Resumen

Una de las principales preocupaciones de los centros de datos actuales es maximizar la utilización de los servidores. En cada servidor o procesador multinúcleo se ejecutan simultáneamente varias aplicaciones, lo que aumenta la eficiencia de los recursos gracias a la compartición de los mismos. Sin embargo, el rendimiento y la equidad dependen en gran medida de la proporción de recursos que recibe cada aplicación, lo que provoca que su tiempo de ejecución sea imprevisible. El creciente número de núcleos (y de aplicaciones ejecutándose al mismo tiempo) con cada nueva generación de procesadores hace que crezca la preocupación por el efecto causado por las interferencias en los recursos compartidos.

La compartición de recursos se ha abordado típicamente en la computación de alto rendimiento (HPC); sin embargo, debido a la creciente importancia de la computación en la nube, los problemas que suelen tratarse en HPC se han trasladado a este ámbito. Esta tesis se centra en mitigar la interferencia en los recursos compartidos cuando diferentes aplicaciones se consolidan en un mismo procesador desde dos perspectivas: HPC y computación en la nube.

En el contexto de HPC, para reducir la interferencia causada por la ejecución concurrente de múltiples aplicaciones, en esta tesis se proponen políticas de gestión para dos de los recursos más críticos del sistema: la caché de último nivel (LLC) y los núcleos del procesador. La LLC desempeña un papel clave en las prestaciones del sistema con los procesadores multinúcleo actuales ya que reducen considerablemente el número de accesos de alta latencia a la memoria principal. Se proponen estrategias de particionado de la LLC tanto para cachés inclusivas como no inclusivas, ya que ambos diseños están presentes en la actualidad en los procesadores para servidores. Para los dos esquemas de caché, se identifican y detectan eficientemente nuevos comportamientos problemáticos en lo que se refiere a la LLC. Esto permite asignar un mayor espacio de caché a aquellas aplicaciones que hacen un uso eficiente del mismo. En cuanto a los núcleos del procesador, muchas aplicaciones paralelas, como las aplicaciones de grafos, no escalan bien a medida que se incrementa el número de hilos/procesos debido a problemas *hardware* y *software*. Sin embargo, el planificador de Linux, que aplica una estrategia de tiempo compartido, no ofrece buenas prestaciones cuando se ejecutan aplicaciones de grafo, ya que, a diferencia de otras aplicaciones científicas, procesan grandes cantidades de datos. Para maximizar la utilización del sistema, esta tesis propone ejecutar múltiples aplicaciones de grafo simultáneamente en el mismo procesador, asignando el número óptimo de núcleos

a cada una. Adaptando el número de hilos creados por las aplicaciones en tiempo de ejecución, es posible cambiar el número de núcleos asignados para satisfacer los requisitos de las aplicaciones de manera dinámica.

Para estudiar el impacto de los recursos compartidos del sistema en la computación en la nube, esta tesis aborda tres grandes retos: la compleja infraestructura de los sistemas en la nube, las características de las aplicaciones que se ejecutan en la nube y el impacto de la interferencia entre máquinas virtuales (MV) en el rendimiento de éstas. En primer lugar, esta tesis presenta la plataforma experimental desarrollada con los principales componentes de un sistema en la nube (*hardware* y *software*) para realizar estudios representativos del rendimiento de la nube. En segundo lugar, se presenta un amplio estudio de caracterización sobre un conjunto de aplicaciones de latencia crítica representativas, ya que muchas cargas de trabajo importantes en la nube deben cumplir estrictos requisitos de calidad de servicio (QoS) para brindar una experiencia de usuario satisfactoria. El objetivo de los estudios es identificar las cuestiones que los proveedores de servicios en la nube deben tener en cuenta para mejorar el rendimiento y la utilización de los recursos. Por último, se realiza una propuesta que, de manera dinámica, permite detectar y estimar de forma precisa la interferencia entre MV en escenarios en los que se ejecutan múltiples MV con aplicaciones de latencia crítica. El enfoque se basa en métricas que pueden monitorizarse fácilmente en la nube pública, ya que las MV deben tratarse como "cajas negras". Toda la investigación descrita se lleva a cabo respetando las restricciones y cumpliendo los requisitos para ser aplicable en entornos de producción en la nube pública.

En resumen, esta tesis aborda la contención en los principales recursos compartidos del sistema en el contexto de la consolidación de servidores, tanto en entornos de altas prestaciones como en entornos de nube. Los resultados experimentales muestran importantes ganancias de prestaciones sobre el planificador del sistema operativo Linux al reducir las interferencias en los recursos compartidos. En los procesadores con LLC inclusiva, el tiempo de ejecución (TT) se reduce en más de un 40 %, mientras que se mantiene (e incluso mejora) el IPC en más de un 3 %. En los sistemas con LLC no inclusiva, la equidad y el TT mejoran en un 44 % y un 24 %, respectivamente, al mismo tiempo que se obtiene una mejora del rendimiento de hasta en un 3,5 %. Al distribuir los núcleos del procesador de forma eficiente, se alcanza una equidad casi perfecta (de media un 94 %), y el TT puede reducirse hasta un 80 %. En entornos de computación en la nube, la degradación del rendimiento debido a la contención en los recursos compartidos puede estimarse con un error de un 5 % en la predicción global. Todas las propuestas presentadas en esta tesis han sido diseñadas para ser aplicadas en procesadores de servidores comerciales sin requerir ninguna información previa. Las decisiones se toman dinámicamente en tiempo de ejecución utilizando los datos recogidos de los contadores de prestaciones *hardware*.