

# WLCG-HSF Systems Performance and Cost Modeling Working Group

HSF-WLCG Workshop two weeks later

Markus Schulz (CERN)

# Assumption

- Most of the people attending the GDB have been at the workshop
- A very brief orientation and summary should be sufficient
- The GDB is probably a good place to:
  - to clear up some confusion (on what we try to do)
  - discuss what we need to do in addition to the current plan



# Orientation

- WLCB MB endorsed the creation of the group by 14 November 2017
  - **Joint WLCG HSF working group**
  - 35 active members → [wlcg-SystemPerformanceModeling@cern.ch](mailto:wlcg-SystemPerformanceModeling@cern.ch)
  - Workload, workflow, framework developers, people who plan, engineer and operate IT sites
- February GDB Pepe Flix gave an update on our activities
  - [https://indico.cern.ch/event/651350/contributions/2878070/attachments/1600371/2537084/08022018\\_Cost\\_Model\\_GDB\\_JFlix.pdf](https://indico.cern.ch/event/651350/contributions/2878070/attachments/1600371/2537084/08022018_Cost_Model_GDB_JFlix.pdf)
- Joint HSF-WLCG Workshop 2 sessions
  - Wednesday 1:30 and 2:00 long
  - First giving an overview, second more detail oriented
  - Summaries:
    - <https://indico.cern.ch/event/658060/contributions/2907221/attachments/1624803/2586970/summary.pdf>
    - <https://indico.cern.ch/event/658060/contributions/2940581/attachments/1625234/2587855/Summary-CostPerformanceModeling.pdf>

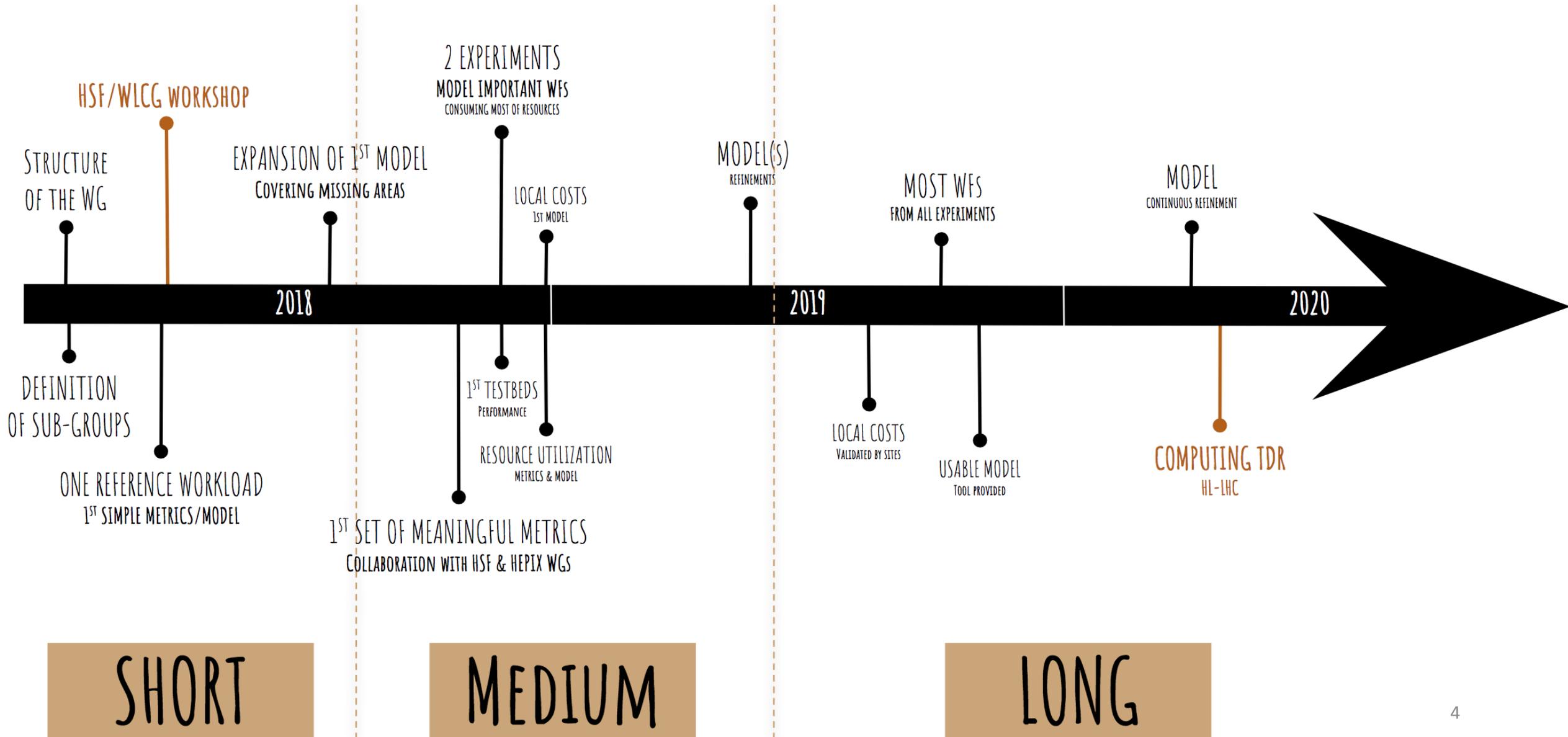
# Main Activities:

RED : what we want

BLACK: what we need to do to get RED

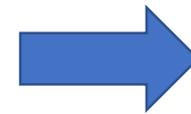
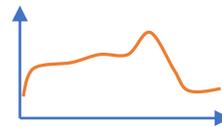
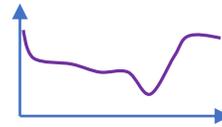
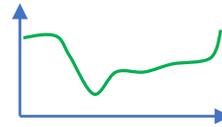
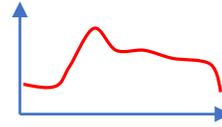
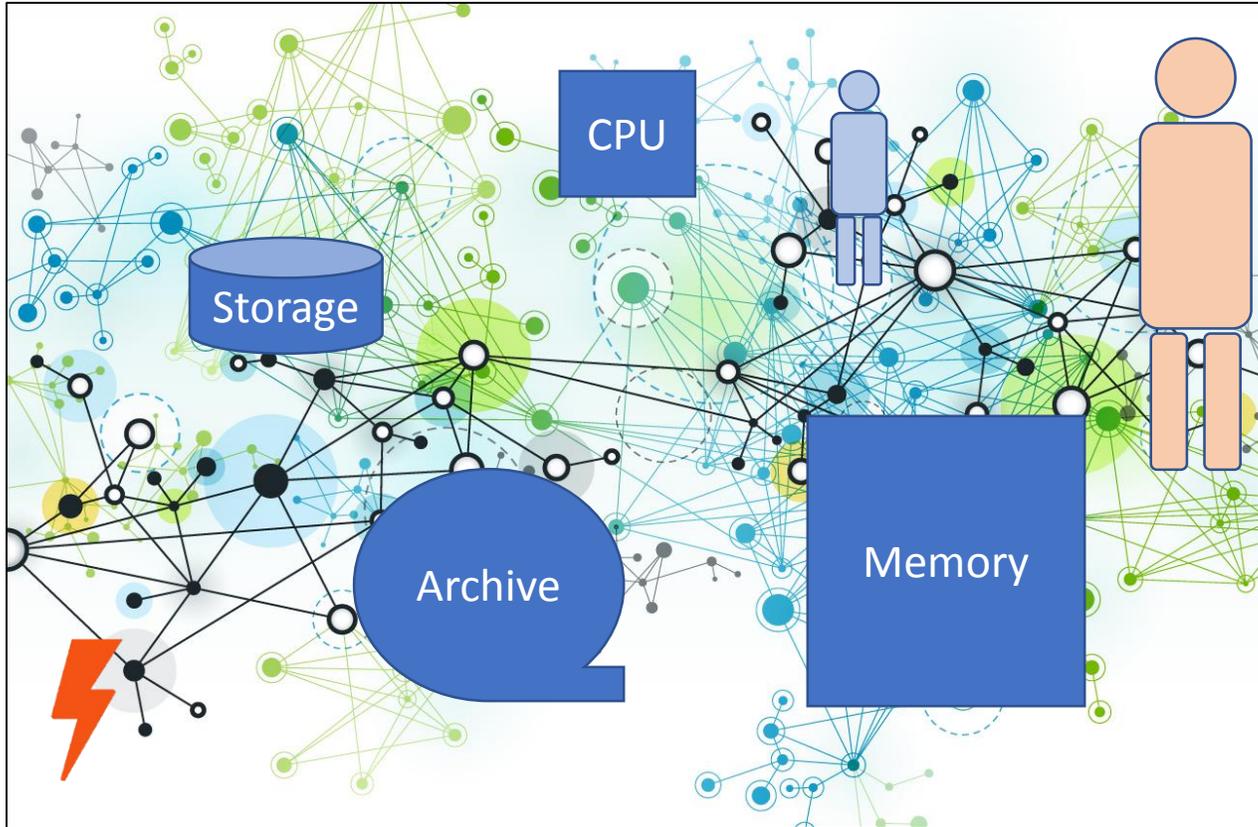
- Revision of most important workloads for each experiment
  - To create a set of reference workloads accessible to study
  - Common taxonomy
- Packaged versions of the most important workloads
  - To run them easily
- **Definition of properties that best characterise a workload (Metrics)**
  - And measure these
- **Draft a cost evaluation process**
  - Mapping Metrics of workloads to fabrics to local costs
  - Creating a feedback loop to optimize the amount of Physics we can do within budget
- Compile a list of relevant performance analysis tools
  - And make them usable for the average developer/site admin
- Set up a distributed testbed to run tests
  - For laboratory condition measurements
- **Start a simple generic resource calculation model**
  - From spreadsheet to code, transparent

# Roadmap [preliminary]



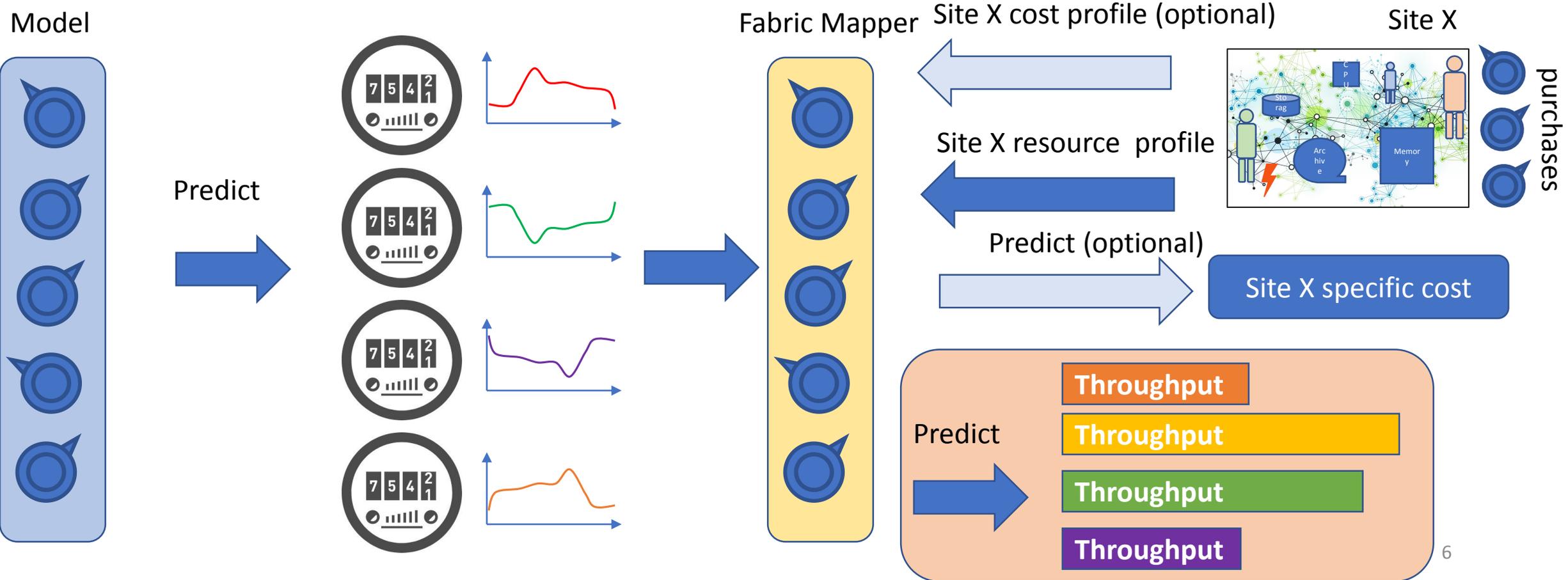
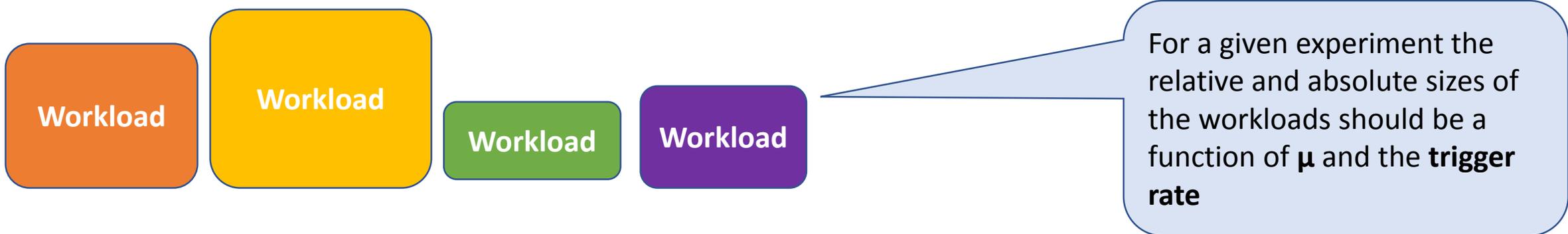


What we intend do 1)



Model





# Status

- **All activities have shown considerable progress**
- Some are now “straightforward” (but still require lots of work)
  - Revision of most important workloads for each experiment
  - Packaged versions of the most important workloads
- Defining metrics moved from definition to building a potential framework for measuring them
  - *prmon*
  - Stimulating discussions on how gaps could be filled
    - Especially communication from within the workload to external tools
      - Common workload taxonomy ?
    - Role of batch systems

# Status

- **Mapping of resource profiles to local costs**
  - Exercise with 4 sites showed that there are massive differences what sites consider as cost.
  - Some confusions on goal of a cost model for sites
    - What is it for? ← not to compare sites!
    - Understanding how many events/year a infrastructure can process given the characteristics of workloads and sites. Guiding purchases and developments.
      - → feedback loop between computing models/workloads/fabrics
  - Importance of awareness of complete costs
    - Power infrastructure, racks etc.....
    - Impact of increased complexity on human effort
- **A first version of a spreadsheet for calculating costs exists**
  - Renaud Vernet (IN2P3)
  - Already useful for sites

# Generic resource calculation model

- **Good starting point**
  - Working from Ken Bloom's Python framework
- More parameters from the application metric needs to be added
  - Tuning knobs needed
- **Networks need to added**
- Should evolve into a tool that can answer "What-If" questions to explore the future
- **Current focus: understand what we do today**

# Identified (big) Gaps

- **Currently we deal with existing workloads and infrastructures**
  - Things that exist can be measured relatively easily
- **For HL-LHC neither the workloads nor the infrastructures are well defined**
  - But several ideas exist how they might look
    - New techniques: Machine learning, accelerators, .....
    - Scale of the computing needs
    - Models: Data Lake, Analysis Farms .....
  - How to model these ideas ?
  - Some first ideas have been discussed
    - Microkernels representing the new approaches
    - Working with the demonstrators/prototypes ...

# Next (obvious) Steps

- **Continuation of the current activities**
  - They are all on a good path
- Start an activity to explore how we can model HL-LHC conditions
  - Volunteers welcome
- Link the performance measurement tools assessment with HSF activities
  - One catalogue and one cook book !
- See how we can profit from activities like the NICA computer centre model
  - Based on discrete event simulation
- **Recruit more active participants**

Are we on a clear path to a performance and cost model?

# Many open questions!

- **How to describe impact of complexity?**
  - Human effort, complex systems are less reliable.....
- **Current approach to Metrics: Measure all that can be measured !**
  - Ideal: orthogonal, minimal set of quantities that characterise the workloads
  - Have to carefully watch which quantities are used for building the models
  - How can we turn time series into compact parameters?
    - That can be understood intuitively
- **Methodical approach to active probing**
  - Throughput dependency on resource restrictions
    - memory, network, storage latency
      - Running in y GB we loose x%
      - Reading remote data with bandwidth z latency k we reduce throughput by x%.....

# More more Gaps

- **How to use the model to identify limiting factors/inefficiencies**
  - Tuning workloads and fabrics .....
- **HL-LHC: How to model impact of new computing / storage models**
  - Can we use current data popularity data to make quantitative statements on disk/tape impact?
  - .....
- **How to model scheduling inefficiencies? ← not small!**
  - Batch, workflow management etc. ,
- **How to assess overall efficiency?**
  - Which fraction of the capabilities of our systems do we exploit?
    - Operations/Cycle etc. , disk occupancy and activity ....
  - Does this make sense? Or is processed events per time the only metric that matters?

# Final words.....

- Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning.
  - Winston Churchill