

# Cluster of GPUs

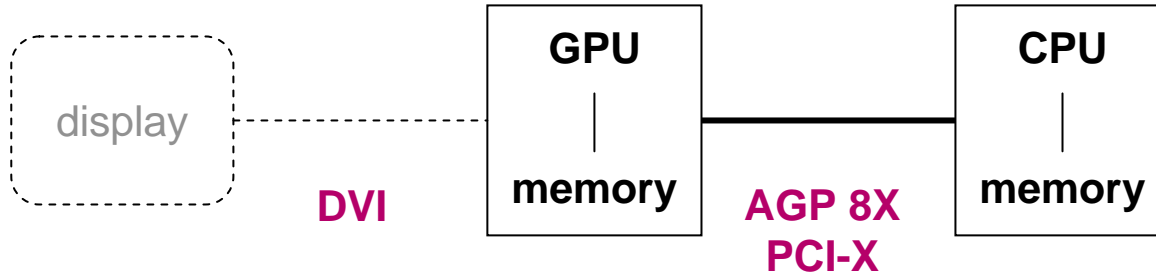
Jan Prins

Research Fair – Aug 2004



# Single-GPU

- System-level description

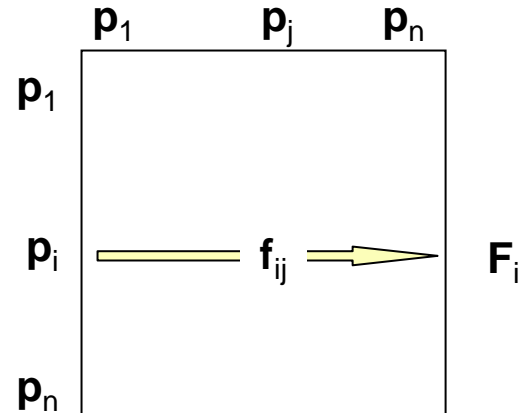
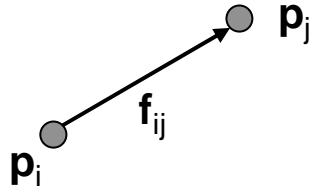


- Sample application: all-pairs n-body simulation

– n bodies:  $\mathbf{p}_1 \dots \mathbf{p}_n$

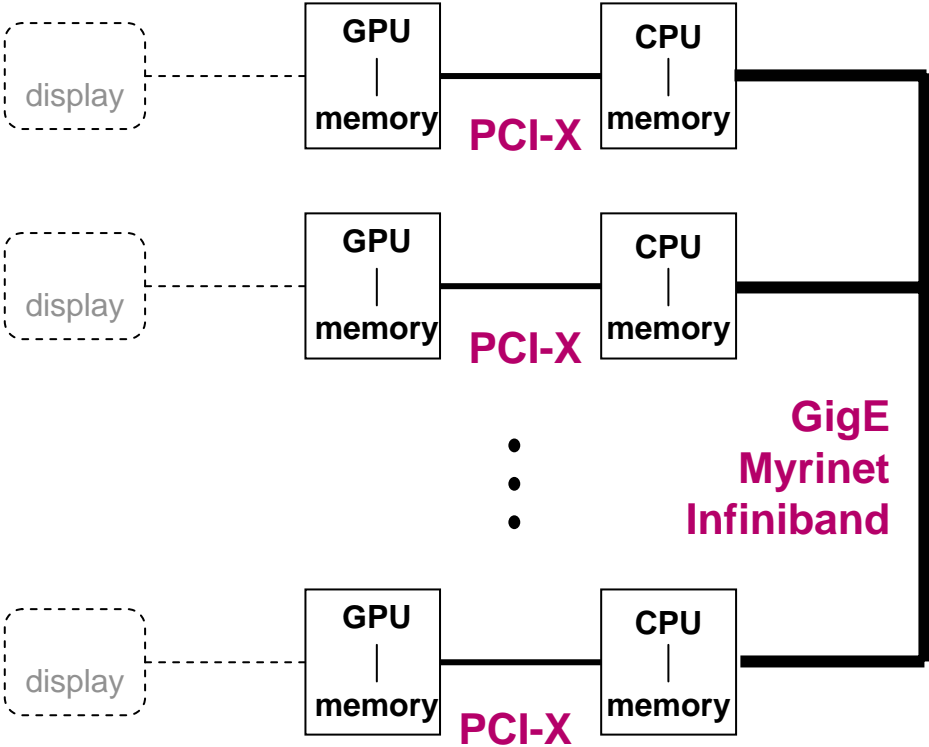
$$\mathbf{f}_{ij} = -G \cdot \frac{m_i \cdot m_j}{r_{ij}^2} \cdot \frac{\mathbf{p}_i - \mathbf{p}_j}{r_{ij}}$$

$$\mathbf{F}_i = \sum_{j \neq i} \mathbf{f}_{ij}$$



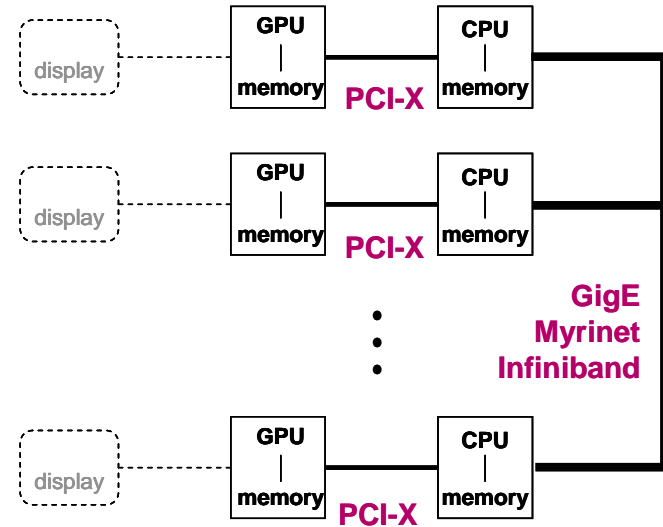
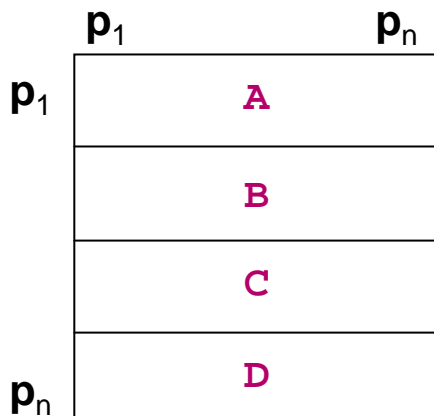
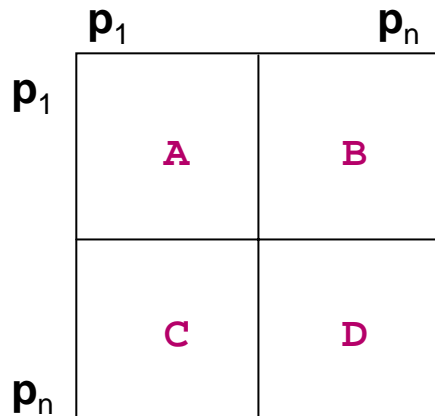
# Cluster of GPUs

- System-level description



# Decomposition of n-body computation

- Strategies for 4 GPUs **A, B, C, D**



- Local force evaluation
- Combine forces from GPUs on same row
- Update local body positions
- Send updated positions to other GPUs



# How do we program such applications?

---

- **At the GPU level**

- OpenGL or DirectX calls
- Cg
- Brook

- **At the CPU level**

- C + MPI (Message Passing Interface)
  - » Single program augmented with calls to the communications library
  - » Provides efficient implementation of collective communication operations
- UPC (Unified Parallel C)
  - » Single program with explicit declaration of shared and local variables
  - » Flexible access to shared state, but less effective for collective communication



# The RA position

---

- **We are interested in exploring the following questions**
  - What is the best systems-level realization of a cluster of GPUs
    - » Investigate options and choose wisely
  - What is a suitable programming model for a cluster of GPUs
    - » Parallelize some sample applications
    - » Evaluate programmability and analyze performance in detail
- **What skills are needed**
  - Familiarity with Linux
  - Courage in the face of early release software and drivers
  - Interest in GPUs and parallel computing
- **Who is involved**
  - Naga Govindaraju, Dinesh Manocha, Jan Prins, Dan Reed

