

ST0263 - Tópicos Especiales en Telemática

Profesor: Edwin Montoya Múnera – emontoya@eafit.edu.co

2018

Inspira Crea Transforma

UNIVERSIDAD
EAFIT[®]

Descripción del curso

- Infraestructura TI, DevOps y Desarrollo de Software para Sistemas Escalables, Intensivos en Procesamiento e Intensivos en Datos.

Sistemas Escalables

- Sistema en Producción => DevOps
- Escalabilidad -> crecimiento de la demanda de recursos en los sistemas como el incremento en el número de usuarios, mayores cantidades de almacenamiento, mejores tiempos de respuesta, y sistemas más seguros entre otros.
- La escalabilidad está soportada por las buenas prácticas, patrones de arquitectura, tácticas y herramientas tanto a nivel de aplicación como a nivel de infraestructura TI.

Sistemas Escalables

- DevOps

- Todas las prácticas, tácticas, tecnologías que permiten ir del **Desarrollo** en equipos ágiles hasta la entrega continua en producción de software probado en ambientes de **Producción** ya sea On-premise (datacenter propio) o Cloud (nube en diferentes opciones: IaaS, PaaS, SaaS, DaaS, etc)

Computación de alto desempeño

- El Alto desempeño en computación hace referencia a la necesidad de contar con infraestructura tecnológica para el **procesamiento paralelo y distribuido** que ejecutan aplicaciones y programas que resuelven problemas complejos, los cuales requieren la articulación de **muchos procesadores y computadores** para lograr en tiempos razonables, la ejecución de aplicaciones que ambientes tradicionales demorarían semanas, meses o inclusive años.
- Estos sistemas de alto desempeño comúnmente son llamados Computación de Alto Desempeño o **High-Performance Computing (HPC)**.

Big Data

- infraestructura tecnológica que permitan no solo almacenar este volumen tan alto de datos, sino que permita procesarlos en tiempos razonables.
- Es por ello, que surge todo un campo de Ciencia de los Datos e Ingeniería de los Datos, éste último más relacionado con los Ingenieros de Sistemas.

Objetivo General

- Desarrollar Sistemas y Aplicaciones Escalables, bajo DevOps, intensivos en datos e intensivos en procesamiento que consideran diferentes atributos de calidad con particularidades propias de los tipos de sistemas a considerar (transaccionales, intensivos en datos e incentivos en procesamiento).

Objetivos de aprendizaje

- **Comprender** la fundamentación de los sistemas distribuidos, arquitectura, características, modelos y tipos y los diferentes tipos de sistemas (transaccionales, intensivos en datos e intensivos en computación).
- **Desarrollar** sistemas y aplicaciones transaccionales escalables con atributos de calidad relacionados a Disponibilidad, Rendimiento y Seguridad.
- **Desarrollar** sistemas y aplicaciones bajo el enfoque DevOps
- **Desarrollar** sistemas y aplicaciones intensivos en datos que utiliza una infraestructura TI adecuada para Big Data.
- **Desarrollar** sistemas y aplicaciones intensivos en computación (HPC) que utiliza una infraestructura adecuada para HPC ya sea en Cluster o en Cloud Computing.

Competencias Generales

- **Instrumentales:**
 - Capacidad de análisis y síntesis
 - Capacidad de organización y planificación
 - Comunicación oral y escrita
 - Conocimiento de una lengua extranjera
 - Resolución de problemas
 - Toma de decisiones
- **Personales:**
 - Trabajo en equipo
 - Habilidades en las relaciones interpersonales
 - Razonamiento Crítico
- **Sistemáticas:**
 - Aprendizaje autónomo
- **Otras:**
 - Capacidad de aplicar los conocimientos teóricos en la práctica
 - Experiencia previa

Competencias Específicas

- **Especifica** requisitos no funcionales relacionados con los atributos de calidad de Disponibilidad, Rendimiento y Seguridad en sistemas informáticos escalables.
- **Diseña** aplicaciones y sistemas computacionales que tienen en cuenta atributos de calidad como Disponibilidad, Rendimiento y Seguridad.
- **Selecciona** la arquitectura de sistema adecuada.
- **Aplica y usa** mejores prácticas y tecnologías para DevOps
- **Desarrolla** sistemas y aplicaciones computacionales que consideran diferentes atributos de calidad teniendo en cuenta la disponibilidad de tecnologías de hardware, software, redes.
- **Gestiona** la tecnología subyacente de Big Data
- **Adquiere** una metodología básica de desarrollo de proyectos big data (ciclo de vida)
- **Desarrolla** aplicaciones en las diferentes etapas de un proyecto big data específicamente
- **Gestionar** la infraestructura TI para ejecutar aplicaciones intensivas en computación.
- **Desarrolla** aplicaciones paralelas y distribuidas intensivas en cómputo utilizando una metodología específica de Algoritmos Paralelos como PCAM.

Contenidos

- **UNIDAD No 1 - SISTEMAS ESCALABLES (6 semanas, 18 horas)**
- **UNIDAD No 2 - COMPUTACIÓN Y BIG DATA (6 semanas, 18 horas)**
- **UNIDAD No 3 – COMPUTACIÓN DE ALTO DESEMPEÑO (4 semanas, 12 horas)**

UNIDAD No 1 –

SISTEMAS ESCALABLES

- **Competencias específicas:**

- **Especifica** requisitos no funcionales relacionados con los atributos de calidad de Disponibilidad, Rendimiento y Seguridad en sistemas informáticos escalables.
- **Diseña** aplicaciones y sistemas computacionales que tienen en cuenta atributos de calidad como Disponibilidad, Rendimiento y Seguridad.
- **Selecciona** la arquitectura de sistema adecuada
- **Utiliza** acercamiento y tecnologías DevOps
- **Desarrolla** sistemas y aplicaciones computacionales que consideran diferentes atributos de calidad teniendo en cuenta la disponibilidad de tecnologías de hardware, software, redes.

UNIDAD No 1 –

SISTEMAS ESCALABLES

- Temas:
 - Ambientes de Despliegue (on-premise y cloud)
 - DevOps
 - Escalabilidad
 - Disponibilidad
 - Rendimiento
 - Seguridad

UNIDAD No 2 – COMPUTACIÓN Y BIG DATA

Competencias Específicas:

- Identifica, selecciona y gestiona la infraestructura TI para ejecutar aplicaciones Big Data.
- Desarrolla aplicaciones distribuidas intensivas en datos basadas en Map/Reduce y Spark en diferentes lenguajes de desarrollo.

UNIDAD No 2 – COMPUTACIÓN Y BIG DATA

Resultados de aprendizaje:

- Gestiona la tecnología subyacente de Big Data
- Adquiere una metodología básica de desarrollo de proyectos big data (ciclo de vida)
- Desarrolla aplicaciones en las diferentes etapas de un proyecto big data específicamente:
 - Fase de ETL
 - Fase de Procesamiento con técnicas de analítica de datos sencillas, más énfasis en los modelos de procesamiento y programación Big Data.

UNIDAD No 2 – COMPUTACIÓN Y BIG DATA

Temas:

- * Infraestructura TI para Big Data
- * Procesamiento Distribuido para Big Data
- * Introducción a Analítica de datos
- * Proyectos de Big Data

UNIDAD No 3 – COMPUTACIÓN DE ALTO DESEMPEÑO

- **Competencias específicas:**

- Gestiona una infraestructura de TI para ejecutar aplicaciones intensivas en Computación
- Desarrolla aplicaciones paralelas y distribuidas intensivas en computo

UNIDAD No 3 – COMPUTACIÓN DE ALTO DESEMPEÑO

- **Resultados de aprendizaje:**

- Desarrolla sistemas y aplicaciones de alto desempeño con las principales tecnologías para clústeres de alto desempeño, tanto desde la infraestructura como desde las herramientas de desarrollo de software.

UNIDAD No 3 – COMPUTACIÓN DE ALTO DESEMPEÑO

Temas:

* Infraestructura para HPC

- Introducción al HPC
- Arquitecturas de Supercómputo
- Centros de Datos para HPC
- Hardware para HPC
- Software para HPC

* Desarrollo de software paralelo

- Lenguajes de Programación y Compiladores para HPC
- Metodología PCAM
- MPI
- Coprocesamiento: GPU / CUDA y MIC

Evaluación:

- 10% MOOC Intro a DevOps
- 10% MOOC sobre Big Data
- 10% MOOC sobre HPC
- 10% Proyecto 1 – Desarrollo sw y DevOps
- 15% Proyecto 2 – Sistemas escalables y DevOps
- 10% Proyecto 3 – Big Data
- 10% Proyecto 4 - HPC
- 15% Examen Final
- 10% Seguimiento

Evaluación:

- **UNIDAD DIDACTICA No 1 - SISTEMAS ESCALABLES – 44%**
 - 10% MOOC Intro a DevOps
 - 10% Proyecto 1
 - 15% Proyecto 2
 - 5% contribución al Examen Final
 - 4% contribución al Seguimiento
- **UNIDAD DIDACTICA No 2 - COMPUTACIÓN Y BIG DATA – 28%**
 - 10% Desempeño comprensivo (Realización de MOOC)
 - 10% Proyecto 3
 - 5% contribución al Examen Final
 - 3% contribución al Seguimiento
- **UNIDAD DIDACTICA No 2 – COMPUTACIÓN DE ALTO DESEMPEÑO – 28%**
 - 10% Realización de un MOOC sobre HDP
 - 10% Proyecto 4
 - 5% contribución al Examen Final
 - 3% contribución al Seguimiento