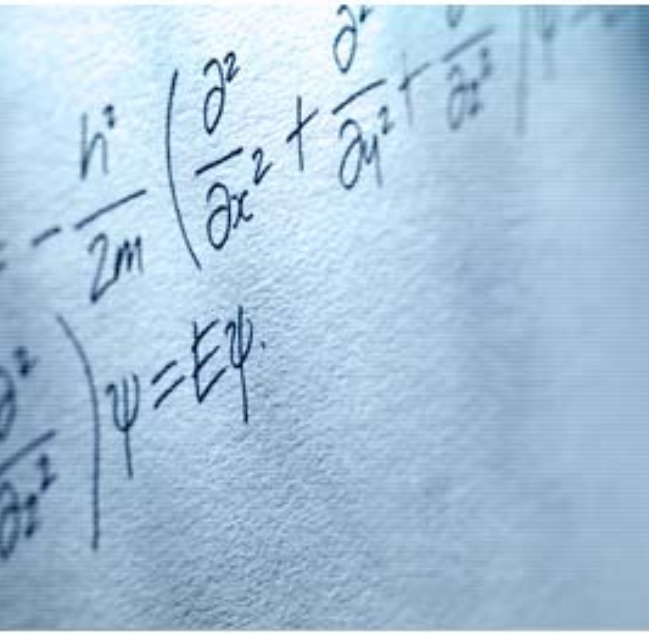
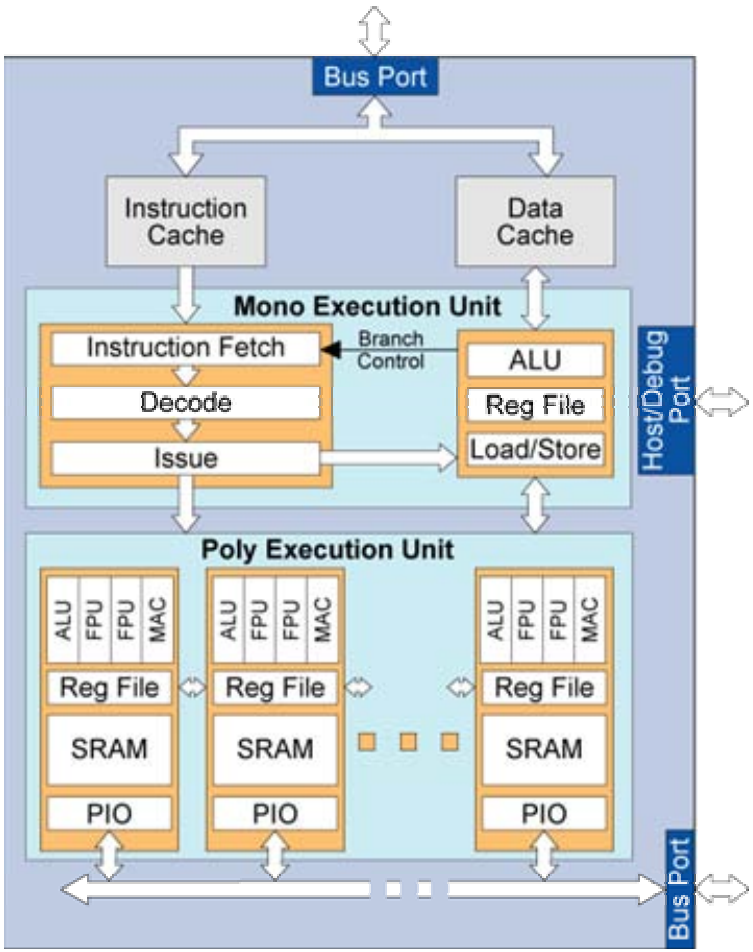


ClearSpeed's CS301:
The World's First Commercially-Available Stream Processor

Architecture, Algorithms and Benchmark Results

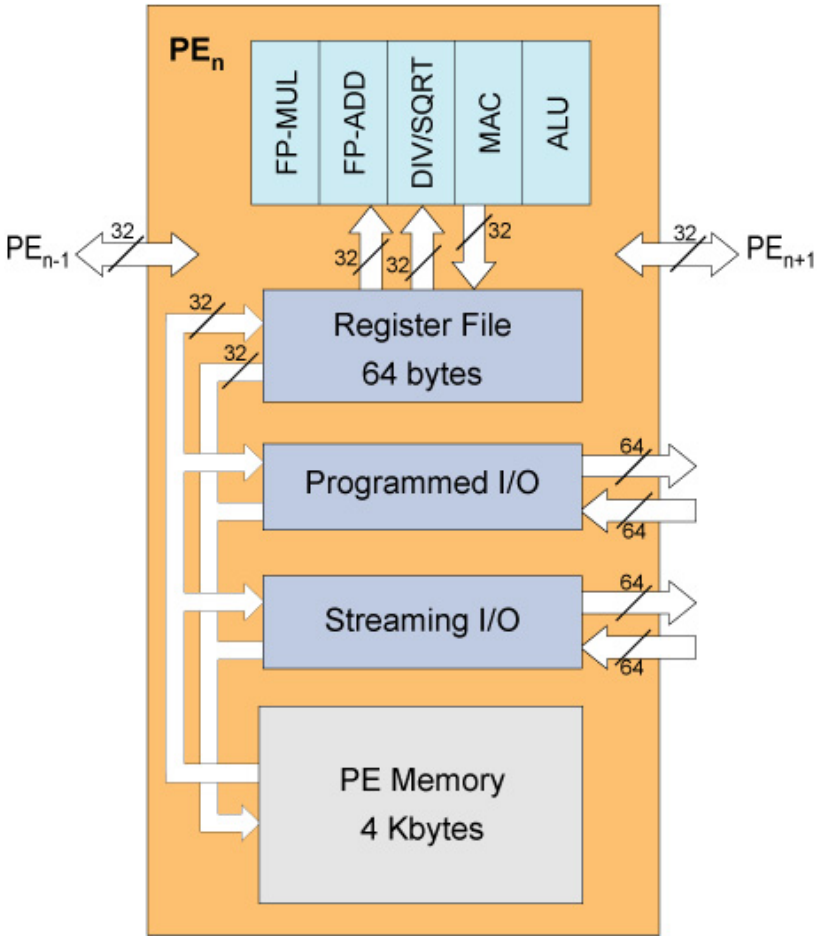
Simon McIntosh-Smith simon@clearspeed.com
Dairsie Latimer dairsie@clearspeed.com
Ron Bell ron.bell@awe.co.uk
Stephen Hudson stephen.hudson@awe.co.uk

CS301 Processor



- Multi-threaded Array Processing
  - Programmed in high-level languages
  - Hardware multi-threading
    - Enables simultaneous data streaming and computation for latency tolerance
  - Run-time extensible instruction set
- Array of Processors Elements
  - PEs are VLIW cores
  - Flexible data parallel processing
  - Built-in PE fault tolerance, resiliency
- High performance, low power
  - 10 GFLOPS/Watt
- Multiple high bandwidth I/O channels

CS301 Processing Elements

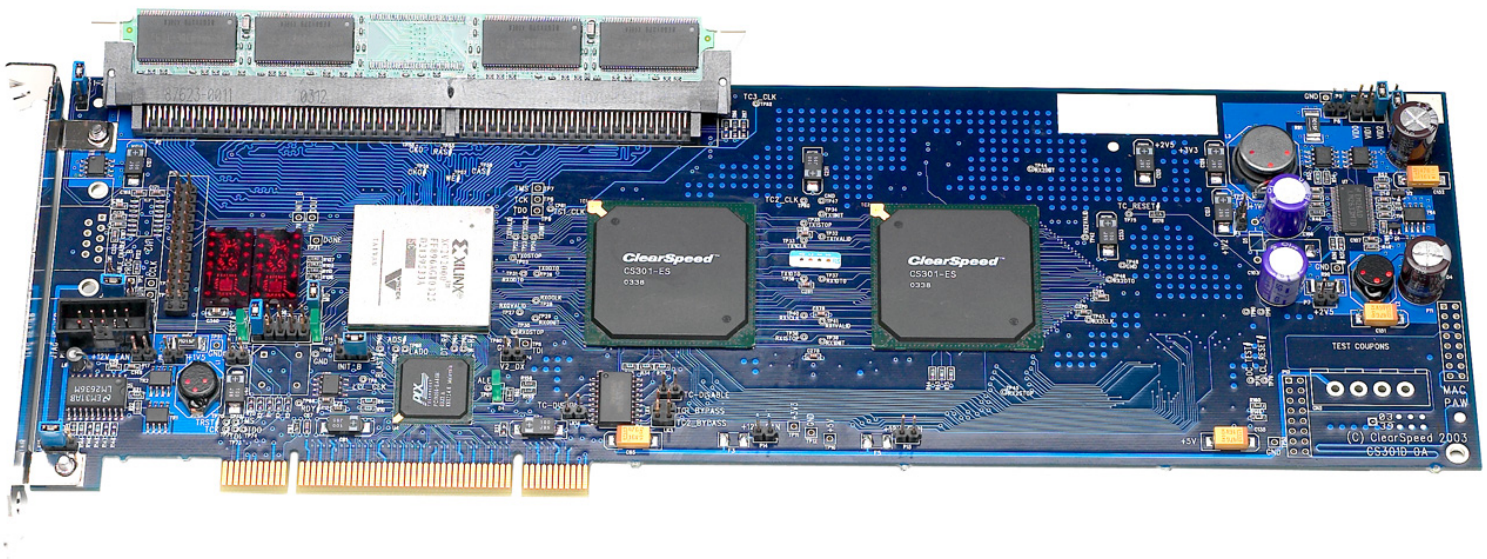


Each PE is a VLIW processor:

- Multiple execution units
  - Floating point adder
  - Floating point multiplier

  } 32-bit IEEE 754
  - Divide/square root unit
  - Fixed point MAC 8x8->16+48
  - Integer ALU with shifter
  - Load/store
- High-bandwidth, 5-port register file (3r, 2w)
- Closely coupled 4KB SRAM for data
- High bandwidth per PE load/store (PIO)
- Per PE address generator
  - Complete pointer model, including parallel pointer chasing and vectors of addresses

CS301-based development board



- 2 chip board – 50 GFLOPS peak @ 10W total
- 200K FFTs/s (1K complex single precision IEEE754)
- Up to 1GB DRAM for local processing
- Shipping since 1Q04
- Single slot width full-size PCI card

What Applications Can Be Accelerated?

Any applications with significant *data parallelism*:

- Fine-grained – vector operations
- Medium-grained – unrolled independent loops
- Coarse-grained – multiple simultaneous data channels/sets

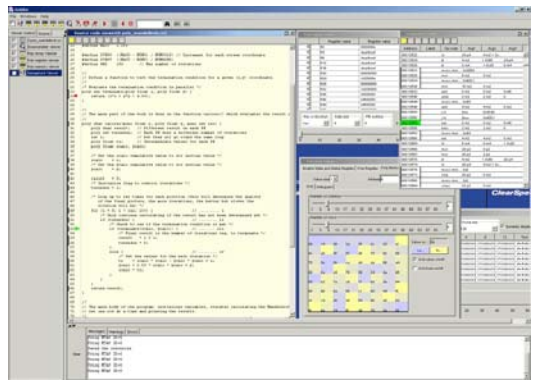
Example applications and libraries include:

- Math libraries – BLAS, LAPACK (→ Matlab, Maple, …)
- Chemistry – GROMACS, CHARMM, BLAST, DLPOLY, …
- Computational finance – Monte Carlo, genetic algorithms
- Intelligent systems – artificial neural networks
- Signal processing – FFT (1D, 2D, 3D), FIR
- Simulation – CFD, N-body, Finite Element
- Image processing – filtering, image recognition, DCTs

Software Development Environment

Software Development Kit (SDK)

- C compiler, assembler, libraries, visual debugger, etc.
- CS301-based development boards
- Available for Linux and Windows

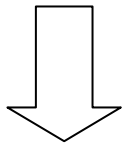


Applications and libraries under development

- Math – L3 BLAS, LAPACK
- DSP – FFTs (1D, 2D, 3D)
- Bio/Chemistry – GROMACS, DLPOLY, DockIt
- Financial – random number generation, Monte Carlo

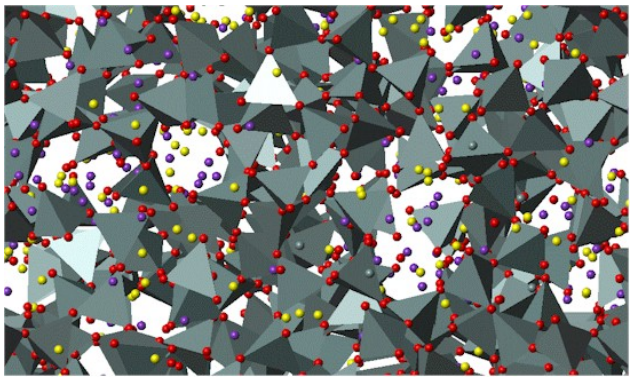
Porting Code


```
void daxpy(double *c, double *a, double alpha, uint N) {
  uint i;
  for (i=0; i<N; i++)
    c[i] = c[i] + a[i]*alpha;
}
```



```
void daxpy(double *c, double *a, double alpha, uint N) {
  uint i;
  poly double cp, ap;
  for (i=0; i<N; i+=num_pes) {
    memcpym2p(&cp, &c[i+pe_num], sizeof(double));
    memcpym2p(&ap, &a[i+pe_num], sizeof(double));
    cp = cp + ap*alpha;
    memcpyp2m(&c[i+pe_num], &cp, sizeof(double))
}
```

ClearSpeed-AWE Joint Investigation



- Chemistry codes: DLPOLY (Molecular Dynamics)
  - Owned by UK Daresbury Lab, heavily used at AWE
  - Widely used in academia and industry
  - 91% of CPU in 5 relatively small routines
  - One of these (forces) calls the other 4 to compute forces on all atoms
  - “forces” called once per time step
  - Data needing to be returned by “forces” from CS to host relatively small
  - Calculation for each atom is independent

- Matrix Multiply Benchmark (SGEMM)
  - CS301 single precision code started at ~20% efficiency
  - AWE helped CS restructure code to give 12 GFLOPS – 47%
  - Performance verified by AWE on CS301 hardware
  - Next-generation processor from ClearSpeed significantly increases this performance – “Avebury”