

วีซีซีพี: ระบบซอฟต์แวร์โคออดิเนตเช็คพอยต์แบบโปร่งใสสำหรับคลัสเตอร์เสมือน

VCCP: A Transparent, Coordinated Checkpointing Software System
for Virtual Clusters

ณัฐพล สาระกุล , กษิตศ ชาญเขียว

บทคัดย่อ

เทคโนโลยีเวอร์ชวลไลเซชันให้ความสามารถสร้างเครื่องเสมือนที่มีฟังก์ชันสร้างเช็คพอยต์ (checkpoint) และไม่เกรซัน (migration) ได้โดยโปร่งใสต่อแกสโอเอสและโปรแกรมประยุกต์ แต่อย่างไรก็ตามมันไม่ใช่เรื่องเล็กๆ ในการเตรียมฟังก์ชันสร้างเช็คพอยต์ในระดับความโปร่งใสแบบเดียวกันสำหรับโปรแกรมประยุกต์แบบขนาน (parallel application)ที่กำลังทำงานอยู่ด้านบนสุดของเครื่องเสมือน ดังนั้นวิทยานิพนธ์ฉบับนี้จึงเสนอสถาปัตยกรรมคลัสเตอร์เสมือน (Virtual Cluster Architecture: VCA) และระบบเช็คพอยต์คลัสเตอร์เสมือน (Virtual Cluster CheckPointing: VCCP) ซึ่งเป็นซอฟต์แวร์เลเยอร์หนึ่งสำหรับทำเช็คพอยต์และรีโคเวอรี่คลัสเตอร์เสมือนในระดับไฮเปอร์ไวเซอร์ (hypervisor) โดยไม่ต้องดัดแปลงแกสโอเอสและโปรแกรมประยุกต์ที่อยู่ภายใน เพราะว่าระบบวีซีซีพีใช้โปรโตคอลการเช็คพอยต์และรีโคเวอรี่แบบใหม่ผู้วิจัยจึงได้ทำการวิเคราะห์ความถูกต้องเพื่อแสดงให้เห็นว่าวีซีซีพีไม่ได้ทำให้เฟรมข้อมูลสูญหายหรือผิดพลาดดับ ไปจนถึงการกล่าวถึงข้อจำกัดบางประการของมันและแสดงให้เห็นว่าวีซีซีพีควรสามารถทำงานได้กับโปรแกรมประยุกต์ที่ทำงานได้กับการเช็คพอยต์ดั้งเดิมที่ทำในระดับโพเรซ ผู้วิจัยทำการพัฒนาระบบวีซีซีพีขึ้นบนคีมูละและทำการวัดสมรรถนะของระบบด้วยเอ็นพีบี (NPB: Nass Parrallel Benchmark) จากการทดลองแสดงให้เห็นว่าการรวมระบบวีซีซีพีเข้าไปบนคีมูละนั้นสร้างโอเวอร์เฮด (overhead) ต่ำ แต่อย่างไรก็ตามในบางแอสเคอร์เนลทำให้ผู้วิจัยค้นพบว่าเครือข่ายวีดีโอายังต้องได้รับการปรับให้มีประสิทธิภาพมากยิ่งขึ้น ดังนั้นผู้วิจัยจึงวางแผนการพัฒนาเครือข่ายเสมือนแบบใหม่ขึ้นในอนาคต รวมไปถึงยังค้นพบว่าสมรรถนะของการเช็คพอยต์ขึ้นอยู่กับเวลาที่เครื่องเสมือนใช้ไปกับการบันทึกสแตท เวลาในการชะล้างช่องสื่อสาร และเวลาที่เครื่องเสมือนต้องรอคอยให้เครื่องอื่น ๆ ทำเช็คพอยต์ของตัวเองให้เสร็จ สุดท้ายผู้วิจัยอยากกล่าวว่างานวิจัยนี้มีเป้าหมายเพื่อนำเสนอประเด็นเรื่องของการเช็คพอยต์และรีโคเวอรี่แบบโปร่งใส สำหรับเรื่องของสมรรถนะจะเป็นงานพัฒนาในอนาคตต่อไป

Abstract

Virtualization technology provides ability for creating a virtual machine that can checkpoint and migrate its internal state transparently to guest operating systems and applications. However, it is nontrivial to provide checkpointing mechanisms with the same level of transparency for parallel applications running on top of a cluster of virtual machines. This paper proposes a layering Virtual Cluster Architecture (VCA) and the Virtual Cluster CheckPointing (VCCP) system, a thin-layer software system that performs checkpointing and recovery of a virtual cluster at the hypervisor-level without modifications to guest OS and applications. The VCCP system implements novel hypervisor-base Coordinated Checkpointing and Recovery protocols. Our correctness analysis shows that VCCP does not cause message lost or reordering. We also outline its limitations and show that VCCP should be applicable to applications that work with traditional process-level checkpointing. We have also implemented the VCCP prototype base on QEMU system and measured its performance on the NAS parallel benchmark. The experiments show the incorporation of the VCCP protocols to QEMU only generates low execution overheads. However, for some NAS kernel, we found that the execution times of some NAS kernels do not decrease as the number of VM increases due to the inefficiency of VDE switch network. We plan to implement a new network infrastructure in our future work. We also find that checkpointing performance depends on the time VM spend to save local VM state, flush time, and the time VM wait one another for local checkpointing to finish. Finally, we would like to highlight here that the emphasis of this work, at this initial stage, is to promote the transparency issue, while leaving performance to the future implementations of VCCP.