Contribution ID: **89**   Type: **Group assignment**

Group assignment for Track 1: Technologies and Platforms

Monday, 14 June 2021 14:50 (2h 10m)

The goal of this exercise is to provoke you into thinking about some of the key choices in computing.

The scenario
Modern scientific experiments are massive producers of data. Imagine that you're a computing manager for one such experiment, which produces 100 terabits of raw data per second and has no computing infrastructure yet. Your task is to use your current knowledge to conceptualize data processing for your experiment and, in the process, to uncover important choices to make.

The challenge
Focus on key aspects of compute and software, and less so on networks, accelerators or data flows. What kind of considerations, tradeoffs and assumptions would you have to take into account?

What kind of equipment would you use, where would you put it and why? What kind of software would you run? What do you think would be the rough purchase and maintenance cost and effort? Can you identify gaps in your current knowledge that you would need to fill in?

What we expect
You're not expected to have all the answers! In many cases already listing the important questions can be helpful. Seasoned professionals can spend even 10 years of their careers making such a plan for a single experiment.

Try to answer the challenge in conceptual terms, and using rough estimates. When faced with unknowns, you can make assumptions –make sure to clearly specify when that's the case. It's best if you present your solution on the basis of a 1-slide diagram illustrating key concepts and components, but it's not a requirement.

Summary

Primary author: NOWAK, Andrzej

Track Classification: Track 1: Technologies and Platforms