

論文概要

九州工業大学大学院情報工学府 情報創成工学専攻

学生番号	18675010	氏名	金本 颯将
論文題目	GPUDirect RDMA を用いた高信頼かつ柔軟な障害検知機構		

1 はじめに

近年、情報システムの複雑化に伴い、システム障害を避けるのが難しくなっている。システム障害は迅速かつ正確に検知する必要があるが、従来手法では信頼性の高い障害検知は難しかった。そこで、監視対象ホストに搭載された GPU を用いて障害検知を行う GPU Sentinel [1] が提案されている。GPU Sentinel では、GPU 上で自律的に動作する OS 監視システムがメインメモリ上の OS データを解析し、ビデオメモリにデータを直接書き込むことで障害情報を画面に出力する。しかし、詳細な障害情報を通知するには OS の通信機能を利用する必要があり、障害の影響を受けやすい。

本研究では、OS を介さずに GPU と直接ネットワーク通信を行い、詳細な障害情報をリモートホストに通知するシステム GRASS を提案する。

2 GRASS

GRASS は GPUDirect RDMA を用いることで、図 1 のようにリモート監視システムと GPU 上の OS 監視システムの間で直接通信を行う。リモート監視システムが RDMA Write を用いて GPU メモリに障害検知要求を書き込むと、OS 監視システムはポーリングによりその要求を取得する。OS 監視システムが障害の種類や原因を GPU メモリに格納すると、リモート監視システムは RDMA Read を用いたポーリングによりその情報を取得する。

GRASS を用いることで、GPU とリモートホストを連携させて様々な障害検知が可能になる。例えば、リモート監視システムからの要求受信時に障害検知を行うインライン検知を用いることで、システムへの負荷を抑えることができる。OS 監視システムが定期的に障害検知を行うバックグラウンド検知を用いることで

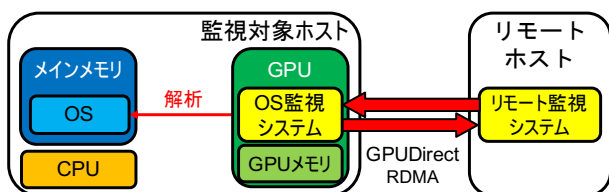


図 1: GRASS のシステム構成

表 1: リモートホストでの障害検知

障害	F1	F2	F3	F4	F5	F6	F7
GRASS	✓	✓	✓	✓	✓	✓	✓
GPU+OS						✓	
OS カーネル		✓	✓		✓	✓	

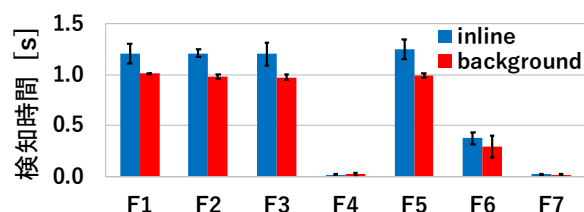


図 2: GRASS における障害検知時間

リモート監視システムからの要求に即座に応答することができる。また、リモート監視システムが OS データを取得して障害検知を行うこともでき、必要に応じてインタラクティブに監視対象ホストの内部状態を調べることができる。

3 実験

GRASS の有用性を確かめるために、OS に 7 種類の障害を挿入してリモートホストでの障害検知の可否と検知時間を調べる実験を行った。比較として、GPU Sentinel と OS の通信機能を組み合わせたシステムおよび、OS カーネル内で障害検知と通知を行うシステムを用いた。表 1 および 図 2 に示す結果から、GRASS では十分に短い時間で様々な障害を検知できることが確認できた。

4 まとめ

本研究では、OS を介さずに GPU と直接ネットワーク通信を行い、障害に関する詳細な情報をリモートホストに通知するシステム GRASS を提案した。今後の課題は、障害発生時にメインメモリ上の OS データを書き換えることで障害からの部分的な復旧を行えるようにすることである。

参考文献

- [1] Y. Ozaki, S. Kanamoto, H. Yamamoto, and K. Kourai. Detecting System Failures with GPUs and LLVM, APSys 2019.