Federico Panichi

Technical Manager · Performance Engineer (HPC & AI)
federico.panichi@studio.unibo.it · +44 7955 193980 - +39 352 04 80 649
LinkedIn: linkedin.com/in/federico-panichi

Profilo professionale

Technical Manager e Performance Engineer con esperienza consolidata nella progettazione, implementazione e ottimizzazione di sistemi computazionali ad alte prestazioni (HPC). Esperinza professinale in compagnie internazionali per 7 anni: AVID, NAG, AMD. Dottorato di ricerca in Fisica con competenze nella fisica computazionale, tecniche di analisi statistica e Machine Learning (ML). Specializzato nel profiling e nel tuning di codici in parallelo (MPI, OpenMP, hybrid), gestione di cluster per simulazioni scientifiche, applicazioni industriali e integrazione di soluzioni AI nei workflow di ricerca universitaria e sviluppo industriale. Certifciato Google Cloud Architect 2021. Capacità dimostrate nella leadership tecnica, mentoring, e nella conduzione di progetti cross-disciplinari in contesti accademici e industriali.

Esperienza in High Performance Computing (HPC)

Esperienza pluriennale nell'implementazione e ottimizzazione di codici in parallelo su CPU via OpenMP, MPI e hybrid (OpenMP+MPI). Gestione e tuning di cluster: PlGrid in Polonia durante il mio dottorato; Marenostrum 5 in Spagna nell'ambito del progretto POP; attualmente lavoro con cluster privati AMD EPYC (Milan, Genoa a Turin). Containerizzazione di applicazioni industriali: Apptainer/Singularity, conda, mamba, pixy, virtualenv. Utilizzo di tecniche di calcolo avanzato e profiling per l'analisi delle performance: uProf, Vtune, perf, HPCToolkit, Extrae, valgrind.

Competenze tecniche principali

- Linguaggi: Python, C/C++, Fortran, Rust, Julia, Bash.
- Parallelismo e architetture: MPI, OpenMP, multithreading, hardware-specific optimizations (NUMA, affinity, hugepages).
- Profiling & Tuning: uProf, valgrind, Extrae, perf, VTune, gprof, likwid, TAU, perfetto debugging e analysis di performance a livello stack (compilatore, runtime, I/O).
- Librerie & Math: cuBLAS, rocBLAS, oneMKL, BLAS, LAPACK, MKL, AOCL, FFTW; ottimizzazione di kernel numerici.
- Sistemi & DevOps: gestione cluster HPC, SLURM, Kubernetes per workload scientifici, containerizzazione (Docker, Apptainer/Singularity).
- ML/AI & LLM: PyTorch, TensorFlow; Esperienza nel utilizzo di agenti IA per automatizzazione del profiling.
- Agentic AI & Tools LLM: utilizzo operativo di Claude CoWork, Claude Code, GitHub copilot con gli ultimi modelli LLM (Claude Opus 4.6 e GPT 5.3); progettazione e deployment di agentic AI per automatizzare interi workflow (intero team Python e Rust).

Progetti rilevanti & contributi

- Progetto PyZSS: definizione e implementazione di pipeline Python ottimizzate per migliorare le performance di applicazioni di ML/HPC su architetture AMD; mentoring di colleghi junior.
- Implementazione di framework per il profiling di PyTorch, NumPy, SciPy e codici HPC (Gromacs, VASP, Quantum-ESPRESSO) e la valutazione delle performance delle toolchain Intel, GNU e AMD.
- Esperienza ed implementazione di modelli RAG, CoVe, e Agentic Loops per supportare processi di ricerca interna, analizzare le performance dei modelli di ML e mitigare il tasso di allucinazioni.
- Come technical manager del team Python ho sviluppato e supervisionato l'ottimizazzione di PyTorch per CPU con miglioramenti fino al +37% su server EPYC e workload reali.

Strumenti LLM, RAG e Automazione

Utilizzo operativo di Claude CoWork per collaborazione e assistenza LLM-driven; progettazione e deployment di agentic AI che automatizzano interi flussi di lavoro: automazione dell'intero team Python (task orchestration, code generation, review automation) e automazione end-to-end del workflow per Rust (CI, linting, testing, release). Esperienza pratica nell'adozione di RAG per arricchire modelli con knowledge-base e nell'uso di NotebookLM per analisi e documentazione interattiva.

Esperienza professionale (selezione)

Technical Manager / Performance Engineer – AMD, Manchester (UK) · Apr 2022 – Presente

- Guida tecnica e responsabilità nel definire, implementare e gestire progetti di ottimizzazione delle performance su architetture server (AMD EPYC), con focus su workload HPC e ML.
- Progettazione e coordinamento di attività di profiling sistematico per identificare di bottlenecks a livello di compilatore, librerie, runtime e I/O.
- Implementazione di patch e proof-of-concept per miglioramenti delle librerie numeriche (AOCL) e Python.
- Introduzione di pipeline containerizzate ed automatizzate con agenti IA per test di regressione delle performance per CI/CD in Jenkins e GitHub actions.

Performance Engineer – NAG (Numerical Algorithms Group), Manchester (UK) · Feb 2020 – Apr 2022

- Tuning di applicazioni scientifiche e industriali: analisi delle prestazioni, rewriting di hotspot e ottimizzazione algoritmica.
- Sviluppo di strumenti di benchmarking e suite di test per valutare impatto di diverse toolchain e flag di compilazione.
- Collaborazione con team di ricerca per integrare soluzioni IA nella pipeline di analisi dei dati e per utilizzare RAG nella documentazione tecnica.

Software Engineer – AVID Technology, Szczecin (PL) · Oct 2018 – Feb 2020

- Sviluppo di componenti ad alte prestazioni per software multimediale (C++/Python); analisi e ottimizzazione dell'uso della memoria e I/O.
- Responsabile tecnico di progetto per l'integrazione di tool di conversione e gestione pipeline, con attenzione alla riproducibilità e scalabilità.

Formazione

PhD in Fisica computazionale — Università di Stettino (Polania).

- Ulteriori corsi: High Performance Computing, Advanced Numerical Methods, Parallel Programming. Machine learning and Statistical Analysis.

Laurea Magistrale in Astrofisica (105/110) — Università di Bologna (Italia)

- Sviluppo di metodi numerici e statistici, analisi dati.

Leadership & gestione tecnica

Coordinamento diretto di team tecnici (fino a 7 persone) e gestione cross-funzionale di team estesi (fino a 30 persone). Pianificazione ed organizzazione con orizzonte a 18 mesi, gestione delle priorità (RACI), definizione di KPI tecnici, mentoring e valutazione delle competenze.

Lingue & Disponibilità

Italiano (madrelingua), Inglese (fluente). Disponibilità: on-site / ibrido / remoto. Preferenza: posizioni in ambito HPC, ricerca e sviluppo ML/AI.