The \textit{\textsc{anton} 3} ASIC: a Fire-Breathing Monster for Molecular Dynamics Simulations

Hot Chips 33
24 August 2021
The Anton 3 hardware team


† Work conducted while at D. E. Shaw Research; author’s affiliation has subsequently changed.
Molecular dynamics (MD) simulation

• Understand biomolecular systems through their motions
• Numerical integration of Newton’s laws of motion
  – Model atoms as point masses
  – Compute forces on every atom based on current positions
  – Update atom velocities and positions in discrete time steps of a few femtoseconds
• Force computation described by a model: the force field
Biomolecular force fields

\[ F = \text{Bonded} + \text{Near electrostatic, Van der Waals} + \text{Distant electrostatic} \]
Meet ANTON
ANTON 2: The defending champion

High-Throughput Interaction Subsystem (HTIS)

Flex Tile

- Dispatch Unit
- 256KB SRAM
- Network Interface
- Geometry Core
- Geometry Core
- Geometry Core
- Geometry Core
- 16KB D$ 64KB I$ 16KB D$ 64KB I$ 16KB D$ 64KB I$ 16KB D$ 64KB I$

Mini-flex tile

- 64 KB SRAM
- Network Interface
- 128 queues
- output
- Dispatch Unit
- Geometry Core

Pairwise Point Interaction Module (PPIM)

- PPIM 0
- PPIM 1
- PPIM 17
- PPIM 18
- PPIM 19
- PPIM 20
- PPIM 36
- PPIM 37

D E Shaw Research
How do we make it better?

• Increase computational throughput
  – Pairwise force computation
  – Programmable cores

• Address exposed bottlenecks
  – Bond computation
  – Communication bandwidth

• Improve utility
  – Maximum simulation capacity
  – Programmability
  – Flexibility

• Manage design complexity
Concentrated compute: The Core Tile

- **Evolutionary changes**
  1. Support additional functional forms
  2. Increase memory capacity
  3. Tune instruction set for MD application
  4. Increase code density

- **Revolutionary changes**
  4. Co-locate compute resources
  5. Specialize bonded force computation
  1. Double effective density of pairwise interaction calculation
  2. Implement fine-grained synchronization within memory and network
Bond calculator

<table>
<thead>
<tr>
<th>Term</th>
<th>Stretch</th>
<th>Angle</th>
<th>Dihedral / Torsion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atoms</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Coordinate ($q$) | $\ell$ | $\theta$ | $\phi$ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential ($V$)</td>
<td>$k(\ell - \ell_0)^2$</td>
<td>$k(\theta - \theta_0)^2$</td>
<td>$\sum_{n\leq 6} k_n \cos (n\phi - \phi_0)$</td>
</tr>
</tbody>
</table>

Positions: $r_1, \ldots, r_N$

Parameters: $(k, \ell_0, \ldots)$

Forces: $F_1, \ldots, F_N$
Near versus far

**Volume within radius** $r$

- $N \propto r^3$

**Electrostatic force at distance** $r$

- $F \propto \frac{1}{r^2}$

---

BIG

small
Efficient communication: The Edge Tile

- **Evolutionary changes**
  ① Increase SERDES data rate
  ② Reduce hop latency

- **Revolutionary changes**
  ③ Separate edge network
  ④ MD-specific compression
  ⑤ Novel interaction method
Laying tiles
Physical design

- Channel-less, abutted layout
- Few unique blocks
- Global, low-skew clock mesh
- Engineered global routing
- Column-level redundancy
- Robust power delivery
## The evolution of ANTON

<table>
<thead>
<tr>
<th></th>
<th>ANTON</th>
<th>ANTON 2</th>
<th>ANTON 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tape-out</td>
<td>2007</td>
<td>2012</td>
<td>2020</td>
</tr>
<tr>
<td>CPU cores</td>
<td>8+4+1</td>
<td>66</td>
<td>528*</td>
</tr>
<tr>
<td>PPIMs</td>
<td>32</td>
<td>76</td>
<td>528*</td>
</tr>
<tr>
<td>Flex SRAM</td>
<td>0.125 MiB</td>
<td>4 MiB</td>
<td>66 MiB*</td>
</tr>
<tr>
<td>Atoms / node</td>
<td>460</td>
<td>8,000</td>
<td>110,000*</td>
</tr>
<tr>
<td>Clock frequency</td>
<td>0.485/0.970 GHz</td>
<td>1.65 GHz</td>
<td>2.8+ GHz</td>
</tr>
<tr>
<td>Channel bandwidth</td>
<td>0.607 Tbps</td>
<td>2.7 Tbps</td>
<td>5.6+ Tbps</td>
</tr>
<tr>
<td>Process node</td>
<td>90 nm</td>
<td>40 nm</td>
<td>7 nm</td>
</tr>
<tr>
<td>Transistors</td>
<td>0.2 G</td>
<td>2.0 G</td>
<td>31.8 G</td>
</tr>
<tr>
<td>Die size</td>
<td>299 mm²</td>
<td>410 mm²</td>
<td>451 mm²</td>
</tr>
<tr>
<td>Power</td>
<td>30 W</td>
<td>190 W</td>
<td>360 W</td>
</tr>
</tbody>
</table>

* 22/24 columns
29 September 2020: chips arrive
MD running (water) < 9 h later

30 September 2020: 1\textsuperscript{st} protein run
Faster @ 250 MHz than Anton 2

31 October 2020: Multi-node
Node board

- 48 VDC, torus links, Ethernet, USB
- 500+ W custom voltage regulator
- Node control complex
- ASIC interface FPGA
- Data processor (64-bit Linux)
- Control processor (no OS)
Scale up

8×8 nodes

2×64 nodes

512 nodes
Network

Complete 512-node, 3D torus

X-dimension connected entirely in backplane

Z-dimension cabled (blue and yellow)

Y-dimension split across cables (green) and backplane
Taming (cooling) the beast

\[ T_J < 65 \, ^\circ\text{C} \, @ \, 500 \, \text{W} \]
MD performance

![Graph showing MD performance with log scale and >100× faster annotation.](image-url)
Acknowledgements

- System software group for machine bring-up
- Embedded software group for creating and tuning the application
  - Ken Mackenzie for performance results and figures
- Systems group for support and infrastructure
  - And lots of photos!
- Chemistry team for putting Anton to good use
  - Kevin Yuh for MD simulation videos
Performance references

• **GPU performance results**
    [https://deshawresearch.com/publications.html](https://deshawresearch.com/publications.html).
  - “Gromacs/NAMD Multi-GPU Scaling”, unpublished internal benchmarking.

• **Supercomputer performance results**

• **Anton performance results (original publications; improved performance used in comparisons)**