SAND2015-6715C

Exceptional service in the national interest

Sandia National Laboratories

NNSA Advanced Architecture Test Beds

James H. Laros III, Gwen Voskuilen,
James Ang, Simon Hammond

ADVANCED ARCHITECTURE TEST BEDS

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

Advanced Architecture Test Beds

- *Be a scout for future computing architectures*
  - **FOR OUR NNSA APPLICATIONS!**
- Test beds act as a conduit for collaboration and co-design
  - With Labs (ASC and ASCR), Vendors, Universities, etc.
- Testing early hardware for both ATS and CTS programs
  - **And now ATDM!**
  - Early MIC/Phi and Haswell) -> ATS-1
  - Early Power -> CORAL
- Exploratory R&D
  - Alternative Programming Models
  - Architecture Aware Algorithms
  - Advanced Memory sub-system development, and use
  - Energy Efficient Hardware, Runtime, Systems Software
  - **Ultimately Preparing and Steering our Applications**

Advanced Architecture Test Beds

- **APPLICATION FOCUSED**
- Reduce impact on mission [illegible] a rapidly changing technol[illegible] environment.
  - Significant **PRODUCTION** rewrite/modification [illegible] required
  - Ensure that wh[illegible] the change" it is th[illegible] for code longevity, p[illegible]forts, performance *etc.*
  - Go through the pain up front so the transition for full codes is easier

    Eliminate or reduce missteps

Mini-Apps

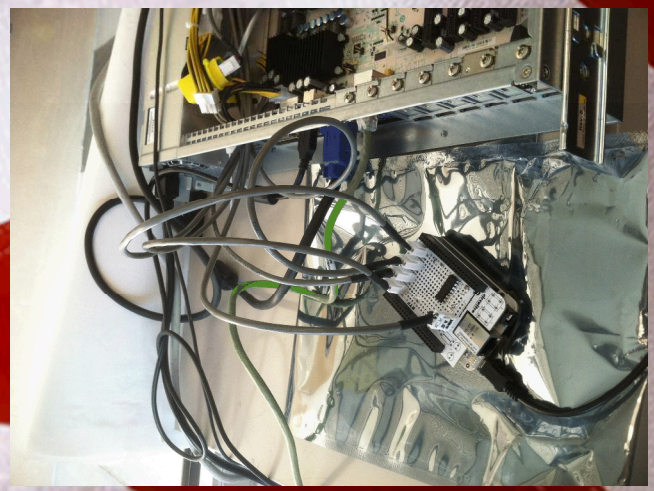
Production Codes

Advanced Architecture Test Beds

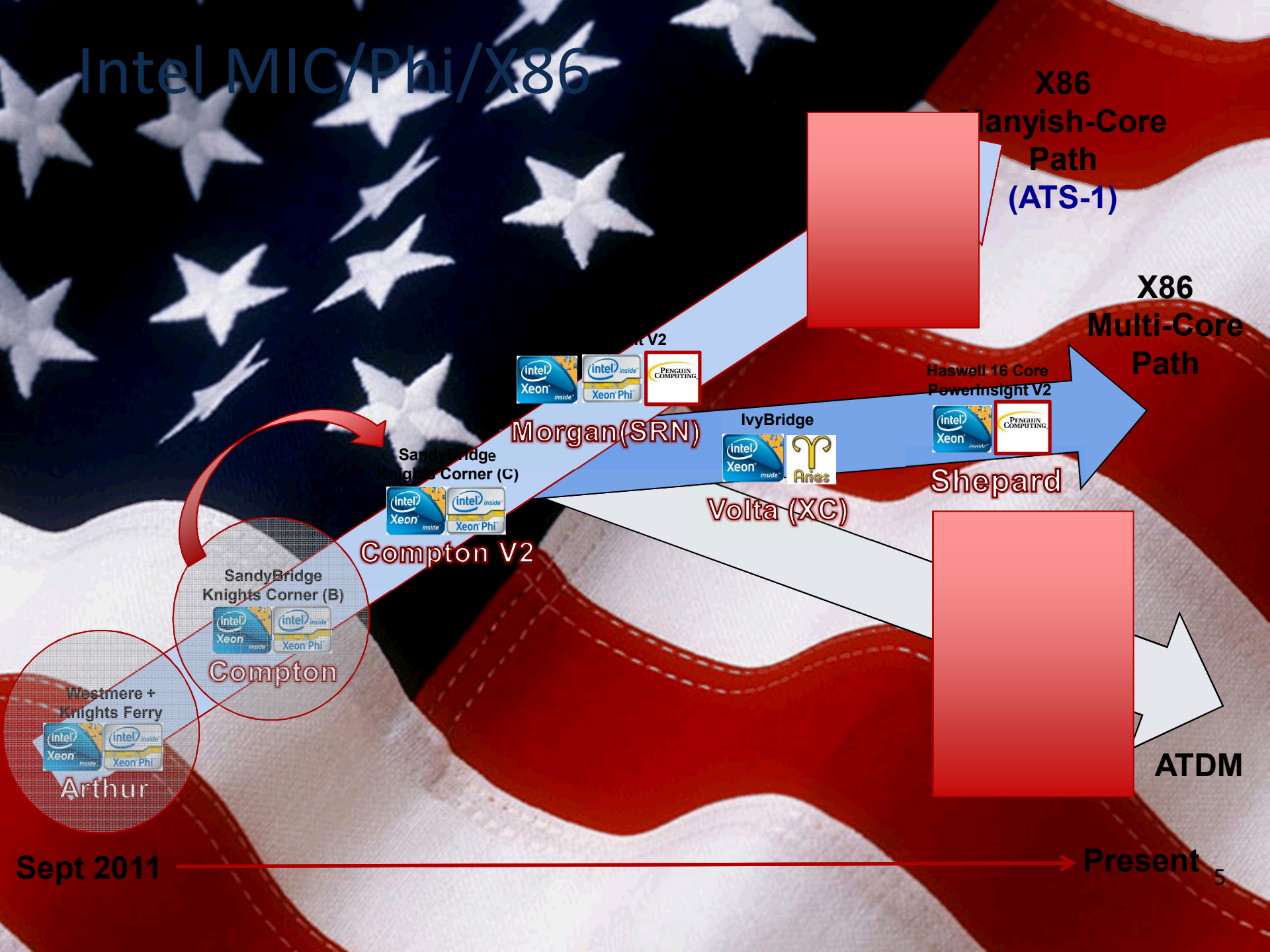
A Co-Design Tool

- Philosophy
  - Hardware and Software integration (as proved to be) to be highly dynamic
  - INTENTIONALLY closed environments than production
  - Systems are NOT for long capability/capacity cycles
  - Priority is to explore broad and diverse set of emerging architectural alternatives ☺
- First Rule of Test Beds
  - NEVER say no
    - At least to any reasonable request ☺

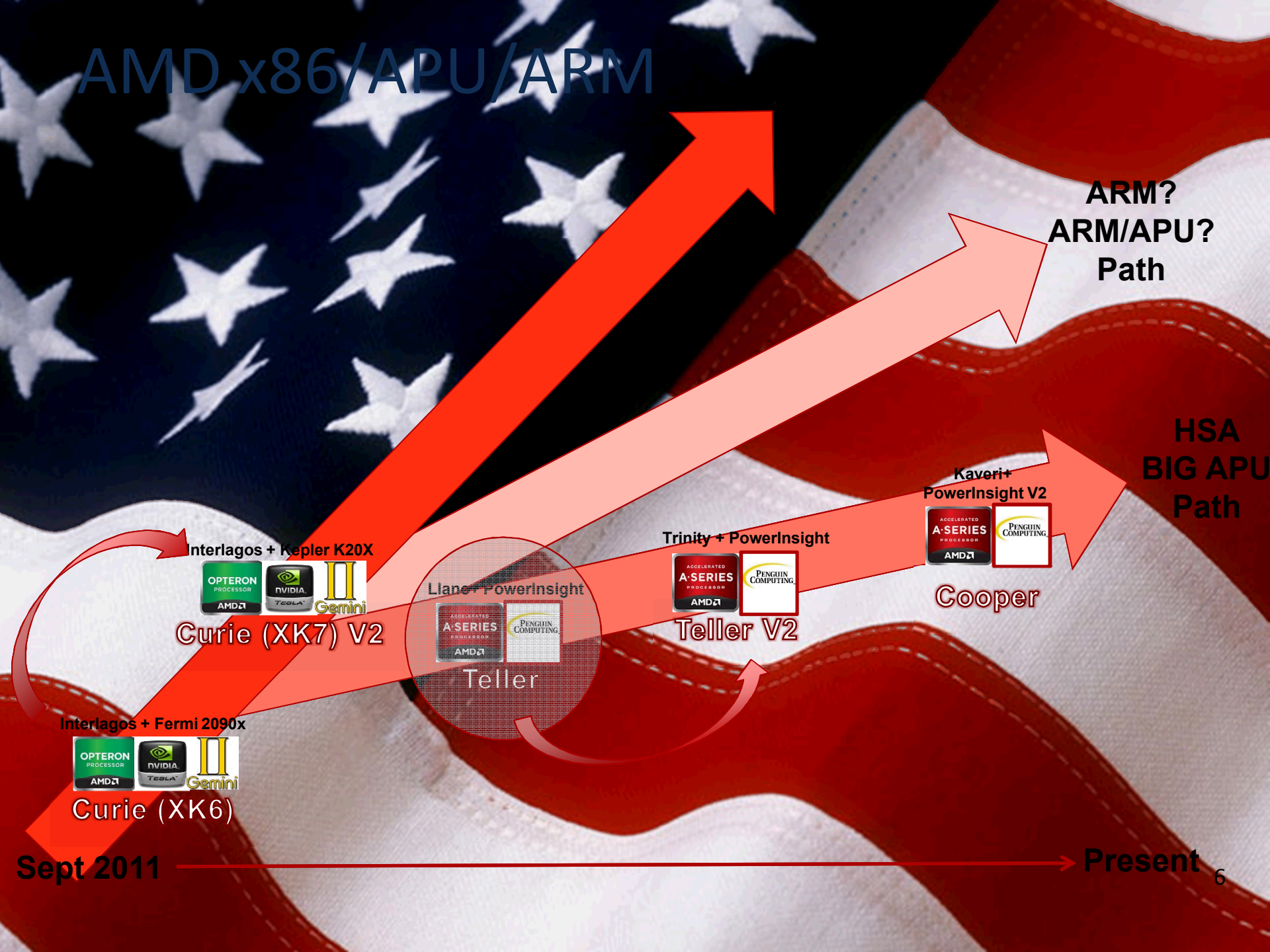
Numerous custom hardware changes
Even more custom software configurations

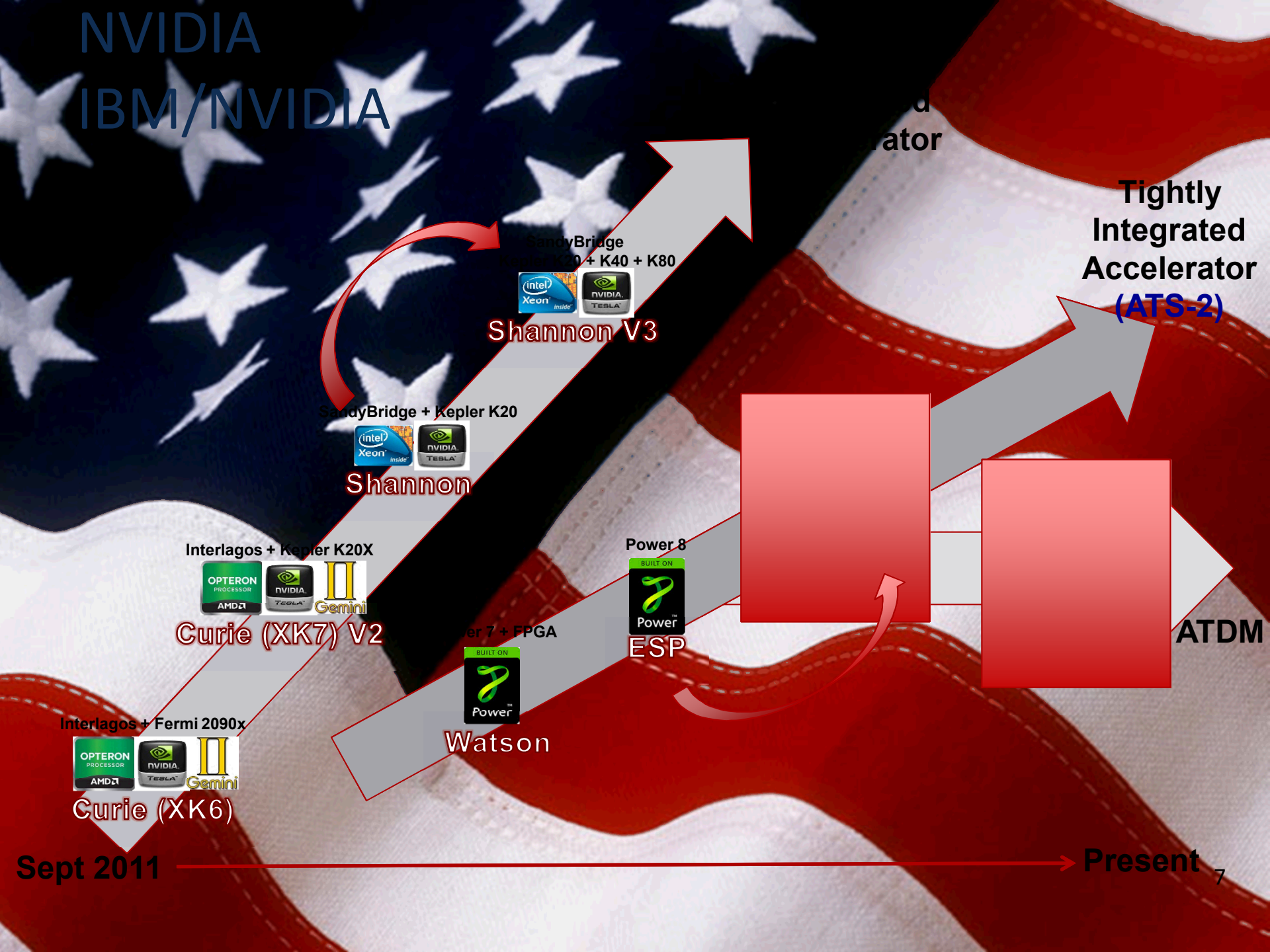


Intel MIC/Phi/X86

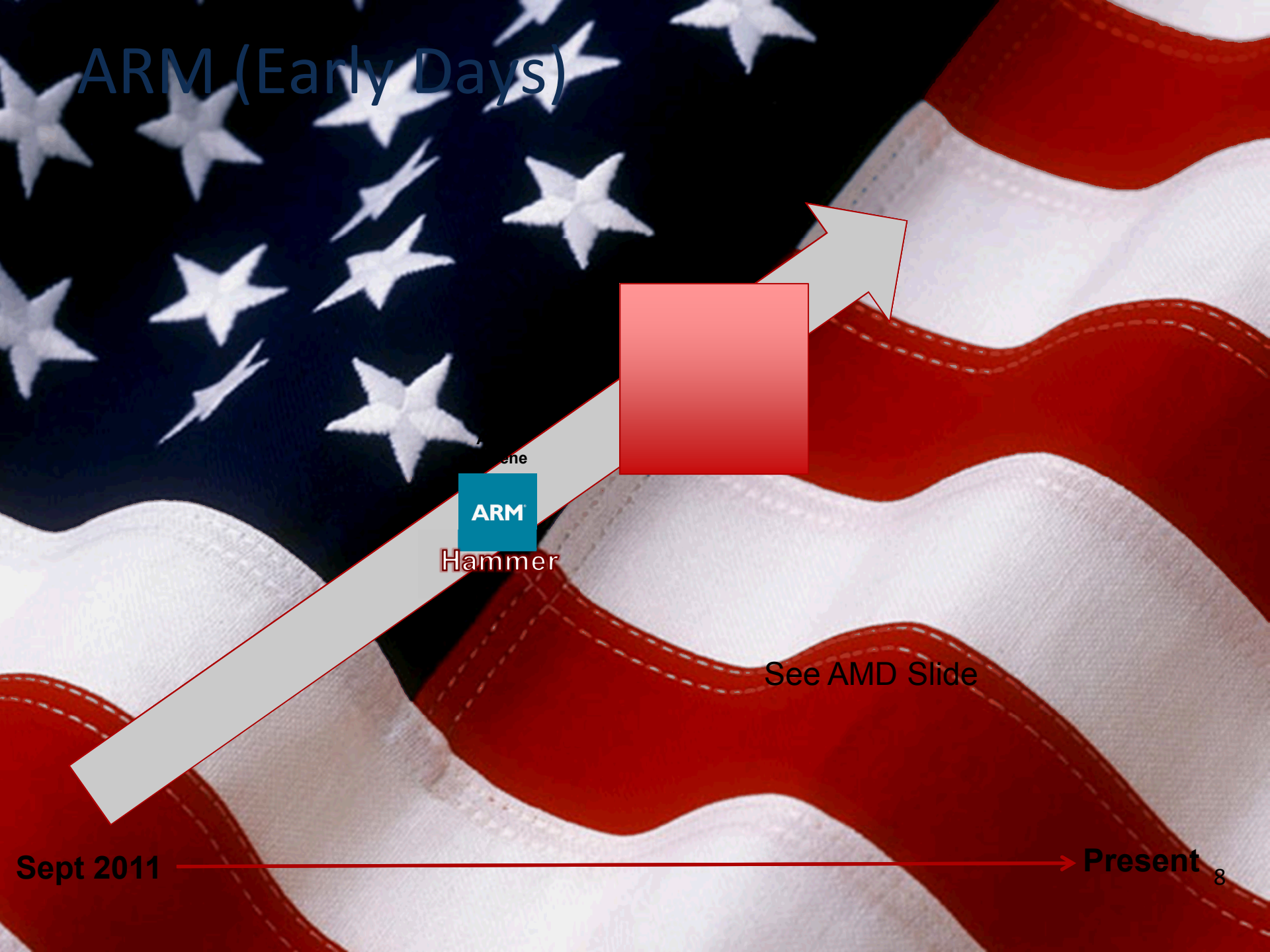


AMD x86/APU/ARM




NVIDIA
IBM/NVIDIA
ator
Tightly
Integrated
Accelerator
(ATS-2)
SandyBridge
+ K40 + K80
Shannon V3
SandyBridge + Kepler K20
Shannon
Interlagos + Kepler K20X
Curie (XK7) V2
Power 8
ESP
er 7 + FPGA
Watson
ATDM
Interlagos + Fermi 2090x
Curie (XK6)
Sept 2011
Present

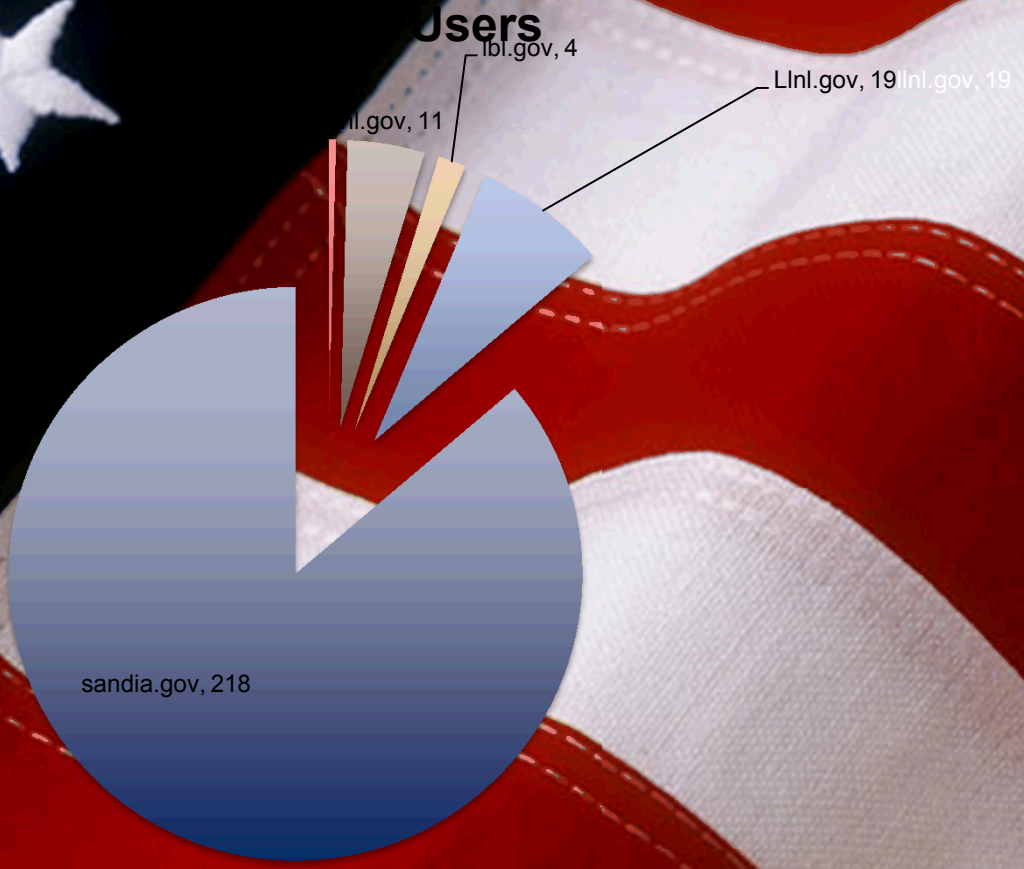
ARM (Early Days)



The Test Bed Community 8/4/2016

- [illegible] unique users – was 195 on [illegible] 24, 2014
- Systems by Type with user [illegible]
  - AMD Integrated GPUs
    - Teller 142
    - Cooper 24
  - ARM
    - Hammer 31
  - IBM Power +NVIDIA
    - Ride 10
    - White 77
  - Intel X86 + MIC
    - Morgan 84
    - Compton 192
  - Trinity + Power Research
    - Shepard 91
    - Volta 144

  X86 + NVIDIA
    - Curie 172
    - Shannon 173

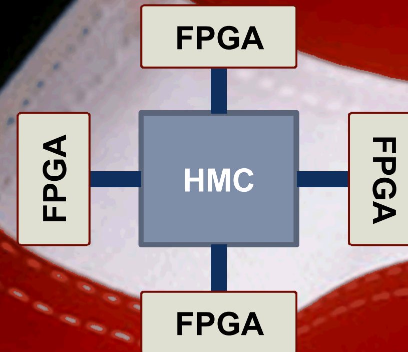


67 - Vendor, University and other users

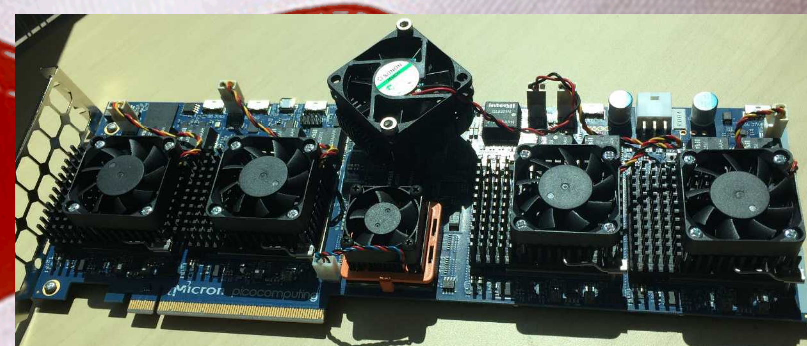
HMC and Micron Test Beds

- 3D stacked memory is a huge opportunity for bandwidth-limited applications
  - Massive bandwidth – 15x
  - 10x performance improvement over DDR3 for GUPS (random access)
- Many questions
  - What data should be placed in HMC vs. conventional DDR?
  - How should data be laid out?
  - How should applications make use of special instructions like atomics, bit-writes, etc.
- Test Beds used to characterize HMC
  - Characterize HMC performance as it varies with configuration, contention, memory access pattern, etc.

Micron-Pico Test Bed

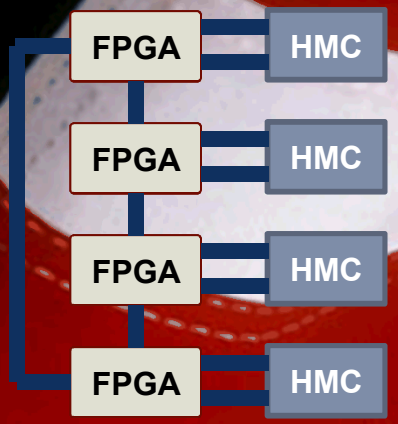


- [illegible]GAs + 1 2GB 4-link HMC
- [illegible]n-board DDR
- Experiments
  - Address patterns (ra[illegible] spmv, etc.)
  - Traces from mini-app[illegible]
- Results from Ra[illegible]near access patterns
- Other experi[illegible]hold while boards are at Micron for upgrades & re[illegible]
  - Had some issues with socket reliability
  - Testing has revealed some controller bugs that Micron is working on



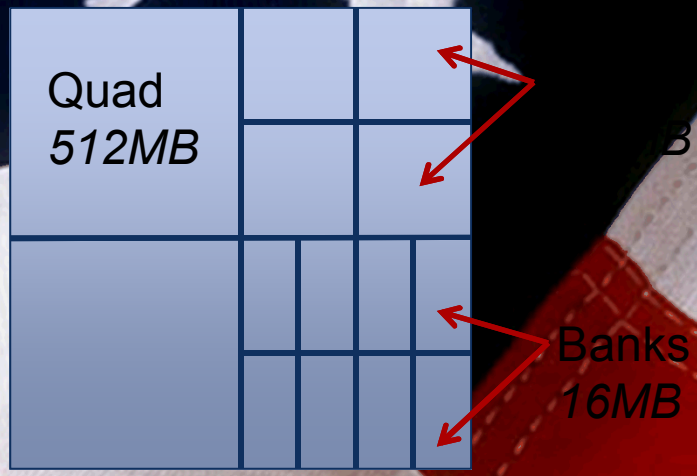
Micron-Convey Test Bed

- [illegible]nt arrival, no results yet
- [illegible]wo architectures
  - "multi": 4 FPGAs + 4 2
    - FPGAs also interco
  - "mini": 1 FPGA + 1 2
- Powerful, shared [illegible] co-processor model enables tight interaction b[illegible]t & board
  - Ability to de[illegible]m instructions which run on board
  - Host applicati[illegible]n call these instructions
  - Also, CAPI-enabled – future avenue for research



Initial studies on Pico Test Bed

- [illegible] GUPS (random accesses)
- [illegible]udy: Impact of contention on achievable bandwidth
  - Increase contention by [illegible] to access smaller and smaller areas within the HMC

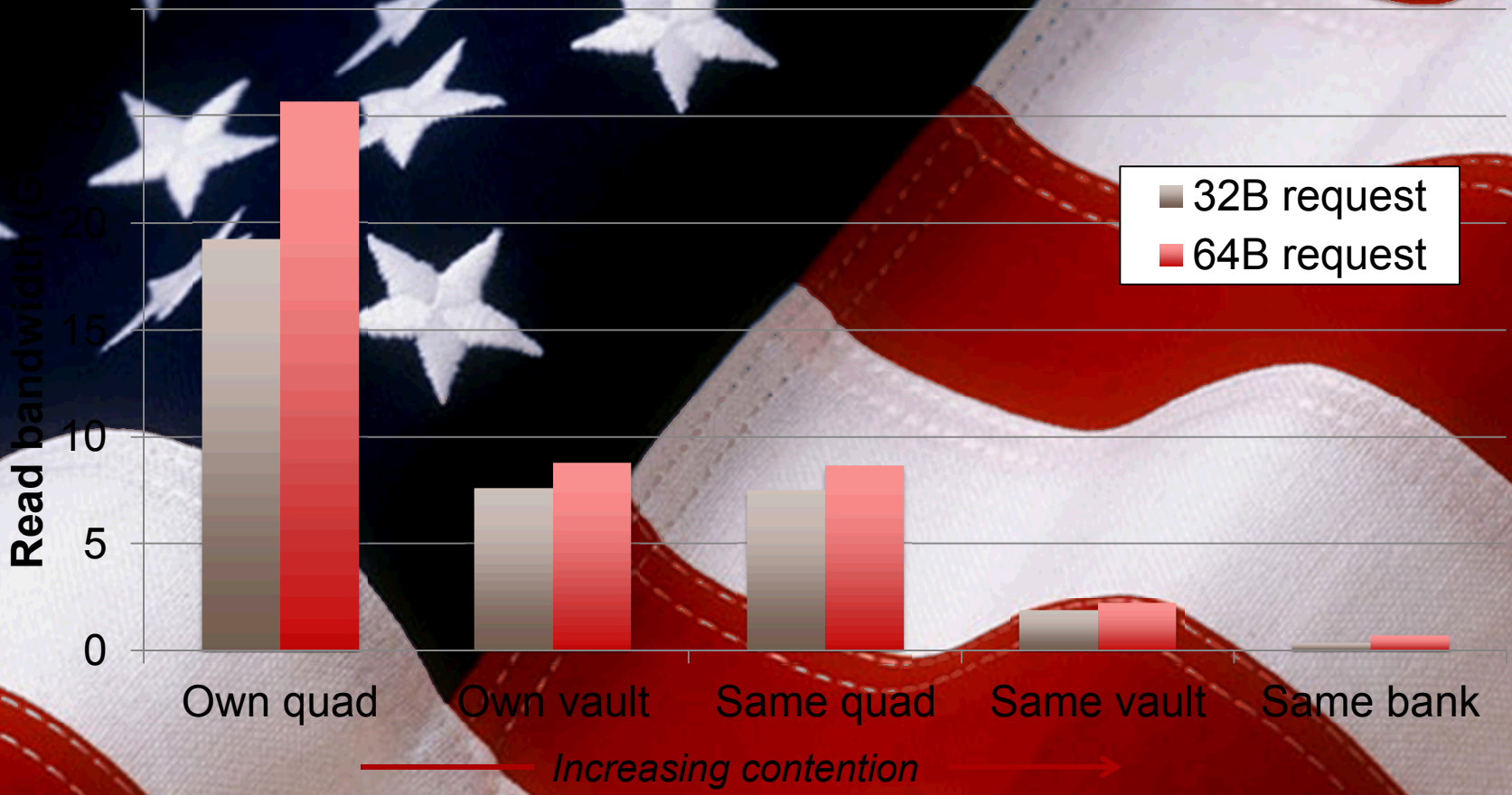


Access patterns
- Own quad: links access their local quad
- Own vault: links access one vault in their local quad
- Same quad: links access a single quad
- Same vault: links access a single vault
- Same bank: links access a single bank

- Study: Impact of request size on achievable bandwidth
  - Request size: 16B-128B (traditional DDR requests are 64B)
  - Compare read-only to read-write accesses

Impact of contention on bandwidth

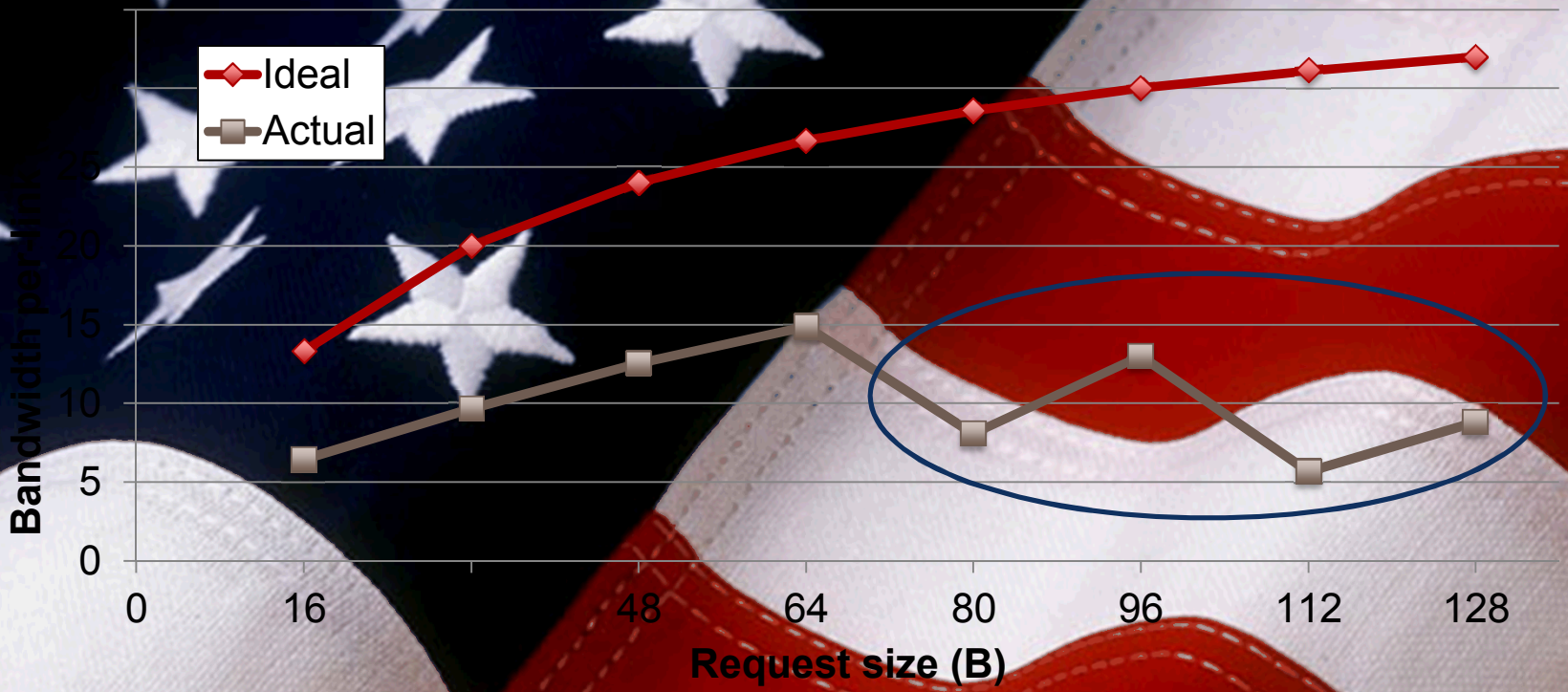
**Reported bandwidths are for all links*



- Significant degradation in bandwidth with increasing contention
  - Implication: Layout memory to spread accesses across vaults

Read-write bandwidth vs. request size

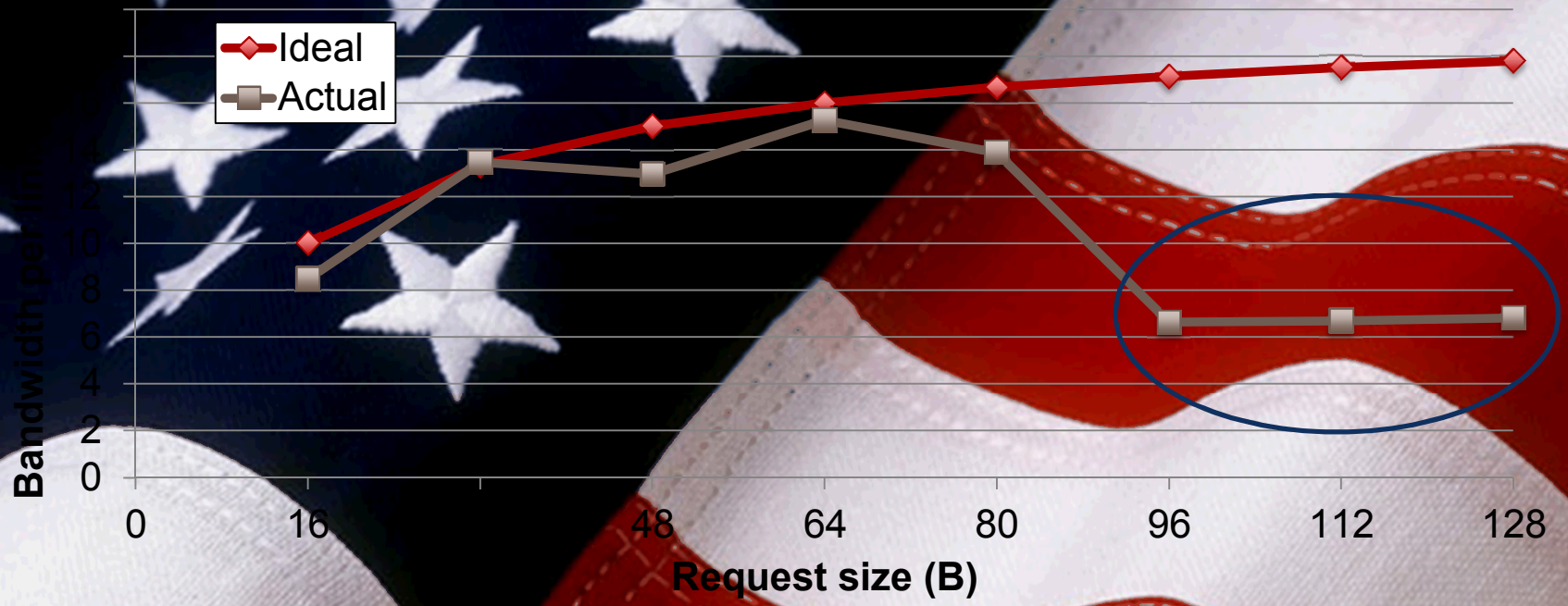
** Reported bandwidths are for a single link*



- Measured trend tracks *ideal* for requests smaller than 64B
- Controller bugs cause decreased throughput for larger requests
  - Micron is working on some solutions for these

Read-only bandwidth vs. request size

** Reported bandwidths are for a single link*



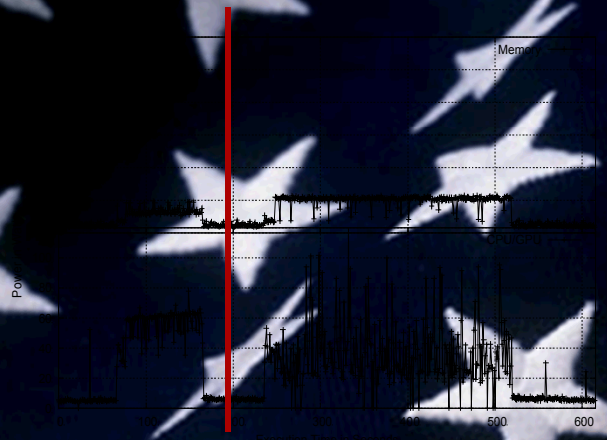
- Read data bandwidth very closely tracks ideal to 64B requests
- Implication: Read/write ratio will be important to achievable bandwidth

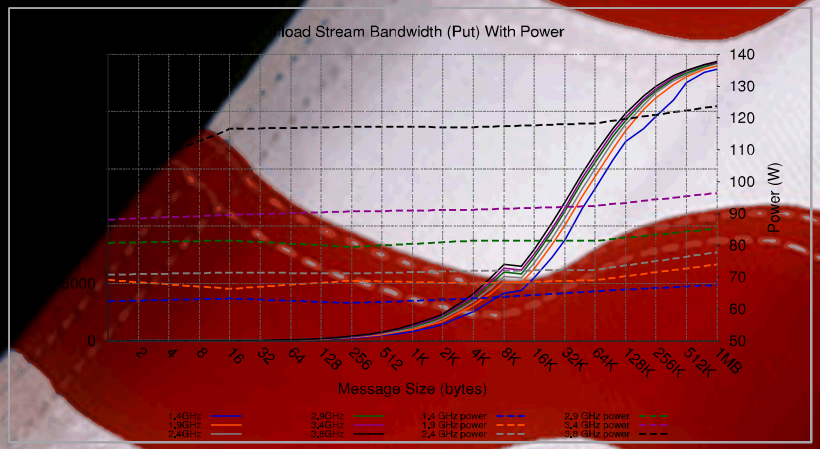
Questions?

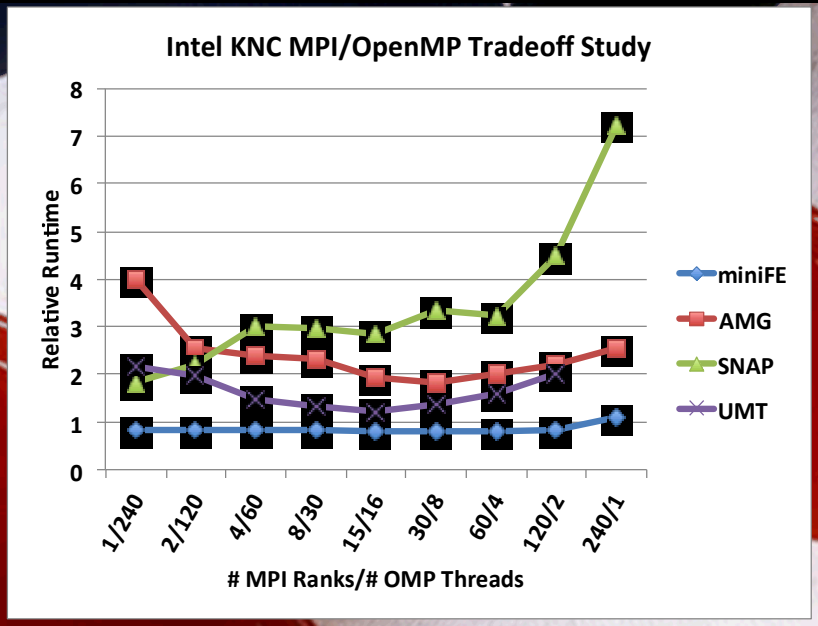
Exceptional service in the national interest

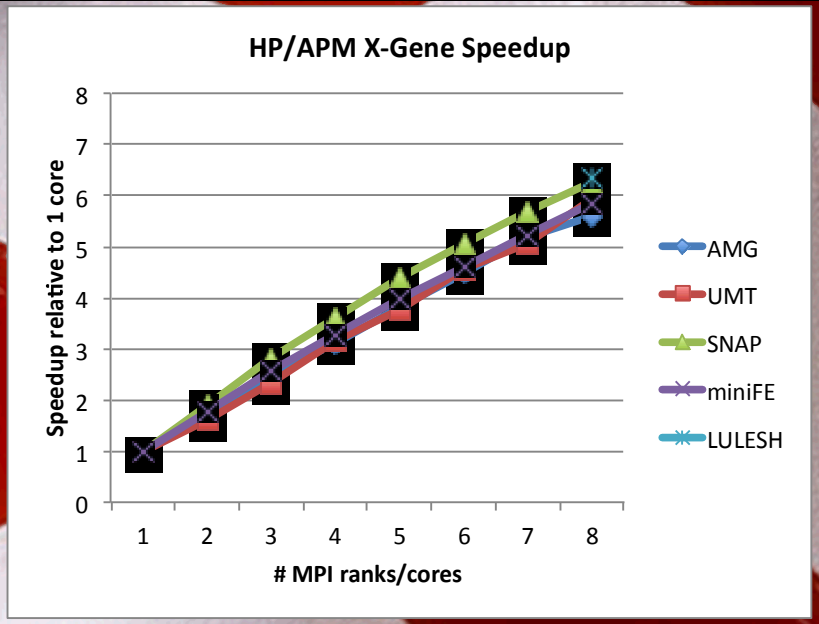
Backup

Example Analysis using Test Beds









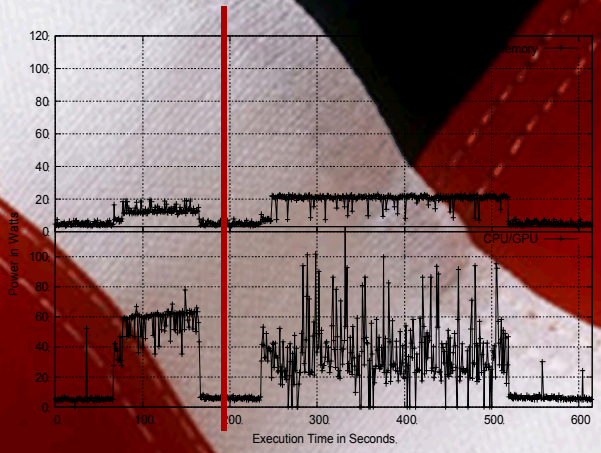
Sandia/Penguin Collaboration

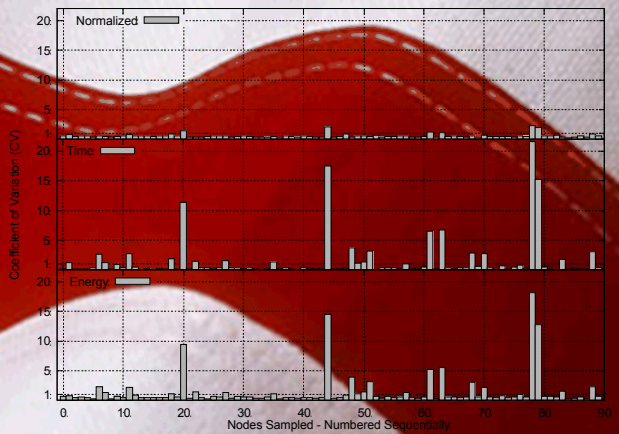
small example of co-design

Commodity

ustom PCB

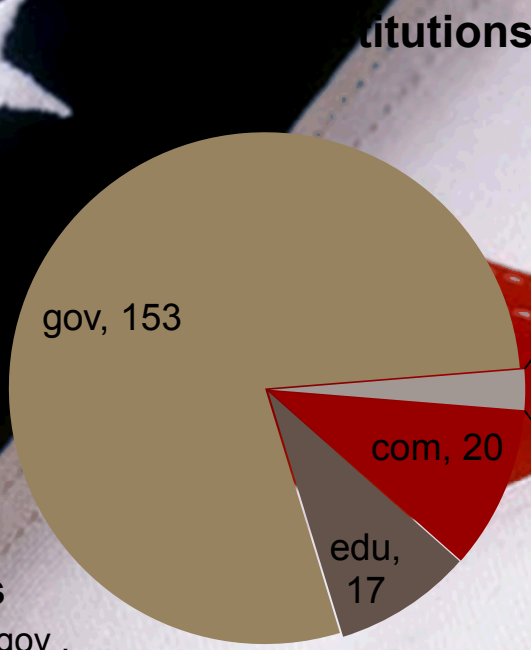
PowerInsight

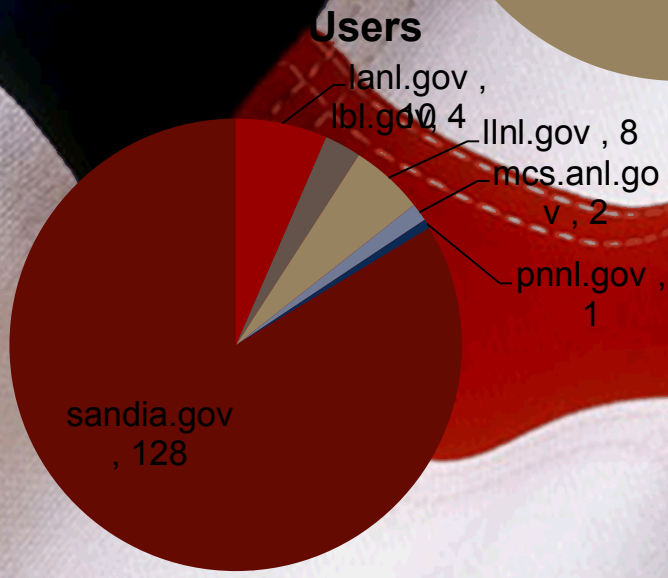




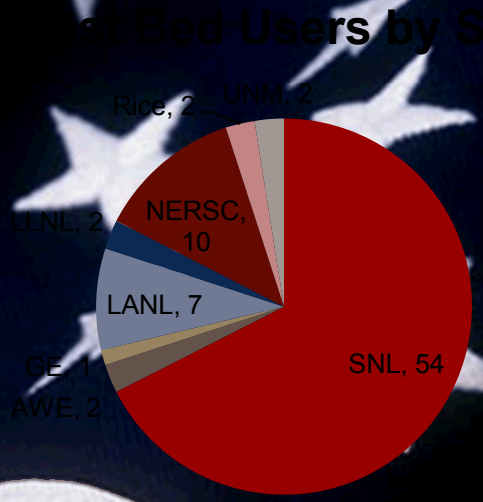
The Test Bed Community 2/24/2014

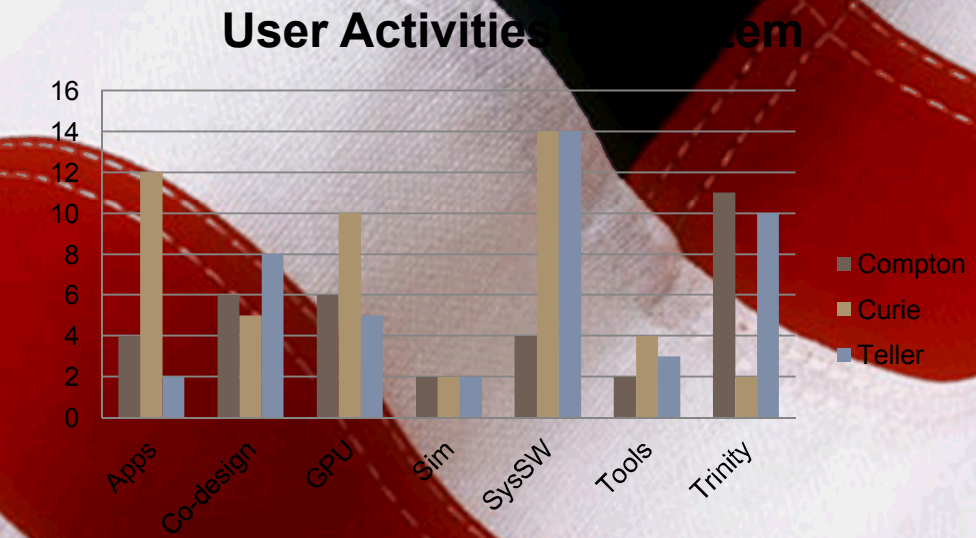
- Unique users – was on Feb 8, 2013
- Users by system:
  - Compton 123
  - Curie 112
  - Shannon 101
  - Teller 120
  - Volta 45





The Test Bed Community 2/8/2013

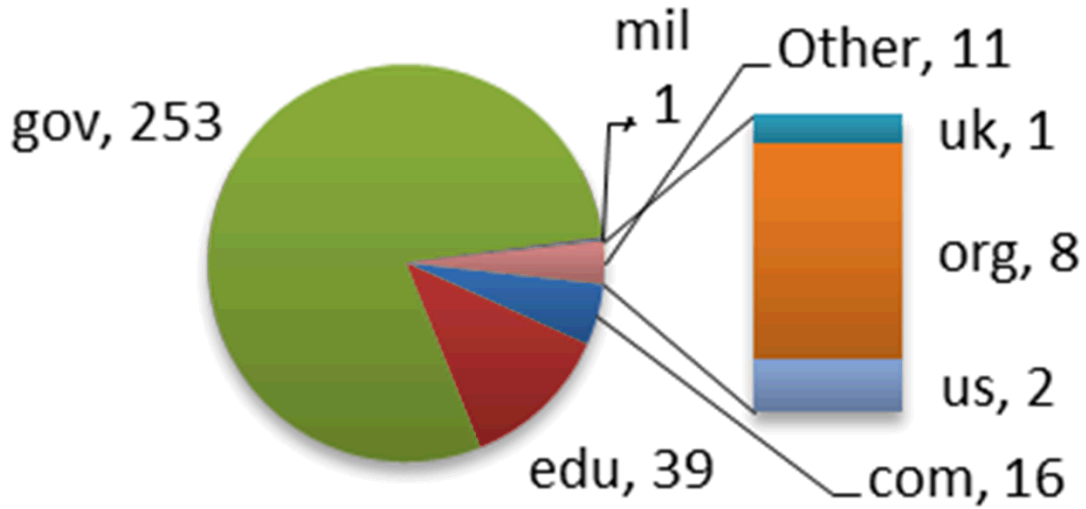


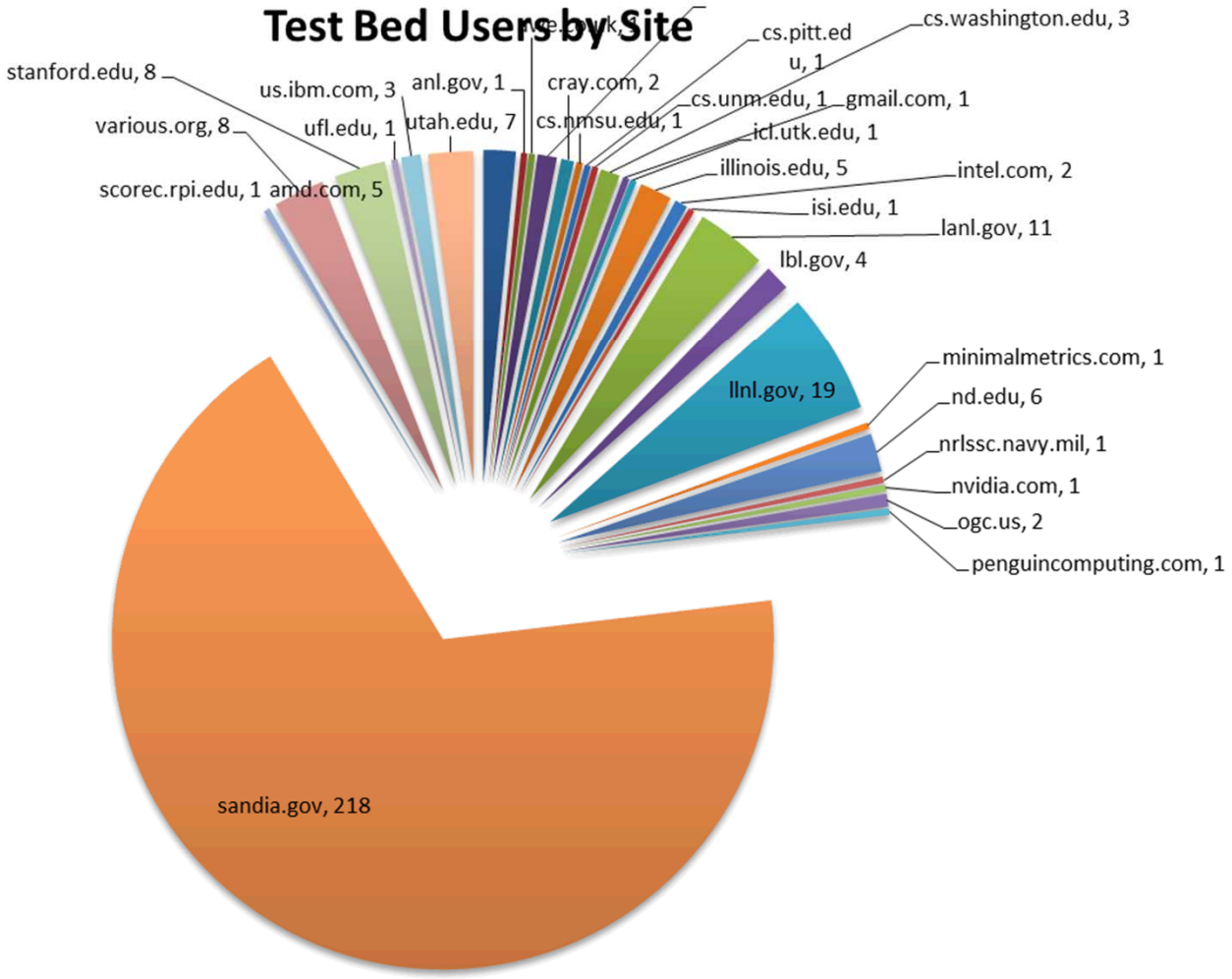


...e justification submissions:

- ...codes for large scale PDE ...strained optimization ...roblems
- Evaluating programming model support for Intel MIC
- Resilience testing for exascale systems
- Supporting Trinity Benchmarks and acceptance tests
- Experiment with AWE applications
- Performing initial assessments of using accelerators with CTH.
- Experiments with KNC to see if it will work with our RDMA work.

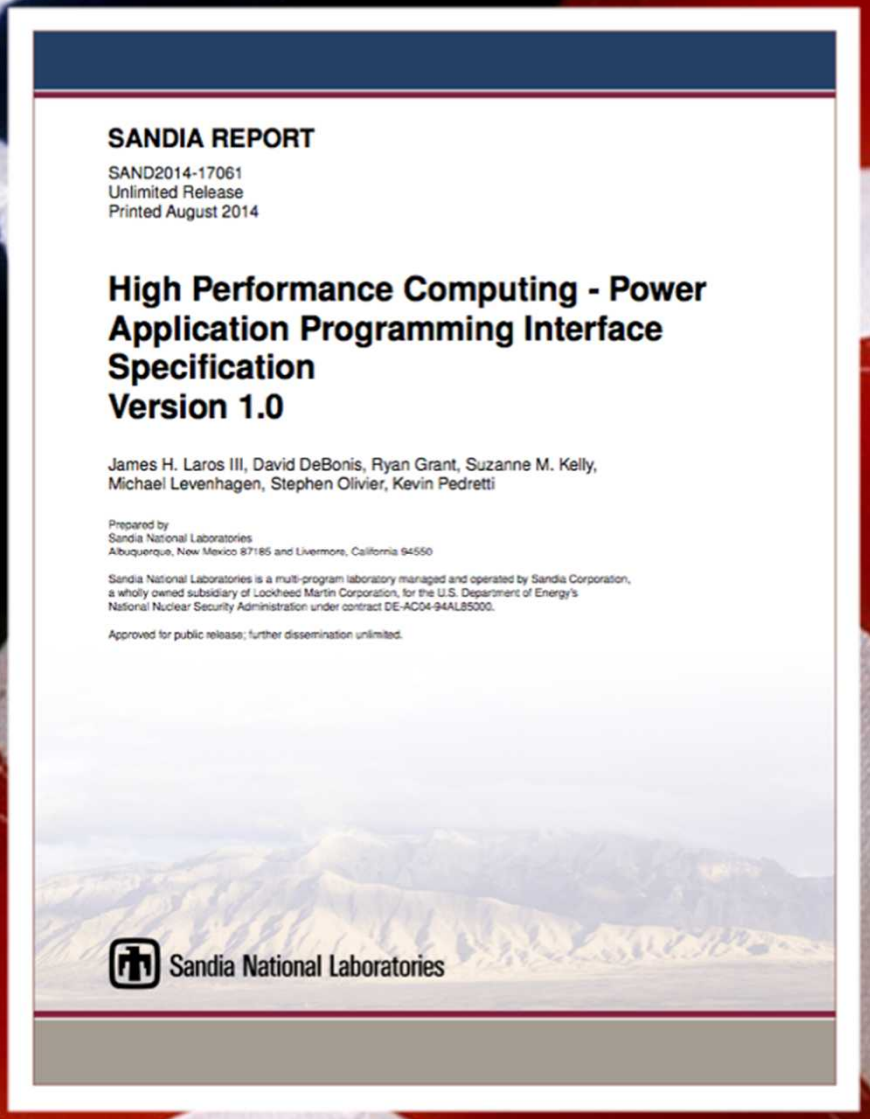
Users Institution by Type




Test Bed Users by Site
stanford.edu, 8
various.org, 8
scorec.rpi.edu, 1
us.ibm.com, 3
ufl.edu, 1
amd.com, 5
anl.gov, 1
utah.edu, 7
cray.com, 2
cs.nmsu.edu, 1
cs.pitt.edu, 1
cs.unm.edu, 1
cs.washington.edu, 3
gmail.com, 1
icl.utk.edu, 1
illinois.edu, 5
intel.com, 2
isi.edu, 1
lanl.gov, 11
lbl.gov, 4
llnl.gov, 19
minimalmetrics.com, 1
nd.edu, 6
nrlssc.navy.mil, 1
nvidia.com, 1
ogc.us, 2
penguincomputing.com, 1
sandia.gov, 218

Version 1.0 Delivered August 26th 2014

- Community needed a portable API for **measuring** and **controlling** power and energy
- Sandia developed PowerAPI specification to fill this gap
- Provides portable power **measurement** and **control** interfaces
- Covers full spectrum facility to compone
- First production implementation will be Trinity (ATS1)
- Continued (increasing) community involvement and influence



SANDIA REPORT

SAND2014-17061
Unlimited Release
Printed August 2014

High Performance Computing - Power Application Programming Interface Specification Version 1.0

James H. Laros III, David DeBonis, Ryan Grant, Suzanne M. Kelly, Michael Levenhagen, Stephen Olivier, Kevin Pedretti

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories