# Parallel Functional Programming Lecture 8 Data Parallelism II

Mary Sheeran (with thanks to Ben Lippmeier for borrowed slides)

http://www.cse.chalmers.se/edu/course/pfp

# Data parallelism

Perform *same* computation on a collection of *differing* data values

examples: HPF (High Performance Fortran) CUDA

Both support only flat data parallelism

Flat : each of the individual computations on (array) elements is sequential

those computations don't need to communicate parallel computations don't spark further parallel computations

#### **Regular, Shape-polymorphic, Parallel Arrays in Haskell**

Gabriele Keller<sup>†</sup>

Simon Peyton Jones<sup>‡</sup> Ben Lippmeier<sup>†</sup>

Roman Leshchinskiy<sup>†</sup>

<sup>†</sup>Computer Science and Engineering, University of New South Wales {keller,chak,rl,benl}@cse.unsw.edu.au <sup>‡</sup>Microsoft Research Ltd, Cambridge simonpj@microsoft.com

Manuel M. T. Chakravarty<sup>†</sup>

API for purely functional, collective operations over dense, rectangular, multi-dimensional arrays supporting shape polymorphism

**ICFP 2010** 

# Ideas

Purely functional array interface using collective (whole array) operations like map, fold and permutations can

- combine efficiency and clarity
- focus attention on structure of algorithm, away from low level details

Influenced by work on algorithmic skeletons based on Bird Meertens formalism

Provides shape polymorphism not in a standalone specialist compiler like SAC, but using the Haskell type system

# terminology

#### **Regular arrays**

dense, rectangular, most elements non-zero

#### shape polymorphic

functions work over arrays of arbitrary dimension

# terminology

**Regular arrays** 

dense, rectan

**shape polym** functions wo note: the arrays are purely functional and immutable

All elements of an array are demanded at once -> parallelism

P processing elements, n array elements => n/P consecutive elements on each proc. element

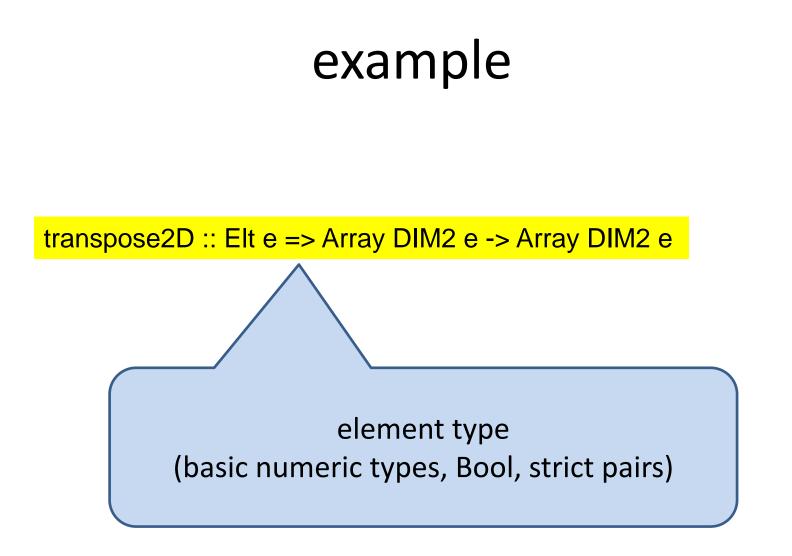
## version

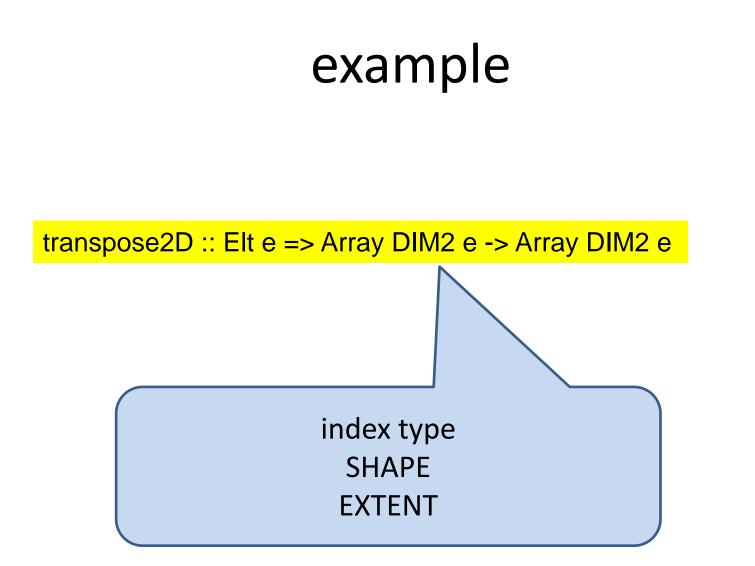
I use Repa 2.1.1.5 (which works with the GHC that you get with the current Haskell platform)

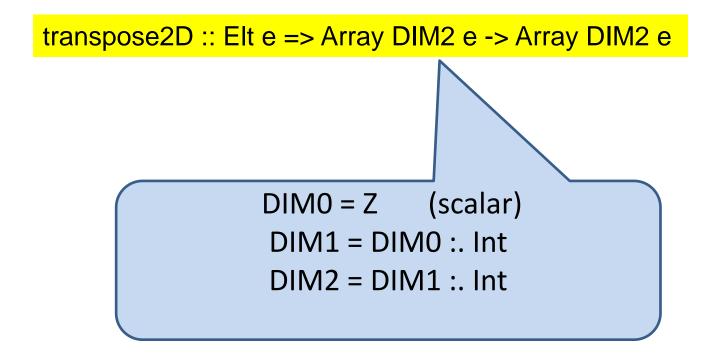
If you have GHC 7.4 installed, you can use a later Repa, which has more array types (and doubtless better performance)

import Data.Array.Repa as A

transpose2D :: Elt e => Array DIM2 e -> Array DIM2 e







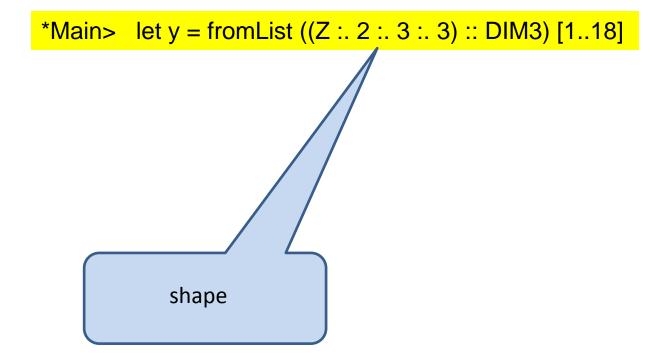
# snoc lists

```
Haskell lists are cons lists
1:2:3:[] is the same as [1,2,3]
```

Repa uses snoc lists at type level for shape types and at value level for shapes

DIM2 = Z :. Int :. Int is a shape type

Z :. i :. j read as (i,j) is an index into a two dim. array



\*Main> let y = fromList ((Z :. 2 :. 3 :. 3) :: DIM3) [1..18]

the type of the shape needs to be there, otherwise get very annoying error messages

\*Main> let y = fromList ((Z :. 2 :. 3 :. 3) :: DIM3) [1..18]

```
*Main> y
Array (Z :. 2 :. 3 :. 3) [1.0,2.0,3.0,4.0,5.0,6.0,7.0,8.0,9.0,10.0,11.0,12.0,13.
0,14.0,15.0,16.0,17.0,18.0]
```

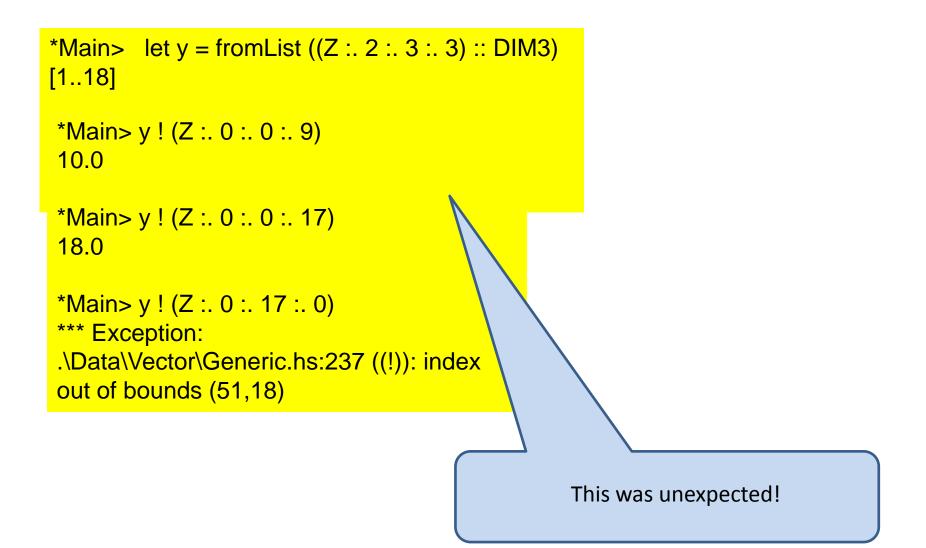
\*Main> extent y ((Z :. 2) :. 3) :. 3

```
*Main> let y = fromList ((Z :. 2 :. 3 :. 3) :: DIM3) [1..18]
*Main> y ! (Z :. 0 :. 0 :. 0)
1.0
*Main> y ! (Z :. 1 :. 1 :. 1)
14.0
```

#### \*Main> let y = fromList ((Z :. 2 :. 3 :. 3) :: DIM3) [1..18]

\*Main> y ! (Z :. 0 :. 0 :. 20) \*\*\* Exception: .\Data\Vector\Generic.hs:237 ((!)): index out of bounds (20,18)

bounds checking is done at RUN TME

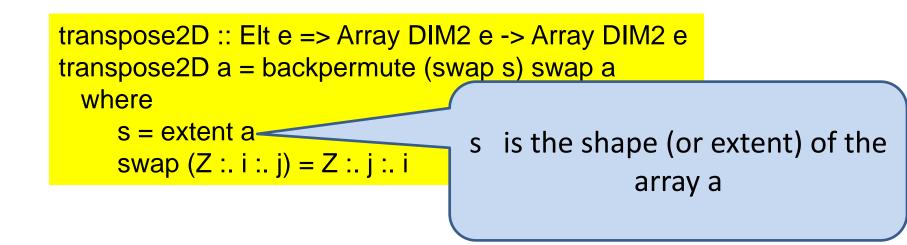


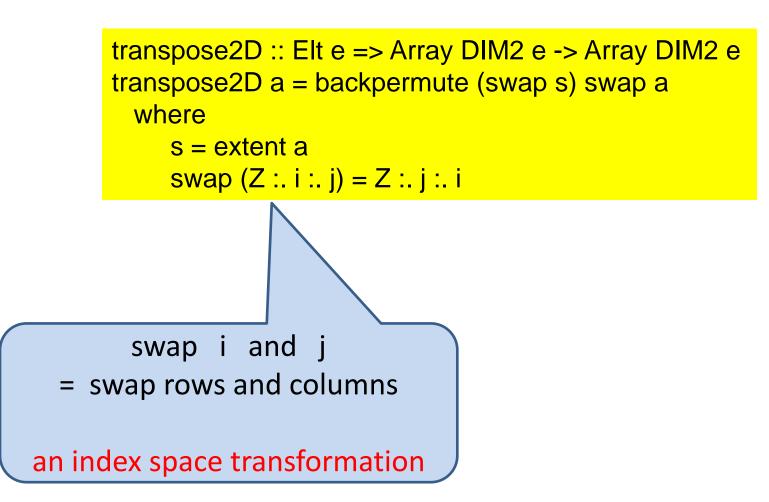
\*Main> let z = fromList (Z :. 2 :. 3 :: DIM2) [1..6] \*Main> transpose2D z Array (Z :. 3 :. 2) [1.0,4.0,2.0,5.0,3.0,6.0]

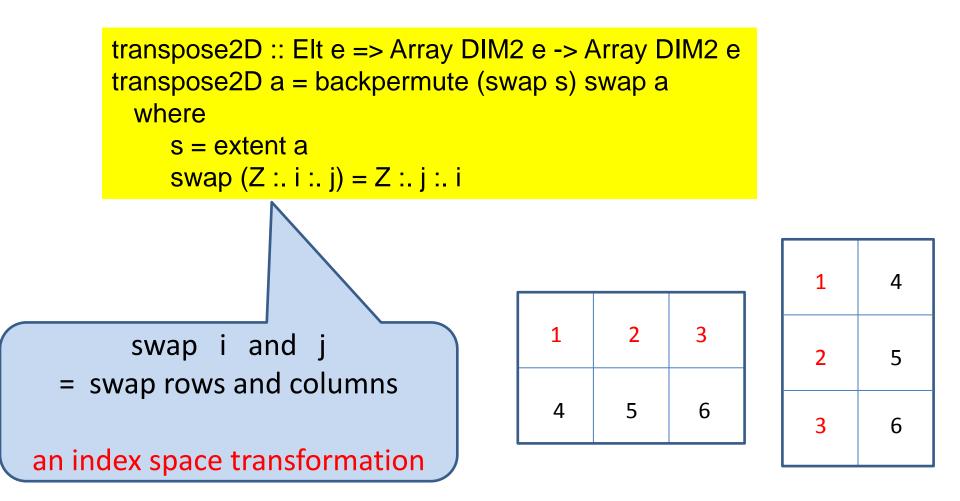
1	2	3
4	5	6

1	4
2	5
3	6

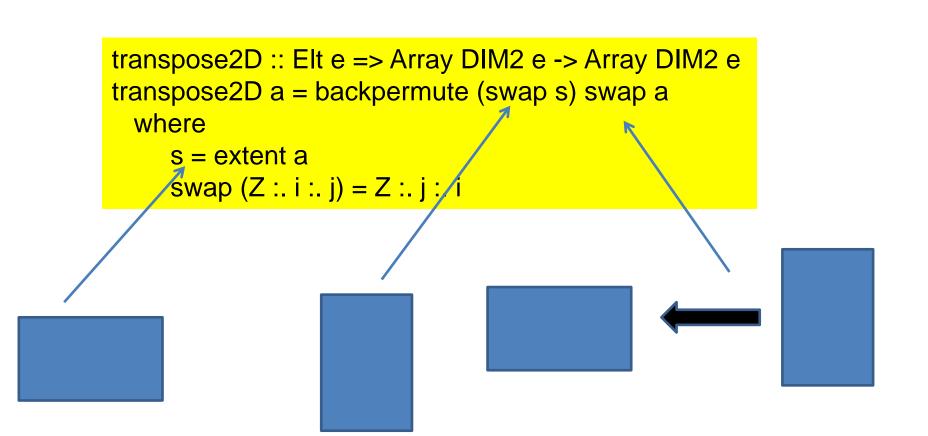
```
transpose2D :: Elt e => Array DIM2 e -> Array DIM2 e
transpose2D a = backpermute (swap s) swap a
where
    s = extent a
    swap (Z :. i :. j) = Z :. j :. i
```







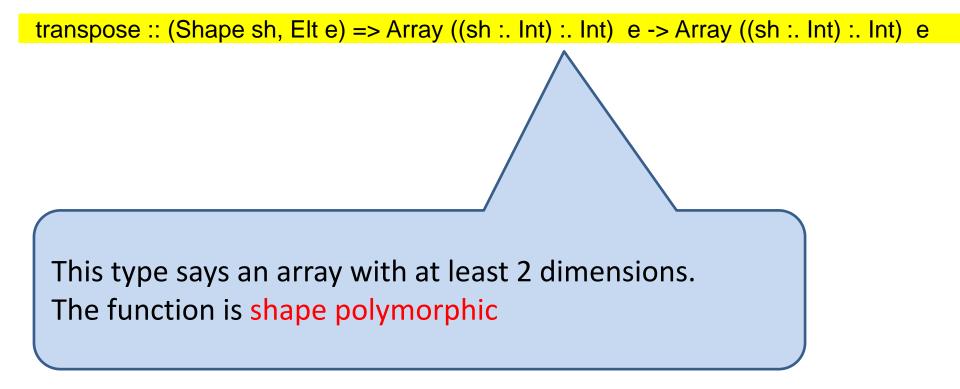
backpermute :: (Shape shin, Shape shout, Elt a) => shout -> (shout -> shin) -> Array shin a -> Array shout a



# more general transpose (on inner two dimensions)

transpose :: (Shape sh, Elt e) => Array ((sh :. Int) :. Int) e -> Array ((sh :. Int) :. Int) e

#### more general transpose (on inner two dimensions) is provided



#### more general transpose (on inner two dimensions) is provided

transpose :: (Shape sh, Elt e) => Array ((sh :. Int) :. Int) e -> Array ((sh :. Int) :. Int) e

Functions with at-least constraints become a parallel map over the unspecified dimensions (called rank generalisation)

Important way to express parallel patterns

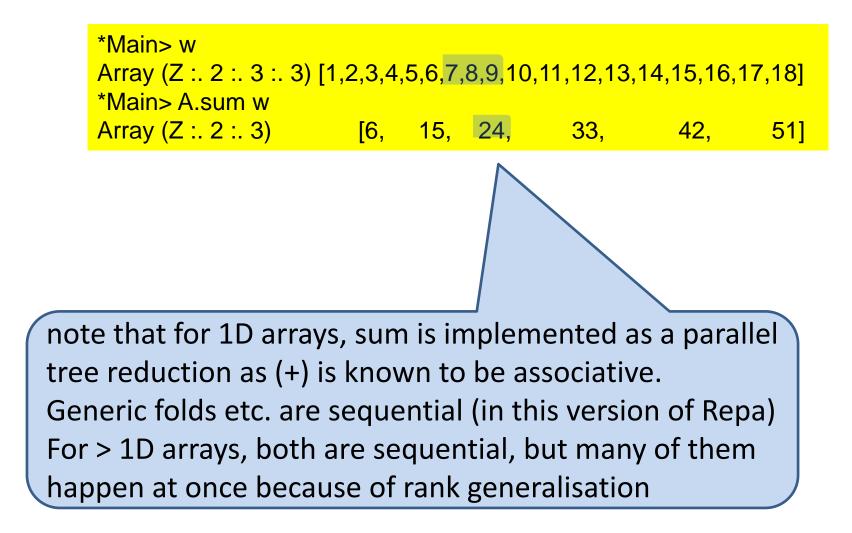
\*Main> let w = fromList (Z :. 2 :. 3 :. 3 :: DIM3) [1..(18 ::Int)]

\*Main> A.transpose w Array (Z :. 2 :. 3 :. 3) [1,4,7,2,5,8,3,6,9, 10,13,16,11,14,17,12,15,18]

# A.sum :: (Shape sh, Elt a, Num a) => Array (sh :. Int) a -> Array sh a

reduces shape by one dimension

\*Main> w Array (Z :. 2 :. 3 :. 3) [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18] \*Main> A.sum w Array (Z :. 2 :. 3) [6, 15, 24, 33, 42, 51]



#### backpermute can change the shape

```
selEven :: (Shape sh, Elt e) => Array (sh:.Int) e -> Array (sh:.Int) e
{-# INLINE selEven #-}
selEven !arr = force $ backpermute new_shape expand arr
where
    (ns :.n) = extent arr
    new_shape = ns :.((n+1) `div` 2)
    expand (is :.i) = is :.(i * 2)
```

### backpermute can change the shape

```
selEven :: (Shape sh, Elt e) => Array (sh:.Int) e -> Array (sh:.Int) e
{-# INLINE selEven #-}
selEven !arr = force $ backpermute new_shape expand arr
where
    (ns :.n) = extent arr
    new_shape = ns :.((n+1) `div` 2)
    expand (is :.i) = is :.(i * 2)
```

Note how the new shape depends only on the old shape and not on the data in the array (My def. differs slightly from that in the paper.)

#### backpermute can change the shape

```
selEven :: (Shape sh, Elt e) => Array (sh:.Int) e -> Array (sh:.Int) e
{-# INLINE selEven #-}
selEven !arr = force $ backpermute new_shape expand arr
 where
  (ns :.n) = extent arr
  new shape = ns :.((n+1) div 2)
  expand (is :.i) = is :.(i * 2)
selOdd :: (Shape sh, Elt e) => Array (sh:.Int) e -> Array (sh:.Int) e
{-# INLINE selOdd #-}
selOdd !arr = force $ backpermute new_extent expand arr
 where
  (ns :.n) = extent arr
  new_extent = ns :.(n `div` 2)
  expand (is :.i) = is :.(i * 2 + 1)
```

```
*Main> let w = fromList (Z :. 2 :. 3 :. 3 :: DIM3) [1..(18 ::Int)]
*Main> selEven w
Array (Z :. 2 :. 3 :. 2) [1,3,4,6,7,9,10,12,13,15,16,18]
*Main> selOdd w
Array (Z :. 2 :. 3 :. 1) [2,5,8,11,14,17]
```

# filter?

filter :: (Elt e) => (E -> Bool) -> Array DIM1 e -> Array DIM1 e

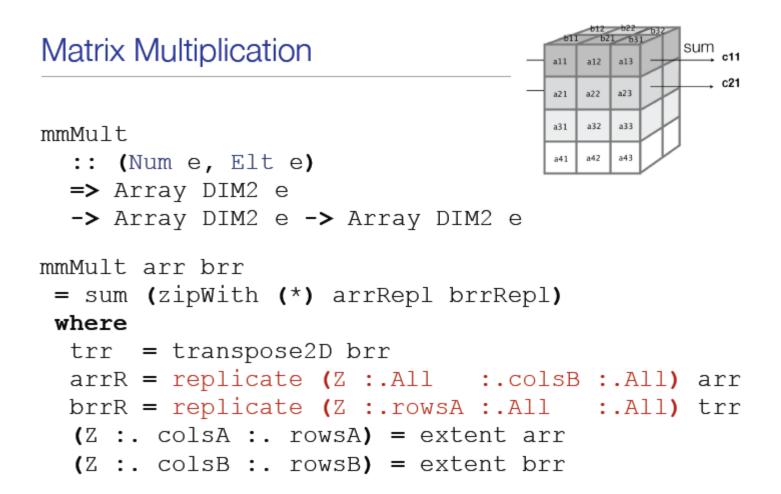
can't be shape polymorphic

the shape of the output depends on the value of the input

filtering rows in a matrix might give different lengths (but we only deal with rectangular arrays)

#### Matrix Multiplication $(A.B)_{i,j} = \sum_{k=1}^{n} A_{i,k} \cdot B_{k,j}$

<b>a</b> 11	<b>a</b> 12	<b>a</b> 13					<b>c</b> <sub>11</sub>	<b>c</b> <sub>12</sub>
3	3	a		$b_{11}$	$\mathbf{b}_{12}$	=		
<b>a</b> 21	<b>u</b> 22	<b>a</b> 23	•	<b>b</b> <sub>21</sub>	<b>b</b> 22		<b>C</b> 21	C22
<b>a</b> <sub>31</sub>	<b>a</b> 32	<b>a</b> 33					$c_{31}$	<b>C</b> <sub>32</sub>
			1	<b>D</b> 31	<b>b</b> <sub>32</sub>		<b>C</b> 41	<b>C</b> <sub>42</sub>
$a_{41}$	$a_{42}$	<b>a</b> 43					<b>U</b> 41	<b>C</b> 42



#### Fusion

- It's nice to program with bulk operations
   .. but we usually want them to be fused.
- We imagine replicating the source arrays being replicated when writing the program, but we don't want this at runtime.
- Fusion eliminates the intermediate arrays and the corresponding memory traffic.

Manifest and Delayed Arrays

```
data Array sh e
  = Manifest sh (UArr e)
  | Delayed sh (sh -> e)
```

- Manifest wraps a bona-fide unboxed array. Bulk-strict semantics. Forcing one element forces them all.
- Delayed wraps an element producing function, perhaps an index transformation that references some other array.
- Delayed functions are inlined and fused by the existing GHC optimiser (and lots of rewrite rules).



data Array = Manifo | Delayo

 Manifest wraps a bor Bulk-strict semantics.

worker-wrapper transformation, hoisting etc.

End up with the index transformations nicely composed

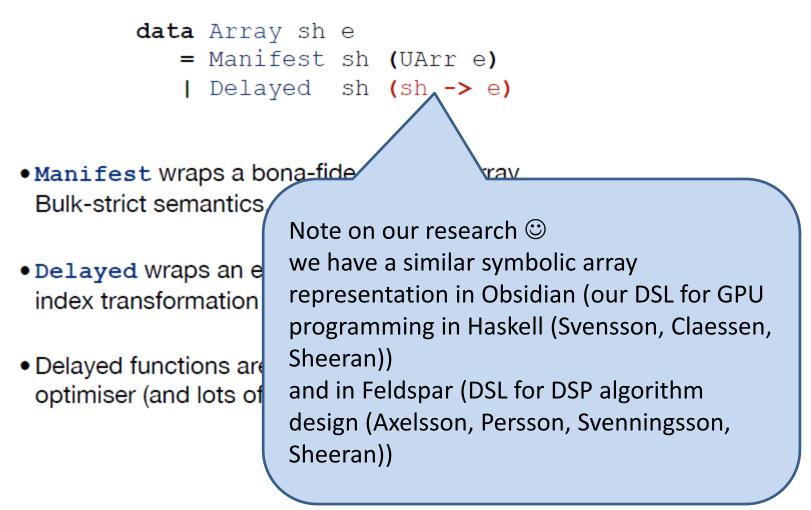
This is what gives tight loops in the resulting code (and good performance)

• Delayed wraps an element index transformation that rec

some other array.

 Delayed functions are inlined and fused by the existing GHC optimiser (and lots of rewrite rules).

#### Manifest and Delayed Arrays



#### Sharing and force

```
let arr = ...
brr = map f arr
in mmMult brr brr
```

#### Sharing and force

data Array sh e
= Manifest sh (UArr e)
| Delayed sh (sh -> e)

let arr = ...
brr = map f arr
in mmMult brr brr

#### Sharing and force

```
force :: Array sh e
    -> Array sh e
    data Array sh e
    = Manifest sh (UArr e)
    | Delayed sh (sh -> e)
let arr = ...
    brr = force (map f arr)
in mmMult brr brr
```

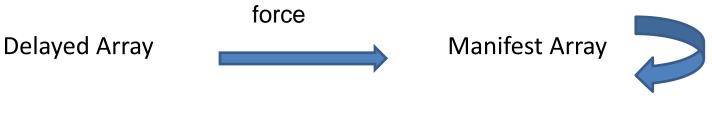
- For Manifest arrays, force is the identity.
- For Delayed arrays, it evaluates all the elements in parallel, producing a manifest array.
- The programmer must add force manually.

#### force :: (Shape sh, Elt a) => Array sh a -> Array sh a



#### evaluate all elements in parallel

#### force :: (Shape sh, Elt a) => Array sh a -> Array sh a

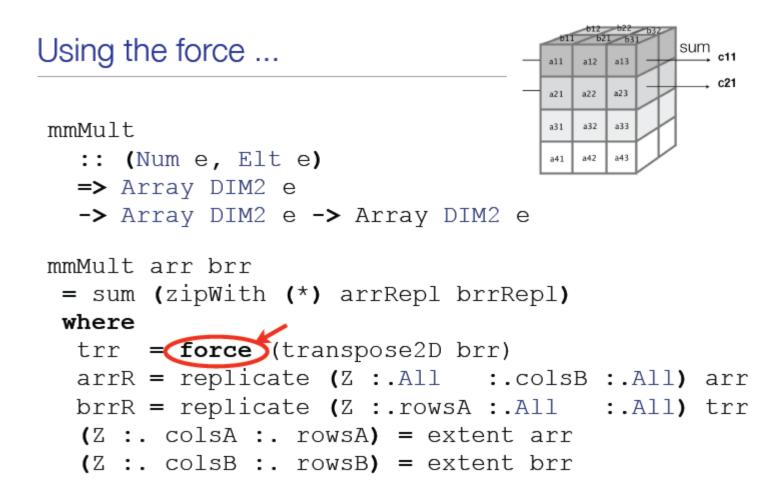


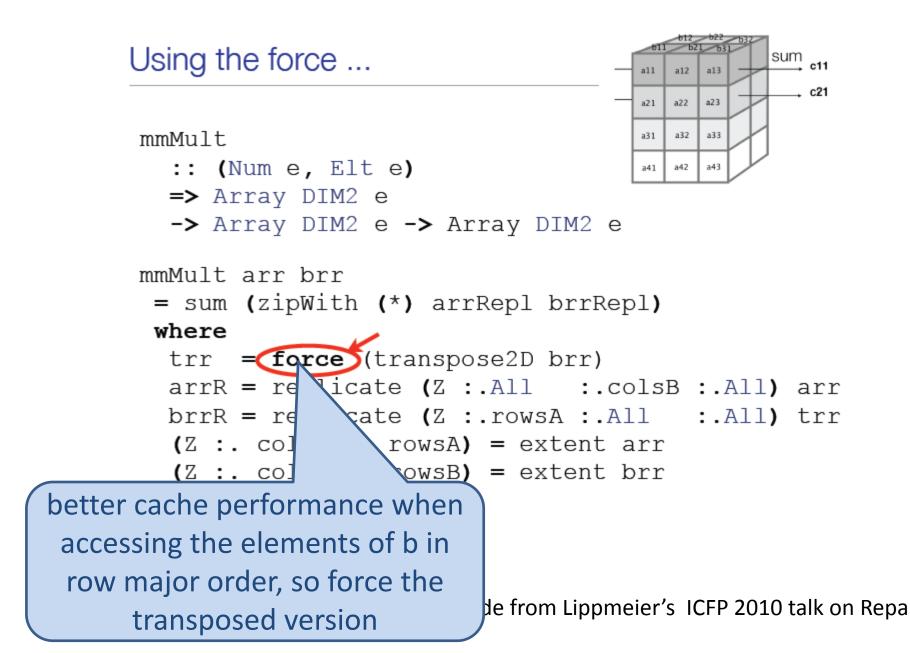
evaluate all elements in parallel

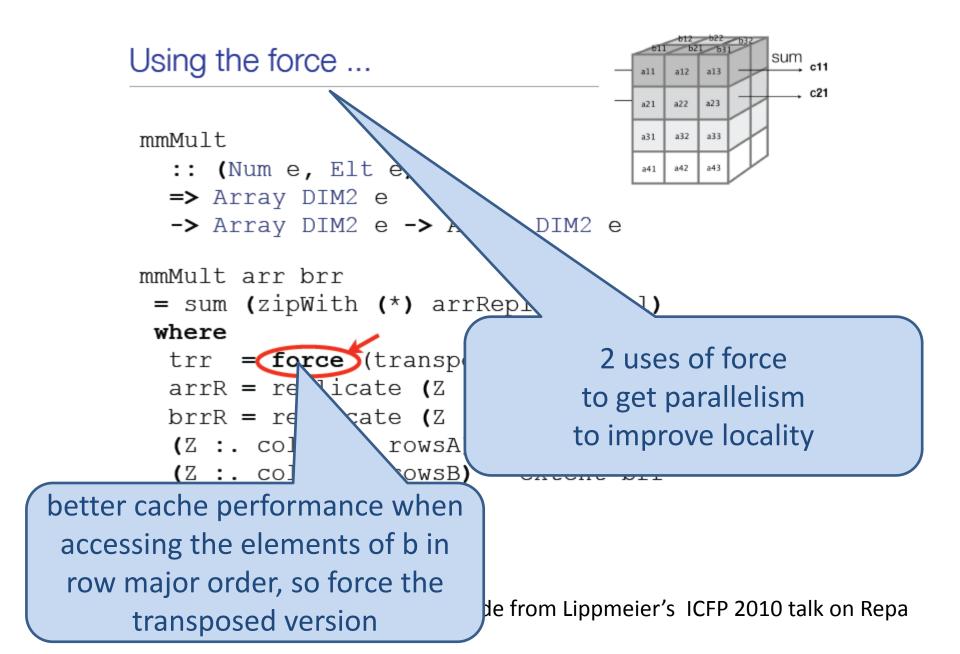
#### force :: (Shape sh, Elt a) => Array sh a -> Array sh a

**Delayed Array** 

if you index into a delayed array without forcing it first, then each indexing operation costs a function call. It also *recomputes* the value of the array element at that index.







## prescan in Repa (clear version)

## same prescan in Repa (my fastest so far)

```
-- assumes input of length a power of 2

prescan :: (Elt a) => (a -> a -> a) -> a -> (Array (Z :. Int) a) -> (Array (Z :. Int) a)

{-# INLINE prescan #-}

prescan f !i !as = sc as

where

sc as | size (extent as) == 1 = force $ fromList (Z :. (1 :: Int)) [i]

sc as | otherwise =

let es = force $ selEven as

os = force $ selEven as

ss = force $ selOdd as

ss = force $ sc (A.zipWith f es os)

in as `deepSeqArray` interleave2M ss (A.zipWith f ss es)
```

5 or 6 times faster for sumAll . prescan (+) (0::Int) on 2^20 inputs still 3-4 times slower than scanl1 ☺ but good speedup on 2 cores -N4 and hopefully on more

#### more operations

map :: (Shape sh, Elt a, Elt b) = (a -> b) -> Array sh a -> Array sh b

Doesn't care about shape of array. Just applies the function to each element.

plain Haskell foldl :: (a -> b -> a) -> a -> [b] -> a

Repa foldl :: (Shape sh, Elt a, Elt b) => (a -> b -> a) -> a -> Array (sh :. Int) b -> Array sh a

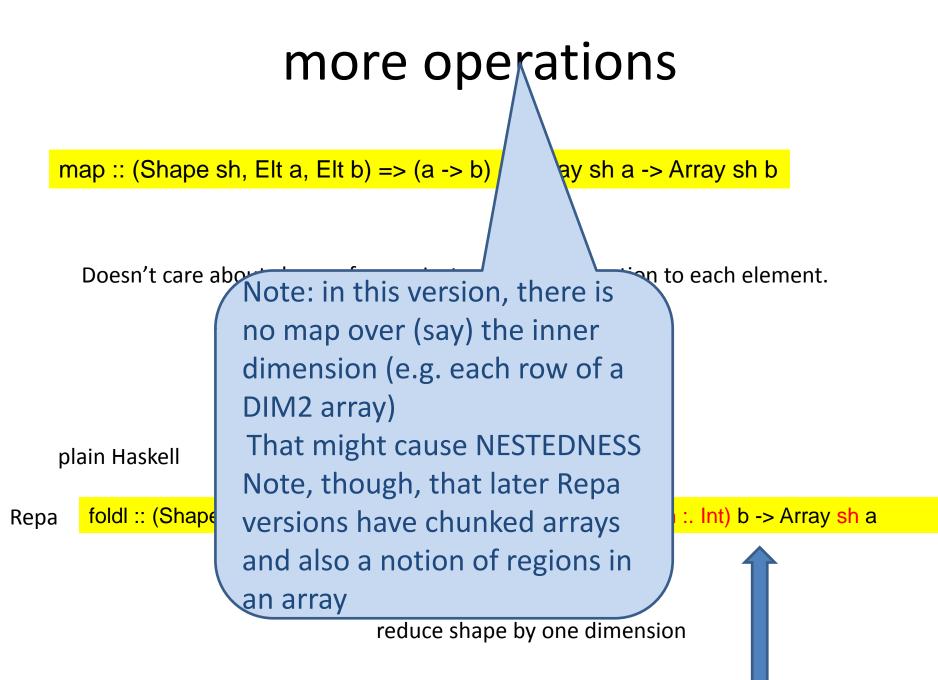
reduce shape by one dimension

```
*Main> let y = fromList ((Z :. 2 :. 3 :. 3) :: DIM3) [1..18]
*Main> A.fold (+) 0 y
Array (Z :. 2 :. 3) [6.0,15.0,24.0,33.0,42.0,51.0]
```

```
*Main> A.transpose y
Array (Z :. 2 :. 3 :. 3)
[1.0,4.0,7.0,2.0,5.0,8.0,3.0,6.0,9.0,10.0,13.0,16.0,11.0,14.0,17.0,12.0,15.0,18.0]
```

```
*Main> A.fold (+) 0 (A.transpose y)
Array (Z :. 2 :. 3) [12.0,15.0,18.0,39.0,42.0,45.0]
```

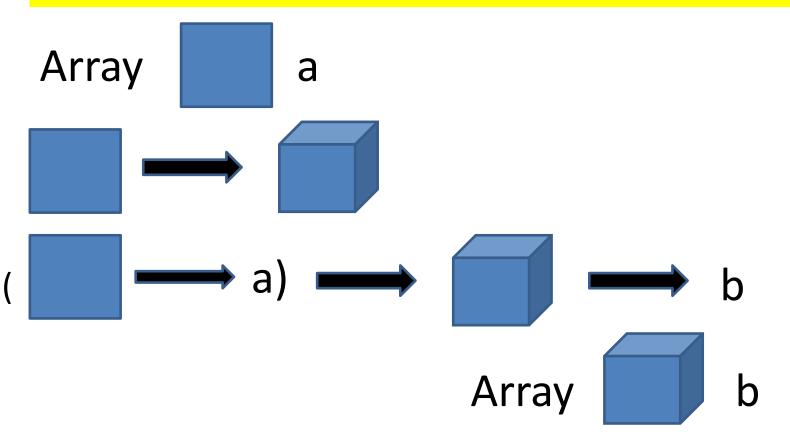
each fold is sequential, but they are all done at once



traverse

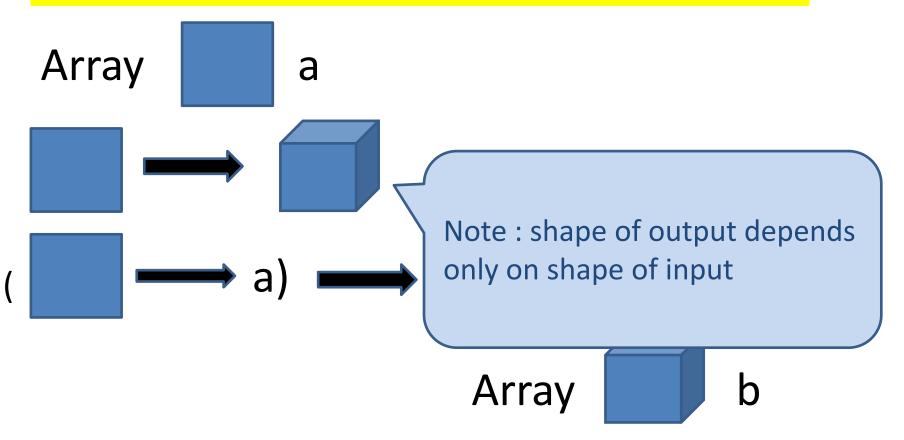
:: (Shape sh', Shape sh, Elt a) =>

Array shin a -> (shin -> shout) -> ((shin -> a) -> shout -> b) -> Array shout b



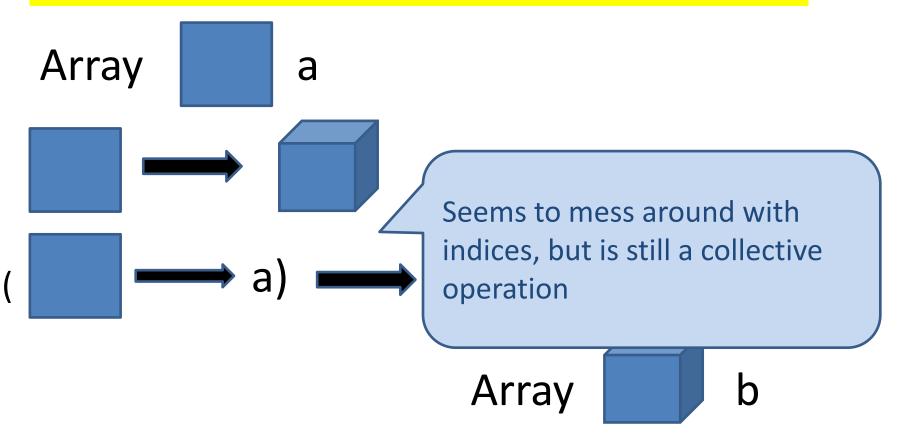
traverse

- :: (Shape sh', Shape sh, Elt a) =>
  - Array shin a -> (shin -> shout) -> ((shin -> a) -> shout -> b) -> Array shout b



traverse

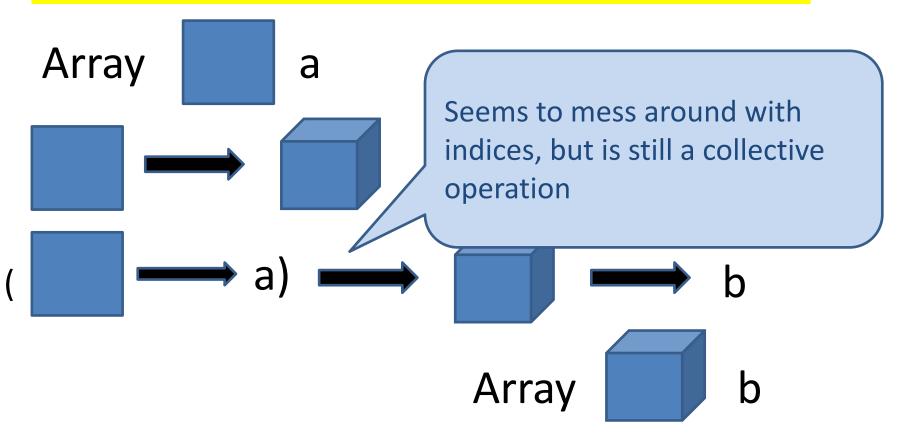
- :: (Shape sh', Shape sh, Elt a) =>
  - Array shin a -> (shin -> shout) -> ((shin -> a) -> shout -> b) -> Array shout b



traverse

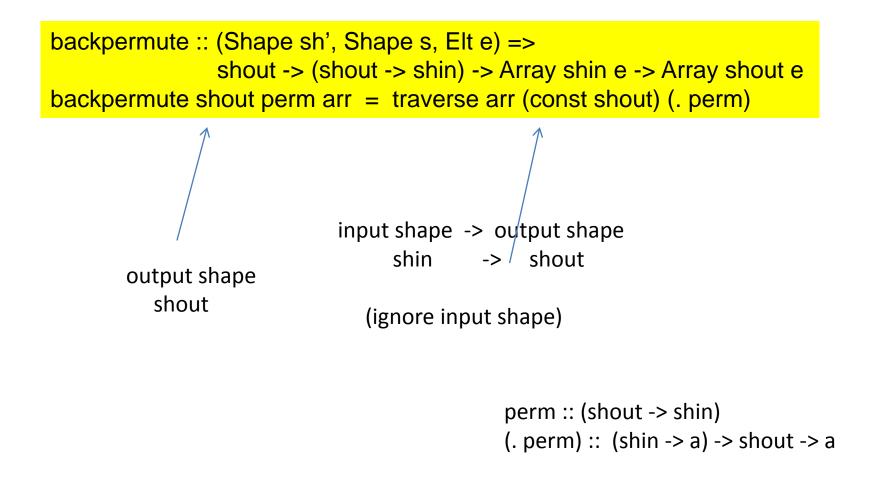
:: (Shape sh', Shape sh, Elt a) =>

Array shin a -> (shin -> shout) -> ((shin -> a) -> shout -> b) -> Array shout b



there is also a version called unsafeTraverse that skips bounds checking and so is faster

## use of traverse



### use of traverse

A.map :: (Shape sh, Elt b, Elt a) => (a -> b) -> Array sh a -> Array sh b

map f arr = traverse arr id (f .)

### unsafeTraverse

## unsafeTraverse

```
{-# INLINE bfly #-}
bfly !k !as
 = unsafeTraverse as id (f(s :. i) \rightarrow let a = f(s :. i)
                                               b = f(s :. (flipBit i k))
                                            in if (testBit i k) then (b-a) else (a+b))
{-# INLINE twids #-}
twids !k !as
  = let k2 = 2<sup>k</sup>
        k2' = 2*k2 in
    unsafeTraverse as id (f(s:.i) \rightarrow let a = f(s:.i))
                                                t = tw (i `mod` k2) k2'
                                            in if (testBit i k) then t*a else a)
```

```
{-# INLINE interleave2M #-}
interleave2M arr1 arr2
= arr1 `deepSeqArray` arr2 `deepSeqArray`
 unsafeTraverse2 arr1 arr2 shapeFn elemFn
where
         shapeFn dim1 dim2
         | sh :. len1 <- dim1
         , sh :. len2 <- dim2
         = sh :. (len1 + len2)
         elemFn get1 get2 (sh :. ix)
         = case ix `mod` 2 of
                   -> get1 (sh :. ix `div` 2)
                  0
                          -> get2 (sh :. ix `div` 2)
                  1
```

```
*Main> let w = fromList (Z :. 2 :. 3 :. 3 :: DIM3) [1..(18 ::Int)]
*Main> selEven w
Array (Z :. 2 :. 3 :. 2) [1,3,4,6,7,9,10,12,13,15,16,18]
*Main> selOdd w
Array (Z :. 2 :. 3 :. 1) [2,5,8,11,14,17]
```

```
*Main> interleave2 (selEven w) (selOdd w)
Array (*** Exception: Data.Array.Repa.interleave2: arrays must
have same extent
*Main> interleave2M (selEven w) (selOdd w)
Array (Z :. 2 :. 3 :. 3)
[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18]
```

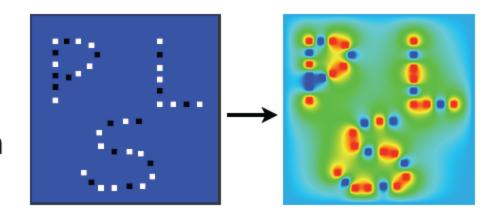
## unsafeTraverse

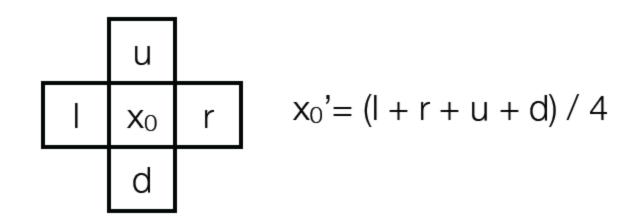
```
{-# INLINE bfly #-}
bfly !k !as
 = unsafeTraverse as id (f(s :. i) \rightarrow let a = f(s :. i)
                                              b = f(s :. (flipBit i k))
                                           in if (testBit i k) then (b-a) else (a+b))
{-# INLINE twids #-}
twids lk las
  = let k^2 = 2^k
        k2' = 2*k2 in
    unsafeTraverse as id (f(s:.i) \rightarrow let a = f(s:.i))
                                               t = tw (i `mod` k2) k2'
                                           in if (testBit i k) then t*a else a)
```

{-# INLINE fft4 #-} fft4 !n !as = foldr1 (.) [ force . twids k . bfly k | k <- [0..(n-1)]] as

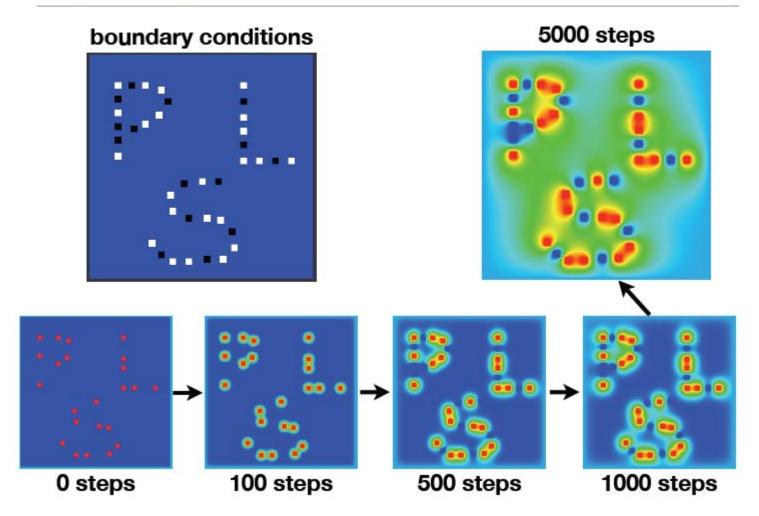
#### **Example Applications**

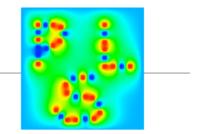
Solving the Laplace Equation





#### Laplace Equation



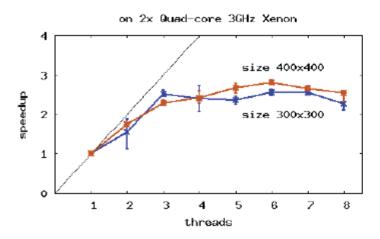


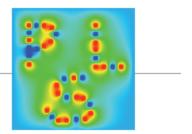
```
Laplace Equation
```

stencil :: Array DIM2 Double

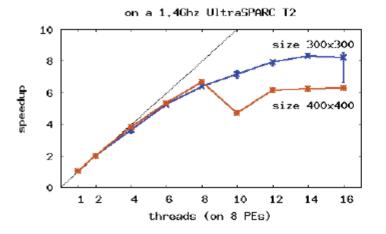
```
-> Array DIM2 Double
stencil arr
= traverse arr id update
where
  :. height :. width = extent arr
  update get d@(sh :. i :. j)
   = if isBoundary i j
      then get d
      else (get (sh :. (i-1)) :. j)
         + get (sh :. i :. (j-1))
         + get (sh :. (i+1)) :. j)
         + get (sh :. i :. (j+1))) / 4
```

#### Laplace Equation





	GCC	single thread	fastest parallel
Xenon	0.70	1.7s	0.68s
т2	6.5s	32s	3.8s



- GHC native code generator does no instruction reordering on SPARC. No LLVM 'port.
- Single threaded on T2 is slow

## Conclusions

Based on DPH technology

Good speedups!

Neat programs

Good control of Parallelism

BUT CACHE AWARENESS needs to be tackled

# Next lecture (Monday)

I would like to have a couple of student talks next wednesday (having talked to a couple of you earlier). Please contact to me again to confirm!

Student talks on topics related to the course would be most welcome!