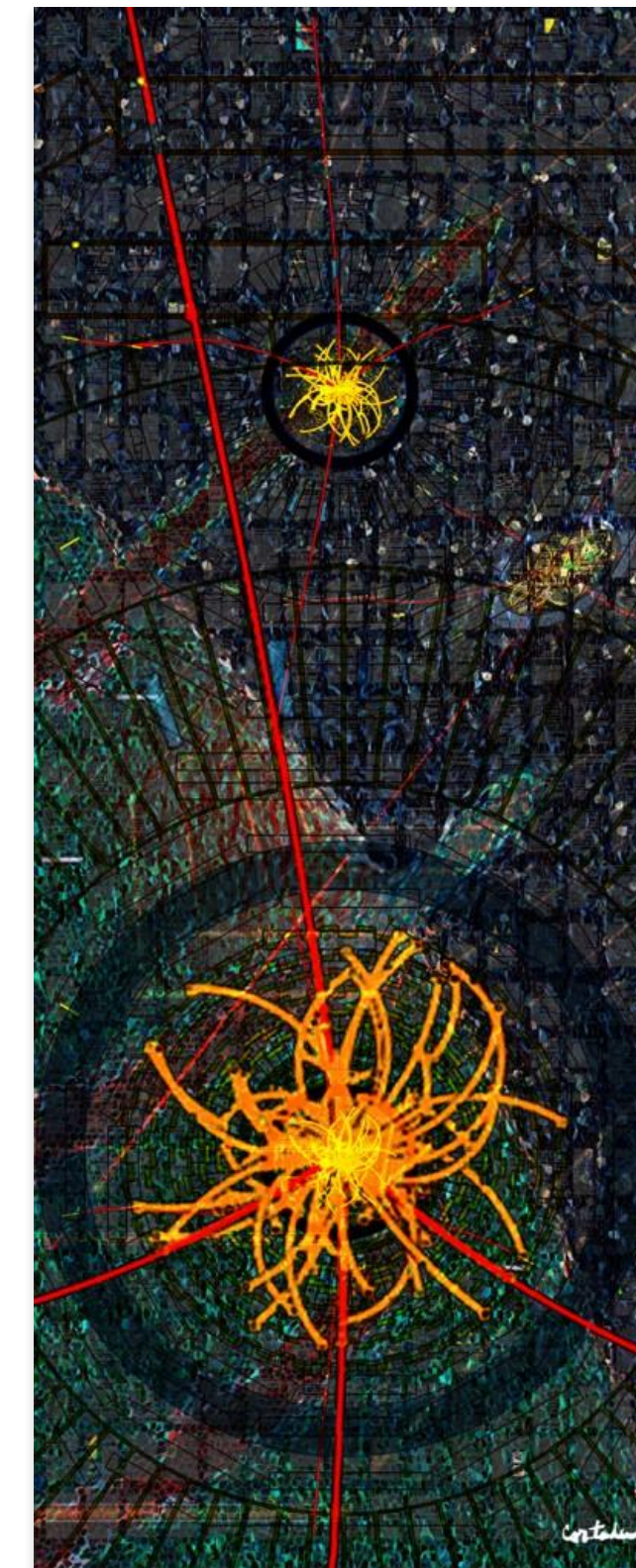
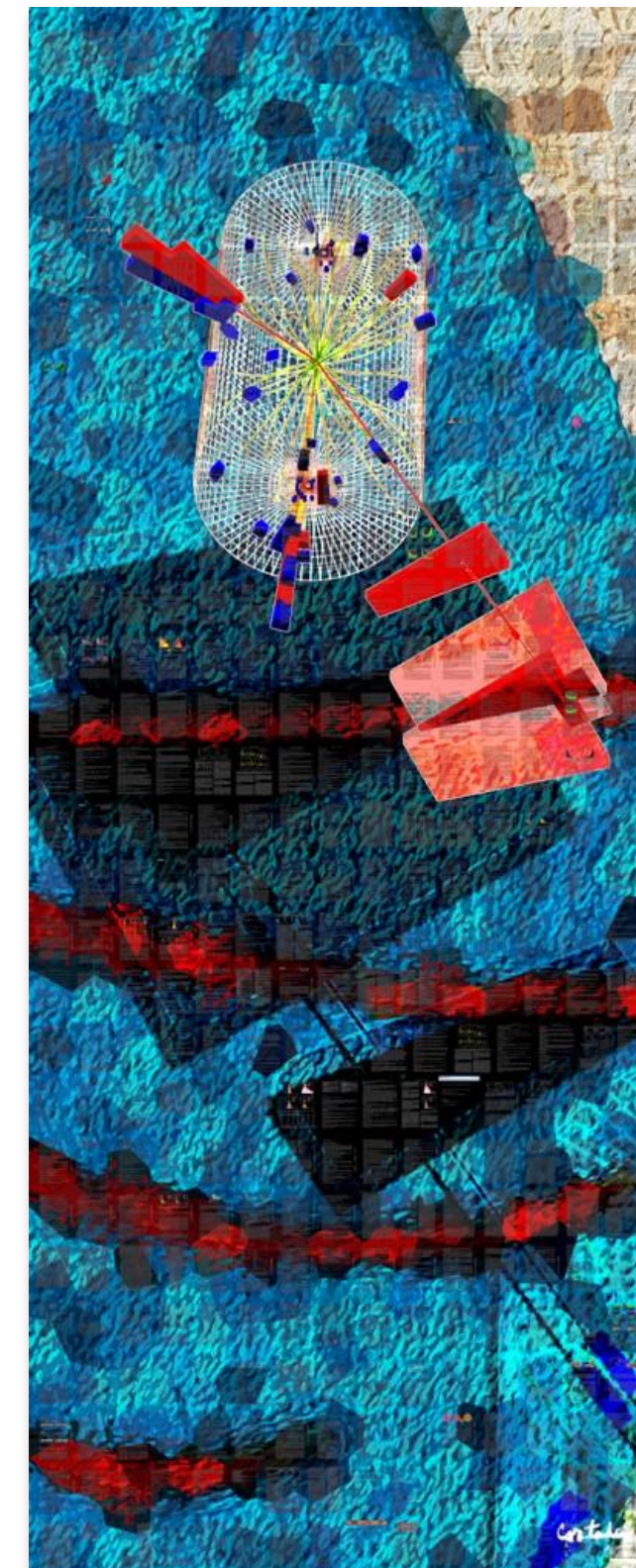
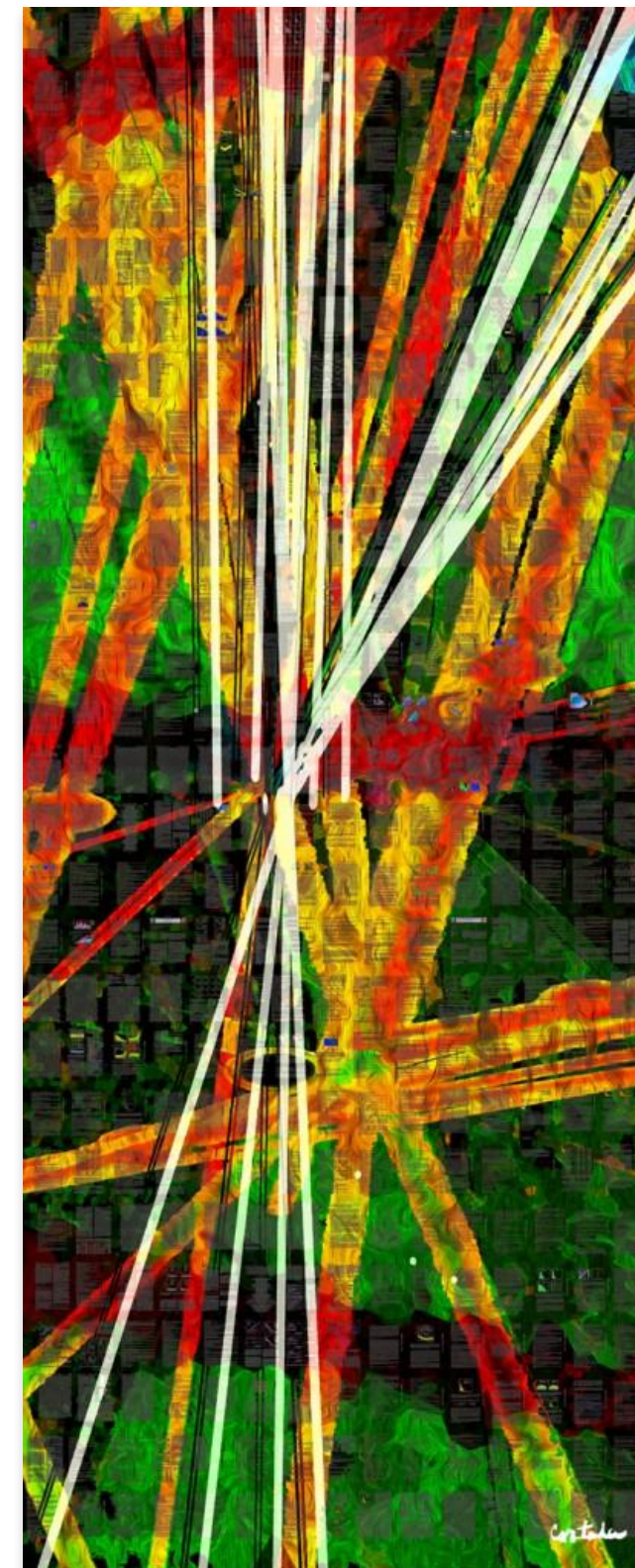
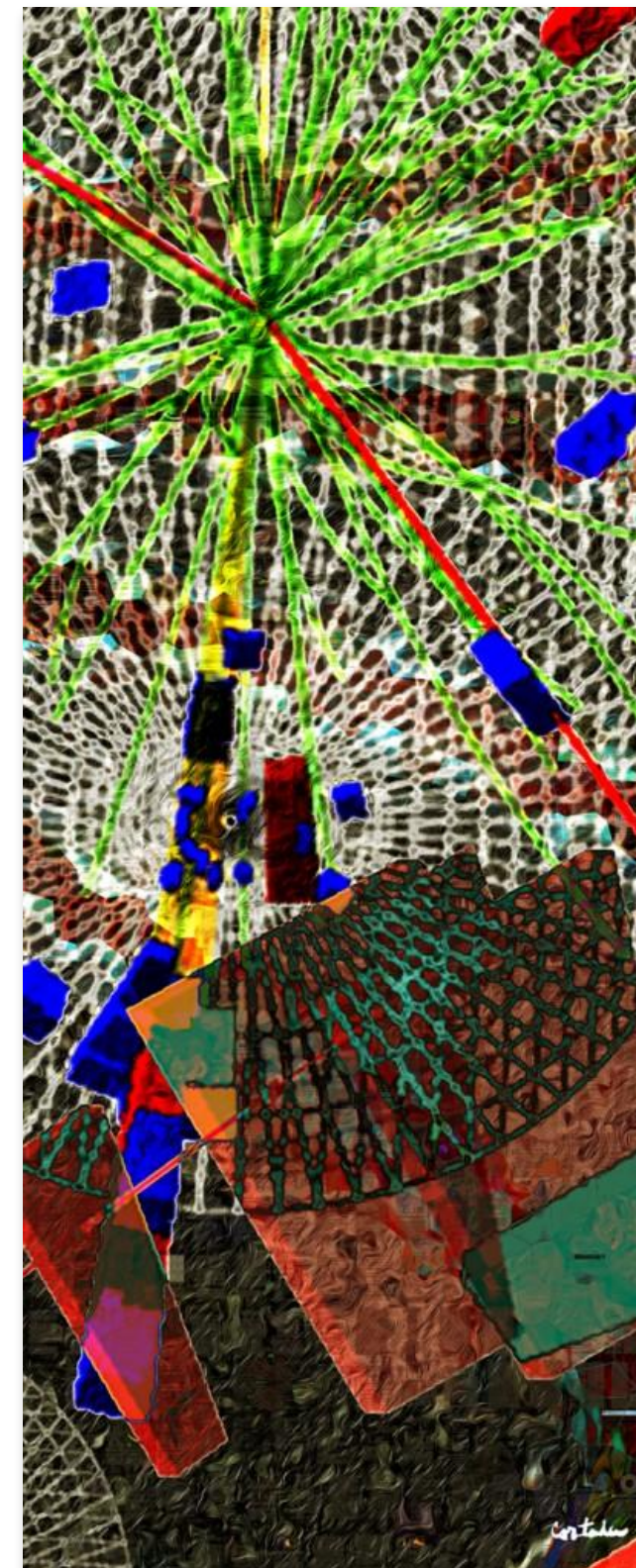
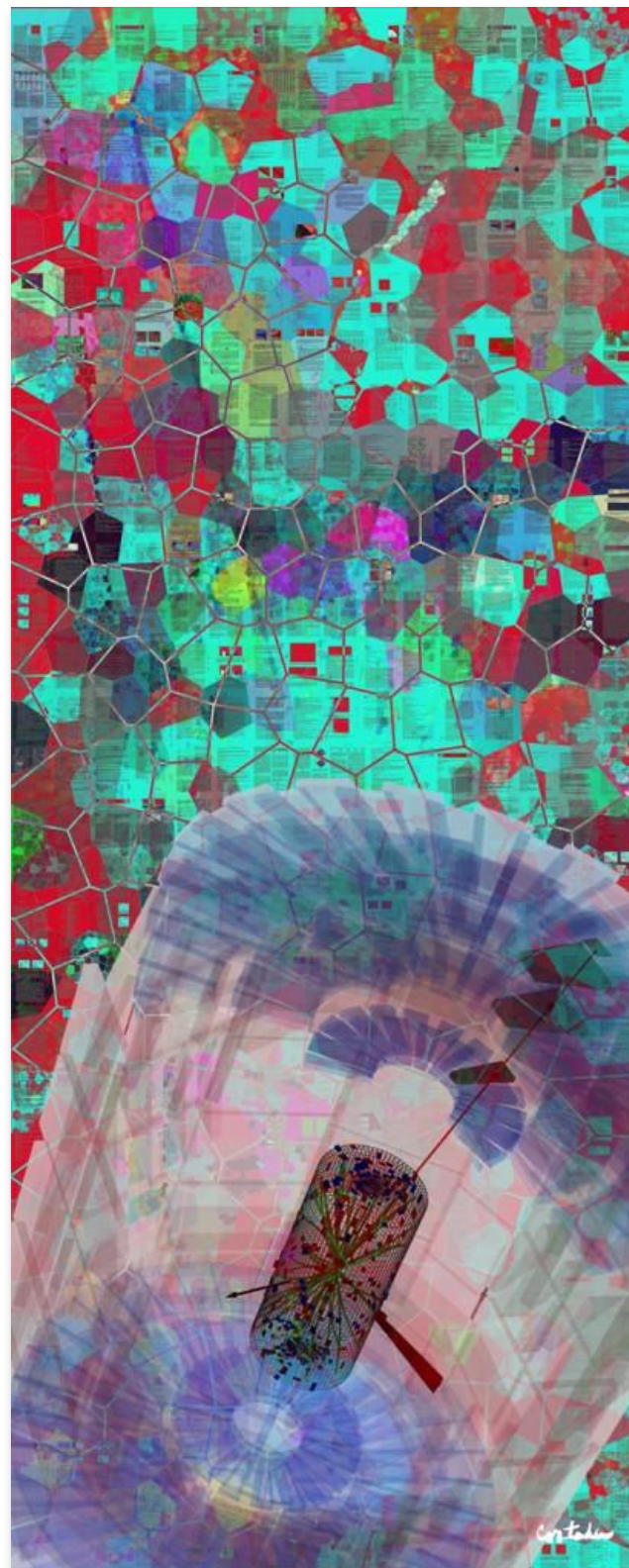


FNAL FCRSG CMS

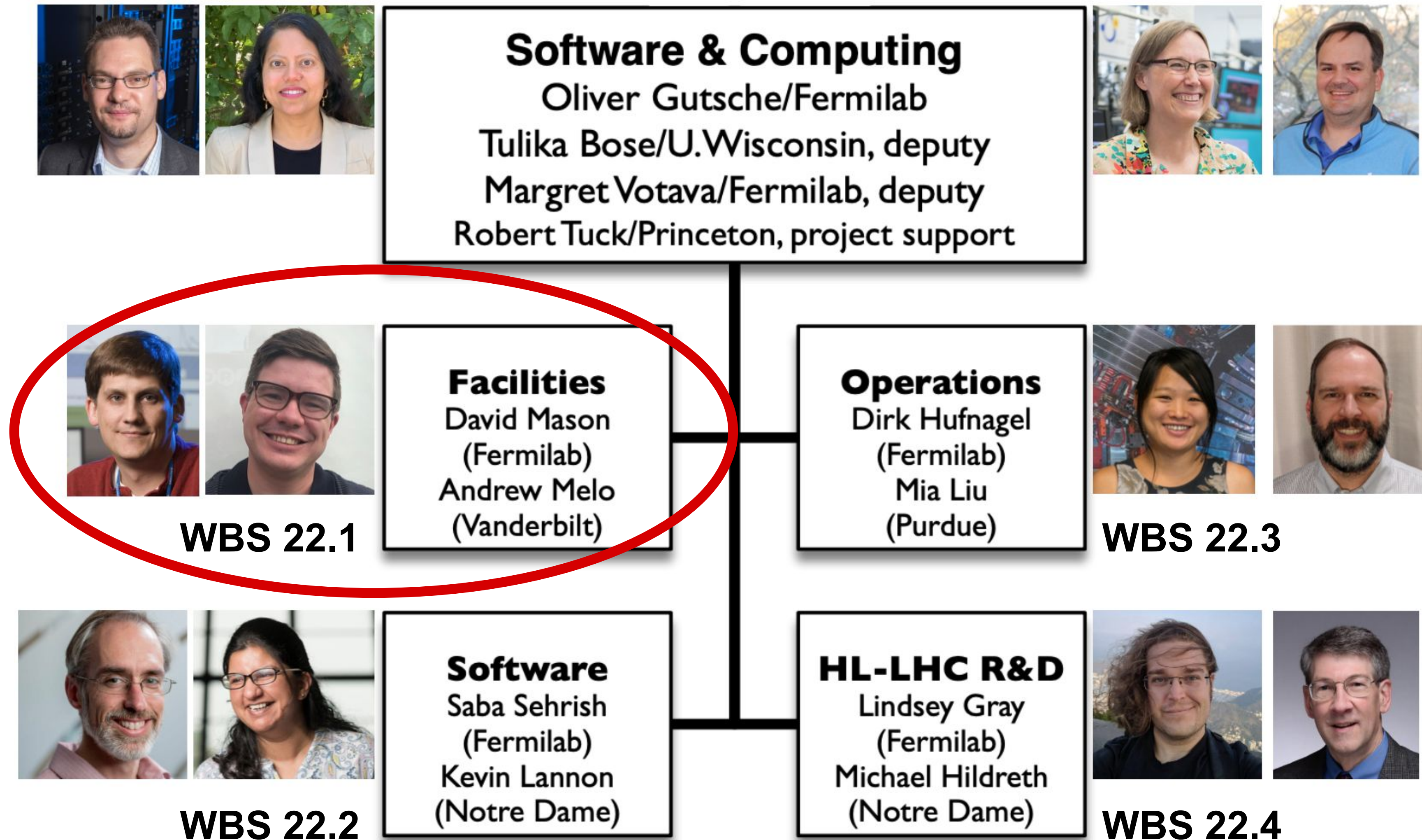
D. Mason for U.S. CMS Software and Computing Operations Program



- USCMS Ops program
 - This talk will be **completely facility focused**
- Project Planning and Risks
 - Brief 2022 Summary/Milestones achieved
- 2023 Resources
 - Forecasted resource needs into HL-LHC era
 - Brief drill down on tape, including WLCG Network/Tape Challenges
 - Run 3 and HL-LHC estimates, schedule change effect
 - HEPCloud and EAF

U.S. CMS S&C Operations Program Organization

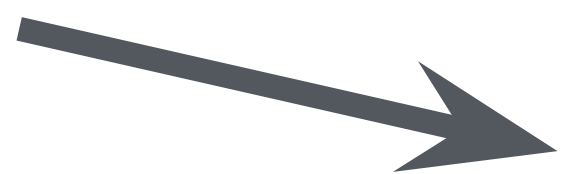
The U.S. CMS S&C Operations Program Execution Team (PET)



- The USCMS Fermilab Facility is a piece of the FNAL Institutional Cluster, and consists of the Tier 1 and LPC
- USCMS Provides computing and storage resources to CMS via FNAL Tier 1 and University based Tier 2 and 3 computing facilities.
 - **FNAL Tier 1 provides high availability computing and custodial data storage, as well as the analysis facility supporting the LPC.**
 - USCMS Tier 2's provide reliable compute and storage to CMS for both user analysis and central production
- The Operations Program budgets for USCMS Tier 1 and 2 hardware and support.
- In the following we summarize our process and estimates for estimating hardware needs and resulting budgets into the HL-LHC era (currently through 2030) for the Tier 1 ~~and 2's~~:

Project planning & risks

- Milestones and performance goals are tracked using ProjectManager.com
- Risks tracked in FNAL IPPM – major risks expected in 2023:
 - Cost increases and delivery delays
 - CPU purchase from 2021 had >6 month delivery delay
 - Have seen ~10% cost increases on first 2022 purchases (network and tape drives)
 - Russia-Ukraine conflict
 - Likely that JINR Tier 1 can become unavailable
 - 20% of CMS Tier 1 capacity (FNAL is 40%)
 - Have so far accepted 650 TB of JINR data, included with 2022 tape pledge
 - Have agreed to additional tape storage with CMS
 - 17 PB deleted from tape this past year, more expected
 - Heavy Ion run that we expected in 2022 will happen in 2023 instead



S&C ID	TASK NAME	DURATION	PLANNED ...	PLANNED ...	PERCE...	KPI ...
SC-M-WBS1-22-3	Decision Point: Should Tier-2s use Erasure-Coded Storage	126 days	1/6/2022	6/30/2022	100%	
SC-M-WBS1-23-3	Tier 1 Pledge Deployment	567 days	1/28/2021	3/31/2023	60%	
	Deploy 2023 WLCG T1 Pledge	262 days	3/31/2022	3/31/2023		
	Report WLCG Pledge Deployed	262 days	3/31/2022	3/31/2023		
SC-M-WBS1-21-4	Deploy 2022 WLCG T1 Pledge	306 days	1/28/2021	3/31/2022	100%	
	Update 5+ year T1 resource plan/forecast	65 days	6/2/2021	8/31/2021	100%	
	Extract Run 3 resource plan from LPC heads	88.96 days	1/28/2021	6/1/2021	100%	
	Ops program internal review	0 days	9/1/2021	9/1/2021	100%	
	PB of Retired HW from T1 into LPC EOS	242 days	1/28/2021	12/31/2021	100%	
	Report WLCG Pledge Deployed	1 day	3/31/2022	3/31/2022	100%	
	Tier 1 Purchases	633 days	1/28/2021	7/3/2023	13%	
	Purchase 2021 CPU	226 days	2/19/2021	12/31/2021	100%	
	Funds moved to EQ	1 day	2/19/2021	2/19/2021	100%	
	RFP submitted	85 days	2/22/2021	6/18/2021	100%	
	Vendor Responses Received	19 days	6/21/2021	7/15/2021	100%	
	PO Awarded	14 days	7/16/2021	8/4/2021	100%	
	Hardware Delivered	9 days	11/19/2021	12/1/2021	100%	
	HW Deployed	22 days	12/2/2021	12/31/2021	100%	
	Purchase 2021 Disk	163.04 days	2/19/2021	10/6/2021	100%	
	Req submitted	1 day	2/19/2021	2/19/2021	100%	
	Disk Arrays Deployed	64 days	2/19/2021	5/19/2021	100%	
	Servers Deployed	132.04 days	4/5/2021	10/6/2021	100%	
	Purchase 2022 CPU	567 days	1/28/2021	3/31/2023		
	Deploy	567 days	1/28/2021	3/31/2023		
	Purchase 2022 Disk	619 days	1/28/2021	6/13/2023	50%	
	RFP process started	567 days	4/12/2021	6/13/2023	100%	
	Hardware Delivered	567 days	1/28/2021	3/31/2023		
	Spectra Additional 20 Tape Drive Purchase 2022	371 days	1/31/2022	7/3/2023		
	LPC GPU Replacement 2022	566.96 days	1/28/2021	3/31/2023		
	Move T1 to LTO9 Media (Begin Writing)	1 day	1/2/2023	1/2/2023		
	T1 M8 to LTO9 Migration	522 days	6/5/2023	6/3/2025		
SC-P-WBS1-1	Performance Goal: Meet CMS Site Readiness Metrics at T1	328 days	10/1/2021	1/3/2023		99.5
	Performance Goal: Meet CMS Site Readiness Metrics at T1 (...	64 days	1/3/2022	3/31/2022		99.5
	CY22Q1	64 days	1/3/2022	3/31/2022		99.5
	Performance Goal: Meet CMS Site Readiness Metrics at T1 (...	328 days	10/1/2021	1/3/2023		
	CY22Q2	66 days	4/1/2022	7/1/2022		
	CY22Q3	66 days	7/1/2022	9/30/2022		
	CY22Q4	328 days	10/1/2021	1/3/2023		
SC-P-WBS1-2	Performance Goal: LPC Availability	261 days	1/3/2022	1/2/2023	30%	99.2
	LPC Interactives dual stack	22 days	3/2/2022	3/31/2022	100%	
	CY22Q1	63.96 days	1/3/2022	3/31/2022	100%	99.2
	CY22Q2	65 days	4/1/2022	6/30/2022		
	CY22Q3	66 days	7/1/2022	9/30/2022		
	CY22Q4	66 days	10/3/2022	1/2/2023		
	Performance Goal: Provide Necessary Tier 1 Services 2022	990 days	3/18/2019	12/30/2022	30%	
	Batch System Dual Stack	19 days	2/1/2022	2/25/2022	100%	
	Tier 1 IPv4/6 dual stack	1 day	3/18/2019	3/18/2019	100%	
	Storage	1 day	3/18/2019	3/18/2019	100%	
	2022 WLCG March Tape Challenge	5 days	3/7/2022	3/11/2022	100%	
	Dcache upgrade to 7x series	44 days	3/2/2022	5/2/2022	100%	
	Disk	44 days	3/2/2022	5/2/2022	100%	
	Tape	44 days	3/2/2022	5/2/2022	100%	
	Placeholder Random Thing to Make this Last The Year	260 days	1/3/2022	12/30/2022		
SC-M-WBS1-22-10	Migrate T1 to OSG 3.6	135 days	2/24/2022	8/31/2022	31%	
	Migrate XRootD to OSG 3.6	135 days	2/24/2022	8/31/2022		
	Migrate HTCondor-CE to OSG3.6	129.96 days	2/24/2022	8/24/2022	100%	
	Batch Farm to OSG 3.6	1 day	3/2/2022	3/2/2022	100%	
	Storage to OSG 3.6	131 days	3/2/2022	8/31/2022		

2022 Milestones

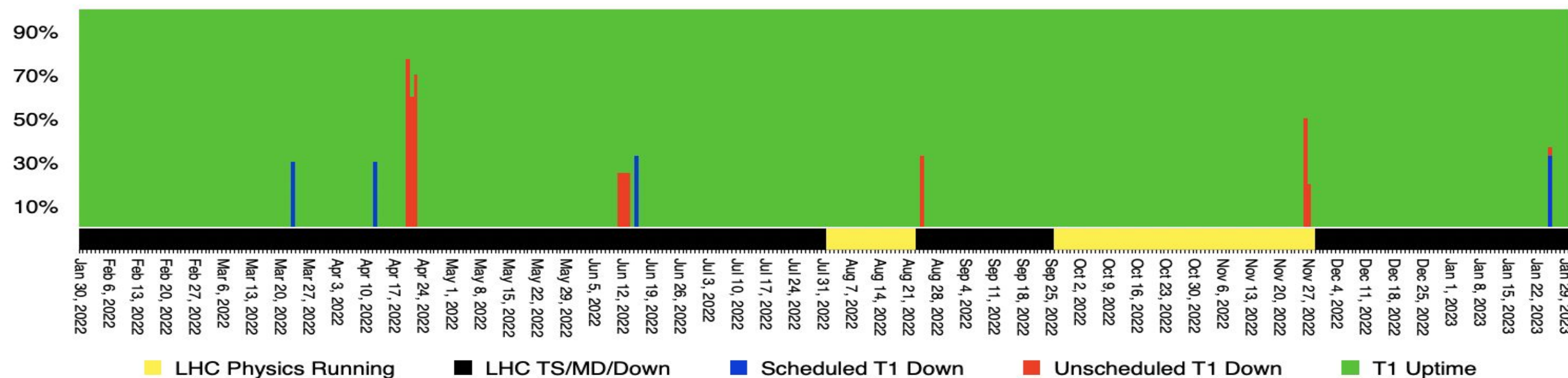
- Readiness metric for the year met – **98.9%**
- 2023 Disk pledge expected to be late
 - Though completed last summer, several month delay in issuing PO + supply chain delays. Delivery estimate is March 31, Pledge is due April 1. Believe this first time not making pledge in time.
- OSG 3.6 migration complete, Global pool moved to IDTokens

Milestones for 2023

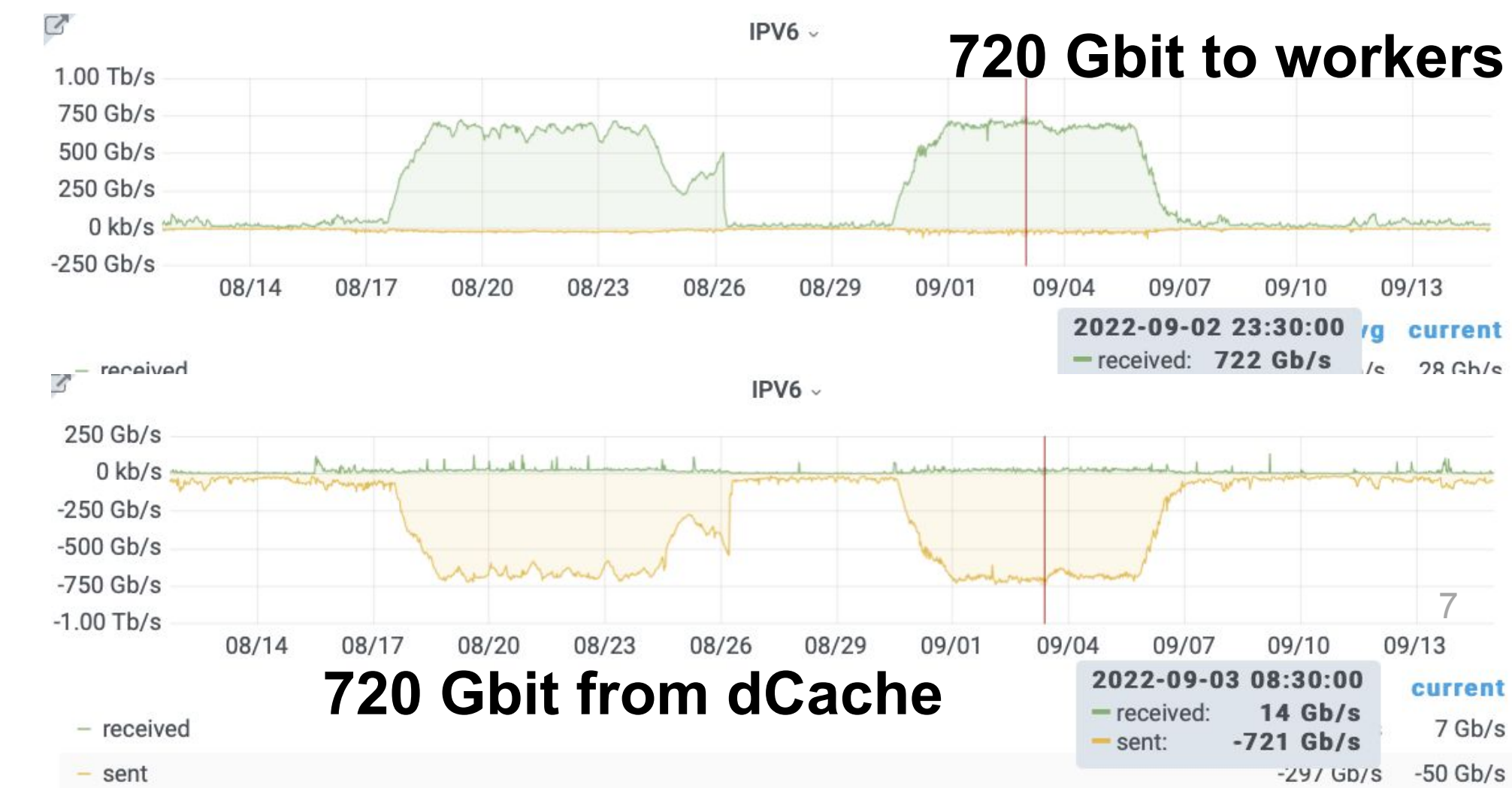
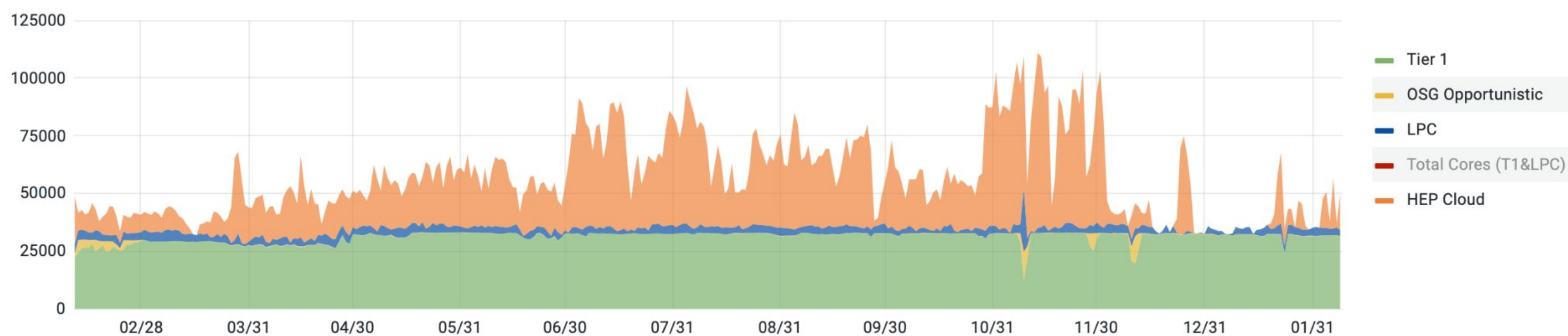
- Purchases for 2023 WLCG Pledge
 - Large amount of CPU and Disk needed (more later)
 - Need to replace tape buffer (Budgeted \$ for NVME)
- EOS upgrade to EOS 5, begin investigating erasure coding viability
- Testing and planning for transition to CTA in ~2024
- M8→LTO9 migration to start to free up library slots in IBM Tape Library
- Run 3 datataking, including large heavy ion run

Tier 1 Performance

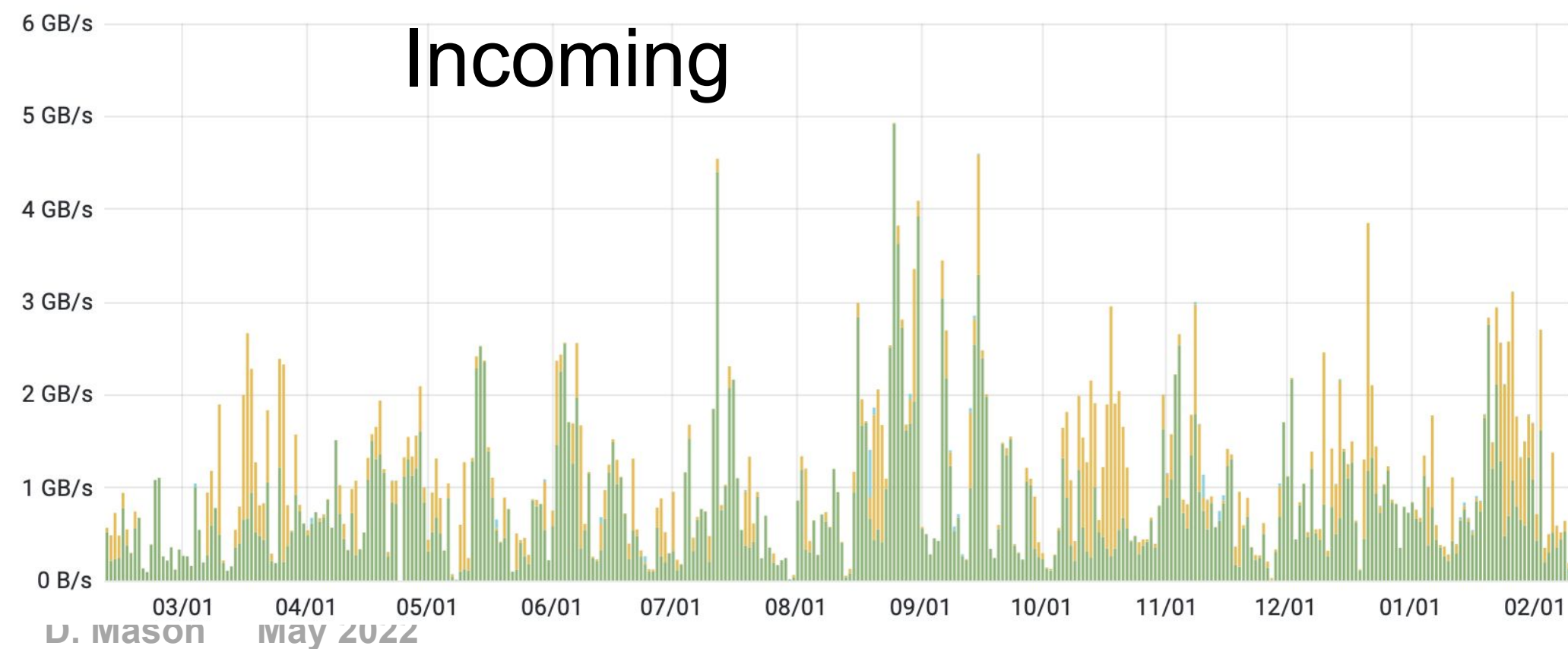
- Required to meet >98% readiness
- 98.9% over the past year



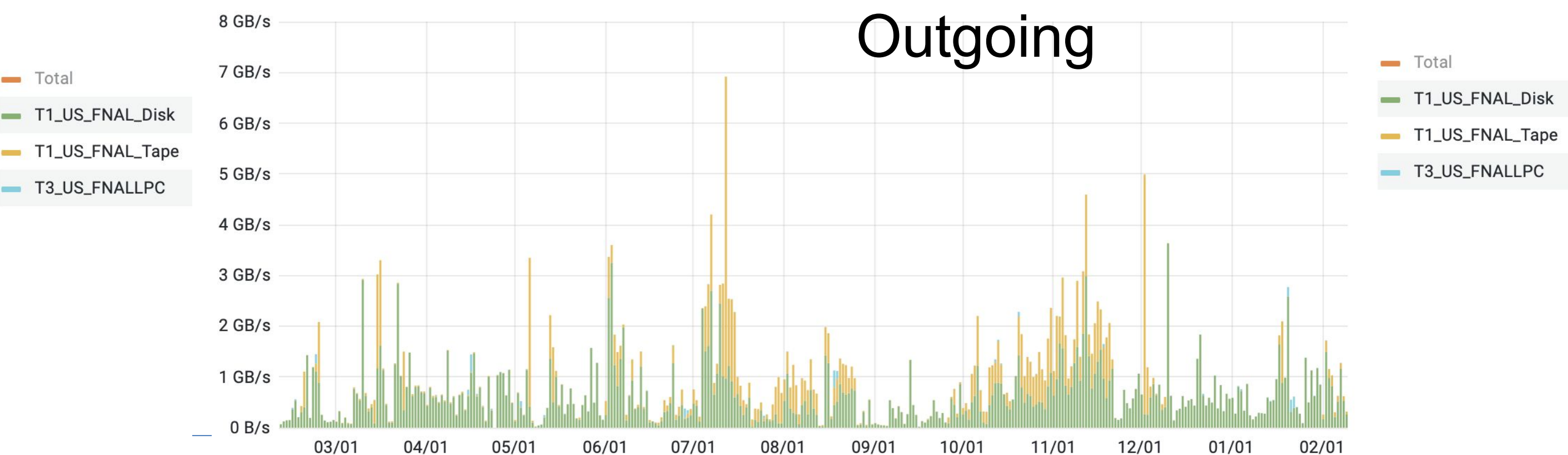
Claimed Cores on the Combined Tier 1 and LPC plus CMS on HEP Cloud



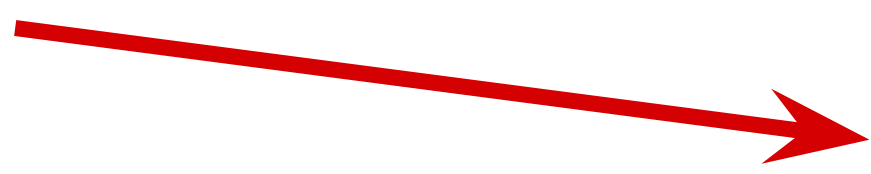
Transfer Throughput



Transfer Throughput



- LPC provides a central physics hub for uscms, but also cms in general
- **Mid January – first in-person CMSDAS school since 2020**
 - **Very successful**, about 50 students learning all things CMS analysis.



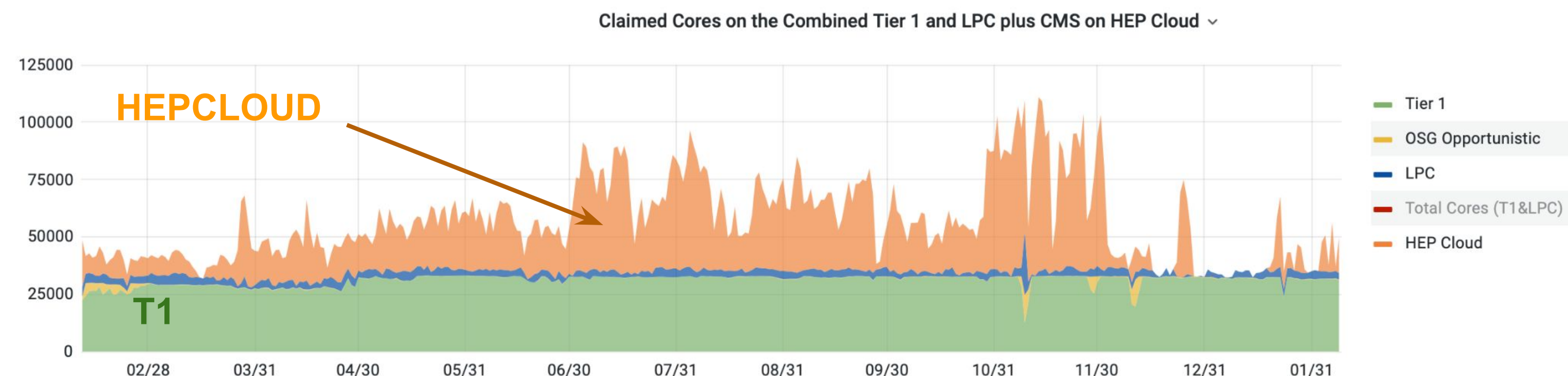
LPC Has

- Hundreds of active users
- About 100 interactive VM's
- Allocated 4k cores from the CMS FNAL Facility
- 10 PB (replicated, so 5 PB real) EOS storage
- NFS space (360TB Nexsan server)



EAF and HEPCloud

- EAF is a production thing, for some definition of production
- **Continues to make very good progress**
 - LPC users use this – used some amount during the recent cmsDAS along with AF at Vanderbilt
 - This year we want to continue to understand the operations model for this resource, and synergies between the FNAL EAF and AF’s CMS uses at Tier 2 centers
- **HEPCloud managed to exhaust primary allocation and several additions again in 2022!**
 - As well as several XSEDE NSF allocations
 - During past year:
 - 196M core-hours from HEPCloud
 - 141M opportunistic from OSG
 - Predominantly from Syracuse
 - **337M Combined!**
 - **Compare to 274M from T1**



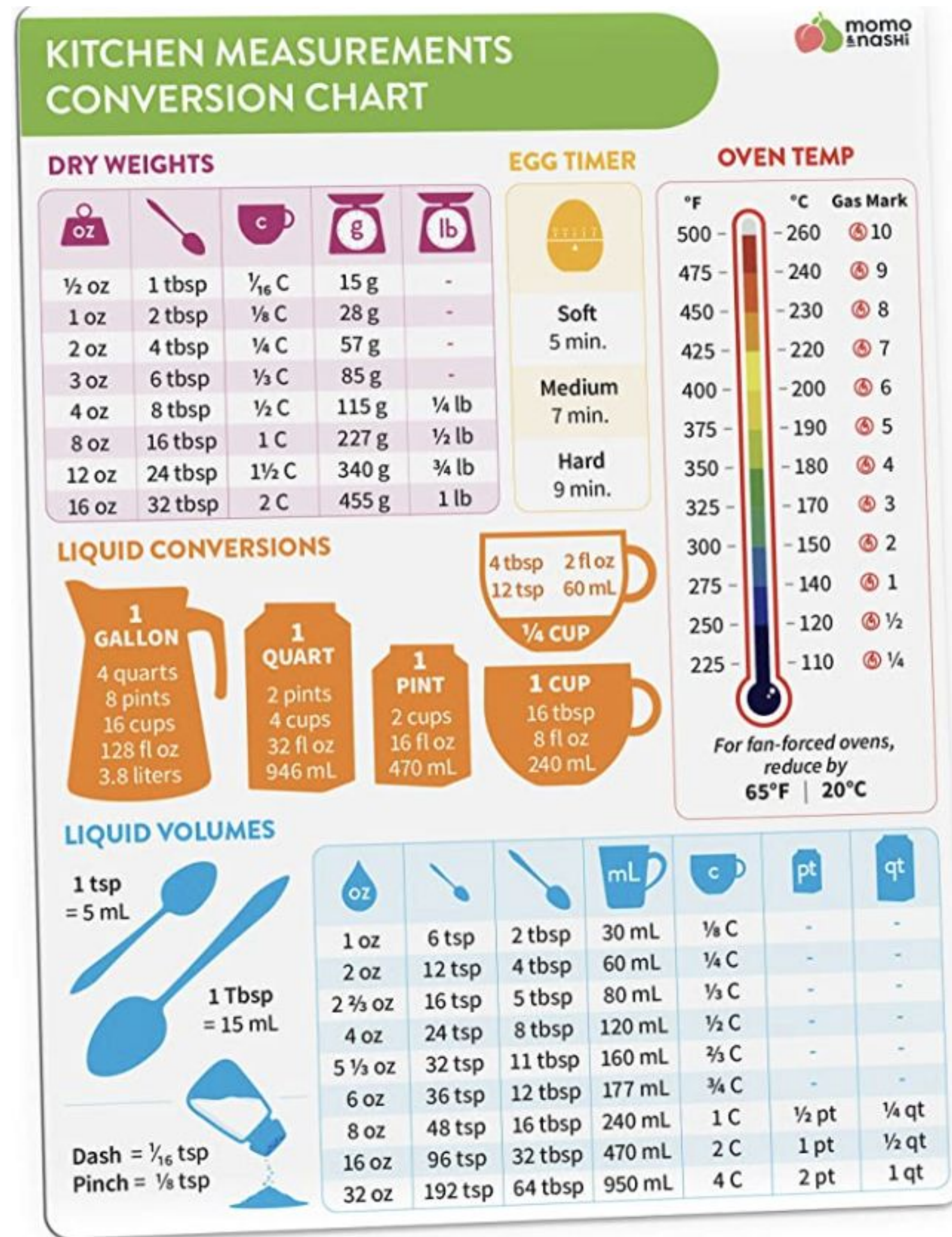
For the last decade and a half, WLCG has specified CPU in kHS06*-years.

With the mix of CPU's we currently have in the T1:

$$\text{Core-hr/year} = (\text{kHS06}) / 12.1 * 365 * 24$$

I.e. the 2022 pledge of 320kHS06 share from FNAL is about 30k cores, or 263Mcore-hr. (at high availability)

*WLCG moving to HEP Score over next few years



KITCHEN MEASUREMENTS CONVERSION CHART

DRY WEIGHTS

oz	spoon	C	g	lb
1/2 oz	1 tbsp	1/16 C	15 g	-
1 oz	2 tbsp	1/8 C	28 g	-
2 oz	4 tbsp	1/4 C	57 g	-
3 oz	6 tbsp	3/8 C	85 g	-
4 oz	8 tbsp	1/2 C	115 g	1/4 lb
8 oz	16 tbsp	1 C	227 g	1/2 lb
12 oz	24 tbsp	1 1/2 C	340 g	3/4 lb
16 oz	32 tbsp	2 C	455 g	1 lb

EGG TIMER

- Soft: 5 min.
- Medium: 7 min.
- Hard: 9 min.

OVEN TEMP

°F	°C	Gas Mark
500	-260	10
475	-240	9
450	-230	8
425	-220	7
400	-200	6
375	-190	5
350	-180	4
325	-170	3
300	-150	2
275	-140	1
250	-120	1/2
225	-110	1/4

For fan-forced ovens, reduce by 65°F | 20°C

LIQUID CONVERSIONS

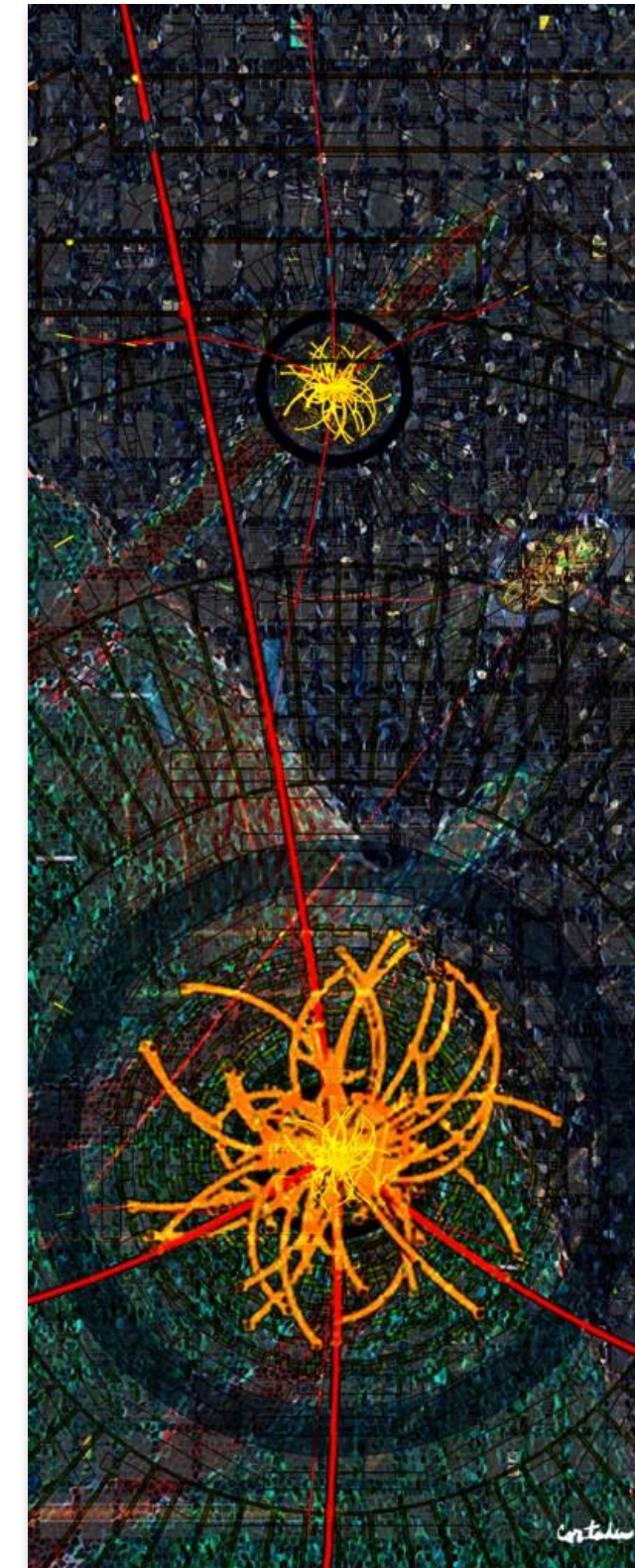
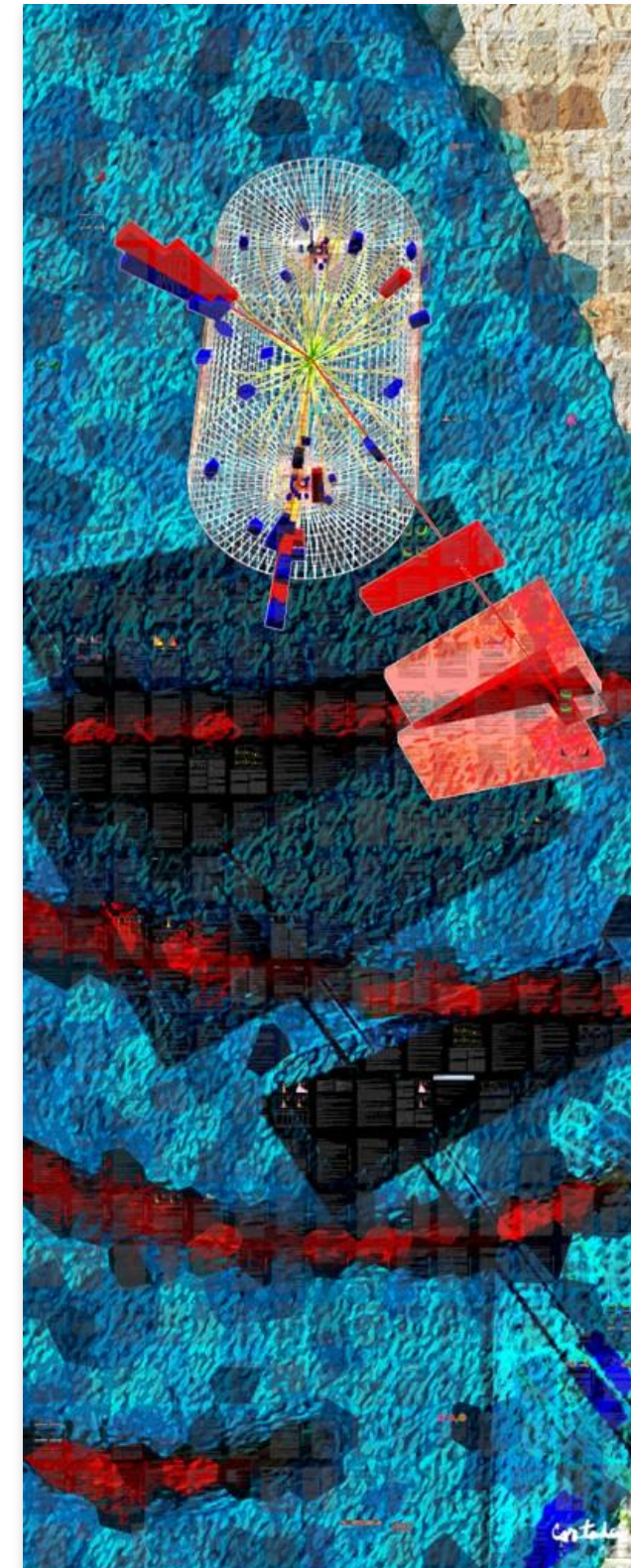
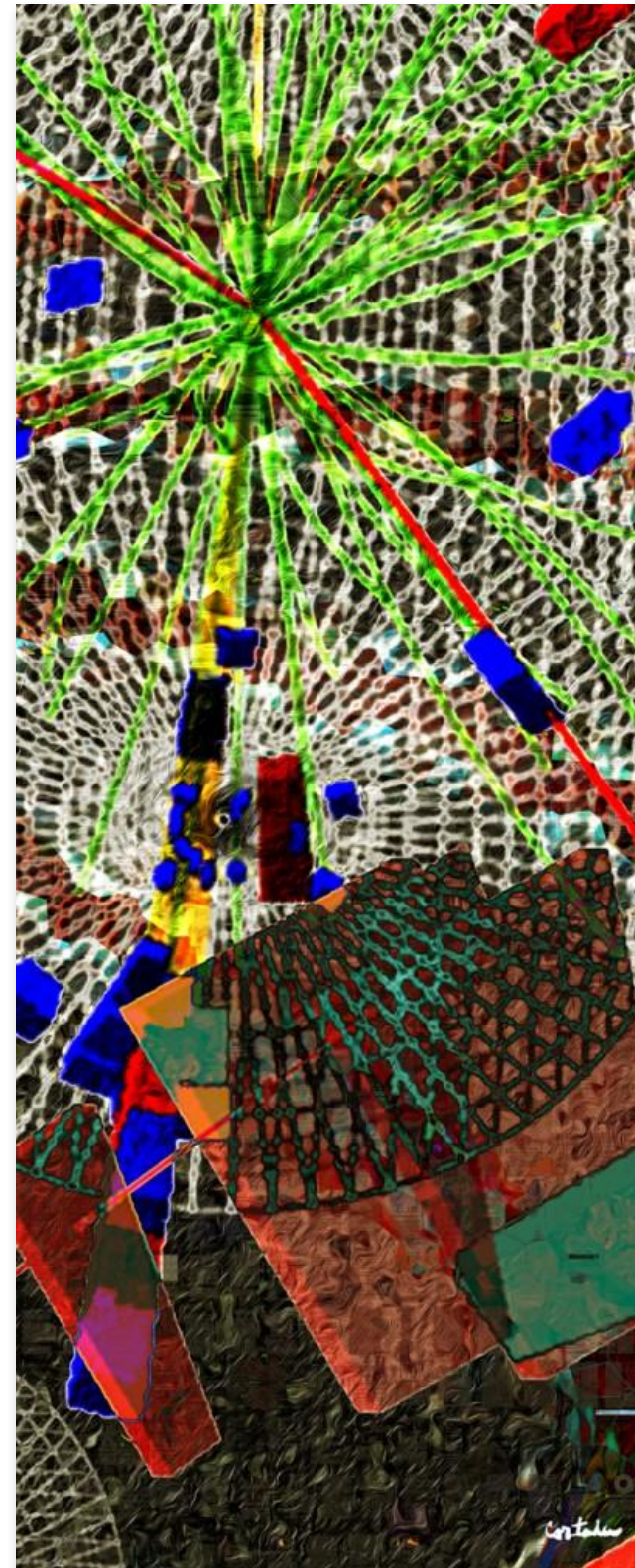
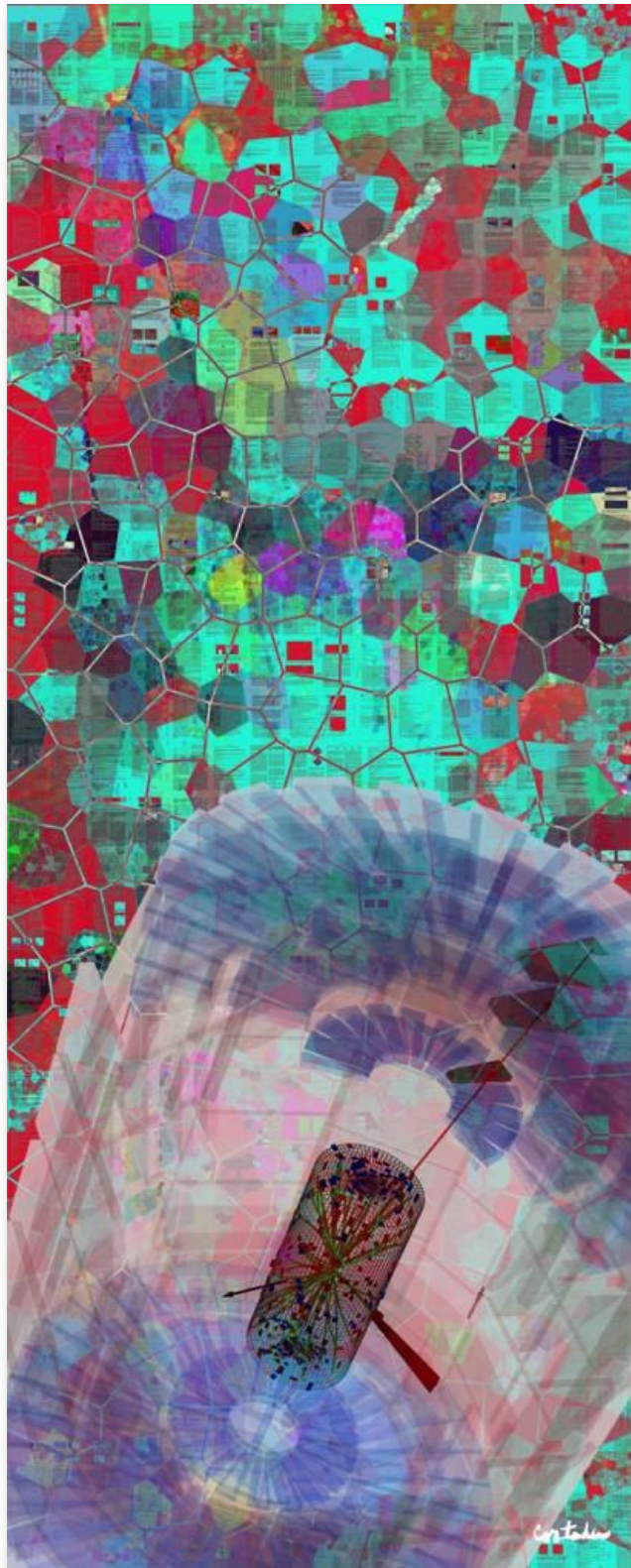
- 1 GALLON: 4 quarts, 8 pints, 16 cups, 128 fl oz, 3.8 liters
- 1 QUART: 2 pints, 4 cups, 32 fl oz, 946 mL
- 1 PINT: 2 cups, 16 fl oz, 470 mL
- 1/4 CUP: 4 tbsp, 2 fl oz, 12 tsp, 60 mL
- 1 CUP: 16 tbsp, 8 fl oz, 240 mL

LIQUID VOLUMES

oz	tsp	tbsp	mL	C	pt	qt
1 oz	6 tsp	2 tbsp	30 mL	1/8 C	-	-
2 oz	12 tsp	4 tbsp	60 mL	1/4 C	-	-
2 2/3 oz	16 tsp	5 tbsp	80 mL	1/3 C	-	-
4 oz	24 tsp	8 tbsp	120 mL	1/2 C	-	-
5 1/3 oz	32 tsp	11 tbsp	160 mL	2/3 C	-	-
6 oz	36 tsp	12 tbsp	177 mL	3/4 C	-	-
8 oz	48 tsp	16 tbsp	240 mL	1 C	1/2 pt	1/4 qt
16 oz	96 tsp	32 tbsp	470 mL	2 C	1 pt	1/2 qt
32 oz	192 tsp	64 tbsp	950 mL	4 C	2 pt	1 qt

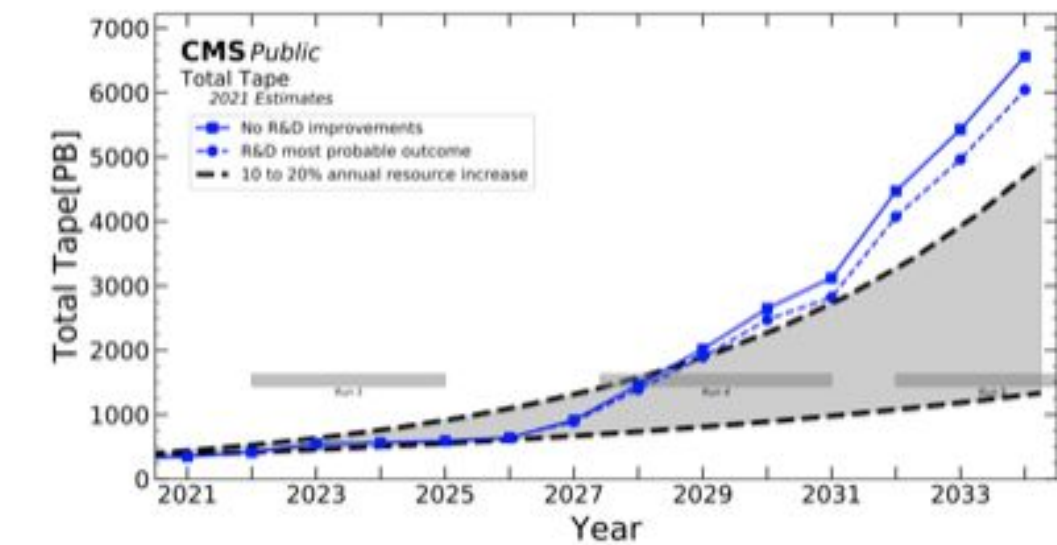
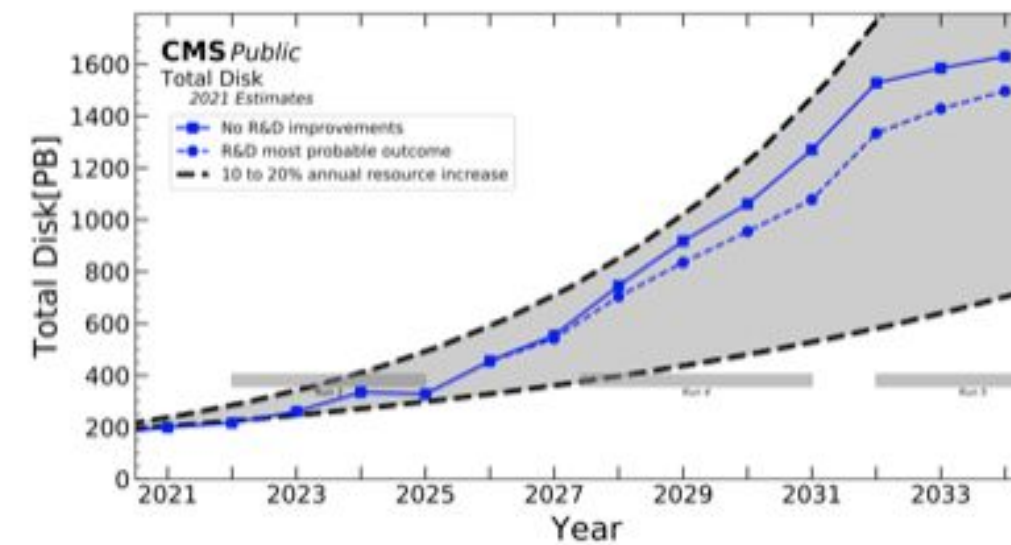
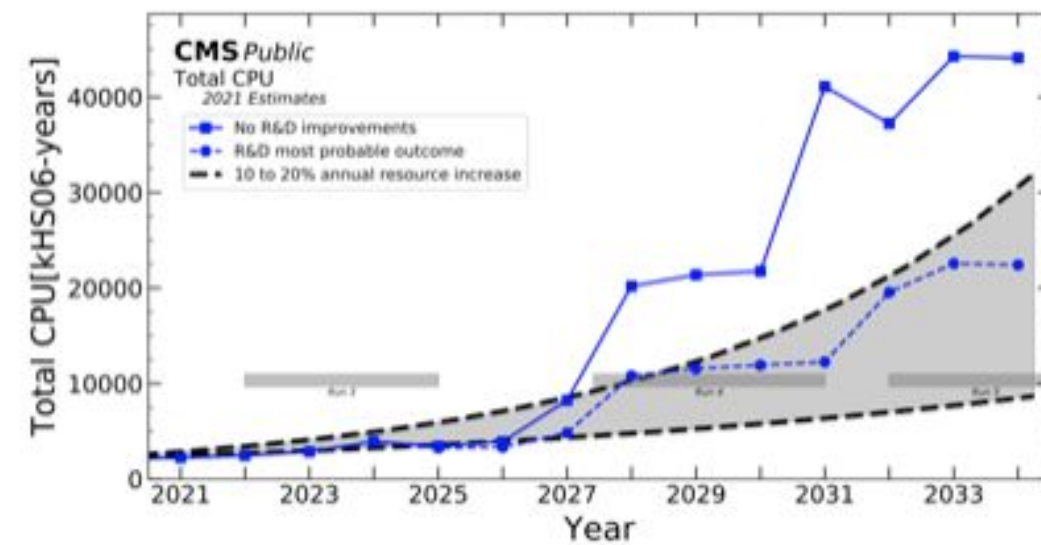
1 tsp = 5 mL
1 Tbsp = 15 mL
Dash = 1/16 tsp
Pinch = 1/8 tsp

First: How Much Stuff Do We Need?



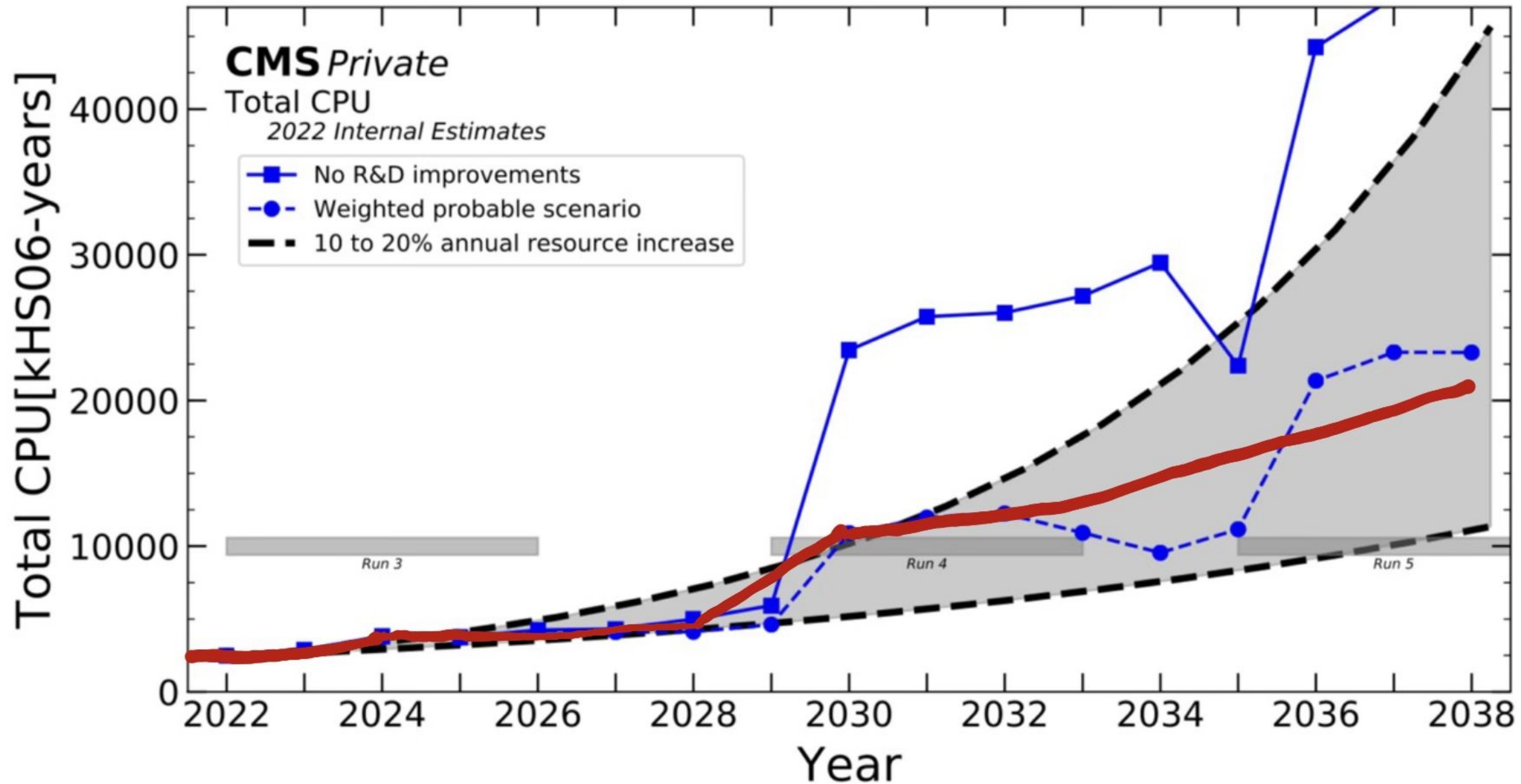
Resource requirement modeling – CPU

CPU time, disk and tape time projected requirements estimated to be required annually for CMS processing and analysis needs.



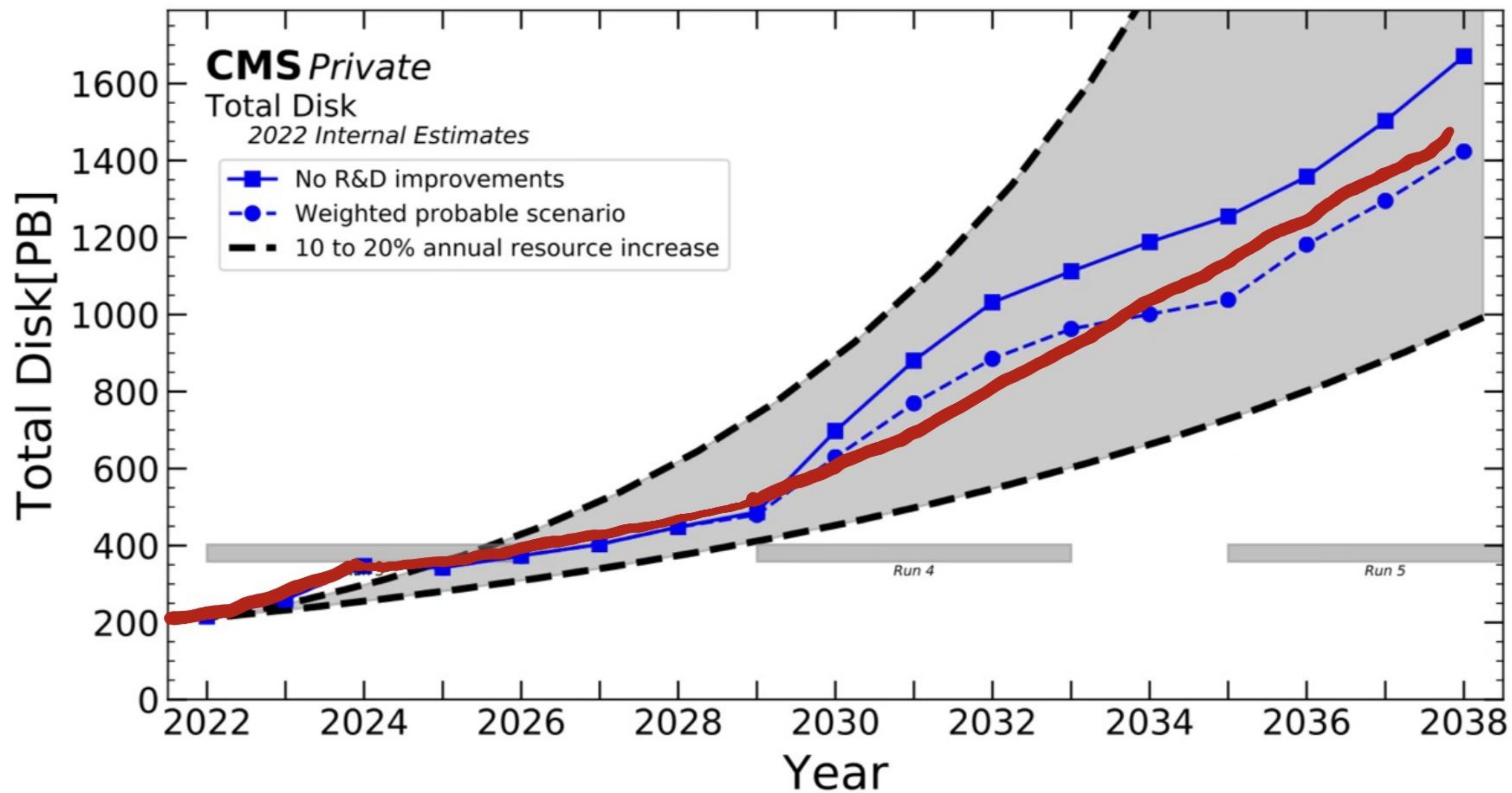
- Above are (last year's) CMS model forecast results into HL-LHC for CPU, Disk and Tape.
- The blue lines are model forecasts, with dotted being most likely, after R&D goals realized
- The grey band follows 5 and 10% hardware/\$ improvement
- Differences within the two sets give an indication of their uncertainties
 - For blue lines, its to what extent CMS is able to achieve expected improvements in software and capability
 - For grey its to what extent hardware cost/resource improves over time.
- With these forecasts and uncertainties in mind, in what follows we generally “bet” on CMS succeeding in making expected improvement, and aim to simplify.

CMS CPU modeling



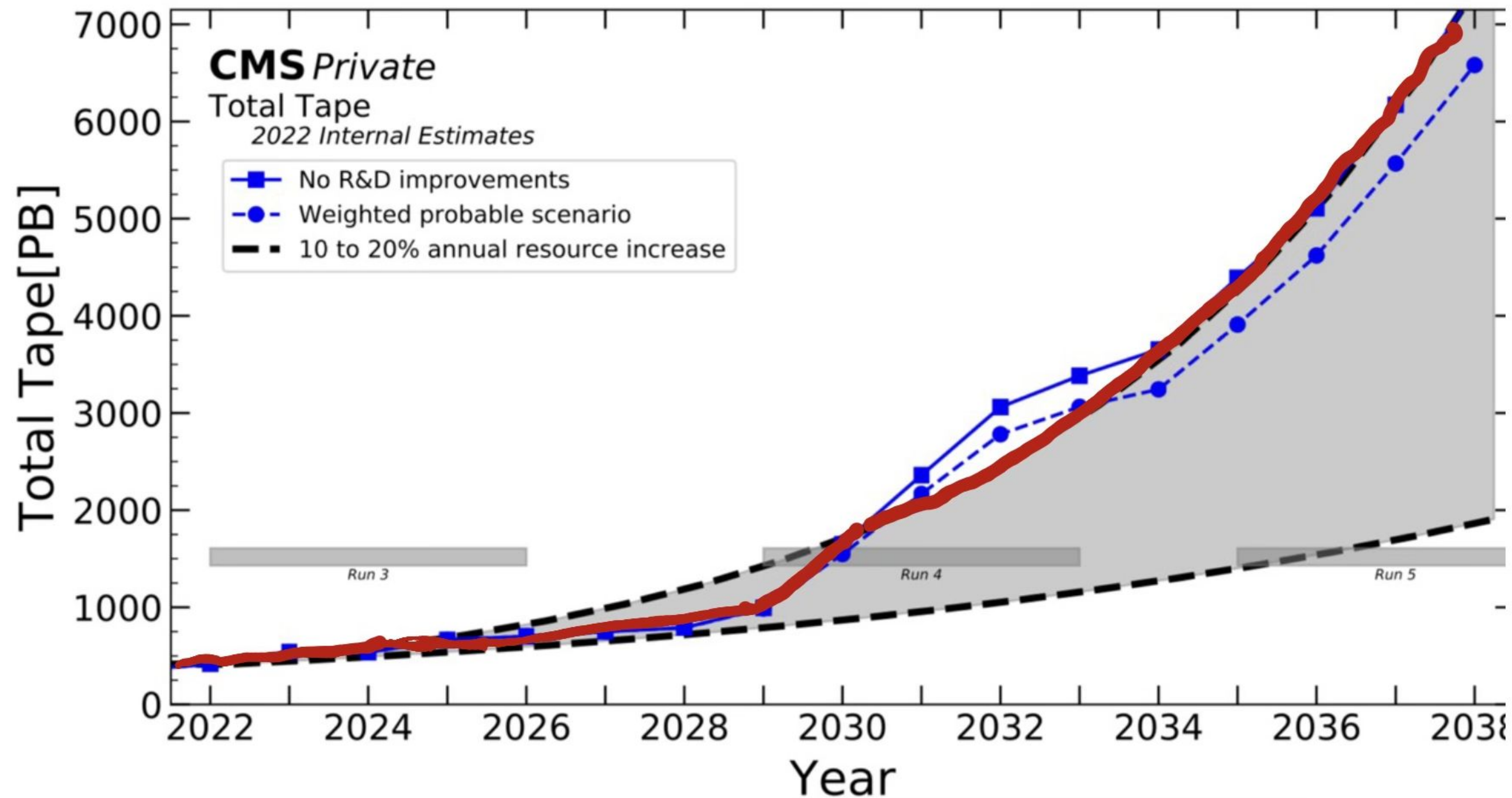
- Based on David Lange's latest model projections in June
- For CPU after 2024, we stay with 5% increase assumption to meet 10% line in 2028
- Catch up to 20% in 2028-30 (bump that doubles CPU capacity from 2028)
- Stay with 10% following.

CMS Disk Modeling



- For disk, after 2024 we stay with 10% increases until 2029, 15% following.

CMS Tape modeling



- For tape we stay with the ~10% /year increase through 2029, then catch up to 20% curve by 2030. After that 20% (but thats beyond 2030 budgets we entertain here)

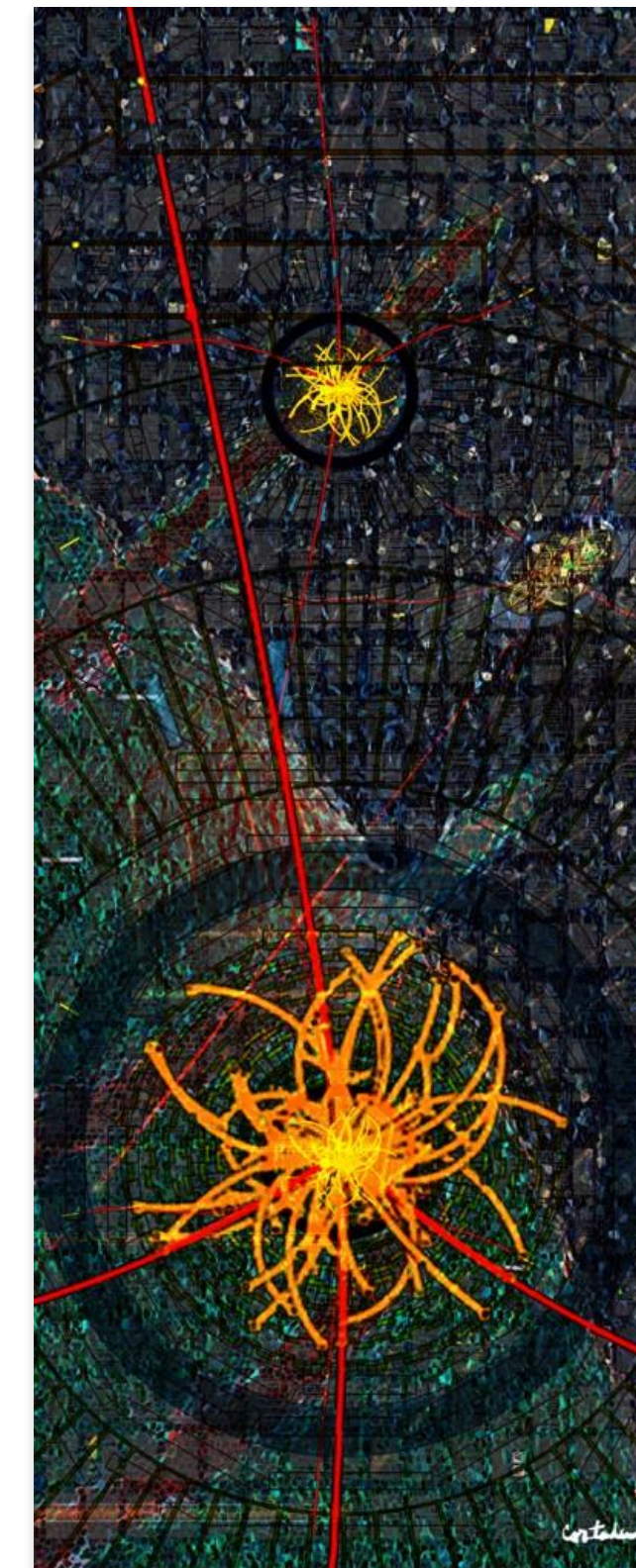
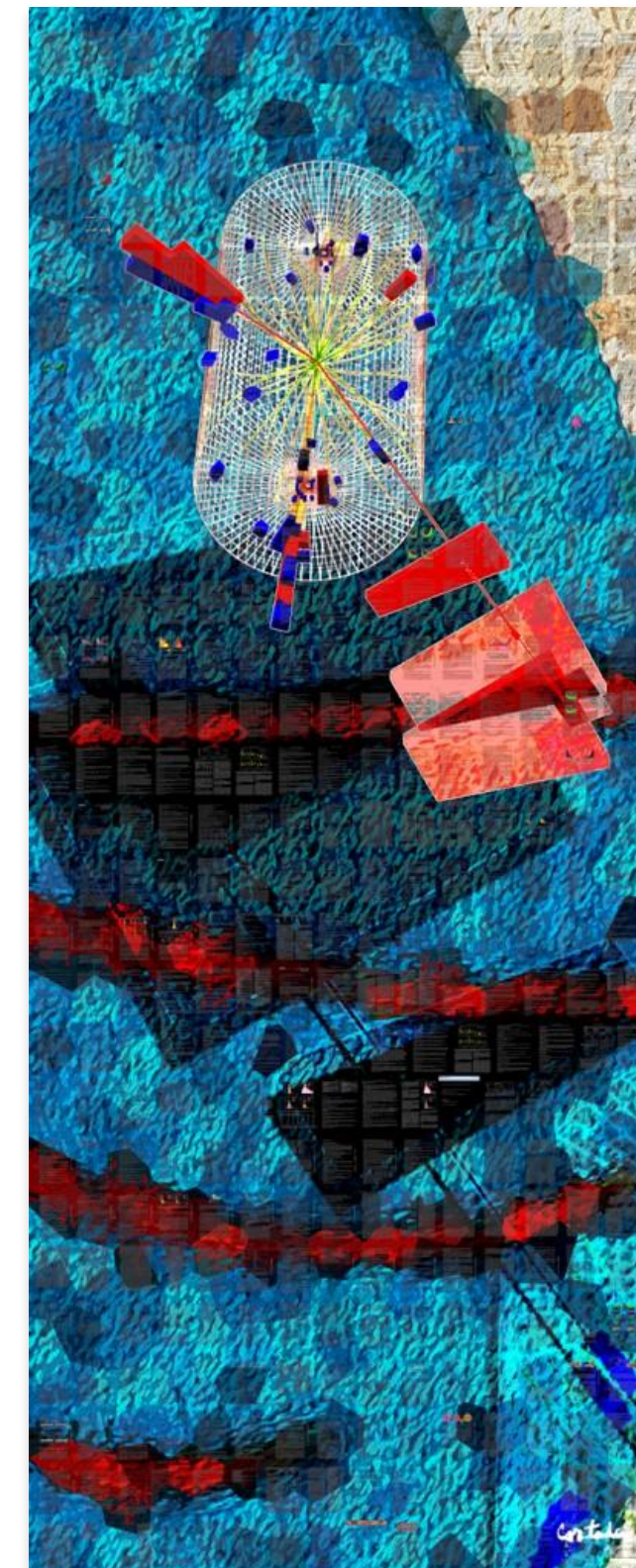
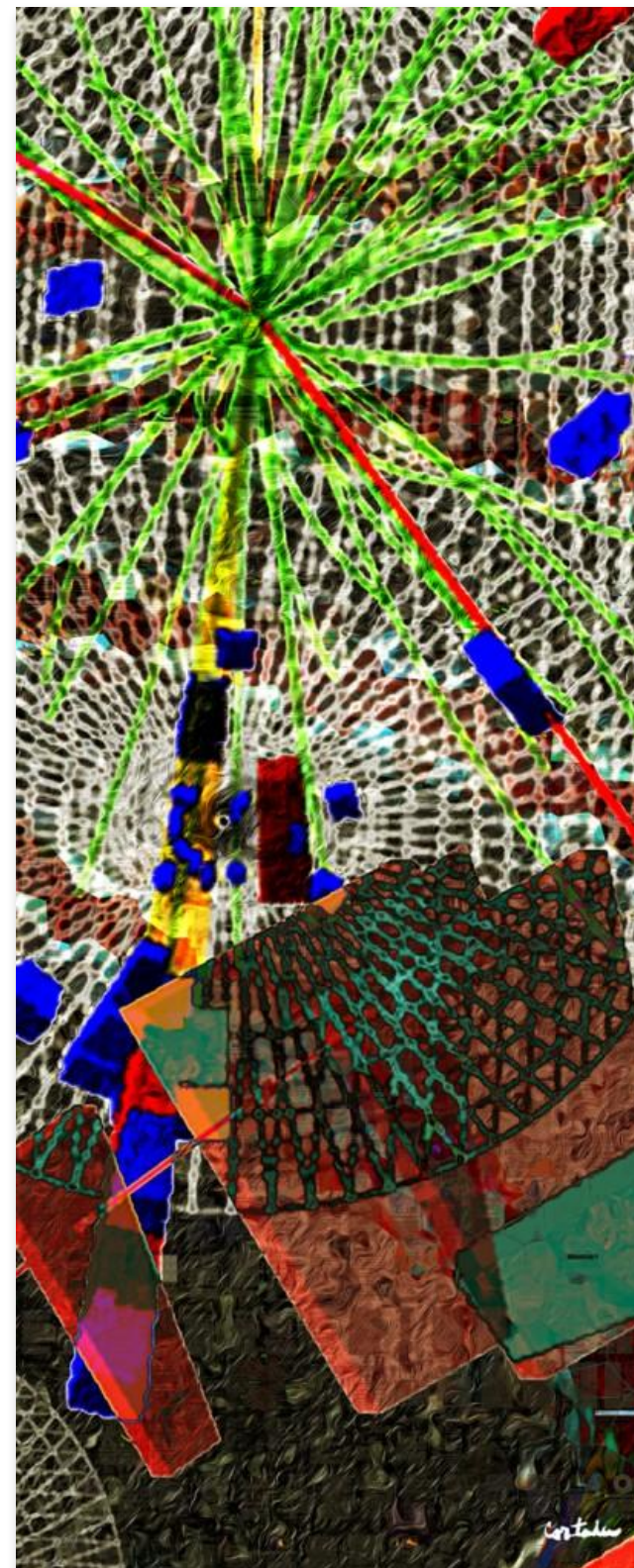
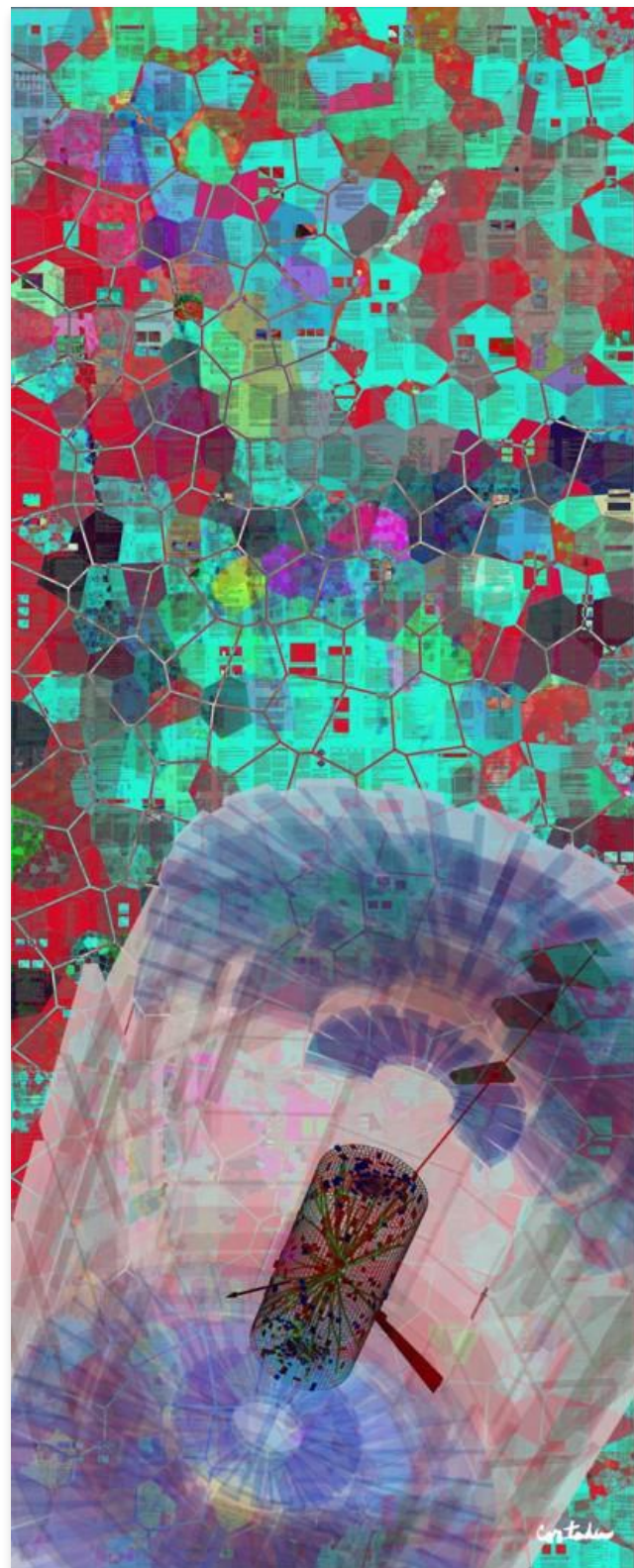
U.S. CMS Part of Estimated Resource Requests

Year	Run 3				LS3			Run 4				LS4		Run 5			
	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038
T1 CPU (kHS06)	292	320	336	352	370	388	407	693	1178	1295	1425	1567	1724	1897	2086	2295	2524
Increase %	9.00%	9.60%	5.00%	4.80%	5.00%	5.00%	5.00%	70.00%	70.00%	10.00%	10.00%	10.00%	10.00%	10.00%	10.00%	10.00%	10.00%
T1 Disk (PB)	33.2	39.2	44	48.8	53.7	59	65	71.4	82.2	94.5	108.7	125	143.7	165.3	190.1	218.6	251.3
Increase %	8.00%	18.10%	12.20%	10.90%	10.00%	10.00%	10.00%	10.00%	15.00%	15.00%	15.00%	15.00%	15.00%	15.00%	15.00%	15.00%	15.00%
T1 Tape (PB)	104	126.4	150.4	174.4	191.8	211	232.1	255.3	449.1	538.9	646.7	776	931.2	1117.5	1341	1609.2	1931
Tape Increase PB	12	22.4	24	24	17.4	19.2	21.1	23.2	193.8	89.8	107.8	129.3	155.2	186.3	223.5	268.2	321.8
T2 CPU (kHS06)	300	338	354	372	391	410	431	732	1,245	1,369	1,506	1,657	1,823	2,005	2,205	2,426	2,668
% increase	12.1%	12.5%	5.0%	5.0%	5.0%	5.0%	5.0%	70.0%	70.0%	10.0%	10.0%	10.0%	10.0%	10.0%	10.0%	10.0%	10.0%
T2 Disk (TB)	25	29	31	34	38	42	46	50	58	67	77	88	101	116	134	154	177
% increase	6.5%	19.4%	6.8%	10.0%	10.0%	10.0%	10.0%	10.0%	15.0%	15.0%	15.0%	15.0%	15.0%	15.0%	15.0%	15.0%	15.0%

Year	Run 3				LS3		
	2022	2023	2024	2025	2026	2027	2028
T1 CPU (kHS06)	292	320	336	352	370	388	407
Increase %	9.00%	9.60%	5.00%	4.80%	5.00%	5.00%	5.00%
T1 Disk (PB)	33.2	39.2	44	48.8	53.7	59	65
Increase %	8.00%	18.10%	12.20%	10.90%	10.00%	10.00%	10.00%
T1 Tape (PB)	104	126.4	150.4	174.4	191.8	211	232.1
Tape Increase PB	12	22.4	24	24	17.4	19.2	21.1
T2 CPU (kHS06)	300	338	354	372	391	410	431
% increase	12.1%	12.5%	5.0%	5.0%	5.0%	5.0%	5.0%
T2 Disk (TB)	25	29	31	34	38	42	46
% increase	6.5%	19.4%	6.8%	10.0%	10.0%	10.0%	10.0%

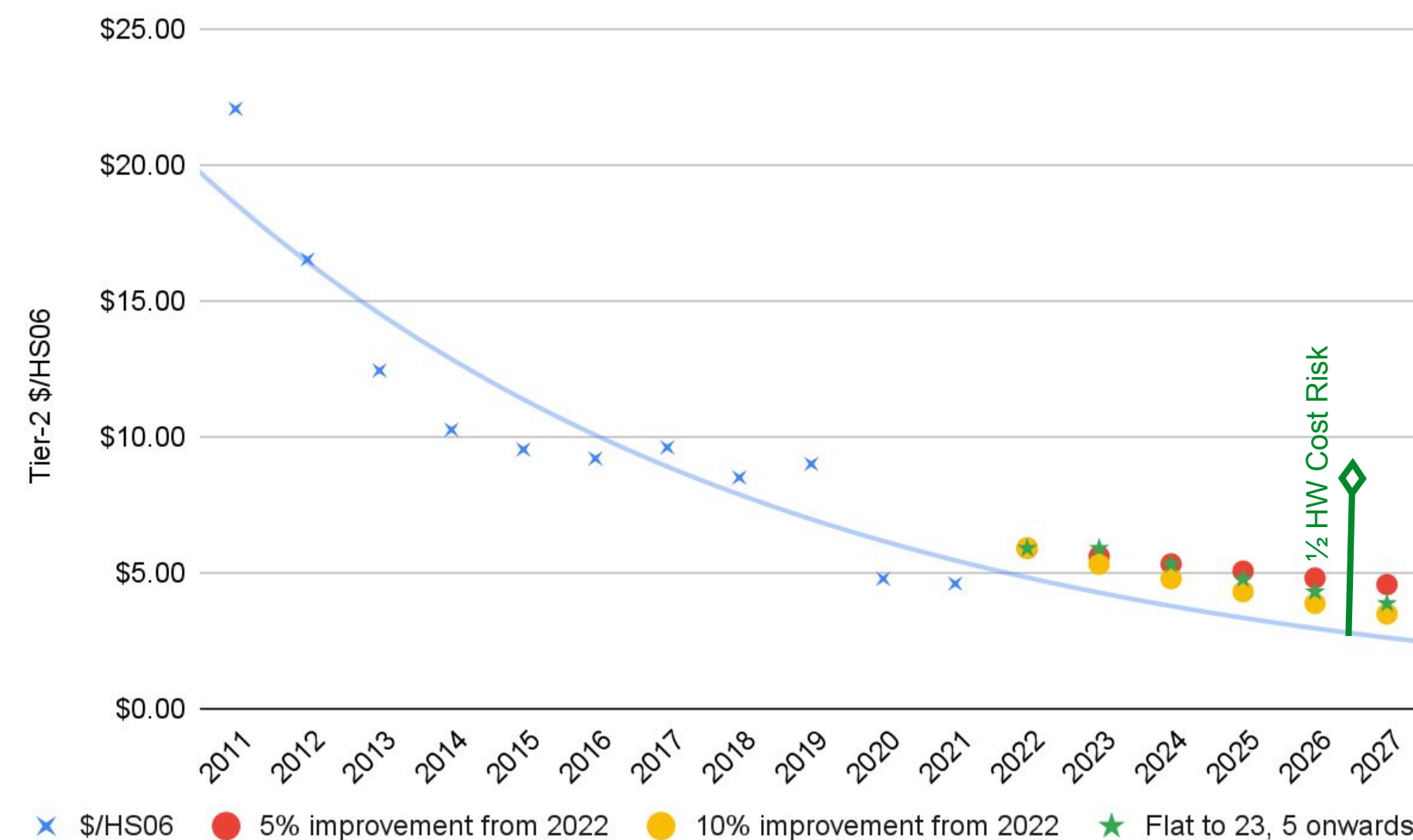
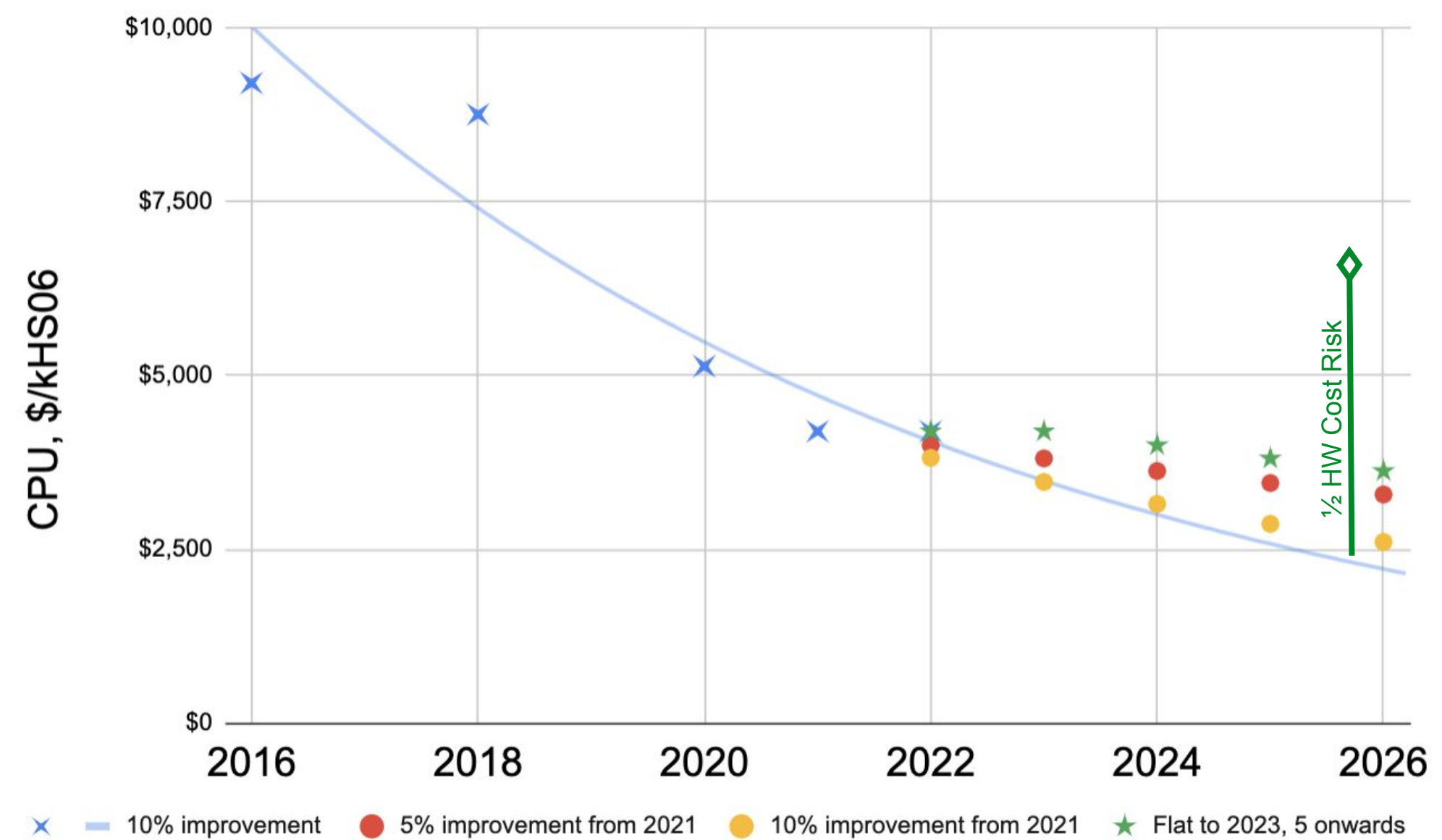
- Increases are cumulative beginning with estimated '24 pledges
- Top is resource estimates through 2038
 - Below is zoom through end of LS3
- With this we have how much we think we need to provide, next toward how much it costs

Second: How Much Will It Cost?



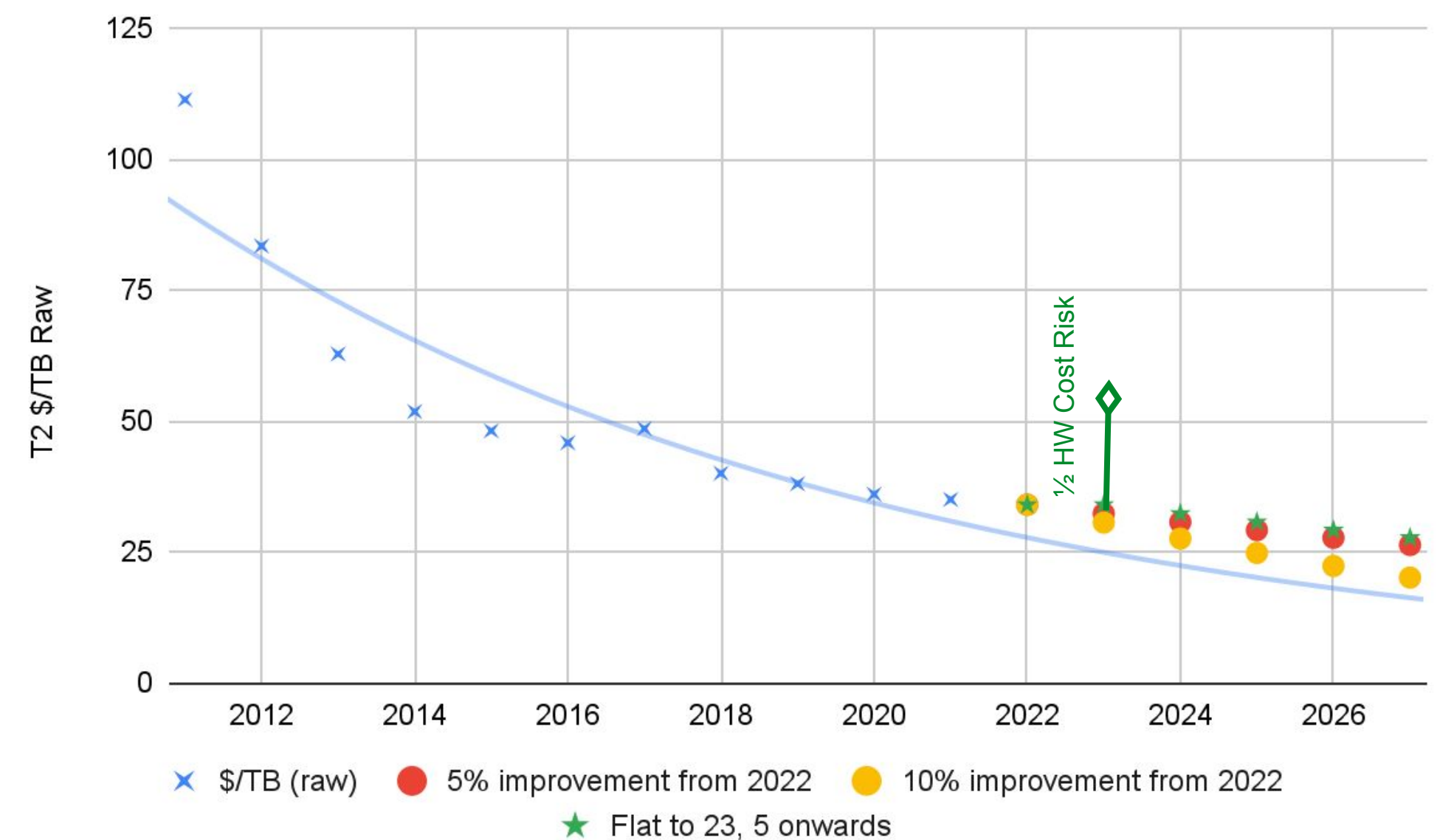
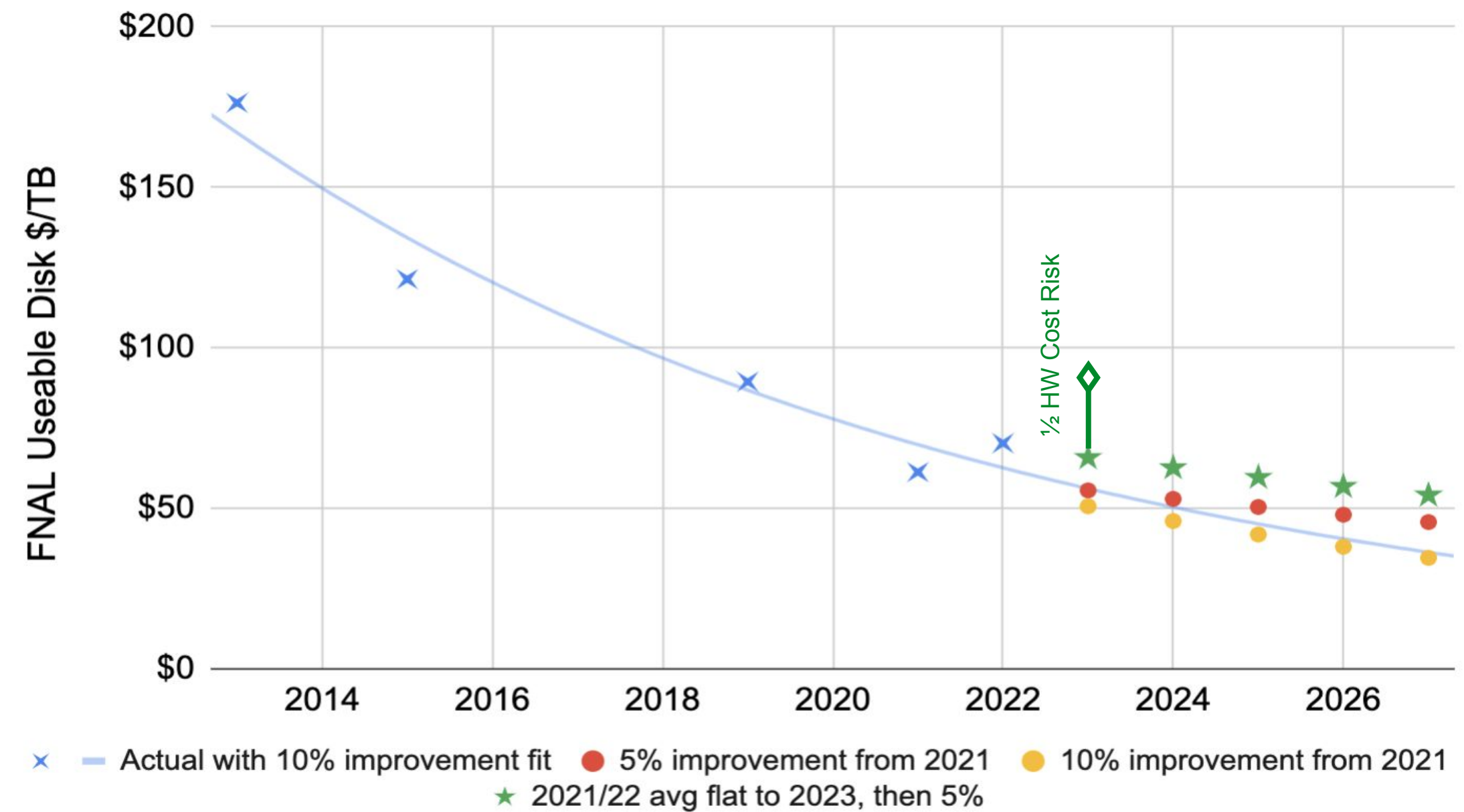
Hardware cost prognostications – CPU

- Blue points are real numbers from past purchases
- Red (5%), Yellow (10%) are extrapolations from 2021/2022 (T1/T2)
- Green stars are what is used in estimates, \$4.20/HS06 for T1, \$5.90 for T2.
- NB: FNAL was able to re-use '21 pricing in '22, so they don't yet see the '22 price increases
- Betting on costs staying flat through 2023, covering increases by risks.
- Afterwards resuming conservative 5% improvement.
- ½ of \$750k hardware cost risk value indicated on 2023 as a vertical bar

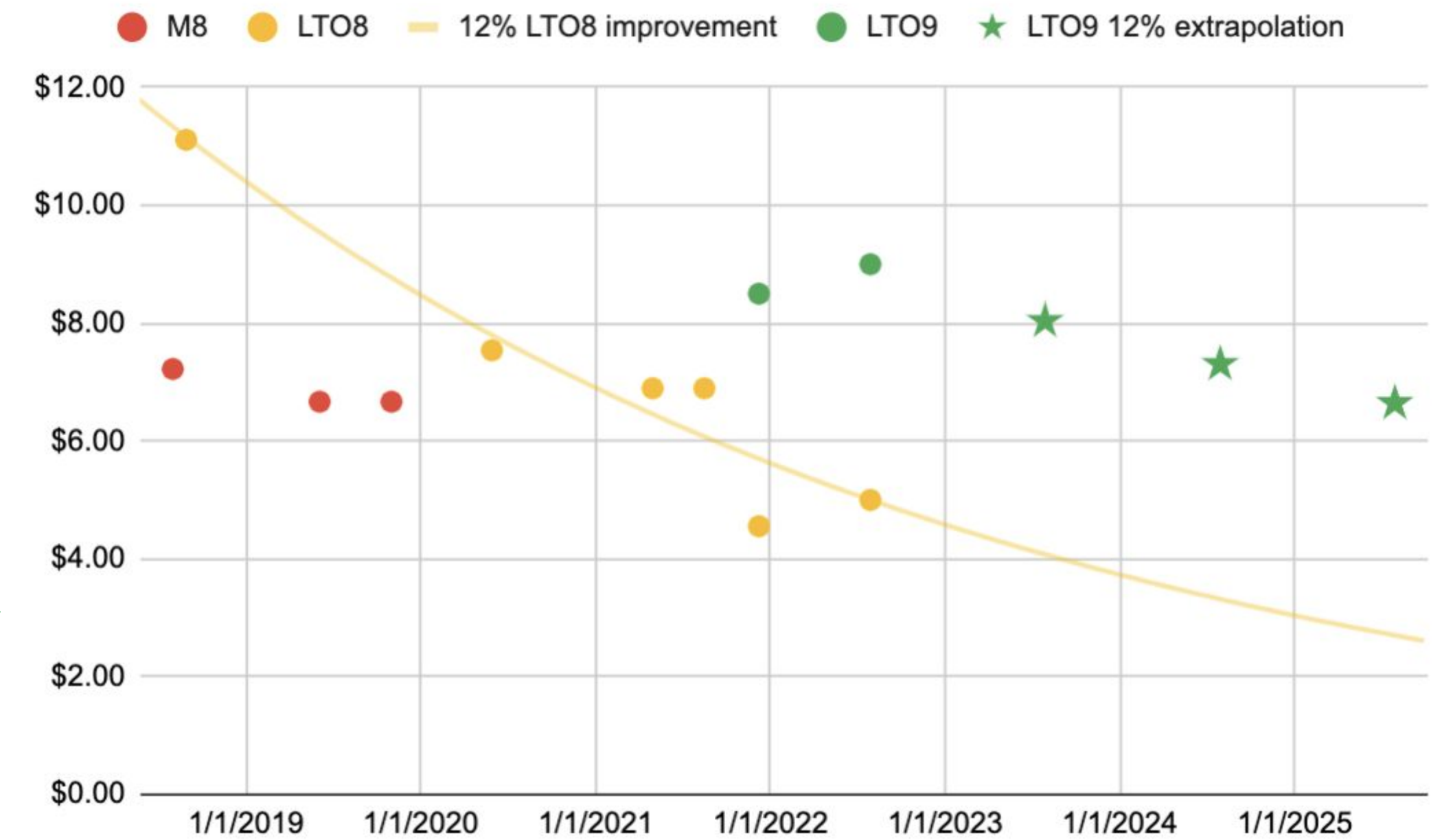
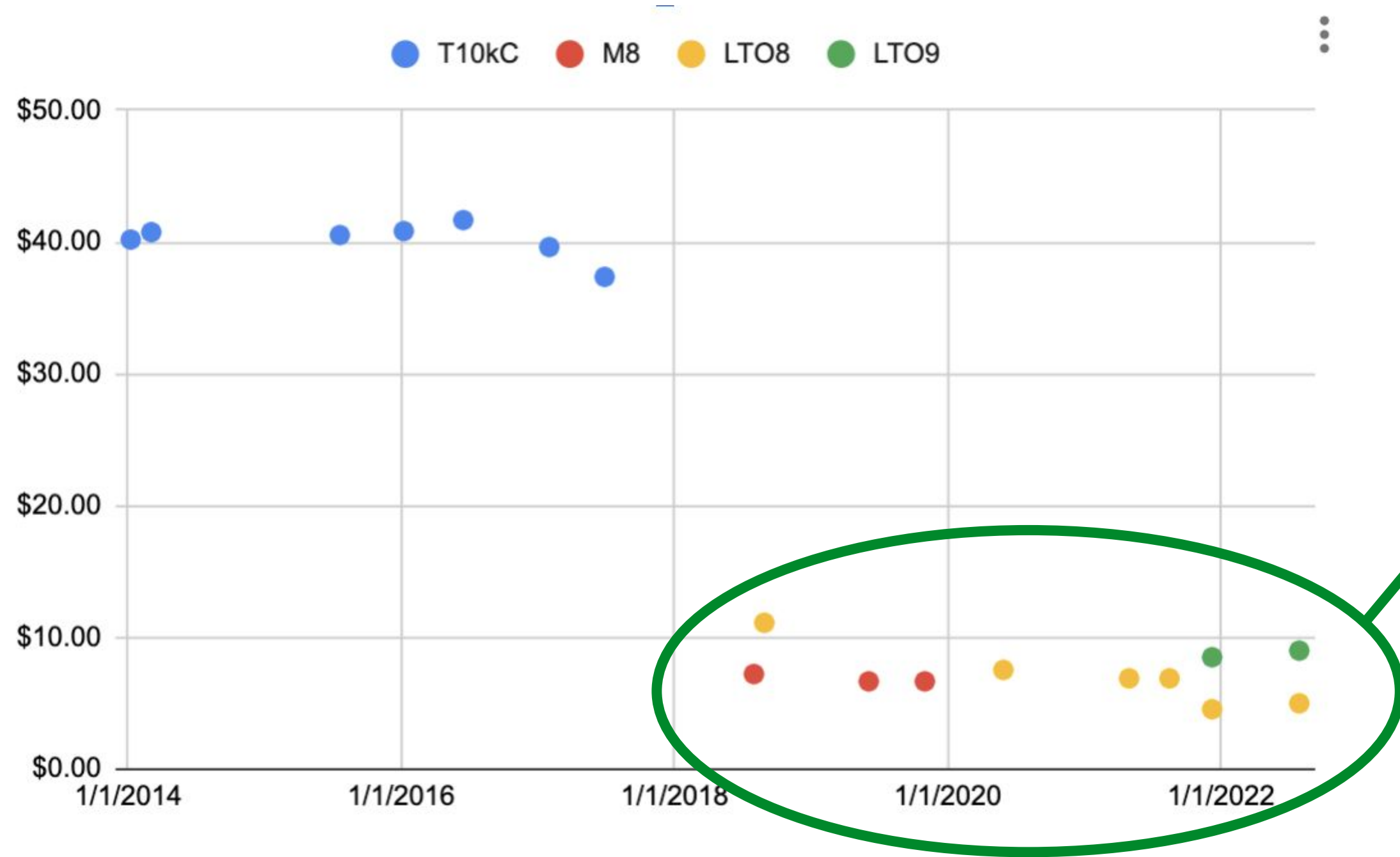


Hardware Cost Prognostications – Disk

- As for CPU, blue points are past purchase numbers
- These are **useable/raw** (T1/T2) disk numbers, HW raid factor and server cost included
- Green stars indicate what we're using in estimates
- 2023 = average of 2021, 2022 (\$66/TB RAW (T1), \$34.5/TB Usable (T2)), after usual conservative 5% improvement
- ½ of hardware cost risk shown as vertical bar in 2023



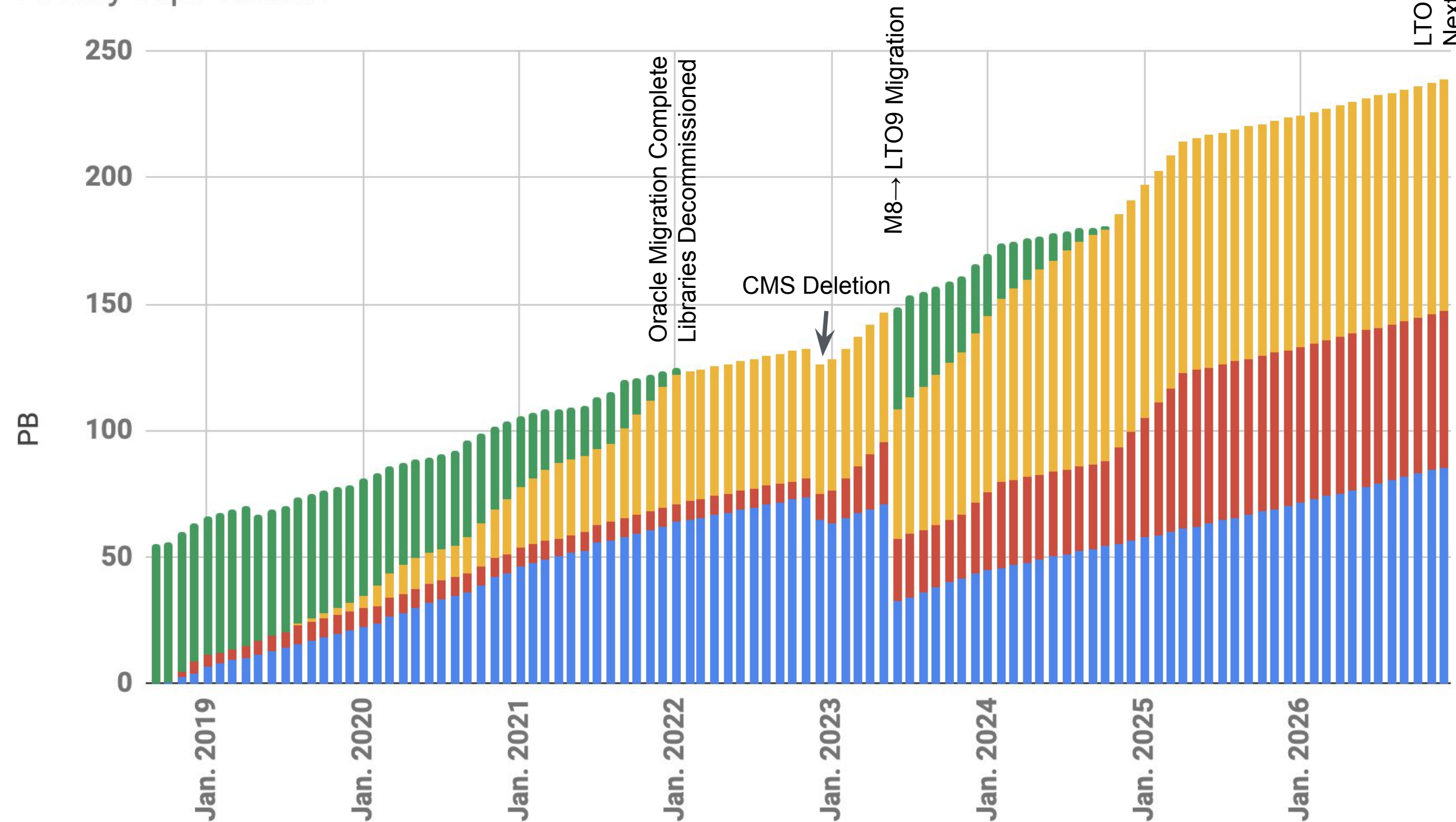
Hardware Cost Prognostications Tape Media



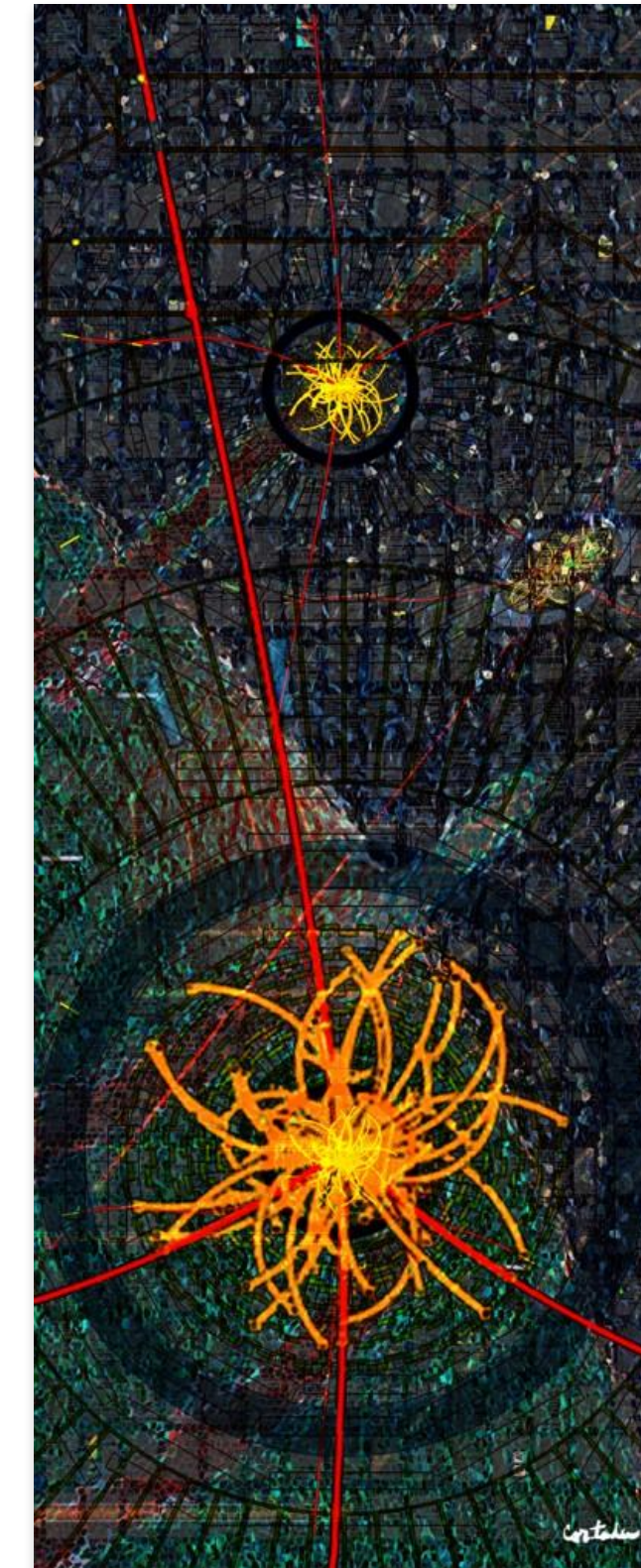
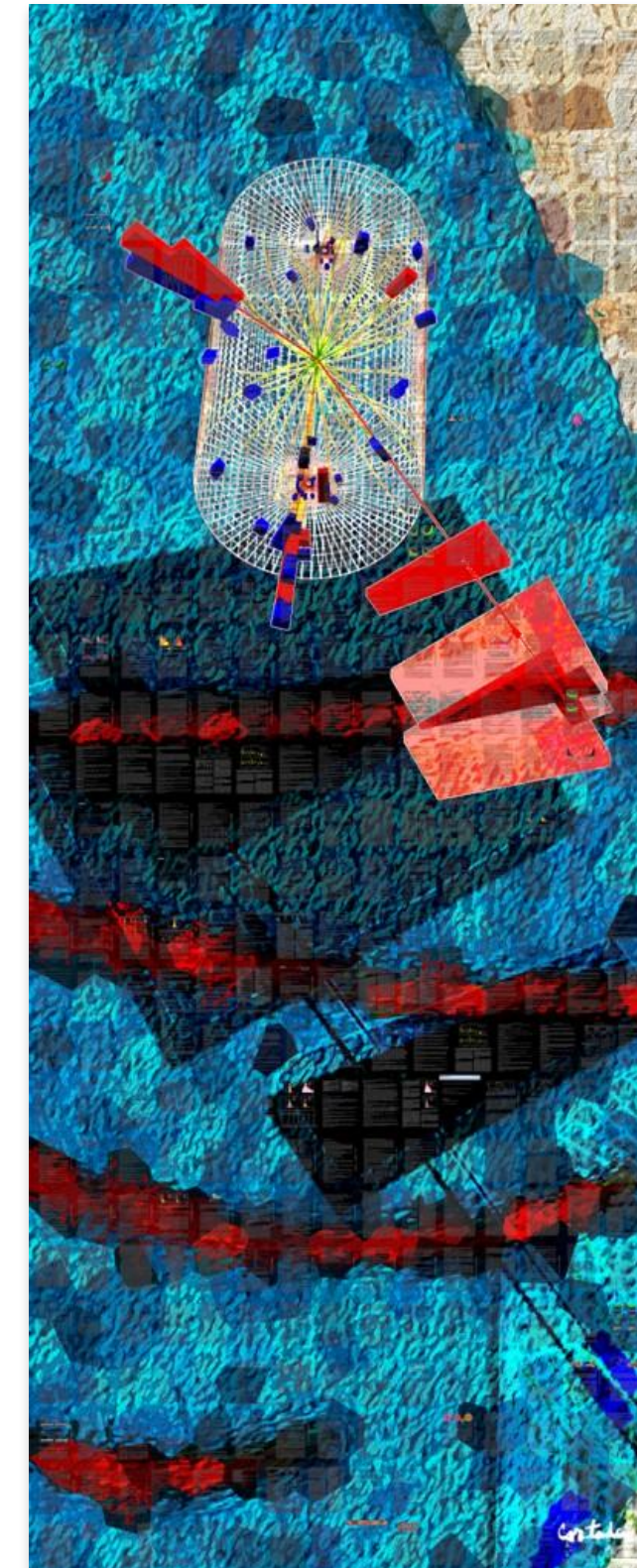
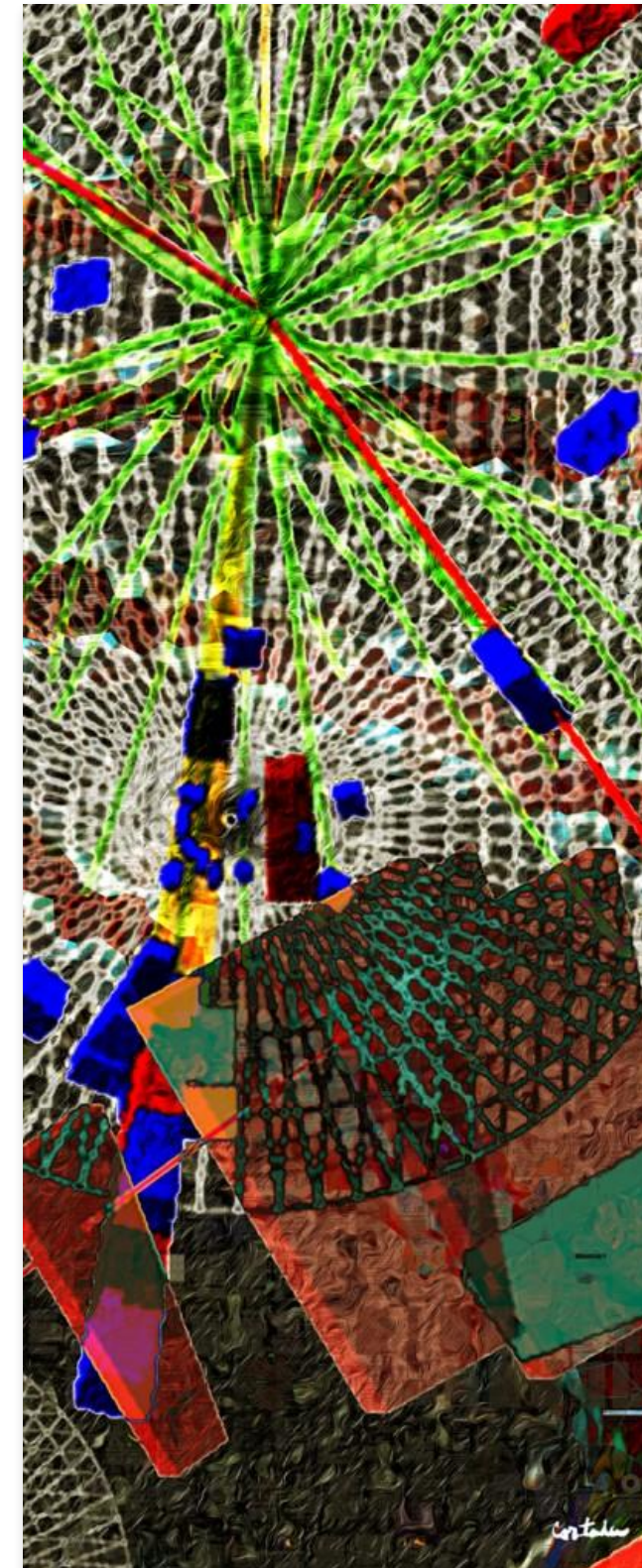
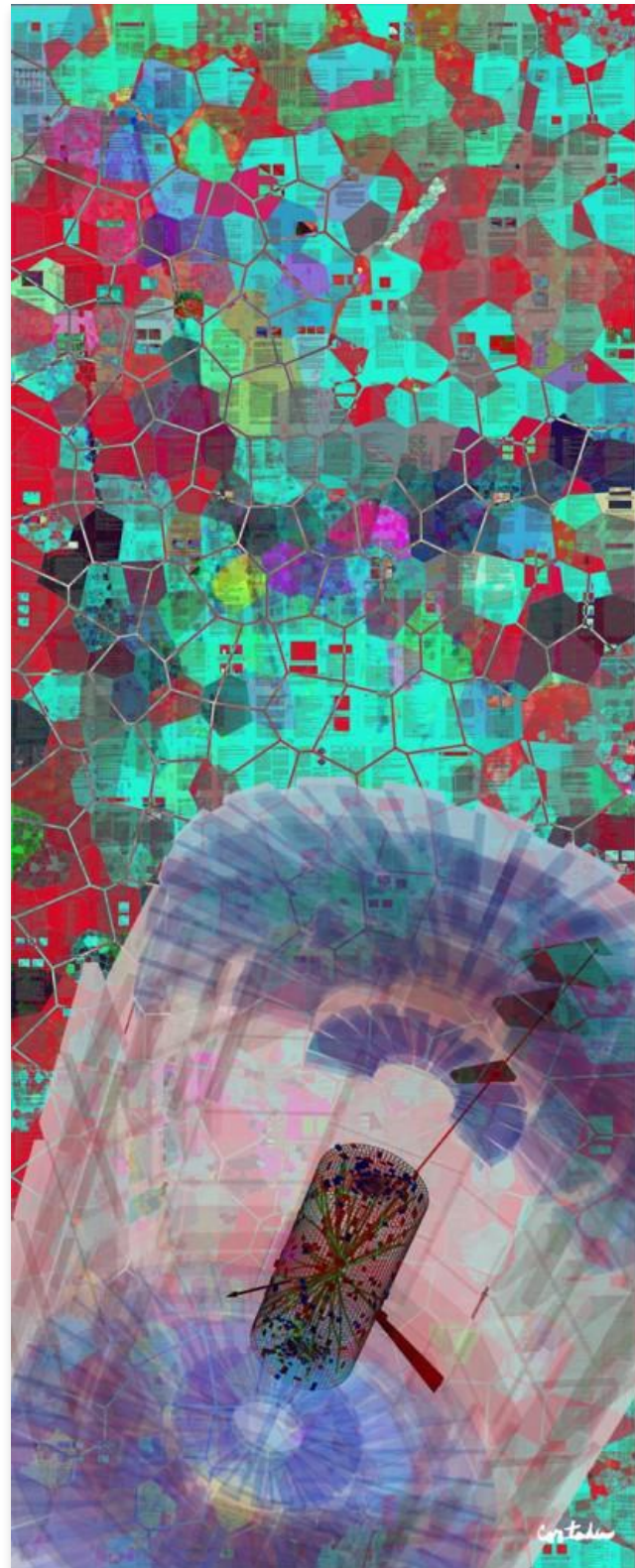
- Currently using LTO8 in both libraries, LTO sees 14% y-y improvement, we see 12%
 - As in we see what the market has been seeing
- **expect move to LTO9 in 2023, depending on market**
- “Jump” when LTO9 hits \$7-8/TB range, but buying some now for M8 migration

- On right shows FNAL tape planning (actual through Aug 2022, projections after)
- Blue is pp additions, Yellow is migrated to new media,
 - Red is HI (not included in budgets here)
- **CMS Deleted 17 PB at FNAL, amount recovered showed in plot, more comes with squeezing in 2023**
- Migration from M8 to LTO9, predominantly from IBM→Spectra library to balance libraries
- **Rumor that LTO10 (2024) WILL NOT be backward compatible for first time**
- **Likely it will make sense to skip LTO10 , new library with LTO11 in 2028.**
 - **Tape drive throughput has basically plateaued – hope for 400MB/s to 600MB/s with LTO11.**
- **When LTO11 available, migrate LTO8 to new format**

Facility Tape Written



Profit!



The Whole Shebang

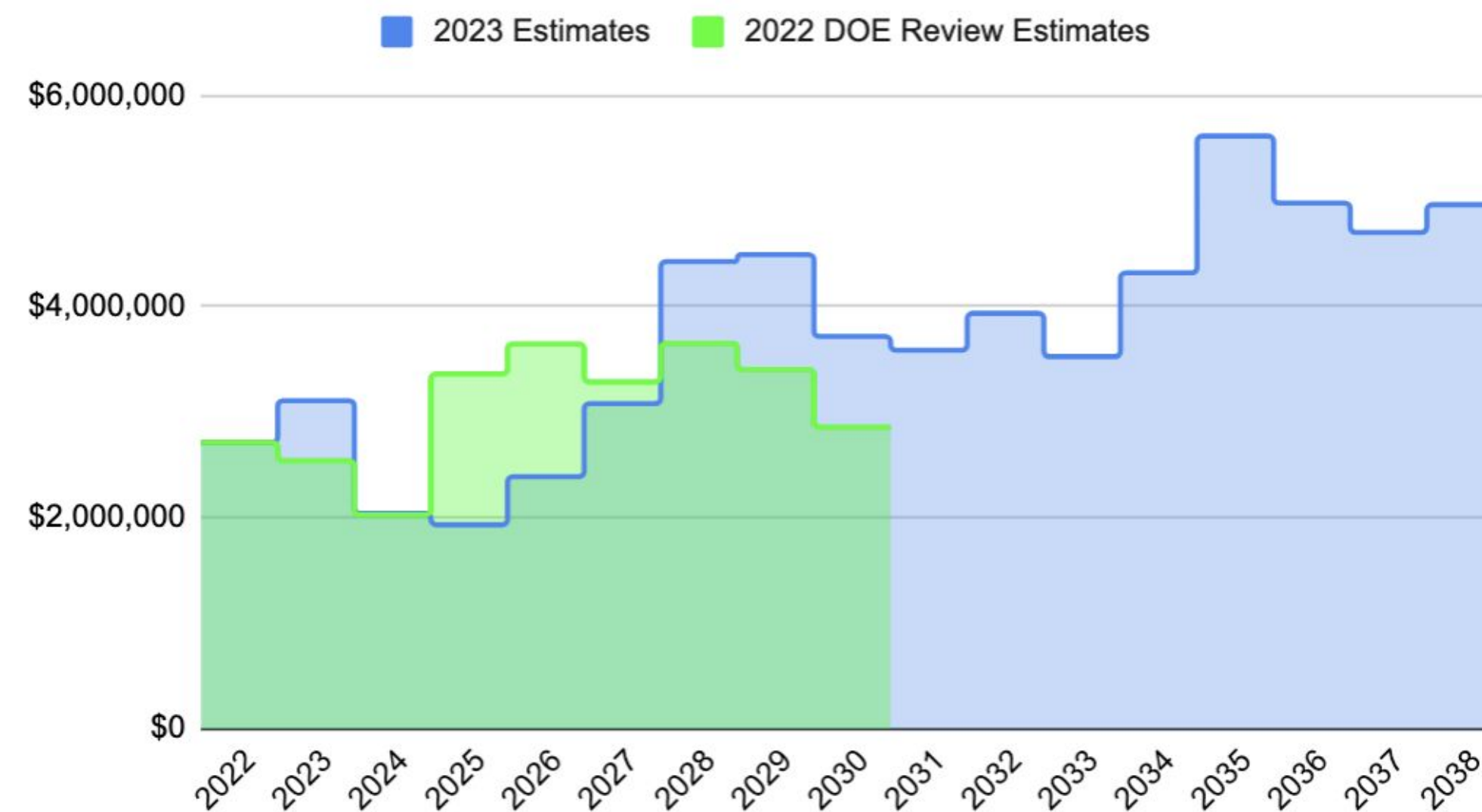
	Run 3		LS3			Run 4				LS4		Run 5					
Year	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038
CPU	\$628,616	\$650,307	\$347,695	\$257,366	\$384,671	\$492,505	\$1,726,509	\$2,035,279	\$902,279	\$725,403	\$681,421	\$793,070	\$891,888	\$1,792,111	\$2,035,895	\$1,236,199	\$1,157,217
Disk	\$1,074,938	\$963,951	\$735,187	\$558,869	\$834,636	\$1,206,936	\$467,118	\$1,048,245	\$1,407,267	\$1,452,104	\$1,367,405	\$1,324,596	\$1,594,765	\$1,937,909	\$1,536,681	\$2,038,653	\$1,400,280
Tape	\$656,329	\$1,095,186	\$691,137	\$744,301	\$744,301	\$956,961	\$1,807,593	\$983,543	\$1,196,201	\$983,543	\$1,462,024	\$983,543	\$1,408,859	\$1,462,024	\$983,543	\$983,543	\$983,543
Network	\$318,986	\$425,315	\$318,986	\$425,315	\$425,315	\$425,316	\$425,316	\$425,316	\$425,316	\$425,316	\$425,316	\$425,316	\$425,316	\$425,316	\$425,316	\$425,316	\$425,316
Total	\$2,713,677	\$3,134,759	\$2,093,005	\$1,985,852	\$2,388,924	\$3,081,718	\$4,426,535	\$4,492,384	\$3,931,062	\$3,586,367	\$3,936,166	\$3,526,525	\$4,320,828	\$5,617,360	\$4,981,334	\$4,703,711	\$3,966,356

Preliminary
Not for
Distribution

	Run 3				LS3		
Year	2022	2023	2024	2025	2026	2027	2028
CPU	\$628,616	\$650,307	\$347,695	\$257,366	\$384,671	\$492,505	\$1,726,509
Disk	\$1,074,938	\$963,951	\$735,187	\$558,869	\$834,636	\$1,206,936	\$467,118
Tape	\$656,329	\$1,095,186	\$691,137	\$744,301	\$744,301	\$956,961	\$1,807,593
Network	\$318,986	\$425,315	\$318,986	\$425,315	\$425,315	\$425,316	\$425,316
Total	\$2,713,677	\$3,134,759	\$2,093,005	\$1,985,852	\$2,388,924	\$3,081,718	\$4,426,535

- **2023 will be higher than past forecasts predominantly due to addition of new tape buffer**
 - Lesser effect is cost inflation
- **Can see effect of schedule change in push of Cost from 2025-6 into later years**

2022 vs 2023 Budget Forecasting comparisons



- **FNAL Facility is Performed Well for first year of RUN 3**
 - Continues to perform well and set the bar for other CMS sites ([thank you!](#))
 - Year 2 will be compressed and include a lot of data fast (especially during HI datataking in October)
- CMS Schedule has been updated, additional Run 3 year in 2025, HL-LHC now begins in 2029.
- 2023 Physics datataking begins in April
- Heavy Ion Run expected late September-October
- 2023 Run ends on October 30 (earlier than past years)
- Risks to watch (still) are supply chain and Russia/Ukraine situation
- Purchasing process for 2024 started, large disk and CPU purchases
- Working towards CTA, preparing for data challenge in 2024 will be focuses this year (beyond Run 3 data-taking).