The power of parallel computation

D. J. Bernstein

Thanks to: University of Illinois at Chicago NSF CCR-9983950 Alfred P. Sloan Foundation

How fast is sorting?

Input: array of *n* numbers. Each number in $\{1, 2, ..., n^2\}$, represented in binary.

Output: array of *n* numbers, in increasing order, represented in binary; same multiset as input.

A machine is given the input and computes the output. How much time does it use?

on

is at Chicago 0

undation

How fast is sorting?

Input: array of n numbers. Each number in $\{1, 2, ..., n^2\}$, represented in binary.

Output: array of *n* numbers, in increasing order, represented in binary; same multiset as input.

A machine is given the input and computes the output. How much time does it use?

Summarize scalabi by reporting expor $n^{o(1)}$ means log n $100n^{5/\log\log n} + \sqrt{2}$ $n^{1+o(1)}$ means n of $n \log n$ or $n(7(\log$ (Definition: o(1) r function of n that e.g. $5n = n^{1 + (\log \xi)}$ $(\log 5)/\log n \operatorname{conv}$ At this level of det how fast is the ma

How fast is sorting?

Input: array of *n* numbers. Each number in $\{1, 2, ..., n^2\}$, represented in binary.

Output: array of *n* numbers, in increasing order, represented in binary; same multiset as input.

A machine is given the input and computes the output. How much time does it use?

Summarize scalability by reporting exponent of n. $n^{o(1)}$ means log n or $(\log n)^3$ or $100n^{5/\log \log n} + \sqrt{1/n}$ or ... $n^{1+o(1)}$ means n or 5n or $n \log n$ or $n(7(\log n)^3 + 8)$ or ... (Definition: o(1) means any function of n that converges to 0. e.g. $5n = n^{1+(\log 5)/\log n}$; $(\log 5)/\log n$ converges to 0.) At this level of detail, how fast is the machine?

<u>5</u>?

numbers. 1, 2, . . . , *n*²}, ary.

າ numbers, ,

ary;

nput.

n the input

output.

pes it use?

Summarize scalability by reporting exponent of n. $n^{o(1)}$ means log *n* or $(\log n)^3$ or $100n^{5/\log \log n} + \sqrt{1/n}$ or . . . $n^{1+o(1)}$ means n or 5n or $n \log n$ or $n(7(\log n)^3 + 8)$ or ... (Definition: o(1) means any function of n that converges to 0. e.g. $5n = n^{1+(\log 5)/\log n}$; $(\log 5)/\log n$ converges to 0.) At this level of detail, how fast is the machine?

The answer depen how the machine v Possibility 1: The

- "1-tape Turing ma using selection sor
- Specifically: The r
- a 1-dimensional ar
- containing $n^{1+o(1)}$
- Each cell stores n^{α}
- Input and output a stored in these cel

Summarize scalability by reporting exponent of n.

 $n^{o(1)}$ means log n or $(\log n)^3$ or $100n^{5/\log \log n} + \sqrt{1/n}$ or . . .

 $n^{1+o(1)}$ means n or 5n or $n \log n$ or $n(7(\log n)^3 + 8)$ or ...

(Definition: o(1) means any function of n that converges to 0. e.g. $5n = n^{1 + (\log 5)/\log n}$: $(\log 5)/\log n$ converges to 0.)

At this level of detail, how fast is the machine?

The answer depends on how the machine works. Possibility 1: The machine is a "1-tape Turing machine using selection sort." Specifically: The machine has a 1-dimensional array containing $n^{1+o(1)}$ "cells." Each cell stores $n^{o(1)}$ bits. Input and output are stored in these cells.

lity nent of *n*.

or $(\log n)^3$ or $\sqrt{1/n}$ or . . .

or 5n or $n)^3 + 8)$ or . . .

neans any

converges to 0. $\frac{5}{\log n}$

erges to 0.)

cail,

chine?

The answer depends on how the machine works.

Possibility 1: The machine is a "1-tape Turing machine using selection sort."

Specifically: The machine has a 1-dimensional array containing $n^{1+o(1)}$ "cells." Each cell stores $n^{o(1)}$ bits.

Input and output are stored in these cells.

The machine also "head" moving th Head contains $n^{o(}$ Head can see the its current array p perform arithmetic move to adjacent Selection sort: He looks at each array picks up the larges moves it to the en picks up the secon etc.

The answer depends on how the machine works.

Possibility 1: The machine is a "1-tape Turing machine using selection sort."

Specifically: The machine has a 1-dimensional array containing $n^{1+o(1)}$ "cells." Each cell stores $n^{o(1)}$ bits.

Input and output are stored in these cells.

The machine also has a "head" moving through array. Head contains $n^{o(1)}$ cells. Head can see the cell at its current array position; perform arithmetic etc.; move to adjacent array position. Selection sort: Head looks at each array position, picks up the largest number, moves it to the end of the array, picks up the second largest, etc.

ds on *w*orks.

machine is a achine

t."

machine has

ray

"cells."

 $p^{(1)}$ bits.

are

ls.

The machine also has a "head" moving through array. Head contains $n^{o(1)}$ cells.

Head can see the cell at its current array position; perform arithmetic etc.; move to adjacent array position.

Selection sort: Head looks at each array position, picks up the largest number, moves it to the end of the array, picks up the second largest, etc.

Moving to adjacer takes $n^{o(1)}$ second Moving a number takes $n^{1+o(1)}$ seco Same for comparis Total sorting time $n^{2+o(1)}$ seconds. Cost of machine: $n^{1+o(1)}$ dollars for $n^{1+o(1)}$ cells. Negligible extra co

The machine also has a "head" moving through array. Head contains $n^{o(1)}$ cells.

Head can see the cell at its current array position; perform arithmetic etc.; move to adjacent array position.

Selection sort: Head looks at each array position, picks up the largest number, moves it to the end of the array, picks up the second largest, etc.

Moving to adjacent array position takes $n^{o(1)}$ seconds. Moving a number to end of array takes $n^{1+o(1)}$ seconds. Same for comparisons etc. Total sorting time: $n^{2+o(1)}$ seconds. Cost of machine: $n^{1+o(1)}$ dollars for $n^{1+o(1)}$ cells. Negligible extra cost for head.



has a rough array. ¹⁾ cells.

cell at

osition;

etc.;

array position.

ad

y position,

st number,

d of the array,

d largest,

Moving to adjacent array position takes $n^{o(1)}$ seconds.

Moving a number to end of array takes $n^{1+o(1)}$ seconds.

Same for comparisons etc.

Total sorting time: $n^{2+o(1)}$ seconds.

Cost of machine: $n^{1+o(1)}$ dollars for $n^{1+o(1)}$ cells.

Negligible extra cost for head.

Possibility 2: The "2-dimensional RA using merge sort." Machine has n^{1+o} in a 2-dimensional $n^{0.5+o(1)}$ rows, n^0 Machine also has Merge sort: Head sorts first $\lfloor n/2 \rfloor$ n sorts last $\lceil n/2 \rceil$ n merges the sorted Moving to adjacent array position takes $n^{o(1)}$ seconds.

Moving a number to end of array takes $n^{1+o(1)}$ seconds.

Same for comparisons etc.

Total sorting time: $n^{2+o(1)}$ seconds.

Cost of machine: $n^{1+o(1)}$ dollars for $n^{1+o(1)}$ cells.

Negligible extra cost for head.

Possibility 2: The machine is a "2-dimensional RAM using merge sort." Machine has $n^{1+o(1)}$ cells in a 2-dimensional array: $n^{0.5+o(1)}$ rows. $n^{0.5+o(1)}$ columns. Machine also has a head. Merge sort: Head recursively sorts first |n/2| numbers; sorts last $\lceil n/2 \rceil$ numbers; merges the sorted lists.

t array position s.

to end of array onds.

sons etc.

ost for head.

Possibility 2: The machine is a "2-dimensional RAM using merge sort." Machine has $n^{1+o(1)}$ cells in a 2-dimensional array: $n^{0.5+o(1)}$ rows. $n^{0.5+o(1)}$ columns. Machine also has a head. Merge sort: Head recursively sorts first |n/2| numbers; sorts last $\lceil n/2 \rceil$ numbers; merges the sorted lists.

Merging requires η to "random" array Average jump: n^0 to adjacent array | Each move takes *i* Total sorting time $n^{1.5+o(1)}$ seconds. Cost of machine: $n^{1+o(1)}$ dollars.

Possibility 2: The machine is a "2-dimensional RAM using merge sort." Machine has $n^{1+o(1)}$ cells in a 2-dimensional array: $n^{0.5+o(1)}$ rows. $n^{0.5+o(1)}$ columns.

Machine also has a head.

Merge sort: Head recursively sorts first |n/2| numbers; sorts last $\lceil n/2 \rceil$ numbers; merges the sorted lists.

Merging requires $n^{1+o(1)}$ jumps to "random" array positions. Average jump: $n^{0.5+o(1)}$ moves to adjacent array positions. Each move takes $n^{o(1)}$ seconds. Total sorting time:

 $n^{1.5+o(1)}$ seconds.

Cost of machine: once again $n^{1+o(1)}$ dollars.

machine is a M

- $^{(1)}$ cells
- array: .5+o(1) columns.
- a head.
- recursively
- umbers;
- umbers;
- lists.

Merging requires $n^{1+o(1)}$ jumps to "random" array positions. Average jump: $n^{0.5+o(1)}$ moves to adjacent array positions. Each move takes $n^{o(1)}$ seconds.

Total sorting time: $n^{1.5+o(1)}$ seconds.

Cost of machine: once again $n^{1+o(1)}$ dollars.

Possibility 3: The "pipelined 2-dimer using radix-2 sort. Machine has n^{1+o} in a 2-dimensional Each cell in the ar network links to the cells in the same c Each cell in the to network links to the cells in the top row

Merging requires $n^{1+o(1)}$ jumps to "random" array positions.

Average jump: $n^{0.5+o(1)}$ moves to adjacent array positions.

Each move takes $n^{o(1)}$ seconds.

Total sorting time: $n^{1.5+o(1)}$ seconds.

Cost of machine: once again $n^{1+o(1)}$ dollars.

Possibility 3: The machine is a "pipelined 2-dimensional RAM using radix-2 sort." Machine has $n^{1+o(1)}$ cells in a 2-dimensional array. Each cell in the array has network links to the 2 adjacent cells in the same column. Each cell in the top row has network links to the 2 adjacent cells in the top row.

 $n^{1+o(1)}$ jumps y positions.

 $^{.5+o(1)}$ moves positions.

 $n^{o(1)}$ seconds.

once again

Possibility 3: The machine is a "pipelined 2-dimensional RAM using radix-2 sort."

Machine has $n^{1+o(1)}$ cells in a 2-dimensional array. Each cell in the array has network links to the 2 adjacent cells in the same column. Each cell in the top row has network links to the 2 adjacent cells in the top row.

Machine also has attached to top-le CPU can read/wri sending request th Does not need to before sending nex CPU can read an of $n^{0.5+o(1)}$ cells in $n^{0.5+o(1)}$ second Sends all requests,

then receives response

Possibility 3: The machine is a "pipelined 2-dimensional RAM using radix-2 sort."

Machine has $n^{1+o(1)}$ cells in a 2-dimensional array. Each cell in the array has network links to the 2 adjacent cells in the same column. Each cell in the top row has network links to the 2 adjacent cells in the top row.

Machine also has a CPU attached to top-left cell. CPU can read/write any cell by sending request through network. before sending next request. CPU can read an entire row of $n^{0.5+o(1)}$ cells in $n^{0.5+o(1)}$ seconds. Sends all requests, then receives responses.

- Does not need to wait for response

machine is a nsional RAM

- $^{(1)}$ cells
- array.
- ray has
- ne 2 adjacent
- olumn.
- p row has
- ne 2 adjacent
- V.

Machine also has a CPU attached to top-left cell.

CPU can read/write any cell by sending request through network. Does not need to wait for response before sending next request.

CPU can read an entire row of $n^{0.5+o(1)}$ cells in $n^{0.5+o(1)}$ seconds. Sends all requests, then receives responses.

Radix-2 sort: CPL shuffles array using even numbers befo $3\ 1\ 4\ 1\ 5\ 9\ 2\ 6\mapsto$ 4 2 6 3 1 1 5 9. Then using bit 1: 4 1 1 5 9 2 6 3. Then using bit 2: 1 1 9 2 3 4 5 6. Then using bit 3: 1 1 2 3 4 5 6 9.

etc.

Machine also has a CPU attached to top-left cell.

CPU can read/write any cell by sending request through network. Does not need to wait for response before sending next request.

CPU can read an entire row of $n^{0.5+o(1)}$ cells in $n^{0.5+o(1)}$ seconds. Sends all requests,

then receives responses.

Radix-2 sort: CPU shuffles array using bit 0, even numbers before odd. $31415926 \mapsto$ 4 2 6 3 1 1 5 9.

Then using bit 1: 4 1 1 5 9 2 6 3.

Then using bit 2: 1 1 9 2 3 4 5 6.

Then using bit 3: 11234569.

etc.

a CPU ft cell.

te any cell by rough network. wait for response at request.

entire row

ds.

onses.

Radix-2 sort: CPU shuffles array using bit 0, even numbers before odd. $3\ 1\ 4\ 1\ 5\ 9\ 2\ 6\mapsto$ $4\ 2\ 6\ 3\ 1\ 1\ 5\ 9.$

Then using bit 1: 4 1 1 5 9 2 6 3.

Then using bit 2: 1 1 9 2 3 4 5 6.

Then using bit 3: 1 1 2 3 4 5 6 9.

etc.

Each shuffle takes $n^{1+o(1)}$ seconds.

 $n^{o(1)}$ shuffles.

Total sorting time: $n^{1+o(1)}$ seconds.

Cost of machine: $n^{1+o(1)}$ dollars.

Radix-2 sort: CPU shuffles array using bit 0, even numbers before odd. $31415926 \mapsto$ 4 2 6 3 1 1 5 9.

Then using bit 1: 4 1 1 5 9 2 6 3.

Then using bit 2: 1 1 9 2 3 4 5 6.

Then using bit 3: 1 1 2 3 4 5 6 9.

etc.

Each shuffle takes $n^{1+o(1)}$ seconds.

 $n^{o(1)}$ shuffles.

Total sorting time: $n^{1+o(1)}$ seconds.

Cost of machine: once again $n^{1+o(1)}$ dollars.

g bit 0, pre odd. Each shuffle takes $n^{1+o(1)}$ seconds.

 $n^{o(1)}$ shuffles.

Total sorting time: $n^{1+o(1)}$ seconds.

Cost of machine: once again $n^{1+o(1)}$ dollars.

Possibility 4: The "2-dimensional me using Schimmler s Machine has n^{1+o} in a 2-dimensional Each cell has netw to the 4 adjacent Machine also has attached to top-le CPU broadcasts in to all of the cells, cells do most of th

Each shuffle takes $n^{1+o(1)}$ seconds.

 $n^{o(1)}$ shuffles.

Total sorting time: $n^{1+o(1)}$ seconds.

Cost of machine: once again $n^{1+o(1)}$ dollars.

Possibility 4: The machine is a "2-dimensional mesh using Schimmler sort." Machine has $n^{1+o(1)}$ cells in a 2-dimensional array. Each cell has network links to the 4 adjacent cells. Machine also has a CPU attached to top-left cell. **CPU** broadcasts instructions to all of the cells, but cells do most of the processing.

once again

Possibility 4: The machine is a "2-dimensional mesh using Schimmler sort." Machine has $n^{1+o(1)}$ cells

in a 2-dimensional array. Each cell has network links to the 4 adjacent cells.

Machine also has a CPU attached to top-left cell. CPU broadcasts instructions to all of the cells, but cells do most of the processing.

Sort row of $n^{0.5+c}$ in $n^{0.5+o(1)}$ second Sort each pair in p $\underline{3\ 1}\ \underline{4\ 1}\ \underline{5\ 9}\ \underline{2\ 6} \mapsto$ 13145926 Sort alternate pair $1 \underline{31} \underline{45} \underline{92} 6 \mapsto$ 1 1 3 4 5 2 9 6 Repeat until numb equals row length. Sort *each* row, in in $n^{0.5+o(1)}$ second

Possibility 4: The machine is a "2-dimensional mesh using Schimmler sort." Machine has $n^{1+o(1)}$ cells

in a 2-dimensional array. Each cell has network links to the 4 adjacent cells.

Machine also has a CPU attached to top-left cell. **CPU** broadcasts instructions to all of the cells, but cells do most of the processing.

Sort row of $n^{0.5+o(1)}$ cells in $n^{0.5+o(1)}$ seconds: Sort each pair in parallel. $3\ 1\ 4\ 1\ 5\ 9\ 2\ 6\mapsto$ 13145926 Sort alternate pairs in parallel. $1 \ 3 \ 1 \ 4 \ 5 \ 9 \ 2 \ 6 \mapsto$ 11345296 Repeat until number of steps equals row length. Sort *each* row, in parallel, in $n^{0.5+o(1)}$ seconds.



machine is a esh

ort."

 $^{(1)}$ cells

array.

ork links

cells.

a CPU

ft cell.

structions

but

ne processing.

Sort row of $n^{0.5+o(1)}$ cells in $n^{0.5+o(1)}$ seconds:

Sort each pair in parallel. $3 1 4 1 5 9 2 6 \mapsto$ 1 3 1 4 5 9 2 6

Sort alternate pairs in parallel. $1 \underline{31} \underline{45} \underline{92} 6 \mapsto$ 1 1 3 4 5 2 9 6

Repeat until number of steps equals row length.

Sort *each* row, in parallel, in $n^{0.5+o(1)}$ seconds.

Schimmler sort: Recursively sort quin parallel. Then for the sort each column Sort each row in parallel sort each row in parallel sort each column

With proper choice left-to-right/rightfor each row, can that this sorts who Sort row of $n^{0.5+o(1)}$ cells in $n^{0.5+o(1)}$ seconds:

Sort each pair in parallel. $\underline{3\ 1\ 4\ 1\ 5\ 9\ 2\ 6}\mapsto$ 13145926

Sort alternate pairs in parallel. $1 \underline{31} \underline{45} \underline{92} 6 \mapsto$ 11345296

Repeat until number of steps equals row length.

Sort *each* row, in parallel, in $n^{0.5+o(1)}$ seconds.

Schimmler sort: Recursively sort quadrants in parallel. Then four steps: Sort each column in parallel. Sort each row in parallel. Sort each column in parallel. Sort each row in parallel. With proper choice of left-to-right/right-to-left for each row, can prove

that this sorts whole array.

o⁽¹⁾ cells ds:

barallel.

s in parallel.

per of steps

parallel, ds. Schimmler sort: Recursively sort quadrants in parallel. Then four steps: Sort each column in parallel. Sort each row in parallel. Sort each column in parallel. Sort each row in parallel.

With proper choice of left-to-right/right-to-left for each row, can prove that this sorts whole array.

For example, assumption 8×8 array is

3	1	4	1	5	9
5	3	5	8	9	7
2	3	8	4	6	2
3	3	8	3	2	7
0	2	8	8	4	1
1	6	9	3	9	9
5	1	0	5	8	2
7	4	9	4	4	5
5					

Schimmler sort: Recursively sort quadrants in parallel. Then four steps: Sort each column in parallel. Sort each row in parallel. Sort each column in parallel.

Sort each row in parallel.

With proper choice of left-to-right/right-to-left for each row, can prove that this sorts whole array. For example, assumption 8×8 array is

3	1	4	1	5
5	3	5	8	9
2	3	8	4	6
3	3	8	3	2
0	2	8	8	4
1	6	9	3	9
5	1	0	5	8
7	4	9	4	4

		hat ells:
9	2	6
7	9	3
2	6	4
7	9	5
1	9	7
9	3	7
2	0	9
5	9	2

uadrants

- our steps:
- in parallel.
- arallel.
- in parallel.
- arallel.
- e of
- to-left
- prove
- ole array.

For example, assume that this 8×8 array is in cells:

3	1	4	1	5	9	2	6
5	3	5	8	9	7	9	3
2	3	8	4	6	2	6	4
3	3	8	3	2	7	9	5
0	2	8	8	4	1	9	7
1	6	9	3	9	9	3	7
5	1	0	5	8	2	0	9
7	4	9	4	4	5	9	2

Rec	curs	ive	ly s	ort	qı
top	\rightarrow	, bo	otto	m	\leftarrow
1	1	2	3	2	2
3	3	3	3	4	5
3	4	4	5	6	6
5	8	8	8	9	9
1	1	0	0	2	2
4	4	3	2	5	4
7	6	5	5	9	8
9	9	8	8	9	9

For example, assume that this 8×8 array is in cells:

3	1	4	1	5	9	2	6
5	3	5	8	9	7	9	3
2	3	8	4	6	2	6	4
3	3	8	3	2	7	9	5
0	2	8	8	4	1	9	7
1	6	9	3	9	9	3	7
5	1	0	5	8	2	0	9
7	4	9	4	4	5	9	2

Recursively sort quadrants, top \rightarrow , bottom \leftarrow :

1	1	2	3	2
3	3	3	3	4
3	4	4	5	6
5	8	8	8	9
1	1	0	0	2
4	4	3	2	5
7	6	5	5	9
9	9	8	8	9

2	2	3
5	5	6
6	7	7
9	9	9
2	1	0
4	4	3
8	7	7
9	9	9

me that in cells:



Recursively sort quadrants, top \rightarrow , bottom \leftarrow :

1	1	2	3	2	2	2	3
3	3	3	3	4	5	5	6
3	4	4	5	6	6	7	7
5	8	8	8	9	9	9	9
1	1	0	0	2	2	1	0
4	4	3	2	5	4	4	3
7	6	5	5	9	8	7	7
9	9	8	8	9	9	9	9

Sort each column in parallel:

8	-		-		-
1	1	0	0	2	2
1	1	2	2	2	2
3	3	3	3	4	4
3	4	3	3	5	5
4	4	4	5	6	6
5	6	5	5	9	8
7	8	8	8	9	9
9	9	8	8	9	9

Recursively sort quadrants, top \rightarrow , bottom \leftarrow :

1	1	2	3	2	2	2	3
3	3	3	3	4	2 5	5	6
3	4	4	5	6	6	7	7
5	8	8	8	9	9	9	9
1	1	0	0	2	2	1	0
4	4	3	2	5	4	4	3
7	6	5	5	9	8	7	7
9	9	8	8	9	9	9	9

Sort each column in parallel:

1	1	0	0	2
1	1	2	2	2
3	3	3	3	4
3	4	3	3	5
4	4	4	5	6
5	6	5	5	9
7	8	8	8	9
9	9	8	8	9

2	1	0
2	2	3
4	4	3
5	5	6
6	7	7
8	7	7
9	9	9
9	9	9

uadrants,

•



Sort each column in parallel:

1	1	0	0	2	2	1	0
1	1	2	2	2	2	2	3
3	3	3	3	4	4	4	3
3	4	3	3	5	5	5	6
4	4	4	5	6	6	7	7
5	6	5	5	9	8	7	7
7	8	8	8	9	9	9	9
9	9	8	8	9	9	9	9

Sort each row in part each row in part

3 2 2 2 2 2 3 3 3 3 3 4 6 5 5 5 4 3 4 4 4 5 6 6 9 8 7 7 6 5 7 8 8 8 9 9						
3 3 3 3 3 4 6 5 5 5 4 3 4 4 4 5 6 6 9 8 7 7 6 5 7 8 8 8 9 9	0	0	0	1	1	1
6 5 5 5 4 3 4 4 4 5 6 6 9 8 7 7 6 5 7 8 8 8 9 9	3	2	2	2	2	2
4 4 4 5 6 6 9 8 7 7 6 5 7 8 8 8 9 9	3	3	3	3	3	4
9 8 7 7 6 5 7 8 8 8 9 9	6	5	5	5	4	3
7 8 8 8 9 9	4	4	4	5	6	6
	9	8	7	7	6	5
999999	7	8	8	8	9	9
	9	9	9	9	9	9

Sort each column in parallel:

1	1	0	0	2	2	1	0
1	1	2	2	2	2	2	3
3	3	3	3	4	4	4	3
3	4	3	3	5	5	5	6
4	4	4	5	6	6	7	7
5	6	5	5	9	8	7	7
7	8	8	8	9	9	9	9
9	9	8	8	9	9	9	9

1	2	2
2	1	1
4	4	4
3	3	3
6	7	7
5	5	5
9	9	9
9	8	8







Sort each column in parallel:

0	0	0	1	1	1
3	2	2	2	2	2
3	3	3	3	3	3
4	4	4	5	4	4
6	5	5	5	6	5
7	8	7	7	6	6
9	8	8	8	9	9
9	9	9	9	9	9
Sort each row in parallel, alternately \leftarrow , \rightarrow :

0	0	0	1	1	1	2	2
3	2	2	2	2	2	1	1
3	3	3	3	3	4	4	4
6	5	5	5	4	3	3	3
4	4	4	5	6	6	7	7
9	8	7	7	6	5	5	5
7	8	8	8	9	9	9	9
9	9	9	9	9	9	8	8

Sort each column in parallel:

0	0	0	1	1	
3	2	2	2	2	
3	3	3	3	3	
4	4	4	5	4	
6	5	5	5	6	
7	8	7	7	6	
9	8	8	8	9	
9	9	9	9	9	

1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	7	7
9	8	8
9	9	9

arallel,



Sort each column in parallel:

		-			-	-	
0	0	0	1	1	1	1	1
3	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	5	4	4	4	4
6	5	5	5	6	5	5	5
7	8	7	7	6	6	7	7
9	8	8	8	9	9	8	8
9	9	9	9	9	9	9	9

Sort each row in p \leftarrow or \rightarrow as desired

0	0	0	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	7	7	7	7
8	8	8	8	8	9
9	9	9	9	9	9

Sort each column in parallel:

0	0	0	1	1	1	1	1
3	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	5	4	4	4	4
6	5	5	5	6	5	5	5
7	8	7	7	6	6	7	7
9	8	8	8	9	9	8	8
9	9	9	9	9	9	9	9

		-
1	1	1
2	2	3
3	3	3
4	4	5
5	6	6
7	7	8
9	9	9
9	9	9







Sort one row in $n^{0.5+o(1)}$ second

All rows in parallel $n^{0.5+o(1)}$ seconds.

Total sorting time: $n^{0.5+o(1)}$ seconds.

Cost of machine: $n^{1+o(1)}$ dollars. Sort each row in parallel, \leftarrow or \rightarrow as desired:

Sort one row in $n^{0.5+o(1)}$ seconds.

All rows in parallel: $n^{0.5+o(1)}$ seconds.

Total sorting time: $n^{0.5+o(1)}$ seconds.

Cost of machine: once again $n^{1+o(1)}$ dollars.

arallel,

1:



Sort one row in $n^{0.5+o(1)}$ seconds.

All rows in parallel: $n^{0.5+o(1)}$ seconds.

Total sorting time: $n^{0.5+o(1)}$ seconds.

Cost of machine: once again $n^{1+o(1)}$ dollars.

Some philosophica

1-tape Turing mad RAMs, 2-dimensio compute the same

Prove this by prov each machine can

computations on t

- (We believe that e
- reasonable model
- can be simulated l
- 1-tape Turing mac "Church-Turing th

Sort one row in $n^{0.5+o(1)}$ seconds.

All rows in parallel: $n^{0.5+o(1)}$ seconds.

Total sorting time: $n^{0.5+o(1)}$ seconds.

Cost of machine: once again $n^{1+o(1)}$ dollars.

Some philosophical notes

1-tape Turing machines, RAMs, 2-dimensional meshes compute the same functions.

Prove this by proving that each machine can simulate computations on the others.

(We believe that *every* reasonable model of computation can be simulated by a 1-tape Turing machine. "Church-Turing thesis.")

ds.

|-

once again

Some philosophical notes

1-tape Turing machines,RAMs, 2-dimensional meshescompute the same functions.

Prove this by proving that each machine can simulate computations on the others.

(We believe that *every* reasonable model of computation can be simulated by a 1-tape Turing machine. "Church-Turing thesis.")

1-tape Turing mad RAMs, 2-dimensio compute the same in polynomial time at polynomial cost Prove this by prov simulations are po (Is this true for ev reasonable model Quantum compute factor in polynomi Can Turing machi Can quantum com

Some philosophical notes

1-tape Turing machines, RAMs, 2-dimensional meshes compute the same functions.

Prove this by proving that each machine can simulate computations on the others.

(We believe that *every* reasonable model of computation can be simulated by a 1-tape Turing machine. "Church-Turing thesis.")

1-tape Turing machines, RAMs, 2-dimensional meshes compute the same functions in polynomial time at polynomial cost. Prove this by proving that simulations are polynomial. (Is this true for every reasonable model of computation? Quantum computers can factor in polynomial time. Can Turing machines do that? Can quantum computers be built?)

l notes

chines, nal meshes functions.

ing that

simulate

he others.

every

of computation

руа

chine.

esis.")

1-tape Turing machines, RAMs, 2-dimensional meshes compute the same functions in polynomial time at polynomial cost.

Prove this by proving that simulations are polynomial.

(Is this true for everyreasonable model of computation?Quantum computers canfactor in polynomial time.Can Turing machines do that?Can quantum computers be built?)

1-tape Turing made RAMs, 2-dimensional do not compute the same functions within, e.g., time nand cost $n^{1+o(1)}$.

Example: 1-tape $\bar{}$ cannot sort in n^{1+} Too local.

Example: 2-dimention cannot sort in $n^{0.5}$ Too sequential. 1-tape Turing machines, RAMs, 2-dimensional meshes compute the same functions in polynomial time at polynomial cost.

Prove this by proving that simulations are polynomial.

(Is this true for every reasonable model of computation?

Quantum computers can factor in polynomial time.

Can Turing machines do that?

Can quantum computers be built?)

1-tape Turing machines, RAMs, 2-dimensional meshes *do not* compute the same functions within, e.g., time $n^{1+o(1)}$ and cost $n^{1+o(1)}$ Example: 1-tape Turing machine cannot sort in $n^{1+o(1)}$ seconds. Too local. Example: 2-dimensional RAM

cannot sort in $n^{0.5+o(1)}$ seconds. Too sequential.

chines,

nal meshes

functions

2

•

ing that Iynomial.

ery

of computation?

ers can

al time.

nes do that?

puters be built?)

1-tape Turing machines, RAMs, 2-dimensional meshes do not compute the same functions within, e.g., time $n^{1+o(1)}$ and cost $n^{1+o(1)}$.

Example: 1-tape Turing machine cannot sort in $n^{1+o(1)}$ seconds. Too local.

Example: 2-dimensional RAM cannot sort in $n^{0.5+o(1)}$ seconds. Too sequential. Review of sorting measured in secon machine costing *n*

- $n^{2.0+o(1)}$: 1-tape
- $n^{1.5+o(1)}$: 2-dimen
- $n^{1.0+o(1)}$: pipeline $n^{0.5+o(1)}$: 2-dimen

Why does anyone sorting time is n^{1-} Why choose third Silly! Fourth mack

1-tape Turing machines, RAMs, 2-dimensional meshes *do not* compute the same functions within, e.g., time $n^{1+o(1)}$ and cost $n^{1+o(1)}$

Example: 1-tape Turing machine cannot sort in $n^{1+o(1)}$ seconds. Too local.

Example: 2-dimensional RAM cannot sort in $n^{0.5+o(1)}$ seconds. Too sequential.

Review of sorting times, measured in seconds, for machine costing $n^{1+o(1)}$ dollars:

 $n^{2.0+o(1)}$: 1-tape Turing machine. $n^{1.5+o(1)}$: 2-dimensional RAM.

 $n^{1.0+o(1)}$: pipelined RAM.

 $n^{0.5+o(1)}$: 2-dimensional mesh.

Why does anyone say that sorting time is $n^{1+o(1)}$? Why choose third machine? Silly! Fourth machine is better!

chines, nal meshes

 $n^{1+o(1)}$

Furing machine -o(1) seconds.

sional RAM 5+o(1) seconds.

Review of sorting times, measured in seconds, for machine costing $n^{1+o(1)}$ dollars: $n^{2.0+o(1)}$: 1-tape Turing machine. $n^{1.5+o(1)}$: 2-dimensional RAM. $n^{1.0+o(1)}$: pipelined RAM. $n^{0.5+o(1)}$: 2-dimensional mesh. Why does anyone say that sorting time is $n^{1+o(1)}$? Why choose third machine? Silly! Fourth machine is better!

Warning: o(1) is a Speedup factor su might not be a spe for small values of When n is small, RAM might seem sensible machine c But, once n is large having a huge mer waiting for a single

is a silly machine of

Review of sorting times, measured in seconds, for machine costing $n^{1+o(1)}$ dollars:

 $n^{2.0+o(1)}$: 1-tape Turing machine. $n^{1.5+o(1)}$: 2-dimensional RAM. $n^{1.0+o(1)}$: pipelined RAM. $n^{0.5+o(1)}$: 2-dimensional mesh.

Why does anyone say that sorting time is $n^{1+o(1)}$? Why choose third machine? Silly! Fourth machine is better! Warning: o(1) is asymptotic. Speedup factor such as $n^{0.5+o(1)}$ might not be a speedup for small values of n.

When *n* is small, RAM might seem to be a sensible machine design.

But, once *n* is large enough, having a huge memory waiting for a single CPU is a silly machine design. times, ds, for ,^{1+o(1)} dollars:

Turing machine. nsional RAM. ed RAM.

nsional mesh.

say that ⊦o(1)<u>?</u>

machine?

nine is better!

Warning: o(1) is asymptotic. Speedup factor such as $n^{0.5+o(1)}$ might not be a speedup for small values of n.

When *n* is small, RAM might seem to be a sensible machine design.

But, once *n* is large enough, having a huge memory waiting for a single CPU is a silly machine design.

Myth: Parallel computati improve price-perf p parallel compute may reduce time b but increase cost k Reality: Can often a large serial com into p small parall so cost does *not* increase by factor

Warning: o(1) is asymptotic. Speedup factor such as $n^{0.5+o(1)}$ might not be a speedup for small values of n.

When n is small, RAM might seem to be a sensible machine design.

But, once n is large enough, having a huge memory waiting for a single CPU is a silly machine design.

Myth: Parallel computation cannot improve price-performance ratio; *p* parallel computers may reduce time by factor pbut increase cost by factor p. Reality: Can often convert a *large* serial computer into *p* small parallel cells, so cost does *not* increase by factor p.

asymptotic. ch as $n^{0.5+o(1)}$ eedup

n.

to be a lesign.

ge enough,

mory

e CPU

design.

Myth:

Parallel computation cannot improve price-performance ratio; p parallel computers may reduce time by factor pbut increase cost by factor p.

Reality: Can often convert a *large* serial computer into *p small* parallel cells, so cost does *not* increase by factor *p*. Myth: Designing a cannot produce me small constant-fac compared to, e.g., What matters is s streamlining, such instruction-decodin

Reality: In 1997, I was 1000 times fai

set of Pentiums at

What matters is p

Myth:

Parallel computation cannot improve price-performance ratio; *p* parallel computers may reduce time by factor pbut increase cost by factor p.

Reality: Can often convert a *large* serial computer into *p* small parallel cells, so cost does *not* increase by factor p.

Myth: Designing a new machine cannot produce more than a compared to, e.g., a Pentium. What matters is special-purpose streamlining, such as reducing instruction-decoding costs. Reality: In 1997, DES Cracker

was 1000 times faster than a What matters is parallelism.

- small constant-factor improvement
- set of Pentiums at the same price.

on cannot

ormance ratio;

ers

by factor p

by factor p.

convert

outer

el cells,

p.

Myth: Designing a new machine cannot produce more than a small constant-factor improvement compared to, e.g., a Pentium. What matters is special-purpose streamlining, such as reducing instruction-decoding costs.

Reality: In 1997, DES Cracker was 1000 times faster than a set of Pentiums at the same price. What matters is parallelism.

Future computers massively parallel Look at o(1) detai we've reached larg Computer designe today's RAM-style just as we laugh a a 1-tape Turing m Algorithm experts

today's dominant algorithm analysis count CPU "opera view memory acce Myth: Designing a new machine cannot produce more than a small constant-factor improvement compared to, e.g., a Pentium. What matters is special-purpose streamlining, such as reducing instruction-decoding costs.

Reality: In 1997, DES Cracker was 1000 times faster than a set of Pentiums at the same price. What matters is parallelism.

Future computers will be massively parallel meshes. Look at o(1) details to see that we've reached large enough n.

Computer designers will laugh at today's RAM-style machines, just as we laugh at a 1-tape Turing machine.

Algorithm experts will laugh at today's dominant style of algorithm analysis, where we count CPU "operations" and view memory access as free.

- a new machine
- ore than a
- tor improvement
- a Pentium.
- pecial-purpose
- as reducing
- ng costs.
- DES Cracker ster than a
- the same price. arallelism.

Future computers will be massively parallel meshes. Look at o(1) details to see that we've reached large enough n.

Computer designers will laugh at today's RAM-style machines, just as we laugh at a 1-tape Turing machine.

Algorithm experts will laugh at today's dominant style of algorithm analysis, where we count CPU "operations" and view memory access as free.

Collision search

Common cryptana Find collision in *H*

Input: Program to at high speed. *H* is a function from

256-bit strings to

256-bit strings.

Output: 256-bit structure such that $x_1 \neq x_2$ and $H(x_1) = H(x_2)$ Future computers will be massively parallel meshes. Look at o(1) details to see that we've reached large enough n.

Computer designers will laugh at today's RAM-style machines, just as we laugh at a 1-tape Turing machine.

Algorithm experts will laugh at today's dominant style of algorithm analysis, where we count CPU "operations" and view memory access as free.

Collision search

Common cryptanalytic problem: Find collision in H.

Input: Program to compute H at high speed. H is a function from 256-bit strings to 256-bit strings.

Output: 256-bit strings x_1, x_2 such that $x_1 \neq x_2$ and $H(x_1) = H(x_2)$.

will be meshes.

Is to see that e enough n.

rs will laugh at machines,

t

achine.

will laugh at

style of

where we

ntions" and

ss as free.

Collision search

Common cryptanalytic problem: Find collision in *H*.

Input: Program to compute *H*at high speed.*H* is a function from256-bit strings to256-bit strings.

Output: 256-bit strings x_1, x_2 such that $x_1 \neq x_2$ and $H(x_1) = H(x_2)$.

For any 256-bit r: Compute H(r), H(r)until finding a stri that begins with 4 (A "distinguished Call that string ZOops, Z(r) might But usually it does Computing Z(r) t involves $\approx 2^{40}$ inp

Collision search

Common cryptanalytic problem: Find collision in H.

Input: Program to compute H at high speed. H is a function from 256-bit strings to 256-bit strings.

Output: 256-bit strings x_1, x_2 such that $x_1 \neq x_2$ and $H(x_1) = H(x_2)$.

For any 256-bit r: Compute H(r), H(H(r)), ... until finding a string that begins with 40 zero bits. (A "distinguished point.") Call that string Z(r). Oops, Z(r) might not exist. But usually it does. Computing Z(r) typically involves $\approx 2^{40}$ inputs to *H*.

lytic problem:

o compute H

m

trings x_1, x_2

2).

For any 256-bit r: Compute H(r), H(H(r)), ... until finding a string that begins with 40 zero bits. (A "distinguished point.") Call that string Z(r).

Oops, Z(r) might not exist. But usually it does.

Computing Z(r) typically involves $\approx 2^{40}$ inputs to H.

Choose random r_1 Compute $Z(r_1)$, Z Uses $\approx 2^{40}n$ input r_1 , $H(r_1)$, $H^2(r_1)$, r_2 , $H(r_2)$, $H^2(r_2)$, r_n , $H(r_n)$, $H^2(r_n)$ "Birthday paradox $pprox 2^{79} n^2$ input pai chances for a collis

For any 256-bit r: Compute $H(r), H(H(r)), \ldots$ until finding a string that begins with 40 zero bits. (A "distinguished point.") Call that string Z(r).

Oops, Z(r) might not exist. But usually it does.

Computing Z(r) typically involves $\approx 2^{40}$ inputs to *H*.

Choose random r_1, r_2, \ldots, r_n . Uses $\approx 2^{40}n$ inputs to *H*: $r_1, H(r_1), H^2(r_1), H^3(r_1), \ldots$ $r_2, H(r_2), H^2(r_2), H^3(r_2), \ldots$ $r_n, H(r_n), H^2(r_n), H^3(r_n), \dots$ "Birthday paradox": $\approx 2^{79} n^2$ input pairs, so $\approx 2^{79} n^2$ chances for a collision in H_{\cdot}

Compute $Z(r_1), Z(r_2), \ldots, Z(r_n)$.

(*H*(*r*)), . . . ng 0 zero bits. point.") (*r*).

not exist.

5.

ypically uts to *H*.

Choose random r_1, r_2, \ldots, r_n . Compute $Z(r_1), Z(r_2), ..., Z(r_n)$. Uses $\approx 2^{40}n$ inputs to *H*: $r_1, H(r_1), H^2(r_1), H^3(r_1), \ldots$ $r_2, H(r_2), H^2(r_2), H^3(r_2), \ldots$ $r_n, H(r_n), H^2(r_n), H^3(r_n), \ldots$ "Birthday paradox": $pprox 2^{79} n^2$ input pairs, so $pprox 2^{79} n^2$ chances for a collision in H.

Say there's a collis $H^{161}(r_2) = H^{190}(r_2)$ $Z(r_2)$ is after H^{16} $Z(r_7)$ is after H^{19} and $H^{160}(r_2) \neq H$ Then $Z(r_2) = Z(r_2)$ Recognize this by $Z(r_1), Z(r_2), \ldots,$ and comparing ad Backtrack to find Oops, may have m backtracking can I But usually not a

Choose random r_1, r_2, \ldots, r_n . Compute $Z(r_1), Z(r_2), \ldots, Z(r_n)$. Uses $\approx 2^{40}n$ inputs to *H*: $r_1, H(r_1), H^2(r_1), H^3(r_1), \ldots$ $r_2, H(r_2), H^2(r_2), H^3(r_2), \ldots$ $r_n, H(r_n), H^2(r_n), H^3(r_n), \ldots$

"Birthday paradox": $pprox 2^{79} n^2$ input pairs, so $pprox 2^{79} n^2$ chances for a collision in H.

Say there's a collision: e.g., $H^{161}(r_2) = H^{190}(r_7)$ where $Z(r_2)$ is after $H^{161}(r_2)$, $Z(r_7)$ is after $H^{190}(r_7)$, and $H^{160}(r_2) \neq H^{189}(r_7)$. Then $Z(r_2) = Z(r_7)$. Recognize this by sorting $Z(r_1), Z(r_2), \ldots, Z(r_n)$ and comparing adjacent outputs. Backtrack to find collision. backtracking can be expensive. But usually not a problem.

- Oops, may have multiple collisions;

$$(r_2), \ldots, r_n$$
.
 $(r_2), \ldots, Z(r_n)$.

ts to H: $H^{3}(r_{1}), \ldots$ $H^{3}(r_{2}), \ldots$

 $, H^{3}(r_{n}), \ldots$

": rs, so $pprox 2^{79} n^2$ sion in *H*.

Say there's a collision: e.g., $H^{161}(r_2) = H^{190}(r_7)$ where $Z(r_2)$ is after $H^{161}(r_2)$, $Z(r_7)$ is after $H^{190}(r_7)$, and $H^{160}(r_2) \neq H^{189}(r_7)$. Then $Z(r_2) = Z(r_7)$. Recognize this by sorting $Z(r_1), Z(r_2), \ldots, Z(r_n)$ and comparing adjacent outputs.

Backtrack to find collision.

Oops, may have multiple collisions; backtracking can be expensive. But usually not a problem.



Mesh computer is about *n* times fast not much more ex Say there's a collision: e.g., $H^{161}(r_2) = H^{190}(r_7)$ where $Z(r_2)$ is after $H^{161}(r_2)$, $Z(r_7)$ is after $H^{190}(r_7)$, and $H^{160}(r_2) \neq H^{189}(r_7)$.

Then $Z(r_2) = Z(r_7)$. Recognize this by sorting $Z(r_1), Z(r_2), \ldots, Z(r_n)$ and comparing adjacent outputs. Backtrack to find collision.

Oops, may have multiple collisions; backtracking can be expensive. But usually not a problem.

Serial computer: $\approx 2^{40}n$ evaluations of H: $\approx n \log n$ sorting steps; $\approx 256n$ bits of RAM. 2-dimensional mesh computer with n parallel processors: $\approx 2^{40}$ evaluations of *H*: $\approx 8\sqrt{n}$ sorting steps; $\approx n$ small cells.

Mesh computer is about *n* times faster, not much more expensive.

sion: e.g.,
$$r_7$$
) where
 ${}^1(r_2)$,
 ${}^0(r_7)$,
 ${}^{189}(r_7)$.

 $r_7).$

sorting

 $Z(r_n)$

acent outputs.

collision.

nultiple collisions;

be expensive.

problem.

Serial computer: $\approx 2^{40}n$ evaluations of *H*; $\approx n \log n$ sorting steps; $\approx 256n$ bits of RAM.

2-dimensional mesh computer with n parallel processors: $\approx 2^{40}$ evaluations of H; $\approx 8\sqrt{n}$ sorting steps; $\approx n$ small cells.

Mesh computer is about *n* times faster, not much more expensive.

Using collision sea for "discrete logar Want to figure our P, kP on an ellipti Define H(x, y) = xFind collision in H usually reveals k. **Price-performance** $q^{1/2+o(1)}$ dollar-se if curve has q poir Fancier methods, s rho, kangaroo, etc

Serial computer: $\approx 2^{40}n$ evaluations of *H*; $\approx n \log n$ sorting steps; $\approx 256n$ bits of RAM.

2-dimensional mesh computer with n parallel processors: $\approx 2^{40}$ evaluations of *H*; $\approx 8\sqrt{n}$ sorting steps; $\approx n$ small cells.

Mesh computer is about *n* times faster, not much more expensive.

Using collision search for "discrete logarithms": Want to figure out k given P, kP on an elliptic curve. Define H(x, y) = xP + ykP. Find collision in *H*; usually reveals k. Price-performance ratio: $q^{1/2+o(1)}$ dollar-seconds if curve has q points. Fancier methods, same 1/2: rho, kangaroo, etc.