

論文 / 著書情報
Article / Book Information

題目(和文)	最適なチェックポイント・リスタートのための設計及び実装
Title(English)	Design and Implementation for Optimal Checkpoint/Restart
著者(和文)	佐藤賢斗
Author(English)	Kento Sato
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第9422号, 授与年月日:2014年3月26日, 学位の種別:課程博士, 審査員:松岡 聡,遠藤 敏夫,増原 英彦,三好 直人,脇田 建
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第9422号, Conferred date:2014/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

論文審査の要旨及び審査員

報告番号	甲第		号	学位申請者氏名		佐藤 賢斗	
論文審査 審査員		氏 名	職 名		氏 名	職 名	
	主査	松岡 聡	教授	審査員	脇田 建	准教授	
	審査員	遠藤 敏夫	准教授				
		増原 英彦	教授				
		三好 直人	教授				

論文審査の要旨（2000 字程度）

本論文は「Design and Implementation for Optimal Checkpoint/Restart」と題し、大規模スーパーコンピュータを対象とし、チェックポイント・リスタートを中心とした耐障害技術を提案するもので、英語で全6章から構成される。

第1章「Introduction」では、本研究の背景として、大規模スーパーコンピュータの耐障害性における技術的課題を述べ、本研究の位置づけを示したあと、提案手法である Asynchronous checkpointing system, Multi-tier storage design with burst buffer, Fault tolerant messaging interface という3つの概要と、その貢献を簡潔に述べている。最後に、本稿の構成について説明している。

第2章「Fault Tolerance on HPC Systems and the Challenges」では、大規模スーパーコンピュータにおける障害について、定量的データと共に、より詳細に技術的課題を述べたあと、本研究で対象としているチェックポイント・リスタート技術について、関連研究とともに詳細に説明し、既存手法の問題点を示している。

第3章「Asynchronous Checkpointing System」では、一つ目の提案である Asynchronous checkpointing system（非同期型チェックポインティング）について述べられている。この手法は、Multi-level Checkpoint/Restart 技術における、下位の並列ファイルシステムへのチェックポイントを、上位のチェックポイントやアプリケーションの実行と非同期的に行うことで、大幅にオーバーヘッドを削減している。また、この非同期チェックポインティングの複合的確率モデルを構築し、既存の同期型チェックポイントと比較した評価でその優位性を示している。

第4章「A User-level Infiniband-based File System and Checkpoint Strategy for Burst Buffer」では、二つ目の提案である Multi-tier storage design with burst buffer について述べられている。この提案では、信頼性を向上させるために、Burst Buffer を用いた大規模スーパーコンピュータのストレージデザインに着目し、この Burst buffer 用に最適化された I/O インタフェース（IBIO）を提案している。また、このストレージデザインにおいて、既存の様々なチェックポイント戦略を評価するためのモデルを構築し、高信頼型大規模スーパーコンピュータを構築するためのストレージデザインや、各々のチェックポイント戦略のトレードオフについて、定量データに基づいて議論している。

第5章「FMI: Fault Tolerant Messaging Interface」では、三つ目の提案である Fault tolerant messaging interface（FMI）について述べられている。この提案では、In-memory diskless checkpointing やスケラブルな障害検知技術を通信用ライブラリに統合することにより、障害発生時においてもアプリケーションやプログラマから透過的なリスタート機構を実現している。特に、ベンチマークを用いた評価では、平均障害間隔が1分という大変短い状況において28%のオーバーヘッドでアプリケーションを実行できることを、実際に Fault injection を行うことにより示している。

第6章「Conclusions and Future Work」では、本研究の総括と貢献を述べると共に、今後の方向性を示している。

以上のように、本論文は大規模スーパーコンピュータのための、チェックポイント・リスタートを中心とした耐障害技術を提案し、またその有効性を示しており、理学的に貢献するところ大である。よって本論文は博士（理学）の学位論文として十分価値があるものと認める。