



Contribution ID: 21

Type: **Presentation long (25 mins)**

Next-generation Green Computing

Wednesday, 20 October 2021 11:15 (25 minutes)

Green Computing is dedicated to the reduction of energy and material consumption in information and communication technologies. Although serious energy-saving measures are put in place already today, they are being overtaken by accelerated advances of digitization.

The steady rise of the Internet of Things (IoT) results in a drastic grow of the number of sensors. In the same way, the resolution of measuring devices continues to increase. As a result, not only industrial and scientific domains but also everyday life have to cope with ever-increasing data volumes that need to be processed and, thereby, lead to dramatic swells of the energy consumption. As published in Nature [1], the worldwide power consumption due to IT needs in 2030 will be approx. 20%.

In our contribution, we will emphasize that substantial data reduction is indispensable to cope with future power demands. Sooner or later, everyone comes to the point that almost no raw data can be stored in the long term anymore, but only a comparatively tiny fraction of “relevant information” that needs to be extracted automatically from huge data streams by applying suitable machine learning methods.

The observatory Square Kilometre Array (SKA) has to overcome all these challenges - its thousands of antennas will produce more data than the world-wide internet (see our last year's contribution [2]). “Traditional” data compression is not sufficient, but it is a matter of extracting “relevant information” already during data acquisition in near-realtime. Due to the time constraint, information loss is inevitable and not reversible. To minimize the resulting “data irreversibility”, we propose a “Dynamic Life Cycle” (DLC) that extends existing Big Data life cycle models by introducing two feedback loops: one between the sensors and a nearby computing center, and one between world-wide distributed data archives and the sensors - both of which constantly optimize the sensors' control systems.

DLC may contribute significantly to sustainability in compute ecosystems. Realizing DLC, however, is technically extremely demanding. Its core, the data reduction processes, is to be described in detail by metadata to ensure the FAIR principles and, ultimately, reproducibility of scientific results in view of non-existent raw data. Hereby, the metadata can be larger in volume than the archived data itself, since the permanently changing parameters of the sensors' controls and the states of the workflows for extracting the “relevant information” must be constantly recorded. The understanding how data irreversibility affects data reduction changes over time, meaning that the quality of archives must be steadily monitored (by comparison with simulation data). Searching huge data for “rare events” or “unknown signals” needs an efficient massively parallel computing, which is in its infancy in image processing and machine learning.

Our presentation addresses the impact of DLC on Green Computing. New ways of working together are indicated for a machine learning-based “next-generation Green Computing”.

[1] N. Jones: How to stop data centres from gobbling up the world's electricity, Nature 561, 163 (2018).

[2] H. Heßling, M. Kramer, S. Wagner: Data Challenges at the Square Kilometre Array (SKA), EGI Conference 2020.

Speaker bio:

Hermann Hessling studied Physics at the Universities of Münster, Goettingen, and Hamburg. He received the Ph.D. (Dr. rer. nat.) in Theoretical Physics and was appointed a postdoctoral research fellow at Deutsches Elektronen-Synchrotron (DESY) Hamburg (1993-1996). Since 2000 he has been Professor of Applied Informatics at the University of Applied Sciences (HTW) Berlin. His scientific interests include distributed high-performance computing and, in particular, extracting knowledge out of large-scale data in real-time.

Most suitable track

Envisioning the future

By submitting my abstract, I agree that my personal data is being stored in accordance to conference Privacy Policy

Primary authors: HESSLING, Hermann (Verein für datenintensive Radioastronomie (VdR), and University of Applied Sciences (HTW) Berlin); Prof. KRAMER, Michael (Verein für datenintensive Radioastronomie (VdR), and Max Planck Institute for Radio Astronomy (MPIfR) Bonn); Prof. WAGNER, Stefan (Verein für datenintensive Radioastronomie (VdR), and Zentrum für Astronomie Heidelberg (ZAH))

Presenter: HESSLING, Hermann (Verein für datenintensive Radioastronomie (VdR), and University of Applied Sciences (HTW) Berlin)

Session Classification: Envisioning the Future - Presentations