



ClearSpeed Technology: Enabling the Petascale Era

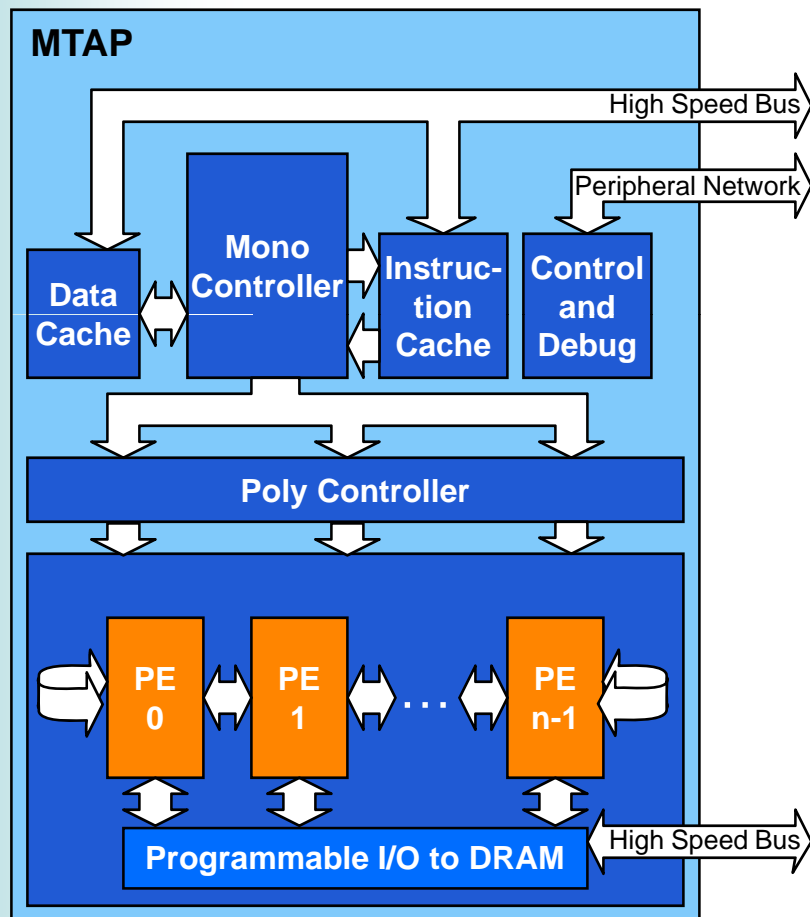
Simon McIntosh-Smith, simon@clearspeed.com
VP of Customer Applications

ISC08, June 16th 2008

ClearSpeed Introduction

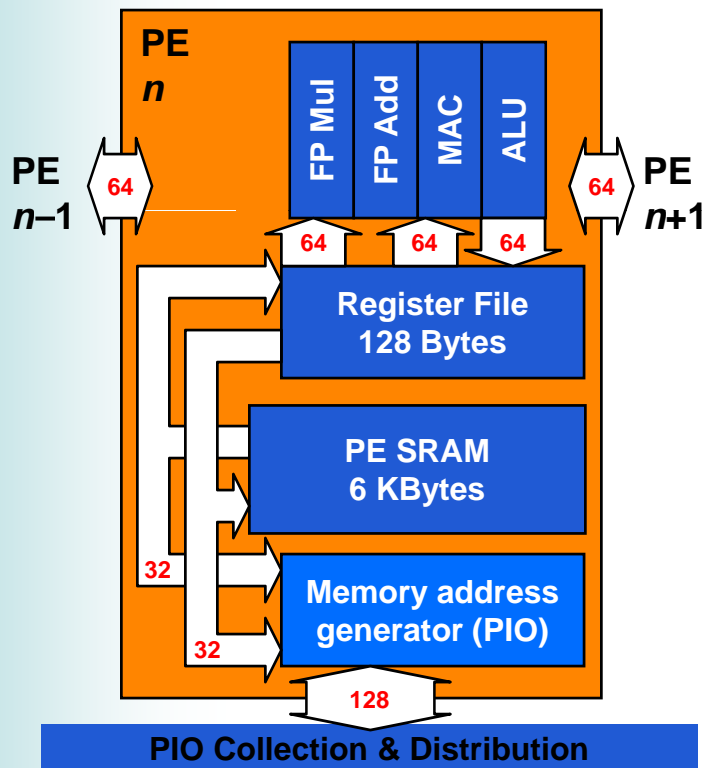
- **Founded in 2001 we are a European company, based in Bristol, UK with offices in San Jose, CA and Tokyo**
- **We are driven by the principle that to deliver HIGH PERFORMANCE you have to deliver HIGH RELIABILITY and HIGH POWER EFFICIENCY as systems become larger and more space, power and cooling constrained**
- **Therefore we deliver the world's most power efficient, high-performance processors, with supporting subsystems, software development tools, libraries and applications**
- **We provide solutions for both the High Performance Computing (HPC) and embedded systems markets**
- **We are partnered with HP, IBM, SGI, Sun, Bull and other OEMs to deliver accelerated systems**

ClearSpeed's accelerator architecture



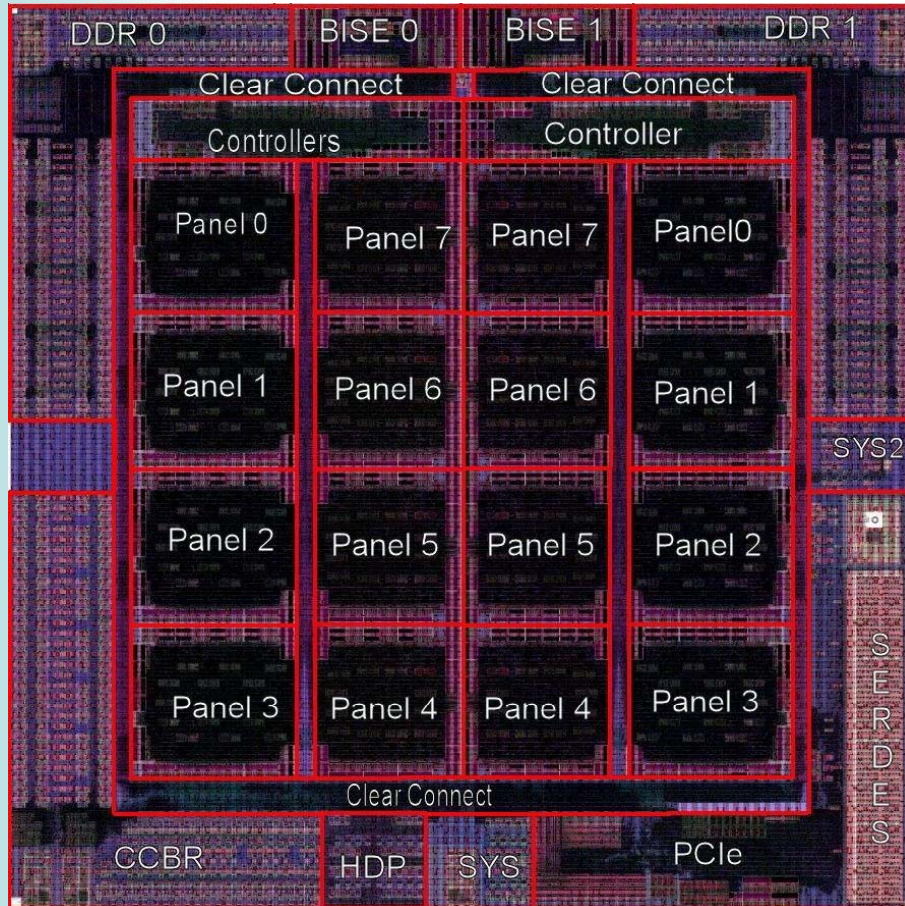
- Architecture designed for Coarse-Grained Data Parallel Processing:
 - Achieves high performance, low power
 - Multi-threading enables asynchronous, overlapped I/O with compute
 - Scalable array of many Processor Elements (PEs)
 - Includes enterprise-class reliability features necessary for HPC, such as ECC on memories, spare PEs etc.
- Programmed in an extended version of ANSI C called Cⁿ:
 - Rich expressive semantics
 - Single “poly” data type modifier

Processing Element (PE) architecture



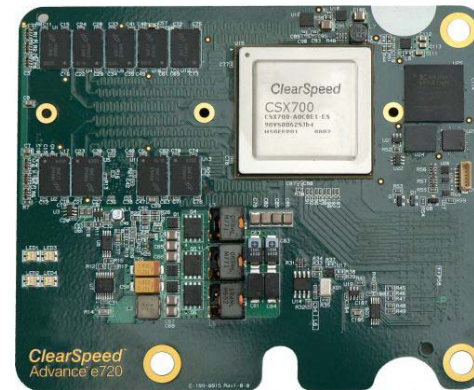
- Multiple execution units per PE
 - Floating point adder } 32 & 64-bit
 - Floating point multiplier } IEEE 754
 - Fixed-point MAC 16x16 → 32+64
 - Integer ALU with shifter
 - Load/store
- High-bandwidth, 5-port register file per PE
- Fast inter-PE communication path (swazzle)
- Closely coupled ECC protected SRAM for data
 - Keeping data close is key to **low power**
- Per PE address generators & DMA (PIO)
 - Complete pointer model, including parallel pointer chasing and vectors of addresses
 - Key for gather/scatter and sparse operations

The CSX700 – “Callanish”



- **Processor Cores:**
 - 192 Processor Elements (2x96)
 - 96 double precision GFLOPS
 - 250MHz
 - Error Correction (ECC) on all internal memories
- **SoC details:**
 - Integrated PCI Express x16
 - 2x integrated ECC DDR2 memory controller + scrubber
- **Design details:**
 - IBM 90nm process
 - 256 million transistors
- **12W Max (Power Managed)**
- **Launched at ISC08**

The ClearSpeed Advance™ e710 & e720 accelerators



- **Enterprise-class HPC accelerators**
- **Designed to fit into existing servers such as 1U & blades**
 - Low power consumption – 25W max; small, light
- **Designed for high reliability (MTBF)**
 - All memory is error protected; no moving parts (e.g. fans) are required
- **96 Double Precision (D.P.) IEEE 754 GFLOPS peak**
 - ~4 GFLOPS per watt double precision
- **Over 2 GBytes/s between accelerator and host**
- **No extra power connectors, cooling or space (slots) required**
- **Launched at ISC08**

Powerful software development environment

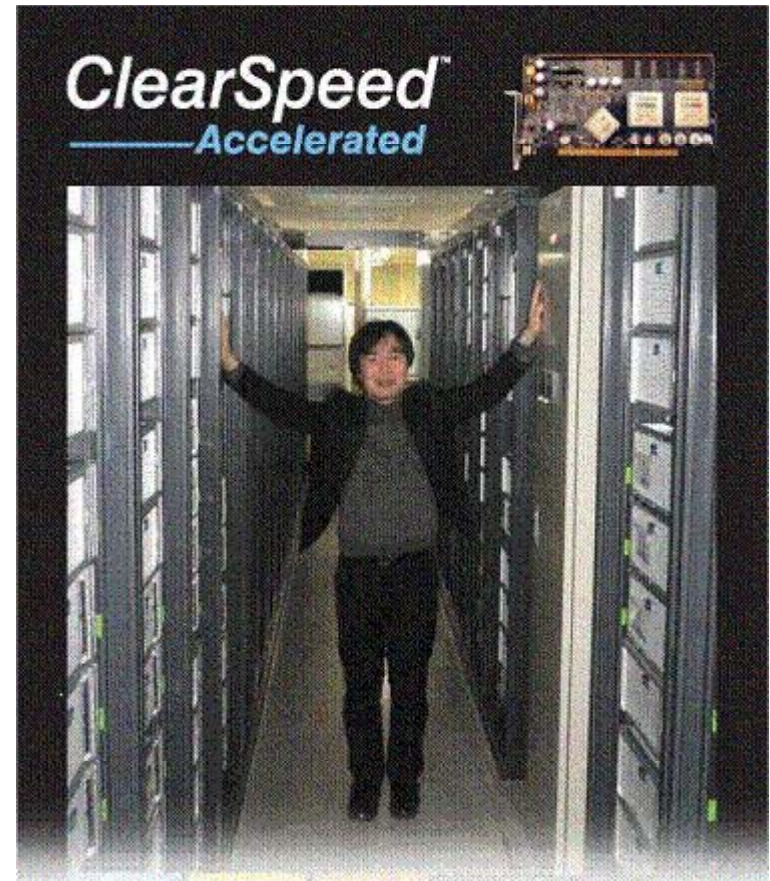
- **Version 3.10 beta release at ISC**
- **Binary compatible across all ClearSpeed products**
 - The same executables run on the CSX700 and CSX600
- **Cⁿ optimising compiler**
 - C with **poly** extension for SIMD datatypes
 - Uses mature ACE CoSy compiler development system
 - Generates performance within 90% of hand optimized performance on many codes
- **Debugger – a port of gdb**
 - *Runs on ClearSpeed's hardware at full speed*
- **Profiling – csprof**
 - Heterogeneous, system-wide visualisation of an accelerated application's performance while running on both a multi-core host and ClearSpeed accelerators. *ClearSpeed's hardware can be profiled in real-time*
- **Libraries (BLAS, RNG, FFT, more...) & High level APIs**
- **Preview of an ECLIPSE IDE**



Accelerators in the Top500

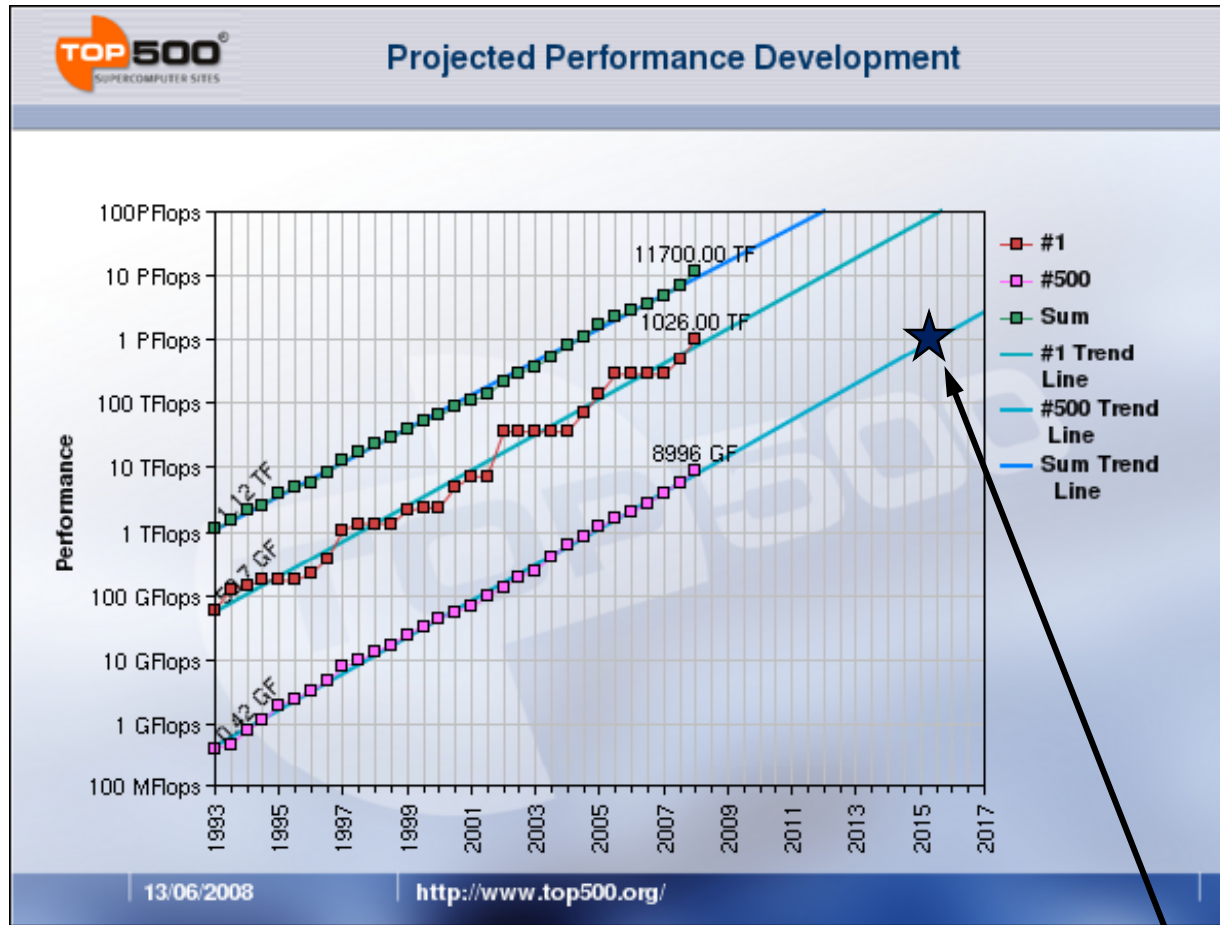
ClearSpeed enabled the *first* accelerated system in the Top500: TSUBAME at Tokyo Tech (announced October 9th 2006).

- Upgraded twice since and now includes 648 ClearSpeed Advance x620 cards
- Performance increased by 68% for just a 2% increase in power consumption, ~5% increase in cost, no increase in size or cooling
- #9 in November 2006 Top500



Professor Matsuoka standing beside TSUBAME at Tokyo Tech

Petascale – why does it matter?



- **Within 7 years *everything* will be Petascale!**

How do we get to Petascale systems?

- **Not by just having more multi-core CPUs**
- **Many now believe accelerators are the only way to scale to Petaflops**
- **“My prediction: High performance computing will soon be dominated by accelerator-based systems.” – Michael Wolfe, The Portland Group**
- **Accelerators can be optimized for a narrower range of applications than mainstream CPUs**
- **Thus accelerators can better exploit the inherent parallelism in most HPC applications**

Making Petascale Possible

- **ClearSpeed designs accelerators specifically for HPC and to scale to Petaflops:**
 - Highly reliable – all memories error protected, no moving parts
 - Very power efficient – industry leading GFLOPS per watt
 - Focused on 64-bit IEEE754 floating point
 - Programmed in standard languages with simple extensions
 - Full hardware support for run-time debugging and profiling
 - Sophisticated software tools for system-wide performance tuning
- **This focus enables Petascale systems to be built with:**
 - Useful uptimes and delivering reliable results
 - Feasible power consumption, space and cooling requirements
 - Much smaller carbon footprints

The ClearSpeed Accelerated Terascale System

CATS-700 being announced at ISC08



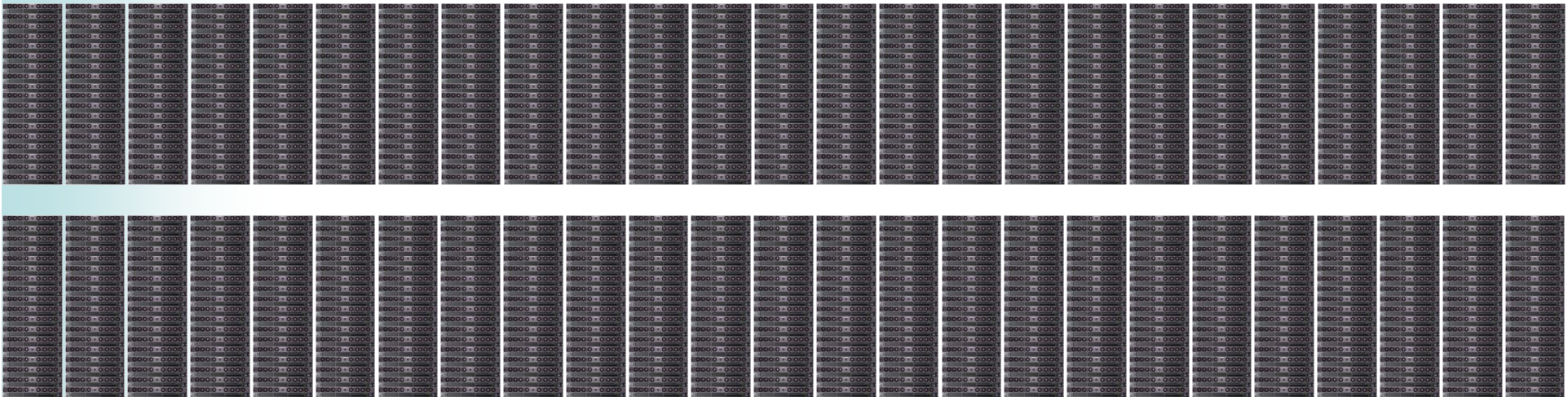
- **Enterprise-class reliability – error correct/detect on *all* memories**
- **1.152 TFLOPS double precision (64-bit) in 1U**
- **12 Advance™ e710 accelerators**
- **24 GBytes of DDR2 DRAM with ECC**
- **96.0 GBytes/s of DRAM bandwidth**
- **3.192 Giga Addresses/s DRAM address bandwidth**
- **400 watts typical power consumption**
- **Two PCI Express x8 connections to the host (up to 3m long)**
- **Up to 41 TFLOPS double precision peak in a single rack**
 - From 36 CATS-700 1U nodes
- **10X greater performance than the fastest dual socket quad-core servers (3GHz x86) at the *same* power consumption**

A rack of CATS



- Enterprise-class reliability – ECC on *all* memories, both on- and off-chip
- From 18 CATS-700:
 - 20.7 TFLOPS double precision
 - 432 GBytes of DDR2 with ECC
 - 1.73 TBytes/s of DRAM bandwidth
 - 7.2 KW typical power consumption
- From 18 3GHz quad core hosts:
 - 1.8 TFLOPS double precision
 - 7.2 KW typical power consumption
- **22.5 TFLOPS double precision total**
- **14.4 KW total power consumption**
- No silent software errors

A Petaflop of CATS



- Enterprise-class *reliable* Petaflop
- 1.125 PFLOPS double precision peak
- 50 racks
- 750KW typical power consumption
- 21.6 TBytes of CATS DRAM memory – all with ECC
- 86.4 TBytes/s of CATS DRAM memory bandwidth
- Can be delivered *this year!*
- Roadrunner is 6X the size and 5X the power consumption

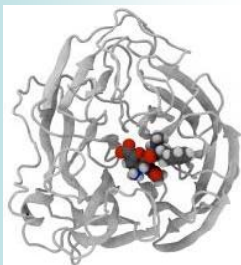


Real Science
Real Numbers
Real Software

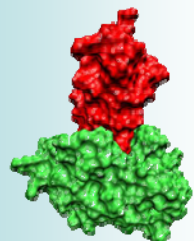
Petascale applications

Example Petascale applications on CATS

- **Computational chemistry:**



- The Density Functional Theory component of the Molpro Quantum Chemistry code has been ported to CATS and sees a 10X speedup wallclock time
 - Enables QM/MM codes to apply very high accuracy methods to much larger scale molecules than ever before – e.g. neuraminidase inhibitors in the search for better flu virus vaccines



- The high-accuracy Molecular Mechanics-based docking code BUDE has been ported to CATS and also sees a 10X speedup
 - A paper describing this work will be presented at ISC08
- These computational chemistry methods need petascale performance to solve the **grand challenges** in nanotechnology, energy (photon/electron transport), materials, rational drug design, rational enzyme design and solid/liquid phase change simulations amongst others

Example Petascale applications on CATS

- **Computational Electromagnetics (CEM):**
 - Similar to LINPACK in application characteristics
 - Most of the computational cost is in LAPACK and L3 BLAS functions such as ZGETRF and ZGEMM
 - Applications include radar cross section, antenna design and electromagnetic wave propagation when not in free space
 - Seeing up to a 5X speedup with e710 and e720
- **Computational Fluid Dynamics:**
 - Especially turbulent fluid flow
 - Needs orders of magnitude more performance than available today
- **Monte Carlo based methods**
 - E.g. Computational finance
 - Seeing up to a 20X speedup on 64-bit codes with CATS



Summary

- **ClearSpeed accelerators including CATS were designed specifically for HPC from the ground up**
- **Have focused on *Enterprise-class* features to enable reliable Petascale systems**
- **Real applications starting to prove accelerators can enable new real science**
 - **E.g. real compounds synthesized from BUDE simulations, better Neuraminidase inhibitors being found with Molpro/Sire**
- **CATS enables Petascale systems and beyond, in smaller form factors than previously possible and within existing infrastructure constraints**
 - **E.g. 100 TFLOPS double precision in 4 CATS racks *today*, small enough to fit in a department, and only 60KW**
- **Very soon *everything* will be Petascale!**

ClearSpeed™

