TEXAS ADVANCED COMPUTING CENTER

A UNIVERSITY HPC CENTER PERSPECTIVE ON HPC AND CLOUD PROVIDERS

April 25th, 2018

Dan Stanzione, Executive Director, TACC

Associate VP for Research, UT-Austin

TEXAS ADVANCED COMPUTING CENTER

HPC VS. THE CLOUD: IN THE END THERE CAN BE ONLY ONE(?)

April 25th, 2018

Dan Stanzione, Executive Director, TACC

Associate VP for Research, UT-Austin

OUTLINE

- HPC vs. the Cloud – is this a debate?
- HPC and the Cloud
- HPC is the Cloud?

TACC AT A GLANCE

Personnel

160 Staff (~70 PhD)

Facilities

12 MW Data center capacity
Two office buildings, Three Datacenters, two visualization facilities, and a chilling plant.

Systems and Services

~35,000 users in ~3,000 projects on fifteen production platforms
200+ Data collections in 60+ PB

HPC - Stampede-2, Lonestar 5, Hikari
Data - Wrangler
VIS/ML - Maverick
Cloud/Interactive – Chameleon, Jetstream, Roundup
Storage Stockyard, Corral, Ranch
Experimental – Fabric, Catapult, etc.

Hikari

- 380V D
system
NEDO
cores.
al
o
/s

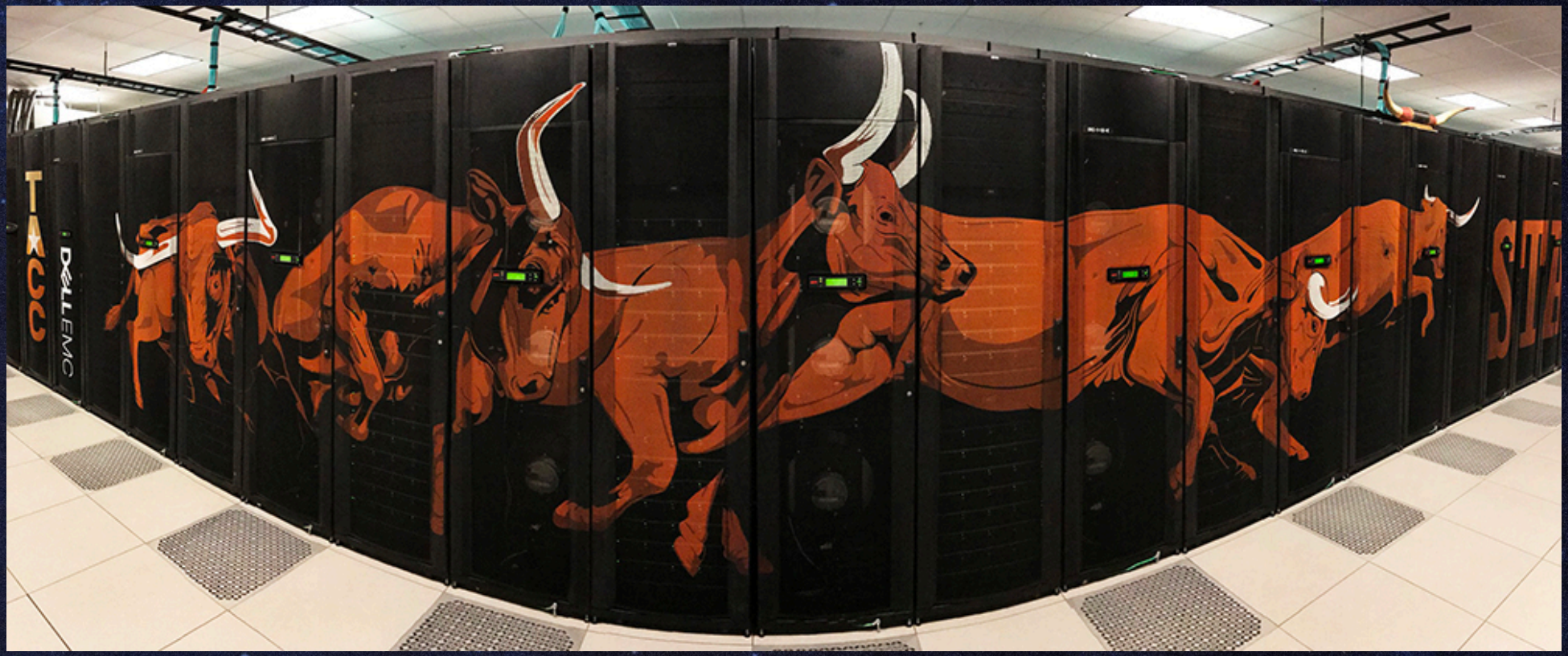
TACC TEXAS The University of Texas at Austin

STAMPEDE 2

- Funded by NSF as a renewal of the original Stampede project.
- The largest XSEDE resource (and largest university-based system).
- Follow the legacy of success of the first machine as a supercomputer for a *broad* range of workloads, large and small.
- Install without ever having a break in service – in the same footprint.

A FIRST CRITICAL DISTINCTION

SUPERCOMPUTERS ARE BEAUTIFUL,



SUPERCOMPUTERS ARE BEAUTIFUL,

AND CLOUD SYSTEMS ARE REALLY UGLY

(GET YOUR ACT TOGETHER, FOLKS)

HPC VS. THE CLOUD

- People often equate HPC and Cloud Computing
  - HPC has a big datacenter full of servers.
  - The Cloud has a big datacenter(s) full of servers
    - In this sense, both Accounting and Particle Physics involve Math, so must be the same.
  - We both like the word "scalable" though we use it in very different ways.

HPC VS. THE CLOUD

- This basic commonality actually does give us a LOT in common.
  - Cloud and HPC providers care about:
    - Power, cooling, and datacenter efficiencies.
    - Costs of hardware, and reliability of components
    - Standardization of hardware platforms, provisioning, etc.

BUT ALSO A LOT THAT'S DIFFERENT

- (This is a given for this crowd, but just in case. . .
- Usually, when an HPC person says "Scalable", they mean:
  - A Large number of synchronous, cooperating, tasks, working together to perform a single large computation.
  - Think global weather, colliding blackholes, jet airframe simulation
  - So what matters? Network *latency*; reproducible performance; vectorization, memory access times, I/O rate to a single file, etc.
- Usually, when a Cloud person says "Scalable" they mean:
  - More asynchronous transactions against what is likely a shared database.
  - Think scaling web site hits, search engines, airline reservation systems.
  - What matters? *Average* performance, cost per transaction, aggregate I/O, uptime, etc.

IS COST THE DIFFERENTIATOR?

- Because of these simlarities, the comparison of "HPC vs. the Cloud" or "HPC in the Cloud" is "Which costs less per core hour?"
- This is sort of a silly comparison. An inflatable raft is cheaper than a 200' yacht, but I know which one I would rather cross the Pacific in.
- But, since I brought it up. . .
  - HPC is probably cheaper.
  - Because we need to tightly couple applications, we might have a large00x performance advantage *per hour* anyway.
  - On a certain large genomics pipelines, we are an order of magnitude less than commercial clouds.
- This isn't a good measure because… *(Hang on, 3 slide tangent)*

A WHOLE LOT OF GOOD STUFF HAS LEAKED INTO HPC FROM THE CLOUD, AND WE SHOULDN'T UNDERSTATE IT.

- Containers. Not so much VMs, but singularity.
- "As a service".
  - You bastards. I think we have cloud-marketing-as-a-service (CMaaS) assailing us constantly.
    - (A quote from a nameless cloud marketing slide:
      - ***"Cloud Computing . . . Requires no special knowledge of computing"***
    - Spoken like someone who sells compute time **by the hour*)*
  - But REST APIs, and consuming HPC through services is on the rise.

A WHOLE LOT OF GOOD STUFF HAS LEAKED INTO HPC FROM THE CLOUD, AND WE SHOULDN'T UNDERSTATE IT.

- Socializing pay as you go
  - Remember "You can't charge for cycles" as accepted truth?
  - "I don't need my own hardware". is gaining acceptance. Even in the government 😃. Except maybe the Post Office. But I hear they interface with Amazon in other ways. . .
- Object stores — more on this later.

AND SOME STUFF FROM HPC HAS CLEARLY INFLUENCED THE MODERN CLOUD

- Things that may have been said at cloud companies since the initial offerings came out 10+ years ago:
  - "Crap, maybe we need a scheduler"
  - "Crap, maybe we need a filesystem"
  - "Crap, maybe we need an interconnect of reasonable latency"

THINK ABOUT THE SOFTWARE

- Many cloud applications generate revenue – they are worth a lot, and are developed aggressively.
- Many scientific applications generate a *few* science results – they are developed *grudgingly*, by graduate students, over decades.
  - ***I don't really mean the "big" DOE applications here, I mean more the single PI academic code – of which we run ~4,000 different codes.***
  - So, if you say "Let's replace POSIX with an Object store" then an airline reservation system will spend a few million dollars recoding.
  - But lots of science would stop, or at least slow way down.

THINK ABOUT THE SOFTWARE

- The Cloud solution to scaling POSIX-compliant parallel filesystems is “don’t”. That’s a fair answer in that space. Not a good HPC answer.
- When colliding black holes, verification and validation is kind of hard. We can now get one experimental verification, for $2B.
  - For 15 years, LIGO codes had to produce *BIT FOR BIT* identical answers to change hardware platforms.

A LITTLE MORE ON THAT OBJECT STORE THING

- If we can convince app writers to use object stores — this truly solves our biggest weakness (in terms of reliability).
  - Will they ever really give up POSIX?
  - Can we make simulated POSIX hurt enough?
  - If using the non- or semi- POSIX interface is 5x faster, will users finally switch?
    - My guess: No, but we can hope.

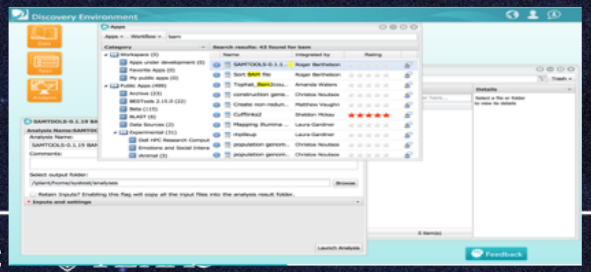
ULTIMATELY, THIS IS A SILLY COMPARISON

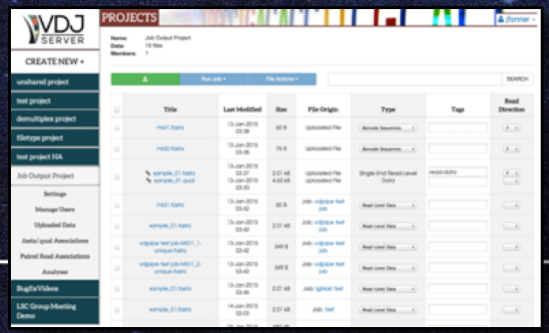
- OK, so HPC is good at delivering synchronous, tighly coupled simulations.
- The cloud may not be in all cases (certainly works in some cases).
  - So what? Why do we argue over the cost of using square pegs or round pegs?
- All of scientific computation is *MUCH BIGGER* than just the simulation.
- When I do a project, I also use Evernote, Trello, Slack, Box, etc. I don't try and run colllaboration tools on a supercomputer – and I've never done a cost analysis of why.

THE EVOLUTION OF A CYBERINFRASTRUCTURE

- Ten years ago, cyberinfrastructure was largely about building the hardware and networks to support large scale science.

- Today, it's about **new interfaces** to support **data analysis**, **collaboration** and **sharing**,**reproducibility** as well as easy access to simulation



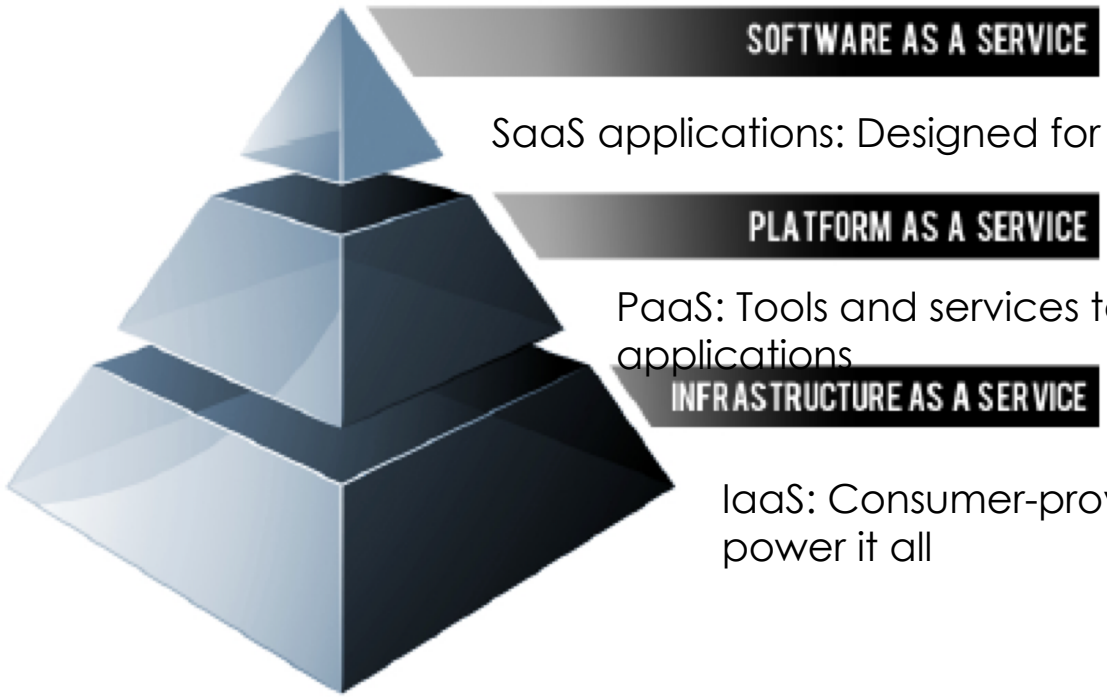


FOR THEIR RESEARCH, USERS NEED TO CONFIGURE…

- Interactive graphical research environments
- Orchestrated, heterogeneous science workflows
- Their own web applications, DBMS, AMQP services, etc.
- Specific operating system profiles
- Network architectures
- Newly emergent hardware

TACC aims to deliver the same quality of service for these use cases as for mainstream HPC

THE CLOUD COMPUTING STACK

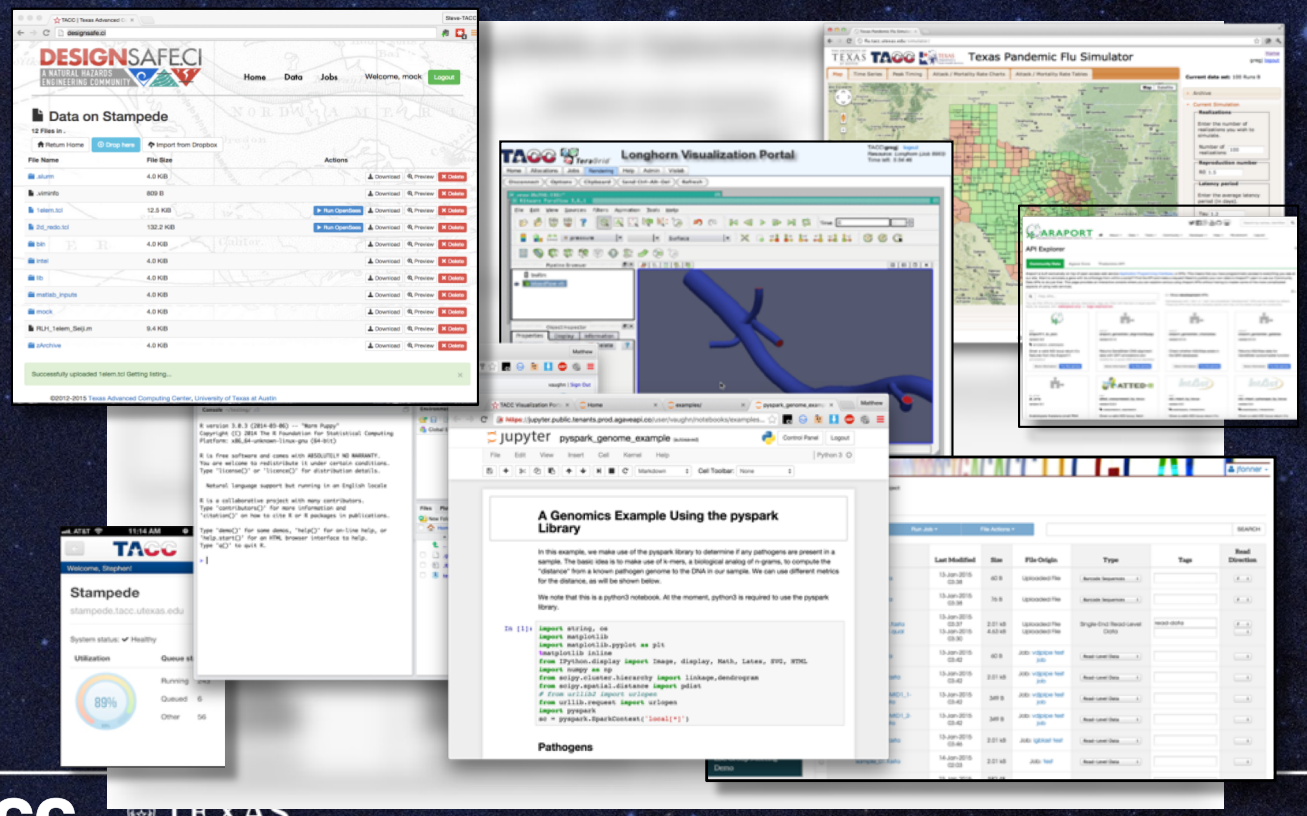


SaaS applications: Designed for end-users, delivered over the web

PaaS: Tools and services to help code and deploy web applications

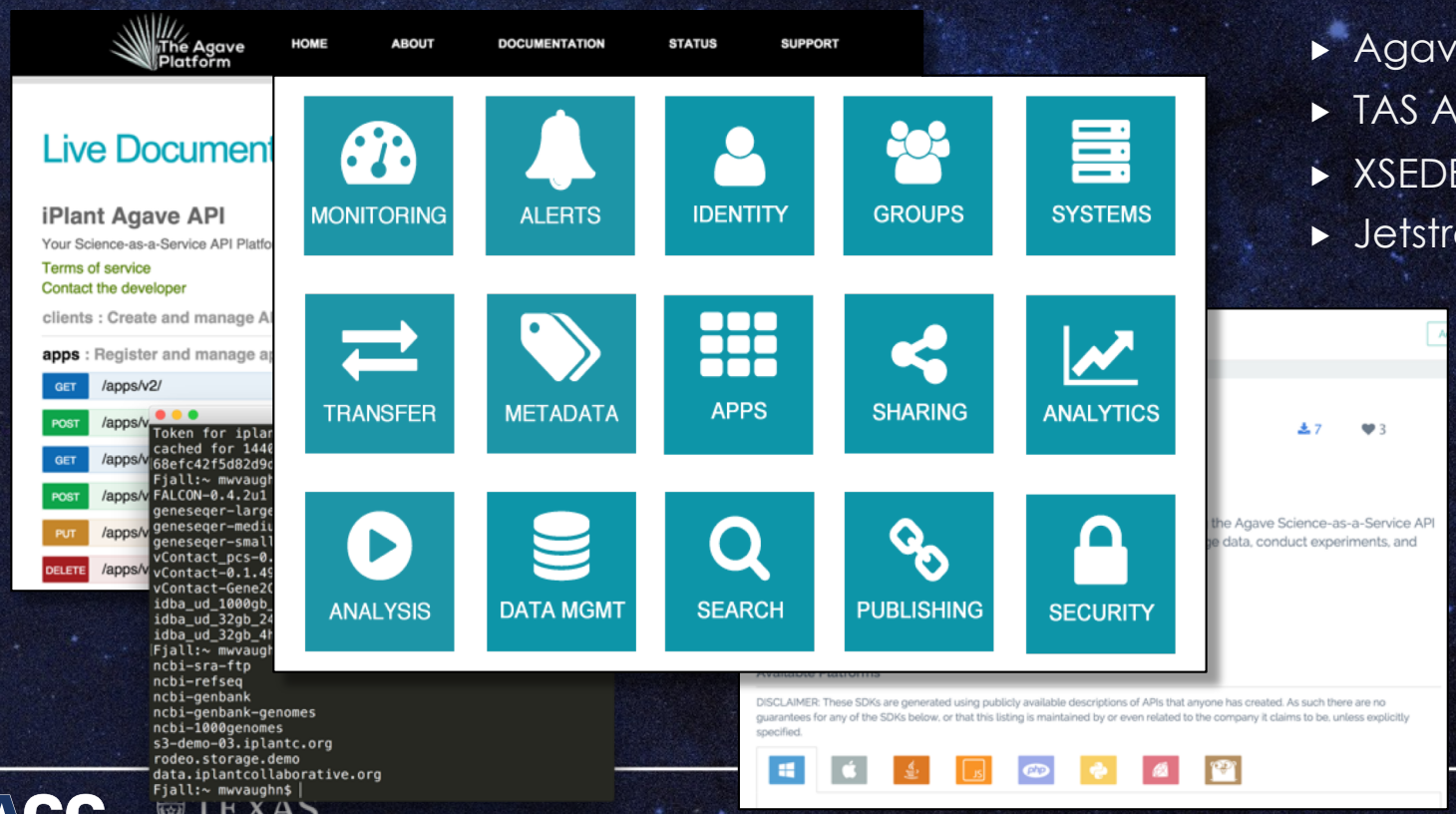
IaaS: Consumer-provisioned hardware and software to power it all

SAAS EXAMPLES



- DesignSafe Discovery Workspace
- Araport API Explorer
- VDJserver
- JupyterHub
- Rstudio
- TACC Vis Portal

PAAS: WEB SERVICE APIS EVERYWHERE



- Agave API
- TAS API
- XSEDE API
- Jetstream API

IAAS: FLEXIBLE HARDWARE PROVISIONING

Roundup
VMware Internal Cloud (TACC production)

Chameleon
OpenStack Public Cloud (CISE research)

Rodeo
OpenStack Internal Cloud (TACC development)

Jetstream
OpenStack Public Cloud (XSEDE research)

SO, IT'S NOT ABOUT HPC OR THE CLOUD

- Turns out big datacenters can be used in more than one way.
- It also turns out, to do science, we need Cloud *and* HPC.
  - We use HPC techniques when we need performance, and cloud techniques when we need APIs, failover, etc.
- In the end, to do science – we need both.

THERE ARE MORE SYNERGIES

- Recently, we have been able to scale "deep learning" (or as I like it to call it, "HPC where the answers don't have to be right"), to over 4,000 coupled nodes.
- Deep learning across nodes, it turns out, is doing a bunch of matrix operations with vector units, and synchronizing across them before advancing to the next layer.
  - This is remarkably like scaling up partial differential equation solvers.
- It's likely the performance features *HPC* needs will be the *future* performance features clouds need.

AND IF WE MISS THE SYNERGIES, THIS COULD HAPPEN. . .

AND NO ONE WANTS THAT…

TEXAS ADVANCED COMPUTING CENTER

THANKS!

DAN@TACC.UTEXAS.EDU