

El BSC lanza la version 3.0 de PyCOMPSs/COMPSs

El Barcelona Supercomputing Center ofrece a la comunidad HPC un conjunto de herramientas que ayudan a los desarrolladores a programar y ejecutar sus aplicaciones de forma eficiente en infraestructuras computacionales distribuidas.



En particular, PyCOMPSs ofrece la interface para aplicaciones Python que es utilizada como herramienta para la convergencia de la supercomputación, inteligencia artificial y big data.

Esta versión de COMPSs incluye varias funcionalidades nuevas, tales como soporte transparente a checkpointing, creación automática de información de data provenance y soporte para aplicaciones MPMD MPI.

Para las aplicaciones en Python, PyCOMPSs incluye una compilación Mypy del código y se ha desarrollado una nueva CLI para unificar las ejecuciones de aplicaciones en distintos entornos.

El grupo [Workflows and Distributed Computing](#) del Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC) está orgulloso de anunciar el lanzamiento de la versión 3.0 (nombre en clave Lavender), del entorno de programación [COMPSs](#).

Esta versión de COMPSs, disponible a partir de hoy, actualiza el resultado del trabajo del equipo en los últimos años en la provisión de un conjunto de herramientas que ayudan a los desarrolladores a programar y ejecutar sus aplicaciones de forma eficiente en infraestructuras computacionales distribuidas, como clústers, nubes y clústeres gestionados con contenedores. COMPSs es un modelo de programación basado en tareas capaz de mejorar notablemente el rendimiento de aplicaciones a gran escala mediante la paralelización de manera automática de su ejecución.

COMPSs ha estado disponible en los últimos años por los usuarios del supercomputador MareNostrum y en la Red Española de Supercomputación (RES) y ha sido usado en varios proyectos de investigación como EUBra-BIGSEA, MUG, EGI, ASCETIC, TANGO, NEXTGenIO, I-BiDaaS, mF2C, CLASS, ExaQUTE, ELASTIC, BioExcel CoE, LANDSUPPORT, EXPERTISE ETN y el proyecto de innovación Edge Twins HPC FET. En estos proyectos se ha utilizado COMPSs para desarrollar casos de uso proporcionados por diferentes comunidades de diversas disciplinas como la biomedicina, la ingeniería, la biodiversidad, la química, la astrofísica, finanzas, telecomunicaciones, procesos de fabricación y ciencias de la tierra. Actualmente también está en extensión y en uso en aplicaciones en los proyectos AI-SPRINT, PerMedCoE y CAELESTIS. También ha sido aplicado a un caso de uso del ChEESE CoE. Una mención especial es el proyecto eFlows4HPC coordinado por el grupo, iniciado en enero de 2021, que tiene como objetivo desarrollar una pila de software para workflows donde uno de los componentes principales es el entorno PyCOMPSs / COMPSs.

La nueva versión incluye nuevas funcionalidades que amplían la tolerancia a fallos de las aplicaciones y sus características de reproducibilidad.

Se ha añadido un soporte transparente de *checkpointing* basado en tareas que permite recuperar ejecuciones fallidas. El sistema propuesto aprovecha el determinismo de las tareas para evitar la reejecución de tareas en caso de fallo, copiando automáticamente su salida a medida que avanza la ejecución. La realización de estas copias comporta una sobrecarga importante en las operaciones de entrada/salida y en la red; el equilibrio óptimo para ese compromiso entre resiliencia y rendimiento depende de cada ejecución y de las preferencias del usuario final. Con este objetivo, el sistema propuesto es el resultado de la combinación de diferentes mecanismos que seleccionan sistemáticamente qué valores de salida es necesario realizar *checkpoint*, y permite la personalización de estas decisiones incorporando mecanismos para definir nuevas políticas. Además de las copias sistemáticas, el sistema también proporciona a los desarrolladores de aplicaciones un método para configurar puntos específicos del código de la aplicación donde realizar el *checkpoint* de la ejecución.

Los entornos COMPSs también se han ampliado con un mecanismo que registra automáticamente la procedencia de los datos de la ejecución de las aplicaciones de las COMPSs (data provenance). Esto permite la reproducibilidad y la replicabilidad de las aplicaciones COMPSs y los algoritmos de dislib. El enfoque se basa en la especificación RO-Crate 1.1, que ofrece una buena integración con las herramientas y marcos existentes. El motor de ejecución del COMPSs se ha modificado para generar RO-Crates. Un sistema registra accesos únicos a los archivos para identificar automáticamente las entradas y salidas de la aplicación. Un postproceso extrae la información necesaria de los registros para generar el RO-Crate.

En el marco del proyecto eFlows4HPC se han realizado otros desarrollos para permitir la convergencia entre Computación de Alto Rendimiento (CAR), Inteligencia Artificial (IA) y análisis de datos. Se proporciona un nuevo soporte para las aplicaciones MPMD MPI como tareas, así como soporte para el epílogo y el prólogo de las tareas, y un soporte genérico para descripciones reutilizables de ejecuciones de software externo dentro de una tarea COMPSs. Para unificar la ejecución de la aplicación en distintos entornos informáticos, se ha diseñado e implementado una nueva interfaz de línea de comandos (CLI).

Otras mejoras ofrecidas son la compilación Mypy del enlace Python para acelerar las ejecuciones de Python, la integración con DLB DROM para mejorar la afinidad en las tareas OpenMP y el soporte RISC-V de 64 bits.

Además, la versión 3.0 de COMPSs viene dotada con otras funcionalidades menores, extensiones y resolución de errores.

COMPSs ha tenido alrededor de **1000 descargas** durante el año pasado y es utilizado por unos **20 grupos** en aplicaciones reales. Recientemente, COMPSs ha atraído el interés de áreas como la ingeniería, el reconocimiento de imágenes, la genómica y sismología, áreas en las que se han realizado cursos específicos y acciones de divulgación.

Los paquetes y la lista completa de funciones están disponibles en la página de [descargas](#). En esta página también se puede encontrar una imagen de Docker que permite probar las funcionalidades de COMPSs a través de un tutorial paso a paso que guía al usuario en el desarrollo y ejecución de un conjunto de aplicaciones de ejemplo.

Además del software, hay un conjunto de guías de usuario y administrador, y artículos publicados en conferencias y revistas relevantes.

- Para más información, visite nuestra página web: <http://www.bsc.es/compss>

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 7 Sep 2024 - 14:10): <https://www.bsc.es/es/noticias/noticias-del-bsc/el-bsc-lanza-la-version-30-de-pycompsscompss>