

ARM SUPERCOMPUTER



2019 ASC PI Meeting Lessons Learned on Astra

Presented by
James H. Laros III

Vanguard Program: Advanced Technology Prototype Systems

- **Prove viability of advanced technologies for NNSA integrated codes, at scale**
- Expand the HPC-ecosystem by developing emerging yet-to-be proven technologies
 - Is technology viable for future ATS/CTS platforms supporting ASC mission?
 - Increase technology AND integrator choices
- Buy down risk and increase technology and vendor choices for future NNSA production platforms
 - Ability to accept higher risk allows for more/faster technology advancement
 - Lowers/eliminates mission risk and significantly reduces investment
- Jointly address hardware and software technologies
- First Prototype platform targeting Arm Architecture

Tri-lab collaboration is integral to Vanguard

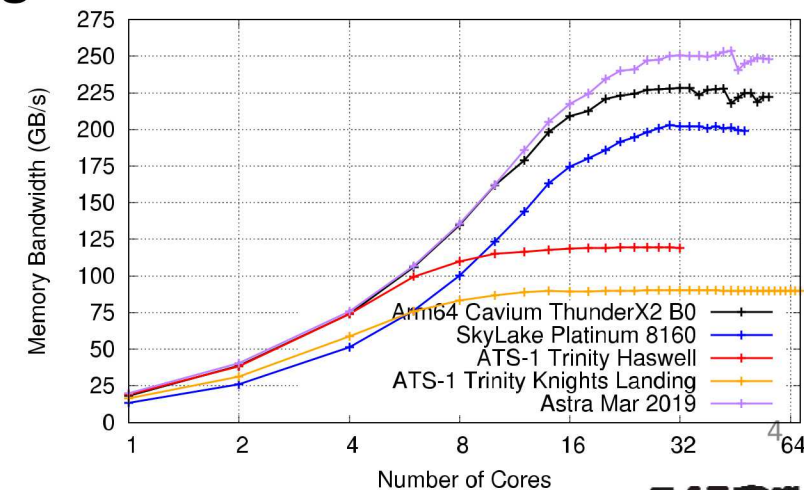
Vanguard Astra: At a Glance

- 2,592 HPE Apollo 70 compute nodes
 - 5,184 CPUs, 145,152 cores, 2.3 PFLOPs (peak)
- Cavium Thunder-X2 ARM SoC, 28 core, 2.0 GHz
- Memory per node: 128 GB
 - 16 x 8 GB DDR DIMMs
 - Aggregate capacity: 332 TB, 885 TB/s (peak)
 - 247 GB/s per node STREAM
- Mellanox IB EDR, ConnectX-5
 - 112 36-port leaf, 3 648-port spine switches
- ATSE software stack
 - TOSS Base Operating system
- HPE Apollo 4520 All-flash Lustre storage
 - Storage Capacity: 403 TB (usable)
 - **Upgrade to 3X memory in preparation for move to classified**
 - Storage Bandwidth: 250 GB/s
 - 400 GB/s stunt mode, 432 GB/s peak



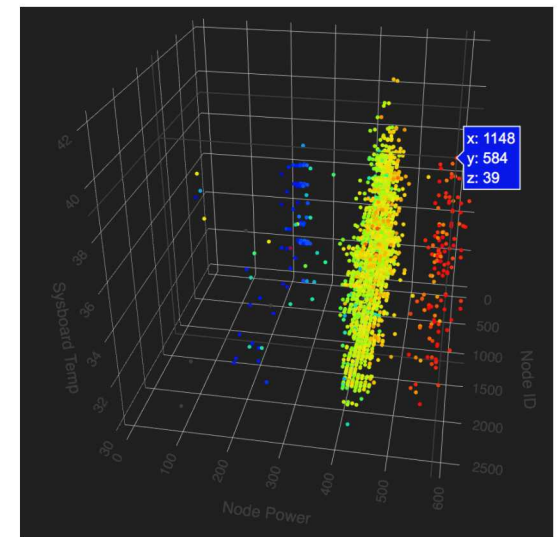
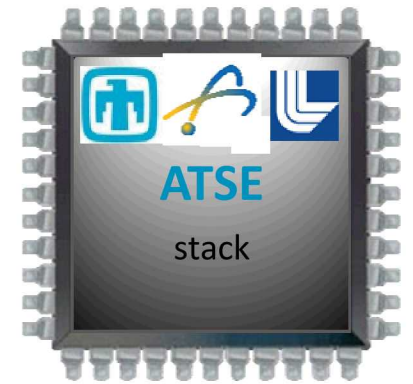
Vanguard Astra: Lessons Learned or Reasons to Prototype new Technologies

- “Right sizing” technical review team worked well - Smaller can be better
 - reinforced at Crossroads TAT
- “Right sizing” system to meet goals
 - Depends on the technology targeted
 - Depends on budget
 - At scale reasonably shows will run on both CTS and ATS
 - Test entire software environment, systems management included, at scale
 - Vendor interest (depends on vendor)
- Unless you are into hard hats and safety vests consider finishing the building before system delivery
- Problems are not observed in isolation
- Initial implementation of Thermal solution required tweaking
 - Met our target of 21-22 C water, 24-25 inlet air temp
 - Lowered water temp to 20 C - led to increased stability
- Delivered processor part originally did not meet SOW memory bandwidth requirements – Does now



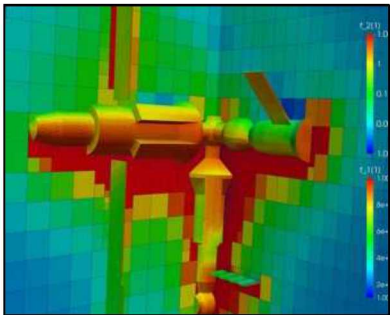
Vanguard Astra: Lessons Learned *or* Reasons to Prototype new Technologies

- Similar to CTS, NALU and other applications were forcing out of spec voltage swings
 - In this case memory bus
- Pesky fabric instability
 - Lots of hands on nodes probable cause
- Pioneering new systems management solution (HPCM) with vendor
 - Combined with new software stack (ATSE)
- At scale testing reveals previously unseen issues
 - kworker bug not seen on Comanche or x86 platforms
- Systems monitoring is CRITICAL (debug/analyze many of above)
- Early hardware requires frequent and quick iterations of software stack
 - Tension with accelerated move to classified where this is a challenge
- Keeping system in sync (updating software) a challenge -> future work with containers



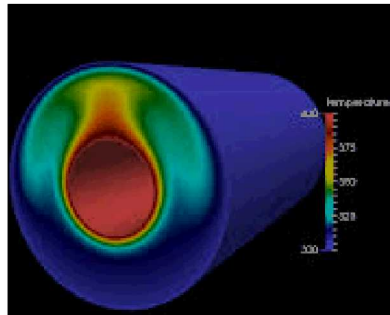
Early Results from Astra

- ThunderX2 is less reliant on vectorization to utilize available memory bandwidth.
 - Cores can consume available memory bandwidth without vectorized code.
 - Downside: vector units are small so compute-dense code may run slower, extra cores help offset this when comparing node-to-node
- Most of our complex solver libraries and applications compile with GCC or Arm compilers without significant issues.
 - Functional portability for broad code portfolio without significant code rework (NALU, SPARC, CTH, etc.)
 - Acid test is getting the performance out of generated code
- Cache performance will likely impact some of our codes that have reasonable locality
 - Suspect that caches simply perform slower on TX2 versus Xeon
 - Lack of support for gather operations
 - Most packages ported and running on the platform, ATSE environment has worked out well



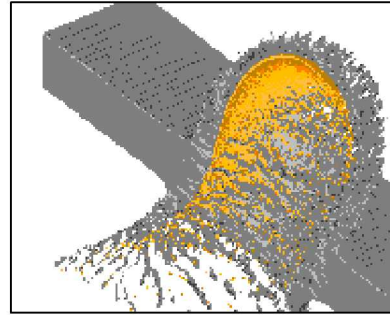
Monte Carlo

1.60X



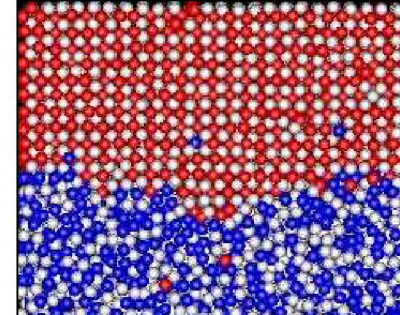
CFD Models

1.45X



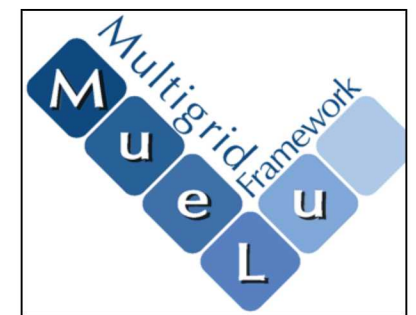
Hydrodynamics

1.30X



Molecular Dynamics

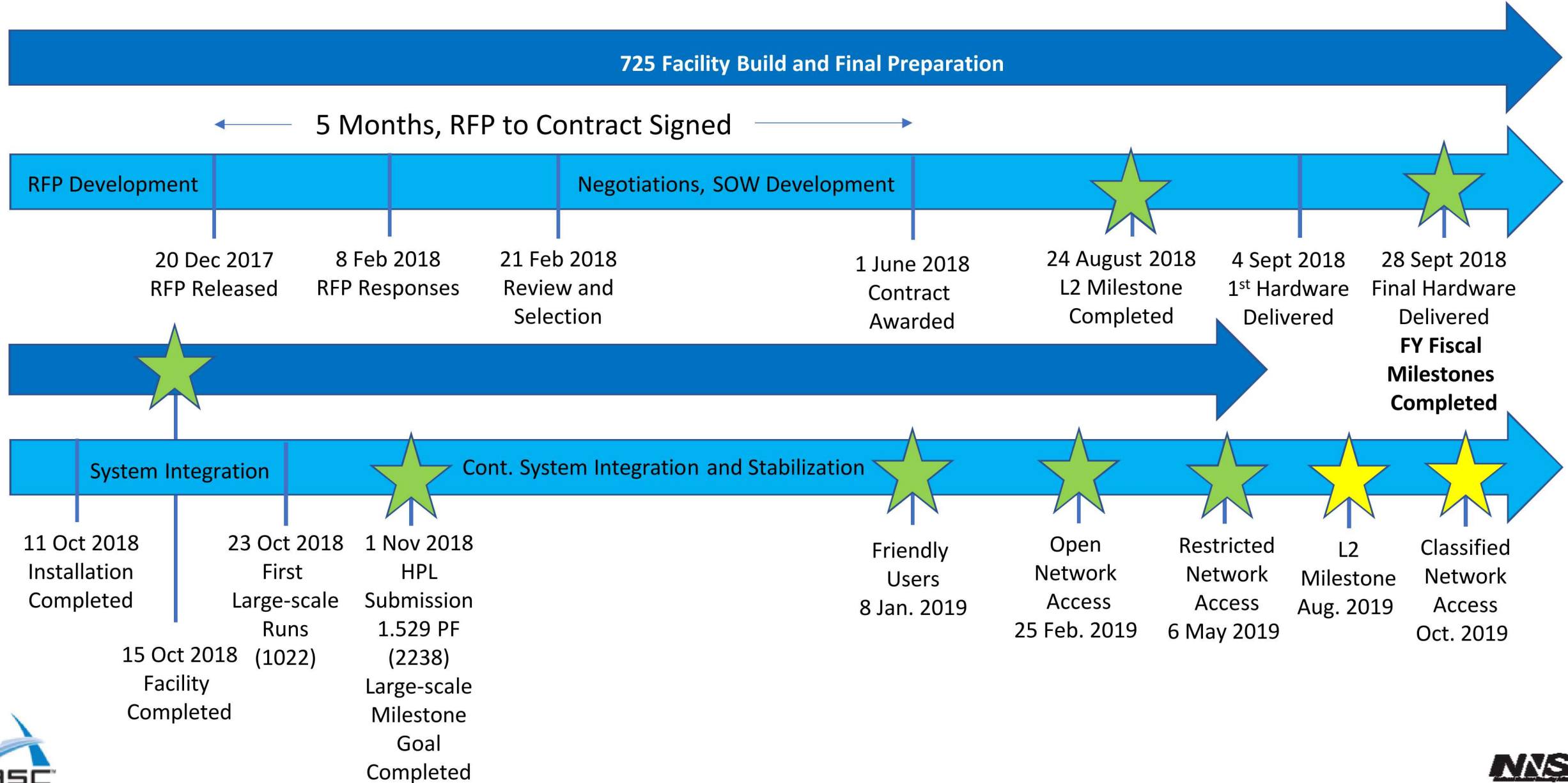
1.42X



Linear Solvers

1.87X

Vanguard-Astra: Timeline



Vanguard Astra

- Arm processor – first time used for HPC platform at scale
- All-flash lustre storage – not a first but a first with Arm clients
- HPE - New integrator, lots of experience but little specific to DOE and these types of platforms
- Vanguard program will certainly have lowered risk for all for follow on platforms
- New technology and integrator possibilities for NNSA

The Arm logo, consisting of the word "arm" in a bold, blue, lowercase sans-serif font.The Hewlett Packard Enterprise logo, featuring a green rectangular icon above the text "Hewlett Packard Enterprise" in a bold, black, sans-serif font.



Exceptional Service in the National Interest