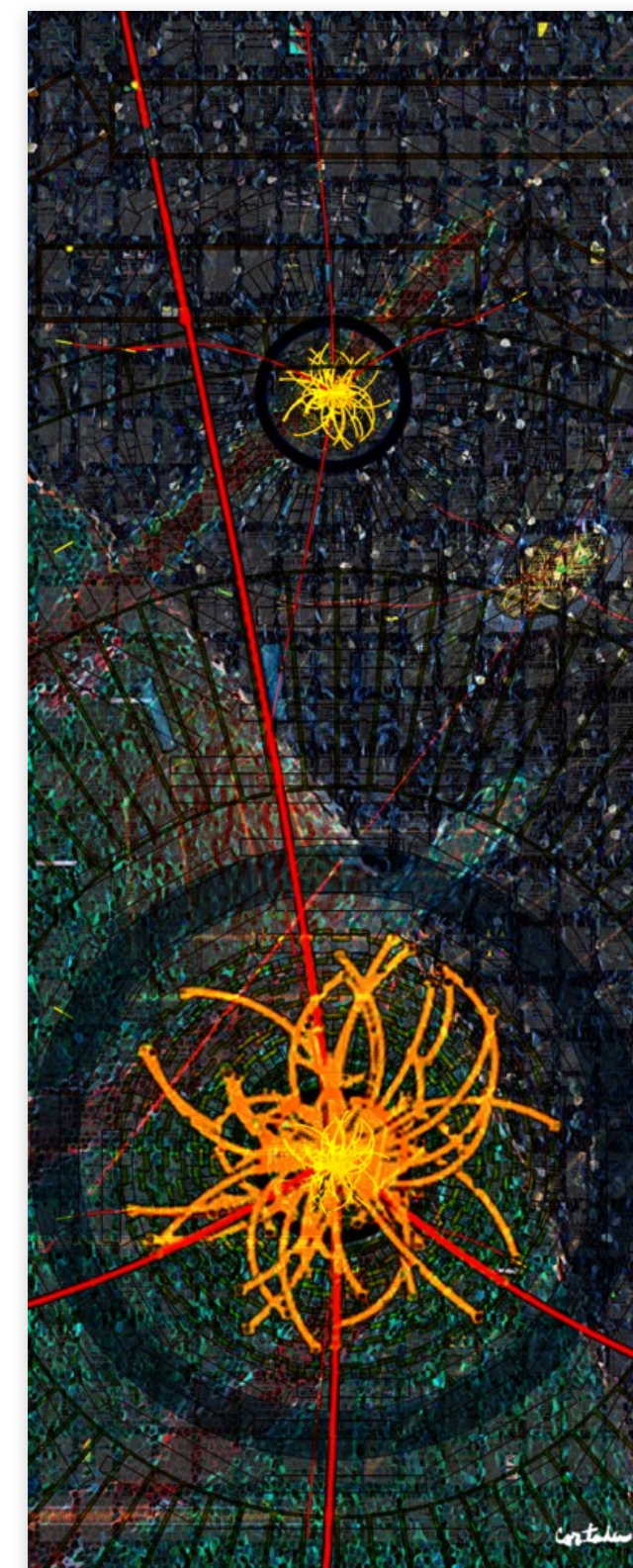
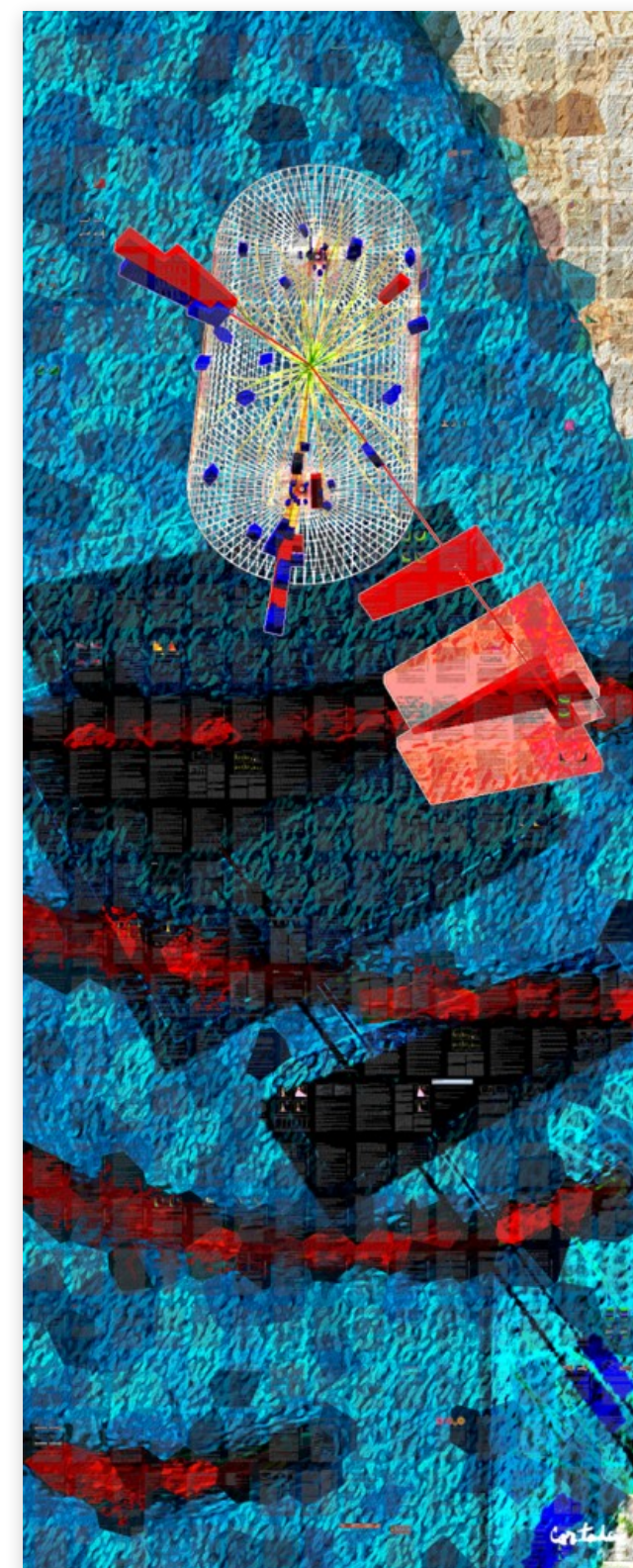
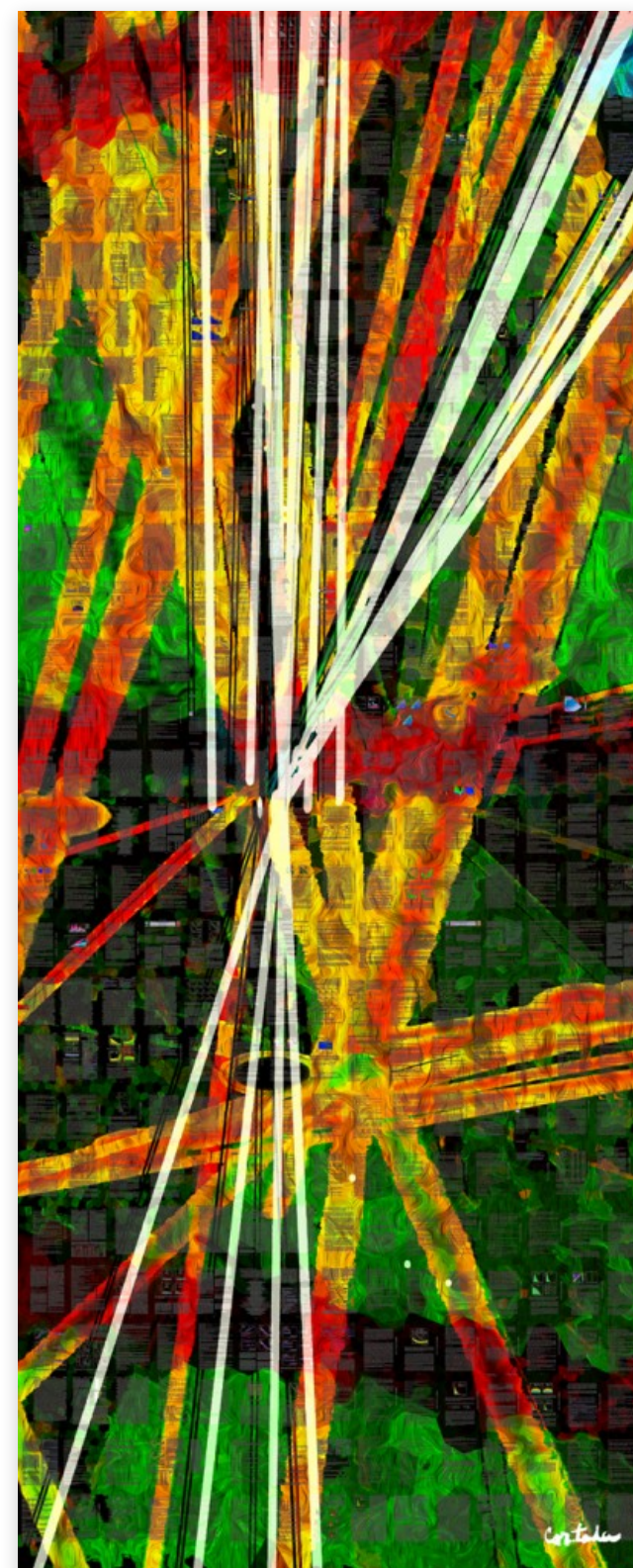
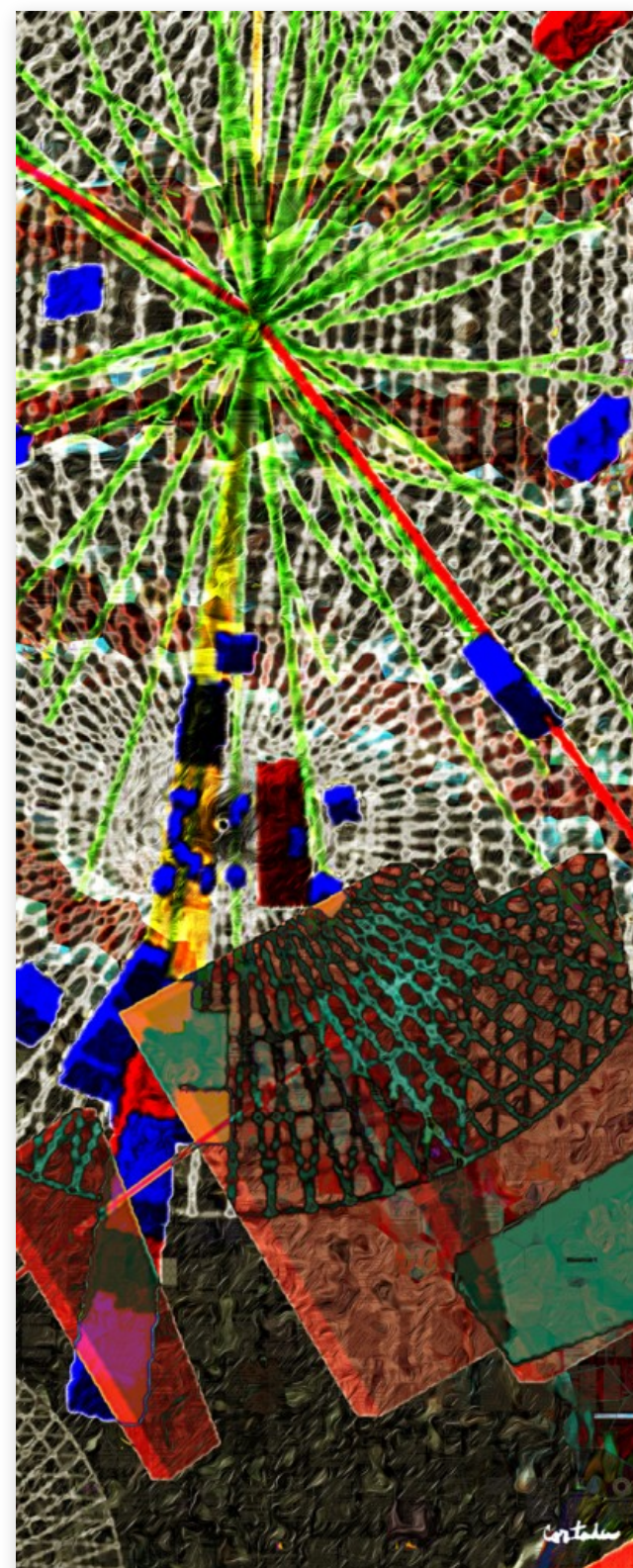
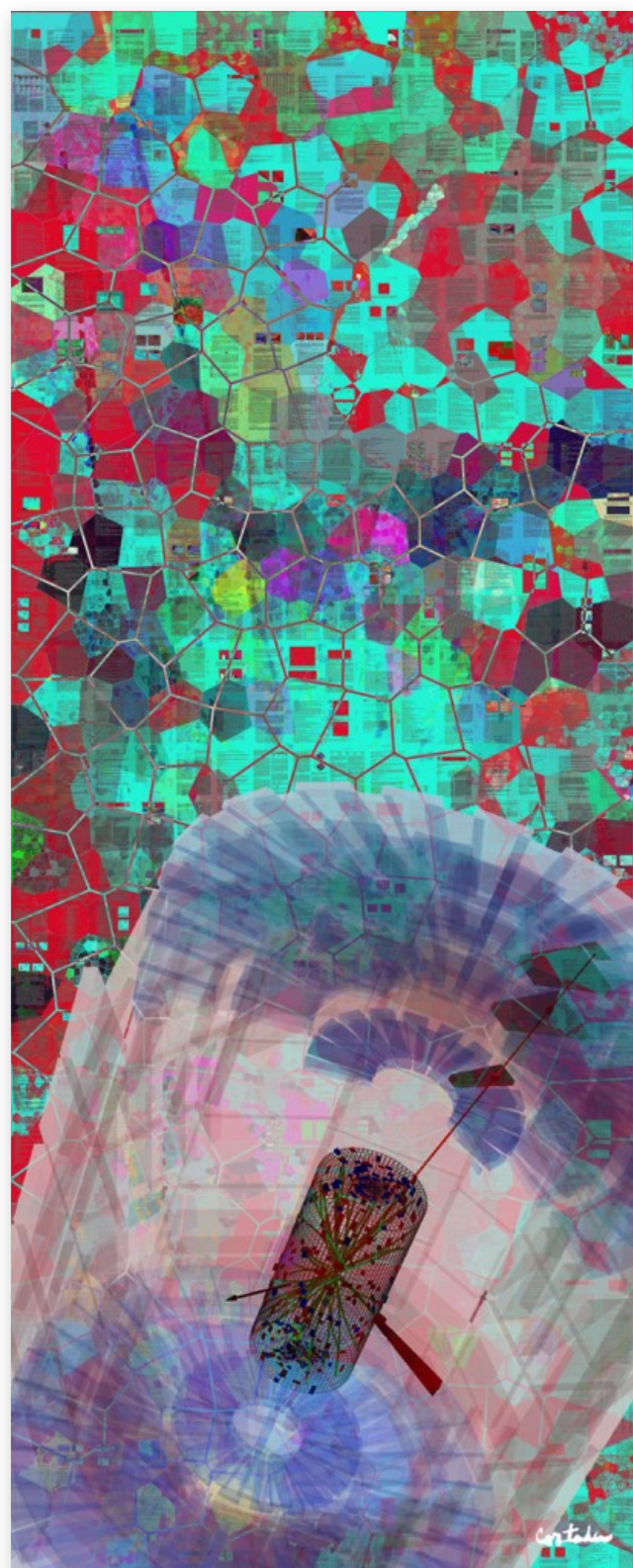


CMS Rucio Experience

Eric Vaandering — Fermilab

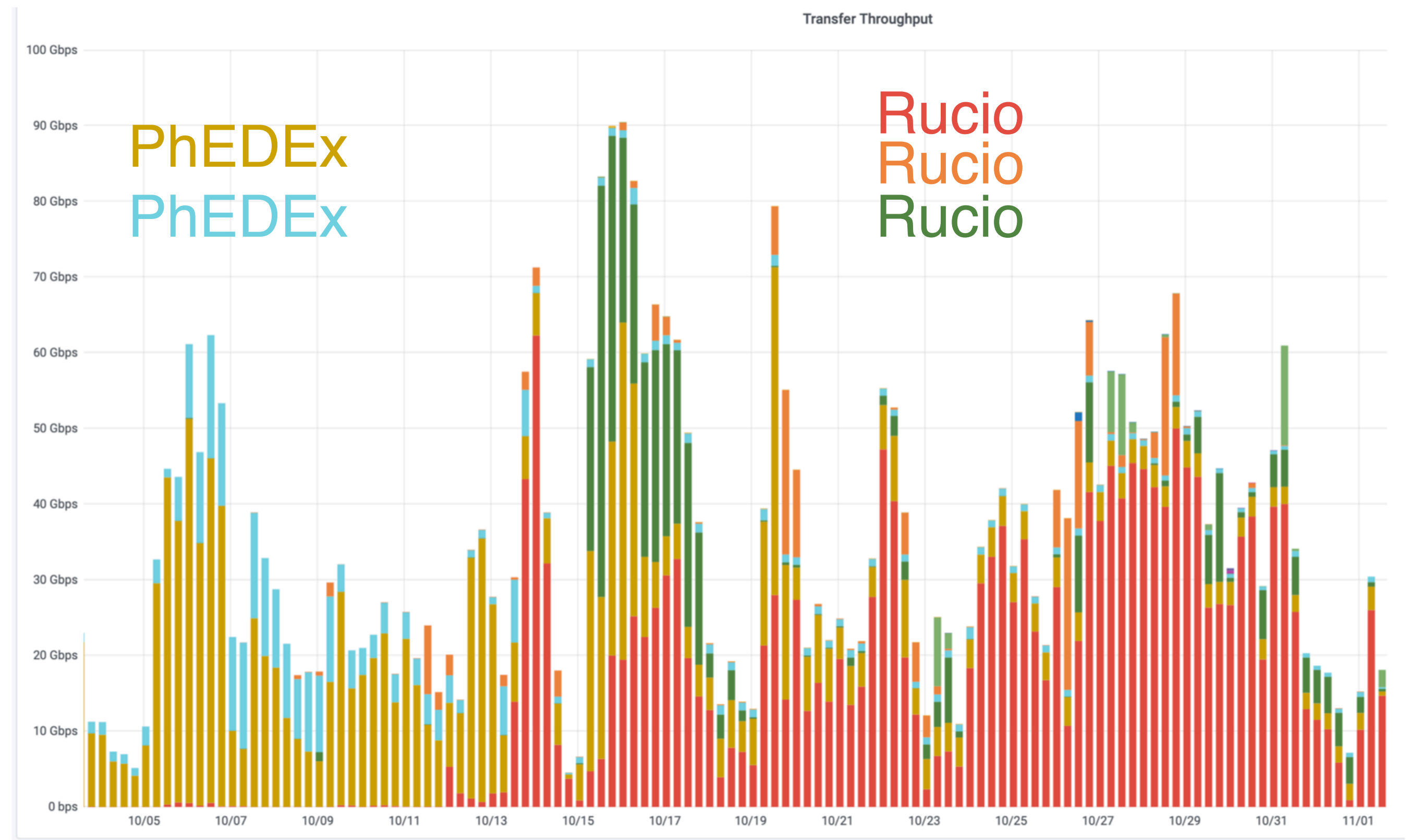


- In 2017 CMS begins exploring a replacement of our data management software
- Late 2017 a small group of us begin an evaluation of Rucio to meet CMS needs
 - Data models are a little different, but Rucio is flexible enough to adopt our strict 1:N 3-tier model
 - Processing workflow is different too (read directly from site storage) but this too is fine
 - Lots of increasing scale tests and syncing of metadata from legacy system into Rucio
- After a competitive selection review in summer 2018 CMS decided to adopt Rucio
 - Transition during LS2 to be ready by Run3
 - A few items identified as missing
 - ★ Consistency checking adaptable to CMS workflow
 - ★ Verification that a file is on tape (or secure storage) before removing last disk copy
 - ★ Lack of a “dry run” or simulation mode for removals
 - CMS often does data removal campaigns

Transition to Rucio

- June 2020 we transitioned our smallest NanoAOD to Rucio; almost no one noticed
 - This required a subset of what we need for a full Rucio transition and taught us valuable experience
 - Gave operators much needed experience in operating Rucio
- Full transition in November 2020
- Done with zero downtime of the production system.
 - Obviously non-trivial development to make this happen in other CMS software

FTS bandwidth at right shows a rapid transition with no let-up



Rucio experience

- Overall we are very happy with our experience with Rucio. Both with the product and with the development team
 - Confident that FNAL as a Rucio stakeholder will satisfy our concerns with the governance model
- How we operate Rucio
 - 2-3 dedicated transfer teams operators (also look after our xrootd federation)
 - Lots of dev-ops by a few of us
 - We don't have a great test-bed yet, so we delay adoption of major versions for few weeks
 - Running on CERN Oracle DB backend

What we're working on now

- Flux-GitOps (put into production this week)
 - Very nice, makes configuration changes available to all, tracks history and state
 - Upgrades nearly instantaneously
- Consistency checking with dashboard
 - We aim to make this something usable by the community (in the form of a toolbox)

CMS Data Consistency

sort by: RSE | CC run (+) (-) | unmerged run (+) (-)

RSE	Consistency Enforcement										Unmerged Files					
	History	Last run	Status	Dark files				Missing files			History	Last run	Elapsed time	Status	Files	Size
				Detected	Confirmed	Acted	Action status	Detected	Acted	Action status						
T2_CH_CERN		2021/09/28	started									2021/09/24	51m11s	done	2954259	517.1T
T2_UA_KIPT		2021/09/28	done	7				1				2021/09/18	10.34s	done	4785	4.2T
T2_IN_TIFR		2021/09/27	done	33		0	done	0	0	done		2021/09/18	54.80s	done	16274	12.3T
T2_BR_UERJ		2021/09/27	done	1		0	started	0				2021/09/26	42.35s	done	26768	8.7T
T2_CN_Beijing		2021/09/27	done	6		1	done	36413	0	aborted		2021/09/21	11m9s	done	69742	4.7T
T2_US_Florida		2021/09/27	done	39824		0	started	0				2021/09/25	21m29s	done	205565	180.3T
T2_RU_JINR		2021/09/27	done	103		7	done	0	0	done		2021/09/27	14m4s	done	21629	45.7T
T2_IT_Bari		2021/09/27	done	125		0	done	173	0	done		2021/09/25	87.20s	done	20493	33.9T
T2_DE_DESY		2021/09/27	done	66		7	done	0	0	done		2021/09/09	81.69s	done	52222	83.2T
T2_UK_London_IC		2021/09/26	done	276		5	done	0	0	done		2021/09/26	51m30s	done	42380	96.1T
T2_BE_IHE		2021/09/26	done	240		2	done	14096	0	aborted		2021/09/28	6m16s	done	31921	45.8T
T2_EE_Estonia		2021/09/26	done	109		8	done	2	0	done		2021/09/17	43.03s	done	21892	12.3T
T2_UK_London_Brunel		2021/09/26	failed									2021/09/22	0.54s	failed		
T2_US_UCSD		2021/09/26	done	35899		35714	done	3	0	done		2021/09/26	45.89s	done	28240	13.7T
T1_ES_PIC_Disk		2021/09/25	done	134		0	done	0	0	done		2021/09/24	1m47s	done	55948	50.0T
T3_US_Rutgers		2021/09/25	done	0		0	done	0	0	done		2021/09/16	2.66s	done	39	720.0
T2_UK_SGrid_RALPP		2021/09/25	done	34		2	done	0	0	done		2021/09/21	8m12s	done	91954	90.4T
T2_IT_Pisa		2021/09/25	done	139		19	done	768	0	done		2021/09/26	34.60s	done	99250	27.9T

- Our team is small, but we have still made some contributions to making Rucio a community project
 - CMS has been running on kubernetes since sometime in our evaluation process
 - ★ *Everything* except central services (database, message brokers, monitoring) is running in kubernetes
 - ★ Includes CMS specific components and necessary 3rd party components, e.g. Prometheus
 - ★ Hosted on CERN open stack
 - ★ So we've contributed quite a few extensions and updates as Rucio changes
 - ★ We are extremely pleased with this mode of operations
 - Monitoring probes
 - ★ Were ATLAS specific and centric. We've been systematically evaluating them, moving the relevant ones to common code base, and adding "missing" probes
 - Archived to tape verification
 - ★ Requested by our review, added to FTS, functionality added to Rucio to implement
 - Consistency checking
 - ★ Still a work in progress. CMS uses a different model so we refactored the entire process and are developing a toolkit which others can use to implement a consistency checks

Technology updates enabled by Rucio

- Having a flexible framework where we can collaborate on technology changes is a huge help
- These are things which would have been difficult in our old system.
- Some are done, some ongoing, some are still to come
 - CTA replacing CASTOR
 - HTTPS replacing GridFTP
 - Globus Online transfers for US HPCs
 - Tokens replacing X509 certificates
 - Quality of service transitions

Other areas to improve

- Thinking of what we could make use of in the near term...
- Multi-hop FTS to Globus Online transfers
- Simulation mode
 - How much space is reclaimed by removing a list of rules?
 - Which datasets are complete removed from a list of RSEs?
 - Which datasets are unaffected because of overlapping rules?
- Continue standardizing on Prometheus monitoring vs. Statsd

Conclusions

- This transition has gone very well for CMS
- Obviously, replacing a major part of our production system would expose *some* issues
- We are very satisfied with the collaborative aspect of this so far
 - Changes to the governance would ensure this going forward
- We look forward to running with Rucio in Run3 and helping ensure it's suitable for Run4 as well