

Benchmarking
workshop

WELCOME

Edge Technologies

**i.e. HW/SW platforms based on efficient Neuromorphic,
artificial and spiking neural networks, solutions**



Benchmarking workshop

**If you're not benchmarking your performance against
your competitors, you're just playing with yourself.**

Al Paison





WHY ARE WE HERE ?

- Benchmarking for resource constrained systems is eclectic ?
- What are the methods used today ?
- Why do they fail ?
- How can they be standardised ?
- What are different perspectives of the users, industrial, academic, fabs ?
- What should we be working on as a group to enable the benchmarking ?

HOW CAN WE LEVERAGE THE ANDANTE COMMUNITY TO STANDARDIZE
BENCHMARKING IN EDGE COMPUTING

AGENDA



08:30 Welcome & start

- Introduction *Andrea Dunbar*
- Standard benchmarking for machine learning
Siavash Bigdeli
- Useful benchmarking for commercial neuromorphic hardware *Dylan Muir*

09:05 Track 1: Tools & methodologies *Stefano Traferro*

- Simulation framework for energy and latency in multi-core neuromorphic architectures *Stefano Traferro*

09:25 Track 2: Benchmarking categories *Dylan Muir*

- Consideration of the real world in use case-based benchmarking *Kay Bierzynski*
- Challenges of using medical datasets *Diana Cojocaru*

10:05 Track 3: Metrics *Rodrigo Martín Fernández*

- Impact of HW optimizations on KPIs *Petar Jokic*

10:25 Break

10:35 Track 3: Metrics (continuation)

- Early power extraction with Cadence tools *Erfan Azarkhish*
- KPI-aware optimization and design *Maen Mallah
Ferdinand Pscheidl*

11:15 Outlook

- Future of benchmarking *Simon Narduzzi*

11:30 Discussion split into the three tracks

What are the main challenges?

What are the current tools, what work is being done?

What direction should ANDANTE project take with respect to these challenges?

12:10 Interactive Questionnaire & Report & feedback

12:30 End of the meeting

Benchmarking
workshop

BREAK



Benchmarking
workshop

DISCUSSION

You will be assigned to your track rooms now



Questions

Go on the Mentimeter Link in Chat or here:

Link: www.mentimeter.com

Code: 28 19 15 9



Benchmarking
workshop

Track 1: Tools & methodologies

Moderator: Stefano Traferro



TRACK-1: TOOLS & METHODOLOGIES - TOPICS WHY ARE WE HERE ?



[Q] Which is the optimal level of abstraction to run benchmarking simulations?

- What does optimal mean? Design faster? Algorithm/model faster?
 - Use case dependent.
 - Relative performance looks more important during the development; maybe for commercial use the absolute figures are more important.
 - Different level of precision are depending on the goal of the estimation
 - Can we think to have a *hierarchical* precision?
 - Can we characterize the lower levels and push it up?
- How can I expand “my” tool(s) to better recognize the metrics/KPIs?
 - Possibility of collaboration in Andante?

TRACK-1: TOOLS & METHODOLOGIES - TOPICS WHY ARE WE HERE ?



[Q] Which kind of simulations are required? What does the EDA provide? Why do design teams develop their own tools?

- For (analog) circuit generation you need structure -> Important to have circuit generators
- Can we join the effort in the Andante to make a “broader” platform and avoid re-inventing the wheel on common problems?
 - Avoid overlaps and commonalities
- Asynchronous design?
 - Synthesis -> timing driven placement -> power/energy estimation
 - Currently, huge lack of tools
 - How to export to higher level of abstractions?
 - Are people/companies interested in join the effort to tackle this problem?

TRACK-1: TOOLS & METHODOLOGIES - TOPICS WHY ARE WE HERE ?



[Q] The mapping problem: does it make sense to develop a generic tool (sort of compiler) that can have a wide use, instead of having proprietary ad hoc solutions? Is there already anything available in that direction?

- Common (open) frameworks are available, but do they cover our (Andante) needs?
- General considerations:
 - Can we share tools across companies to improve synergy, efficiency and get better, unified overall quality and usability?
 - Can we specify data exchange formats to improve collaboration across companies?
 - Each company could work on different part(s) and/or level of abstraction.

Benchmarking
workshop

Track 2: Benchmarking categories

Moderator: Dylan Muir





TRACK-2: COMMERCIAL BENCHMARKING

Which are the most important / convincing KPIs for NM benchmarks?

- Depends a lot on use case
- e.g. accuracy vs latency
- * Depends a lot on environment
 - * e.g. Germany versus Middle East — different sensing environment
- * To work together as a community, we need some common set of KPIs to target
 - * E.g. for HW design - how to optimise HW if everything is use-case dependent?
- * Optimise general HW versus optimise for a use-case



TRACK-2: COMMERCIAL BENCHMARKING

How do we get a benchmark comparable with a TPU?

- We don't want to be benchmarking irrelevant data sets
- We can choose datasets and benchmarks, create them if necessary
 - These need to be public ideally
- We need to run benchmarks on GPUs/TPUs ourselves



TRACK-2: COMMERCIAL BENCHMARKING

Need to focus on SW tooling

- Given a dataset, what is the optimal HW?
- Need to focus on usability by ML developers
 - Esp. for SNNs, need to translate to ML developers

TRACK-2: COMMERCIAL BENCHMARKING



What concrete actions should ANDANTE take? 2–3 things done by the community?

- Emphasize the differences between GPU/TPU, CPU accelerators and NM accelerators
 - 10x Power efficiency, Focus on power for KPIs
 - Should always compare with SotA
 - New use cases / show cases for NM HW
 - HW benchmarks that are specific for SNN HW
- Provide joint publications on which KPIs to use
- “Beyond human ability” show-case applications
- Need to show that our work has value
 - Something they can achieve that cannot achieve otherwise

Benchmarking
workshop

Track 3: Metrics

Moderator: Rodrigo Martín Fernández





TRACK-3: METRICS – WORKSHOP GOALS

- Understand the SW and HW tools used today
- Identify the scenarios in which these tools are not working
- Envision the future of Benchmarking (main focus: low power neuromorphic HW, i.e., edge devices)
- Ultimate goal: Standardise benchmarking
 - Proposed features for such a standard: metrics, consideration of target use cases, etc.
 - Possible within ANDANTE?
 - Should the partners work together on such an optimization?

TRACK-3: METRICS – DISCUSSION TOPICS



SUGGESTIONS

General:

- What are the main challenges of Benchmarking?
- What are the current tools? What work is being done?
- What direction should ANDANTE project take with respect to these challenges?
- Use-case-based benchmarking
- Standardisation of Benchmarking

Metrics

- *Which metrics should be used?*
- *Which kind of simulations are required? What does the EDA (Electronic Design Automation) provide? Why do design teams develop their own tools?*
- *The mapping problem: does it make sense to develop a generic tool (sort of compiler) that can have a wide use, instead of proprietary ad hoc solutions? Is there already anything available in that direction?*
- *Which is the optimal level of abstraction to run benchmarking simulations?*
- *Given a fixed data set, what KPIs should be used, is this useful?*

TRACK-3: METRICS – OUTCOME

- **Discussion Focus:** Standardisation of benchmarking metrics
- There are already some activities like TinyML
- Use-case-based vs NN model based benchmarking
- Proposals for ANDANTE
 - Ensure use-case-based (not NN model based) benchmarking becomes standard
 - Definition and implementation of specific benchmarking framework
 - ✓ Open source /public
 - White papers:
 - ✓ Exactly explain ALL the metrics
 - ✓ 2nd Step: Define metrics for selected use cases (e.g. Keyword Spotting)

Thank you

Rodrigo Martín Fernández
Maen Mallah
Ferdinand Pscheidl
Stefano Traferro
Amirreza Yousefzadeh
Kay Bierzynski
Diana Cojocar

Fraunhofer (IIS)
Fraunhofer (IIS)
Fraunhofer (EMFT)
IMEC-NL
IMEC-NL
Infineon
Philips

Dylan Muir
Andrea Dunbar
Erfan Azarkhish
Siavash Bigdeli
Simon Narduzzi
Petar Jokic
Stéphanie Derron

SynSense
CSEM
CSEM
CSEM
CSEM
CSEM
CSEM

