OOKAMI PROJECT APPLICATION

Date: **02/11/2021**

Project Title: Graph Analytics and Combinatorial Kernels (GRACE)

Usage:

☒ Testbed

☐ Production

Principal Investigator:

Name:	Sayan Ghosh
University/Company/Institute:	Pacific Northwest National Laboratory
Mailing address including country:	MSIN: J4-30, P.O. BOX 999, RICHLAND, WA 99352
Phone number:	(509)-372-6323
Email:	sayan.ghosh@pnnl.gov

Names & Email of initial project users:

PNNL – Sayan Ghosh (sayan.ghosh@pnnl.gov), Marco Minutoli (marco.minutoli@pnnl.gov)

BNL – Dossay Oryspayev (doryspaye@bnl.gov) **LBL –** Oguz Selvitopi (roselvitopi@lbl.gov)

Usage Description:

We are working on several applications in the domain of combinatorial scientific computing, spanning infection modeling, numerical analysis, machine learning and multi-omics analytics. Since the input data are in the order of Gigabytes or Terabytes, most of our codes are written to exploit distributed-memory systems. A number of graph/combinatorial applications are also challenging to efficiently scale (both strong and weak scaling cases) due to the inherent irregular memory-access intensive computation patterns. Hence, scaling up to larger inputs may require developing new algorithms or heuristics. In addition, relatively high memory requirements and communication overheads often necessitate increasing the number of nodes and identifying mechanisms for communication avoidance usually leads to improved performance.

Therefore, it is vital to identify the trade-offs that can be exposed such that a user can select options that can yield the desired performance at the expense of quality/accuracy. Thus, a machine allocation award will lead to a number of cutting-edge software/algorithmic innovations in the area of parallel combinatorial scientific computing, which can enhance the performance of diverse scientific workloads.

Computational Resources:

Total node hours per year: 10,000

Size (nodes) and duration (hours) for a typical batch job: 30-75 minutes

Disk space (home, project, scratch): 20 TB

Personnel Resources: Prospective users in our project are already conversant with cluster computing, but one or two general training sessions per calendar year would be appreciated.

Required software:
MPI, Python, HPCToolkit

If your research is supported by US federal agencies:

Agency: DOE ECP

Grant number(s): 17-SC-20-SC

The research is in parts supported by the U.S. DOE ExaGraph project.
https://sites.google.com/lbl.gov/exagraph/

Production projects:

Production projects should provide an additional 1-2 pages of documentation about how
(a) the code has been tuned to perform well on A64FX (ideally including benchmark data comparing performance with other architectures such as x86 or GPUs)
(b) it can make effective use of the key A64FX architectural features (notably SVE, the high-bandwidth memory, and NUMA characteristics)
(c) it can accomplish the scientific objectives within the available 32 Gbyte memory per node