

An Overview of Chapel: a productive parallel programming language

Sung-Eun Choi

Cray, Inc.

KIISE-KOCSEA HPC SIG Joint Workshop @ SC'12

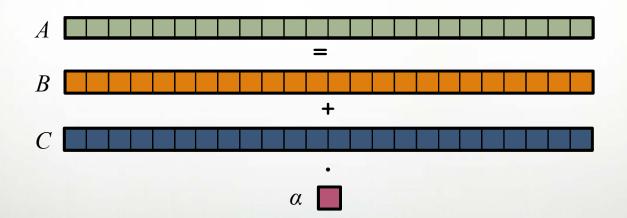




Given: *m*-element vectors *A*, *B*, *C*

Compute: $\forall i \in 1..m, A_i = B_i + \alpha \cdot C_i$

In pictures:



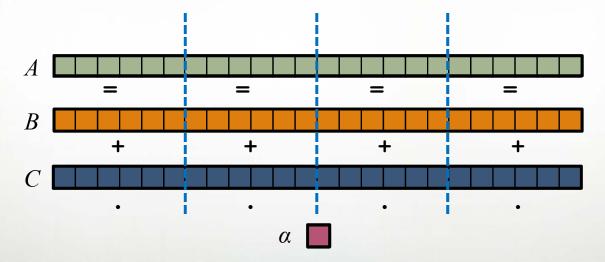




Given: *m*-element vectors *A*, *B*, *C*

Compute: $\forall i \in 1..m, A_i = B_i + \alpha \cdot C_i$

In pictures, in parallel:



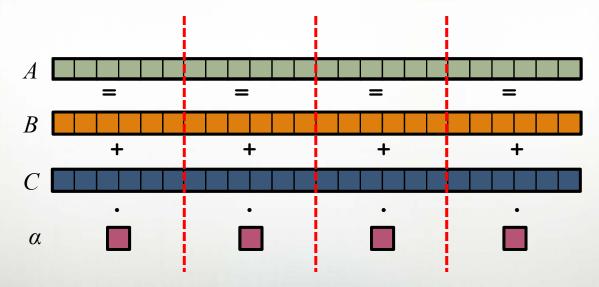




Given: *m*-element vectors *A*, *B*, *C*

Compute: $\forall i \in 1..m, A_i = B_i + \alpha \cdot C_i$

In pictures, in parallel (distributed memory):



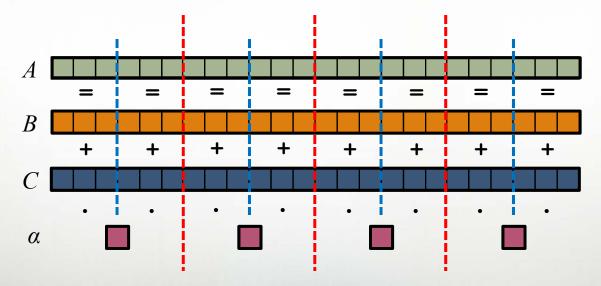




Given: *m*-element vectors *A*, *B*, *C*

Compute: $\forall i \in 1..m, A_i = B_i + \alpha \cdot C_i$

In pictures, in parallel (distributed memory multicore):





STREAM Triad: MPI

#include <hpcc.h>





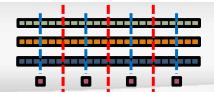
```
MPI
```

```
static int VectorSize;
static double *a, *b, *c;
int HPCC_StarStream(HPCC_Params *params) {
 int myRank, commSize;
 int rv, errCount;
 MPI Comm comm = MPI COMM WORLD;
 MPI_Comm_size( comm, &commSize );
 MPI Comm rank( comm, &myRank );
 rv = HPCC_Stream( params, 0 == myRank);
 MPI_Reduce( &rv, &errCount, 1, MPI_INT, MPI_SUM,
   0, comm );
 return errCount;
int HPCC_Stream(HPCC_Params *params, int doIO) {
 register int j;
 double scalar;
 VectorSize = HPCC LocalVectorSize( params, 3,
   sizeof(double), 0 );
 a = HPCC_XMALLOC( double, VectorSize );
 b = HPCC XMALLOC( double, VectorSize );
 c = HPCC_XMALLOC( double, VectorSize );
```

```
if (!a || !b || !c) {
  if (c) HPCC free(c);
  if (b) HPCC free(b);
  if (a) HPCC_free(a);
  if (doIO) {
    fprintf( outFile, "Failed to allocate memory
  (%d).\n", VectorSize );
    fclose( outFile );
  return 1;
for (j=0; j<VectorSize; j++) {</pre>
  b[j] = 2.0;
  c[j] = 0.0;
scalar = 3.0;
for (j=0; j<VectorSize; j++)</pre>
  a[j] = b[j]+scalar*c[j];
HPCC free(c);
HPCC free(b);
HPCC_free(a);
return 0;
```



STREAM Triad: MPI+OpenMP





MPI + OpenMP

```
#include <hpcc.h>
                                                       if (!a || !b || !c) {
#ifdef OPENMP
                                                          if (c) HPCC free(c);
#include <omp.h>
                                                         if (b) HPCC free(b);
#endif
                                                         if (a) HPCC_free(a);
static int VectorSize;
                                                         if (doI0) {
                                                            fprintf( outFile, "Failed to allocate memory
static double *a, *b, *c;
                                                         (%d).\n", VectorSize );
int HPCC_StarStream(HPCC_Params *params) {
                                                            fclose( outFile );
 int myRank, commSize;
 int rv, errCount;
                                                         return 1;
 MPI Comm comm = MPI COMM WORLD;
 MPI Comm size ( comm, &commSize );
                                                      #ifdef OPENMP
 MPI Comm rank( comm, &myRank );
                                                      #pragma omp parallel for
                                                      #endif
 rv = HPCC_Stream( params, 0 == myRank);
                                                       for (j=0; j<VectorSize; j++) {</pre>
 MPI_Reduce( &rv, &errCount, 1, MPI_INT, MPI_SUM,
                                                         b[j] = 2.0;
   0, comm );
                                                         c[i] = 0.0;
 return errCount;
                                                       scalar = 3.0;
int HPCC_Stream(HPCC_Params *params, int doIO) {
                                                      #ifdef _OPENMP
 register int j;
                                                      #pragma omp parallel for
 double scalar;
                                                      #endif
                                                       for (j=0; j<VectorSize; j++)</pre>
 VectorSize = HPCC LocalVectorSize( params, 3,
                                                         a[j] = b[j]+scalar*c[j];
   sizeof(double), 0 );
                                                       HPCC free(c);
 a = HPCC XMALLOC( double, VectorSize );
                                                       HPCC free(b);
 b = HPCC XMALLOC( double, VectorSize );
                                                       HPCC_free(a);
  c = HPCC_XMALLOC( double, VectorSize );
                                                       return 0;
```





STREAM Triad: MPI+OpenMP vs. CUDA

MPI + OpenMP

```
#include <hpcc.h>
#ifdef _OPENMP
#include <omp.h>
#endif
static int VectorSize;
static double *a, *b, *c;
int HPCC StarStream(HPCC Params *params)
 int myRank, commSize;
 int rv, errCount;
 MPT Comm comm = MPT COMM WORLD:
 MPI_Comm_size( comm, &commSize );
 MPI_Comm_rank( comm, &myRank );
 rv = HPCC_Stream( params, 0 == myRank);
 MPI_Reduce( &rv, &errCount, 1, MPI_INT, MPI_SUM, 0, comm );
 return errCount;
int HPCC_Stream(HPCC_Params *params, int doIO) {
 register int i;
 double scalar;
 VectorSize = HPCC_LocalVectorSize( params, 3, sizeof(double), 0 );
 a = HPCC_XMALLOC( double, VectorSize );
 b = HPCC XMALLOC( double, VectorSize );
 c = HPCC_XMALLOC( double, VectorSize );
 if (!a || !b || !c) {
   if (c) HPCC_free(c);
   if (b) HPCC_free(b);
   if (a) HPCC_free(a);
   if (doIO) {
     fprintf( outFile, "Failed to allocate memory (%d).\n", VectorSize );
     fclose( outFile );
   return 1;
#ifdef OPENMP
#pragma omp parallel for
#endif
 for (j=0; j<VectorSize; j++) {
   b[i] = 2.0;
   c[j] = 0.0;
 scalar = 3.0;
#ifdef OPENMP
#pragma omp parallel for
#endif
 for (j=0; j<VectorSize; j++)</pre>
  a[j] = b[j]+scalar*c[j];
 HPCC free(c);
 HPCC_free(b);
 HPCC_free(a);
 return 0;
```

CUDA

```
2000000
#define N
int main() {
  float *d_a, *d_b, *d_c;
  float scalar;
  cudaMalloc((void**)&d a, sizeof(float)*N);
  cudaMalloc((void**)&d b, sizeof(float)*N);
  cudaMalloc((void**)&d c, sizeof(float)*N);
 dim3 dimBlock(128);
 dim3 dimGrid(N/dimBlock.x );
 if( N % dimBlock.x != 0 ) dimGrid.x+=1;
 set_array<<<dimGrid,dimBlock>>>(d_b, .5f, N);
 set_array<<<dimGrid,dimBlock>>>(d_c, .5f, N);
  scalar=3.0f;
 STREAM_Triad<<<dimGrid,dimBlock>>>(d_b, d_c, d_a, scalar, N);
  cudaThreadSynchronize();
  cudaFree(d_a);
 cudaFree(d b);
  cudaFree(d_c);
global void set array(float *a, float value, int len) {
 int idx = threadIdx.x + blockIdx.x * blockDim.x;
  if (idx < len) a[idx] = value;
_global__ void STREAM_Triad( float *a, float *b, float *c,
                              float scalar, int len) {
 int idx = threadIdx.x + blockIdx.x * blockDim.x;
 if (idx < len) c[idx] = a[idx]+scalar*b[idx];</pre>
```



STREAM Triad: MPI+OpenMP vs. CUDA

MPI + OpenMP

```
#include <hpcc.h>
#ifdef _OPENMP
#include <omp.h>
#endif
static int VectorSize;
static double *a, *b, *c;
int HPCC StarStream(HPCC Params *params)
 int myRank, commSize;
 int rv, errCount;
 MPT Comm comm = MPT COMM WORLD:
 MPI_Comm_size( comm, &commSize );
 MPI_Comm_rank( comm, &myRank );
 rv = HPCC_Stream( params, 0 == myRank);
 MPI_Reduce( &rv, &errCount, 1, MPI_INT, MPI_SUM, 0, comm );
```

CUDA

```
#define N
                2000000
int main() {
 float *d_a, *d_b, *d_c;
 float scalar;
 cudaMalloc((void**)&d a, sizeof(float)*N);
 cudaMalloc((void**)&d b, sizeof(float)*N);
 cudaMalloc((void**)&d c, sizeof(float)*N);
```

HPC suffers from too many distinct notations for expressing parallelism and locality

```
register int is
 double scalar;
 VectorSize = HPCC_LocalVectorSize( params, 3, sizeof(double), 0 );
 a = HPCC_XMALLOC( double, VectorSize );
 b = HPCC XMALLOC( double, VectorSize );
 c = HPCC_XMALLOC( double, VectorSize );
 if (!a || !b || !c) {
   if (c) HPCC_free(c);
   if (b) HPCC_free(b);
   if (a) HPCC_free(a);
   if (doIO) {
     fprintf( outFile, "Failed to allocate memory (%d).\n", VectorSize );
     fclose( outFile );
   return 1;
#ifdef OPENMP
#pragma omp parallel for
#endif
 for (j=0; j<VectorSize; j++) {
   b[i] = 2.0;
   c[j] = 0.0;
 scalar = 3.0;
#ifdef OPENMP
#pragma omp parallel for
#endif
 for (j=0; j<VectorSize; j++)
  a[j] = b[j]+scalar*c[j];
 HPCC free(c);
 HPCC_free(b);
 HPCC_free(a);
 return 0;
```

```
set_array<<<dimGrid,dimBlock>>>(d_b, .5f, N);
 set_array<<<dimGrid,dimBlock>>>(d_c, .5f, N);
 scalar=3.0f;
 STREAM_Triad<<<dimGrid,dimBlock>>>(d_b, d_c, d_a, scalar, N);
 cudaThreadSynchronize();
 cudaFree(d_a);
 cudaFree(d b);
 cudaFree(d_c);
global void set array(float *a, float value, int len) {
 int idx = threadIdx.x + blockIdx.x * blockDim.x;
 if (idx < len) a[idx] = value;
global void STREAM Triad( float *a, float *b, float *c,
                              float scalar, int len) {
 int idx = threadIdx.x + blockIdx.x * blockDim.x;
 if (idx < len) c[idx] = a[idx]+scalar*b[idx];</pre>
```



STREAM Triad: Chapel

```
MPI + OpenMP
                                                            Chapel
#include <hpcc.h>
#ifdef _OPENMP
#include <omp.h>
#endif
static int VectorSize;
                          config const m = 1000,
static double *a, *b, *c;
                                                  alpha = 3.0;
int HPCC StarStream(HPCC Params *par
 int myRank, commSize;
 int rv, errCount;
 MPI_Comm comm = MPI_COMM WORLD;
                                                                                                             the special
 MPI_Comm_size( comm, &commSize );
                          const ProblemSpace = {1..m} dmapped ...;
 MPI_Comm_rank( comm, &myRank );
 rv = HPCC_Stream( params, 0 == myF
                                                                                                             sauce
 MPI_Reduce( &rv, &errCount, 1, MPI
                          var A, B, C: [ProblemSpace] real;
 return errCount;
int HPCC_Stream(HPCC_Params *params
 register int i;
                          B = 2.0;
 double scalar;
                                                                                                         N);
 VectorSize = HPCC_LocalVectorSize(
                          C = 3.0;
                                                                                                         N);
 a = HPCC_XMALLOC( double, VectorS:
 b = HPCC XMALLOC( double, VectorS:
 c = HPCC_XMALLOC( double, VectorSi
                          A = B + alpha * C;
 if (!a || !b || !c) {
                                                                                                         _c, d_a, scalar, N);
  if (c) HPCC_free(c);
  if (b) HPCC_free(b);
  if (a) HPCC_free(a);
  if (doIO) {
------
                              ------
 scalar = 3.0;
#ifdef OPENMP
                                                        global void STREAM Triad( float *a, float *b, float *c,
#pragma omp parallel for
#endif
                                                                                         float scalar, int len) {
                                                         int idx = threadIdx.x + blockIdx.x * blockDim.x;
                                                         if (idx < len) c[idx] = a[idx]+scalar*b[idx];</pre>
```



STREAM Triad: Chapel

MPI + OpenMP Chapel #include <hpcc.h> #ifdef _OPENMP #include <omp.h> #endif config const m = 1000, static int VectorSize; static double *a, *b, *c; alpha = 3.0;int HPCC StarStream(HPCC Params *par int myRank, commSize; int rv, errCount; MPI Comm comm = MPI COMM WORLD; the special MPI_Comm_size(comm, &commSize); const ProblemSpace = {1..m} dmapped ...; MPI_Comm_rank(comm, &myRank); rv = HPCC_Stream(params, 0 == myF sauce MPI_Reduce(&rv, &errCount, 1, MPI var A, B, C: [ProblemSpace] real; return errCount; int HPCC_Stream(HPCC_Params *params register int i; B = 2.0;double scalar; N); VectorSize = HPCC_LocalVectorSize(C = 3.0;N); a = HPCC_XMALLOC(double, VectorS: b = HPCC XMALLOC(double, VectorS c = HPCC_XMALLOC(double, VectorS: A = B + alpha * C;if (!a | !b | !c) { _c, d_a, scalar, N); if (c) HPCC_free(c); if (b) HPCC_free(b); if (a) HPCC_free(a); if (doTO) Philosophy: Good language design can tease details of locality and #ifdef OP #pragma #endif parallelism away from an algorithm, permitting the application scientist and HPC expert to each focus on their strengths.

Outline



- ✓ Motivation
- ➤ Chapel Background and Themes
- Tour of Chapel Concepts
- Project Status



What is Chapel?



- An emerging parallel programming language
 - Design and development led by Cray Inc.
 - in collaboration with academia, labs, industry
 - Initiated under the DARPA HPCS program
- Overall goal: Improve programmer productivity
 - Improve the programmability of parallel computers
 - Match or beat the performance of current programming models
 - Support better portability than current programming models
 - Improve the robustness of parallel codes
- A work-in-progress







- Being developed as open source at SourceForge
- Licensed as BSD software

Target Architectures:

- Cray architectures
- multicore desktops and laptops
- commodity clusters
- systems from other vendors
- in-progress: CPU+accelerator hybrids, manycore, ...







- 1) General Parallel Programming
- 2) Global-View Abstractions
- 3) Multiresolution Design
- 4) Control over Locality/Affinity
- 5) Reduce HPC ↔ Mainstream Language Gap







- 1) General Parallel Programming
- 2) Global-View Abstractions
- 3) Multiresolution Design
- 4) Control over Locality/Affinity
- 5) Reduce HPC ↔ Mainstream Language Gap





1) General Parallel Programming

Recall from our STREAM example..

Style of HW Parallelism	Programming Model	Unit of Parallelism
Inter-node	MPI	executable
Intra-node/multicore	OpenMP	iteration/task
GPU/accelerator	CUDA	SIMD function/task





1) General Parallel Programming

With a unified set of concepts...

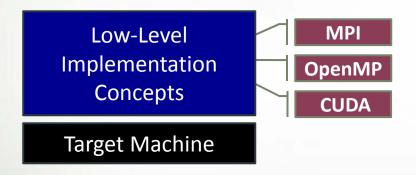
- ...express any parallelism desired in a user's program
 - Styles: data-parallel, task-parallel, concurrency, nested, ...
 - Levels: model, function, loop, statement, expression
- ...target all parallelism available in the hardware
 - Types: machines, nodes, cores, instructions

Style of HW Parallelism	Programming Model	Unit of Parallelism
Inter-node	Chapel	executable/task
Intra-node/multicore	Chapel	iteration/task
GPU/accelerator	Chapel	SIMD function/task





3) Multiresolution Design: Motivation



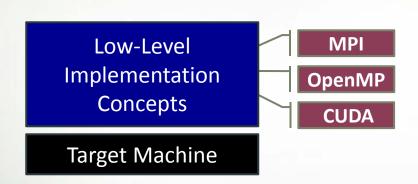
"Why is everything so tedious/difficult?"

"Why don't my programs port trivially?"





3) Multiresolution Design: Motivation



"Why is everything so tedious/difficult?"

"Why don't my programs port trivially?"



Target Machine

"Why don't I have more control?"



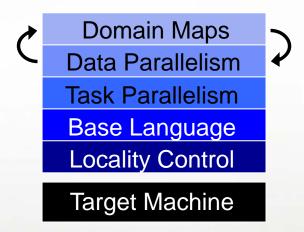




Multiresolution Design: Support multiple tiers of features

- higher levels for programmability, productivity
- lower levels for greater degrees of control

Chapel language concepts



- build the higher-level concepts in terms of the lower
- permit the user to intermix layers arbitrarily



Outline

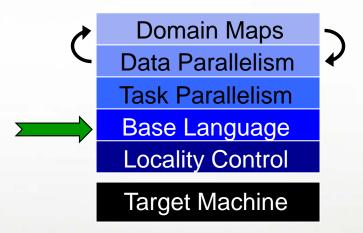


- ✓ Motivation
- ✓ Chapel Background and Themes
- > Tour of Chapel Concepts
- Project Status





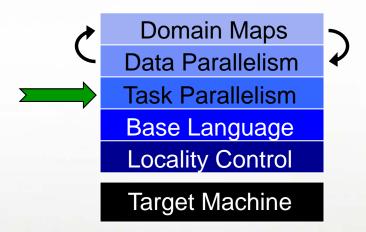
Base Language Features







Task Parallel Features







Task Parallelism Terminology

Task: a unit of (parallel) work in Chapel

- All parallelism is implemented using tasks
- main() is the only task when a program begins





Begin Statements

```
begin writeln("Hello from task 0 of 4");
begin writeln("Hello from task 1 of 4");
begin writeln("Hello from task 2 of 4");
begin writeln("Hello from task 3 of 4");
writeln("All tasks done");
```

```
Hello from task 0 of 4
All tasks done
Hello from task 1 of 4
Hello from task 3 of 4
Hello from task 2 of 4
```





Cobegin statements

```
cobegin {
  writeln("Hello from task 0 of 4);
  writeln("Hello from task 1 of 4);
  writeln("Hello from task 2 of 4);
  writeln("Hello from task 3 of 4);
}
writeln("All tasks done");
```

```
Hello from task 1 of 4
Hello from task 2 of 4
Hello from task 3 of 4
Hello from task 0 of 4
All tasks done
```







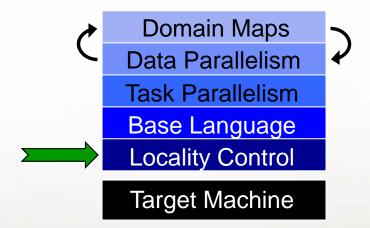
```
coforall t in 0..numTasks-1 do
  writeln("Hello from task ", t, " of ", numTasks);
writeln("All tasks done");
```

```
Hello from task 2 of 4
Hello from task 0 of 4
Hello from task 3 of 4
Hello from task 1 of 4
All tasks done
```















Definition:

- Abstract unit of target architecture
- Supports reasoning about locality
- Capable of running tasks and storing variables
 - i.e., has processors and memory
- Can be queried for characteristics like amount of memory or number of cores

Typically: A multi-core processor or an SMP





Controlling Locality

- Users specify the number of locales to use at program launch
 - The locale variables are available to the user in an built-in array called Locales
- On-clauses support placement of computations:

```
writeln("on locale 0");

on Locales[1] do
   writeln("now on locale 1");

writeln("on locale 0 again");
```



Important Note #1



Parallelism and Locality are distinct concepts

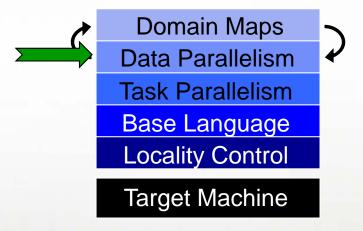
- e.g., begin statements create tasks
 - new task will run on the current locale
- e.g., on-clauses place computation
 - no parallelism introduced

Composing these concepts can be very powerful





Data Parallel Features

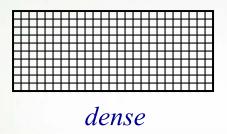


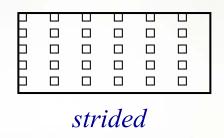


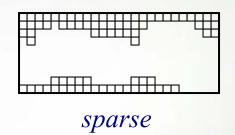




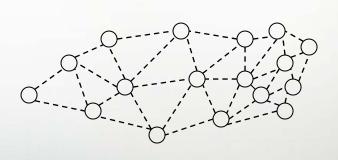
Chapel supports several types of domains (index sets):











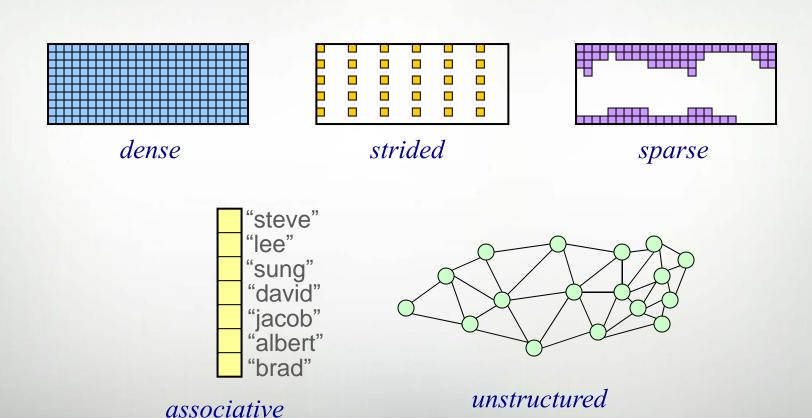
unstructured







Each domain type can be used to declare arrays:







Chapel Domain/Array Operations

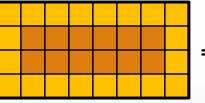
Parallel and Serial Iteration

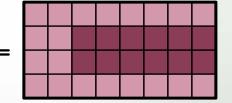
```
for a in A do a = 0.0;
forall (i,j) in D do A[i,j] = i + j/10.0;
```

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8
2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8
3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8
4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8

Array Slicing; Domain Algebra

$$A[InnerD] = B[InnerD+(0,1)];$$





Promotion of Scalar Operators and Functions

$$A = B + alpha * C;$$

$$A = \exp(B, C);$$

 And several others: indexing, reallocation, set operations, remapping, aliasing, queries, ...







Operations on arrays are the same regardless of the domain that is used to declare the arrays

$$A = B + alpha * C;$$

 A, B, and C could be dense, strided, sparse, associative, or unstructured



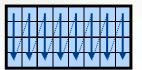


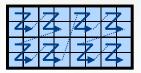
Data Parallelism Implementation Qs

Q1: How are arrays laid out in memory?

Are regular arrays laid out in row- or column-major order? Or...?







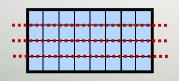


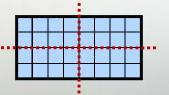
...?

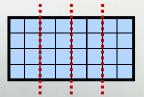
How are sparse arrays stored? (COO, CSR, CSC, block-structured, ...?)

Q2: How are arrays stored by the locales?

- Completely local to one locale? Or distributed?
- If distributed... In a blocked manner? cyclically? block-cyclically? recursively bisected? dynamically rebalanced? ...?









...?



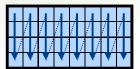


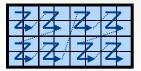
Data Parallelism Implementation Qs

Q1: How are arrays laid out in memory?

Are regular arrays laid out in row- or column-major order? Or...?









...?

How are sparse arrays stored? (COO, CSR, CSC, block-structured, ...?)

Q2: How are arrays stored by the locales?

- Completely local to one locale? Or distributed?
- If distributed... In a blocked manner? cyclically? block-cyclically? recursively bisected? dynamically rebalanced? ...?

A: Chapel's *domain maps* are designed to give the user full control over such decisions



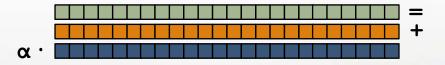


STREAM Triad: Chapel (multicore)

const ProblemSpace = {1..m};



var A, B, C: [ProblemSpace] real;

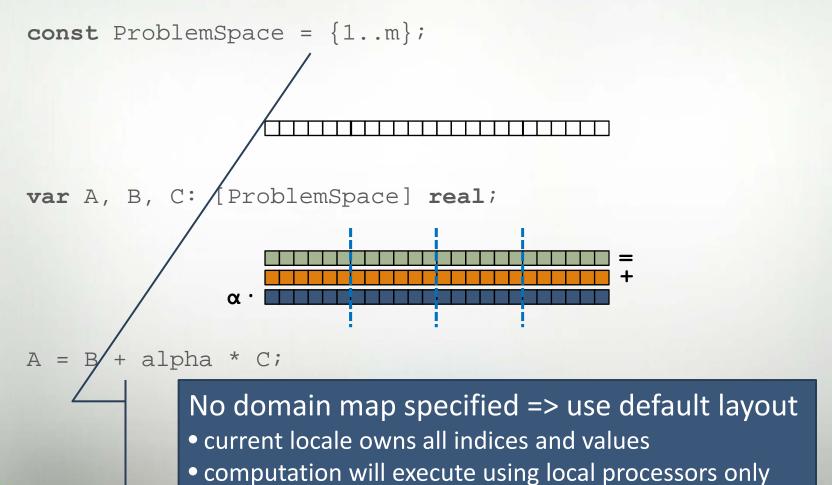


A = B + alpha * C;



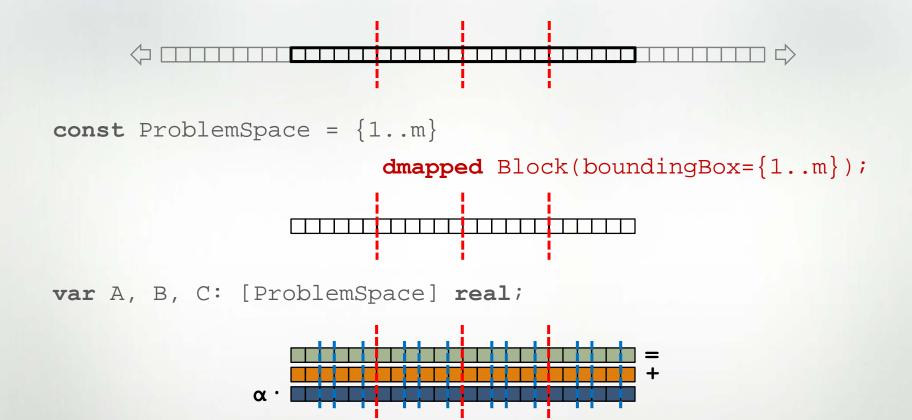


STREAM Triad: Chapel (multicore)





STREAM Triad: Chapel (multilocale, blocked)

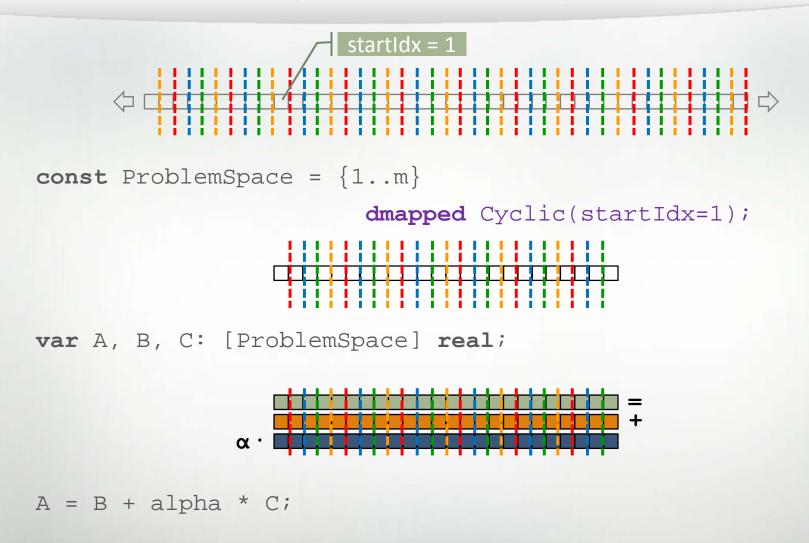




A = B + alpha * C;



STREAM Triad: Chapel (multilocale, cyclic)

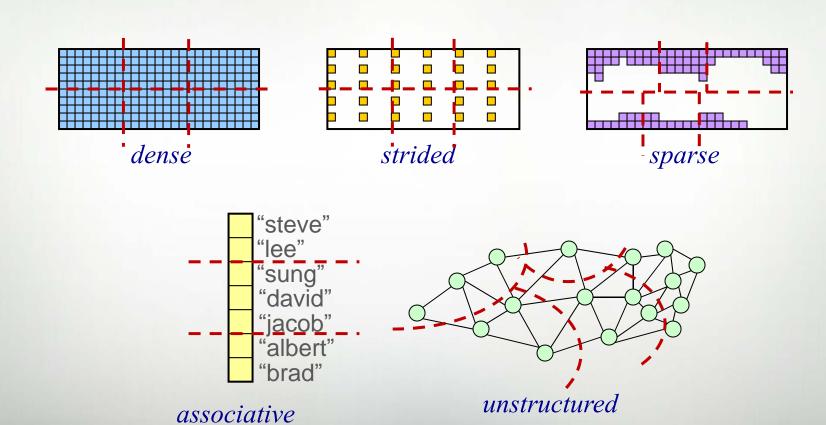








All Chapel domain types support domain maps









Completed domain maps:

- Block, Cyclic, Replicated
- Sparse COO and CSR
- Quadratic probing associative

In the works:

- Block-Cyclic, 2D dimensional
- Distributed associative and sparse







Users can write their own domain maps

- GPU computations
- Dynamically load balanced domains/arrays
- Resilient data structures
- in situ interoperability with legacy codes
- out-of-core computations

•



Summary



Chapel is a new parallel programming language aimed at drastically improving programmer productivity

- Chapel avoids locking crucial implementation decisions into the language specification
 - Separates the roles of domain scientist and HPC expert
 - Results in much cleaner, maintainable code



Outline



- ✓ Motivation
- ✓ Chapel Background and Themes
- √ Tour of Chapel Concepts
- Project Status





Implementation Status – Version 1.6.0

In a nutshell:

- Most features work
- Many performance optimizations remain

This is a good time to:

- Try out the language and compiler
- Give us feedback on the language
- Use Chapel for parallel programming education
- Use Chapel for non-performance-critical projects







- Grow the set of architectures we can target effectively
- Grow the set of codes we are evaluating
- Performance optimizations
- Evolve from prototype- to production-grade







Chapel project page: http://chapel.cray.com

overview, papers, presentations, language spec, ...

Chapel SourceForge page: https://sourceforge.net/projects/chapel/

release downloads, public mailing lists, code repository, ...

Mailing Lists:

- chapel_info@cray.com: contact the team
- chapel-users@lists.sourceforge.net: user-oriented discussion list
- chapel-developers@lists.sourceforge.net: dev.-oriented discussion
- chapel-education@lists.sourceforge.net: educator-oriented discussion
- chapel-bugs@lists.sourceforge.net: public bug forum
- chapel_bugs@cray.com: private bug mailing list





THE SUPERCOMPUTER COMPANY