Ain Shams University
Faculty of Science
Department of Mathematics

Designing Fast Algorithms Using Parallel Models

Thesis
Submitted for the partial fulfillment of
the requirements of the M.Sc. degree
in computer science

Submitted by
Ibrahim Mohamed Abd Elmokatder Gad

Supervised by

Ass. Prof.
Sameh Sami Daoud
Department of Mathematics,
Faculty of Science,
Ain Shams University.

Ass. Prof.
Hazem Mohamed Bahig
Department of Mathematics,
Faculty of Science,
Ain Shams University.

Submitted to
Department of Mathematics,
Faculty of Science,
Ain Shams University,
Cairo-Egypt

2014

Contents

List of Tables

List of Figures


Acknowledgment

Acknowledgements

Praise be to Allah who favored me with capability and patience to complete this work.

I would like to thank **Prof . Dr. Sameh S. Daoud** Ass. Professor Emeritus of Mathematics, Computer Science Division, Faculty of Science, Ain Shams University for guiding me throughout the thesis work. I am grateful for his kind help and valuable discussions. I would like to express my sincere thank and great gratitude to **prof. Dr. Hazem M. Bahig** Associate professor, Department of Mathematics, Faculty of Science, Ain Shams University for his continuous help, valuable discussion and supervisor. I always admire with his experience, quiet manner, and life logic.

I am also grateful *my family* specially my parents for their patience, understanding and encouragement. All appropriations for the academic staff of Computer Science, Department of Mathematics, Faculty of Science, Ain Shams University. Many thanks also go to the staff members of Department of Mathematics, Faculty of Science, Tanta University.


Abstract

Abstract

RSA is one of the most important public key cryptosystems for information security. The security of RSA depends on integer factorization problem, it relies on the difficulty of factoring large integers. The General Number Field Sieve algorithm (GNFS) is currently the best known method for factoring large numbers over than 100 digits. The algorithm consists of five main steps. The third step, which is sieving step, takes the most time consumed in GNFS algorithm as proved in previous researches.

In this thesis, we improve the running time of this step by parallels it on a cluster system. We proposed the different techniques to improve the running time of this step. The experimental results have shown that the algorithm has achieved a good speedup and can be used for factoring a large integers.


Summary

Summary

Factoring is very important in the field of cryptography, specifically in the RSA cryptosystem. RSA algorithm is one of the most important public key cryptosystems for information security. RSA is used in real world applications such as: internet explorer, email systems, online banking, digitally signing, and encrypting email.

Due to the importance of this algorithm, cryptanalysts have been working for decades to identify weaknesses in the algorithm. The security of RSA depends on integer factorization problem, preciously relies on the difficulty of factoring large integers. Till now the problem of factoring integers is still not solved practically, which means there is no deterministic polynomial time algorithm for factoring integers. Much research has gone into the problem of factoring a large number. Due to advances in factoring algorithms and computing hardware the size of the number that can be factorized increases year by year.

There are many integer factorization algorithms used to factor large numbers. For example: Trial division, Pollards p-1 algorithm, Lenstra Elliptic Curve Factorization (ECM), Quadratic Sieve (QS) and General Number Field Sieve (GNFS) algorithm. The General Number Field Sieve algorithm (GNFS) is currently the best known method for factoring large composite numbers over than 100 digits.

Although the GNFS algorithm is efficient, it still takes a long time to factor a large integer such as an integer with 150-digits or larger. In order to reduce the execution time, one natural solution is to distribute jobs to parallel computers. The GNFS algorithm contains several time consuming steps. The most time consuming step is the sieving step which is used to generate enough relations. This step is very suitable for parallelization because the relation generations are independent. This thesis describes an implementation of the General Number Field Sieve (GNFS) algorithm using C language. This thesis, presents a new two methods for a parallel sieving step on GNFS implemented on a BA-cluster.

The thesis consists of five chapters, two appendices, and the bibliography.

Chapter One: It briefly shows the importance of cryptography and the RSA cryptosystem.

Chapter Two: This chapter introduces the basic concepts related to parallel computing. We give an overview for an introduction about why we need parallel computing, definition of parallel computers, motivation of parallel computing, applications area for parallel computing, and the keys that effect on parallelism. We give a brief description about the parallel architectures, also an introduction to parallel algorithms and the complexity measures for any parallel algorithm, the different types of parallel programming models. We also give a brief description about the library MPI that is used in the thesis.

Chapter Three: The aim of this chapter is to give an overview about integer factorization algorithms. We will focus on one of these algorithms that is used to factor large integers. We give mathematical definitions related to factorization problem. We give the importance of factorization problem in RSA, and different algorithms for factorization. We introduce, in details, one of the best algorithm for integer factorization which is the general number field sieve (GNFS). We give an example of the steps of GNFS algorithm.

Chapter Four: This chapter gives the two improved algorithms for sieving step in a cluster system. We give the reasons for selecting the sieve step as a main object for our research, and an overview for serial sieve step. We give an overview for previous works for parallel sieve step. We propose new method for parallel sieve step on a cluster system. Also we describe the configuration of hardware and software, used to implement the parallel sieving step on the cluster system, the experimental results for the proposed methods are given.

Chapter Five: This chapter gives the conclusions and the expected future work that will improve the performance of GNFS.

Appendix A: It includes algorithms for division.

Appendix B: It includes the steps for how to run a parallel program at BA cluster.