Петербургский институт ядерной физики им. Б. П. Константинова
Национального исследовательского центра «Курчатовский институт»

Monitoring of large scale technical systems

Reporter: Andrey Ye. Shevel

The crew:
Sergey **Gorelov**, Alexey **Naikov**, Anatoly **Oreshkin**, Alexey **Shvetsov**, Anatoly **Titov** (young)

Outline

- Large scale technical installations intro
- Monitoring systems
  - monitoring component
  - Monitoring checkup
  - developments
  - examples
  - Artificial Neural Networks approach in monitoring

(Super) Large scale technical systems/installations features

- **Complexity**
  - Many hundreds or hundreds of thousands or even more network of Interconnected components: nodes, processes, users.
    - Geographically distributed all the above stuff.
    - Complicated interactions between components of installations and involved employee/organization and users.
- **Monitoring system**
  - Complicated architecture;
  - Generate vast amount of data including huge volume of monitoring data, i.e. Big Data
- **Scalability**
  - The number and type of components usually increasing during the life time.
- **Reliability**
  - The ability to continue the functioning despite adversity.

Monitoring examples taken into account

- There are many examples of monitoring in HEP (pls see Reference slide [1-2])
- Own experience:
  - Monitoring experimental HEP computing infrastructure at home and abroad.
  - Monitoring large scale geographically distributed (many hundreds of Km) quantum communication line
    - http://government.ru/news/42449/
    - https://www.1tv.ru/news/2021-06-08/407818-mezhdu_moskvoy_i_sankt_peterburgom_proveli_pervyy_seans_videokonferentssvyazi_po_kvantovoy_magistrali.

Monitoring system

- In any large scale installation the monitoring system is the main window to watch what is going on.
- The data in monitoring are coming from sensors.
  - All the sensors deliver the generated data into monitoring computer network.
    - The computer network architecture does matter.
  - The data from sensors are in fact the projections of the internal processes in the large installation.

Sensors

- The data about the large scale installation is gathered by a number of sensors/detectors (may be by tens of thousands or/and millions).
  - Which type of sensors are required (what has to be measured)?
  - Which a number of sensors/parameters is enough?
  - Which a sensor's precision is enough?
- Sensor data is often represented as time series data, which is a sequence of data points collected and saved in a database.

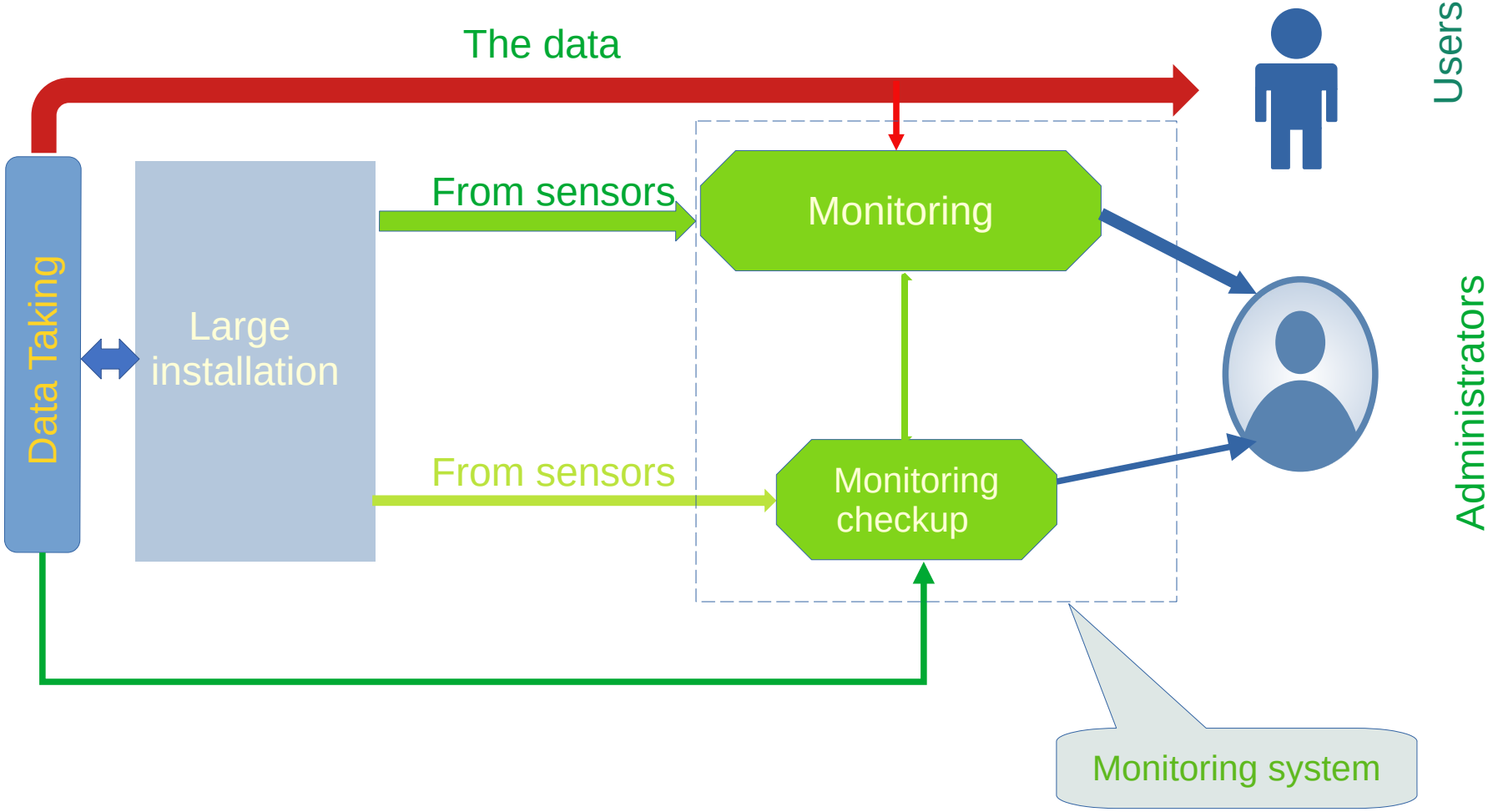
Monitoring network main features

- The monitoring network architecture must include:
  - Enough network capacity (transfer speed).
  - Reliability.
  - Storage to store all the gathered data.
  - Visualization facilities.

Monitoring system components

- Taking into account the reliability the monitoring system must have pair independent components:
  - The monitoring component has to present overall the functioning of the large installation.
  - The monitoring checkup has to monitor the health of the monitoring component itself and partly monitored large installation.
- In huge technical installation is possible to consider many mentioned pairs.

Monitoring system and around



Monitoring system development

- **Monitoring component** is developed on the base of the *existing monitoring frameworks* (monitoring program systems).
  - Most labor and time consuming part of the development is the *configuration* to tailor the monitoring architecture to the concrete requirements.
- **Monitoring checkup** is developed on the base of the monitored installation requirements and depends on monitoring system architecture as well.

Existing Monitoring frameworks

- There are hundreds existing monitoring systems frameworks (free of charge and proprietary). Here are popular examples:
  - Nagios.org
  - Zabbix.com
  - Prometheus.io
  - Many others (there is a number of surveys on the topic).
- The process of choice for monitoring for specific large installation is quite sensitive.

Monitoring checkup

- The monitoring checkup depends on the architecture of the monitoring and monitored large installation. It is often developed from scratch. That is because large installation monitoring consists of many specific sensors and other active components.
- The monitoring checkup has to
  - show current status of the monitoring system;
  - send alarms to administrators when the monitoring system is out of order;
  - store all the data about the health of the monitoring system itself and large installation components important for monitoring.

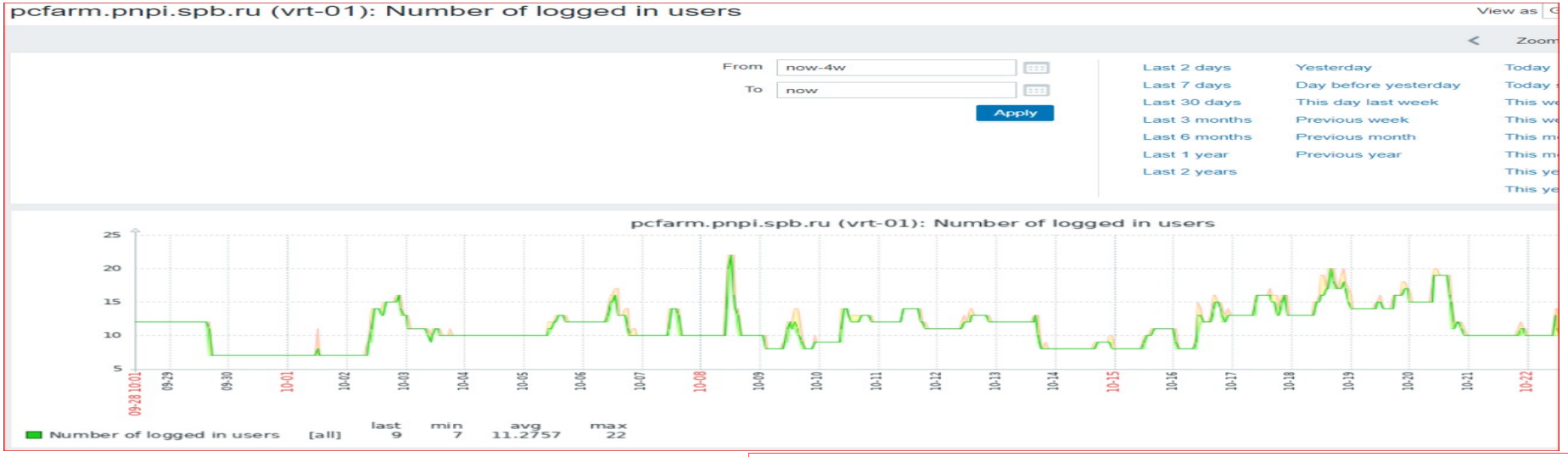
The monitoring data analysis

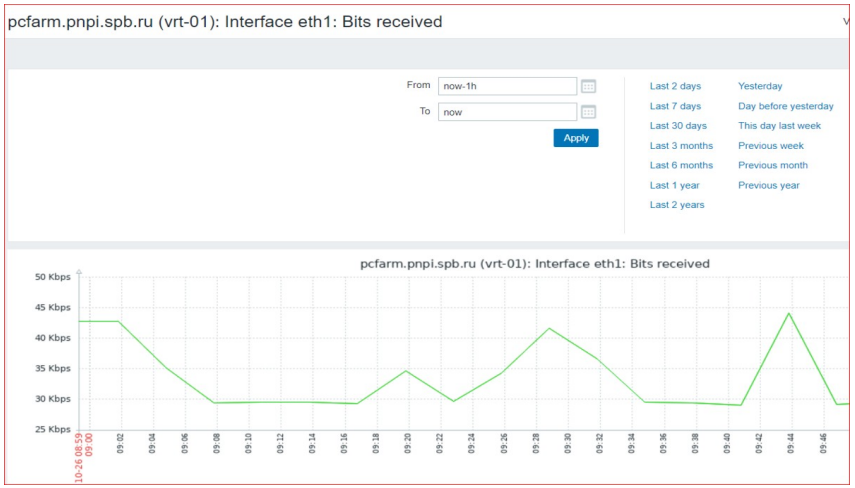
- Analyze the system logs of monitoring system itself.
- Analyze the specific period of time and/or specific parameter and/or correlation between parameters of monitoring system itself.
- Discover anomaly among the gathered monitoring data.
  - It is done in various ways including methods of artificial intelligence.

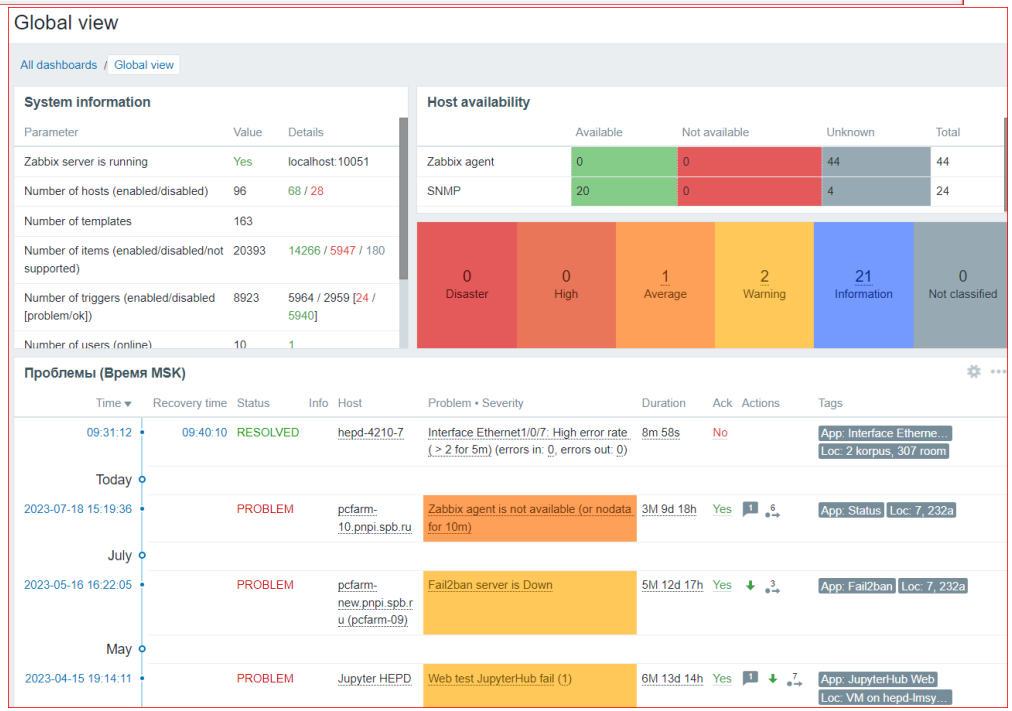
The monitoring system at HEPD

- Current example is based on the real monitoring for HEPD computing infrastructure (free of charge framework Zabbix).
  - Zabbix server
  - Zabbix sensors
- Zabbix permits to gather and show monitoring data for large computer networks
  - Gather of the sensor data with monitoring agents.
  - Database and analysis toolkit.
  - Possibility to react on specific events (perform script).
  - Visualization.
  - Sending alarms to the administrator.
- The monitoring checkup in this example is not large: it consists of just several test scripts with sending alarm to administrator in case of problems.

Examples of the pictures from Zabbix



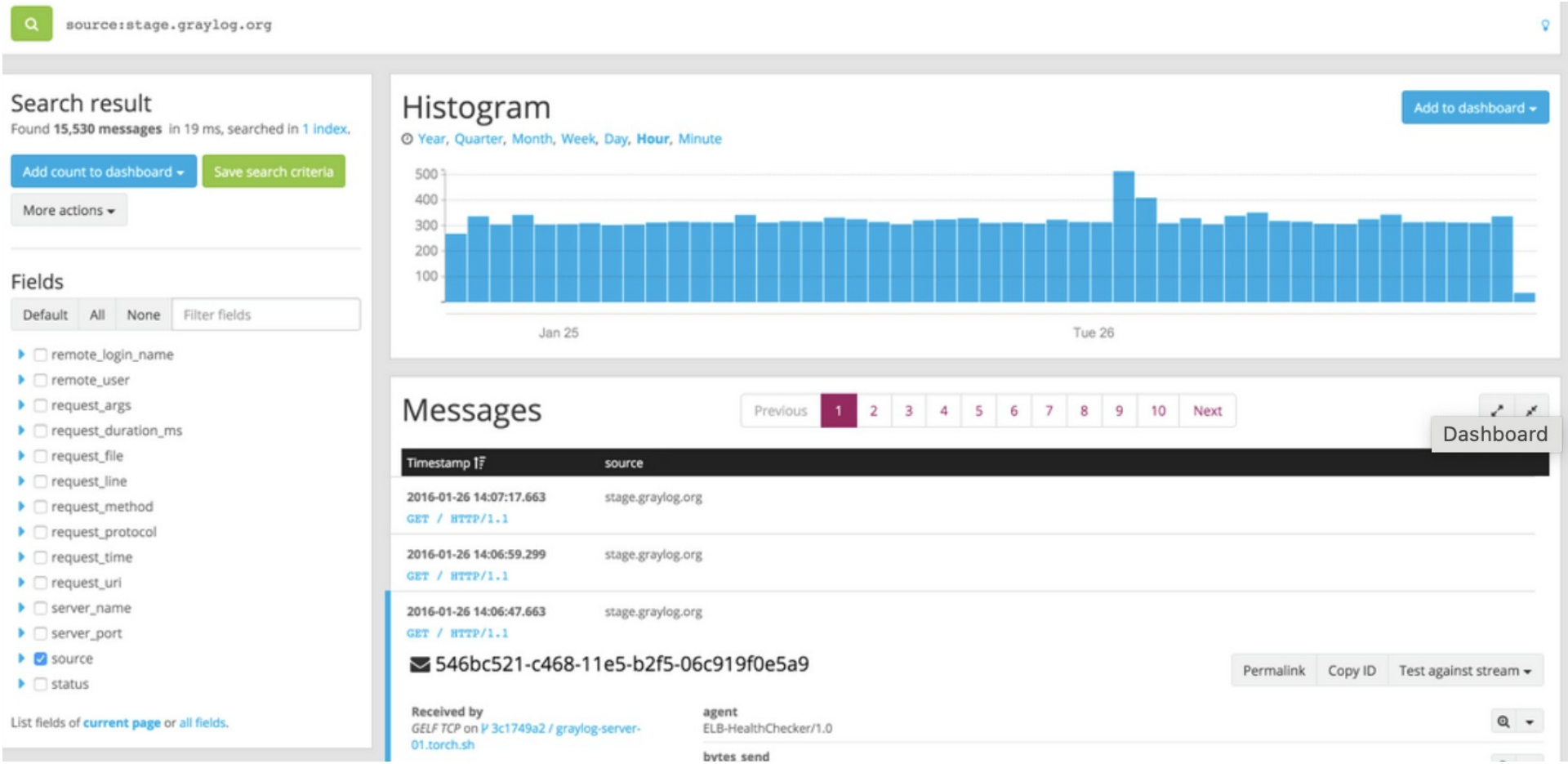




Advanced monitoring system

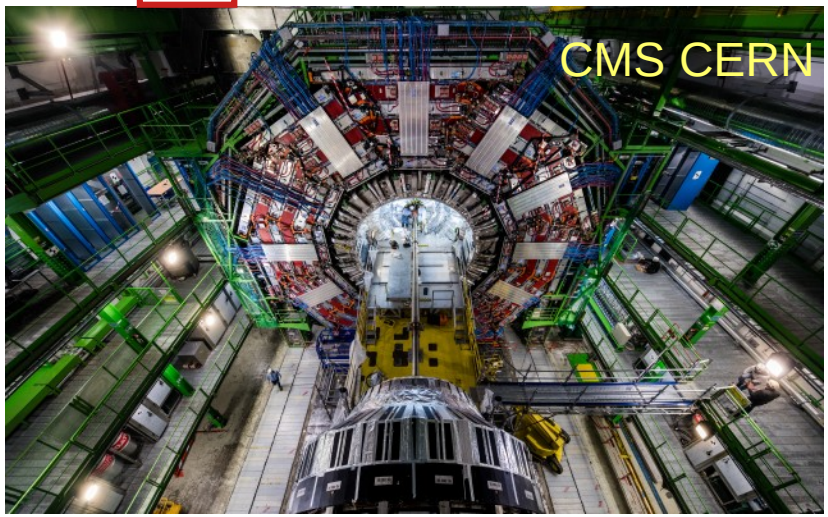
- Graylog - the Graylog software centrally captures, stores, and enables real-time search and log analysis against terabytes of machine data from any component in the IT infrastructure and applications. The software uses a three-tier architecture and scalable storage based on network gathering toolkit **OpenSearch** and database **MongoDB**.
  - Three tier architecture: gathering, logic, presentation.
- Those all are for technical administrators.

Example view for Graylog pages



Technical administrator’s roles

Physics and monitoring data



Manuals, logbooks guides, recommendations

Administrator roles in large scale installations

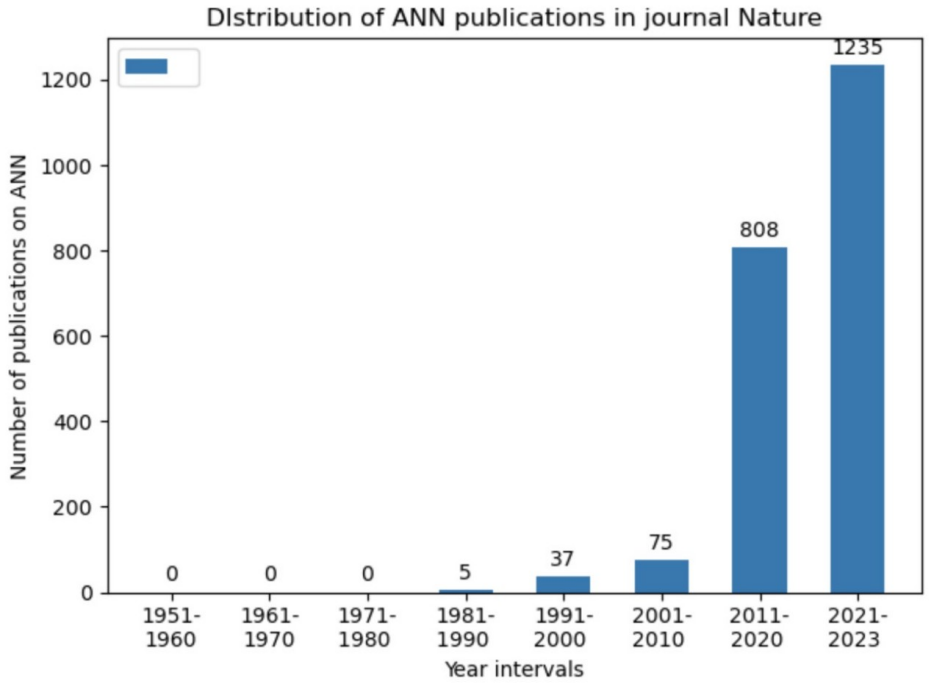
- The monitoring system for large technical installations is quite large and complicated, i.e. administrator's activity is unavoidable.
- All that means that system administrators who are responsible for whole installations must take into account huge volume of the data:
  - The installation status including logs;
  - The monitoring system status including logs;
  - Many pages of manuals, instructions, orders, recommendations, LogBooks, etc.
  - Quite often shifted administrator has just initial knowledge on the administration.
- In the past the team of experts spend enough time to study how the installation is functioning and got the common experience how to understand and react for system events. Team members did and do assist each others.
  - Even after that the understanding what is happened with large installation might require the time and a lot of labor.
- It is not bad idea to automate the assistance and experience gathering for administrators.

Automate assistance for administrators

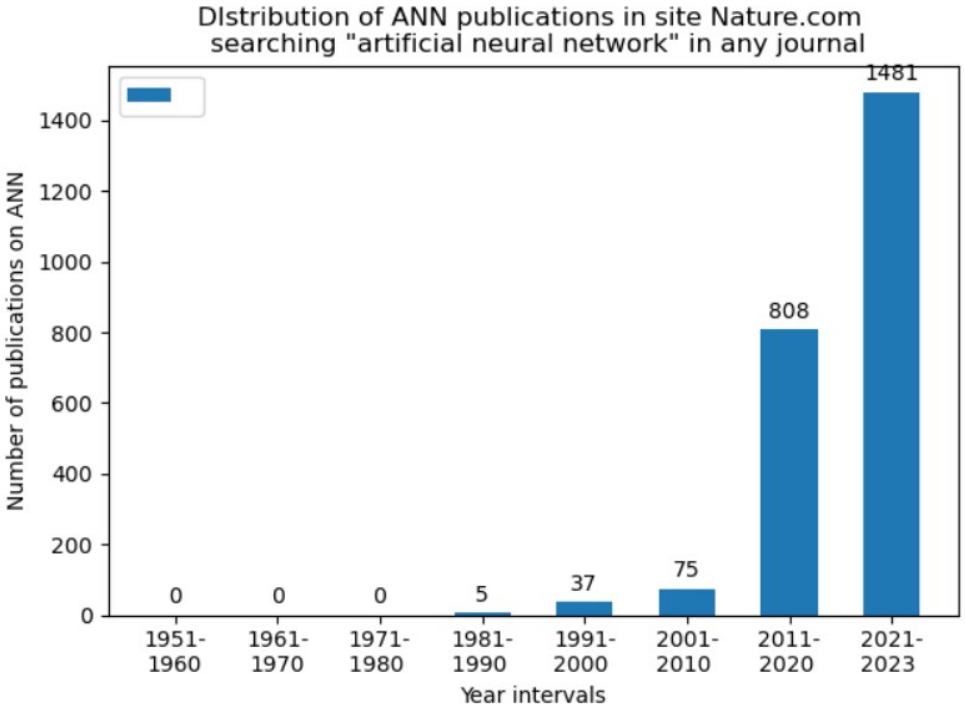
- Great progress last 10 or so years in artificial intelligence (AI) and in particular in the administrator's assistance for large system monitoring might be implemented with Artificial Neural Networks [the survey in Ref 4] approach to valuable reduce the time and labor requirements.
  - One of the approach is to develop the chat-bot to assist the administrator (to assist is NOT do all job instead the administrator in EACH case).
- There are hundreds publicly available AI assistance in the Internet
  - There are limitations to use existing trained public chat-bots in the Internet due to security reasons.
  - That means the chat-bot has to be created and trained with local documents, logs, other data and be integral part of the installation. Also we need to take into account vernacular technical language used to describe the installation.
    - The example of such the effort in the past is available at Reference slides [3] (2012).
    - Future HEP experiments related docs/papers are in References slides [5-7] (2022-2023).

The Interest score to Artificial Neural Networks

2023-07-05



2023-11-06



The functioning example of ANN assistant

- The AI assistant (i.e. chat-bot) is able to talk with the administrator in form of question-answer (QA) manner. Question: from administrator, Answer: from chat-bot. E.g.
  - Q1: why channel 24 has many errors?
  - A1: due to available data it is probably because malfunction of block 98 in rack 69. Instruction how to check that block is available at URL ...
    - Q1.2: I did not see the block 98 in that place. Why it is happen and who can help in the situation?
    - A1.2: Most probably the block 98 is in reparation area. The responsible person Noname Noname phone xxxxxxxx, email: noname_NN@yyy.yy.xx.
  - Q2: Where to find logs from the detector ZZZZ for last two days?
  - A2: Please go to the URL …, enter username and password, click button LOGS, in drop down menu chose required detector and dates.
  - Q3: Please send alarm when the sensor … generate the data …
  - A3: The alarm for the sensor … condition … has been set.

ANN assistant development and limitations

- Training process for the AI chat-bot is very sensitive:
  - Preparation of the input data (remove wrong and/or duplicated data) – data quality.
  - Volume of the input data (the more the better).
    - Let us remind fundamental rule: garbage in input will give garbage in output.
      - Compare with real world: converting waster into valuable material is not possible without preliminary sorting and cleaning.
- Assistant AI tool needs to be maintained during life time of the large installation.
- Finally
  - AI toolkit might do faster and with less involved experts than the usual human team only in logical inferences on the base of available data.
  - In the case where the available expert team is unable to solve the problem – do not expect the AI is able to solve it better.

Possible areas of application for assisting ANN chat-bots

- Coming large experimental setup in HEP experiments.
- Future or renewed particle accelerator.
- Future or renewed Nuclear Reactor.
- Future large computing/communication system.
- Conclusion
  - We need to find a way to implement the multi user computing platform to create technical possibilities for the chat-bots development for a number of tasks.

Minimum configuration for creating ANN chat-bots

- Possible minimum start configuration for example LLaMA family of appropriate scale
  - 1xCPU with up to date throughput and other parameters;
  - RAM is around 128 GB or more;
  - One or more NVIDIA GeForce RTX 3090 (24GB VRAM) or NVIDIA A100 (40GB VRAM);
  - Available disk storage is about 10TB.
    - Price estimation for hardware is around 2M Rubles.
- Software toolkits start examples
  - Ray - ray.io, https://github.com/ray-project/ray
  - Rasa - https://github.com/RasaHQ/rasa
- Possible start technology: Retrieval Augmented Generation.

Which might be estimations

- The development of the plan for concrete assistant platform (experimental setup and/or computer/communication system) depends on interests and sources of financial support.
- The possible timescale and man/power estimation:
  - ~24 months would be required for initial production version.
  - ~72 men/month would be required.
  - ~3-4 students might be involved.

Thank you!

Spare slides are following

References-1

1. C. Chavez-Barajas et al // DQM4HEP - A Generic Online Monitor for Particle Physics Experiments // https://cds.cern.ch/record/2296574/files/arXiv:1801.10414.pdf
2. Miguel Rubio-Roy et al 2017 J. Phys.: Conf. Ser. 898 032009 // https://inspirehep.net/files/6a3ed0750fda8df5ae2fcf6d96c9ac97
3. Artificial intelligence in the service of system Administrators // C Haen et al 2012 J. Phys.: Conf. Ser. 396 052038
4. Quick survey for ANN // https:///indico.jinr.ru/event/3505/contributions/21545/

References-2

5 Department of EnergyAnnounces $10 Million for Artificial Intelligence Researchfor High Energy Physics (March 2022)

- ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR AUTONOMOUS OPTIMIZATION AND CONTROL OF ACCELERATORS AND DETECTORS // FUNDING OPPORTUNITY ANNOUNCEMENT (FOA) NUMBER: DE-FOA-0002875 FOA TYPE: INITIAL // CFDA NUMBER: 81.049 FOA Issue Date: November 9, 2022 // Submission Deadline for Applications: January 11, 2023, at 11:59 PM Eastern Time

6 Artificial Intelligence for the Electron Ion Collider (AI4EIC) (more than 80 authors from 63 universities and labs) // arXiv:2307.08593v1 [physics.acc-ph] 17 Jul 2023

7 Department of Energy Announces $16 Million for Research on Artificial Intelligence and Machine Learning (AI/ML) for Nuclear Physics Accelerators and Detectors // AUGUST 17, 2023. Today, the U.S. Department of Energy (DOE) announced $16 million for fifteen projects that will implement artificial intelligence methods to accelerate scientific discovery in nuclear physics research. https://www.energy.gov/science/articles/department-energy-announces-16-million-research-artificial-intelligence-and

Ultra-large-scale system (ULSS)

- Ultra-large-scale system (ULSS) is a term used in fields including Computer Science, Software Engineering and Systems Engineering to refer to software intensive systems with unprecedented amounts of hardware, lines of source code, numbers of users, and volumes of data. The scale of these systems gives rise to many problems: they will be developed and used by many stakeholders across multiple organizations, often with conflicting purposes and needs; they will be constructed from heterogeneous parts with complex dependencies and emergent properties; they will be continuously evolving; and software, hardware and human failures will be the norm, not the exception. The term 'ultra-large-scale system' was introduced by Northrop and others to describe challenges facing the United States Department of Defense. https://en.wikipedia.org/wiki/Ultra-large-scale_systems

Definitions

- Britannica Dictionary definition of INTELLECT
  - the ability to think in a logical way
- Britannica Dictionary definition of INTELLIGENCE
  - the ability to learn or understand things or to deal with new or difficult situations
  - secret information that a government collects about an enemy or possible enemy
- Britannica Dictionary definition of ARTIFICIAL INTELLIGENCE
  - an area of computer science that deals with giving machines the ability to seem like they have human intelligence

ANN for large experimental setups

1. The problem
    1. A large volume of documentation (technical descriptions, administrative orders, operating manuals, etc), as well as a volume of logs (automatic and semi-automatic log records) about the functioning of the entire system.
    2. A meaningful analysis (obtaining an answer to a specific question based on all available data) of such a large amount of data (hundreds of GB or more) is a non-trivial task, which in many cases turns out to be labor- and time-consuming.
2. Possible development
    1. It seems reasonable to undertake the development of a special Artificial Expert Assistant using ANN technology, which could provide the operator (system administrator) with effective assistance in the described task.
3. Possible implementation
    1. The using of local computing facilities with appropriate model with regular re-training to take into account new log records.