Mikhail Kurnosov

Email: mkurnosov@gmail.com
Web: www.mkurnosov.net
Location: Novosibirsk, Russia

Positions and Experience

Siberian State University of Telecommunications and Information Sciences

Computer Systems Department, Novosibirsk, Russia
2005 – present Professor

- **Performance Analysis of HPC Applications in Presence of Network Congestion** 2021-2023
  - Development of software tools for analyzing MPI communication patterns and MPI traffic characteristics of open-source HPC applications, mini-apps and benchmarks in presense of network congestion: LAMMPS, NAS Parallel Benchmarks, Mantevo mini-apps, HPCG (C99, x86-64, ARMv8)
  - Profiling tools for MPI 3.1 applications (C/C++, Fortran), message size statistic, communication matrices for P2P and RMA operations (C99, Python, gnuplot, graphviz)
  - Traffic generation software for congesting multi-stage InfiniBand networks (C99, MPI, incast P2P/RDMA, all-to-all)
  - Congestion monitoring tools for InfiniBand multi-switch network (Python, ibverbs, InfiniBand PerfCounters)
- **Shared Memory Based Collective Communication Algorithms** 2020-2023

  Development of shared memory based algorithms for MPI collective communication operations: broadcast, allreduce/reduce, barrier: copy-in-copy-out, ZeroCopy using Linux CMA and kernel-assisted XPMEM/KNEM (Open MPI, C99, HWLOC, PMIx, GNU/Linux, x86-64)
- **Collective Operations for Clusters with Multi-stage Networks** (Huawei cooperation) 2022-2023

  Development of topology-aware collective communication algorithms for multi-level network topologies (fat tree, spine-leaf, kD torus topologies) and memory hierarchies: L2/L3 caches, sockets/packages, NUMA nodes (Open MPI, C99, HWLOC, PMIx, GNU/Linux, ARMv8)
- **Topology-aware Collective Operations MPI 3.1** (Huawei cooperation) 2021—2022

  Component with collective algorithms for Open MPI library. Key features: dynamic construction of collective schedules for nonblocking operations MPI 3.1 (directed acyclic graphs), asynchronous execution of send/recv/reduce operations on data readiness (C99, InfiniBand/RoCE, x86-64, ARMv8)
- **Hierarchical Algorithms for MPI Collective Opeartions** (Huawei cooperation) 2018—2021

  OpenMPI-based library of MPI blocking collective operations. Two-level topology: intra-node shared memory and inter-node communication network (C99, HWLOC, PMIx, InfiniBand, GNU/Linux, x86-64, ARM)
- **The Open MPI Project** (https://github.com/open-mpi/ompi) 2016—2018

  Contribution to open source high performance message passing library — implementation of MPI collective communication algorithms: blocking, non-blocking operations (C99, GNU/Linux)
- **Performance Analysis and Optimization** 2008—2017
  - Experimental analysis of autovectorizers of C/C++ compilers on x86-64 and Xeon Phi (Intel C/C++, GCC C/C++, LLVM/Clang, PGI C/C++, modified Extended Test Suite for Vectorizing Compilers, Intel RAPL)
  - MPIPerf project (https://github.com/mkurnosov/mpiperf) — benchmark for MPI 3.0 collective

and point-to-point communication routines (C99, GNU/Linux)
  - MPI Process Placement Tool (TopoMPI): optimizes the placement of MPI processes across cluster processor cores in order to localize communications through node's system memory (graph parititioning, communication profiling? C99, MPI Profiling Interface)

- **Designing and Administration of HPC Clusters** 2005—2010

  NIS/LDAP, NFS, SLURM/TORQUE, InfiniBand, custom GNU/Linux distro for diskless boot

- **Teaching Courses** (http://www.mkurnosov.net/teaching) 2006—present

  Software performance optimization courses (branchless code, ILP, caches, TLB, huge-pages, NUMA balancing and memory policy, code vectorization: AVX2/AVX512), parallel programming (MPI, OpenMP), distributed computing (MapReduce), data structures and algorithms

  Advising: bachelors, masters and PhD students — three candidate of science (PhD)

Huawei
Kunpeng Math Library, Novosibirsk, Russia
2023 — present Expert
(part-time)
Performance optimization of scalable sparse solver (MPI, OpenMP, SIMD, ARM64)
Performance analysis and optimization of HPC applications

Rzhanov Institute of Semiconductor Physics Siberian Branch of Russian Academy of Sciences, Computer Systems Laboratory, Novosibirsk, Russia
2006 – present Senior Research Scientist

Design and Optimization of Parallel Programs (C, MPI, OpenMP, CUDA, SIMD-vectorization): molecular dynamics (custom software package)

Yandex, School of Data Analysis, Novosibirsk, Russia
2014 – 2016 Lecturer

Parallel and Distributed Computing Course (C++11 threads, OpenMP, MPI, MapReduce, Hadoop, distributed algorithms)

Intel, Threading Tools, Nizhny Novgorod, Russia
2006 Software Engineer (Intern)

Porting GNU/Linux version of Intel Thread Profiler runtime library to the Intel Pin – dynamic binary instrumentation tool (C/C++, Pin probes, POSIX threads, GNU/Linux)

Republican Center of Children's Creativity, Gorno-Altaysk, Russia
1998 – 2005 Software Engineer

- **Educational Programming Language Rapira++** (https://github.com/mkurnosov/rapiraxx)

  Development of interpreter for educational programming language Rapira++ with: modifiable syntax (Russian and Altay keywords), basic OOP constructions, IDE with visual programing (Delphi, Windows)

- Client-Server Software for Monitoring and Remote Control of Application on Windows-based Workstations (C, DLL injection, Win32 API, network sockets, MySQL)
- Teaching (high school students): Rapira++, Visual Basic, Delphi, JavaScript; D-Link network technologies (TCP/IP, Ethernet switches: VLANs, Spanning Tree, QoS; WiFi protocols)

Education

2016	**Doctor of Science**, Siberian State University of Telecommunications and Information Sciences, Novosibirsk, Russia
2005 – 2008	**Candidate of Science (PhD)**, Siberian State University of Telecommunications and Information Sciences, Novosibirsk, Russia
2000 – 2005	**Diploma in Mathematics**, Gorno-Altaysk State University, Gorno-Altaysk, Russia

Awards

2023	Award “Master Of Communications”, Ministry of Digital Development, Communications and Mass Media of the Russian Federation
2021	Diploma of the Minister of Digital Development, Communications and Mass Media of the Russian Federation
2019	Certificate of Honor of the Governor of the Novosibirsk Region
2017	Diploma of city administration of Novosibirsk
2012	Award of the Government of the Russian Federation in the field of education
2009	Award of Administration of Novosibirsk Region
2008	Intel Scholarship in Recognition of Academic Progress and Active Scientific Work
2007	Alcatel-Lucent Scholarship
2007	Scholarship of the Government of the Russian Federation
2004	Scholarship of the Russian Federation President

Attended Schools and Workshops

2009	Architecture of High-Performance Computer Clusters, Institute for System Programming of RAS, Moscow, Russia
2007	Intel Multicore Programming for Academia, Intel, Nizhny Novgorod, Russia
2007	Java Programming, Sun Java Academy, Sun Microsystems, Novosibirsk, Russia
2006-2008	Russian-German Schools on Parallel Programming and High-Performance Computing Systems, Institute of Computational Technologies SB RAS, Novosibirsk, Russia

Selected Publications (mostly in Russian)

ORCID: orcid.org/0000-0002-7808-1635 Scopus Author ID: 23667793600
ResearcherID: C-9586-2016

Book Chapters

1. Khoroshevsky V., Kurnosov M. et al. Computational Methods, Algorithms and Hardware and Software Tools for Parallel Modelling of Natural Processes. Chapter 2, SB RAS, 2012. – 355 p. (in Russian, ISBN 978-5-7692-1237-6).
2. Kurnosov M. Introduction to Data Structures and Algorithms. – Novosibirsk, 2015. – 179 p. (in Russian, ISBN 978-5-9906983-4-5)
3. Kurnosov M., Paznikov A. Theory of Distributed Computer Systems Functioning Organization. – Novosibirsk, 2015. – 52 p. (in Russian, ISBN 978-5-9906983-5-2).

Dissertation (thesis)

1. Kurnosov M. *Algorithms for Functioning Organization of Hierarchical Distributed Computer Systems* (Doctor of Science), Siberian State University of Telecommunications and Information Sciences, Novosibirsk, Russia, October 2016.
2. Kurnosov M. *Models and Algorithms of Mapping Parallel Programs into Distributed Computer Systems*, Ph.D. Dissertation (Candidate of Science), Siberian State University of Telecommunications and Information Sciences, Novosibirsk, Russia, December 2008 (Advisor: Corresponding Member of RAS Prof. V.G. Khoroshevsky).

Selected Refereed Journal Articles

- Romanyuta A., Kurnosov M. *Shared memory based MPI Reduce and Bcast algorithms* // Journal of Numerical Methods and Programming. Vol. 24, Issue 4. 2023. https://doi.org/10.26089/NumMet.v24r424s
- Kurnosov M. *MPI Reduction and Broadcast Algorithms for Computer Clusters with Multistage Interconnection Networks* // Journal "Vestnik SibGUTI", 2023, No. 3, 13 p. (in Russian).
- Kurnosov M. *Barrier Synchronization Hierarchical Algorithm for Multicore Shared-memory Systems* // Journal "Vestnik SibGUTI", 2022, No. 2, pp. 4-11 (in Russian).
- Kurnosov M., Tokmasheva E. *Barrier Optimization on Asymmetrical NUMA Subsystems* // Journal "Vestnik SibGUTI", 2021, No. 1, 15 p. (in Russian).
- Kurnosov M., Tokmasheva E. *Shared Memory based MPI Broadcast Algorithms* // Journal "Vestnik SibGUTI", 2020, No. 1, pp. 42-59 (in Russian).
- Kurnosov M. *Analysis and Optimization of Pipelined Broadcast Algorithms* // Journal "Vestnik SibGUTI", 2019, No. 2, pp. 43-56 (in Russian).
- Peryshkova E., Kurnosov M. *Modeling Network Contention Effects on Process Allocation in Computer Systems* // Journal "Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitelnaja tehnika i informatika" (Tomsk State University Journal of Control and Computer Science), 2019, No. 47, pp. 93-101 (in Russian).
- Kurnosov M. *Analysis and Optimization of a k-chain Reduction Algorithm for Distributed Computer Systems* // Journal "Numerical Methods and Programming", 2017. – Vol. 17. – pp. 318-328 (in Russian).
- Kulagin I., Kurnosov M. *Instrumentation and Optimization of Transactional Sections Execution in multithreaded Programs* // Proc. of the Institute for System Programming. – 2015. – Vol. 27 (6). – pp. 135-150 (in Russian).
- Kulagin I., Paznikov A. Kurnosov M. *Heuristic Algorithms for Optimizing Communications in Parallel PGAS-programs* // Journal "Vestnik SibGUTI", 2014, No. 3, pp. 52-66 (in Russian).
- Pavsky K., Kurnosov M., Polyakov A. *Software Tools for Optimizing Parallel Modeling of Nanostructures with Quantum Dots* // Journal "Avtometria", 2014, Vol. 50(3), pp. 56-61 (in Russian).
- Kurnosov M., Paznikov A. *Heuristic Algorithms for Mapping Parallel MPI Programs into Multicluster Computer and Grid Systems* // Journal "Vichislitelnie metodi I programmirovanie", 2013, Vol. 14(2), pp. 1-10 (in Russian).
- Kurnosov M. *MPIPerf: a Toolkit for Benchmarking MPI-libraries*. Journal "Vestnik NNGU", 2012, No. 5(2), pp. 385-391 (in Russian).
- Kurnosov M., Paznikov A. *Modelling of Decentralized Algorithms for Scheduling Jobs in Grid Systems* // Journal "Problemi informatiki", 2012, No. 2, pp. 45-54 (in Russian).
- Kurnosov M., Paznikov A. *Decentralized Algorithms for Scheduling Parallel Tasks in Geographically-distributed Computer Systems* // Journal "Vestnik TGU. Upravlenie, vichislitelnaya tehnika i informatika", 2012, No. 1(18), pp. 133-142 (in Russian).
- Kurnosov M. *Allgather Algorithms for Hierarchical Distributed Computer Systems* // Journal "Vestnik Komputernih i Informacionnih Tehnologiy", 2011, No. 5, pp. 27-34 (in Russian).

- Kurnosov M. *Optimization of Collective Communications Routines in Computer Systems with Hierarchical Networks* // Journal "Vestnik TGU. Upravlenie, vichislitelnaya tehnika i informatika", 2011, No. 2(15), pp. 61-71 (in Russian).
- Kurnosov M., Paznikov A. *Algorithms and Software Tools for Decentralized Scheduling of MPI Programs in Multicluster Computer Systems*. Journal "Vestnik TGU. Upravlenie, vichislitelnaya tehnika i informatika", 2011, No. 3(16), pp. 78-85 (in Russian).
- Kurnosov M. *Structure-oriented Method for Optimizing MPI Collective Communications in Distributed Computer Systems* // Journal "Vestnik SibGUTI", 2010, No. 2(10), pp. 54-65 (in Russian).
- Kurnosov M., Paznikov A. *Decentralized Scheduling of Parallel Tasks in Geographically-distributed Computer Systems* // Journal "Vestnik SibGUTI", 2010, No. 2(10), pp. 79-86 (in Russian).
- Khoroshevsky V.G., Kurnosov M.G. *Algorithms for Assigning Parallel Program Branches to Computer System Processor Cores* // Optoelectronics, Instrumentation and Data Processing. – 2008. – Vol. 44, No. 2. – P. 135-143.

Conference/Workshop Proceedings

- Romanuta A., Kurnosov M. *Shared Memory-based Algorithms for MPI Irregular Collective Operations* // Proc. of the IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON), 2022.
- Kurnosov M., Tokmasheva E. *Optimizing Barrier Algorithms on Asymmetric Subsystems of NUMA Machines* // Proc. of the IEEE Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT-2021), online, 13-14 May, 2021, 5 p.
- Kurnosov M., Tokmasheva E. *Shared Memory based MPI Broadcast Algorithms for NUMA Systems* // Russian Supercomputing Days: Proceedings of the International Conference, 2020. – P. 1-12.
- Kurnosov M., Berlizov D., Tkacheva T., Tokmasheva E. *Analysis and Optimization of Pipelined Broadcast Algorithms on Gigabit Ethernet and InfiniBand Networks* // Proc. of the 15th Int. Asian School-Seminar Optimization Problems of Complex Systems (OPCS), 2019. – pp. 86-91. ISBN 978-1-7281-2986-0.
- Peryshkova E., Kurnosov M. *Experimental Study of Network Contention Effects on All-to-All Operation* // Proc. of the 14th International Scientific-Technical Conference Actual Problems of Electronic Instrument Engineering (APEIE-2018), 2018. – Vol. 6 – P. 506-510.
- Moldovanova O., Kurnosov M. *Automatic SIMD Vectorization of Loops: Issues, Energy Efficiency and Performance on Intel Processors* // Russian Supercomputing Days: Proceedings of the International Conference, 2017. – P. 55-66.
- Paznikov A., Kurnosov M., Kupriyanov M. *Algorithms of Collective Operations for Distributed Arrays in Partitioned Global Address Space* // 2017 IEEE II International Conference on Control in Technical Systems, Saint Petersburg, Russia, 2017, pp. 5-8.
- Moldovanova O., Kurnosov M. *Auto-Vectorization of Loops on Intel 64 and Intel Xeon Phi: Analysis and Evaluation* // Proc. of the 14th International Conference on Parallel Computing Technologies (PaCT-2017), 2017. – Springer LNCS 10421. – P. 143-150.
- Kurnosov M. *Dynamic Mapping of All-to-All Collective Operations into Hierarchical Computer Clusters* // Proc. of Int. scientific-technical conference on Actual Problems of Electronic Instrument Engineering (APEIE-2016), 2016. – Vol. 1, Part 2. – 475-478.
- Kulagin I., Kurnosov M. *Optimization of conflict detection in parallel programs with transactional memory* // Proc. of 10th Annual International Scientific Conference on Parallel Computing Technologies (PCT-2016). – pp. 582-594.
- Kulagin I., Paznikov A., Kurnosov M. *Heuristic Algorithms for Optimizing Communications in Parallel PGAS-programs* // Proc. of the 13th International Conference on Parallel Computing Technologies, 2015. – Springer Lecture Notes in Computer Science. Vol. 9251. – pp. 405-409.
- Kurnosov M., Paznikov A. *Efficiency Analysis of Decentralized Grid Scheduling with Job Migration and Replication* // 7th International ACM Conference on Ubiquitous Information Management and Communication (ICUIMC-2013), Malaysia, 2013. – 7 p.

- Khoroshevsky V., Kurnosov M. *Mapping Parallel Programs into Hierarchical Distributed Computer Systems* // Proceedings of 4th International Conference "Software and Data Technologies (ICSOFT 2009)", - Sofia: INSTICC, 2009. - Vol. 2. - P. 123-128.
- Kurnosov M.G. *MPIPerf: a Toolkit for Benchmarking MPI-libraries* // Proc. of International conf. "Parallel Computational Technologies", Novosibirsk, Russia, 2012, pp. 212-223 (in Russian).
- Kurnosov M.G. *Topology-aware Collective Communication Algorithms for Distributed Computer Systems* // Proc. of Conference "Supercomputer technologies: development, programming, application" (SCT-2010), Divnomorskoe, Russia, 2010, Vol. 2, pp. 62-66 (in Russian).
- Kurnosov M.G. *Structure-oriented Subsystems Allocation in Computer Systems* // Proc. of conference "High-performance parallel computing on clusters", Kazan, Russia, 2008 (in Russian).
- Khoroshevsky V., Kurnosov M. *Modelling of Algorithms for Mapping Parallel Applications into Structures of Computer Systems* // Proc. of international conf. "Simulation-2008", Kiev, Ukraine, 2008, Vol. 2, pp. 435-440 (in Russian).
- Kurnosov M. *Parallel Algorithm for Mapping Communication Graph of MPI Task into Computer System* // Proc. of international conf. "Parallel Computational Technologies", Chelyabinsk, Russia, 2008 (in Russian).
- Kurnosov M. *Experience in Building Computer Clusters with a Remote Diskless Boot* // Proc. of conference "High-performance parallel computing on clusters", Nizhny Novgorod, 2005, pp. 149-154 (in Russian).

Research Grants (Principle investigator)

1. Topology Representation and Collective Communication Algorithms in HPC Heterogeneous Clusters Networks (R&D contract), *Huawei*, 2022-2023.
2. Concurrent Traffic Model and Performance Optimization (R&D contract), *Huawei*, 2021-2022.
3. Topology-aware MPI collectives for Huawei ARMv8-based systems (R&D contract), *Huawei*, 2019-2020.
4. Models and methods of analyzing and organizing of multiprogram execution of parallel programs on large-scale computer systems, *Russian Foundation for Basic Research*, 2018-2020
5. Models, algorithms and software for optimizing PGAS-programs, *Russian Foundation for Basic Research*, 2015-2017
6. Algorithms and system software for optimizing functioning of hierarchical computer systems, *Russian Foundation for Basic Research*, 2015-2016
7. Models, Methods and Software for Efficient Execution of Parallel Programs on Multiarchitectural Computer Systems, *Russian Foundation for Basic Research*, 2011-2013
8. Topology-aware Algorithms and Software for Functioning Organization of Distributed Computer Systems, *Russian Foundation for Basic Research*, 2008-2010
9. Development of Tools for Mapping Parallel MPI Programs into Multicore Computer Clusters, *Foundation for Assistance to Small Innovative Enterprises*, 2008-2009
10. Grant of Novosibirsk's Administration, 2009