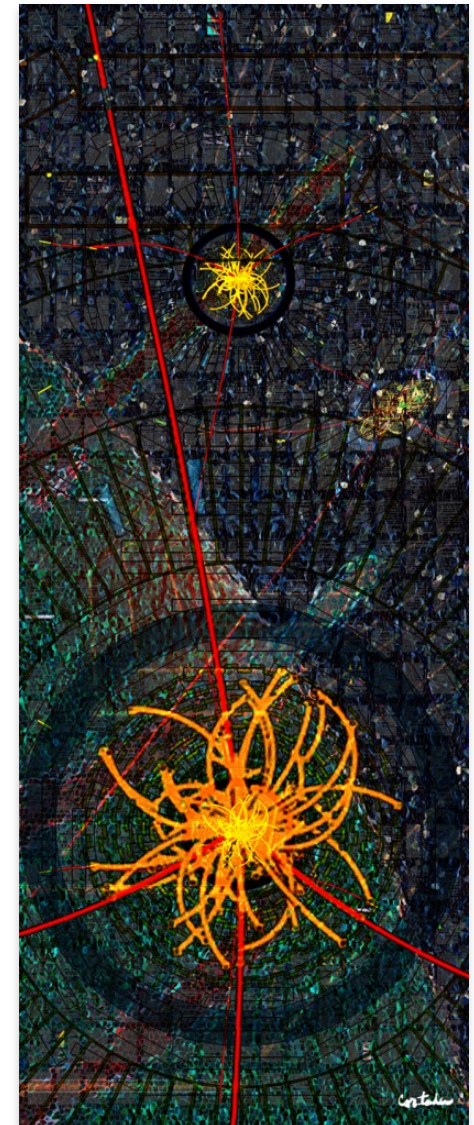
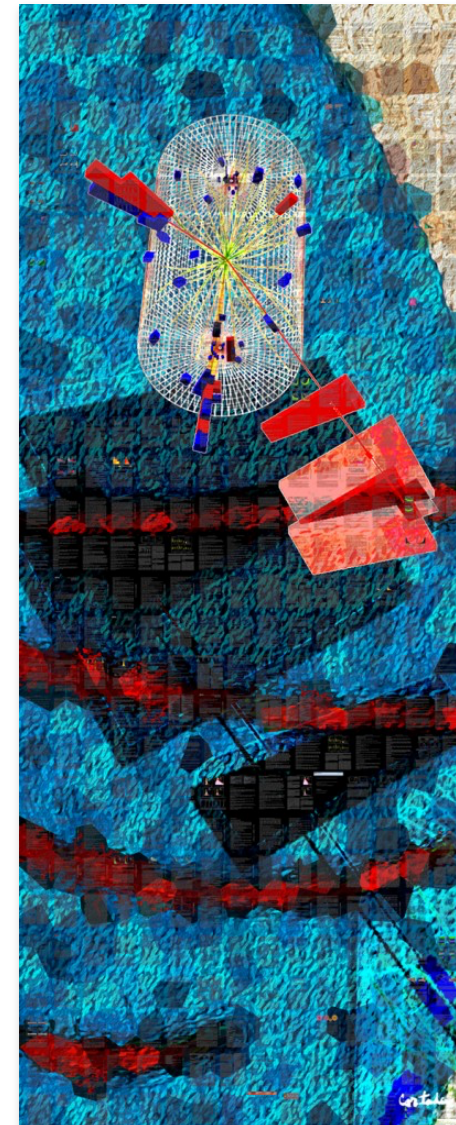
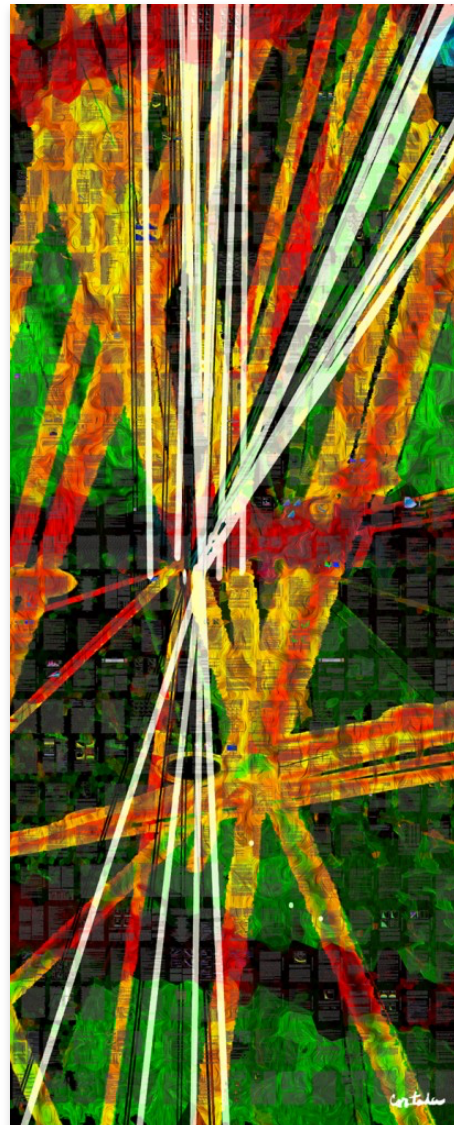
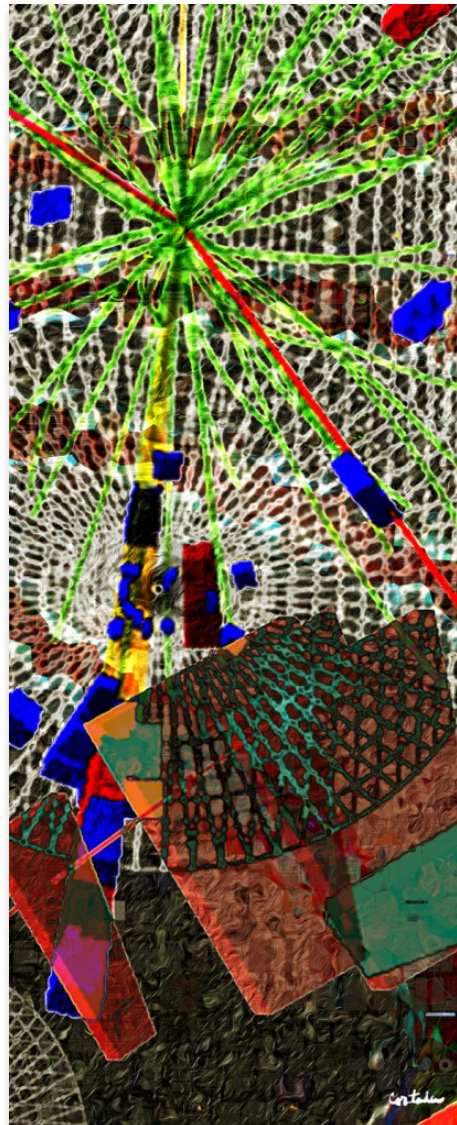
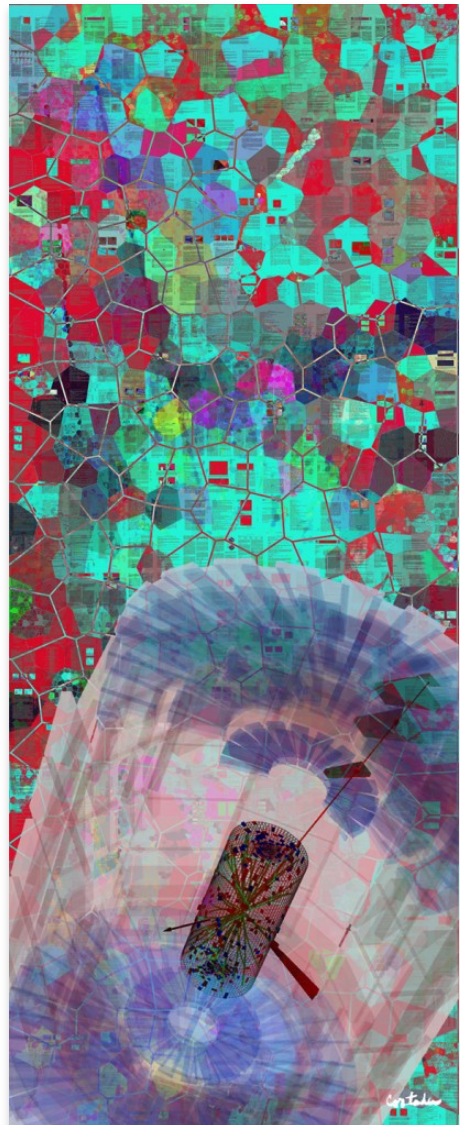




# Analysis Facilities – US CMS Ops

Brian Bockelman

Co-lead for HL-LHC R&D





# Why does Ops care?

- Why does the USCMS Operations program care about future analysis systems?
  - In Run 1 & 2, analysis systems were largely considered out-of-scope. CMS S&C helped users create their ntuples on the grid – but that’s where analysis begins, not ends!
  - USCMS provided *infrastructure* (such as CMSLPC or CMS Connect) to use for analysis but not necessarily *services*. Users are expected to write their own batch scripts and interact with the storage – with assistance from the staff.
- Why is HL-LHC different?
  - **Necessity:** if events go up by 2 orders magnitude, it may become impossible to get competitive analyses done on the hodgepodge of scripts & services used today.
  - **Opportunity:** We believe some of the new techniques may make analysis simpler for users (costing less in personnel time) or faster – overall making the collaboration more competitive.
  - **Sustainability:** By using larger, non-HEP specific ecosystems we look forward to reducing the burden to only the HEP-specific pieces.
    - ★ Assume fewer people will do analysis with more data to comb through.

# Thoughts on ~~analysis~~ facilities

- First and foremost, we already have analysis facilities.
  - What we are discussing is how we evolve facilities: what new services do we add and resources do we manage to meet HL-LHC's needs?
  - We have a large central facility, CMSLPC at FNAL.
  - T2s also serve as facilities for analysis: there's available disk space, CPU and, at some institutions, interactive login clusters.
- In fact, there's little likelihood we have a brand-new facility for analysis.
  - What's possible is some specialized infrastructure at an existing facility (or two).
  - That is, possible models considered can be a Kubernetes "add-on" or allow a site to specialize for high-IO needs.
    - ★ Specialization, of course, takes awhile: it might take 5 years to replace the majority of the storage subsystem.
  - What's fairly unlikely is "add a new site for analysis-only".
- Accordingly, whatever we do must largely overlay on top of today's facilities.
  - This may include evolving today's facilities!
- We do not see significant funding to run on the cloud at this time.
  - Can always change! Currently, hardware growth is only seen via HPC.
  - We are sympathetic to the 'professor has \$10k to run on GCP' but this does not drive the operations program's decision making process.



- Analysis for CMS is internal competition of ideas.
  - that means that an analysis infrastructure has to allow/enable everyone to execute their ideas (with reasonable restrictions).
- ‘NANOAOD’ is seen as the key to sustainability of analysis in HL-LHC.
  - This format is very small & lightweight (2KB/event) and relatively simple (no heavyweight C++ objects or requiring CMS libraries).
  - Even in HL-LHC, we can reasonably have a complete dataset in NANOAOD format at a single site.
    - ★ Having a way to augment a local copy of NANOAOD with researcher-derived objects from MINIAOD is seen as desirable.
      - Worldwide, we can *probably* host a single copy of MINIAOD on disk.
      - AOD will probably never be usable to general CMS researchers and will require close coordination.
  - Without NANOAOD, we are doubtful that we can enable the competition of ideas.



- I'm not here to say what services are needed to do analysis.
  - I'm not a physicist!
- Largely, I think of these as falling in a few categories:
  - Services we (Ops, facilities) run:
    - ★ Existing examples: batch system, storage, HTTP caches.
    - ★ Potential new services (implementation):
      - notebook/browser interface (JupyterHub),
      - Task services (Dask-as-a-service),
      - fitting service (???)
      - column services (SkyHook DM).
    - ★ It costs a *lot* of time / effort (i.e., \$\$\$) to mature a service.
    - ★ Yes, new infrastructure techniques (Kubernetes) can decrease the cost to deploy and operate these. Does not decrease to zero.
  - Providing the ability to run microservices:
    - ★ Users can “saddle up” to the site and run a particular service.
    - ★ I think we tend to massively overestimate user's ability to do this.
      - If they can run a microservice, they are probably on the registrant list for this workshop.



- What properties are we looking for in our services?
  - Multi-user. Our scope is, at the minimum, all of USCMS.
  - Integrates / leverages the rest of the infrastructure. Examples:
    - ★ Do not BYO authz infrastructure.
    - ★ Do not ignore the fact that 90% of the experiment's computing and storage resources are external to any individual US facility.
  - (Insert your favorite list of good development habits here ... no need to repeat).
- We do not expect any service to be delivered 100% complete. Hence the need for R&D!
  - However, we expect a roadmap to production.
- My personal observations of missing pieces from yesterday:
  - System-oriented discussion: many discussions about individual pieces. How can we make more pieces fit together? Do we have a “Coffee vertical” and a “hep-tables vertical”.
  - How data gets into the analysis system from the wider infrastructure.
  - How to move community-building to the forefront.



- We are growing an analysis ecosystem for HL-LHC.
  - In a healthy ecosystem, any individual component is replaceable.
  - In fact, over a sufficiently long period of time, all components are replaced.
- Each analysis system integrates pieces together from the ecosystem into a coherent story.
  - We do not need to have a single analysis system
  - But neither do we have the effort to do a bespoke one for each user. What's the right level of granularity?
- An ecosystem must be sustainable.
  - The operations program can invest strategically but cannot shoulder the cost of everything.
  - HEP does not have the purview or talent to do everything by itself. There's a reason why no HEP site develops its own batch system in-house (anymore).
    - ★ The less generic pieces we work on the more time is left for physics.
  - Part of R&D is selecting for survival and leaving other things behind.
    - ★ You see some of this in the evolution of python histogram libraries...