NSF LEADERSHIP-CLASS COMPUTING FACILITY

Horizon: A MUG Shot

Dan Stanzione
Executive Director, TACC
Associate Vice President for Research, UT-Austin

August, 2026

TACC TEXAS ADVANCED COMPUTING CENTER

TEXAS The University of Texas at Austin

TACC @ 25

- TACC turned 25 this month!
- Since June of 2001:
  - >125,000 users.
  - 4 of those users got Nobel prizes
- Last few years:
  - New user job submission every 7 seconds
    - On Frontera, Stampede, Lonestar, and Vista
    - ~16,000 nodes, 1M CPU cores, 2k GPUs.
  - >10k tickets per year.
  - >20k users active each year.

MVAPICH at TACC

- I say this is every year, but I can't overstate how important it is to have the MVAPICH team on board.
- MVAPICH is part of the stack on every Intel, AMD, NVIDIA system we have, including all the CPU ones and all the GPU ones!
- Our strategy for every system is to have two compilers and two MPI stacks.
  - Most vendors would prefer you have one!
  - Our next system will have MVAPICH and OpenMPI
- Not only do we get a tuned stack for our systems, and early insight into potential network issues, we also get to keep the vendors honest and get an invaluable tool for debugging!
- We've been working with the OSU team for 20+ years, have them as funded partners through the lifetime of the Leadership Facility, and hope it goes on forever!

The NSF Leadership Class Computing Facility

- A more capable follow-on to the current (aging) NSF leadership systems
- A more holistic and collaborative view of how we support “leadership applications”.
- A long term investment to match the science missions we support.

The NSF Leadership Class Computing Facility

- The most complicated deployments we've ever done.
- In the worst possible market conditions
- To do the most complex range of science and engineering we've ever done.
- With the most complicated administrative structure we have ever had.

Distributed Centers

- AUCC – Diverse pathways to Data Science and Computing
  - Leverage five HBCUs, that are co-investing in data science as a focus area.
- NCSA - Understand how to exploit new chips for AI for our application mix
  - And in particular how to improve I/O to them, often an accelerator afterthought.
- PSC – Data Intensive Computing
  - Data Mirror for published archives, a focus on protected data and FAIR access.
- SDSC – Testbeds focused on a subset of Science Cases:
  - ML Inference in scientific workflow
  - Exabyte size instrument data workflows
  - Democratization of access (via PATH/OSG).

And Other Partners

- OSU Continues to provide our open source network stack!
  - New version tuned for (some) Blackwell just released in beta (MVAPICH+ 5.1.0b)
  - Used our NVL-72 cabinet.
- Cornell continues to partner for online training content.
- Vendors:
  - NVIDIA, Dell, VAST, Cambridge, Sabey Datacenters
  - And many others with smaller systems.

The Markets are, of course, crushing us.

- It’s good that this was the first NSF project where we could carry a contingency budget. We needed it.
- This is really difficult for all of us, but especially smaller centers.
  - For 20 years our “real” price per node for a big cluster was ~$6k-8k.
  - An NVL-4 node in Blackwell is more like $120k
- Your user experience is very different if instead of 150 nodes, you have 8.

QUARTER	30TB TLC SSD	30TB QLC SSD	QLC VS HDD	30TB HDD
Q2'2025	$3,062	$2,450	4.9x	$495
Q3'2025	$3,460	$2,566	5.2x	$495
Q4'2025	$7,765	$6,212	10.7x	$580
Q1'2026	$10,950	$9,461	14.2x	$668
Q2'2026	$13,140	$11,353	14.8x	$768

A Note on Markets

- Memory – 500% of last year's price
- Storage – 700% of last year's price
- If this sounds abstract, the RAM for the CPU portion of Horizon went from $9.5M to $95M in 8 months.
- A "typical" 2 GPU server was $20k in 2021; today it's about $150k, and needs $500k of new cooling infrastructure.
- This is very simply a supply/demand imbalance – about $100B was spent worldwide on servers in 2020. The big 4 alone are spending $750B on AI infrastructure this year. Production can't grow that fast.

This is More than a Price Increase
It is an Existential Threat for Academic Computing

- The state of funding is already somewhat precarious (masked by the state/university willingness to invest in AI).
- But, assuming you have an in-room water system to support chilling units, running **one rack** of what the clouds are running (Vera-Rubin) will be a ~$9-10M investment in 2027.
  - With likely a 9-18 month lead time.
  - And you will not get spare parts when things fail in a timely manner.
- You can get a mostly crippled RTX server with x86 – necessitating a different software stack and different optimizations – for maybe $200k (and you don't want to get within 50 feet of it in your datacenter).
- Orders under a Billion dollars essentially don't matter – you will be at the back of the line (all the majors are tens of billions behind on orders to the CSPs).
- Why worry about scaling to 1,024 nodes with your network software? Thousand node clusters are for people whose hardware budget is over $100M. (Or can pay $10k/hour for 1,024 GPUs of NVL-72 at Amazon).
- Running a datacenter does not make you very popular right now…

For HPC…

- Almost no Blackwell systems made the Top 500, anywhere in the world.
- Yet shipments started at over a thousand cabinets a week many months ago.
  - Need some? Use the cloud, they have all of them.
- At the same time – tried sharing these stacks?
  - I had *one lab* come to me and say running their bot – with export controlled code they couldn't put on the cloud – needed 1.5TB of NVLINK memory and to stay resident for about 18 months.
    - That's a million dollars of hardware you can't touch once it runs their model… for 1.5 years.
- So, when your customers come to you, even with, say, a $100k buy in:
  - You can't get much.
  - You don't know when it will show.
  - The price will change after you order it, so you will get less.
  - It won't go far when you get it.
  - Now that's a model for success…
- Rumor has it **more than 50%** of HPC procurements are ending with **cancellations**.
- More on this later, back to what we've done this year.

The New Archive System is in Production.

- This Ranch will take us past an Exabyte.
- If you are a Ranch user – the old one is still mounted through December.
  - And you've been getting emails about moving data for almost a year – please do so!
- The last Ranch was to do 100PB (2018). The one before, about 10PB. The one in 2006 was to do about 1PB.
- And people ask why I don't worry about deleting old data.
  - Data from last ~3 years: 90% of all data
  - All data generated by humanity since the dawn of time until ~3 years ago: 10% of all data.

Early Access Systems

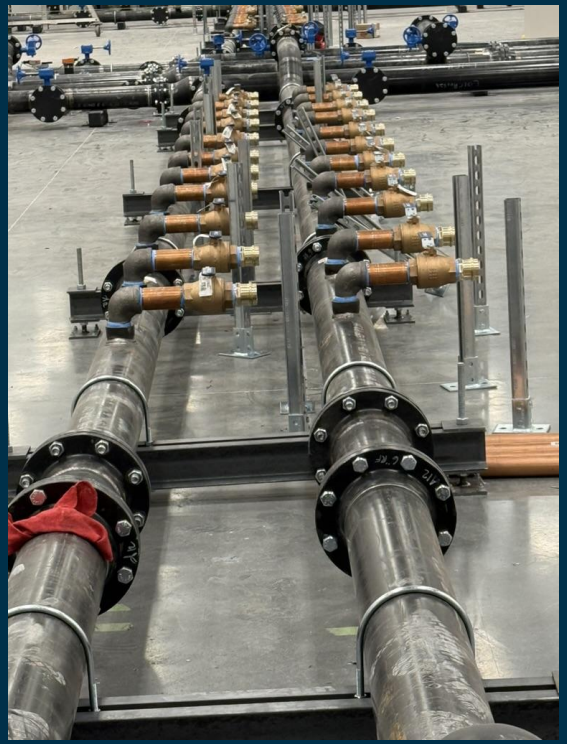
- In summer 2024, we deployed Vista:
  - 600 Grace-Hopper (GH200) nodes (from Gigabyte).
  - 250 Grace-Grace ARM CPU nodes.
- Instrumental in porting software, final performance projections, validating the ARM switch, etc.
  - All has been really successful.
- Six months ago, we added a GB200 NVL-72 rack to Vista.
  - Basis for the latest MVAPICH releases, early work on app tuning for Blackwell, etc.
  - First rack where we juggled NVLINK vs. IB (and NCCL vs. MPI) for transport.
- With Vista and Stampede3, we've been running VAST storage in production, swapping out Lustre in production, at modest scales (3k total clients).
- We've also spent the last year+ working with SambaNova on inference platform prototypes.

Horizon Deployment

- Horizon will replace Frontera (39PF CPU, 2019)
  - 8,368 CPU compute nodes
  - 90 GPU (360GPUS) compute nodes
- Frontera used 6MW, Horizon will be double that.
- So the past year was about building a new home, which is now complete.
- Partnered with Sabey Datacenters in Round Rock, TX to build out a hall to our specs.

Our New Facility: Built to support 250KW/cabinet

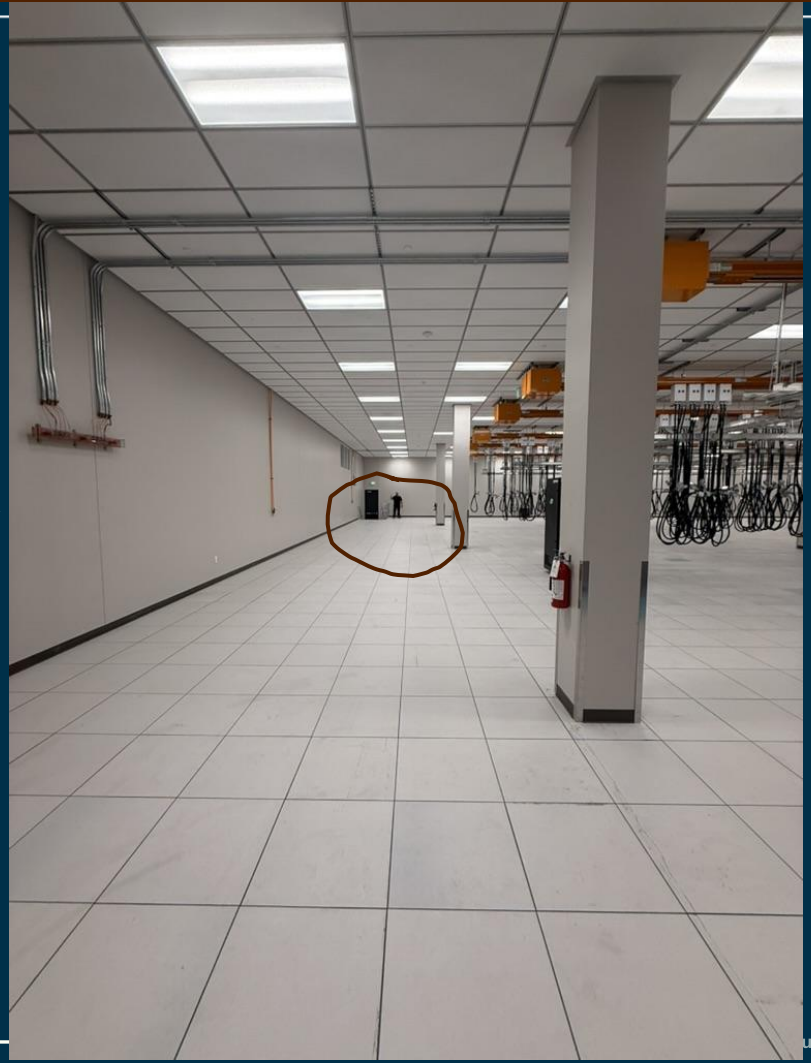
NOTE: Verified by City of Round Rock, the entire Sabey campus currently uses the water of 14 average households.



September 2025
Underfloor cooling install

October 2025
Raised floor install

- Still lots of space to fill
  - For scale, that's Nathaniel ----------→



A Side Note (Back to Markets)

- Everything hard about getting computers? It applies at least equally to every part in a datacenter supply chain.
  - Every fuse, panel, connector, hose, gasket, busway… you name it.
  - We 3D Printed a connector for a PG-25 fill hose… and actually used it.
  - We needed some sanitary flanges for the underfloor pipe. Talked to a vendor:
    - Had 20,000 available at 10AM.
    - At noon, had none available and back-ordered 4 months.
- Any one part can set you back *months* while your rare, expensive compute depreciates.

Horizon (The First LCCF System)

- Accelerated: 2,000 GPU nodes with a Grace socket and two Blackwell GPUs, 800Gb Infiniband (non-blocking).
- CPU: 4,750 Vera-Vera nodes, with two 88-core Vera sockets, 400Gb Infiniband.
- Primary compute capability will be :
  - ~200-300PF double precision (10x Frontera) (15-20x for most apps).
    - More with emulation??
  - ~10 EF ”TF32” precision.
  - ~20 EF bfloat16 precision
  - ~80EF INT4/TF4, if you are into that kind of thing.
- All nodes built in the new Dell Integrated Rack infrastructure
- Shared filesystem, Infiniband connected, roughly 400 usable PB, 100% Solid State. (8/16TB/s W/R).

A Little more on the Network

- 2,000 port non-blocking Fat Tree for GPUs at 800Gbps.
- 4,750 port non-blocking Fat Tree for CPUs at 400Gbps.
- We don’t see the need for full non-blocking *between* those, as jobs spanning those partitions should be rare (if ever, should history hold).
- The “Core”, network (3rd level of tree) bridges in those two networks, with an additional ~800 ports at 400Gbps for storage nodes.
  - Design target of 20TB/s to the storage, and between the fabrics.
- Storage also has a back end IB network… between the filesystem compute servers and the actual Flash arrays, more than 1k additional nodes.
- Total roughly 9k endpoints, 15k+ cables.

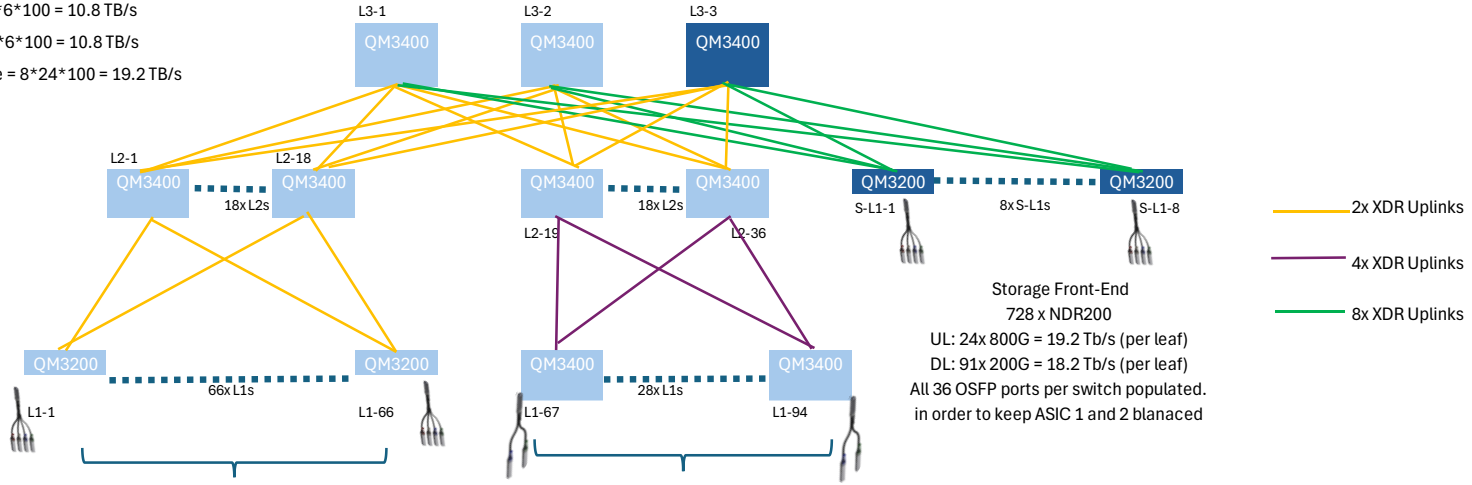
TACC Horizon - Three Level Design

L2 to L3 BW from Islands

VV=18*6*100 = 10.8 TB/s

GB=18*6*100 = 10.8 TB/s

Storage = 8*24*100 = 19.2 TB/s



Storage Front-End
728 x NDR200
UL: 24x 800G = 19.2 Tb/s (per leaf)
DL: 91x 200G = 18.2 Tb/s (per leaf)
All 36 OSFP ports per switch populated.
in order to keep ASIC 1 and 2 blanaced

VV Nodes
4750x XDR400
Non-Blocking
UL: 36x 800G = 28.8 Tb/s (per leaf)
DL: 72 x 400G = 28.8 Tb/s (per leaf)

GB Nodes
2000x XDR (800G) Socket Direct
Non-Blocking
UL: 72x 800G = 57.6 Tb/s (per leaf)
DL: 72x 800G = 57.6 Tb/s (per leaf)

NOTES:

Compute Switches : 66x QM3400s; 66x QM3200s

Storage Switches : 1x QM3400s; 8x QM3200s

VV and GB Single Rack SU with 1 leaf per rack

Storage is 28 racks. Storage L1s will be one every 3.5 racks

GB Nodes will require Socket Direct due to PCIe gen 5

Open Ports (XDR Logical Ports)

- All Compute L1 = 0; Storage L1 : 0
- VV L2 = 2 per L2; GB L2 = 26 per L2
- L3 = 8 per L3

Storage Back End Networking designed and provided by VAST

Storage front-end are **QSFP112** NDR200

Horizon Racks

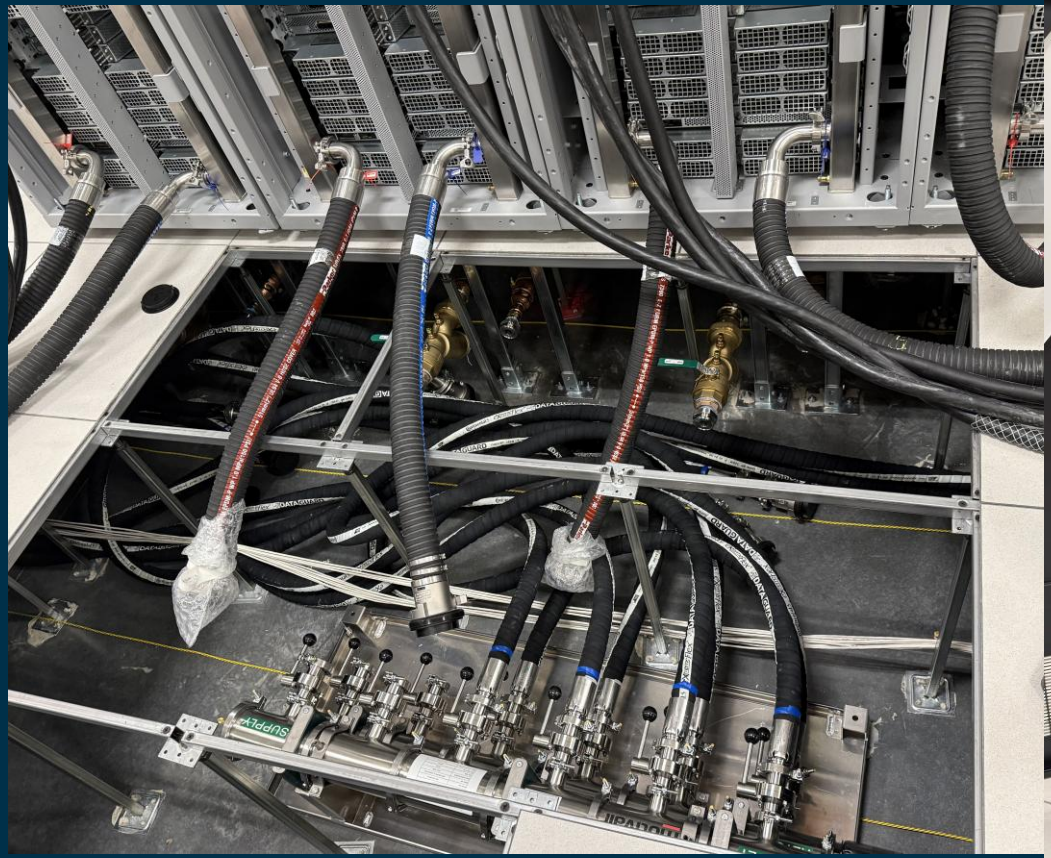
- As we are using the NVIDIA Blackwell “NVL-4” shelf, we aren’t limited to 72 GPUs per cabinet.
  - Also, they will weigh 6,000lbs with liquid and rear doors. (~2 Honda Civics, depending on your trim package).
- Horizon GPU racks will have 144 GPUs, and come in at about 215KW per rack.
- Horizon CPU (Vera) racks will come a little later, but in roughly the same form factor – 72 dual-socket nodes per cabinet, >100KW per rack.
- 28 GPU Racks, 66 CPU racks, 34 storage, switching, management racks, 128 cabinets total (smaller than Stampede – but a lot more power).
- Estimating ~13MW

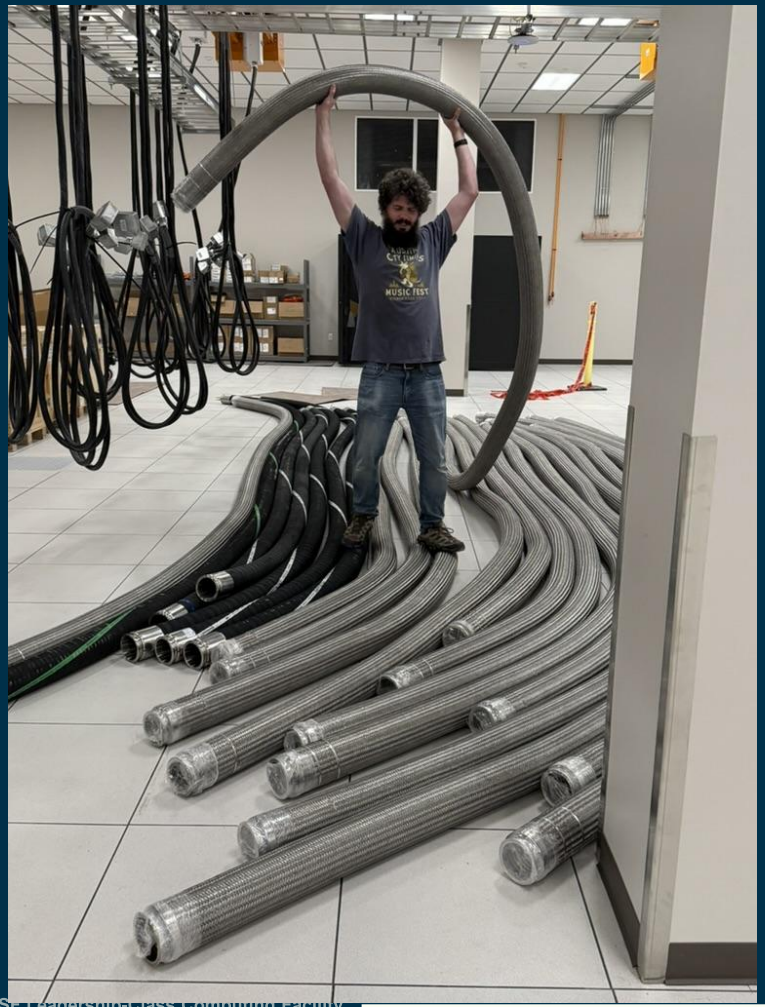
We are currently deploying

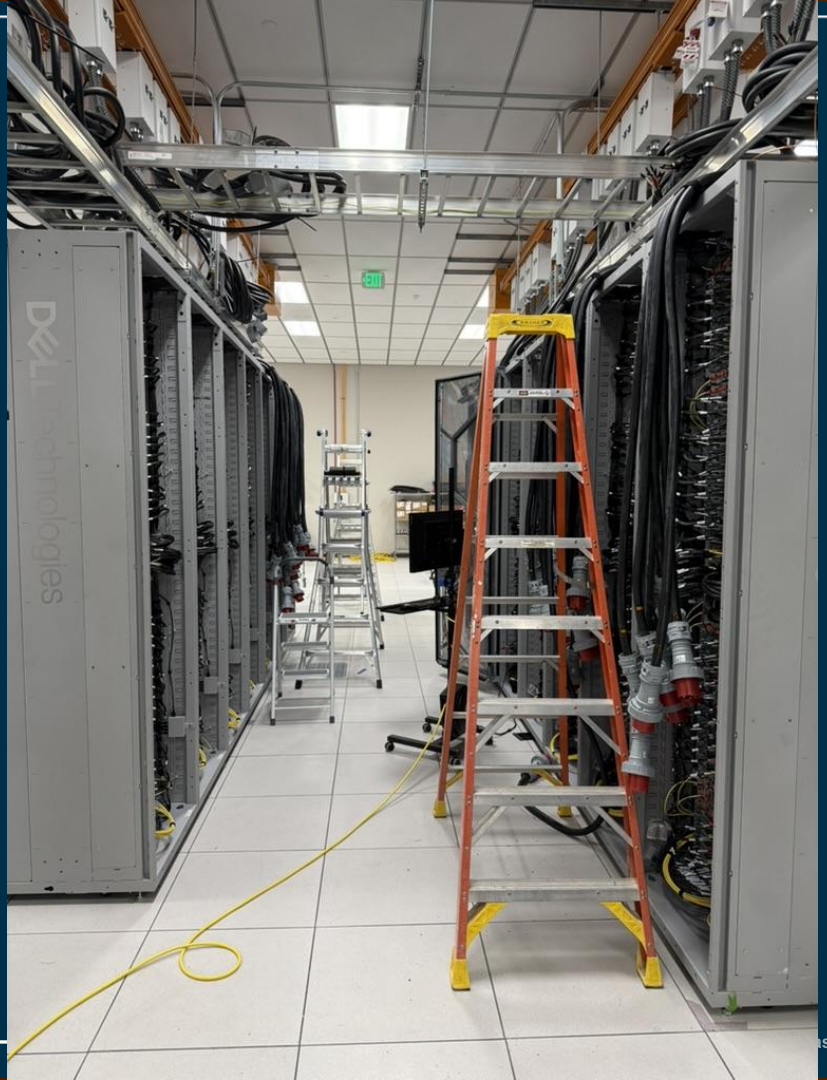
- As of January 1. . .

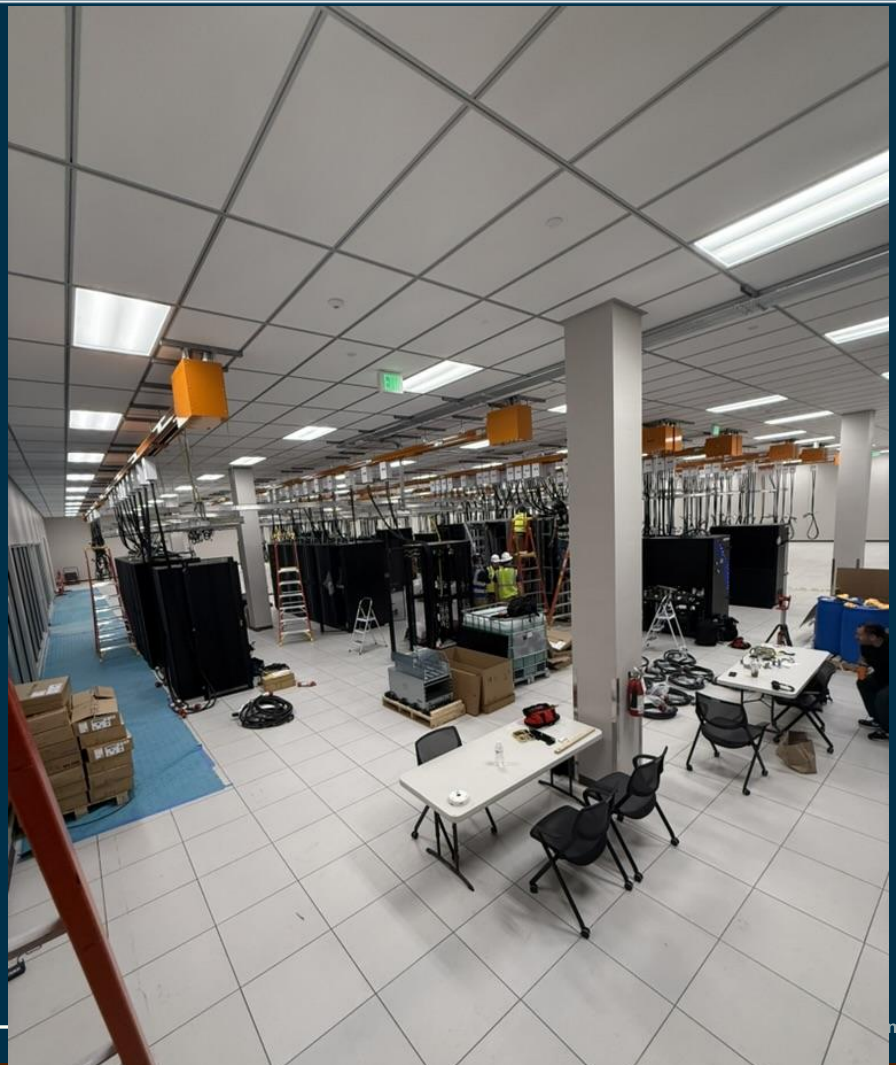

DELL Technologies

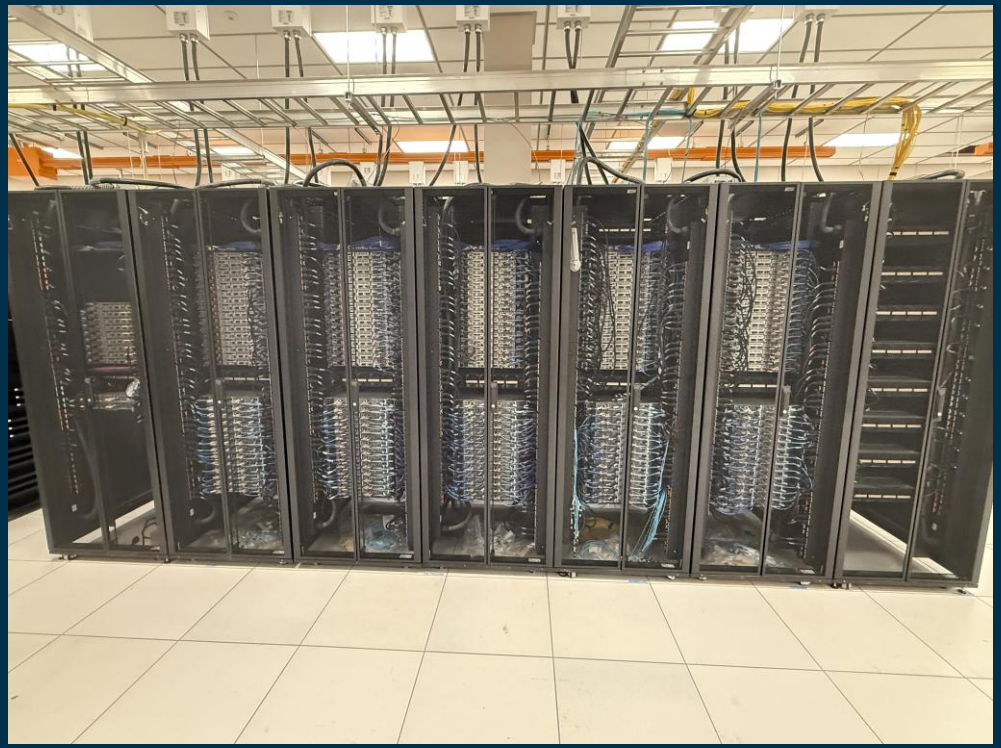
22 Racks on the Floor 4/1/26

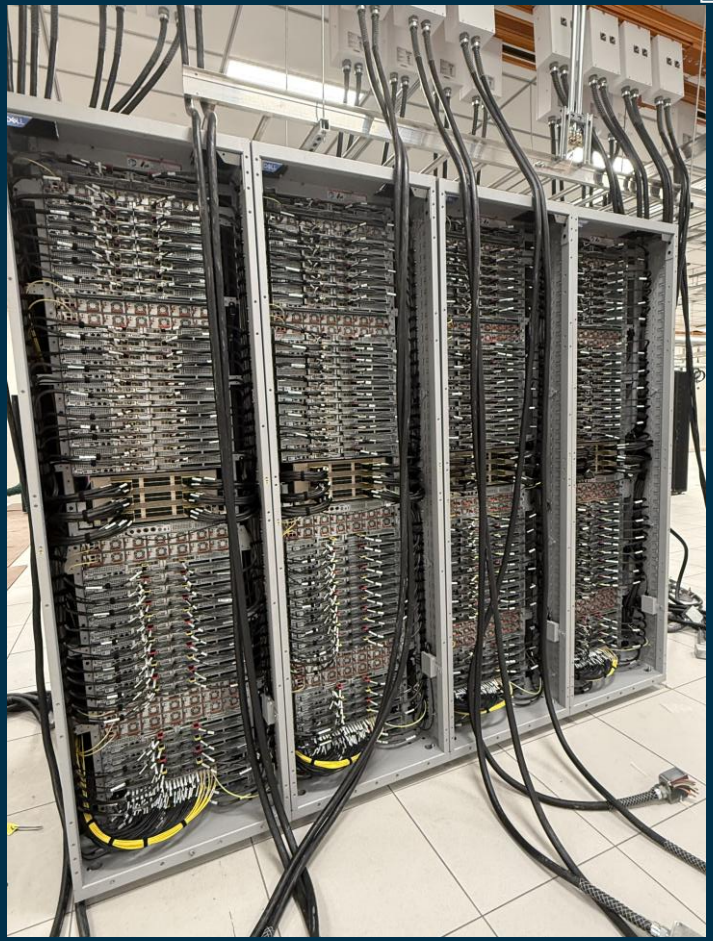




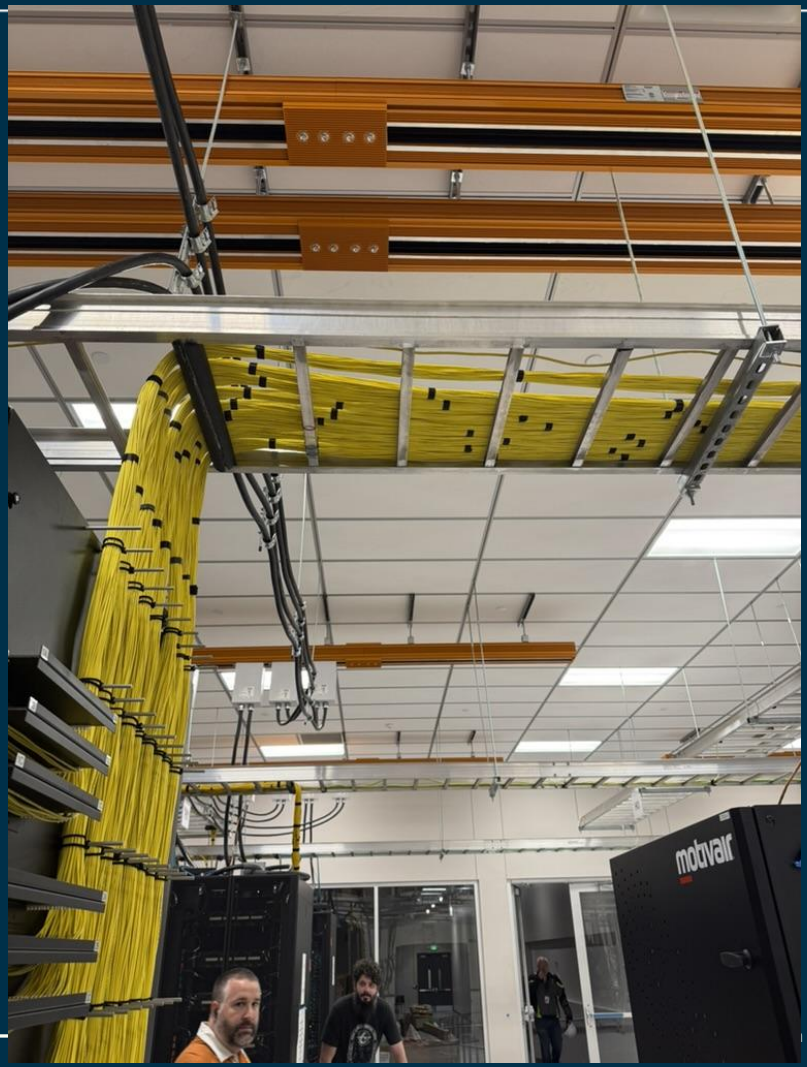









motivair
CAUTION
CUIDADO

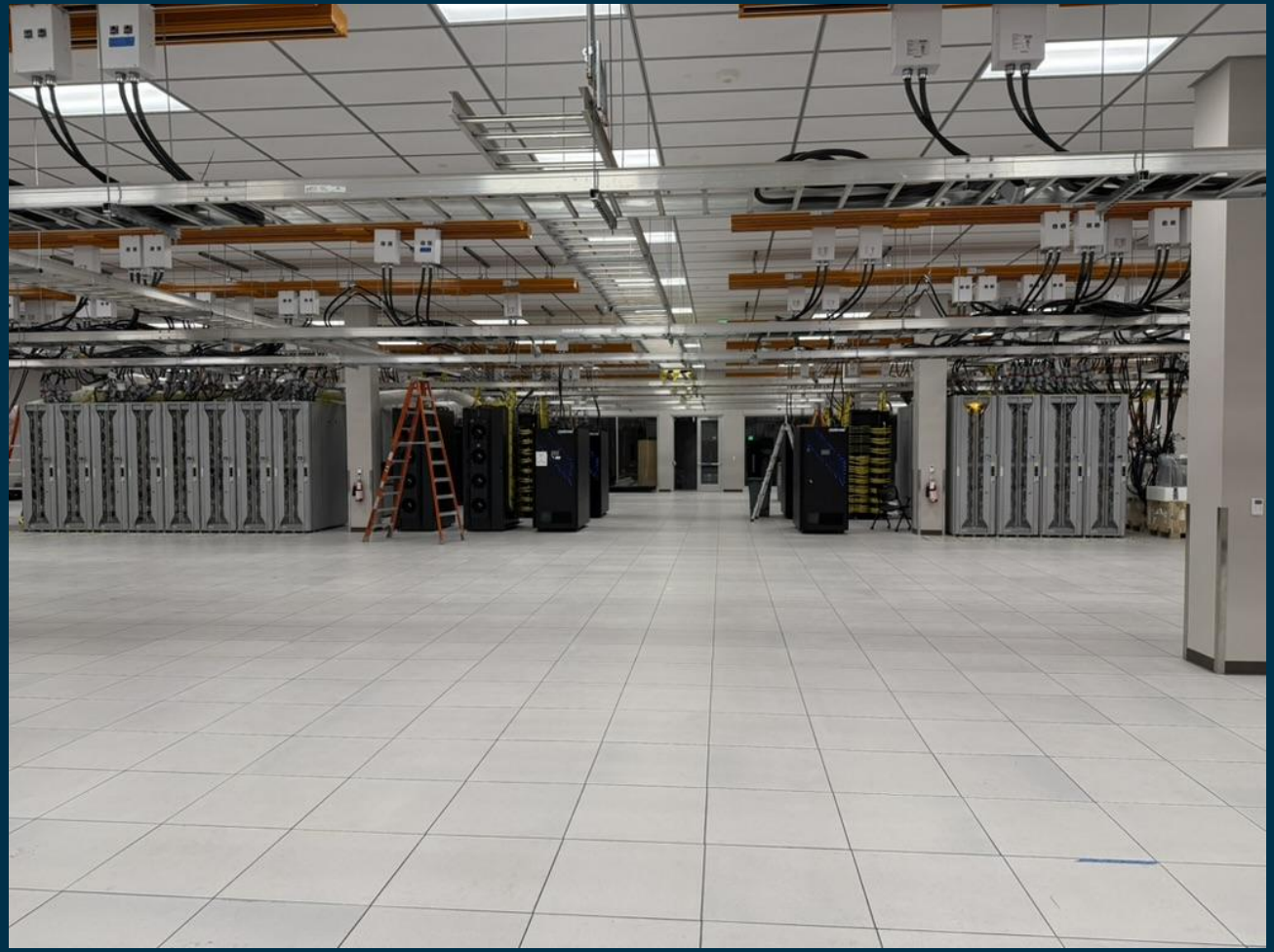

motivair

Horizon Updates

- Not all progress is linear.
- Some of the install team has needed some… basic science lessons.

Horizon Updates

- Not all progress is linear.
- Some of the install team has needed some… basic science lessons.
- Installing 28 doors (admittedly Dell's first ever of these) took 10 weeks.



Status 8/10/26

We have 4,000 GPUs on the Floor.

- 28 GPU Racks
- 21 Storage/management racks
- 3 Switch racks
- 7 CDUs
- Random rented chillers to cover parts shortages.

3,856 currently have power, in-rack cooling, and rear door cooling active.

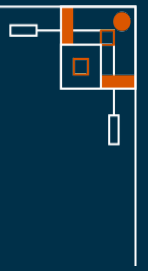
3,072 have been tested with HPL and Stream and are healthy.

- >120PF achieved (native)
- >300PF (emulated FP64)
- >60EF (FP4)

All 400PB of storage installed; 100PB of storage up and tested (2PB mounted on clients).

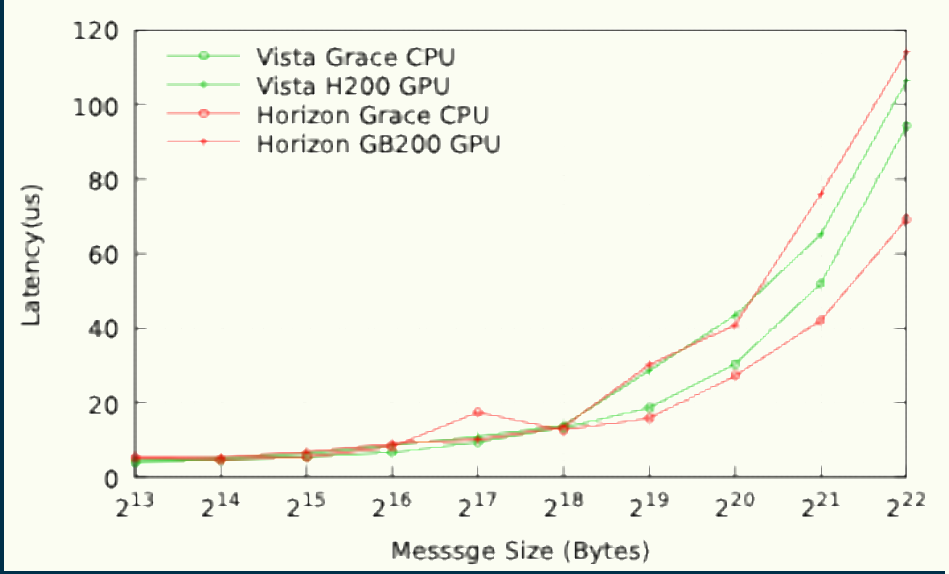
What can it do?

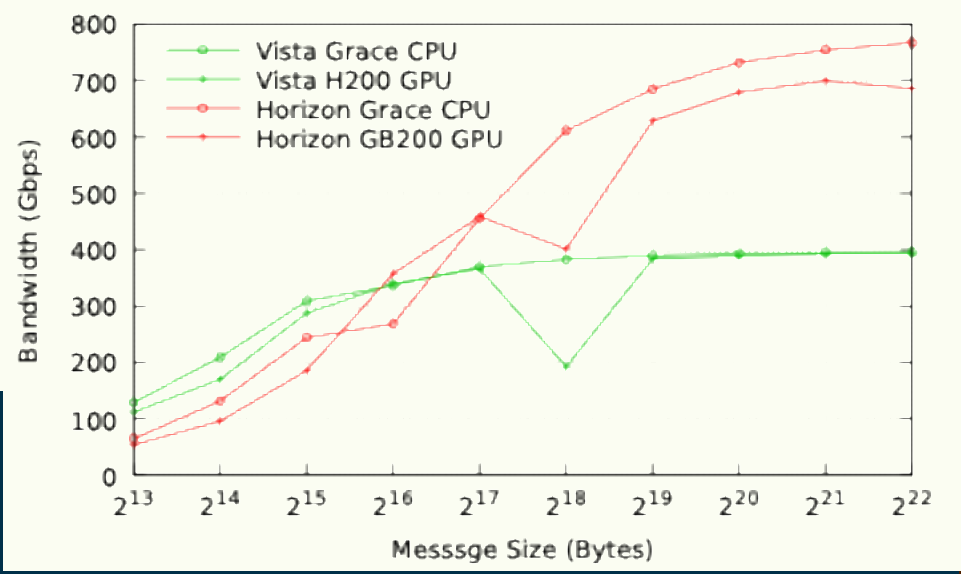
- We have a fair number of low level benchmarks, and some early app numbers.
- Just letting the earliest of early users on
  - (Normally, we'd do more, but cooling issues mean we need to nuke some of the filesystems one more time before production).
- Tuning is going to be a lot of work…
  - UCX settings matter… *a lot*. Particularly at scale. Like 10-1 perf difference matter.
  - Getting things to traverse NCCL vs. MPI and on the right network, with the right driver, pinned to the right socket, memory, and network card, seems to be the number one challenge (we will set a lot of this in your SLURM environment when we get to a sane state).

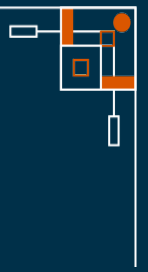


Benchmarks

You start with some low level stuff to see if anything works… and if it's faster than last time.





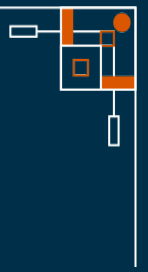


HPL will tell you if your hardware is healthy and consistent
But there are lots more versions than there used to be…

Native HPL:

June 26th, first run to crack 100PF (3,072 GPUs) :

```
================================================================================
T/V                N    NB     P     Q               Time                 Gflops (   per GPU)
--------------------------------------------------------------------------------
WC0          7421952  1024    64    48            2686.58             1.015e+08 ( 3.302e+04)
--------------------------------------------------------------------------------
||Ax-b||_oo/(eps*(||A||_oo*||x||_oo+||b||_oo)*N)=   0.000088729743 ...... PASSED
```



HPL will tell you if your hardware is healthy and consistent
But there are lots more versions than there used to be…

Native HPL:

3700 GPUs, replace the NCCL Library, up to 124.3PF :

```
================================================================================
T/V                N    NB     P     Q               Time                 Gflops (   per GPU)
--------------------------------------------------------------------------------
WC0           8355840  1024    74    50            3130.05             1.243e+08 ( 3.358e+04)
```

Benchmarks like this are typically accompanied by statements like:

“stupid openmpi doesn't propagate all the right variables for the UCX settings, so had to specify them to be exported on the mpirun command line,”

There is HPL, then there is HPL.

- For Native HPL, we will hit about 130PF on the GPU partition (>84% efficiency).
  - That's about 34TF per GPU, **or 4.9PF per rack!!!**
- For *emulated* HPL via Ozaki2, we get about 2.5x the per rack number.
  - 85TF per GPU, or ~12.5PF per rack. Just as good a residual. Here's 22PF on 256 GPUs:
  - `WC0  2097152  2048  16  16  271.14  2.268e+07 ( 8.859e+04)`
  - At 3k, we have runs well over 200PF, should get to~300PF.
- For *Mixed Precision* HPL (HPL-MxP), we get more like 15-20x the native number… more than **80PF per rack** – trying to get a 2EF number right now.
- Of course, the AI Flops number is much higher… so how fast do you tell someone your system is? (Maybe by what it does?)

And all of that helps, but does not insure, applications run well.

- Fortunately, they mostly do… but I/O is a problem for us at the moment, and lots of things start out badly until we figure out the right set of UCX hints/parameters. And the MVAPICH team fixes it.
- But early app numbers look good, if requiring some translation.
  - For Instance:

SeisSol		
baseline		
Machine	**Frontera**	**Horizon**
# of nodes	4000	100
Node type	Cascade Lake CPU, 56 cores	2 Grace CPUs + 4 Blackwell GPUs
Simulation time (secs)	394.19	394.26

Translation:
For a (formerly CPU) seismic code, in the same fixed time, we can run a problem that used 50% of Frontera on 10% of the Frontera GPUs. *(which are ~half the machine, by investment).*

More App Cases

AWP		
Baseline		
	16 times bigger model on Horizon than Frontera	
Machine	**Frontera**	**Horizon**
# of nodes	3200	350
Node type	Cascade Lake CPU, 56 cores	2 Grace CPUs + 4 Blackwell GPUs
nx	3000	6000
ny	1372	2744
nz	800	1600
# of steps	20000	40000
Simulation time (secs)	2407	2920

Translation:

(Disclaimer: 10%+ of the runtime measured on Horizon will disappear when the full filesystem is there).

AWP is a 3D Finite Difference Code to simulate multi-rupture earthquakes.

In roughly the same time, 35% of Horizon GPUs can run a 16x bigger problem than 40% of Frontera.

Next Steps on LCCF Deployment

- Probably in early 2027, we will add the CPU component.
  - Less nodes than we planned, but with more memory per node– thanks market.
  - Guessing still in the 4k nodes/8k sockets/700k cores.
- Once we have Horizon online, we will steal some of Vista to build the interactive/Kubernetes partition for Horizon.
- Some more of Vista will join some Sambanova racks and fast KV and RAG storage in a new NSF inference system, ***Sagebrush***.
- We still need to figure out node sharing settings on Horizon – but the minimum allocation unit as a $200k node-hour doesn't seem to make much sense.
- Oh yeah, and start operations!
- As an aside, you will probably see some (or a lot of) old Frontera compute nodes turn up on **Stampede3** (and the oldest Skylake partition to go away). We will run it until it dies… CPU time is still valuable, and hard to replace.

Is this a bubble? (Dan's opinion)

- Some rough math:
  - $2T in capital investment *so far*. What's the interest?
    - Good Corporate bonds: 5%/year
    - VC money: 10/12% a year.
    - Let's use 7%.
    - **$140B** a year needed to pay the *interest* on the money.
  - 100GW of power.
    - 1MW of average power load for a year costs *me* $700k.
    - So 100GW is **$70B** a year. 500GW is **$350B** a year.
  - Depreciation
    - 2/3 of the spend is IT gear – the GPUs, servers, and storage. Lasts maybe 4 years as commercially viable?
    - Depreciation -- **$350B/year.**
  - **OK, so break even on the infrastructure needs $560B/year, growing to >$1T/year in 2030.**
  - **Just infrastructure – no salaries, no taxes, no profit, no marketing, no anything else.**
  - **OpenAI 2025 revenue - $10B.**
- Yes, this is a bubble.

A Note on Data Center Impacts (Dan's Opinion)

There are real concerns and a lot of misinformation mixed together.

- Power usage is a *very* real concern, that merits attention.
  - On site generation and storage could be a boon to the grid – with proper regulation and incentives.
- Water usage does not *have* to be a concern.
  - Responsibly built datacenters don't need to use that much -- Sabey is a good example.
  - A little regulation on what "responsible" means would go a long way.
- Most Noise/Health concerns don't stand up to scrutiny.
  - Stand outside the back wall by the datacenter, here or Round Rock – you can't tell if it's running or not.
  - There are outliers where water is managed poorly and chemical additives *may* be a concern, but most keep treated water in a closed loop.
- Land use… won't end up being that much.
- Promises of jobs need careful review – most need very few people on site after construction is finished.

More Opinions – What *We* Need to Do.

- OK, so the funding context is not really pretty.
- If you do get money, you probably can't buy much, and it won't ever show up.
  - Until the bubble bursts.
- That's the reality – but there are things we can and should do in the meantime:

More Opinions – What *We* Need to Do.

- Squeeze every ounce of performance out of what we have – that means software and algoritms (mixed precision, emulation, etc.).
- Squeeze every ounce of life out of hardware – figure out what/where you can run without support, or creatively re-purpose.
  - Frankly, new nodes are so much more power, the environmental argument for turning off old nodes is mostly gone!
- We have few workable vendor options for most components – support good startups /alternatives in making more.
  - Not just buying there stuff, but showing them the real barriers, and helping – again, with good software – to get things to work on multiple platforms!
- (Continue to) Embrace the Clouds – but make sure our users maximize the use of every paid hour on them!
- Leverage all of our shared resources!

NSF LEADERSHIP-CLASS COMPUTING FACILITY

Thanks!

dan@tacc.utexas.edu

TACC TEXAS ADVANCED COMPUTING CENTER

TEXAS The University of Texas at Austin