Combined micro-architectural and physical side-channel security

A Data Management Plan created using DMPonline.be

Creator: Jesse De Meulemeester

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: 11PFE24N

ID: 204288

Start date: 01-11-2023

End date: 31-10-2027

Project abstract:

To streamline the development of modern processors, designers typically make abstractions of the lower building blocks on which they build their layers. When developing micro-architectural components, for instance, the physical effects are abstracted away. This abstraction is also reflected in the security evaluation of these processors; micro-architectural security and physical side-channel security are typically considered separately. However, the combined security should also be considered due to the close interaction of these micro-architectural components and the underlying physical implementation. This holistic approach may allow for new insights and may unveil new vulnerabilities.

This project aims to bridge the gap between research into micro-architectural security and physical side-channel security. It aims to investigate this combined security and develop efficient countermeasures to protect next-generation processors. To facilitate these goals, we will consider the combined security from both a physical and a micro-architectural perspective.

Last modified: 18-04-2024

Combined micro-architectural and physical side-channel security
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

1. Side-channel measurements: Files containing the raw side-channel (e.g., power, EM, ...) measurements. New data. Up to 10 TB over the whole project.
2. Code files: e.g., proof-of-concept implementations, tools, and post-processing scripts. New data. Up to 10 GB over the whole project.
3. Diagrams, figures, and plots (up to 10 GB over the whole project).
4. Papers and presentations (up to 1 GB over the whole project).

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

All research data will be stored on storage which is provided, managed, and maintained by KU Leuven. Following the KU Leuven guidelines, this data will be stored for at least 10 years. The side-channel traces will be stored on a high-capacity file server specifically designed for this type of data. All other files will be stored in a GitLab repository, which is supervised by a DMP officer and gets a daily backup. KU Leuven additionally provides a Research Data Repository (RDR) to share large amounts of data publically, up to 50GB per year. This will be used to share data, including side-channel measurements, complementing our publications.

What’s the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

NA

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

NA

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

NA

Combined micro-architectural and physical side-channel security
FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Measurement results	Files containing the raw side-channel measurements (power, EM).	Generate new data.	Digital	Experimental	.hdf5, .csv, .txt, ...	<10TB	
Code files	Proof-of-concept implementations, tools, and post-processing scripts.	Generate new data.	Digital	Software	Code files (.c, .h, .v, .py, ...)	<10GB	
Diagrams, figures, and plots	Diagrams, figures, and plots	Generate new data.	Digital	Compiled/aggregated data	.tex	<10GB	
Papers and presentations	Papers and presentations	Generate new data.	Digital	Compiled/aggregated data	.tex	<10GB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

NA

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

All code projects will be accompanied by a readme file describing the software and hardware prerequisites required to (re-)use the code. Additionally, all code will be extensively documented (both at the source code level and at the module level) such that any single part of the code can also be easily re-used.
All datasets containing traces will contain a detailed overview of the experimental setup, including all preparatory steps and the exact hardware and software used. This information will also be provided in a readme file.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- Yes

All experimental data (traces, ...) will be made available in KU Leuven RDR, thus we will use the DataCite standard.

3. Data storage & back-up during the research project

Where will the data be stored?

Data used to produce project results will be stored in a GitHub/GitLab repository. In a later phase, and after the publication of the research paper concerned, the relevant data will be made available in the RDR repository. KU Leuven RDR is hosted on KU Leuven servers maintained by the KU Leuven IT team and located in the KU Leuven data centers. In parallel, we will store the same data in our departemental home directories, hosted by a server at our research department.
Larger data, such as side-channel traces, will be stored on COSIC's server dedicated to storing side-channel traces.

How will the data be backed up?

We replicate the data sets among at least two department computers. The GitLab repositories are hosted by a server at our research department and get a daily backup. GitHub is a well-established, commercial repository service provider, which we assume has industry-standard data availability measures in place (such that repositories will not be lost unless deleted by authorized GitHub users). RDR backups are handled by KU Leuven.
The larger data stored on COSIC's file server is not backed up. These are easy-to-reproduce data of a size (10GB—1TB) that is impractical to backup. In the case of this big data, we will keep 1 million traces of data as a sample together with a detailed description of the measurement. The description will be stored on GitLab or RDR.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

- Yes

There is sufficient storage & backup capacity during the project. We have many Terabytes of storage space available on the department machines we use to replicate our data sets. The department's Gitlab instance and GitHub are offered services and are thus ensured to have enough spare capacity.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Access to our department machines is managed by access to the university network and standard-practice user permissions on the machine (only users with the right set of permissions can access the data, our department controls the management of these users).
The GitLab repository has access control and is maintained by a DMP officer who keeps the structure of the repository in place and manages the access control. Unauthorized GitHub modifications are prevented through their industry-standard user management.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

The costs of the gitlab maintenance are covered by our research group. These costs are integrated into the general IT costs. The use of GitHub is free of charge. The costs of RDR are covered by KU Leuven.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

Datasets are stored on GitHub/GitLab and RDR repositories and will be retained for at least 10 years after the end of the project.

Where will these data be archived (stored and curated for the long-term)?

To the extent feasible in terms of the volume of collected data, the data will be stored in the Gitlab, GitHub, and RDR repositories. If we consider it necessary to purchase archival storage capacity for the entirety of the research project's artifacts (e.g., due to non-reproducibility of some data sets and limitations on the Gitlab and GitHub services), we will do so appropriately and in time before the project concludes.

Options include the storage services offered by KU Leuven or dedicated storage hardware to be kept long-term at our department.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

The costs of the gitlab maintenance are covered by our research group. These costs are integrated into the general IT costs. The use of GitHub is free of charge. The costs of RDR are covered by KU Leuven.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository

Relevant data used to produce project results will be made available. Relevant data can consist of software or hardware scripts and code, algorithms, protocols, manuscripts, figures or others. Publications will be made available in pdf.

If access is restricted, please specify who will be able to access the data and under what conditions.

NA

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

To the extent feasible due to the volume of the collected data sets, data used for published articles will be stored in Gitlab repositories with access control and publicly accessible GitHub repositories. Publications will be made available in open-access repositories. All source code, protocols, and algorithms used for published articles will be made available as open-source software on GitHub.

When will the data be made available?

Upon publication of research results.

Which data usage licenses are you going to provide? If none, please explain why.

Data will be licensed under CC-BY 4.0 (for data) and MIT or GPLv3 (for code).

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

What are the expected costs for data sharing? How will these costs be covered?

No costs are to be expected.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

The PhD student, Jesse De Meulemeester, funded by this grant, is responsible for data documentation and metadata.

Who will manage data storage and backup during the research project?

In the first place, the PhD student, Jesse De Meulemeester, funded by this grant, is responsible for data storage and backup. The DMP-officer maintaining the Gitlab repositories will share in the responsibility for data storage and backup.

Who will manage data preservation and sharing?

The PhD student, Jesse De Meulemeester, funded by this grant, is responsible for data preservation and reuse.

Who will update and implement this DMP?

The PhD student, Jesse De Meulemeester, funded by this grant, is responsible for updating and implementing this DMP.