

**If I knew then what I know now, I would have never
designed it this way: Approaches to CI
conceptualization and design**

Frank Würthwein
Director, San Diego Supercomputer Center

January 18th 2023





Focus on time invariant characteristics of your science



- The LHC program is a multi-generational global enterprise
 - 200++ institutions across 40++ countries for 50++ years
- The accelerator provides $O(1e11)$ mutually independent beam crossings per year
 - The compute & storage per collision must be very small to make project affordable
- Data produced is immutable
 - There exists no use case for read-modify-write



Multi-generational Global Enterprise



- Optimize architecture for in-kind contributions
- Optimize architecture for openness & expandability such that people can live anywhere in the world and still participate meaningfully
- Today's friends may be tomorrow's enemies
 - Iran, Russia, Ukraine, Israel, China, Lebanon, Taiwan, Saudi Arabia, India, Pakistan, ... are all engaged in LHC program
- Try to avoid hiding fundamentals behind project specific tools because your work force is transient
 - Students should learn something useful for their future careers instead tools so specific that they are meaningless in the larger world

-
- This is a science that is fundamentally high throughput computing
 - It is parallelizable by its very nature
 - There is nothing here that requires low latency multi-node communications
 - Multi-threading: yes, multi-node: no

-
- CMS produces official datasets for the collaboration that support FAIR principles, if you restrict its meaning to the ~3k collaboration members.
 - Data is written locally, and committed to archive after a file is closed, when it becomes immutable.
- => We can implement distributed data infrastructures that are void of cache coherence, and thus a priori scalable.
- => Massively simplifies “global data infrastructure”.**

-
- We all are influenced by fashions of our trade
 - Try to remind yourself regularly of the basic time invariant characteristics of your scientific enterprise to avoid chasing fashions unnecessarily.
 - Initial designs as well as change should be grounded in the time invariant characteristics of your scientific enterprise.
 - Define metrics grounded in the science, and decide between differing visions/options based on that.
 - Large projects have large budgets and thus have a tendency to produce bespoke interfaces and tools.
 - Doing so is a disservice to their workforce