Driving the next generation of Extremely Large Telescopes using Adaptive Optics with GPUs

Damien Gratadour
LESIA, Observatoire de Paris
Université Paris Diderot

ANR grant ANR-12-MONU-0022 of the French Ministry of Research
Outline

- Adaptive Optics
- The European Extremely Large telescope
- Challenges and enabling technologies
- System design and high-performance simulations with GPUs
  - Benefit from CUDA optimized libraries
- Adaptive optics real-time control with GPUs
  - Benefit from GPUdirect RDMA
- Towards a GPU accelerated extremely large telescope
Adaptive Optics

- Observing the Universe through atmospheric turbulence
  - Earth atmosphere = mixture of air at different temperatures (local variations < 1°C) = « air bubbles » with different refractive index
  - Crossing the atmosphere distort the optical wavefront = reduce image quality
  - No matter the telescope diameter : resolution is limited by the atmosphere (equiv. to a 20cm telescope on best astronomical sites, e.g. Hawaii, Chile)

- Compensating dynamically evolving aberrations
  - Using one or several deformable mirrors to « reshape » the wavefront : adaptive optics
  - Applications not limited to astronomy : biomedical imaging, high-power laser focusing (telecom, manufacturing, etc ..), satellite / space debris tracking, military applications, etc ...
Adaptive Optics

**Principle**

- Analyze wavefront disturbances: wavefront sensor (WFS)
- Compute (in real-time) the compensation to apply using a real-time controller (RTC)
- Apply the compensation using a deformable mirror (DM)
- Close-loop system, real-time (1ms latency) operations
- A beam splitter shares the light between the AO system and the science instrument
European Extremely Large Telescope (E-ELT)

- 40m diameter telescope!

- Primary mirror will be made of ~800 segments of 1.4m diameter

- Theoretical resolution in the near-Infrared: 10 milliarcseconds, i.e. $2 \times 10^5$ smaller than the full moon (~30 arcminutes)

- About 1000 $m^2$ of collecting area, i.e. 15 times more sensitive than the largest state of the art professional telescopes currently in operation
Extremely Large!

- 100m dome, 2800 tones structure rotating @ 360°, seismic safe (Chile)
- 1.2 G€ project, first light foreseen in 2022
- Construction led by ESO (European Southern Observatory), international organisation funded by 15 European countries
- Telescope components + science instruments built by european research labs + industrial partners
- Non-European partners: Chile (site) and maybe Brazil (under negotiations to become an ESO funding partner)
E-ELT and adaptive optics

• Past and present: adaptive optics as a facility instrument
  • For current 8-10m telescopes, not using AO does not prevent from using the telescope (but cannot reach diffraction limit, i.e. best resolution)

• Future: the E-ELT, an adaptive telescope
  • Because of its size the E-ELT requires AO for routine operations (compensate at minimum for structure flexures, wind shake, etc ...)
  • Current design includes a 5-mirrors main optical path
    • 4th mirror is a deformable mirror (~5000 actuators)
    • 5th mirror is mounted on a tip-tilt stage for image stabilization
  • Adaptive optics was at the core of the telescope design studies
  • Telescope design includes 4 to 8 high-power lasers for adaptive optics guiding to get 100 % sky coverage
Challenges and enabling technologies

- AO system complexity scales as the square of the telescope diameter
  - x5 in diameter (as compared to 8m VLT) => x25 in complexity!
  - Because E-ELT = adaptive telescope, strong requirements on AO:
    - Reliability: 100% uptime
    - Versatility: 100% sky coverage (Laser guiding)

- ESO E-ELT Instrument roadmap identifies several enabling technologies developments related to AO
  - New WFS concepts, including fast low noise detectors
  - DM with proper diameter, number of actuators and stroke
  - RTC able to handle this amount of computation in « real-time » (> 300 GMACs)

System design and numerical simulations

- Numerical simulations = first step of a scientific instrument design
  - Trade-off studies = iterate and constrain the instrument design w.r.t. scientific requirements
  - Need an efficient numerical simulation platform
AO simulations

- Multiple physics simulation from atmospheric turbulence models to close-loop systems control theory

- Generate turbulence
AO simulations

- Multiple physics simulation from atmospheric turbulence models to close-loop systems control theory
- Generate turbulence
- Wavefront sensor model
AO simulations

- Multiple physics simulation from atmospheric turbulence models to close-loop systems control theory
- Generate turbulence
- Wavefront sensor model
- Integrate real-time controller
- Deformable mirror model
AO simulations

- Multiple physics simulation from atmospheric turbulence models to close-loop systems control theory
- Generate turbulence
- Wavefront sensor model
- Integrate real-time controller
- Deformable mirror model
- Corrected image model
High-performance simulations

• Complex numerical simulations
  • Design, debug, use, visualize: require an interpreted language with interactive display capabilities, and comprehensive user interface

• Compute intensive simulations
  • Require optimized mathematical libraries and efficient implementation
  • Should be scalable to run on large platforms

• AO simulations involve: FFT, raytracing, MVM, RNG
  • Optimized libraries available on GPUs thanks to CUDA toolkit
  • Significant speedup as compared to CPU libs already demonstrated

• Idea: bind an interpreted language with CUDA
  • Challenge: keep the user friendly interface AND the optimized execution time
High-performance simulations

- YoGA : Yorick with GPU Acceleration

Why Yorick ?
- Free : saves us several 10k€ / year in software licence
- C-based, Open-source : expandable
- Nice display features, array manipulations, C syntax, etc ..
- Many AO simulation tools and general libraries already existing, great community contributions

How it works
- Create persistent objects on the GPU and manipulate from within the Yorick interpreter
- Transparent memory management, link to optimized CUDA libs
High-performance simulations

- COMPASS: COMputing Platform for Adaptive optics SystemS
  - An extension of YoGA: specialized C++ classes for the simulation of AO components
  - Funded by the French Ministry of Research (~1M€ grant)
  - 6 partner labs in France gathering most of the French AO community + HPC specialist (MdlS)
High-performance simulations

COMPASS : COMputing Platform for Adaptive optics SystemS
High-performance simulations

- COMPASS: COMputing Platform for Adaptive optics SystemS
- Execution profile (in ms) on a single Tesla M2090
  - x5 to x10 speedup as compared to CPU sequential code
  - Real-time controller component execution speed is compatible with real-time operations in real life: GPUs as an enabling technology for the E-ELT
  - Demonstrated scalability: w.r.t. GPU used and also simulated system scale

<table>
<thead>
<tr>
<th>Telescope diam.</th>
<th>Turbu generation</th>
<th>Raytracing turbu</th>
<th>Raytracing DM</th>
<th>WFS</th>
<th>COG</th>
<th>Control</th>
<th>DM shape computation</th>
<th>Raytracing target</th>
</tr>
</thead>
<tbody>
<tr>
<td>4m</td>
<td>0.107</td>
<td>0.008</td>
<td>0.008</td>
<td>0.138</td>
<td>0.013</td>
<td>0.019</td>
<td>0.137</td>
<td>0.008</td>
</tr>
<tr>
<td>8m</td>
<td>0.192</td>
<td>0.022</td>
<td>0.023</td>
<td>0.459</td>
<td>0.031</td>
<td>0.060</td>
<td>0.562</td>
<td>0.023</td>
</tr>
<tr>
<td>20m</td>
<td>0.550</td>
<td>0.135</td>
<td>0.136</td>
<td>3.07</td>
<td>0.079</td>
<td>0.363</td>
<td>3.22</td>
<td>0.137</td>
</tr>
<tr>
<td>30m</td>
<td>0.927</td>
<td>0.299</td>
<td>0.300</td>
<td>6.73</td>
<td>0.168</td>
<td>0.915</td>
<td>7.39</td>
<td>0.302</td>
</tr>
<tr>
<td>40m</td>
<td>1.44</td>
<td>0.526</td>
<td>0.525</td>
<td>11.9</td>
<td>0.320</td>
<td>2.263</td>
<td>13.62</td>
<td>0.527</td>
</tr>
</tbody>
</table>
Unified framework

• Challenge: unify numerical simulations and real-time control frameworks
  • Reduce development cost, increase reliability and upgradability
  • Achieved performance compatible with such a unified approach
Unified framework

- **Challenge**: unify numerical simulations and real-time control frameworks
  - Reduce development cost, increase reliability and upgradability
  - Achieved performance compatible with such a unified approach
Adaptive Optics real-time control

- AO RTC simulator

![Diagram showing CPU, GPU, Pixel data, Subap Nphot, X centroids, Y centroids, and DM commands]
Adaptive Optics real-time control

- AO RTC simulator: achieved performance on a Tesla M2090
- Compatible with first light instrument (MICADO) specifications: > 500 FPS
AO real-time control

- How do we do that in real life?
- PRANA: Prototype Real-time Architecture for Next generation Ao
AO real-time control

- How do we do that in real life?
- PRANA: Prototype Real-time Architecture for Next generation Ao

- Before GPUdirect RDMA: multiple copies of pixel data
AO real-time control

- How do we do that in real life?

- PRANA: Prototype Real-time Architecture for Next generation Ao

  - With GPUdirect RDMA: direct transfer of pixel data from camera to GPU
AO real-time control

- PRANA: Prototype Real-time Architecture for Next generation Ao
  - Using GPUdirect RDMA from NVIDIA
  - PLDA Stratix V PCIe development board
  - QSFP+ to PCIe
  - QuickPCIe IP core from PLDA for DMA
AO real-time control

- **PRANA**: Prototype Real-time Architecture for Next generation Ao
  - Using GPUdirect RDMA from NVIDIA
  - PLDA Stratix V PCIe development board
  - QSFP+ to PCIe
  - QuickPCIe IP core from PLDA for DMA

42 Gb/s demonstrated in the lab from PLDA board to GPU (data generated on the board) : to be compared to 64Gb/s max bandwidth of PCIe Gen3
AO real-time control

- PRANA: Prototype Real-time Architecture for Next generation Ao
  - Using GPUdirect RDMA from NVIDIA
  - PLDA Stratix V PCIe development board
  - QSFP+ to PCIe
  - QuickPCIe IP core from PLDA for DMA
  - QuickUDP IP core from PLDA
AO real-time control

- **PRANA**: Prototype Real-time Architecture for Next generation Ao
  - Using GPUdirect RDMA from NVIDIA
  - PLDA Stratix V PCIe development board
  - QSFP+ to PCIe
  - QuickPCIe IP core from PLDA for DMA
  - QuickUDP IP core from PLDA

8.8 Gb/s per 10 Gbe link demonstrated in a loopback mode (UDP stack + DMA through PCIe)
AO real-time control

- **PRANA: Prototype Real-time Architecture for Next generation Ao**
  - Building a full-scale prototype
  - 1rst commercial 10 Gbe Camera from Ermergent Vision Technologies
  - Up to 1.5 kFPS in 256x256 mode
    (1024x2048 @ > 300 FPS)
  - GigeVision protocol (GVCP / GVSP) integrated in the FPGA design
  - Camera to be integrated on an optical bench including:
    - WFS optics
    - Realistic turbulence generation through phase screens
    - « Truth » WFS for performance analysis
AO real-time control

- PRANA: Prototype Real-time Architecture for Next generation Ao
- FPGA design on PLDA board
Towards a GPU accelerated ELT

- GPUs seem to provide an adequate solution to various technological challenges for the design and construction of the E-ELT:
  - Enabling fast end-to-end simulations for the design phases of first light instruments
  - Enabling adaptive optics real-time control at required framerate

- Beyond first light instruments: first generation instruments
  - Several WFS / several DMs: even more degrees of freedom (up to 100k!)
  - End-to-end simulations not a good approach for preliminary design studies (just starting): parameter space is too large
  - Switching to pseudo-analytical approach: collaboration with Extreme Computing group @ KAUST (contact H. Ltaief)
  - Can we achieve low-latency high framerate real-time control for adaptive optics with a distributed architecture?
The team

- Arnaud Sevin
- Denis Perret
- Julien Brulé
- Bertrand Leruyet
- Maxime Lainé
- Florian Ferreira

and collaborators at GEPI, IPAG, LAM, ONERA, MdIS, ESO, KAUST, INFN
Thank you