# Escuela Internacional de Informática

### Procesamiento y Modelado de Aplicaciones de Datos Masivos

**Docentes Responsables:**

Dra. Marcela Printista (UNSL)

Dra. Verónica Gil Costa (UNSL)

**DESTINATARIOS**

Dirigido a estudiantes e investigadores interesados en herramientas de modelado y programación para el procesamiento de datos masivos.

**REQUISITOS**

Nociones avanzadas de lenguaje C.

**OBJETIVOS**

El objetivo del curso es caracterizar los problemas de procesamiento sobre datos masivos en paralelo desde el punto de vista de una arquitectura de multi-core y de los paradigmas de programación aptos para ella. Se abordarán conceptos relacionados con las herramientas de desarrollo, criterios de diseño, selección del hardware según la aplicación final y se realizarán prácticas con plataformas basadas en el modelo de threads, Map-Reduce y Stream Processing (como Hadoop y Storm) con el objetivo de brindar a los asistentes los elementos necesarios para formar su criterio al momento de realizar actividades aplicadas y de investigación bajo esta temática.

**CONTENIDOS**

Tecnologías Multicore. Procesamiento de datos masivos: Caso de estudio la Web. Recuperación de la información en la Web. Motores de Búsqueda. Real Time Big Data. El nuevo desafío: Capacity Planning. Modelo de computación paralela: BSP y Map-Reduce.

Modelo de programación orientado a actores (stream processing): Sistemas para procesamiento de flujos de eventos. Sistemas para procesamiento de eventos complejos. Colaboración entre sistemas de procesamiento de eventos. Aplicaciones colaborativas clusters-móviles

**SISTEMA DE EVALUACION**

Examen teórico-práctico