

Opleiding : Python, parallel programmeren en gedistribueerd computergebruik

Praktijkcursus - 4d - 28u00 - Ref. PYP

Prijs : 2100 € V.B.



4,4 / 5

Het succes van Python voor wetenschappelijke toepassingen (Data science, Big Data, Machine Learning, enz.) vereist steeds meer rekencapaciteit. In deze cursus maak je kennis met het parallelle/distributed computing paradigma, van de basisconcepten tot de meest geavanceerde technieken en bibliotheken in het Python ecosysteem.



Pedagogische doelstellingen

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

- ✓ De concepten van parallel programmeren verwerven
- ✓ De delen van een programma identificeren die geparallelliseerd kunnen worden
- ✓ Een duidelijke visie hebben op het parallelle computing ecosysteem voor Python
- ✓ Ontwikkelen van parallelle toepassingen (asynchroon programmeren, multithreading, multiprocessing, distributed computing)
- ✓ Weten hoe je berekeningen uitvoert op de GPU's van grafische kaarten
- ✓ Hoe een workflow van taken uitvoeren in de Cloud

Doelgroep

Ontwikkelaars, datawetenschappers, data-analisten, projectmanagers.

Voorafgaande vereisten

Goede kennis van de taal Python en, indien mogelijk, de wetenschappelijke bibliotheken Numpy, Scipy en Pandas.

DEELNEMERS

Ontwikkelaars, datawetenschappers, data-analisten, projectmanagers.

VOORAFGAANDE VEREISTEN

Goede kennis van de taal Python en, indien mogelijk, de wetenschappelijke bibliotheken Numpy, Scipy en Pandas.

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.

Praktische modaliteiten

Leer methodes

70% van de tijd wordt besteed aan het in praktijk brengen van de gepresenteerde concepten en bibliotheken. Het gebruik van Jupyter notebooks en het uitvoeren van code in de Cloud zorgen voor echte interactiviteit.

Opleidingsprogramma

1 Parallellisme en het ecosysteem van Python

- De verschillende vormen van parallellisme en hun architecturen (CPU, GPU, ASIC, FPGA, NUMA, OpenMP, MPI, enz.).
- Beperkingen en limieten.
- Het parallelle rekenecosysteem voor Python.

Praktisch werk

Een programma profileren (cProfile, Kcachegrind en pyprof2calltree). Een C-programma compileren met SIMD-instructies. Numpy correct installeren: hoe een snelheidstoename van 40x te bereiken.

2 De basis: asynchroon programmeren, multithreading en multiprocessing

- Asynchroon programmeren: generatoren en asynchronie.
- Multithreading: gelijktijdige toegang, sloten, enz.
- De grenzen van multithreading in Python.
- Multiprocessing: gedeeld geheugen, procespools, voorwaarden, enz.
- Eerste gedistribueerde computercluster met managers en proxy.

Praktisch werk

Creatie van een enkele gegevensverwerkingsketen met elk model en een gedistribueerd rekencluster tussen de machines van de deelnemers.

3 Gedistribueerd computergebruik : Celery, Dask en PySpark

- Concepten en configuratie.
- Implementatie van elke bibliotheek.

Praktisch werk

Verschillende oefeningen komen aan bod (matrixberekeningen, beeld/tekstverwerking, Bitcoin, Machine Learning, enz.) Gebruik van Zeppelin-notitieblokken.

4 GPU-rekenen

- GPU-architecturen: kernels, geheugen, threads, enz.
- OpenCL- en CUDA-bibliotheken.
- Implementatie van de Scikit-cuda, PyCUDA en Numba bibliotheken.

Praktisch werk

Matrixberekening en beeldverwerking. Machine Learning met de mxnet bibliotheek: Neurale kunst. Just In Time compilatie.

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.

5 Andere parallele programmeerbibliotheken

- Message Passing Interface met MPI4py.
- PyOpenCL: code implementeren met heterogene systemen.
- Joblib: Lichtgewicht pijplijnen.
- Greenlets: naar betere multithreading.
- Pythran : Compileer uw Python-programma's op multicore- en vectorarchitecturen.

Praktisch werk

Basisoefeningen bij elke bibliotheek.

6 Taakworkflows maken

- Primitieven beschikbaar met Celery, Dask en PySpark.
- Creëer en beheer workflows met de Luigi en Airflow bibliotheken.

Praktisch werk

Aanmaken van gegevensverwerkingspijplijnen met elke bibliotheek.

7 Berekeningen uitvoeren in de cloud

- Overzicht van internetdiensten voor de cloud.
- Een cluster beheren met Ansible.

Praktisch werk

Berekeningen uitvoeren in de Cloud.

Data en plaats

KLAS OP AFSTAND

2026 : 16 juni, 15 sep.

PARIS LA DÉFENSE

2026 : 16 juni, 15 sep.