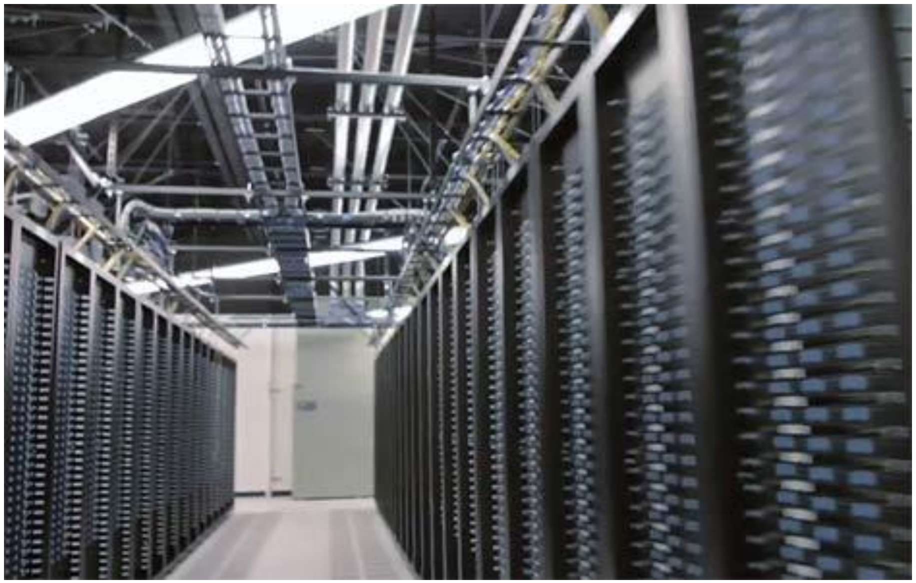
Energy-efficient Cluster Computing with FAWN: Workloads and Implications

Vijay Vasudevan, David Andersen, Michael Kaminsky*, Lawrence Tan, **Jason Franklin**, Iulian Moraru

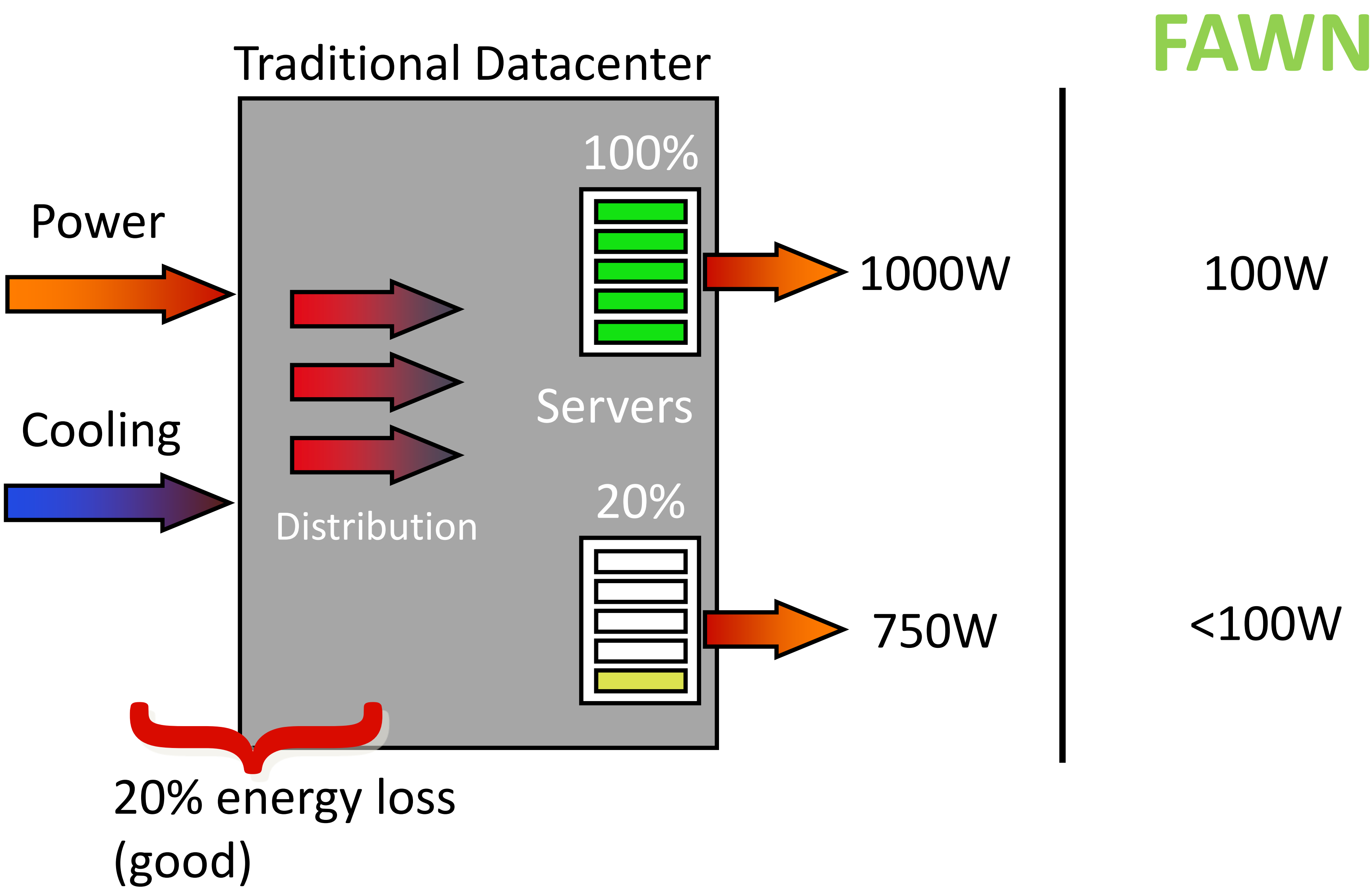
Carnegie Mellon University, *Intel Labs Pittsburgh

Energy in Data Centers

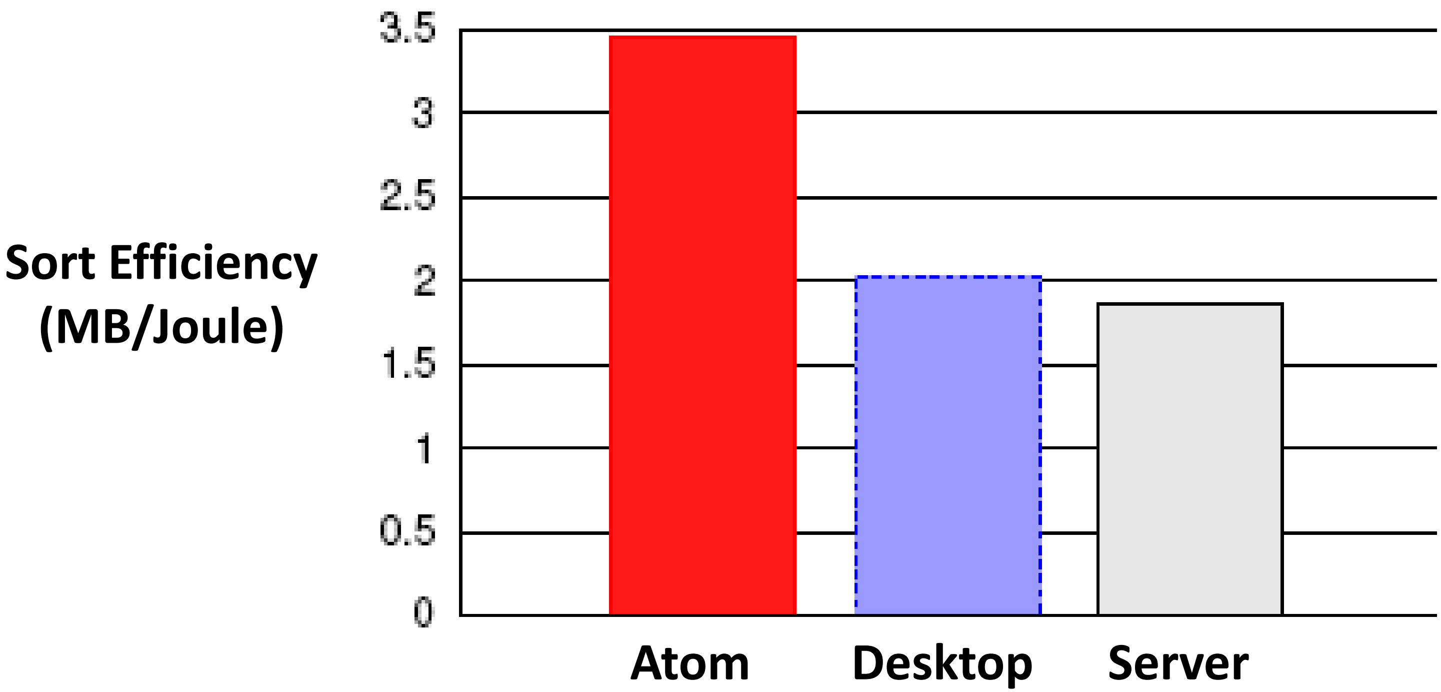


- US data centers now consume 2% of total US power
- Energy has become important metric of system performance
- Can we make data intensive computing more energy efficient?
  - Metric: Work per Joule

Goal: reduce peak power



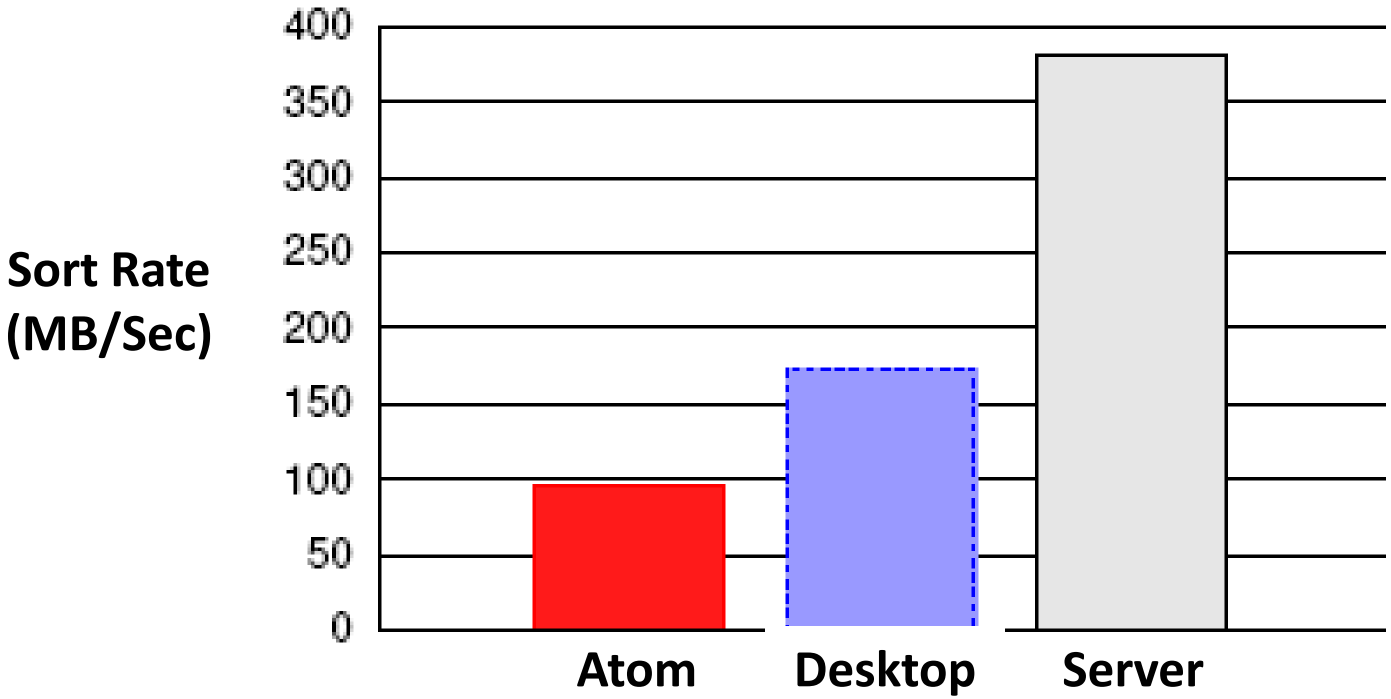
Wimpy Nodes are Energy Efficient



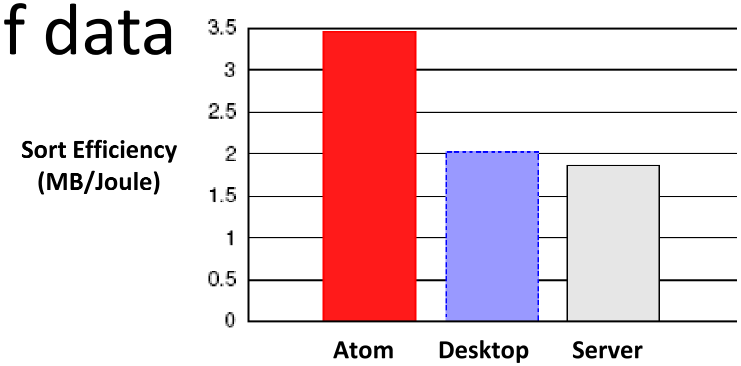
Sorting 10GB of data

Wimpy Nodes are Energy Efficient

...but slow

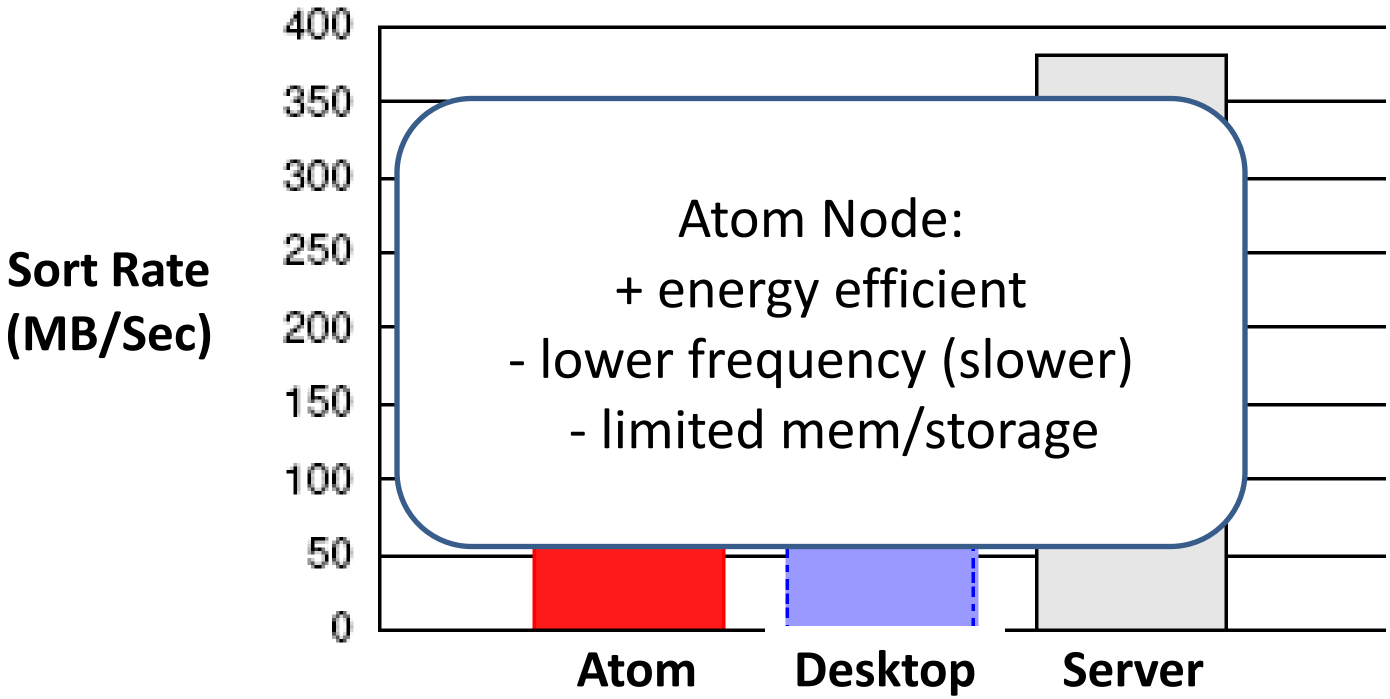


Sorting 10GB of data



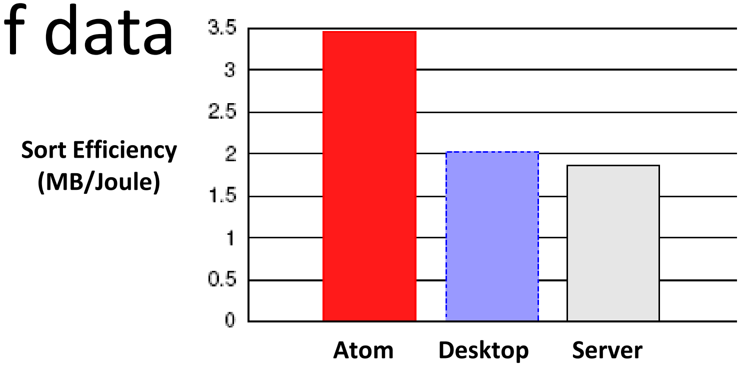
Wimpy Nodes are Energy Efficient

...but slow



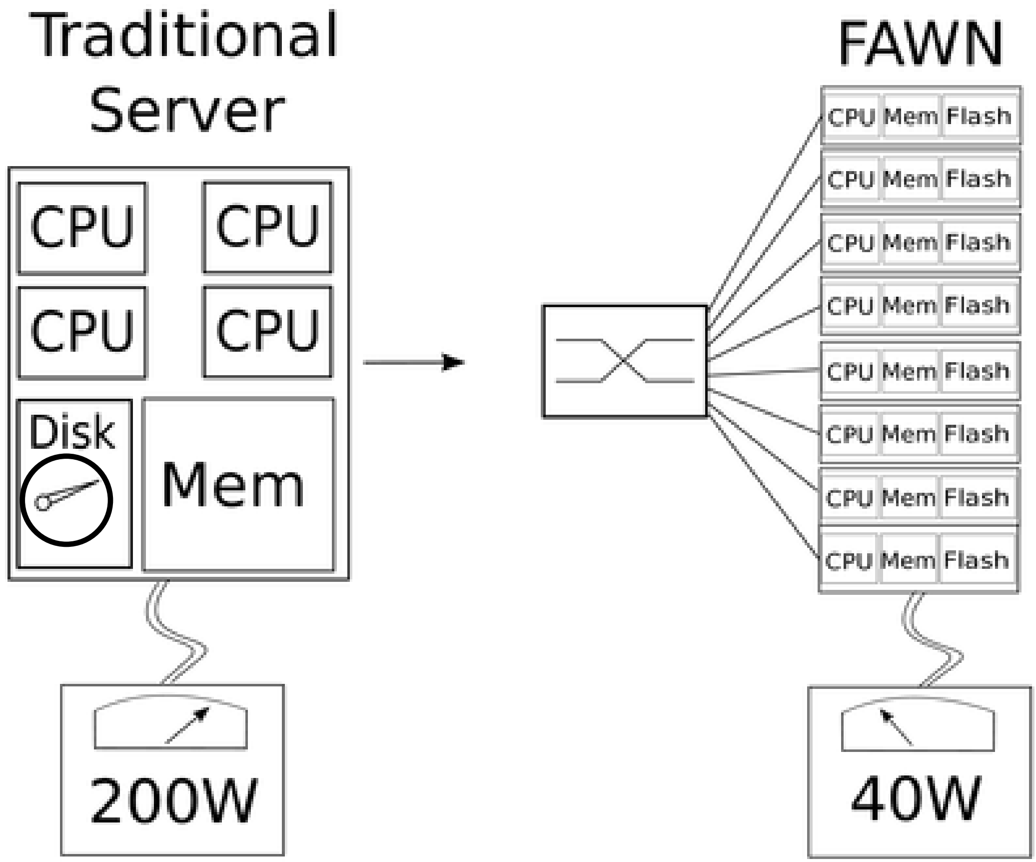
Atom Node:
+ energy efficient
- lower frequency (slower)
- limited mem/storage

Sorting 10GB of data



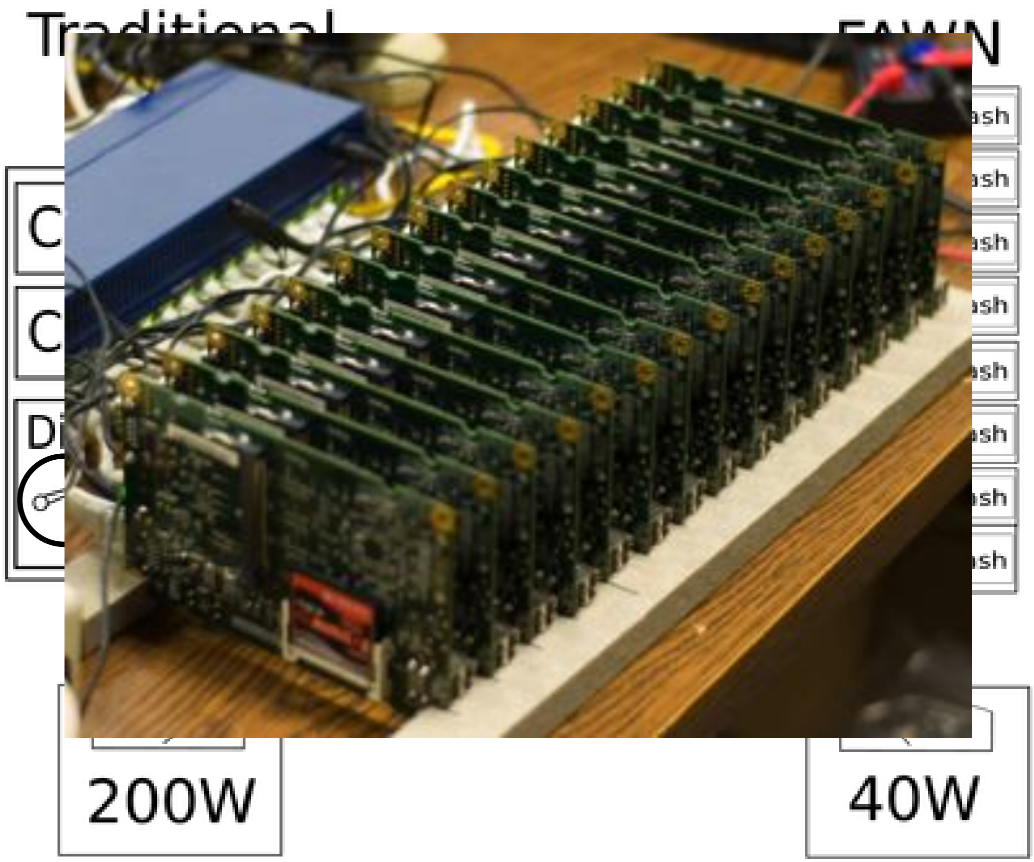
FAWN - Fast Array of Wimpy Nodes

Leveraging parallelism and scale out to build eEfficient Clusters



FAWN - Fast Array of Wimpy Nodes

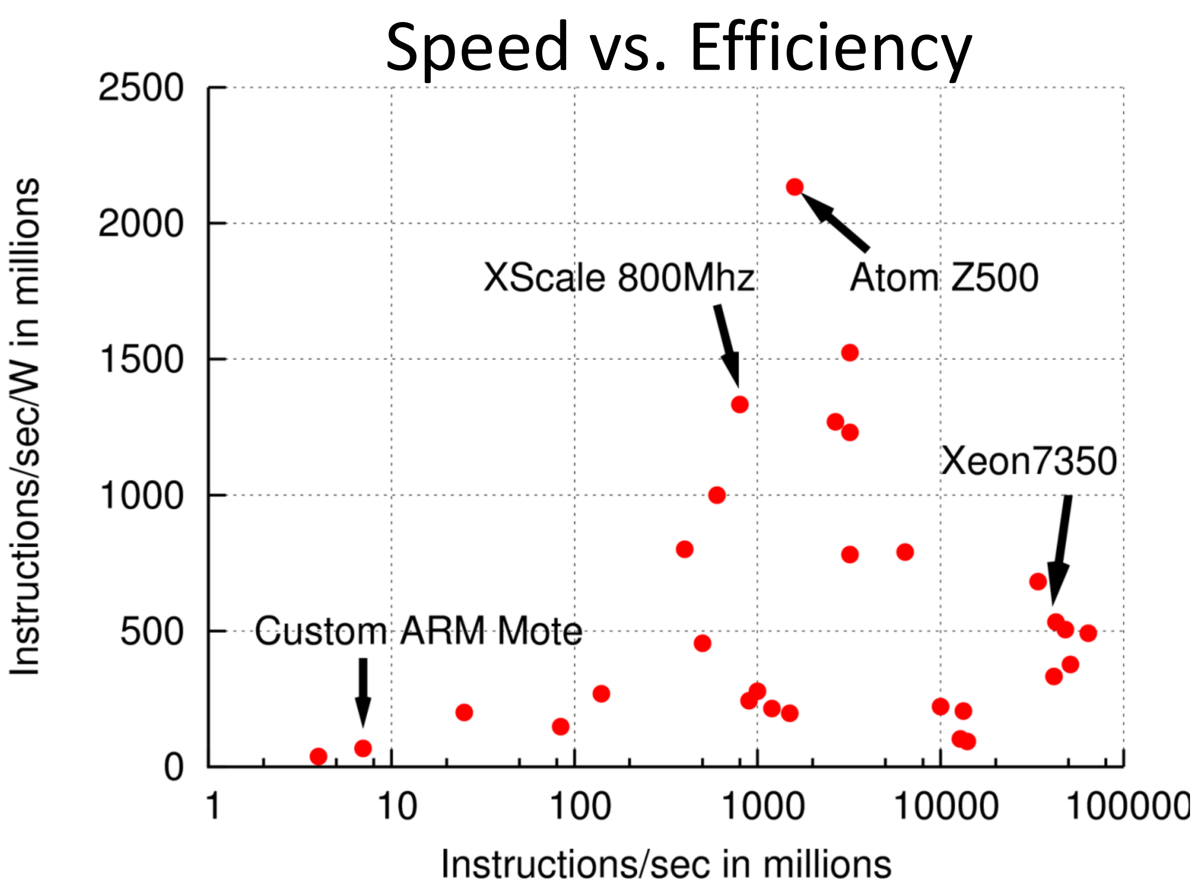
Leveraging parallelism and scale out to build eEfficient Clusters



FAWN in the Data Center

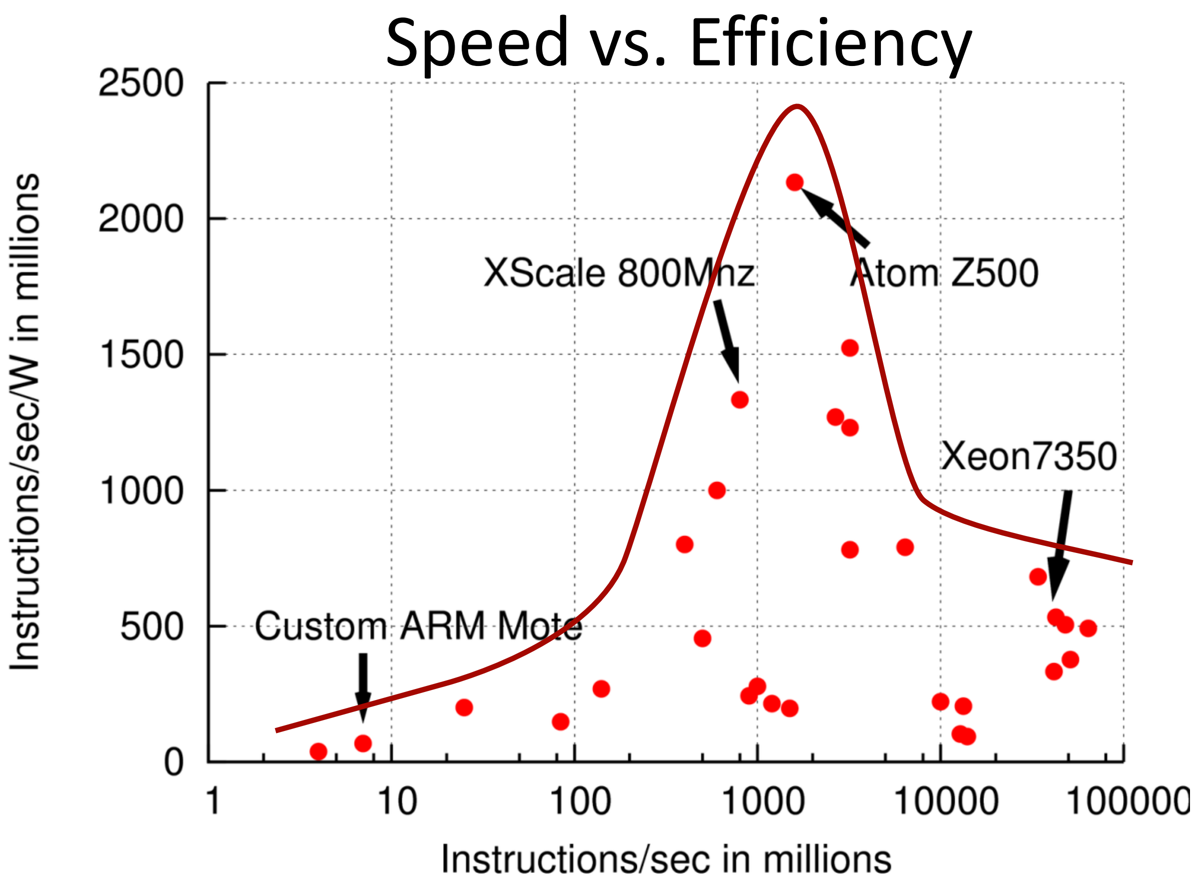
- **Why** is FAWN more energy-efficient?
- **When** is FAWN more energy-efficient?
- **What** are the future design implications?

CPU Power Scaling and System Efficiency



* Efficiency numbers include 0.1W power overhead

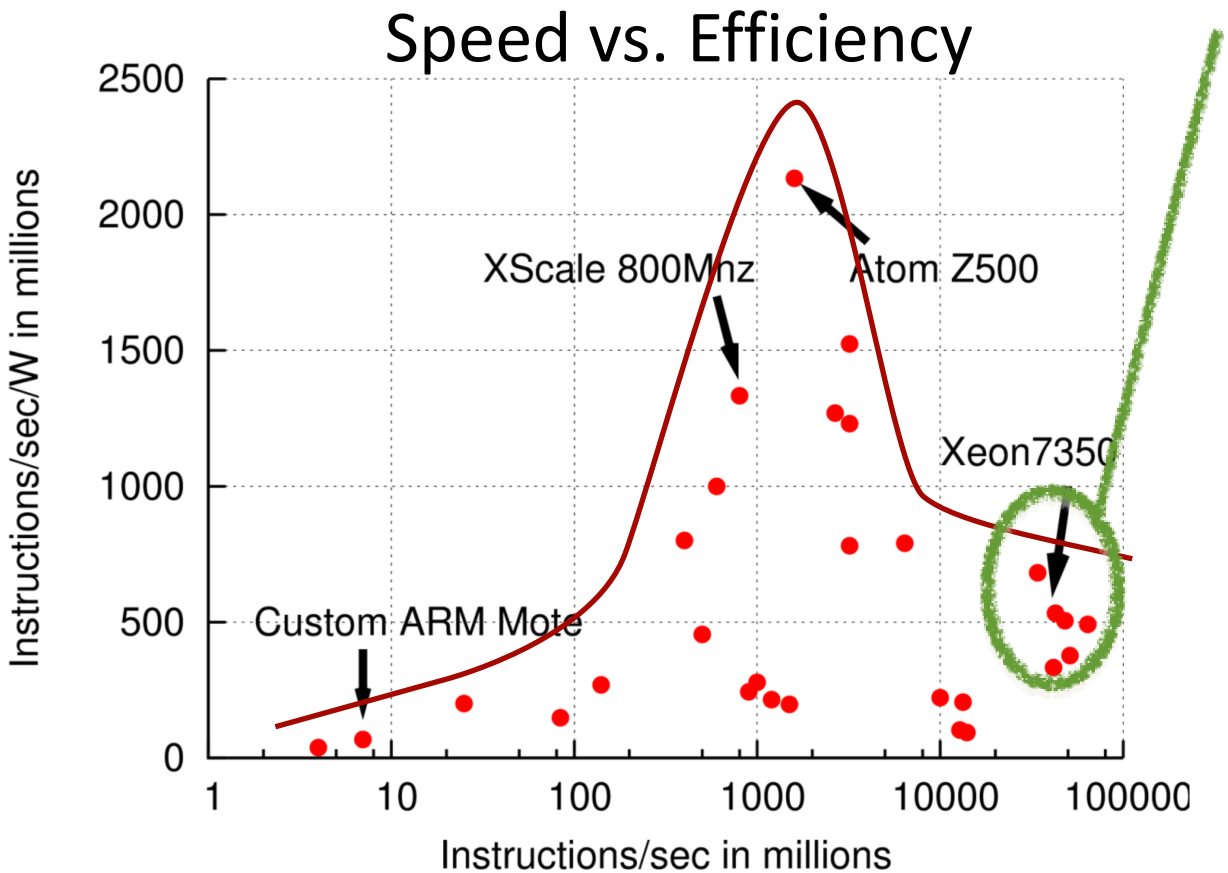
CPU Power Scaling and System Efficiency



* Efficiency numbers include 0.1W power overhead

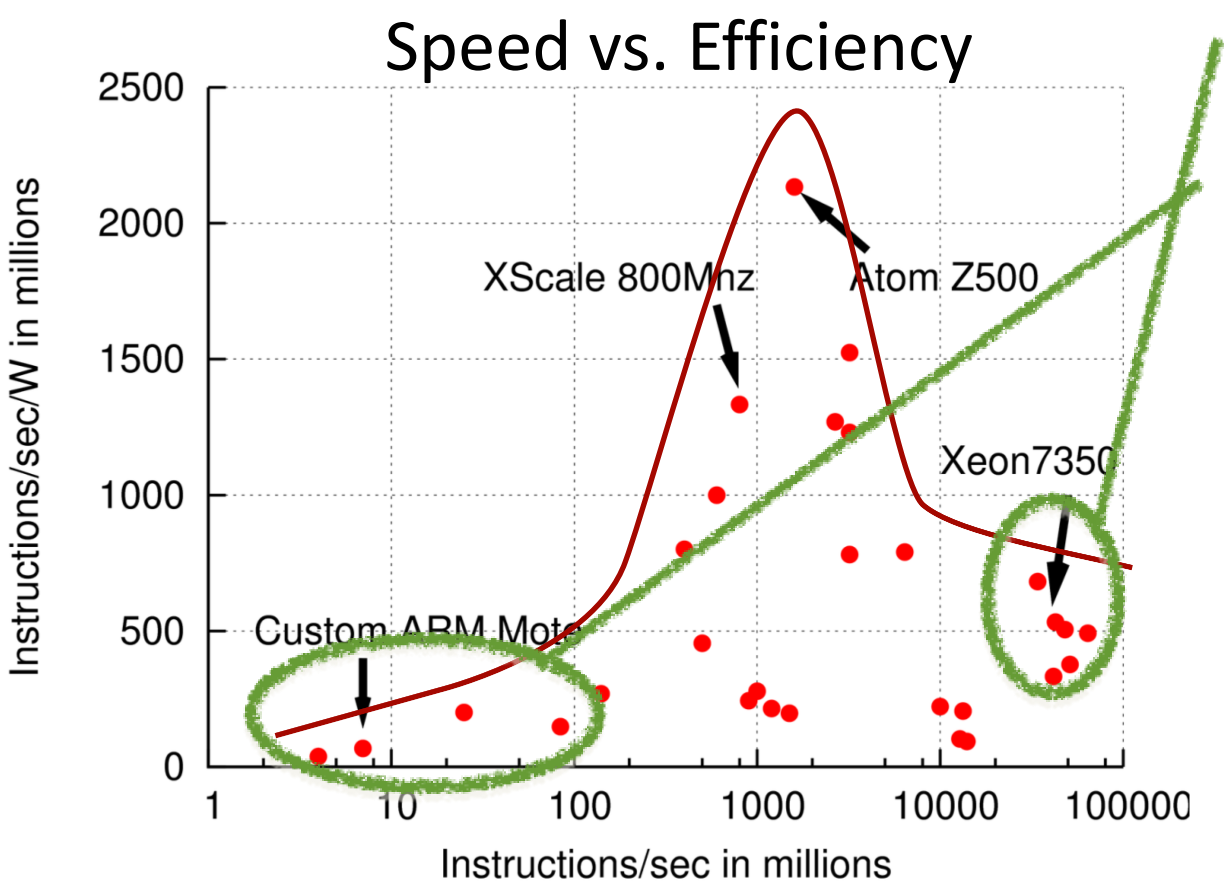
CPU Power Scaling and System Efficiency

Fastest processors exhibit superlinear power usage



* Efficiency numbers include 0.1W power overhead

CPU Power Scaling and System Efficiency

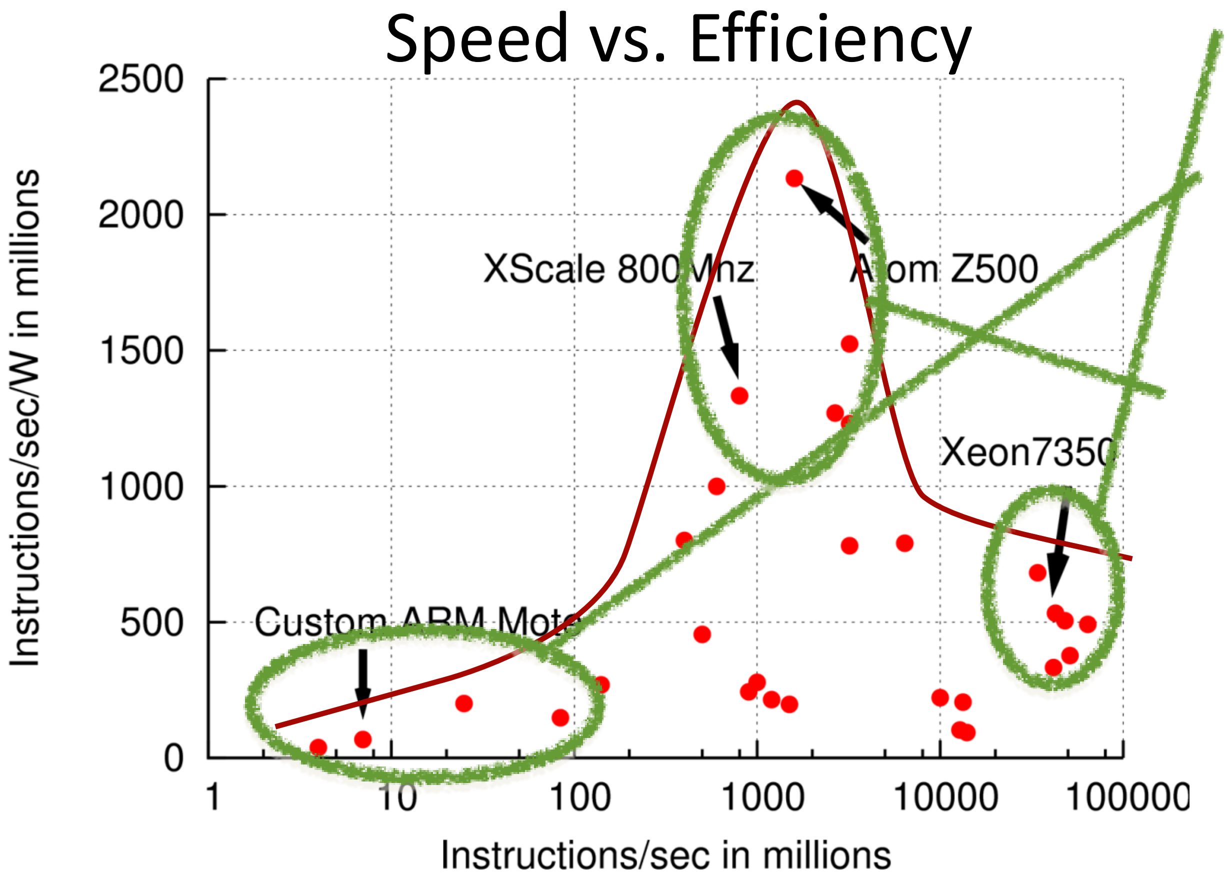


Fastest processors exhibit superlinear power usage

Fixed power costs can dominate efficiency for slow processors

* Efficiency numbers include 0.1W power overhead

CPU Power Scaling and System Efficiency



Fastest processors exhibit superlinear power usage

Fixed power costs can dominate efficiency for slow processors

FAWN targets sweet spot in system efficiency when including fixed costs

* Efficiency numbers include 0.1W power overhead

FAWN in the Data Center

- **Why** is FAWN more energy-efficient?

- **When** is FAWN more energy-efficient?

When is FAWN more efficient?

Core i7-based Desktop (Stripped down)

- Single 2.8GHz quad-core Core i7 860
- 2GB of DRAM
- 40W – 140W (idle – peak)

Modern Wimpy FAWN Node

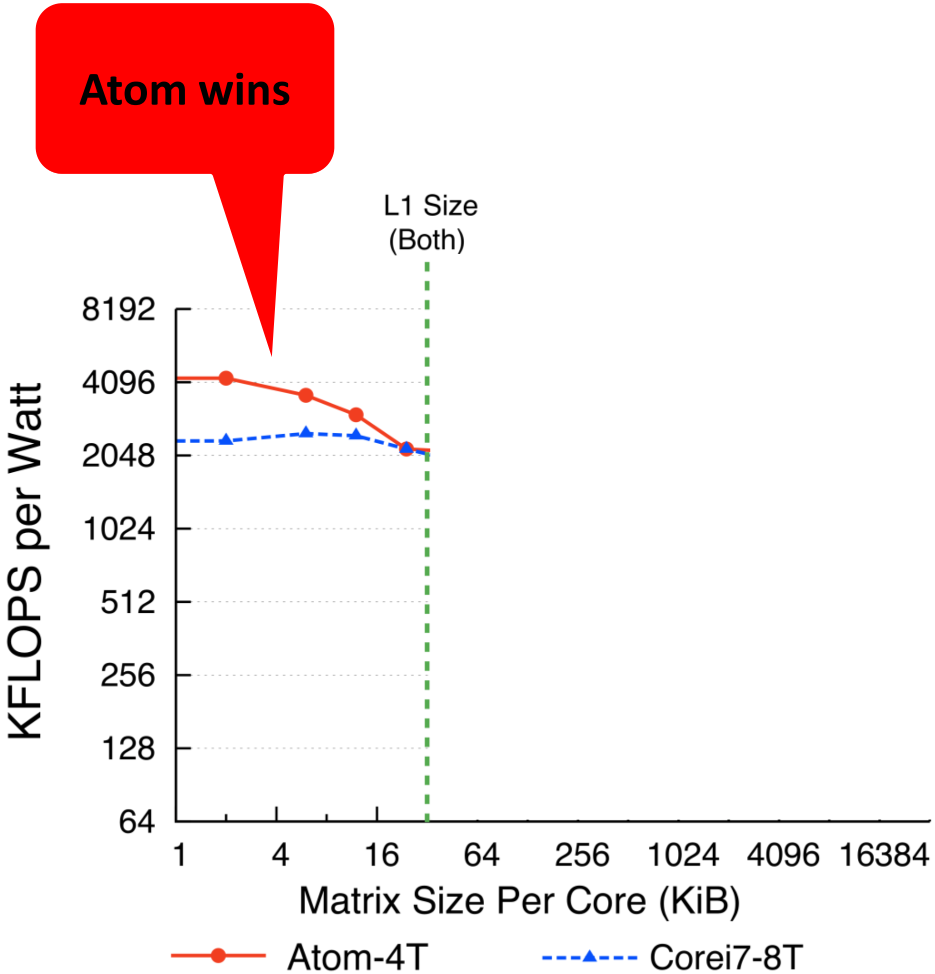
- Prototype Intel "Pineview" Atom
- Two 1.8GHz cores
- 2GB of DRAM
- 18W -- 29W (idle – peak)

Data-intensive computing workloads

1. I/O-bound – Seek or scan
2. Memory/CPU-bound
3. Latency-sensitive, but non parallelizable
4. Large, memory-hungry

FAWN's sweet spot

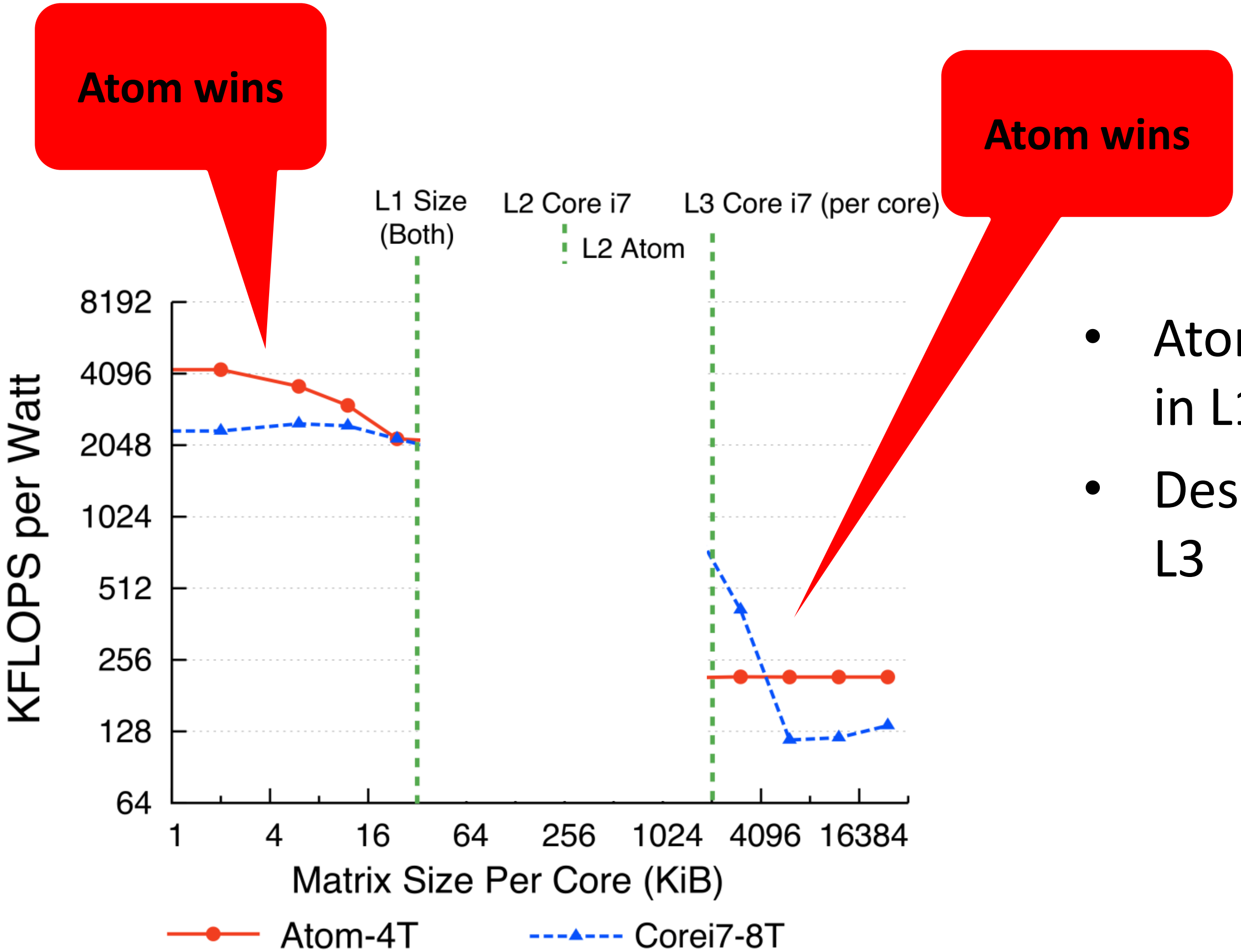
Memory-bound Workloads



Efficiency vs. Matrix Size

- Atom 2x as efficient when in L1 and DRAM
- Desktop Corei7 has 8MB L3

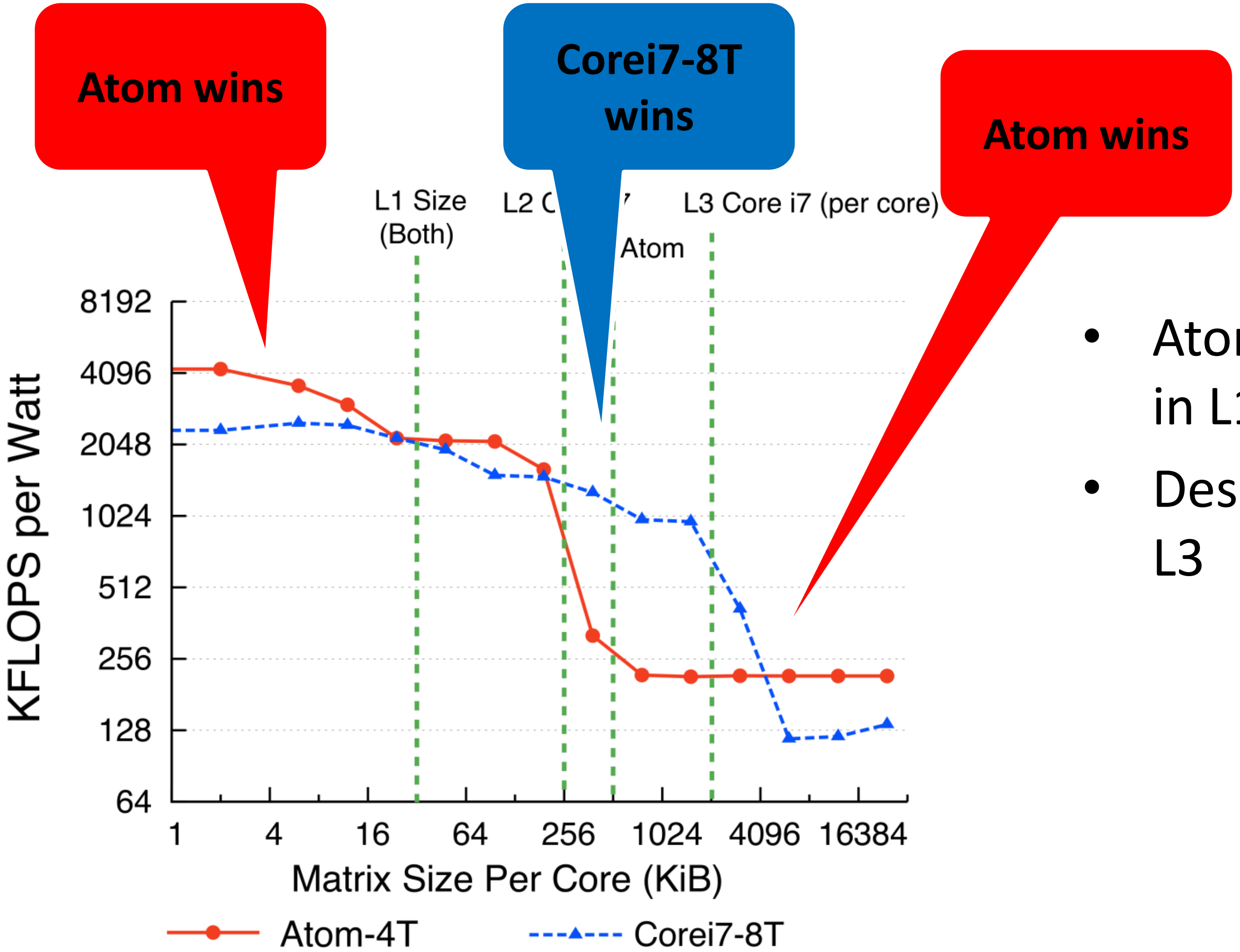
Memory-bound Workloads



Efficiency vs. Matrix Size

- Atom 2x as efficient when in L1 and DRAM
- Desktop Corei7 has 8MB L3

Memory-bound Workloads



Efficiency vs. Matrix Size

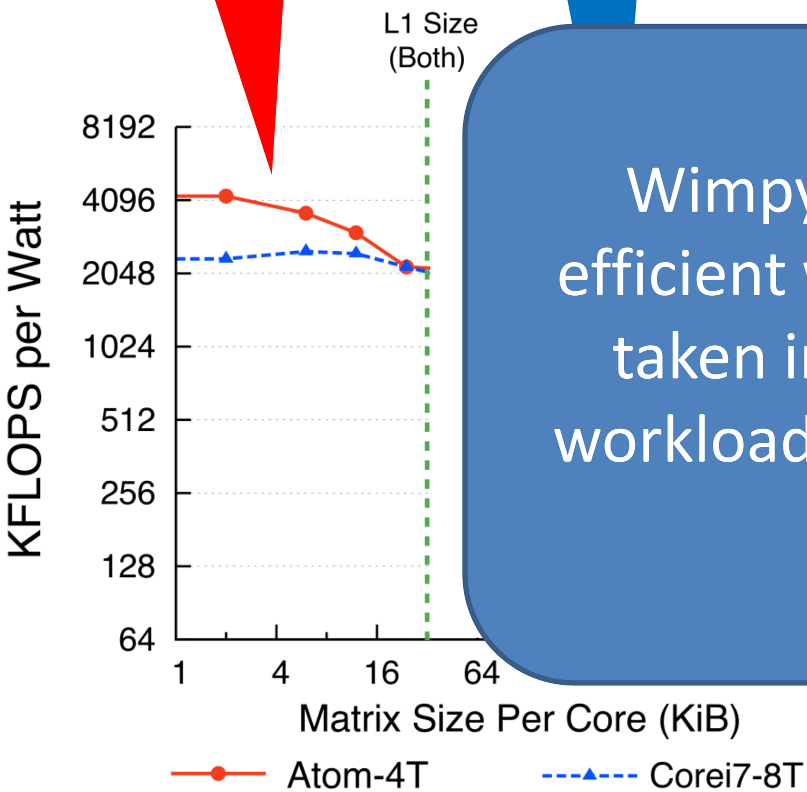
- Atom 2x as efficient when in L1 and DRAM
- Desktop Corei7 has 8MB L3

Memory-bound Workloads

Atom wins

Corei7-8T wins

Atom wins



Efficiency vs. Matrix Size

cient when

M

has 8MB

Wimpy nodes can be more efficient when cache effects are taken into account, for your workloads it may require tuning of algorithms

CPU-bound Workload

- Crypto: SHA1/RSA
- Optimization matters!
  - Unopt. C: Atom wins
  - Opt. Asm:
    - Old: Corei7 wins!
    - New: Atom wins!

	Old-SHA1 (MB/J)	New-SHA1 (MB/J)	RSA-Sign (Sign/J)
Atom	3.85 →	5.6	56
i7	4.8 =	4.8	71

CPU-bound operations can be more energy efficient on low-power processors

- Crypto:
- Optimization matters!
  - Unopt. C: Atom wins
  - Opt. Asm:
    - Old: Corei7 wins!
    - New: Atom wins!

		New-SHA1 (MB/J)	RSA-Sign (Sign/J)
Atom	3.85 →	5.6	56
i7	4.8 =	4.8	71

CPU-bound operations can be more energy efficient on low-power processors

- Crypto:
- Optimization matters!
  - Unop
  - Opt.
    - Old
    - Ne

However, code may need to be hand optimized

		New-SHA1 (MB/J)	RSA-Sign (Sign/J)
		5.6	56
		4.8	71

Potential Hurdles

- Memory-hungry workloads
  - Performance depends on locality at many scales
    - E.g., prior cache results, on or off chip/machine
  - Some success w algo. changes e.g., virus scanning
- Latency-sensitive, non-parallelizable
  - E.g., Bing search, strict latency bound on processing time
    - W.o. software changes, found atom too slow

FAWN in the Data Center

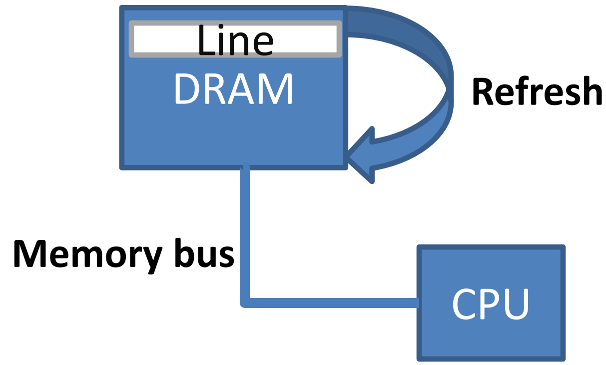
- **Why** is FAWN more energy-efficient?

- **When** is FAWN more energy-efficient?

- **What** are the future design implications?
  - With efficient CPUs, memory power becomes critical

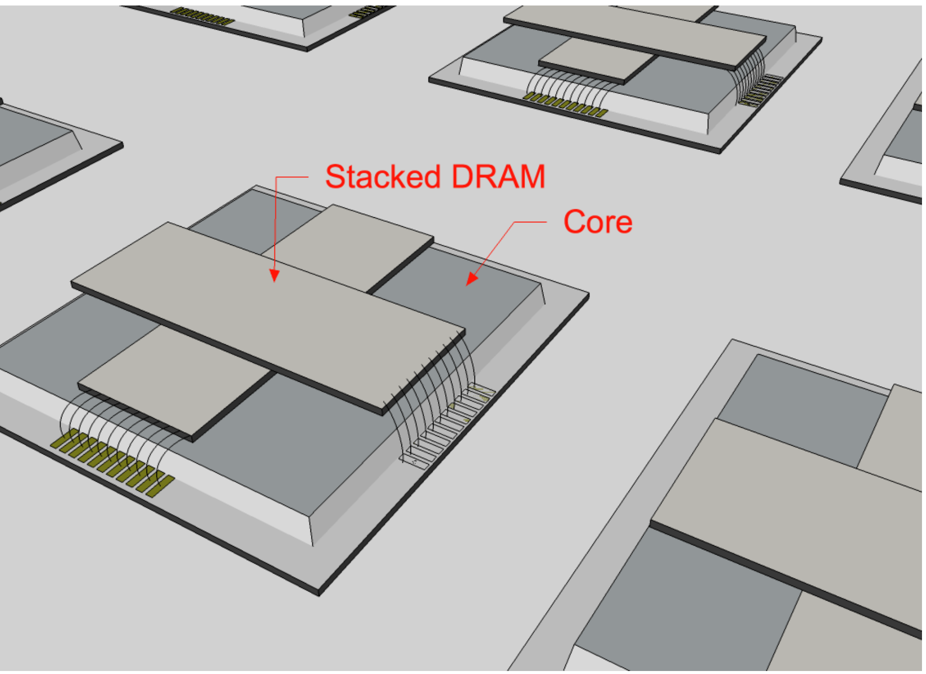
Memory power also important

- Today's high speed systems: mem. ~= 30% of power
- DRAM power draw
  - Storage:
    - Idle/refresh
  - Communication:
    - Precharge and read
    - Memory bus (~40% ?)



- CPU to mem distance greatly affects power
  - Point-to-point topology more efficient than bus, reduces trace length
    - +Lower latency, + Higher bandwidth, + Lower power cons
    - - Limited memory per core
  - Why not stack CPU and memory?

Preview of the Future



FAWN RoadMap

- Nodes with single CPU chip with many low-frequency cores
- Less memory, stacked with shared interconnect
- Industry and academia beginning to explore
  - iPad, EPFL Arm+DRAM

To conclude, FAWN arch. more efficient, but…

- Up to 10x increase in processor count
- Tight per-node memory constraints
- Algorithms may need to be changed
- Research needed on…
  - **Metrics**: Ops per Joule?
    - Atoms increase workload variability & latency
    - Incorporate quality of service metrics?
  - **Models:** Will your workload work well on FAWN?

To con … icient,

Questions?

www.cs.cmu.edu/~fawnproj

- Up to 10x increase in processor count
- Tight per-node memory constraints
- Algorithms may need to be changed
- Research needed on…
  - **Metrics**: Ops per Joule?
    - Atoms increase workload variability & latency
    - Incorporate quality of service metrics?
  - **Models:** Will your workload work well on FAWN?

Related Work

- System Architectures
  - JouleSort: SATA disk-based system w. low-power CPUs
  - Low-power processors for datacenter workloads
    - Gordon: Focus on FTL, simulations
    - CEMS, AmdahlBlades, Microblades, Marlowe, Bluegene
  - IRAM: Tackling memory wall, thematically similar approach
- Sleeping, complementary approach
  - Hibernator, Ganesh et al., Pergamum