



III – LIDI

Instituto de Investigación en Informática LIDI

Lineas de Investigación en Paralelismo en el III-LIDI

Ing. Armando E. De Giusti

Director del Instituto de Investigación en Informática LIDI

Investigador Principal CONICET

Dr. Adrián Pousa.

Dra. Victoria Sanz.

Investigadores del III-LIDI.

Profesores de la Facultad de Informática UNLP



- Una visión general del III-LIDI. Proyectos.
- Proyecto relacionado con Paralelismo.
- Temas conceptuales.
- Relación con el Postgrado en el área de Paralelismo.
- Análisis por temas de I + D +I en Paralelismo, vistos como 3 Subproyectos.
- Un caso de estudio: *"Comparison of Hardware and Software Implementations of AES on Shared-Memory Architectures"*

Recursos Humanos



- 25** Investigadores Formados.
- 23** Investigadores en Formación / Becarios Doctorales
Tesis de Postgrado.
- 12** Becarios Alumnos de iniciación a la Investigación /
Tesis de Grado.
- 4** Profesionales de Apoyo.
- 22** Investigadores formados de otras Universidades
participan en los proyectos de I+D+I del Instituto.

Este núcleo constituye uno de los Institutos de Investigación
en Informática más relevantes en Argentina.

Proyectos Investigación



Metodologías,
Técnicas y
Herramientas de
Ingeniería de
Software en
Escenarios
Híbridos. Mejora de
Proceso.



Computación de
Alto Desempeño:
Arquitecturas,
Algoritmos, Métricas
de rendimiento y
Aplicaciones en
HPC, Big Data,
Robótica, Señales y
Tiempo Real



Sistemas
Inteligentes.
Aplicaciones en
Reconocimiento de
Patrones, Minería
de Datos y Big Data

Proyectos Investigación



Metodologías, Técnicas y Herramientas de Ingeniería de Software en Escenarios Híbridos. Mejora de Proceso.

Ingeniería de Software para escenarios híbridos.

Gobernanza Digital. Mejora de Procesos.

Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos.



**Computación de Alto Desempeño:
Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real**

Arquitecturas Multiprocesador en Computación de Alto Desempeño.

Algoritmos paralelos en computación de alto desempeño. Fundamentos, construcción y evaluación de aplicaciones.

Procesamiento para problemas de Tiempo Real / Robótica. Cloud, Fog & Edge Computing.

Procesamiento de señales y reconocimiento de patrones.



Sistemas Inteligentes. Aplicaciones en Reconocimiento de Patrones, Minería de Datos y Big Data

Diseño e implementación de modelos predictivos y descriptivos.

Minería de Textos. Caracterización y categorización de documentos.

Técnicas para minado de grandes volúmenes de datos.

Aprendizaje automático aplicado al reconocimiento de patrones.



Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real

Arquitecturas Multiprocesador en Computación de Alto Desempeño.

Algoritmos paralelos en computación de alto desempeño. Fundamentos, construcción y evaluación de aplicaciones.

Procesamiento para problemas de Tiempo Real / Robótica. Cloud, Fog & Edge Computing.

Procesamiento de señales y reconocimiento de patrones.

- Caracterización de arquitecturas multiprocesador para CAD, analizando técnicas para el desarrollo de código.
- Caracterización de arquitecturas multiprocesador para CAD, analizando técnicas para el desarrollo de código.
- Multicore, many-core. GPU, FPGA, MIC, procesadores de bajo costo (tipo Raspberry). Arquitecturas híbridas.
- Planificación de procesos en asimétricos.
- Rendimiento prestacional y energético.
- Tolerancia a fallos.
- Modelos de E/S para sistemas de HPC.

Proyectos Investigación



Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real

Arquitecturas Multiprocesador en Computación de Alto Desempeño.

Algoritmos paralelos en computación de alto desempeño. Fundamentos, construcción y evaluación de aplicaciones.

Procesamiento para problemas de Tiempo Real / Robótica. Cloud, Fog & Edge Computing

Procesamiento de señales y reconocimiento de patrones.

- Paralelización de soluciones en multiprocesadores.
- Lenguajes y técnicas para computación paralela. Costo de desarrollo de las soluciones.
- Métricas de evaluación de rendimiento (prestacional y energético) sobre diferentes plataformas
- Aplicaciones científicas, búsquedas, simulaciones, bioinformática, big data, criptografía.
- Ambientes para la enseñanza de computación paralela.

Proyectos Investigación



Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real

Arquitecturas Multiprocesador en Computación de Alto Desempeño.

Algoritmos paralelos en computación de alto desempeño. Fundamentos, construcción y evaluación de aplicaciones.

Procesamiento para problemas de Tiempo Real / Robótica. Cloud, Fog & Edge Computing

Procesamiento de señales y reconocimiento de patrones.

- Software para sistemas de tiempo real.
- Sistemas operativos de tiempo real.
- Sistemas embebidos. Microcontroladores. Sensores.
- Robótica – IoT.
- Edge Computing. Fog Computing.
- Cloud Robotics.

Proyectos Investigación



Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real

Arquitecturas Multiprocesador en Computación de Alto Desempeño.

Algoritmos paralelos en computación de alto desempeño. Fundamentos, construcción y evaluación de aplicaciones.

Procesamiento para problemas de Tiempo Real / Robótica

Procesamiento de señales y reconocimiento de patrones.

Algoritmos paralelos aplicados a:

- Reconocimiento de patrones.
- Clasificación supervisada y no supervisada.
- Procesamiento y análisis de imágenes.
- Sistemas de posicionamiento, navegación y localización.

Relación Postgrado



En el Postgrado de la Facultad de Informática se imparte una Maestría y una Especialización directamente relacionadas con el proyecto.

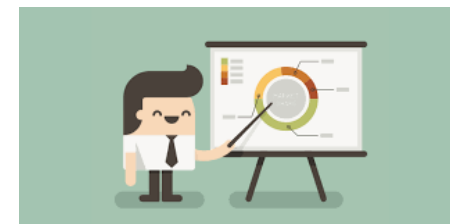
Asimismo se dirige el Doctorado en Ciencias Informáticas de la Facultad.

Doctorado en
Ciencias Informáticas

Maestría en Cómputo
de Altas Prestaciones

Especialización en
Cómputo de Altas
Prestaciones

El III-LIDI participa de diferentes eventos científicos (congresos, jornadas, charlas, conferencias) de carácter nacional e internacional relacionados a sus temas de investigación



Desde el año 2013 **anualmente organiza** las Jornadas de Cloud Computing- Big Data & Emergent Topics



Proyectos Investigación



**Computación de Alto
Desempeño:
Arquitecturas,
Algoritmos, Métricas
de rendimiento y
Aplicaciones en HPC,
Big Data, Robótica,
Señales y Tiempo Real**

Arquitecturas Multiprocesador en
Computación de Alto Desempeño.

Algoritmos paralelos en
computación de alto desempeño.
Fundamentos, construcción y
evaluación de aplicaciones.

Procesamiento para problemas de
Tiempo Real / Robótica. Cloud, Fog
& Edge Computing.

Procesamiento de señales y
reconocimiento de patrones.

Vamos a analizar el Proyecto en tres grandes
Subproyectos.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Objetivo General

- El eje de esta línea de I/D lo constituye el estudio de las arquitecturas multiprocesador que integran sistemas distribuidos y paralelos.

Incluye como temas centrales:

- Arquitecturas many-core, híbridas y asimétricas.
- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas.
- Estudio y optimización de código heredado.
- Desarrollo y evaluación de estrategias de resiliencia.
- Modelado y simulación de E/S en HPC.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Tipos de Arquitecturas y Problemas estudiados GPU

- GPUs y Cluster de GPUs Las GPUs son el acelerador dominante en la comunidad de HPC hoy en día por su alto rendimiento y bajo costo de adquisición. En la actualidad, tanto NVIDIA como AMD trabajan especialmente en mejorar la eficiencia energética de sus placas y disminuir el alto costo de programación.
- La combinación de GPUs con otras plataformas paralelas como clusters y multicores, brindan un vasto conjunto de posibilidades de investigación en arquitecturas híbridas.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Tipos de Arquitecturas y Problemas estudiados

A partir de diferentes combinaciones como son:

- Máquinas multicore con más de una GPU, que combinan herramientas de programación paralela como OpenMP/CUDA o Pthread/CUDA.
- Cluster de máquinas multicore cada una con una o más placas de GPU, lo que permite combinar OpenMP/MPI/CUDA o Pthread/MPI/CUDA.

Los desafíos que se plantean son múltiples, sobre todo en lo referido a distribución de datos y procesos en tales arquitecturas híbridas a fin de optimizar el rendimiento de las soluciones.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Tipos de Arquitecturas y Problemas estudiados MIC

- MIC En forma reciente Intel brinda una alternativa a partir de la arquitectura MIC (Many Integrated Core Architecture).
- Esta arquitectura permite utilizar métodos y herramientas estándar de programación con altas prestaciones. (a diferencia de las GPU)
- De esta forma, se reduce el entrenamiento y se focaliza en el problema a resolver.
- Recientemente, Intel ha lanzado Knights Landing (KNL), la segunda generación de Xeon Phi. A diferencia de sus predecesores que operaban como coprocesador a través del puerto PCI, los procesadores KNL pueden operar en forma autónoma. Además, integran las nuevas extensiones vectoriales AVX-512 y tecnología de memoria 3D, entre otras características avanzadas.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Tipos de Arquitecturas y Problemas estudiados FPGA

- FPGAs Una FPGA (Field Programmable Gate Array) es una clase de acelerador basado en circuitos integrados reconfigurables.
- La capacidad de adaptar sus instrucciones de acuerdo con la aplicación objetivo le permite incrementar la productividad de un sistema y mejorar el rendimiento energético para ciertos tipos de aplicaciones.
- Tradicionalmente fueron utilizadas para el procesamiento digital de señales. Sin embargo, en los últimos años, existen dos tendencias claras para extender su uso a otros dominios.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Tipos de Arquitecturas y Problemas estudiados FPGAs

- En primer lugar, el establecimiento de alianzas estratégicas entre fabricantes de procesadores y de FPGAs para integrar estos dispositivos en arquitecturas híbridas (Intel con Altera; IBM con Xilinx) ha extendido su empleo.
- En segundo lugar, el desarrollo de nuevas herramientas de programación para FPGAs empleando estándares familiares para HPC, con las cuales se espera reducir los tradicionales tiempos y costos de programación .
- Por último, la incorporación de FPGAs a los servicios de Cloud abre oportunidades para el uso de esta clase de aceleradores..



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Tipos de Arquitecturas y Problemas estudiados TPU

- Las unidades de procesamiento tensorial (TPU) son una clase de circuitos integrados personalizados específicos de aplicaciones (ASIC) desarrolladas por Google con el propósito de acelerar las cargas de trabajo de aprendizaje automático que requieren las aplicaciones desarrolladas en su framework TensorFlow.
- Su uso provee una alternativa a otras arquitecturas ya conocidas como CPUs, GPUs y MICs. En ese sentido, interesa analizar las tasas de aceleración y eficiencia energética provistas por esta nueva arquitectura, en comparación con el resto.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Arquitecturas y Problemas estudiados. Eficiencia Energética

- La mejora de la eficiencia energética es una de las principales preocupaciones en la informática actual, principalmente a partir de las plataformas con gran cantidad de procesadores.
- Muchos esfuerzos están orientados a tratar la eficiencia energética como eje de I/D, como métrica de evaluación, y también a la necesidad de metodologías para medirla.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Arquitecturas y Problemas estudiados. Eficiencia Energética

- Análisis de metodologías y herramientas para medir y optimizar el consumo energético.
- Estudio de técnicas para reducir el consumo energético en aplicaciones de HPC de acuerdo con las arquitecturas utilizadas.
- Evaluación de eficiencia energética de diferentes algoritmos y plataformas paralelas. Optimización de la eficiencia energética.
- Empleo de los contadores hardware para definir estrategias de programación que lleven a reducir el consumo, manteniendo a su vez el rendimiento en valores aceptables.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Arquitecturas y Problemas estudiados. Código Heredado

- La mayoría de los programas de cálculo y simulación numérica empleados hoy en día fueron desarrolladas cuando las arquitecturas paralelas no existían.
- Es por lo que este conjunto de aplicaciones presenta la oportunidad de desarrollar técnicas y herramientas que permitan optimizar el código, tanto desde el punto de vista computacional como desde la ingeniería de software, paralelizando los algoritmos originales.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Arquitecturas y Problemas estudiados. Resiliencia

- En la actualidad, lograr sistemas resilientes resulta un verdadero desafío considerando el creciente número de componentes, la cercanía a los límites físicos en las tecnologías de fabricación y la complejidad incremental del software.
- La corrección de las aplicaciones y la eficiencia en su ejecución se torna más importante en HPC debido a los extensos tiempos de ejecución.
- En ese sentido, resulta interesante/importante desarrollar estrategias de detección y recuperación de fallos, especialmente a través de librerías de software.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Arquitecturas y Problemas estudiados. E/S Paralela.

- A pesar de los avances tecnológicos, las operaciones de E/S en los centros de supercómputo siguen siendo un cuello de botella para determinadas aplicaciones HPC.
- El rendimiento de un sistema dependerá de la carga de trabajo (patrones de E/S de las aplicaciones) y de su configuración (hardware y software).
- Contar con herramientas que permitan modelar y predecir el comportamiento de este tipo de aplicaciones en HPC resulta fundamental para mejorar su rendimiento



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Otras Arquitecturas estudiadas.

- Dispositivos de bajo costo con capacidades para cómputo paralelo
- En la actualidad se comercializan placas de bajo costo como Raspberry PI u Odroid que poseen múltiples núcleos simples. Asimismo, existen diversos dispositivos móviles con capacidades similares.
- Es de interés estudiar como explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Líneas de Investigación y Desarrollo

- Arquitecturas many-core (procesadores MIC, GPU y TPU) y FPGA.
- Análisis de este tipo de máquinas y de técnicas para desarrollar código optimizado. - Arquitecturas híbridas (diferentes combinaciones de clusters, multicores, manycores y FPGAs).
- Diseño de algoritmos paralelos sobre modelos de arquitectura.
- Técnicas de resiliencia.
- Exploración de lenguajes y modelos de programación para HPC.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Líneas de Investigación y Desarrollo

- Consumo energético en las diferentes arquitecturas paralelas, en particular en relación con los algoritmos paralelos y la configuración de la arquitectura. Análisis de metodologías y herramientas de medición.
- Modelado y estimación del consumo de potencia de arquitecturas HPC.
- Análisis y desarrollo de modelos e implementación de simuladores de la pila de software de E/S en HPC.



SP1 Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Algunos resultados obtenidos. Investigación experimental.

- Desarrollo y evaluación de algoritmos paralelos sobre diferentes arquitecturas multiprocesador. Analizar rendimiento, eficiencia energética y costo de programación.
- Analizar capacidades de lenguajes no convencionales para procesamiento paralelo.
- Desarrollo de nuevos planificadores de tareas para multicores asimétricos.
- Desarrollar técnicas de tolerancia a fallas que permitan aumentar la resiliencia de sistemas paralelos y distribuidos.
- Desarrollar técnicas de modelado y simulación de E/S en HPC.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Algunos resultados recientes obtenidos (publicados)

- Se analizaron técnicas de optimización para código heredado relacionado a cómputo numérico.
- Se optimizaron soluciones paralelas para el procesamiento de pattern matching en arquitectura Xeon Phi KNL.
- Se optimizó una solución al problema de los N cuerpos computacionales con atracción gravitacional (versión directa) sobre diferentes arquitecturas Intel basadas en instrucciones AVX-512.
- Se diseñó la herramienta SEDAR y evaluó su desempeño para detección y recuperación de fallos transitorios.
- Se desarrolló y validó un modelo estadístico para consumo de potencia en placas RPi de diferentes generaciones.
- Se realizaron y analizaron modificaciones al framework TensorFlow para permitir la maleabilidad de hilos.



SP1: Arquitecturas Multiprocesador: Software de Base, Modelos y Aplicaciones.

Algunos resultados obtenidos (publicados)

- Se realizó una comparación entre las arquitecturas Intel Xeon Phi KNL y NVIDIA Pascal estudiando el problema de caminos mínimos en grafos
- Se analizó el rendimiento y consumo energético requerido para simulaciones basadas en agentes sobre arquitecturas de GPUs
- Se avanzó en el modelado y simulación inicial del sistema de archivos paralelos PVFS2 y se implementó por medio del paradigma de agentes en un simulador.
- Se desarrolló una plataforma distribuida, escalable, híbrida y redundante que permite aprovechar recursos de cómputo basados en arquitecturas ARM a través de un servicio de Cloud.
- Se desarrolló un modelo para predecir el consumo energético de un sistema y diferentes estrategias para reducirlo cuando ocurre una falla.
- Se compararon los enfoques de desarrollo HLS y HDL en FPGA en el ámbito del procesamiento de imágenes.

Proyectos Investigación



**Computación de Alto
Desempeño:
Arquitecturas,
Algoritmos, Métricas
de rendimiento y
Aplicaciones en HPC,
Big Data, Robótica,
Señales y Tiempo Real**

Arquitecturas Multiprocesador en
Computación de Alto Desempeño.

Algoritmos paralelos en
computación de alto desempeño.
Fundamentos, construcción y
evaluación de aplicaciones.

Procesamiento para problemas de
Tiempo Real / Robótica. Cloud, Fog
& Edge Computing.

Procesamiento de señales y
reconocimiento de patrones.

Vamos a analizar el Proyecto en tres grandes
Subproyectos.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Objetivos

- El eje de esta línea de I/D lo constituye el estudio de la integración de arquitecturas distribuidas que van desde el nivel de sensores (Edge Computing) a la capa de procesamiento en la nube (Cloud Computing), con una posible capa intermedia que se conoce como Fog Computing.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Objetivos y Temas centrales

- Distribución de la capacidad de procesamiento en cada nivel.
- Análisis de performance en las comunicaciones, según el grado de distribución del procesamiento.
- Administración de recursos en cada nivel.
- Migración de “inteligencia” al nivel “Edge” para reducir consumo y comunicaciones.
- Desarrollo y evaluación de aplicaciones que integran niveles de procesamiento.
- Análisis de eficiencia en tiempo, consumo y comunicaciones.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Modelos de procesamiento estudiados

Cloud Computing

- En este proyecto el modelo de procesamiento distribuido integra desde sensores a la nube.
- Cloud Computing se reserva para el procesamiento centralizado de algoritmos complejos, con gran volumen de datos.
- Parte de estos datos podrán haber sido “preprocesados” en las capas “Edge” o “Fog”.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Modelos de procesamiento estudiados

Fog Computing

- El modelo de Fog Computing surge como respuesta al crecimiento de los desarrollos relacionados con el Internet de las Cosas. Estos desarrollos requieren procesamiento en la nube, pero presentan características que hacen complejo su armado haciendo uso exclusivo de tecnologías de Cloud Computing.
- Fog Computing consiste entonces en una plataforma intermedia que provee procesamiento, almacenamiento y servicios de comunicación en red entre los dispositivos “Edge” que adquieren los datos en el modelo IoT.
- Las características de Fog Computing son las de una capa intermedia orientada a diferentes aplicaciones.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Modelos de procesamiento estudiados

Fog Computing. Aplicaciones.

- Aplicaciones de tiempo real distribuidas.
- Gran número de nodos, que puede escalar dinámicamente.
- Heterogeneidad de los nodos.
- Movilidad.
- Red predominante de comunicaciones inalámbricas.
- Baja latencia y conocimiento de la ubicación.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Modelos de procesamiento estudiados

Edge Computing

- El modelo de “Edge” computing está impuesto por el crecimiento exponencial del número de dispositivos sensores con inteligencia local disponibles. Internet “de las cosas” (IoT) crece y hoy hablamos de 50.000 millones de dispositivos conectados a Internet, con un tráfico del orden de 800 Zbytes.
- Se requiere entonces capacidad de procesamiento cerca de los sensores, integración de datos locales y geográficamente ubicados y posibilidad de pre procesar los mismos y enviarlos con un menor overhead de comunicaciones a las capas superiores (Fog o Cloud).



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Modelos de procesamiento estudiados

Edge Computing. Ventajas.

- Respuestas automáticas al usuario, en tiempo real, utilizando su capacidad local.
- Disminución del tráfico de datos.
- Mayor seguridad por el procesamiento local.
- Disminución del consumo energético, por las características de los componentes.
- Mejora de la eficiencia global para un sistema distribuido, débilmente acoplado.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Líneas de Investigación y Desarrollo

- Arquitecturas aplicables en Edge Computing. Modelos de referencia.
- Arquitecturas aplicables en Fog Computing. Modelos de referencia.
- Administración de recursos y datos en Edge y Fog computing. Integración de los mismos.
- Seguridad en los datos.
- Manejo de aplicaciones móviles y relación entre las capas Edge y Fog.
- Vinculación de las capas Edge y Fog con el Cloud. Servicios requeridos.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Líneas de Investigación y Desarrollo

- Distribución óptima de tareas entre capas, para aplicaciones complejas.
- Escalabilidad en aplicaciones Edge-Fog-Cloud.
- Métricas de eficiencia considerando tiempo de respuesta / costo comunicaciones / consumo energético.
- Aplicaciones: robots / drones / vehículos autónomos.
- Algoritmos colaborativos en tiempo real integrando las tres capas.
- Aplicaciones: sistemas inteligentes distribuidos para reducir el consumo energético. Migración de “inteligencia” en sistemas embebidos.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Resultados. Investigación experimental.

- Análisis comparativo de tecnologías empleadas en Edge/Fog Computing.
- Análisis comparativo de los servicios para arquitecturas Edge y Fog.
- Estudio de protocolos de comunicación con el Cloud, desde el nivel de Edge y/o Fog.
- Análisis y despliegue de plataformas de sensores/módulos de sensado.
- Análisis de performance de comunicaciones, en función del escalado de la arquitectura de sensores.
- Estudios de consumo energético en aplicaciones de Edge/Fog/Cloud.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Resultados. Investigación experimental.

- Estudio comparativo de la distribución de carga de procesamiento en aplicaciones distribuidas (en particular de tiempo real).
- Integración de aplicaciones móviles y relación entre las capas Edge y Fog.
- Aplicaciones: robots / drones colaborativos trabajando en comunicación con una capa Fog y con el Cloud.
- Aplicaciones: Sistema de tableros inteligentes para optimización de consumo energético en edificios distribuidos.



SP2: Arquitecturas Edge – Fog y Cloud en Procesamiento Distribuido.

Algunos Resultados obtenidos (publicados)

- Se han analizado plataformas orientadas a la gestión de recursos en Edge y Fog Computing.
- Se han estudiado protocolos y tráfico de comunicaciones en aplicaciones distribuidas en tiempo real que requieren interactuar con el Cloud.
- Se desarrolló un modelo de tablero inteligente para reducción del consumo y su integración con la capa Fog y comunicación con la nube.
- Se está estudiando el tema de consumo distribuido, en el caso de los tableros inteligentes, de los robots distribuidos y del trabajo con drones.

Proyectos Investigación



**Computación de Alto
Desempeño:
Arquitecturas,
Algoritmos, Métricas
de rendimiento y
Aplicaciones en HPC,
Big Data, Robótica,
Señales y Tiempo Real**

Arquitecturas Multiprocesador en
Computación de Alto Desempeño.

Algoritmos paralelos en
computación de alto desempeño.
Fundamentos, construcción y
evaluación de aplicaciones.

Procesamiento para problemas de
Tiempo Real / Robótica. Cloud, Fog
& Edge Computing.

Procesamiento de señales y
reconocimiento de patrones.

Vamos a analizar el Proyecto en tres grandes
Subproyectos.



SP3: Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Objetivo

- El eje central de la línea de I/D es investigar en temas de cómputo paralelo y distribuido de alto desempeño: fundamentos y construcción, evaluación y optimización de aplicaciones en arquitecturas multiprocesador.
- Se aplican los conceptos en problemas numéricos y no numéricos de cómputo intensivo y/o sobre grandes volúmenes de datos con el fin de obtener soluciones de alto rendimiento.
- También incluye la construcción de ambientes para la enseñanza de la programación concurrente y paralela



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Líneas de Investigación y Desarrollo

- Lenguajes, modelos y paradigmas de programación paralela (puros e híbridos a distintos niveles).
- Caracterizar diferentes modelos de arquitecturas paralelas y sus comunicaciones.
- Asignación de procesos a procesadores optimizando el balance de la carga de procesamiento.
- Métricas de evaluación de complejidad y rendimiento: speedup, eficiencia, escalabilidad, consumo energético, costo de programación.
- Analizar y desarrollar ambientes para la enseñanza de programación concurrente y paralela.
- Aplicar los conceptos a problemas de diferente índole.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados).

- Se experimentó la paralelización en arquitecturas híbridas, estudiando el impacto del mapeo de datos y procesos, así como de los lenguajes y librerías empleadas.
- Aceleración de aplicaciones con cómputo colaborativo CPU-GPU. Se propuso un modelo híbrido para estructurar código a ser ejecutado sobre un sistema heterogéneo con múltiples cores y 1 GPU. A partir de este modelo se desarrolló un algoritmo paralelo de pattern matching para sistemas heterogéneos CPU-GPU. Los resultados revelaron que este algoritmo supera en rendimiento a trabajos previos, desarrollados para sistemas multicore y GPUs, para datos de tamaño considerable.
- Se presentó una solución al problema de pattern matching que aprovecha toda la potencia computacional de los procesadores Intel Xeon Phi KNL 7230 mediante el uso de SIMD y paralelismo de hilos. Se mostró que el algoritmo propuesto alcanza aceleraciones significativas.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados).

- Algoritmos paralelos para alineamiento de secuencias biológicas. Esta operación consiste en comparar dos o más secuencias biológicas, como pueden ser las de ADN o las de proteínas, y resulta fundamental en investigaciones de la bioinformática y la biología molecular.
- El algoritmo de Smith-Waterman es considerado el método de alineamiento más preciso. Este algoritmo resulta costoso debido a su complejidad computacional cuadrática mientras que la situación se agrava aún más a causa del crecimiento exponencial de datos biológicos en los últimos años.
- El reciente surgimiento de aceleradores en HPC (GPU, Xeon Phi, FPGA, entre otros) da la oportunidad de acelerar los alineamientos sobre hardware comúnmente disponible a un costo accesible.
- A futuro, interesa explorar las fortalezas y debilidades del uso de nuevas tecnologías de software para arquitecturas paralelas como, por ejemplo, oneAPI.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados).

- Algoritmos paralelos de cálculo de los caminos mínimos, un problema “clásico” de la teoría de grafos que tiene aplicación en el dominio de las comunicaciones, del ruteo de tráfico, de la bioinformática, entre otros. En particular el algoritmo de FloydWarshall (FW) permite computar la distancia mínima entre todos los pares de un grafo. Además de poseer una alta demanda de ancho de banda, FW resulta costoso computacionalmente al ser $O(n^3)$.
- Se desarrollaron implementaciones optimizadas para dos arquitecturas HPC recientes -como son Intel Xeon Phi KNL y NVIDIA Pascal- y se analizó comparativamente su rendimiento y eficiencia energética (teórica) en diferentes escenarios.
- Como trabajo futuro, interesa incorporar otros modelos de las arquitecturas elegidas al estudio para robustecer el análisis.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados).

- Simulaciones utilizando modelos basados en agentes. El objetivo de esta línea son las simulaciones basadas en agentes sobre infraestructuras de altas prestaciones con modelos más cercanos a la realidad que ayuden a la toma de decisiones a científicos de otras disciplinas, y no expertos en el área de la informática, (biología, ecología, física, etc.).
- Se han desarrollado algoritmos y simuladores como soluciones HPC que sean eficientes y escalables y con el menor consumo posible de energía.
- Dado que son simulaciones con un gran número de ejecuciones por escenario, no solo es necesario que la solución tenga buenas prestaciones sino que sea posible predecir, mediante un modelo energético, la energía necesaria para llevar a cabo diferentes escenarios de simulación.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados).

- Simulación de enfermedades infecciosas transmitida por vectores utilizando modelos basados en agentes. El Dengue, Zika y Chikungunya, son las enfermedades reemergentes de mayor preocupación a nivel mundial.
- Se implementó un modelo basado en agentes en GPU para la evaluación de la reproducción del vector *Aedes aegypti*, orientado a la toma de decisiones.
- El modelo ha sido validado con datos de Santo Tomé, Corrientes, dando excelentes resultados.
- Se realiza un estudio y análisis de performance en otras arquitecturas paralelas basadas en GPU. Además, se evalúa el consumo energético para determinar las relaciones entre escalabilidad, consumo y eficiencia.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados).

- Aplicaciones en Big Data utilizando HPC. Existen distintos tipos de herramientas para trabajar estos problemas y la mayoría de ellas se basan en el paradigma de programación llamado MapReduce, propuesto por Google en los inicios de esta gran explosión de datos. La más ampliamente utilizada en la actualidad, es conocida como Apache Spark y, entre sus más importantes características se encuentran los mecanismos eficientes de tolerancia a fallos y el uso intensivo de memoria RAM que emplea, siendo una de las más rápidas.
- Además, dispone de una amplia variedad de librerías que posibilitan utilizar algoritmos de Machine Learning, procesamiento en Streaming, etc.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados).

- En esta línea de trabajo, se están desarrollando nuevas técnicas para el preprocesamiento de datos para problemas de clasificación en Big Data, utilizando el framework Apache Spark sobre Cómputo de Altas Prestaciones.
- Como parte del trabajo realizado se lleva a cabo el estudio de la calidad de los datos actualmente disponibles de manera pública en el ambiente científico relacionados a este tipo de problemas de clasificación.
- El estudio se basa en el análisis de las características intrínsecas de los conjuntos de datos disponibles donde se analizan aspectos tales como redundancia, outliers, áreas sobrepuestas del problema, porcentaje de desbalance entre clases, entre otros factores que, estando presentes, pueden generar una degradación en el rendimiento del clasificador a utilizar.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados)

- Simulación de N cuerpos computacionales con atracción gravitacional. Su propósito es aproximar en forma numérica la evolución de un sistema de cuerpos en el que cada uno interactúa con todos los restantes.
- Si bien existen diferentes métodos para procesar la simulación de los N cuerpos, en todos los casos se requiere alta demanda computacional. Se ha utilizado este problema para analizar diferentes aspectos y/o arquitecturas como: cluster de GPUs homogéneas; análisis de diferentes distribuciones de trabajo en cluster de GPUs heterogéneas; desarrollo de modelos de estimación de tiempo de ejecución y de energía para GPU, cluster de GPU y multiGPU.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados)

- Simulación de N cuerpos computacionales con atracción gravitacional. Además, se estudió la paralelización de la versión directa de esta simulación sobre diferentes arquitecturas Intel considerando procesadores de última generación Xeon y Xeon Phi.
- Se focalizó en cómo diferentes técnicas de optimización logran mejorar el rendimiento final y en la comparación de rendimiento y eficiencia energética entre ambas arquitecturas.
- A futuro, interesa extender el estudio incorporando arquitecturas de GPUs.



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados).

- Problemas de optimización de simulación de sistemas dinámicos complejos mediante heurísticas. La búsqueda de un conjunto de parámetros de entrada que optimicen el funcionamiento de un simulador de un sistema físico es un proceso de alto costo computacional que puede considerarse intratable y requiere de heurísticas que permitan disminuir el tiempo de ejecución.
- Se ha desarrollado una metodología eficiente que se ha aplicado en la simulación de inundaciones de ríos de llanura con las que se llevaron adelante las experiencias (Río Paraná en particular). Se utilizaron clusters de procesadores como plataforma (que puede ser colaborativa).



Algoritmos Paralelos y Evaluación de Rendimiento en plataformas de HPC

Algunos resultados obtenidos. (publicados)

- Ambientes para la enseñanza de concurrencia. Se desarrolló el entorno multirrobot CMRE para la enseñanza de programación concurrente y paralela a partir de cursos iniciales en carreras de Informática.
- Incluye un entorno visual que representa una ciudad en la que pueden definirse varios robots que interactúan. Combina aspectos de memoria compartida y distribuida mediante instrucciones específicas y el concepto de pasaje de mensajes a través de primitivas de envío y recepción.
- Se tratan los conceptos de heterogeneidad y consumo energético en las primitivas de los robots. También se lo combina con robots físicos como demo.

Proyectos Investigación



**Computación de Alto
Desempeño:
Arquitecturas,
Algoritmos, Métricas
de rendimiento y
Aplicaciones en HPC,
Big Data, Robótica,
Señales y Tiempo Real**

Arquitecturas Multiprocesador en
Computación de Alto Desempeño.

Algoritmos paralelos en
computación de alto desempeño.
Fundamentos, construcción y
evaluación de aplicaciones.

Procesamiento para problemas de
Tiempo Real / Robótica. Cloud, Fog
& Edge Computing.

Procesamiento de señales y
reconocimiento de patrones.

- Referencia a las publicaciones.
<http://weblidi.info.unlp.edu.ar/wp/#>
- Formación de Recursos Humanos.
- Líneas de Cooperación nacional e internacional

Proyectos Investigación



**Computación de Alto
Desempeño:
Arquitecturas,
Algoritmos, Métricas
de rendimiento y
Aplicaciones en HPC,
Big Data, Robótica,
Señales y Tiempo Real**

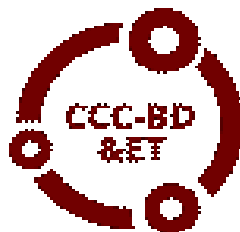
Arquitecturas Multiprocesador en
Computación de Alto Desempeño.

Algoritmos paralelos en
computación de alto desempeño.
Fundamentos, construcción y
evaluación de aplicaciones.

Procesamiento para problemas de
Tiempo Real / Robótica. Cloud, Fog
& Edge Computing.

Procesamiento de señales y
reconocimiento de patrones.

Consorcio de I+D+I en Cloud Computing, Big Data & Emerging Topics



CONSORCIO DE I+D+I
CLOUD COMPUTING,
BIG DATA &
EMERGING TOPICS

- Una visión general del III-LIDI. Proyectos.
- Proyecto relacionado con Paralelismo.
- Temas conceptuales.
- Relación con el Postgrado en el área de Paralelismo.
- Análisis por temas de I + D +I en Paralelismo, vistos como 3 Subproyectos.



PREGUNTAS ??