

DATACENTER SUSTAINABILITY PANEL : PERSPECTIVE FROM TACC

Dan Stanzione

Executive Director, TACC

Associate Vice President for Research, UT-Austin

September 2023



TACC - 2023



LEADERSHIP-CLASS
COMPUTING FACILITY

TACC

TEXAS ADVANCED COMPUTING CENTER

TACC RESOURCES

- ▶ We operate the Frontera, Stampede-2, Jetstream, and Chameleon systems for the National Science Foundation
- ▶ Longhorn and Lonestar-6 for our Texas academic and industry users.
- ▶ Altogether, ~20k servers, >1M CPU cores, 1k GPUs
- ▶ Typical power ~6MW
 - ▶ Max 9.5MW
- ▶ Adding 20MW of datacenter capacity for LCCF (30MW total) 2025.



COOLING STRATEGIES

- ▶ Stampede 1 and 2 – In-row Chillers enclosed hot aisles (2012 build out).
- ▶ Frontera (2019) Stampede 3 (2023) - Direct Liquid Cooling of processors.
- ▶ Frontera RTX (2019), Lonestar-6 (2021) – Immersion cooling.
- ▶ Next datacenter – we will definitely have (probably warmer) water to each rack location, the rest is somewhat TBD
- ▶ We also employ chilled water storage to offload the power grid at peak demand.
- ▶ We employ roughly 200kw of direct solar, and by wind credits for about 20% of the remainder.
 - ▶ New datacenter will be 100% wind offsets.



SUSTAINABILITY AND DATACENTERS

- ▶ Obviously, sustainability is a priority.
- ▶ But the mission - providing the best computational resources – is the highest priority.
 - ▶ We are both the cause of and solution to many of these problems 😊.
- ▶ Datacenters are still a tiny fraction of usage compared to, say, transportation.
 - ▶ And our datacenters help design batteries, carbon capture and storage, better photovoltaic materials, remediation for plastics and chemicals, etc, etc.
 - ▶ A better use of power than the much larger datacenters for X/Twitter, Cat Videos, and generating targeted ads.
- ▶ If we had a green power grid, not only would our datacenters not be a problem, a lot of other stuff wouldn't be either – but we can't change that unilaterally.

A FEW BITS OF OUR SUSTAINABILITY PLANS:

- ▶ We continue to run experiments to improve the efficiency of our datacenter operations:
 - ▶ We have an experimental Hydrogen fuel cell being put in our current datacenter power loop.
 - ▶ We are working with several startups on novel cooling technologies.
 - ▶ We continue to work with our vendors to be able to raise inlet temperatures for water – while maintaining a high enough delta-T to keep chillers running efficiently.
 - ▶ We are in Texas, we are probably going to still need chillers, even if water temps reach 35C.
 - ▶ Going to 100% wind credits for a 7% markup – willing to pay that.
- ▶ Similarly, we are working to improve how power is managed:
 - ▶ Capping power at modules (e.g. Grace-Hopper cards, and future versions with potentially more components) rather than at the server level will reduce the datacenter build out for “max power”.
 - ▶ We will be below 9MW in our current projected design for Horizon, the “10x” replacement for the Frontera system in 2025.
- ▶ But. . .

BUT...

- ▶ Datacenters (at least good ones) are already pretty efficient.
 - ▶ Any good student of Amdahl's law knows there are diminishing returns on working on the already-tuned parts of any problem.
 - ▶ At most, we can squeeze out another 10-15% from datacenter efficiency measures.
- ▶ Real change means tackling **software**. . .

SOFTWARE AND SUSTAINABILITY

- ▶ We know, for instance, that per “peak” FLOP, we get a 5-6x multiple moving to GPUs.
 - ▶ But outside of AI, a large fraction of codes don't run on GPUs.
 - ▶ (And arguments can be made on yield of peak flops across architectures).
 - ▶ 5x is more than 15%.
- ▶ We also know, but don't really talk about, that most actual app runs get a single digit percentage of peak performance.
 - ▶ Which means code efficiency offers the potential for an order of magnitude improvement.
 - ▶ Yes, more efficient code uses somewhat more instantaneous power – but shorter runtimes help a lot.
- ▶ The problems is software is hard, diverse, and often beyond our reach. . .
 - ▶ But a crappy job on software, with 1,000% potential, is probably better than a great job on datacenter, with 10% potential.

SOFTWARE AND SUSTAINABILITY

- ▶ We are sampling performance data every few minutes on every job to keep a profile of efficiency
 - ▶ This is one of the ways we target consultants.
- ▶ Pushing the user base (somewhat) towards increasing GPU usage.
 - ▶ Just added GPU monitoring; anecdotally, there is massive inefficiency there.
- ▶ A problem we have is *incentives* -- users just want the fastest answer – no incentive to get a slower answer that uses less power (we saw this a lot on Stampede 2).
- ▶ Perhaps we change our charging units from wall clock hours to total Joules consumed??
- ▶ We hope to start reporting energy usage to users next year – not sure when/if we will go to energy-based charging.
 - ▶ Incentivize more efficient codes.
 - ▶ Maybe incentivize moving loads to optimal power cost times? (West Texas wind power can be somewhere between free and negative a fair number of hours per year).

FRONTERA

TACC



TEXAS

