Comparing Parallel Functional Languages: Programming and Performance *

H-W. Loidl[1], F. Rubio[2], N. Scaife[3], K. Hammond[4], S. Horiguchi[3], U. Klusik[5], R. Loogen[5], G.J. Michaelson[1], R. Peña[2], S. Priebe[5], Á.J. Rebón Portillo[4] and P.W. Trinder[1]

[1] *Department of Computing and Electrical Engineering*
Heriot-Watt University, Edinburgh EH14 4AS, Scotland
({`hwloidl,greg,trinder`}`@cee.hw.ac.uk`*)*

[2] *Dpto. Sistemas Informáticos y Programación,*
Universidad Complutense de Madrid, 28040 Madrid, Spain
({`fernando,ricardo`}`@sip.ucm.es`*)*

[3] *Japan Advanced Institute for Science and Technology,*
1/8 Asahidai, Tatsunokuchi, Nomigun, Ishikawa, 923-1211
({`norman,hori`}`@jaist.ac.jp`*)*

[4] *School of Computer Science,*
University of St Andrews, KY16 9SS, Scotland
({`kh,alvaro`}`@dcs.st-and.ac.uk`*)*

[5] *Fachbereich Mathematik und Informatik*
Philipps-Universität Marburg, D-35032 Marburg, Germany
({`klusik,loogen,priebe`}`@mathematik.uni-marburg.de`*)*

Abstract. This paper presents a practical evaluation and comparison of three state-of-the-art parallel functional languages. The evaluation is based on implementations of three typical symbolic computation programs, with performance measured on a Beowulf-class parallel architecture.

We assess three mature parallel functional languages: PMLS, a system for implicitly parallel execution of ML programs; GPH, a mainly implicit parallel extension of Haskell; and Eden, a more explicit parallel extension of Haskell designed for both distributed and parallel execution. While all three languages employ a completely implicit approach to communication, each language takes a different approach to specifying and controlling parallelism, ranging from explicit identification of processes as language constructs (Eden) through annotation of potential parallelism (GPH) to automatic detection of parallel skeletons in sequential code (PMLS).

We present detailed performance measurements of all three systems on a widely available parallel architecture: a Beowulf cluster of low-cost commodity workstations. We use three representative symbolic applications: a matrix multiplication algorithm, an exact linear system solver, and a simple ray-tracer. Our results show how moderate speedups can be achieved with little or no changes to the sequential code, and that parallel performance can be significantly improved even within our high-level model of parallel functional programming by controlling key aspects of the program such as load distribution and thread granularity.

Keywords: Parallel Computation, Functional Programming, Skeletons, Implicit Parallelism, Automatic Task Decomposition, Load Balancing, Haskell, ML.

1. Introduction

The potential advantages of purely functional programming languages for prototyping and developing parallel programs have long been recognised (Burge, 1975). The high level of programming abstraction simplifies the task of programming, fosters code reuse and facilitates the development of substantially architecture-independent programs. The absence of side-effects avoids the unnecessary serialisation which is a feature of most conventional programs. A comprehensive discussion of these issues is given by Hammond and Michaelson (1999).

Realising this potential in an effective manner has proved an elusive goal, however. Reducing or eliminating programmer control places considerable emphasis on sophisticated automatic systems for detecting and controlling parallelism, making such systems fairly rare and often only available on a few parallel architectures. A comparison of different implementations of such automatic resource management mechanisms, as presented in this paper, is even rarer — to our knowledge this is the first head-to-head performance comparison of several parallel functional languages on the same parallel architecture.

For this paper, five research groups have cooperated to produce the comparisons. We assess three parallel functional languages: Eden and GPH, both extensions of the standard non-strict functional language Haskell (Peyton Jones and Hughes, 1999), and PMLS, a parallel implementation of the strict functional language ML (Milner et al., 1997). The languages all have high-level coordination, represent a range of language and implementation alternatives and are three of the relatively few robust parallel functional language implementations available. Assessment is made on both language and performance levels. We compare the language features available to express parallel coordination, in particular we focus on how parallel tasks are identified and created. In the case of Eden (Breitinger et al., 1997b), task identification and creation

* This work is primarily supported by the Austrian Academy of Sciences (APART fellowship 624), the Japan Society for the Promotion of Science (Postdoctoral fellowship P00778), and UK's Engineering and Physical Sciences Research Council (grant nos. GR/L 93379, GR/M 32351 and GR/L 42889).

are explicit. In the case of GpH (Trinder et al., 1996), potential parallelism is identified through new language primitives, with tasks created automatically during program execution on the basis of load. In the case of PMLS (Michaelson et al., 2001), parallel tasks are identified by automatically detecting instantiations of certain higher-order function templates, skeletons. On the performance level we use three representative symbolic applications that have also been widely studied in the general parallel programming community: a matrix multiplication algorithm, an exact linear system solver, and a simple ray-tracer.

The remainder of this paper is structured as follows: Section 2 discusses general concepts of parallel programming and their importance in the context of a functional language. Section 3 presents a detailed comparison of the three languages we are considering. This section separates the user-visible language constructs that are needed for expressing parallelism from the implementation of these constructs. Section 4 presents measurements of all three systems for the three example programs mentioned above. We discuss the ease of implementing these programs, the support for performance tuning, and the overall performance achieved on a 32-node Beowulf cluster. Section 5 relates our languages to other parallel functional programming languages. Finally, Section 6 concludes.

2. Parallel Functional Programming

2.1. Why Parallel Functional Programming?

Parallel programming is inherently harder than sequential programming. Traditionally the programmer must not only describe *what* to compute, i.e. a correct algorithm, but also *how* to organise the subcomputations on the target architecture, i.e. effective parallel coordination. Contemporary functional languages have three key properties that make them attractive for parallel programming: they have powerful mechanisms for abstracting over both computation and coordination; they eliminate unnecessary dependencies; and their high-level coordination achieves a largely architecture-independent style of parallelism.

2.1.1. *Abstraction.*

Functional languages have excellent abstraction mechanisms that can be applied to both computation and coordination (Hughes, 1989). Two important abstraction mechanisms are *function composition* and *higher-order functions.* Function composition allows complex problems to be decomposed into simpler sub-functions. Higher-order functions, ones that manipulate other functions, allow new control constructs to be defined as required. Through use of powerful mechanisms such as these, functional programs are typically much shorter than their imperative or object-oriented equivalents.

The principle of abstraction can be carried through to parallel programming, where higher-order functions may be used to form the basis of new parallel programming constructs. Typically, parallel functional programs will abstract over details such as process placement, the timing and volume of communication, and synchronisation issues. More effort can thus be devoted to improving parallel algorithms. High level abstraction of parallel constructs encourages experimentation with alternative parallelisations, which often leads to improved solutions for novel parallel problems.

2.1.2. *Elimination of unnecessary dependencies.*

The absence of side-effects makes it relatively straightforward to identify potential parallelism. Since the natural method of program construction is by composing functions to the depth required rather than by sequential composition, accidental sequential dependencies are not introduced into the source program. The only source of sequential dependency is that the arguments to a function must be evaluated before they can be used. That is, dependencies are identified solely on the basis of use. Since values do not change once they have been computed, dataflow analysis is not needed to determine usage patterns, even at an inter-procedural level.

2.1.3. *Architecture-independence.*

Good parallel abstractions encourage high-level portability by abstracting over lower level issues. In extreme cases, this abstraction may hide all details of the parallel implementation yielding a model of implicit parallelism. As the low level issues often depend on properties of a specific architecture, a high-level approach is significantly less architecture-dependent than lower-level approaches. The architecture-independence is bought at the price of elaborate language processors: either the compiler or the runtime system or a combination of both must adapt the high-level parallelism for the underlying architecture. By using, at the runtime-system level, standards like PVM (PVM, 1993) or MPI (MPI, 1997), languages can abstract over architecture characteristics. Unlike imperative languages, functional languages enable a high degree of abstraction over such standards through higher order functions and polymorphism.

2.2. Tasks, Processes and Threads.

Parallel programming involves the identification and creation of sub-tasks that collectively perform the overall task of the program. These sub-tasks must be allocated to (*placed on*) processors that will execute them in some order. Depending on the system, load balancing may occur by *migrating* sub-tasks between processors at execution time.

In this paper, we will distinguish two levels of parallel tasks: *processes*, relatively heavyweight tasks whose behaviour is often revealed to the programmer; and *threads*, which are implicit, and which form part of a process.

Task Identification and Granularity. Tasks may be identified either explicitly by the programmer using some language construct, or implicitly by the system identifying potentially parallel parts of the program. In some cases, the identification may be assisted by the use of *annotations*: programmer instructions that may or may not be exploited by the language implementation. The *granularity*, i.e. the size of the computation, of tasks may thus be determined by the programmer, the compiler, the runtime system or a combination of these.

Task Creation. Tasks may be created either statically at initialisation or dynamically during execution of the program. In the latter case, they may be created either immediately they are identified (*eager task creation*) or delayed until they are deemed to be required by the runtime system (*lazy task creation*). When a task is created, it is allocated resources that allow it to execute independently on some parallel processor. In some cases a task may return resources to the system while being suspended, i.e. while it waits for the availability of required data.

Task Placement. When a task is created, it is placed on a processor that will execute it. This placement may either be on the basis of static information determined before execution by the compiler, or dynamically, perhaps in response to load information. Static placement usually gives a good balance for regular task structures, in cases where the communication pattern can be determined in advance. Dynamic placement is more appropriate in situations where the task structure is irregular, cannot be pre-determined, or where the structure changes during program execution.

Scheduling and Load Management. Scheduling is needed to manage the execution of multiple tasks on a single processor. Such scheduling may be required to be *fair*, i.e. guaranteed to execute every available thread eventually. Dynamic rebalancing of workload may also be required, especially for irregular task structures on high-latency systems. Rebalancing is usually achieved by *migrating* tasks, but alternatives are to employ *task subsumption*, in which smaller tasks are merged into larger ones, or to maintain a work-pool of potential tasks, which can be communicated between processors at lower cost than tasks which are already executing. Rebalancing may occur as a result of creating excess work on a single processor, or as a consequence of starvation on some processors, in which case a *task stealing* mechanism may be used.

2.3. Communication

Communication is fundamental to executing parallel tasks. In traditional parallel programming, communication is explicit: the programmer uses explicit message-passing primitives, or communicates through explicitly shared variables, which must normally be protected against concurrent modifications. In the more implicit approaches advocated here, communication occurs as a consequence of shared data dependencies between tasks. The systems use either message passing or shared-memory, as appropriate, and automatically protect the data against concurrent modifications, as required.

Code or Data. Traditional parallel systems usually only support data transmission. In a functional setting, it is natural for functions to be transmitted between parallel tasks, and in a non-strict setting, this may extend to partially evaluated or completely unevaluated forms. Although this is no conceptual limitation, the parallel systems discussed here do not perform code migration. Only code pointers are transmitted, as the whole code is usually supposed to reside in all processors. This is sometimes characterised as an SPMD, single program multiple data, approach.

Push or Pull. Data may be transmitted either on demand (a pull mechanism) or when produced (a push mechanism). Pulling has the advantage of transmitting only the data that is required, but pushing will require fewer packets to be communicated if most of the data that is transmitted is required, and will reduce the amount of synchronisation that is needed. In some cases, however, large data structures may be transmitted unnecessarily. This leads to *speculative work*, since not all of the data structure may be needed to compute the result value. The optimal approach is application-dependent, but in general a combination of push and pull appears to be ideal.

Communication Topology. In the more implicit approaches the topology of processes changes dynamically in response to load balancing demands. In this case, the topology is transparent to the programmer, and it might differ between identical program executions. In more explicit approaches, the programmer can control the topology by connecting processes in the desired way. Topologies such as rings or tori can be explicitly programmed. In contrast to such dynamic approaches some systems use a static topology with the exact number of processes fixed at compile time. Such a static approach is common with libraries for parallel programming or skeletons (see Section 5.1). Note that we make no attempt at matching the topology of the architecture to the topology of the processes, since this would introduce an architecture-dependent aspect to program development.

Data Marshalling. Sophisticated data marshalling techniques are employed to automatically pack complex data structures. In some cases, this marshalling extends to graphs as well as hierarchical data structures, and may involve the packing of unevaluated as well as fully evaluated forms (see Section 4.2.4).

Synchronisation. Most systems also employ implicit task synchronisation, when values produced by one task are required by another. A task that requires an uncomputed value may suspend execution awaiting delivery of that value. The task is resumed when the value becomes available. Unlike conventional language approaches, such synchronisation is entirely transparent to the functional programmer, and is handled internally by the runtime system. That is, no explicit communication is required, and no other action is required from the programmer.

3. Language Comparison

This section compares the three parallel functional languages PMLS, GPH, and Eden, a comparison of a wider range of functional languages can be found at Loogen (1999). The three languages have been chosen for the following reasons. Firstly to be consistent with a high-level computation language we select languages with high-level coordination and exclude languages with imperative or low-level coordination. Secondly the languages represent a range of language designs, e.g. both eager and lazy languages, and with coordination ranging from almost entirely implicit (PMLS) to a language (Eden) in which processes can be manipulated by the programmer. Thirdly, the languages represent a range of implementation designs, e.g. both those with predominantly static coordination (PMLS) and those with predominantly dynamic coordination (GpH). Finally we have selected three of the relatively few robust parallel functional languages available.

In this section we introduce the underlying notions of skeleton- thread- and process-based approaches to parallelism, classify our languages, discuss the user-visible language constructs and the implementations of these languages.

3.1. Language

In this section we introduce the parallelism constructs in each language and compare them in terms of expressiveness and paradigm.

3.1.1. *PMLS*

Parallel ML with Skeletons (PMLS) is a parallelising compiler for strict Standard ML, that realises parallelism in higher order functions (HOFs) as algorithmic skeletons. The PMLS system is based on a purist interpretation of the skeletons "credo", seeking to minimise programmer involvement in identifying and exploiting parallelism.

Skeleton-based approaches define a set of parallel templates or *skeletons* (Cole, 1989). The programmer writes the program using these skeletons as appropriate. A parallelising compiler can then exploit the rules provided for each skeleton in order to produce an efficient parallel implementation of the program on the target architecture.

From the functional programmer's perspective, a skeleton is simply a normal higher-order function (HOF). Each HOF is mapped to a different abstract parallel process topology, with parameters specifying details of the tasks that are to be performed.

Since the only parallel constructions that are available to the programmer are the HOFs that have been provided by the language, programmers must design parallel algorithms by adapting the sequential source to these HOFs. The compiler and runtime system are jointly responsible for setting up the corresponding process topologies, and for mapping processes to processors in the best possible way.

HOFs may be given different behavioural interpretations when compiling for different target architectures. This allows a single HOF to abstract over a range of possible parallel behaviours, which are selected on the basis of concrete details such as communication latency, or the granularity of the tasks to which the HOF is applied. In essence, skeletons modify behaviours but not values.

As an example Figure 1 shows an implementation of the common higher order function `map` in PMLS. It applies the function `f` to all

```
fun map f [] = [] |
    map f (h::t) = f h::map f t
val map = fn : ('a -> 'b) -> 'a list -> 'b list
```

Figure 1. Parallel map in PMLS

the elements of the list `(h::t)`. If `f` converts something of type `'a` to type `'b` then `map f` converts an `'a list` to a `'b list`. If we unfold `map f` across a list `[e1,e2...eN]`, the effect is the evaluation of `[f e1,f e2,...,f eN]`. There is no interaction between the evaluation of each element, so in principle these evaluations may be carried out in arbitrary order, in particular in parallel.

A common approach to parallelising `map` is to construct a *task farm* skeleton consisting of a farmer processor controlling worker processors pre-loaded with `f`. Given an initial list, the farmer:

- records all workers as free;
- repeatedly:
  - sends an unprocessed list element to a free worker and records it as busy;
  - receives a processed list element from a busy worker and records it as free;
- until all list elements have been processed;
- assembles the processed list in the appropriate order.

This approach is self-balancing: no workers sit idle so long as there are more list elements to be processed, and variations in the times to process different elements have minimal impact.

There are various topologies for task farms, for example the linear chain where each processor has a bi-directional connection to its predecessor and successor. The farmer passes unprocessed data down the chain of busy workers to the first free worker, and processed data is passed back up the chain to the farmer. Here, the farmer need not keep track of free and busy workers, and may assemble the final list as processed elements become available.

Constructs. The PMLS compiler generates parallel code solely from calls to `map` and `fold`. No other SML constructs are provided or exploited for parallelism. However, the system enables the introduction of new HOFs with new skeletons. In some cases, like `fold`, a proof obligation is put on the programmer to ensure correctness of the parallel code: in the case of `fold` the binary operation must be associative.

Methodology. The programmer need have no conception of parallelism. The compiler will try to exploit parallelism in explicit uses of `map` and `fold`.

```
(* original function *)
fun inc [] = [] |
    inc (h::t) = h+1::inc t

(* first synthesised function, using map *)
val inc1 = map (fn h => h+1)

(* second synthesised function, using foldr *)
val inc2 = foldr (fn h => fn t => h+1::t) []
```

Figure 2. Program Synthesis in PMLS

A pre-processor may also be used to synthesise higher-order functions in programs that lack them, using proof planning driven by middle out reasoning (Cook, 2001). For example Figure 2 shows how, given the function `inc`, this pre-processor can synthesise both `inc1`, defining `inc` in terms of `map`, and `inc2`, defining `inc` in terms of `fold`.

3.1.2. GpH

GpH (Trinder et al., 1998) is a modest conservative extension of Haskell98 (Peyton Jones and Hughes, 1999) realising a thread-based approach to parallelism. *Thread-based approaches* to parallelism allow parallel threads to be created, but do not provide mechanisms to control those threads. Threads are thus managed entirely under runtime-system control. By combining simple thread primitives with higher-order functions, high-level abstractions can be constructed, such as the evaluation strategy approach (Trinder et al., 1998).

Constructs. GpH provides parallel (`par`) and sequential (`seq`) composition as coordination primitives (see Figure 3). Denotationally, both compositions are projections onto the second argument. Operationally `seq` causes the first argument to be evaluated before the second and `par` indicates that the first argument may be executed in parallel. The latter operation is called the “sparking” of parallelism and is used in different variants in many parallel languages. The runtime-system, however, is free to ignore any available parallelism. In this model the programmer only has to expose expressions in the program that can usefully be evaluated in parallel. The runtime-system manages the details of the parallel execution such as thread creation, communication etc.

```
-- basic constructs
par :: a -> b -> b              -- parallel composition
seq :: a -> b -> b              -- sequential composition

type Strategy a = a -> ()       -- type of evaluation strategy

using :: a -> Strategy a -> a   -- strategy application
using x s = s x `seq` x

rwhnf :: Strategy a             -- reduction to weak head normal form
rwhnf x = x `seq` ()

class NFData a where            -- class of reducible types
      rnf :: Strategy a         -- reduction to normal form
```

Figure 3. Basic Coordination Constructs in GPH

Experience of implementing non-trivial programs in GPH shows that the unstructured use of `par` and `seq` can lead to rather obscure programs. This problem can be overcome with *evaluation strategies*: lazy, polymorphic, higher-order functions controlling the evaluation degree and the parallelism of a Haskell expression. Evaluation strategies provide a clean separation between coordination and computation. The driving philosophy is that it should be possible to understand the computation specified by a function without considering its coordination. Figure 3 shows the basic operations over strategies. A strategy on a value of type `a` is a function from `a` to the nullary value `()` executed purely for effect, and the null value is returned to indicate completion. The `using` construct applies a strategy to a Haskell expression. The basic strategy `rwhnf` reduces an expression to weak head normal form (WHNF), the default in Haskell. The overloaded strategy `rnf` reduces an expression to normal form (NF), i.e. containing no reductions. As there are types that are not reduced to normal form in Haskell, e.g. function types, `rnf` is restricted to types that are reduced to normal form by the `NFData` class which is instantiated for all major types. Because strategies are simply functions they can be combined, or passed as parameters using standard language capabilities.

For example the `parList` strategy in Figure 4 is higher-order, applying the argument strategy `strat` to every element of a list in parallel. This strategy is then used in the GPH implementation of parallel map (`parMap`). Note how the algorithmic code is cleanly separated from the strategy, using the sequential code of `map f xs` unmodified when introducing parallelism.

Methodology. GPH programs are developed with an integrated suite of sequential and parallel software tools, based on the Glasgow Haskell Compiler (GHC) (Peyton Jones et al., 1993). The tools for sequential

```
parList :: Strategy a -> Strategy [a]
parList strat []     = ()
parList strat (x:xs) = strat x `par` (parList strat xs)

parMap :: Strategy b -> (a -> b) -> [a] -> [b]
parMap strat f xs = map f xs `using` parList strat
```

Figure 4. Parallel map in GpH

software development include: the Hugs interpreter, for fast development, the GHC compiler and sequential runtime system for optimising compilation to sequential code; and sequential time and space profilers integrated into GHC (Sansom and Peyton Jones, 1995). The tools for parallel software development include: the GranSim parameterisable parallel simulator (Hammond et al., 1995) for flexible and accurate simulation of the parallel behaviour on a range of parallel machines; the GHC compiler and GUM parallel runtime system for parallel execution on multiprocessors; a set of visualisation tools for both GranSim and GUM, visualising the activity of a parallel machine in several levels of detail; prototypes of more detailed parallel profilers (King et al., 1998).

3.1.3. *Eden*

Eden (Breitinger et al., 1997b) extends the lazy functional language Haskell by syntactic constructs to explicitly define and instantiate processes. In contrast to the previous techniques, *process-based approaches* like Eden expose parallel tasks at the language level. The programmer must then manage the tasks using the control mechanisms provided in the language. Eden is explicit about process creation and about the communication topology, but implicit about other control issues such as sending and receiving messages, and process placement. Granularity is under the programmer's control because he/she decides which expressions must be evaluated as parallel processes, and also some control of the load balancing is possible at the programmer's level.

Constructs. Eden provides process abstractions and process instantiations for coordination as shown in Figure 5. The new expression `process x -> e` of a predefined polymorphic type `Process a b` defines a *process abstraction* having formal parameter `x::a` as input and expression `e::b` as output. Process abstractions of type `Process a b` can be compared to functions of type `a -> b`, the main difference being that the former, when instantiated, are executed in parallel. Additionally, when the output or input expression is a tuple, a separate concurrent thread is created for the evaluation of each tuple element. We will refer to each of them as a channel.

```
newtype Process a b = ...

-- process abstraction (language construct)
process x -> e :: Process a b

-- process instantiation
(#) :: (Transmissible a, Transmissible b) => Process a b  -> a -> b

-- non-deterministic merge process
merge :: Process [[a]] [a]
```

Figure 5. Basic Coordination Constructs in Eden

A *process instantiation* is achieved by using the predefined infix operator `(#)`. The context `Transmissible` is needed to guarantee that the elements can be sent through the channels. Each time an expression `e1 # e2` is evaluated, a new process is created to evaluate the application of `e1` to `e2`. We will refer to the latter as the *child* process, and to the owner of the instantiation expression as the *parent* process. The instantiation semantics specifies in which processes these expressions shall be evaluated: (1) Expression `e1` together with its whole environment is *copied* in the current evaluation state to a new processor, and the child process is created there to evaluate the application of `e1` to `e2`, where `e2` must be remotely received. (2) Expression `e2` is eagerly evaluated in the parent process. The resulting full normal form data is communicated to the child process as its input argument.

Once processes are running, only fully evaluated data objects are communicated. The only exception are lists: they are transmitted in a *stream*-like fashion, i.e. element by element. Each list element is first evaluated to full normal form and then transmitted. Processes trying to access input not yet available are temporarily suspended. This is the only synchronising mechanism in Eden.

Figure 6 presents a simple parallel `map` skeleton in Eden, in which a different process is created for every element of the input list. Strategies are used in Eden to influence the evaluation order. In this example, the `spine` strategy is used to eagerly evaluate the spine of the process instantiation list. In this way all processes are immediately created. More sophisticated parallel implementations of `map` have been developed in Eden (Klusik et al., 2002; Klusik et al., 2000) and some will be discussed in Section 4.

Methodology. Like GpH, Eden is based on the Glasgow Haskell Compiler, and can use the same sequential profiling utilities. For parallel profiling Eden provides a simulator called Paradise (Hernández et al.,

```
map_par :: (Transmissible a, Transmissible b) =>
           (a -> b) -> [a] -> [b]
map_par f xs = [pf # x | x <- xs] `using` spine
               where pf = process x -> f x

spine :: Strategy [b]
spine []     = ()
spine (_:xs) = spine xs
```

Figure 6. Parallel map in Eden

2000) which is based on GRANSIM, so that tuning the performance of an Eden program is a similar process to that in GPH.

Parallel programming in Eden can be done by explicitly defining and instantiating a process topology. This would be equivalent to sequential functional programming with explicit recursion. Sometimes this is appropriate, but an experienced functional programmer will try to use higher-order functions, i.e. skeletons, as much as possible in order to reduce the amount of work and the possibility of making mistakes. In a complex application both methods may be simultaneously needed. The main difference between Eden and more traditional skeleton-based languages, such as PMLS, is the fact that skeletons can be specified within Eden itself. Thus, Eden serves both as a computation and coordination language, yielding a high degree of flexibility for the programmer.

3.2. IMPLEMENTATION

In this section we compare the implementations of the languages on arbitrary parallel architectures.

3.2.1. *PMLS*

The PMLS approach is based on:

- maximising compile-time activity to minimise run-time overheads;
- configuring the virtual topology of the target system to reflect closely the HOF hierarchy in the source program.

While this is relatively inflexible, for example making exploitation of parallelism across condition branches difficult, it often results in very efficient code.

Compile Time. The PMLS compiler front end parses, elaborates and type checks SML to produce an abstract syntax tree (AST). The ML Kit is used as the front end. The AST is traversed to extract an abstract network showing the nesting hierarchy of HOFs. Free variable lifting, or defunctionalisation, is performed to simplify passing free variable bindings to skeletons, and to avoid runtime transmission of closures. The AST and abstract network are traversed to identify HOFs to be realised as skeletons and to generate skeleton network code and MPI registration in C. The resulting AST is finally translated into Objective Caml for linkage by the OCaml and GNU C compilers with the appropriate skeletons, and skeleton network and MPI registration code.

PMLS skeletons are written in C with MPI. The `map` function is implemented as a task farm and `fold` as a divide-and-conquer network. The skeletons are hybrid and may be run either in parallel or sequentially. Skeletons are coordinated at runtime by generic “Pskel” nodes which can switch their hybrid modes. Otherwise, skeletons are linked statically with no runtime change of topology. Adopting an SPMD approach, all processors are pre-loaded with all skeletons and functions.

The use of Objective Caml and GNU C to generate native code enables a high degree of portability. PMLS has been ported to a Fujitsu AP3000, IBM SP/2, Cray T3E, networks of UltraSparc workstations, SUN Enterprise and Beowulf, displaying consistent performance across all platforms. For further details see Scaife et al. (2001).

Run Time. PMLS generates code to link static skeletons through Pskel nodes. The Objective Caml run-time environment provides garbage collection and appropriate libraries. At run-time, the Pskel nodes at each level determine their behaviour from the skeleton network. In particular, intermediate Pskel nodes in the hierarchy will switch between parent and child operation if initiated in parallel mode. There is no movement of code or closures at runtime. For further details of the compiler design and implementation see Michaelson et al. (2001).

The single processor efficiency of PMLS has been measured for the `raytracer` as 86% on our Beowulf cluster. The main sources of overhead are slight inefficiencies introduced in program transformation stages, such as extra function calls, and the need to propagate additional information that is used as arguments to the skeletons used for exploiting the parallelism. In a multi-processor setup the worker nodes of the skeleton used in `linSolv` exhibit an efficiency of 84%. In this case the main source of overhead is some idle time introduced by blocking communications between nodes in this skeleton. An implementation of a more efficient version, using non-blocking communication wherever possible, is currently in development.

Early versions of PMLS were hampered by inefficiencies in the translation process from SML to Objective Caml. More recent versions employ a set of optimising transformations allowing fairly similar performance between the output from PMLS and hand-coded Objective

Caml. For example, the ray tracer has also been implemented in Objective Caml and gives a sequential runtime of 195 seconds on similar data to that used in this paper. The output of the PMLS compiler takes 241 seconds on the same data. A slowdown of around 20% is acceptable and is attributable to the remaining inefficiencies in the translation process.

3.2.2. *GpH*

Compile Time. The two additional language constructs of GPH, `par` and `seq`, are treated as built-in functions by the compiler. They are implemented as system-calls in the GUM runtime-system. GPH programs are compiled using almost all of the sequential optimisations in GHC, although care must be taken to preserve `par` and `seq`.

Run Time. The GUM runtime-system for GPH realises a parallel graph-reduction machine built on top of GHC's sequential STG-machine. To synchronise multiple threads, a thread locks the node of the graph when starting its evaluation, and other threads requesting that data will be added to a blocking queue attached to the locked closure. Access to remote closures is managed by new FetchMe nodes, representing global indirection. On requesting the contents of such a node a message will be sent to the target processor and the requesting thread will be added to a blocking queue. The details of these synchronistation and communication mechanisms are discussed in Loidl (1998)[Chapter 2].

Being integrated into GHC, GUM makes use of existing analyses and optimisations for efficient sequential execution. A discussion of the design and implementation of GUM is given in Trinder et al. (1996). In summary, the additional features to enable parallel execution are:

- sparking of threads, i.e. identified program expressions may be evaluated as independent threads or they may be inlined by other threads, achieving dynamic granularity control as in the lazy task creation mechanism (Mohr et al., 1991);

- multi-threading, i.e. independent threads of control are executed in an interleaved fashion thereby enabling an overlap of computation and communication on each processor;

- virtual shared heap, i.e. the physically distributed heap is treated as a shared heap with global pointers to remote processors, with transparent communication on access of non-local data;

- automatic marshalling of data and work, i.e. when data or work is needed on another processor, a graph structure is automatically serialised, sent to another processor, and again unpacked into a graph structure;

- distributed garbage collection, i.e. weighted reference counting is used to garbage collect global pointers that are not used any more.

In order to assess the overheads of the different systems we have measured key parameters of the runtime-system. One important parameter is the single-processor efficiency, i.e. the sequential runtime divided by 1 PE runtime in percent. For GUM we have previously measured 80%–93% on simple programs (Trinder et al., 1996), and now 77% on `linSolv` as used in Section 4.3. In a multi-processor execution it turns out that maintaining a virtual shared heap on a distributed memory machine is most expensive. In particular the management of a hash table mapping local heap addresses to global heap address accounts for up to 3.8% of the total execution time, in earlier version, pre-dating recent improvements in GUM even up to 8%. In comparison, the costs for packing graph structures and communication play only a minor role in the total runtime: less than 1% for this program. The costs for creating parallelism are, by design, very small: creating a spark requires only adding a pointer to an array, and threads are very light (14 bytes for the thread descriptor) with initially small, tunable stacks (1kB).

A detailed discussion of these overheads in GPH is presented in Loidl (2002). This paper separates the overhead into that induced by the thread management, memory management and communication subsystems of GUM. It then focuses on virtual shared memory management, which turns out to be the most expensive part, and uses both `linSolv` and `raytracer` as example programs. Several improvements of the basic load balancing mechanism, that we exploit in these measurements, are presented in Loidl (2001).

3.2.3. *Eden*

Compile Time. Eden extends the optimising Glasgow Haskell Compiler with a few modifications. In Eden, lazy evaluation is changed to eager evaluation in two cases. Firstly, processes are eagerly instantiated when the expression under evaluation demands the creation of a closure of the form `o = e1 # e2`. Secondly, instantiated processes eagerly produce their output expressions and communicate them on channels. These modifications of the standard Haskell semantics are aimed at increasing the degree of parallelism and at speeding up the distribution of the computation, and they are implemented by compile-time provided by Eden, i.e. process abstractions, process instantiations, dynamic channels and `merge` instantiations, are translated into runtime-system calls.

Run Time. The design of DREAM (Breitinger et al., 1997a), the parallel abstract graph-reduction machine implementing Eden, is largely similar to GUM. We focus on the differences to GUM:

- In DREAM, the concept of a virtual shared heap does not exist. Each process evaluates its outputs autonomously with respect to other processes. The entire graph needed by a newly instantiated process is copied into its heap *before* it starts running. In some cases, this can even lead to some duplication of work, but it reduces the communication overhead of DREAM. Moreover, global garbage collection reduces to the sending of terminate messages to processes whose output has been detected to be garbage during a local garbage collection.

- In contrast to GPH, Eden threads are mandatory. Processes in DREAM and threads in GUM are related as follows: A process in DREAM is implemented by several threads, which directly correspond to threads in GUM. These threads run concurrently on the same processor, so that different output values can be independently produced. Threads synchronize on shared graph nodes as in GPH. In addition, special queue-me closures represent input from remote processes which is not available yet. On requesting the contents of such a closure a thread will blocked until the input arrives.

- Process placement in Eden is controlled by the runtime-system in two different modes that can be set-up at the beginning of the execution: (1) round-robin mode, in which processes are instantiated in consecutively numbered processors, or (2) random mode, where processes are instantiated in randomly chosen processors.

As Eden shares parts of GUM's thread management and communication subsystem, the runtime-system overheads are similar. However, Eden overheads are smaller, as it is not necessary to maintain a virtual shared graph. The single processor performance for `linSolv` as used in Section 4.3 has been 89%. In general, the main bottlenecks in Eden are due to the packing and unpacking routines, which are not yet optimized. For instance, packing a 600×600 matrix of integers takes 1% of the time required for multipying it. Moreover, as there are not yet multicasting facilities in Eden, once a packet has been sent to a process, it cannot be reused the packing effort in order to send the same packet to other process. See Rubio (2001) for a more detailed description of Eden overheads.

Table I. Language Comparison

	Eden	GpH	PMLS
Language			
Approach	process-based	thread-based	skeleton-based
Constructs	proc. abstraction proc. instantiation	`par/seq`	HOFs
Programming Abstraction	skeletons	eval. strats	—
Methodology	define topology and/or skeletons, simulation	simulate, execute, visualise	—
Implementation			
Compile-time support	force strict evaluation of channel data	—	synthesise HOFs, process network, link skeletons
Run-time support	graph-red. over distributed heap	graph-red. over shared heap	skeleton library over distributed heap

3.3. Summary

Table I summarises the language and implementation features of the PMLS, GpH, and Eden. On the language level it shows the higher level of abstraction for PMLS, using a skeleton-based approach, which does not require language extensions for parallelism at all, whereas GpH adds combinators to expose parallelism and Eden adds a construct for explicit process creation. On the implementation level PMLS performs sophisticated static analysis and program synthesis in order to generate a sufficient amount of parallelism. Both GpH and Eden rely mostly on a sophisticated runtime-system with dynamic resource management.

To achieve good single processor performance all systems use state-of-the-art sequential compilers for functional languages: GpH and Eden use GHC, and PMLS uses OCaml. Using `linSolv` (Section 4.3) as benchmark we achieve single processor efficiencies of 77% for GpH, mainly due to using a two-space garbage collector in the current implementation, 89% for Eden (using a better garbage collector), and 84%

for PMLS which uses a two-generation garbage collector. In measuring the overheads in multi-processor executions we identified in GPH the maintenance of hash tables in the virtual shared memory management, and in PMLS the usage of blocking communication at certain stages and the single-master, multiple-worker parallel model to be the most costly components. The details of these runtime-system measurements for GPH and PMLS, including data obtained from Beowulf and SunSMP machines, will be published in separate paper (Loidl et al., 2002).

4. Experimental Results

This section describes the results we have obtained using three programs: `matMult`, a matrix multiplication algorithm, `linSolv`, an exact linear system solver, and `raytracer`, a simple ray tracer. The parallel algorithms themselves have been explained in more detail in previous papers. In this section we focus on a comparison of the implementations in and the performances achieved with PMLS, GPH, and Eden.

Although rather simple in nature, these programs represent a range of applications we are interested in. In previous studies on developing parallel applications in GPH (Loidl et al., 1999) we have identified the class of symbolic applications, with complex data structures and irregular parallelism, as the most interesting application domain. For pragmatic reasons we had to keep the program size down: ensuring that all three versions implement the same algorithm and produce comparable dynamic structures was the main engineering constraint. Of the 3 programs in this section the linear system solver, with its multiple homomorphic images approach, fits these characteristics best, with the other programs focussing on different aspects of the execution.

More specifically, matrix multiplication is a well-studied parallel program and serves to relate our approach to that of imperative languages (with concrete language and performance comparison in Section 6). The linear equation solver exhibits a structure typical for a class of symbolic applications, which is quite different from conventional iteration-based techiques. It also performs a high amount of heap consumption and creates less regular parallelism, and is therefore closest to typical symbolic applications. The ray tracer is an example of a data-parallel application, and issues of task and computation granularity become important in this context.

4.1. Experimental Framework

All measurements have been performed on a 32-node Beowulf cluster (Ridge et al., 1997) at Heriot-Watt University, consisting of Linux RedHat 6.2 workstations with a 533MHz Celeron processor, 128kB cache, 128MB of DRAM and 5.7GB of IDE disk. The workstations are connected through a 100Mb/s fast Ethernet switch with a latency of $142\mu s$, measured under PVM 3.4.2.

4.2. Matrix Multiplication

4.2.1. *Problem Description*

Given two square matrices of arbitrary precision integers $A, B \in \mathbb{Z}^{n\times n}, n \in \mathbb{N}$ find their product, i.e. a matrix $C \in \mathbb{Z}^{n\times n}$ such that $C_{i,j} = \sum_{k=1}^{n} A_{i,k} * B_{k,j}$.

4.2.2. *Parallel Algorithms*

We start with a sequential algorithm directly implementing the above specification of matrix multiplication, shown in Figure 7. By using an algebraic datatype `Matrix a` to represent matrices as lists of lists we can overload standard arithmetic operations such as multiplication. The main function is `multMatT`, which takes A and B^T, i.e. the transposed matrix B as input. It computes $A * B$ in a double nested list comprehension, computing the rows of the result matrix in the outer comprehension and the elements of a row in the inner comprehension. The function `multVec` computes the sum in the specification above for two vectors of length n.

```
data (Num a) => Matrix a = M [[a]]

multMat :: (Num a) => Matrix a -> Matrix a -> Matrix a
multMat (M m1) (M m2) = M (multMatT m1 (transpose m2))

multMatT :: (Num a) => [[a]] -> [[a]] -> [[a]]
multMatT m1 m2T = [ [multVec row col | col <- m2T] | row <- m1]

multVec :: (Num a) => [a] -> [a] -> a
multVec v1 v2 = sum (zipWith (*) v1 v2)
```

Figure 7. Sequential `matMult` (Haskell version)

4.2.2.1. *Naive data parallelism:* Since each element of the result matrix can be computed independently, we can exploit data parallelism by generating one parallel task for each element in the result matrix. However, the excessive number of small computations leads to a very poor performance in general. For example, the GpH implementation of this *naive data parallel* version yields a speedup of about one up to 16 processors. We do not consider this version any further.

4.2.2.2. *Row clustering:* The granularity of the naive parallel algorithm can be increased by computing an entire row of the result matrix by one task. Assuming square matrices of size $n \times n$ with integers of average size l in its internal representation, and assuming that integers are multiplied by using the algorithm of Karatsuba and Ofman (1962), the computational complexity for each task is $\mathrm{O}(n^2 * l^{log_2 3})$, while the total communication complexity, i.e. the amount of data (in machine words) to be sent, is $\mathrm{O}(n^3 * l)$. The latter complexity is due to the fact that each task requires the whole second matrix to compute one final row, and n tasks are created. In order to effectively improve parallel performance, the granularity of the tasks has to be increased by computing as many elements as possible inside each task and the communication has to be minimised.

We can improve the granularity further by computing many rows of the resulting matrix by each task. With perfect load distribution, if p processors are available, p tasks should be created, each one evaluating n/p rows of the resulting matrix. Using such a *row clustering* approach the communication complexity of the main process is $\mathrm{O}(n^2 * p * l)$ whereas the computational complexity of each process is $\mathrm{O}(n^3 * l^{log_2 3}/p)$. For large values of n better speedups can be expected, since the computation-communication ratio increases.

4.2.2.3. *Block clustering:* An alternative form of clustering the data is to partition the input matrices into blocks, performing *block-clustering*, and then perform the basic arithmetic over these blocks rather than over simple integer values. Figure 8 depicts this partitioning, and indicates that for the computation of one block in the result matrix, only one row of the partitioned matrix A and one column of the partitioned matrix B is needed. In this version the computational complexity of each process is still $\mathrm{O}(n^3 * l^{log_2 3}/p)$ but its communication complexity is only $\mathrm{O}(n^2 * l/\sqrt{p})$ as the processors do not require the whole second matrix.

4.2.2.4. *Torus topology:* All parallel versions so far rely on a broadcast of all data at the beginning of the computation with a communication complexity of $\mathrm{O}(n^2 * l * \sqrt{p})$. Therefore, the main process tends to become a bottleneck especially for large numbers of processors. To avoid such a bottleneck we can use a *torus topology* as depicted in Figure 11. Initially each process in the torus receives only its own blocks from matrices A and B. In each step the processor computes the product of

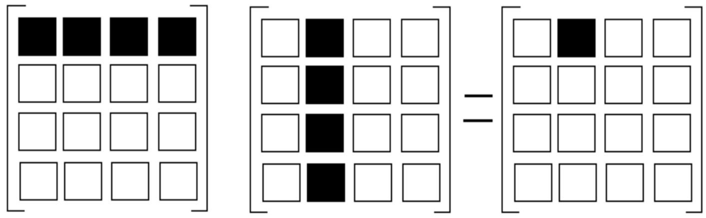

Figure 8. Structure of block-clustering `matMult`

both blocks, adds the product to the intermediate result computed so far, and then obtains the next blocks from its neighbours. As shown in Figure 11 the blocks of the first matrix are transmitted from left to right in the torus, while those of the second matrix are transmitted top down. This algorithm is well-known in the literature as Gentleman's algorithm (Quinn, 1994). In this version the communication complexity of the main process is $O(n^2 * l)$, i.e. it does not depend on the number of processors, and the communication between the processors is spread over the entire execution of the program. The main drawback of this approach is that it requires a perfect square number of processes to form a torus topology.

4.2.3. *Implementations*

```
multMatTParRow :: (Num a,Transmissible a) =>
                  Matrix a -> Matrix a -> Matrix a
multMatTParRow (M m1) (M m2) = M (concat result)
 where result = map_par multMatT (zip (splitIntoN noPe m1) (repeat m2))
```

Figure 9. Row-clustering `matMult` (Eden Version)

4.2.3.1. *Eden:* The *row-clustering* version in Eden creates as many processes as processors available with each of them computing $\frac{n}{p}$ rows of the product matrix. This version, as shown in Figure 9, uses the built-in variable `noPe`, representing the number of available processors. The function `splitIntoN n xs` splits the list `xs` into `n` nearly equal size sublists (see Appendix A for the definition of `splitIntoN` and other auxiliary functions used in this section).

The *block-clustering* version creates `size`×`size` processes, each of them computing a block of the product matrix. In order to reduce the total amount of communication, the typical value of `size` will be $\lfloor\sqrt{\texttt{noPe}}\rfloor$. The main difference to the row-clustering version is the way

```
multMatTParBlock :: (Num a, Transmissible a) =>
                    Int -> Matrix a -> Matrix a -> Matrix a
multMatTParBlock size m1 m2 = decluster size result
  where result = map_par multMatT (zip (clusterLeft size m1)
                                       (clusterRight size m2))
```

Figure 10. Block-clustering `matMult` (Eden Version)

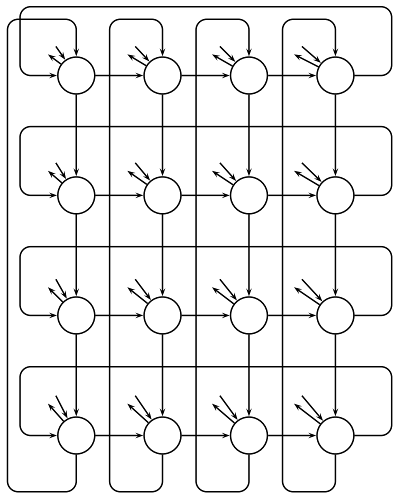

Figure 11. Process topology generated using a torus

in which the matrices are split, which is encoded in the `clusterLeft` and `clusterRight` functions. The first function splits matrix A into a list of rows, the second function splits matrix B into a list of columns.

The *torus version* of the algorithm can be expressed in Eden in terms of its general torus skeleton (Peña and Rubio, 2001). The main argument of the torus skeleton is the function to be performed by each node in the torus topology (see Figure 11). Each node has three input parameters: one from the parent; one from the left neighbour; and one from the top neighbour. It produces three values: one to the parent; one to the right neighbour; and one to the neighbour below.

With this torus skeleton, the matrix multiplication algorithm `multMatPar` shown in Figure 12 takes the size of the torus, `torusSize`, splits the matrices `m1` and `m2` into blocks `m1ss` and `m2ss`, respectively, thereby pairing them appropriately, and calls the torus skeleton `torus` with the function `multMatPar'` to be applied by the node processes of the torus. The per-node function performs a list of matrix multiplications `sms` — one for each pair of blocks it receives — and sums all

```
torus :: (Transmissible a,Transmissible b,Transmissible c,Transmissible d) =>
             ((c,a,b)->(d,a,b)) -- Main function in each process
             [[c]] ->           -- Inputs from parent to children
             [[d]]              -- Outputs from children to parent
torus f m = ...

multMatPar :: (Num a, Transmissible a) =>
              Int -> Matrix a -> Matrix a -> Matrix a
multMatPar torusSize m1 m2 = combine results
 where results = torus (multMatPar' torusSize) (zipWith zip m1ss m2ss)
       m1ss    = splitMatrix1 torusSize m1
       m2ss    = splitMatrix2 torusSize m2

-- Function performed by each worker
multMatPar' :: (Num a, Transmissible a) => Int ->
               ((Matrix a, Matrix a), [Matrix a], [Matrix a]) ->
               (Matrix a, [Matrix a], [Matrix a])
multMatPar' size ((sm1,sm2),sm1s,sm2s) = (result,toRight,toBottom)
  where toRight  = take (size-1) (sm1:sm1s)
        toBottom = take (size-1) (sm2':sm2s)
        sm2'     = transpose sm2
        sms      = zipWith (curry multMat2) (sm1:sm1s) (sm2':sm2s)
        result   = foldl1' addMatrices sms
```

Figure 12. Torus version of `matMult` in Eden

products to obtain the `result` which is returned to the parent. Note that the first pair, `(sm1,sm2)`, is received directly from the parent, whereas the other pairs are received from the left and right neighbours as part of `sm1s` and `sm2s`, respectively.

4.2.3.2. GpH*:* Figure 13 shows a *row-wise clustering* version of `multMatPar` in GpH. This version uses the sequential matrix multiplication, `multMat`, as shown in Figure 7 without change. All parallelism is defined by a strategy attached to `multMat`. The strategy first evaluates both input matrices, in order to avoid competition for unevaluated data during the evaluation, and then uses the predefined strategy `parListChunk z rnf m` to fully evaluate chunks of `z` elements in the matrix `m` in parallel.

```
multMatPar :: (Num a, NFData a) => Int -> Matrix a -> Matrix a -> Matrix a
multMatPar z m1 m2 = multMat m1 m2
                      `using` \ (M m) -> rnf m1 `seq`
                                         rnf m2 `seq`
                                         parListChunk z rnf m
```

Figure 13. Row-clustering `matMult` (GpH version)

The *block-wise clustering* GpH version in Figure 14 implements the algorithm sketched in Figure 8. In contrast to the purely strategic row-clustering version, it uses explicit functions for clustering and declustering the input and result matrices. Note that the code used to multiply the clustered matrices, `multMatT`, is the sequential matrix multiplication overloaded to work on matrices of matrices. The strategy attached to the clustered result matrix, guarantees that every block in the clustered result matrix is evaluated in parallel. Such separation of data-layout from computation and reuse of sequential code greatly improves the productivity in our languages, and is in contrast to sophisticated C-based block-clusterings, where extensive code restructuring is needed to obtain very efficient parallel programs (Frens and Wise, 1997).

Based on experiences with different cluster functions, we have developed a generic mechanism for clustering arbitrary user-defined data structures, using formal program transformation to derive data parallel code such as this from the sequential code (Loidl et al., 2001).

```
multMatPar :: (Num a, NFData a) => Int -> Matrix a -> Matrix a -> Matrix a
multMatPar z m1 m2 =
  decluster (multMatT (cluster z m1) (cluster z (transposeMat m2))
            `using` \ (M m) -> rnf m1 `seq`
                               rnf m2 `seq`
                               parList (parList rnf) m )
```

Figure 14. Block-clustering `matMult` (GPH version)

4.2.3.3. *PMLS:* The PMLS implementation uses nested lists for representing matrices, and Objective Caml's arbitrary-precision integer arithmetic library for the operations over the matrix elements. There is no general overloading of the basic arithmetic functions for matrices as in Haskell.

```
(* vector product call *)
fun inner row col = multVec row col

(* inner map - parallel *)
fun outer BT row = map (inner row) BT

(* outer map - parallel *)
fun multMat A B = map (outer (transpose B)) A
```

Figure 15. Row-clustering `matMult` (PMLS version)

A straightforward sequential SML algorithm, that uses `map` instances instead of Haskell's list comprehensions, is shown in Figure 15. Since this code uses one of the HOFs that is implemented as a parallel skeleton, it can be directly parallelised by the PMLS compiler resulting in a pair of nested `map` skeletons. The outermost `map` in `multMap` computes a list of matrix-vector products by mapping the matrix-vector operation, called `outer`, over the rows of matrix A. The `outer` function computes a list of dot products by mapping the `multVec` function over the columns of matrix B. Note that the entire matrix B is free in `multMat`. The compiler's free-value analysis phase detects this property and generates code to transmit B to the workers prior to running the outer farm.

The parallel `map` skeleton has clustering of data built into it. The clustering size is global to the whole program and set manually, at present. With a clustering parameter of one this algorithm corresponds to the *naive data parallel* version mentioned above. In non-nested mode, with clustering set to a suitable value, the behaviour is identical to the *row-wise clustering* version. In nested mode, with both `map` skeletons implemented in parallel, the matrix B is only transmitted to the intermediate processors.

```
(* Block map over outer product *)
fun BMmult (A,B) =
  let
    val rows = length A
    val outerAB = outer_product (A,transpose B)
    val AB = map Mdotprod outerAB
  in
    split rows AB
  end
```

Figure 16. Block-clustering `matMult` (PMLS version)

Figure 16 shows an approximation of a block-wise clustering version. The blocks are generated by the `map`'s implicit clustering mechanism. Since PMLS does not provide a user-level mechanism for enforcing absolute placement of data, the quality of the code depends on the ratio of processors to blocks. The best results are achieved if the number of blocks is a multiple of the number of processors. Overall, this method is slightly less communications-bound than the row-wise clustering method since the entire matrix B is not transmitted to all the processors.

4.2.4. *Performance Results*

The measurements presented in this section are based on two 200×200 matrices of arbitrary precision integers, none of which is larger than $2^{16} - 1$, i.e. one machine word. For the row- and block-clustering versions Eden uses as many blocks as processors, whereas GPH uses a chunk size of 40. For the row-clustered version PMLS uses blocks of 3 rows, while for the block-clustered version it uses blocks of size 40×40.

The results presented here will be related to the performance of parallel versions implemented in C with PVM and GMP in Section 4.2.5 and in the conclusions (Section 6.2).

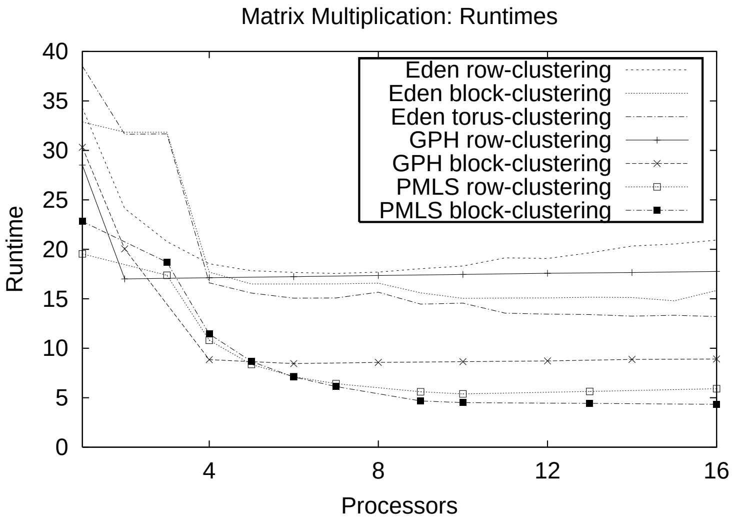

Figure 17. Runtimes of `matMult` on the Beowulf (in seconds)

Figure 17 summarises the runtimes and Figure 18 the speedups of all versions on our Beowulf cluster. The sequential performance of the strict language, PMLS, is noticeably better than that of the lazy languages, Eden and GPH, with variations of about 26% between the versions of the latter languages.

For all versions the performance tails off fairly early with an increasing number of processors. In general, this is due to the high ratio of communication to computation as elaborated in Section 4.2.2. In Eden the torus topology behaves better than the block clustering version, which in turn is better than row clustering. The torus version shows a small increase in performance even for large numbers of PEs. This is in contrast to e.g. the block-clustering GPH version, which shows good speedups up to 4 PEs but tails off after that. In PMLS the difference in performance between the simple row-clustered and the refined block-clustered version, due to reduced communication, is most pronounced. The amount of communication can be directly linked to the free occurrence of B in the row- (Figure 15) but not in the block-clustered version (Figure 16). Furthermore, PMLS uses a task farm skeleton, as presented in Section 3.1.1, for implementing `map` in parallel. This model achieves a good load balance but limits the scalability of the system because the master process becomes a communication bottleneck for large numbers of processors.

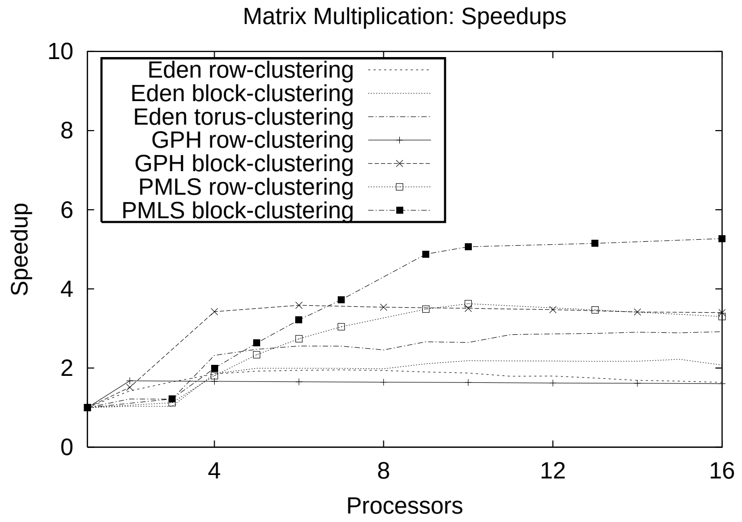

Figure 18. Speedups of `matMult` on the Beowulf

One important difference between the implementations of the three languages is the way that data items are packed in order to send them to other processors. In PMLS a generic serialisation routine is used, whereas GPH implements its own graph packing algorithm. As a result, the PMLS version is more portable, but the GPH and Eden versions are in general more efficient. Graph packing could be improved even further by developing specialised packing routines for commonly used data structures, such as lists, thereby reducing packet size and packing time. On a high-latency architecture such as the Beowulf and for communications-bound algorithms such as `matMult` this should yield significant performance improvements.

In summary this example shows how Eden's richer coordination constructs, compared to GPH and PMLS, can be used to improve parallel performance, without having to resort to mechanisms of explicit synchronisation. The higher level of abstraction in GPH and PMLS reduces programming effort for the initial version, but also reduces the amount of programmer control. Although we describe Eden as having the most explicit coordination in this comparison, it must be emphasised that it is far more implicit than most conventional parallel programming languages.

Table II. Performance Results for C+PVM matmult programs on the Beowulf (runtimes in seconds)

	row-parallel		block-parallel		torus-parallel	
# PEs	RT	Spdup	RT	Spdup	RT	Spdup
1	5.75	1	5.75	1	5.75	1
4	2.00	2.87	2.00	2.87	1.93	2.98
9	1.36	4.23	1.18	4.87	1.08	5.32
16	1.34	4.29	1.03	5.58	0.79	7.28
25	1.83	3.14	0.97	5.93	0.68	8.46

4.2.5. *Comparison with C*

The three parallel matrix multiplication algorithms have also been implemented in C+PVM using the GMP (Gnu Multi-Precision) library to cope with arbitrary sized integers. The program sizes differ substantially from the parallel functional programs. The sequential C matrix multiplication program using the GMP library consists of 156 lines of code (excluding blank lines and comments), while the parallel programs comprise 378 lines of code for the row-parallel algorithm, 436 lines for the block-parallel version and 457 lines for the torus algorithm. The parallel C+PVM programs are a factor of 4 to 6 longer than our parallel functional programs. Table II shows some runtimes and speedups of the different parallel C+PVM programs for 200×200 matrices of arbitrary precision integers.

The most involved torus-parallel program yields the best parallel runtimes and speedups. While the sequential runtime is a factor of 4 to 6 better, the speedup values progress in a similar way as for the functional programs.

4.3. LinSolv

4.3.1. *Problem Description*

The `linSolv` algorithm discussed in this section finds an exact solution of a linear system of equations of the form $Ax = b$ where $A \in \mathbb{Z}^{n \times n}, b \in \mathbb{Z}^n, n \in \mathbb{N}$. In contrast to more common numerical algorithms, which usually produce an approximate solution over floating point numbers for a given accuracy, the algorithm presented here finds an exact solution and works over arbitrary precision integers.

4.3.2. *Parallel Algorithm*

To find an exact solution for a given system of equations, `linSolv` uses a *multiple homomorphic images* approach (Lauer, 1982). This is a common computer algebra approach and consists of the following three stages:

1. map the input data into several homomorphic images,
2. compute the solution in each of these images, and
3. combine the results of all images to a result in the original domain.

Figure 19 depicts this structure for the implementation of `linSolv`. This structure is particularly useful for operations on arbitrary precision integers. In this case the original domain is $\mathbb{Z}$, the set of all integer values, and the homomorphic images are $\mathbb{Z}$ modulo p, written $\mathbb{Z}_p$, with p being a prime number. If the input numbers are very big and each prime number fits into one machine word the basic arithmetic in the homomorphic images is cheap fixed precision arithmetic. Only in the combination phase, when applying a fold-based Chinese Remainder Algorithm (CRA) (see Lipson (1971)), expensive arbitrary precision arithmetic has to be used to construct the result values. A detailed discussion of several variants of this algorithm is given in Loidl (1997).

The basic parallel structure of the algorithm is one of performing all computations in the homomorphic images in parallel. The Haskell code for the top-level function, which is unchanged for the parallel GPH version, is shown in Figure 20. It uses LU-decomposition followed by forward and backsubstitution to compute the solution `pmx` in the homomorphic image (Press et al., 1992). The main difficulties in the parallel algorithm are two-fold. Firstly, we have to make sure that new results are computed if primes turn out to be "unlucky", i.e. if the determinant of the input matrix A in the homomorphic image generated by this prime number is zero. This can be done either using demand-driven evaluation (GPH) or adding explicit code to handle that case (Eden, PMLS). Secondly, we have to avoid a sequential bottleneck in the combination phase at the end. In earlier papers we have experimented with a tree-based CRA routine to reduce this bottleneck. However, an analysis of the CRA code (Loidl, 1997) reveals that a tree-based CRA algorithm performs much more total computation than a list-based one, due to the more expensive computations at each node of the tree, and we use a list-based CRA in the parallel algorithm.

4.3.3. *Implementations*

4.3.3.1. GPH*:* The parallel GPH version attaches the strategy shown in Figure 21 to the top level expression of the sequential code in the last

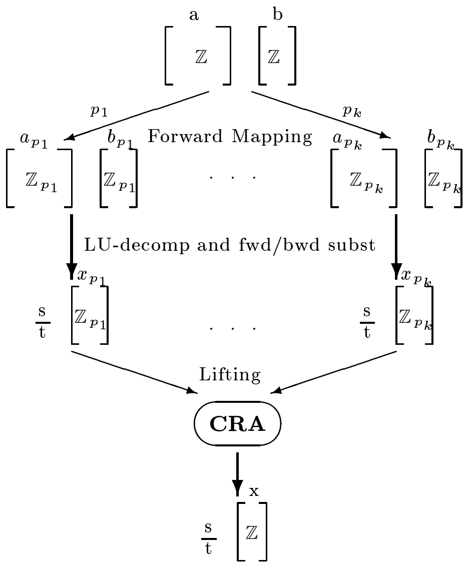

Figure 19. Structure of the `linSolv` algorithm

line of Figure 20. We use an infinite list `xList` representing the results of all homomorphic images together with the prime number, as the basis of the image, and the value of the determinant of A in that image. The strategy guesses the number of primes needed to compute the overall result (`noOfPrimes`) and uses a `parListN` strategy to generate data parallelism over that segment of `xList`. Using `parList` inside the `par_sol_strat` strategy, which is applied to the solution in every image, causes each component of the result to be evaluated in parallel. We need to check whether the determinant is zero to avoid redundant computation. This check is done here, rather than when computing `noOfPrimes` to minimise data dependencies in the algorithm. If some prime numbers turn out to be unlucky the `list_cra` will evaluate the additional results by demanding as-yet-unevaluated list elements. The final strategy application `parList rnf x` specifies that all elements of the result should be combined in parallel.

4.3.3.2. *Eden:* Even though computation in Eden is lazy, communication is eager, except for stream-like lists. Thus, care has to be taken not to send the whole list. To ensure a demand-driven evaluation of homomorphic solutions we use a task farm skeleton as outlined in Section 3.1.1. More specifically, we use the replicated work-

```
linSolv :: SqMatrix Integer ->                -- nxn matrix A
           Vector Integer ->                  -- n vector b
           (Vector Integer, Integer, Integer) -- n vector x s.t. A*x=b
linSolv a b =
  let
   {- Step1: forward mapping -}
    ...
   {- Step2: Computation of solutions in Z/p -}
    ...
    -- Infinite list of hom. solutions of a*x=b in Z_p
    xList = map get_homSol primes

    get_homSol :: Integer -> [Integer]
    get_homSol p  =
      let
        b0 = toHom p b
        a0 = toHom p a
        modDet = toHom p (determinant a0)
        pmx = -- inlined version of: homsolv0 p a0 b0
              let
                lua = lu p a0
                (l,u) = split_lu p lua
                y = fwd_subst p l b0
                x = bwd_subst p u y
              in
                x
      in
      p : modDet : pmx

   {- Step3: lifting via list-based CRA -}
    ...
    primeList = projection 0 xList -- primes (bases for the hom ims)
    detList   = projection 1 xList -- dets in all hom ims
    det       = snd (list_cra pBound primeList detList detList)
    x_i i     = snd (list_cra pBound primeList x_i_List detList)
                where x_i_List = projection (i+2) xList
    -- overall solution:
    x = vector (map x_i [0..n-1])
    ...
  in
  x 'using' strat
```

Figure 20. Top level code of the sequential `linSolv` algorithm (Haskell version)

ers paradigm (Lester, 1993). A manager and a set of worker processes are created, and two tasks are initially released to each of the workers. As soon as any worker finishes a task, it sends the result to the manager, and a new task is delivered to the worker. The computation in the manager is demand-driven and triggered by the availability of result values. As soon as the manager has all the needed results it terminates all the worker processes. Notice that in this *speculative version* the workers may be working speculatively on useless tasks, but only when the useful tasks have already been consumed and hence the degree of speculation is tightly limited. More details about the replicated workers skeleton can be found in Klusik et al. (2002). Figure 22 shows the Eden code for the speculative version of `linSolv`. The only modification

```
strat =
 \ res ->
   rnf noOfPrimes                                      `seq`
   parListN noOfPrimes par_sol_strat xList             `par`
   parList rnf x
   where par_sol_strat :: Strategy [Integer]
         par_sol_strat = \ (p:modDet:pmx) -> rnf modDet `seq`
                                             if modDet /= 0
                                               then parList rnf pmx
                                               else ()
```

Figure 21. Parallel strategy for `linSolv` (GPH version)

```
xList_all = map_rw get_homSol primes

xList = filter lucky xList_all
```

Figure 22. Parallel `linSolv` (Eden speculative version)

to the sequential code is the use of a parallel replicated workers map `map_rw` instead of a sequential map over the infinite list of primes.

To avoid the potential waste of resources due to speculation we can implement a *conservative version* as shown in Figure 23. In this version the prime numbers are divided into those known to be needed (`p_needed`) and those which are only needed if some of the earlier primes are unlucky (`p_spec`). The function `additional` adds for each unlucky prime a new prime number to the task list `primes'`. Note in the definition of `additional` that due to the demand-driven evaluation the availability of unlucky primes in `xs` triggers the generation of one result element in `ys`.

4.3.3.3. *PMLS:* The PMLS implementation has been developed from the sequential Haskell implementation. Arbitrary length integers are provided by Objective Caml's `num` library, whereas GPH and Eden use the GNU gmp library. Replacing the default arithmetic for SML with these arbitrary precision routines exposes some limitations of SML's overloading scheme. In direct comparison this step was easier in Haskell.

The main problem in the PMLS implementation, shown in Figure 24, is the handling of *unlucky primes*. Because SML is strict, new primes cannot simply be demanded during the evaluation of the `map` skeleton. There are two possible solutions to this problem. Either the homomorphic solution function could generate a new prime upon detecting an unlucky one, as it is done in the conservative Eden version, or the forward-mapping phase could be made iterative with the number of valid solution vectors as a convergent. The second of these was implemented since there are problems with generating unique primes within

```
xList_all = map_rw get_homSol primes

xList = filter lucky xList_all
xList_unlucky = filter (not.lucky) xList_all

(p_needed, p_spec) = splitAt ( 1 + toInt noOfPrimes) primes
primes' = p_needed ++ (additional xList_unlucky p_spec)

additional :: [Integer] -> [Integer] -> [Integer]
additional xs ys = zipWith (\ x y -> y) xs ys
```

Figure 23. Parallel `linSolv` (Eden conservative version)

the `map` instance function. Unfortunately this solution to the problem of unlucky primes results in less efficient parallelism for two reasons.

Firstly, in the iterative solution we introduce sequential synchronisation points at the end of each iteration to exchange data between the processors. This is required to guarantee that all processors, computing an element of the result vector, terminate on the same iteration. This nesting of parallelism inside an iterative structure is a general problem with our methodology. To overcome this problem it would be possible to either broadcast the convergent, introducing additional communication, or to define a special *iterative* skeleton, as it is done in the SKIPPER system (Serot, 2001). However, we choose a solution that is more general albeit also more costly.

Secondly, the amount of parallelism is drastically reduced by the `map` call during the first iteration of the `getSols` function. Usually, only one or two unlucky primes are found for modest sizes of problems. If the number of unlucky primes is a multiple of the number of processors (including zero) then there is no parallel performance penalty, otherwise there is a minimum of one homomorphic solution time as an overhead. Additionally, the optimal granularity of the `map` call will be different between the iterations, the first phase more efficient with coarser granularity (since there will be the total number of estimated primes to decompose over), the latter with minimal granularity (since there will only be a small number of unlucky primes). We can set the granularity at runtime but this, currently, requires explicit programmer input. An alternative would be to have dynamic behaviour in our skeletons.

4.3.4. *Performance Results*

As inputs for the performance measurements we use a dense 62×62 matrix of arbitrary precision integers. No element in the matrix is larger than $2^{16} - 1$ and the density of the matrix is higher than 90%. The sequential runtimes show PMLS to achieve best single processor performance with 190.8s, followed by GPH with 381.8s, and Eden with 491.7s. We attribute this fairly large difference mainly to algorithmic differences in the code: The PMLS version uses a more efficient for-

```
(* Solve ax = b modulo p *)
fun gen_xList a b p =
  let
    val (a0,b0) = (matHom p a,vecHom p b)
    val modDet = modHom p (determinant a0)
    val ((iLo,jLo),(iHi,jHi)) = matBounds a
    val pmx =
      fxlist jLo (jHi-jLo+L1)
        (fn j => modHom p (determinant (replaceColumn j a0 b0)))
  in
    p::modDet::pmx
  end

(* Iterative forward mapping phase *)
fun getSols xList [] = xList
  | getSols xList primes =
  let
    val xList' = map (gen_xList aN bN) primes
    val noUnlucky = countUnlucky xList'
    val xList' = filter (not o isUnlucky) xList'
    val primes' = additionalprimes primes noUnlucky
  in
    getSols (xList@xList') primes'
  end
val xList = getSols [] (primesuptomaxprod pBound)

(* Combination via CRA *)
val detList = projection 1 xList
val det = list_cra pBound primes detList detList
fun x_i i =
  let
    val x_i_List = projection (i+2) xList
  in
    list_cra pBound primes x_i_List detList
  end
val x = seqmap x_i (fxlist 0 n (fn x => x))
```

Figure 24. Parallel `linSolv` (PMLS version)

ward substitution after LU decomposition in the homomorphic solution phase. This difference, in combination with the lazy evaluation mechanism used in GpH and Eden, leads to a higher heap consumption resulting in higher overall runtime. Furthermore, due to implementation limitations GpH currently has to use a two-space garbage collector, which is known to be less efficient than the generational garbage collector used by GHC for sequential execution (see below). Finally, the difference between Eden and GpH is due to the fact that Eden uses an older version of GHC.

Figure 25 shows the runtimes and Figure 26 shows the relative speedups for the Eden, GpH, and PMLS implementations of `linSolv` for up to 16 PEs on the Beowulf cluster. For the input data used in these measurements a sufficient number of lucky primes are generated to utilise all processors in the machine. Since these top-level threads

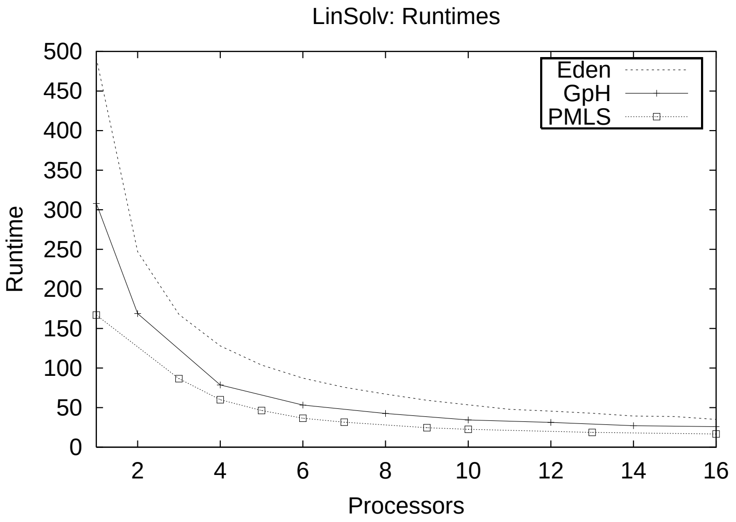

Figure 25. Runtimes of `linSolv` on the Beowulf

can compute their results independently, they perform relatively little communication and the parallel overhead is relatively small giving good parallel efficiency.

A direct comparison of the different languages shows that Eden achieves the best overall speedup on 16 PEs: 14.0, compared to both PMLS and GpH at 11.9. However, since Eden has far higher sequential execution time, the PMLS version is the fastest one on 16 PEs. An examination of the activitiy profiles reveals that PMLS's skeleton maintains more parallelism while collecting the data, whereas in GpH this final stage is mostly sequential.

The Eden measurements use the speculative version with the replicated workers skeleton that dynamically sends work to processes. This approach achieves dynamic load distribution without relying on a potentially expensive implementation of a virtual shared heap, as used in GpH, and the measurements show good speedups even beyond 16 PEs.

In examining the dynamic memory management of all systems, we observe that the total heap allocation on all PEs is highest for PMLS: 1052MB, whereas GpH allocates only 618MB. However, due to higher maximal heap residency in GpH, processors spend on average 19.0% of the total execution time on garbage collection, whereas in PMLS this percentage is only 11.2%. Measuring the heap fragmentation of both systems as the standard deviation of allocation on each processor we obtain similar values for both systems, 147MB for PMLS and 157MB

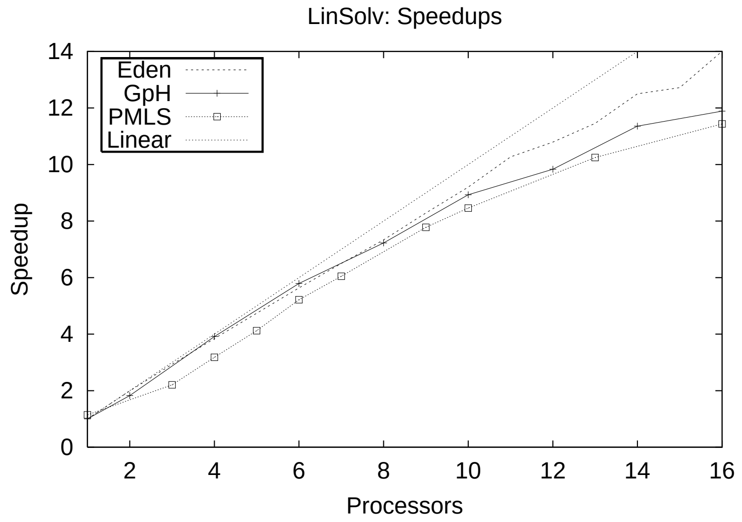

Figure 26. Speedups of `linSolv` on the Beowulf

for GPH. This indicates that in **linSolv**, GPH's dynamic memory management does not dramatically increase heap fragmentation.

As these numbers indicate, GPH's garbage collector seems to generate higher overheads than that in PMLS. The main reason is the current usage of a one-generation copying collector, rather than a real generational collector as supported by GHC for sequential compilation (Sansom and Peyton Jones, 1993). Furthermore, Objective Caml's two-generation collector, as used by PMLS, provides cheap incremental collection for the young generation, which better exploits the additional heap space provided by multiple PEs. The implementation of a mark-and-sweep collector for the older generations is known to be very efficient, too (Doligez and Leroy, 1993). Another potential reason for this overhead is the weighted reference counting on global pointers in GPH, although this overhead allows PEs to collect local garbage independently, avoiding global synchronisation.

In summary the `linSolv` example demonstrates that for some applications lazy evaluation can reduce the amount of coordination required. Both the conservative Eden and the PMLS versions had to introduce additional coordination to model GPH's demand-driven generation of parallelism and to handle unlucky prime numbers. In Eden the speculative version proved to be faster than the conservative version, but in general such an approach bears the danger of wasting resources. Although the static partitioning and mapping of PMLS is generally less flexible than the approach taken in GpH, the re-use of well-tuned parallel skeletons can compensate for the loss in flexibility in this case. It also induces smaller runtime-overheads e.g. for garbage collection. In terms of speedup the skeleton-based versions in Eden and PMLS are more efficient in collecting the results and achieve the following speedups on 16 PEs: 14.0 (Eden), 11.9 (PMLS), 11.9 (GpH), with PMLS having the fastest seuqential execution.

4.4. Ray Tracer

4.4.1. *Problem Description*

The `raytracer` program calculates a 2D image of a scene of 3D objects by tracing all rays in a grid, or window. In tracing a ray, the intersections with the objects are computed. When an intersection is found, the ray is reflected and the colour of the intersection point is computed based on the strength of the ray and on the texture of the object's material. The code is based on the Id version that was published as a part of the Impala suite (Impala, 2001) of parallel benchmark programs.

4.4.2. *Parallel Algorithm*

Figure 27 shows the top-level function of the sequential Haskell algorithm. The function `ray` takes the size of the window in `x` and `y` dimension and the `world`, represented as a list of spheres, as input. The computation proceeds as two nested maps, with the outer map operating over the lines of the grid and the inner map, `do_line`, applying the `tracepixel` function to every point in the grid, represented by the coordinates `(i,j)`, returning a vector representing the colour.

```
ray :: Int -> Int -> [Sphere] -> [[((Int, Int), Vector)]]
ray x y world = map do_line sizes_y
    where do_line :: Int -> [((Int,Int), Vector)]
          do_line i = map (\ j -> ((i,j), f i j)) sizes_x
          sizes_x = [0..x-1]
          sizes_y = [0..y-1]
          f i j = tracepixel world lights i j firstray scrnx scrny
          (firstray, scrnx, scrny) = camparams x y
```

Figure 27. Sequential `raytracer` (Haskell version)

We consider two parallel versions of this program. Both versions exploit data parallelism but differ slightly in the way the data is initially distributed.

4.4.2.1. *Parallel map:* Because the computation to be performed on each pixel, `tracepixel`, is fairly cheap, we do not exploit parallelism in the inner map but instead execute only the outer map in parallel.

To achieve good granularity in the outer loop, the computation over several lines are collected into chunks and processed together.

4.4.2.2. *Direct map:* The direct map version exploits the same kind of data parallelism but differs in its initial distribution of data. Each process is given all necessary data and extracts its own portion of the data by selecting lines in the grid. To improve the granularity of the communication, sub-sequences of pixels are collected into packets. Typically as many tasks as available processors (`noPe`) are generated. To improve the load-balance, task `i` ($0 \leq$ `i` $\leq$ `noPe-1`) computes all result lines `i+j*noPe` with `j` ≥ 0. Note that in this version no dynamic distribution of tasks is required after the startup phase. Compared to the parallel map version this should achieve a faster startup of the parallel processes and a better load distribution.

4.4.3. *Implementations*

4.4.3.1. *PMLS.* The PMLS implementation in Figure 28 uses a parallel map and has been developed from the sequential Haskell version in Figure 27.

```
fun ray x y world =
  let
    val (firstray, scrnx, scrny) = camparams x y
    fun do_pixel ij =
      let
        val (i,j) = (ij div 1000,ij mod 1000)
      in
        ((i,j),tracepixel world lights (real i) (real j) firstray scrnx scrny)
      end
    val ind = indxs 0 (x - 1) 0 (y - 1)
  in
    map (map do_pixel) ind
  end
```

Figure 28. Parallel `raytracer` (PMLS version)

The function `do_pixel` initiates ray tracing at pixel co-ordinate `(i,j)` and is mapped over the index image `ind`. The same considerations regarding granularity control apply to PMLS. However, the choice whether to do the outer or inner `map` in parallel is determined by the characteristics of the PMLS runtime-system. When two skeletons are direct arguments to each other, as in the example here `(map (map do_pixel))`, there is no advantage in nested implementation since granularity control is performed automatically inside the `map` skeleton. In addition, the PMLS compiler requires all free variables in the function position of a `map`, in this case `do_pixel`, to be sent to the individual processors when initialising the skeleton instance. Since the inner `map` has free values which could (potentially) change between successive calls they have to be re-transmitted upon each call. This means that the total amount of data transmitted is significantly less if the outer map is implemented in parallel.

```
ray :: Int -> Int -> Int -> [Sphere] -> [[((Int, Int), Vector)]]
ray chunk x y world = map do_line sizes_y
                      `using` parListChunk chunk rnf
```

Figure 29. Parallel `raytracer` (GPH version)

4.4.3.2. GPH. The GPH implementation is based on the parallel map version and uses an additional explicit parameter `chunk` to control the size of the chunks. The code in Figure 29 shows the body of the function `ray` (the local definitions are unchanged), with an evaluation strategy implementing granularity control via clustering. We use the same `parListChunk` strategy as in the row-clustered matrix multiplication.

```
ray :: Int -> Int -> [Sphere] -> [[((Int, Int), Vector)]]
ray x y scene = shuffleN outps
    where outps = [ (process i -> f_dm i) # void | i <- [0..noPe-1]]
                      `using` seqList r0
          f_dm n _  = map do_line (takeEach noPe (drop n sizes_y))
```

Figure 30. Parallel direct map version of `raytracer` (Eden version)

4.4.3.3. *Eden.* The Eden implementation uses the direct-map version and is shown in Figure 30. The function `f_dm` represents the work to be executed by one process. In the direct-map version this includes the extraction of its own portion of the input data using the `takeEach` function to combine every n-th line of the grid into one chunk.

The processes are created in a list of process instantiations (`outps`). The sequential strategy `seqList r0` is used to drive an eager process creation, creating the processes before the outport values are needed.

4.4.4. *Performance Results*

The measurements in Figures 31 and 32 use a 350×350 image with a chunk size of 10 and a scene consisting of 640 spheres as input. The sequential runtimes are: 176.99s for Eden, 163.31s for GPH, and 172.10s for PMLS. For this application the sequential performance of all three versions is fairly similar with a variation of less than 10%. This is mainly due to the fact that `raytracer` does not make use of the laziness in the language: all parts of the picture are indeed computed and since there

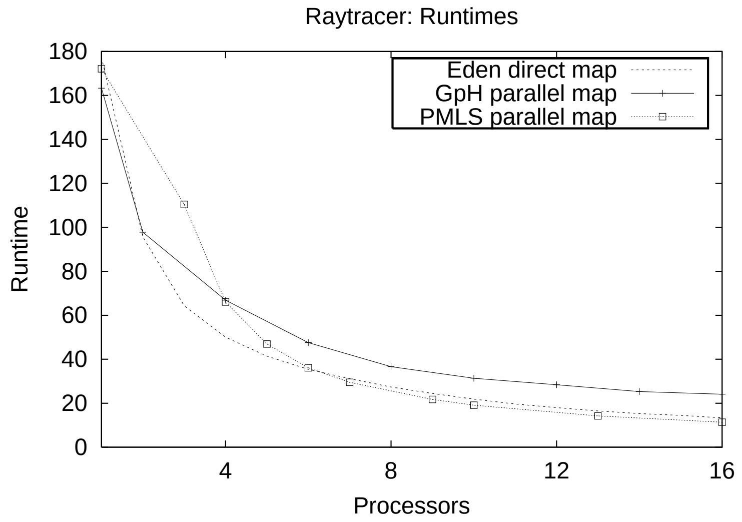

Figure 31. Runtimes for `raytracer` on the Beowulf (in seconds)

is no interaction they can be computed eagerly. This dynamic program characteristic manifests itself in similar garbage collection overheads for PMLS and GpH: 3.3% and 3.1% as mean over all processors.

For PMLS initial sequential results showed significantly poorer performance than the GpH and Eden versions. This is due to a known limitation of the PMLS compiler. The results reported here required some minimal user interaction during the compilation process. The PMLS group is currently adding an appropriate sequential optimisation step to the compiler.

The rather simple and regular structure of the computation lends itself to a static approach such as the static task farm in PMLS or the direct-map in Eden. The partitioning of the program can be achieved statically and the distribution of work is carried out only once at the beginning of the program. Since the work is fairly evenly distributed, no sophisticated dynamic load balancing is necessary. On the other hand enforcing a fixed data distribution is easier in Eden than in GpH. In general, the more dynamic facilities of GpH are not used in this application. We have experimented with GpH versions that model the Eden approach more closely, but they did not yield any significant performance improvements.

Not surprisingly for an application with a fairly regular structure of parallelism, PMLS performs best in terms of speedup as well as absolute runtime. On 16 processors the runtime is 11.4s, corresponding

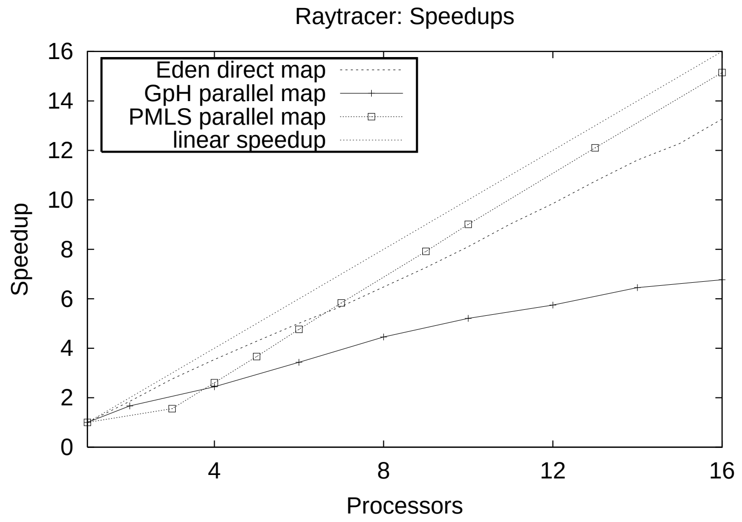

Figure 32. Speedups for `raytracer` on the Beowulf

to a relative speedup of 15.2. The results for Eden, with its slightly more dynamic resource management, are similar: parallel runtime of 13.4s with a speedup of 13.3. GPH pays a higher cost for its dynamic resource management, resulting in a comparatively poor speedup of 6.8 on 16 processors and a parallel runtime of 24.1s.

Another problem we have observed in the GPH version is a potentially poor load distribution where few processors monopolise the entire available parallelism. This is due to a combination of factors: in this program all parallelism is generated on the main processor at the beginning of the computation, and on the Beowulf start-up times between PEs may vary significantly, moreover the GPH runtime-system does not currently allow tasks to migrate from a loaded PE to an idle PE. Hence the fastest processor(s) sometimes obtain all available work before the slower processors have a chance to send their first work requests. It is possible to crudely control the work distribution by imposing an upper limit on the number of threads that may be alive on one PE, and that is what we used in these measurements.

In contrast, for PMLS load-balancing is assumed to be a property of the skeleton implementation. The parallel map skeleton used by all applications has a degree of implicit load-balancing as a result of the processor farming model. This works well in cases like `raytracer` but requires manual tuning for particular instances which can change as execution proceeeds (for example in `linSolv` different balancing strategies are used for the initial and the additional results). Eden's replicated worker skeleton `map_rw` as used in the `linSolv` example provides implicit dynamic load balancing based on the master worker paradigm. Surprisingly, this skeleton is outperformed for the raytracer by a static work distribution where the work list is sent to all processors and the work packages are selected locally within each process.

In summary, the results for `raytracer` underline a general trend in these measurements for Eden and GPH, namely the impact of dynamic resource management overheads on scalability. Eden, which has a lower overhead, performs almost as well as PMLS. However GPH has to maintain a virtual shared heap, and this diminishes parallel performance for larger numbers of processors. In some cases we have observed an overhead of up to 16% of the total execution time, although typical percentages are 3-8% (Loidl, 2002). We are also investigating refined load balancing mechanisms, which show better performance.

5. Related Work

For comprehensive overviews on parallel functional programming we refer to Hammond and Michaelson (1999) and Trinder et al. (2002). In this section we focus on comparing our approaches with other implemented systems. Only few implementations have overcome a purely experimental status and concrete head-to-head comparisons of different languages on the same architecture are even rarer. To our knowledge this paper is the first such systematic comparison.

5.1. Skeleton-based Approaches

The prospect of implicit parallelism with the use of skeletons has spurred the development of several skeletons-based systems. HDC (Herrmann, 2000) is a strictly-evaluated subset of Haskell with skeleton-based coordination, in particular support for `fold` and `map`, and several forms of divide-and-conquer. For the Karatsuba algorithm for polynomial multiplication HDC achieves a relative speedup of 363 on 729 processors of a 1024-processor transputer-based Parsytec machine.

A system closely related to GPH is Caliban (Kelly, 1989; Taylor, 1996) in which `moreover` clauses, similar to GPH's `using`, can be attached to sequential program source in order to specify parallel behaviours. Expressions are annotated to indicate tasks to be created, and the linkage between the tasks can be specified using normal functions. In the current implementation, the process network is static, with moreover clauses being resolved at compile-time and processes being statically mapped to the target topology. A simple raytracer, introduced by Kelly (1989), has been measured on a 128 processor Fujitsu AP1000, achieving speedups of up to 24 on 35 processors.

Other prominent skeleton-based systems are SCL (Darlington et al., 1996a) and P3L (Bacci et al., 1995). Both use separate coordination languages with small sets of basic skeletons that can be freely nested. The most mature implementation of SCL, SPF, uses Fortran as computation language. Substantial applications such as a Barnes-Hut algorithm have been implemented in SPF (Darlington et al., 1996b) and measured on a Fujitsu AP1000. Pelagatti (2002) presents performance results for P3L on four applications, including a parallel raytracer, obtained with the `SkIE` prototype environment for P3L on a 24-node Meiko CS-2 and an 8-PC Linux cluster.

An active research area in the skeletons community is the nesting of skeletons (Hamdan, 2000). In particular, with support for nesting it is possible to construct complex parallel applications by composing and transforming skeletons using given transformation rules and compositional cost models for performance prediction as developed by Pepper (1993) and Bacci et al. (1999).

5.2. Thread- and Process-based Approaches

Para-functional programming, as introduced by Hudak (1986), is the general approach of adding control directives to a functional program in order to specify parallel execution. These control directives allow the programmer to describe detailed schedules of the execution as well as a particular mapping of threads to processors. First-class schedules (Mirani and Hudak, 1995) extend para-functional programming to Haskell, using monads to separate expressions and control directives. These annotations usually describe potential parallelism, in the sense of GpH's `par`, and therefore represent a thread-based approach. Its implementation builds on the concept of futures, as used in MultiLisp (Halstead, 1985). First-class schedules have been implemented by compiling Haskell to the MultiLisp-based operating system Sting. Preliminary performance results on a 16 processor Silicon Graphics Challenge shared-memory machine show good speedups for a parallel Barnes-Hut algorithm for solving the n-body problem (Mirani, 1996).

ALFL (Goldberg, 1988) is an LML-like, lazy, implicitly-parallel functional language, implemented on a distributed-memory Intel Hypercube as well as on a shared-memory Encore machine, with performance comparisons between the two architectures.

Concurrent Clean (Plasmeijer et al., 1999; Nöcker et al., 1991) is a lazy language with parallelism annotations. In his PhD thesis Kesseler

(1996) quotes performance results for three systems: Concurrent Clean on the ZAPP abstract machine; Concurrent Clean on the PABC abstract machine; and a Miranda-like, implicitly-parallel, lazy language, implemented on the abstract HDG machine (Kingdon et al., 1991). All measurements have been performed on (different) transputer networks. In contrast to this paper, no detailed comparison of languages or systems is given. Kesseler (1996) reports good speedups for small programs such as nqueens (5.6 on 8 processors) but poorer results for a raytracer (3.9 on 16 processors) in his implementation.

5.3. Other Approaches

One of the most successful data parallel functional languages is NESL (Blelloch, 1996). NESL is a strict, strongly-typed, data-parallel language with implicit parallelism and implicit thread interaction. It has been implemented on a range of parallel architectures, including vector machines. A wide range of algorithms have been parallelised in NESL, including a Delaunay algorithm for triangularisation (Blelloch et al., 1996), several algorithms for the n-body problem (Blelloch and Narlikar, 1997), and several graph algorithms. The focus in these papers, however, is on the comparison and improvement of algorithms rather than speedup measurements or a comparison with other languages. Two data-parallel extensions of Haskell have been partially implemented: Data Field Haskell (Holmerin and Lisper, 2000) and Nepal (Chakravarty et al., 2001). No performance results are available.

SISAL (Cann, 1992) is a first-order, strict functional language with implicit parallelism and implicit thread interaction. Its implementation is based on a dataflow model and it has been ported to a range of parallel architectures. Good absolute performance in comparison to Fortran code is quoted in LANL (2001).

The pHLuid system (Flanagan and Nikhil, 1996) is a parallel implementation of Id on networks of workstations. It uses a dataflow model of computation in order to achieve implicit parallelism. Id is polymorphic, higher-order and has a non-strict semantics, implemented via lenient or parallel eager evaluation. A fusion of Id and Haskell, called pH, has been proposed (Nikhil and Arvind, 2001) but no implementation is available, yet. Flanagan and Nikhil (1996) report near-linear (relative) speedups on a workstation cluster for simple programs such as nqueens and matrix multiplication. Hammes et al. (1995) present a rare language and performance comparison of implicitly parallel Id with sequential Haskell on a realistic benchmark program.

6. Conclusions

We have compared three state-of-the-art parallel functional programming systems (PMLS, GpH, and Eden) and evaluated their performance on a Beowulf architecture using three symbolic applications: several matrix multiplication algorithms using arbitrary precision arithmetic (`matMult`); an exact linear system solver (`linSolv`); and a simple ray-tracer (`raytracer`).

PMLS, GpH and Eden all aim to support parallel symbolic computations at low programmer cost. While it is relatively straightforward to achieve good (often linear) speedups for *regular, numerical* parallel computations, it can be much harder, or even impossible, to achieve the same results for *irregular, symbolic* computations, especially those with complex data structures or irregular task structures (Loidl et al., 1999). Relatively small performance improvements may thus be of much greater significance to users of such systems. At the same time, symbolic application programmers are usually domain experts rather than computer scientists, and are often unwilling or unable to invest major effort in recoding for parallelism. In this section, we will evaluate the three systems in terms of language features, performance, and productivity. We will consider them in order of anticipated programmer effort: namely PMLS, GpH, and Eden.

6.1. Summary Language Comparison

All three functional languages aim to provide higher-level models of parallelism, with the objective of reducing programmer overhead. All three abstract over low-level details of communication timing, data structure marshaling (including cyclic graph structures) and synchronisation that must be specified in e.g. C+PVM. Moreover, in all three languages, details of task/thread creation and program decomposition are delegated to the compilation system.

PMLS provides a convenient model of implicit parallelism using *skeletons* — a set of pre-defined higher-order functions with associated parallel behaviours. Since skeletons are partitioned into parallel components and mapped to processing units statically, this approach has the lowest runtime overhead of the three considered here, and where the application structure fits the pre-defined skeletons perfectly, it will also have the lowest programmer overhead. However, such an approach is less flexible than the dynamic approaches taken by Eden and GpH. This is apparent in less regular or longer-running applications, such as `linSolv`, where a regular static structure cannot be determined from the program source.

GpH has a similar philosophy to that of PMLS, aiming to require minimal programmer input in order to achieve acceptable parallel performance. However, it provides more control (if required) over evaluation order, strictness and parallelism, allowing programmable evaluation strategies to be developed. This approach trades low programmer overhead for a variety of programming styles for a potentially high dynamic overhead. This cost is most apparent in regular applications, where a simple static process to processor mapping could be determined either manually or automatically. In such a case, manual tuning may be needed to extract good parallel performance for GpH, where PMLS might automatically find such a mapping, or it might be straightforward to program such a mapping in Eden. `raytracer` is an example of such simple static mapping.

Finally, of the three languages studied here, Eden provides the greatest control over parallelism, and thus requires the greatest programmer effort. Control is provided over task decomposition, allocation to virtual processors and communication channels. Given sufficient tuning effort, it is possible to develop more sophisticated parallel algorithms, as with the torus version of `matMult` (Section 4.2.2). As with GpH, all PMLS skeletons can be easily replicated (Klusik et al., 2000; Peña and Rubio, 2001), with a similar mapping effect. Since all load management details must be explicitly programmed, however, and there is no support for lazy communication there will be situations where GpH mechanisms cannot be easily replicated, such as using a potentially infinite number of homomorphic images in `linSolv`.

Recognising the value of the skeletons approach for suitable applications, all three languages provide support for such a style. PMLS naturally provides the most direct support, with static process mapping and cost modeling as part of the compilation process. GpH provides a full set of standard skeletons written in Haskell, and using a dynamic cost model and mapping (Hammond and Rebon Portillo, 1999). Haskell's constructor classes are used to abstract over machine models and alternative data structures. Finally, a rich set of skeletons, including some novel branch-and-bound skeletons, has been developed using Eden constructs and used on several parallel machines (Klusik et al., 2000; Peña and Rubio, 2001; Du Bois et al., 2002).

6.2. Performance Comparison

It is received wisdom that eager evaluation (used for strict function calls) will outperform lazy evaluation (used for non-strict function calls) due to the overhead of recording partial results in the latter case. It follows that fully strict languages should outperform non-strict languages

Table III. Comparative Performance (Seq RT: runtime on a 1 PE parallel machine; Par RT: runtime on a 16 PE parallel machine; Spdup: Speedup on 16 PEs calculated as $\frac{\text{Seq RT}}{\text{Par RT}}$)

	Eden			GPH			PMLS		
	Seq RT	Par RT	Spdup	Seq RT	Par RT	Spdup	Seq RT	Par RT	Spdup
`matMult`	38.5s	13.2s	2.9	30.3s	8.9s	3.4	22.8s	4.3s	5.2
`linSolv`	491.7s	35.1s	14.0	307.9s	25.9s	11.9	190.8s	16.1s	11.9
`raytracer`	177.4s	13.4s	13.3	163.3s	24.1s	6.8	172.1s	11.4s	15.2

(experimental results suggest that this can be over a factor of 10 in the worst benchmark cases (Hartel et al., 1996)).

Similarly, it is argued that full communication should outperform lazy communication, since fewer messages are required in the former case if an entire data structure is communicated. Given that PMLS is fully strict, with strict communication, Eden is non-strict, with strict communication and GPH is non-strict with lazy communication, we would consequently expect PMLS to outperform Eden which should outperform GPH[1]. We would also expect the same ordering on the basis of runtime overheads, but with the possibility of similar overheads for Eden and PMLS. The performance results summarised in Table III are therefore somewhat surprising.

For all three benchmarks PMLS achieves the smallest execution times. In the case of `linSolv`, Eden's speedup is higher but sequential execution time is higher, too. GPH achieves similar speedup as PMLS with sequential time between the other two versions. In the case of the `raytracer` (the most regular of the three benchmarks we have considered) PMLS shows even better speedups than Eden or GPH. While mirroring earlier results almost exactly (Hammond and Rebon Portillo, 1999), the GPH performance for the `raytracer` is distinctly disappointing. This has subsequently led us to improve the GPH load distribution mechanism (Loidl, 2001).

For comparison, we have re-implemented the `matMult` benchmark in C+PVM using the Gnu Multi-Precision library for arbitrary precision arithmetic (Table IV) and the GNU C compiler on the same parallel machine. For the block-parallel version (the only one implemented in all four systems), the speedup results using full C optimisation (-O2)

[1] This does discount the maturity of the optimising Glasgow Haskell Compiler, which forms the basis for the Eden and GPH implementations.

Table IV. Comparative Performance of Matrix Multiplication in C (Seq RT: runtime on a 1 PE parallel machine; Par RT: runtime on a 16 PE parallel machine; Spdup: Speedup on 16 PEs calculated as $\frac{\text{Seq RT}}{\text{Par RT}}$)

	C			Eden		GpH		PMLS	
	Seq RT	Par RT	Spdup	Seq RT	Spdup	Seq RT	Spdup	Seq RT	Spdup
`matMult`(rows)	5.75s	1.34s	4.3	34.3s	1.7	–	–	19.5s	3.3
`matMult`(block)	5.75s	1.03s	5.6	32.9s	2.1	30.3s	3.4	22.8s	5.3
`matMult`(torus)	5.75s	0.79s	7.3	38.5s	2.9	–	–	–	–

are comparable with those for PMLS: 5.6 on 16 processors. The base sequential performance is, however, a factor of 3-6 better than for the functional languages. This is a smaller factor than might have been expected for functional versus imperative code, especially given the use of list structures rather than in-place arrays in the functional versions. We anticipate that the difference could be further reduced by using e.g. monadic techniques to allow in-place array updates, but at some cost in source code legibility/programmer time.

6.3. Productivity Comparison

Measuring programmer productivity is notoriously difficult, due to differences in individual ability, prototyping effects, etc. We have therefore chosen to use lines of code as a reasonable approximation[2]. It is accepted that the number of lines of code produced by any given trained programmer is roughly constant regardless of the programming language used or the general ability of the programmer. Whilst not wishing to overstate our findings and accepting that small variations may not be significant, we therefore believe that this provides a fair assessment of expected productivity for the applications studied.

Table V gives the number of lines of code for each of the three programs that have been studied here, plus corresponding figures for

[2] We were unable to satisfactorily separate application development time from experimentation and system development, as we would have liked, so have omitted these figures. Even with profiling and tuning (which, as we have shown is essential to achieving good performance), it is unlikely that these would exceed 1-2 days effort for each of the functional programs. In comparison, the development time for the (very simple) parallel C program was more than 1 week, without undertaking performance tuning or optimisation, and with the benefit both of prototyping from the functional code and of code reuse from other parallel applications.

Table V. Productivity Comparison (in lines-of-code)

	Eden		GpH		PMLS		C	
	Seq Size	Par Size	Seq Size	Par Size	Seq Size	Par Size	Seq Size	Par Size
`matMult`	68	34 (10)	68	5	85	25 (3)	156	301
`linSolv`	473	8	473	10	751	13	—	—
`raytracer`	453	10	453	7	410	3	—	—

the arbitrary precision matrix multiplication program in C+PVM. The counts exclude comments and white space. The parallel code size represents the number of lines that were either changed in or newly written for the parallel version. As expected, these changes are highly significant for the C program (representing some 65% of the total code size), but are generally insignificant for the functional programs (in the worst case, representing 33% of the total code size, for a highly tuned version of the matrix multiplication algorithm). The sequential functional programs are a factor of 2-3 times shorter than the C equivalent, with the parallel programs being 4 to 6 times shorter. Clearly, a certain amount of the C code could be reused for other applications, but there is equally clearly a very high entry price to parallel programming in C, especially when complex data structures must be communicated. The functional code describing such complex dynamic parallel program structures is very concise. Although not a major difference, the sequential Haskell code is generally slightly shorter than the SML code. This is mainly a consequence of better standard library support for Haskell, though high-level language features such as overloading and list comprehensions have also been exploited.

Since PMLS is the most implicit approach of the three languages studied, and Eden the most explicit, we would anticipate that PMLS should require least changes with Eden requiring the most changes. While this is generally true, the figures are distorted to some extent by the performance tuning that has been carried out. Although the initial version of the `matMult` in PMLS required only 3 lines, the final tuned version required 25. The corresponding Eden figures are 10 lines and 34 lines, respectively. The GpH code was not tuned, however, and therefore only 5 lines in total were changed. The `linSolv` application showed a reversal of the general result, with Eden requiring fewest changes. This may reflect the poor match between the irregular parallel structure of this application and the standard skeletons/strategies used by the other two systems. It is worth noting that the total number of changed lines is generally small, and that our comparisons must therefore be correspondingly tentative.

We conclude that while C may offer better performance than unoptimised functional code, the difference is less than might be expected. Moreover, the high level features available in functional code mean that programmer productivity is likely to be much greater than in C.

6.4. Maturity and Usability

All three functional language systems discussed here can be rated as *mature research systems*, running a range of parallel benchmark applications on a variety of parallel architectures. Work on GpH began in 1994, and it since has been applied to numerous programs, including the 40,000 line Lolita natural language engineering system (Loidl et al., 1997). To assist program development, it offers a sophisticated set of profiling tools (Hammond et al., 2000), including ideal and realistic simulation. GpH is publicly available in OpenSource form as part of the GHC compiler project (WWW-GPH, 2001).

The Eden system is a later development, sharing underlying parallel scheduling and communication infrastructure with the earlier GpH system. It has been tested on a variety of small and medium benchmark applications, but has not yet been applied to large-scale applications, such as Lolita. Both GpH and Eden provide low-level portability by compiling through either C+PVM or C+MPI.

In contrast to the two GHC-based systems, the PMLS system has a more heterogeneous structure, exploiting state-of-the-art implementation technology from several sources. The core system uses the Objective Caml compiler for sequential compilation and calls C+MPI routines for implementing the parallel skeletons. Up-to-date versions of Eden and PMLS are available from the developers on request.

7. Future Work

All three systems are under active development. For PMLS the current objectives are to provide a more expressive set of algorithmic skeletons, to optimise the performance of the existing skeletons and to automatically identify skeleton structures in arbitrary code. This work will exploit both dynamic profiling-based performance prediction (which has been found to give good predictions within a narrow range of program characteristics) and automatic program transformation techniques.

The main research direction for GPH is to improve architecture-independence by refining the mechanisms for load balancing and data distribution in order to deal with high-latency machines such as Beowulf clusters. Based on these refinements, research will focus on the development of an adaptive runtime-system capable of automatically tuning its behaviour to suit the characteristics of the parallel machine.

Finally, following the upgrade to conform to the latest sequential GHC compiler, work in Eden will focus on optimisations to reduce communication costs.

Acknowledgements

The authors would like to thank the following organisations for funding the work presented here: the Austrian Academy of Sciences (APART fellowship 624), the Japan Society for the Promotion of Science (Postdoctoral fellowship P00778), the Spanish Project TIC2000-0738, and UK's Engineering and Physical Sciences Research Council (grants EPSRC GR/L 93379, EPSRC GR/M 32351 and EPSRC GR/L42889). Our cooperative work would not have been possible without the travel funding from the following organisations: the British Council and the DAAD (grant 1097), and the Spanish-British Acción Integrada HB 1999-0102.

References

Bacci, B., M. Danelutto, S. Orlando, S. Pelagatti, and M. Vanneschi: 1995, 'P^3L: A Structured High Level Programming Language and its Structured Support'. *Concurrency — Practice and Experience* **7**(3), 225–255.

Bacci, B., S. Gorlatch, C. Lengauer, and S. Pelagatti: 1999, 'Skeletons and Transformations in an Integrated Parallel Programming Environment'. In: *PACT'99 — Intl. Conf. on Parallel Architecture and Compilations Techniques*, Vol. LNCS 1662. pp. 13–27, Springer.

Blelloch, G.: 1996, 'Programming Parallel Algorithms'. *Communications of the ACM* **39**(3), 85–97.

Blelloch, G., G. Miller, and D. Talmor: 1996, 'Developing a Practical Projection-Based Parallel Delaunay Algorithm'. In: *Symp. on Computational Geometry.* Philadelphia, PA, May 24–26, pp. 186–195, ACM.

Blelloch, G. and G. Narlikar: 1997, 'A Practical Comparison of N-Body Algorithms'. In: *Parallel Algorithms*, Vol. 30 of *Series in Discrete Mathematics and Theoretical Computer Science.* American Mathematical Society.

Breitinger, S., U. Klusik, R. Loogen, Y. Ortega Mallén, and R. Peña Marí: 1997a, 'DREAM - the DistRibuted Eden Abstract Machine'. In: *IFL '97 — Intl. Workshop on the Implementation of Functional Languages 1997*, Vol. 1467 of *LNCS.* Univ of St Andrews, Scotland, pp. 250–269, Springer.

Breitinger, S., R. Loogen, Y. Ortega Mallén, and R. Peña Marí: 1997b, 'The Eden Coordination Model for Distributed Memory Systems'. In: *HIPS'97 — Workshop on High-level Parallel Programming Models*. Geneva, Switzerland, pp. 120–124, IEEE Computer Science Press.

Burge, W.: 1975, *Recursive Programming Techniques*. Addison-Wesley.

Cann, D.: 1992, 'Retire Fortran? A Debate Rekindled'. *Communications of the ACM* **35**(8), 81–89.

Chakravarty, M., G. Keller, R. Lechtchinsky, and W. Pfannenstiel: 2001, 'Nepal — Nested Data-Parallelism in Haskell'. In: *EuroPar'01 — European Conf. on Parallel Processing*, Vol. 2150 of *LNCS*. August 28–31, Manchester, U.K., pp. 524–534, Springer.

Cole, M. I.: 1989, *Algorithmic Skeletons: Structured Management of Parallel Computation*, Research Monographs in Parallel and Distributed Computing. Cambridge, MA: The MIT Press.

Cook, A.: 2001, 'Transformation and Proof in a Parallelising Compiler'. Ph.D. thesis, Dept. of Computing and Electrical Engineering, Heriot-Watt University.

Darlington, J., Y. Guo, and H. To: 1996a, 'Structured Parallel Programming: Theory Meets Practice'. In: R. Milner and I. Wand (eds.): *Research Directions in Computer Science*. Cambridge University Press.

Darlington, J., Y. Guo, H. To, and J. Yang: 1996b, 'SPF: Structured Parallel Fortran'. In: *PCW'96 — Intl. Parallel Computing Workshop*. Kawasaki, Japan.

Doligez, D. and X. Leroy: 1993, 'A concurrent, generational garbage collector for a multithreaded implementation of ML'. In: *POPL'93 — Symposium on Principles of Programming Languages*. pp. 113–123, ACM Press.

Du Bois, A., R. Pointon, H.-W. Loidl, and P. Trinder: 2002, 'A Declarative Parallel Bottom-Avoiding Choice'. In: *SBAC-PAD 2002 — Symp on Computer Architecture and High Performance Computing*. Vitoria, Brazil. Submitted for publication.

Flanagan, C. and R. Nikhil: 1996, 'pHluid: The Design of a Parallel Functional Language Implementation on Workstations'. In: *ICFP'96 — Intl. Conf. on Functional Programming*. Philadelphia, Pennsylvania, pp. 169–179, ACM Press.

Frens, J. and D. Wise: 1997, 'Auto-blocking matrix multiplication, or tracking BLAS3 performance from source code'. *PPoPP'97 — Symp on Principles and Practice of Parallel Programming* **32**(7), 206–216.

Goldberg, B.: 1988, 'Multiprocessor Execution of Functional Programs'. *Intl. J. of Parallel Programming* **17**(5), 425–473.

Halstead, R.: 1985, 'Multilisp: A Language for Concurrent Symbolic Computation'. *ACM Transactions on Programming Languages and Systems* **7**(4), 106–117.

Hamdan, M.: 2000, 'A Combinational Framework for Parallel Programming Using Algorithmic Skeletons'. Ph.D. thesis, Dept. of Computing and Electrical Engineering, Heriot-Watt University.

Hammes, J., O. Lubeck, and W. Böhm: 1995, 'Comparing Id and Haskell in a Monte Carlo Photon Transport Code'. *J. of Functional Programming* **5**(3), 283–316.
<URL:http://www.cs.colostate.edu/~hammes/documents/final1.ps.Z>.

Hammond, K., D. King, H.-W. Loidl, A. Rebón Portillo, and P. Trinder: 2000, 'The HasPar Performance Evaluation Suite for GpH: a Parallel Non-Strict Functional Language'. *Software – Practice and Experience*.
<URL:http://www.cee.hw.ac.uk/~dsg/gph/papers/ps/spe.ps.gz>.

Hammond, K., H.-W. Loidl, and A. Partridge: 1995, 'Visualising Granularity in Parallel Programs: A Graphical Winnowing System for Haskell'. In: *HPFC'95*

— *Conf. on High Performance Functional Computing*. Denver, CO, pp. 208–221. <URL:http://www.cee.hw.ac.uk/~hwloidl/publications/Glasgow/HPFC95.ps.gz>.

Hammond, K. and G. Michaelson (eds.): 1999, *Research Directions in Parallel Functional Programming*. Springer.

Hammond, K. and A. J. Rebon Portillo: 1999, 'HaskSkel: Algorithmic Skeletons for Haskell'. In: *IFL'99 — Intl. Workshop on the Implementation of Functional Languages*, Vol. 1868 of *LNCS*. Lochem, The Netherlands, Springer. <URL:http://www-fp.dcs.st-and.ac.uk/publications/1999/haskskel.ps.gz>.

Hartel, P., M. Feeley, M. Alt, L. Augustsson, P. Baumann, M. Beemster, E. Chailloux, C. Flood, W. Grieskamp, J. van Groningen, K. Hammond, B. Hausman, M. Ivory, R. Jones, J. Kamperman, P. Lee, X. Leroy, R. Lins, S. Loosemore, N. Röjemo, M. Serrano, J.-P. Talpin, J. Thackray, S. Thomas, P. Walters, P. Weis, and P. Wentworth: 1996, 'Benchmarking Implementations of Functional Languages with "Pseudoknot", a Float-Intensive Benchmark'. *Journal of Functional Programming* **6**(4).

Hernández, F., R. Peña, and F. Rubio: 2000, 'From GranSim to Paradise'. In: *SFP'00 — Scottish Functional Programming Workshop*. pp. 11–19, Intellect.

Herrmann, C.: 2000, 'The Skeleton-Based Parallelization of Divide-and-Conquer Recursions'. Ph.D. thesis, University of Passau.

Holmerin, J. and B. Lisper: 2000, 'Development of Parallel Algorithms in Data Field Haskell'. In: *EuroPar'00 — European Conf. on Parallel Processing*. Munich, Germany, pp. 762–766, Springer.

Hudak, P.: 1986, 'Para-Functional Programming'. *IEEE Computer* **19**(8), 60–70.

Hughes, R.: 1989, 'Why Functional Programming Matters'. *The Computer Journal* **32**(2), 98–107. <URL:http://www.cs.chalmers.se/~rjmh/Papers/whyfp.ps>.

Impala: 2001, 'Impala – (IMplicitly PArallel LAnguage Application Suite)'. <URL:http://www.csg.lcs.mit.edu/impala/>.

Karatsuba, A. and Y. Ofman: 1962, 'Multiplication of Multi-digit Numbers on Automata'. *Soviet. Phys. Dokl.* (7), 595–596.

Kelly, P.: 1989, *Functional Programming for Loosely-Coupled Multiprocessors*, Research Monographs in Parallel and Distributed Computing. MIT Press.

Kesseler, M.: 1996, 'The Implementation of Functional Languages on Parallel Machines with Distributed Memory'. Ph.D. thesis, University of Nijmegen.

King, D., J. Hall, and P. Trinder: 1998, 'A Strategic Profiler for Glasgow Parallel Haskell'. In: *IFL'98 — Intl. Workshop on the Implementation of Functional Languages*, Vol. 1595 of *LNCS*. University College London, UK, pp. 88–102.

Kingdon, H., D. Lester, and G. Burn: 1991, 'The HDG-machine: a Highly Distributed Graph-Reducer for a Transputer Network'. *Computer Journal* **34**(4), 290–301.

Klusik, U., R. Loogen, S. Priebe, and F. Rubio: 2000, 'Implementation Skeletons in Eden — Low-Effort Parallel Programming'. In: *IFL'00 — Intl. Workshop on the Implementation of Functional Languages*, Vol. 2011 of *LNCS*. Aachen, Germany, pp. 71–88, Springer.

Klusik, U., R. Peña, and F. Rubio: 2002, 'Replicated Workers in Eden'. In: *CMPP'00 — Constructive Methods for Parallel Programming*. Ponte di Lima, Portugal, Nova Science Books and Journals. To appear.

LANL: 2001, 'Sisal Performance Data'. WWW page. <URL:http://www.llnl.gov/sisal/PerformanceData.html>.

Lauer, M.: 1982, 'Computing by Homomorphic Images'. In: B. Buchberger, G. E. Collins, R. Loos, and R. Albrecht (eds.): *Computer Algebra — Symbolic and Algebraic Computation*. Springer, pp. 139–168.

Lester, B.: 1993, *The Art of Parallel Programming*. Prentice-Hall.

Lipson, J. D.: 1971, 'Chinese Remainder and Interpolation Algorithms'. In: *SYMSAM'71 — Symp. on Symbolic and Algebraic Manipulation*. pp. 372–391, Academic Press.

Loidl, H.-W.: 1997, 'LinSolv: a Case Study in Strategic Parallelism'. In: J. O'Donnell (ed.): *Glasgow Workshop on Functional Programming*. Ullapool, Scotland, September 15–17. <URL:`http://www.dcs.gla.ac.uk/fp/workshops/fpw97/Loidl.ps`>.

Loidl, H.-W.: 1998, 'Granularity in Large-Scale Parallel Functional Programming'. Ph.D. thesis, Department of Computing Science, University of Glasgow. <URL:`http://www.cee.hw.ac.uk/~hwloidl/publications/Glasgow/PhD.ps.gz`>.

Loidl, H.-W.: 2001, 'Load Balancing in a Parallel Graph Reducer'. In: K. Hammond and S. Curtis (eds.): *SFP'01 — Scottish Functional Programming Workshop*, Vol. 3 of *Trends in Functional Programming*. Bristol, UK, pp. 63–74, Intellect. <URL:`http://www.cee.hw.ac.uk/~dsg/gph/papers/drafts/sfp01-gum.ps.gz`>.

Loidl, H.-W.: 2002, 'The Virtual Shared Memory Performance of a Parallel Graph Reducer'. In: H. Bal, K.-P. Löhr, and A. Reinefeld (eds.): *CCGrid 2002 — International Symposium on Cluster Computing and the Grid*. Berlin, Germany, pp. 311–318, IEEE Press. <URL:`http://www.cee.hw.ac.uk/~dsg/gph/papers/ps/dsm02.ps.gz`>.

Loidl, H.-W., R. Morgan, P. W. Trinder, S. Poria, C. Cooper, S. L. Peyton Jones, and R. Garigliano: 1997, 'Parallelising a Large Functional Program Or: Keeping LOLITA Busy'. In: *IFL '97 — 9th Intl. Workshop on the Implementation of Functional Languages 1997*, Vol. 1467 of *LNCS*. Univ of St Andrews, Scotland, pp. 198–213, Springer. <URL:`http://www.cee.hw.ac.uk/~dsg/gph/papers/ps/lolita.ps.gz`>.

Loidl, H.-W., N. Scaife, G. Michaelson, and P. Trinder: 2002, 'Design Decisions in Implementing Parallel Functional Languages'. In: *IFL'02 — Intl Workshop on the Implementation of Functional Languages*. Madrid, Spain, September 16–18. In preparation.

Loidl, H.-W., P. Trinder, and C. Butz: 2001, 'Tuning Task Granularity and Data Locality of Data Parallel GpH Programs'. *Parallel Processing Letters* **11**(4), 471–486. <URL:`http://www.cee.hw.ac.uk/~dsg/gph/papers/ps/hlpp01.ps.gz`>.

Loidl, H.-W., P. Trinder, K. Hammond, S. Junaidu, R. Morgan, and S. Peyton Jones: 1999, 'Engineering Parallel Symbolic Programs in GPH'. *Concurrency — Practice and Experience* **11**, 701–752. <URL:`http://www.cee.hw.ac.uk/~dsg/gph/papers/ps/cpe.ps.gz`>.

Loogen, R.: 1999, 'Programming Language Constructs'. In: K. Hammond and G. Michaelson (eds.): *Research Directions in Parallel Functional Programming*. Springer, pp. 63–91.

Michaelson, G., N. Scaife, P. Bristow, and P. King: 2001, 'Nested Algorithmic Skeletons from Higher Order Functions'. *Parallel Algorithms and Applications* **16**, 181–206. Special Issue on High Level Models and Languages for Parallel Processing.

Milner, R., M. Tofte, R. Harper, and D. MacQueen: 1997, *The Definition of Standard ML (Revised)*. Cambridge, MA: MIT Press.

Mirani, R.: 1996, 'High-level Abstractions for Parallel Functional Programming'. Ph.D. thesis, Yale University.

Mirani, R. and P. Hudak: 1995, 'First-Class Schedules and Virtual Maps'. In: *FPCA'95 — Conf. on Functional Programming Languages and Computer Architecture*. La Jolla, California, pp. 78–85, ACM Press.

Mohr, E., D. Kranz, and R. Halstead Jr.: 1991, 'Lazy Task Creation: A Technique for Increasing the Granularity of Parallel Programs'. *IEEE Transactions on Parallel and Distributed Systems* **2**(3), 264–280.

MPI: 1997, 'MPI-2: Extensions to the Message-Passing Interface'. Technical report, University of Tennessee, Knoxville.

Nikhil, R. and Arvind: 2001, *Implicit Parallel Programming in pH*. Morgan Kaufmann Publishers. ISBN 1-55860-644-0.

Nöcker, E., J. Smetsers, M. van Eekelen, and M. Plasmeijer: 1991, 'Concurrent Clean'. In: *PARLE'91 — Parallel Architectures and Languages Europe*. Veldhoven, The Netherlands, pp. 202–219, Springer.

Pelagatti, S.: 2002, 'Task and Data Parallelism in P3L'. In: F. Rabhi and S. Gorlatch (eds.): *Patterns and Skeletons for Parallel and Distributed Computing*. Springer. To appear.

Peña, R. and F. Rubio: 2001, 'Parallel Functional Programming at Two Levels of Abstraction'. In: *PPDP'01 — Intl. Conf. on Principles and Practice of Declarative Programming*. Firenze, Italy, September 5–7, pp. 187–198.

Pepper, P.: 1993, 'Deductive Derivation of Parallel Programs'. In: R. Paige, J. Reif, and R. Wachter (eds.): *Parallel Algorithm Derivation and Program Transformation*. Kluwer Academic Publishers, pp. 1–53.
<URL:`http://uebb.cs.tu-berlin.de/papers/published/TR92-23.ps.gz`>.

Peyton Jones, S., C. Hall, K. Hammond, W. Partain, and P. Wadler: 1993, 'The Glasgow Haskell Compiler: a Technical Overview'. In: *Joint Framework for Information Technology Technical Conference*. Keele, U.K., pp. 249–257.
<URL:`http://www.dcs.gla.ac.uk/fp/papers/grasp-jfit.ps.Z`>.

Peyton Jones, S. and J. Hughes: 1999, 'Haskell 98: A Non-strict, Purely Functional Language'. Available at `http://www.haskell.org/`.

Plasmeijer, R., M. van Eekelen, M. Pil, and P. Serrarens: 1999, 'Parallel and Distributed Programming in Concurrent Clean'. In: K. Hammond and G. Michaelson (eds.): *Research Directions in Parallel Functional Programming*. Springer, pp. 323–338.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery: 1992, *Numerical Recipes in C: The Art of Scientific Computing*, Chapt. LU Decomposition and Its Applications. Cambridge University Press, 2nd edition. ISBN 0-521-43108-5.

PVM: 1993, 'Parallel Virtual Machine Reference Manual, Version 3.2'. University of Tennessee.

Quinn, M.: 1994, *Parallel Computing*. McGraw-Hill.

Ridge, D., D. Becker, P. Merkey, and T. Sterling: 1997, 'Beowulf: Harnessing the Power of Parallelism in a Pile-of-PCs'. In: *IEEE Aerospace Conference*. pp. 79–91.
<URL:`http://www.beowulf.org/papers/AA97/aa97.ps`>.

Rubio, F.: 2001, 'Programación Funcional Paralela Eficiente en Eden'. Ph.D. thesis, Universidad Complutense de Madrid (Spain). In Spanish.

Sansom, P. and S. Peyton Jones: 1993, 'Generational Garbage Collection for Haskell'. In: *FPCA'93 — Functional Programming Languages and Computer Architecture*. Copenhagen, Denmark, pp. 106–116, ACM.

Sansom, P. and S. Peyton Jones: 1995, 'Time and Space Profiling for Non-Strict, Higher-Order Functional Languages'. In: *POPL'95 — Symp. on Principles of Programming Languages*. pp. 355–366, ACM.

Scaife, N., G. Michaelson, and S. Horiguchi: 2001, 'Comparative Cross-Platform Performance Results from a Parallelizing SML Compiler'. In: *IFL'01 — Intl. Workshop on the Implementation of Functional Languages*, Vol. 2312 of *LNCS*. Stockholm, Sweden, Sep 24–26, pp. 138–154.

Serot, J.: 2001, 'Tagged-token Data-flow for Skeletons'. *Parallel Processing Letters* **11**(4), 377–392.

Taylor, F. S.: 1996, 'Parallel Functional Programming by Partitioning'. Ph.D. thesis, University of London.

Trinder, P., K. Hammond, H.-W. Loidl, and S. Peyton Jones: 1998, 'Algorithm + Strategy = Parallelism'. *J. of Functional Programming* **8**(1), 23–60. <URL:http://www.cee.hw.ac.uk/~dsg/gph/papers/ps/strategies.ps.gz>.

Trinder, P., K. Hammond, J. Mattson Jr., A. Partridge, and S. Peyton Jones: 1996, 'GUM: a Portable Implementation of Haskell'. In: *PLDI'96 — Programming Language Design and Implementation*. Philadephia, USA, pp. 78–88, ACM Press. <URL:http://www.cee.hw.ac.uk/~dsg/gph/papers/ps/gum.ps.gz>.

Trinder, P., H.-W. Loidl, and R. Pointon: 2002, 'Parallel and Distributed Haskells'. *Journal of Functional Programming* **12**(4–5), 469–510. <URL:http://www.cee.hw.ac.uk/~dsg/gph/papers/ps/jfp01.ps.gz>.

WWW-GPH: 2001, 'Glasgow Parallel Haskell'. WWW page. <URL:http://www.cee.hw.ac.uk/~dsg/gph/>.

Appendix

A. Auxiliary Functions

This appendix summarises auxiliary functions we have used in the body of the paper. Most of these functions modify a data structure so as to define parallelism over this modified data structure. The code itself, however, is sequential.

Figure 33 presents the code for some predefined strategies used in the body of the paper. The strategy `parListChunk c s xs` specifies the evaluation of segments of size `c` of the list `xs` in parallel, applying the strategy `s` to every list element.

```
-- sequentially applies a strategy to the first n elements of a list
seqListN :: (Integral a) => a -> Strategy b -> Strategy [b]
seqListN n strat []     = ()
seqListN 0 strat xs     = ()
seqListN n strat (x:xs) = strat x `seq` (seqListN (n-1) strat xs)

-- applies a strategy to (sequential) chunks of a list in parallel
parListChunk :: Int -> Strategy a -> Strategy [a]
parListChunk n strat [] = ()
parListChunk n strat xs = seqListN n strat xs `par`
                          parListChunk n strat (drop n xs)
```

Figure 33. Predefined evaluation strategies

Figure 34 summarises the functions for splitting lists into segments of (almost) equal size and merging them again. This is used by the **matMult** and **raytracer** examples to achieve "data clustering". The function `splitIntoN n xs` splits the list `xs` into `n` segments of the same size, whereas the function `splitAtN n xs` splits a list into segments of the size `n`. The function `takeEach` extracts each n-th element from a given list. It is used in `unshuffleN` to produce a list of lists of every n-th element, starting with 0-th, 1-st, etc element. Thus, `unshuffleN` is an alternative form of clustering, observing the following identity for all n that divide the length of the input list:

$$\texttt{shuffleN} \ . \ (\texttt{unshuffleN n}) \ == \ \texttt{id}$$

```
-- Auxiliary functions for splitting and merging lists
bresenham :: Int -> Int -> [Int]
bresenham n p = take p (bresenham1 n)
                where bresenham1 m = (m`div`p):bresenham1 ((m`mod`p)+n)

-- split list into n sublists of (almost) same size
splitIntoN :: Int -> [a] -> [[a]]
splitIntoN n xs = f bh xs
                  where bh = bresenham (length xs) n
                        f [] [] = []
                        f (t:ts) xs = hs : (f ts rest)
                                      where (hs,rest) = splitAt t xs

-- split list into blocks of size n
splitAtN :: Int -> [a] -> [[a]]
splitAtN n [] = []
splitAtN n xs = ys : splitAtN n zs
                where (ys,zs) = splitAt n xs

-- pick every n-th element from a list, starting from 0th elem
takeEach :: Int -> [a] -> [a]
takeEach n [] = []
takeEach n (x:xs) = x : (takeEach n (drop (n-1) xs))

-- list of lists of every n-th element, starting from 0th, 1st, ...
unshuffleN :: Int -> [a] -> [[a]]
unshuffleN n xs = [takeEach n (drop i xs) | i <- [0..n-1]]

-- combine a list of lists generated by unshuffleN
shuffleN :: [[b]] -> [b]
shuffleN ([]:_)  = []
shuffleN xss     = map head xss ++ shuffleN (map tail xss)
```

Figure 34. Functions for splitting and merging lists