in relation to the ongoing scans, see the listing below.

Please note that the list of scanned networks has been removed in this report to preserve space; For this content, please see Table 1 in Section 1.

```
Description
These scans are executed to check the current configuration and
security state of systems belonging to the Max-Planck Society.
Please see ’networks in scope’ for a list of networks we are
scanning.

If these scans cause operational issues, please contact Tobias
Fiebig at tfiebig@mpi-inf.mpg.de or via phone: +xx xxx xxxx.

Results of these scans will be made available to you as soon as
the evaluation has been concluded.

For further information, you can also reach out to the CISO of
the MPG: Guntram Rupp, guntram.rupp@gv.mpg.de

Executed measurements
We are executing the following daily measurements:


Hourly: ICMP connect scan; This measurement is in place to identify
        networks mostly used for clients, which should be firewalled.


07:00 (UTC): Full scan (ICMP, SYN Port 1:12000 TCP, Port 53,80,111,
             123,137,138,139,443,11211 UDP, banner grabbing on open
             ports, curl on all known hostnames for HTTP ports,
             certificate retrieval for TLS/STARTTLS ports with all
             known names for SNI.


13:00 (UTC): Full scan (ICMP, SYN Port 1:12000 TCP, Port 53,80,111,
             123,137,138,139,443,11211 UDP, banner grabbing on open
             ports, curl on all known hostnames for HTTP ports,
             certificate retrieval for TLS/STARTTLS ports with all
             known names for SNI.


19:00 (UTC): Full scan (ICMP, SYN Port 1:12000 TCP, Port 53,80,111,
             123,137,138,139,443,11211 UDP, banner grabbing on open
             ports, curl on all known hostnames for HTTP ports,
             certificate retrieval for TLS/STARTTLS ports with all
             known names for SNI.


01:00 (UTC): Full scan (ICMP, SYN Port 1:12000 TCP, Port 53,80,111,
             123,137,138,139,443,11211 UDP, banner grabbing on open
             ports, curl on all known hostnames for HTTP ports,
             certificate retrieval for TLS/STARTTLS ports with all
             known names for SNI.

These scans are in place to identify exposed services in network
segments, identify orphaned or forgotten systems, and identify
unfirewalled access networks.

Networks in scope
The following networks are in scope for these scans. This list has
been generated from DNS data and internal documentation. Please
reach out to tfiebig@mpi-inf.mpg.de if your address or netblocks
have been falsely attributed to the MPG and included in this list.
Please see the next section for explicitly excluded more specifics.

{
  ”ASXXXXXX”: [
      ”XXX.XXX.XXX.XXX/XX”,
      ...
  ]
}

Networks excluded
The following networks have been manually excluded upon request, if no
more specific network is included above:
{
  ”ASXXXXXX”: [
    ”XXX.XXX.XXX.XXX/XX”,
    ...
  ]
}
```

Table 2: Overview of used measurement system.

Hostname/rDNS	IPv4	IPv6	Threads	Memory	OS	Purpose
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]

2.2 Used Software

Here, we briefly list the major upstream software components used in the final toolchain used for scanning the MPG's infrastructure. All components were integrated using bash scripting and GNU parallel

parallel 20210822+ds-2 GNU parallel is a framework for running compute jobs in parallel, also allowing creation of parallel jobs across multiple machines. In the scanning campaign, GNU parallel was used on a shared NFS storage, provided by [redacted].

```
% parallel --version
GNU parallel 20210822
Copyright (C) 2007-2021 Ole Tange, http://ole.tange.dk and Free Software
Foundation, Inc.
License GPLv3+: GNU GPL version 3 or later <https://gnu.org/licenses/gpl.html>
This is free software: you are free to change and redistribute it.
GNU parallel comes with no warranty.

Web site: https://www.gnu.org/software/parallel

When using programs that use GNU Parallel to process data for publication
please cite as described in the manpage.
```

nmap 7.91+dfsg1+really7.80+dfsg1-2ubuntu0.1 Nmap is one of *the* traditional port scanning utilities available, and integrates a multitude of additional banner collection and interaction modules. We selected nmap over a more traditional high bandwidth scanning framework as, e.g., zMap [8][8], as we wanted to limit the total packets per second to below 200k pps, see Section 2.3, and nmap allowed fine-grained control of scan settings within the utilized and paralellized toolchain, spreading load equally across networks of the MPG. Hence, limitations of nmap in comparison to zMap were not relevant for our use-case.

```
% nmap --version
Nmap version 7.80 ( https://nmap.org )
Platform: x86_64-pc-linux-gnu
Compiled with: liblua-5.3.6 openssl-3.0.2 nmap-libssh2-1.8.2 libz-1.2.11 \
               libpcre-8.39 libpcap-1.10.1 nmap-libdnet-1.12 ipv6
Compiled without:
Available nsock engines: epoll poll select
```

curl 7.81.0-1ubuntu1.10 Curl is a standard library and framework for executing protocol requests, especially for HTTP(S)2. We use curl to obtain index pages for discovered open ports that indicated an HTTP related protocol being in use in prior nmap scans of a host.

[8] zMap obtains benefits in throughput for scans by removing statekeeping from scanning.

```
% curl --version
curl 7.81.0 (x86_64-pc-linux-gnu) libcurl/7.81.0 OpenSSL/3.0.2 zlib/1.2.11 \
    brotli/1.0.9 zstd/1.4.8 libidn2/2.3.2 libpsl/0.21.0 (+libidn2/2.3.2) \
    libssh/0.9.6/openssl/zlib nghttp2/1.43.0 librtmp/2.3 OpenLDAP/2.5.14
Release-Date: 2022-01-05
Protocols: dict file ftp ftps gopher gophers http https imap imaps ldap ldaps \
           mqtt pop3 pop3s rtmp rtsp scp sftp smb smbs smtp smtps telnet tftp
Features: alt-svc AsynchDNS brotli GSS-API HSTS HTTP2 HTTPS-proxy IDN IPv6 \
          Kerberos Largefile libz NTLM NTLM_WB PSL SPNEGO SSL TLS-SRP \
          UnixSockets zstd
```

OpenSSL 3.0.2-0ubuntu1.9 OpenSSL is a standard TLS library for Linux systems. We use OpenSSL's `s_client` implementation to retrieve certificates of remote systems if either a TLS tunnel has been detected[9] or using the protocol specific STARTTLS invocation if a protocol commonly featuring STARTTLS has been found by nmap[10].

```
% openssl version
OpenSSL 3.0.2 15 Mar 2022 (Library: OpenSSL 3.0.2 15 Mar 2022)
```

2.3 Scanner Composition and Configuration

Scanning follows a multi-stage process, see Figure 1. We selected this architecture, as it gives us fine-grained control over the number of in-flight scanned hosts, while also allowing us to distribute the scanning load comparatively evenly across hosts of the MPG[11], with larger networks naturally being it more frequently. As due dilligence, prior to executing scans against all systems within the MPG, we executed several test-runs with increasing pps counts against [redacted] in coordination with the network operators of that institute. [redacted] the network was able to handle inbound loads of over 200k PPS when we executed the toolchain outlined below, saturating all workers simultanously with /28s from [redacted]. Hence, we concluded that–when spread over all networks of the MPG–a load of less than 200k pps should be feasible. For exceptions we encountered in practice, please see Section 2.8.

Stage 1: In the first stage, we take the input networks, see Table 1, and split them by /28, i.e., into blocks of at most 16 IPv4 addresses. If an input network is smaller than a /28[12], we add it to the list of /28s as-is.

Subsequent stages are then applied to each entry of our scan list. For this, the following stages are executed in parallel, with 32 jobs per worker node, i.e., up to 12×32=384 active jobs in parallel.

Stage 2: For each network scanned by a node, we first execute an ICMP echo request scan. For this, up to ten (nmap default) ICMP echo requests are sent to each IPv4 address. If at least one ICMP echo reply is received, the host is flagged as reachable. Furthermore, we obtain the reverse DNS entry, if present, for each scanned IP address at this stage. Please note that this limits our visibility, see Section 2.7.

Stage 3: For each name identified via reverse DNS in Stage 2, we try to obtain an IPv6 address by executing a AAAA DNS request. If we are able to obtain an IPv6 address, we add it to the subsequent scan stages for this host.

[9] As found for, e.g., HTTPS.

[10] For example, common in SMTP.

[11] For example, the [redacted] holds a full /16.

[12] See, for example, [redacted] .

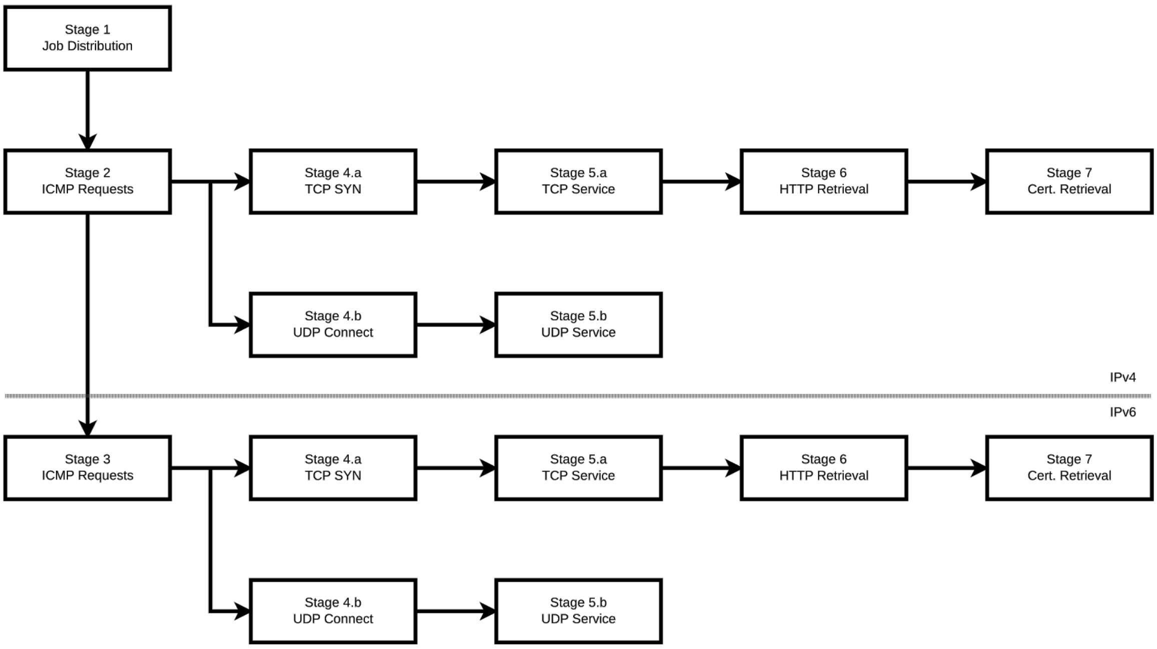


Figure 1: Overview of the individual stages in the scanning process.

Stage 4.a (IPv4|IPv6): For each reachable address, we execute a TCP SYN scan of ports 1-12,000 with up to 2 retires, at least 16 requests in parallel, at a rate of 500 packets per second. Given that this process is sequential per address per /28, and the maximum number of inflight jobs being 384, this means that the upper bound of generated pps is 192k, i.e., below 200k, while in practice this value is usually not reached, as it would require all workers to be in the TCP SYN stage at the same time. We decided to limit the port range for SYN scans to the first 12,000 ports instead of scanning ports 1-65,535 to reduce the scan-time for all MPG networks without increasing the pps load[13]. This again limits our visibility. Again, please see Section 2.7 for a discussion on this limitation.

Stage 5.a (IPv4|IPv6): For the list of ports determined to be open, we then execute a full service scan with nmap[14]. This includes OS detection based on timings and information conveied by running services, service detection, and banner-grabbing. Please note that we did not execute service discovery for hosts for which we found the majority of ports to be reacting to SYNs due to tcpwrapper being used.

Stage 4.b (IPv4|IPv6): In addition to a TCP SYN scan, we also try to detect open UDP ports. However, given that UDP is a stateless protocol, the absence of a reply to a sent packet does not necessarily indicate that no service is listening on a given UDP socket. Hence, while ICMP messages indicating that a port is closed provide conclusive information about the state of said port, and protocol compliant replies indicate that a port is open, the absence of a reply does not carry any information as to whether a service is listening or not. For that reason, we limited our UDP discovery to common and commonly misconfigured protocols and services, specifically:

[13] Using, e.g., zMap here would not have improved runtime, as that improvement in runtime would have increased the pps load.

[14] `nmap -A`

53 DNS

80 Quick[15]

111 RPC mapper

123 NTP

137 Netbios

138 Netbios

139 Netbios

443 Quick, as an evolution of HTTP

11211 memcached in-memory key-value store

Concerning the limitations this introduces, please see Section 2.7.

Stage 5.b (IPv4|IPv6): As with TCP, we then perform full service discovery for ports detected to be open.

Stage 6 (IPv4|IPv6): If nmap reports an HTTP like protocol being used for an open port, we then retrieve the index page presented by the server. We do this, regardless of the port being initially detected as plain HTTP or HTTPS, for HTTP and HTTPS, as some servers actually implement plain and HTTPS enabled connections on the same port[16]. In addition to retrieving the index page without providing an HTTP Host header, essentially accessing the default vhost on the remote system, we also issue a request for each FQDN we encountered in relation to an IP address, as well as localhost. Especially the latter is important, as in certain cases a server may have bound globally, while running an internal service on a virtual host explicitly named 'localhost', under the assumption that the host is only reachable with requests from the machine itself. However, as the vhost is usually determined based on the HTTP Host header sent by a client, this is not the case, and we may thereby evade ACLs configured on a remote system.

Stage 7 (IPv4|IPv6): Finnaly, if nmap indicates that a port uses TLS in tunnel mode, or the protocol supports STARTTLS (smtp, pop3, imap, ftp, xmpp, xmpp-server, irc, postgres, irc, mysql, lmtp, nntp, sieve, ldap), we retrieve the certificate and certificate chain related to that certificate from the remote system. Given that for many protocols SNI has been implemented, we again do this for an empty SNI indicator (default certificate), 'localhost', and all names known to us from the scan itself of utilized databases.

2.4 Scanning Schedule

We decided to execute repeated/regular scans of the MPG infrastructure over two weeks using the aforementioned toolchain. The reasoning behind this, in contrast to executing scans that cover, e.g., a larger port range or more UDP ports, was that we assumed, due to the relative abundance of public IPv4 addresses in research networks[17] we assumed that several research and work related components, e.g., lab equipment, embedded devices, or workstations, only become available during limited timeframes, despite being generally publicly reachable. As we assume that such devices are often not designed to be run fully exposed to the Internet, we prioritized regular scanning.

Ultimately, considering that a full scan through all stages outlined in Section 2.3 takes around two hours to complete, we implemented the following schedule:

[15] Unlikely, as quick usually runs on port 443

[16] This can, for example, be implemented using the stream server feature of NGINX, see `http://nginx.org/en/docs/stream/ngx_stream_core_module.html`

[17] While several MPG institutes and similar research organizations have an abundance of IPv4 addresses allocated to them, the IPv4 run-out [25] lets, e.g., new ISPs have only a fraction of that. Freedom Internet b.v., for example, only has 1/5th (54 /24) of the IPv4 addresses of several institutes holding a /16 (255 /24) available.

ICMP request scan (Stage 1 & 2 only): Hourly

Full scan: 07:00UTC, 13:00UTC, 19:00UTC, 01:00UTC

Executing this schedule over the duration of two weeks also ensures that we capture weekend/holiday vs. weekday activity. Furthermore, as the order of /28 is randomized in Stage 1, it also means that each host is scanned at various points throughout the day.

2.5 Levels of Criticality

Throughout this report, we will be using four different levels of criticality in relation to detected events. Specifically we distinguish, from critical to low, between:

Critical: *Immediate Mitigation:* This category includes events that are at least incidents. Furthermore, these incidents must either allow unauthorized access to a system, allowing further access to privileged information[18]. Similarly, cases where PII of third parties is revealed, or physical harm is enabled fall in this category.

High: *Critical Findings:* Critical issues must be mitigated as soon as possible. They usually pose a high likelyhood of leading to an incident but do not signify an immediately ongoing incident; This also includes cases of non-specific threats that are easily mitigated.

Medium: *Urgent Findings:* These issue are those that pose no immediate threat but should be addressed in the near future, and might benefit lateral movement, or are indicative of limitations in operational procedures.

Low: *Chores:* Findings of low criticality that should be addressed/picked up in regular maintenance. This includes issues like, for example, updating out-dated libraries outside of security support, for which no vulnerabilities are known yet.

This classification has been created to be more accessible than standard frameworks for event criticality[19].

2.6 Assessment of Criticality

To assess the criticality, we initially employed an open coding process assigning tags and criticality levels based on the definitions in Section 2.5. All 3541 hosts that had open ports during the duration of the scans were manually inspected by a first analyst[20] and received tags if an issue was observed. Subsequently, all hosts with a criticality above 'Low' were re-coded by a second coder[21] to ensure reliability[22]. This second round of coding included additional manual verification, e.g., [redacted], annotated with newly introduced tags. In total, [redacted] changes were made by the second coder, where [redacted] relates to the criticality of [redacted] in [redacted] at [redacted]. Apart from that, mostly tags for [redacted] that replied to client queries were added.

2.7 Limitations

As remote scanning always has to consider trade-offs, see above, our results have limitations which have to be considered in their interpretation. Specifically:

[18] Please see the issue [redacted] in Section 2.8 for an example right at the border of this definition.

[19] See, for example, CVSS

[20] [redacted] from [redacted].

[21] [redacted] from [redacted].

[22] Additionally, [redacted] systems were excluded, as their assessment is straight-forward.

Fully external scan: Our scans were conducted from a network not affiliated with the MPG or a large research network. Hence, we might have missed instances of misconfigured services that are exposed to an unreasonably large number of networks, e.g., the whole DFN. However, given that the objective of this scan is establishing a baseline and identifying systematic issues, we consider this acceptable. Nevertheless, follow-up work, see Section 5, should address this aspect.

Limited network scope: While our scans spanned a seizable part of the MPG's networks, it did not cover all networks, while partially also including networks not used by the MPG, also see Section 2.8. This situation occurred due to unclear documentation of prefixes in public databases and our aggregation approach. Naturally, follow-up work should address the issue of underscanning. Again, as outlined above, though, limitations in visibility are not a major issue for our objective, as issue we *do* observe exist. Furthermore, given precautions taken to not overload networks by limiting overall PPS, and by quickly reacting to corresponding exclusion requests, we should not have caused unreasonable load on collateral networks included.

Limited port range for TCP: We consciously limited the port-range we scanned to the first 12,000 ports. While this covers most ports services commonly run on, it also means that we might have missed specific or uncommon services, which can be expected to exist in an organization like the MPG, especially in the context of special embedded devices. Still, as before, the major implication is that the actual security state might be worse than detected, i.e., we are still able to provide a base-line estimate.

Limited port range for UDP: We limited our UDP scans to a few specific ports. A notable, accidental, omission is port 161, SNMP. As before, this limits our visibility, and must be considered when discussing our results.

High volume scanning: Despite limiting ourselves to below 200k pps while scanning, we did not take any measures to hide our activities. We followed measurement best-practices in making our scanning infrastructure identifiable and did not, for example, use a distributed approach which would more easily evade automated detection and blocklisting systems. Given that, over the course of the scanning campaign, we were contacted by three organizations, we assume that the impact of this is limited, especially also given that we used several machines (despite being from the same /24) and ensured that requests to target networks would be spread over the duration of a scan.

Restriction to ICMP responsive hosts: By restricting our scans to ICMP responsive hosts, institutes restricting ICMP[23] might not have seen all their systems scanned. This should be addressed by future activities by conducting scans that simply assume hosts to be online, likely using an even lower pps rate over a significantly longer time frame. Nevertheless, for the study at hand, this has been a tradeoff between coverage and the objective to identify networks where hosts are end-user machines and/or not permanently online.

Limited number of retries for TCP SYN: We limited the number of retries for TCP SYN scans to two, to reduce the overall number of packets sent to MPG networks, especially given our repetition rate. This means that services that are reachable, but have high latency or where we experience packet loss on-path, might have been missed. However, given we regularly repeat our measurements, this risk is mitigated.

[23] While commonly recommended for security reasons, this is a problem technically solved by generating more pps when scanning, i.e., assuming hosts to be online, and hence falls into the category of security by obscurity.

Exclusion of tcpwrapper hosts: We excluded several hosts that run tcpwrapper and reply to TCP SYN on all ports. This, again, limits visibility. However, besides hosts in [redacted][24], only the [redacted] has a total of six, the [redacted] one, and [redacted] seven hosts showing this behavior. Hence, the overall impact of this is limited.

Software patch level not structurally evaluated: Our methodology did not specifically investigate the patch-level of hosted software. We made this decission as we a) assumed software running in the MPG to regularly be custom, and b) because fine granular vulnerability detection based on version information alone is not necessarily reliable. Instead, while version numbers can be an indication, the presented version number may deviate from the actually used version[25], or security patches might have been backported. Hence, to keep the scope of this evaluation feasible within the available personnel resources, a full-scale evaluation of patch levels was not included in our survey. Nevertheless, we used clear indicators of outdated software, e.g., software that has been discontinued for an extended period of time[26], or versions for which no long-term support edition still receiving security updates is available[27] as additional information when considering a system to be unmaintained. Follow-up work should cover this aspect in a structured manner as well, see Section 5.

2.8 Abuse Reports and Interactions with Institutes

During the execution of the scans, several events occurred that either required immediate mitigation due to the detection of an incident, were investigated as an incident, or were abuse complaints by individual institutes about the ongoing scans. To be transparent about the turn of events, we document them in this section.

Observed Incident: Compromised System at [redacted]: On 2023-04-18, a first explorative analysis of the initial full scans was conducted to verify the functionality of the running measurement toolchain. During this exploration, a host with the name [redacted] and the IPv4 address [redacted] was found exposing an HTTP server on tcp/8888, as well as NTP (udp/123), Netbios (udp/137-139), SSH (tcp/22), and RPC BIND (udp/111). The HTTP service on tcp/8888 presented a directory listing. The directory contained multiple directories indicating various versions of the same research related software, specifically [redacted][28] [redacted]

In addition, a file [redacted] could be found in the webroot. When accessed, this file provided a web-shell allowing the direct execution of system commands. The vulnerability was verified [redacted]

A notification was sent [redacted] summarizing the finding and recommending immediate mitigation, i.e., immediate incident response, including replacing the system with a new, documented and monitored, setup. [redacted] Hence, overall, mitigation was done quickly, [redacted]. [redacted]

[24] [redacted]

[25] See Section 2.2 where the actual version of nmap differs from the one reported by dpkg.

[26] For example, a system whose vendor has been out-of-business since 2008.

[27] For example, systems running and exposing Python 2.5, for which support ended in May 2011.

[28] [redacted]

Abuse Complaint: [redacted] On [redacted] requested further information about the ongoing scans, as a machine from the scan cluster [redacted] in the firewall logs for [redacted]. [redacted] inquired why the ongoing scans were showing up [redacted]. A response was provided [redacted] within 10 minutes, detailing the scanning process and highlighting that this report would provide actionable recommendations later on. Additionally, it was offered to provide an unprocessed summary of issues in the specific network ahead of time. This summary was requested. Three immediate mitigation issues were reported ahead of time, specifically:

- An FTP server allowing anonymous connections and writes while exposing local usernames[redacted].
- An industrial control system [redacted], exposing an HTTP interface without TLS enabled. Despite no weak credentials being in place, it was recommended to not have this category of systems Internet reachable.
- An outdated [redacted] installation[redacted]. As, at the time of the exchange, [redacted] is out of LTS support for nearly five years, it was recommended that this system sould be upgraded. As an alternative option, adding either OpenID or HTTP authentication [redacted] was suggested. Additionally, a monitoring script [redacted] was shared with the team to enable real-time monitoring of the service.

Besides this, several major and minor issues were reported, see Appendix II, ranging from orphaned and exposed systems to setups hosting dissemination sites for published papers, with an unclear maintenance states.

Abuse Complaint: [redacted]: On [redacted] a complaint was sent [redacted] by [redacted]. The message noted that the in-scope network [redacted] contains systems hosted for non-MPG third parties. [redacted] requested that hosted system would be excluded from the scans, and provided a list of relevant networks[29]. This change request was implemented immediately, and a reply indicating compliance with the request was sent [redacted] within 10 minutes of receipt of the complaint.

Abuse Complaint: [redacted]: [redacted] a complaint [redacted] noted [redacted] the scanning network, requesting an explanation. A reply was sent [redacted] within 90 minutes, explaining the observed volume [redacted] per scan [redacted]. A follow-up [redacted] indicated that the number of requests in excess [redacted] pertained to the whole scan period. This was clarified, while also noting the projected end-date of the scans with an additional message [redacted]. No further follow-up occurred.

Investigated Event: [redacted]: During an exploration of the dataset after the conclusion of the scans, [redacted], it was noticed that the host [redacted] has an HTTP cloneable git repository in its web-root. The host presents a certificate [redacted], indicating that the host is part of the institutes' [redacted] infrastructure. Similarly, the page-title [redacted] supports this assumption. The exposed git repository allows cloning of the running application's source-code[redacted].

[29] See Table 1, which lists these exclusions.

Additionally, the `.git/config` file revealed authentication information for a deploy user [redacted].

To assess the criticality of this event, the repository was cloned from the webroot [redacted] using the found authentication information. An audit of the repository's contents indicated, that the repository does not reveal confidential information. [redacted] However, it was noted that the repository path indicates that the repository is within the namespace of an individual operator, violating best practices of locating deploy repositories in infrastructure namespaces. Similarly, authentication for deploy systems pulling remote repositories should use restricted deploy keys instead of username password combinations. [redacted]

To ascertain whether code could be committed using the found authentication information, a whitespace change was created and committed. However, the user was not authorized to push to the deploy repository. Hence, despite revealing parts of a code base and credentials [redacted] for a [redacted] system, the likelihood of this directly causing an incident is limited[30].

Based on these observations it was decided [redacted] not notify [redacted], and instead only include the event in this report as a critical finding.

Preemptive Notification: [redacted]**:** On 2023-06-11, CVE-2023-27997 was released. CVE-2023-27997 documents an unauthenticated remote code execution vulnerability in SSL VPN Gateways produced by Fortinet. Due to the high criticality of this vulnerability, the CISO was informed via mail at [redacted]. The notification included a reference to [redacted], noted that affected systems should be updated to a patched version[31] as soon as possible, and a reference to prior CVEs of similar criticality from the same vendor[32].

Furthermore, it noted potentially affected systems, see Table 3.

Table 3: Overview of notified systems.

IP	Hostname	Institute
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[30] [redacted]

[31] 7.2.5, 7.0.12, 6.4.13, 6.2.15, and 6.0.17.

[32] CVE-2018-13379, CVE-2019-11510, CVE-2022-40684, and CVE-2022-42475.

2.9 Access to Measurement Data:

Access to the collected data has been available to the main author of this report. Beyond the main author, access to the collected data has been shared with two experts withing MPI [redacted]. Specifically, [redacted], the [redacted] of MPI [redacted] received access [redacted]. Similarly, [redacted] received access [redacted]. In both cases, access was provided to mitigate the single point of failure caused by a single main investigator in case of prolonged or permanent unscheduled unavailability, and to ensure a second opinion on complex cases can be easily obtained.

Apart from that, individual institutes could receive a copy of their data if requested. This was done twice, once for the MPI [redacted], which received a full copy [redacted], and once for MPI [redacted], where a copy of data for this institute was sent [redacted].

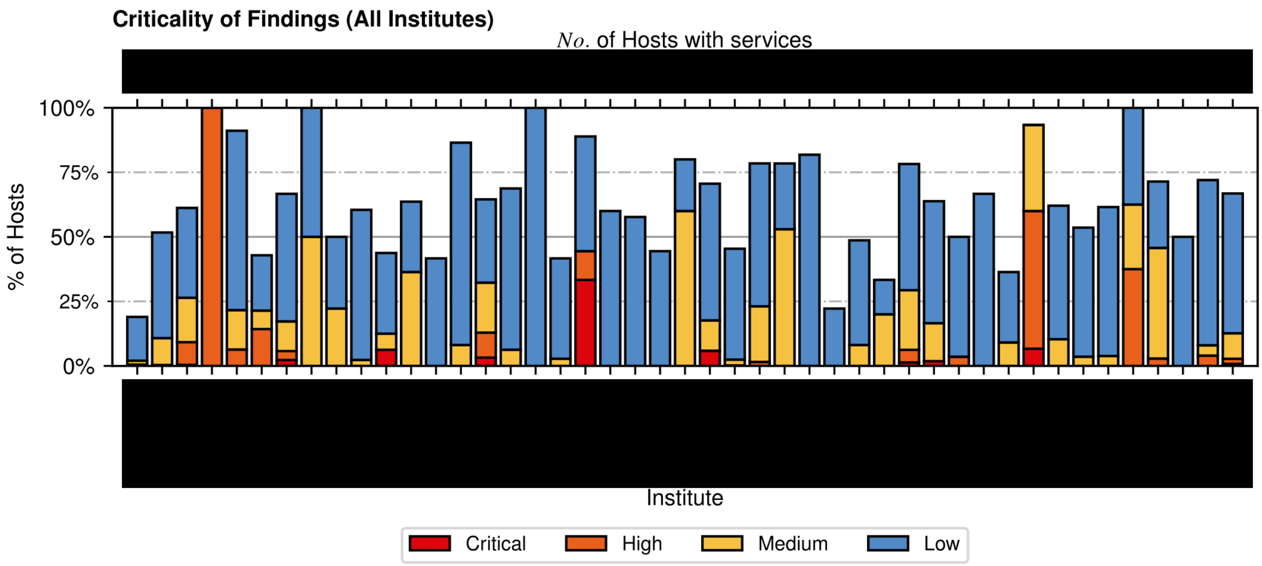


Figure 2: Overview of hosts with findings by criticality per institute. Networks without active hosts have been excluded. Please not that the number of hosts offering services highly differs per institute.

3 Findings

In this section, we give a high level aggregated overview of our findings across MPG networks[33].

3.1 Overview of MPG Networks

The nature of network and their security posture of individual institutes differs significantly. Naturally, some institutes, as for example the [redacted], contain a significantly higher number of hosts ([redacted]) than, for example, institutes with [redacted], as for example [redacted] with ([redacted]). At the same time, the share of critical findings also differs, with the [redacted] or [redacted] having mostly observations rated as low, and presenting observations for less than 50% of hosts. In contrast, the [redacted] exhibits findings considered to have medium or higher impact in half of all systems providing services, see Figure 2.

Turning towards the distribution of open ports across all institutes, see Figure 3. As can be expected, the most commonly found ports are [redacted]. Similarly, [redacted] can be frequently found to be open to the Internet. The most common ports that should at least be reconsidered are CallBook services (tcp/2000)[34] and SIP (tcp/5060), configuration ports for [redacted], and port tcp/8008 opened by Fortinet devices. For all these ports it will have to be critically evaluated whether they have to be open for the corresponding service.

If not, ports should be closed whenever they are not needed, with the following order of preference, ideally combining multiple steps:

1. If the service is not used, disable the listening daemon.
2. If the service is used internally, limit access to the port to hosts that need access

[33] For a detailed perspective on individual Institutes, please see Appendix I

[34] Technically, the mail filter configuration protocol SIEVE can also be found on this port. However, this was not the case for the concerned hosts, as we found co-located SIP related ports and no SIEVE banners.

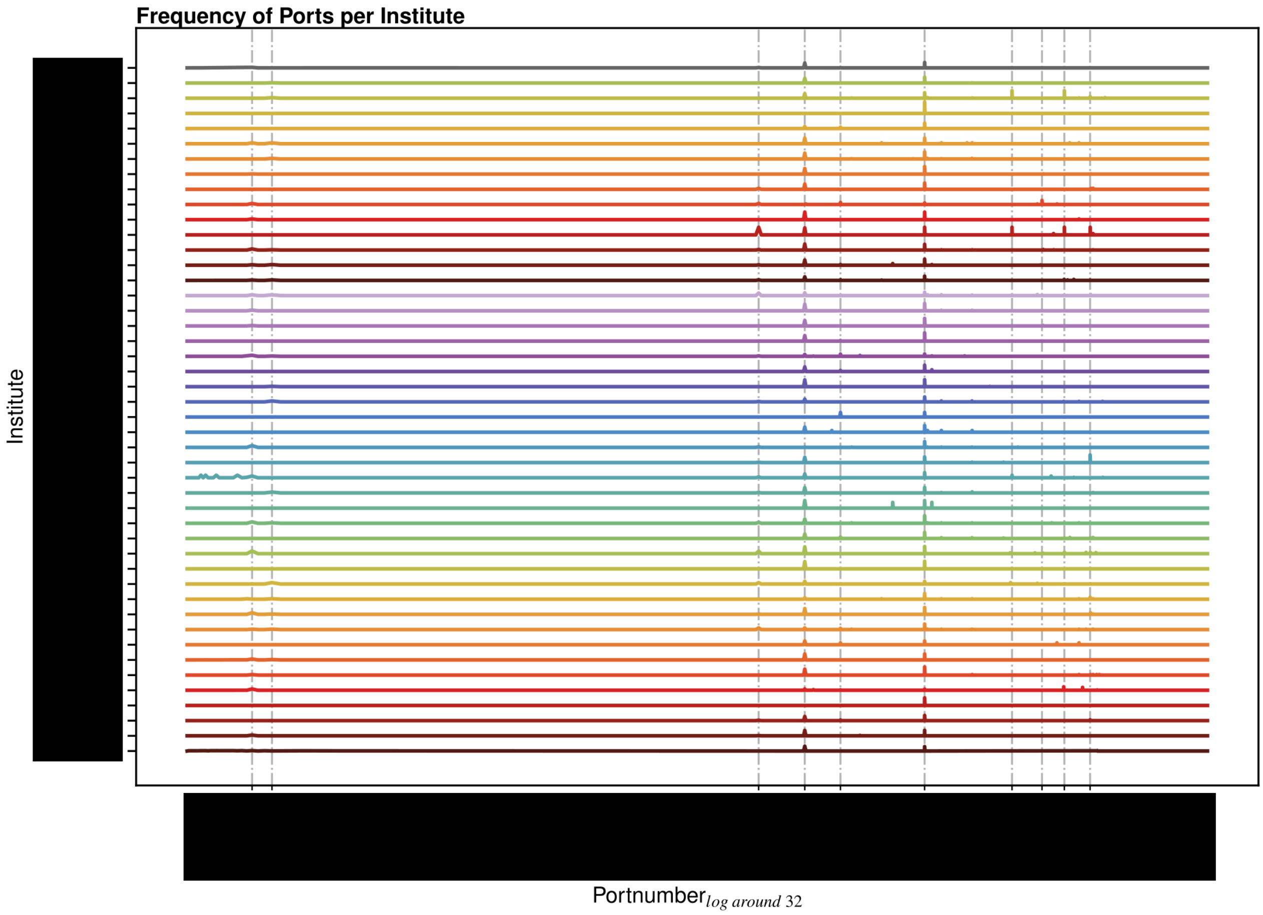


Figure 3: Overview of open ports seen per institute. The frequency is based on the number of hosts in that institute seen with the respective port being 'open'.

   a. Via a packet filter on the host, and/or,
   b. Via a packet filter at the network *segment*'s border

3. Only allow-list a known set of ports inbound at the network's demarcation points

Especially for the [redacted] related ports, their global availability indicates that the devices are not utilized in full conformance to best practices, i.e., if the device is not used as intended, or if it has been configured in a 'best effort' manner. Port [redacted] usually serves as a port to deliver block/filtering notifications for clients, especially internal when the device is used to restrict outbound connectivity to sites permitted by corporate policy. Following the points above, we would recommend against increased exposure by permitting external hosts to connect to [redacted] systems. Even though, technically, inbound connections should not pose a risk, given [redacted], see Section 2.8, this may differ in practice.

Moving to an aggregate perspective of the security state of MPIs, see Figure 4, we find that 1997/3541 hosts exposing services (56.40%) warrant further investigation. Of these 1997 hosts, 1514 (42.76% of all hosts exposing services/75.81% of hosts with findings) show a pattern which warrants further investigation, but should not lead to immanent incidents, or if it should have limited impact (low), 378 (10.67% of all hosts exposing services/18.93% of hosts with findings) have findings that warrant expedited

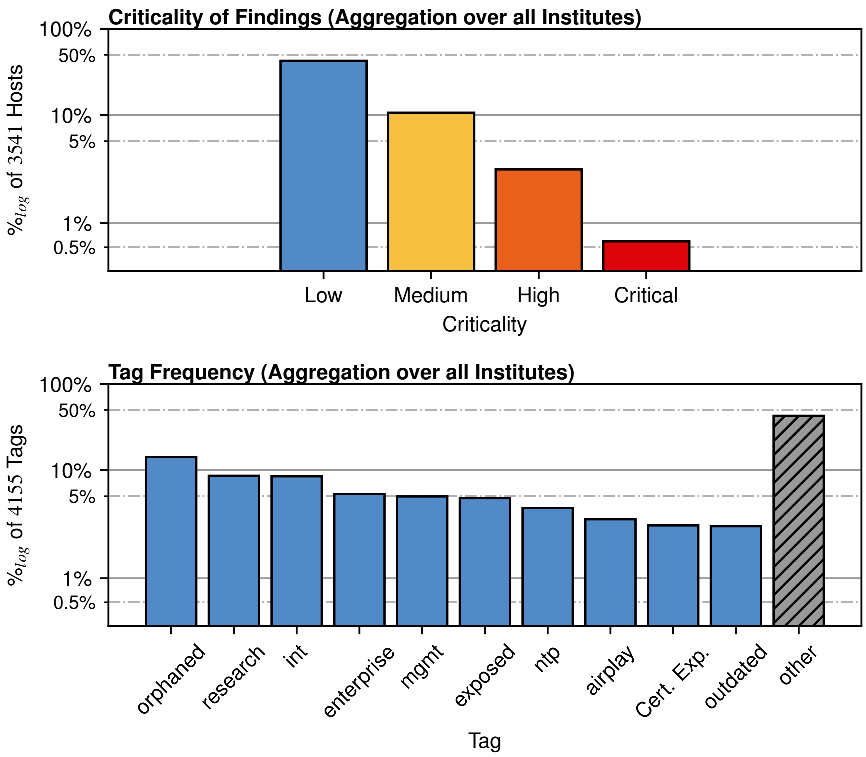


Figure 4: Aggregation of criticality ratings and tag frequency across all hosts. Please see the corresponding appendixes for information on individual institutes.

investigation and/or mitigation (medium), 83 (2.34% of all hosts exposing services/4.16% of hosts with findings) show issues that require immediate investigation and/or mitigation, while only 22 (0.62% of all hosts exposing services/1.10% of hosts with findings) show issues likely to cause an incident in the near future.

Considering the types of issues, again, see Figure 4, we find orphaned (research related) infrastructure, exposed (internal) systems, and misconfigurations–especially for enterprise systems–to be frequent. However, overall, the nature of issues is as diverse as institutes themselves, as signified by the heavy-tail distribution of our tags. Please see Section 3.2 for a more extensive discussion of issue-clusters we observed, including specific examples.

Summary Hence, despite a wide-spread occurrence of issues warranting investigation or being indicative of limitations in the operational structure[35], the general situation is comparable to what can be found across most organizations [21].

However, we also observe that institutes are not homogeneous when it comes to digital infrastructure, and individual institutes may experience a higher frequency of High/Critical issues. Please note that this report does not make specific statements on the root-causes of these differences in *individual institutes* beyond general observations on the mechanics within the organization as a whole, see Section 4.

[35] For example, asset management, monitoring, software/service life-cycle management, etc.

3.2 Vulnerability Clusters

Here, we give an overview of clusters and themes of misconfigurations and vulnerabilities encountered while observing networks of the MPG. Each case will be illustrated with a specific example that highlights the interaction effects of the corresponding clusters.

3.2.1 Exposed Research Infrastructure

A common pattern throughout MPI networks are various forms of exposed research infrastructure. This reaches from the case already described in Section 2.8, where [redacted], to systems and applications used in the dissemination of research work, also see the example below. While the latter does not necessarily have a high potential for security impact besides enabling potential lateral movement. The two underlying patterns for these cases are:

a) Systems that fall out of the standard set of applications commonly operated in IT environments, have special requirements, or need domain knowledge, while being essential to the core functions of an MPI, or,

b) Systems developed by, e.g., a PhD candidate or otherwise non-permanent staff for dissemination or collection of research results.

Example I: *Paper Specific Dynamic Site:* As an illustrative example, we are using [redacted]. The application provides information on research published in two papers, appearing in 2010 and 2011. The site carries a copyright statement of 2011, i.e., indicating that the site is over 10 years old. Various resources currently do not load, and the data service–back then provided via FTP–is no longer available.

The underlying issue here is that this site–like developed by the concerned researchers or procured externally–did not receive consistent maintenance, and was not discontinued or transformed to a static site. However, the underlying mechanic is most likely that ECRs/PhD candidates where in charge of running the site; In absence of an orderly transfer of infrastructure which tends to happen upon ECRs moving on to the next stage of their career, these resources became orphaned. Maintaining such infrastructure then is a close to impossible task for an IT team. Not providing the service is similarly not an option, given that it would just motivate external Shadow IT.

Example II: [redacted] *Storage System:* Another example of research needs leading to Shadow IT like deployments, thereby creating security and reliability issues, is [redacted]. This system is part of MPI [redacted], and also known as [redacted], with the generative naming scheme indicating that it is located in an access network. Furthermore, the system presents a certificate with the CommonName [redacted]. [redacted] is a dynamic DNS provider, and we conjecture that [redacted].

While, technically, this device could also be from [redacted] Enterprise lineup, the dynamic DNS setup and generative naming scheme provide indications more towards an under-the-desk solution. We conjecture that this system was put in place when a researcher quickly needed access to a large storage pool, which could not be provided by the institute in the required timeframe, as system requirements in the specific field of that MPI usually do not entail storage needs of the necessary scale.

The solution in place at the moment, however, creates a difficult situation in terms of reliability–as the backup and monitoring situation of the device is unlikely to be optimal if it is a consumer device and/or Shadow IT–and a liability in terms of security, as the device is fully exposed on the Internet.

3.2.2 Orphaned and Outdated Systems

The theme of orphaned systems is, in general, one of the most common security misconfigurations [7]. Throughout MPG networks, various forms of orphaned and outdated systems beyond the aforementioned research infrastructure exist. These range from the interaction with research objectives mentioned above, to more fundamental services like, for example, firewalls.

Systems quickly become orphaned, if responsible personnel departs, the system has not been documented, or provides an encapsulated service that keeps functioning without disruptions, or without updates being necessary.

Example I: *Outdated VCS:* A good example for an outdated system without specific research association is [redacted], a system related to MPI [redacted]. This system presents a TLS certificate with a validity range from [redacted] to [redacted], i.e., it expired over three years ago. The software running at [redacted] is [redacted], a version control system, in version [redacted], which has been released [redacted], i.e., over five years ago. [redacted] Furthermore, currently, [redacted] names presented in the certificate as DNSAltNames no longer resolve, and of the remaining [redacted] only [redacted] still resolves to [redacted]. The system was likely decommissioned, but then never shut off.

Example II: *Outdated Security Appliance:* The host [redacted] displays an error page for a [redacted], a discontinued product from [redacted] providing security services. Support for this product ended [redacted], while extended support was available until [redacted][36], i.e. there is no support for this product, in the best case, for over three years. Furthermore, the product relies on [redacted], which is similarly out-of-support.

These systems are run by the MPI [redacted]. Given that such an institute likely handles PII of the highest criticality, i.e., [redacted], use of outdated products must be discontinued. While the display of a more recent [redacted] on [redacted] for the address might indicate an ongoing migration, this will have to be accelerated.

3.2.3 Exposed Management Services

Throughout networks of various MPIs, management service are exposed to the Internet, even though these should be restricted to an internal set of connecting addresses. In general, exposure of management ports is not a significant issue in itself. However, exposing these ports usually increases exposure against automated scans, increases the need for a quick reaction time in case of vulnerabilities in these management interfaces[37], and may aid targeted attacks in lateral movement. Especially industrial control systems also regularly do not find themselves with management interfaces implemented using best practices, i.e., often run with outdated software or have a limited set of authentication options available.

[36] [redacted]

[37] See, for example, the recent VMWare remote code execution vulnerabilities

Management services are often accidentally exposed when the–often appliance based–device or server does not properly separate management and service provisioning to multiple interfaces or at least VLANs. Especially in the case of firewalls, exposing management interfaces makes it questionable whether the firewall itself has been configured with sufficient care to provide protection beyond a firewall being in place.

Example I: *Firewall Management Interface:* On [redacted], a [redacted] Firewall and VPN Gateway device exposes its management interface via port [redacted], i.e., at [redacted]. The device belongs to a network segment allocated to [redacted], announced by [redacted]. The exact institute [redacted] cannot be conclusively determined. However, based on an adjacent IP address [redacted] a connection to the MPI [redacted] is likely.

While the exposed interface does not use the default credentials known for [redacted][38], this device likely logs PII on web access in the concerned institute. Hence, a compromise would be a significant incident for the institutes user base, while potentially also allowing lateral movement.

Example II: *Exposed ICS/IIoT Devices:* On [redacted] under the name [redacted], an [redacted] IIoT device listens globally with its [redacted] management interface [redacted]. Even though no default credentials are in place, and the specific use of the device can not be established, the affiliation of the device to the MPI [redacted] in connection with the naming not indicating non-research use warrants investigation on whether this device is related to any experimental infrastructure. [redacted] In any case, exposure of IIoT configuration interfaces via the Internet is not advisable and should be mitigated.

3.2.4 Exposed Infrastructure Services

Besides management interfaces of appliances, a multitude of exposed ports that do not have to be reachable from the Internet can be found in various MPG affiliated networks. This ranges from administrative ports for [redacted], to infrastructural services like [redacted] and [redacted]. Exposure of these ports, if the underlying service is properly configured, is usually not an issue, as potentially privileged access should either have been deconfigured or protected by authentication and authorization measures. However,as access to these ports is not necessary from outside networks, it is recommended to prevent random access, on the one hand to reduce the impact of Internet background noise, and on the other hand to hinder lateral movement and compromises in case a new vulnerabilities becomes known.

Example I: *Exposed [redacted] Services*: An example for exposed infrastructural services are [redacted], who all expose port [redacted], i.e., [redacted]. While, in general, an exposed [redacted] port is not an issue, it could be abused in attacks on systems' availability. While, again, this is uncommon, reducing global reachability of this service does not impede its functionality, as–usually–only a limited set for [redacted] is supposed to connect to the service.

Example II: *Exposed [redacted] Application Server:* [redacted] is an application server that allows the execution of [redacted] applications to create web-applications. The application server, by default, also exposes a management interface on a dedicated port, often [redacted]. In standard scenarios all administrative functionality there should be disabled (especially from non-local hosts). However, in general, as a matter of best practices, it is advisable to prevent access to this port from the outside, and ideally binding it just to localhost.

[38] [redacted]

3.2.5 Exposed Internal Services

Within most MPIs, internal services can be found, that have been exposed to the Internet. This relates to services, where external parties do not need access. For example, lab notebooks, room booking systems, or internal documentation/wiki systems. While, in general, these applications often require user accounts, the application logic itself is still exposed to internet background noise. Again, restricting access to a limited set of users is not a feasible security mechanic. However, by ensuring only internal users can access these applications, one might mitigate especially mass-exploitation events in case other precautions–as regular updates–are not executed.

If no external users have to access these systems, multiple options exist to reduce their exposure:

- IP based access restrictions to internal hosts; However, this may be cumbersome for users, if then they have to, e.g., always use a VPN to access the resource.
- Basic HTTP authentication; Placing standard HTTP AUTH in front of a site also reduces the volume of internet background noise interacting with the application. Furthermore, if a central authentication source is being used, HTTP AUTH can be tied to, e.g., LDAP. Applications also often provide features to leverage HTTP AUTH for account creation and syncronization.
- Using OAuth/OpenID/SAML; MPG institutes already have a SAML/OpenID solution in-place to federate authentication for, e.g., [redacted]. Using this in front of internal applications has a comparable effect to HTTP AUTH, while allowing features like 2FA to be integrated.

Example: *User Documentation/Wiki:* An installation of [redacted] runs at [redacted]. The wiki is generally access restricted via the application's authentication and authorization features. This indicates that the application itself should not be generally accessible. Note, that [redacted] runs in version [redacted][39], released [redacted] to mitigate [redacted]. There are no known vulnerabilities for this version of [redacted][40]. Nevertheless, development on this application may not be active; Hence, using either SSO or HTTP AUTH instead of exposing the application directly would reduce the impact a vulnerability discovered in the future could have.

3.2.6 Misconfigurations

Security misconfigurations are a common, inherently simple, yet hard to mitigate issue [7]. While the term, in general, covers all accidentally caused preventable security issues due to the configuration state of systems, and many items also presented in other clusters are thereby covered, we will focus on more specific instances here. What security misconfigurations have in common is that the system usually attains the desired functionality, while the misconfiguration exposes data, enables lateral movement, or even allows a system to be compromised.

A good example for these specific misconfigurations are `.git/` folders exposed in the webroot of applications. As the investigated event in Section 2.8 already provides an example of this, we will not further elaborate on this point here. Please note that security misconfigurations are likely to occur if time pressure or urgency are applied. [redacted]

[39] [redacted]

[40] [redacted]

Mitigation for security misconfigurations can be performed by implementing operational best practices, regular scans and audits, a ‘just culture‘-approach[41], and following an approach where monitoring ensures that security misconfigurations observed once are detected if they occur again.

3.2.7 Insufficiently Operated Services

Throughout the MPG, several organizations still employ either outdated services, e.g., still use FTP for file sharing instead of relying on either a web-based solution[42] or migrating to newer protocols like SFTP. Another frequent instance of this issue is the use of HTTP proxies. The HTTP proxy protocol works well over HTTPS; However, most documentation discusses TLS only in the context of TLS interception, and browsers usually require a proxy auto configuration file to use authenticated proxies with TLS. Not using HTTPS for a web proxy means that credentials used for the proxy are transferred in plain text, and can therefor be considered compromised.

Example: *Ephemeral Proxy Passwords:* The MPI [redacted] runs a service [redacted], enabling users to request a temporary password for accessing their web proxy. We connect this to their [redacted] proxy [redacted] running [redacted] on the same machine not using TLS, even though the [redacted] webserver [redacted] presents a valid certificate[43]. Instead of running a plaintext web proxy, TLS should be employed. Additionally, the proxy could be–combined with a PAC–authenticated via SAML/OpenID, thereby creating an easy to configure remote access method, which could also incorporate 2FA[44].

3.2.8 Complex Security Solutions

IT security is often misunderstood as an absolute property. As such, teams–facing requirements like ‘2FA has to be rolled out *now*’, together with a strong spirit of having to be as secure as possible, develop and implement solutions that are highly complex, and place additional overhead on the systems’ users, while not necessarily generating a tangible security benefit. In the worst case, such approaches may even limit the overall security status of a system.

Reasons for the latter are, for example, the introduction of vulnerabilities and misconfigurations in the complex solution, or simply users resistance because the added overhead of the solution is an impediment to their work. In that case, a complex security solution may prompt users to implement forms of Shadow IT that circumvent mechanics in place, reduce the operators’ insights into what is happening on their network, and may quickly become permanent temporary solutions.

Example: *Complex access generation framework:* The MPI [redacted] runs a service which allows users to obtain remote access to internal resources at [redacted], see also Figure 5. Users can request a new ephemeral SSH key, which is generated on the server, and the private key then sent via email to the requesting user. Instructions on the page indicate that a login with the newly generated credentials is possible for up to 24 hours. The same page also has a button enabling a temporary IMAP password, which is valid for up to 15 minutes.

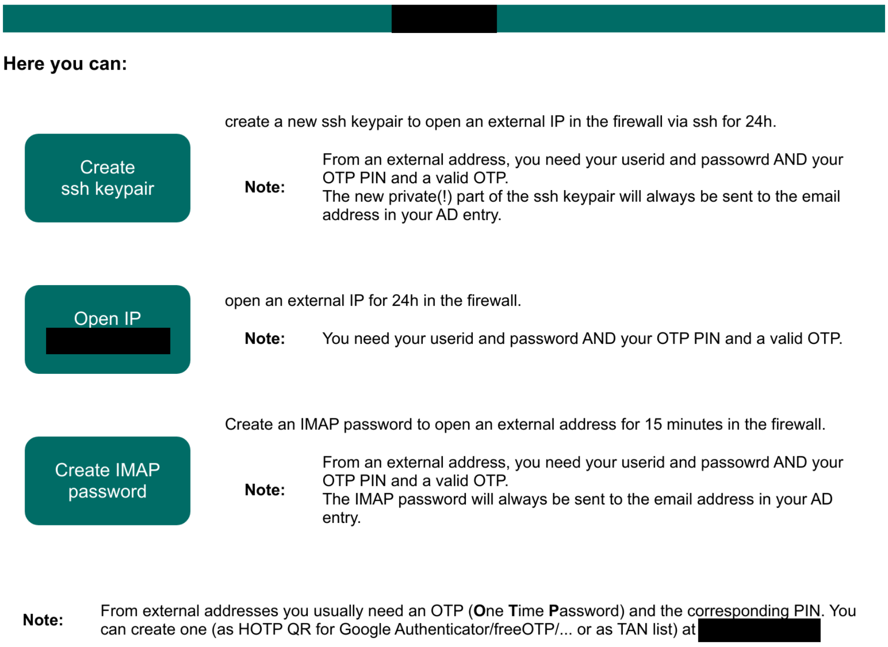
While there is no evidence that this system is already in production, the security benefit of this approach is questionable. The process in itself is a) frequent, and b) comparatively complex, i.e., likely to get into a user’s way during their normal (remote) work activities. Hence, it is likely, that users look

[41] The opposite of a ‘blame-culture’, focusing on ‘What went wrong?’ instead of appointing (individual) blame, without neglecting accountability, see Dekker & Breakey [6]. The approach is rooted in the realization that negative outcomes are often caused by systemic issues, and improving safety requires the identification and improvements to those underlying systemic factors. Focusing on (blaming) individuals hides systemic factors.

[42] For file uploads, Nextcloud offers a feature to create upload-only folders.

[43] At the same time, this setup thereby technically qualifies for a complex security solution, see the next section.

[44] The recommendation to use TLS was already provided to this institute, see Section 2.8.



[redacted]

Here you can:

create a new ssh keypair to open an external IP in the firewall via ssh for 24h.

Note: From an external address, you need your userid and passowrd AND your OTP PIN and a valid OTP. The new private(!) part of the ssh keypair will always be sent to the email address in your AD entry.

[redacted]

open an external IP for 24h in the firewall.

Note: You need your userid and password AND your OTP PIN and a valid OTP.

Create an IMAP password to open an external address for 15 minutes in the firewall.

Note: From an external address, you need your userid and passowrd AND your OTP PIN and a valid OTP. The IMAP password will always be sent to the email address in your AD entry.

Note: From external addresses you usually need an OTP (**O**ne **T**ime **P**assword) and the corresponding PIN. You can create one (as HOTP QR for Google Authenticator/freeOTP/... or as TAN list) at [redacted]

Figure 5: Webinterface for generating ephemeral access credentials.

for alternate access options. One such way would be to place a small embedded device at their desk which connects outbound to a VPN concentrator, allowing them uncomplicated access to the institute's infrastructure[45].

To highlight the limited effect of this intervention in comparison to other measures: The same MPI also runs several systems in their network ([redacted]), which display a questionable security state, see also Appendix II. For example, [redacted] runs [redacted], a version found on [redacted], an OS version that is likely out of support [redacted][46]. The system itself seems to be orphaned, with the latest [redacted] stemming from [redacted]. The system lacks TLS support, and the installed software [redacted] has version [redacted]. Since then, multiple releases of this software appeared, with the latest being [redacted] from [redacted]. Similarly, the same network contains [redacted] exposed management interfaces.

Hence, we would recommend to use a simple process for SSH, e.g., using SSH jump-hosts with mandatory SSH key authentication and/or the added use of 2FA on each login (not using key based authentication). Instead of complicating the SSH login process, efforts should be spent on identifying and discontinuing orphaned systems.

3.3 Summary

In summary, we observe a diverse assortment of vulnerabilities across various Max-Planck institutes. Naturally, the number of events differs between institutes, as it is also dependent on the size of the local infrastructure, as well as on the available staff. However, despite observing several instances of concerning and partially critical situations, we did not encounter any events or incidents that would make a compromise of the MPG *as a whole* likely[47].

[45] A service enabling such a feature would be https://tailscale.com/, a service with which end-users of average technical adeptness can easily realize such a setup, e.g., using a Rasberry Pi: https://tailscale.com/download/linux/rpi

[46] Technically, extended support is available as a paid option [redacted]. However, given other parameters, we consider it unlikely that this version is being used.

[47] This is partially rooted in the MPG's general structure, see Section 4

4 Analysis

In this section, we will hold an analytical lense at the summary of findings from Section 3. Our main objective will be to identify assets and challenges in terms of organizational mechanics. Furthermore, we will contrast the result of our analysis with common best practices in enterprise networks, and discuss in how far these practices are applicable and/or compatible to the MPG under the expicit objective of conducting research inherrent to the organization.

4.1 Organizational Mechanics

Organizational mechanics here describes parameters, circumstances, and interaction effects within the organization, which can either be claimed based on data collected in Section 3, inferred from the public organizational structure, or are known to the author of this report based on their interaction with staff in operational functions within the MPG. If the latter is the case, this will be explicitly marked for derived mechanics.

4.1.1 Role and Impact of Decentralization

The findings in Section 3 present a clear picture on the state of standardization and centralization within the MPG:

- There is no standard in terms of operating systems, basic services (network, DNS, Email, authentication, remote access).
- IT infrastructure between institutes is not integrated within an overarching management chain, and institutes act autonomously.
- There is no central monitoring of operational practices.

As such, the assortment of institutes cannot be seen as ‘the Infrastructure’ of the MPG. Instead, organizationally, each institute is its own island of digital infrastructure, independent from the rest of the organization.

While this decentralized nature of digital infrastructure in the MPG essentially prevents a security-by-control stance[48] in operations, it also exhibits protective features. As interaction between different institutes, apart from a few components mostly relevant for administrative staff centralized around the GV, are not unlike interactions with any other external party, institutes attain relative resilliance against issues in one part of the organization. More colloquially speaking, a major security incident in one institute does not necessarily impact any other institute, as long as no central and/or shared component[49], or the compromised institute provides services for another institute[50].

Given that institutes handle data of highly varying criticality[51], this compartmentalization ensures that oversights in one part of the organization do not threaten critical infrastructure in other parts. Furthermore, given diversification, a security-by-control approach could be implemented where possible, while avoiding to create noise by including institutes where such an approach is not necessary[52].

[48] We summarize the standard approach to IT security as ‘security-by-control’, see Section 4.2.2 for a detailed description.

[49] One of the few would be [redacted].

[50] See, e.g., IT around MPI [redacted], [redacted], or [redacted].

[51] Medical data of patients is, for example, more critical than pictures from an excavation.

[52] See ‘Notification Fatigue’ [29].

4.1.2 Research Infrastructure and User Requirements

Related to observations on decentralization are observations on user requirements. We encountered several types of issues that relate to researchers needing specific IT support for their work. The diversified structure of the MPG enables this, i.e., teams local to their users will have a better grasp of their needs. Nevertheless, see, e.g., Section 2.8, we also observed cases where researchers contributed systems to be run on the premises of their MPI, which might not have been tightly integrated with their host institution.

The organizational mechanic that has to be considered here is that users will seek solutions to their problems, and in case of doubt users within the MPG will prioritize accomplishing their research work above, e.g., IT security or maintainability[53]. This mechanic is further supported if organizational requirements lead to complex and overhead heavy security solutions[54].

4.1.3 Employee Churn and Orphaned Systems

Employee churn is an inherent aspect of the MPG, especially among the layer of research most concerned with digital infrastructure for research support, i.e., PhD students and post-docs. Both groups are legally[55] and academically[56] prone to depart the organization after 2-6 years. Still, especially PhD students will frequently be the ones in charge of, e.g., creating promotional or open-data/open-artifact pages for their research papers, setting up research equipment which integrates with digital infrastructure, creating tools and systems that support their daily working with research systems and lab equipment, and creating automation around common scientific tasks. More senior researchers, like tenured and tenure-track group leaders, as well as directors, usually have a diversified role-profile consisting of organizational, managerial, and strategic tasks, which makes it unlikely that they will maintain systems created by the aforementioned groups. This, in turn, benefits a mechanic where systems become orphaned, especially if the service they run continues to function without a person being responsible, i.e., without regular maintenance being necessary.

4.1.4 Enterprise Toolchains and Infrastructure

Throughout various MPIs, we could observe enterprise toolchains integrated for specific tasks. Commonly, this includes threat detection and prevention in the form of Firewall Appliances, VPN appliances, and Intrusion Detection Systems. We conjecture, that these systems are put in place if a specific competency, e.g., integrating open source components, is not available within the local MPI's IT staff, or available staff lacks sufficient time to focus on that task. Nevertheless, during interactions with staff from one MPI contacted during the scans, see Section 2.8, they noted–not without rather justified pride–that their Firewall system had been built in-house and the total CapEx of the system over time was around two orders of magnitude less than a commercial solution we observed in several MPIs. At the same time, the commercial solutions we did observe were found due to exposed management ports and comparable misconfiguration, indicating that they might not be utilized to the fullest extent of their feature set, and might even induce additional exposure for a network.

We argue that in both cases, i.e., when using and combining open source components vs. leveraging commercial options, sufficient organizational capabilities have to be available–via consultants or internal expertise–to ensure that these systems are properly operated[57].

[53] More colloquially put: 'Users will find a way.'

[54] See the 'Complex Security Solutions' example in Section 3.2, [redacted].

[55] See 'Wissenschaftszeitvertragsgesetz'

[56] One is expected to change their host organization after a PhD or post-doc.

[57] See also Section 4.1.6 and Section 5.2.2

4.1.5 Systematic Diversification and Specialization

Related to the previous point, we also observed that close to all MPIs require a full stack IT services department. While, for basic services like basic networking, printing, client administration and support, this is common, MPIs also run services like virtualization infrastructure, applications like Nextcloud, authentication systems and Email[58]. Please note[59] that centralizing these services is not feasible within the MPG, as it would make the organization more susceptible towards attacks, while creating major migration overhead, to put a system in place many researchers are likely to work around, while also causing discomfort with existing expertise[60].

4.1.6 Capability Retention

With the main objective of the MPG being to conduct *independent* research, it is critical that the MPG retains the ability to do so independently. This also means that the MPG must retain its ability to build and operate–especially research critical–infrastructure, especially in sectors were a mono or oligopolization of the market can be observed, e.g., in the area of computational infrastructures[61].

This, in turn, ties in with the retention of expertise, i.e., employees tasked with the operation of systems, who often have gathered extensive non-technical knowledge on their local user base and that user base's specific needs and requirements. At the same time, the MPG orients itself around TVöD; From a monetary perspective, and [redacted], it is unlikely that the salary is the sole motivator for employees in system operations.

While, at first, this does not seem like a security critical observation, it ultimately is. Common aspects valued are working climate/culture, freedom, authority and responsibility, and identification with ones work. If a strategy to improve IT security is implemented that conflicts with one or multiple of these points, the result can be a cascading reduction/churn in workforce. That process would lead to technical capability loss, increase the likelihood of orphaned systems, and destroy years of institutional knowledge on specific MPI's user bases, which is essential to balance researchers' needs vs. infrastructural measures and changes. Hence, even if the head-count could be replenished, infrastructure support–by lacking the required care component–might lead to more cases where 'users find a way'.

4.2 Comparison to Enterprise Networks

In this section, we compare the perspective of organizational mechanics mentioned above to the situation commonly found in enterprise networks.

4.2.1 Operational Requirements

Enterprise networks, in general, have a predictable set of operational requirements. Basic services like printing, editing and sharing documents, network access, remote network access for traveling employees, and the use of special in-house applications like, for example, SAP if rolled out by an organization. In addition, departments regularly have specialized applications, if these have not been integrated into a framework like SAP, for example for HR/Payroll management, procurement, reimbursements, etc.

[58] Especially email is notoriously known for being complex, interdependent, and critical for users, i.e., an unreliable email system has a high visibility towards users, while also severely impacting their work.

[59] See also Section 4.2.

[60] See Section 4.1.6.

[61] See [10].

However, these tools are regularly well understood[62], workloads are predictable, and new software is usually introduced in a (semi-)structured manner. The application stack used is decided top-down, and employees preferences in terms of specific tools can be heard, but there is no inert requirement to satisfy these preferences. Furthermore, despite local variations, requirements of employees vary only coarsely based on their general role. Security incidents can be quantified monetarily by calculating the fiscal impact of direct damages and costs of recovery, losses due to halted production workloads, losses in reputation, and potential contractually agreed damages.

In contrast to that, as outlined before, a research network serves, first and foremost, the purpose of enabling researchers work. As in all modern organizations, this includes standard workflow tools, from editing documents, over collaborative work to (remote) access, and printing. However, at the same time, researchers tend to be particular about the toolchains they use for their work, which often may even include a specific operating system. New applications may become necessary without prior planning or consultation with the IT department, for example, if a new microscope is bought. In that process, the instrument's efficacy tends to be the deciding factor for procurement, while standard questions like infrastructural fit of IT components may not be considered.

Furthermore, in lived practice, the delineation between private and professional life is traditionally poor in science[63]. This means that, contrary to a bank teller who logs into a machine, uses a corporate application for eight hours, and then logs out, a researcher may conduct a variety of private activities on their professional machine. Preventing that, either with technical measures, or by policy, will lead to resistance and a reduction in productivity–as energy is spent on 'finding ways', and a morale impact may be observed.

This, however, does not mean that within a research organization no parts exist that behave exactly like a traditional enterprise. Major parts of the administration, for example, are operationally indistinguishable from an enterprise environment. Furthermore, while the careful and secure operation of digital infrastructure is imperative for all institutes, a special duty of care for confidentiality, integrity, and availability may apply to individual parts of the organization. This pertains to, e.g., medical research facilities and nuclear research facilities, which handle sensitive PII and/or fall under critical infrastructure. Especially when different parts of the organization with such varying circumstances have to interface, challenges emerge.

Finally, losses in terms of research data may have a varying impact–ranging from low in cases where researchers work with open data and an incident only restricts their work–to disastrous–when medical data is concerned. In either case, a monetary quantification of costs is not straight-forward beyond the direct costs of recovery.

4.2.2 Applicability of 'Security-by-Control'

With 'Security-by-Control', we describe the set of best-practices commonly understood to be ideal in terms of managing security for an enterprise network. While technically simple, these are often not implemented (sufficiently) in practice. However, as a guide-line, they summarize an overall approach which can be useful for managing an enterprise's network. Specifically, this entails:

[62] There are side parameters, especially in the context of SAP; However, challenges here regularly stem from a procedural misunderstanding, i.e., an organization trying to adapt large generalist frameworks to their processes instead of the other way around.

[63] See, e.g., [27, 4, 26]; While the matter of work-life-balance requires a nuanced discussion, for the matter of creating secure infrastructure it is a reality which has to be accounted for.

Use of a cattle approach: Within system administration, the 'cattle approach' refers to an approach where variation within a system is reduced to ensure a consistent outcome[64]. The term 'cattle' stems from the analogy, in which individually managed systems are called 'pets'. By using a cattle approach, supported by an automation framework, an organization can ensure that only those things that are permitted do occur on their systems, while reducing the overall effort needed to ensure systems are, for example, monitored, create backups, and receive software updates[65]. However, a cattle approach always requires making things the same that inherently are not.

Reduction of the TCB: As a subset of the cattle approach, an enterprise environment should strife to reduce its TCB, or 'Trusted Code/Computing Base'[66]. Besides the benefits it provides in terms of the cattle approach, ensuring that there is one type (or some types from one vendor) of printers means only one drive has to be updated. Restricting users' choices for webbrowsers to one, ideally the one paired with the OS, means less overhead. Using only one operating system, means only one codebase has to be trusted. Using one vendor for GPUs in workstations from one vendor means only one more software package to account for in terms of GPU drivers, and ideally only one management software for all workstations.

Prevention of using unmonitored services/private use: In the same notion as with the TCB reduction, an enterprise will usually strife to limit private use and use of un-sanctioned services for security (and reliability) reasons. Accessing an injured users' mailbox is significantly easier–in terms of privacy legislature–if there is policy (and ideally also technical solutions) in place that prevent the private use of one's corporate email account. The likelihood of receiving an infected file via Facebook is significantly reduced by preventing Facebook access all together[67].

Centralization: In addition to the above measures, centralization is usually a key-item in enterprise systems. This includes authentication and authorization, thereby enabling a central place for auditing access policies, but also more general monitoring and revision-safe auditing/logging systems. For example, in a corporate network without any permissible private use, centralized intrusion detection following a strict allow-listing approach based on the limited set of allowed applications is feasible. Similarly, centralizing core services like email ensures that there is one place where expertise to run this service is needed[68]. Similarly, centralized lifecycle management can ensure that systems do not become orphaned.

4.2.3 Implications of Research Networks

When comparing the enterprise approach to security to the requirements of a research network, we note that:

- A cattle approach, including a centralization of services and functions has significant risks, due to:
  - Inhibiting research activities due to a reduction in the organization's ability to cater towards researchers' individual needs.
  - Risks in terms of employee satisfaction, both for researchers and operations staff.
  - Security risks, as–despite centralization–researchers taking evasive actions and deploying Shadow IT increases attack surfaces, while centralization aids lateral movement.

[64] Not necessarily the best. See Limoncelli et al. [21], who compare this approach with fastfood; It is not the best, but always the same.

[65] Note, though, that a compromised IaC/Automation setup in this approach entails a full compromise of the whole organization.

[66] We acknowledge that, in practice, especially enterprise software regularly suffers from the opposite mechanic.

[67] Here, the 'users find a way' aspect nevertheless plays an important role. While users can usually be convinced that, e.g., in a nuclear research facility, no phones may be brought in or used, and no Facebook browsing may take place, this will significantly differ for other critical environments. For example, a night-shift ICU nurse would likely develop a surprising creativity if they could no longer browse a little when they might catch five minutes of break throughout the night. Hence, situation dependent, it can be advisable to create a dedicated and separated relieve 'valve'.

[68] Please note that the MS Exchange Product common in enterprise environments is a strong candidate for outsourcing, as–due to its complexity–even the BSI is not convinced that a sustainably secure self-hosted operation is feasible.

Nevertheless, a cattle approach can be useful to handle different parts of the organization with the same set of requirements to relieve workload from local staff, ensure consistency, and improve the maintenance status of the setup.[69]

- Points where research components of the infrastructure interface with more enterprise like parts are critical.
- Security requirements across institutes differ.
- Capability building and retention is critical in securely handling the diversified infrastructure created by research requirements.

Hence, we have to derive recommendations that create an enabling structure–both for researchers and operators in institutes–aiding them in a non-paternalistic way, ideally leveraging bottom-up mechanics. This includes building in-house topic-expertise, and making that expertise available to the rest of the organization.

Interventions should focus on *reducing* overhead and effort for operators and researchers while addressing the root-causes of the vulnerability clusters identified in Section 3.2. At the same time, it has to be established which components and institutes could benefit from a more control-focused approach to security, and it has to be ensured that these parts of the system interface with the rest in a way that ensures potential harm does not spill over, i.e., sufficient fencing between segments must be in place to ensure that individual–ultimately not preventable–compromises do not lead to a compromise of the whole organization.

Similarly, structures must be in place that assist recovery, preparation for recovery, and ensuring that root causes of incidents are identified and mitigated throughout the organization. Finally, given that several vulnerability clusters related to relatively low criticality ‘chores’, structures must be in place that assist in handling these chores, in a way that *reduces* workload.

[69] Consider, for example, a part of the organization with explicit expertise in the operation of email setups providing standardized email hosting to institutes which to not have the personnel resources and/or expertise to self-host.

5 Recommendations

In this section, we provide recommendations and refer to best practices which could benefit the security state of the MPG. For all recommendations we provide a brief contextualization of how they integrate within the requirement set of the MPGs network, i.e., how their implementation should be executed. Please note that, while the technical implementation of several recommendations is straight forward, we caution against an implementation solely based on the abbreviated descriptions below. Instead, we provide a note regarding the implementation report that can be requested for the corresponding recommendation, if describing the full implementation scope exceeds the space available in this report.

5.1 Technical Recommendations

Technical recommendations pertain to the introduction of new products, mechanics, and the implementation of topological and logical changes to network infrastructure. Please note that, while technical in nature, several of these recommendations must integrate with recommendations in Section 5.2 to be effective. We note this explicitly.

5.1.1 Continuous External Monitoring

Recommendation 1

Implement continuous external scans similar to the methodology used in this report as an in-house service, integrated with a local team as described in Section 5.2.4.

The underlying scans for this report, see Section 2, have been comparatively straight-forward. Nevertheless, they found several critical issues, and highlighted a wide range of medium and low criticality issues/chores. Mitigating these will take time.

Additionally, we note that–especially for research networks–requirements change. Digital infrastructure is never done; New systems get deployed, old decommissioned, and what used to be considered secure may become orphaned, causing events.

Conforming to Section 2.7, these scans should use a significantly lower packet per second rate, while also expanding the range of scanned ports to all ports for TCP and a larger set of ports for UDP[70]. Furthermore, scans should no longer rely on ICMP reachability of hosts, and the list of networks in scope should be more closely maintained to represent the actually used networks of the MPG. We recommend a rate with which each host receives a full scan between once a week and once a month.

The gathered data can then be made available to operators of network segments via an API or analysis interface, ideally enriched with additional recommendations, similar to the Appendix of this document. This enables operators to integrate these scans into their monitoring, ensuring quick mitigation of the inevitably occurring issues, while also being able to resolve false-positives and acknowledging/documenting low-priority issues.

[70] All UDP ports might also be an option.

When implementing this recommendation, it should be done in coordination with the community of system operators within the MPG, i.e., as a self-hosted in-house service, to ensure that this service caters towards the unique requirements and situation of the MPG. Furthermore, it should be integrated with the team described in Section 5.2.4. That team should actively monitor the gathered data to identify new issues, and should actively engage with institute's operations teams to assist them in resolving chores, and reduce overhead encountered by these teams.

Based on the infrastructure created for the initial scans, we estimate the initial CapEx for creating this service between EUR5,000 and EUR15,000[71], with an annual OpEx around EUR6,000, excluding personnel costs.

Please request the report on 'Continuous External Scanning Infrastructure' for a full description of the implementation and design guidelines for this recommendation.

5.1.2 Evaluation of TI Feeds

Recommendation 2

Start monitoring available TI (Threat Intelligence) feeds, integrated into the infrastructure of Recommendation 1.

TI, or threat intelligence, feeds are data sources provided by external parties to provide information on potential threats originating from or targeted at a network. Given the distributed nature of the MPG with a highly diversified platform and mechanics that cater towards orphaned systems, it might be useful to obtain access to threat intelligence feeds that document attack behavior from networks[72].

The objective of using such a service would be monitoring whether hosts within the MPG participate in, e.g., automated malicious activity on the Internet due to an overlooked compromise. The use of TI feeds could be integrated with the recommendation in Section 5.1.1 and, especially for continuous exploratory analysis to identify new and emerging issues, the team described in Section 5.2.4.

5.1.3 Monitoring of Asset Management and Patch status

Recommendation 3

Regularly audit all current infrastructure/systems for documentation, monitoring, and patching status.

We strongly encourage all institutes to continuously audit their current infrastructure and its integration into ongoing real-time monitoring. Specifically, for *each* component connected to a network, the following questions should be regularly answered[73]:

- Is this device documented in the internal documentation/asset management?
  - Which software runs on this device?
  - How do we update this device?
  - If the device fails, can we restore it?[74]

[71] CapEx might be reduced depending on the implementation of the recommendation in Section 5.2.4.

[72] Greynoise.io, for example, offers such a service. Similar offers are available from, for example, the ShadowServer Foundation. However, [redacted] in-house options and expertise may also be available.

[73] This process can be supported by automation that triggers regular, e.g., every six months, reminders for reconsidering these points.

[74] For backups, see below.

  - Does the documentation list the responsible ‘owner’?
  - What is the purpose of this device?
- Who is the owner of the device, and the owner of the service the device provides?
- Are backups created for this device, and if so, can these be used to restore functionality?
- Is the purpose this device fulfills still current?
- Is the device itself (base OS patch-level, hardware) and installed software (patch state) monitored, i.e., are notifications automatically sent to a person able to install updates or fix hardware if software gets outdated, and is this escalated if updates are not installed?
- Are patches still made available for the device itself and services running on the device?
- Does this system have to be externally reachable?

With these questions combined, ideally supported by modelling them in local real-time monitoring, a major fraction of identified chores and underlying root-causes for higher criticality events observed during this survey should be mitigated.

5.1.4 Network Segmentation and Fencing

Recommendation 4

Analyze institutes’ security and operational requirements in a bottom-up fashion, considering operational practice and researchers’ needs. Based on this, ensure proper fencing of institutes/infrastructures with higher security and/or centralization requirements (central administration/medical/critical infrastructure).

As outlined in Section 4, operational and security requirements differ throughout the MPG. While, for example, institutes providing health-care services may have a higher level of assurance needs, e.g., in MPI [redacted], an open network is a fundamental component of enabling researchers’ work. Similarly, while research driven networks in institutes may have operational requirements conflicting with enterprise practices, components related to the central administration may behave more ‘enterprise-like’.

As such, we recommend to carefully assess the operational requirements of individual networks, and to ensure that more open networks are sufficiently separated from more restrictive systems. Please note that this entails carefully assessing users’ needs in terms of open network access and efficacy, i.e., in some cases a more stringent lock-down in terms of security for core functionality might necessitate the introduction of parallel infrastructure to ensure users to not evade isolation mechanics, see Section 4.2.2, Prevention of using unmonitored services/private use.

In any case, this assessment *must* be conducted in collaboration with, and ideally in responsibility of local operations teams, i.e., bottom-up instead of top-down. This is essential to not create a situation in which responsibility is delegated to teams in institutes, while authority is centrally allocated. This separation of authority and responsibility is a common path to disaster in centralized system operations, leading to a disconnect from responsibility among those left without authority over their work due to their lack of agency, while the party holding authority considers responsibility to be delegated, ultimately leading to negative outcomes as neither party considers themselves responsible.

Please request individual reports on ‘Segmentation and Infrastructure Design Recommendations’ per Institute, or a report on ‘Implementing Collaborative System Assessment and Redesign’ for a collaborative approach[75].

5.2 Organizational and Governance Recommendations

Here, we will take a step back and discuss governance mechanics that should be considered when trying to improve the overall security state of the MPG.

5.2.1 Catering Towards Decentralization

Recommendation 5

Preserve the current state of decentralized administration and operation to mitigate lateral movement in case of compromises given the diversified nature of requirements and practices across institutes, while ensuring that operators are enabled to cater infrastructure to their users’ needs.

As discussed in Section 4.1.1, decentralization is a major asset of the MPG, ensuring that local security incidents remain local. It is essential that this asset is not destroyed by the introduction of centralized security measures, or by centralizing parts of the organization, in an attempt to then be in a better position to exert centralized security/system management and/or reduce overhead.

We argue that, given the underlying mechanics of a research network across a highly diversified organization, such measures would introduce the threat landscape of a centralized approach in terms of lateral movement by attackers, without providing the benefits of a standardized environment due to the diversified nature of requirements.

While arguable, one might consider [redacted] In turn, we observed one misconfiguration[76] and one complex security solution related to 2FA[77]. [redacted]

Hence, we argue, that future approaches should aim at strengthening the distributed and collaborative operations environment in the MPG, leaving authority and responsibility local, while strengthening knowledge exchange, joint efforts, and planning.

Please note that, for individual institutes, a more centralized approach might be viable. I.e., if the nature of an institute requires extremely limited services, it might be advisable if a method is found that allows semi-centralized provisioning of services, also see Section 5.2.2. Still, such a project should not be forcefully imposed on institutes, and will require a careful and considerate approach valuing local expertise.

[75] If implemented as a continuous process, mediating this responsibility would fit well within the scope of the team described in Section 5.2.4.

[76] See the investigated event in Section 2.

[77] See the example in Section 3.2.8.

5.2.2 Encouragement of Distributed Specialization

Recommendation 6

Encourage, facilitate, and leverage existing specific technological expertise, encourage the bottom-up formation of competence centers, and invest into internal capability building and sharing.

As discussed in Sections 4.1.4-4.1.6, capabilities are a major challenge within the MPG. On the one hand, we observe some institutes exhibiting outdated or ill managed enterprise systems, e.g., outdated and/or expensive firewalling systems. Basic services, like email, have grown extremely complex over the past decades, and maintaining deliverability is an issue[78].

On the other hand, the MPG has–due to its size–a significant asset of highly skilled system operators with expertise in a wide range of systems and applications. For example, considering the complexity of email and related protocols, the MPI [redacted] employs [redacted][79]. [redacted][80]. It would be prudent for the MPG to:

- Leverage such in-house expertise[81], by creating opportunities to enable personnel to apply specific expertise outside of their own MPI, thereby reducing the need for expertise in all aspects of system administration to be present in *each* MPI[82]. Besides potential higher efficacy due to more specific expertise of in-house personnel, this approach would, due to the organization's scale, also likely be more cost-effective in the long run.
- Encourage such activities, given considerations in Section 4.1.6 and the positive impact of such activities on the public picture among operators on the non-monetary benefits of the overall organization, i.e., due to the positive impact it has on capability acquisition.

We hence recommend to encourage an interest driven bottom-up[83] specialization among service staff within the MPG, e.g., the creation of champions for certain service types, to a degree independent of implementations. Note that, in leveraging such a structure, it must be ensured that distributed expertise is not used as full-scale outsourcing, creating a centralized structure. Instead, distributed experts should–as external 'consultants'–assist other institutes in the deployment of specific solutions and provide training to enable at least managed operation (over time) of regular tasks, see Section 5.2.3.

Please request a report on 'Base Service Capability Requirement Analysis' for a report detailing basic services and service types that would benefit from creating specializations, an exploration of opportunities within the MPG to leverage such expertise, and suggestions on a governance structure to enable distributed access to expertise. Furthermore, please also see Section 5.2.4.

[78] See, for example, [14].

[79] [redacted], see [14].

[80] See, [redacted].

[81] Please note that, especially in the given example, the specific in-house expertise *significantly* exceeds expertise commonly found in the industry.

[82] See also the outsourcing considerations based on frequency of activities [21, Chapter 48.8]. While rare tasks, e.g., setting up a backup system, only occur every couple of years in an organization, making in-house expertise not economically viable, the MPG effectively consists of a distributed set of organizations, making accumulating in-house expertise viable again.

[83] This is essential for motivation.

5.2.3 Introduction of an Operational Excellence Framework

> *Recommendation 7*
>
> Introduce an operational excellence framework following Limoncelli et al. [21, Chapters 55 and 56] in all institutes. An operational excellence framework *must not* be used to derive KPIs.

A common best practice to improve the efficacy and general operational performance of (IT) teams is the introduction of an operational excellence framework, see, for example, the implementation guidelines by Limoncelli et al. [21, Chapters 55 and 56], on which we base this section. An operational excellence (OE) framework utilizes, for digital infrastructure, eight broad categories of operational responsibilities (OR), in which a specific service or system-operations in general is evaluated. Specifically:

Regular tasks (RT) Standard (non-emergency) task performance.

Emergency response (ER) Response to unexpected events, handling incidents, breaches and outages.

Monitoring and metrics (MM) How well historic and real-time monitoring are implemented, and whether metrics reflect/quantify performance appropriately.

Capacity planning (CP) Performance in estimating and assessing future resource requirements, including HR, supply-chain, and capacity.

Change management (CM) Performance of the change management process, including configuration management and updates (hard-/software).

New product introduction and old product removal (NPI/OPR) Performance in introducing new products and services, as well as discontinuation of old services.

Service deploy and decommission (SDD) Handling of service deployment for *existing* services.

Performance and efficiency (PE) Process of deriving insights from monitoring and capacity planning data to assess and operation's efficiency.

These OR are assessed in commonly five levels, even though an organization may deviate from this approach to better adjust the model to operational needs, usually reducing the number of levels especially when introducing the approach. The five default levels are, per Limoncelli et al. [21]:

Level 1, Initial The initial level is–far too often–the 'natural' state of digital infrastructure. If they exist at all, usually they are improvised/ad-hoc. There is a high reliance on 'individual heroics', see also the 'cowboy' persona [21, Appendix B]. New processes start here.

Level 2, Repeatable In this level, the process reached a sufficient level of documentation to enable different individuals to execute the same process and reach the same result.

Level 3, Defined In addition to a repeatable process/OR, it is clear *who* is responsible (which role) to execute the process, and those roles are aware of these responsibilities.

Level 4, Managed A managed process/OR requires metrics to be collected (frequency of process execution, process efficacy, etc.) and made available to stakeholders, e.g., via a dashboard. Decisions about the process are rooted in that data, and deviations from normal operations are analyzed and used for further refinement.

Level 5, Optimizing In addition to a managed process, this level requires active use of measured metrics to refine and improve the process.

For guidelines on assessing OR levels, please see Limoncelli et al. [21, Chapter 56]. Also note that these levels relate to launch categorizations in service deployment and automation/IaC assessments [21, Chapter 20.1.2].

Implementation Considerations When considering the introduction of an OE in an organization, it is imperative to follow best practices similar to a DevOps approach, most importantly 'Pull, don't push.' and 'Build community' [21, Chapter 1, 4]. Furthermore, it ***must*** be understood by organizational leadership, that introducing an OE framework is an *enabling* mechanic for teams, i.e., works bottom up to *enable* teams to self-improve and is ***not*** a tool for controlling and/or a KPI. In fact, using an OE, while it *reflects* operational readiness/performance as a metric or performance analysis lever, or a KPI in all relevant matters, e.g., in the question of promotions and/or salary scaling, will *prevent the OE from being effective*. If an OE is not used as a self-improvement tool for teams in a 'just culture'-environment, but instead as a KPI or if negative consequences are drawn for 'low' level ratings, the OE will not be treated with the necessary honesty, and–like any other metric–be gamed [21, Chapter 55]. Instead, only positive reinforcement can be used, e.g., by making new and interesting challenges available to teams with high operational excellence, and providing additional room for self-improvement and professional development, see also Section 4.1.6. Finally, it has to be considered that Levels 4 and 5 are not always necessary for all services.

Please request an additional report on an 'Implementation Strategy for an Operational Excellence Framework' for a detailed strategy on how a participatory introduction of OE across the MPG can be enabled top-down.

5.2.4 Introduction of an Overaching Support Service

Recommendation 8

Introduce a distributed team supporting operators in handling ad-hoc infrastructure requirements, providing continuous scanning, encouraging distributed expertise, supporting incident response, and evangelizing for service improvements in the organization.

The recommendations already outlined in this section, as well as conclusions from our analysis in Section 4 contain inter-dependencies which cannot be resolved by isolated tactical decisions and improvements. Instead, we suggest to develop a decentralized but dedicated team focused at addressing key-challenges, specifically:

- Handling the process of turning ad-hoc digital infrastructure requirements due to research activity into services on at least Level 3 of the OE, see Section 5.2.3.
- Owning a continuous scan service, see Section 5.1.1, providing a real-time monitoring interface to institutes.
- Performing continuous observations of continuous scan data and TI feeds, following up on observed issues, specifically also including 'low' criticality tasks (chores), by providing comprehensive guidance to operators and leveraging community effects.
- Providing incident response capabilities to the MPG, including disaster recovery assistance and security incidence response.

- Providing a point-of-contact for distributed capabilities, see 5.2.2.
- Provide 'evangelizing'[84] for technology and process improvements, including trainings, e.g., on implementing OE, new technologies, and security in general.

Ad-Hoc Infrastructure Handling The major root cause of Shadow IT is a mismatch between organizationally provided infrastructure and users actual needs. For example, requiring all users to use a specific email platform might lead to issues, if the client integration does not conform to users' work flows, e.g., as no compatible clients are available for a user's operating system, or if mandatory proprietary clients break with the users established workflow[85]. Instead of using the organizationally provided service, users then start to utilize their private or dedicated privately acquired email accounts for professional communication.

More connected to the examples from Section 2.8 and 3.2, as discussed in Section 4.1.2, this problem also occurs if, for example, researchers acquire experimental equipment which needs a very specific software environment for its control software. If that stack does not have a good infrastructural fit with an organization's remaining infrastructure, the operations team might not be willing to perform an integration of that stack and take over operational responsibility. Similarly, a local operations team might not be able to provide support for a complex and new service as quickly as needed by researchers for the acquisition to be useful[86]. This also occurs in the case of, e.g., having to provide a valorization/value-added service accompanying a paper's publication, e.g., when a piece of code authored by a researcher[87] to provide access to accompanying material or implementations has to be run, ideally for 'eternity'. Finally, a common pattern would be requirements for a dynamic website for, e.g., a conference or similar activity. It is unlikely that an institutes team has the spare resources to provide assistance in the implementation and operation of such services on the short notice demanded by the realities of scientific work.

Hence, in all of these cases, users may:

- Use available resources to improvise a service, e.g., using an access network with public IP addresses to host a service using dynamic DNS and consumer hardware, see Section 3.2.1.
- Use low-barrier resources in self-administration to provide the service, yet failing to ensure continuous maintenance, see Sections 3.2.1 and 4.1.3.
- Use external services, e.g., by renting webhosting or a virtual machine from a third party[88].

This, in turn, creates liabilities in terms of reliability and security. Especially the case of orphaned research artifacts is notorious within science, as a major inhibitor of open and reproducible science.

Hence, we suggest that the strategic team can be requested by users and IT teams to handle such 'unusual' requests, while also working on ways[89] to improve the OE of the overall processes, i.e., identifying new organizational services to resolve clusters of these issues on a high OE level. The team then can either provide solutions as available, assist in the implementation of solutions, or suggest to the requesters that they should go the path usually taken with Shadow IT to 'make things work', but–as soon as functionality is attained–take over the service and ensure an OE level of at least three

[84] This, or rather 'evangelist' as a person-description, is a common term in digital infrastructure operations for institutionalized advocates of best-practices and technology.

[85] This is an anecdotally common mechanic observed upon Office 365 introductions in research-heavy environments.

[86] Consider new lab equipment bought to handle a pressing research question given revision requirements for a paper with an upcoming deadline, or having to buy a replacement for broken equipment from another vendor, potentially again under the common deadline requirements/-pressure found in science.

[87] This considers the assumption that researchers, in general, are not trained software engineers, hence exhibiting a limited code quality in their outputs. We acknowledge that individual cases may be an exception here.

[88] These activities were not in scope for this report, but the comparison to the [redacted] report in Section 6 indicated several instances of this mechanic.

[89] Not necessarily implementing themselves.

across all ORs. Ultimately, the objective should be to hand over the service in a maintainable state so the distributed operational team can perform–at least–regular tasks/OR.

Please note that the services of this team *must* be cost-neutral within the institutes' context, i.e., using this service should not incur an ICC (intra company charge). Otherwise, the risk persists that the service is not used to avoid costs, defeating the objective of improving security within the organization. We suggest to fund the service from centralized budget, given the objectives potentially even from security related funds.

Furthermore, we note that this tasks necessitates the availability of equipment, e.g., servers/workstations/network equipment, to be able to react to a versatile set of requirements. This equipment should a) not suffer from supply-chain issues, i.e., delivery times delaying availability, and b) should be as cheap as possible. Here, we suggest that institutes can[90] offer decommissioned hardware to the strategic team, which builds a repository ensuring availability if requests arise.

This approach is effective, as–given depreciation cycles of 3-4 years–decommissioned hardware is often still sufficient to run services at the scale of observed special requests, and spare-parts can be–in absence of support contracts–provided from the repository itself. Furthermore, this approach is environmentally conscious, reducing e-waste. The concern that production workloads should not run on out-of-SLA systems [21] is secondary in the first establishment of a maintainable process. In fact, ensuring a local team has to perform a service migration, e.g., via a non-incident backup-restore, is a good way to ensure a sufficient OE level has been attained, i.e., the OR have been transferred, see also Dietrich et al. [7]. Please note that this approach requires careful crafting of a governance framework which combines low overhead with sufficient incentives and mechanics to avoid a 'tragedy of the commons' [13], e.g., institutes using the team/service as a 'cheap' way of outsourcing regular tasks or acquisitions.

Scan Infrastructure In Section 5.1.1, we suggest to conduct a continuous, external monitoring of all infrastructure, comparable to what has been done in preparation of this report. For this, ideally, a scanning infrastructure with attached data collection and analysis infrastructure similar to the one described in Section 2 would be advisable. Data collection for this approach could be outsourced, and pricing for this service at the scale of the MPG[91] starts in excess of $1,000[92]. In comparison, replicating the scanning infrastructure from Section 2 could be implemented at around EUR6,000 in annual OpEx, with CapEx varying based on the planned extend of the infrastructure, and if existing systems can be repurposed, see Section 5.1.1.

We argue that, given that the effective load of services outlined in the previous paragraph will not be continuous, the proposed strategic team would be well suited to own such an in-house service. Furthermore, owning this service would allow the team to further refine scanning, and tailor it to specific needs of the MPG. For example, it might be reasonable to develop sensors to be placed within the local networks of institutes to collect data relevant in the context of defense-in-depth.

Continuous Observations As discussed in Section 3, a major portion of (security) issues within MPG networks falls into the category of 'chores', i.e., issues that do not pose an immediate risk, might conceivably be useful to an attacker for lateral movement, or iterative privilege escalation, while often

[90] Potentially: Should be required.

[91] Between a /12 and a /13 in IPv4, i.e., between 524,288 and 1,048,576 IPv4 addresses.

[92] E.g., https://account.shodan.io/billing at $1,099 for 327,680 IPs by shodan.io, with several competitors like, e.g., Censys requiring sales contacts for pricing estimations.

needing repetitive effort–especially due to their frequency–to be mitigated. Similarly, what is considered an issue can change over time, and that change process has to be maintained. The same holds for processing information from external TI sources, which have to be processed, analyzed, and then all items for mitigation handled.

Handling if events from these sources then needs approaching local teams, discussing the issue, and–most crucially–providing suggestions on how the observed issues can be mitigated while taking the local operational requirements into account. For example, outdated software may be a specific requirement of self-developed research infrastructure or necessary for interacting with a specific scientific instrument. In that case, mitigation requires to examine infrastructure and requirements, devising a solution which maintains research functionality while limiting the potential impact of an event. Similar provisions hold in cases of insufficiently operated systems, where local teams have to be enabled to implement more optimal solutions. An example here would be mitigating complex security solutions by assisting teams in designing more directed implementations, see Section 3.2.8.

The suggested strategic team could provide this consistent monitoring with a corresponding mandate from the MPG. Additionally, they could conduct with RnD level engagements with the infrastructure in terms of active threat analysis[93], especially to identify issues not apparent from scans alone or issues for which new interactions with targets during scans are required.

Incidence Response Independent of the type of incident, responding to an incident does require specific expertise. Even though, technically, SOPs[94] should have been devised for each institute to ensure business continuity, along with, e.g., a backup *and* restore plan, an ongoing incident might exceed local capabilities, as well as resources of distributed teams. No matter if it is a security incident[95] or another risk materializing[96], each incident start with a chaos phase directly after detection. During this phase, situational awareness has to be created, ensuring that further steps can be rooted in concious decisions, including the choice of the correct SOP and handling of non-technical user interactions [17, 21].

However, if local capabilities are insufficient in relation to the dimensions of the incident, or additional parameters hinder the response, the chaos phase might prolong. In either case, bringing in external support can improve incident response, if that external support either fills capability gaps due to incident-specific specializations[97], provides additional personnel resources, or can also provide infrastructural resources necessary for incident handling.

By building these capabilities in-house within the proposed team, an integration of these resources in incident SOPs is more easily possible than with, e.g., a consulting retainer. At the same time, using an in-house team allows integrating it in the existing community, further aiding implicit coordination[98] which can be essential during incident response [17].

Expertise Point-of-Contact We suggested to create distributed expertise and specializations within the MPG in Section 5.2.2. Effective curating of existing expertise and promoting interaction between teams and dissemination of information which expertise is available there is a continuous, non-technical task. At the moment, the MPG already has self-coordinating structures enabling this process, e.g., with an annual conference of operations personnel within the MPG and an existing community structure.

[93] This might also be an interesting appeal for (temporarily) joining this team, as it provides variation to the overall workflow.

[94] Standard Operating Procedures split a process into into individual steps which should be iteratively followed to handle an event in a standardized way. By improving reliability and consistency, while providing clear instructions, SOPs are an invaluable element in quality assurance. Check-lists as used in (especially) aviation are the most commonly known example of how SOPs are implemented.

[95] For example, a large-scale infection with ransomware.

[96] Consider a fire destroying major parts of an institute's digital infrastructure.

[97] For example, familiarity and more frequent experience/trainings with security incident handling SOPs.

[98] Essentially, leveraging better coordination due to people already knowing each other.

However, with other activities suggested for the support team, adding a non-technical role to the team to encourage would leverage synergies between, e.g., frequent contact with different MPIs due to other activities and dissemination/coordination. Hence, we suggest to consider integrating such a non-technical role/component into the team. However, in the context of this role, it is instrumental to consider also some additional requirements, see 'Challenges and Opportunities' below.

Evangelizing The process of 'evangelizing' or 'being the champion' is an integral function in the seamless introduction of new practices or technology to an organization. It is a common concept in security [2, 11], especially in the context of SDL [15]. Anecdotal observations indicate the importance of this mechanic in other aspects of digital infrastructure operations as well, for example, in the introduction of IPv6. Similarly, Limoncelli et al. suggest catering towards champions in the user-base to aid migrations [21].

To leverage this mechanic, we suggest to enable the support team to engage in evangelizing for technology and security improvements, specifically, if these technologies are not covered by distributed expertise, see above. Similar to the SDL approach, this would provide a measure to the MPG to improve technology adoption (including operational security and an OE framework), for new technology[99] and ideal tools to address issues encountered. Specifically, [redacted] improve the ultimate outcome of this intervention, and strengthen the overall security posture of the organization, while also catering towards user acceptance of the change.

Challenges and Opportunities Overall, the suggested support team could be an effective measure to address key operational challenges of running a research network. Even though most challenges the team is supposed to directly address are not security-specific per-se, they regularly pertain to the root-causes of security issues[100]. Similarly, the dynamic nature of work in this team might make it an ideal lever for recognition of well-executed operation or the creation of distributed expertise[101]. For example, if a high OE level permits, local experts could temporarily join the team to evangelize for their technology of expertise, effectively focusing on improving, e.g., following the example from Section 5.2.2, focusing on encouraging teams across the MPG to improve their [redacted].

Nevertheless, we also see challenges in the implementation of such a support team. As such a team would be a semi-central entity, organizationally grouping it with existing organizational units might lead to a perception of a paternalistic/top-down approach among operators in MPIs. This would significantly inhibit the effectiveness of the team. Similar effects might occur if non-technical roles for community building do not engage the community with sufficient care. Addressing this issue, i.e., preventing these risks from materializing is imperative for a successful implementation. Hence, we suggest to approach an implementation of these suggestions in a participatory way, ensuring bottom-up involvement.

Furthermore, we do see risks in terms of talent acquisition for roles within such a team. Despite existing expertise being available within the MPG which can partially fill some roles, this personnel is usually already bound in day-to-day operations. Tensions may arise if roles in the support team provide

[99] For example, IPv6, for which we saw limited adoption during our scans.

[100] For example, in the context of research infrastructure.

[101] Limoncelli et al. suggest using 'interesting projects' as a reward mechanic.

non-monetary appeal[102], either leading to issues when personnel would (permanently) move to such a position, leaving work behind within their original team, or if such roles are filled with personnel not already working in the MPG. Hence, appropriate measures must be considered to mitigate this risk.

Similarly, the required versatility, technical capabilities, and 'out-of-the-box' mindset necessitates a profile which can usually attain compensation in excess of what the TVoeD allows. Hence, non-monetary incentives will have to play a major role in talent acquisition. This includes the ability for remote work[103], ability to contribute to open source projects, following projects out of interest[104], conference attendance, research activity[105], and–more generally–interesting work[106].

Ultimately, the correct implementation of the suggested support team is a complex process. For a more comprehensive recommendation on the introduction of such a team, please request an additional report on 'Establishing a Centralized Operational Support Team'.

5.3 Summary

In summary, we make several recommendations towards improving the security posture of the MPG. Based on our analysis in Section 4, the majority of these recommendations is not directly targeted at security effects, but underlying processes and mechanics that lead to security issues as symptoms. As such, we note that the technical recommendation we provide significantly rely on the introduction of accompanying governance mechanics and organizational actions to be effective.

Furthermore, given Section 4.1.1 and Section 4.2, in conjunction with two events in relation to [redacted][107], we caution against the implementation of a standard 'enterprise level' security approach utilizing (service) centralization, separation of authority and responsibility, top-down/'push' management practices, and forced standardization of environments and software stacks. We argue that, given the operational requirements of the MPG and how they diverge from standard enterprise environments, such an approach carries significant risks:

- Inhibiting research activity by creating overhead for researchers as they have to align with centralized infrastructure and services[108].
- Forfeiting the integral isolation of assets in the current decentralized infrastructure, i.e., enabling easier lateral movement beyond individual institutes after a breach.
- Might be impacting the scientific integrity of the MPG by creating implicit dependency relationships with suppliers of centralized and outsourced services [10].
- Increasing staff turnover in operations teams, as a centralized approach changes the working environment in a way that reduces non-monetary incentives of working for the MPG, while compensation cannot compete with that found in the industry offering similar working environments (lacking non-monetary incentives).
- Decreasing security similar to the research inhibiting mechanics, as users and/or distributed teams utilize alternate solutions (Shadow IT) to reach their goals, if these are not provided in a straight-forward way by centralized infrastructure, or if the service migration in the course

[102] See Section 4.1.6.

[103] Given scope and purpose of the team, a location-specific implementation may not be organizationally required anyway.

[104] See the 80/20 rule in the early days of Google, which enabled employees to spend up to 20% of their time on self-motivated projects.

[105] Ranging from academic research in CS related fields to more applied security/systems work.

[106] See, for example, the matter of a mandate to engage with the organization's networks in a non-destructive way.

[107] Note that we did not establish a direct causal relationship, even though we conjecture it to be likely.

[108] See the 'cattle approach', and the connection between centralization and a requirement to make things the same that are not [9].

> of centralization breaks existing workflows. Similarly, such a service migration likely creates a significant debt of 'snowflake'[109] systems which can not be directly centralized, and hence are kept as 'permanent temporary solutions' [21].

We acknowledge that ensuring IT security for the digital infrastructure of the MPG is a significant task, and note that following our recommendations might be a more difficult approach to improving IT security. However, we argue, that our recommendations accept security as an effect of proper operation and addresses root-causes instead of focusing on security-as-a-goal, thereby creating a more sustainable organizational change towards a secure digital infrastructure.

This goes beyond the common concept of 'security-by-design', which still compartmentalizes 'security' as a *goal* during the development of systems and moves the responsibility for 'security' into the design and development stage. Thereby, security-by-design leaves out the aspect of continuous maintenance necessary for secure operation, as well as the operational requirements that emerge during the use of a system, e.g., consider a securely designed system which is perceived as operationally complex in practice and, hence, users evading that system.

[109] This term in digital infrastructure refers to systems that are each unique, similar to snowflakes. Also, similar to snowflakes, one is beautiful, while many are a blizzard, see Limoncelli et al. [21].

6 Comparison to the [redacted] Report

In this section, we will compare our results, scope, methodology, analysis, and recommendations to the report based on a scan by the [redacted] project previously received by the president of the MPG in early 2023.

6.1 Dataset Description

We received a copy of a ZIP file containing information provided by [redacted] on [redacted]. Please see Table 4 for an overview of the files in this archive. This archive has been received via [redacted]. [redacted] OCR has been executed on supplied PDF files and XLS files have been converted to CSV files, with one CSV file per sheet. [redacted] We base our analysis on [redacted] the files marked as ‘[redacted]’ in Table 4 [redacted]

Table 4: Overview of received files.

File	SHA256	Origin
./Action_Items_Report_mpgesellschaft.pdf	[redacted]	[redacted]
./Action_Items_Report_mpgesellschaft.txt	[redacted]	[redacted]
./analyze.sh	[redacted]	[redacted]
./Cert_analyzer.pdf	[redacted]	[redacted]
./Cert_analyzer.txt	[redacted]	[redacted]
./Cert_analyzer.searchable.pdf	[redacted]	[redacted]
./Zugangsdaten/mpg.de_[redacted].csv	[redacted]	[redacted]
./[redacted] Study and Deliverables.pdf	[redacted]	[redacted]
./mpg.xlsx	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Network tests.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.External connections.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Action Items.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.External assets.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Internal connections.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Web tests.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.DNS tests.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Vulnerabilities.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Organizational FQDNs.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Mail tests.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Internal Cloud assets.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Login pages.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Certificate tests.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.TLS tests.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.External Cloud assets.csv	[redacted]	[redacted]
./mpg.xlsx-csv-Files/mpg.Port scan.csv	[redacted]	[redacted]
./Readme.md	[redacted]	[redacted]
./Vuln_analyzer.pdf	[redacted]	[redacted]
./Vuln_analyzer.txt	[redacted]	[redacted]
./Vuln_analyzer.searchable.pdf	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

./Action_Items_Report_mpgesellschaft.pdf This document lists [redacted] immediate action items of a criticality between [redacted] and [redacted]. Of these, [redacted] are hosts supporting SSLv2, [redacted] are outdated WordPress instances, and [redacted] are sites that allow logging in without using HTTP.

./Action_Items_Report_mpgesellschaft.txt A plain-text (OCR) version of ./Action_Items_Report_mpgesellschaft.pdf.

analyze.sh A bash script testing whether FQDNs from mpg.xlsx-csv-Files/mpg.Vulnerabilities.csv resolve, and if they do, what the status of the associated domain is.

./Cert_analyzer.pdf This document lists certificates presented by MPG hosts, for which a variety of issues has been established, see Table 5.

Table 5: Overview of reported certificate issues.

Category	Domains	Importance
Expired certificates	[redacted]	[redacted]
Certificates mixing CNs/DNSAltNames below mpg.de and not below mpg.de	[redacted]	[redacted]
Weak signature algorithm	[redacted]	[redacted]
Certificate expiring in less than a week		[redacted]
Certificate contains at least one CN/DNSAltName of a 'vulnerable domain'	[redacted]	[redacted]
Certificate expiring in less than a month		[redacted]
Wildcard certificate mismatch		[redacted]
Certificate mixing CNs/DNSAltNames from different subtrees of mpg.de	[redacted]	[redacted]
A 'vulnerable domain' uses a certificate also valid for this domain	[redacted]	[redacted]
Using a Let's Encrypt Certificate	[redacted]	[redacted]
Certificate expiring in less than 60 days	[redacted]	[redacted]
Certificate expiring in less than 90 days	[redacted]	[redacted]
Certificate is the first issued for this domain	[redacted]	[redacted]
Certificate issuer used for the first time	[redacted]	[redacted]
New certificate issued during the last month	[redacted]	[redacted]
Presented certificate is not the newest certificate for that domain	[redacted]	[redacted]

./Cert_analyzer.txt A plain-text version of ./Cert_analyzer.pdf, i.e., an OCR of the PDF.

./Cert_analyzer.searchable.pdf A version of ./Cert_analyzer.pdf which includes text, i.e., is searchable, as ./Cert_analyzer.pdf contained all pages as images.

./Zugangsdaten/mpg.de_[redacted].csv This is a CSV file listing [redacted] credentials. For each entry, several data points are provided:

ip An IPv4 address.

url An (HTTP) URI, in all cases for services below mpg.de.

login A username.

password A redacted string, presumably supposed to contain a password, but redacted by replacing all characters with [redacted].

computer_name The name of a host.

operating_system The operating system installed on a host.

malware_path An absolute path on a windows file-system.

date A non-ISO8601 date string without timezone information.

Based on these contents we assume that this file contains credentials received via a threat intelligence feed that, for example, collects data retrieved from malware credential collection/control systems [3]. The remaining data then corresponds to additional metadata that could be collected about the infected hosts. In total, information from [redacted] hosts identified by their IP address is listed, concerning [redacted] unique URLs on [redacted] unique domains and [redacted] unique logins. Even though some critical domains are present ([redacted]), the majority of entries belongs to public services, e.g., application portals, and public research data shares.

./[redacted] Study and Deliverables.pdf A letter by [redacted] titled 'Study of Vulnerabilities', describing the scope of the work that forms the basis of the report, and describing the deliverables. The letter states that [redacted] performed [redacted] Information collected in that stage is then used for analyzing the security posture of an organizations, [redacted]

The listed deliverables are:

"action items report" An overview of immediate action items. We assume that this refers to ./Action_Items_Report_mpgesellschaft.pdf.

"Certificates issues summary" Issues with certificates, including a ranking of their severity. We assume this refers to ./Cert_analyzer.pdf.

"Vulnerabilities summary" A list of DNS related issues and vulnerabilities. We assume this refers to ./Vuln_analyzer.pdf.

'An Excel File' A file with details of all issues found [redacted]. Given that this refers to a single file, we assume it references ./mpg.xlsx.

Overall, this letter provides limited information on the actual methodology used. Instead it remains general, even though common place statements are underlined with buzz-words or terms likely known to decision makers. For example, the section on [redacted] starts with [redacted] a statement too general to carry meaning, while focusing on well-known terms [redacted] thereafter, i.e., [redacted], without providing a connection between these terms [redacted]. Similarly, the [redacted] section states that [redacted]. i.e., refers to vulnerabilities likely to have received attention ([redacted]), without making a statement on what [redacted] entails.

./mpg.xlsx An Excel file containing various datasets. The individual sheets correspond to files in ./mpg.xlsx-csv-Files/. Hence, for a detailed description, we refer to these individual paragraphs.

./mpg.xlsx-csv-Files/mpg.Network tests.csv This file lists port-scan results for [redacted] different hosts attributed to the MPG. For each host, it provides [redacted]. Over these [redacted] hosts it finds [redacted] unique open ports ranging [redacted] services, see Table 6.

Table 6: **Overview of open ports.**

Port	Service	Count	Port	Service	Count	Port	Service	Count
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]		[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

./mpg.xlsx-csv-Files/mpg.External connections.csv This file lists [redacted] 'external connections'.

Hyperlink [redacted] occurrences.

Authoritative nameserver [redacted] occurrences, of which the Top 3 entries are [redacted]. These Top 3 are responsible for [redacted] ([redacted]).

NS DNS record [redacted] occurrences, of which the Top 3 entries are [redacted]. These Top 3 are responsible for [redacted] ([redacted]).

Script inclusion [redacted] occurrences, of which most are common CDN locations for such assets ([redacted]).

CSS inclusion [redacted] occurrences, of which most are common CDN locations for such assets ([redacted]).

MX DNS record [redacted] occurrences, of which [redacted] are [redacted].

Image inclusion [redacted] occurrences.

Iframe inclusion [redacted] occurrences.

WHOIS email [redacted] occurrences.

Font [redacted] occurrences.

CNAME record [redacted] occurrences, of which the Top 3 destinations are [redacted], accounting for [redacted] of all entries.

SOA DNS record [redacted] occurrences, of which the Top 1 ([redacted]) accounts for [redacted] of all entries.

XML HTTP request [redacted] occurrences.

Ping request [redacted] occurrences.

Redirect [redacted] occurrences.

Script inclusion (Redirected) [redacted] occurrences.

Image inclusion (Redirected) occurrences.

CSS inclusion (Redirected) [redacted] occurrences.

Media inclusion [redacted] occurrences.

Media inclusion (Redirected) [redacted] occurrences.

Based on these statistics, we assume that 'external connections' refers to all HTTP or DNS connections that are not strictly tied below the main domain of the MPG (mpg.de).

./mpg.xlsx-csv-Files/mpg.Action Items.csv This file lists [redacted] 'action items'. Based on the number of entries and naming scheme, we assume that this file is a superset of ./Action_Items_Report_mpgesellschaft.pdf. The categories in this file are [redacted].

The general classes of issues are:

Fix Web issue [redacted] occurrences; This category consists of missing [redacted] headers.

Fix Mail server issue [redacted] occurrences; This category is dominated by missing [redacted] records. Furthermore, minor issues [redacted][110][redacted] are listed.

Vulnerable application [redacted] occurrences; Here, sub-categories exist, ranging from web frameworks [redacted], over webservers [redacted] and server-side execution environments [redacted], to web-applications [redacted]. Criticality reaches from [redacted]

Fix PKI issue [redacted] occurrences; This category includes [redacted] sub-categories, technically forming a sub-set of the categories found in ./Cert_analyzer.pdf. However, the occurrences between both files do not match. Here, we find:

- 'Certificate will expire within 7 days' (urgency 7)
- [redacted] 'External domain mismatch' (urgency 7)
- [redacted] 'Internal domain mismatch' (urgency 5)
- [redacted] 'Self-Sign Certificate' (urgency 7)
- [redacted] 'Weak certificate algorithm' (urgency 7)
- 'Symantec certificate' (urgency 7)
- [redacted] 'Certificate will expire within 30 days' (urgency 6)
- [redacted] 'Expired certificate' (urgency 5)
- [redacted] 'Wildcard domain mismatch' (urgency 4)
- [redacted] 'Old certificate' (urgency 2)

Fix DNS issue [redacted] occurrences; The category contains [redacted] subcategories with a criticality ranging from [redacted] to [redacted]. The subcategories are:

- [redacted] 'Inconsistent resolution from declared nameservers' (urgency 6)
- [redacted] 'Inconsistent resolution' (urgency 6)
- [redacted] 'Nameserver is inconsistently resolved' (urgency 6)
- [redacted] 'Unresolved authoritative nameserver' (urgency 6)
- [redacted] 'The domain is resolved to reserved IP.' (urgency 6)
- [redacted] 'Single authoritative nameserver' (urgency 5)
- 'Single nameserver in NS record' (urgency 5)
- [redacted] 'Authoritative nameservers are not geo-separated' (urgency 4)
- [redacted] 'Nameservers are not geo-separated' (urgency 4)

Fix Network issue [redacted] occurrences; The category contains [redacted] subcategories with a criticality ranging from to . The subcategories are:

- [redacted] 'MySQL accepts connection' (urgency 8)
- [redacted] 'MS SQL accepts connection' (urgency 8)
- 'Postgres accepts connection' (urgency 8)
- [redacted] 'Postgres no SSL' (urgency 7)

[110] [redacted]

- [redacted] 'SFTP weak authentication method' (urgency 2)
- [redacted] 'FTP weak authentication method' (urgency 2)
- [redacted] 'SMTP deprecated port' (urgency 2)

Login over HTTP is possible [redacted] occurrences without further diversification ([redacted]).

Dangerous hyperlink connection [redacted] occurrences without further diversification ([redacted]).

SSLV2 [redacted] occurrences without further diversification ([redacted]).

Wordpress user enumeration [redacted] occurrences ([redacted]).

Login only over HTTP [redacted] occurrences without further diversification ([redacted]).

Deprecated application [redacted] occurrences; All of these relate to instances of [redacted] ([redacted]).

Dangerous script inclusion occurrence with an [redacted], concerning [redacted].

./mpg.xlsx-csv-Files/mpg.External assets.csv This file lists [redacted] 'External assets'. For each item, an [redacted] are specified. However besides the self-descriptive values for [redacted], no information on the nature of these values is provided. We assume that the [redacted] refers to MPG controlled systems referencing [redacted]. We have no information concerning the nature of the [redacted] field, even though it implies that [redacted].

Overall, we find [redacted] 'Hyperlinks', with the remaining items ranging from DNS dependencies to script and web-asset inclusions.

./mpg.xlsx-csv-Files/mpg.Internal connections.csv This file lists [redacted] 'internal connections'. For each item, a [redacted]. [redacted] seem to be similar to ./mpg.xlsx-csv-Files/mpg.External assets.csv.

Here, the two largest categories are [redacted] 'Hyperlinks', and [redacted] 'Authoritative nameservers'.

./mpg.xlsx-csv-Files/mpg.Web tests.csv This file contains [redacted] records related to basic web configuration issues, for example, missing HSTS headers, or a missing referrer-policy. For each entry, [redacted] is provided, along with a 'Grade info'. The latter seems to be automatically generated from an automated security analysis as, e.g., typos (doubled dots at the end of sentences) repeat throughout the file.

./mpg.xlsx-csv-Files/mpg.DNS tests.csv This file lists [redacted] items related to [redacted], providing [redacted] categories ([redacted]). In total, [redacted] entries do not have any additional information beyond [redacted]. The remaining items are split over several common [redacted][111].

./mpg.xlsx-csv-Files/mpg.Vulnerabilities.csv The contents of this file relates to the 'Vulnerable Applications' category from ./mpg.xlsx-csv-Files/mpg.Action Items.csv. However, in contrast to the abstract 'urgency' in ./mpg.xlsx-csv-Files/mpg.Action Items.csv, this file now states a numeric field as being a CVSS score[112].

[111] See, for example, 'private' addresses as defined in RFC1918 [24].

[112] See [12].

./mpg.xlsx-csv-Files/mpg.Organizational FQDNs.csv This file lists [redacted], providing further metadata like [redacted]. However, there is no apparent distinction between hostnames/FQDNs and zones.

./mpg.xlsx-csv-Files/mpg.Mail tests.csv This file lists [redacted] findings for email configuration related issues. The tags for individual systems are [redacted] Furthermore, [redacted]. The file does not provide further elaboration on why a domain was included. Furthermore, it does not provide any information on whether mail sending or receiving is concerned.[113] Finally, of [redacted] entries, seem to be web-hosts (www.), and [redacted] entries have a null MX, common practice to signal that email receipt is disabled for a name[114].

./mpg.xlsx-csv-Files/mpg.Internal Cloud assets.csv This file lists [redacted] hostnames of internal cloud assets. Overall, it contains information on [redacted] cloud providers, specifically: [redacted]. It is not clear why [redacted] was included, given comparable entities [redacted], which are also used by the MPG, were not included[115].

./mpg.xlsx-csv-Files/mpg.Login pages.csv This file lists [redacted] login pages of MPG systems over [redacted] categories, specifically [redacted]. In addition, for each item, information on whether [redacted] is provided.

./mpg.xlsx-csv-Files/mpg.Certificate tests.csv This file lists [redacted] entries. Each entry consists of [redacted]. Information in this file corresponds to the information found in ./Cert_analyzer.pdf. Of the [redacted] entries in this file, [redacted] have no findings.

./mpg.xlsx-csv-Files/mpg.TLS tests.csv This file contains [redacted] TLS related observations. Each entry consists of [redacted], and a [redacted] may be listed multiple times with different [redacted]. Tests range over a variety of known [redacted], ranging from supporting outdated versions of [redacted], and common attacks like [redacted]. Additionally, [redacted] entries do not have a finding. Please note that this file does not list the port and type (TLS vs. StartTLS) for any of the findings,

./mpg.xlsx-csv-Files/mpg.External Cloud assets.csv This file lists [redacted], of which [redacted] are hyperlinks. The [redacted] is a superset of what we found in ./mpg.xlsx-csv-Files/mpg.Internal Cloud assets.csv. Specifically, now we find: [redacted].

./mpg.xlsx-csv-Files/mpg.Port scan.csv This file contains [redacted] entries, presumably port scans. Each entry contains [redacted]. In total, [redacted], which expose [redacted]. The two most frequent ones are [redacted]. No information on [redacted][116] is provided.

./Readme.md This file contains further information on the shared version of the report, and has likely been created by [redacted] the MPG [redacted].

./Vuln_analyzer.pdf This file lists an aggregated view of the vulnerability information listed in other parts of the report.

./Vuln_analyzer.txt A plain-text version of ./Vuln_analyzer.pdf.

[113] Please see Holzbauer et al. for a comprehensive discussion on email sending measurements [14]; Furthermore, an online test is available at https://email-security-scans.org/.

[114] See, RFC7505 [20].

[115] For a comprehensive perspective on cloud use in academic contexts, please see [10].

[116] See Section 2.

./Vuln_analyzer.searchable.pdf A searchable (OCR included) version of ./Vuln_analyzer.pdf.

[redacted]

6.2 Comparison on Methodology

The [redacted] report does not provide any information on its methodology. This means that we cannot compare it to our methodology, see Section 2.

6.2.1 Estimation and Assumptions on [redacted] Methodology

Even though [redacted] did not share methodological information, based on the dataset description in Section 6.1, we can make specific assumptions about the used methods:

Automation As noted in Section 6.1, the structure of the report, the volume of information, and repeated typos lend themselves to assuming a semi-automated or fully automated process. This also aligns with, e.g., our observation regarding hosts with zero MXes being included in the report.

Focus on HTTP/Web Even though the report contains a port-scan in ./mpg.-xlsx-csv-Files/mpg.Port scan.csv, the majority of reported vulnerabilities is web-related, with the exception of specific TLS, DNS, and Mail findings.

Initial Asset Discovery The file ./mpg.xlsx-csv-Files/mpg.Organizational FQDNs.csv contains a list of names related to the MPG. Given the remainder of the evaluation focuses on *'the organizational domain'*, it is conceivable that the initial discovery focused on assets below 'mpg.de'. Furthermore, some FQDNs are incomplete, i.e., they start with a delimiter (.example.com). We assume that the file has been (initially) gathered using certificate transparency logs, and further hosts in relation to the MGP were then discovered, e.g., using traceroutes, RIR data, or crawling. Please see Vermere et al. [30] for a comprehensive overview of asset discovery procedures.

Crawling The report, as well as ./[redacted] Study and Deliverables.pdf emphasize the exploration of dependency trees in digital infrastructure. Based on the report assets and the high frequency of hyperlinks as a connection type across several data files, we hence assume that, from the initial list of potential assets, the major technique used was web crawling.

Focus on Known Vulnerability Identification Considering the list of identified vulnerabilities, see, e.g., ./Action_Items_Report_mpgesellschaft.pdf and ./mpg.xlsx-csv-Files/mpg.Vulnerabilities.csv, and in conformance with our assumption regarding an at least partially automated process, we assume that the bulk of vulnerability analysis was performed using an automated framework which compared presented version strings from delivered sites and port banners to a CVE and/or vulnerability database. The urgency is then looked up based on CVE(s) assigned to the identified software and version combination. In specific cases, e.g., when dealing with certificates, supported TLS versions, and DNS/mail related services, it is likely that an active measurement approach was used to retrieve concerned resources and/or test whether, e.g., [redacted].

This framework may have been self-developed or an off-the-shelf solution like, for example, Nessus[117]. However, from the data itself, we cannot make a conclusive statement on the used software.

Frequency and Time of Evaluation The report by [redacted] does not contain indications of the exact time at which the assessment was conducted, or whether multiple scans/evaluations were conducted. The file ./Zugangsdaten/mpg.de_[redacted].csv places the data in early [redacted]. However, ./Cert_analyzer.pdf lists [redacted] domains with certificates expiring within a week. The earliest expiring listed certificate expires on [redacted], the last expiring one expires [redacted]. This places the time at which data for ./Cert_analyzer.pdf was gathered and/or analyzed on [redacted]. At the same time, ./[redacted] Study and Deliverables.pdf dates from [redacted]. In summary, this means that we can place the scans having taken place between [redacted] and [redacted], but are unable to establish whether multiple scans were conducted to revisit and verify earlier results.

6.2.2 Similarities and Differences in Methodology

Here, we compare our approach to that we assume to have been in place for [redacted] based on our assumptions above. Please note, that the publication of ground-truth information on [redacted]'s methodology may demonstrate our assumptions to be, at least partially, inaccurate, therefore also influencing the outcome of our comparison below. Until such further information becomes available, we will however make this assessment based on these assumptions.

Automation While both, our scan and the process used by [redacted] start out from a (semi) automated data collection process, we additionally performed a qualitative analysis of potential vulnerabilities, see Section 2 using two independent analysts, following qualitative research best practices [22]. This allows us to identify configuration issues and vulnerabilities beyond well-known issues. Especially given issues with self-developed software, mishandled enterprise systems, and non-common or sector specific outdated applications, this enables us to identify more potential issues. Furthermore, our additional manual analysis process enables us to provide more detailed guidance on mitigation actions, and enables an analysis of security posture beyond individual vulnerabilities.

Focus on HTTP/Web Our approach follows a TCP/Service-Level focused approach, in which we start from network segments, identify open ports and then analyze running services. [redacted]'s approach appear to be web-centric, which limits perspectives on vulnerabilities mostly to web-related services, especially those running on standard web ports. This means that findings we have, for example concerning outdated network appliances[118], do not show up as action items in [redacted]'s methodology, despite being of high importance. Please also see the section on 'Scope' below.

Initial Asset Discovery & Crawling Both, [redacted] and us start out from FQDNs collected from public and non-public data feeds. However, while we also additionally leverage RIR data to identify network segments utilized by MPIs, we filter this data to ensure we limit those address spaces to MPIs' networks. While this reduces our visibility and may exclude networks not used by MPIs, it also reduces the amount of collateral damage.

[117] https://www.tenable.com/products/nessus/nessus-professional

[118] [redacted]

In contrast to our work, [redacted] explores web dependencies and the use of external cloud assets. Furthermore, the focus on web-dependencies means that [redacted] fully explores site-maps for MPG related web presences, reaching far beyond the organization itself, with a majority of observations being related to major CDN based library inclusions and external hyperlinks/resource embeddings. However, this also means that the scope for [redacted] is significantly less discrete, see Section 6.3 below.

Focus on Known Vulnerability Identification While [redacted] focused on the identification of well-known vulnerabilities, we took a more nuanced approach in line with our objective to assess general security posture, see Section 2. While our approach also partially considers software versions to establish whether a system is orphaned, we do not perform a detailed analysis of software versions, especially not software and library versions deeper in the web-stack, see Section 2.

While this may lead us to missing crucial vulnerabilities, e.g., in outdated [redacted] instances, we consciously decided to limit our scope here, as it reduces the amount of potential false-positives, thereby benefiting our objective of estimating overall security posture reaching beyond updates itself. In more plain terms, we decided to reduce noise by, e.g., false positives due to static site exports (web applications) or long-term-support versions (see, for example, stable versions of Linux distributions backporting security fixes without changing version numbers, in favor of gaining a better exploration of security posture overall, e.g., by identifying orphaned systems, structural exposure issues etc.

Both approaches are viable and issues discovered in either should be mitigated. However, the approach by [redacted] requires additional follow up work for verification and analysis, which has not taken place in the report by [redacted], judging from the amount of findings and reporting for these findings remaining automated, see the detailed file descriptions above.

Frequency and Time of Evaluation The scan conducted by [redacted] has been conducted in [redacted], while results became available in [redacted]. Our scans have been conducted from [redacted] to [redacted], with the report becoming available in late [redacted]. As such, the overall processing time between the two reports is similar.

Concerning the frequency of scans, no information is available that [redacted] conducted repeated scans, potentially limiting the perspective on non-permanent hosts, e.g., the [redacted] we discuss in Section 3 and for individual institutes in the appendix.

6.3 Comparison on Scope

[redacted] did not provide any information on the selected scope of their engagement. Hence, a ground-truth comparison is not possible. We therefore base our comparison again on assumptions we make based on what we can infer from the data provided in the [redacted] report.

6.3.1 Estimation and Assumptions on [redacted] Methodology

Included Hosts The file ./mpg.xlsx-csv-Files/mpg.Port scan.csv lists portscans for [redacted] hosts, of which [redacted] have open ports. We used bttf-whois [28] to attribute these addresses to ASNs as they were announced on [redacted].

Of those [redacted] hosts, [redacted] addresses are in networks announced by [redacted], [redacted] have been announced by [redacted], and by other European research networks ([redacted]). This leaves [redacted] addresses, of which [redacted] were not announced at the time ([redacted]), [redacted] are RFC1918 [24] addresses, and three are reserved (0.0.0.0, 127.0.0.1 (localhost [5]), and 192.0.2.1 (TEST-NET-1 [1])). This leaves [redacted] hosts of third parties, including [redacted].

Of the [redacted] hosts with open ports, [redacted] addresses are in networks announced by [redacted], [redacted] have been announced by [redacted], and [redacted] by other European research networks ([redacted]). This leaves [redacted] hosts of third parties, including [redacted].

Included Names There are [redacted] different names in mpg.Organizational FQDNs.csv, of which [redacted] have a direct relation to an MPG related domain name ([redacted]). This stretches over [redacted] PSL public suffixes, with the most common ones being [redacted]. The majority of indirectly related domains seems to concern event- and collaboration/research specific sites[119].

External assets include [redacted] unique names, of which at least [redacted] have been exposed to a vulnerability assessment of some sorts based on these sites having received additional remarks indicating a vulnerability. The external assets which have been subjected to a vulnerability assessment cluster over [redacted] different PSL private suffixes and [redacted] different PSL public suffixes. The top private suffix is [redacted] with [redacted] different external assets, followed by [redacted], [redacted] and [redacted]. In general, likely due to the inclusion of hyperlinks as external assets, the top-hits here include mostly research organizations and German universities. However, we also find national and international government organizations like [redacted]. Furthermore, taking a look on public suffixes of domains subjected to a vulnerability scan, naturally, the top hit is [redacted], followed by [redacted] and [redacted].

However, we also find that apparently also names under potentially sensitive public suffixes/ccTLDs—especially given the estimated time of the scans—have been included, specifically [redacted] and [redacted]. Sites in [redacted] that were subjected to a vulnerability assessment include [redacted], the [redacted], which—following the archiving data on archive.org[120]—has seen limited reachability from [redacted]. This may be related to an increase in [redacted]. Furthermore, the news site [redacted] has been subjected to a vulnerability assessment. For [redacted], we find [redacted] academic sites [redacted] as well as one [redacted].

Even though the methodology of the [redacted] project has not been documented, i.e., we are not able to determine the addresses from which the scans by [redacted] originated, it is conceivable that vulnerability assessments are considered an attack, and the origin of those attacks could have been traced back to [redacted].

[119] See Section 4 why this is to be expected.

[120] See [redacted].

6.3.2 Similarities and Differences in Scope

Here, we compare the scope of our assessment with the scope we estimated for [redacted] based on the provided data.

Included Hosts As documented in Section 2, we used a set list of networks which we scanned, see Table 1. Even though [redacted] does not provide such a list, we can assume that hosts/addresses noted in the report were in fact included in the scope. This means that we can compare our scope to that of [redacted] within four parameters:

- The number of hosts included in [redacted]'s port-scan, but not in ours.
- The number of hosts included in our report, but not in [redacted]'s.
- The number of hosts noted as out-of-scope for our report[121], but that were included by [redacted].
- The number of hosts included in both reports.

Additionally, we can also evaluate these parameters over the findings in terms of open ports.

In total, we find that [redacted] included [redacted] hosts, of which [redacted] are also included in our set, while [redacted] hosts are included in the [redacted] scan while not being part of our evaluation. Of those, [redacted] hosts included in the [redacted] report are part of networks which were later excluded from our evaluation, see Section 2. Our evaluation, in total, spans networks covering [redacted] hosts.

The [redacted] found open ports for [redacted] hosts, while we found open ports for [redacted] hosts[122]. Overlap between these two sets is, however, limited, as only [redacted] hosts with open ports are in both sets, of which we find the same open ports for hosts, while [redacted] finds more open ports in [redacted] cases, and we find more open ports in [redacted] cases. In addition, our set contains [redacted] hosts with open ports not found by the [redacted], while [redacted] found [redacted] hosts with open ports not showing up in our scans, of which [redacted] are in networks which were later excluded from our evaluation, see Section 2. Of the remaining [redacted] hosts with open ports found by [redacted], [redacted] were part of the addresses we scanned. Concerning the [redacted] we found and [redacted] did not find, [redacted] are also listed in ./mpg.xlsx-csv-Files/mpg.Port scan.csv. A possible explanation for this reduced visibility on our part would be that these hosts did not respond to our initial ICMP requests, or our conscious choice of a tight scope in terms of included networks to reduce collateral damage, also see the limitations we outline in Section 2.7. As we do not have any information on the actual scan methodology and request frequency of the [redacted] project, we can only speculate on the reasons for hosts we observe to have open ports, even though they did not. Possible options include that we scanned a larger/different list of ports, hosts were only intermittently online[123], were deployed after the [redacted] scan, the scanning frequency of the [redacted] project caused packet-loss, and/or [redacted] used a too low retry value to be resilient against packet loss.

Included Names Our scans did not focus on names. Even though our evaluation did recover various names, this was not part of our objective. However, by using a more restrictive scope with a set list of networks to scan, we can be certain to not have accidentally scanned networks of foreign entities. This is not the case for the methodology used by [redacted].

Furthermore, by documenting our methodology and following measurement

[121] See Section 2 on networks that were excluded.

[122] Please note that the [redacted] report does not distinguish between UDP and TCP. We hence assume all ports to be TCP.

[123] Hence, were only observable in our repeated scans.

best-practices, operators could inform us about issues caused by our scans and out-of-scope systems, see Section 2. We have no information on how the scanning infrastructure operated by [redacted] was configured, and if best-practices for abuse handling were followed. Furthermore, we do not have any information on whether abuse requests by scanned networks were received and/or appropriately handled.

6.4 Comparison on Findings

Our comparison on scope follows from the assumptions on the [redacted] report outlined above. None of our critical findings were covered by the items listed in ./Action_Items_Report_mpgesellschaft.pdf. While some high severity findings (extremely orphaned systems still/only supporting SSLv2), high impact events were missed by [redacted]. This includes the compromised system found in our study, several outdated Enterprise applications[124], see Section 2.8.

Instead, findings and [redacted] from the [redacted] report focus on web based vulnerabilities, especially of front-facing web-services[125]. Web facing systems, of course, hold potential for reputation loss, if they are defaced[126], might be used to distribute malware, expose internal credentials if these are shared across systems, or allow lateral movement via co-located applications on the same system. However, usually, these systems are at least isolated via a DMZ from core systems holding research data/PII and internal infrastructure. As such, the compromise of a web presence is not comparable to the impact of, e.g., internal infrastructure, like research systems, groupware solutions, or building automation systems being compromised or used in any form of extortion[127], being compromised. Especially structural issues[128] have to be identified and mitigated to sustainably improve security posture.

Please note that this does not mean that web related issues should not be mitigated, or that a web server being compromised is not an issue. However, the organizational setup should function in a way that ensures systems do not become orphaned/outdated, while also ensuring that a compromise—if it takes places—does not allow lateral movement or subsequent compromises, and is quickly detected and mitigated.

6.5 Comparison on Recommendations and Analysis

The report of the [redacted] project does not contain an analysis of findings going beyond individual observations. Besides a list of 'action items', no recommendations are provided. Hence, we cannot compare our analysis and recommendations to the [redacted] report.

6.6 Summary

In summary, we find that the [redacted] report provides limited information in terms of used methodology. It leaves its scope undefined, and an analysis of the provided data indicates that the scope may have included systems of foreign entities. It does not provide an in-depth security analysis and misses important vulnerabilities while focusing its evaluation on a limited pre-determined set of mostly web-related vulnerabilities. This means that the scope and methodology is unsuited for assessing the security posture of diversified research infrastructure as found for, e.g., the MPG[129], as it will miss major issue clusters in the form of (misconfigured) enterprise systems, orphaned systems, and sector specific/in-house applications. Yet, despite using this

[124] [redacted]

[125] [redacted] are [redacted] vulnerabilities.

[126] An attacker changing a web presence to demonstrate they were able to compromise a system.

[127] See ransomware; Even though the term is commonly bound to attacks that encrypt information, e.g., building automation systems allow different extortion attacks as well.

[128] See our analysis in Section 4.

[129] See Section 4.

narrow perspective, analyzing the port-scans in the [redacted] report indicates a comparable (and sometimes exceeding) coverage of the MPG infrastructure, i.e., would have allowed for similar conclusions as ours, if the scan data would have been analyzed. However, the report does not provide such an analysis. Especially evaluations for off-port services and general security posture, as well as the major issue clusters we identified, are not part of the report.

Furthermore, automated analysis without manual refinement lead to a report that states large numbers in terms of detected vulnerabilities, while staying vague on the specific relationship to security, implications, and overall risk despite citing CVSS values. We did not find indications of manual verification of vulnerabilities, i.e., whether they are false-positives, and if this was performed, no information on this was included in the report.

Potential Impact If provided to decision makers, such a report could negatively inflict on the relationship between operations and management[130], or could trigger spontaneous activities in a bid to 'improve security'[131]. The reporting of high numbers for high criticality events (e.g., CVSS scores over 8) creates immediate negative visibility for operations, while the report does not include a contextualizing analysis discussing organizational mechanics and potential root-causes, effectively taking a 'blame focused' approach at issue remediation[132]. Similarly, spontaneous activities carry the risk of unintended consequences[133]. Furthermore, spontaneous activities may have a negative impact on non-monetary incentives, see Section 4.1.6, if pressure is applied[134], or if changes are rolled out that disregard the unique nature of research networks, again, see Section 4.1.2. In those cases, a report with limited contextualization and analysis as the one provided by [redacted] may negatively impact the security posture of an organization. Finally, in line with prior work on notification fatigue [29], a large report with many potential false-positives that are rated highly on the CVSS score also holds the potential of reducing operators willingness to engage with subsequent security evaluations, or take critical vulnerability assessments seriously, see also criticism of the CVSS score [23].

Please note that we did not assess whether similar effects materialized in response to the [redacted] report. If we describe anecdotes of such effects, these are observations based on a relative concurrence to the [redacted] report and do not imply a statement on a causal relationship.

[130] See Limoncelli et al. on 'visibility and perception' [21].

[131] See Dietrich et al. [7]

[132] See Kaur et al. for perspectives on just culture in infrastructure operations and lessons learned from the safety sciences [19, 18].

[133] Please see our observations on 2FA related issues found in the MPG in Section 3.

[134] [redacted]

7 Conclusion

Here, we briefly summarize our findings regarding our objectives and highlight key action items/immediate next steps. The objectives of our investigation were:

- Identify security issues that require immediate or immanent mitigation due to a high likelihood of causing events or incidents.
- Assess the overall security posture of the MPG, focusing on structural and organizational shortcomings.
- Identify action items and strategies to sustainably improve the security posture of the organization.
- Evaluate findings of this internal investigation against the information that was provided by the [redacted] project.

Findings In our scans, we find several critical issues and a plethora of other issues of varying severity. These must be mitigated. For this, we provide detailed information and mitigation recommendations in the appendix of this document, grouped per institute.

Security Posture Overall, we find that the security posture varies between different institutes, similar to the extend of infrastructure used by institutes. Even though some institutes demonstrate an increased number of critical or high severity issues, the overall security posture, especially in institutes with larger infrastructures is comparable to sector competition. Furthermore, by clustering our findings, we identify organizational root-causes in the nature of the digital infrastructure in the MPG as a research organization, proliferating the emergence of these issue-clusters. Our further analysis finds organizational mechanics relevant in maintaining secure digital infrastructure for the MPG. Specifically:

- A highly decentralized environment.
- Unique requirements of research infrastructure and researchers as users.
- Employee churn among scientific employees.
- Use of enterprise toolchains and infrastructure that is not tailored to the organization's requirements.
- Capability retention in digital infrastructure operations.

Especially these characteristics as a research network create a unique environment that is distinct from standard enterprise networks, restricting the applicability of standard enterprise practices to improve IT security. However, we also find that the high degree of decentralization in terms of operation is a major asset for the organization's security status, as it makes lateral movement after the compromise of an individual institute less likely.

Recommendations Based on our analysis, we suggest to improve IT security within the MPG by addressing root-causes, mostly focusing on bottom-up governance[135] and organizational improvements. A corner stone of these improvements is leveraging and expanding existing distributed expertise, improving network and service segmentation, the introduction of an operational excellence framework, and the formation of a team that actively addresses operational requirements of research infrastructure, ensuring high levels of operational excellence.

[135] Governance here does not mean a strict top-down approach as it is often understood. Instead, we refer to bottom-up governance, i.e., the process of nurturing emergent (self) organization and enabling self-improvement in an organization, see also D2.1 of CS4E [16].

Comparison to the [redacted] report Finally, we compare our results to the report by the [redacted] project received in early 2023. We find that the [redacted] report lacks a documentation of its methodology, and does not provide an analysis of its results. Based on the data provided in the report, we conjecture that it utilizes a mostly automated methodology, focusing on pre-determined web related vulnerabilities, leaving misconfiguration and individual instances of issues out of scope. Web related issues are scored based on version matching, prompting high CVSS scores for issues without reported manual verification of severity, while underestimating security issues outside this narrow perspective. Hence, the used methodology is unsuited for investigating the security state of non-standardized infrastructures with many sector and in-house applications, as found in research networks like those of the MPG.

Comparing our results, we find that the [redacted] report failed to identify the bulk of our critical and high-severity findings in its 'action items', including the [redacted] discussed in Section 2. We also find that the report includes port-scans that would–if analyzed–have allowed the identification of the critical/high severity issues we identified. Furthermore, despite lacking ground-truth on the utilized scope, we find indications in the provided data that the [redacted] report included foreign entities in its vulnerability assessment.

Finally, in conjunction with our observations on automated evaluation and the lack of analysis, we note that the report–due to the way the information was analyzed and presented–carries the risk of causing harm to the security posture of an organization, mainly due to sozio-organizational effects in the context of operator visibility, inner-organizational trust, 'just culture', and notification fatigue.

Therefore, in summary, we conclude that, while likely well intentioned, the report holds risks for causing harm (out-of-scope scans, organizational effects), takes a narrow and automated perspective that is methodologically unsuited for investigating the security posture of research infrastructure (web-focus, automated analysis). Furthermore, the report requires extensive follow-up work to assess the severity of findings, which have to be manually assessed and validated, without providing guidance on that process.

Key Action-Items

Remediation of issues: The findings and mitigation recommendations in the appendix should be distributed in the organization and addressed following the associated severity ratings. Appendix I provides information on how to execute this process.

Evaluation of [redacted] findings: Despite the [redacted] report having technical limitations, findings from that report should be manually verified.

Implementation of recommendations: We recommend to consider implementing our organizational suggestions, especially concerning preserving decentralization and not utilizing a standard enterprise approach, while creating appropriate support structures for operators to handle challenges of a research environment and establishing an operational excellence framework within the organization.

Institute specific support and analysis: As outlined in Section 3, security posture varies between institutes, and some institutes collected significant technical debt. Support should be provided in conformance to the organizational recommendations made in this report.

Reference List

[1] J. Arkko, M. Cotton, and L. Vegoda. *IPv4 Address Blocks Reserved for Documentation*. RFC 5737. IETF, Jan. 2010. URL: `http://tools.ietf.org/rfc/rfc5737.txt`.

[2] Ingolf Becker, Simon Parkin, and M Angela Sasse. "Finding security champions in blends of organisational culture". In: *Proc. USEC* 11 (2017).

[3] Leyla Bilge et al. "Disclosure: detecting botnet command and control servers through large-scale netflow analysis". In: *Proceedings of the 28th Annual Computer Security Applications Conference*. 2012, pp. 129–138.

[4] Rossella Bozzon et al. "Work–life interferences in the early stages of academic careers: The case of precarious researchers in Italy". In: *European Educational Research Journal* 16.2-3 (2017), pp. 332–351.

[5] M. Cotton and L. Vegoda. *Special Use IPv4 Addresses*. RFC 5735. IETF, Jan. 2010. URL: `http://tools.ietf.org/rfc/rfc5735.txt`.

[6] Sidney WA Dekker and Hugh Breakey. "'Just culture:'Improving safety by achieving substantive, procedural and restorative justice". In: *Safety science* 85 (2016), pp. 187–193.

[7] Constanze Dietrich et al. "Investigating system operators' perspective on security misconfigurations". In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. 2018, pp. 1272–1289.

[8] Zakir Durumeric, Eric Wustrow, and J Alex Halderman. "ZMap: Fast Internet-wide Scanning and Its Security Applications." In: *USENIX Security Symposium*. Vol. 8. 2013, pp. 47–53.

[9] Tobias Fiebig and Doris Aschenbrenner. "13 propositions on an internet for a" burning world"". In: *Proceedings of the ACM SIGCOMM Joint Workshops on Technologies, Applications, and Uses of a Responsible Internet and Building Greener Internet*. 2022, pp. 1–5.

[10] Tobias Fiebig et al. "Heads in the Clouds? Measuring Universities' Migration to Public Clouds: Implications for Privacy & Academic Freedom". In: *Proceedings on Privacy Enhancing Technologies Symposium*. Vol. 2023. 2. 2023.

[11] Trevor Gabriel and Steven Furnell. "Selecting security champions". In: *Computer Fraud & Security* 2011.8 (2011), pp. 8–12.

[12] Seth Hanford. "Common vulnerability scoring system, v3 development update". In: *Technical report, Forum of Incident Response and Security Teams (FIRST)*. 2013.

[13] Garrett Hardin. "The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality." In: *science* 162.3859 (1968), pp. 1243–1248.

[14] Florian Holzbauer et al. "Not that Simple: Email Delivery in the 21st Century". In: *2022 USENIX Annual Technical Conference (USENIX ATC 22)*. 2022, pp. 295–308.

[15] Michael Howard and Steve Lipner. *The security development lifecycle*. Vol. 8. Microsoft Press Redmond, 2006.

[16] Natalia I. Kadenko et al. *CyberSec4Europe D2.1 Governance Structure v1.0*. 2020. URL: `https://cybersec4europe.eu/wp-content/uploads/2020/02/D2.1-Governance-Structure-final-Submitted.pdf`.

[17] Mannat Kaur et al. "" I needed to solve their overwhelmness": How system administration work was affected by COVID-19". In: *Proceedings of the ACM on Human-Computer Interaction* 6.CSCW2 (2022), pp. 1–30.

[18] Mannat Kaur et al. "" Oh yes! over-preparing for meetings is my jam:)": The Gendered Experiences of System Administrators". In: *Proceedings of the ACM on Human-Computer Interaction* 7.CSCW1 (2023), pp. 1–38.

[19] Mannat Kaur et al. “Human factors in security research: Lessons learned from 2008-2018”. In: *arXiv preprint arXiv:2103.13287* (2021).

[20] J. Levine and M. Delany. *A ”Null MX” No Service Resource Record for Domains That Accept No Mail*. RFC 7505. IETF, June 2015. URL: `http://tools.ietf.org/rfc/rfc7505.txt`.

[21] Thomas A Limoncelli, Christina J Hogan, and Strata R Chalup. *The Practice of System and Network Administration: Volume 1: DevOps and other Best Practices for Enterprise IT*. Vol. 1. Addison-Wesley Professional, 2016.

[22] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. “Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice”. In: *Proceedings of the ACM on human-computer interaction* 3.CSCW (2019), pp. 1–23.

[23] Nuthan Munaiah and Andrew Meneely. “Vulnerability severity scoring and bounties: Why the disconnect?” In: *Proceedings of the 2nd International Workshop on Software Analytics*. 2016, pp. 8–14.

[24] Y. Rekhter et al. *Address Allocation for Private Internets*. RFC 1918. IETF, Feb. 1996. URL: `http://tools.ietf.org/rfc/rfc1918.txt`.

[25] Philipp Richter et al. “A primer on IPv4 scarcity”. In: *ACM SIGCOMM Computer Communication Review* 45.2 (2015), pp. 21–31.

[26] Nicholas J. Russel et al. “Max Planck PostdocNet Survey Report 2022”. In: *Max-Planck Society Post-Doc Network* (2022). URL: `https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_3507886`.

[27] Minna Salminen-Karlsson, Andrea Wolffram, and Nina Almgren. “Excellence, masculinity and work-life balance in academia: Voices from researchers in Germany and Sweden”. In: *International Journal of Gender, Science and Technology* 10.1 (2018), pp. 52–71.

[28] Florian Streibelt et al. “Back-to-the-Future Whois: An IP Address Attribution Service for Working with Historic Datasets”. In: *International Conference on Passive and Active Network Measurement*. Springer. 2023, pp. 209–226.

[29] Anthony Vance et al. “The fog of warnings: how non-essential notifications blur with security warnings”. In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 2019, pp. 407–420.

[30] Mathew Vermeer et al. “SoK: a framework for asset discovery: systematizing advances in network measurements for protecting organizations”. In: *2021 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE. 2021, pp. 440–456.

I Findings per Institute

I.1 Description

Here, you can find a summary of findings for each institute. This summary includes a list of findings and recommendations specific to the institute beyond the general recommendations in Section 5. Findings are classified in four categories highlighting the urgency for addressing them:

Critical: *Immediate Mitigation:* This category includes events that are at least incidents. Furthermore, these incidents must either allow unauthorized access to a system, allowing further access to privileged information. Similarly, cases where PII of third parties is revealed, or physical harm is enabled fall in this category.

High: *Critical Findings:* Critical issues must be mitigated as soon as possible. They usually pose a high likelyhood of leading to an incident but do not signify an immediately ongoing incident; This also includes cases of non-specific threats that are easily mitigated.

Medium: *Urgent Findings:* These issue are those that pose no immediate threat but should be addressed in the near future, and might benefit lateral movement, or indicate limitations in operational procedures.

Low: *Chores:* Findings of low criticality that should be addressed/picked up in regular maintenance. This includes issues like, for example, updating out-dated libraries outside of security support, for which no vulnerabilities are known yet.

Please note that the urgency we attribute to an observation might not correlate to how critical it actually is upon closer inspection. However, we set these urgencies to provide guidance on the order in which items should be investigated. This also means that items we consider critical may not require mitigation upon closer investigation, or issues rated lower may require more immediate attention upon closer investigation. However, without detailed knowledge of the corresponding infrastructure, this can not be easily determined. Hence, we leave it in the hands of local staff.

To guide this process we provide recommendations and further steps for each finding we document. Please note, again, that these recommendations may not always be applicable and it depends on the local situation on how applicable recommendations are. Operators can reach out to the report's author in case further input is needed.

Please also note that findings are provided for each host defined by an IP address exposed to the Internet, as it is not always possible to determine whether two addresses ultimately belong to the same host. Correspondingly, finding and recommendations may be provided multiple times for what is ultimately the same host.

I.2 Access to Machine-Readable Data

All institutes can request a copy of the data that was collected about them in machine readable form the author. Please see the 'Contact' section on the first page of this document. When contacting the author, please use your official MPG email address and include a list of prefixes/addresses you are responsible for. Data will be provided as an archive of JSON files.

I.3 How to Use the Appendix

Central Management/Leadership For central management in charge of distributing the findings of this report, we recommend the following approach:

- Provide each institute with a copy of the redacted version of this report, along with the Appendix specific to that institute.
- Request feedback on a timeline for first feedback on the findings in the report, ideally within one week.
- The timeline should include milestones for:
  - Assessing all 'Critical' findings and providing feedback on the individual findings. The focus of that report should be on the needed support to mitigate those issues, or–if presented issues were determined to not be an issue–a documentation of that.
  - Assessing all 'High' findings and providing feedback on the individual findings. The focus of that report should be on the support that is needed to mitigate those issues, or–if presented issues were determined to not be an issue–a documentation of that.
  - Feedback on the plan to assess 'Medium' and 'Low' criticality findings, and whether it is feasible within the workload of day-to-day operations.

Individual Institutes For individual institutes, upon receipt of the report, we recommend the following approach:

- Elect an internal incident coordinator who is in charge of coordinating communication. This does not have to be a technical person. It can be useful to select somebody in a good position to nudge people to follow-up on issues; Team-assistants often have a talent to do that in a non-authoritative way.
- Have the coordinator print the institute's appendix[136].
- In the team, jointly go over the 'Critical' and 'High' findings; Share your general impression, for which the coordinator takes notes.
- Dispatch each 'Critical' and 'High' finding to a person in charge of verifying it (jointly decide who).
- Discuss and agree on how long it will take to assess 'Medium' and 'low' issues, if each team member–for example–will use 15 minutes each day to look at at least three 'Medium' or 'Low' findings.
  - The handle notes down the person on their authoritative print out.
  - the person receives an extra copy of the pages for the finding(s) they are responsible for.
- Starting with 'Critical' findings first, followed by 'High' findings, investigate each issue. Note down on the report-copy:
  - Re-assess the criticality of the issue given internal knowledge of the infrastructure.
  - Read the mitigation recommendations and determine whether a feasible option is included (Document on print out, e.g., tick-marking steps in lists.).
  - Document whether the finding will be risk accepted, or which mitigation will be implemented.
  - Document the estimated time-to-completion and resource needs for the mitigation.
  - As soon as the process has been completed, hand the printout of each completed task to the coordinator.

[136] This *could* be done in a ticket system; However, entering the data may ultimately cost more time than it saves to have all issues in a ticket system. Ultimately, paper vs. tickets is a personal choice of teams. We recommend paper.

- The coordinator transfers key items from individual findings to their authoritative copy;
  - Issue criticality assessment (Matches rating from report/Lower/Higher).
  - Issue mitigation plan.
  - Issue mitigation resource needs/timeline.
- The team meets and drafts a response on the report, including:
  - Feedback on 'Critical' and 'High' findings (based on information recorded in the authoritative print-out).
  - Mitigation timelines/plans.
  - Mitigation resource/support needs.
  - Timeline for revisiting 'Medium' and 'Low' findings.
  - The coordinator communicates the results to central.
- Mitigation for 'Critical' and 'High' findings proceeds following the timelines the team established; Progress is feed back to the local coordinator, who can sent, e.g., monthly aggregate progress reports to central if desired by them.
- The team starts processing 'Medium' and 'Low' findings, updating the coordinator once an item has been assessed.
  - For all items, focus on whether their rating (with internal information) should have been different and notify the coordinator.
- As time permits, plan regular sessions revisiting especially 'Critical' and 'High' rated issues. Furthermore, you can use these meetings to try to identify 'patterns' in 'Medium' and 'Low' findings.
  - Ensure that–whatever is discussed in these meetings–no repercussions are launched against people who might have had a role in decisions or actions that contributed to an issue. This is a matter of trust and culture.
  - It can be useful to adopt a 'Chatham House' rule for these meetings, i.e., you are allowed to discuss *what* was said outside of the meeting, but not *who* said it (this might be somewhat difficult, especially in smaller teams).
  - It can improve the situation, if–before discussing issues–a round is made in which everyone notes a (major) technical mishap they had, no matter if it was caught in time or lead to an incident.
  - Assess what (ultimately) caused the issue.
  - Assess whether external (out of your control) factors contributed to the issue as well, or contribute to actions and behavior that lead to the issue.
  - Assess whether the issue fits a pattern with other issues.
  - Assess whether you can make changes to your infrastructure that make it less likely that the issue occurs again in the future. This might be (more) monitoring, redesigning your network/separation of services, use of a different vendor/setup, etc.
  - If an improvement is not feasible (limited resources, i.e., time/staff or funding, not compatible with the user-base/users' needs, etc.), first discuss it independent of those constraints.
  - Document results and action items; Especially if external constraints contribute to issues, communicate them upwards (see 'Managing your Manager' in [21]).
- In general, please see Section 5.2.3, and consider whether that approach (see the full description by Limoncelli et al. [21]) might work out for the team.

II Institute Summaries

The redacted version of this report does not include the summaries for individual institutes.