

Parallel Applications Workshop, Alternatives to MPI+X

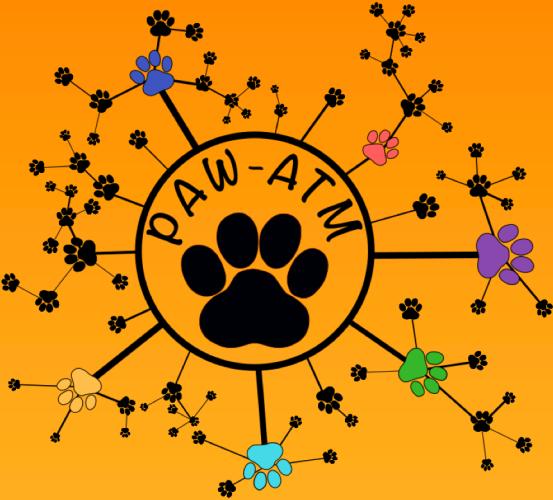
November 17th, 2019

Held in conjunction with SC19: The International Conference for
High Performance Computing, Networking, Storage, and Analysis



In cooperation with:





Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

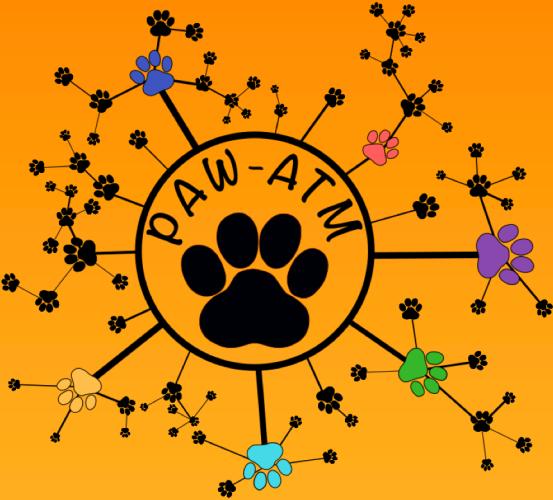
Denver, CO

MPI: HPC's *de facto* programming model

MPI (Message Passing Interface)

[MPI Forum, ~1992–]

- a library for inter-process communication and synchronization
- run a process per compute node, core, or NUMA domain
- call routines for point-to-point messaging, collectives, barriers, ...



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO



MPI+X: Adding intra-node parallelism

MPI+X

- X = something supporting finer-grain parallelism
- typically used for multi-core or hybrid computing (e.g., GPUs)

X could be...

...**Pthreads**

[POSIX, ~1995–]

...**OpenMP**

[OpenMP ARB, ~1997–]

...**OmpSs**

[BSC, ~2010–]

...**CUDA**

[Nvidia, ~2007–]

...**OpenCL**

[Khronos Group, ~2009–]

...**OpenACC**

[OpenACC.org, ~2012–]

...**Kokkos**

[Sandia, ~2012–]

...**RAJA**

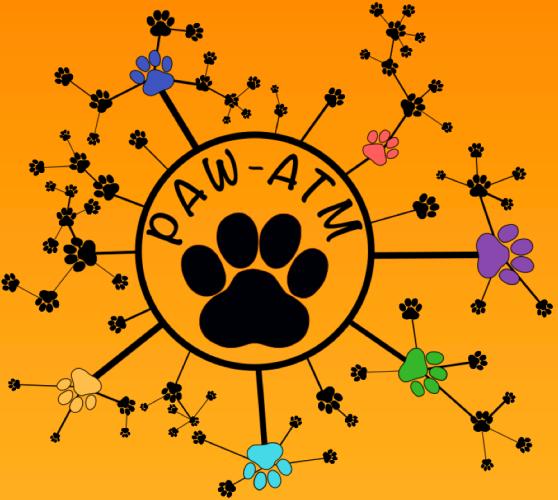
[LLNL, ~2016–]

...

• Included in Today's Talks

• Active

• Dormant



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO

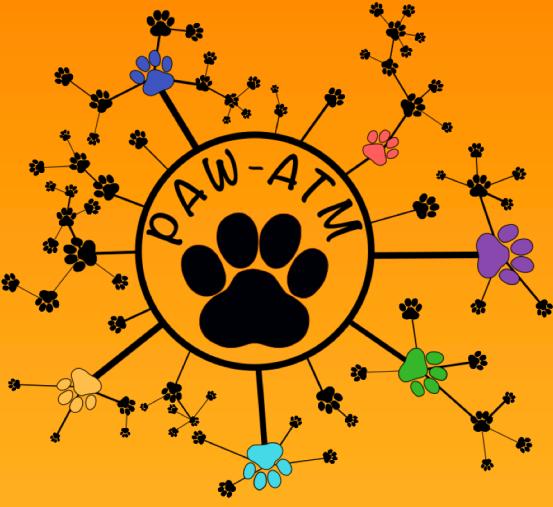
MPI+X: Pluses and Minuses

Pluses:

- general-purpose
- adopted / ubiquitous
 - MPI is at least; some choices of X are less so
- explicit:
 - if the system can do something, you probably can too

Minuses:

- explicit:
 - detail-oriented
 - relatively low-level
 - not very insulated from architectural details and changes



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO

A Historical Analogy with Fortran

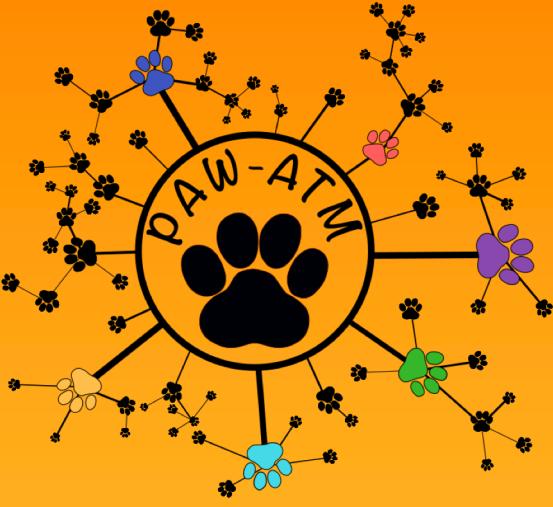
In Fortran's early days, programmers were concerned about switching to it from assembly language, due to:

- loss of explicit control
- fear of performance impact

However, Fortran succeeded, providing many productivity benefits

- Its success led to other successful high-level languages:
 - e.g., C, C++, Java, Python, Go, Rust, Swift, ...
- Today, we rarely worry about things like register allocation
 - besides, can always drop down to assembly if required

In HPC, we've been less successful at creating—and broadly adopting—higher-level alternatives to MPI+X to reap similar benefits



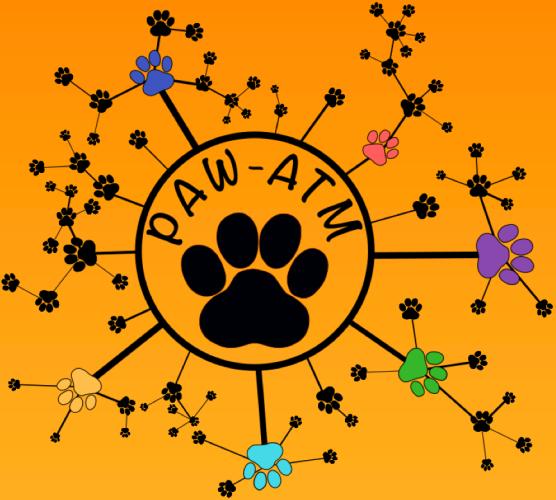
Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO

PAW-ATM's Goals

- 1) To provide a forum to discuss **Alternatives to MPI+X**
 - capable of targeting HPC-scale **Parallel** systems
- 2) To highlight **Applications** from real users to focus on concrete and applied cases



Parallel Applications Workshop
Alternatives to MPI+X

November 17th, 2019

Denver, CO

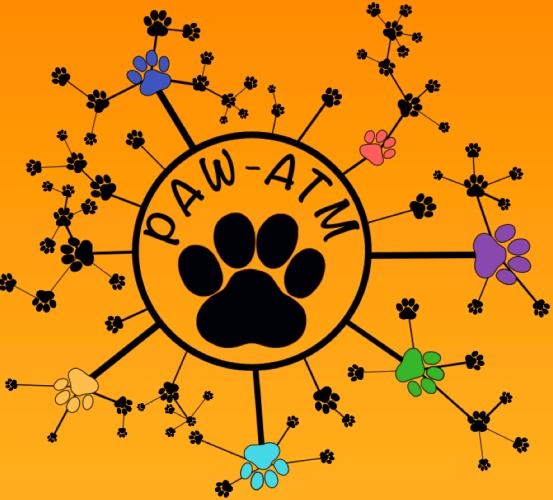
This Talk: Goals and Disclaimers

Goals:

- ✓ Give context and motivation for PAW-ATM 2019
- Provide a mental framework for today's talks

Disclaimers:

- This will be a rapid summary of a complex topic
 - Technologies will be dramatically oversimplified
 - Others will be bypassed altogether
- My apologies for any big mistakes or omissions



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO



Early Alternatives to MPI

- High Performance Fortran
- ZPL
- NESL
- Charm++
- ...

[HPF Forum, ~1991–2002]

[U. Washington, ~1992–2005]

[CMU, ~1993–1997]

[UIUC, ~1993–]



PGAS Approaches

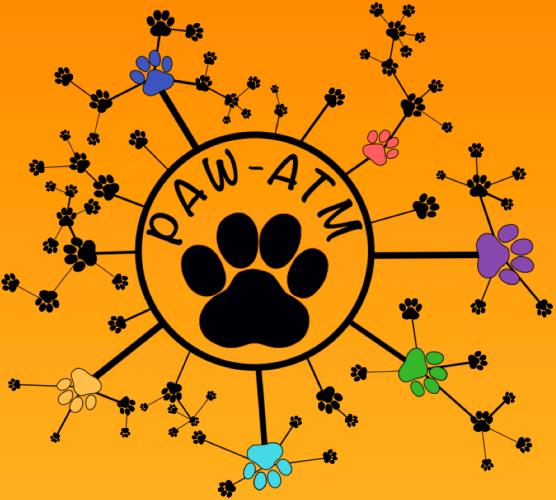
PGAS: Partitioned Global Address Space

- **unifying theme:** refer to remote values by naming them
- this “name” could be...
 - ...a remote address
 - ...the name of a variable stored remotely
 - ...a remote element of a distributed array
 - ...

Parallel Applications Workshop
Alternatives to MPI+X

November 17th, 2019

Denver, CO



Parallel Applications Workshop
Alternatives to MPI+X

November 17th, 2019

Denver, CO



Original PGAS Approaches

PGAS-oriented Libraries:

- **[Open]SHMEM**
- **Global Arrays**
- **ARMCI**
- **GASNet[-EX]**

[OpenSHMEM project / Cray, ~1993–]

[PNNL, ~1994–]

[PNNL, ~1999–2010]

[LBL / UC Berkeley, ~2002–]

PGAS languages

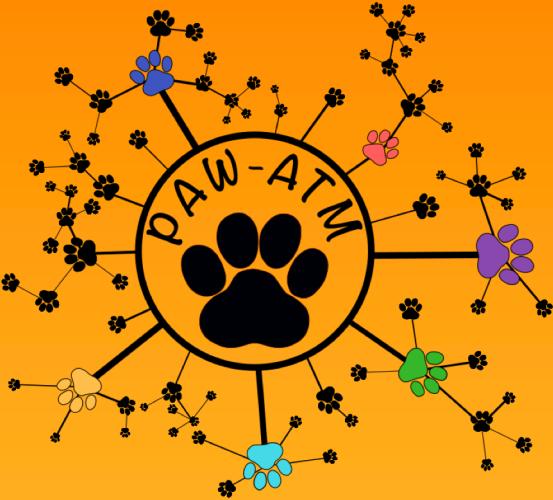
- **Fortran 2008 / 2018**
- (previously **Co-Array Fortran, F--**)
- **Titanium**
- **Unified Parallel C (UPC)**

[Fortran Working Group, ~1997–]

[Cray, et al., ~1997–]

[UC Berkeley, ~1998–2005]

[UPC Consortium, ~1999–]



Parallel Applications Workshop
Alternatives to MPI+X

November 17th, 2019

Denver, CO



Next-Generation PGAS Approaches

HPCS Languages

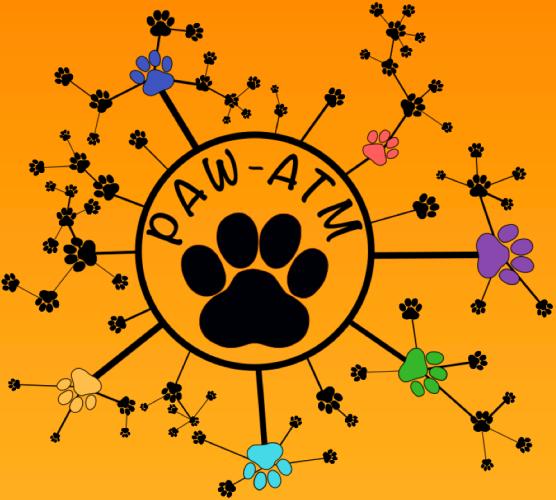
- **Chapel** [Cray, ~2003–]
- **X10** [IBM, ~2004–]
- **Fortress** [Sun/Oracle, ~2005–2012]

Others

- **XcalableMP** [XMP Working Group, ~2008–]
- **PCJ** [Nicolaus Copernicus U. / U. Warsaw, ~2012–]
- ...

C++ Template Libraries

- **Coarray C++** [Cray, ~2013–]
- **UPC++** [UC Berkeley, ~2014–]



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO



Task-based Alternatives to MPI

Runtime Systems

- **HPX**
- **Legion**
- ...

[Stellar Group, ~2012–]

[Stanford / SLAC / LANL / NVIDIA, ~2012–]

Programming Models / Languages

- **Charm++**
- **PaRSEC**
- **[Py]COMPSSs**
- **Regent**
- **Pygion**
- ...

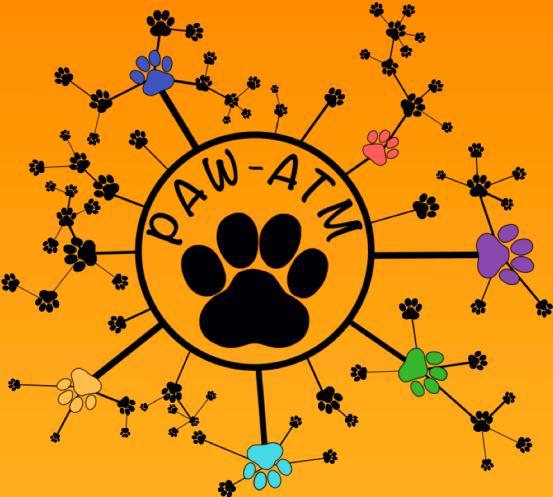
[UIUC, ~1993–]

[U. Tenn. Knoxville, ~2009–]

[BSC, ~2011–]

[Stanford / SLAC, ~2015–]

[Stanford / SLAC, ~2019–]



Parallel Applications Workshop Alternatives to MPI+X

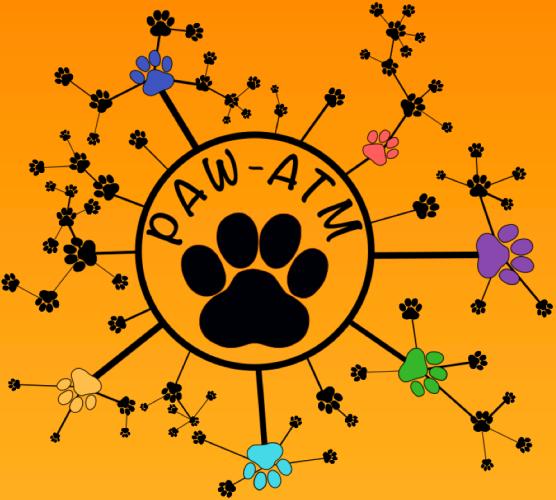
November 17th, 2019

Denver, CO



Analytics- / NumPy-based Approaches

- Hadoop [Apache Software Foundation, ~2006–]
- Spark [UC Berkeley / ASF, ~2009–]
- Flink [Apache Flink Committee, ~2014–]
- Dask [Dask Dev. Community / NumFOCUS, ~2015–]
- Arkouda [DOD, ~2019–]
- Legate [NVIDIA, ~2019–]
- ...



Parallel Applications Workshop
Alternatives to MPI+X

November 17th, 2019

Denver, CO

Which Alternatives to Consider?

Regent

Chapel

XcalableMP

X10

Fortran 2018

UPC

Legate

Dask

Arkouda

PaRSEC

COMPSS

Pygion

PyCOMPSS

Hadoop

Flink

Spark

Charm++

Coarray C++

UPC++

Global Arrays

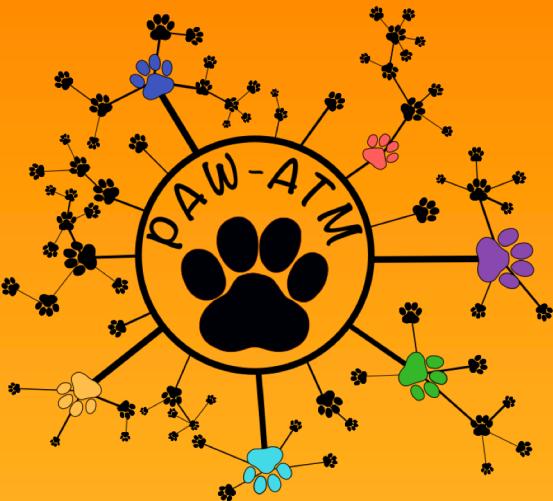
PCJ

HPX

Legion

GASNet-EX

OpenSHMEM



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO

Which Alternatives to Consider? And Why?

Regent

Chapel

XcalableMP

X10

Language vs. Language Extension vs. Library?

Fortran 2018

UPC

Compiled vs. Interpreted?

Legate

Dask

Arkouda

Type(s) of parallelism: SPMD vs. data- vs. task-parallel?

PARSEC

COIMPSS

PvISION

PVCOIMPSS

Explicit vs. Implicit Communication, Synchronization, Parallelism?

Hadoop

Flink

Spark

Support for distributed data structures? Local- or global-view?

Charm++

Coarray C++

UPC++

Global Arrays

PCJ

Compatible / Interoperable with Language of Choice?

HPX

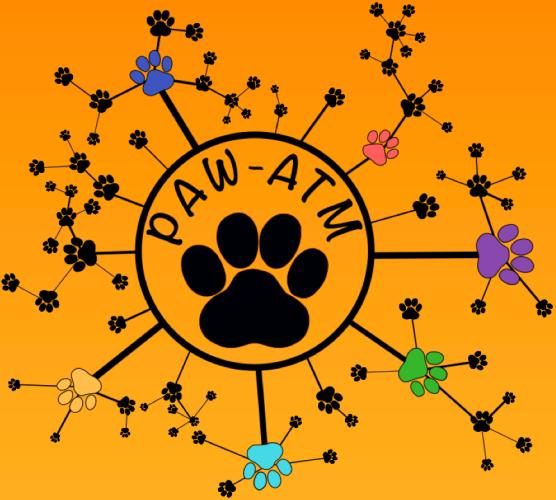
Legion

Portable? GPU Support?

GASNet-EX

OpenSHMEM

...?



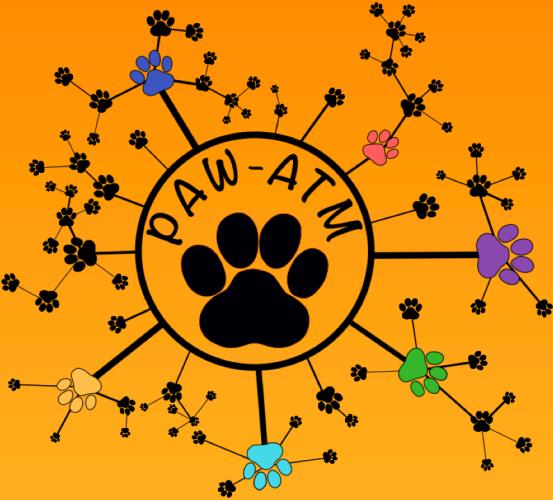
Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO

Program

- 9:00: Welcome and Introduction
- 9:10: **Keynote:** *Computer Science Challenges to Imaging the Universe with the SKA Radio-Telescope*
Peter J. Braam, *University of Oxford*
- 10:00: **Break (catered)**
- 10:30: **Session 1:** Scientific Simulations
- 11:30: **Session 2:** Applying MPI+X Alternatives: from Shallow Water to Deep Learning
- 12:30: **Lunch (on your own)**
- 2:00: **Session 3:** Communication Libraries and APIs
- 3:00: **Break (catered)**
- 3:30: **Session 4:** Integrating Python into HPC
- 4:10: **Panel Discussion:** Applications in Alternatives to MPI+X
- 5:30: **Wrap-up**



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO

Program

9:00: Welcome and Introduction

9:10: **Keynote:** *Computer Science Challenges to Imaging
the Universe with the SKA Radio-Telescope*

Peter J. Braam, *University of Oxford*

Regent

Legion Kokkos

Chapel

UPC++

PaRSEC Charm++

OpenSHMEM

GASNet

Pygion

Arkouda

10:00: **Break** (catered)

10:30: **Session 1:** Scientific Simulations

11:30: **Session 2:** Applying MPI+X Alternatives: from
Shallow Water to Deep Learning

12:30: **Lunch** (on your own)

2:00: **Session 3:** Communication Libraries and APIs

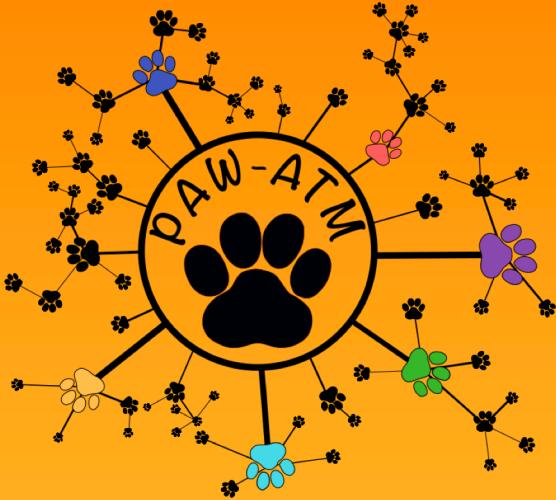
3:00: **Break** (catered)

3:30: **Session 4:** Integrating Python into HPC

4:10: **Panel Discussion:** Applications in Alternatives to MPI+X

5:30: **Wrap-up**

Pygion Legion Regent
Chapel [Py]COMPSS Fortran 2018



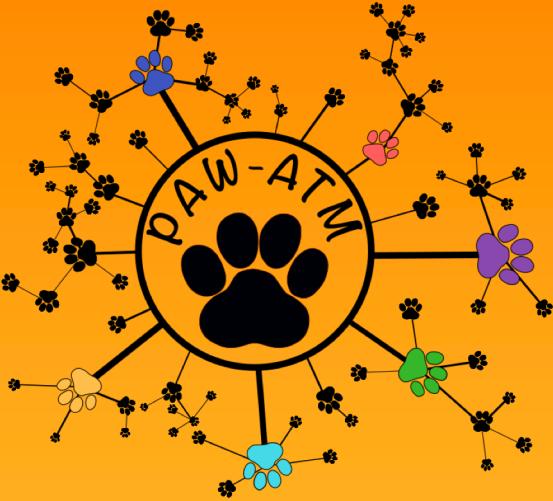
**Parallel Applications Workshop
Alternatives to MPI+X**

November 17th, 2019

Denver, CO

PAW-ATM 2019: Program Committee

- Olivier Aumage - Inria
- Vicenç Beltran - Barcelona Supercomputing Center
- John Biddiscombe - CSCS Swiss National Supercomputing Centre
- Salvatore Filippone - Cranfield University
- Marta G. Gasulla - Barcelona Supercomputing Center
- Magne Haveraaen - University of Bergen
- Costin Iancu - Lawrence Berkeley National Laboratory
- Laxmikant Kale - University of Illinois
- Bill Long - Cray, a Hewlett Packard Enterprise company
- Swaroop S. Pophale - Oak Ridge National Laboratory
- Jason Riedy - Georgia Institute of Technology
- Francesco Rizzi - NexGen Analytics
- Mitsuhsia Sato - RIKEN Advanced Institute for Computational Science
- Elliott Slaughter - SLAC National Accelerator Laboratory



Parallel Applications Workshop

Alternatives to MPI+X

November 17th, 2019

Denver, CO



SC19
Denver, CO | hpc is now.

TCHPC

PAW-ATM 2019: Organization

Workshop Chair

- Karla Morris
- Sandia National Laboratories

Organizing Committee

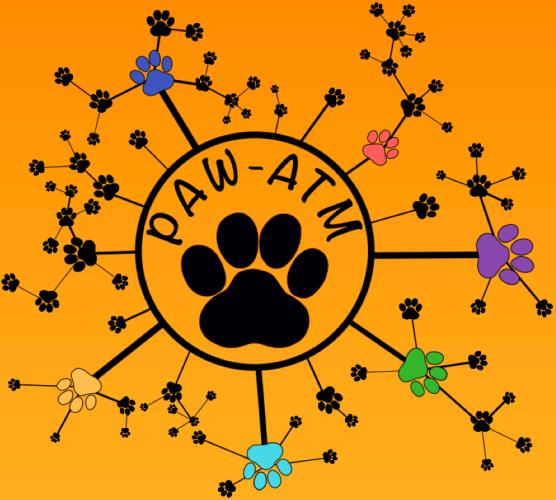
- Rosa M. Badia
- Barcelona Supercomputing Center
- Bradford L. Chamberlain
- Cray, a Hewlett Packard Enterprise company
- Sean Treichler
- NVIDIA

Program Committee Chairs

- Bill Long
- Cray, a Hewlett Packard Enterprise company
- Francesco Rizzi
- NexGen Analytics

Advisory Committee

- Salvatore Filippone
- Cranfield University
- Katherine A. Yellick
- Lawrence Berkeley National Laboratory
- Damian W. I. Rouson
- Sourcery Institute



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

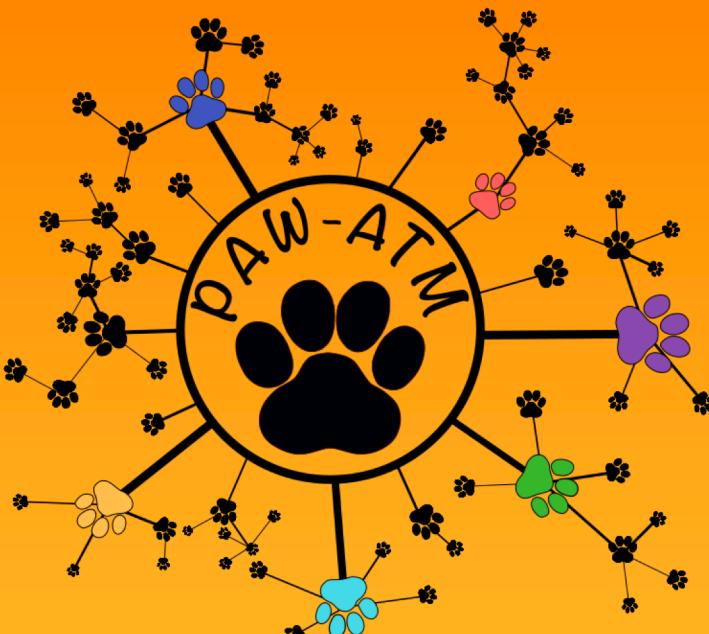
Denver, CO



Please send in your feedback!

After the workshop, *please* send in your feedback:
<https://sc19.supercomputing.org/session/?sess=sess108>

The screenshot shows the SC19 conference website. At the top, there's a navigation bar with links for PROGRAM, EXHIBITS, SCINET, ATTEND, SUBMIT, and a purple "REGISTER →" button. Below the navigation is a decorative banner with abstract geometric shapes in various colors. The main content area has a header "Session" with a back-to-home icon. Below the header, there are links to FULL PROGRAM, CONTRIBUTORS, ORGANIZATIONS, SEARCH PROGRAM, FLAGGED, HAPPENING NOW, and MAPS. The main title of the session is "Parallel Applications Workshop, Alternatives to MPI+X". Underneath the title, it says "Session Chairs: Karla Morris - Sandia National Laboratories, Brad Chamberlain - Cray Inc, Rosa M. Badia - Barcelona Supercomputing Center, Sean Treichler - Nvidia Corporation, Bill Long - Cray Inc, Francesco Rizzi - Independent". To the right of this text is a red-outlined button labeled "give feedback". Further down the page, there are sections for "Event Type: Workshop", "Registration Categories" (with a single red box containing a white letter "W"), "Tags" (with four small colored boxes: cyan, light blue, teal, and light green), "Time: Sunday, 17 November 2019, 9am - 5:30pm", and "Location: 507".



Welcome to PAW-ATM 2019!



In cooperation with:





Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO



Session 1

• Scientific Simulations

10:30 - *Soleil-X: Turbulence, Particles, and Radiation in the Regent Programming Language*

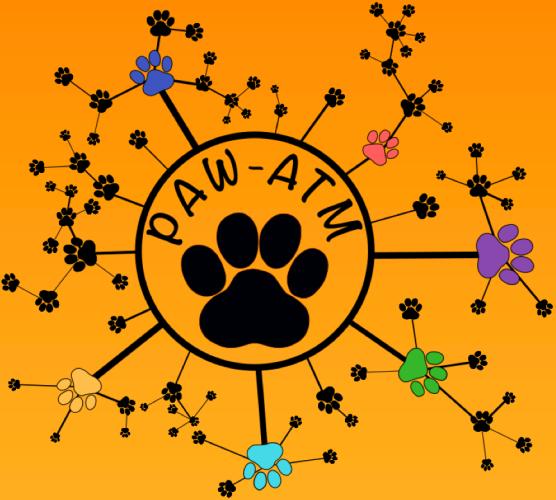
Hilario C. Torres, Manolis Papadakis, Lluis Jofre, Wonchan Lee, Alex Aiken and Gianluca Iaccarino

10:50 - *Exploring the Use of Novel Programming Models in Land Surface Models*

Ethan T. Coon, Wael Elwasif, Himanshu Pillai, Peter E. Thornton and Scott L. Painter

11:10 - *Simulating Ultralight Dark Matter with Chapel: An Experience Report*

Nikhil Padmanabhan, Elliot Ronaghan, Jovana L. Zagorac, and Richard Easter



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO



Session 2

- Applying MPI+X Alternatives: from Shallow Water to Deep Learning

11:30 - *A UPC++ Actor Library and Its Evaluation On a Shallow Water Proxy Application*

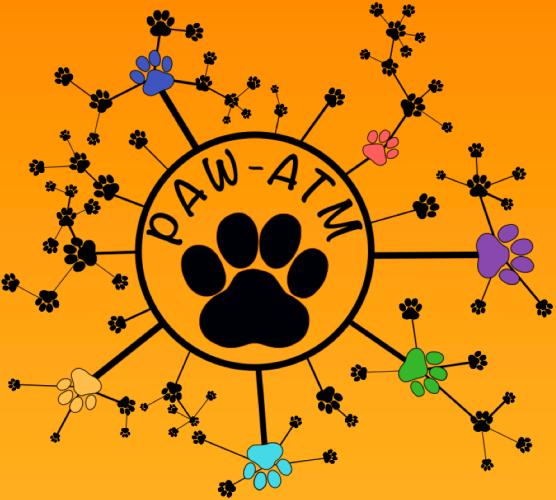
Alexander Pöppl, Scott Baden, and Michael Bader

11:50 - *Evaluation of Programming Models to Address Load Imbalance on Distributed Multi-Core CPUs: A Case Study with Block Low-Rank Factorization*

Yu Pei, George Bosilca, Ichitaro Yamazaki, Akihiro Ida, and Jack Dongarra

12:10 - *Scalable Machine Learning with OpenSHMEM*

Gerard Taylor, James Dinan, Md. Wasi-ur Rahman, and David Ozog



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO

Session 3

- Communication Libraries and APIs

2:00 - *Designing, Implementing, and Evaluating the Upcoming OpenSHMEM Teams API*

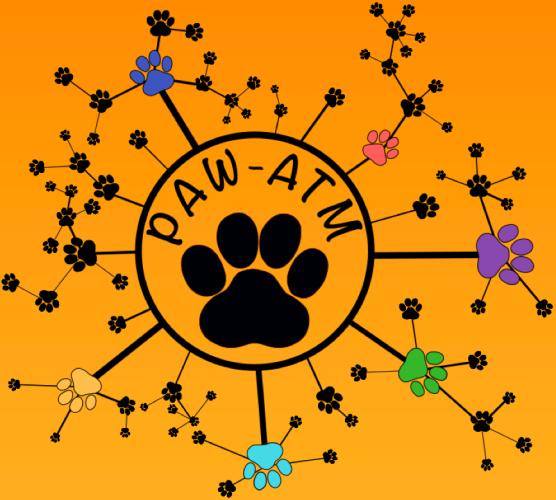
David Ozog, MD Wasi-ur- Rahman, Gerard Taylor, and James Dinan

2:20 - *Enabling Low-Overhead Communication in Multi-threaded OpenSHMEM Applications using Contexts*

Wenbin Lu, Tony Curtis, and Barbara Chapman

2:40 - *Efficient Active Message RMA in GASNet Using a Target-Side Reassembly Protocol*

Paul H. Hargrove and Dan Bonachea



Parallel Applications Workshop
Alternatives to MPI+X

November 17th, 2019

Denver, CO

Session 4

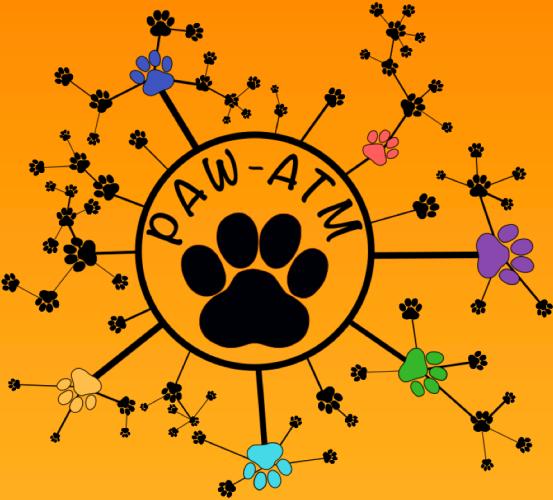
- Integrating Python into HPC

3:30 - *Pygion: Flexible, Scalable Task-Based Parallelism with Python*

Elliott Slaughter and Alex Aiken

3:50 - *Arkouda: NumPy-like arrays at massive scale backed by Chapel*

Michael H. Merrill, William Reus, and Timothy Neumann



Parallel Applications Workshop Alternatives to MPI+X

November 17th, 2019

Denver, CO

Panel Discussion

- Applications in Alternatives to MPI+X

- Panel Moderator:
 - Patrick McCormick
 - Los Alamos National Laboratory

- Panelists:
 - Alex Aiken
 - Stanford University
 - Rosia M. Badia
 - Barcelona Supercomputing Center
 - Michael Ferguson
 - Cray, a Hewlett Packard Enterprise company
 - Damian W. I. Rouson
 - Sourcery Institute
 - Mike Merrill
 - U.S. Department of Defense