NATSA: A Near-Data Processing Accelerator for Time Series Analysis

Ivan Fernandez[§] Ricardo Quislant[§] Christina Giannoula[†]
Mohammed Alser[‡] Juan Gómez-Luna[‡] Eladio Gutiérrez[§]
Oscar Plata[§] Onur Mutlu[‡]
[§]*University of Malaga* [†]*National Technical University of Athens* [‡]*ETH Zürich*

Abstract— **El análisis de series temporales es una técnica clave para extraer y predecir sucesos en ámbitos tan diversos como la epidemiología, la genómica, la neurociencia, las ciencias medioambientales o la economía, entre otros. El perfil de matriz o** *(matrix profile)*, **es uno de los algoritmos más avanzados para realizar análisis de series temporales, y permite calcular la subsecuencia más similar para una subsecuencia de consulta dada dentro de una serie temporal cortada. El perfil de matriz tiene una baja intensidad aritmética, pero suele trabajar con grandes cantidades de datos de series temporales. En los sistemas informáticos actuales, estos datos deben moverse entre las unidades de memoria externas y las unidades de cálculo internas para realizar el perfil matricial. Esto provoca un importante cuello de botella en el rendimiento, ya que el movimiento de datos es extremadamente costoso tanto en tiempo de ejecución como en energía.**

En este trabajo, presentamos *NATSA*, **la primera solución para el análisis de series temporales en procesado cercano a memoria. La idea clave es explotar las modernas memorias de gran ancho de banda (HBM) apiladas en 3D, para permitir un cálculo especializado eficiente y rápido cerca de la memoria, que es donde residen los datos de las series temporales. NATSA ofrece tres ventajas fundamentales 1) calcula rápidamente el perfil de la matriz para una amplia gama de aplicaciones construyendo unidades de procesamiento aritmético de coma flotante especializadas y energéticamente eficientes cerca de la HBM, 2) mejora la eficiencia energética y el tiempo de ejecución reduciendo la necesidad de movimiento de datos a través de buses lentos y de alto consumo energético entre las unidades de cálculo y las unidades de memoria, y 3) permite el análisis de datos de series temporales a escala explotando el acceso a memoria de baja latencia, alto ancho de banda y energéticamente eficiente que proporciona la HBM. Nuestra evaluación experimental muestra que NATSA mejora el rendimiento hasta 14,2× (9,9× de media) y reduce la energía hasta 27,2× (19,4× de media), con respecto a la implementación multinúcleo de última generación. NATSA también mejora el rendimiento en 6,3× y reduce la energía en 10,2× con respecto a una plataforma NDP de uso general con 64 núcleos en orden.**