Scientific publications

date	publication	content	reference
March 2014	PanEuropean Networks "Science and technology" issue 10	paragraph "green computing node" about EUROSERVER	http://www.paneuropeannetworks.com/ST10/
28 March 2014	Press release	Europe invests realising next-generation green computing for micro-servers and scalable compute	Spanish press release
October 2014	EETimes	article on the HiPEAC microservers session and thus EuroSERVER project	http://www.eetimes.com/author.asp?section_id=36&doc_id=1324294
January 2015	EETimes	European server project promotes ARM on FDSOI	http://www.electronics-eetimes.com/en/european-server-project-promotes-arm-on-fdsoi.html?cmp_id=7&news_id=222923411&page=0
summer 2015	HiPEAC newsletter	Socket over RDMA and shared peripherals for ARM microservers	http://www.hipeac.net/assets/public/publications/newsletter/hipeacinfo43.pdf
October 2015	Date'16	EUROSERVER: Share-Anything Scale-Out Micro-Server Design	

Hybrid and Semantic Bit-Field Compression in Caches

Authors: Angelos Arelakis, Fredrik Dahlgren, Per Stenstrom

Abstract: We introduce a compressed cache design that integrates two new ideas: 1) a novel floating-point data compression scheme that compresses the various semantic bit-fields in isolation using statistical compression; and 2) a heuristic-based hybrid compression mechanism that offers compressibility and superior performance gains robustly across diverse data types either manipulated by a single application or by many programs running in a multi-core system.

Link to the publication web-site. Under submission

HyComp: A Hybrid Cache Compression Method for Selection of Data-Type-Specific Compression Methods

Authors: Angelos Arelakis, Fredrik Dahlgren, Per Stenstrom

Abstract: Proposed cache compression schemes make design-time assumptions on value locality to reduce decompression latency. For example, some schemes assume that com- mon values are spatially close whereas other schemes assume that null blocks are common. Most schemes, however, assume that value locality is best exploited by xed-size data types (e.g., 32-bit integers). This assumption falls short when other data types, such as oating-point numbers, are common. This paper makes two contributions. First, HyComp { a hybrid cache compression scheme { selects the best- performing compression scheme, based on heuristics that predict data types. Data types considered are pointers, integers, oating-point numbers and the special (and trivial) case of null blocks. Second, this paper contributes with a compression method that exploits value locality in data types with predened semantical value
elds, for example as in the exponent and the mantissa in oating-point numbers. We show that HyComp, augmented with the proposed oating-point-number compression method, oers superior performance in comparison with prior art.

Link to the publication web-site. In 48th ACM/IEEE Int. Symp. on Microarchitecture (not yet available from ACM digital library)

Adaptive Row Addressing for Cost-Efficient Parallel Memory Protocols in Large-Capacity Low-Latency Memories

Authors: Dmitry Knyaginin, Vassilis Papaefstathiou, Per Stenstrom

Abstract: Modern commercial workloads drive a continuous demand for larger and still low-latency main memories. JEDEC member companies indicate that parallel memory protocols will remain key to such memories, though widening the bus to address larger capacities would cause multiple issues ultimately reducing the speed (the peak data rate) and cost-eciency of the protocols. Thus to stay high speed and cost-ecient, parallel protocols should address larger capacities using the available number of pins, which suggests Multi-Cycle Addressing (MCA). However, additional address-transfer cycles can signicantly worsen performance and energy eciency. This paper contributes with the concept of adaptive rowa ddressing comprising row-address caching to reduce the number of address-transfercycles enhanced by row-address prefetching and an adaptive row-access priority policy to improve state-of-the-art memory schedulers. The paper shows that an MCA protocol using the proposed concept is as ecient in terms of performance and energy as an idealistic (issue-free) protocol of the same speed but with enough pins for single-cycle addressing.

Link to the publication web-site. Under submission

A Methodology for Modeling Dynamic and Static Power Consumption

Authors: Bhavishya Goel, Saly A. McKee

Abstract: As the number of cores on chip grows, system designers and application programmers increasingly need to consider trade-offs between performance and power consumption, whether to stay within a given power budget or to prolong battery lifetime. In particular, application and runtime developers need information about the power consumed by different processor components to understand the energy effects of application scaling or of techniques like Dynamic Voltage and Frequency Scaling (DVFS). We present a methodology to model the static and dynamic power consumption of individual cores and of uncore components. We validate our power model for both single-threaded and parallel benchmarks on an Intel Haswell platform. We show that the power model can be used to gain

insights (some non-intuitive) about energy efficient scaling for parallel workloads. (1) We show that uncore energy consumption can be up to 80% of total energy. Especially, the uncore static energy can be as high as 67% of total energy and is a major source of energy inefficiency. (2) We find that the frequency at which application expends lowest energy cannot be determined statically without taking into account the degree to which an application is memory-bound. (3)We demonstrate that the serial fraction of the program determines the level of concurrency at which least energy is expended. The insights we present are specific to Haswell, but our approach can easily be used to model power consumption for other microarchitectures.
Link to the publication web-site. Under submission

Jericho: Achieving Scalability Through Optimal Data Placement on Multicore Systems

Authors: S. Mavridis, Y. Sfakianakis, A. Papagiannis, M. Marazakis and A. Bilas

Abstract: Achieving high I/O throughput on modern servers presents significant challenges. With increasing core counts, sever memory architectures become less uniform, both in terms of latency as well as bandwidth. In particular, the bandwidth of the interconnect among NUMA nodes is limited compared to local memory bandwidth. Moreover, interconnect congestion and contention introduce additional latency on remote accesses. These challenges severely limit the maximum achievable storage throughput and IOPS rate. Therefore, data and thread placement are critical for data-intensive applications running on NUMA architectures. In this paper we present Jericho, a new I/O stack for the Linux kernel that improves affinity between application threads, kernel threads, and buffers in the storage I/O path. Jericho consists of a NUMA-aware filesystem and a DRAM cache organized in slices mapped to NUMA nodes. The Jericho filesystem implements our task placement policy by dynamically migrating application threads that issue I/Os based on the location of the corresponding I/O buffers. The Jericho DRAM I/O cache, a replacement for the Linux page-cache, splits buffer memory in slices, and uses per-slice kernel I/O threads for I/O request processing. Our evaluation shows that running the FIO microbenchmark ona modern 64-core server with an unmodified Linux kernel results in only 5% of the memory accesses being served by local memory. With Jericho, more than 95% of accesses become local, with a corresponding 2x performance improvement.
Link to the publication web-site: http://www.storageconference.us/2014/

Vanguard: Increasing Server Efficency via Workload Isolation in the Storage I/O Path

Authors: Y. Sfakianakis, S. Mavridis, A. Papagiannis, S. Papageorgiou, M. Fountoulakis, M. Marazakis and A. Bilas

Abstract: Server consolidation via virtualization is an essential technique for improving infrastructure cost in modern datacenters. From the viewpoint of datacenter operators, consolidation offers compelling advantages by reducing the number of physical servers, and reducing operational costs such as energy consumption. However, performance interference between co-located workloads can be crippling. Conservatively, and at significant cost datacenter operators are forced to keep physical servers at low utilization levels typically below 20%), to minimize adverse performance interactions. In this paper, we focus on addressing the issue of performance interference on a virtualized server operating at high utilization levels. In our work, we find that eliminating interference in the I/O path is critical for achieving good performance on consolidated servers. We present Vanguard, a device driver stack that implements a full I/O path in the Linux kernel that provisions competing workloads with dedicated resources. We focus on two key resources: in-memory buffers for the filesystem, and space on SSD devices that serve as a transparent cache for block devices. Our approach effectively mitigates performance interference, for several mixes of transactional, streaming and analytical processing workloads. We find that with our approach a server can run more workloads close to their nominal performance level as compared to the unmodified Linux I/O path, by careful allocation of I/O path resources to each workload. At excessive load levels, i.e. when the aggregate load exceeds the capabilities of the server, our approach can still provide isolated slices of the I/O path for a subset of the co-located workloads yielding at least 50% of their nominal performance. In addition, Vanguard is shown to be 2.5x more efficient in terms of service resource usage for a 4-workload mix, taking into account utilization and power consumption. With an I/O-heavy mix 6-workload mix, Vanguard is 8x more efficient than the unmodified baseline Linux system.

Euroserver

Link to the publication web-site: http://dl.acm.org/citation.cfm?id=2670998

Towards Operating System Support for Remote Memory Usage on ARM Microservers

Authors: J. Velegrakis, M. Marazakis, I. Mavroidis, A. Bilas, J. Goodacre and M. Katevenis

Abstract: This document describes an overview of our work in the research for the Euroserver project that deals with Operating System support for multicore platforms using ARM processors, the management of the unevenly scattered resources and especially remote memory. In particular, we describe two methods of using remote memory without changing the userspace processes' perception of the system: use of remote memory as a main memory extension and as a remote swap device. Data transfers to the remote memory can be done from userspace in a load/store or DMA transfer manner.

Link to the publication site: ACACES 2014

10GigE Virtualized NIC on ARM based FPGAs

Authors: K. Harteros, I. Mavroidis, G. Kalokerinos, V. Papaefstathiou and M. Katevenis

Abstract: The current data centers offer advanced software devices, shared storage and virtualization. In order to deliver such services, high speed networks and virtualized environment is required. Within this context, we have designed a virtualized 10GigE NIC for data centers, shared among low power ARM cores. Our novel design fully utilized the bandwidth of 10Gbit NIC using dedicated route for each destination/source node and by performing only local memory accesses. For the implementation we used Xilinx Zynq 7000 and Virtex7 FPGAs.

Link to the publication site: ACACES 2014

EUROSERVER: Energy Efficiency Node for European Micro-servers

Authors: Y. Durand, P. Carpenter, S. Adami, A. Bilas, D. Dutoit, A. Farcy, G. Gaydadjiev, J. Goodacre, M. Katevenis, M. Marazakis, E. Matus, I. Mavroidis, J. Thomson

Abstract: EUROSERVER is a collaborative project that aims to dramatically improve data centre energy-efficiency, cost, and software efficiency. It is addressing these important challenges through the coordinated applications of several key recent innovations: 64-bit ARM cores, 3D heterogeneous silicon-on-silicon integration, and fully-depleted silicon-on-insulator (FD-SOI) process technology, together with new software techniques for efficient resource management, including resource sharing and workload isolation. We are pioneering a system architecture approach that allows specialized silicon devices to be built even for a low-volume markets where NRE costs are currently prohibitive. The EUROSERVER device will embed multiple silicon "chiplets" on an active silicon interposer. Its system architecture is being driven by requirements from three use cases: data centres and cloud computing, telecom infrastructures, and high-end embedded systems. We will build two fully integrated full-system prototypes, based on a common micro-server board, and targeting embedded servers and entreprise servers.

Link to the publication site: IEEE 2014 / 2014 17th Euromicro Conference on Digital System Design (DSD 2014)

EUROSERVER: Share-Anything Scale-Out Micro-Server Design

Authors: M. Marazakis, J. Goodacre, D. Fuin, P. Carpenter, J. Thomson, E. Matus, A. Bruno, P. Stenström, J. Martin, Y. Durand and I. Dor

Abstract: This paper provides a snapshot summary of the trends in the area of micro-server development and their applications in the broader enterprise and cloud markets. Focusing on the technology aspects, we provide an understanding of these trends and specifically the differentiation and uniqueness of the approach being adopted by the EUROSERVER FP7 project. The unique technical contributions of EUROSERVER range from the fundamental system *compute unit* design architecture, through to the implementation approach both at the *chiplet* nanotechnological integration, and the *everything-close* physical form factor. Furthermore, we offer optimizations at the virtualization layer to

exploit the unique hardware features, and other framework optimizations, including exploiting the hardware capabilities at the run-time system and applications layers.

Link to the publication web-site: http://www.date-conference.com/conference/session/7.2

Centralized Parallel Multi-Path Multi-Slot Allocation Approach for TDM NoCs

Authors: Y. Chen, E. Matus, G.P. Fettweis

Link to the publication site: 29th Annual IEEE Canadian Conference on Electrical and Computer Engineering (CCECE'16)

Trellis-search based Dybamic Multi-Path Connection Allocation for TDM-NoCs

Authors: Y. Chen, E. Matus, G.P. Fettweis

Link to the publication site: IEEE/ACM Great lakes Symposium on VLSI (GLSVLSI'16)

Date	Event	Where	Type	Partner	Who	What
Nov 22, 2013	10th workshop on virtualization in high peformance cloud computing (VHPC SC13)	Denver (USA)	workshop	ARM	JG	Invited talk
Dec 10, 2013	Next Generation Computing Systems: components and architectures for a scalable market	Brussels (BE)	workshop	CEA, ST	DD	Project overview+ 3D technology
Jan 20, 2014	HiPEAC 2014	Vienna (AT)	conference	EuTECH, CHA, CEA, FORTH, BSC	6+	Project presentation poster
Jan 28, 2014	Final workshop of the Next Generation Computing Roadmap Study	Brussels (BE)	workshop	CEA	YD	Technology overview
Feb 9-13, 2014	ISSCC	San Francisco (USA)	conference	ARM	JG	Invited talk
Apr 14, 2014	Semba	Pont-en-Royans (FR)	workshop	CEA	YD	Invited pres. Technology overview
May 15, 2014	HiPEAC Computer system week	Barcelona (ES)	thematic session	FORTH	IM	Organization of "Microserver and Virtualization" thematic session
May 19, 2014	Xen project Hackathon	London (UK)	conference	OnApp	JT	Promoted EuroSERVER, Microserver concepts
May 31 - June 8, 2014	MSST 14	Santa Clara (USA)	conference	FORTH	MM	Paper presentation "Jericho Achieving Scalability Through Optimal Data Placement on Multicore System"
Jun 14-18, 2014	ACM/IEEE ISCA'14	Minneapolis (USA)	conference	CHALMERS	AA & PS	Paper presentation "SC2: a statistical compression cache scheme"

Jun 18-21, 2014	Workshop on Parallel I/O Optimization	Hamburg (DE)	workshop	FORTH	MM	Invited talk "Jericho Achieving Scalability Through Optimal Data Placement on Multicore System"
Jul 7-11, 2014	MPSoC 14	Margaux (FR)	conference	TUD	EM	Invited talk "Datalow models and runtime environment for CRAN applications"
Jul 7-11, 2014	MPSoC 14	Margaux FR)	conference	FORTH	IM	Invited talk "NUMA-like Achitecture for Microservers"
Jul 7-11, 2014	MPSoC 14	Margaux FR)	confererence	ARM	JG	Invited keynote
July 13-19, 2014	ACACES 2014	Fiuggy (IT)	conference	FORTH	KH & JB	Poster presentation "Towards Operating System Support for Remote Memory Usage on ARM Microprocessers"
Aug 17, 2014	DSD 2014 : Special Session on European Projects (EPDSD)	Verona IT)	conference	BSC/CEA/…	YD	Project technical presentation
Oct 1, 2014	CFAED Research festival	Dresden (DE)	event	TUD	EM	Demonstration of TUD's "Dataflow multi-core platform for telecom"
Oct 9, 2014	HiPEAC computer system week	Athens (GR)	confererence	ARM	JG	Talk
Oct 9, 2014	HiPEAC computer system week	Athens (GR)	conference	FORTH	IM	Presentation "Euroserver: Fast, Energ-efficient Microserver communication in the EuroSERVER project"
Oct 9-10, 2014	Vodafone Innovation Days	Dusseldorf (DE)	event	TUD	EM	Demonstration of TUD's "Dataflow multi-core platform for telecom"
Oct 22, 2014	26th International symposium on computer architecture and high performance computing (SBAD-PAD)	Paris (FR)	symposium	ARM	JG	Invited talk

Oct 27-29, 2014	SOC'14	Tampere (FI)	conference	TUD	OA & GF	Presentation "Adaptative Runtime Management of heterogeneous MPSoC: Analysis, acceleratio and silicon prototype"
Nov 3-5, 2014	ACM SoCC'14	Seattle (USA)	conference	FORTH	MM	Poster + paper presentation "Vangard Increasing Server fficiency via Workload isolation in the storage IO Path"
Dec 2014	Electronic Times interview	international	interview	ARM	JG	interview
Dec 2, 2014	IEEE 3DIC Conference	Dublin (IR)	conference	ARM	JG	Keynote
Jan 20, 2015	HiPEAC 2015	Amsterdam (NL)	conference	all	BSC, FORTH, Chalmers, OnAPP + 2 invited speakers (B. Grot, A. Motakis)	Workshop oganization "Greencomputing Node for European Micro-Servers"
Jan 20, 2015	HiPEAC / EuoSERVER workshop	Amsterdam (NL)	conference	FORTH	DP	Paper presentation "Sofware mechanisms for ARM microserver internal/external communication using memory and RDMA"
Jan 20, 2015	HiPEAC / EuroSERVER workshop	Amsterdam (NL)	conference	OnApp	JC	Paper presentation "A Hypervisor architecture for microservers"
Jan 21, 2015	HiPEAC 2015	Amsterdam NL)	conference	all		Poster presentation
Feb 13, 2014	UC Berkeley award ceremony 'IEEE' milestone for the RISC porject	Berkeley USA)	award ceremony	FORTH	MM	Invited talk "Interprocessor communication and its interface to the memory hierarchy"
Mar 9-11, 2015	ARTEMIS/ITEAC co-summit 2015	Berlin (DE)	conference	BSC/CEA/FORTH	YD, ID, PC	Project presentation poster
Mar 19, 2015	UKDF	Manchester (UK)	conference	ARM	JG	Invited talk

Mar 23, 2015	ARM Internal strategy workshop	Cambridge (UK)	workshop	ARM	JG	Workshop participation
Mar 25-25, 2015	NGMN Industry Conference & Exhibition 2015	Frankfurt (DE)	conference	TUD		Paper presentation "Dataflow for CRAN applications"
Apr 10, 2015	Innovate UK EEC SIG AGM	Manchester UK)	confererence	ARM	JG	Invited talk
Apr 15, 2015	Coolchips	Yokohama (JP)	conference	ARM	JG	Keynotes
May 7-8, 2015	IEEE/ACM ANCS'15	Oakland, CA (USA)	conference	FORTH	A. Psathakis, V. Papaefstathiou , N. Chrysos, F. Chaix, E. Vasilakis, D. N. Pnevmatikatos, M. Katevenis	Paper Presentation "A Systematic Evalution of Emerging Mesh-like {CMP} NoCs"
July, 13-17, 2015	MPSoC'15	Ventura (USA)	conference	ARM	J. Goodacre	Keynote presentation "Starting a push towards European Exascale: exaNODE/exaNEST/ecoSCALE and mores"
July 16, 2015	MPSoC Forum 2015	Ventura (USA)	conference	TUD	E.Matus	Presentation: Data Plane Framework for Software Defined Radio Access Networks (https://www.mpsoc-forum.org/previous/2015/program.html)
Sept 3-4, 2015	UK eFutures workshop	Leeds (UK)	workshop	ARM	J. Goodacre	Presentation about shared and positioned EUROSERVER activities and applicability in a workshop (http://efutures.ac.uk/sensors-systems-event-34th-september-2015)

Sept 15, 2015	ICTenergy, Workshop on the Future Energy in ICT Research Agenda	Bristols (UK)	workshop	ARM	J. Goodacre	Presentation about position of EUROSERVER for energy usage in ICT (https://www.evenbrite.co.uk/e/workshop-on-the-future-energy-in-ict-research-agenda-tickets-16972659682)
Sept 21, 2015	HiPEAC Systems Computing Week	Milano(IT)	thematic session	FORTH	N. Chrysos, F. Chaix	organization of a thematic session on "System-level interconnects for exascale-class Datacenters and HPC"
Oct 21, 2015	PhD Dissertation	Chalmers, Goteborg, Sweden	PhD	Chalmers	Angelos Arelakis and Per Stenstrom	PhD dissertation of Angelos Arelakis: Statistical Compression Cache Design
Oct 26-29, 2015	LCN'2015	Clearwater Beach, Florida (USA)	IEEE conference	BSC	Renan Fischer e Silva, Paul Carpenter	Exploring Interconnect Energy Savings Under East-West Traffic Pattern of MapReduce Clusters
Oct, 2015	PhD Dissertation	Dresden (DE)	PhD	TUD	B.Noethen	Investigation of communications mechanisms in heterogenous multi-processor systems
Nov 24, 2015	BigStorage ETN Project meeting		technical session	FORTH	Maniolis Marezakis	technical presentation about 'Storage Solutions' (the EuroServer features shared I/O devices); http://bigstorage-project.eu/
Dec 8, 2015	IEEE Micro Conference	Honolulu, Hawaii	conference	Chalmers	Angelos Arelakis and Per Stenstrom	Conference presentation
Dec, 2015	Morgan Claypool Synthesis Lectures	N/A	Book	Chalmers	A. Arelakis and P. Stenström w/ S. Sardashti and D. Wood (Univ of Wisconsin, USA)	Tutorial textbook on compression in the memory hierarchy

Dec, 2015	Morgan Claypool Synthesis Lectures	N/A	Book	Chalmers	A. Arelakis and P. Stenström w/ S. Sardashti and D. Wood (Univ of Wisconsin, USA)	Tutorial textbook on compression in the memory hierarchy
Dec 5-9, 2015	IEEE Micro Conference	Honolulu, Hawaii	conference	Chalmers	Angelos Arelakis and Per Stenstrom	Two Papers - "HyCOMP: A Hybrid Cache Compression Method for Selection of Data-Type-Specific Compression Methods" - "A scalable Address Extension of DDR4 for Large Capacity Memories"
Jan 18-20, 2016	HiPEAC 2016	Praga (CZ repubic)	conference	All partners		Project Poster presentation+ Invited talk "*Survey about power management in the microserver context*" in the worskhop "*Microserver: energy efficient servers*"
Mar 14-18, 2016	Date 2016	Dresden (DE)	conference	All partners		Paper "Euroserver: Share-Anything Scale-Out Micro-Server Design" submitted + Project presentation booth shared with ExANODE, ExANEST, Eurolab-4-HPC and Ecoscale
May 15-18, 2016	The 29th annual IEEE Canadian conference on electrical and computer engineering (CCECE'16)	Vancouver (CA)	IEEE conference	TUD		Paper "Centralized parallel multi-path multi-slot allocation approach for TDM NoCs"
May 18-20, 2015	IEEE/ACM Great Lakes Symposium on VLSI (GLSVLSI'16)	Boston, Massachusetts, USA	IEEE conference	TUD		Paper "Trellis-search based Dynamic multi-path connection allocation or TDM-NoCs"
Jun 28 - Jul 1, 2016	IEEE conference on dependable systems and networks (DSN 2016)	Toulouse (FR)	conference	TUD		Paper "Fault Tolerant Routine Algorithms for the Hexagonal Network-on-Chip based on the Turn Model"

POSTERS

Software mechanisms for ARM based Microserver Internal/external Communication using shared memory on RDMA

Authors: KD. Poulios, J. Velegrakis and M. Marazakis

Link to the publication web-site: https://www.hipeac.net/2015/amsterdam/workshop-and-tutorials

TECHNICAL REPORTS (derived from theses)

OS support for using remote memory and network interface on a system with an RDMA-capable interconnect

FORTH/TR-452, April 2015

Authors: J. Velegrakis

Abstract: Recent efforts towards performance and power optimization in large-scale Data Centers have brought the use of Microservers in the forefront. Compared to traditional architectures, Microservers consist of smaller, less power-hungry Compute Units (CUs) compared to traditional architectures. The key concept is that by integrating such smaller CUs in high numbers, the resulting many-core system can achieve high multi-threaded performance, while maintaining a low power profile. In such an environment, expensive resources must be shared among CUs, since it is costly and impractical to dedicate one to each CU. However, the management of these resources requires to implement sharing mechanisms in the Operating System (OS) and the software stack. In this work, we investigated and implemented OS and user space software mechanisms that are necessary for the deployment of ARM-based CUs in such large-scale systems. We address the sharing of remote resources by fully exploiting the underlying hardware features, such as Remote Direct Memory Access (DMA) and Remote Load/Store. In particular, we have implemented software mechanisms to (1) enable access to remote memory and (2) allow usage of a shared virtualized 10 Gbps Network Interface Card (NIC) by several CUs simultaneously. Remote memory access is implemented in three different ways: (1) As an extension of the local DRAM of the CUs, (2) As a remote Swap Device, and (3) as an I/O character device accessed directly from user space. We demonstrate that remote memory can be used effectively without performance penalty in a system running a full OS. The sharing of a virtualized 10Gbps NIC is achieved by a kernel network driver, that we have implemented, which enables utilization of the customized hardware as a standard Ethernet Device. This allows legacy applications that use Berkeley Sockets to run unmodified. The network driver makes use of scatter-gather DMA for fast zero-copy packet transmission and reception and operates in Full-Duplex mode by using two independent DMA channels. Additionally, it supports Interrupt Coalescing and management of the MAC and PHY hardware blocks using the Management Data Input/Output (MDIO) interface. In conclusion, this work shows that we can indeed utilize a system built upon ARM-based CUs that were not originally designed to operate in such an environment, by the sharing of remote and shared resources by the Linux OS and its user space environment. We expect this work to become even more relevant with upcoming 64-bit ARM-based platforms, targeting large-scale servers for Data Centers.

Link to the publication web-site:
http://www/carv/index_main.php?pub_title=&author=&author_fn=&pub_ctg=6&pub_subctg=0&yearFrom=&yearTo=&lab=CARV&submitted=true&Submit=Search&pbl=CARV&l=e

Sockets-based communication over an RDMA-capable interconnect

FORTH/TR-455, April 2015

Authors: D. Poulios

Abstract: In recent years, changes in the server market have brought power and space efficient server designs, like the Microserver. Such designs utilize large numbers of lightweight compute nodes bundled together to serve scale-out data center workloads. Unfortunately, scalability can often be limited by the quality of internal communication among running nodes, where low throughput and, even more critically, high latency can lead to poor performance. In this work, we explore the efficiency of a Remote Direct Memory Access (RDMA) capable internal network in a Microserver environment. Applications commonly use the standard Socket API for interprocess communication across networks. Therefore, to take advantage of the aforementioned internal network without modifying existing applications, socket-related system calls have to be intercepted. We implement system call interception in user space, using a modified Standard C Library, in order to bypass the kernel TCP / IP stack. A kernel driver has also been developed to securely perform data transfers via RDMA operations, which require physical addresses. Remote completion notifications of RDMA operations are triggered by a custom hardware Mailbox Mechanism, which also handles communication among nodes, necessary to initiate and close local connections. By combining these user and kernel space elements, we direct local TCP traffic through our internal network. Evaluation results show a 3x to 5x improvement to the latency, using our system compared to a typical ethernet configuration.
Link to the publication web-site:
http://www/carv/index_main.php?pub_title=&author=&author_fn=&pub_ctg=6&pub_subctg=0&yearFrom=&yearTo=&lab=CARV&submitted=true&Submit=Search&pbl=CARV&l=e

date1	date	Partner	Phd/Master	student	topic
J-14	Jan 2014	BSC	PhD	Renan Fischer e Silva	Heterogeneous architecture Interconnect Energy Saving on Microserver workloads
F-14	Feb 2014	CHALMERS	MSc	Li Kang	Memory compression
M-14	Mar 2014	TUD	PhD	Yue Zhen Wen	MPSoC runtime system
J-14	Jun 2014	BSC	PhD	Luis Garrido	Heterogeneous architecture Virtualization techniques for the exploitation of resources across coherence islands
J-14	Jun 2014	CHALMERS	Lic Eng	Dmitry Knyaginin	Hybrid (DRAM/NVM) memory systems
J-15	Jan 2015	TUD	PhD	Friedrich Paulus	Data management, optimization for telecom MPSoC
J-15	Jan 2015	TUD	PhD	Sebastian Haas	Multiprocessor, achitecture for data intensive applications
A-15	Apr 2015	FORTH	MSc	Yannis Velegrakis	OS support for using remote memory and network interface on a system with an RDMA-capable interconnect
A-15	Apr 2015	FORTH	MSc	Dimitris Poulios	Sockets-based communication over an RDMA-capable interconnect
M-15	May 2015 (expected)	CHALMERS	MSc	Jonas Andersson	Memory compression
M-15	May 2015 (expected)	CHALMERS	MSc	Niklas Doverbo	Memory compression
S-15	Sept 2015 (expected)	CHALMERS	PhD Eng	Angelos Arelakis	Memory compression
S-15	Sep 2015	TUD	PhD	Nairuhi Grigorian	Software Defined RAN Architecture