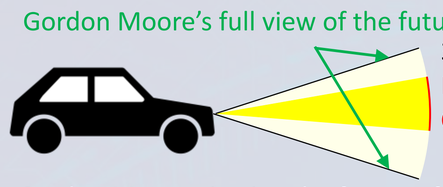
SAND2017-8646C

Superstrider and the 3rd Path Beyond Moore's Law

Sandia National Laboratories and Georgia Institute of Technology
Erik P. DeBenedictis and Jeanine Cook, Center for Computing Research, Albuquerque, NM
Sriseshan Srikanth and Thomas M. Conte, School of Computer Science, Atlanta, GA

Problem

1. Nearly all "Beyond Moore's law" efforts focus on quantum computing and neural networks
2. What is the equivalent for supercomputing, particularly if Moore's law doesn't end?
   - Moore's 1965 article looked ahead and saw the future of computing broadly
   - Over time industry narrowed its meaning to just microprocessor throughput and DRAM size.
   - Can the narrowing be reversed?
3. Can we adapt industry's 3D advances to supercomputing – like parallel computing in 1980s+?

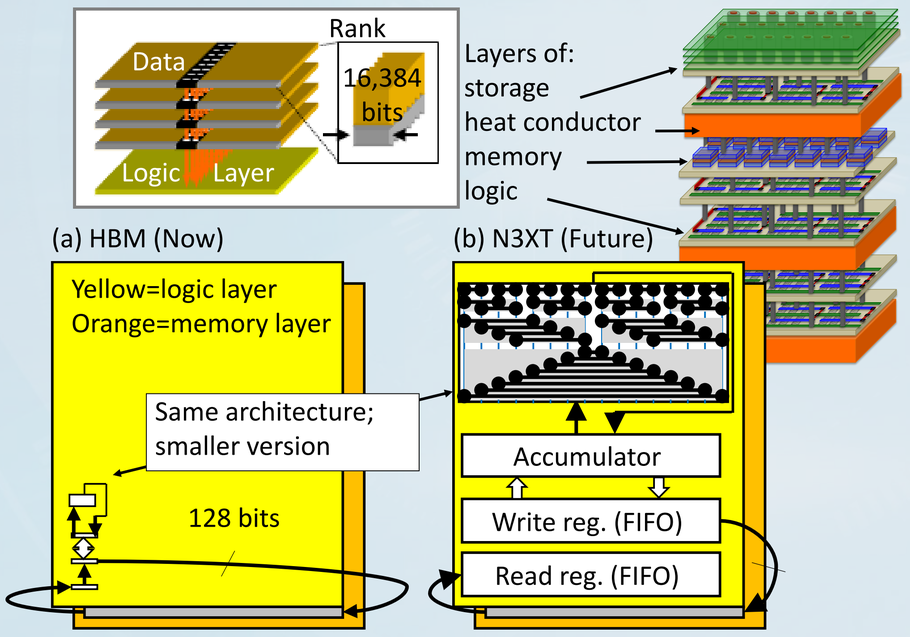


a. Gordon Moore viewing the future of the computing, but only able to see so much

Approach

1. We need a scaling theory for new 3D chip technology.
2. Government took responsibility for parallel algorithms and software; now for 3D.
3. Example:
   - Sorting network example like Batcher's bitonic $O(\log^2 n)$ well known in chip design
   - Best sort on microprocessor is Quicksort $O(n \log n)$, but it is much slower
   - A 3D chip could run a Batcher bitonic sort due to greater memory bandwidth.

(a) and (b) show a current product (HBM) and a future vision (N3XT). Industry is roadmapping many steps between the two, differing by the tightness of *z*-direction bandwidth.
(c) and (d) show the Superstrider architecture that scales by degree of 3D sophistication. (e) shows the scale up sequence.



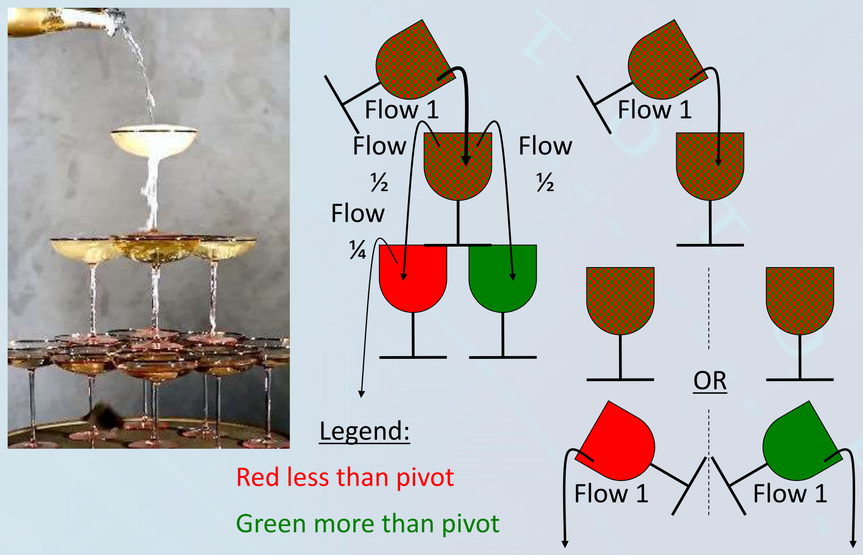
(c) HBM design (now) (d) N3XT design (future)

HBM Row: 16,384 bits = Interface width ×	Cycles	Possible network size
128 bits (HBM)	128	8 = 4 × 2
256	64	24 = 8 × 3
512	32	64 = 16 × 4
1,024	16	160 = 32 × 5
16,384 (N3XT)	1	4,608 = 512 × 9

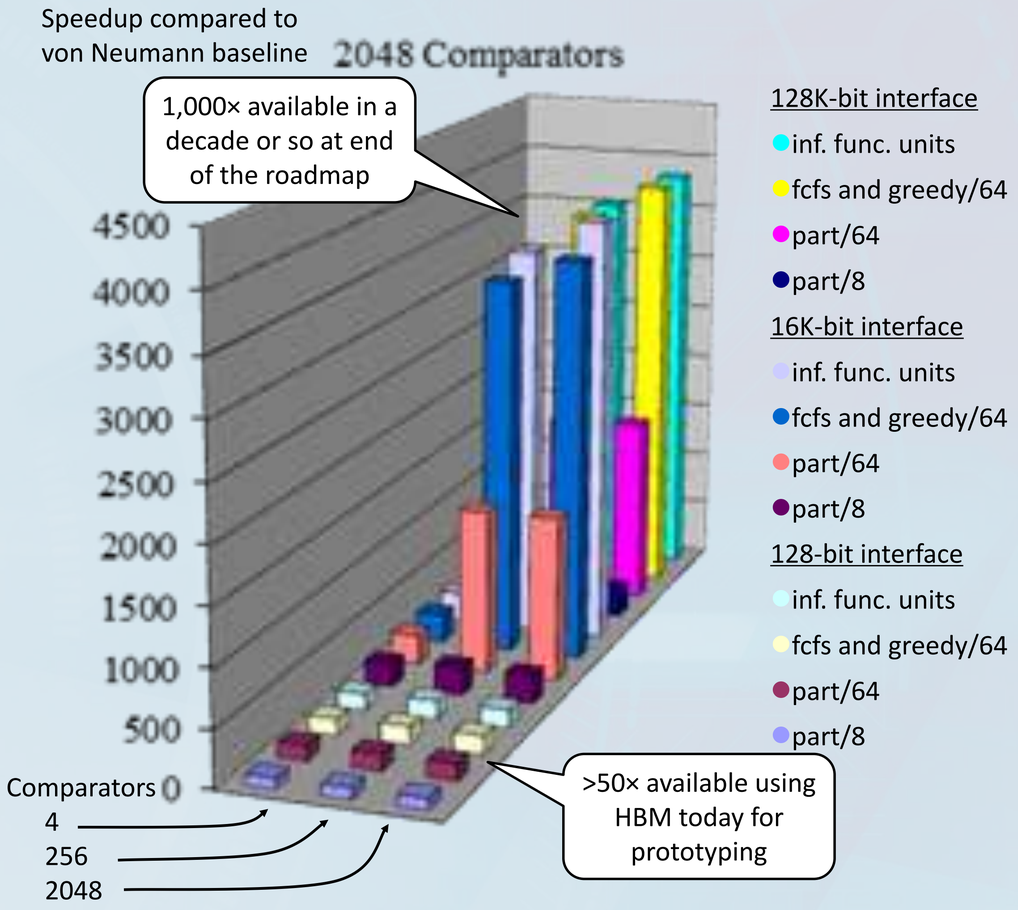
(e) Scaling scenario

Results

- Superstrider algorithm performs "sparse-to-dense stream conversion"
- Unique tree data structure for "accumulating" sparse matrix results
- Enabled by speed of a sorting network.



(a) Champagne tower (b) Superstrider



Significance

1. Superstrider generalizes to a larger strategy
2. Leverages infrastructure and would not require retraining programmers
3. Can start with announced FPGAs with HBM on package
4. Simulation shows 1,000× performance boost over the long run
   - Enough to be a continuation of Moore's law rather than a "generation"
5. Solution characteristics
   - Leverages industry's technology developments
   - Adapts scale up to 3D chips and proposes development of algorithms
   - Initial Superstrider "sparse to dense stream converter"
6. Funded by LDRD and ASC and with assistance from IEEE Rebooting Computing and IRDS (roadmap)

U.S. DEPARTMENT OF ENERGY
National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under Contract DE-NA0003525.

SAND 2017-xxxP

Sandia National Laboratories