



UNIVERSIDAD COMPLUTENSE
MADRID

AVISO DE CONFERENCIA

Modelos de programación para arquitecturas heterogéneas on-chip

Prof. Rafael Asenjo Plaza
Depto. de Arquitectura de Computadores, Univ. de Málaga

Facultad de Informática
Sala de Grados • 24 de octubre de 2014 • 12:30
Entrada libre hasta completar el aforo

Resumen

En esta presentación abordaremos el abanico de alternativas disponibles actualmente para maximizar la utilización de los recursos en chips acelerados por GPU, prestando especial atención al consumo de energía. Se discutirán, desde los aspectos HW que hacen especiales estas arquitecturas heterogéneas "on-chip" respecto de las basadas en GPUs discretas, hasta las distintas aproximaciones SW para programarlas y evaluar el rendimiento obtenido en términos de prestaciones por unidad de energía. En cuanto al soporte HW, resumiremos el estado del arte al respecto de las GPUs integradas, los mecanismos de comunicación entre GPU y CPU, además de los más recientes avances para proporcionar coherencia de datos entre los dos tipos de dispositivos. También se mencionarán las alternativas existentes para evaluar en tiempo real el consumo de energía de estos procesadores. En esta línea, se utilizarán como ejemplos arquitecturas heterogéneas on-chip como Intel Haswell, AMD Kaveri, Samsung Exynos o Qualcomm Snapdragon. Por otro lado, en cuanto al soporte de programación para estas arquitecturas, incluiremos primero un repaso de modelos tradicionalmente usados con GPUs discretas, como StarPU, OmpSs o XKaapi y su posible aplicación con GPUs integradas, para posteriormente abordar las más recientes propuestas de la mano de Qualcomm e Intel, como son MARE y Concord, respectivamente. Así mismo, introduciremos las estrategias al efecto que hemos propuesto en nuestro grupo de investigación. Más precisamente describiremos brevemente las extensiones de TBB que hemos desarrollado para las plantillas de `parallel_for` y `pipeline` con objeto de conseguir un uso energéticamente eficiente de arquitecturas heterogéneas on-chip mediante planificación dinámica, particionado adaptativo de datos y mapeado de la carga computacional en los dos tipos de dispositivos.

Sobre Rafael Asenjo

Rafael Asenjo Plaza es Ingeniero de Telecomunicación (1993) y Doctor Ingeniero de Telecomunicación (1997) por la Univ. de Málaga. Actualmente es Profesor Titular de Universidad en el Departamento de Arquitectura de Computadores (2001) y lidera un equipo ("Parallel Programming Models and Compilers") formado por 8 jóvenes doctores que trabajan en los campos de programación para arquitecturas de altas prestaciones y aumento de la productividad en la paralelización de códigos irregulares. Sirve como General Chair del PPOPP'16 y ha sido miembro del comité organizador de HPCA'04, ParCo'05, PPOPP'11 y HPCS'14, así como miembro del comité de programa de PME'A'09 y '10, SBAC-PAD'12, IPDPS'13 e IPDPS'14. Sus principales intereses están relacionados con los modelos de programación paralela, paralelización de códigos irregulares y arquitecturas heterogéneas y multiprocesador. Colaboró en el compilador IBM XL-UPC en 2008 y ha contribuido al desarrollo del runtime del compilador del lenguaje Chapel de Cray desde el 2011.