Proposal for Theme 1119 Extension during 2020–2023

1. **Title of the theme:** 05-6-1119-2014/2023, "Methods, Algorithms and Software for Modeling Physical Systems, Mathematical Processing and Analysis of Experimental Data"
   Priority: 1 Status: Extended
   Leaders: Gheorghe Adam, Petr V. Zrelov
   Deputies: Ján Buša, Ochbadrakh Chuluunbaatar
   Theme beginning: 2014
   Participating JINR Laboratories: LIT, VBLHEP, BLTP, FLNR, FLNP, DLNP, LRB, UC
   Participating Countries, Institutes and International Organizations:
   Armenia, Belarus, Belgium, Brazil, Bulgaria, Canada, CERN, China, Czech Republic, France, Georgia, Germany, Greece, Israel, Italy, Japan, Kazakhstan, Moldova, Mongolia, Poland, Romania, Russia, Slovakia, South Africa, Switzerland, Tajikistan, USA, Vietnam.

2. **Abstract of the theme**
   Carrying out paramount advanced research in computational mathematics and physics, directed to the creation of new mathematical methods, algorithms, and software for the numerical or symbolic-numerical solution of topics arising in experimental and theoretical physics studies. This subject area includes a wide spectrum of investigations approved for completion in JINR within the seven year period 2017–2023 in high energy physics, nuclear physics, physics of condensed matter and of nanostructures, biophysics, information technologies, the solution of which is inseparable from the use of computing. Such subject matters of the outmost importance in JINR are the NICA project, the neutrino program, the superheavy and exotic nuclei physics, the neutron based investigations. The needed numerical or symbolic-numerical computing will be done on the Multifunctional Information and Computing Complex (MICC), primarily the HybriLIT heterogeneous computing platform (which involves the training and testing cluster HybriLIT and the "GOVORUN" supercomputer) and the emerging Big Data distributed infrastructure. The research teams include both experienced scientists with outstanding scientific achievements and enthusiastic young scientists and engineers. The requested financing will cover salaries, participations in scientific conferences, scientific visits and the acquisition of a minimal number of personal computers and licenses, within the approved resources for LIT-JINR. A distinctive feature of this research is the close cooperation of the Laboratory of Information Technologies (LIT) with research groups from all JINR laboratories and Member State institutions.

3. **Introduction**
   A distinctive feature of the overwhelming part of the scientific projects approved to be carried on during the seven year period 2017–2023 in the Joint Institute for Nuclear Research (JINR) (http://www.jinr.ru/roadmap-en/) is the need of solving a broad spectrum of computing intensive tasks of ever increasing complexity. The significant increase of the installed power for high performance computing (HPC) through the heterogeneous supercomputer "GOVORUN" widens the realm of available computational complexity of the numerical and symbolic-numerical problems to be solved in LIT for the overall progress of the JINR scientific research.

   The extension of the Theme 1119 proposed for the four-years 2020–2023 is primarily motivated by the continuation of the demands coming from the most important JINR projects found under development during 2017–2023 (NICA, Baikal GVD, the creation of superheavy nuclei).

   The characterization of the activity done within the Theme 1119 as *unity in diversity* is preserved. The active participation of LIT scientists in over 40 JINR scientific projects points to the diversity of the solved tasks.

   Unity through *unifying principles* comes from the existence of a *common mathematical background* of all these projects. The strongest requirement following from this common background is to get *reliable* solutions within a *rapidly evolving hardware – software environment*.

   While the numerical and symbolic-numerical algorithms developed for the solution of the physical models rest on the *discretization* of their definition domains, the enforcement of the *reliability* of the numerical solutions asks for supplementary demands, apart from the classical *stability* requirement:

(i1) A discretization scheme is to *inherit all algebraic properties* of the continuous mathematical model. This inheritance is of crucial importance in the solution of *strongly non-linear* models.

(i2) A discretization scheme is to show *insensitiveness* with respect to the approximation of the infinite set of the real numbers by the finite set of the machine floating point numbers. More often than not, the implementation of this feature may ask for the development of *multiscale-adapted algorithms*. The multi-scaling heavily depends on the nature of the problem described by the mathematical model.

(i3) The *computational complexity* of the derived numerical solutions is to be as low as possible. To get low complexity algorithms characterized by increased accuracy and quality, it is often necessary to develop *new principles of approximation*. The derivation of new methods based on such principles needs deep culture, talent, inspiration, innovative thinking.

The enforcement of the reliability of the statistical analysis of the *rare events* (occurring, e.g., in the production of superheavy nuclei, or the identification of the tiny fraction of useful information present in the accumulated Big Data) asks for new methods and approaches which were absent in the previous quantitative models.

The implementation of *long lasting* numerical software is severely impeded by the *rapidly changing hardware environment*. The last hardware revolution, starting about a decade and a half ago, resulted in the replacement of the *single-processor* chips with a variety of *multi-core processor* dies, *manycore processor* dies, GPU *accelerators* in conjunction with multi-core processors, etc. In connection with this radical change of the hardware, the renowned *Gartner Institute* (GI) – specialized in long term predictions in the information technology (IT), has estimated that the design and implementation of *effective software*, able to fully exploit the parallel hardware architectures, is expected to have global outputs not earlier than by the end of the next decade. The huge variety of reports in this direction confirms the veracity of the GI prediction.

These hardware driven exceptional circumstances promote the *parallel computing* as the most necessary future development in LIT-JINR, under various programming paradigms grasped by the LIT staff. In this respect, the heterogeneous computing platform HybriLIT (http://hlit.jinr.ru/en) is the main resource for high performance computing (HPC) in JINR.

Nowadays, all over the world, more and more importance is attached to the creation of systems for processing and analysis of Big Data, machine learning and artificial intelligence for solving problems of the natural sciences. Among the most relevant and cutting-edge areas is the creation of solutions for combining Big Data technologies with the high-performance computing. This makes it possible the solution of certain classes of tasks at different stages of a physical experiment at a qualitatively new level – from modeling detectors and systems for data storage and analysis to event reconstruction, physical analysis, monitoring the detector functioning, and predicting possible emergency situations, studying the dissemination of topical scientific information, etc.

Another challenge to be faced in the near future is the *quantum computing*, the implementation of which will ask for a fundamentally new approach to the computing philosophy.

4. State-of-the-art of the research field within the theme 1119

The present extension proposal of the theme 1119 during the next four years, 2020–2023, is heavily based on the existing JINR culture characterized by symbiotic synergy between different research groups from all the JINR Laboratories and research groups in LIT. The first ones ask for the solution of specific mathematical topics raised by and needed for the fulfillment of their research projects. The LIT groups contribute to the development of mathematical methods, design of algorithms, and their implementation into software directed to the solution of the defined mathematical topics. Instances of such collaboration are spread throughout the proposal.

The viability of such an interaction scheme heavily depends on the good relationships of the group coordinators of the two sides and, last but not least, on the decisions of the JINR and JINR Laboratories management leaderships. From the perspective of all interested parties, these relationships can be characterized as excellent, with the highest chances of enhancement and fruitful continuation in the future.

One of the high advantages of this scheme is its *versatility*: the existing mathematical computing expertise in LIT frequently accommodates *several* concurrent needs coming from different JINR teams, hence overall harmonization of a broad range of scientific interests thus avoiding waste coming from unrelated developments of a same scientific expertise.

Even under the occurrence of excellent human premises for collaborative work, there are many underwater stones which have to be identified and overcome.

Ideally, a staff member working within the theme 1119 has to have a multifaceted culture in several, distant from each other, fields.

(a) *Thinking as a computer scientist*, which assumes expert knowledge of several modern programming techniques, is a prerequisite for successful implementation of numerical and/or symbolic-numerical HPC packages on present day supercomputers like "GOVORUN".

(b) A *deep mathematical expertise* is asked for the correct formulation of the problems under consideration, irrespective of their origin, in experimental or theoretical physics, the right definition of their features, like well-posedness, the bold advancement in the terra incognita of the Big Data.

(c) A *high education level in numerical analysis* enables the correct computational complexity tackling, the avoidance of the pitfalls coming from the operation with machine numbers instead of the real/complex numbers underlying the continuous mathematical models.

(d) The *profound understanding of the physics side of the problems* enables correct assessment of the orders of magnitude of the different quantities entering the investigated phenomena and use of the Weiss principle in the formulation of the hierarchically ordered numerical approximations.

Gathering a team with a firm grip on all these directions is a great challenge. Moreover, the diversity of the tasks to be solved raises supplementary specific issues, the most important of which are briefly discussed below.

- *The software support of a scientific experiment extends over all the life cycle of the experiment*. While there is a predefined program of the work to be done, there are many factors which ask for the change of the supporting software of a given experiment. A few instances mentioned below briefly illustrate the various changing circumstances.

(1) With the addition of position sensitive detectors to the YuMO setup, the most used detector facility at the IBR-2M reactor, the automation of the data acquisition and analysis through the development of primary data treatment software entered a new phase which was partially solved during the last years for the isotropic pattern scattering case. The orders of magnitude more complicated case of the anisotropic scattering is an open task to be solved in the present proposal.

(2) The work done at several IBR-2M diffractometers (like, e.g., the high resolution Fourier diffractometer (HRFD)) includes experiments on the study of irreversible processes during which the spectroscopic data are quickly changing. The question was quite recently raised on the maximally possible automation of the data processing (the interactive definition of the initial data, visualization of the acquired data, etc.), within different conditions of such experiments.

(3) Participation in the solution of technical problems enabling very high precision work of the Time-Projection Chamber (TPC), the main tracking detector of the experiment MPD at the NICA collider. A laser calibration system is being prepared for monitoring drift velocity and for taking electric field distortion into account. Such calibration system that can reproduce fiducial tracks with well-defined absolute positions is needed to monitor the TPC performance. In the long run, the implementation of such a system and its maintenance will be of high importance within the lifetime of the MPD/NICA experiments.

(4) The measurements made at the BM@N setup under changing detector configurations at different working sessions during the last two years ask for maximally adapted software able to take operatively into account the arising changes.

(5) The periodic renewal of the big detectors (ATLAS, ALICE, CMS, LHCb) during the LHC shutdowns asks for corresponding upgrade/modification of the software. The LIT team is best qualified for developing software support and maintenance of the CSC module of the CMS setup.

(6) The unique worldwide recognized discoveries of new superheavy elements, which completed the 7-th period of the Mendeleev's periodic table, will be pursued at a larger scale during the next four years. The analysis of the future FLNR results and their comparison with those of the concurrent groups will need further theoretical investigations concerning, e.g., reliable estimates of the radioactive nuclei lifetime and cross sections under small statistics.

- The availability of powerful computing tools supporting various parallelizing options (the heterogeneous "GOVORUN" supercomputer) make possible *new investigation methods concerning both data analysis and large scale computing experiments*.

(i) The data analysis based on the design of deep neural networks was recently successfully started in LIT-JINR in connection with several topics which aim at reliable pattern recognition under (slightly) changing circumstances. The continuation of such studies is very promising: the availability of sufficient resources for the deep learning starting period will result in unsurpassed performance of such algorithms at the predictive stage.

(ii) The already mentioned Big Data analytics, raised by the very large scale experiments will be a new direction of investigation which asks for the grasp of new ways of thinking.

(iii) A last example of unusual usage of the "GOVORUN" supercomputer concerns the run of very large scale numerical experiments. Motivated by the Memorandum of Understanding signed in November 2018 in Bucharest and foreseeing reciprocal support of the NICA and ELI-NP projects, this kind of undertaking has shown that the run of very big packages on a supercomputer is an issue of unprecedented complexity. The implementation of the PIConGPU package on the "GOVORUN" supercomputer is intended to understand and to characterize the phenomena related to the laser based acceleration of elementary particles, which promises to result in revolutionary changes of the principles guiding the buildup of the future new accelerators. While the available PIConGPU package was developed for CUDA graphic accelerators, its run on the "GOVORUN" supercomputer acutely raises the problem of its optimization on the Volta 100 tensor graphics accelerators.

- While the bulk of the activity during the next four years will be devoted to the creation, maintenance and upgrade of software tools supporting the flagship JINR experiments, *a number of independent developments will be also be pursued*, with the aim at discovering fundamentally new approaches to the solution of new topics, or of existing, but as yet unsolved, ones. In the long term run, it is estimated that such research will provide unique new computing tools, primarily for the scientific research done in JINR.

We have to mention the quantum information and quantum computing as a possible alternative to the processing of the large amounts of data expected to be accumulated at the future NICA experiments (MPD, SPD).

New effective algorithms are highly needed for already existing problems, within alternative approaches: molecular dynamics investigations based on new principles of code developments, new extensions of the three-point basic element method for the solution of smoothing, interpolation, or extrapolation problems, the accurate solutions of bound states, quasi-bound or scattering states in few body problems, multistage multiscale solution of integrals for the theoretical calculations of the observables, etc.

- Last but not least, the JINRLIB software library maintained both in English and in Russian, will have a significant role in providing convenient user access to thoroughly tested codes (especially parallel codes) and to good practice guidelines for the preparation of parallel programs.

5. Description of the proposed research

The research done within theme 1119 can be conveniently divided into four kinds of activities. Besides the mentioned theme leaders, the coordinators of these activities are reputed LIT scientists, energetic and with proved leadership qualities.

(1) Mathematical and computation methods for simulation of complex physical systems.

Leaders: Gh. Adam, J. Buša, I.V. Puzynin

(2) **Software complexes and mathematical methods for processing and analysis of experimental data.** Leaders: P.V. Zrelov, V.V. Ivanov

(3) **Numerical methods, algorithms and software for multicore and hybrid architectures and Big Data analytics.** Leaders: Gh. Adam, O. Chuluunbaatar, O.I. Streltsova, V.V. Korenkov, P.V. Zrelov

(4) **Methods, algorithms and software of computer algebra and quantum computing.** Leader: V.P. Gerdt

There is an intimate relationship between the theme 1119 and the development of the information-computing infrastructure of JINR, done within the theme 05-6-1118-2014/2023 and its project on the development of the Multifunctional Information and Computing Complex (MICC). On one side, the MICC provides the necessary computing power for the run of both sequential and parallel programs. On the other side, the team involved in the theme 1119 is the main contributor to the creation of new computer packages in JINR and the maintenance of existing ones. Even the professional use of outer (open or licensed) codes themselves needs creative work for their tuning (adaptation) to the tasks to be solved in the Institute and to the enhancement of their efficiency on the heterogeneous HybriLIT platform (HybriLIT cluster and "GOVORUN" supercomputer).

The output of the LIT scientists' activity within theme 1119 is twofold. One consists in the *creation of computational tools*, the other one in the *publication of scientific papers*.

The computational tools encompass three broad destinations, each of which asking for specific abilities and expertise.

(a) The straight solution of tasks emerging from projects conducted in JINR or done with JINR participation refers, as a rule, to the information-computing support of various experiments.

Since the *online data collection* is tuned to the detector details, the created software support necessarily depends on the details of the detector. Every modification of the detector design obligatorily asks for new contributions from the personnel securing the software support. As a consequence, *the software support lasts over the whole lifetime of the experiment*. A detail of practical importance is that the software support provision is a *discontinuous* process which liberates in the meantime the energies of the software specialists for other endeavors. The five examples presented in section 4 illustrate these ideas.

(b) The software developed for offline data processing and analysis is under two strong pressure factors imposing new developments.

The first factor, already mentioned at the end of the section 3, comes from the multi-core (manycore) processor structures, or from the presence of graphic accelerators (GPU), both asking for *parallelization* as a means to increase the efficiency of the computations.

The second factor comes from the changes, by many orders of magnitude, of the parameters at which the new experiments are done. Heuristic assumptions holding at lower energies, which entered software designed decades ago and becoming common places by repetition, were shown in LIT scrutiny to be unreliable under the parameters of the new high energy physics experiments. As a result, bugs in the widely used Higgins Monte Carlo generator were removed. Vital modules of the Geant4 package were included in the new 2010–2017 Geant4 releases. The last Geant4 major revision was defined with important co-authorship of V.V. Uzhinsky (LIT).

(c) Finally, we have to add to this incomplete enumeration the solution of *difficult mathematical problems* resisting so far to the previous attempts to get guaranteed output under controlled accuracy. The quality of such results, obtained within select international collaborations, is best illustrated by:

- Election of O. Chuluunbaatar as full member of the Mongolian Academy of Sciences in 2018 based on the scientific results got in LIT JINR;

- Official approval (during 2017) of the upgrade of the Cathode Strip Chamber software of the CMS setup (implemented by V. Palichik and N. Voytishin) as the standard module for the muon reconstruction;

- Essential development of the QGS and FTF models of Geant4 package has been achieved. These models are included in the favorite Physics List used by all LHC Collaborations. (V.V. Uzhinsky);

- The paper by V.P. Gerdt et al. presented at the International Symposium on Symbolic and Algebraic Computation (ISSAC 2017) was recognized as the best one and awarded the prize of the Association of Computer Machinery (ACM) (http://www.issac-conference.org/2017/awards.php);

- JINR Encouraging Prize 2017: «Determination of the decay time of scintillators and investigation of space correlation of nuclear radiation by the autocorrelation method».

Authors: V. Morozov, N. Morozova, *V. Zlokazov*

- The *Adiabatic Representation for Atomic Dimers and Trimers in Collinear Configurations,* implemented in a very efficient numerical package, was able to yield accurately asymptotic expansions of basis functions, effective potentials, fundamental solutions of the second-order ordinary differential equations and corresponding asymptotic scattering states, as well as the resonance scattering, metastable and bound states. (A.A. Gusev, S.I. Vinitsky, O. Chuluunbaatar et al., Physics of Atomic Nuclei, 2018, Vol. 81, No. 6, pp. 945–970)

A summary of the *scientific publications* coming from activity within theme 1119 of the LIT personnel during the last three years points to 205 co-authorships within the CMS project and other 294 articles published in refereed journals; 33 publications in periodical volumes; 35 invited lectures and 111 oral presentations at international conferences; 63 electronic publications. A list of selected papers is provided in the Report on theme on 2017–2019.

(1) Mathematical and computation methods for simulation of complex physical systems

This activity comprises the development and use of mathematical and computing methods for modeling new experimental facilities, accelerating complexes and their elements, nuclear-physical processes, complex physical systems. New mathematical methods will be developed and, where suitable, existing ones will be extended with the aim to take into account the main features of the physical processes and mathematical models: non-linearity, multi-parametric behavior, the existence of critical modes and phase transitions. The model refinement, the investigation of the possibilities of their use and comparison with experimental data will be mainly done by means of the development of parallel algorithms and their implementation in software packages tuned for the present day hardware architectures, primarily the HybriLIT heterogeneous computing platform.

• Calculations of magnetic fields for different facilities:

- Three-dimensional computer simulation of magnetic systems in the framework of the NICA (JINR) and FAIR (GSI) projects. Calculation of the required characteristics of the magnetic field in the working areas of the magnets.

- Developing fast FEM algorithms for 3D simulation of the magnetic field distributions in COMSOL Multiphysics software.

- Development of methods and algorithms for efficient highly accurate three-dimensional modeling of magnets and computations targeted to the creation of superconducting cyclotrons for proton therapy, in collaboration with DLNP.

- Optimization of beam dynamics programs in cyclotrons aimed at increasing speed, efficiency, and accuracy of calculations.

• The investigation of mathematical models of complex physical processes will be pursued in the frame of quantum-field and molecular-dynamics approaches:

- Development of effective QCD-motivated models for describing properties of nuclear matter at NICA energies:

(a) Generalization of the model Nambu – Jona-Lasinio with Polyakov loop and numerical analysis of the behavior of the masses and coupling constants of the quarks, diquarks, mesons and nucleons at finite temperature and density;

(b) Completion of the construction of a model for the numerical solution of hydrodynamic equations using new state equations and its combination with the kinetic description of the initial state and the state after freezeout;

(c) Simulation of specific properties of equations of state, in particular, the dependence of physical quantities on the magnetic field strength at low temperatures and large values of the chemical potential;

(d) Based on the hybrid model and using the developed codes combining the properties of the kinetic and hydrodynamic approaches, analysis of the available experimental data on the collision of heavy ions in the energy region of the future NICA accelerator.

- Development of new molecular dynamics algorithms aimed at increasing accuracy and significantly reducing the computing time.

- Development of models of interaction of ion beams with targets in order to explain various effects: long-range and description of structural changes of materials under heavy ion and nanocluster irradiation, finding threshold values of energy loss in irradiated materials, leading to structural changes and through tracks in thin targets.

- Development of mathematical models, algorithms and programs for the study of processes of local nonequilibrium created by the impact of ultrashort laser pulses on materials. Formulation and solution of systems of the heat conduction equations for the study of laser ablation, carrying out numerical experiments in the framework of molecular dynamics. Validation of theoretical models by comparing the obtained numerical results with experimental data.

• New mathematical methods will be developed and, where suitable, existing ones will be extended with the aim to take into account specific features of the physical processes:

- Development of numerical methods for describing equilibrium and nonequilibrium properties of mesoscopic systems of trapped atoms.

- Numerical investigation of optical and self-assembled atomic lattices, with emphasis on the control of their properties by external fields. The extension of optimized perturbation theory and self-similar approximation theory for these systems.

- Investigation of nonlinear multiparameter processes in complex physical systems under external fields, including models of superconducting structures, localized states in condensed media, simulation of gas-hydrodynamic processes in porous media.

- Numerical investigation of nuclear-physical processes based on a hybrid model of microscopic potential, including reactions with light exotic nuclei.

- Development of methods for modeling the reflection of neutrons from layered nanostructures. Studies of magnetic films as structures consisting of vector micro-objects.

- Studies of black hole and wormhole models, both in astrophysics and in cosmology, are aimed at understanding the results of new experiments.

• The continuous LIT information-computing support is instrumental for the improvement of the working regimes of several experimental facilities:

- Group of small-angle neutron scattering, FLNP (YuMO spectrometer at the IBR-2 reactor):

(a) Maintenance of the SAS primary processing program for the YuMO spectrometer;

(b) Development of a program for working with position sensitive detectors for isotropic and anisotropic scattering samples and for ring detectors;

(c) Development and maintenance of a parallel version of the Fitter program.

- For the BAIKAL project, DLNP: step-by-step creation of an alert system.

- Modeling of electromagnetic cascade showers in the field of ultrahigh energies and numerical study of their characteristics in experimental data obtained in such astrophysical projects as IceCube, Antares, Baikal.

• Development of reliable scale-adapted algorithms of reduced computational complexity:

- Development of an extrapolation method of the sixth order with the aim at improving the algorithm efficiency for the numerical solution of a wide range of tasks.

- Development of methods and algorithms for processing and analysis of the neutron noise of the IBR-2M reactor.

- Multistage multiscale approach to the Bayesian automatic adaptive quadrature.

- Development of algorithms for the numerical simulation of the evolution of a liquid crystal in a pulsed electric field, as well as under the influence of an orienting structured surface.

- Simulation of peculiarities of the absorption-emission and photon density of states of a cholesterol liquid crystal with isotropic defect inside.

- Quantum-chemical cluster approach to electronic systems with strong spin-orbit interactions.

- Carrying out calculations of electromagnetic and thermal quantities characterizing physical processes in MgB2 superconductors.

(2) Software complexes and mathematical methods for processing and analysis of experimental data

This activity is directed at the solution of the following main tasks: derivation of new mathematical methods for the extraction of the useful information from the raw data obtained in experiments done in JINR or with the JINR participation; development of algorithms and implementation of program packages for the solution of problems arising in the high energy physics – including the data got at the accelerator facilities LHC, NICA, FAIR as well as at the experimental facilities of the JINR neutrino program, the nuclear physics, the condensed matter physics and the physics of radiation biology. The development of deep learning neural network algorithms will become a significant part of this activity.

There are four classes of JINR undertakings covered with LIT participation.

• Development of mathematical methods, algorithms and software for reliable simulation and interpretation of the experimental data:

- Within the Geant4 package: simulation of the generation of charmed particles and their transport in matter; improvement of the algorithm for calculation of excitation energies of nuclear residuals and improvement of the description of nuclear fragment formation; improvement of FTF and QGSM models and their application for modeling the conditions of various experiments (PANDA, NICA/MPD, NICA/SPD).

- Development of software support, processing and analysis of data collected at the NUCLEON and COMBAS experiments.

- Investigation of the structure and properties of vesicular systems of phospholipids and medical drugs based on the phospholipid transport nanosystem by means of analysis of the neutron and X-ray small-angle scattering data, depending on external factors and chemical composition.

• Development of deep learning neural network algorithms:

- Development of a framework for the analysis and management of biological monitoring data for controlling and forecasting the status of the environment.

- Development of mathematical methods for the determination of fine structures in the distribution of nuclear reaction products by mass and energy.

- Development of software for the reconstruction of elementary particle tracks based on deep learning methods in the processing of experimental information from modern high-energy physics track detectors.

• Software-information support of JINR projects:

- Improvement of information systems for online and offline data processing of experimental facilities of the NICA complex: Database developments for the tasks of experiments BM@N and MPD.

- Software support for the BM@N experiment: Development and implementation of algorithms for modeling, processing and analysis of data for the BM@N track system consisting of gas and semiconductor detectors with microstrip information acquisition (GEM, SILICON, CSC) and their subsequent integration into the BMNRoot.

- Development and refinement of the DQGSM model by comparison with experimental data of the BM@N.

- The BM@N experiment:

(a) trajectory reconstruction of charged particles in outer tracking detectors: MWPC, Si, CSC, DCH, GEM;

(b) Particle identification in the TOF700 time-of-flight detector;

(c) Search for strange hyperons on data from the Nuclotron.

- The MPD experiment: Participation in the realization of a laser calibration system for detector alignment, monitoring drift velocity and for taking electric field distortion into account inside the Time-Projection Chamber (TPC) of the MPD central barrel.

- Further development of statistical methods for the analysis of experimental data under small statistics and incomplete observation of the studied processes: testing the likelihood of hypotheses, estimates of the half-life of the nucleus and of the reaction cross section and of their accuracy, noise filtering and buildup of the optimal experimental design.

- Development of batch processing of neutron diffraction spectra measured in real-time in situ mode (HRFD at IBR-2, FLNP).

• Software information support of large scale outer experiments done with JINR participation:

- Software support of ATLAS experiment: (a) Development and maintenance of configuration and management of ATLAS TDAQ; (b) Development and support of modules of the EventIndex project; (c) Development of modules of Project ATLAS Condition DB (in preparation for RUN3).

- Software support of CMS experiment:

(a) Development, testing and implementation into the official CMS release of algorithms for the separation of overlapping signals and for building track segments in cathode strip chambers (CSC);

(b) Evaluation of the CSC spatial resolution and effectiveness on experimental data with LHC;

(c) Study of the effects of CSC ageing on muon beam tests with radiation source (CERN Gamma Irradiation Facility (GIF++));

(d) Background rates estimation in CSCs with LHC data.

- Software support of CBM experiment:

(a) Development of methods for the selection of rare processes: models, methods, algorithms and software;

(b) Development and maintenance of the complex of databases.

(3) Numerical methods, algorithms and software for multicore and hybrid architectures and Big Data analytics

A. Developments for Multicore and Hybrid Architectures

The mainstream of this activity is aimed at the development of numerical methods, algorithms and software packages developed on the basis of parallel programming techniques using OpenMP, MPI, CUDA/OpenCL, machine learning methods and deep learning (ML/DL), intended to the effective use of multicore and hybrid architectures for the solution of massively parallel, resource-intensive problems of theoretical and experimental physics. The worked out methods and algorithms will take into account the trends in the development of the computational architectures and of the IT-technologies allowing implementing the necessary functionality for different high-performance computing tools and to significantly speed up the solution of a wide range of tasks facing JINR.

To provide opportunities for the development of mathematical models and algorithms as well as for resource-intensive calculations, including graphics accelerators which significantly reduce the computing time, an ecosystem for data analysis and ML/DL tasks was set up and is under active development within the heterogeneous computing platform HybriLIT. The created ecosystem includes two components: the first one is designed for resource-intensive, massively parallel learning tasks of neural networks using NVIDIA graphics accelerators (Fig. 1a); the second is devoted to the design and implementation of models and algorithms based on JupyterHub – a multi-user platform for working with Jupyter Notebook (Fig. 1b).

• The emerging ecosystem is expected to serve to the development of ML/DL based algorithms:

- using recurrent and convolutional neural networks with deep learning for solving problems of fast recognition of multiple tracks in particle physics experiments, including the NICA megaproject and the neutrino program;

- using the neural network approach for tasks of analyzing and classifying medical and biological data.

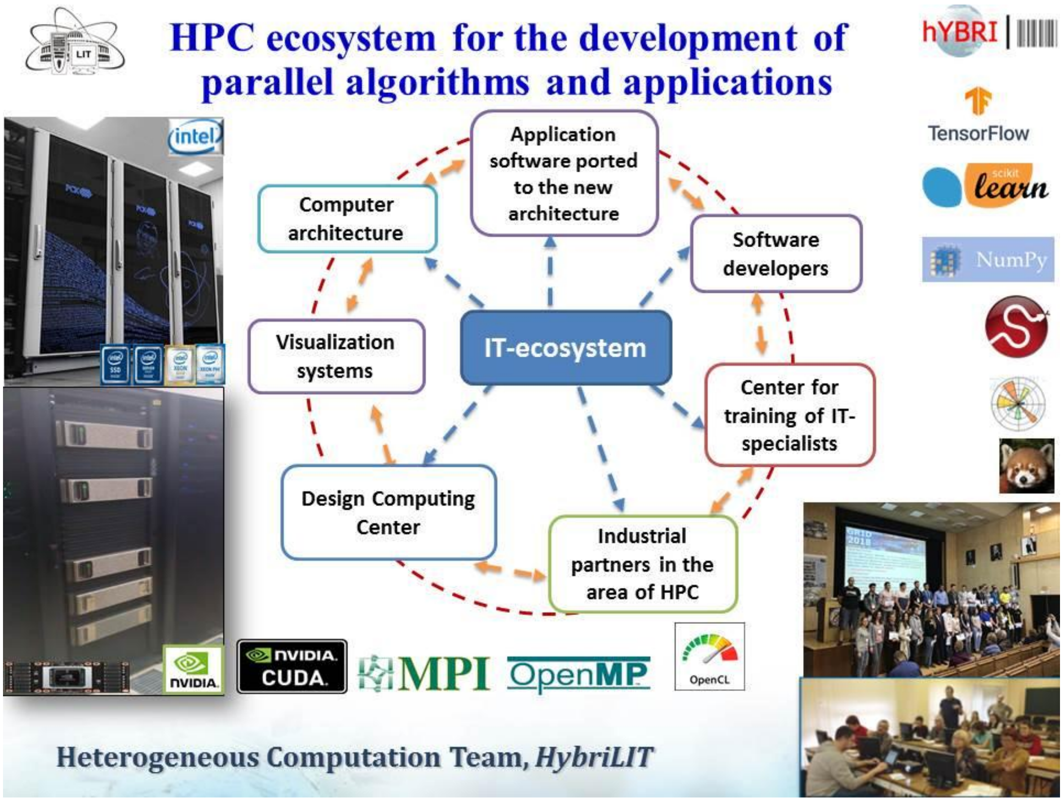

Fig. 1a

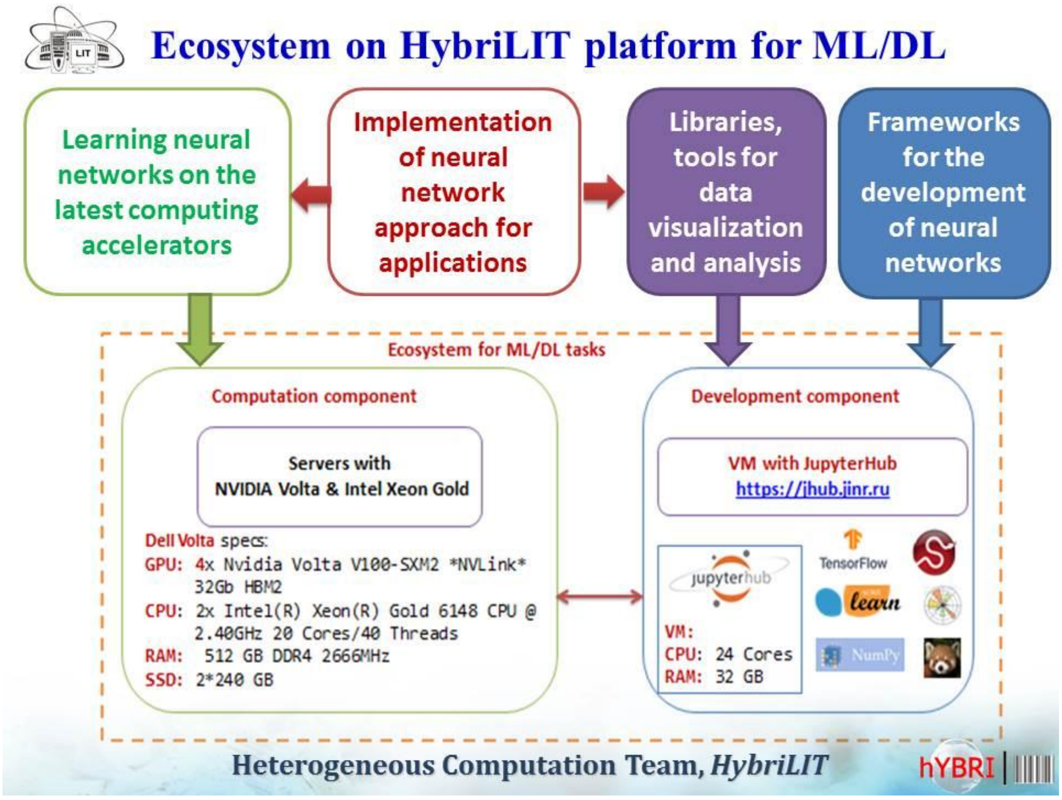

Fig. 1b

• Large scale numerical experiments will be done, within the recently signed Memorandum of Scientific Collaboration (MoSC), for the project ELI-NP, specifically by the implementation and refinement of the PIConGPU code for the laser acceleration of elementary particles.

• Support will be extended to the solution of applied engineering and physics problems on the HybriLIT platform using software packages adapted for parallel computations, e.g., selected COMSOL Multiphysics modules, LAMMPS, GAMESS, etc.

• New parallel implementations of software packages will be developed for solving a number of current and new specific tasks, namely:

- parallelization of the most time-expensive functions using tools such as OpenMP, MPI, CUDA/OpenCL, in the event modeling and reconstruction in NICA experiments;
- calculations on heterogeneous computing platforms for the study of long Josephson junctions;
- the numerical study of multidimensional models based on evolutionary equations;
- solving equations of motion of the molecular dynamics and equations of continuous media;
- calculations of multiple integrals arising from theoretical studies related to the NICA megaproject and entering the study of various physical processes;
- solution of multi-physical problems in accelerator design, dosimetry and radiation safety;
- study of the possibility of increasing the performance of data analysis in ROOT through the use of graphics processors, the comparative analysis of parallelization using PROOF and OpenCL (for the application of the obtained results in the work on the NICA project);
- diagonalization of large matrices in the theory of random matrices for the description of the decay properties of giant multipole resonances of heavy and medium nuclei;
- computations of non-standard problems of the magnetostatics;
- solving optimization problems aimed at getting the best parameters of superdense nuclear matter models in the simulation of heavy ion collisions and in astrophysical applications;
- development of LDPC-like efficient decoding algorithms based on a genetic approach;
- development of finite element method based computational schemes for the description of the quadrupole oscillations of the collective model of the nucleus;
- development of computational schemes for the triangular three-center Coulomb problem and application to the H_3^+ molecule ion and H_2O molecule;
- modeling processes of single and multiple ionization/photoionization of biomolecules by the method of separable potentials;
- study of single ionization of atoms and molecules by fast proton impact in different kinematic regimes.

• The most important developed parallel program packages will be implemented in the JINRLIB library.

B. Big Data Analytics

• Development of the concept and step-by-step implementation within the Big Data approach of a scalable software-analytical platform for the collection, storage, processing, analysis, retrieval of relevant information and visualization of results for the MPD, SPD and BM@N experiments at the NICA accelerator complex and for experiments within the JINR neutrino program.

• Development of methods and software for effective use of Big Data analytics for resource-intensive computations on coprocessors and graphics processors for modeling, reconstruction and processing of experimental data (installations at the NICA accelerator complex, JINR neutrino program experiments, LHC experiments).

• Development of techniques for event reconstruction and intellectual monitoring of detectors using methods and technologies of Big Data and machine learning. The use of powerful high-level Big Data libraries, including the Spark MLlib machine learning library, the GraphX operations library for performing graphs and parallel operations with them, the Spark Streaming library for processing real-time streaming data, and the SparkSQL library. Use of new mathematical methods such as DQC (Dynamic Quantum Clustering), Complex Networks, etc.

• Development of time series analysis methods using Big Data technologies for intelligent monitoring of distributed computing systems. Application of machine learning and artificial intelligence to the optimization of the functionality of the distributed computing for physical experiments.

(4) Methods, algorithms and software of computer algebra and quantum computing

This activity foresees the development of methods of computer algebra and quantum computing for simulation of quantum information processes; creation of algorithms and programs for symbolic-numerical solution of problems arising in experimental and theoretical studies, using the latest computational hardware resources, including the heterogeneous platform HybriLIT.

• Advances in computer algebra methods for modeling quantum systems and quantum information processes:

- Development of methods of quantum information processing by means of dipolar and spinor atomic systems, derivation of optimal methods for spintronics.
- Modeling of intelligent networks performing operations based on quantum decision theory. Numerical analysis of artificial intelligence models.
- Investigation of methods and algorithms for quantum computing and quantum informatics focused on machine learning (quantum machine learning) in order to adapt them to the task of track reconstruction and data processing for the NICA collider.
- Comparison and analysis of the computational efficiency of the adiabatic and "universal" quantum computing on real equipment available through the cloud.
- Development of constructive algorithmic methods based on the computational theory of finite groups for studying the entanglement of many-particle quantum systems.
- Development of quantum algorithms for the study of phase transitions in lattice QCD models at finite density in the framework of tensor network methods.

• Development of computer algebra methods with application to symbolic-numerical solution of differential equations

- Improvement of algorithms and programs for the study and solution of systems of nonlinear algebraic and differential equations arising in the simulation of quantum computing and other physical processes, based on the capabilities of the heterogeneous HybriLIT cluster and the “GOVORUN” supercomputer.

• Development of algorithms requested by special topics

- Development of a new computational scheme for constructing a quasi-probability Wigner distribution for composite systems.
- Development of a relativistic version of the Lindblad equation for open systems based on the Bogolyubov equations for the state vector in relativistic quantum field theory.
- Application of the proposed versions of the quasi-probabilistic Wigner distributions and the relativistic Lindblad equations to the study of composite systems interacting with an intense laser field.
- Development of algebraic methods for the computation of radiation corrections, depending on many kinematic variables and masses, based on functional equations and recurrence relations.

• Development and implementation of special computational tools

- Embedding into the system of computer algebra of the Axiom module of field-theoretical calculations in high-energy physics, which is part of the FORM system – the most efficient among the computer algebra systems in terms of the speed of calculations of this type.

6. **Personnel** (**limited to 1 page**).

The present proposal for the extension of theme 1119 for 2020–2023 explicitly provides the names of the LIT staff and of the main co-workers (group leaders) from other JINR Laboratories, who bring important contributions in the formulation of specific problem to be solved as well as in the interpretation of the obtained results. The strong interaction with the MICC team is emphasized at activity 3.

List of Activities:

	Activity or experiment Laboratory or other Division of JINR	Leaders Main researchers
1.	**Mathematical and computation methods for simulation of complex physical systems**	**Gh. Adam** **J. Buša** **I.V. Puzynin**
	LIT	S. Adam, P.G. Akishin, I.V. Amirkhanov, E.A. Ayrjan, A.S. Ayriyan, I.V. Barashenkov, M.V. Bashashin, A.A. Bogolubskaya, I.L. Bogolubsky, A.M. Chervyakov, N.D. Dikusar, H. Grigorian, Yu.L. Kalinovsky, T.V. Karamysheva, D.S. Kulyabov, K.V. Lukyanov, N.V. Makhaldiani, E.G. Nikonov, K. Oganesyan, T.P. Puzynina, B. Saha, N.R. Sarkar, I. Sarkhadov, Z.A. Sharipov, N.Yu. Shirikova, A.G. Soloviev, T.M. Solovieva, Yu.B. Starchenko, L.A. Siurakshina, Z.K. Tukhliev, A.V. Volokhova, O.O. Voskresenskaya, A. Wojczechowski, R.M. Yamaleev, E.P. Yukalova, E.V. Zemlyanaya, E.I. Zhabitskaya
	VBLHEP	A.Yu. Boytsov, E.E. Donets, S. Gevorkyan, H.G. Khodzhibagiyan
	BLTP	D.B. Blashke, V.V. Braguta, D.E. Castilio Alvares, A.V. Friesen, M. Hnatic, A.S. Hvorostuhin, V.K. Lukyanov, V.D. Toneev, V.V.Voronov, V.I. Yukalov, V.Yu. Yushankhai
	FLNR	A. Oleinichak, Yu.E. Penionzhkevich, R.A. Rymzhanov, V.A. Skuratov
	FLNP	E.B. Askerov, A.V. Belushkin, A.I. Kuklin, A.I. Ivankov, Yu.N. Pepelyshev
	DLNP	L.G. Afanasiev, I.A. Belolaptikov, G.A. Karamysheva, O. Karamyshev, I.N. Kiyan, B.A. Shaibonov, G.D. Shirkov
2.	**Software complexes and mathematical methods for processing and analysis of experimental data**	**P.V. Zrelov** **V.V. Ivanov**
	LIT	E.P. Akishina, E.I. Aleksandrov, I.N. Aleksandrov, D.A. Baranov, M.V. Bashashin, S. Belogurov, O.Yu. Derenovskaya, I.A. Filozova, A.A. Kazakov, A.I. Kazymov, B.F. Kostenko, G.E. Kozlov, L.Yu. Kruglova, S.A. Lebedev, T.I. Mikhailova, M.A. Mineev, G.J. Musulmanbekov, A.V. Nechaevsky, G.A. Ososkov, V.V. Palichik, D.I. Pryakhina, V.S. Rikhvitsky, T.F. Sapozhnikova, V.N. Shigaev, S.K.Slepnyov, A.N. Sosnin, A.V. Uzhinsky, V.V. Uzhinsky, N.N. Voitishin, A.V. Volokhova, A.V. Yakovlev, E.V. Zemlyanaya, E.I. Zhabitskaya, V.B. Zlokazov

VBLHEP	B.V. Batyunya, A.V. Bychkov, A.S. Galoyan, S.R. Gevorkyan, K.V. Gertsenberger, M.N. Kapishin, A.O. Kechechyan, V.P. Ladygin, V. Lenivenko, A.I. Malakhov, S.P. Merts, S.V. Razin, O.V. Rogachevsky, V.N. Zhezher
BLTP	V.D. Toneev
FLNR	A.G. Artukh, B. Erdemchimeg, A.S. Fomichev, Yu.E. Penionzhkevich, Yu.V. Pyatkov, Yu.M. Sereda, Yu.G. Sobolev, Yu.S. Tsyganov, V.K. Utenkov
FLNP	A.M. Balagurov, I.A. Bobrikov, M.A. Kiselev, D.P. Kozlenko, M.V. Frontasyeva
DLNP	I.V. Bednyakov, V.A. Bednyakov, A.G. Olshevsky, L.G. Tkachev, A.S. Zhemchugov
3. Numerical methods, algorithms and software for multicore and hybrid architectures and Big Data analytics	**Gh. Adam** **O. Chuluunbaatar** **O.I. Streltsova** **V.V. Korenkov** **P.V. Zrelov**
LIT	E.I. Aleksandrov, A.S. Ayriyan, E.A. Ayrjan, M.V. Bashashin, S.D. Belov, D.V. Belyakov, J. Buša, Jr., A.M. Chervyakov, O. Chuluunbaatar, I.A. Filozova, A.A. Gusev, I.S. Kadochnikov, M.A. Matveev, I.S. Pelevanyuk, D.V. Podgainy, R.V. Polyakova, L.V. Popkova, T.P. Puzynina, A.A. Sapozhnikov, T.F. Sapozhnikova, N.R. Sarkar, I. Sarkhadov, R.N. Semyonov, S.I. Serdyukova, Z.A. Sharipov, A.G. Soloviev, T.M. Solovieva, Z.K. Tukhliev, N.N. Voytishin, A.V. Volokhova, O.I. Yuldashev, M.B. Yuldasheva, E.V. Zemlyanaya, E.I. Zhabitskaya, M.I. Zuev
LIT-MICC	V.V. Korenkov, V.V. Mitsyn, T.A. Strizh
VBLHEP	K.V. Gertsenberger, A.D. Kovalenko, A.A. Moshkin, E.E. Perepelkin, O.V. Rogachevsky, V.V. Voronyuk,
DLNP	V.A. Bednyakov, D.V. Naumov, A.G. Olshevskiy, O. Samoilov, E.A. Yakushev
FLNR	P.Yu. Apel, S.V. Mitrofanov, V.A. Skuratov
BLTP	P.M. Krassovitskiy, R.G. Nazmitdinov, Yu.V. Popov, Yu.M. Shukrinov, S.I. Vinitsky
FLNP	N. Kučerka, V.V. Novitsky + 3 pers.
4. Methods, algorithms and software of computer algebra and quantum computing	**V.P. Gerdt**
LIT	N. Abbasly, V. Abgaryan, A.A. Bogolubskaya, A.M. Khvedelidze, V.V. Kornyak, E.A. Kotkova, A.M. Raportirenko, I.A. Rogozhin, K.K. Sharma, O.V. Tarasov, A.G. Torosyan, D.A. Yanovich, E.P. Yukalova
BLTP	A.V. Czhizhov, P. Fiziev, A.I. Titov, V.I. Yukalov
VBLHEP	O.V. Rogachevsky
FLNR	B.N. Gikal

7. **Budget estimation (limited to 1 page).**
The core of the LIT investments is related to the MICC project and theme 1118. The theme 1119 secures a convenient superstructure with expenses summarized in the following table.

k$

	Chapters of the JINR budget	Previsions 2020-2023	Yearly distribution			
			2020	2021	2022	2023
1	2	3	4	5		6
1	Salaries	8761.8	1887.9	2076.7	2284.4	2512.8
2	Unified social tax	2646.1	570.1	627.2	689.9	758.9
3	Social national fund	569.5	122.7	135.0	148.5	163.3
4	International cooperation expenses	238.8	59.7	59.7	59.7	59.7
	a) business trips to Member States	60.8	15.2	15.2	15.2	15.2
	b) business trips to Non-Member States	72.0	18.0	18.0	18.0	18.0
	c) business trips in Russia	28.0	7.0	7.0	7.0	7.0
	d) arrivals to JINR	50.0	12.5	12.5	12.5	12.5
	e) Conferences, protocol expenses	28.0	7.0	7.0	7.0	7.0
5	Materials	104.0	20.0	24.0	28.0	32.0
6	Equipment	365.0	80.0	90.0	95.0	100.0
7	Electro energy costs					
8	Heating and water supply costs					
9	Payment of commissioning and adjustment works					
10	Service payments to scientific-research organizations					
11	Scientific and information support					
15	Payment of communication services					
	TOTALS	12685.2	2740.4	3012.6	3305.5	3626.7

8. **Concise SWOT analysis. Each of the items 3–7 should clearly indicate the major strengths and weaknesses of the project/theme for each specific issue.**

• ***Strong sides of theme* 1119**

- With 23 DSc and 47 PhD the theme 1119 staff represents a force of very highly qualified scientific level.
- The interest to get PhD or DSc degree is constantly high. The JINR's right to independently issue academic degrees, which was fixed by the Decree No. 1792-r of August 23, 2017, of the Government of the Russian Federation is, by the beginning of April 2019, in the stage of fixing technical details. The foreseen new LIT Dissertation Council will preserve the highest target level of quality which characterized the previous LIT Dissertation Council.
- The elder staff possesses highest class scientific knowledge acquired by hard work.
- The honest presentation of the opinions is a plus.
- Scientific papers of outstanding quality are published as illustrated at the beginning of Sec. 5.
- A variety of means is supported for the increase of the level education of all the staff, especially young (training courses, tutorials, dedicated conferences and workshops for young scientists and specialists, group and individual consultations, scientific seminars, etc.).

These requirements are of critical importance for the efficient use of the new HPC facilities. Both the HybriLIT cluster and the "GOVORUN" supercomputer raise high demands to their users. They ask, among others, for the grasp of the newest computing paradigms, for the ability to define the optimum computing environment which can result in the highest performance increase under the existing hardware reality. In connection with the HPC computing, we have to notice the hard work done for the creation of an adapted software information environment of new online communication tools allowing fast distant contact and quick exchanges of ideas among the users.

- Of special importance for the theme 1119 community are the periodic LIT *conferences and workshops*.

The International MMCP Conferences (Mathematical Modeling and Computational Physics) have a long history in the LIT. Organized once every two years in cooperation by LIT and research institutions from Slovakia and Romania, they have got a solid international reputation. Three editions deserve special mention:

- MMCP 2011 (72 participants, 41 accepted contributions) – with proceedings published in 2012 at Springer as vol. LNCS 7125, a real bestseller according to the Springer market reports;
- MMCP 2015 (90 participants, 58 accepted contributions) – with proceedings published in 2016 as vol. 108 of the European Physical Journal Web of Conferences (EPJ-WoC);
- MMCP 2017 (257 participants, 100 accepted contributions) – with proceedings published in 2018 as vol. 173 of EPJ-WoC.

The next edition, MMCP 2019, will be organized in Stará Lesná, Slovakia, on July 1–5, 2019. While topics of perennial interest were present at all the MMCP editions, each new MMCP Conference has brought interesting innovations which reflected the new existing worldwide computing landscape. This attracted an increasingly larger number of participants, reaching, in 2017, the ceiling of maximum possible accommodation.

A feature which is to be noticed is the increasing proportion of young scientists present at the MMCP Conferences and at the accompanying satellite events.

- The educational program "Big Data Analytics" within the framework of the "International IT - School on Big Data", organized under the auspices of JINR, will play a very important role in the training of specialists for JINR, including computing for mega-projects (such as the experiments at the NICA Collider). The Big Data Analytics curriculum includes lectures and practical exercises on methods and technologies for processing and analyzing Big Data, distributed computing, and machine learning. The goal of the program is to train specialists in the field of Data Science.

The course of study provides an understanding of the underlying principles and technologies of Big Data, it provides knowledge about the methods and tools asked by the work with Big Data, it teaches how to extract hidden knowledge, to identify patterns, to make predictions. In the frame of the "International IT-School on Big Data", work will be continued on the organization of national and international schools for students, young scientists and specialists.

- ***Weak sides of theme* 1119**

The day-to-day reality points to the occurrence of difficulties in gathering ideal teams:

- The *professional flexibility* of the mature scientists, as a rule, is lower than desired.

The major factor to be blamed is the too rapid pace of expansion of the hardware-software environment. The occurrence of several major changes of the computing paradigms during the last three decades was hard to digest.

The senior scientists did not migrate to the basic today object-oriented programming paradigm. This hinders their work with the latest packages and creates difficulties in the supervision of the work of the young scientists.

- On the other side, the depth of the *initial mathematical training* of the young scientists is relatively low. Often, the present day graduates miss solid knowledge of many chapters of applied mathematics and numerical analysis, together with a basic grasp of the physical phenomena and models. This initial drawback is gradually overcome through hard work on specific problems.

There is, however, an empirical fact which deserves consideration:

- The occurrence of a high *rate of failure* of the attempts of young people to develop successful carriers in computer mathematics and computational physics. Only the strong ones pass the endurance test required by the grasp of the mastership in this field. Often, this difficulty is associated with the *salary gap*: people working in IT business get significantly higher salaries under an intellectual effort which fits their education level. However, the remainder of this churning process is hard characters who like their work and want to become more and more skillful in its exercise.

Collaboration

Country or International Organization	**City**	**Institute or Laboratory**
Armenia	Yerevan	YSU
		IIAP NAS RA
		RAU
		Foundation ANSL
Belarus	Minsk	IM NASB
	Gomel	GSTU
Belgium	Louvain-la-Neuve	UCL
Brazil	Sao Carlos, SP	IFSC USP
Bulgaria	Sofia	IMI BAS
		INRNE BAS
		SU
	Plovdiv	PU
Canada	Toronto	IBM Lab
CERN	Geneva	CERN
China	Hefei	IPP CAS
Czech Republic	Prague	CTU
France	Metz	UL
Georgia	Tbilisi	UG
		TSU
		GTU
		RMI TSU
Germany	Darmstadt	GSI
	Dresden	HZDR
		IFW
	Frankfurt/Main	Univ.
	Freiberg	TUBAF
	Giessen	JLU
	Kassel	Uni Kassel
	Karlsruhe	KIT
	Munich	LMU
	Rostock	Univ.
Greece	Thessaloniki	AUTH
Israel	Tel Aviv	TAU
Italy	Turin	INFN
	Bari	UniBa
	Catania	LNS INFN

Japan	Saitama	SU
Kazakhstan	Almaty	INP
Moldova	Chisinau	IAP ASM
Mongolia	Ulaanbaatar	NUM
		IPT MAS
Poland	Krakow	NINP PAS
	Lublin	UMCS
	Otwock-Swierk	NCBJ
	Warsaw	WUT
	Wroclaw	UW
Romania	Bucharest	IFA
		IFIN-HH
		IFIN-HH – ELI-NP
		ISS
		UB
	Cluj-Napoca	INCDTIM
	Timişoara	UVT
Russia	Moscow	ITEP
		NNRU “MEPhI”
		KIAM RAS
		MSU
		MSRU
		PFUR
		RCC MSU
		GPI RAS
	Moscow, Troitsk	INR RAS
	Dubna	Dubna Univ.
	Gatchina	NRC KI PNPI
	Perm	PSNRU
	Protvino	IHEP
	Puschino	IMPB RAS
		ITEB RAS
		IPR RAS
	Saratov	SSU
	St. Petersburg	SPbSU
		NIIEFA
	Tomsk	TSU
	Tver	TvSU
Slovakia	Kosice	IEP SAS
		TUKE
		PJSU
	Presov	PU
	Banska Bistrica	UMB
South Africa	Cape Town	UCT
	Stellenbosch	SU
Switzerland	Zurich	ETH
Tajikistan	Dushanbe	TNU

		PHTI ASRT
	Khujent	KSU
USA	Cambridge, MA	MIT
	Davis, CA	UCDavis
	Los Angeles, CA	UCLA
	Madison, WI	UW-Madison
	San Diego, CA	UCSD
Vietnam	Hanoi	VNU