

SFC



increases locality

"2-D locality sensitive hash"

Map Reduce

data structures key-value

SQL



sequential
Java

MT
Java

- global shared memory
- global heap
- global vars.

~~synchronization~~

- safe

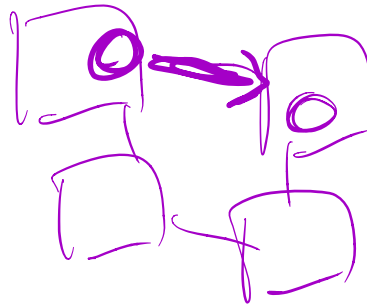


- deadlock

fine-grained



6x!



message passing

MPI

Sockets

Dijkstra

seq: ~ 20 LOC

MT: $\sim 30-40$ LOC

MP: divide the heap (partitioning)

load balancing

~~10,000 LOC~~ granularity

5,000 LOC message passing infra.

deadlock

asynchronous

minimizing
state xfr

\Rightarrow fewer
msgs

fewer
bytes

"callback"
hell

~~A~~
~~B~~
~~C~~
~~D~~

~~A~~ B

B C

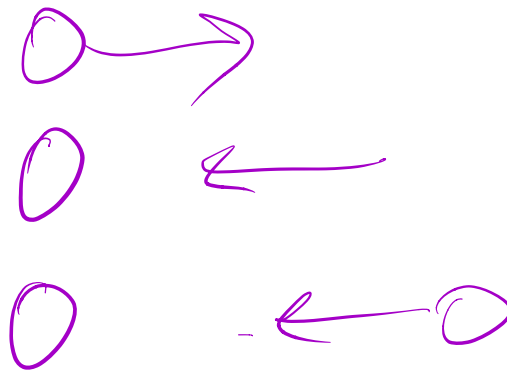
C D

Continuations

closure

distributed

Synchronous checkpoint



(logging \rightarrow) every message

STATE OF THE ART

MSGSEND 25 96 TYPE

...

MP Infrastructure

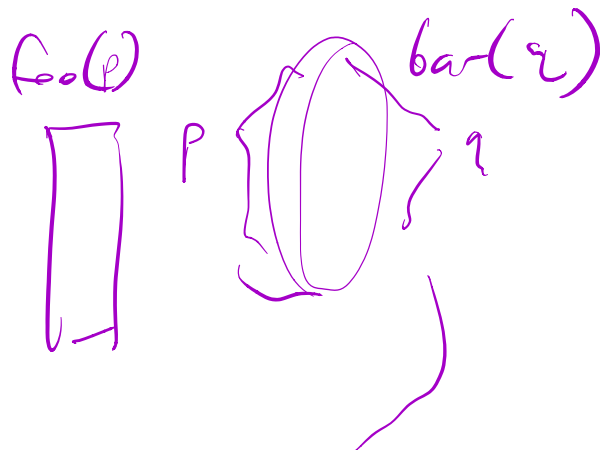
"narrow"
APIs
DSL

graph-analysis
map reduce
SQL ...

"few dream"
automatic //ization

compiler conservative
in direct version w/ //ism

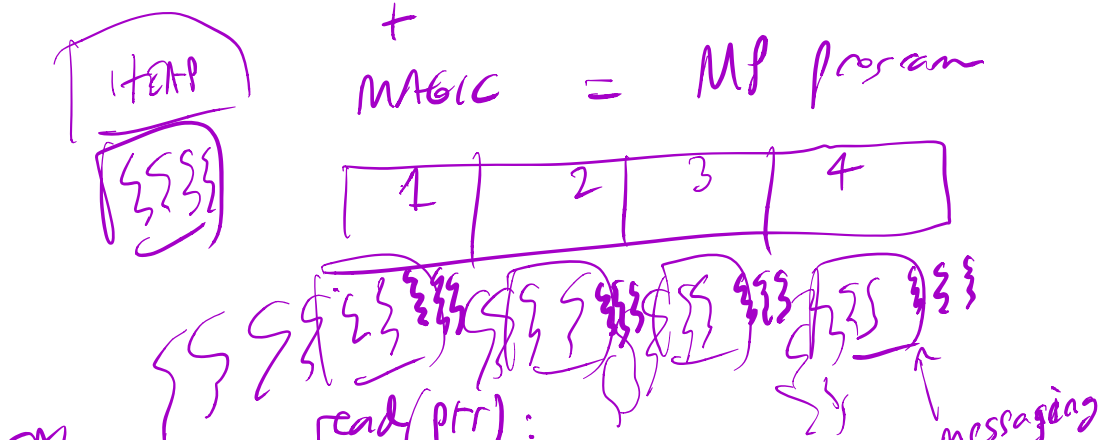
aliasing



Distributed Shared Memory

MT program

+
MAGIC = MP program



PERC
read
~ 10 cycles

if ptr in 1
read it 2 return val

to
manage
heap

MORE THREADED
HIDE
LATENCY

if not
find region
send message "HEY"
await response
return it

LOCKS ?

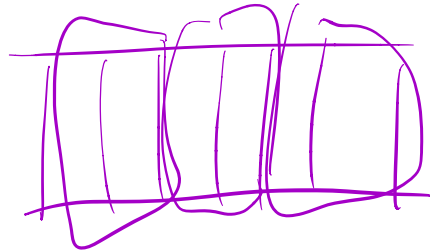
LAZY/
PREFETCHING

wide latency

COST OF
MESSAGES

Treadmarks

MT program

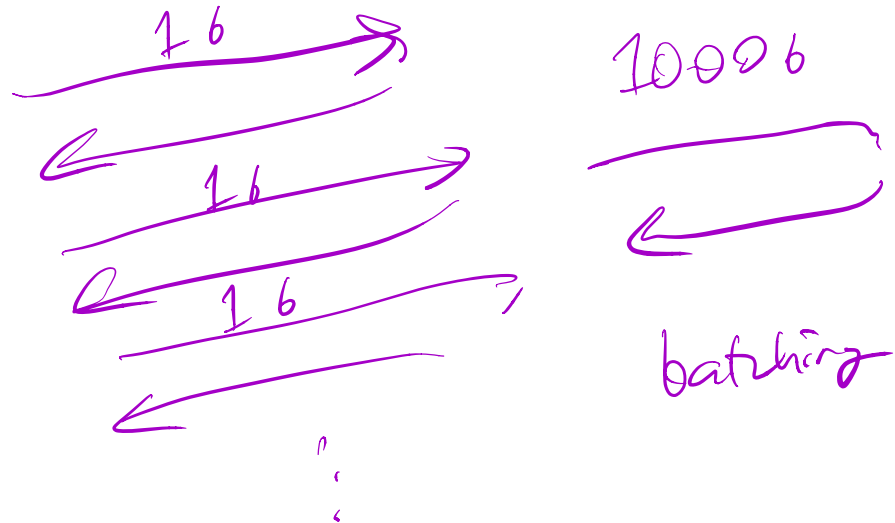


direct

FACE

- rkel

CC myprog.c -ltreadmarks



lazy release consistency

locally cached versions + updates
acquire lock on master

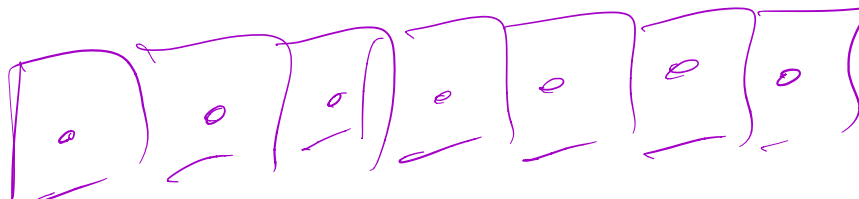
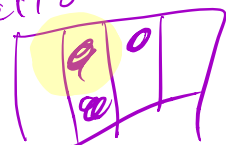
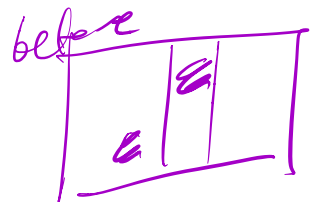
⋮

local

on your own after
copy

release lock

send updates to everyone



DSM:

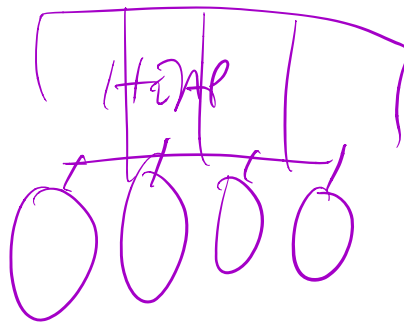
ordinary DSM



MT



MP



"MT"



MP

PGAS

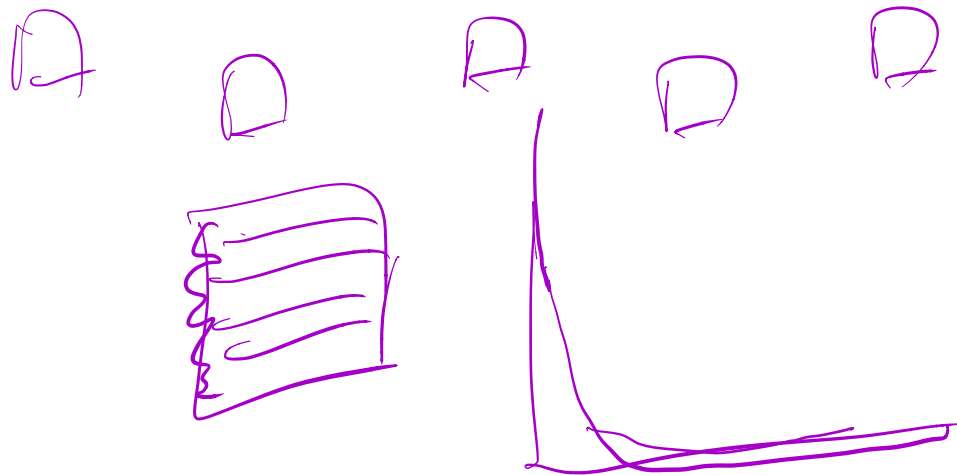
Partitioned
Global
Address
Space

"Grappa"

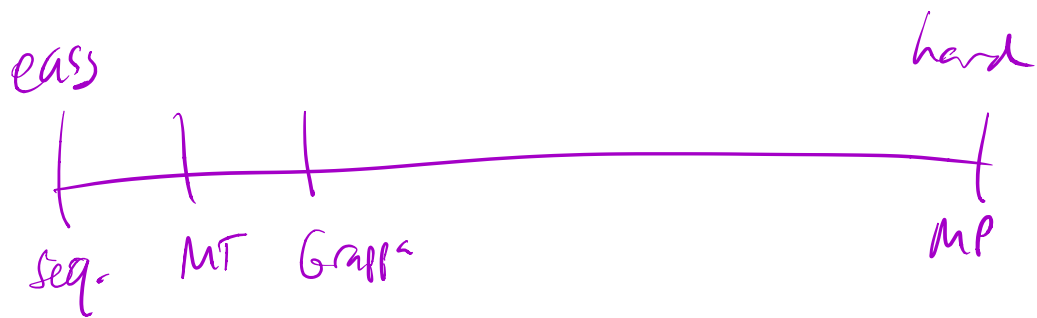
local memory

global memory - partitioned

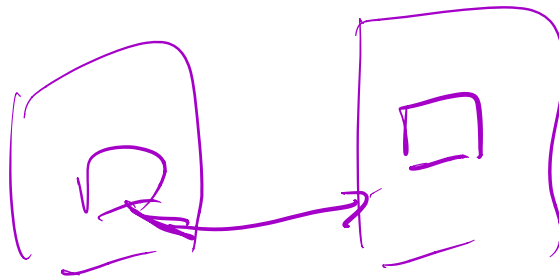
InfiniBand



point in
the design space



RDMA



user-level Thread System

Cooperative scheduling

explicit yields

int x; invoke lambda

$[&](0) \{ x_{code} \};$

\equiv — read only
 $\&$ — by reference
 $\equiv x, y, \dots$

READ

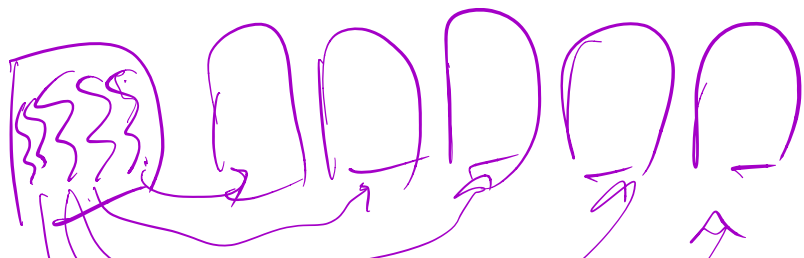
ADMA

→ no context switch

delegate — send "lambda" (Code)
instead of the data

work stealing

load balancing





randomized work stealing

