NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.
The Design And Implementation Of
A Systolic Array Silicon Compiler

Xiaoling Sun

A Thesis
in
The Department
of
Computer Science

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Computer Science at
Concordia University
Montréal, Québéc, Canada

November 1987

© Xiaoling Sun, 1987
Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilm l'exposé et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-41644-0
ABSTRACT

The Design And Implementation Of
A Systolic Array Silicon Compiler

Xiaoling Sun

The Systolic Array Silicon Compiler (SASC) accepts systolic designs described at algorithm level and is intended to generate behavioral simulation results for validation and VLSI layouts for fabrication. The design considerations and implementation issues of SASC are discussed. SASC uses bit-serial target architecture, but can accept systolic algorithms for other architectures. It performs two-level pipelined implementation to improve system throughput and keeps all timing synchronization user transparent.

A high-level description language (Systolic Array Silicon Compiler Language (SASCL)) which reflects the properties of systolic architecture is designed. This data flow language provides a simple and hierarchical way of describing systolic designs. SASCL descriptions are first translated into an intermediate form (Systolic Array Silicon Compiler Intermediate Language (SASCIL)). Each SASCIL statement can then be easily translated into the corresponding hardware. Both SASCL and SASCIL are technology independent. The generation of layout from SASCIL is technology dependent and currently can be done for the Northern TeleCom CMOS1B (5 micron) technology.

SASC uses a number of (CMOS1B) leaf cells which can be used to compose other function blocks. These leaf cells and function cells can be automatically placed and routed. SASC can easily be adapted to other technologies by redesigning the leaf cells and appropriately modifying the technology file.

Currently, SASC can generate layouts from SASCIL descriptions. The
philosophies and translation rules for SASCL to SASCIL are proposed. The simulation part of SASC as well as the SASCL to SASCIL translation is left for future implementation.
ACKNOWLEDGEMENTS

I would like to express my greatest gratitude to my thesis supervisor, Dr. H. F. Li for his guidance and encouragement. He outlined the thesis project and made great contributions in its development.

I would also like to thank Dr. R. Jayakumar for his criticism and help on final completion of this thesis, and Dr. C. Lam for his advice, support and proofreading.

I am also grateful to Mr. Luc Morin for his helpful discussions, Mr. Derek Pao for his work on the general router and Mrs. Lily Lam and Mr. Paul Gill; Mrs. Pauline Dubois, Mr. Dave Hargreaves, Mr. Girish Patel and Mr. Marco Zelada for their technical expertise, cooperation and help.

Finally, I would like to extend my thanks to the Shenyang Institute of Computing Technology, Chinese Academy of Science for providing me with a study-abroad scholarship.
# Table of Contents

Abstract .................................................................................................................. iii

Acknowledgements .................................................................................................. v

Table Of Contents ..................................................................................................... vi

List Of Figures .......................................................................................................... ix

List Of Appendices ................................................................................................... xi

1 INTRODUCTION ...................................................................................................... 1

1.1 Systolic Array ..................................................................................................... 2

1.2 Silicon Compilation ............................................................................................. 4

1.3 About The Thesis ................................................................................................. 5

2 THE SYSTOLIC ARRAY SILICON COMPILER ..................................................... 9

2.1 The Design Objectives ......................................................................................... 10

2.2 The Design Considerations ................................................................................. 10

2.3 The SASC Environment: An Overview ............................................................... 15

3 THE SYSTOLIC ARRAY SILICON COMPILER LANGUAGE ......................... 18

3.1 Systolic Architecture Vs Data Flow of Computation ......................................... 18

3.2 Data Flow Languages and Description Languages ............................................. 19

3.3 The Design Objectives And Approaches ............................................................ 21

3.4 The Language Features ...................................................................................... 23

3.5 The Language Description .................................................................................. 28

3.6 The Semantic Considerations ............................................................................ 30

3.7 Specifying An Algorithm To The Compiler ...................................................... 39
5.5.2 The Routing Environment .......................................... 86
5.5.3 Generating Layout Geometry ......................................... 87
5.6 Generation Of The Layout From The Intermediate Code .............. 91
  5.6.1 The Interpreter .......................................................... 91
  5.6.2 Generation Of Layout From The SASCIL ......................... 96
5.7 Generating Layout For The Example Algorithm ....................... 100

6 Conclusion ........................................................................ 117

References ........................................................................... 120
List of Figures

Fig. 1.1.1 General Structure of Systolic Array ........................................ 3
Fig. 1.2.1 Silicon Compilation Process ....................................................... 6
Fig. 2.2.1 Two Phase Clock Convention .................................................... 12
Fig. 2.3.1 Overall Structure Of The SASC ............................................... 17
Fig. 3.5.1 An Example Of Two Types Of PE's ........................................... 31
Fig. 3.5.2 The Higher Level Design Hierarchy ......................................... 31
Fig. 3.5.3 The Main Logic Transmission Directions .................................. 33
Fig. 3.7.1 The Systolic Array For The FIR Algorithm ................................ 41
Fig. 4.3.1 The Communication Strategies Of Target Architecture ............ 51
Fig. 4.3.2 The Target Architecture Of The FIR Algorithm ..................... 58
Fig. 5.1.1 VLSI Design Hierarchy .......................................................... 60
Fig. 5.1.2 Layout Generation Hierarchy ................................................... 61
Fig. 5.2.1 List of Leaf Cells ................................................................. 64
Fig. 5.2.2 Cell Bounding and Supply/Clock wires ................................... 67
Fig. 5.2.3 An Example Leaf Cell .............................................................. 68
Fig. 5.3.1 List of Primitives ................................................................. 72
Fig. 5.3.2 Block Diagram Of Composition Routine ................................. 74
Fig. 5.4.1 CMOS Domino Logic Basic Gate .......................................... 79
Fig. 5.4.2 An Example Of Variable/Product Table ................................ 81
Fig. 5.4.3 An Example Of N-logic Block Layout .................................... 83
Fig. 5.4.4 Block Diagram Of Logic Compiler ....................................... 84
# List of Appendices

<table>
<thead>
<tr>
<th>Appendix I: The BNF Notation</th>
<th>133</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix II: The Syntax Definition Of SASCL</td>
<td>134</td>
</tr>
<tr>
<td>Appendix III: The System Functions Of SASCL</td>
<td>138</td>
</tr>
<tr>
<td>Appendix IV: The Syntax Definition Of SASCIL</td>
<td>143</td>
</tr>
<tr>
<td>Appendix V: The Primitive Library</td>
<td>147</td>
</tr>
<tr>
<td>Appendix VI: The Leaf Cell Library</td>
<td>171</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

When Von Neumann and his colleagues designed the first computer forty
years ago, they partitioned the design into two separate parts, namely, processor
and memory. This is called the Von Neumann architecture. Even though basic
research and improvements in fabrication technology have motivated the
development of four generations of computers, still most of today's computers are
built using the Von Neumann architecture.

In the last decade, improvements in hardware technology for conventional
computer architecture seem to have met the upper bound of physical constraints.
Example of these are speed of light, power-cooling ratios etc, which are
achievable by a single processor. Meanwhile, advances in integrated circuit
technology are rapidly increasing the density and complexity of circuits which is
doubled every one or two years. Now, one micron technology is practical,
allowing the integration of one million transistors on a single chip. So, more and
more people begin to challenge the Von Neumann architecture and computation
model. The classical computer architecture suffers from two main problems: (1)
Since a processor is separated from its memory by communication paths, no
matter how fast the processor is, the duration of the computation is determined
by the time required to move data between processor and memory. (2) The single
processor architecture only supports sequential data processing. The innovation
of computer architecture includes two major aspects: the Von Neumann
computation model vs Non-Von Neumann computation model and general-
purpose computer architecture vs special-purpose computer architecture. The
development of VLSI technology opened the door for the design of innovative
computer architectures and algorithms. The potential power of VLSI is the large
amount of concurrency that it may support. Many new approaches to computer architecture have been investigated, designed and built [1]-[41]. Systolic architecture is one of them.

1.1 The Systolic Array

The systolic array concept was proposed by H. T. Kung at CMU in 1979 [20] as a realistic alternative to conventional Von Neumann architecture. Systolic computer is a special purpose massively parallel processor for solving compute-bound problems, for example, signal processing, matrix multiplication, FFT, etc. Conceptually, systolic architecture is a general methodology for mapping high level computations into hardware structures [35]. In a systolic system, data flow from computer memory in a rhythmic fashion, passing through many processing elements before returning to the memory. Fig. 1.1.1 shows the general structure of a systolic array. The main properties of systolic architecture are:

(1) Make multiple use of each input data flow. Systolic systems can achieve very high throughput by having each input travel through an array of processing elements.

(2) Modular expansibility.

In a pure systolic system, only local communications and controls are involved. Global communications are restricted to system clocks, power and ground lines. So, it is theoretically infinitely expandable, though a large clock skew will slow down the global clock.

(3) Extensive concurrency.

Concurrency in a systolic system is obtained by:

(a) Pipelining the stages involved in the computation of each single result.
Fig. 1.1 General Structure of Systolic Array.
(b) Multiprocessing many results in parallel.

(c) It is sometimes possible to introduce two-level of pipelining by allowing the operations inside the processing elements themselves to be pipelined.

(4) Use only a few types of simple cells.

(5) Use simple and regular data and control flow.

Many systolic computers and algorithms have been designed, built and tested [26]-[41]. Systolic algorithms use parallel pipelined processing with localized control and signal flows and exhibit extensive parallelism with minimum communication. They form a new promoting area of computer innovation.

1.2 Silicon Compilation

A common model of VLSI design partitions the design process into functional, logic block, circuit, layout and mask level. Design of VLSI systems is complicated by additional refinements needed to layout the geometry masks. Although manual design of chip masks could usually achieve high packing density, the complexity involved makes the design of custom VLSI circuits time consuming, tedious, inflexible, error prone, and very expensive. Research in the area of silicon compilation started in late 1970's [53]. It is an automatic synthesis process which takes care of design translation and verification. The correctness of a design is ensured. A silicon compiler, [51]-[80] is a software system which is capable of accepting language or symbolic descriptions and producing a chip mask geometry automatically. It could also simulate a system design at the level of description and allow a rapid validation and implementation of VLSI systems. The language or symbolic description could be at different levels of design hierarchy decided by tradeoffs between application and system design requirements. Generally speaking, a silicon compiler is more concerned with
functional implementation, design cost and correctness rather than optimal packing density. Fig. 1.2.1 shows a silicon compilation process.

Silicon compilers \([42]-[87]\), \([82]-[87]\) have many common features with software compilers. They both take as input a program written in a programming language and produce as output a program in another language (the code or target language). The "compiled code" is only good for a particular machine (in the case of software compilation) or a particular architecture (in the case of silicon compilation). Talking about language features, for example, computations could be abstracted at different levels; information hidden is one of the main decisions to be made, etc. The main difference between the two compilers is that the output of a silicon compiler is not directly executable by a computer. A language, called CIF (CalTech Intermediate Form) which abstracts fabrication process into a number of conceptual layout levels that represent the physical features one observes in the final silicon wafer, is used as a standard medium to interface with fabrication technology. A silicon compiler should provide all of the control between different sections of a program since there is no operating system in the environment to handle controls as in a software compiler. There is also a need for a simulator to validate the system design.

1.3 About The Thesis

Although silicon compilation is a new area of exploitation, it grows very rapidly. A number of research interests, for example, partitioning, placement, routing, testing, design verification and simulation etc., are under this broad topic. It seems impossible to take care of all these aspects in this thesis. What we intend to do is to design a silicon compiler for a particular application area, namely systolic array, with some novelties. The thesis analyzes systolic architecture properties and exploits the natural similarity between systolic array
Fig. 1.2.1 Silicon Compilation Process.
and data flow model of computation and sets up reasonable design objectives for both the silicon compiler and its description language. In the design process, efforts are also made on raising the level of abstraction at which a designer may think, reducing the amount of details that a designer must indulge in and adding intelligence and flexibility to the system by careful design of a high level description language as well as the compiler itself.

The thesis is organized as follows: We present in chapter 2 a discussion of previous work on silicon compilation and design the compiler of our own. In chapter 3, we pay attention to the high level description language design which is one of the main achievements of the compiler. Then, the design of the intermediate language is given in chapter 4. We discuss the layout generation process in chapter 5 and give the conclusion in chapter 6.

For this thesis work, a high level data flow language, called Systolic Array Silicon Compiler Language, and an intermediate form of data description, named Systolic Array Silicon Compiler Intermediate Language, are proposed. We present a set of philosophies and translation rules for the translating the high level description into the intermediate code. The automation of this part is beyond this thesis. We perform hand-compilation for the high level translation process in an example given.

For the translation from the intermediate code into mask geometry, we design and implement both leaf cell and primitive libraries. Leaf cell library is a set of KIC files which are symbolic descriptions of minimum logic elements supporting the target architecture. Primitive library is implemented as a set of composition routines which compose a group of logically related leaf cells together and give the initial placement of calling cells. A general router used at all levels of composition, a CIF generator adapting the intermediate data files into the standard CIF code, a logic compiler being capable of converting boolean
type of data into a random logic structure implementation, and a interpreter conducting the syntax analysis of the intermediate code are designed and implemented. The incomplete part in the low level translation is the implementation of a placer as well as final chip floor plan generation. The behavioral simulator is also left for future work.
CHAPTER 2

THE SYSTOLIC ARRAY SILICON COMPILER

Silicon compilation has been a very interesting topic in computer science and technology. It covers the studies on design description, translation, simulation and verification as well as placement, routing, testing, etc. There are many existing silicon compilers. They were developed for a wide range of application areas, such as signal processing [47], data flow computers [61], finite-state machines [43] or PLA [46], etc, and support many fabrication technologies. The levels of data description cover topology, device, block, architecture and behavior. The target realizations used include standard cells [47], [57], [86], [87], gate arrays [52], [84], [85], random logic and unconstrained custom integrated circuits [88]. More recently, some silicon compilers [61] incorporate behavioral simulation, design verification and other advanced features. Although some of the existing silicon compilers can be used to design chips for systolic arrays, there are a number of problems which are crucial. The MacPtts [50], [60], [61], for instance, was designed for general signal processing applications. Therefore, systolic arrays are only a subset of the architecture supported. It is general enough, but may not be efficient enough to reflect the properties of systolic computation. For example, systolic arrays include a small set of network topologies, and the logic synchronization mechanism is simpler. On the other hand, the existing compilers may not support the fabrication technology available to us. The only way to overcome these difficulties is to design our own silicon compiler -- the Systolic Array Silicon Compiler (SASC).
2.1 The Design Objectives

The main design objectives of the SASC are as follows:

(1) The SASC is aimed at developing a powerful tool which accepts a high level language description and produces chip mask geometry automatically in order to achieve a rapid implementation of systolic arrays. System and functional level simulations will be carried out if they are required.

(2) The SASC should ensure logical correctness and design rule satisfaction of a design.

(3) It should be able to generate designs of complexity compatible with those realized by a human layout designer using manual methods.

(4) It should ensure design integrity in existing design database environments.

(5) The high level language should help users to produce correct design and be user friendly.

(6) The SASC should provide a quick, simple means to incorporate design rule changes.

2.2 The Design Considerations

Primarily, the SASC is based on the structural design methodology. In this methodology, a digital system could be regarded as a combination of register-to-register data transfer paths and finite-state machines. The SASC can be partitioned into two subsystems: data path compiler (for data path generation) and logic compiler (for FSM generation). Each of them runs from high level specification to low level mask geometry. Different methodologies are chosen to solve the problems involved in each of them. We use standard cell methodology
for data path generation and give a special treatment to FSM which is to be discussed later.

The main system assumptions we make are based on the properties of systolic array, simplicity, regularity and global system clock driven as well as some implementation constraints:

(1) Algorithms are synchronous. If we go down to the structure level of a design, all logic elements are under the control of system clocks.

(2) The two-phase non-overlapping clocking scheme is used and shown in Fig. 2.2.1. Data for each function block is received at clock phase one and evaluations are to be done at clock phase two.

(3) Only bit-serial system structure is supported.

(4) The simplicity of processing element is defined as follows:

(a) A systolic system module is considered to be implemented on a single chip (this restriction will be removed in the second stage of the project).

(b) A system may consist of many processing elements and each of them should contain a simple logic circuit: A Boolean expression may not have more than either 9 variables with any number of products or 9 products with any number of variables (this limitation is due to the logic compiler implemented).

(5) The technology supported is the CMOS1B of Northern TeleCom. The SASC strongly supports hierarchical way of structuring a design. The system structure of the SASC is a result of overall considerations to achieve the main design objectives, as discussed below.

(1) *High level language* --> *intermediate code* --> *object code translation*: We first decided to keep the design description to the compiler at the algorithm
Fig. 2.2.1 Two Phase Non-overlapping Clock Convention.
level. Later we discovered that it is too complicated to go through a translation directly from behavior to mask layout in a single step. An intermediate language called Systolic Array Silicon Compiler Intermediate Language (SASCIL) is designed to describe the system at a suitable level lying somewhere between the above two levels. The intermediate code separates the detail of low level issues from users and make the high level language completely technology independent.

(2) **Hierarchical structure of the database system:** The center of the SASC is a database—a set of CIF codes which represents certain hardware logic elements. Primitive Cell Library implements a set of predefined primitive operators that can be parameterized. Leaf Cell Library consists of a set of minimum function elements which can be used to compose different primitives according to the user's specification. The above hierarchical structure forms lower level of system support. The higher level design description is formed by nesting the SASCL primitive functions. Appendices V and VI provide detailed information about these libraries.

(3) **Domino logic structure in implementing finite-state machines:** One of the main system assumptions we make in the compiler design is the simplicity of processing elements. CMOS domino logic [81] is a kind of random logic structure which is one alternative to CMOS complementary logic implementation. Even though speed achieved by complementary logic is usually better than that of others [43], its silicon area cost is higher if the standard library methodology is used. We use domino logic implementation since speed is not crucial for a FSM with less than ten products or variables and a good packing density could be obtained. For example, for a boolean equation with $m$ products and $n$ variables, at most $m \times n + 4$ transistors are needed in domino logic implementation and $m \times n \times 2$ for complementary logic ones.

(4) **The hiding of libraries:** The SASC intends to free users from dealing with
specific instantiations of low level cells. Both primitive and leaf cell details are completely transparent to users.

(5) *Bit-serial implementation*: The advantages of bit-serial implementation over parallel ones are the simplicity, the flexibility and less human efforts needed for layout, circuit, logic and function designs of primitive operators. Automatic transformation from a non-bit-serial algorithm to the bit-serial target architecture represents an important aspect of intelligence and flexibility of the SASC. The correctness of the transformation should be ensured by the compiler, without user knowledge and interaction.

(6) *Two-level pipelined architecture*: Kung, Ruane and Yen have showed how two-level pipelined arithmetic units could be used to form high-performance systolic arrays [27]. The additional level of pipelining can greatly enhance the system throughput with only a small increase in hardware cost. This scheme is especially suitable for VLSI implementations.

The SASC user may have no idea about two-level pipelining at all and the compiler could generate the proper system structure, invoke the corresponding primitive operators, set up the necessary wire connections, and synchronize the logical events that should take place bit-serially.

(7) *Automatic floorplan generation*: The SASC is responsible for the placement and routing of the primitive/logic modules without human guidance. Special signals, e.g. power, ground and system clocks, and special components, e.g. I/O pads, buffers and frames, should not be specified by users. The SASC will generate a complete floor plan for the components and wire connections.

(8) *To accommodate design rule change*: The SASC addresses the following issues in order to accommodate design rule changes:

(a) By using intermediate language as a medium to separate the SASCL from lower level "code generation" details, the SASCL is
made technology independent;

(b) By using a technology file which drives system modules, such as the router and logic compiler, technology changes can be tolerated by these utilities.

(c) By using a hierarchical structure of the databases and potential bit-serial realization, library redesign when technology changes is reduced to a minimum.

(9) Behavioral simulations: A simulator translates the SASCL program into behavioral description form and produces a textual file of the corresponding results.

2.3 The SASC Environment: An Overview

The highest level of the system is a high level data flow language which provides a simple and hierarchical manner of structuring the VLSI design of a systolic array.

The language compiler translates the high level description into intermediate form which deals with specific instantiations of low-level cells and has all connectivity defined explicitly.

The interpreter translates the intermediate code of a design into invoked primitives and all logic connections are transformed into physical ones. It also acts as a low level manager which calls the composition routines, logic compiler, placer and router, etc, and processes the intermediate data files and finally generates layout.

The composition routines call specific library operators and links them together according to predefined primitive definitions. A function block is a collection of logically related primitive cells. K-homogeneous (K<=5) function
blocks form target system.

Logic compiler is an automated synthesis subsystem for finite-state machines. It accepts the intermediate code of boolean data type and generates the corresponding layout.

The router could be used at both the primitive and the higher functional block levels. It estimates channel width, then determines the exact coordinate positions and contact cuts for all cells, and finally connects I/O ports of various modules.

The compiler also translates a high level description into a behavioral description which is then used by the simulator.

The final floor plan of a chip could be generated automatically. Fig. 2.3.1 is the overall system structure of the SASC.
Fig. 2.3.1 The Overall Structure Of The SASC.
CHAPTER 3

THE SYSTOLIC ARRAY SILICON COMPILER LANGUAGE

In this chapter we discuss the high level description language called Systolic Array Silicon Compiler Language (SASCL). SASCL is a data flow language designed with the aim of having a high level of data abstraction in order to free the user from the description details as much as possible.

3.1 Systolic Architecture Vs Data Flow Of Computation

Data flow architectures [12], [71]-[73] are new approaches to computer architecture in pursuit of high performance computers. They are based on the data flow model [70] of computation which is fundamentally different from the sequential one-instruction-at-a-time model. It is based on two basic principles:

1. Operations execute when the required operands are available, that is, asynchronous execution.

2. Operations have no side effects.

These principles say that, in other words, the only sequential constraints imposed by the data flow model are the data dependencies. It allows all the parallelism in a computation to be exploited. Secondly, since there is no need for a global updatable memory, no side effects are produced.

Recalling the discussions of chapter 1, it is easy to see that the computation performed by systolic arrays follow the above two principles. In particular, property (1) and (2) (see page 2) imply principle (1), and property (3) implies principle (2). The general structure of systolic array in Fig. 1.1.1 describes a typical ring type data flow architecture.
In summary we can conclude that systolic architecture is based on the data
flow model of computation and forms a subclass of data flow architecture.

3.2 Data Flow Architecture and Description Languages

The data flow model is a fundamentally different approach in solving
problems; it requires a new language to efficiently describe the data flow
architectures. Data flow languages can be designed to reflect the model at
different levels. It would seem that the natural way of expressing a directed data
flow graph (DFG) would be via a graphical language [89]. High level data flow
languages are desirable when applications are getting bigger and complicated and
DFG's become unreadable and error-prone [68], [74], [76], [77].

Many general-purpose functional and data flow languages have been
proposed for describing non-Von-Neumann computation. For example, the VAL
(Value Algorithm Language) [88] is a well defined and tested high level data flow
language developed at MIT. It embodies all the design principles and tries to
achieve most of the design objectives of these data flow languages. It also takes a
step further in supporting concurrency and machine independence. The weakness
of VAL is the lack of general I/O facilities and recursion. The LAU [89] parallel
system is a contribution to the development of the software concept of Single
Assignment to parallel programming languages and architecture. It was designed
as a tool for expressing parallel algorithms without first transforming them into
equivalent sequential ones. More recently, because of the growing demand for
best facilities in the investigation and implementation of novel digital system
architectures and algorithms, more powerful languages, called description
languages, have been developed for input description to an automatic synthesis
program which can take care of design translation and verification at different
design levels [58]-[63]. The FIRST [47], UNIT version-2 [57], MacPitts [61],
ICEWATER [58] and XI [49] are the typical ones. Mapping a language to certain kinds of computer architecture is one of the common characteristics of these languages. Although some of these could be used for describing systolic processors, they are not very efficient and powerful enough in a silicon compilation environment which is more specialized.

The FIRST, MacPlitts, and XI are designed for applications in digital signal processing systems. In particular, the FIRST is suitable for implementation of real-time signal transformation (bit-serial only) systems and the MacPlitts particularly emphasizes on data path and FSM architecture. The circuit model supported by XI is synchronized, clocked elements and feedback paths associated with clocked elements. The UNIT version-2 can be used to describe an entire microcomputer system and supports both bit-serial and bit-parallel systems. The ICEWATER has the widest application area because of its symbolic description feature. The levels of data abstraction of these languages vary from algorithm level to symbolic ones. A hierarchical manner of describing a design was supported by all of them. Talking about the language features, XI is an extension of the algorithmic language C. It takes advantage of standard concepts of programming languages and environments in which they run that seem to be directly applicable to the circuit design process. It supports the full C programming language and encourages structured thinking and design. MacPlitts is a good circuit description language. The data type of it, therefore, is less structured than that of XI. Both the FIRST and ICEWATER are simple, block structured languages. They deal with function block library operators and require explicit connectivity specification. Some of the language expressions used by these languages directly reflect the target architectures, for example, the extended data type of "register" in XI. The concepts of "bundle" introduced in XI and the "conceptual bus" in UNIT version-2 are good at representing a simple
group of wire connections. The "COND" form in the MacPitts is convenient to express FSM.

3.3 The Design Objectives and Approaches

Any of the description languages discussed in previous section could be used to describe systolic designs. However, a common feature of all of them is their machine-dependency. MacPitts, for instance, was implemented in LISP; Xi is an extension of C programming language and ICEWATER is based on an UNIX-like environment, etc.

Secondly, technology dependency of the low level compilation processes creates another problem. They all support NMOS fabrication process except the MacPitts which supports both MOS1S and NMOS technology. ICEWATER is the best one to face technology changes but pay a penalty of achieving the lowest level of design hierarchy. The technology supported by SASC at the present time is the Northern TeleCom CMOS1B, so that none of these compilers could be used under our circumstances.

Recall the design objectives of the SASC mentioned in the previous chapter. We aimed at having a high level of data abstraction in order to free users from the description details as much as possible. The levels described by FIRST, UNIT version-2 and ICEWATER are too low to match our goals. There is no doubt that MacPitts and Xi are behavioral ones. One of the system assumptions made in these two languages is that users have the ability to express their circuit behaviors by an algorithm description, rather than one that describes the circuit structure. Although any control/data flow graph can be directly specified and high concurrency in data path operations are promoted by MacPitts, it is not powerful enough to reflect the characteristics of systolic computation. For example, the regularity of systolic arrays may allow some defaults (like clock,
power and ground) to be used instead of specifying them explicitly; logic
synchronization might be abstracted at a high level to help users avoid giving all
timing details, etc. We also set up the novelties of our compiler as the automatic
algorithm transformation into a bit-serial one and the two level pipelined
implementation of computation without users' knowledge and interaction. In
addition, behavioral simulation of the system is also incorporated in the compiler.
Thus, the compiler requires a language with options for both layout generation
and system simulation. It is obvious that none of the existing languages meets all
our requirements. A new language is definitely required.

The Systolic Array Silicon Compiler Language (SASCL) intends to tailor the
scopes of the existing description languages. It reflects the properties of systolic
computation and concentrates on achieving the SASC design objectives. The
design objectives of the SASCL are as follows:

(1) *To provide a simple and hierarchical manner of describing a systolic
algorithm.*

SASCL meets this by using one of the most important features of functional
languages: a higher order function may take another function as an argument. It
makes the program remarkably short and succinct.

(2) *To provide implicit concurrency and synchronization.*

This is the most important characteristic of SASCL as a data flow
language. SASCL meets this by allowing operations which could be executed
independently to be defined without explicit language notations.

(3) *To allow implicit parameters, such as system clocks, to be default operands.*

System clocks, power and ground are the common default variables. Two
phase system clocking scheme is used by default. SASCL hides all these and any
retiming details (in case of 2-level pipelining or bit-serial conversion). A user
working on the algorithm level should know how a data is "sampled" and processed, but he need not be bothered to define them.

(4) To support Single Assignment Rule (SAR).

A variable in a program may be assigned values by one statement only. Systolic architectures support SAR naturally due to their locality of references. Sometimes, a global variable may be used to hold certain value to be used in several computations. A modification of a variable is not allowed. This allows maximum concurrency and parallelism to be exploited since the sequentiaility of a program is constrained by data dependency only.

(5) To be user-friendly.

SASCL meets this by:

(a) Choosing a textual form which is similar to that of well-known programming languages. The declaration part of a SASCL program is similar to that of C and PASCAL and the program body is similar to that of LISP;

(b) Making the data types, the data structures and the language expressions close to what systolic system designers normally use to describe their designs;

(c) Giving SASC as much intelligence as possible, so that user's responsibility to a design is limited to the algorithmic level description and probability of design errors is reduced by the subsequent automation.

3.4. The Language Features

The design of language features is one of the key points which may affect the merits of the target architecture and lead to clear, simple and reliable
programming concepts when it is used.

(1) Separated applications

SASC supports two main applications: layout generation and system simulation (behavioral).

During the design process in the real world, people usually work on these two parts separately: they simulate the system first, then go on to lay out the design. This is because, due to the complexity of a VLSI design, translation from an algorithm down to a final layout may take a few months. So, people should ensure correctness of their algorithm before actually attempting the layout. However, concurrent systems are difficult to prove correct formally, and simulation is a valid means of validating them.

SASC allows a very fast validation of a design. The user having a validated design expressed in the form of a program can then generate the corresponding layout. If the generated layout is acceptable, and the simulation reveals no design error, the design process is completed. Otherwise he has to go back to modify his algorithm design.

The layout and simulation parts of SASC are actually separated. The data types used for these two purposes are completely different. A user working on systolic system design regards his variables as signal or primitive operators, while in a simulation, he would use value attributes, for example, integer, character or boolean. Actually, the data types in these two applications refer to different objects. For example, a signal in a parameter-list of the layout description refers to a physical entity, such as an I/O port, while a variable in the simulator describes the value of a signal carried by that port.

In conclusion, SASCL allows layout generation and simulation of a design to be done in two parts of a program that share some "syntactic" parts of the design. A linkage between the two could be established by a system function
called "link".

(2) Data Types

SASCL supports two classes of data types, one for layout generation and the other for simulation.

Signal, flag and register are the data types for layout generation. Among these, flag and register are stored data types and signal is not. A variable having register data type could be used to hold values, numbers and characters. A variable having flag data type may have a boolean value true or false.

The data types supporting simulation are integer, character and logic.

(3) Data Structures

The data structure "bits" is used to describe signals, flags or registers which have more than one bit. The following are legal declarations (capital letters represent the reserved words of SASCL):

\[ x: \text{BITS 4 OF SIGNAL}; \]
\[ y: \text{BITS n OF REGISTER}; \]

The data structure "elements" could be used at both layout generation and system simulation. In the case of layout generation, for example, an "elements" may represent a set of variables:

\[ rw: \text{BITS 4 OF REGISTER}; \]
\[ w: \text{ELEMENTS 3 OF rw}; \]

defines a variable of register data type, four bits for each register.

(4) Values

SASCL does not allow any identifiers to be defined as a global updatable variable. It obeys SAR. SASCL allows an identifier bound to a certain value to be accessed as a global variable only under the following case: a variable has to be
globally initialized before systolic computations start. The problem for layout
generation is that a global variable of register data type should be declared and
the wire connections between the register and the loading I/O pads should be
generated. Because I/O pads only could be called at the topmost level of design,
variables of register data type should be declared as global ones.

(5) Parallelism

SASCL exploits maximum concurrency implicitly and explicitly. A user
should know what concurrency the algorithm tolerates, but he does not take part
in the exploitation of these concurrency in the implementation. Computations
which do not have data dependency could be performed concurrently. For
example, in the following piece of program:

c = (ADD (ADD (a d) ADD (b d)))

the definition implies that the functions "ADD (a d)" and "ADD (b d)" could
be performed concurrently.

Function "FSM" is designed for a direct description of finite state machines.
Conditions in a FSM should be evaluated in parallel.

SASCL supports multi-dimensional parallelism. A number of system
functions (see Appendix III) provide some means of exploiting the concurrency of
algorithms. All functions for describing network topology in a layout generation
program specify explicit concurrency also. For example,

```
LINEAR ( pe (xin xout) 4);
```

will generate 4 independent tasks, each of them performing the computation defined in the function "pe". As a result, two levels of parallelism could be exploited:

(a) function LINEAR defines the top level concurrency explicitly: the four pe's work in parallel.

(b) Inside each pe, computations without data dependency could be done in parallel (implicitly).

(5) **Sequentialism**

Even though SASCL is aimed at exploiting maximum concurrency of an algorithm, in some cases sequential executions are still unavoidable. Loops, for example, are the most general cases where computations should be done sequentially. SASCL supports this by using "FOR" statement. The format of it is similar to that of C programming language. An implicit sequentiality in SASCL is the data dependency of computations.

(7) **Synchronization of Computation**

As we mentioned in chapter 1, systolic array architecture is a special case of data flow computers. Sequentiality, in a data flow computer, is based on data dependency of computations. Synchronization is always required in a systolic array in order to achieve proper timing described in an algorithm. SASCL supports this kind of synchronization by two predefined functions, "synch" and "event".

"Synch" is a primitive function used to synchronize signals listed in its parameter list.
"Event" is a library function designed for describing synchronization of a set of computations where both inputs and outputs should be synchronized. The concept of "event" here is a high level timing abstraction which could be regarded as a time period. We assume that an event starts at some time and ends at some time afterwards. What an user should specify is what computations should take place inside this event, rather than actual starting and ending time. By calling this function, an user could simply put the input and output variables of the computation into the parameter list of the function event and separate the two parts by a semicolon.

(8) Reflection on Systolic Architecture

SASCL achieves this by the following main aspects:

(a) Some data types supported by SASCL have hardware correspondences. For example, a variable of signal data type represents a storage device, etc.

(b) A user-defined function at the highest level represents a processing element in a systolic array.

(c) Some system functions are defined to describe systolic features: for example, functions "LINEAR" and "SQUARE" for network topology and "FSM" for combinational logic, etc.

3.5 The Language Description

In this section, we give a general description of the syntax definitions of SASCL. Appendix II is a complete list of the syntax definition.

An algorithm description of a systolic array could be divided into two parts: processing element description and global algorithm description. A systolic computer usually contains a few types of processing element (PE). Each PE could
be regarded as a black box with an interface and an algorithm describing the local computation. The applications of SASCL program for layout generation and simulation are separated. A SASCL program consists of four parts: program header, global declaration, function declaration and main program code, as in the following syntax.

\[ \text{<applications> ::= <layout> | <simulation>} \]

\[ \text{<layout> ::= "layout" <array name> ";"} \]

\[ \text{<global declaration>} \]

\[ \text{<function definition>} \]

\[ \text{<main code>} \]

Global declaration is decomposed into two parts as shown below.

\[ \text{<global declaration> ::= \{<const declaration>\} <var declaration>} \]

Constant declaration binds numbers to identifiers, and variable declaration defines the variables to be used.

The syntax of the two declarations is given below:

\[ \text{<const declaration> ::= "const" <const list>\}
\]

\[ \text{<const list> ::= <constant> \{ <constant> \}\}
\]

\[ \text{<constant> ::= <const name> "=" <numerics> ";"\}
\]

\[ \text{<const name> ::= <Identifier>} \]

\[ \text{<var declaration> ::= "var" <var list>\}
\]

\[ \text{<var list> ::= <variables> \{ <variables> \}\}
\]

\[ \text{<variables> ::= <name list> ";" <var type> ";"\}
\]

\[ \text{<name list> ::= <var name> \{ "," <var name> \}\}
\]

\[ \text{<var name> ::= <Identifier>} \]

\[ \text{<var type> ::= <single type> | <bits type> | <elements type>\}
<single type> ::= "signal" | "register" | "flag"

<bits type> ::= "bits" <digit> {<digit>} | <const name> "of"

<single type>

<elements type> ::= "elements" <digit> {<digit>} | <const name> "of"

<var name>

Here is an example of global declaration:

    CONST
    word = 8;

    VAR
    x : SIGNAL;
    y : BITS word OF REGISTER;
    z : ELEMENTS 4 OF y;

The SASCL encourages hierarchically structuring a design by using both user defined and library functions. It allows a higher order function to take another function as argument. For example, consider a systolic array with two types of PE's, PE1 and PE2 which are connected alternately. Fig. 3.5.1 shows the interconnection network. Hierarchically we may define a function called PE which consists of PE1 and PE2. The new network structure is shown in Fig. 3.5.2. Then, a library function, called LINEAR, could be called to construct the whole array.

A user defined function is defined by the following syntax:

<function definition> ::= <function head>

    <ports declaration>

    <function body>

<function head> ::= <function name> <parameter list>

<function name> ::= <identifier>
Fig. 3.5.1 An Example Of Two Types Of PE's.

Fig. 3.5.2 The Higher Level Design Hierarchy.
<parameter list> ::= "(" <logic direction> {<logic direction>} ")"

<logic direction> ::= <direction> ":" {<I/O pairs> |<inport name> | <outport name>}

<direction> ::= <main sides> "-" | "\" <secondary sides>

<main sides> ::= "L" | "T" | "TL" | "TR"

<secondary sides> ::= "R" | "B" | "BL" | "BR"

<I/O pairs> ::= <I/O pair> {<I/O pair>}

<I/O pair> ::= <inport name> "," <outport name> ";"

<inport name> ::= <identifier>

<outport name> ::= <identifier>

I/O signals are declared at the head of a function, and the logical data flow directions are defined in the parameter list. There are two types of I/O signals: single bit or multi-bit signal. An input and an output signal pair forming a generic data flow is separated from others by a semicolon, and an input signal is separated from an output signal by a comma. Logic data flow direction is proposed for the convenience of describing systolic algorithms. However, this generic information is useful as a reference for placing and routing low level hardware entities.

The logic data flow direction could be defined by four main directions: Top->Bottom (T->B), Left->Right (L->R), TopLeft->BottomRight (TL->BR) and TopRight->BottomLeft (TR->BL). The opposite directions are described by changing the direction sign "->" into "<-". Fig. 3.5.3 shows the four main directions. Here is an example of function definition:
Fig. 3.5.3 The Main Logic Transmission Directions.
multiplier ( L->R: a1,a0; T->B:bl,bo; TL->BR:cl,co);

a1,a0,bl,bo,cl,co: BITS 8 OF SIGNAL;

BEGIN
    .......
END.

By using the above function a library function SQUARE will generate a mesh structure with diagonal connection between PEs (or a 2-D hexagonal network). A 3 by 3 systolic array for multiplication could be defined by the following function call:

SQUARE (multiplier (ain aout bin bout cin cout) 3 3).

The port type definitions are given by:

<ports declaration> ::= <ports name> "::" <ports type>
<ports name> ::= <port name> {<port name>}
<port name> ::= <lnport name> | <outport name>
<ports type> ::= "signal" | <mult bits>
<mult bits> ::= "bits" <digit> {<digit>} | <const name> "of" "signal"

In the example given above, we can see that port type declaration is a special case of global declaration — only data type "signal" and data structure "bits n of signal" are allowed.

The syntax definitions of a function body are:

<function body> ::= "begin" <function list> "end" ";;"
<function list> ::= <local declaration> <statement list>
<local declaration> ::= {<const declaration>} <var declaration>
<statement list> ::= <statement> {<statement>}

A function body contains a set of assignment statements and function calls. Assignment statements mostly look like:

```c
xout = x;
```

State assignment is for naming states of a FSM. For example, to declare a finite state machine with four states, U, V, W and X we use:

```c
VAR
s : FLAG;
st : ELEMENTS 4 OF s;
```

In the function body, the state assignment statement should be:

```c
st = U, V, W, X;
```

It says that the four states of "st" are called U, V, W and X, respectively.

A function can call either a system function or a user defined function according to the following syntax.

```c
<function call> ::= <function names> "(" <argument list> ")"
<function names> ::= <function name> | <sysfunction name>
```
\[
<\text{argument list}> ::= \{<\text{function call}> | <\text{continued symbol}> | \\
<\text{const name}> | <\text{digit}>\}.
\]

A higher order function may use another function as argument. For example, the function "pe" in the following example is the argument of the function LINEAR.

\[
\text{LINEAR ( pe (x xout y yout) n);}
\]

A set of primitive functions and library functions is defined and it contributes to the main power of SASCL. Primitive functions are a set of functions to describe arithmetic, logic and comparison operations. Library functions reflect systolic algorithm features. The definition of these predefined functions are given in Appendix III.

3.6 The Semantic Considerations

In this section, the main semantic considerations guiding the language design are presented.

(1) \textit{Continued symbols}.

In a systolic design, continued numbers and letters are used quite often. We use double dot sign, "..", to denote this kind of continuation. For example, if we have a variable declaration as shown below:

\[
\text{VAR} \\
x : \text{REGISTER}; \\
y : \text{ELEMENTS 4 OF X};
\]

an assignment in a function might be
\[ y = 1.4; \]

During the compilation process, 1, 2, 3 and 4 will be assigned to \( y1, y2, y3 \) and \( y4 \) respectively.

(2) \textit{Defaults}

By using hierarchical design methodology, a systolic array could be abstracted to have unique types of processing elements. A network contains nothing but a number of duplications of the defined PEs and a set of wires to connect them together.

To reflect this, SASCL allows a user to define I/O ports of a top-level processing element as global variables and the compiler will take care of the default names of the duplicated ones. For example, in

\[
\text{LINEAR ( pe (x xout y yout) 5 );}
\]

signals \( x \) and \( xout \) will represent \( x1, x2, ..., x5 \) and \( xout1, xout2, ..., xout5 \), etc. respectively. Default wires connecting the I/O ports are also implied by the above definition. I/O ports at the boundary of the array will be connected to I/O pads which will be assigned the same name.

For all the predefined functions (primitive and library functions), power, ground and system clocks are the only common default signals.

Subscript default is used in algorithm description. For example, if a variable is defined as a multi-bit register:

\[
\text{VAR}
\]

\[
x : \text{BITS 4 OF REGISTER;}
\]

the variable name associated with a subscript is allowed to be referred directly.
For example, $x[1]$ refers the first bit of register $x$.

(3) *Algorithm conversion*

According to the structure of data to be processed, we may classify systolic algorithms into four categories: bit-serial, bit-parallel, word-serial and word-parallel [41]. SASC not only supports bit-serial implementation, but also accepts other types of algorithms and transforms them into equivalent bit-serial ones. The first two types of algorithms could be directly supported by bit-serial library instances invoked by the compiler. The last two categories include algorithms requiring parallel processing operators. Two kinds of conversions are required:

\[
\text{word parallel } \longrightarrow \text{ bit parallel} \\
\text{word serial } \longrightarrow \text{ bit serial}
\]

SASCL provides ways to describe all these kinds of algorithms and the compiler has the ability to identify them and carry out proper conversions, transforming an algorithm into a bit-serial equivalent one with suitable retiming (data skewing). Since both transformations listed above could be done without changing the network topologies, the need of a transformation could be detected by checking the variable declaration made for the ports. The correctness of the transformation are ensured by the compiler, without user knowledge and other interaction.

(4) *Two-level pipelining conversion*

Beyond the maximum concurrency defined in a user program, one more level of pipelining may be added as a consequence of internal hardware details to enhance the system throughput. Users may not be aware of two-level pipelining. The compiler generates the proper system structure, invokes the corresponding primitive operators, sets up the necessary wire connection, and synchronizes the logical events. We assume that if a function contains arithmetic operators only,
then a conversion from one to two-level pipelining structure will be conducted without user interaction.

3.7 Specifying An Algorithm To The Compiler

A number of systolic algorithm design process have been studied. Generally, the following approach is preferable to describe an algorithm in SASCL. Considering the way a systolic array works, the processing elements could be regarded as black boxes which are capable of performing certain computation. Data with fixed flow pattern are being pumped into the boxes. Information flow between PE's in a pipelined fashion, and communication with the outside world occurs only at the "boundary" PE's. We would like to separate the problem of a system design into three parts:

(1) Processing element description:

Assume that a systolic array consists of a few kinds of PE's, each of them can be represented by a procedure (or subroutine). The I/O ports of the PE's are the windows of the black boxes, connecting to their neighbors and represented by the parameters passed to the procedure. The body of a procedure is the algorithmic description of the computation performed. Variable declarations should be made at the beginning of a procedure. A set of primitive functions ("add", "mult" and "shift", etc.) can be used to describe the computation. The data flow directions through the PE should be defined in the parameter list.

(2) System timing and data flow description:

Logic synchronization is one of the most crucial parts of digital system design. The functions "delay", "synch" and "event" provide means to describe timing synchronization at different levels.

(3) The network topology:
Interconnections in systolic arrays are very regular. SASCL provides a set of library functions ("linear", "square", etc.) to represent network topologies. These library functions may take a procedure as a parameter like other functional programming languages.

Beyond the above three, a hierarchical manner of structuring a design is strongly encouraged. We assume that there is only one type of PE at the topmost level.

The following is a digital filtering algorithm called Finite Impulse Response (FIR) filter. Mathematically, a FIR filtering problem is defined as follows:

Given: the weights \( \{w_1, w_2, ..., w_h\} \)

the initial values \( \{y_0, y_{-1}, ..., y_{-k+1}\} \) and

the input data \( \{x_1, x_2, ..., x_n\} \).

Compute: the output sequence \( \{y_1, y_2, ..., y_{n+1-h}\} \) defined by

\[
y_i = w_j \times x_{i+j-1}.
\]

The equation can be interpreted as below if \( h = 4 \):

\[
\begin{align*}
y_1 &= w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 + w_4 \times x_4 \\
y_2 &= w_1 \times x_2 + w_2 \times x_3 + w_3 \times x_4 + w_4 \times x_5 \\
y_3 &= w_1 \times x_3 + w_2 \times x_4 + w_3 \times x_5 + w_4 \times x_6 \\
y_4 &= w_1 \times x_4 + w_2 \times x_5 + w_3 \times x_6 + w_4 \times x_7
\end{align*}
\]

The systolic array performing the FIR algorithm is shown in Fig. 3.7.1. The algorithm can be described as following:

1. During system initialization (INIT(w)), the weights are loaded in wr's.
2. For each system clock cycle, the operation of a cell is:
Fig. 3.7.1 The Systolic Array For The FIR Algorithm.
(a) Receive $x_1$ and $y_1$ from two input ports and set two internal 
registers $r_y$ and $r_x$ to $x_1$ and $y_1$, respectively.

(b) Send previous $y$ to the output port $y_0$, send $x_1$ to $x_0$ and compute 
$r_y = w_r \cdot x_1 + y_1$.

The data travel by moving through one cell per clock cycle. Following is one 
of the possible SASCL programs for the FIR.

/* layout-generation program for FIR filtering algorithm */
LAYOUT fir;

/* constant declaration */
CONST
  word = 4;
  n = 4;

/* variable declaration */
VAR
  xln, yln, xout, yout : BITS word OF SIGNAL;
  wr : BITS word OF REGISTER;
  w : ELEMENT n OF wr;

/* function definition -- processing element "pe" */
pe (L->R : x1, x0; y1, y0);

/* declare I/O signals */
  x1, x0, y1, y0 : BITS word OF SIGNAL;

BEGIN

/* local variables */

VAR
\text{rx, ry : BITS word OF REGISTER;}

\text{/* synchronize I/O signals */}
\text{EVENT ((xl yl) (xo yo));}
\text{rx = xl;}
\text{/* ry = wr * x + ry */}
\text{ry = ADD (MUL (wr rx) ry);}
\text{xo = rx;}
\text{yo = DELAY (ry 1);}
\text{END;}

\text{/* main program begin */}

\text{BEGIN}
\text{/* Initialize w */}
\text{INIT (w);}
\text{/* generate the linear network */}
\text{LINEAR (pe (x1n x1t y1n y1t) n);}
\text{/* end of the main program */}
\text{END.}
CHAPTER 4

DERIVING SYSTOLIC ARRAY FROM ALGORITHM DESCRIPTION

When designing a silicon compiler, one may immediately find that directly going from behavior to layout is too complicated to be completed in a single translation step. Technology-dependency is a common feature of silicon compilers since they are always designed to support certain fabrication process. Today, advances in VLSI technology change the fabrication process significantly and frequently. In order to face this challenge, it is required to minimize the work involved in adapting to technology changes and separate it from other parts of the compiler. In SASC we achieve this by using a two stage translation. SASC, in the first stage, translates the SASCL description into an intermediate language description called the Systolic Array Silicon Compiler Intermediate Language (SASCIL), which is then translated into the final layout or behavioral description during the second stage.

4.1 The Design Objectives And Considerations

The key point of the SASCIL design is to select a proper level (between the high level language and the low level object code) where SASCIL should stay. Being at that level, SASCIL should make the two transformations of SASC (SASCL to SASCIL and SASCIL to object code) easy to deal with.

In the high level transformation process from SASCL to SASCIL, SASC performs most of the technology independent translations. For example, the parallel to serial algorithm transformation, the two-level pipelining transformation and automatic composition of timing and simulation details, etc. are carried out during this time. In the intermediate form, library functions of
SASCL become a set of primitive calls and logic connections in the SASCIL formats.

In the transformation process from SASCIL to the object code, logic representation of primitive operators and wires should be translated into physical ones. Besides the information carried by SASCIL, there are some other unknowns at this level. For example, the dimension of an user defined function block or the locations of special signals, e.g. Vdd, Gnd and Clocks, should be given by the placer or router. We classify the information to be passed between the two transformations into several categories and give a representation to each of them in SASCIL.

The design objectives of SASCIL are listed in the following:

(1) SASCIL is aimed as a simple and effective tool which isolates the technology-dependent parts of the compiler from the high level language completely in order to support technology independency of SASCL and obtain good system maintenance. SASCIL meets this by:

(a) Abstracting the level of data description at the logical one. For example, the function call in SASCL

\[ c = \text{add}(a \ b); \]

may describe the computation of adding signals a and b, and putting the sum into the register c. The corresponding intermediate code

\[
[ \text{p1 lib 0 Adder ... t ... a b ...} ]
\]

\[
[ \text{p2 register ... c ...t ...} ]
\]

has nothing related to the technology dependent parts of SASC (for instance, physical features of the calling primitive, etc.). It only describes the logic relations of the signals to be used. The isolation of
the technology dependent parts from the high level translation process
supports a better system maintenance.

(b) Choosing proper data types to reduce the complexities of the two
level translation as well as the amount of data to be described by
the intermediate form. SASCIL divides the information for layout
generation into seven categories: signal, port, temp(orary), regis-
boolean and libr(ary). The data types we use are intended to make
a conceptually clear boundary between the two levels of
compilation without increasing the complexities. The types signal,
register, boolean and primitive library functions have data type or
language construct correspondences in SASCL. They are easy to be
identified and translated. The type port is selected in order to give
a clear distinction between signals at I/O port and somewhere else.
The low level translation will benefit from the use of the port data
type which provides the logical location of I/O signals as references
to the placer. “Temp” is another data type without direct language
correspondence in SASCL. It gives convenience to the high level
translation. The data type “lib” contains library primitives (except
the register type) and user-defined functions. We take data type
register out of type “lib” since it is a special case in the primitive
library.

(c) Using regular form of language expressions. Only two kinds of
statements are used: assignment statement and function call
statement. It supports the simplicity and regularity of the
language.

(2) SASCIL should provide a complete set of information needed for
generating low level object code. SASCIL achieves this by:
(a) Providing two sets of language constructs, one for layout generation and the other for behavioral simulation.

(b) Using data types which convey the complete information required for layout generation to the low level translation process.

(3) SASCIL should be readable and has a simple and clear format to help debugging.

This is one of the important features of SASCIL. It is required since: (a) Technology dependency is one of the crucial problems faced by silicon compiler designers. When technology advances, the low level database should be updated and debugging may be required from the intermediate code to the final output. (b) SASCIL is also an interface to the simulator. The readability of SASCIL will help the debugging in both the above cases. This objective is achieved by:

(a) Having data types and primitive operator names in a readable text form. If signals are explicitly defined in SASCL programs, use the same names for the signals in the SASCIL description.

(b) Clearly distinguishing user-defined functions which are included in a pair of brackets "{" and "}".

(c) Reducing the number of unreadable representations as much as possible and let boolean type of data have an equation form, for example,

\[
[ \text{pp boolean ... a \ast b + c \ast d} ]
\]

4.2 The Language Description And Semantic Considerations

We divide the low level primitive operators and the intermediate information into two categories which are defined by the following syntax:
\[ \text{<statements> ::= \{<statement1> | <statement2>\}} \]
\[ \text{<statement1> ::= <items>"="<items> | <boolean type>} \]
\[ \text{<items> ::= <simple type> | <register type>} \]
\[ \text{<simple type> ::= <signal-temp> | <port>} \]
\[ \text{<signal-temp> ::= "[" <identifier> "signal" | "temporary" "]"} \]

\[ \text{<statement2> ::= <lib type> | <event type>} \]

The first category consists of four subtypes, "signal", "temporary", "register", "port" and "boolean". All these data types except the temporary data type (which is going to be removed during the low level translation) have hardware correspondences. For example, a "signal" corresponds to a wire and a "register" may refer to one of the different types of registers in the hardware primitive library.

Logic operations, function FSM and condition statements of SASCL are going to be translated into "boolean type" in a form defined as follows:

\[ \text{<boolean type> ::= "[" <boolean head> <equation> "]"} \]
\[ \text{<boolean head> ::= <identifier> "boolean" <#variable> <#product> <#member> <speed factor>} \]
\[ \text{<#variable> ::= <digit> \{<digit>\}} \]
\[ \text{<#product> ::= <digit> \{<digit>\}} \]
\[ \text{<#member> ::= <digit> \{<digit>\}} \]
<speed factor> ::= "f" | "m" | "s"
<equation> ::= { <Identifier> <logic symbol> } <Identifier>
<logic symbol> ::= "+" | "−"

During the low level translation process, the interpreter translates boolean type intermediate codes into a two-dimensional product/variable table, then calls the logic compiler to generate the corresponding layout geometry.

In the second category, "library type" could be decomposed into "primitive library" and "user defined library". "Primitive library" here is predefined functions in the SASCL and "user defined library." may contain primitive calls, user defined functions and "simple type" definitions, etc.

"Event type" contains timing information which will be used for simulation purposes. At different levels of the design hierarchy, all paths an event may pass could be traced. Main attributes of an event could be found immediately following the event identifier.

The complete syntax definition of the intermediate language is given in Appendix IV.

4.3 Deriving The Systolic Array From The Algorithm Description

Deriving the systolic array form the described algorithm is a complicated process. In this section we give a detailed discussion on the high level transformation process with the FIR algorithm described in section 3.7, bringing out many of the novelties of SASC.

Conceptually the high level translation process should be carried out in the following aspects:

(1) Identify the type of the algorithm. Recall the discussion we have in
section 3.6. The algorithm transformation performed by SASC converts parallel operators and wire connections to bit-serial ones without changing the network topology. A bit-parallel algorithm is detected by checking the variable declarations at the I/O ports. Signals are transmitted either bit-serially on single wires or bit-parallelly on parallel bus. We assume that the way of data processing is same as the way of data transmission. In other words, SASCL does not support algorithms having bit-serial communication and bit-parallel operators or vice versa. Fig. 4.3.1 shows the communication strategies supported by SASC target architecture. The FIR algorithm employs bit-parallel operators as well as bit-parallel communication strategy.

(2) List primitives and signals that are explicitly defined, for example, "add", "sub", etc. This is the prime information of the target architecture.

(3) Add the defaults. The SASCL allows some implicit descriptions of systolic algorithms. The defaults which should be derived at this stage are:

(a) Primitives at I/O ports, for example receivers, senders and corresponding signals, etc.

(b) Default interconnections of library function calls.

(c) Default nodes of primitive function calls, for example the statement

\[ s = \text{add(a b c)}; \]

may be interpreted as:
(a) Bit-serial.

(b) Bit-parallel.

Fig. 4.3.1 The Communication Strategies of Target Architecture.
\[ t = \text{add}(a \ b); \text{ and} \]
\[ s = \text{add}(t \ c); \]

The "t" represents a node of temporary data type. The common default variables, power, ground and Clocks, are not added at this level since they are physically dependent on the placement and routing.

(d) Default I/O ports: a pair of I/O ports are required when a register data type variable needs to be initialized. We assume that initialization is done bit-serially.

(4) Calculate the timing required by the algorithm. Data paths traveled by signals should be traced. Proper time delays are added according to the primitives used and the interpretation of the time functions.

(5) Transform the given algorithm to a bit-serial equivalent one, if it is necessary. An automatic algorithm transformation should be conducted in order to generate the target architecture supported. The transformations are carried out according to the following rules:

(a) Replace each K-bit data wires by one-bit data wire.

(b) Replace each word-operator by a bit-operator.

(c) Replace each K-bit register by a K-bit shift register.

(d) Change the timing specified in the algorithm according to the serial architecture implementation.

(6) Conduct one to two level pipelining transformation. There is no explicit language feature for representing the one to two-level pipelining conversion since it is assumed that the conversion is automatically carried out by the compiler without user awareness. The purpose of introducing two-level pipelining is to improve system throughput for
computation-intensive problems. Algorithms with pure arithmetic computations are selected to have such conversion. The number of stages of the second level pipelining is chosen by the compiler. The correct timing should be ensured corresponding to the bit-serial implementation of the original algorithm (if an algorithm transformation is done).

Now, let's look at the FIR algorithm given in chapter 3 again and illustrate details of the translation. The key word LAYOUT denotes that it is a layout generation program. Global variables are declared under the CONSTants and VARiables. The processing element is defined by the user defined function "pe". Variables in the parameter list of "pe" represent multi-bit signals across boundary and the data flow, xl -> xo and yl -> yo, travel from the left to the right hand side. The computations that take place are multiplication and addition only. This is a parallel algorithm with pure arithmetic operation.

The primitives explicitly declared in the algorithm are:

Signals:   xl, xo, yl, yo (for function pe)

xln, xout, yln, yout, win, wout (for function fir)

Registers:   rx, ry, rw

Primitive-functions:   add, mul

Library-functions: init, event, linear.

The defaults involved in the algorithm description are:

I/Oports:   w1, wo (for function pe).

Primitives:   xl, yl, w1 (receivers).

xo, yo, wo (senders).
Signals: for default primitives.

for network connections:

\[ \text{xin}_1, \text{xin}_2, \text{xin}_3, \text{xin}_4, \]
\[ \text{yin}_1, \text{yin}_2, \text{yin}_3, \text{yin}_4, \]
\[ \text{win}_1, \text{win}_2, \text{win}_3, \text{win}_4, \]
\[ \text{xout}_1, \text{xout}_2, \text{xout}_3, \text{xout}_4, \]
\[ \text{yout}_1, \text{yout}_2, \text{yout}_3, \text{yout}_4, \]
\[ \text{wout}_1, \text{wout}_2, \text{wout}_3, \text{wout}_4. \]

Nodes: for function call: \( \text{mul(} \text{wr \ rx)}; \)

User-defined-functions: \( \text{pe}_1, \text{pe}_2, \text{pe}_3, \text{pe}_4. \)

The statement "\( \text{yo} = \text{DELAY (} \text{ry} \ 1) \)" says that signal \( \text{yo} \) is obtained by delaying \( \text{ry} \) one time period (which is the time of one word delay for the word-serial algorithm). The first bit of signal \( \text{yo} \) should appear at the output port four cycles after the first bit of the signal gets out of register \( \text{ry} \) since the algorithm transformation is done.

After the transformation, signals \( \text{xin}, \text{yin}, \text{xout} \) and \( \text{yout} \) should be one-bit wide and, register \( \text{wr}, \text{rx} \) and \( \text{ry} \) are 4-bit shift registers in the FIR algorithm.

A pair of storage registers is required for a pair of assignment statements shown below:

\[ \text{rx} = \text{xl}; \]
\[ \text{xo} = \text{rx}; \]

Register \( \text{rx} \) could be regarded as a receiver which takes input from signal \( \text{xl} \), and a default register (e.g. named \( \text{rxs} \)) is required as a sender and takes the output of \( \text{rx} \) as input data. The output of register \( \text{rxs} \) is connected to signal \( \text{xo} \). The \( \text{rx} \) and \( \text{rxs} \) work at clock phase one and two, respectively. If an algorithm transformation
is conducted (like in the example), either rx or rxs should be replaced by a shift register. The number of bits to be added is dependent on the word-length of the systolic array and the timing specification of the original algorithm. Since signals xi and xo are synchronized with the signals yi and yo, respectively (defined by function EVENT), the time delay required should be calculated according to the longest data path, that is, the data path for yi and yo. Two cycle delay is required (one for multiplication and one for addition). Assume that register rxs is a shift register. It should be \((n+c)\) bits long where \(c\) is the number of cycles that the computations take minus one.

Function INIT indicates a register type of variable has to be initialized through terminals of D-flip-flops. The first parameter of the function is the name of the variable to be initialized and the second one defines the terminals by which the initialization is done. In particular, the default is terminal "d" (like in the given example); "0" and "1" specify terminals "reset" and "set", respectively. Signal wln is delayed \(n\) cycles before it reaches the port wout.

Function LINEAR defines a linear array of four elements performing the computation specified in the function "pe". This translation is done in the low level translation process since it is unnecessary to generate four-pe's. They can be obtained by duplicating the pe four times when it is composed in the low level translation process. The network interconnections are specified explicitly by means of assignment statements. Default is used to represent the signal names of the I/O ports of each processing element. For instance, the global signals xin and xout are a pair of signals corresponding to the signals xi and xo in the function pe. As a consequence a set of intermediate code is derived for the given FIR algorithm and it is shown below. The comments and line numbers used here are for explanation purpose only, they are not part of the intermediate form. Symbol "I" means that the corresponding terminal of that primitive is not used. It
happens when (1) a signal of a primitive has more than one I/O ports located on different sides of the layout boundaries or (2) some terminals of primitive operators like "set" or "T" of a D-flip-flop, are logically not used. Physically the terminal "set" should be connected to Vdd in the low level translation process.

/* define the fir array */
[ fir lib 1 pe 0 4 2 6 ( xln1 xln4 ) ( yln1 yln4 ) ( win1 win4 ) ]

/* define the network connections */
[xin2 signal] == [xout1 signal]
[xin3 signal] == [xout2 signal]
[xin4 signal] == [xout3 signal]
[yin2 signal] == [yout1 signal]
[yin3 signal] == [yout2 signal]
[yin4 signal] == [yout3 signal]
[win2 signal] == [wout1 signal]
[win3 signal] == [wout2 signal]
[win4 signal] == [wout3 signal]

/* define the processing element "pe" */
[ pe lib 1 3 1 4 2 6 ( xln xout ) ( yln yout ) ( win wout ) ]
{ [s1 signal] == [win port 1 1]
[s2 signal] == [yln port 1 1]
[s3 signal] == [xln port 1 1]
[s4 signal] == [xout port o r]
[s5 signal] == [yout port o r]
[s6 signal] == [wout port o r]
[ p3 register 0 1 2 n1 s6 ! ]
[ p3 register 0 1 2 n1 s6 ! ] == [ p1 register 5 4 1 s1 n1 ! ]
[ p1 register 5 4 1 s1 n1 ! ] == [ s1 signal ]
Lines 14-34 are the intermediate code corresponding to function "pe" in the FIR algorithm. The port declarations of the "pe" are represented by lines 15-20. If a signal is declared as an I/O port, an assignment statement like line 16 should be given in the intermediate form. The type "port" is assigned to the declared signal and the type "signal" (which represents wire connection) is given to a new name provided by the high level translation process. Lines 22-33 describe the computations the "pe" performs. For instance, line 30 says that p5 is a predefined library function working at clock phase two in level 1 of the design hierarchy. The two input and one output signals used are n2, n4 and n5, respectively. The intermediate code inside the inner pair of brackets ("{" and "}") define the function pe. The systolic array "fir" is described in lines 1-14, and 35 which describe nothing but the network interconnections. The target array corresponding to the above intermediate code is shown in Fig. 4.3.2.
Fig. 4.3.2 The Target Architecture of The FIR Algorithm.
CHAPTER 5

THE COMPONENTS OF SASC

We discussed in the previous chapter how a systolic array is derived from algorithm description and how the high level translation process (from the SASCL to the SASCIL) is conducted as well as the intermediate language design. In this chapter, we will show how a layout is generated from the intermediate form.

5.1 The Layout Generation Hierarchy

A VLSI design is hierarchical in nature. Special system structures are usually required. A system could be partitioned into function blocks each performing a subset of the computation. Then the function blocks could be either divided into some sub-function blocks or contain a set of computations described by the library functions provided by the description language. Fig. 5.1.1 shows the hierarchy of the design methodology supported by SASCL and indicates that at least three levels of hierarchy are included in a single design.

In contrast to the top-down design methodology supported by the high level description language, the composition of a layout is carried out bottom-up. The layout generation hierarchy is given in Fig. 5.1.2.

The lowest level composition is for primitive operators. A primitive operator can be parameterized and formed by calling a fixed set of leaf cells and linking them together according to the predefinition. The initial placement of some primitives are predefined manually and generated by a set of composition procedures. Other primitive operators may have variable parameters. For example, shift registers may have parameters such as word length and type of
Fig. 5.1.1 VLSI Design Hierarchy.
Fig. 5.1.2 Layout Generation Hierarchy.
storage device used, etc. The corresponding composition routines are capable of taking into account these parameters and generating proper layout. The physical size of a primitive is not known until its composition process is completed.

The middle level composition is designed for function blocks. A function block is a collection of logically related primitive cells. Even though the main objective of SASC is to reduce engineering effort and design turn around time rather than minimizing chip area utilization, there are still some tradeoffs between the extent of automation and the penalty paid in silicon area cost. The complexities of the primitive operators varies from a transistor pair (e.g. INPAD) to some powerful function blocks having hundreds of transistors inside, such as comparator, multiplier, etc. The size of primitive operators varies from a hundred by a hundred square microns to thousands by thousands square microns. Moreover, processing elements of systolic arrays usually have very simple functionality and a few communication ports. If a PE is specified in 5 levels of design hierarchy and the placement and routing is done for each of them separately, some gaps between large rigid building blocks are difficult to avoid. Great efforts we make in manual design of leaf cells in order to reduce the silicon area cost would be worth nothing because of the area wastage due to large gaps which may appear in each level during the design. The solution to this is to allow user to specify their designs using as many design levels as they want, but performing the middle level placement and routing only once. In other words, all primitive operators used in a PE will be placed and routed at the level next to the topmost one. The small building blocks (primitive operators) could take into account the global placement and routing requirements, tailor the shape of the PE, and avoid large gaps generated between large blocks. The traffic in the routing channel will not be a problem due to the simplicity of systolic arrays and the general complexity of primitive cells. Achieving a better placement of all
primitive operators, of course, is harder than that for leaf cells, but it is not a very difficult and computation intensive task under our situation of compiling systolic arrays.

The high level composition is for final floor plan generation of a chip. The placement and routing are taken care of by the SASC automatically according to the placement algorithm employed and the data generated by the lower level composition without user knowledge and interaction.

5.2 The Leaf Cells

The leaf cells are the basis of the central database to support the entire SASC system.

5.2.1 Selection of Leaf Cells

Leaf cells are the building blocks of primitive functions. As soon as primitive operators are selected, the functionalities of leaf cells could be decided. A number of standard leaf cells are defined and listed in Fig. 5.2.1. The programmable layout generation for all combinational logic modules greatly simplifies the construction of the library.

5.2.2 Leaf Cell Conventions

In order to guide the automation process and obtain a good library maintenance, leaf cell conventions were carefully defined. The decisions made are mainly based on the system assumptions meeting the design objectives, the horizontal symbolic layout rules [80] and the technology constraints, and the estimation of silicon area cost of the minimum logic element.

(1) Two-phase clocking scheme was chosen as the logic implementation
<table>
<thead>
<tr>
<th>NAME</th>
<th>FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>driver</td>
<td>inverted driver</td>
</tr>
<tr>
<td>ndriver</td>
<td>non-inverted driver</td>
</tr>
<tr>
<td>dlatch</td>
<td>D-latch</td>
</tr>
<tr>
<td>drs</td>
<td>D-flip-flop with R/S</td>
</tr>
<tr>
<td>fba</td>
<td>full binary adder</td>
</tr>
<tr>
<td>multiplexer</td>
<td>4--&gt;1 switch</td>
</tr>
<tr>
<td>demultiplexer</td>
<td>1--&gt;4 switch</td>
</tr>
<tr>
<td>logic compiler</td>
<td>logic compilation</td>
</tr>
<tr>
<td>inpad</td>
<td>input pad</td>
</tr>
<tr>
<td>outpad</td>
<td>output pad</td>
</tr>
<tr>
<td>vddpad</td>
<td>power pad</td>
</tr>
<tr>
<td>gndpad</td>
<td>ground pad</td>
</tr>
<tr>
<td>frames</td>
<td>frames</td>
</tr>
<tr>
<td></td>
<td>driving</td>
</tr>
<tr>
<td></td>
<td>driving</td>
</tr>
<tr>
<td></td>
<td>bit-delay</td>
</tr>
<tr>
<td></td>
<td>bit-delay</td>
</tr>
<tr>
<td></td>
<td>arithmetic</td>
</tr>
<tr>
<td></td>
<td>control</td>
</tr>
<tr>
<td></td>
<td>control</td>
</tr>
<tr>
<td></td>
<td>combinational logic generator</td>
</tr>
<tr>
<td></td>
<td>generation</td>
</tr>
<tr>
<td></td>
<td>generation</td>
</tr>
<tr>
<td></td>
<td>chip generation</td>
</tr>
<tr>
<td></td>
<td>chip generation</td>
</tr>
<tr>
<td></td>
<td>chip generation</td>
</tr>
<tr>
<td></td>
<td>chip generation</td>
</tr>
<tr>
<td></td>
<td>chip generation</td>
</tr>
</tbody>
</table>

Fig.5.2.1 List Of Leaf Cells.
structure. The supported CMOS fabrication technology determines the use of complementary clocks in each phase in order to drive both P and N types of transistor.

(2) Leaf cells in the library have variable lengths and standard heights. The standard heights are used in order to simplify the placement and routing algorithms. There are three possible heights of leaf cells (for the CMOS1B technology):

\[ 1H = 165 \text{ micron (11 ports on left/right sides)} \]
\[ 1.5H = 261 \text{ micron (17 ports on left/right sides)} \]
\[ 2H = 341 \text{ micron (22 ports on left/right sides)} \]

(3) Each leaf cell is considered to be placed on a grid. The origin of the leaf cell is taken to be its left bottom corner. The separation of grid lines decided by the minimum port to port spacing in the given technology. The horizontal grid lines start from the origin and are located on the positions of multiples of 16 (minimum poly contact cut dimension 11 microns plus minimum poly to poly spacing 5 microns). The vertical grid lines sit on the positions of

\[ 1 + 16 \times x \ (x = 0, 1, 2 \ldots) \]

The 1 is a variable used to avoid the overlapping of the first vertical and the horizontal I/O ports. For the technology supported at the present time, 1 is equal to 8 microns (minimum poly to poly spacing 5 microns plus 3 microns contact cut extension).

(4) There are two layers available for external connections of leaf cells, namely, metal and polysilicon. In general, metal wires go horizontally and polysilicon run vertically. The input ports should appear on the top or left sides and the output ports are on the bottom or right sides. All
the I/O ports must sit on grid lines.

(5) Same signal may have more than one I/O ports. These ports should be internally connected, so that an external connection can be made to any of them.

(6) The origin of the I/O ports is defined as the left edge for vertical wires and the bottom edge for horizontal wires.

(7) The cell bounding and supply/clock wiring are defined as shown in Fig. 5.2.2. The relative positions at which the power and clock lines enter the cell are fixed. The line width of power and clock rails are 10 and 5 microns respectively through out the whole cell. No under cut could be made on them except for internal supplies of the leaf cell.

Fig. 5.2.3 shows a layout of a standard leaf cell.

5.2.3 The Library Description files

All leaf cells (except combinational logic ones) are laid out manually and represented in KIC format. A file called "leafdesc" gives block description of the physical characteristics of each cell and is defined as the follows:

<leaf cell name> <height> <width> <# of I/O ports>
<port 1> <signal name> <side> <offset from origin>
<port 2> <signal name> <side> <offset from origin>
<port n> <signal name> <side> <offset from origin>
Fig. 5.2.2 Cell Bounding and Supply/Clock Wires.
THE QUALITY OF THIS MICROFICHE IS HEAVILY DEPENDENT UPON THE QUALITY OF THE THESIS SUBMITTED FOR MICROFILMING.

UNFORTUNATELY THE COLOURED ILLUSTRATIONS OF THIS THESIS CAN ONLY YIELD DIFFERENT TONES OF GREY.

LA QUALITE DE CETTE MICROFICHE DEPEND GRANDEMENT DE LA QUALITE DE LA THESSES SOUMISE AU MICROFILMAGE.

MALHEUREUSEMENT, LES DIFFERENTES ILLUSTRATIONS EN COULEURS DE CETTE THESES NE PEUVENT DONNER QUE DES TEINTES DE GRIS.
Fig. 5.2.3 An Example Leaf Cell.
The cell name and signal name should not have more than 10 characters. Unique port numbers are used to identify all the ports. Different port numbers may have the same signal name but different offsets which indicate the relative physical location of each port. The numbers 0, 1, 2 and 3 are defined to represent the left, right, bottom and top side of a cell respectively.

A leaf cell index file, called "leafindex", is set up for quick retrieval of the required leaf cell from the file "leafdesc". Leafindex is searched by cell name which are stored in sorted order. A leaf cell name in leafindex is followed by the starting address of its data and the number of bytes occupied in file "leafdesc". By knowing these, the file leafdesc can be directly accessed.

5.2.4 Technology Dependency

The main problem of the standard library approach is the technology dependency. When technology changes, the cell library has to be redesigned manually. Finite State Machines (FSM) and boolean expression are complex and random in nature. VLSI designers spend a lot of time in designing them. Quite a number of leaf cells are needed if a full standard cell library strategy is chosen. In order to minimized the amount of work to be involved in the case of a technology change, we use a semi-standard library approach -- using a programmable random logic implementation for FSM and combinational logic modules and a standard library implementation for data path, which reduces the complexity of the library. As a consequence, technology dependency of the leaf cells is reduced to a minimum.

5.3 The Primitive Operators

5.3.1 Selection Of Primitive Operators
The main objective of the primitive library design is to select a minimum set of operators which could be used to construct the systolic arrays. To achieve this goal, a study of existing systolic algorithms was pursued. In many cases, arithmetic operations have to be performed simultaneously, such as those expressed by the mathematical symbols "Sigma" for addition and "Pi" for multiplication. These are compute-bound computations that can benefit from the systolic approach. Existing systolic designs may be classified into following categories:

(1) **Signal & Image Processing**
   - FIR, IIR filtering and 1-D convolution
   - 2-D convolution and correlation
   - discrete fourier transform
   - interpolation
   - 1-D and 2-D median filtering
   - geometric warping

(2) **Matrix Arithmetic**
   - matrix-vector multiplication
   - matrix-matrix multiplication
   - matrix triangularization
   - QR-decomposition
   - solution of triangular linear systems

(3) **Non-numerical Applications**
   - data structure: stack and queue, searching, priority queue and sorting
   - graph algorithms: transitive closure, minimum spanning trees and connected components
   - language recognition: string matching and regular
expression

dynamic programming

encoders

relational data-base operations.

Primitive operators selected for the SASC should be able to cover the above applications. Arithmetic and storage operators are intensively used in many of the systolic systems. Hence, adder, subtractor and multiplier are selected as the basic arithmetic operators. Two types of storage devices are needed: buffering and buffering with initial set and reset capabilities. A full programmable realization of combinational logic modules is chosen based on the simplicity of processing elements of systolic arrays. We will give detailed discussion about the design and implementation of the "logic compiler" in section 5.4. The logic compiler greatly simplifies the construction of two levels of libraries and reduces the amount of work to design leaf cells manually. Bit comparator and switch-box are the basic cells for sorting, string matching and network switching problems. The use of CMOS technology brings in the necessity of adding a primitive operator, Clockgen, to generate complementary, clocking signals. Drivers are definitely required if the speed of a system is crucial. In addition, a group of primitives, (for example, I/O pads and frames), are selected according to the requirements of generating final chip layout. Fig. 5.3.1 shows a list of primitives selected.

5.3.2 Primitive Composition

Composition of primitive operators is carried out by "composition routines". A composition routine is a program module capable of grouping low level objects (leaf cells) into a primitive floor plan and generating the layout of a parameterized primitive operator. The CIF (CalTech Intermediate Format), an
<table>
<thead>
<tr>
<th>Arithmetic</th>
<th>Storage</th>
<th>Control</th>
<th>Boolean Expression</th>
<th>Driving</th>
<th>Chip Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adder</td>
<td>Dlatch</td>
<td>Switchbox</td>
<td>Logic Compiler</td>
<td>Driver</td>
<td>Inpad</td>
</tr>
<tr>
<td>Subtractor</td>
<td>Drs</td>
<td>Multiplexer</td>
<td></td>
<td>Ndriver</td>
<td>Outpad</td>
</tr>
<tr>
<td>Multiplier</td>
<td>Shiftregd</td>
<td>Demultiplexer</td>
<td></td>
<td>Clockgen</td>
<td>Vddpad</td>
</tr>
<tr>
<td>Add-Multiplier</td>
<td>Shiftregdrs</td>
<td>Comparator</td>
<td></td>
<td></td>
<td>Gndpad</td>
</tr>
</tbody>
</table>

Fig. 5.3.1 List Of Primitives.
"embedded" language, is used to enable objects to be defined and instantiated hierarchically.

In the low level translation process, a set of intermediate code describing a high level design are analyzed and a group of parameters are kept in a text file named caliprl which is the input data file of the composition routines. It contains the following information:

\[
\text{<primitive name> <# of bits> <word length>}
\]

The name of the primitive is unique. It is a concatenation of the name of the calling primitive in the primitive library and the name of the cell in the design. For example, in Adderc2-p5, Adderc2 is the library operator used. p5 is the name of adder in the system and assigned by the high level compilation process. The second and the third parameters are digits. Furthermore, a one-to-one mapping from a primitive name to the corresponding layout geometry is conducted. Fig. 5.3.2 shows the composition routine environment. The outputs of a composition routine are:

1. The placement file named *.pla specifies the initial cell placement. Its format format is:

\[
\text{<hr>}
\]

\[
\text{<total # of cells>}
\]

\[
\text{<cell name> <cell #> <x, y coordinates of the origin>}
\]

where \text{<hr>} is an important parameter which guides the routing algorithm used and we will discuss it in section 5.5. The \text{<cell #>} should correspond to the order in which the cell instance was invoked. The \text{y} coordinate of cells are quantized to integer multiples of \text{hr}. 
Fig. 5.3.2 Block Diagram Of Composition Routine.
(2) *.net is a netlist file which gives explicit description of the internal wire connections and external ports of a primitive. The head of a *.net file is defined as below:

\[
\text{<# of exter-port>}
\]

\[
<\text{side}> \quad <\text{# of exter-ports}> \quad <\text{exter-port #, signal name}>, \ldots
\]

There are some constraints brought by the router: (a) The external ports on each side should be specified as a group and listed in a sorted order according to their physical locations starting from the origin of the block. (b) The groups of signals should appear in the following sequence: left, right, bottom and top. If there is no port on a side, the corresponding specification should be defined as

\[
<\text{side}> 0
\]

to declare the special case. The body of a net list is:

\[
c \quad <\text{cell#, port#}> \quad t \quad <\text{cell#, port#}> \quad \ldots \quad t \quad <e, \text{side, port#}, \text{order}>
\]

[Note]:

- \text{c} ---- connect
- \text{t} ---- to
- \text{e} ---- external
- \text{side} ---- 0 : left
  \quad 1 : right
  \quad 2 : bottom
  \quad 3 : top

(3) *.pdes is a block description file for primitive operators which has the same definition as calling leaf cells except for their names. For example, consider the primitive "Drivers" and the leaf cell "driver8". In this case,
Initial placement and routing is not required. The file *.pdes functionally corresponds to a data file, named mod1.dat, generated by the router for other primitives. We will introduce file mod1.dat in section 5.5.

(4) A KIC file named *.kic shows the primitive floor plan generated.

(5) A text file called *.deb which gives information as the files leafdescr, leafindex and pridescr are accessed. The files *.kic and *.deb are used for debugging purposes only.

5.3.3 Primitive Description File

A primitive description file, called "pridescr", contains information about all the primitive operators. It is defined as follows:

<primitive name> <# of leaf cells used>

<leaf cell 1>

<leaf cell 2>

<leaf cell n>

<primitive name> <# of leaf cells used>

Leaf cells under each primitive name are those which are predefined to construct the primitive operator. The sequence of leaf cell names listed in each primitive description segment implies the sequence of initial placement to be done in the
corresponding composition routine.

5.3.4 Expandability Of The Libraries

The primitive and leaf cell libraries are extensible. We leave both of them open and allow new members to be added. To extend the leaf cell library, complete information about the cell being added should be given at the end of the leaf cell description file, leafdesc. Then a system routine, LEAFINDEX. GEN, is called to automatically find out the position of the new cell in file leafdesc, modify the corresponding index file and sort the file by cell name.

The primitive library actually is a set of composition routines as we mentioned in the previous section. To extend the primitive library is to design the primitive operators to be added, write corresponding composition programs, append to the file "composition" and finally add the required information into the files "pdesc" and "*.net".

Each leaf cell or primitive operator in the system has a unique name. Leaf cells and primitive operators are distinguished by the first character of their name. The first character of a primitive cell name is a capital letter and a name of leaf cell is all in lower case. Some primitive cells, for example, Driver, may have exactly the same definition with the corresponding leaf cell, driver, since the leaf cell 'driver' might be used to construct some primitive operators and the primitive 'Driver' may be called to build a function block of a target system. Some leaf cells have more than one version of layout geometry. For example, a bit-delay (d latch) may work at clock phase 1 or 2 depending on the high level description. In each case, the corresponding layout is fixed. We append a postfix, c1 or c2, to the name of the leaf cell in order to distinguish the two. Each version of the layout has its own leaf cell description.
5.4 The Logic Compiler

The logic compiler is a layout program which deals with intermediate code of boolean type and generates corresponding layout geometry. It services both the primitive and the leaf cell libraries. A combinational logic module may be a leaf cell used in constructing a certain primitive operator or a primitive operator within a function block.

5.4.1 Assumptions

In order to simplify our problem, a few system assumptions are established. They are:

1. The boolean equations provided by the user are optimized. This assumption could be removed later, if an optimizer becomes available.

2. The size of the boolean equations could be:
   
   (a) any number of products, but not more than \( 9 \) variables, or
   
   (b) any number of variables but not more than \( 9 \) products.

   The value of \( 9 \) is determined by the leaf cell conventions of section 5.2.2.

3. Three types of speed factors can be used. They are fast, slow, and medium. For each layout of a combinational logic module, only one speed factor is supported.

5.4.2 Domino Logic Structure

Domino logic is chosen as the logic structure of implementation. It allows a single clock to precharge and evaluate a cascaded set of dynamic logic blocks. This involves incorporating a static CMOS buffer into each logic gate as shown in Fig. 5.4.1. During the precharge time, the output node of the dynamic gate is
Fig. 5.4.1: CMOS Domino Logic Basic Gate.
precharged high and output of the buffer is low. As the subsequent logic stages are fed from this buffer, transistors in subsequent logic blocks will be turned off during the precharge phase. When the gate is evaluated, the output will conditionally discharge, causing the output of the buffer to conditionally go to high. The limitations of the structure are (a) only non-inverting structures are possible; (b) each gate must be buffered. However, these limitations could be compensated. For example, by cascading a inverted logic block to remove restriction (a). Comparing with complementary CMOS logic implementation, domino logic has lower power dissipation. Area cost can be estimated to a first order by the number of transistors. For an M input gate, a complementary CMOS gate uses \((2 \times M)\) transistors whereas only \((M + 4)\) transistors are required for domino implementation.

5.4.3 Logic Compilation

The logic compiler is called when boolean type of intermediate code is encountered. Let's look at an example. Assume we have a boolean type of intermediate code as shown below:

\[
[F \text{ boolean } 7 4 13 1 a \cdot b + c \cdot d + e \cdot f + g].
\]

In the low level translation process, first of all lexical analysis is applied to the intermediate form. If a boolean type of data is detected, a procedure named "boolgen" is invoked. It converts the intermediate form into a two-dimensional table, called the variable/product table. The rows in this table represent the variables and columns represent the products. If a variable appears in a product, "\(1\)" should be set in the corresponding row and column. For the above example, Fig. 5.4.2 shows the corresponding variable/product table. Then, the table is mapped into a n-logic block of domino logic implementation. The vertical device wells represent the products and the horizontal poly wires are corresponding to
<table>
<thead>
<tr>
<th>Var.</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

\( P = A \cdot B + C \cdot D + E \cdot F + G \)

Fig. 5.4.2 An Example Of Variable/Product Table.
the variables. Fig. 5.4.3 shows the layout of the n-logic block for the example.

We can see that the layout in Fig. 5.4.3 is not optimal. If we let metal wires cross the cell and put contact cuts at the left/right sides, and then connect them together, some redundant contact cuts (one with a circle) could be removed. A layout optimization procedure removes redundant contact cuts and the space originally occupied by redundant contact cuts are saved. Finally, a compact layout geometry is generated. Fig. 5.4.4 presents the block diagram of the logic compiler.

5.5 Placement, Routing and CIF File Generation

This section is organized in three parts. In section 5.5.1, we discuss the basic assumptions made for the placement and routing algorithms. Then, a general channel router [90] is introduced. Next, in section 5.5.2, we present a global picture of the routing environment. In section 5.5.3, we show how a composition process is ended by generating mask geometry of a function block.

5.5.1 Placement And Routing

Routing is one of the classical problems related to design automation. Many routing algorithms have been proposed and implemented. Due to the properties of systolic array and the hierarchical layout methodology, the routing problems in the SASC environment is much simpler. The routing scheme used is basically a general channel routing method [91],[92]. The router was designed for use in all the three levels (primitive cell, function block and the final floor plan) of a chip. It can handle cells with variable heights and width. I/O ports are allowed to appear on all the four sides.

Basic assumptions:

We assume that (a) initial placement of routing cells is done before the
Fig. 5.4.3 An Example Of N-logic Block Layout.
Fig. 5.4.4 Block Diagram Of Logic Compiler.
router is called. The composition procedures give the initial placement of primitive operators and a place is to be used for the two top level placement; (b) there are two separate layers (one metal and one poly) available for making wire interconnections. In a routing channel, wires in the two layers should run orthogonally. The leaf cell conventions we mentioned in section 5.2.2 are the basic rules in layout generation and are also followed by the router.

Routing grid:

Each functional block to be routed may be regarded as a set of black boxes placed on a grid each with the origin in the left-bottom corner. The grid is separated by equal spaces in both x and y coordinates (leaf cell convention 3). Grids are used to define tracks in which wires can run. All I/O ports must sit on the grid lines. Particularly, the left and right hand side I/O ports should be located on horizontal grid lines while the top and bottom side I/O ports should be on vertical grids.

The important integer hr:

For cells having different heights, there exists some value, hr, such that the heights of the cells are roughly equal to some integer multiple of hr. For example, 90 might be chosen as the value of hr for the leaf cells in the library so that

\[
\begin{align*}
1H &= 165 \text{ micron} \approx 2 \times hr \\
1.5H &= 251 \text{ micron} \approx 3 \times hr \\
2H &= 341 \text{ micron} \approx 4 \times hr
\end{align*}
\]

By using this assumption, cells can be placed in rows of height hr. A cell may occupy 2, 3, or 4 rows according to its height. When a channel between two adjacent cells is to be routed, cells on the right or top will be relocated to make room for the wires and contact cuts. Each channel could be regarded to have infinite number of tracks to house all the wires passing through it.
The routing scheme:

The routing algorithm used could be divided into two parts: path assignment and track assignment. Let's look at an example. Assume it is desired to connect two points A and B. There may be more than one path to connect the two. The objective of path assignment is to find out a path having minimum number of turns and a minimum wire length. We assumed before that all cells are laid down in rows. According to the initial placement information given, the horizontal adjacency relationship between the cells can be derived and described by an "adjacency graph". The location of the cells are adjusted after aligning each cell. Channels are modeled by a channel graph where a node represents the intersection of a horizontal and a vertical channels. An edge represents the channel between intersections. The nodes in the channel graph are labeled and by using an algorithm similar to Lee's algorithm, the best path is selected. Finally, a path is represented by introducing dummy ports at the intersection of channels.

The algorithm used to do the track assignment is basically the left-edge algorithm. In this pass, the channel graph is broken down into individual vertical and horizontal channels. These channels are routed one after another. When assigning a wire to a track, we take into account the constraints imposed on it by all other wires of the channel.

5.5.2 The Routing Environment

The router is called in each level of composition. It accepts a placement file (*.pla), a netlist file (*.net), an index file and a description file as inputs for leaf cells or function blocks used. There are four output files produced by the router:

1. The file mod.dat contains the description of the block generated. It has the same format as the file "leafindex" defined in section 5.2 and is used as an input when higher level composition is carried out.
(2) The file mod2.dat gives the placement information of each cell inside the block after the wire routing is done. It has the same format as the file ".pla" defined in section 5.3.

(3) The file wire.dat describes all the wire connections in the block, such as the origin and the relative position of each wire in the channel, the physical attributes of each wire (the name of the layer that a wire is going to run, the length, width and contact cuts), etc. The format of the file is:

```
<# of channel>
channel <channel #> <x,y coordinate>
```

```
* net <net #> *
[layer/contact-cut] <length> <width> <> <<
```

(4) The file "result" is intended for debugging. It records all intermediate information generated when the router is run. Fig. 5.5.1 shows the routing environment.

5.5.3 Generating Layout Geometry

The CIF generator is a program module to produce the CIF file of a routed
Fig. 5.5.1 The Routing Environment.
primitive or function block. Data generated by the router (in file wire.dat) is local --- specifying the relative position of wires in the channel. Transformation from the local information to a global floor plan is conducted. The basic components, lower level cell layout in KIC format, and the wire connections and contact cuts are converted into the CIF format. Fig. 5.5.2 shows the block diagram of the CIF generator.
Fig. 5.5.2 Block Diagram Of CIF generator.
5.6 Generation Of The Layout From The Intermediate Code

In previous sections we have introduced the main modules involved in the low level translation process and shown how the compilation proceeds in each of them. We now give a brief discussion on the design and implementation of the interpreter first. Then we will show how these modules are integrated and how a layout is generated from the intermediate form.

5.6.1 The Interpreter

The Interpreter is a program which works as a lexical analyzer of SASCIL. As we mentioned in chapter 4, there are seven kinds of data types in the intermediate form. They are: signal, port, temporary, register, boolean, library and event. Also, there are two types of statements: assignment statement and library/function call. An assignment statement is a simple statement which may apply to signal, port, temporary, register and boolean data types. A function call is a set of statements which may contain both assignment statement and library/function calls. The main objective of the Interpreter is to map logic description of data into corresponding physical ones. For all meaningful combinations of the data types, corresponding translation procedures are designed and implemented. The formats of a set of data files are defined. Fig. 5.6.1 is the list of procedures designed for the interpretation of the assignment statement. There are seven data files to be created during an interpretation process. They are:

1. File callpri contains all calling primitives in the function block of the current level. It is used as an input data file accessed by the composition routines as we defined in section 5.3.2.

2. File frecord keeps a record of all data files generated during the
<table>
<thead>
<tr>
<th>Left</th>
<th>Signal</th>
<th>Port</th>
<th>Temp</th>
<th>Register</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right</td>
<td>Signal</td>
<td>Port</td>
<td>Temp</td>
<td>Register</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5.6.1 List Of Procedures For Assignment Statements.
Interpretation. It is designed for the low level management and has the following format:

<file-name 1>
<file-name 2>

<file-name n>

(3) File connectout is one of the prime data files to be used to generate the netlist file of a function block. It consists of the internal connections of cells in the block except the common default signal Vdd, Gnd and Clocks. It is of the following form:

c <cell-#> <port-#> t ... t <cell-#> <port-#>
c <cell-#> <port-#> t ... t <+ / ->
c <cell-#> <port-#> t ... t <var-name>: signal/port
c <var-name 1>: signal/port t <var-name 2>: signal/port

[Note]: + -- connected to Vdd
-- connected to Gnd

(4) File priports is another prime data file designed for the generation of the netlist file of function blocks. The information contained in this file describes the statues of the calling cells except the common default ones. The format definition of this file is:
(5) File databook is an intermediate data bookkeeping file and contains variable names, types, primitive/function names and their cell numbers.

Consider, for example, the following assignment statement

\[
[yout \text{ port r }] = [ \text{R3 register 7 4 2 n8 n9 n10 n10}] 
\]

It implies the following:

(a) The output of register R3 is connected to an output port called "yout" which is located at the right-hand side of the function block.

(b) Register R3 is a 4-bit shift register (shifts to the right and works under clock phase 2) and the storage devices used are D-flipflop with set and reset functions.

(c) The input, output and the reset signals of the primitive are named n8, n9 and n10 respectively.

When the interpreter is called, it analyzes the grammar of the statement and invokes a procedure called "signal-reg" to pursue the interpretation. The type of storage device "drs" and corresponding primitive library operator "Shiftdrsc2" is selected. The mapping between the logic description of R3 to the physical description of primitive Shiftregdrsc2 is performed. The I/O ports of R3 are the
Internal nodes connecting to other primitive cells or signal wires in the cell composition level. The I/O ports of Shiftregdrsc2 are predefined. The first 12 ports are reserved for Vdd, Gnd and system clocks. Ports 13 to 21 are assigned to the signal n8, n9, and n10 of R3 accordingly.

The data generated in file pports is:

Shiftregdrsc2_R3 5 13 n8 14 n9 15 + 16 n9 17 n9 18 + 19 + 20 n10 21 n10

It says that the primitive Shiftregdrsc2 is to be called to generate cell R3. The cell number of R3 is 5. Port 15 (q-representing \( \overline{q} \)), 18 (set) and 19 (set) are not used. The unused ports 18 and 19 have to be connected to Vdd since the terminal "set" of a D-flip-flop is not allowed to be in uncertain status.

The data generated in file connectout describes the I/O signals to go to outside of the block. In the present example, it should be:

```
c 5 15 t out port r.
```

We interpret it as the following: connecting the port 15 of cell 5 to the port "out" which is on the right hand side of the block.

The content of file callprl declares the 4-bit shift register and to be used by the composition routines later on:

```
5 Shiftregdrsc_R3 4.
```

The netlist file of primitive R3 will be generated when the corresponding composition routine is called.

In conclusion, the interpreter is capable of doing:

(a) Lexical-analysis of the intermediate form.

(b) Generation of a complete set of data for
the current level:
  * netlist
  * calling primitives
the higher level:
  * cell integration
  * logic description of ports
the lower level:
  * physical description of ports.

5.8.2 Generation Of Layout From The SASCIL

Generation of a layout from the intermediate form is a complicated automation process. A number of program modules are involved. The central database is accessed and the intermediate data files are generated. This translation process may be viewed in different ways. From the view of data representation, it can be partitioned into three layers: logical representation, physical description of the data, and the layout geometry. Fig. 5.8.2 shows the three levels.

We give the flowchart of Fig. 5.8.3 as an alternative of the profile of the low level compilation. It shows the main functionalities of the program modules to be used. In order to give a clear explanation, we would like to partition it into three parts: the compilation of the primitive operators, function block/final chip floor plan and combinational logic.

(1) *Compiling the primitive operators:* A primitive operator is generated by going through the solid line of the flowchart. Invocation of a primitive operator, for example the primitive "Adder", is available in the SASCIL file. The interpreter first of all accesses the SASCIL file, conducts the lexical analysis of the code and produces a set of intermediate data files (file caliprl, priports, etc.,
Fig. 5.6.2 The Levels Of Data Representation.
introduced in previous section). Then, the primitive composition module is invoked. By accessing the central database, two leaf cells, "fba" and "drsc2" are used to compose the primitive "Adder". It reads the file "callprl" and invokes the corresponding subroutine to generate the initial placement data of the primitive. Further placement and routing in turn are carried out by the router. The descriptions of the primitive including the replacement of leaf cells, the wiring and the block descriptions are given in files mod1.dat, wire.dat and the mod2.dat, respectively. The module "clfgen" converts these files into the CIF format. We assume that the primitive Adder is going to be used as a cell in a function block. In other words, a KIC format of the primitive is required for higher level composition (a layout only at the topmost level should be presented in the CIF form). So, the program "clftokic" is invoked, the primitive Adder in KIC format is generated. Now, we come to the end of the primitive composition.

(2) Compiling function blocks/chips: As we mentioned in section 5.8.1, the composition of function blocks/chips is from bottom to top. SASC goes to the lowest level of the design hierarchy where there is no user defined functions to be involved. By going to the left branch of the dashed line (see Fig. 5.8.3), and switching to the solid line several times, the translation process for all primitive operators at this level can be finished. The corresponding KIC and data description files are generated. They will be used by the higher level composition. Then, SASC goes up one level of the hierarchy, travels along the right branch of the dashed line, and switches to solid line again. The process is repeated until the topmost level layout is generated.

(3) Compiling combinational modules. The process is same as that we discussed in section 5.4.3 and shown by the dot-dashed line of the flowchart.
Fig. 5.6.3 Block Diagram Of Low Level Compilation Process.
5.7 Generating Layout For the Example Algorithm

In this section, we demonstrate the low level translation process with an example. The intermediate form of the function pe in the FIR algorithm (given in chapters 3) is shown in the end of chapter 4.

The interpreter generates the following data files:

(1) File çalipri contains eleven primitive calls:

1 Dlatchc2_p3 1
2 Rotatec1_p1 4
3 Shiftregdc2_p8 4
4 Shiftregdc2_p7 4
5 Shiftregdc2_p9 5
6 Dlatchc1_p4 1
7 Shiftregdc1_p2 2
8 Adderc2_p5 1
9 Multiplierc1_p6 1
10 Clockgen30c1_p10 1
11 Clockgen30c2_p11 1

(2) File priports:

1 Dlatchc2_p3 13 n1 14 s6 15 !
2 Rotatec1_p1 13 s1 14 n1 15 !
3 Shiftregdc2_p8 13 n6 14 s5 15 !
4 Shiftregdc2_p7 13 n5 14 n6 15 !
5 Shiftregdc2_p9 13 n3 14 s4 15 !
6 Dlatchc1_p4 13 s3 14 n3 15 !
8 Adderc2_p5 13 n2 14 n4 15 ! 16 n5 17 + 18 + 19 n2 20 n4 21 ! 22 + 23 +
0 Multiplier 13 n1 14 n3 15 ! 16 + 17 + 18 n1
19 n3 20 ! 21 + 22 + 23 n4 24 !
10 Clockgen30c1_p10 13 IN1 14 C1 15 C1-
11 Clockgen30c2_p11 13 IN2 14 C2 15 C2-

(3) File connectout:

c 1 14 t s6 signal
c 2 14 t 1 13
c 2 13 t s1 signal
c 3 14 t s5 signal
c 4 14 t 3 13
c 5 14 t s4 signal
c 6 14 t .5 13
c 7 13 t s2 signal
c 7 13 t s3 signal
c 10 13 t port

(4) File frecord:

data/Intermed/callprl
data/Intermed/connectout
data/Intermed/priports
data/Intermed/databook
data/Intermed/pe_0.pdes
data/intermed/frecord

(5) File databook:

p3 register Dlatchc2_p3 s6 1
p1 register Rotateshiftdc1_p1 n1 2
s5 signal
p8 register Shiftregdc2_p8 s5 3
s4 signal
p9 register Shiftregdc2_p9 s4 5
p2 register Shiftregdc1_p2 n2 7
p4 register Dlatchc1_p4 n3 7
p5 lib Adderc2 8 0 1 1 2 11
p6 lib Multiplier 9 0 1 1 2 12
p10 lib Clockgen30c1 10 0 1 1 1 3
p11 lib Clockgen30c2 11 0 1 1 2 3

(6) File pe.predescr:

p1 910 181 15
1 c1 0 176
2 c1 1 176
3 c2 0 180
4 c2 1 160
5 vdd 0 144
6 vdd 1 144
7 gnd 0 48
8 gnd 1 48
9 c2 0 32
10 c2 1 32
The composition module accesses the file callpri and automatically generates the initial placement files for all the primitives except the multiplier (which is not available in the library at the present time. The "multiplier" used is designed only for the demonstration. It consists of a full binary adder and a D-flipflop). Here are the *.pla files.

p2.pla:
The *.net files automatically generated by the composition routines are for primitives p1, p4, p7, p8 and p9. Routing is not required for p2 and p3 since they contain one leaf cell only. File p4.net, p10.net and p11.net are given manually. They are part of the primitive library. The following shows the .net file for p5 (Adderc2).

```plaintext
1 6 11 c1-
1 6 9 c2
1 6 7 gnd
1 6 5 vdd
1 6 3 c2-
1 6 1 c1
r 6 12 c1-
r 6 10 c2
r 6 8 gnd
r 6 6 vdd
```
r 6 4 c2-
 r 6 2 c1
 b 6 13 a
 b 6 14 b
 b 6 15 cout
 b 6 16 sum
 b 6 17 set
 b 6 18 reset
 t 5 19 a
 b 5 20 b
 t 5 21 cout
 t 5 22 set
 t 5 23 reset
 c i 20 t 2 13 t e t 21 3 n
 c 1 17 t 2 1 n
 c 1 14 t e t 1 9 1 n
 c 1 16 t e t 2 0 2 n
 c 2 18 t e t 2 2 4 n
 c 2 20 t e t 2 3 5 n
 c 1 13 t e b 1 3 1 n
 c 1 15 t e b 1 4 2 n
 c 1 19 t e b 1 5 3 n
 c 1 21 t e b 1 6 4 n
 c 2 19 t e b 1 7 5 n
 c 2 21 t e b 1 8 6 n
The data for p2 and p3 in file *.pdes is produced automatically when the composition routine is called. File p2.pdes is a part of the cell description file and the source file for higher level composition and routing.

```
p2 430 165 15
1 c1 0 160
2 c1 1 160
3 c2- 0 144
```
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>c2-</td>
<td>1</td>
<td>144</td>
</tr>
<tr>
<td>5</td>
<td>vdd</td>
<td>0</td>
<td>128</td>
</tr>
<tr>
<td>6</td>
<td>vdd</td>
<td>1</td>
<td>128</td>
</tr>
<tr>
<td>7</td>
<td>gnd</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>gnd</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>9</td>
<td>c2</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>c2</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>11</td>
<td>c1-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>c1-</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>d</td>
<td>0</td>
<td>98</td>
</tr>
<tr>
<td>14</td>
<td>q</td>
<td>1</td>
<td>98</td>
</tr>
<tr>
<td>15</td>
<td>q-</td>
<td>1</td>
<td>112</td>
</tr>
</tbody>
</table>

The debugging files *.dir and *.deb, are shown in Fig. 5.7.1.

Now, let's see an example of replacement and routing for primitive p1. The router takes the above data files and generates the files mod1.dat (replacement), wire.dat (wiring) and mod2.dat (block description) shown below:

File mod1.dat:

```
p1 4 7

d latchc1 16 16
:d latchc1 1136 16
:d latchc1 2256 16
:d latchc1 3376 16

channel 1 0 16 length 160 width 1
channel 2 1136 16 length 160 width 0
```
(a) pl.kic.

Primitive name: Shiftregdcl_pl wordlength 4
*** Looking for primitive = Shiftregdcl 4 ***

dlatchcl 206 165 15

dlatchcl 206 165 15

dlatchcl 206 165 15

dlatchcl 206 165 15

dlatchcl 206 165 15

(b) pl.deb

Fig. 5.7.1 Files pl.kic and pl.deb.
channel 3 2256 16 length 160 width 0
channel 4 3376 16 length 160 width 0
channel 5 4496 16 length 160 width 1
channel 6 0 0 length 4496 width 1
channel 7 0 192 length 4496 width 0

The wire.dat:

** net 1 **
p 0 0 5 96
c 0 96
m -18 96 0 101
m \ 0 96 -2 101
c 0 96

** net 2 **
m -18 112 -2 117

** net 3 **
m -18 0 -2 5

channel 3

** net 1 **
m -18 0 0 5

** net 2 **
** net 3 **
m -18 29 0 39

m 0 0 4496 5
c 0 0
c 4496 0
p 0 0 5 18
p 4496 0 4501 18

File mod2.dat:

p1 4494 181 15
1 c1 0 176
2 c1 1 176
3 c2- 0 160
4 c2- 1 160
5 vdd 0 144
6 vdd 1 144
7 gnd 0 48
8 gnd 1 48
9 c2 0 32
10 c2 1 32
The first two files associated with KIC files used in the leaf cell library are called by the module "cifgen" and the corresponding CIF file, pl.cif, is generated in the following format.

pl.cif:

(CIF file of symbol hierarchy rooted at p1);

DS 2 1.1;
9 dlatchcl1;
L CNG;
B 19800 5600 10300 4500;
L CPG;
B 19000 5000 10300 4500;
L CPW;
B 18000 4000 10300 4500;
L CNT;
B 2300 1500 8050 3150;

DF;
DS 3 1 1;
@ diatchc1;

DF;

C 1;
E

A KIC representation of primitive p1 is required since it is an operator in the function "pe". The file p1 is generated after the "clftokin" transformation. Fig. 5.7.2 shows the layout of p1. (rotated shift register). By repeating the above process 11 times, all the 11 primitives are created and ready to be used.

Before the function block composition can proceed further, the following data files should be properly prepared:

1) The data description file which is a collection of file mod2.dat given by the router and the file *.pdes produced by the composition routine. Each time mod2.dat or *.pdes file is generated, append them to the end of the existing file.

2) The index file is generated by calling a program LEAFINDEX.GEN. The LEAFINDEX.GEN accesses the data description file, gives the physical position of each cell in the data description file and put them in sorted order in the index file.
(3) All primitive operators to be used must be composed and presented in the KIC format.

Now, we are ready to go to the higher level of the layout generation hierarchy -- to generate the function block "pe". At present, there are two black boxes in the function block layout generation: the design and implementation of a fully automatic placer might not be an easy task and the works involved in the block generator, therefore, are very straightforward: (a) keep track of the layout generation hierarchy; (b) access the data files produced in the lower level composition and (c) generate a netlist file for the router. In the given example, we provide the file pe.pla and pe.net manually. Then, the router and the CIF generator is called to generate the layout of function "pe" which is shown in Fig. 5.7.3. Repeating the function block generation process, the final chip floor plan of the given example can be obtained and is demonstrated in Fig. 5.7.4.
Fig. 5.7.4 Layout of The FIR Systolic Array.
CHAPTER 6

CONCLUSION

The SASC is an automatic design synthesis program specially designed to reduce engineering efforts on investigation and implementation of systolic architectures and algorithms. It accepts systolic systems described in a high level data flow language called SASCL. SASCL is specially designed for SASC and it allows systolic designs to be described at an algorithm level. It produces (1) behavioral simulation results to help validate the algorithms and (2) the corresponding VLSI layout for fabrication in Northern Telecom CMOS1B (5 micron) process. The target architecture supported by SASC is the bit-serial systolic arrays. However, it accepts other types of algorithms (like bit-parallel, word-serial, and word-parallel) and is capable of transforming them into a bit-serial equivalent ones automatically. In order to improve the system throughput, an automatic two-level pipelined implementation (if it is possible) is carried out without user knowledge and interaction. SASC also keeps all timing synchronization user transparent. In addition, other features of SASC such as automatic chip floor plan generation, use of default for power, ground and system clocks, etc. add intelligence and flexibilities to the compiler.

Many of the description languages of existing silicon compilers could be used to describe systolic designs. Unfortunately, some common features of these languages (for example, machine and technology dependency, the generality of data types, data structures and language constructs) make them not ideal for this purpose. On the other hand, the novelties of our compiler, for example, automatic algorithm transformation and separate behavioral simulation and layout generation may require special language constructs. Since none of the existing languages meets all our requirements, a new language, named SASCL, is
proposed.

SASCL reflects the properties of systolic architecture at a high level of data abstraction. It provides a simple and hierarchical manner of describing a systolic design. It supports implicit concurrency and implicit/explicit logical synchronization. Being a data flow language, SASCL follows the basic principles of data flow model of computation: Single Assignment Rule (SAR) and allowing a higher order function take another function as argument. User friendliness is one of the main goals of SASCL design, and it is achieved by special on selection of data types, data structures, language constructs and textual form of the program, etc. The responsibility of a user is restricted to the algorithm level and the probability of design errors is reduced.

In the compilation process, algorithms or behavioral descriptions are translated into silicon or simulation results respectively. It is accomplished in two levels of translation which involve three levels of design:

1. Algorithmic description (SASCL)
2. Logic description (SASCIL)

The high level translation process (from SASCL to SASCIL) may be viewed as an ordered sequence of interpretations. Each step refines the design by adding more details: syntax analysis of a systolic algorithm (a set of instruction in SASCL) is performed and a mapping from non-bit-serial architecture to a bit-serial equivalent is conducted if a non-bit-serial design is given. Then, a decision is made on whether two-level pipelined implementation could be applied to the design. The output of this pass is a set of intermediate code (SASCIL).

SASC is supported by a central database system which contains two levels of library cells (primitives and leaf cells), a technology file (which gives the-current
fabrication technology parameters) and a set of data description files (for example, the description of physical characteristics of leaf cells). Primitive operators are composed by leaf cells and can be parameterized according to the high level algorithm description. The low level translation process (from SASCIL to layout description) is an ordered sequence of interpretations with some loops because of the hierarchical way of building a VLSI design. It, first of all, translates the intermediate code into primitive invocations and logical connections of wires; Then, the composition program accesses the central database, invokes corresponding composition routines, converts logical description of the design into physical ones. Physical information such as the initial placement and netlist files are passed to the router which performs further placement and routing. A module, called cligenerator, in turn, takes the router's output and produces corresponding layout geometry (in CIF format).

For this thesis, the design of SASC, the definitions of SASCL and SASCIL, and SASCIL to layout generation part of SASC are finished. The design and implementation of the leaf cells and primitive libraries are accomplished. The leaf cells in CIF format are ready to be used and a set of composition routines are capable of generating required primitives. The philosophies and translation rules followed by the SASCL to SASCIL translation process are proposed. The low level translator are designed and implemented. A complete example of a fourth order FIR filtering algorithm is used to illustrate the two levels of compilation. The implementation of behavioral simulator is left for future work.
REFERENCES


14. Jin, Lan and Wei-min Zheng, "Analysis And Modeling Of A Splitted-bus Distributed Multiprocessor System," Inter-


34. Kung, H. T., "Systolic Algorithms And Their Implementation," *Proceeding of Seventeenth Hawaii International Conference on System Science 1984*, Honolulu, HI, USA, pp. 4-6, Jan. 84.


40. Lehman, Philip L., "A Systolic(VLSI) Array For Processing Simple Relational Queries," in *VLSI Systems And*


1982.


APPENDIX I:

BNF NOTATIONS

::= means "is defined as".

{} means "any sequence of 0 or more of the items enclosed".

[] means "either 0 or one concurrence of the items enclosed".

| means "or" in exclusive sense.

<> means "non-terminal symbol is enclosed".

" " means "terminal symbol is enclosed".

Items are either names of classes (e.g. "Identifier") or single symbols of the language's alphabet.
APPENDIX II:

SYNTAX DEFINITION OF SASCL

<applications> ::= <layout> | <simulation>

<layout> ::= "layout" <array name> ";"

<global declaration>

<function definition>

<main code>

<array name> ::= <identifier>

<global declaration> ::= \{<const declaration>\} <var declaration>

<const declaration> ::= "const" <const list>

<const list> ::= <constant> \{<constant>\}

<constant> ::= <const name> "=" <numerics> ";"

<const name> ::= <identifier>

<var declaration> ::= "var" <var list>

<var list> ::= <variables> \{<variables>\}

<variables> ::= <name list> ";" <var type>

<name list> ::= <var name> \{"," <var name>\}

<var name> ::= <identifier>

<var type> ::= <single type> | <bits type> | <elements type>

<single types> t:= "signal" | "register" | "flag".

<bits type> ::= "bits" <digit> \{<digit>\} | <const name> "of"

<elements type> ::= elements <digit> \{<digit>\} | <const name> "of"

<var name>
\begin{verbatim}
<function definition> ::= <function head>
   <ports declaration>
   <function body>

<function head> ::= <function name> <parameter list>

<function name> ::= <identifier>

<parameter list> ::= "(" <logic direction> \{ <logic direction> \} ")"

<logic direction> ::= <direction> ":" \{ <I/O pairs> | <Inport name> | <Outport name> \}

<direction> ::= <main sides> ":" | "<" <secondary sides>

<main sides> ::= "L" | "T" | "TL" | "TR"

<secondary sides> ::= "R" | "B" | "BL" | "BR"

<I/O pairs> ::= <I/O pair> \{ <I/O pair> \}

<I/O pair> ::= <Inport name> "," <Outport name> ";"

<Inport name> ::= <identifier>

<Outport name> ::= <identifier>

<ports declaration> ::= <ports name> ":" <ports type>

<ports name> ::= <port name> \{ <port name> \}

<port name> ::= "signal" | <mult bits>

<mult bits> ::= "bits" <digit> | <const name> "of" "signal"

<function body> ::= "begin" <function list> "end" ";"

<function list> ::= <local declaration> <statement list>

<local declaration> ::= \{ <const declaration> \} <var declaration>

<statement list> ::= <statement> \{ <statement> \}

<statement> ::= <assignment statement> | <function call> ";

   \{ <comments> \}

<assignment statement> ::= <normal assign> | <state assign>
\end{verbatim}
<normal assign> ::= <var name> "=" <var name> | <boolean const> | <function call>

<state assign> ::= <var name> "=" <symbols>

<symbols> ::= <identifier> {"," <identifier>}

<function call> ::= <function names> "(" <argument list> ")"

<function names> ::= <function name> | <sysfunction name>

<argument list> ::= {<function call> | <continued symbol> } | <const name> | <digit>

<main code> ::= "begin" <statement list> "end" "."

<sysfunction name> ::= <primitive function> | <library function>

<primitive function> ::= <arithmetic fun> | <logic fun> | <comparison fun> | <other fun>

<arithmetic fun> ::= "add" | "sub" | "mul" | "div" | "mod"

<logic fun> ::= "and" | "or" | "not"

<comparison fun> ::= "max" | "min" | "equal" | "less" | "great" | "lessequal" | "greatequal"

<other fun> ::= "delay" | "link" *

<library function> ::= <topo description> | <algo description> | <simu description>

<topo description> ::= "linear" | "square" | <triangle>

<triangle> ::= "tritriangle" | "trtriangle" | "bittriangle" | "bittriangle"

<algo description> ::= "fsm" | "if" | "else" | "for"

<simu description> ::= "setnode" | "showat" | "active" | "after"

<identifier> ::= <letter> {<letter> | <digit> | "."}

<letter> ::= "a" | "b" | ... | "z" | "A" | "B" | ... | "Z"

<digit> ::= "0" | "1" | "2" | ... | "8" | "9"

<numerics> ::= <digit> {<digit> | ["+" | "]" | <digit> | <digit>}

[Note]: Functions with * sign are for simulation.
<boolean const> ::= "true" | "false"
<continued symbol> ::= <digit> {<digit>} .. <digit> {<digit>} |
     <letter> .. <letter>
<comments> ::= "/*" {<identifier> <numerics>} "*/"
APPENDIX III:

SYSTE M FUNCTIONS

1. PRIMITIVE FUNCTIONS

(a) Arithmetic Functions

Arithmetic functions apply to at least two integers and produce a result. We have:

- \text{add} (a \ b \ c \ ...) \quad \Rightarrow \quad a + b + c + ...
- \text{sub} (a \ b) \quad \Rightarrow \quad a - b
- \text{mul} (a \ b \ c \ ...) \quad \Rightarrow \quad a \times b \times c \times ...
- \text{div} (a \ b) \quad \Rightarrow \quad a / b
- \text{mod} (a \ b) \quad \Rightarrow \quad a \mod b

(b) Logic Functions

Logic functions allow us to combine logical values. We have:

- \text{and} (a \ b \ c \ ...) \quad \Rightarrow \quad a \land b \land c \land ...
- \text{or} (a \ b \ c \ ...) \quad \Rightarrow \quad a \lor b \lor c \lor ...
- \text{not} (a) \quad \Rightarrow \quad \neg a

(c) Comparison Functions

Comparison functions could be applied to signals, registers, and flags. We have:

- \text{max} (a \ b \ c \ ...) \quad \Rightarrow \quad \text{find the biggest}
- \text{min} (a \ b \ c \ ...) \quad \Rightarrow \quad \text{find the smallest}
equal \( (a \ b) \quad \quad a = b \)
less \( (a \ b) \quad \quad a < b \)
great \( (a \ b) \quad \quad a > b \)
lessequal \( (a \ b) \quad \quad a \leq b \)
greatequal \( (a \ b) \quad \quad a \geq b \)

(d) Delay Functions

delay \( (x \ t) \): delay signal \( x \) by \( t \) cycles.

(e) Initialization and link Functions

init \( (x \ v) \): set up connections between I/O pads and variable \( x \) which is a global variable, either a register or a flag data type. The \( v \) represents terminals by which variable \( x \) to be initialized.

The values of \( v \) could be "r" and "s" which corresponds to terminals "set" and "reset" of a storage register respectively. The default represents terminal "d".

link \( (x) \): linking an external program to the calling program.

2. LIBRARY FUNCTIONS

(a) For Network Topology:

linear\( (\text{fname } n) \): generate a logically linearly connected network (the physical layout may not be linear one) with \( n \) function blocks called \text{fname}.

square\( (\text{fname } m \ n) \): generate a 2-D network with \( m \times n \) function blocks, called \text{fname}. Topologies covered by this function are:
. mesh connected networks
. mesh connected networks with diagonal connections between function blocks
. 2-D hexagonal array.

triangle(fname m n): generate a top-left triangle network with size of m * n.

trtriangle(fname m n): generate a top-right triangle network with size of m * n.
bitriangle(fname m n): generate a bottom-left triangle network with size of 
\[ m \times n. \]

brtriangle(fname m n): generate a bottom-right triangle network with size of 
\[ m \times n. \]

(b) For Algorithm Description:

\[ \text{fsm(state1 (condition)(action))} \]
\[ \text{state2 (condition)(action)} \]

\[ \text{staten (condition)(action))}: \text{fsm is used to define finite state machine.} \]
if((expression)(statement1)
    else (statement2)) : is used to either make a decision or represent a
    hardware switch.

(c) For Simulation:

setnode(fname n1 n2 ...): used to set nodes inside the given function. The
    states of the nodes should be listed
    after each simulation steps.

showat(cycle1,cycle2 ...): used to set up the interrupt points of a
    simulation.

load(varname,value): used to set initial value of the variable.

active(fname,mode): simulate the function according to the active mode.

alter(model1 mode2): alternation of model1 and mode2.
APPENDIX IV:

SYNTAX DEFINITION OF SASCIL

<statements> ::= {<statement1> | <statement2>}
<statement1> ::= <items> "=" <items> | <boolean type>
<items> ::= <simple type> | <register type>
<simple type> ::= <signal-temp> | <port>
<signal-temp> ::= "[" <identifier> "signal;" | "temp" "]"
<port> ::= "[" <port head> <1/o> <location> "]"
<port head> ::= <identifier> "port"
<1/o> ::= "1" | "o"
<location> ::= "i" | "r" | "t" | "b"
<register type> ::= "[" <register head> <short list> | <long list> "]"
<register head> ::= <identifier> "register"
<short list> ::= <register type1> <#bits> <phase> <d> <q> <q>
<register type1> ::= "0" | "2" | "3" | "4" | "5"
<#bits> ::= <digit> {<digit>}
<phase> ::= "1" | "2"
<d> ::= <identifier>
<q> ::= <identifier> | <not used>
<q> ::= <identifier> | <not used>
<not used> ::= "1"
<long list> ::= <register type2> <#bits> <phase> <d> <q> <q> <r> <s>
<register type2> ::= "1" | "6" | "7" | "8" | "9"
\[
\text{lopad} := \text{"Inpad" | "Outpad" } \langle\text{common para}\rangle<\text{in1}\rangle
\]

\[
\text{in1} := \langle\text{identifier}\rangle | \langle\text{not used}\rangle
\]

\[
\text{arithmetic} := \langle\text{add}\rangle | \langle\text{multiply}\rangle
\]

\[
\text{add} := \text{"Adder" | "Subtractor" } \langle\text{common para}\rangle<\text{sum}\rangle<\text{out}\rangle
\]

\[
<\text{in1} \langle\text{in1}\rangle<\text{r}\rangle<\text{s}\rangle
\]

\[
\text{sum} := \langle\text{identifier}\rangle
\]

\[
\text{out} := \langle\text{identifier}\rangle | \langle\text{not used}\rangle
\]

\[
\text{multiply} := \text{"Multiplier" | "Add-multiply" } \langle\text{common para}\rangle<\text{product}\rangle
\]

\[
\langle\text{in1} \langle\text{in1}\rangle<\text{r}\rangle<\text{s}\rangle
\]

\[
\text{product} := \langle\text{identifier}\rangle
\]

\[
\text{others} := \langle\text{switch-comp}\rangle | \langle\text{mult}\rangle | \langle\text{demult}\rangle
\]

\[
\text{switch-comp} := \text{"Switchbox4" | "Comparator" } \langle\text{common para}\rangle<\text{out}\rangle
\]

\[
<\text{out} \langle\text{in}\rangle<\text{in}\rangle
\]

\[
\text{mult} := \text{"Multiplexer" } \langle\text{common para}\rangle<\text{out}\rangle<\text{ctl}\rangle<\text{ctl}\rangle<\text{ctl}\rangle<\text{in}\rangle<\text{in}\rangle<\text{in}\rangle<\text{in}\rangle
\]

\[
\text{ctl} := \langle\text{identifier}\rangle
\]

\[
\text{demult} := \text{"Demultiplexer" } \langle\text{common para}\rangle<\text{out}\rangle<\text{out}\rangle<\text{out}\rangle<\text{out}\rangle<\text{out}\rangle<\text{in}\rangle<\text{in}\rangle<\text{in}\rangle<\text{in}\rangle
\]

\[
\text{users lib} := \text{"[" } \langle\text{userlib name}\rangle\text{ "lib" } 1 \langle\text{identifier}\rangle<\text{common para}\rangle<\text{ports}\rangle\text{"]"
\]

\[
\text{"{" } <\text{statements}> "\}}
\]

\[
\text{userlib name} := \langle\text{identifier}\rangle
\]

\[
<\text{ports}\rangle := \{<\text{(" } \langle\text{identifier}\rangle\langle\text{identifier}\rangle | \langle\text{identifier}\rangle\text{")}  >\}
\]

\[
<\text{event type}\rangle := \text{"[" } <\text{event head}\rangle<\text{event parameter}\rangle\text{"]"
\]

\[
\text{"{" } <\text{pathlist}> "\}}
\]

\[
<\text{event head}\rangle := \langle\text{identifier}\rangle \text{"event"}
\]

\[
<\text{event parameter}\rangle := \langle\text{function name}\rangle<\text{level}\rangle<\#\text{path}\rangle<\#\text{cycles}\rangle
\]
\[\text{<function name> ::= <identifier> •}
\]
\[\text{<#path> ::= <digit> {<digit>}}
\]
\[\text{<#cycles> ::= <digit> {<digit>}}
\]
\[\text{<path list> ::= <path name> {<simple type>}}
\]
\[\text{<identifier> ::= <letter> {<letter> | <letter> | <digit> | ""} \]
\[\text{<letter> ::= "a" | "b" | ... | "z" | "A" | ... | "Z"}
\]
\[\text{<digit> ::= "0" | "1" | ... | "9"}
\]
APPENDIX V:

PRIMITIVE LIBRARY
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>148</td>
</tr>
<tr>
<td>Driver8 (Driver15, Driver30)</td>
<td>150</td>
</tr>
<tr>
<td>Ndriver8 (Ndriver15, Ndriver30)</td>
<td>151</td>
</tr>
<tr>
<td>Clockgen8 (Clockgen15, Clockgen30)</td>
<td>152</td>
</tr>
<tr>
<td>Shiftregdc1 (Shiftregdc2)</td>
<td>153</td>
</tr>
<tr>
<td>Shiftregdrsc1 (Shiftregdrsc2)</td>
<td>154</td>
</tr>
<tr>
<td>Rotateshiftdc1 (Rotateshiftdc2)</td>
<td>155</td>
</tr>
<tr>
<td>Rotateshiftdrsc1 (Rotateshiftdrsc2)</td>
<td>156</td>
</tr>
<tr>
<td>Adderc1 (Adderc2)</td>
<td>157</td>
</tr>
<tr>
<td>Subtractc1 (Subtractc2)</td>
<td>158</td>
</tr>
<tr>
<td>Comparator</td>
<td>159</td>
</tr>
<tr>
<td>Switchbox4</td>
<td>160</td>
</tr>
<tr>
<td>Multiplexer</td>
<td>161</td>
</tr>
<tr>
<td>Demultiplexer</td>
<td>162</td>
</tr>
<tr>
<td>Inpad</td>
<td>163</td>
</tr>
<tr>
<td>Outpad</td>
<td>164</td>
</tr>
<tr>
<td>Vddpad</td>
<td>165</td>
</tr>
<tr>
<td>Gndpad</td>
<td>166</td>
</tr>
<tr>
<td>Frame</td>
<td>167</td>
</tr>
<tr>
<td>Logic</td>
<td>168</td>
</tr>
<tr>
<td>Multiplier</td>
<td>169</td>
</tr>
</tbody>
</table>
Name: Driver8 (Driver15, Driver30)

Function: Inverted driving for 8 (15 or 30) loads.

Leaf-Cell-Used: driver8 (driver15, driver30)

Logic-Symbol:

\[ \text{IN} \rightarrow \text{OUT} \]

I/O-Ports: Input: IN

Output: OUT
Name: Ndriver8 (Ndriver15, Ndriver30)

Function: Non-Inverted driving for 8 (15 or 30) loads.

Leaf-Cell-Used: ndriver8 (ndriver15, ndriver30)

Logic-Symbol:

I/O-Ports: Input: IN

Output: OUT
Name: Clockgen8 (Clockgen15, Clockgen30)

Function: Generate complementary local clocks which can drive 8 (15 or 30) loads.

Leaf-Cell-Used: driver8 (driver15, driver30), ndriver8 (ndriver15, ndriver30)

Logic-Symbol:

\[
\begin{array}{c}
\text{CIN} \\
\hline
\end{array}
\quad
\begin{array}{c}
\text{C} \\
\hline
\text{C'}
\end{array}
\]

I/O-Ports: Input: CIN

Output: C, C'
Name: ShiftRegDCl (ShiftRegDCl)

Function: Shift and store data. It is D-latch type shift register and could be triggered by system clock 1 or 2.

Number-of-Bits: n

Leaf-Cell-Used: dlatchc1 (dlatchc2)

Logic-Symbol:

I/O-Ports:

Input: C1(C2), D

Output: Q1, Q1, Q2, Q2, ..., Qn, Qn
Name: Shiftregdrsc1 (Shiftregdrsc2)

Function: Shift and store data. It is D-flip-flop type shift register and could be triggered by system clock phase 1 or 2.

Number-of-Bits: \( n \)

Leaf-Cell-Used: drsc1 (drsc2)

Logic-Symbol:

I/O-Ports:

Input: C1 (C2), D1, R, S

Output: \( Q_1, \overline{Q_1}, Q_2, \overline{Q_2}, \ldots, Q_n, \overline{Q_n} \)
Name: Rotateshiftdc1 (Rotateshiftdc2)

Function: Store data in D-latch type register. It could be triggered by system clock 1 or 2.

Number-of-Bits: n

Leaf-Cell-Used: d latchc1 (d latchc2)

Logic-Symbol:

I/O-Ports: Input: C1 (C2), D

Output: Q1, Q̅1, Q2, Q̅2, ...... Qn, Q̅n
Name: Rotateshiftdrsc1 (Rotateshiftdrsc2)

Function: Store data in D-flip-flop type register. It could be triggered by system clock 1 or 2.

Number-of-Bits: n

Leaf-Cell-Used: drsc1 (drsc2)

Logic-Symbol:

I/O-Ports: Input: C1 (C2), D1 Reset, Set

Output: Q1, \overline{Q1}, Q2, \overline{Q2}, \ldots, Qn, \overline{Qn}
Name: AdderC1 (AdderC2)

Function: Bit addition clocked by system clock 1 or 2. Set and reset functions to the internal bit delay are provided.

Number-of-Bits: 1

Leaf-Cell-Used: fab1, drsc1 (drsc2)

Logic-Symbol: 

I/O-Ports: Input: A, B, R, S, C1(C2)
Output: COUT, SUM
Name: Subtractc1 (Subtractc2)

Function: Bit subtraction clocked by system clock 1 or 2. Set and Reset functions to the internal bit delay are provided.

Number-of-Blts: 1

Leaf-Cell-Used: fba1, drsc1 (drsc2), driver8

Logic-Symbol:

![Logic Symbol Image]

I/O-Ports:

Input: A, B, R, S, C1(C2)

Output: DIFFER, COUT
Name: Comparator

Function: One-bit compare-exchange. The Rin and Rout act as input and output reset lines respectively. Cin and Yln are the corresponding bits of the two numbers compared and the bit corresponding to the larger number is output at H and smaller number is output at L. The comparator works under two-phase system clock.

Number-of-Bits: 1

Leaf-Cell-Used: driver8, ndriver8, driver30, ndriver30, dlatchc1, dlatchc2, drsc1, Hf*, Lf*, C1*, C2*

Logic-Symbol:

\[ \text{C1} \quad \text{C2} \]

\[ \text{RIN} \quad \text{X} \quad \text{Y} \quad \text{ROUT} \quad \text{H} \quad \text{L} \]

I/O-Ports: Input: Rin, X, Y, C1, C2

Output: Rout, H, L

(*) cells are generated by Logic Compiler.
Name: Switchbox4

Function: The four-function switchbox performs data transfer under independent box control:

<table>
<thead>
<tr>
<th>D1</th>
<th>D2</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>straight</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>exchange</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>lower broadcast</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>upper broadcast</td>
</tr>
</tbody>
</table>

Leaf-Cell-Used: dlatec1, dlatec2, Hfunction*, Lfunction*

LogIC-Symbol:

I/O-Ports: Input: RIN, X, Y, D1, D2, C1, C2

Output: ROUT, H, L

[*]: cells are generated by Logic Compiler.
Name: Multiplexer

Function: Two control signals (C0 and C1) are used to selectively connect one of the four inputs (A0-A3) to the output W (Y). Signal EN enables the output.

Leaf-Cell-Used: multiplexer4-1

Logic-Symbol:

I/O-Ports: Input: C0, C1, EN, A0, A1, A2, A3

Output: W, Y
Name: Demultiplexer

Function: Performing the inverse function of a multiplexer: it distributes the input signal D to output lines (F1-F4) designated by the state of input A and B. Signal ST is used to strobe the output.

Leaf-Cell-Used: demultiplexer1-4

Logic-Symbol:

I/O-Ports: Input: ST, D, A, B

Output: F1, F2, F3, F4
Name: Inpad

Function: Receive signals from outside of the chip.

Number-of-Blts: 1

Leaf-Cell-Used: Inpad

Logic-Symbol:

I/O-Ports:

Input: IN

Output: OUT
Name: Outpad

Function: Send internal signals to outside of the chip.

Number-of-Bits: 1

Leaf-Cell-Used: outpad

Logic-Symbol:

\[\text{IN} \rightarrow \text{OUT}\]

I/O-Ports:

Input: IN

Output: OUT
Name: Vddpad

Function: Power pad.

Number-of-Bits: 1

Leaf-Cell-Used: vddpad

Logic-Symbol:

Ports: VDD, GND
Name: Gndpad
Function: Ground pad
Number-of-Blts: 1
Leaf-Cell-Used: gndpad
Logic-Symbol:

Ports: VDD, GND
Name: Frame

Function: Chip frame.

Leaf-Cell-Used: frameA, frameB, frameC, frameD or free size frame.

Logic-Symbol:

[†]: Free size frame is program-generated.
Name: Logics

Function: Generate boolean expression mask geometry according to user specification.

Logic-Symbol:

I/O-Ports: Input: C, V1, V2, ..., Vn
Output: F

[*] Limitations on the number of products and variables in a boolean equation are: if P is infinite, V should be less than or equal to 9, and vice versa.
Name: Multiplier

Function: Bit multiplication:

Product = A * B.

Number-of-Bits: 1

Leaf-Cell-Used: 

Logic-Symbol:

I/O-Ports: Input: A, B, Clock

Output: Product
Name: Add-multiplier

Function: Bit add and multiplication.

Number-of-Bits: 1

Leaf-Cell-Used:

Logic-Symbol:

I/O-Ports: Input: A, B, C

Output: OUT
APPENDIX VI:

LEAF CELL LIBRARY
Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>172</td>
</tr>
<tr>
<td>Driver8 (Driver15, Driver30)</td>
<td>173</td>
</tr>
<tr>
<td>Ndriver8 (Ndriver15, Ndriver30)</td>
<td>174</td>
</tr>
<tr>
<td>Diatchc1 (Diatchc2)</td>
<td>175</td>
</tr>
<tr>
<td>Drsc1 (Drsc2)</td>
<td>176</td>
</tr>
<tr>
<td>Fba1</td>
<td>177</td>
</tr>
<tr>
<td>Multiplexer4-1</td>
<td>178</td>
</tr>
<tr>
<td>Demultiplexer1-4</td>
<td>179</td>
</tr>
<tr>
<td>Inpad</td>
<td>180</td>
</tr>
<tr>
<td>Outpad</td>
<td>181</td>
</tr>
<tr>
<td>Vddpad</td>
<td>182</td>
</tr>
<tr>
<td>Gndpad</td>
<td>183</td>
</tr>
<tr>
<td>Frame</td>
<td>184</td>
</tr>
<tr>
<td>Logic</td>
<td>185</td>
</tr>
</tbody>
</table>
Name: Driver8 (Driver15, Driver30)

Logic-Symbol:

Function: \( \text{OUT} = \overline{\text{IN}} \)

Circuit-Diagram:

Bounding-Box: 
- Driver8: 139 x 165
- Driver15: 183 x 165
- Driver30: 216 x 165
Name: Ndriver8 (Ndriver15, ndriver30)

Logic-Symbol:

Function: \[ \text{OUT} = \text{IN} \]

Circuit-Diagram:

Bounding-Box: Ndriver8 119 x 165
Ndriver15 140 x 165
Ndriver30 210 x 165
Name: Dlatch1 (Dlatch2)

Logic-Symbol:

Function: \( Q = D \)

Circuit-Diagram:

Bounding-Box: 206 x 185
Name: Drsc1 (Drsc2)

Logic-Symbol:

Function:

\[ Q = D \]

\[ Q = S \mid s=1 \]

\[ Q = R \mid r=1 \]

Circuit-Diagram:
Name: Fabi

Logic-Symbol:

Function: \( \text{SUM} = A + B \)

Circuit-Diagram:

Bounding-Box: 136 x 261
Name: Multiplexer 4-1

Logic-Symbol:

Function: The 4-input multiplexer selects one of the four inputs to appear on the output. The two bits binary code C0, C1 determines which input will appear on the output. When enable is high, we will get the output from W / Y.

Circuit-Diagram:

Bounding-Box: 411 x 314
Name: Demultiplexer 1-4

Logic-Symbol:

Function: Demultiplexer is the opposite of a multiplexer. In the demultiplexer, a binary input address, A or B, determines which one of the four outputs will go low.

Circuit-Diagram:

Bounding-Box: 370 x 341
Name: Inpad

Logic-Symbol:

Function: Receive signals from outside of a chip and pass them to the VLSI circuit.

Bounding-Box: 375 x 260
Name: Outpad

Logic-Symbol:

Function: Receive signals from VLSI circuit and send them to outside.

Bounding-Box: 396 x 258
Name: Vddpad

Logic-Symbol:

![VDD to GND diagram]

Function: Pad to feed power for the chip.

Bounding-Box: 165 x 260
Name: Gndpad

Logic-Symbol:

Function: The ground pad.

Bounding-Box: 165 x 260
Name: Frame

Logic-Symbol: 

Function: The chip frame.

Bounding-Box: 

Size A: 4510 * 4510

Size B: 4510 * 2250

Size C: 2250 * 2250

Size D: 1490 * 2250

Free size: program generated
Name: Logic

Logic-Symbol:

Function: Generate boolean expression mask geometries of Hf, Lf, C1, C2, Hfunction and Lfunction which are the control parts of primitive "Comparator" and "Switchbox4".

Circuit-Diagram

Bounding-Box On-line generated.