# Saturday February 24th, 2018

	HPCA	CGO		PPoPP	CC
[08:00 - 18:15] <b>Registration</b>					
[08:30 - 10:00] Room: Europa 3	[08:30 - 10:00] Room: Europa 7	[09:15 - 10:00] Room: Europa 2	[09:15 - 10:00] Room: Europa 6	[08:30 - 10:00] Room: Europa 5	[08:30 - 08:45] Room: Europa 1
AACBB: Accelerator Architecture in Computational Biology and Bioinformatics	HIPINEB: High-Performance Interconnection Networks in the Exascale and Big-Data Era	LLVM Performance Workshop	RWDSL'18: 3rd International Workshop on Real World Domain Specific Languages	<u>WPMVP: Workshop on Programming Models for</u> <u>SIMD/Vector Processing</u>	<u>CC: International Conference on Compiler</u> <u>Construction Compiler Construction</u>
Opening Remarks	Opening	How to Evaluate "In-Memory Computing" Performances without Hardware Measurements?	Welcome	Keynote TBA	Opening
Keynote 1: "Accelerating Genome Analysis: A Primer on an Ongoing Journey" Onur Mutlu (ETH, CMU)	Keynote: "The three L's in modern high-performance networking: Low latency, Low cost, Low processing load"		Industrial Experience with the Migration of Legacy Models using a DSL	Vectorization of a spectral finite-element numerical kernel (Application)	[08:45 - 10:00] Room: Europa 1
Exploring Speed/Accuracy Trade-offs					<u>CC Keynote</u>
Accelerating Duplicate Marking In The Cloud					Rethinking Compilers in the Rise of Machine Learning and Al Xipeng Shen (North Carolina State University, USA)
[10:00 - 10:30] Coffee Break with Snack					
[10:30 - 12:10] Room: Europa 3	[10:30 - 12:00] Room: Europa 7	[10:30 - 12:00] Room: Europa 2	[10:30 - 11:50] Room: Europa 6	[10:30 - 12:00] Room: Europa 5	[10:30 - 12:00] Room: Europa 1
AACBB: Accelerator Architecture in Computational Biology and Bioinformatics	HIPINEB Technical Session 1 (research papers)	LLVM Performance Workshop	RWDSL'18: 3rd International Workshop on Real World Domain Specific Languages	WPMVP: Workshop on Programming Models for SIMD/Vector Processing	Session 1: Polyhedral Compilation
Invited Talk: "Next Generation Sequencing: Big Data meets High Performance Computing Architectures" Bertil Schmidt (JGU Mainz)	Analysis and improvement of Valiant routing in low-diameter networks	Optimizing LLVM IR for Guided Vectorization	Saiph: Towards a DSL for High-Performance Computational Fluid Dynamics.	Small SIMD Matrices for CERN High Throughput Computing	Modeling the Conflicting Demands of Parallelism and Temporal/Spatial Locality in Affine Scheduling
GAME: GPU Acceleration of Metagenomics Clustering	Node-type-based load-balancing routing for Parallel Generalized Fat-Trees	Efficient use of memory by reducing size of AST dumps in cross file analysis by clang static analyzer	CFDlang: High-level code generation for high-order methods in fluid dynamics	SIMDization of Small Tensor Multiplication Kernels for Wide SIMD Vector Processors	A Polyhedral Compilation Framework for Loops with Dynamic Data-Dependent Bounds
Exact Alignment with FM-index on the Intel Xeon Phi Knights Landing Processor	Analyzing topology parameters for achieving energy-efficient k-ary n-cubes			MIPP: a Portable C++ SIMD Wrapper and its use for Error Correction Coding in 5G Standard	Polyhedral Expression Propagation
Optimizations of Sequence Alignment on FPGA: A Case Study of Extended Sequence Alignment					
[12:00 - 13:30] Lunch					
[13:30 - 15:10] Room: Europa 3	[13:30 - 15:00] Room: Europa 7	[13:30 - 15:00] Room: Europa 2	[13:30 - 14:50] Room: Europa 6	[13:30 - 15:00] Room: Europa 5	[13:30 - 15:00] Room: Europa 1
AACBB: Accelerator Architecture in Computational Biology and Bioinformatics	HIPINEB Technical Session 2 (research papers)	LLVM Performance Workshop	RWDSL'18: 3rd International Workshop on Real World Domain Specific Languages	<u>WPMVP: Workshop on Programming Models for</u> <u>SIMD/Vector Processing</u>	Session 2: Data-Flow and Pointer/Alias Analysis
Keynote 2: "Automata Processor and its Applications in Bioinformatics" Srinivas Aluru (Georgia Tech)	Evaluating Energy Saving Strategies on Torus, K-Ary N-Tree, and Dragonfly	Cache-aware Scheduling and Performance Modeling with LLVM-Polly and Kerncraft	dsmodels: A Little Language for Dynamical Systems	Ikra-Cpp: A C++/CUDA DSL for Object-Oriented Programming with Structure-of-Arrays Layout	Computing Partially Path-Sensitive MFP Solutions in Data Flow Analyses
Streaming Gap-Aware Seed Alignment on the Cache Automaton	VEF3 traces: towards a complete framework for modelling network workloads for exascale systems	Enabling Automatic Partitioning of Data-Parallel Kernels with Polyhedral Compilation	D'Artagnan: An Embedded DSL Framework for Distributed Embedded Systems	Usuba, Optimizing & Trustworthy Bitslicing Compiler	An Efficient Data Structure for Must-Alias Analysis
Processing-in-Storage Architecture for Large-Scale Biological Sequence Alignment	Improving the Efficiency of Future Exascale Systems with rCUDA			A Data Layout Transformation for Vectorizing Compilers	Parallel Sparse Flow-Sensitive Points-to Analysis
The Genomic Benchmark Suite: Characterization and Architecture Implications					
[15:00 - 15:30] Coffee Break with Snack	•	•	•	•	
[15:30 - 17:50] Room: Europa 3	[15:30 - 17:00] Room: Europa 7	[15:30 - 17:00] Room: Europa 2	[15:30 - 17:00] Room: Europa 6	[15:30 - 17:00] Room: Europa 5	[15:30 - 17:00] Room: Europa 1
AACBB: Accelerator Architecture in Computational Biology and Bioinformatics	Panel Session: "Industrial perspective of high-speed communication technology evolution"	LLVM Performance Workshop	<u>RWDSL'18: 3rd International Workshop on</u> <u>Real World Domain Specific Languages</u>	<u>WPMVP: Workshop on Programming Models for</u> <u>SIMD/Vector Processing</u>	Session 3: Code Generation and Optimisation
Invited Talk: "Addressing Computational Burden to Realize Precision Medicine" Can Alkan (Bilkent University)	Industrial perspective of high-speed communication technology evolution moderated by Prof. Young Cho (University of Southern California), Panelists: Eitan Zahavi, Mellanox Technologies, Israel, Ola Torudbakken, Skala Norge AS, Norway, Cyriel Minkenberg, Rockley Photonics Inc., Switzrland	Tensor Comprehensions	Q#: Enabling Scalable Quantum Computing and Development with a High-level DSL	Investigating automatic vectorization for real-time 3D scene understanding	PAYJIT: Space-Optimal JIT Compilation and Its Practical Implementation
Burrows-Wheeler Short Read Aligner on AWS EC2 F1		LLVM Q&A Panel: Questions Welcome	A Task-Based DSL for Microcomputers	Panel Discussion	Finding Missed Compiler Optimizations by Differential Testing
Towards BIMAX: Binary Inclusion-MAXimal parallel implementation for gene expression analysis			Close		Fast and Flexible Instruction Selection with Constraints
Memory: The Dominant Bottleneck in Genomic Workloads					
Gene Sequencing: Where Time Goes					
Are Next-Generation HPC Systems Ready for Population- level Genomics Data Analytics?					
Closing remarks					
[18:15] Departure of the busses to the Heurigen					
[18:30] Heurigen: Toni & Birgit Nigl					

CGO		PPoPP	СС
	[09:15 - 10:00] Room: Europa 6	[08:30 - 10:00] Room: Europa 5	[08:30 - 08:45] Room: Europa 1
	<u>RWDSL'18: 3rd International Workshop on</u> <u>Real World Domain Specific Languages</u>	WPMVP: Workshop on Programming Models for SIMD/Vector Processing	<u>CC: International Conference on Compiler</u> <u>Construction Compiler Construction</u>
<u>omputing"</u> re Measurements?	Welcome	Keynote TBA	Opening
	Industrial Experience with the Migration of Legacy Models using a DSL	Vectorization of a spectral finite-element numerical kernel (Application)	[08:45 - 10:00] Room: Europa 1
			<u>CC Keynote</u>
			<b>Rethinking Compilers in the Rise of Machine Learning and Al</b> Xipeng Shen (North Carolina State University, USA)

2	[10:30 - 11:50] Room: Europa 6	[10:30 - 12:00] Room: Europa 5	[10:30 - 12:00] Room: Europa 1
	<u>RWDSL'18: 3rd International Workshop on</u> <u>Real World Domain Specific Languages</u>	WPMVP: Workshop on Programming Models for SIMD/Vector Processing	Session 1: Polyhedral Compilation
d Vectorization	Saiph: Towards a DSL for High-Performance Computational Fluid Dynamics.	Small SIMD Matrices for CERN High Throughput Computing	Modeling the Conflicting Demands of Parallelism and Temporal/Spatial Locality in Affine Scheduling
<u>ducing size of AST</u> <u>clang static analyzer</u>	CFDlang: High-level code generation for high-order methods in fluid dynamics	SIMDization of Small Tensor Multiplication Kernels for Wide SIMD Vector Processors	A Polyhedral Compilation Framework for Loops with Dynamic Data-Dependent Bounds
		MIPP: a Portable C++ SIMD Wrapper and its use for Error Correction Coding in 5G Standard	Polyhedral Expression Propagation

	[15:30 - 17:00] Room: Europa 6	[15:30 - 17:00] Room: Europa 5	[15:30 - 17:00] Room: Europa 1
	<u>RWDSL'18: 3rd International Workshop on</u> <u>Real World Domain Specific Languages</u>	WPMVP: Workshop on Programming Models for SIMD/Vector Processing	Session 3: Code Generation and Optimisation
	Q#: Enabling Scalable Quantum Computing and Development with a High-level DSL	Investigating automatic vectorization for real-time 3D scene understanding	PAYJIT: Space-Optimal JIT Compilation and Its Practical Implementation
elcome	A Task-Based DSL for Microcomputers	Panel Discussion	Finding Missed Compiler Optimizations by Differential Testing
	Close		Fast and Flexible Instruction Selection with Constraints

## Sunday February 25th, 2018

	HPCA		CGO				PPoPP				СС
[08:00 - 18:30] <b>Registration</b>			•	-							·
[08:30 - 10:00] Room: Europa 5	[08:30 - 10:0	0] Room: Europa 7	[08:30 - 10:00] Room: Pacific 3	[08:30 - 10:00] Room: Europa 3	[08:30 - 10:00] Room	п: Еигора 2	[08:30 - 10:00] Roo	m: Pacific 1	[08:30 - 10:00] Room: Pa	ific 2 [08:30 - 10:00] Room: Europa 6	[08:45 - 10:00] Room: Europa
<u>WP3: Second Workshop on Pion</u> Processor Paradigms	eering <u>Accelerating</u> <u>Memcached</u>	<u>Big Data Processing with Hadoop, Spark and on Datacenters with Modern Architectures</u>	<u>Tutorial: Improving security</u> with reversibility and session types	PMAM: Workshop on Programming Models and Applications for Multicores and Manycores	GPGPU: Workshop o Processing Using GP	on General Purpose PU	<u>An Introduction to</u> <u>Blocks (Intel® TBB</u> <u>Heterogeneous Pro</u>	Intel® Threading Building ) and its Support for ogramming	<u>Productive parallel</u> <u>programming on FPGA v</u> <u>high-level synthesis</u>	ith <u>Debugging and Profiling Task</u> Parallel Programs with TASKPROF	<u>CC Keynote</u>
Welcome and Introduction Prad	lip Bose Session 1		Session 1	Opening Remarks	Welcome: The Orgar	nizers	Session 1		Session 1	Session 1	Compiler and Language Design for Quantum Computing Bettina Heim (Microsoft Research, USA)
<b>Keynote I: TBD</b> Mikko H. Lipasti (MICRO 2017 Te Award, University of Wisconsin -	est of Time - Madison)			Keynote: "Building the next Generation of MapReduce Programming Models over MPI to Fill the Gaps between Data Analytics and Supercomputers"	Keynote 1: "Initial St GPU a First-Class Cor Sharing and Resourc Jun Yang (William Ke Professor of Electric Engineering, University	mputing Resource: ce Management" epler Whiteford cal and Computer					
[09:40 - 10:00] Room: Europa 5					[09:30 - 10:00] Room	п: Еигора 2					
WP3: Retrospective Survey I					GPGPU Session 1: Pe Structures	ersistent Data					
On the Evaluation of Computer	Architectures				A Case For Persist Ba	arriers in GPUs					
[10:00 - 10:30] <b>Coffee Break wi</b>	ith Snack										
[10:30 - 11:20] Room: Europa 5		0] Room: Europa 7	[10:30 - 12:00] Room: Pacific 3	[10:30 - 12:00] Room: Europa 3	[10:30 - 12:00] Room	n: Еигора 2	[10:30 - 12:00] Roo	m: Pacific 1	[10:30 - 12:00] Room: Pa	ific 2 [10:30 - 12:00] Room: Europa 6	[10:30 - 12:00] Room: Europa
WP3: Invited Talk	Accelerating	Big Data Processing with Hadoop, <u>Spark and</u> on Datacenters with Modern Architectures	Tutorial: Improving security with reversibility and session types	PMAM Session 1: GPU and Accelerator	GPGPU Session 2: Applications/Framew		An Introduction to	Intel® Threading Building ) and its Support for	Productive parallel programming on FPGA v high-level synthesis	Debugging and Profiling Task	Session 4: Compilation for Specialised Domains
40 years since dusk: will hardwa finally make our systems more c Lluis Vilanova (Technion)	re capabilities capable? Session 2		Session 2	Extending ILUPACK with a Task-Parallel Version of BiCG for Dual-GPU Servers	Overcoming the Diff CGH Generation on r	ficulty of Large-scale multi-GPU Cluster	Session 2		Session 2	Session 2	Compiling for Concise Code and Efficient I/O
[11:20 - 12:00] Room: Europa 5				Reduction to Band Form for the Singular Value Decomposition on Graphics Accelerators	Transparent Avoidar Transfer on GPU-ena	nce of Redundant Data abled Apache Spark					Termination Checking and Task Decomposition for Task Based Intermittent Program
WP3: New/Exploratory paradign	ms			Combining PREM compilation and ILP scheduling for high-performance and predictable MPSoC execution	GPU-based Accelera Tissue-Scale Cardiac	tion of Detailed Simulations					A Session Type Provider: Compile-Time API Generation of Distributed Protocols with Refinements in F#
A Multi-component Branch Pred for Low Resource Budget Proce	dictor Design essors										
FFT implementation using mono set computer architecture	o-instruction										
[12:00 - 13:30] Lunch											
[13:20 - 14:20] Room: Europa 5	[13:30 - 15:00] Room: Europ	a 7 [13:30 - 15:00] Room: Pacific 2	[13:30 - 15:00] Room: Pacific 3	[13:30 - 15:00] Room: Europa 3		[13:30 - 14:30] Room: E	Europa 2	[13:30 - 15:00] Room: Pacific	: 1 [13:30 - 1	:00] Room: Europa 6	[13:30 - 15:00] Room: Europa
<u>WP3: Second Workshop on</u> <u>Pioneering Processor</u> <u>Paradigms</u>	<u>PULP: An open hardware</u> <u>platform, the story so far</u>	Turning HPC clusters into High Performance & High Throughput facilities by using remote GPU virtualization	<u>Tutorial: Improving security</u> with reversibility and session types	PMAM Session 2: Fine-grain Parallelism		<u>GPGPU: Workshop on C</u> <u>Processing Using GPU</u>	<u>General Purpose</u>	An Introduction to Intel® Th Building Blocks (Intel® TBB) Support for Heterogeneous	and its High Peri	ormance Distributed Deep Learning: A s Guide	Session 5: Code Translation and Transformation
<b>Keynote II: TBD</b> TBD	PULP concept and goals	[Session 1.1] Presentation of remote GPU virtualization techniques and rCUDA features (50 minutes)	Session 3	Fast and Accurate Performance Analysis of Synchronizat		Keynote 2: "Generating GPU Code using Rewrit Christophe Dubach (Un Edinburgh)	g High Performance te Rules with Lift" hiversity of	Session 3	Session 1		Tail Call Elimination and Data Representation for Functiona Languages on the Java Virtua Machine
[14:20 - 15:00] Room: Europa 5	State of the art of open source hardware design	[Session 1.2] Practical demonstration about how to install and use rCUDA (40 minutes)		Supporting Fine-grained Dataflow Parallelism in Big Dat	a Systems						CAnDL: A Domain Specific Language for Compiler Analysis
WP3: Restrospective Survey II	Summary of PULP systems: PULP, PULPino, PULPissimo			Intra-Task Parallelism in Automotive Real-Time Systems							Semantic Reasoning about t Sea of Nodes
This Architecture Tastes Like Microarchitecture	PULP cores: OR10N, RI5CY, Zero-riscy, Ariane										
Project CrayOn: Back to the future for a more General- Purpose GPU?											

	PPoPP			СС
				·
0] Room: Europa 2	[08:30 - 10:00] Room: Pacific 1	[08:30 - 10:00] Room: Pacific 2	[08:30 - 10:00] Room: Europa 6	[08:45 - 10:00] Room: Europa 1
<u>kshop on General Purpose</u> Jsing GPU	An Introduction to Intel® Threading Building Blocks (Intel® TBB) and its Support for Heterogeneous Programming	Productive parallel programming on FPGA with high-level synthesis	<u>Debugging and Profiling Task</u> Parallel Programs with TASKPROF	<u>CC Keynote</u>
ne Organizers	Session 1	Session 1	Session 1	Compiler and Language Design for Quantum Computing Bettina Heim (Microsoft Research, USA)
Initial Steps toward Making Class Computing Resource: Resource Management" illiam Kepler Whiteford Electrical and Computer , University of Pittsburgh)				
0] Room: Europa 2				
ion 1: Persistent Data				
ersist Barriers in GPUs				
				1
0] Room: Europa 2	[10:30 - 12:00] Room: Pacific 1	[10:30 - 12:00] Room: Pacific 2	[10:30 - 12:00] Room: Europa 6	[10:30 - 12:00] Room: Europa <sup>-</sup>
on 2: /Frameworks	An Introduction to Intel® Threading Building Blocks (Intel® TBB) and its Support for Heterogeneous Programming	<u>Productive parallel</u> programming on FPGA with high-level synthesis	<u>Debugging and Profiling Task</u> <u>Parallel Programs with</u> <u>TASKPROF</u>	Session 4: Compilation for Specialised Domains
the Difficulty of Large-scale tion on multi-GPU Cluster	Session 2	Session 2	Session 2	Compiling for Concise Code and Efficient I/O
Avoidance of Redundant Data				Termination Checking and

[15:30 - 15:50] Room: Europa 5	[15:30 - 17:30] Room: Europa 7	[15:30 - 17:00] Room: Pacific 2	[15:30 - 17:00] Room: Pacific 3	[15:30 - 17:00] Room: Europa 3	[15:30 - 16:30] Room: Europa 2	[15:30 - 17:00] Room: Pacific 1	[15:30 - 17:00] Room: Europa 6	[15:30 - 17:00] Room: Europa 1
WP3: Restrospective Survey III	PULP: An open hardware platform, the story so far	<u>Turning HPC clusters into High Performance &amp; High</u> <u>Throughput facilities by using remote GPU virtualization</u>	<u>Tutorial: Improving security</u> with reversibility and session types	PMAM Session 3: Cache and Pipeline	GPGPU Session 3: Concurrent Kernels	An Introduction to Intel® Threading Building Blocks (Intel® TBB) and its Support for Heterogeneous Programming	High Performance Distributed Deep Learning: A Beginner's Guide	Session 6: Compile- and Run- Time Analysis
45-year CPU evolution: one law and two equations	Advanced PULP silicon implementations	[Session 2] Guided exercises so that the audience uses rCUDA in a cluster located at Technical University of Valencia, Spain	Session 4	Understanding Parallelization Tradeoffs for Linear Pipelines	MaxPair: Enhance OpenCL Concurrent Kernel Execution by Weighted Maximum Matching	Session 4	Session 2	Towards a Compiler Analysis for Parallel Algorithmic Skeletons
[15:30 - 15:50] Room: Europa 5	Acceleration for PULP systems, examples from cryptography and neural networks	Time for attendees to freely exercise with rCUDA in the remote cluster (a set of exercises is proposed)		An Evaluation of Vectorization and Cache Reuse Tradeoffs on Modern CPUs				Generalized Profile-Guided Iterator Recognition
WP3: Panel Session	PULP Programming			VAIL: A Victim-Aware Cache Policy for Improving Lifetime of Hybrid Memory				Efficient Dynamic Analysis for Node.js
<b>Panel TBD</b> Invited Pioneers and speakers plus the retrospective paper authors				[17:00 - 17:05] Room: Europa 3				
[15:30 - 15:50] Room: Europa 5 WP3: Recap/discussion; clossing remarks, action items				Closing Remarks				
Discussion driven by workshop organizers.								
[18:00] HPCA/CGO/PPoPP Welco	8:00] HPCA/CGO/PPoPP Welcome Reception and Poster Session							
[19:45] (Anthony's Bar) Women-	in-Computer-Architecture (WICARCH) get-te	ogether						

## Monday February 26th, 2018

HPCA			CGO	РРоРР
[08:00 - 18:00] <b>Registration</b>				
[08:30 - 08:45] <b>Opening</b>				
[08:45 - 09:55] (Europa 4) HPCA Keynote: What is the role of Architecture and Software Researchers on the Road to C	uantum Supremacy? Margaret Martonosi (Princeton University)			
[09:55 - 10:20] Coffee Break with Snack				
[10:20 - 10:30] Room: Europa 4		[10:20 - 11:45] Room: Europa 2		[10:20 - 11:35] Room: Europa 3
Test of Time Award Session		Session 1: Managed Runti	imes	Session 1: Concurrent Data Structures
HPCA Test of Time Award		SIMD Intrinsics on Manag	ed Language Runtimes	Session chair: Xipeng Shen (North Carolina State University)
[10:30 - 12:00] Room: Europa 4			ework for Efficient and Dynamic Collection	Interval-Based Memory Reclamation
		Selection		Harnessing Epoch-based Reclamation for Efficient Range
Best Paper Session		Analyzing and Optimizing	Task Granularity on the JVM	Queries
Session chair: Josep Torrellas (UIUC)				A Persistent Lock-Free Queue for Non-Volatile Memory
Amdahl's Law in the Datacenter Era: A Market for Fair Processor Allocation				
iNPG: Accelerating Critical Section Access with In-Network Packet Generation for NoC based Many-core	5			
Enabling Efficient Network Service Function Chain Deployment on Heterogeneous Server Platform				
Reducing Data Transfer Energy by Exploiting Similarity within a Data Transaction				
[11:45 - 13:15] Lunch		-		
[13:15 - 14:55] Room: Europa 4	[13:15 - 14:55] Room: Europa 5+6	[13:15 - 14:55] Room: Europa 2		[13:15 - 14:55] Room: Europa 3
Session 2A: Architecture for Neural Network	Session 2B: Cache and Memory	Session 2: Resilience and	Security	Session 2: Compilers and runtime systems
Session chair: Rajeev Balasubramonian (University of Utah)	Session chair: Paul V. Gratz (Texas A&M University)	Automating Efficient Vari Systems	able-Grained Resiliency for Low-Power IoT	Session chair: I-Ting Angelina Lee (Washington University in St. Louis)
Making Memristive Neural Network Accelerators Reliable	A Hybrid Cache Partitioning-Sharing Technique for Commodity Multicores		ndroid Application Repackaging Detection	Juggler: A Dependency-Aware Task Based Execution Framework for GPUs
Towards Efficient Microarchitectural Design for Accelerating Unsupervised GAN-based Deep Learning	SIPT: Speculatively Indexed, Physically Tagged Caches	nAdroid: Statically Detect	ing Ordering Violations in Android Applications	HPVM: Heterogeneous Parallel Virtual Machine
Compressing DMA Engine: Leveraging Activation Sparsity for Training Deep Neural Networks	Domino Temporal Data Prefetcher	SGXElide: Enabling Enclav	e Code Secrecy via Self-Modification	Hierarchical Memory Management for Mutable State
In-situ AI: Towards Autonomous and Incremental Deep Learning for IoT Systems	ProFess: A Probabilistic Hybrid Main Memory Management Framework for			SuperNeurons: Dynamic GPU Memory Management for Training Deep Neural Networks
	High Performance and Fairness			Training Deep Neural Networks
[14:55 - 15:15] Coffee Break with Snack				
[15:15 - 16:55] Room: Europa 4	[15:15 - 16:55] Room: Europa 5+6	[15:15 - 15:25] Room: Europa 2		[15:15 - 16:30] Room: Europa 3
Session 3A: Security	Session 3B: GPU Cache and Memory	Test of Time Award Sessio	n	Session 3: Performance
Session chair: David R. Kaeli (Northeastern University)	Session chair: Bradford M. Beckmann (AMD)	CGO Test of Time Award		Session chair: Milind Chabbi (Baidu Research)
RCoal: Mitigating GPU Timing Attack via Subwarp-based Randomized Coalescing Techniques	Accelerate GPU Concurrent Kernel Execution by Mitigating Memory Pipeline Stalls	[15:25 - 16:55] Room: Europa 2		Bridging the Gap between Deep Learning and Sparse Matrix Format Selection
Are Coherence Protocol States vulnerable to Information Leakage?	LATTE-CC: Latency Tolerance Aware Adaptive Cache Compression Management for Energy Efficient GPUs	Session 3: Best Paper Fina	alists	Optimizing N-Dimensional, Winograd-Based Convolution for Manycore CPUs
Record-Replay Architecture as a General Security Framework	GETM: high-performance GPU transactional memory via eager conflict detection	Poker: Permutation-based Path Encoding	d SIMD Execution of Intensive Tree Search by	vSensor: Leveraging Fixed-Workload Snippets of Programs for Performance Variance Detection
The DRAM Latency PUF: Quickly Evaluating Physical Unclonable Functions by Exploiting the Latency- Reliability Tradeoff in Modern DRAM Devices	Efficient and Fair Multi-programming in GPUs via Effective Bandwidth Management	High Performance Stencil	Code Generation with LIFT	
		Qubit Allocation		
		Dominance-based Duplica	ation Simulation (DBDS): Code Duplication to	
		Enable Compiler Optimiza	ations	
[16:55 - 17:15] <b>Break</b>				
[17:15 - 18:55] Room: Europa 4	[17:15 - 18:55] Room: Europa 5+6	[17:00 - 19:00] Room: Europa 7	[17:15 - 17:45] Room: Europa 3	[17:15 - 17:45] Room: Europa 3
Session 4A: Microarchitecture and Benchmark	Session 4B: Persistent and NVM memory			
Session chair: Benjamin Lee (Duke University)	Session chair: Hai Li (Duke University)	Student Research Competition	CGO & PPoPP Artifact Evaluation	CGO & PPoPP Artifact Evaluation
A Novel Register Renaming Technique for Out-of-Order Processors	Crash Consistency in Encrypted Non-Volatile Main Memory Systems			
Wait of a Decade: Did SPEC CPU 2017 Broaden the Performance Horizon?	Adaptive Memory Fusion: Towards Transparent, Agile Integration of Persistent Memory			
Architectural Support for Task Dependence Management with Flexible Software Scheduling	Efficient Hardware-based Undo+Redo Logging for Persistent Memory Systems		[18:00 - 19:00] Room: Europa 2	[18:00 - 19:00] Room: Europa 3
GDP: Using Dataflow Properties to Accurately Estimate Interference-free Performance at Runtime	Enabling Fine-Grain Restricted Coset Coding Through Word-Level Compression for PCM			
[19:15 - 20:15] Room: Europa 4			CGO Business Meeting	PPoPP Business Meeting
HPCA Business Meeting				

# Tuesday February 27th, 2018

H	PCA	CGO	РРоРР
[08:00 - 17:00] <b>Registration</b>		•	
[08:00 - 09:40] Room: Europa 4	[08:00 - 09:40] Room: Europa 5+6	[08:00 - 09:40] Room: Europa 2	[08:00 - 09:40] Room: Europa 3
Session 5A: GPU	Session 5B: Secure memory	Session 4: Linear Algebra and Vectorization	Session 4: Best Paper Candidates
Session chair: Minsoo Rhu (POSTECH)	Session chair: Rui Hou (Chinese Academy of Science)	The Generalized Matrix Chain Algorithm	Session chair: Idit Keidar (Technion)
Perception-Oriented 3D Rendering Approximation for Modern Graphics Processors	D-ORAM: Path-ORAM Delegation for Low Execution Interference on Cloud Servers with Untrusted Memory	CVR: Efficient Vectorization of SpMV on X86 Processors	Cache-Tries: Concurrent Lock-Free Hash Tries with Constant-Time Operations
Warp Scheduling for Fine-Grained Synchronization	Secure DIMM: Moving ORAM Primitives Closer to Memory	Look-Ahead SLP: Auto-vectorization in the Presence of Commutative Operations	Featherlight On-the-fly False-sharing Detection
WIR: Warp Instruction Reuse to Minimize Repeated Computations in GPUs	Comprehensive VM Protection against Untrusted Hypervisor through Retrofitted AMD Memory Encryption	Conflict-Free Vectorization of Associative Irregular Applications with Recent SIMD Architectural Advances	Register Optimizations for Stencils on GPUs
G-TSC: Timestamp Based Coherence for GPUs	SYNERGY: Rethinking Secure-Memory Design for Error-Correcting Memories		FlashR: Parallelize and Scale R for Machine Learning using SSDs
[09:40 - 10:05] Coffee Break with Snack			
[10:05 - 11:45] Room: Europa 4	[10:05 - 11:45] Room: Europa 5+6	[10:05 - 11:45] Room: Europa 2	[10:05 - 11:45] Room: Europa 3
Session 6A: Novel Architecture	Session 6B: In-Memory Computing	Session 5: Static and Dynamic Analysis	Session 5: Concurrency control and fault tolerance
Session chair: Kei Hiraki (University of Tokyo)	Session chair: Jishen Zhao (UCSD)	Scalable Concurrency Debugging with Distributed Graph Processing	Session chair: Walter Binder (USI)
A Case for Packageless Processors	RC-NVM: Enabling Symmetric Row and Column Memory Accesses for In-Memory Databases	Lightweight Detection of Cache Conflicts	DisCVar: Discovering Critical Variables Using Algorithmic Differentiation for Transient Faults
Extending the Power-Efficiency and Performance of Photonic Interconnects for Heterogeneous Multicores	GraphR: Accelerating Graph Processing Using ReRAM	CUDAAdvisor: LLVM-Based Runtime Profiling for Modern GPUs	Practical Concurrent Traversals in Search Trees
Routerless Networks-on-Chip	GraphP: Reducing Communication of PIM-based Graph Processing with Efficient Data Partition	May-Happen-in-Parallel Analysis with Static Vector Clocks	Communication-Avoiding Parallel Minimum Cuts and Connected Components
HeatWatch: Optimizing 3D NAND Read Operations With Self-Recovery and Temperature Awareness	PM3: Power Modeling and Power Management for Processing-in-Memory		Safe Privatization in Transactional Memory
[11:45 - 13:15] Lunch			
[11:45 - 12:30] (lunch room) <u>Women in Academia and Industry Lunch Session</u>			
[12:35 - 13:10] (Europa 4) <u>Women in Academia and Industry Panel</u>			
[13:15 - 14:25] (Europa 4) CGO Keynote: Biological Computation Sara-Jane Dunn (Microsoft Resear	ch Limited)		
[14:25 - 14:50] Coffee Break with Snack			
[14:50 - 16:30] Room: Europa 4	[14:50 - 16:30] Room: Europa 5+6	[14:50 - 16:30] Room: Europa 2	[14:50 - 16:30] Room: Europa 3
Session 7A: Industry Track	Session 7B: Best of CAL	Session 6: Memory usage Optimisation	Session 6: Models and Libraries
Session chair: Lieven Eeckhout (Ghent University)	Session chair: Dan Sorin (Duke University)	DeLICM: Scalar Dependence Removal at Zero Memory Cost	Session chair: Zoltan Majo (Ergon Informatik AG)
Don't Correct the Tags in a Cache, just Check their Hamming Distance from the Lookup Tag	Resistive Address Decoder	Loop Transformations Leveraging Hardware Prefetching	Making Pull-Based Graph Processing Performant
Reliability-aware Data Placement for Heterogeneous Memory Architecture	Transcending Hardware Limits with Software Out-of-order Processing	Transforming Loop Chains via Macro Dataflow Graphs	An Effective Fusion and Tile Size Model for Optimizing Image Processing Pipelines
SmarCo: An Efficient Many-Core Processor for High-Throughput Applications in Datacenters	Sensing CPU voltage noise through Electromagnetic Emanations	Local Memory-Aware Kernel Perforation	LazyGraph: Lazy Data Coherency for Replicas in Distributed Graph- Parallel Computation
Lost in Abstraction: Pitfalls of Analyzing GPUs at the Intermediate Language Level			PAM: Parallel Augmented Maps
[17:00] Departure of the busses to Palais Liechtenstein			
[18:00] Banquet at Palais Liechtenstein			

## Wednesday February 28th, 2018

	HPCA		
[08:00 - 09:00]	(Europa 4) PPoPP Keynote: From confusion to clarity: hardware concurrency programming model	<b>2008-2018</b> Pet	er Sewell (University of Cambridge)
[09:00 - 09:25]	Coffee Break with Snack		
[09:25 - 11:05]	Room: Europa 4	[09:25 - 11:05]	Room: Europa 5+6
	Session 8A: Industry Track (applications)		Session 8B: Memory
	Session chair: Andrew Putnam (Microsoft)		Session chair: Guangyu Sun (Peking University)
	Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective		ERUCA: Efficient DRAM Resource Utilization and Resource Conflict Avoidance Memory System Parallelism
	Amdahl's Law in Big Data Analytics: Alive and Kicking in TPCx-BB (BigBench)		DUO: Dual Use of On-chip Redundancy for High Reliability
	Memory Hierarchy for Web Search		Memory System Design for Ultra Low Power, Computationally Error Resilient Processor Microarchitectures
	Characterizing Resource Sensitivity of Database Workloads		NACHOS : Software-Driven Hardware-Assisted Memory Disambiguation for Accelerators
[11:05 - 11:20]	Break		
[11:20 - 12:35]	Room: Europa 4	[11:20 - 12:35]	Room: Europa 5+6
	Session 9A: Accelerators		Session 9B: Power
	Session chair: Xuehai Qian (USC)		Session chair: Guru Venkataramani (George Washington University)
	OuterSPACE: An Outer product based SPArse matrix multiplication acCElerator		Power and Energy Characterization of an Open Source 25-core Manycore Proce
	Searching for Potential gRNA Off-Target Sites for CRISPR/Cas9 using Automata Processing across Different Platforms		A Spot Capacity Market to Increase Power Infrastructure Utilization in Multi-Te Data Centers
	Characterizing and Mitigating Output Reporting Bottlenecks in Spatial-Reconfigurable Automata Processing Architectures		GPGPU Power Modeling for Multi-Domain Voltage-Frequency Scaling
[12:35]			

HPCA Closing

CGO	РРоРР
	-

	[09:25 - 11:05]	Room: Europa 2	[09:25 - 11:05]	Room: Europa 3
		Session 7: Program Generation and Synthesis		Session 7: Parallel frameworks and applications
		AutoPA: Automatically Generating Active Driver from Original Passive Driver Code		Session chair: Bernhard Egger (Seoul National University)
e for		Synthesizing an Instruction Selection Rule Library from Semantic Specifications		Efficient Shuffle Management with SCache for DAG Computing Frameworks
		Synthesizing Programs That Expose Performance Bottlenecks		High-Performance Genomics Data Analysis Framework with In- Memory Computing
		Program Generation for Small-Scale Linear Algebra Applications		Griffin: Uniting CPU and GPU in Information Retrieval Systems for Intra-Query Parallelism
				swSpTRSV: a Fast Sparse Triangular Solve with Sparse Level Tile Layout on Sunway Architectures

	[11:20 - 12:35]	Room: Europa 2	[11:20 - 12:10]	Room: Europa 3
		Session 8: Compilation for Specialised Domains		Session 8: Race Detection
		Optimal DNN Primitive Selection with Partitioned Boolean Quadratic Programming		Session chair: Jesper Larsson Träff (TU Wien)
rocessor		Register Allocation for Intel Processor Graphics		VerifiedFT: A Verified, High-Performance Dynamic Race Detector
i-Tenant		A Compiler for Cyber-Physical Digital Microfluidic Biochips		Efficient Parallel Determinacy Race Detection for Two-Dimensional Dags
	[12:35 - 12:45]	Room: Europa 2	[12:10]	
		Best Paper Award Session		
		CGO 2018 Best Paper Award		PPoPP Closing
	[12:45]			
		CGO Closing		

