Cache Oblivious Algorithms

Zhang JiaHui
Neel Kamal

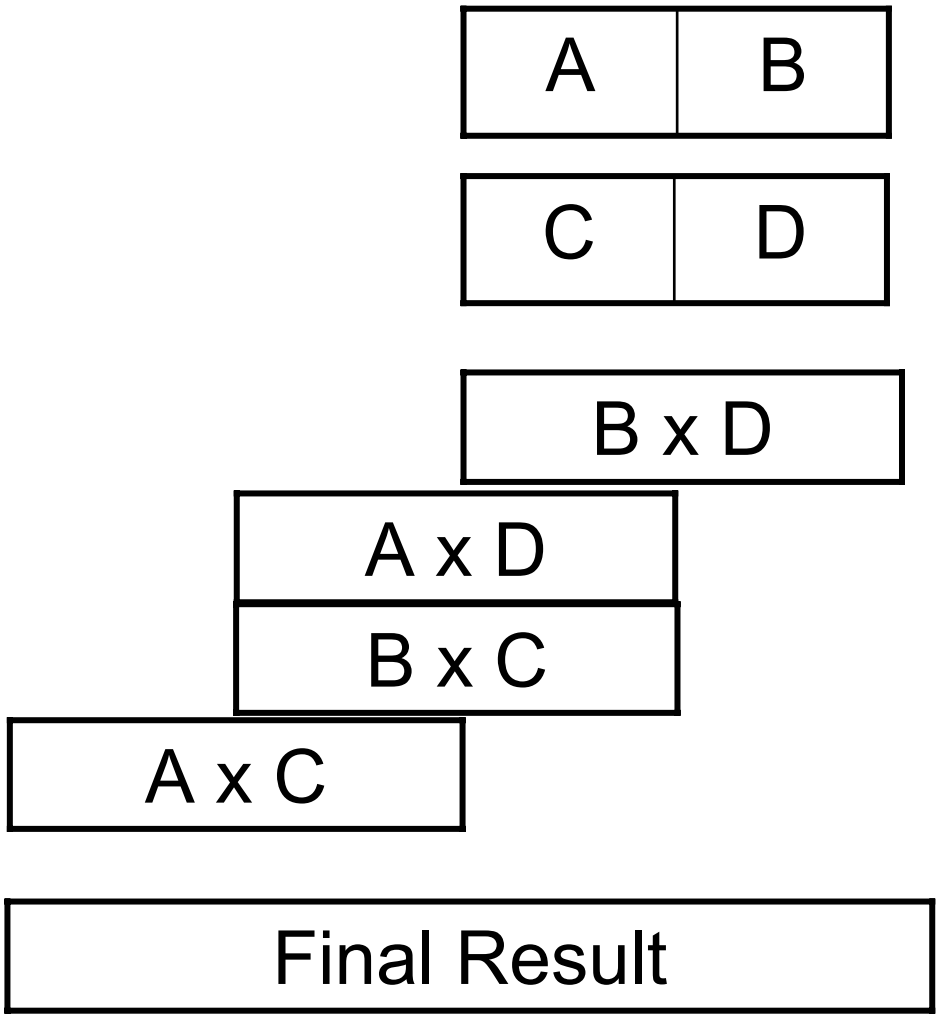
Introduction

- Cache Oblivious vs Cache Aware
- (Z,L) Idea-Cache-Model $Z = \Omega\left(L^2\right)$
- Large Integer Multiplication & RSA
- Dynamic Programming
  - Floyd All-Pair Shortest Paths
  - Longest Common Sequence
- Cache-Behavior Simulator
- Experimental Results

Large Integer Multiplication(1)

- We have two large Integer x and y. x has m digits and y has n digits
- If m>n, append zeros to the left side of n
- If n>m, append zeros to the left side of m
- Suppose m>n, we now have two large Integers both of length m

Large Integer Multiplication(1)



Large Integer Multiplication(1)

m > n

CASE I $m > \frac{\alpha Z}{4}$

$$Q(m) = \begin{cases} \theta\left(\frac{4m}{L}\right) & , if \ \left(m \in \left[\frac{\alpha Z}{8}, \frac{\alpha Z}{4}\right]\right) \\ 4Q\left(\frac{m}{2}\right) + O(1) & , otherwise \end{cases}$$

After k steps, m $\rightarrow \frac{m}{2^k} \in \left[\frac{\alpha Z}{8}, \frac{\alpha Z}{4}\right]$

$$Q(m) = 4^k Q\left(\frac{m}{2^k}\right) = 4^k \theta\left(\frac{4\frac{m}{2^k}}{L}\right) = \theta\left(\frac{16m^2}{LZ}\right)$$

Large Integer Multiplication(1)

m>n

CASE II $m < \frac{\alpha Z}{4}$

$$Q(m) = \theta\left(\frac{4m}{L}\right)$$

If n>m,

$Q(n) = \theta\left(\frac{16n^2}{LZ}\right)$ in CASE I, and $Q(n) = \theta\left(\frac{4n}{L}\right)$

So, combine all the cases, we have

$$Q(n) = \theta\left(\frac{n^2}{LZ} + \frac{m^2}{LZ} + \frac{m}{L} + \frac{n}{L}\right)$$

Large Integer Multiplication(2)

- We do not append zero to the left hand side of the shorter Integer
- CASE I

$$m, n > \alpha Z$$

$$Q(m,n) = \begin{cases} \theta\left(\frac{m}{L} + \frac{n}{L} + \frac{m+n}{L}\right) & ,if\ (m,n \in \left[\frac{\alpha Z}{2}, \alpha Z\right] \quad) \\ 2Q\left(\frac{m}{2}, n\right) + O(1) & ,otherwise, \quad if\ (m > n) \\ 2Q\left(m, \frac{n}{2}\right) + O(1) & ,otherwise \end{cases}$$

Large Integer Multiplication(2)

- After k1 steps m $\rightarrow \frac{m}{2^{k1}} \in \left[\frac{\alpha Z}{2}, \alpha Z\right]$
- After k2 steps n $\rightarrow \frac{n}{2^{k2}} \in \left[\frac{\alpha Z}{2}, \alpha Z\right]$

$$Q(m,n) = 2^{k1}2^{k2}Q\left(\frac{m}{2^{k1}}, \frac{n}{2^{k2}}\right) = 2^{k1}2^{k2}\theta\left(\frac{2\left(\frac{m}{2^{k1}} + \frac{n}{2^{k2}}\right)}{L}\right) = \theta\left(\frac{2\left(m2^{k2} + n2^{k1}\right)}{L}\right) = \theta\left(\frac{mn}{LZ}\right)$$

Large Integer Multiplication(2)

- CASE II $m < \alpha Z$

$$Q(m,n) = \begin{cases} \theta\left(\frac{m}{L} + \frac{n}{L} + \frac{m+n}{L}\right) & , if \ \left(n \in \left[\frac{\alpha Z}{2}, \alpha Z\right] \right) \\ 2Q\left(m, \frac{n}{2}\right) + O(1) & , otherwise \end{cases}$$

$$Q(m,n) = \theta\left(\frac{mn}{LZ} + \frac{n}{L}\right)$$

If $n < \alpha Z$ $\quad Q(m,n) = \theta\left(\frac{mn}{LZ} + \frac{m}{L}\right)$

Large Integer Multiplication(2)

- CASE III $m, n < \alpha Z$

$$Q(m,n) = \theta\left(\frac{m}{L} + \frac{n}{L} + \frac{m+n}{L}\right)$$

- Combine all the cases:

$$Q(m,n) = \theta\left(\frac{m}{L} + \frac{n}{L} + \frac{m+n}{L} + \frac{mn}{LZ}\right)$$

- Total work

$$O(mn)$$

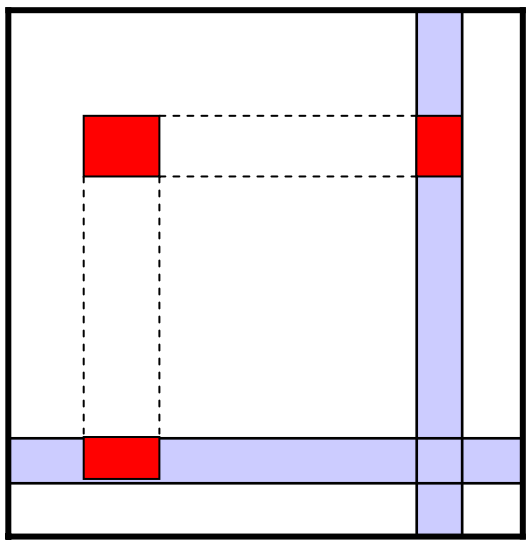
Large Integer Multiplication & RSA

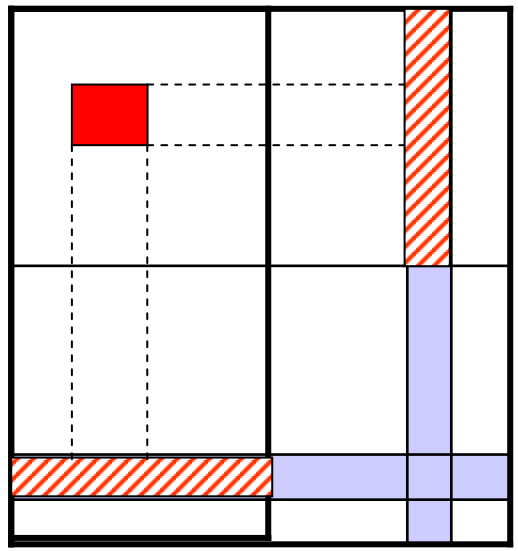
Summary of RSA

- **n = pq** where p and q are distinct primes.
- phi, **φ = (p-1)(q-1)**
- e < n such that gcd(e, phi)=1
- d = e^-1 mod phi.
- c = m^e mod n.
- m = c^d mod n.

All-pair shortest Paths

- Floyd $O(n^3)$
- for k=1 to n
- for i=1 to n
- for j=1 to n
- d[i][j][k]=min(d[i][j][k-1],d[i][k][k-1]+d[k][j][k-1]
- For each k (1..n)





All-pair shortest Paths

- For each iteration of k
- CASE I $n > \alpha\sqrt{Z}$

$$Q(n) = \begin{cases} \theta\left(\frac{n^2}{L}\right) & , if \quad \left(n \in \left[\frac{\alpha\sqrt{Z}}{2}, \alpha\sqrt{Z}\right]\right) \\ 4Q\left(\frac{n}{2}\right) + O(1) & , otherwise \end{cases}$$

- After k steps

$$n \to \frac{n}{2^k} \in \left[\frac{\alpha\sqrt{Z}}{2}, \alpha\sqrt{Z}\right] \qquad Q(n) = 4^k Q\left(\frac{n}{2^k}\right) = 4^k \theta\left(\frac{\left(\frac{n}{2^k}\right)^2}{L}\right) = \theta\left(\frac{n^2}{L}\right)$$

All-pair shortest Paths

- CASE II $n < \alpha\sqrt{Z}$

$$Q(n) = \theta(n)$$

- Combine the cases:

$$Q(n) = \theta\left(n + \frac{n^2}{L}\right)$$

- We have n iterations,

$$Q^{total}(n) = \theta\left(n^2 + \frac{n^3}{L}\right)$$

Longest Common Sequence

- We have 2 long sequences x and y, x is of length m, and y is of length n.
- Try to find the Longest Common Sequence of x and y.
- Dynamic Programming

Longest Common Sequence

		y1	y2	y3	y4	y5	y6
		B	**D**	**C**	**A**	**B**	**A**
x1	**A**	0	0	0	1	1	1
x2	**B**	1	1	1	1	2	2
x3	**C**	1	1	2	2	2	2
x4	**B**	1	1	2	2	3	3
x5	**D**	1	2	2	2	3	3
x6	**A**	1	2	2	3	3	4
x7	**B**	1	2	2	3	4	4

If x[i] == y[j]

c[i][j] = c[i -1][j -1]+1;

else

c[i][j] = max{c[i -1][j];
c[i][j -1];

Longest Common Sequence

		y1	y2	y3	y4	y5	y6
		B	**D**	**C**	**A**	**B**	**A**
x1	**A**	0	0	0	1	1	1
x2	**B**	1	1	1	1	2	2
x3	**C**	1	1	2	2	2	2
x4	**B**	1	1	2	2	3	3
x5	**D**	1	2	2	2	3	3
x6	**A**	1	2	2	3	3	4
x7	**B**	1	2	2	3	4	4

	B	**D**	**C**	**A**	**B**	**A**
A	0	0	0	1	1	1
B	1	1	1	1	2	2
C	1	1	2	2	2	2
B	1	1	2	2	3	3

Longest Common Sequence

- CASE I $m, n > \alpha\sqrt{Z}$

$$Q(m,n) = \begin{cases} \theta\left(\frac{mn}{L}\right) & ,if\ (m,n \in \left[\frac{\alpha\sqrt{Z}}{2}, \alpha\sqrt{Z}\right] \quad) \\ 2Q\left(\frac{m}{2}, n\right) + O(1) & ,otherwise\ , \quad if\ (m > n) \\ 2Q\left(m, \frac{n}{2}\right) + O(1) & ,otherwise \end{cases}$$

Longest Common Sequence

Suppose:

After k1 steps, m $\rightarrow \frac{m}{2^{k1}} \in \left[\frac{\alpha\sqrt{Z}}{2}, \alpha\sqrt{Z}\right]$

After k2 steps, n $\rightarrow \frac{n}{2^{k2}} \in \left[\frac{\alpha\sqrt{Z}}{2}, \alpha\sqrt{Z}\right]$

$$Q(m,n) = 2^{k1}2^{k2}Q\left(\frac{m}{2^{k1}}, \frac{n}{2^{k2}}\right) = 2^{k1}2^{k2}\theta\left(\frac{\frac{m}{2^{k1}}\frac{n}{2^{k2}}}{L}\right) = \theta\left(\frac{mn}{L}\right)$$

Longest Common Sequence

CASE II $m < \alpha\sqrt{Z}$

$$Q(m,n) = \begin{cases} \theta(1+m) & , if \ (n \in \left[\frac{\alpha\sqrt{Z}}{2}, \alpha\sqrt{Z}\right] \ ) \\ 2Q\left(m, \frac{n}{2}\right) + O(1) & , otherwise \end{cases}$$

$$Q(m,n) = \theta\left(\frac{mn}{\sqrt{Z}}\right)$$

In the case when $n < \alpha\sqrt{Z}$

$$Q(m,n) = \begin{cases} \theta(1+m) & , if \ (m \in \left[\frac{\alpha\sqrt{Z}}{2}, \alpha\sqrt{Z}\right] \ ) \\ 2Q\left(\frac{m}{2}, n\right) + O(1) & , otherwise \end{cases}$$

$$Q(m,n) = \theta(m)$$

Longest Common Sequence

CASE III $m, n < \alpha\sqrt{Z}$

$$Q(m,n) = \theta(1+m)$$

Combine all 3 cases:

$$Q(m,n) = \theta\left(1 + 2m + \frac{mn}{\sqrt{Z}} + \frac{mn}{L}\right)$$

Total Work

$$O(mn)$$

Cache Oblivious approaches for Dynamic Programming

Dynamic Programming

- to find an optimal solution
- Sub-problems overlap

Approaches

- bottom up (by recursion usually)
- top down but with a table to memorize earlier solutions

Divide and Conquer method to build the table recursively to make the approach cache oblivious?

Cache Simulator

- With a tall cache assumption
- A fully associative cache

$$Z = \Omega\left(L^2\right)$$

Other Assumptions

- No temporary variable put into the cache
- All input data is assumed to be already present in cache.

Results

Summarizing the theoretical Results:

Large Integer Multiplication $Q(m,n) = \theta\left(\frac{m}{L} + \frac{n}{L} + \frac{m+n}{L} + \frac{mn}{LZ}\right)$

All-pair shortest Paths $Q(n) = \theta\left(n + \frac{n^2}{L}\right)$

Longest Common Sequence $Q(m,n) = \theta\left(1 + 2m + \frac{mn}{\sqrt{Z}} + \frac{mn}{L}\right)$

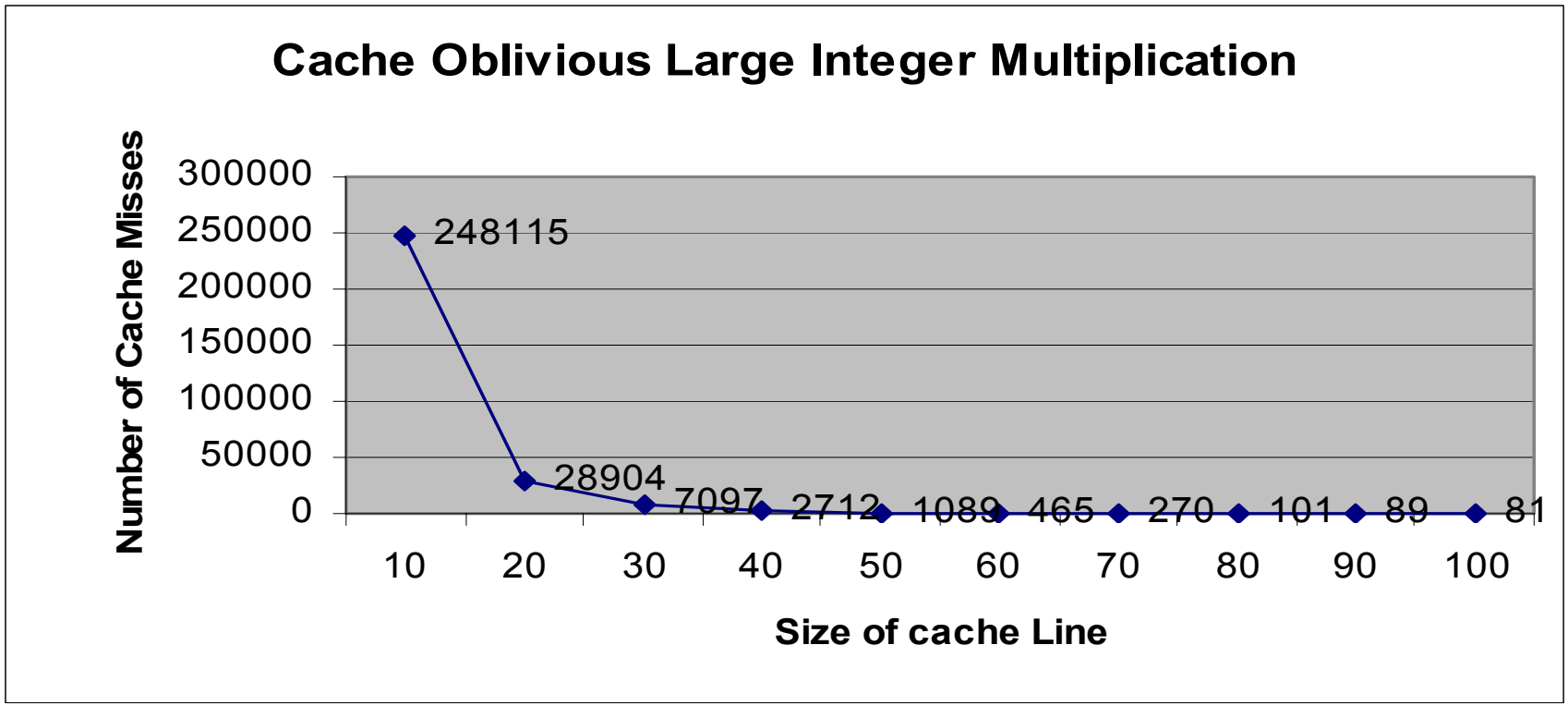
Results from Simulation

Target Machine:

- arch : IA-64
- family : Itanium 2
- CPU MHz : 896.262997
- Cache size : 303312 KB
- OS Linux version 2.4.22
- gcc version 2.96 20000731

We will see that there is a very close match between the theoretical results and the simulation result.

Results



Size of Integer: M = 1000 N = 1000

$$Q(m,n) \approx \theta\left(\frac{mn}{(L)^3}\right)$$

Comparing Results

Case 1: L = 20, Z = 400

Theoretical Result = Θ (1000/8)

Simulator Result = 28904

ratio = 0.0041

Case 2: L=30, Z = 900

Theoretical Result = Θ (1000/27)

Simulator Result = 7097

ratio = 0.0044

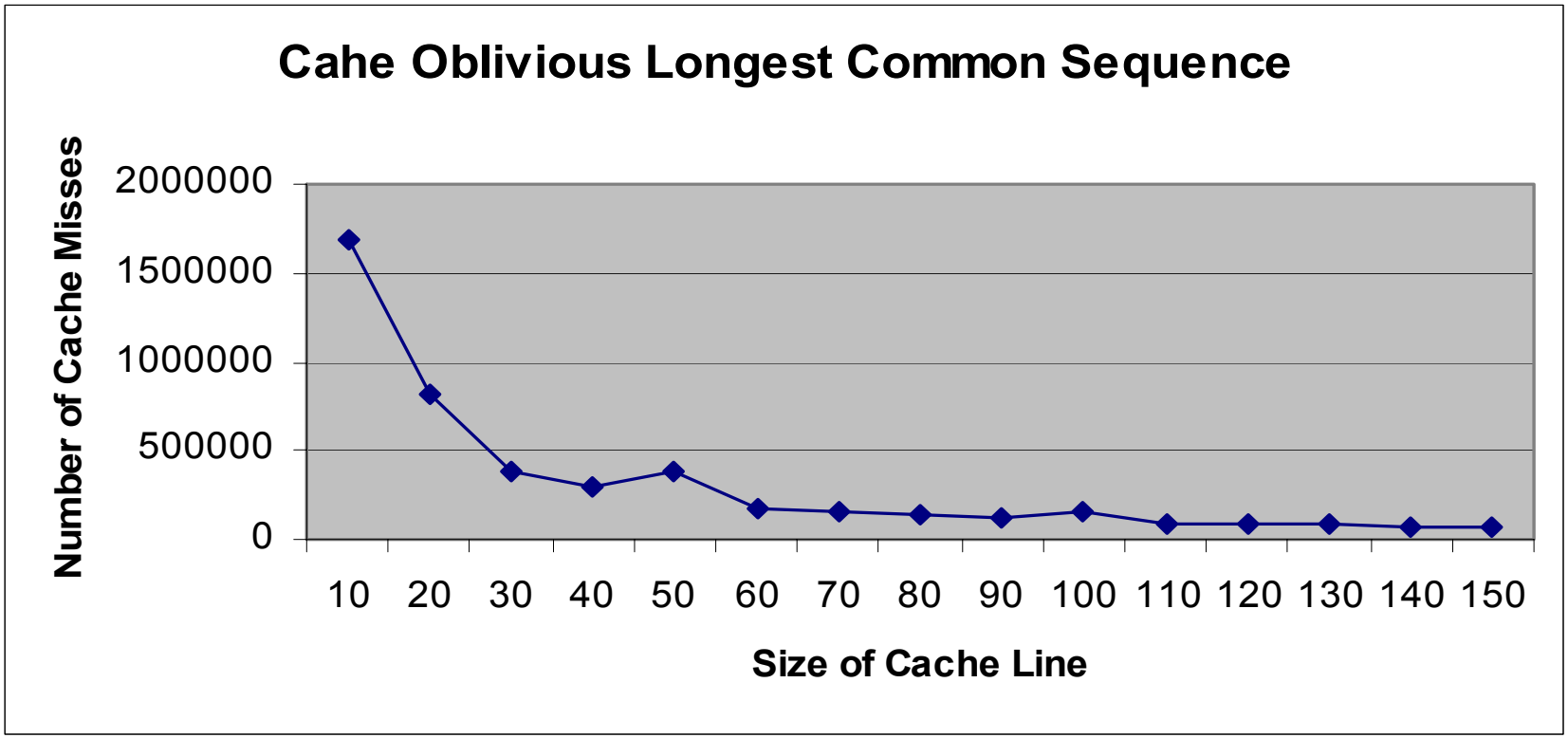
Case 3: L=40, Z = 1600

Theoretical Result = Θ (1000/64)

Simulator Result = 2712

ratio = 0.0052

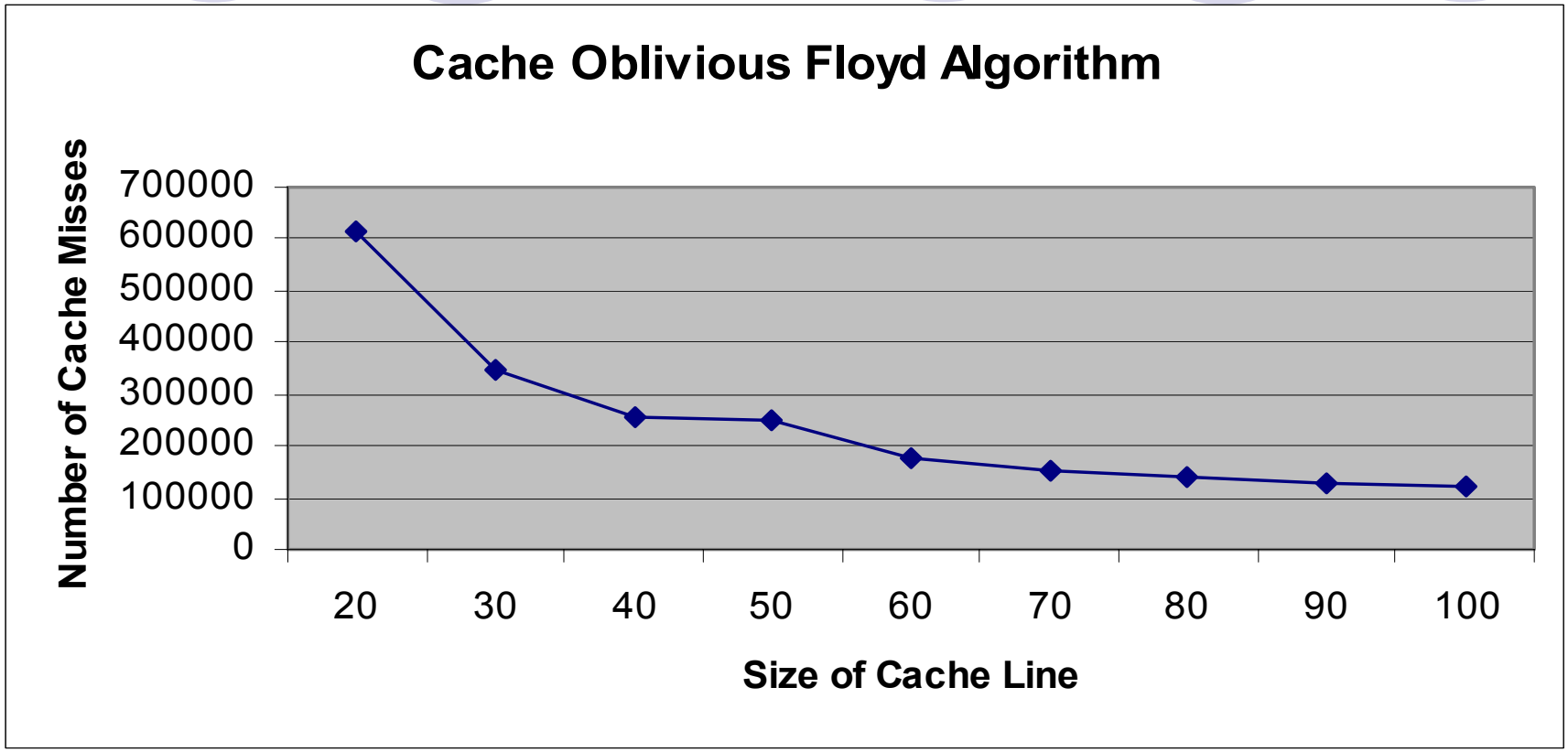
More Results



Size of Sequence: M = 1000 N = 1000

$$Q(m,n) \approx \theta\left(\frac{mn}{L}\right)$$

More Results



Number of Vertices N = 100

$$Q(n) = \theta\left(n + \frac{n^2}{L}\right)$$

Some More Work

We also implemented **Parallel solutions** to each of these problems. We had test results of their performance on CILK.

Thank You