# XDEM: from HPC to the Cloud

Xavier Besseron

LuXDEM Research Centre
http://luxdem.uni.lu/

BusinessClouds 2017
January 24, 2017

# Outline

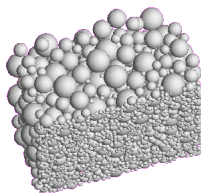# eXtended Discrete Element Method

# What is XDEM?

# What is XDEM?

XDEM software is multi-physics simulation toolbox modeling granular materials and processes
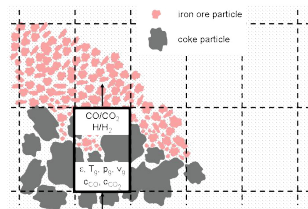


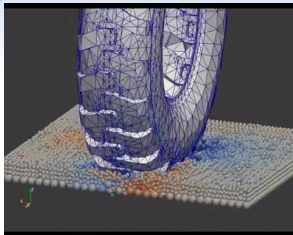## Particle Motion
Snow, Sand, ...
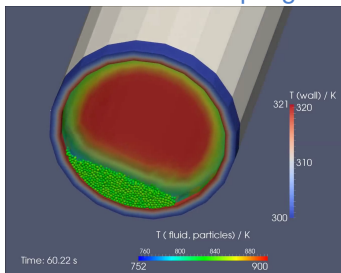


## Chemical Reactions
Coke, Iron ore, Biomass, ...



$+$ coupling external libraries:

CFD with OpenFoam, FEM with Diffpack
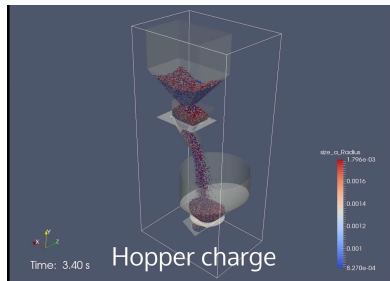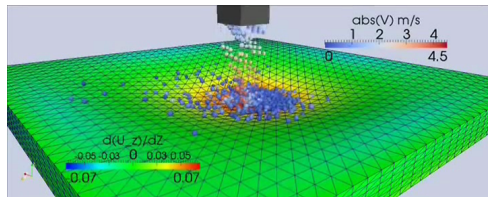
# Examples



Tire on Snow
XDEM–FEM coupling



Charge/Discharge of hoppers



Heat transfer to the walls
of a rotary furnace
XDEM–CFD coupling
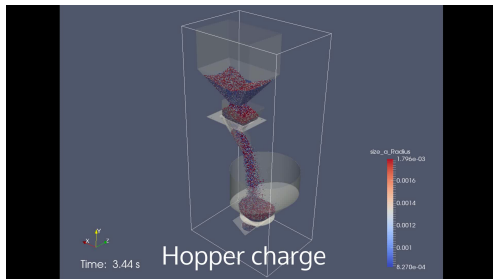


Impacts on an Elastic Membrane
XDEM–FEM coupling

# XDEM-HPC:

# Parallel Implementation of XDEM

# XDEM needs HPC!



Hopper charge

- 427,878 particles
- 6,742 triangles
- Impact
- Rolling friction

## Hopper charge

- 15 s of simulation
- 92 hours with 120 cores
- Est. seq. time > 4 months

## Hopper discharge

- 18 s of simulation
- 120 hours with 144 cores
- Est. seq. time > 6 months

# XDEM needs HPC!



- 427,878 particles
- 6,742 triangles
- Impact
- Rolling friction

## Hopper charge

- 15 s of simulation
- 92 hours with 120 cores
- Est. seq. time > 4 months

## Hopper discharge

- 18 s of simulation
- 120 hours with 144 cores
- Est. seq. time > 6 months

# XDEM needs HPC!



- 427,878 particles
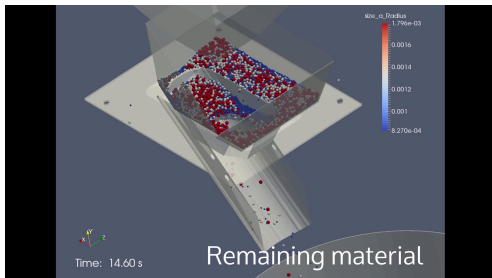- 6,742 triangles
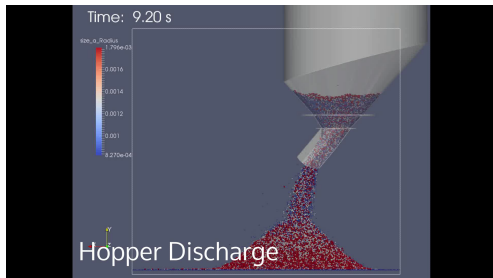- Impact
- Rolling friction

## Hopper charge

- 15 s of simulation
- 92 hours with 120 cores
- Est. seq. time > 4 months

## Hopper discharge

- 18 s of simulation
- 120 hours with 144 cores
- Est. seq. time > 6 months

# Parallel XDEM simulation for HPC

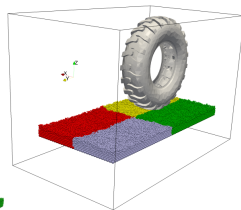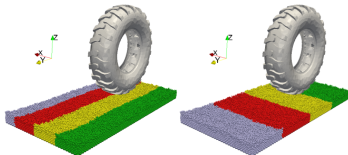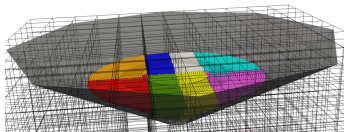## XDEM is computation intensive and uses a lot of memory

⇒ Parallel and distributed implementation

- Aggregates memory of many computing nodes
- Benefits from speedup of many computing cores

## Periodic Load balancing

- based on Orthogonal Recursive Bisection (ORB)
- adapt to changes in particle distribution

## Communications based on MPI

# XDEM Performance (Work-in-Progress)
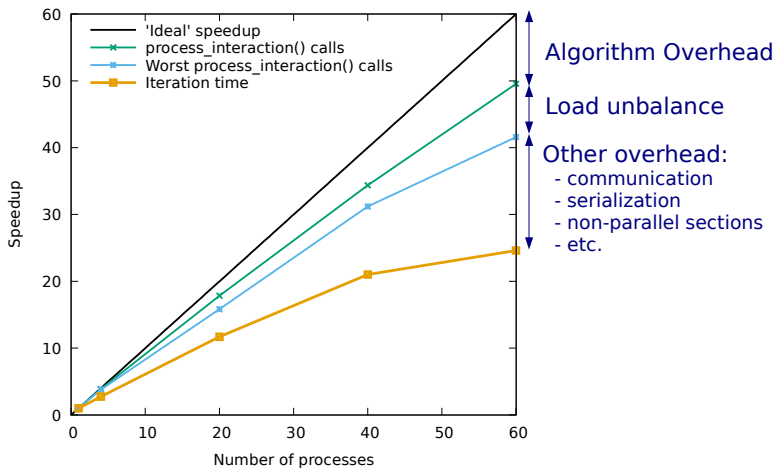
# XDEM-HPC: Work-in-Progress / Future Work

## Better Load balancing algorithms

- Zoltan, RIB, RCB

## Hybrid MPI/OpenMP

- $\Rightarrow$ Avoid intra-node communications

## GPU-based collision detection

- using a third-party library (Bullet)

# XDEM in the Cloud

# XDEMaaS: XDEM as a Service

# XDEMaaS: XDEM as a Service

## What? → Full XDEM Workflow

- Pre-processing with the XDEM GUI
- XDEM parallel execution
- Post-processing with ParaView

## For Who?

- Team members
- Interns, Partners
- ...

## Why? → Hide technical difficulties

- Easy installation and usage
- Controlled and well-tested environment
- No data transfer, data lives 'in the Cloud'
- Can accommodate the different requirements:
  - Large memory, many computing nodes
  - Graphical session with GPU rendering

# XDEMaaS: XDEM as a Service

## How? → Using UL HPC infrastructure

- XCS portal web interface
- Remote visualization based on VirtualGL/TurboVNC
- Pre-configured jobs for XDEM tasks
  - Visualization jobs run on GPU nodes
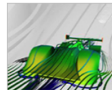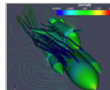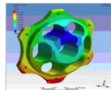  - Parallel jobs on Gaia cluster

# XDEMaaS: XDEM as a Service



XCS web portal

# XDEMaaS: XDEM as a Service



Pre-processing using the XDEM GUI via TurboVNC

# XDEMaaS: XDEM as a Service



Submission of an XDEM parallel job using the XCS webportal

# XDEMaaS: XDEM as a Service
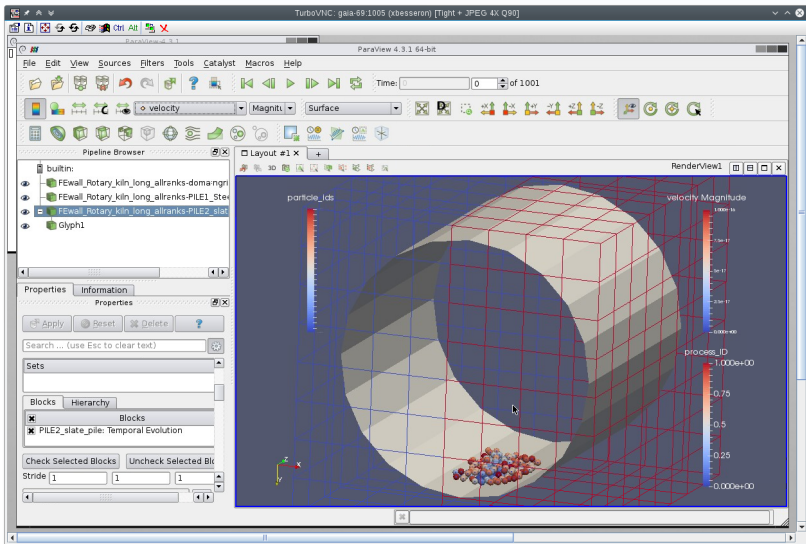


Post-processing with ParaView via TurboVNC

# XDEMaaS: Future Work

## Support for transparent checkpoint/restart

- Hide all the technical issues

## Interactive control and monitoring of XDEM simulations

- Display some metrics in real time
  - iteration time, maximum velocity, total kinetic energy, ...
- Change some parameters, eg simulation end time
- Trigger actions: force checkpoint, write current output

## Real time visualization

- Using ParaView Catalyst?

# Cloud vs In-House HPC:

# A Cost Study using XDEM

# Cost Analysis of an HPC platform [1]

Amortized CAPEX + OPEX → TCO (Total Cost of Ownership)

## CAPital EXpenses

- machines *(4 years)*
- servers *(3 years)*
- storage *(5 years)*
- interconnect *(8 years)*
- room equipment *(15 years)*
- building *(30 years)*

## OPerating EXpenses

- manpower
- energy (power and cooling)
- support
- software licenses

→ Cost are calculated assuming a full utilization of the platform

[1] J. Emeras, S. Varrette, and P. Bouvry. "Amazon Elastic Compute Cloud
(EC2) vs. in-House HPC Platform: a Cost Analysis". In: CLOUD 2016

# Cost Analysis of UL HPC platform [1]

## Chaos cluster

| Node | Cores | Memory | GPUs | Nodes | CPU family | Disk | GFLOPS | Hourly Cost |
|------|-------|--------|------|-------|------------|------|--------|-------------|
| d-cluster1 | 12 | 24 GB | 0 | 16 | Westmere | 250 GB | 108.48 | $ 0.439 |
| e-cluster1 | 16 | 32 GB | 0 | 16 | Sandybridge | 250 GB | 281.60 | $ 0.433 |
| h-cluster1 | 12 | 24 GB | 0 | 32 | Westmere | 250 GB | 108.48 | $ 0.428 |
| r-cluster1 | 32 | 1024 GB | 0 | 1 | Nehalem | 250 GB | 289.28 | $ 1.814 |
| s-cluster1 | 16 | 32 GB | 0 | 16 | Sandybridge | 250 GB | 281.60 | $ 0.433 |

## Gaia cluster

| Node | Cores | Memory | GPUs | Nodes | CPU family | Disk | GFLOPS | Hourly Cost |
|------|-------|--------|------|-------|------------|------|--------|-------------|
| gaia-[1-60] | 12 | 48 GB | 0 | 60 | Westmere | 256 GB | 108.48 | $ 0.453 |
| gaia-[123-154] | 12 | 48 GB | 0 | 32 | Westmere | 256 GB | 147.36 | $ 0.344 |
| gaia-[61-62] | 12 | 24 GB | 1792 | 2 | Westmere | 256 GB | 108.48 | $ 0.641 |
| gaia-[63-72] | 12 | 24 GB | 10240 | 10 | Westmere | 256 GB | 108.48 | $ 0.599 |
| gaia-[75-79] | 16 | 64 GB | 12480 | 5 | Sandybridge | 256 GB | 281.60 | $ 0.577 |
| gaia-[83-122] | 12 | 48 GB | 0 | 40 | Westmere | 256 GB | 140.64 | $ 0.344 |
| gaia-73 | 160 | 1024 GB | 0 | 1 | Sandybridge | 256 GB | 2,560.00 | $ 2.649 |
| gaia-74 | 32 | 1024 GB | 0 | 1 | Sandybridge | 256 GB | 614.40 | $ 1.516 |
| gaia-[155-178] | 24 | 128 GB | 0 | 24 | Haswell | 160 GB | 983.04 | $ 0.458 |

[1] J. Emeras, S. Varrette, and P. Bouvry. "Amazon Elastic Compute Cloud (EC2) vs. in-House HPC Platform: a Cost Analysis". In: CLOUD 2016

uni.lu
UNIVERSITÉ DU
LUXEMBOURG

# Experimental Setup: Execution Platforms

## Amazon EC2: c4.8xlarge instance

- Intel Haswell Xeon E5-2666v3 @ 2.9 Ghz
- 36 virtual cores
- 60 GBytes memory
- 10 Gigabit Ethernet (10 Gb/s)
- → Hourly cost: $ 1.906 (on-demand) or $ 1.2182 (reserved full upfront)

## UL HPC cluster: Dell FC430 nodes

- 2 × Intel Haswell Xeon E5-2680v3 @ 2.5 Ghz
- 24 cores
- 120 GBytes memory
- InfiniBand QDR (40 Gb/s)
- → Hourly cost: $ 0.458

# Experimental Setup: XDEM testcase & Benchmark



Radius (m)

Time: 3.16 s
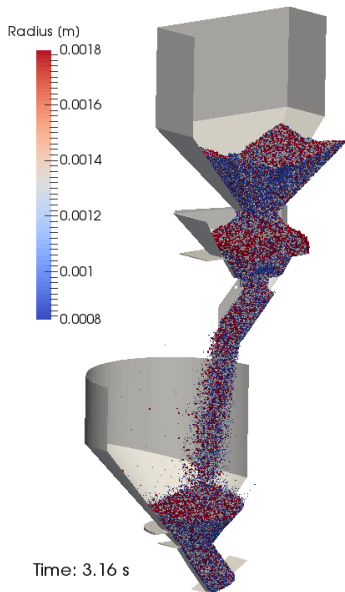
## Discharge of granular materials

- 427,878 spherical particles
- 6,742 triangles
- Impact & Rolling friction
- 15 s simulation
- 750,000 timesteps with $\Delta t = 2e^{-5}$ s
- 30 GBytes output
- Decomposition in 1,470,000 cells
- ORB partitioning
- Sequential time: 3.5 months

$\Rightarrow$ Extracted XDEM benchmark

- 100 representative timesteps

# Performance evaluation

### Average execution time of one XDEM timestep

| | Timestep execution time | |
|---|---|---|
| | Gaia Dell FC430 | EC2 c4.8xlarge |
| Sequential (1 core) | 13.55 s | 13.66 s |
| Parallel 24 cores on 1 node | 0.99 s | 1.36 s |
| Parallel with 240 cores | 0.31 s | 0.40 s |

$\Rightarrow$ Similar sequential performance

$\Rightarrow$ 30% overhead with intra and inter-node communication

# Taking into account performance and cost

Calculating Price per Timestep
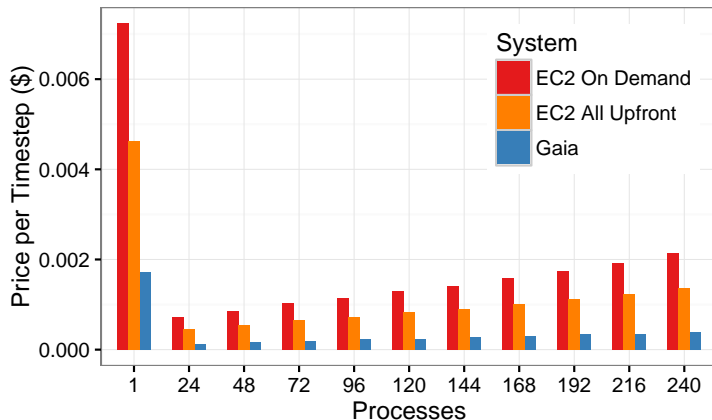


$\Rightarrow$ Amazon EC2 is 3 times to 5 times more expensive

$\Rightarrow$ Sequential execution still requires to pay for one full node

# Cost of running the full XDEM testcase

## Extrapolated duration and price

- Full testcase with 750,000 timesteps
- Trade-off between deadline and cost

| | **Cheapest** (24 processes) | | **Fastest** (240 processes) | |
|---|---|---|---|---|
| Gaia Dell FC430 | 8.61 days | $ 94.64 | 2.69 days | $ 295.29 |
| c4.8xlarge on-demand | 11.81 days | $ 540.55 | 3.50 days | $ 1,602.84 |
| c4.8xlarge reserved | 11.81 days | $ 345.48 | 3.50 days | $ 1,024.44 |

# Conclusions

## XDEM needs HPC

- Important requirements in computing power and memory

## XDEM can provide Cloud services: XDEMaaS

- Accommodate the various requirements of XDEM workflow
- Hide many technical issues

## XDEM can use Cloud platforms (IaaS, PaaS)

- Can be expensive!
- XDEM is an HPC application: performance is not yet satisfactory
- Solutions are on the way: containers, SR-IOV-based MPI, ...

# Thank you!

# Any question?

## LuXDEM Research Centre
http://luxdem.uni.lu/