

HORIZON 2020
TOPIC FETHPC-02-2017
Transition to Exascale Computing



Exascale Programming Models for Heterogeneous Systems
801039

D6.2
**Plan for Dissemination, Training, Exploitation and
Standardization**

WP6: Dissemination and Standardization



Date of preparation (latest version): 31.12.2018
Copyright© 2018 – 2021 The EPiGRAM-HS Consortium

The opinions of the authors expressed in this document do not necessarily reflect the official opinion of the EPiGRAM-HS partners nor of the European Commission.

DOCUMENT INFORMATION

| | |
|--------------------------------|--|
| Deliverable Number | D6.2 |
| Deliverable Name | Plan for Dissemination, Training, Exploitation and Standardization |
| Due Date | 31.12.2018 (PM4) |
| Deliverable lead | KTH |
| Authors | Apostolos Vasileiadis Stefano Markidis |
| Responsible Author | Apostolos Vasileiadis (KTH) e-mail: apovas@kth.se |
| Keywords | Dissemination, exploitation, standardization, training |
| WP/Task | WP6 /Tasks 6.1 and 6.2 |
| Nature | R |
| Dissemination Level | PU |
| Planned Date | 31.12.2018 |
| Final Version Date | 31.12.2018 |
| Reviewed by | Olivier Marsden (ECMWF), Adrian Tate (CRAY), Tim Dykes (CRAY) |
| MGT Board Approval Date | 31.12.2018 |

DOCUMENT HISTORY

| Partner | Date | Comment | Version |
|----------------|-------------|---|----------------|
| KTH | 08.12.2018 | Draft version of the deliverable | 0.1 |
| KTH | 10.12.2018 | Added dissemination | 0.2 |
| KTH | 12.12.2018 | Added exploitation, training and standardization | 0.3 |
| KTH | 12.12.2018 | Reviewed by Stefano Markidis | 0.4 |
| KTH | 17.12.2018 | Event attendance & standardization efforts supplement | 0.5 |
| CRAY | 21.12.2018 | Review by Adrian Tate and Tim Dykes | 0.6 |
| ECMWF | 31.12.2018 | Review by Olivier Marsden | 1.0 |

Executive Summary

This deliverable describes the plan for dissemination, training, exploitation and standardization. These four aspects of research are crucial to establish EPiGRAM-HS as a key player in the latest advances in HPC.

EPiGRAM-HS's dissemination strategy will make sure that all research carried out in EPiGRAM-HS is heard and understood by key groups in the HPC community. The project will use different dissemination channels to make sure that researchers throughout the platform spectrum will have access to all EPiGRAM-HS material, particularly:

- A webpage and a newsletter for general dissemination and repository.
- Attendance of meetings and workshops relevant for high performance exascale computing.
- Publication of scientific papers.
- Organization of workshops as well as joint activities with related projects (focus groups).

In addition, we present an initial exploitation that is an important aspect of any research that is greatly valued within EPiGRAM-HS. All partner organizations have a plan on how to use the results of EPiGRAM-HS using all available resources. Finally, we present participation in standardization bodies, such as MPI forum and GASPI forum.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 6 |
| 2 | Dissemination Plan | 6 |
| 2.1 | Web-based dissemination | 6 |
| 2.1.1 | Website | 6 |
| 2.1.2 | Open Source Code Access | 8 |
| 2.1.3 | Twitter | 8 |
| 2.1.4 | Newsletter | 8 |
| 2.1.5 | YouTube Channel | 9 |
| 2.2 | Other Dissemination Activities | 9 |
| 2.2.1 | Printed material | 9 |
| 2.2.2 | Publication of papers | 9 |
| 2.2.3 | Attendance in Conferences and Workshops | 10 |
| 2.2.4 | Organization of conferences and workshops | 12 |
| 3 | Exploitation, Training and Standardization Plan | 12 |
| 3.1 | Exploitation | 12 |
| 3.2 | Training | 14 |
| 3.3 | Standardization | 15 |
| 4 | Conclusion | 16 |

1 Introduction

The goal of this deliverable is to develop a plan for dissemination activities to promote the research of EPiGRAM-HS, the involvement of the project in standardization and training efforts, and EPiGRAM-HS interaction with other European HPC initiatives. Dissemination and standardization involve all partner organizations within EPiGRAM-HS. It is a joint endeavor to utilize the results acquired by the research.

Different dissemination means have been identified for different target audiences. The main target audiences are the scientific community in the area of exascale computing as well as code owners that might take up EPiGRAM-HS developments.

To ensure adoption and exploitation of the results within the HPC community, the EPiGRAM-HS project members are committed to open standards, open interfaces, and open source software. The goal is to create a world-wide infrastructure with third-party adoption.

The deliverable is organized as follows. Section 2 lays out the dissemination plan for EPiGRAM-HS including website and other online channels as well as publications and organization of events. Section 3 describes the plan for exploitation by all partners, training and standardization efforts. Finally, Section 4 summarizes and concludes the deliverable.

2 Dissemination Plan

Dissemination is the effort to communicate the results of EPiGRAM-HS to key target audiences. There are two main groups that EPiGRAM-HS is interested in:

- HPC community at large
- Application experts in multiple disciplines

To reach those groups EPiGRAM-HS is using a combination of methods and channels to maximize efficiency and impact.

2.1 Web-based dissemination

2.1.1 Website

EPiGRAM-HS is using several dissemination tools to exploit multiple reach capabilities. The main dissemination tool to reach the broader public is the website. The goal of the website is to provide a simple introduction of EPiGRAM-HS for the general public as well as resources such as activities, focusing especially on: conferences and meetings, publications, and news on recent developments relative to EPiGRAM-HS.

The website is set to follow the latest online standards such as clear and modern homepage, mobile functionality, newsletter subscription, information about the partner organizations, researchers and activities section. Sections of the website are using **Google Analytics** as well as **Hotjar** in full compliance with GDPR. Google Analytics helps us gather data such

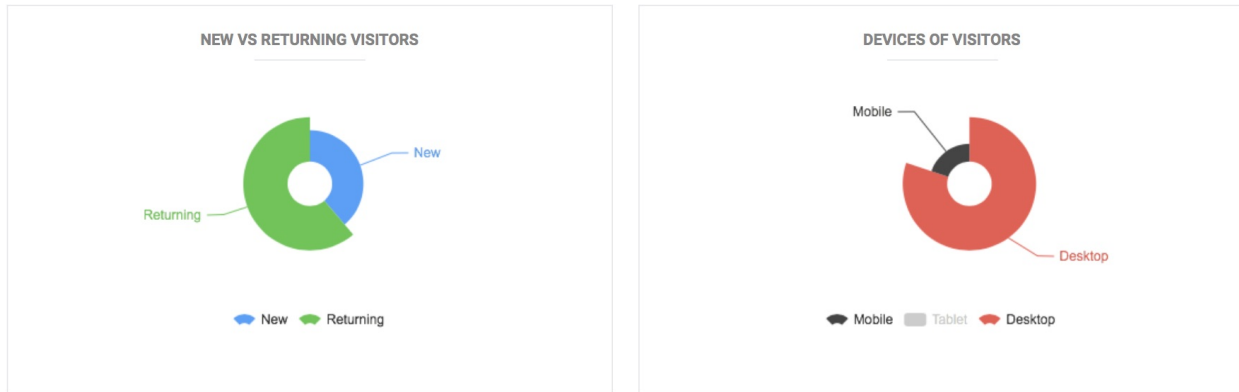


Figure 1: EPIGRAM-HS website analytics. New vs Returning visitors and Mobile vs Desktop visitors (Wordpress plugin Analytify)

as page views, location of visitors, gender and time spent on website, while Hotjar creates visual maps with the most clicked links and buttons on the website. This helps EPIGRAM-HS identify its audience while adjusting the content to adapt to the needs of the visitors.

The website is online since September 15th 2018 and it has been kept updated with recent activities of EPIGRAM-HS ever since. By the end of November 2018 the website has been visited 191 times by 74 visitors with 61% of them being returning visitors (Figure 1).

Current achievements:

1. Added Search Engine Optimization (SEO).
2. Publication of Steven W. D. Chien on Characterizing Deep-Learning I/O Workloads in TensorFlow (<https://epigram-hs.eu/publications/>).
3. Video of Jeffrey Vetter's talk on Exploring Emerging Memory Technologies in Extreme Scale High Performance Computing (https://epigram-hs.eu/vetter_lecture/).
4. News post about EPIGRAM-HS presence in SC18 (<https://epigram-hs.eu/sc18-epigram-hs-presence/>).
5. News post on EPIGRAM-HS researcher Torsten Hoefler being featured in Inside HPC website (<https://epigram-hs.eu/scientific-benchmarking-of-parallel-computing-systems/>).

The EPIGRAM-HS website is continuously updated as new results and activities become available.

EPIGRAM-HS is also active using other online tools that complement the website and more generally the entire dissemination strategy. These include open access online repositories, Twitter, Youtube, and a quarterly newsletter.



Figure 2: EPiGRAM-HS Twitter profile.

2.1.2 Open Source Code Access

Most of the code, developed during the project, will be open source and provided through partner website, public, Git version control repository service (e.g., GitHub or Bitbucket) for exposing research by the EPiGRAM-HS project to the application developer community. This provision will also boost the collaborations with external entities and consolidate the scientific trustworthiness of the results coming from the project.

2.1.3 Twitter

Social media, and especially Twitter, are important so that followers of EPiGRAM-HS can get the recent activity of the consortium, or generally news about HPC and Exascale. The EPiGRAM-HS Twitter account (Figure 2) had 47 followers in the beginning of December 2018. The goal of the Twitter account is to move more traffic on the website and the newsletter.

2.1.4 Newsletter

The newsletter keeps the followers up to date with recent advances on the research and activities of EPiGRAM-HS. A newsletter is an effective way to create networks within the HPC community through conferences and online presence. The newsletter will only be sent to visitors who subscribed willingly and not as a marketing tool. The subscription form is posted on the website and mentioned in conferences and all printed material. Collection and processing of user subscription data is covered by consent; unsubscription from the distribution list is easy and in full compliance with GDPR regulations. The newsletter is administered through **Mailchimp**.

2.1.5 YouTube Channel

EPiGRAM-HS understands that multimedia and quality content creation plays a huge role in all outreach activities. Therefore EPiGRAM-HS is using a **YouTube channel** with focus on lectures, training videos and other events. The channel has also old content from the EPiGRAM project. There will be at least 4 videos published by September 2019.

Current achievements:

1. **Video lecture**, Jeffrey Vetter on Exploring Emerging Memory Technologies in Extreme Scale High Performance Computing. (https://epigram-hs.eu/vetter_lecture/).

2.2 Other Dissemination Activities

The online tools are crucial to create an awareness of EPiGRAM-HS but it is of course not enough if it is not backed up by physical dissemination methods. These are namely printed material, publication of papers in key journals, attendance in conferences, workshops and focus groups, collaboration with other European HPC projects and organization of conferences and workshops.

2.2.1 Printed material

EPiGRAM-HS has printed flyers and a poster (Figure 3) as dissemination material in HPC events to reach general audiences within the HPC community. All printed material will be updated once new results and papers have been published.

2.2.2 Publication of papers

Scientific results of EPiGRAM-HS will be disseminated through scientific papers to conferences and journals. This is the main route to achieving excellence in the HPC field. It is the basis of the research that is being done within the project.

The goal is to publish the results of EPiGRAM-HS in time for selected conferences and workshops. As more results are being achieved, we aim that there will be publications in related journals such as **Transactions on Parallel and Distributed Systems**, **International Journal of High-Performance Computing Applications**, **Journal of Parallel and Distributed Computing**, and **Concurrency and Computation: Practice and Experience**. Publications target the HPC community and a number of suitable conferences have been identified as focus venues (Subsection 2.2.3).

All the publications will be available as green open-access, allowing for a major dissemination of the scientific outcomes from the project.

Current achievements:

1. Characterizing Deep-Learning I/O Workloads in TensorFlow

Steven W. D. Chien, Stefano Markidis, Chaitanya Prasad Sishtla, Luis Santos, Pawel Herman, Sai Narasimhamurthy, Erwin Laure
 PDSW at SC18, <https://arxiv.org/abs/1810.03035>



Figure 3: EPiGRAM-HS poster for conferences and HPC events.

2.2.3 Attendance in Conferences and Workshops

EPiGRAM-HS gives researchers support to attend conferences and workshops. The aim is mainly to present publications on EPiGRAM-HS research and results, although in the case of no available publications, other dissemination material such as flyers and posters will be distributed.

We are targeting high-profile HPC conferences, such as **Supercomputing**, **ISC**, **IPDPS**, **HPDC**, **Cluster**, **EuroPar**, **CCGrid**, **PASC** among others. Convenient workshops for the presentation of our work on large scale heterogeneous systems will be the **AsHES (Accelerators and Hybrid Exascale Systems)** and **RAW (Reconfigurable Architectures**

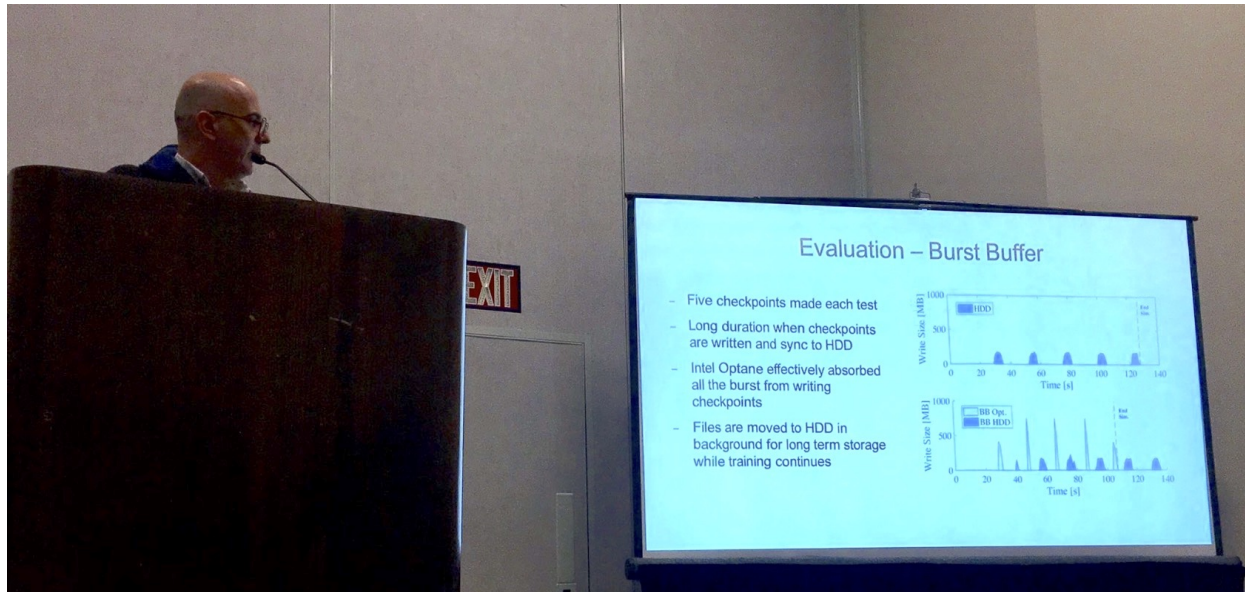


Figure 4: Stefano Markidis’s presentation at SC18, 12th November 2018. Characterizing Deep-Learning I/O Workloads in TensorFlow. Photo: Glenn K. Lockwood

Workshop), typically organized in connection with IPDPS. There is also H2RC (Heterogeneous Computing with Reconfigurable Logic), typically organized at Supercomputing.

Current achievements:

1. **Presentation at UCLA IPAM workshop. 18th Oct 2018, Los Angeles, California**
Adrian Tate, CRAY, Memory- and Data-Centric Abstractions for Management and Optimisation of Data Movement
2. **Presentation at PDSW. SC18, 12th Nov 2018, Dallas, Texas**
Stefano Markidis, KTH, Characterizing Deep-Learning I/O Workloads in TensorFlow
 (Figure 4) <https://arxiv.org/abs/1810.03035>
3. **Presentation at ESPM2. SC18, 12th Nov 2018, Dallas, Texas**
Daniel Holmes, EPCC, Heterogeneous Systems and the Road to Exascale for HPC and AI
4. **Fraunhofer ITWM booth. SC18, 13th-15th Nov 2018, Dallas, Texas**
Booth duty: Valeria Bartsch
Key points for EPiGRAM-HS: discussions on GPI and the involvement of EPiGRAM-HS in FPGAs. Print material was distributed regarding European projects and the involvement of EPiGRAM-HS in exascale research.

5. **Birds of a feather session: Interactivity in HPC. SC18, 13th Nov 2018, Dallas, Texas**

Attendees: Vyacheslav Olshevsky

Key points for EPiGRAM-HS: Organizational and technological challenges of interactive supercomputing were discussed. Special use cases were discussed, in particular a cluster where user quotas allow for high-priority jobs to be launched immediately. Potential impact of cloud computing was mentioned as an alternative approach to interactive HPC.

2.2.4 Organization of conferences and workshops

One of the most important priorities will be organizing international conferences and workshops. It is one of the most important ways we can communicate the research and results of EPiGRAM-HS to the HPC community. EASC was a quite successful conference during the previous years and EPiGRAM-HS intends to repeat it by being one of the main participants. We also aim at organizing workshops to give international visibility to the project. We will target a focus meeting in connection with EuroMPI19.

3 Exploitation, Training and Standardization Plan

To ensure adoption and exploitation of the results within the HPC community, the EPiGRAM-HS project members are committed to open standards, open interfaces, and open source software. In EPiGRAM-HS, we aim to create a world-wide infrastructure with significant third-party adoption.

3.1 Exploitation

There are specific plans for exploitation for each partner organization, namely:

- KTH will exploit the results of the project mainly through the use of the two pilot applications that are widely used in Swedish academia. The project will also give a major push to the national eScience initiative SeRC (Swedish eScience Research Center) that is working on advanced usage of peta to exascale computing in a wide range of disciplines. The results from the EPiGRAM-HS project will also be used in research at the CST department and teaching in the graduate classes at KTH. KTH will exploit and disseminate the results for cooperation with Swedish industrial and academic partners in the exascale projects. In addition, KTH will exploit the results of simulations carried out with **Nek5000** and **iPIC3D**, that will improve competitiveness in the application areas that are most useful for Europe. Nek5000 is a computational fluid dynamics code used to simulate the flow in nuclear reactors (fission energy) and solve fundamental science problems, such as boundary layer physics with aerodynamics application. The iPIC3D code is mainly used for space weather modeling to simulate the effects of solar activity on Earth. Space weather forecasting is a global challenge,

and one of the focuses of the European Space Agency (ESA) since solar storms can damage spacecrafts in space, can increase the radiation doses received by astronauts, and can create black-outs on the Earth.

- UEDIN will use the knowledge and experience from the EPiGRAM-HS project by using it in four areas:
 - in production use
 - in research and development
 - in teaching
 - in technology transfer

Specifically, UEDIN exascale programming models are central to the research activities in EPCC and EPiGRAM-HS will influence and enhance these. Results from EPiGRAM-HS will be fed into UEDIN post-graduate (MSc) teaching on HPC, thereby helping to educate the future European work-force. Where appropriate, the software developed in EPiGRAM-HS will be used in UEDIN day-to-day work as a supercomputing centre and in the work with academic and commercial customers.

- ETH will exploit the results of the project through teaching and research at ETH, and will thus contribute to the education of the future European work-force in parallel and high-performance computing. Specifically, interface design and implementation for exascale message-passing and RMA models are central to the research activities of the ETH group, and EPiGRAM-HS will influence and interact with these. In addition, ETH will seek to exploit the results cooperating with Swiss and European industrial and academic partners in the context of future national and international projects.
- FhG will exploit the results of the EPiGRAM-HS project within a variety of different activities. On the commercial side we will include enhancements of GPI into the production release of GPI to enable industry partners to exploit new future hardware architectures. In the same way the extended capabilities of GPI-Space will be transferred into already existing commercial GPI-Licenses. And finally, we will exploit the new capabilities in upcoming industry cooperations. On the education side we plan to include project results especially from the Deep Learning activities in our course program that we offer for industry and academia. Further exploitation happens in student activities to educate our future HPC work force.
- Cray has the opportunity, using the results from EPiGRAM-HS, to better understand application requirements and demonstrate how new capabilities and optimizations can be provided to address data placement and movement in an increasingly complex memory hierarchy. These insights are a highly valuable business asset and could provide a platform for Cray to design features and components of future systems that are more likely to meet customer requirements, ensuring a product offering that can differentiate Cray in the procurement of advanced systems. The Cray EMEA Research Lab (CERL)

will generate an increased research competency and European IP as a result of this project. Any productization plans, resultant IP or follow-on projects will be managed through CERL and will thus contribute to a deepening and growth of European IP for worldwide consumption.

- ECMWF’s long-term strategy of continued forecasting skill will be greatly influenced from the results of EPiGRAM-HS. Implementation of the strategy will require order-of-magnitude increases in computational effort, which will not be affordable without improvements to computational metrics such as scaling and efficiency. EPiGRAM-HS will allow ECMWF to explore state-of-the-art techniques that could lead to the required scaling and efficiency improvements for the IFS.

3.2 Training

One of the established objectives of EPiGRAM-HS is to offer training opportunities to the scientific community. These are the training activities planned to be conducted, in order of relevance and estimated impact:

- **Tutorials.** Alongside high-profile conferences, such as SC or EuroMPI, hands-on tutorials will be offered to the scientific and industry communities. These tutorials will provide valuable insights about the developments in EPiGRAM-HS. At the same time, it will be an opportunity to engage people towards the contributions in programming models for the exascale given by this project.
- **Courses.** Specialized courses will be organized throughout the duration of the project with the purpose of training scientists on how to take advantage of some of the advanced features developed in MPI and GASPI. These courses will also target HPC application developers and can be materialized as lectures inside other highly-specialized HPC courses, such as the Advanced Computer Architecture and Compilation for High-Performance and Embedded Systems from the HiPEAC initiative or the PDC summer school.
- **Webinars.** Online training opportunities will potentially be offered to allow a large audience to be involved into the EPiGRAM-HS developments. Although, this option depends on the project’s resources, it could give EPiGRAM-HS the opportunity to reach application developers that cannot participate in the other two learning alternatives.

In addition, other training opportunities will be offered. For instance, source code examples on MPI in public web repositories of the project will be given. In addition, how-to tutorials inside the official website and other related activities will also be prepared.

3.3 Standardization

EPiGRAM-HS is expected to make important contributions to both MPI and GASPI standardization. The standardization effort in the MPI Forum will be led by UEDIN with contributions by ETH, KTH and Cray. All four organizations are MPI Forum members. The standardization effort in GASPI will be led by FhG with contributions by KTH. Both FhG and KTH are among the founding members of the GASPI Forum.

The work carried out by the members in the standardization bodies will be presented in the project meetings to receive feedback from all the consortium members. A summary of the standardization meetings will be presented by EPiGRAM-HS members during EPiGRAM-HS meetings and will be made available in the EPiGRAM-HS webpage to promote new proposals in MPI and GASPI.

Current achievements:

1. **MPI Forum Barcelona, Spain, 19th-21st Sept 2018**
Attendees: Daniel Holmes
2. **MPI Forum Virtual, 10th Oct 2018**
Topic: Semantic Terms and Conventions
Attendees: Daniel Holmes
3. **MPI Forum Virtual, 10th Oct 2018**
Topic: MPI_T events
Attendees: Daniel Holmes
4. **MPI Forum Virtual, 7th Nov 2018**
Topic: Sessions
Attendees: Daniel Holmes
5. **MPI Forum Virtual, 7th Nov 2018**
Topic: Sessions
Attendees: Daniel Holmes
6. **MPI Forum San Jose CA, 3rd-6th Dec 2018**
Official agenda: <https://www.mpi-forum.org/meetings/2018/12/agenda>
Attendees: Daniel Holmes

4 Conclusion

In this deliverable, we presented the initial plan for, and first results of, the dissemination activities related to EPiGRAM-HS, as well as a plan for exploitation of the research within the project. Training plans and standardization are also analyzed. This plan will be revised on every deliverable while the research results are being acquired and methods are assessed, in order to provide an agile planning tool.