

Suren A. Chilingaryan, PhD

System Architect | HPC & GPU Computing | Lead of Cloud Computing

📍 76139 Karlsruhe, Germany @ csa@suren.me in linkedin 📧 @csa7fff

System architect with 20 years of experience designing distributed, data-intensive systems for large-scale scientific instrumentation. Strong background in HPC, GPU acceleration, low-latency data paths, and cloud-native platforms. Proven ability to turn research prototypes into reliable production systems and to lead cross-team engineering efforts. Recently led the cloud-native transformation of the KATRIN neutrino experiment, enabling 200 Gb/s data acquisition and integrating thousands of sensors.

KEY COMPETENCES

Technical Leadership	Project management, team leadership, presentations, proposals & reports
System Architecture	Distributed systems, microservice architectures, scalable data-intensive applications, high-bandwidth data acquisition and data-driven control
HPC & Performance	Parallel & approximate algorithms, performance modeling, code optimization
Infrastructure	Kubernetes/OpenShift, distributed storage, TSDB & array databases, CI/CD
Key Technologies	C/C++, POSIX/IPC/Linux Kernel, CUDA/OpenCL, NUMA, Remote-DMA

SELECTED PROJECTS

Since 2018	KATRIN-as-a-Service (KaaS) platform for distributed data acquisition systems › Prototyped platform, secured funding, built a team of 4, and led the rollout › Coordinated with Canonical and internal teams to deploy Kubernetes infrastructure › KaaS cloud is integrating data from fast detectors and thousands monitoring sensors › Cut costs by adapting legacy software for distributed ops and integrating HPC resources › Led development of AF_XDP-based, containerized DAQ framework › Modeled data access patterns and guided research on organization of time-series data › Supervised development of online monitoring framework for the KaaS ecosystem
2017 – 2022	Research and applications of hidden aspects in GPU architectures › Identified performance imbalances and hidden parallelism in NVIDIA and AMD GPUs › Developed method to exploit intra-SM parallelism on NVIDIA (up to 30% speed-up) › Accelerated tomography by 4-6 times by rebalancing of hardware usage › Published paper "Reviewing GPU architectures to build efficient back projection..."
2016 – 2017	GPU-based L1 Trigger for CERN CMS using GPUDirect (6 μs latency budget)
2011 – 2017	Cordinated software work-package and led algorithm paralelization in a project "Ultrafast tomography with online monitoring and image-based control"

EMPLOYMENT

Since 2020	Lead of Cloud Computing at Karlsruhe Institute of Technology: High-throughput cloud technologies for control and data acquisition systems
2013 – 2020	Data Processing and Performance Expert at KIT: high-performance instrumentation for large-scale scientific experiments
2007 – 2012	Postdoctoral Researcher at KIT: data management & software optimization
2005 – 2007	Software Engineer at Yerevan Physics Institute: data acquisition systems
2002 – 2005	Associate Researcher at Forschungszentrum Karlsruhe
2001 – 2002	Software Engineer at Yerevan Physics Institute

PERSONAL DATA

Nationality	Armenian, German Residence Permit (Niederlassungserlaubnis)
Languages	English, Armenian, Russian

EDUCATION

July 2006	PhD in Computer Engineering “Data Exchange Solution for Distributed Data Acquisition Systems and its application for Cosmic Ray Monitor Networks”, National Academy of Science, Armenia
1996 – 2001	MSc in Mathematics “Data Processing using Neural Networks”, Moscow State University, Russia

PROJECTS

2012 – 2017	Remote-DMA based Data Acquisition Platform
›	Developed a driver platform for fast prototyping of PCIe-based electronics with a modular (user-space) DMA engine (up to 12 GB/s) and scripting/debugging support
›	Implemented drivers for in-house electronics, e.g. camera (<i>in use at KIT and Desy</i>)
›	Used RDMA and GPUDirect/DirectGMA to intercommunicate detectors with GPUs
›	Led a case study on applications of GPUs for CMS track trigger, decision on acceptance of a track candidate was made within 6 μ s (data transfer + analysis)
›	GPUs allowed to utilize a more precise algorithm than was possible with classic designs
2010 – 2017	Coordinated student exchange program with Tomsk Polytechnic University
2015 – 2016	Cloud platform for collaborative analysis of tomographic data
›	Led development of web-visualization for large and time-resolved volumes
2011 – 2015	UFO: Ultrafast tomography with online monitoring and image-based control
›	Proposed a scalable architecture for pipelined processing of image streams
›	Coordinated software development with a team in KIT and 3 Russian universities
›	Supervised development of fast reconstruction algorithms (up to 6 GB/s per node)
›	Supervised development of regularized reconstruction methods to compensate low SNR and/or undersampling in case of high-speed tomography
›	Since 2016, the developed system is installed at the KIT synchrotron and enables both high-speed (5 volumes/second) and high-throughput (1000 samples/week) operation
2011 – 2014	ADEI: Advanced Data Extraction Infrastructure
›	Helped to secure funding and initiated collaboration between KIT and YerPhI
›	Supervised a cross-university team of researchers and engineers
›	Developed a platform for exploration and analysis of time-series in terascale archives
›	System is in operation in YerPhI, the SEVAN network, and 7 major facilities at KIT
2009 – 2013	Parallel algorithms and software optimization, 10-100x speed-up
›	Developed parallel algorithms for μ PIV (micro-particle velocimetry)
›	Leveraged the PoweXCell architecture for a MRSES feature selection algorithm
›	Optimized performance of PyHST (ESRF tomographic framework)
›	Implemented a digital image correlation and tracking algorithm for GPUs
2002 – 2008	Distributed systems for data acquisition and slow control
›	Stabilized a slow control system of the KATRIN experiment (<i>in use</i>)
›	Built a distributed data acquisition system for ASEC particle detector network (<i>in use</i>)
›	Developed Linux drivers and control software for AI-accelerator

TECHNICAL SKILLS

Project Management	Team Building, Proposals & Reports, Technical & Financial Planning
Research	Mentoring & Supervision, Presentation & Publication, Math & Statistics
Programming	C/C++, POSIX/IPC, OpenMP/NPTL, NUMA, SIMD, CUDA/OpenCL
System Programming	Linux & Kernel Internals, PCIe Drivers, DMA/GPUDirect/DirectGMA
Networking	Ethernet/Infiniband, EPICS, Sockets/MPI/0MQ, RDMA (RoCe), LibVMA
Software Analysis	gdb, perf/operf, nvvp/nvprof, vtune, valgrind, systap, tcpdump
Data Management	Brokers, Databases (SQL/TSDB/Array/InMem), XML, Web Development
Orchestration	Ansible, Kubernetes/OpenShift, OCI/Docker, LibVirt/KVM, Ceph/GlusterFS