Hyungro Lee

Dept. of Electrical & Computer Engineering
Rutgers, The State University of New Jersey
94 Brett Rd
Piscataway, NJ 08854

phone: +1 (812) 322 - 6252
E-mail: hyungro.lee@rutgers.edu
Web: https://lee212.github.io/

EDUCATION

Indiana University, Bloomington, IN, **Ph.D. in Computer Science**, 2019. Advisor: Prof. Geoffrey C. Fox. Dissertation: *Building Software-Defined Systems with Reproducible Environments*

Indiana University, Bloomington, IN, **M.S. in Computer Science**, 2010. Advisor: Prof. Sun Kim.

Sungkyunkwan University, Seoul, Korea, **B.S. in Computer Education**, 2007

RESEARCH INTERESTS

My research interests are parallel and distributed systems, HPC for deep learning frameworks and large scale scientific applications

EMPLOYMENT

Postdoctoral Research Associate with Prof. Shantenu Jha, Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ October 2018-Present

Teaching Assistant with Prof. Judy Qiu and Prof. Gregor von Laszewski, Department of Intelligent Systems Engineering, Indiana University 2014-August 2018

Research Intern with Prof. Dennis Gannon, Microsoft, Redmond, WA Summer 2013

Research Assistant with Prof. Geoffrey C. Fox, School of Informatics, Computing, and Engineering, Indiana University, 2009-2013

Research Intern with Prof. Kyusang Lee, Samsung Advanced Institute of Technology, Giheung, Korea Summer 2010

Software Engineer, SK Communications, Seoul, Korea 2004-2008

Software Engineer, JPD Internet, Seoul, Korea 2002-2004

PUBLICATIONS

Lorenzo Casalino, Abigail C Dommer, Zied Gaieb, Emilia P Barros, Terra Sztain, Surl-Hee Ahn, Anda Trifan, Alexander Brace, Heng Ma, **Hyungro Lee**, Matteo Turilli, Anthony Bogetti, Syma Khalid, Lillian Chong, Carlos Simmerling, David Hardy, Julio Maia, James Phillips, Thorsten Kurth, Abraham Stern, Lei Huang, John McCalpain, Mahidhar Tatineni, Tom Gibbs, John E Stone, Shantenu Jha, Arvind Ramanathan, Rommie E Amaro. "AI-driven multiscale simulations illuminate mechanisms of SARS-CoV-2 spike dynamics". to appear as winner for the Supercomputing 2020 (SC20) Gordon Bell Special Prize in the *International Journal of High Performance Computing Applications (IJHPCA)*. 2021. doi: 10.1101/2020.11.19.390187.

Austin Clyde, Stephanie Galanie, Daniel W. Kneller, Heng Ma, Yadu Babuji, Ben Blaiszik, Alexander Brace, Thomas Brettin, Kyle Chard, Ryan Chard, Leighton Coates, Ian Foster, Darin Hauner, Vilmos Kertesz, Neeraj Kumar, **Hyungro Lee**, Zhuozhao Li, Andre Merzky, Jurgen G. Schmidt, Li Tan, Mikhail Titov, Anda Trifan, Matteo Turilli, Hubertus Van Dam, Srinivas C. Chennubhotla, Shantenu Jha, Andrey Kovalevsky, Arvind Ramanathan, Martha S. Head, and Rick Stevens. "High Throughput Virtual Screening and Validation of a SARS-CoV-2 Main Protease Non-Covalent

Inhibitor". *Proceedings of the National Academy of Sciences (PNAS)* (submitted). 2021. doi: 10.1101/2021.03.27.437323.

Alexander Brace, **Hyungro Lee**, Heng Ma, Anda Trifan, Matteo Turilli, Igor Yakushin, Todd Munson, Ian Foster, Shantenu Jha, Arvind Ramanathan, "Achieving 100X faster simulations of complex biological phenomena by coupling ML to HPC ensembles", *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC21)* (submitted). 2021. arXiv:2104.04797.

Clyde, A., Wright, D., Ahmed, K., Chodera, J., Karanicolas, J., Kirubakaran, P., **Lee, H.**, Turilli, M., Wan, S., Coveney, P., Jha, S., and Stevens, R., "Integrating High-Performance Simulations and Learning towards Improved Cancer Therapy". *BMC Bioinformatics* (invited), 2021

Lee, H., Merzky, A., Tan, L., Titov, M., Turilli, M., Alfe, D., Bhati, A., Brace, A., Clyde, A., Coveney, P. and Ma, H., 2020. "Scalable HPC and AI Infrastructure for COVID-19 Therapeutics". *Proceedings of the Platform for Advanced Scientific Computing Conference (PASC 21)* (accepted) 2021. arXiv preprint arXiv:2010.10517.

Eric Wilson, John W. Vant, Jacob Layton, Ryan Boyd, **Hyungro Lee**, Matteo Turilli, Benjamin Hernandez, Sean Wilkinson, Shantenu Jha, Chitrak Gupta, Daipayan Sarkar and Abhishek Singharoy, "All You Want to Know About Large System Simulations", *Methods in Molecular Biology.* (Accepted, pending publication), 2021

Norbert Podhorszki, Greg Eisenhauer, Dmitry Ganyushin, Igor Yakushin, Ruonan Wang, Todd Munson, Scott Klasky, Alex Brace, Heng Ma, Matteo Turilli, **Hyungro Lee** and Ian Foster. "Accelerating machine learning workflows with online staging and file-streaming using ADIOS", *Smoky Mountains Computational Sciences and Engineering Conference* (submitted) 2021.

Aymen Al Saadi, Dario Alfe, Yadu Babuji, Agastya Bhati, Ben Blaiszik, Thomas Brettin, Kyle Chard, Ryan Chard, Peter Coveney, Anda Trifan, Alex Brace, Austin Clyde, Ian Foster, Tom Gibbs, Shantenu Jha, Kristopher Keipert, Thorsten Kurth, Dieter Kranzlmüller, **Hyungro Lee**, Zhuozhao Li, Heng Ma, Andre Merzky, Gerald Mathias, Alexander Partin, Junqi Yin, Arvind Ramanathan, Ashka Shah, Abraham Stern, Rick Stevens, Li Tan, Mikhail Titov, Aristeidis Tsaris, Matteo Turilli, Huub Van Dam, Shunzhou Wan, David Wifling. "IMPECCABLE: Integrated Modeling PipelinE for COVID Cure by Assessing Better LEads." arXiv preprint arXiv:2010.06574, 2020.

Eugen Hruska, Vivekanandan Balasubramanian, **Hyungro Lee**, Shantenu Jha, and Cecilia Clementi, "Extensible and Scalable Adaptive Sampling on Supercomputers", *Journal of Chemical Theory and Computation (JCTC)* 2020 16 (12), 7915-7925. doi: 10.1021/acs.jctc.0c00991.

Ma, Heng, Debsindhu Bhowmik, **Hyungro Lee**, Matteo Turilli, Michael Young, Shantenu Jha, and Arvind Ramanathan. "Deep generative model driven protein folding simulations." *In Parallel Computing: Technology Trends*, pp. 45-55. IOS Press BV, 2020. doi: 10.3233/APC200023.

H. Lee, M. Turilli, S. Jha, D. Bhowmik, H. Ma and A. Ramanathan, "DeepDriveMD: Deep-Learning Driven Adaptive Molecular Simulations for Protein Folding," *2019 IEEE/ACM Third Workshop on Deep Learning on Supercomputers (DLS)* in cooperation with TCHPC and held in conjunction with SC19, Denver, CO, USA, 2019, pp. 12-19, doi: 10.1109/DLS49591.2019.00007.

Lee, Hyungro, and Geoffrey Fox. "Big data benchmarks of high-performance storage systems on commercial bare metal clouds." *2019 IEEE 12th International Conference on Cloud Computing (CLOUD). IEEE*, 2019. doi: 10.1109/CLOUD.2019.00014.

Lee, Hyungro, Kumar Satyam, and Geoffrey Fox. "Evaluation of production serverless computing environments." *2018 IEEE 11th International Conference on Cloud Computing (CLOUD). IEEE*, 2018. doi: 10.1109/CLOUD.2018.00062.

Judy Qiu, Supun Kamburugamuve, **Hyungro Lee**, Jerome Mitchell, Rebecca Caldwelly, Gina Bullockz, Linda Hayden, "Teaching, Learning and Collaborating through Cloud Computing Online

Classes", *Proceedings of the Workshop on Education for High Performance Computing (eduHPC) at SC17* September 2017. [Program Link 1] [2] [Article Link]

Lee, Hyungro, and Geoffrey C. Fox. "Efficient Software Defined Systems Using Common Core Components." *2017 IEEE 10th International Conference on Cloud Computing (CLOUD). IEEE*, 2017. doi: 10.1109/CLOUD.2017.59.

Lee, H., Lee, M., Mohammed Ismail, W., Rho, M., Fox, G. C., Oh, S., & Tang, H. (2016). "MGEScan: a Galaxy-based system for identifying retrotransposons in genomes". *Bioinformatics*, 32(16), 2502-2504. doi: 10.1093/bioinformatics/btw157.

Von Laszewski, Gregor, Fugang Wang, **Hyungro Lee**, Heng Chen, and Geoffrey C. Fox. "Accessing multiple clouds with cloudmesh." *In Proceedings of the 2014 ACM international workshop on Software-defined ecosystems*, pp. 21-28. 2014. doi: 10.1145/2609441.2609638.

Chae H, Jung I, **Lee H**, Marru S, Lee S, Kim S. "Bio and Health informatics meets Cloud: BioVLab as an example", *Health Information Science and Systems*, BioMed Central Ltd, 2013, 1, 6. doi: 10.1186/2047-2501-1-6.

Lee H, Yang Y, Chae H, Nam S, Choi D, Tangchaisin P, Herath C, Marru S, Nephew K, Kim S. "BioVLAB-MMIA: A Cloud Environment for microRNA andmRNA Integrated Analysis (MMIA) on Amazon EC2", *IEEE Transactions on NanoBioscience*, 09/2012; 11(3):266-72.
doi: 10.1109/TNB.2012.2212030.

von Laszewski, G., **Lee, H.**, Diaz, J., Wang, F., Tanaka, K., Karavinkoppa, S., Fox, G.C. and Furlani, T., 2012, September. "Design of an accounting and metric-based cloud-shifting and cloud-seeding framework for federated clouds and bare-metal environments." In *Proceedings of the 2012 workshop on Cloud services, federation, and the 8th open cirrus summit* (pp. 25-32). ACM. doi: 10.1145/2378975.2378982.

Lee, Hyungro, et al. "BioVLAB-MMIA: a reconfigurable cloud computing environment for microRNA and mRNA integrated analysis." *Bioinformatics and Biomedicine (BIBM)*, 2011 IEEE International Conference on. IEEE, 2011. doi: 10.1109/BIBM.2011.93.

PRESENTATIONS, POSTERS, and REPORTS

Austin Clyde, Dave Wright, Katya Ahmed, John D. Chodera, John Karanicolas, Palani Kirubakaran, **Hyungro Lee**, Matteo Turilli, Shunzhou Wan, Peter Coveney, Shantenu Jha, Rick Stevens, "Integrating High-Performance Simulations and Learning toward Improved Cancer Therapy", *Fifth Computational Approaches for Cancer Workshop (CAFCW19)*, Denver, CO, 17 November 2019, Presentation by Austin Clyde. [Slides]

Hyungro Lee, Geoffrey Fox, "Event-driven Computing on HPC: Experiments with Scientific Applications", *RIKEN AICS International Symposium* in KOBE, JAPAN, February 2018, Presentation. [Abstract]

Hyungro Lee, Geoffrey Fox, "Software Defined Systems with DevOps Tools and Infrastructure Provisioning", *Ph.D. Forum at IPDPS* conference, Orlando FL May 30-June 2, 2017, Poster. doi: 10.13140/RG.2.2.20831.25762/2.

Badi' Abdul-Wahid, **Hyungro Lee**, Gregor von Laszewski, and Geoffrey Fox, "Scripting Deployment of NIST Use Cases" *Technical Report* January 20 2017. doi: 10.13140/RG.2.2.16201.54881.

Lee, Hyungro, et al. "Towards understanding cloud usage through resource allocation analysis on xsede". *Technical report*, March 25 2014. [Report]

Lee, Hyungro, "Using Bioinformatics Applications on the Cloud.", *Technical Report*, 2014. [Report]

Lee, Hyungro, "Virtualization Basics: Understanding Techniques and Fundamentals.", *Technical*

Report, 2014. [Report]

TEACHING EXPERIENCE

ENGR-E 222: Intelligent Systems II   (Spring 2018)

INFO-I 523: Big Data Applications & Analytics   (Fall 2017)

CSCI-B 649: Cloud Computing   (Spring 2017)

INFO-I 590: Topics in Informatics   (Spring 2016)

INFO-I 524: Big Data Open Source Software and Projects   (Fall 2015)

BUEX-V 594: Business Global Executive Management   (Spring 2015)

Research Experiences for Undergraduate (REU) in Computing   (Summer 2014)

ACTIVITIES & TALKS

Participant at *CANcer Distributed Learning Environment (CANDLE) hackathon*, Argonne National Laboratory, Chicago IL, November 11 - 14 2019.

Presentation at *RIKEN AICS Youth group workshop*, Kobe, Japan, February 5 - 6 2018.

Participant at *AICS International Symposium*, Kobe, Japan, February 7 - 8 2018.

Participant at *SC 17 Student Volunteer Program*, Denver CO, November 11 - 17 2017.

Participant at *IPDPS 2017 PhD Forum*, Orlando FL, May 29 - June 2 2017.

Presentation at *ADMI Symposium on Computing at Minority Institutions*, Virginia Beach, FL March 2017.

Undergraduate Mentor *Summer REU Research Program*, Bloomington, IN, June 2014.

Participant at *SciPy Conference*, Austin, TX, June 2013.

Presentation at *Targets and Tools for Therapeutic Development*, Cambridge Healthtech Institute, Boston, MA March 2013.

PROFESSIONAL SERVICE

Reviewer: Briefings in Bioinformatics, Online ISSN 1477-4054, 2018 - 2019

Reviewer: Concurrency and Computation Practice and Experience, Online ISSN 1532-0634, 2015 - 2018

TECHNICAL SKILLS

Language	C, Python, Bash Shell
Software	Horovod, Tensorflow/Keras, PyTorch DDP, Dask distributed, cuML/cuDF, Ganglia, Nagios, OpenStack, Ansible, Docker, nvprof/nvvp, cProfiler, HDF5, Apache Arrow, TravisCI/codecov/pytest, RabbitMQ, MongoDB, MySQL
Certificate	Cisco CCNA, Engineer Information Processing (Korea)

AWARDS & HONORS

ACM Gordon Bell Special Prize 2020 for High Performance Computing-Based COVID-19 Research [Press Link 1] [Press Link 2]
Title: "AI-Driven Multiscale Simulations Illuminate Mechanisms of SARS-CoV-2 Spike Dynamics", Team: Lorenzo Casalino, Abigail Dommer, Zied Gaieb, Emilia P. Barros, Terra Stzain, Surl-Hee

Ahn, Anda Trifan, Alexander Brace, Anthony Bogetti, Heng Ma, **Hyungro Lee**, Matteo Turilli, Syma Khalid, Lillian Chong, Carlos Simmerling, David Hardy, Julio Maia, James Phillips, Thorsten Kurth, Abraham Stern, Lei Huang, John McCalpin, Mahidhar Tatineni, Tom Gibbs, John Stone, Shantenu Jha, Arvind Ramanathan and Rommie E. Amaro. This work enables AI workflow for molecular systems on HPC resources.

US Department of Energy Secretary's Honor Award: Being part of the National Virtual Biotechnology Laboratory w.r.t. COVID19 therapeutics [Press Link]

IPDPS 2017 Student Travel Grants

SC17 Travel Grant

IU CSCI Grad Fellowship, PhD Award, PhD Student Travel award, ISE (ENGR) Fellowship, 2008, 2010, 2017, 2018

SOFTWARE PORTFOLIO

task_monitoring is a Python library that provides real-time monitoring on HPC systems associated with RADICAL-Cybertools (radical-pilot and radical-entk) by mapping annotations to process (RP_TASK_NAME). Programming Language: Python. https://github.com/lee212/task_monitoring

MGEScan is a computing environment integrated with the Galaxy workflow system for identifying retrotransposons in genome. Programming Language: MPI, Python.
https://github.com/MGEScan/mgescan

SimpleAzure is a Python library for Microsoft Azure Services including Virtual Machine (VM) to provision resources in a simple way. Programming Language: Python.
https://github.com/lee212/simpleazure