



Parallel Architecture and System Software Research Laboratory

<http://passlab.github.io/>
<https://github.com/passlab/>

Department of Computer Science and Engineering
Yonghong Yan

<http://www.cse.sc.edu/~yanyh>

Spring 2019

PASSLab Team

- Faculty
 - Yonghong Yan
- Ph.D. Graduate Students and Starting Semester
 - Kewei Yan, Spring 2018
 - Philip Conrad, Spring 2018
 - Anjia Wang, Summer 2018

 - Yudong Sun, Fall 2018
 - Yaying Shi, Fall 2018
 - Xin Zhao, Fall 2018

 - Xinyao Yi, Spring 2019
- Master Students: TBA
- Undergraduate Students: TBA

- Senior People (postdoc or research scientist): TBA
- Visiting scholar: Welcome <http://passlab.github.io/people/>

Our Research Areas in General

- Address **Performance**, **Power** and **Programmability** challenges of **Parallel** computing
 - Target High Performance Computing (HPC) and emerging applications such as image processing and machine learning
- Advance state of the art of computer systems
 - Parallel programming models and compiler
 - Computer architecture and systems
 - Parallel performance tools
- Support interdisciplinary research for computational and data scientists
 - Developing and optimizing applications for parallel hardware
 - Developing domain-specific programming languages and system support

What We Do:

- Develop intra-/inter-node parallel programming models, compiler, runtime systems and performance tools
 - Work with OpenMP, MPI and domain-specific languages
 - OpenMP compiler based on ROS and runtime based on LLVM/Intel OpenMP runtime
- Develop tools for performance analysis and optimization
 - OpenMP tool via hardware counters (PAPI or perf), OMPT, LTTng, Eclipse TraceCompass
 - Binary analysis and instrumentation using Intel Pin
- Architecture for parallel and domain-specific computing
 - FPGA implementation for kernels of material science and other domains
 - Emerging memory and cache systems
 - RISC-V-based research (NVDLA+Rocket, SMT BOOM, etc)

What We Do:

- Jupyter-related and data management
 - Jupyter-based framework for parallel programming and education
 - Data management framework and integration with other data analysis tools and cloud storage
- Collaborate with scientists to parallelize and optimize real problem
 - Medical imaging applications on GPU using CUDA
 - Deformable image registration for radiation dose calculation
 - SNL for HPCG and mini-/proxy-apps for DoE
 - Material science
- Interact and educate students, peers and community to know the area of parallel computing and HPC
- To support existing and emerging hardware for HPC
 - Multicore CPUs, NVIDIA GPUs, and Intel Xeon Phi
 - Intel/Xilinx FPGA using OpenCL and Mexler dataflow FPGA
 - HBM/HMC memory in GPU and Xeon Phi, and NV RAM
 - Cluster and heterogeneous systems of CPU/GPU/Phi/FPGA

Current Project

- CAREER: Programming the Existing and Emerging Memory Systems for Extreme-scale Parallel Performance
 - NSF, 5 years till 2022
- SHF: Medium: Compute on Data Path: Combating Data Movement in High Performance Computing
 - NSF, one year left till 2019
- SHF: Small: Collaborative Research: Application-aware Energy Modeling and Power Management for Parallel and High Performance Computing
 - NSF, one year left till 2019
- Connecting ROSE to LLVM OpenMP Runtime
 - LLNL, one year till May 2019, extensions based on progress
- **Using Jupyter Notebook and Cloud IDE for the Programming Activities of CSE Courses, new 01/2019 – 12/2019**

Current Project

- Hardware Design and Computer Architecture
 - RISC-V (<https://riscv.org/>) based on Rocket and BOOM chips
 - <https://github.com/freechipsproject/rocket-chip>
 - <https://github.com/ucb-bar/riscv-boom>
 - NVIDIA Deep Learning Accelerator (NVDLA) (<http://nvdla.org/>)
 - FPGA with High Bandwidth Memory from Intel/Xilinx and Maxeler
 - **FPGA development for bio-informatics and material science**
- Data management and and computation integration for medical image data
 - **Work with MD Anderson**

Pending or Upcoming Project

- BIGDATA: F: Polyglot Programming and Execution Environment to Synergize Computational Simulation and Big Data Analytics for Scientific Discovery
 - NSF, Four Years till 2022/2023
- In-situ Performance Analysis and Optimization for Automated Performance Tuning of Extreme-scale Applications:
 - NSF, 3-4 years till 2022/2023
- High Performance Implementation using GPU and FPGA
 - Deformable Image Registration for Cancer Treatment
 - on-hold for a good CUDA development person
 - FPGA/CUDA implementation for material science kernel

Exploration

- Validation and Performance Optimization of AV Software System and Algorithms Using NVIDIA DRIVE Constellation
 - NVIDIA Equipment donation, at least one year till 2019
 - System software and algorithms for AV systems, algorithms, and DL
- Cognitive drone using NVIDIA Jetson
- Deep learning framework and integration with HPC system stack
- New computational paradigm: quantum computing, blockchains, etc

Sponsor and Collaborators

- Sponsor:
 - National Science Foundation
 - DoE/LLNL
 - Industries: NVIDIA/Xilinx/Intel, etc

- Collaborators
 - DoE/LLNL ROSE Team
 - DoE Sandia, Ron Brightwell
 - Virginia Tech, Texas Tech, Stony Brook
 - Industries: via OpenMP ARB

Software and Hardware Resources

- Each of you have a development desktop
- {fornax, carina, pivo}.cse.sc.edu
 - <https://github.com/passlab/passlab.github.io/wiki/Hardware-and-Resources>
 - High-end server with GPUs
- RCI cluster
 - https://www.sc.edu/about/offices_and_divisions/division_of_information_technology/rci/
- XSEDE bridges
 - <https://github.com/passlab/passlab.github.io/wiki/Hardware-and-Resources#nsf-xsede-bridges-from-psc>

<https://github.com/passlab>

- Github organization for almost all of our work
 - People: [passlab team and collaborators](#)
 - Teams/subteams in [passlab](#) for access control
 - Maillist: passlab@googlegroups.com
- <https://github.com/passlab/passlab.github.io> repo:
 - Public website at <http://passlab.github.io/>
 - **Add your name and photo**
 - Internal wiki shared by all the [passlab](#) members
 - <https://github.com/passlab/passlab.github.io/wiki>
 - **You should update often**
- Repos for our development and papers/proposals
- Each of you can have your private repo
 - For your own notes (I can see it if I want to 😊).
 - For creating your public website linked as: <http://passlab.github.io/yanyh>

Development and Tools

- Programming Languages
 - OpenMP, MPI, Python, C/C++/Fortran and CUDA; OpenCL/OpenACC/etc
- Compiler:
 - ROSE, Clang/Flang/LLVM
- Runtime library system
 - LLVM OpenMP runtime, Open MPI, and runtime library for various other languages
- Performance and tools:
 - OMPT, MPI_T, PAPI, Linux perf, LTTng, pintools, (hpctoolkit, TAU, Score-P, etc).
 - SENSEI in-situ, VisIt, Eclipse/Compass, Jupyter notebook
- Hardware and FPGA development
 - RISC-V, Chisel, Verilog, vendor tools

Software Development and Release

- Developing and releasing robust software are VERY IMPORTANT
 - Encourage collaboration
 - Sharing with community
 - Demonstrate professionalism
- We will release our compiler, runtime and tools
- To note:
 - Following coding/naming guideline
 - Write good document
 - Develop easy-to-read code
 - Sharing from github

Development Environment

- Familiar with Linux command line development
 - make/cmake, and autotools
 - Compiler: ROSE, Clang/LLVM, GNU, Intel
 - How to use those performance tools mentioned before
 - IDE: CLion/Eclipse for C/C++ development
- Git and github
 - Git commands for branch, remote, push/pull, merge, pull request, etc
 - github source code/wiki development workflow
- Documentation
 - Wiki in markdown format, tutorial in markdown and Jupyter notebook
 - Presentation using MS powerpoint or Mac Keynote, or google presentation
 - Sheet using MS Excel sheet, or Google sheet
 - Short paper or document: MS words or google doc
 - Technical paper or proposal: Latex

Research Topics and Research Project

- You have a research project that support your study as RA
- Your research topics will be based on the research project that funds you.
 - Likely that you research topics span multiple research projects
 - Likely you will work on more than one research project
- Your development will definitely used for multiple projects.

Education and Outreach

- You will be asked to give lectures or tutorial in some of my course, in our group meeting or for other group in the university, e.g. RCI.
- You may be requested to give talks/tutorials to external groups.
- Subscribe to magazine, mail-list, etc
 - ROSE, LLVM, OpenMP, hpc-announce, hpcwire, nextplatforms
- Participate and volunteer in outreach activities related to engineering and our research area
 - Promote our research
 - Service the community

Meeting and Organization

- Weekly project meeting, concentrate the ongoing work/paper
 - Compiler: Anjia, Xinyao, Yudong, Xin Zhao: 9:45 – 10:45AM Monday, need to reserve a room.
 - Runtime system: Kewei, Yudong, Anjia, Yaying: 11:00 – 12:00PM Monday, need to reserve a room
 - Applications: Xin: 1:30-2:30PM Wednesday
 - Hardware: Yaying: 2:30-3:30PM Wednesday
 - Tools: Philip, Yudong: Wednesday afternoon likely
 - Medical image big data (python, jupyter, ckan): Anjia,
- Weekly 1:1 meeting on Monday or Wednesday
 - Xinyao: Wednesday 9:45 – 10:15
 - Yudong: Wednesday 10:15 - 10:45
 - Anjia: Wednesday 11:00 – 11:30AM
- Group meeting, likely Friday 1:00-2:00PM.
 - Present by each for some topics

Responsibility

- 20 hours/week work officially as full-time RA
 - Concentrate on delivering your work and being productive
 - I expect you to contribute at least after the first semester
- Your work:
 - **Research:** search and re-search related work and write study doc about what others do and their pros/cons
 - **Development:** write code that work
 - **Learning:** study and teach, learn something and teach others
 - **Outreach:** promote, marketing, volunteer, etc.
- Weekly report by Friday of the week
 - 1/3 to half page (~100-300 words), sent to me by email
 - <https://github.com/passlab/passlab.github.io/wiki/Information-About-Weekly-Work-Notes>
 - Keep your worknote on wiki as well

Publications

- Publications are what you put on your resume
- You should set yourself a goal to produce at least one *thing* per semester as first author including
 - Workshop/conference/journal paper, and technical report
 - We set goals at the beginning of the semester
- Check for conferences/workshops and due date
 - <https://github.com/passlab/passlab.github.io/wiki/Conference-for-submissions>

Ph.D. Study and Graduation

- ~ 5 years and it depends
 - Qualifying exam
 - Research, development, publication and others
 - Preliminary defense
 - Dissertation and oral defense
- Department general guideline: 4 years for master and 5 years for undergraduate
- Publications expectation as first author (3 first-author tier 1 paper)
 - > 1 journal
 - 2 tier 1 conference/journal papers
 - >3 workshop papers and joint papers
 - Demonstrate your research scholarship, development skills, presentation, service and maturity in professional level

Travel

- Your travel to attend conference/workshop to present published paper will be covered by our research funding
 - Talk with me before you plan to submit to specific place
 - E.g. We normally do not send you to international place for just presenting a poster
- Let me know as early as possible your personal travel plan during the semester and summer

Internship and External Collaboration

- Internship are encouraged if it is related to your research topic and project
- I will help to find internship opportunity during your study
 - Industries and national lab
- External collaboration are encouraged and required
 - Know the people who work in the area like you
 - Find opportunity to interact with them.

Support for you

- Advising
 - Weekly meeting
 - Group meeting
- Your study plan (Ph.D. or master) and your career plan
- Collaboration and outreach opportunities
 - Internship, attending conference
- Financial support
 - RA and TA, we may put you as TA based on the funding situation and your performance
 - RA Salary will be based on performance
- Other help
 - Let me know

You

- Ambition and goals
 - To make impact
- Achieving excellence through self-motivation and dedication
 - Lack of background or resources is not an issue
 - Work hard and work smart
 - Open-minded and be prepared
- Laziness and dishonesty are not tolerable

Training Needed

- For all (2 to 4 weeks)
 - Linux, C and vim
 - Compilation, linking, make/makefile/cmake/autotools study and practice
 - Git and github (branches, pull request), diff and patch
 - IDE: CLion for C/C++, IntelliJ IDE for Java/Scala, Python IDE
 - Markdown, wiki, ppt, excel, google docs/sheet
 - OpenMP programming
 - PThread programming (assignment for implementing basic OpenMP fork-join/barrier/single/master using pthreads)
- For specific area
 - Compiler: ROSE
 - Runtime: LLVM OpenMP runtime, REX, cilkplus runtime
 - Tools: OMPT, RTune, etc
 - Computer architecture: Scala/Chisel, RISC-V, Rocket-chips, ARM embedded design
 - Applications and benchmarks: AXPY (stream), matvec, mm, LU/Cholesky/QR, Rodinia, mini-apps, proxy-apps
 - Fortran/CUDA/MPI programming
 - Jupyter notebook
 - Spark/Datarex related