Stanford | CONFERENCE 2017

Paul Brest Hall

Day One | TUESDAY, 07 FEBRUARY, 2017

9:00-9:10	Welcome — Steve Jones & Gilad Shainer
9:10-9:55	**Keynote: HPC Meets Big Data: Accelerating Hadoop, Spark, and Memcached** **The Ohio State University** — **DK Panda** This talk will focus on challenges in designing runtime environments for exascale systems with millions of processors and accelerators to support various programming models. We will focus on MPI, PGAS (OpenSHMEM, CAF, UPC and UPC++) and Hybrid MPI+PGAS programming models by taking into account support for multi-core, high-performance networks, accelerators (GPGPUs and Intel MIC), virtualization technologies (KVM, Docker, and Singularity), and energy-awareness. Features and sample performance numbers from the MVAPICH2 libraries will be presented. **DK Panda, Professor, Distinguished Scholar, IEEE Fellow …** DK Panda is a Professor and University Distinguished Scholar of Computer Science and Engineering at the Ohio State University. He has published over 400 papers in the area of high-end computing and networking. The MVAPICH2 libraries with support for MPI and PGAS on IB, Omni-Path, iWARP, RoCE, GPGPUs, Xeon Phis and virtualization (http://mvapich.cse.ohio-state.edu), are currently being used by more than 2,725 organizations worldwide (in 83 countries). More than 407,000 downloads of these libraries have taken place from the project's site. These libraries are empowering several InfiniBand clusters (including the 1st, 13th, 17th and 40th ranked ones) in the TOP500 list. The RDMA packages for Apache Hadoop, Apache Spark and Memcached together with OSU HiBD benchmarks from his group (http://hibd.cse.ohio-state.edu) are also publicly available. These packages are currently being used by more than 205 organizations from 29 countries. More than 19,000 downloads of these packages have taken place from the project's site. High-performance Deep Learning frameworks like Caffe are available from the newly created High-Performance Deep Leaning (http://hidl.cse.ohio-state.edU) project site. He is an IEEE Fellow. More details about Prof. Panda are available at http://www.cse.ohio-state.edu/~panda.
9:55-10:15	**Industry Insights: Architecting Flash for Scale and Performance in HPC** Weka.IO — Liran Zvibel High Performance Computing (HPC) is considered the unlimited class of computing where performance is all that matters. Increasingly, business enterprises are looking to apply the technology and techniques from HPC to

	help them solve their complex business challenges. Weka.IO's CTO, Liran Zvibel, will discuss how affordable HPC class storage performance and scale can be achieved using Flash technology and a hardware independent software architecture. **Liran Zvibel, Co-Founder & CTO** As Co-Founder and CTO, Mr. Liran Zvibel guides longer range technical strategies at Weka.IO. Prior to creating the opportunity at WekaIO, he ran engineering at social startup and Fortune 100 organizations including Fusic, where he managed product definition, design and development for a portfolio of rich social media applications. Liran also held principal architectural responsibilities for the hardware platform, clustering infrastructure and overall systems integration for XIV Storage System, acquired by IBM in 2007. Mr. Zvibel holds a BSc.in Mathematics and Computer Science from Tel Aviv University.
10:15-11:00	**Keynote: Singularity: Containers for Science, reproducibility, and HPC** Lawrence Berkeley National Laboratory Greg Kurtzer Explore how Singularity liberates non-privileged users and host resources (such as interconnects, resource managers, file systems, accelerators …) allowing users to take full control to set-up and run in their native environments This talk explores Singularity how it combines software packaging models with minimalistic containers to create very lightweight application bundles which can be simply executed and contained completely within their environment or be used to interact directly with the host file systems at native speeds. A Singularity application bundle can be as simple as containing a single binary application or as complicated as containing an entire workflow and is as flexible as you will need. **Gregory M Kurtzer, Technology Architect and Developer** Gregory M. Kurtzer is currently the IT HPC Systems Architect and Technology Developer at Lawrence Berkeley National Laboratory. His specialties include Linux (environment, services and deep system internals), open source and development (Perl, C, SQL, PHP, HTML, etc.); HPC applications, administration, automation and provisioning of large scale system architectures. Along with his solid reputation for sparking new trends, the pioneering open source entrepreneur has created, founded, built and contributed to communities with install counts in the millions of users, and numerous breakthrough projects including CentOS Linux, Caos Linux, Perceus, Warewulf and most recently Singularity.
11:00-12:00	**Tutorial: 0 to 60 w/Intel® HPC Orchestrator ~ from Bare Metal to Production HPC** Intel Corporation David Lombard

	Stanford High Performance Computing Center Steve Jones Simplify HPC stack integration, validation and maintenance. This live demonstration will feature the installation of Intel® HPC Orchestrator on a four-node Intel® Xeon Phi™ cluster. This session will highlight the components and feature the ease of installation and include discussion on the benefits of Intel® HPC Orchestrator for clusters ranging from four to tens of thousands of nodes. **Steve Jones, Director** Steve Jones currently runs the High Performance Computing Center at Stanford University, supporting sponsored research for The Department of Energy Advanced Simulation and Computing Program (ASC) Predictive Science Academic Alliance Program (PSAAP). The HPC Center also supports the computational needs of sponsored research for the National Science Foundation (NSF), with a recent award under the American Recovery and Reinvestment Act (ARRA) for the Acquisition of a Hybrid CPU/GPU and Visualization Cluster for Multidisciplinary Studies in Transport Physics with Uncertainty Quantification. Jones is the chair of the annual Stanford High Performance Computing Conference, has designed and administered numerous Top 500 Supercomputers, and speaks regularly about the management of High Performance Computing Clusters. More information can be found at http://hpcc.stanford.edu **David N. Lombard, Senior Principal Engineer** David N. Lombard is a Senior Principal Engineer in Extreme Scale Computing at Intel Corporation; he is also the Chief Architect for HPC Orchestrator. David joined Intel Corporation in 2003 to work on HPC system software development, where he has been involved in all aspects of HPC system design, from their unique BIOS and low-level system management requirements, to scalable system management and control. David began working in HPC on a Cray-1 vector processor in 1985 while at Grumman, moving to parallel computing on an X-MP and then Y-MP. During the 1990's, he led MSC/NASTRAN porting operations at The MacNeal-Schwendler Corporation and became Director of R&D for MSC.Software's HPC Systems Division in 2001.
12:00-13:00	LUNCH
13:00-13:30	**Industry Insights: A Fresh Look at High Performance Computing** Huawei Enterprise Francis Lam High performance computing is rapidly finding new uses in many applications and businesses, enabling the creation of disruptive products and services. Huawei, a global leader in information and communication technologies, brings a broad spectrum of innovative solutions to HPC. This talk examines Huawei's world class HPC solutions and explores creative new ways to solve HPC problems.

	Francis Lam, Director of Product Management Francis brings extensive HPC experience specializing in server systems design and product management. Before joining Huawei Enterprise USA as Director of Product Management, Francis served in Huawei US R&D Center since 2011 as a Server Architect. Prior to joining Huawei, Francis spent 10+ years with multiple global IT system vendors in Silicon Valley.
13:30-14:00	**Best Practices: The Era of Self-Tuning Servers** DatArcs Tomer Morad Servers today have hundreds of knobs that can be tuned for performance and energy efficiency. While some of these knobs can have a dramatic effect on these metrics, manually tuning them is a tedious task. It is very labor intensive, it requires a lot of expertise, and the tuned settings are only relevant for the hardware and software that were used in the tuning process. In addition to that, manual tuning can't take advantage of application phases that may each require different settings. In this presentation, we will talk about the concept of dynamic tuning and its advantages. We will also demo how to improve performance using manual tuning as well as dynamic tuning using DatArcs Optimizer. **Tomer Morad, CEO** Dr. Tomer Morad is a "Startup Postdoc" at the Runway Program of the Jacobs Technion-Cornell Institute of Cornell Tech in NYC. In this context, Tomer co-founded DatArcs, which provides software for boosting performance and energy efficiency of servers. Prior to founding DatArcs, Tomer co-founded and served as CEO of transSpot, a provider of digital advertising solutions for the digital signage market. Before that, Tomer served as CTO and Chairman of transSpot, Chief Security Officer at Horizon Semiconductors, and a technical team leader at an intelligence unit in the Israel Defense Forces. Tomer holds a PhD from the Technion – Israel Institute of Technology, which focused on energy-efficient system resource allocation.
14:00-14:45	**Tutorial: Towards Exascale Computing with Fortran 2015** Sourcery Institute Damian Rouson National Center for Atmospheric Research Alessandro Fanfarillo The Fortran 2008 standard empowers developers to write parallel programs without the source code directly referencing procedures not defined in the Fortran standard (e.g., MPI) and without embedding compiler directives (e.g., OpenMP). The standard accomplishes this by defining a parallel programming model in the form of collective memory allocations in a global address space (coarrays), a straightforward syntax for accessing data in remote processes, and loop-level concurrency. This tutorial will present several features that the draft Fortran 2015 standard [1] introduces to meet challenges that are expected to dominate massively parallel programming in the coming exascale era. The expected exascale challenges include higher hardware- and software-

	failure rates, increasing hardware heterogeneity, a proliferation of execution units, and deeper memory hierarchies. After a high-level overview of the new features and how they address these challenges, we will show performance comparisons between scientific kernels written in Fortran 2008 and 2015. All presented codes will be available as open-source and will compile with the open-source GNU Fortran compiler. Attendees interested in evaluating the codes may compile them inside a Linux virtual machine that is also open-source and available for download [2]. [1] http://j3-fortran.org/doc/meeting/211/16-007r2.pdf [2] http://www.sourceryinstitute.org/store **Damian Rouson, President** Damian Rouson is the President of Sourcery Institute, a California public-benefit nonprofit corporation that operates as a network of independent professionals engaged in computational science research, education, and advisory services. Dr. Rouson has extensive experience in software design and development for multi-physics modeling, including classical, quantum, and magnetohydrodynamic turbulence and multiphase flows. He co-authored the textbook Scientific Software Design: The Object-Oriented Way (Cambridge University Press, 2011) and has taught related tutorials and courses at conferences, national laboratories, corporations, and universities in the U.S. and Europe. He has been a PI or Co-I on research funded by the National Science Foundation, the Office of Naval Research, and the National Institute of Standards and Technology. He has held faculty and instructional staff positions at universities in the U.S. and in Europe and research and technical leadership positions in government laboratories and in the private sector. He holds a B.S. from Howard University and an M.S. and Ph.D. from Stanford University, all in Mechanical Engineering. **Alessandro Fanfarillo, National Center for Atmospheric Research** Alessandro Fanfarillo received his Ph.D. in Computer Science, Control and Geoinformation at University of Rome "Tor Vergata" in March 2016 and joins the National Center for Atmospheric Research in Boulder, Colorado, USA, in April 2016 as a postdoctoral fellow. He is interested in parallel computing, more specifically his research focuses on how to exploit heterogeneous architectures CPU+Accelerators and Partitioned Global Address Space (PGAS) languages (in particular coarray Fortran) for scientific purposes. He is also the lead developer of OpenCoarrays, the open-source library that implements the coarray support in the GNU Fortran compiler. He also holds B.S. and M.S. degrees in Computer Engineering, both from University of Rome "Tor Vergata".
14:45-15:00	BREAK
15:00-15:30	**Best Practices: Large Scale Multiphysics** Cascade Technologies Frank Ham

	A spin-off of the Center for Turbulence Research at Stanford University, Cascade Technologies grew out of a need to bridge between fundamental research from institutions like Stanford University and its application in industries. In a continual push to improve the operability and performance of combustion devices, high-fidelity simulation methods for turbulent combustion are emerging as critical elements in the design process. Multiphysics based methodologies can accurately predict mixing, study flame structure and stability, and even predict product and pollutant concentrations at design and off-design conditions. **Frank Ham, President & CEO** Frank Ham is Cascade's president and CEO. He has over 15 years of experience in the design and development of massively parallel CFD algorithms and software, and leads Cascade's talented and growing team of developers in the production and deployment of the CTI solvers. Before joining Cascade on a full time basis, Frank was a Senior Research Scientist and Chief Software Architect at Stanford University's Predictive Science Academic Alliance Center. Frank received his PhD from the University of Waterloo, Canada, with a thesis on Large Eddy Simulation for complex flows with heat transfer.
15:30-16:00	**Industry Insights: Hot Technology Topics in 2017** OrionX.net Shahin Khan From BitCoins and AltCoins to Design Thinking, Autonomous tech and the changing nature of jobs, IoT and cyber risk, and the impact of application architecture on cloud computing, we'll touch on some of the hottest technologies in 2017 that are changing the world and how HPC will be the engine that drives it. **Shahin Khan, Founder** Founder of StartupHPC and OrionX,.net Shahin Khan is a respected industry veteran. Shahin was interim VP of Marketing at 3Leaf Systems and ScaleArc; VP of Corporate Development, Intellectual Property, and Marketing at Silicon Graphics; VP/GM of the HPC and Visualization products, Chief Competitive Officer, and VP of Product Marketing & Planning at Sun Microsystems; Chief Marketing Officer at Electronics for Imaging, and CMO at Azul Systems. He lived in the UK to lead European Marketing for FPS Computing before its acquisition by Cray Research and started his career at IBM. Shahin is a graduate of The Hun School of Princeton, and Cornell University with Bachelor of Science and Master of Engineering degrees.
16:00-17:00	**Panel: The Exascale Endeavor** Moderator: Gilad Shainer Panel: John Shalf, DK Panda, Frank Ham, Addison Snell **Gilad Shainer, HPC Advisory Council Chairman** Recipient of HPCWire's 2016 Readers Choice Award for Outstanding Leadership in HPC, Gilad Shainer is an HPC evangelist that focuses on high-performance computing, high-speed interconnects, leading-edge technologies and

	performance characterizations. Mr. Shainer holds a M.Sc. degree (2001, Cum Laude) and a B.Sc. degree (1998, Cum Laude) in Electrical Engineering from the Technion Institute of Technology in Israel. He also holds patents in the field of high-speed networking.
17:00	Recap & Day 2 Preview — Steve Jones & Gilad Shainer
17:15-19:00	**Twilight Tutorial: Machine Learning Bootcamp** HPC Advisory Council — Scot Schultz NVIDIA Deep Learning Institute — Julie Bernauer There has been an incredible amount of advancements for machine learning during these last few years. Problems which used to take research teams months or years can now be implemented by even a novice programmer. We will discuss at an introductory level topics such as: what is machine learning, basics of a neural network, the fundamental tools and software used today. We will also cover some of the popular frameworks being leveraged today, the role of accelerated computing through the use of GPUs and intelligent interconnect devices that leverage RDMA, GPUDirect, in-network processing and a more effective mapping of communication to maximize the performance of DNN training. In this practical session, we will cover basic examples and demonstrations of how to get started with mainstream tools to leverage both CPUs and GPUs, frameworks and libraries being leveraged today such as NVIDIA CUDA, OpenCV, Python, NumPy and others. Participants need no prior experience in programming, but will leave the session with confidence on getting started with sample code to experiment with as you continue your journey with machine learning projects in your area of interest. **Scot Schultz, Director, HPC/Technical Computing; Educational Outreach** Scot Schultz is a HPC technology specialist with broad knowledge in operating systems, high speed interconnects and processor technologies. Joining Mellanox as Director of HPC and Technical Computing, Schultz is 25-year veteran of the computing industry. Schultz spent the past 25 years in various engineering, development and leadership roles; including strategic HPC technology ecosystem enablement. Scot has been instrumental with the growth and development of numerous industry standards-based organizations including OpenPOWER Foundation, OpenFabrics Alliance, HPC Advisory Council and many others.

Day Two | WEDNESDAY, 08 FEBRUARY, 2017

9:00	Welcome — Gilad Shainer & Steve Jones
9:05-9:50	**Visionary Perspectives: Where Computing is Going …** Lawrence Berkeley National Laboratory — John Shalf As Deputy Director of Hardware Technology for the Exascale Computing Project (ECP) Shalf will shed light on the multi-lab collaboration chartered with accelerating delivery of a capable Exascale computing ecosystem to provide breakthrough modeling and simulation solutions to address the most critical challenges in scientific discovery, energy assurance, economic competitiveness, and national security. This role goes far beyond the limited scope of a physical computing system. ECP's work encompasses the development of an entire exascale ecosystem: applications, system software, hardware technologies and architectures, along with critical workforce development. **John Shalf, Dept. Head & CTO** John Shalf is CTO for the National Energy Research Supercomputing Center and also Department Head for Computer Science and Data Sciences at Lawrence Berkeley National Laboratory (LBNL). Shalf is a coauthor of over 60 publications in the field of parallel computing software and HPC technology, including three best papers and the widely cited report "The Landscape of Parallel Computing Research: A View from Berkeley" (with David Patterson and others). He also coauthored "ExaScale Software Study: Software Challenges in Extreme Scale Systems," which sets the Defense Advanced Research Project Agency's (DARPA's) information technology research investment strategy for the next decade. He was a member of the Berkeley Lab/NERSC team that won a 2002 R&D 100 Award for the RAGE robot. Before joining Berkeley Lab in 2000, Shalf was a research programmer at the National Center for Supercomputing Applications at the University of Illinois and a visiting scientist at the Max-Planck-Institut für Gravitationphysick/Albert Einstein Institute in Potsdam, Germany, where he co-developed the Cactus code framework for computational astrophysics.
9:50-10:20	**Best Practices: Designing HPC & Deep Learning Middleware for Exascale Systems** The Ohio State University — DK Panda This talk will focus on challenges in designing runtime environments for exascale systems with millions of processors and accelerators to support various programming models. We will focus on MPI, PGAS (OpenSHMEM, CAF, UPC and UPC++) and Hybrid MPI+PGAS programming models by taking into account support for multi-core, high-performance networks, accelerators (GPGPUs and Intel MIC), virtualization technologies (KVM, Docker, and

	Singularity), and energy-awareness. Features and sample performance numbers from the MVAPICH2 libraries will be presented.
10:20-10:50	**Best Practices: Multi-Physics Methods, Modeling, Simulation & Analysis** Stanford University, Center for Turbulence Research Mahdi Esmaily The cycle of modeling high impact applications to finding new solutions is completed by the use of high-performance computing (HPC). I this talk, I will discuss two particular applications which have highly benefitted from HPC. The surgical operation performed on single ventricle heart patients has not been modified in last few decades despite a high rate of mortality. Through multiscale simulation of the circulatory system, it is now possible to model this surgery and optimize it using the state of the art optimization techniques. In-silico analysis has allowed us to test new surgical design without posing any risk to patient's life. I will show the outcome of this study, which is a novel surgical option that may revolutionize current clinical practice. The second application that I will discuss in this talk is related to renewable energy. The particle-based solar receivers operate by collecting radiative energy volumetrically through dispersed particles rather than the conventional approach of absorption via a surface. I will discuss our recent work on the investigation of the operating modes of these devices, where we are exploring the interaction of particles with turbulence, solid boundaries and radiation. **Mahdi Esmaily, Postdoctoral Scholar** Mahdi Esmaily-Moghadam received his B.S. and M.Sc. in Mechanical Engineering from the Sharif University of Technology, Tehran, Iran, and his Ph.D. from the University of California, San Diego, working on the development of multiscale methods for optimization of surgical techniques for single ventricle heart patients. He is currently a postdoctoral scholar at the Center for Turbulence Research at Stanford University, studying particle-based solar receivers. He has an interdisciplinary background in areas of computational and cardiovascular mechanics, particle-laden flows, finite-element analysis, shape optimization methods, high-performance computing, and linear algebraic solvers.
10:50-11:20	**Best Practices: State of Linux Containers** Gaikai Inc. Christian Kniep Linux Containers gain more and more momentum in all IT ecosystems. This talk provides an overview about what happened in the container landscape (in particular Docker) during the course of the last year and how it impacts datacenter operations, HPC and High-Performance Big Data. Furthermore Christian will give an update/extend on the ‘things to explore’ list he presented in the last Lugano workshop, applying what he learned and came across during the year 2016. **Christian Kniep, Release Engineer**

	Christian Kniep seeks application in the nebulous world of DevOps. He is currently enjoying exploration of emerging trends and new technologies, often containerizing them first. Rooted in HPC, supporting CAE applications and VR installations, in Germany's automotive industry Kniep became known as the InfiniBand go-to-guy. Kniep pivoted from operating his 4000 node crash-test cluster to R&D to explore Bull's BXI interconnect. While attending a conference and told by a fellow attendee that HPC "cannot learn anything from the emerging Cloud and Big Data companies" Kniep abandoned more than a decade of traditional HPC, boldly leaping into a fascinating new area and now leads the containerization effort underway at Sony Interactive Entertainment subsidiary Gaikai Inc.
11:20-12:00	**Industry Insights: HPC Computing Trends** Intersect360 Research — Addison Snell Intersect360 Research returns with an annual deep dive into the trends, technologies and usage models that will be propelling the HPC community through 2017 and beyond. Emerging areas of focus and opportunities to expand will be explored along with insightful observations needed to support measurably positive decision making within your operations. **Addison Snell, CEO** Addison Snell is the CEO of Intersect360 Research and a veteran of the High Performance Computing industry. He launched the company in 2007 as Tabor Research, a division of Tabor Communications, and served as that company's VP/GM until he and his partner, Christopher Willard, Ph.D., acquired Tabor Research in 2009. During his tenure, Addison has established Intersect360 Research as a premier source of market information, analysis, and consulting. He was named one of 2010's "People to Watch" by HPCwire. Addison was previously an HPC industry analyst for IDC, where he was well-known among industry stakeholders. Prior to IDC, he gained recognition as a marketing leader and spokesperson for SGI's supercomputing products and strategy. Addison holds a master's degree from the Kellogg School of Management at Northwestern University and a bachelor's degree from the University of Pennsylvania.
12:00-13:00	LUNCH
13:00-13:20	**Best Practices: Application Profiling at the HPCAC High Performance Center** HPC Advisory Council — Pak Lui To achieve good scalability performance on the HPC scientific applications typically involves good understanding of the workload though performing profile analysis, and comparing behaviors of using different hardware which pinpoint bottlenecks in different areas of the HPC cluster. In this session, a selection of HPC applications will be shown to demonstrate various methods of profiling and analysis to determine the bottleneck, and the

	effectiveness of the tuning to improve on the application performance from tests conducted at the HPC Advisory Council High Performance Center. **Pak Lui, Application Performance Manager** Pak is the Application Performance Manager for the HPC Advisory Council. He has been involved in demonstrating application performance on various open source and commercial applications. His main responsibilities involve characterizing HPC workloads, analyzing MPI profiles to optimize HPC applications, as well as exploring new technologies, solutions and their effectiveness on real HPC workloads. Pak also works at Mellanox Technologies. His main focus is to optimize HPC applications on products, explore new technologies and solutions and their effect on real workloads. Pak has been working in the HPC industry for over 15 years. Prior to joining Mellanox Technologies, Pak worked as a Cluster Engineer at Penguin Computing. Pak also worked at Sun Microsystems for over 7 years in Sun's High Performance Computing (HPC) group as a Software Engineer. Pak holds a B.Sc. in Computer Systems Engineering and a M.Sc. in Computer Science from Boston University in the United States.
13:20-13:50	**Best Practices: Containerizing Distributed Pipes** Gaikai Inc. Hagen Toennies In this talk we will present how we enable distributed, Unix style programming using Docker and Apache Kafka. We will show how we can take the famous Unix Pipe Pattern and apply it to a Distributed Computing System. We will demonstrate the development of two simple applications with the focus on "Do One Thing and Do It Well." Afterwards we demonstrate how we make these two programs work to together using Apache Kafka. By encapsulating our applications in containers we will also show how that enables us to go from the limited resources of a development machine to cluster of computers in a data center without changing our applications or containers. **Hagen Toennies, Release Engineer** Hagen Toennies joined Gaikai Inc in 2016 where he is focused on resilient Data Pipeline services to deliver the right data to the right place in the right time. For more than five years, from 2011 until joining Gaikai, Hagen gained experience in search engine scaling and applied his expertise in building custom retrieval systems to deliver search applications for Germany's news publishing sites. His work with the German Federal Statistical Office introduced challenges in scaling database systems and maintaining a clean codebase led to the 2005 launch of Podcast Ready Inc. where the startup co-founder gained practical experience in software development and entrepreneurialism. Highly specialized in Information-Retrieval and Web-Technology Hagen holds a Master of Systems Science from the Bauhaus University in Weimar, Germany. In addition to enjoying most of his free time with his family his

	remaining time is spent dedicated to the Gödel-Escher-Bach book, being a (lisp) fan and a Clojure enthusiast - a JVM based (lisp).
13:50-14:20	**Industry Insights: Deep Learning & HPC: New Challenges for Large Scale Computing** NVIDIA — Julie Bernauer In recent years, major breakthroughs were achieved in different fields using deep learning. From image segmentation, speech recognition or self-driving cars, deep learning is everywhere. Performance of image classification, segmentation, localization have reached levels not seen before thanks to GPUs and large scale GPU-based deployments, leading deep learning to be a first class HPC workload. In this talk, after a short introduction to Deep Neural Networks on GPUs, we will present NVIDIA's platform for deep learning and how new advances in hardware and software integrate in large-scale computing environments. **Julie Bernauer, Team Leader, Pursuit Engineering Solution Architects** Julie Bernauer leads a Pursuit Engineering Solutions Architect team for Machine Learning and Deep Learning at NVIDIA Corporation. She joined NVIDIA in 2015 after fifteen years in academia as an expert in machine learning for computational structural biology. She also teaches GPU computing at Stanford University and DL courses for the NVIDIA Deep Learning institute. She obtained her PhD from Université Paris-Sud in Structural Genomics studying Voronoi models for modelling protein complexes. After a post-doc at Stanford University with Pr. Michael Levitt, Nobel Prize in Chemistry 2013, she joined Inria, the French National Institute for Computer Science. While Senior Research Scientist at Inria, Adjunct Associate Professor of Computer Science at École Polytechnique and Visiting Research Scientist at SLAC, her work focused on computational methods for structural bioinformatics, specifically scoring functions for macromolecule docking using machine learning, and statistical potentials for molecular simulations.
14:20-15:00	**Tutorial: In-Network Computing SHARP Technology for MPI Offloads** Mellanox Technologies — Devendar Bureddy Increased system size and a greater reliance on utilizing system parallelism to achieve computational needs, requires innovative system architectures to meet the simulation challenges. As a step towards a new network class of co-processors intelligent network devices, which manipulate data traversing the data-center network, SHARP technology designed to offload collective operation processing to the network. This tutorial will provide an overview of SHARP technology, integration with MPI, SHARP software components and live example of running MPI collectives. **Devendar Bureddy, Staff Engineer**

	Devendar Bureddy is a Staff Engineer at Mellanox Technologies and has been instrumental in building several key technologies like SHARP, HCOLL, etc. Prior to joining Mellanox, he was a software developer at The Ohio State University in network-Based Computing Laboratory led by Dr. D. K. Panda, involved in the design and development of MVAPICH2, an open-source high-performance implementation of MPI over InfiniBand and 10GigE/iWARP. Devendar received his Master's degree in Computer Science and Engineering from the Indian Institute of Technology, Kanpur. His research interests include high speed interconnects, parallel programming models and HPC software.
15:00-15:15	BREAK
15:15-15:45	**HPC Impact: Using HPC in a Cohort Study of the Health Effects of Handgun Ownership in California** Stanford University School of Medicine & Stanford Law School — Yifan Zhang, *David M. Studdert*
15:45 -16:45	**End Note: Computing of the Future** IBM Research Almaden — Jeffrey Welser **Jeff Welser, VP** Dr. Jeff Welser, Vice President of IBM Research – Almaden in San Jose oversees scientists and engineers performing exploratory and applied research at the home of the relational database and world's first hard disk drive. Today the lab specializes in areas including: Watson technologies, storage systems, data management and analytics, nanotechnology, materials science, Web 2.0 technologies and IBM Smarter Planet projects, such as healthcare informatics, water desalination and electric car batteries. Dr. Welser joined IBM Research in 1995, working on novel silicon devices at the T.J. Watson Research Center, and served as adjunct professor at Columbia University, teaching semiconductor device physics at the same time. Dr. Welser received his PhD in Electrical Engineering from Stanford University. Welser holds 21 US Patents and has published over 75 technical papers and presentations. He is a member of the IBM Academy of Technology, an IEEE Fellow, a member of the American Physical Society, and has served on numerous Federal agency and Congressional panels on advanced semiconductor technology.
16:45	Raffle & Wrap Up — Steve Jones & Gilad Shainer