

Opleiding : Python HPC supercomputer

Praktijkcursus - 5d - 35u00 - Ref. PYC

Prijs : 2610 € V.B.



4,6 / 5

Python est devenu en quelques années le langage de programmation privilégié de toutes les disciplines scientifiques. Bien qu'il soit interprété, ses bibliothèques scientifiques sont particulièrement performantes car écrites dans des langages compilés, comme C/Cython et très bien parallélisées. Aujourd'hui la lenteur du langage n'est plus un frein et il fonctionne sur les plus puissants des supercalculateurs de la planète. Nous vous proposons d'apprendre les concepts de la programmation parallèle appliquée au HPC au travers des meilleures bibliothèques Python utilisables sur ces environnements.



Pedagogische doelstellingen

Aan het einde van de training is de deelnemer in staat om:

- ✓ Een goed begrip hebben van supercomputerconcepten en programmeren
- ✓ Meer informatie over Python-bibliotheken die geschikt zijn voor HPC-rekenen
- ✓ Ontwikkel algoritmen op een supercomputer met behulp van de MPI4Py, Dask, Xarray, Dask+Scikit-Learn en PyTorch bibliotheken...
- ✓ Workflows uitvoeren met Prefect
- ✓ Grote gegevens visualiseren met DataShader

Doelgroep

Ingenieurs, ontwikkelaars, onderzoekers, datawetenschappers, data-analisten en iedereen met een sterke behoefte aan rekencapaciteit met Python.

Voorafgaande vereisten

Python taalvaardigheden, kennis van numpy en pandas bibliotheken.

Praktische modaliteiten

Leer methodes

Het praktische werk wordt uitgevoerd op een supercomputer (type Exaion).

DEELNEMERS

Ingenieurs, ontwikkelaars, onderzoekers, datawetenschappers, data-analisten en iedereen met een sterke behoefte aan rekencapaciteit met Python.

VOORAFGAANDE VEREISTEN

Python taalvaardigheden, kennis van numpy en pandas bibliotheken.

VAARDIGHEDEN VAN DE CURSUSLEIDER

De deskundigen die de cursus leiden zijn specialisten op het betreffende vakgebied. Zij werden geselecteerd door onze pedagogische teams zowel om hun vakkennis als hun pedagogische vaardigheden voor elke cursus die zij geven. Zij hebben minstens vijf tot tien jaar ervaring in hun vakgebied en oefenen of oefenden verantwoordelijke bedrijfsfuncties uit.

BEOORDELINGSMODALITEITEN

De cursusleider beoordeelt de pedagogische vooruitgang van de deelnemer gedurende de gehele cursus aan de hand van meerkeuzevragen, praktijksituaties, praktische opdrachten, ... De deelnemer legt ook van tevoren en naderhand een test af ter bevestiging van de verworven kennis.

Opleidingsprogramma

1 Ontdek supercomputers

- Een overzicht van de allereerste supercomputer tot de krachtigste van vandaag.
- Wat is een supercomputer?
- Fundamentele principes en kenmerken: rekencapaciteit, netwerkcapaciteit en opslagcapaciteit.
- Les différents classements : Top500, Green500, io500.
- Comment se programme un supercalculateur : les ordonnances/gestionnaires de ressources : SLURM, PBS, ...
- Présentation du supercalculateur Exaion sur lequel nous travaillerons.

Praktisch werk

Aan de slag met de Exaion supercomputer: aansluiten, een virtuele omgeving installeren en de eerste jobs uitvoeren met Slurm.

2 MPI programmeren

- Snelle inleiding tot de basisprincipes van parallel rekenen met Python: multithreading, multiprocessing, GIL.
- MPI concepten en de verschillende beschikbare bibliotheken.
- De verschillende primitieven: zenden/ontvangen, verstrooien/verzamenen, broadcast/reduce, procespools, enz.

Praktisch werk

Implementatie van verschillende problemen met de belangrijkste primitieven: een batch afbeeldingen verwerken, de decimalen van PI berekenen, enz.

3 MPI programmeren, toepassingen

- Presentatie van voorbeelden van MPI-toepassingen.

Praktisch werk

Voortzetting van praktisch werk met de belangrijkste primitieven.

4 Dask en zijn ecosysteem

- Aan de slag met dask: basisconcepten, dask array en dataframe.
- De andere onderdelen van dask: vertraagd, futures en zakken.
- Dask sur HPC : Scheduler et workers, créer un cluster dask : Cluster MPI/Slurm...
- Panorama des différentes librairies de l'écosystème Dask.
- NetCDF-bestanden verwerken met XArray.

Praktisch werk

Analyse van tijd- en klimaatreeksen, classificaties en regressies met Dask+Scikit-Learn, cartografische datavisualisatie.

5 Dask en big data

- Grote gegevens visualiseren met DataShader en Xarray.
- Créer des pipelines/workflows avec Prefect.
- DaskML : déployer vos algorithmes de machine learning sur HPC.

Praktisch werk

Verder praktisch werk met gegevensanalyse en visualisatie.

PEDAGOGISCHE EN TECHNISCHE MIDDELEN

- De gebruikte pedagogische middelen en cursusmethoden zijn voornamelijk: audiovisuele hulpmiddelen, documentatie en cursusmateriaal, praktische oefeningen en correcties van de oefeningen voor praktijkstages, casestudies of reële voorbeelden voor de seminars.
- Na afloop van de stages of seminars verstrekt ORSYS de deelnemers een evaluatievragenlijst over de cursus die vervolgens door onze pedagogische teams wordt geanalyseerd.
- Na afloop van de cursus wordt een presentielijst per halve dag verstrekt, evenals een verklaring van de afronding van de cursus indien de stagiair alle sessies heeft bijgewoond.

TOEGANGSMODALITEITEN EN TERMIJNEN

De inschrijving dient 24 uur voor aanvang van de cursus plaatsgevonden te hebben.

TOEGANKELIJKHEID VOOR MINDERVALIDEN

Is voor u speciale toegankelijkheid vereist? Neem contact op met mevr. FOSSE, contactpersoon voor mindervaliden, via het adres psh-accueil@ORSYS.fr om uw verzoek en de haalbaarheid daarvan zo goed mogelijk te bestuderen.

6 GPU-rekenen

- GPU-computerconcepten met Python: hardware, bibliotheken.
- Fonctionnement d'un GPU.
- Dask sur GPU : Créer un cluster CUDA.
- Machine-learning avec PyTorchLightning et RapidsAI.

Praktisch werk

Basisimplementatie met de PyCuda en Cupy bibliotheken. Dataframe manipulatie met Dask-CUDF. Machine learning toegepast op meerdere compute nodes en meerdere GPU's.