

論文の内容の要旨

論文題目 大規模半導体マスクデータのルールチェック高速化に関する研究
氏名 加藤心

1. はじめに

半導体回路の高集積度化を支える微細加工技術において、光リソグラフィで用いられる光源波長よりもはるかに細かな精度でパターンを転写する必要がある、そのための技術として超解像技術 (Resolution Enhancement Technologies: RET) が用いられている。その RET の中でもフォトマスクパターンに大きな影響をもたらすものが光近接効果補正 (Optical Proximity Correction: OPC) 技術である。この OPC 技術の進行に伴い、半導体回路を記述するフォトマスクデータの複雑度が増大している。マスクデータは通常製造前にマスクルールチェック (Mask Rule Check: MRC) と呼ばれる検証工程を経るが、マスクデータ複雑度の増大によるデータボリュームの爆発的な増加に対して有効な検証手段が提案されておらず、MRC 工程の処理時間増大が問題となっている。そこで、本論文では分散並列技術を基盤として、その上で大規模マスクデータの高速検証システムを構築することを研究対象とする。本論文では、一貫して高速化に対する手法を研究対象としており、高速化のための分散アーキテクチャの設計開発と分散スケジューリングの最適化について論じたものである。

2. MRC 分散並列システム設計目標

大規模なデータ処理問題を解く手段として、分散並列処理方式の有効性は昔から論じられており、マスクデータ処理の分野でも Linux プラットホーム上で広く適用されている。一般的な分散並列処理手法では、マスタープロセスが問題分割を担当し(逐次処理)、 n 台のスレーブプロセスが分割された問題を解く(並列処理)。分散並列処理の性能を表す指標としてスケーラビリティがあるが、そのスケーラビリティの特性を表すものとしてアムダールの法則が知られている。アムダールの法則によれば、並列処理プログラムの実行時間 *Runtime* は T を 1 プロセッサでの処理時間、 s を逐次処理部分の割合、 n をプロセッサ数とすると

$$Runtime = sT + \frac{(1-s)T}{n} \quad (式 1)$$

で表される。これをプロットしたのが図 1 である。ここから分かるように、逐次処理部分割合 s を小さくすることが分散並列処理のスケーラビリティを確保する上で重要な要素となる。

そこで一般的な分散並列処理手法 (マスター分割法) でマスクデータ検証実験を行った。結果として、マスクデータは非常にデータボリュームが大きい反面、MRC 処理に求められる図形演算処理は単純であることから、分割処理時間が支配的となり、逐次処理割合 s が 84% から 11% と高い割合を占めることが分かった。そこで、逐次処理割合 s を低く抑える手法が必要であることが示された。

3. SAMS 型分散並列処理システムの提案

MRC 処理の問題の本質はマスクデータマスクデータの分割処理にあり、システム設計上で最も留意すべき部分は逐次処理であるデータ分割処理をいかにして短縮するかであると示された。そこで本研究では新しい試みとして分散スレーブプロセスが同時に入力ファイルにアクセスするという SAMS (Simultaneous Access by Multiple Slaves) 方式を提案し、高速度 MRC システムの構築を目指す(図 2)。

分散並列システムを構成する要素についてモデリングを行った結果、SAMS 型分散並列システムを構築する上で次の 4 つの問題点が予見される。すなわち(1)巨大ファイルアクセスの問題(2)分割境界の問題(3)巨大セルでの冗

長アクセスの問題(4)冗長データ存在の問題である。そしてその 4 つの問題のうち、ファイルアクセスに関する 3 つの問題 (問題 1,3,4) を解決するために、中間ファイルを設定する。具体的には入力データの必要な部分のみを圧縮したキャッシュファイルと、セル存在アドレスを記述したインデックスファイルを導入し、逐次処理部分割合を減らしつつスレーブプロセスから効率的にアクセスするためのシステム設計を行った。SAMS システムにおける処理フローを図 3 に示す。入力ファイル圧縮の手段として OASIS フォーマットを採用し、OASIS フォーマットの特性を解析して圧縮アルゴリズムを提案した。具体的には、フレキシブル整数表現、RECTANGLE レコード、モーダル表現、Repetition 表現およびポイントリストのデルタ表現と呼ばれるデータ圧縮技術について、個々の要素に関する個別アルゴリズムと、全体処理フローアルゴリズムを導出した。さらに圧縮ファイルの内部構造について、ファイルアクセス効率化の観点から巨大セルの分割と、セルおよび図形レコードの配置座標によるソートを行い、スレーブプロセスからのランダムアクセスを極力排除する手法を採った。また同時に分散並列システムの基本要素であるマスタースレーブ間の制御方法について、過去の研究を引用しながらマスクデータの特性を考慮してスケジューリング手続きとしてワークプール法を採用し、問題の分割方法として正方分割を採用することを決定した。

スレーブプロセスの高速化の観点からは、マルチスレッド手法の採用によるファイルアクセス時間の隠蔽について論じた。そして DRC 処理で一般的な図形ベースの処理について問題点を明らかにし、エッジベースのデータモデルの導入を提案して、基礎実験を行いエッジベースのデータモデルの有効性を明らかにした。

4. 分割境界問題の解決

SAMS 型分割並列システムの実現について予見される 4 つの問題のうち、残る一つの問題、つまりスレーブプロセスにおける分割境界の問題に関する解を示す。まず高速な分散並列システムを構築するにはスレーブ間の干渉を排除する必要性を示し、分割境界部分にマージン領域を設ける手法を採用することを決定し、分割方式として正方分割法を選択した。

一般にマージン領域は処理速度に影響を与えるため、可能な限り小さいことが望ましいとされる。そこで本論文ではマスクチェックのルール値とサイジング量から最小マージン領域サイズの理論値を求めるための手法を構築した。マージン領域とエラールール、サイジング値に関する考察から、一般的に MRC 処理に必要なマージン量 M はルールの最大値を *MaxRuleValue*、サイジング量を *SizingValue* とすると、

$$M = \frac{\text{MaxRuleValue}}{2} + \text{SizingValue} \quad (\text{式 2})$$

と導き出される。そして個々のルールに対して必要なマージン量を算出した。例えば、図 4 に示すように、図形の幅方向と高さ方向について、「(幅 $\leq w$ 、長さ $\geq h$)のエラー図形を検出」というルールを用いる場合、必要なマージン量 $Mwidth$ は以下の式で表される。

$$Mwidth = \frac{A}{2} = \frac{\sqrt{h^2 + w^2}}{2} \quad (\text{式 3})$$

5. SAMS 型分散並列システムの検証

本論文で提案された SAMS 型分散並列 MRC 処理システムについて、実データを用いた実験において速度とスケーラビリティを検証した。6 つの実データに対して $n=12$ の条件で実験を行い、従来のマスター分割方式に対して SAMS 型方式が 6 倍から 50 倍高速であることが示された (図 5)。また逐次処理割合 s は 0.1% から 2.3% に抑えることができた。図 6 にプロセッサ数と速度向上比のグラフを示す。そして、1 テラバイトファイルを $n=100$ で処理する際の推定時間として 176 分という値が得られ、「100 分散で 10 時間以内」という目標値を達成した。実験で得られたその他の知見を以下に記す。

- SAMS 型システムの処理時間はデータ規模に比例する。
- 分割数がスレーブ数に対して十分大きくない場合に、粒度の問題が発生しスケーラビリティが悪化する。
- 分割境界マージンサイズの最適化は特にデータ複雑度が高い場合に有効である。

SAMS 型システムの狙いは、スレーブが分割処理を効率的に行えるために良質な中間ファイルを作成しつつ、逐

次処理である中間ファイル生成を簡素化し高速性を確保するというものであるが、本実験によりこれらの目標が達成されたことが示された。

6. マルチジョブ処理時の先行キャッシュファイル作成スケジューリングの提案

本論文ではここまで一つのマスクデータについて分散並列処理を用いることで高速に検証する手法を論じてきた。ここで観点を改めて、連続する複数のマスクデータ検証ジョブを全体として高速化する技法について提案する。まず SAMS 型分散システムの特徴を観察し、次のジョブのキャッシュファイルを先行して作成するという PCG (Preceding Cache Generation) 手法を提案した(図 7)。そして PCG を用いた時の理論的なスケーラビリティを求め、最大有効分散数 N_{max} が逐次処理割合 s から以下のように算出されることを導いた。

$$N_{max} = \frac{1-s}{s} \quad (式 4)$$

図 8 に PCG を用いた際の理論速度向上度グラフを示す($s=0.1$ を想定)。

そして最大サイズのデータに対して実験を行い、PCG 手法により逐次処理割合 s が 2.3% から 0.3% へと改善されることが示された。これにより、1 テラバイトファイルに対して $n=43$ で処理する際の推定時間として 132 分という値が見込まれる。

7. 最後に

マスクデータの複雑化およびファイルサイズ増大は今後も続くと思われ、本論文で示された SAMS 型分散並列 MRC システムの発展、応用の必要性は高くなるはずである。今後の展望としては、DFM (Design for Manufacturing) 技術の適用による MRC 効率化、EDA 分野への応用、そして HPC (High Performance Computing) や GPU (Graphic Processor Unit) という最新コンピューティング技術の適用などが挙げられる。

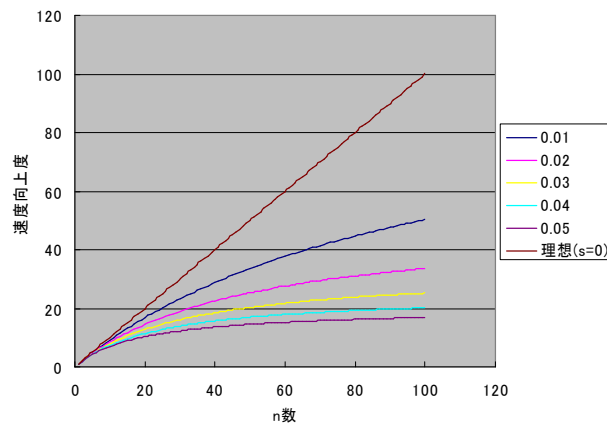


図 1 アムダールの法則 (n 数と速度向上比)

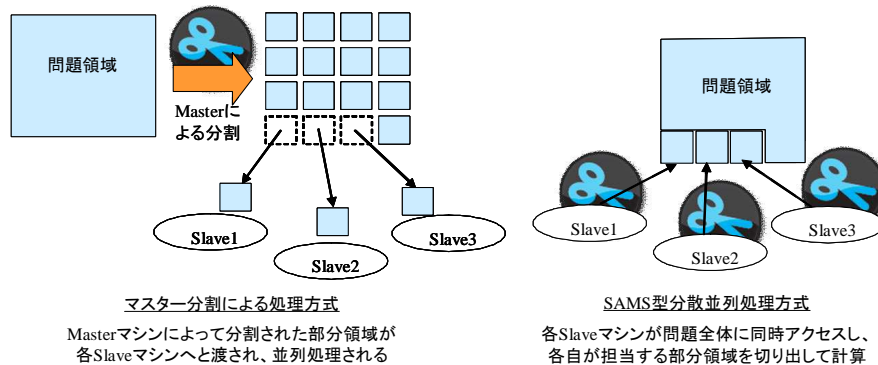


図 2 マスター分割と SAMS 型分散処理方式

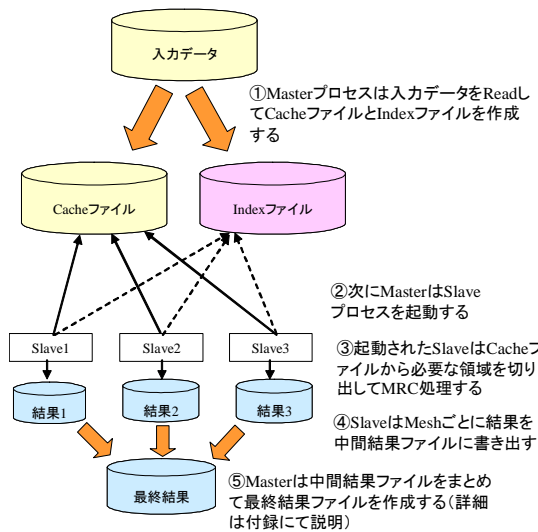


図3 SAMS型分散処理フロー

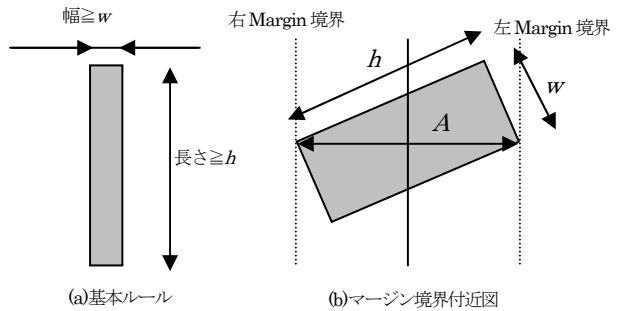


図4 幅チェック時のマージン量

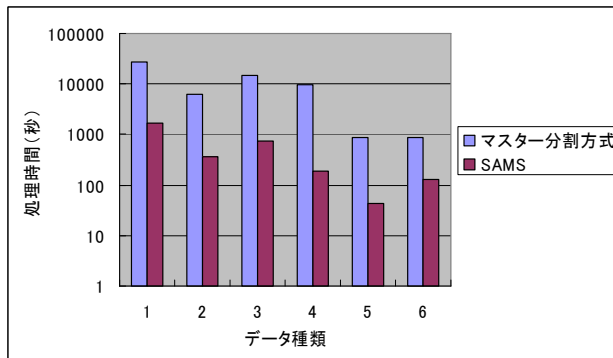


図5 マスター分割方式とSAMSの処理時間比較

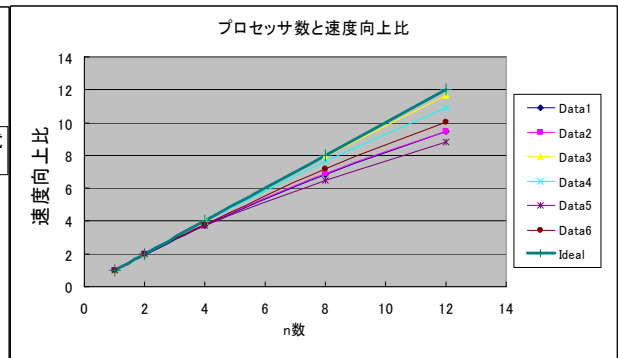


図6 SAMSにおけるプロセッサ数と速度向上比

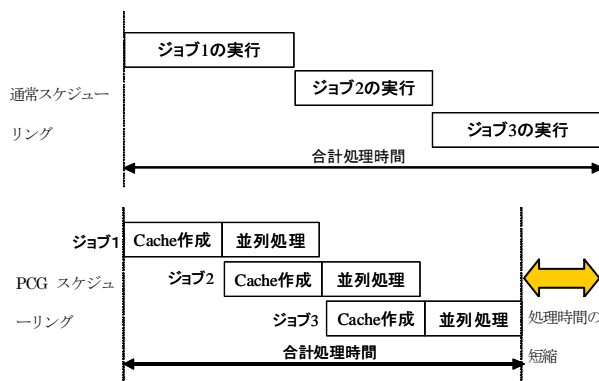


図7 PCGスケジューリング概要

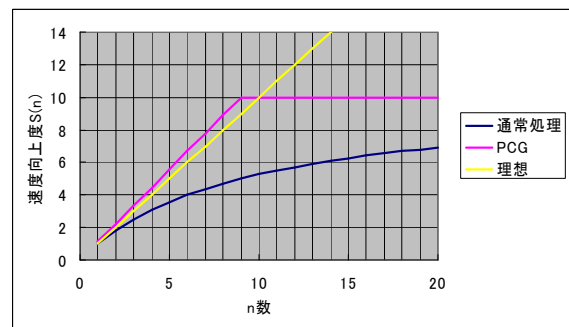


図8 PCGにおける速度向上比(s=0.1)