

Excerpt from Inaugural oneAPI Developer Summit

November 12-13, 2020

<https://webinar.intel.com/oneAPIDeveloperSummit2020>

Ho Leung Ng (Kansas State Univ.)

Viewing Ho Leung Ng's application
Prof_Ho_Intel Nov 2020.pptx

File Home Insert Draw Design Transitions Animations Slide Show Review Help EndNote X8 JChem Foxit PDF

Normal Outline Slide Notes Reading View Sorter Page View Master Views Master Master Master Notes Zoom Fit to Window Color Grayscale Black and White New Window Cascade Move Split Macros

MAIN TARGET AND LEAD Alpha protein GC376

Inverse Design Mapping functionality to molecules

Dreaming of faces

Dreaming of molecules

GENERATIVE AI TO CREATE NEW MOLECULES Computer creativity (inverse design): Can an algorithm propose new molecule with desired properties? Classification and regression in old entries! Chemical space is virtually infinite. Availability of accurate physics/chemistry based cost functions. Alán Aspuru-Guzik (Toronto)

GENETIC ALGORITHMS + DEEP LEARNING arxiv:1909.11655

Generator Genetic Algorithm

Molecules - m

Properties $J(m)$ Evaluation

Fitness eval: $F(m) = J(m) + \beta.D(m)$

Discriminator Neural Network

Dataset Molecules

Discriminator scores $D(m)$

KIRCHEN ANALOGUE

Viewing Ho Leung Ng's application

Prof_Ho_Intel Nov 2020.pptx

File Home Insert Draw Design Transitions Animations Slide Show Review View Recording Help EndNote X8 JChem Foxit PDF

Normal Outline Slide Notes Reading View Sorter Page View Master Views

Ruler Gridlines Guides Zoom Fit to Window New Window Cascade Switch Windows Macros

Color Grayscale Black and White Color/Grayscale Window Macros

Slide Master Handout Master Notes Show Zoom Zoom Window Window

6 Inverse Design Mapping functionality to molecules

7 Dreaming of faces

8 Dreaming of molecules

9 GENERATIVE AI TO CREATE NEW MOLECULES
Computer creativity: inverse design: Can an algorithm propose new molecules with desired properties?
Classification and regression is old science!
Chemical space is virtually infinite.
Availability of accurate physics/chemistry based cost functions.
Alan Aspuru-Guzik (Toronto)

10 GENETIC ALGORITHMS + DEEP LEARNING
Workflow diagram: Input -> Genetic Algorithm -> Deep Learning -> Output

11 PUBCHEM ANALOGS

12 SIMPLE AI + COMPLEX SCORING FUNCTIONS
1. Small datasets available for experimental science
2. Focus on calculation: accuracy
3. Increased availability of computational power

PUBCHEM ANALOGS

The slide displays a 2D chemical structure of a molecule with two phenyl groups, a central sulfur atom, and nitrogen atoms. To the right is a 3D molecular model showing a green molecule interacting with a purple protein surface.



David Hardy—NAMD (UIUC)

Viewing David Hardy's desktop

Hardy's desktop

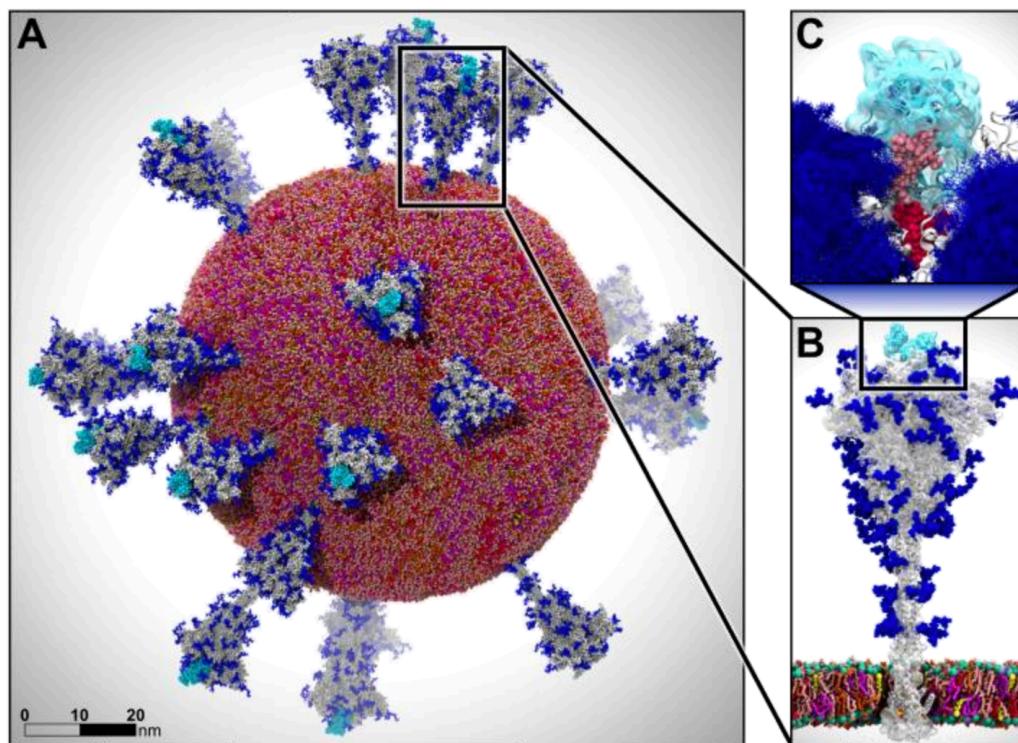


NAMD Simulating SARS-CoV-2 on Frontera and Summit

Collaboration with Amaro Lab at UCSD, images rendered by VMD

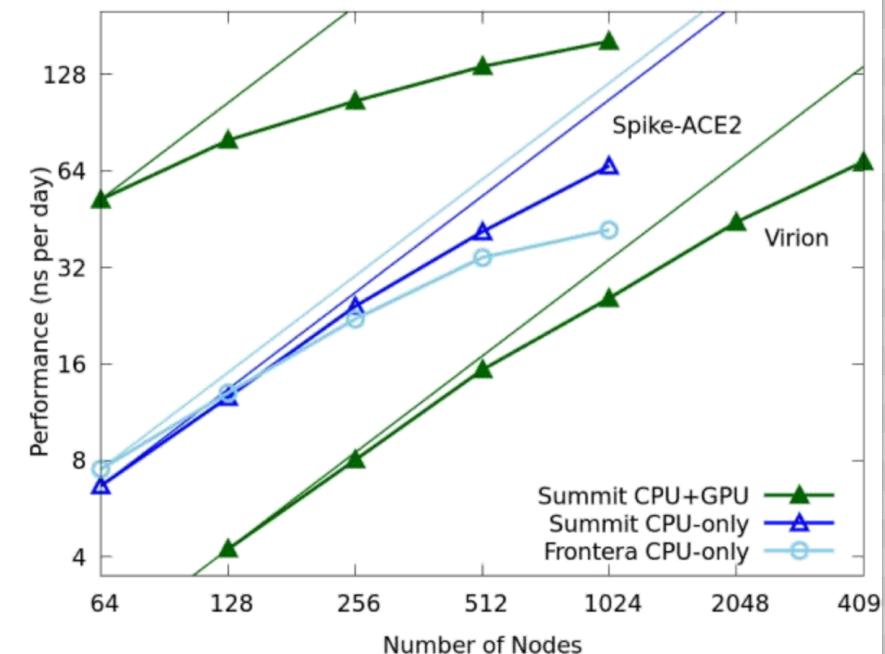
Paper is SC20 Gordon Bell Finalist, involving overall 1.13 Zettaflops of NAMD simulation

(A) Virion, (B) Spike, (C) Glycan shield conformations



Scaling performance:

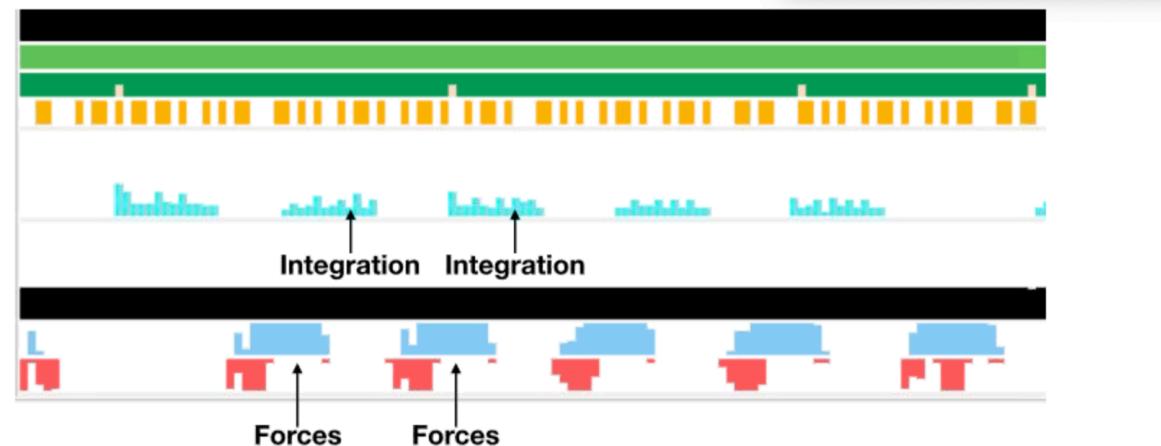
- ~305M atom virion
- ~8.5M atom spike



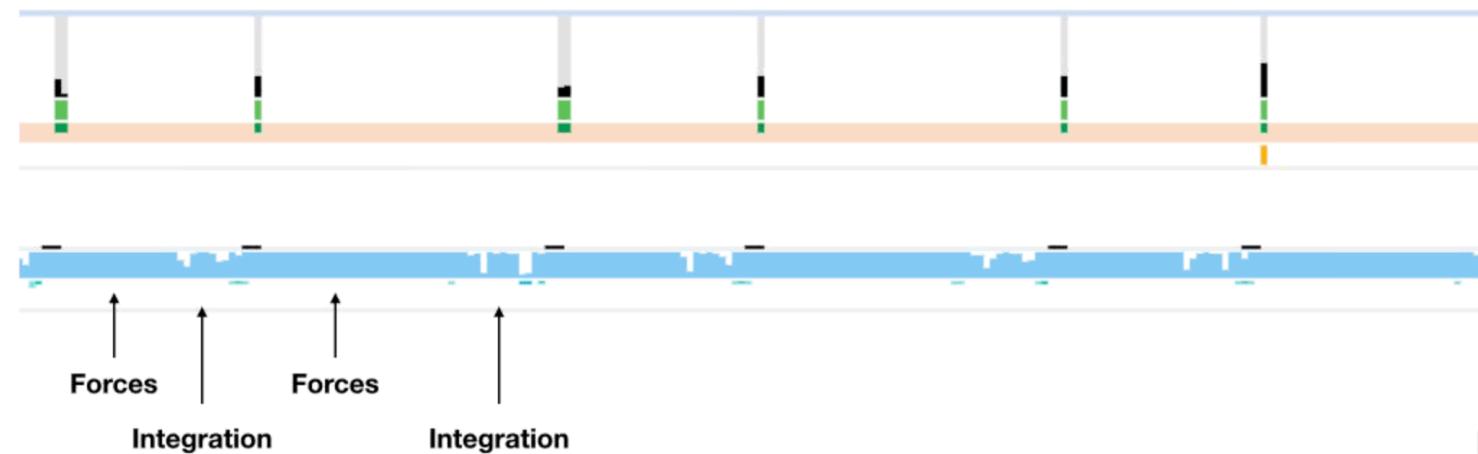


GPU-Resident NAMD Has Better GPU Utilization

NAMD 2.14
Gaps between GPU tasks

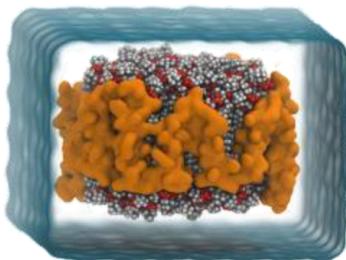


NAMD 3.0
No CPU bottlenecks





GPU-resident NAMD Fully Utilizes GPU-de



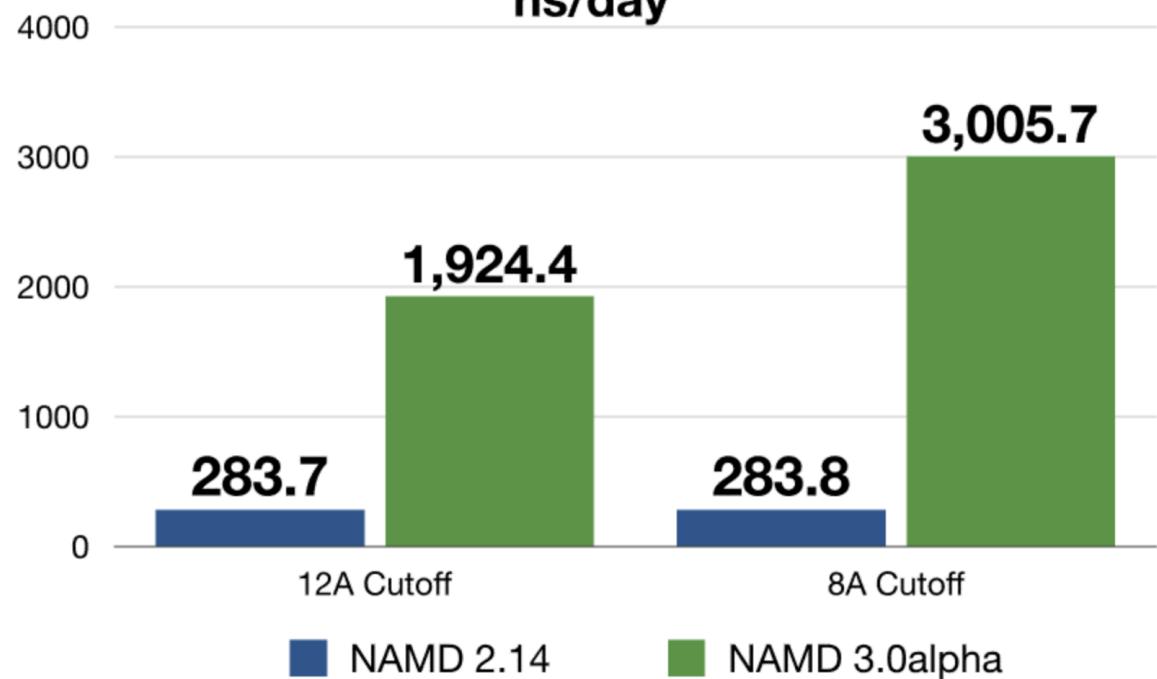
ApoA1
92k atoms



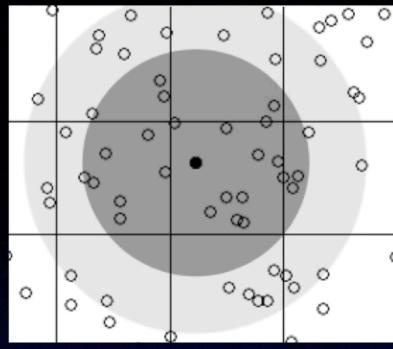
16 Replicas on DGX-2
1 for each NVIDIA V100

We want to run NAMD this way on Aurora!

Measuring aggregate performance
ns/day



Erik Lindahl—GROMACS (KTH)



You Viewing Erik Lindahl's desktop

i=3: 5 6 9 12 15 17 18 25 32 ...

i=4: 7 8 9 11 12 15 17 25 32 43 54 ...

... 8 9 10 11 12 13 19 20 ...



Load 1 particle, then compute 1 force, store it

$$cf. q = \frac{\text{fast computations}}{\text{slow memory I/Os}}$$

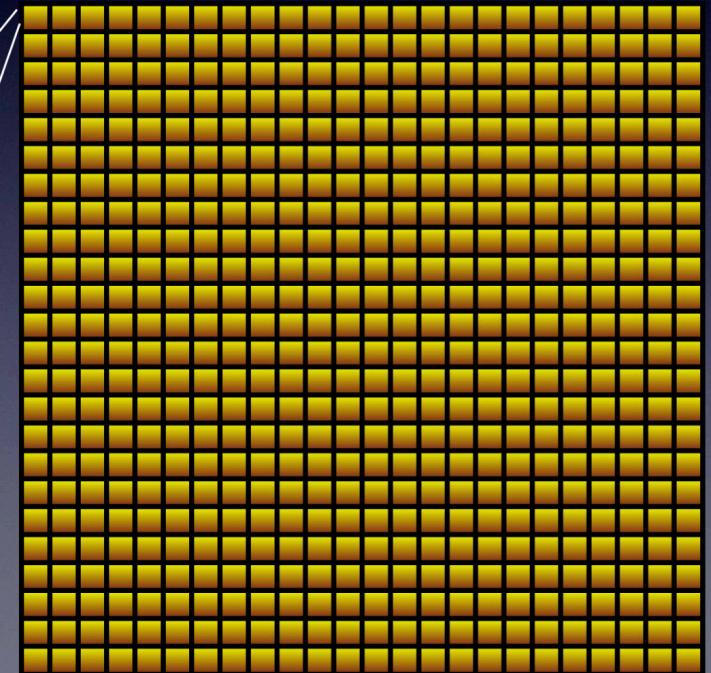
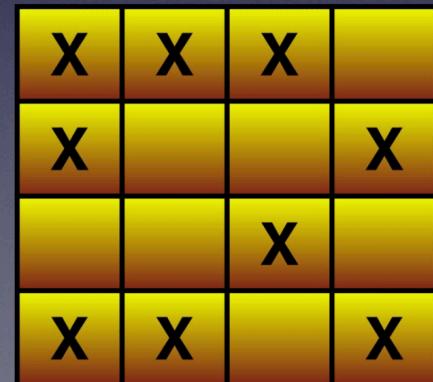
Too much data to send each step, each particle has different neighbors, memory bottleneck:
Won't work well on accelerators.

GPU solution:

Group atoms into “tiles”

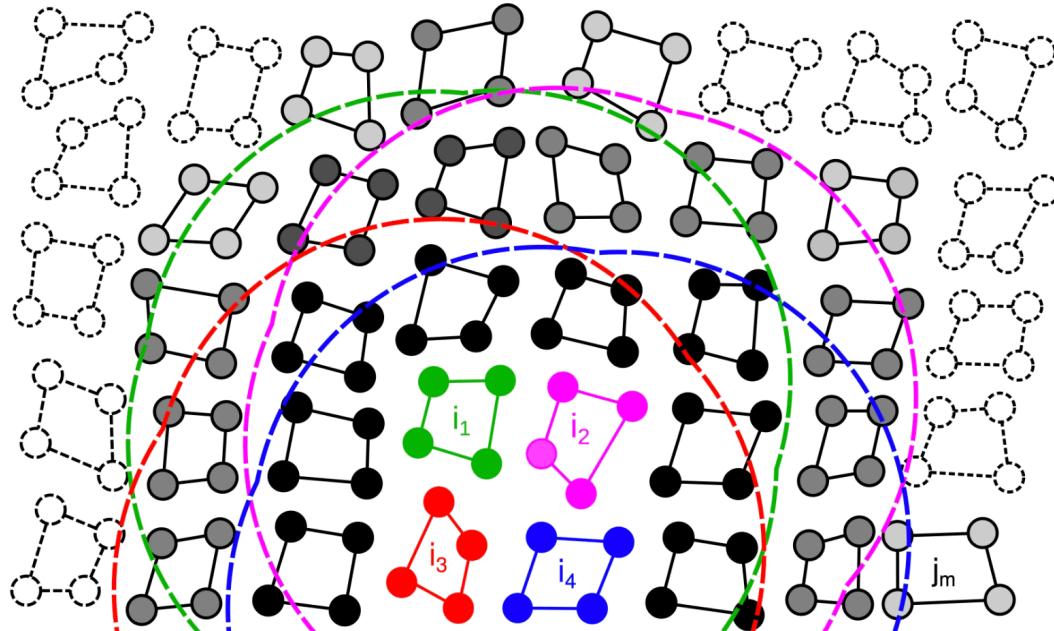
High raw FLOPS-rate,
but low efficiency

Load 4 particles, then
compute 16 forces - only
uses 25% memory bw!



cf. 4-vector SIMDization

Heterogeneous acceleration means fast development - because we can run complex new algorithms on the CPU side using standard tools



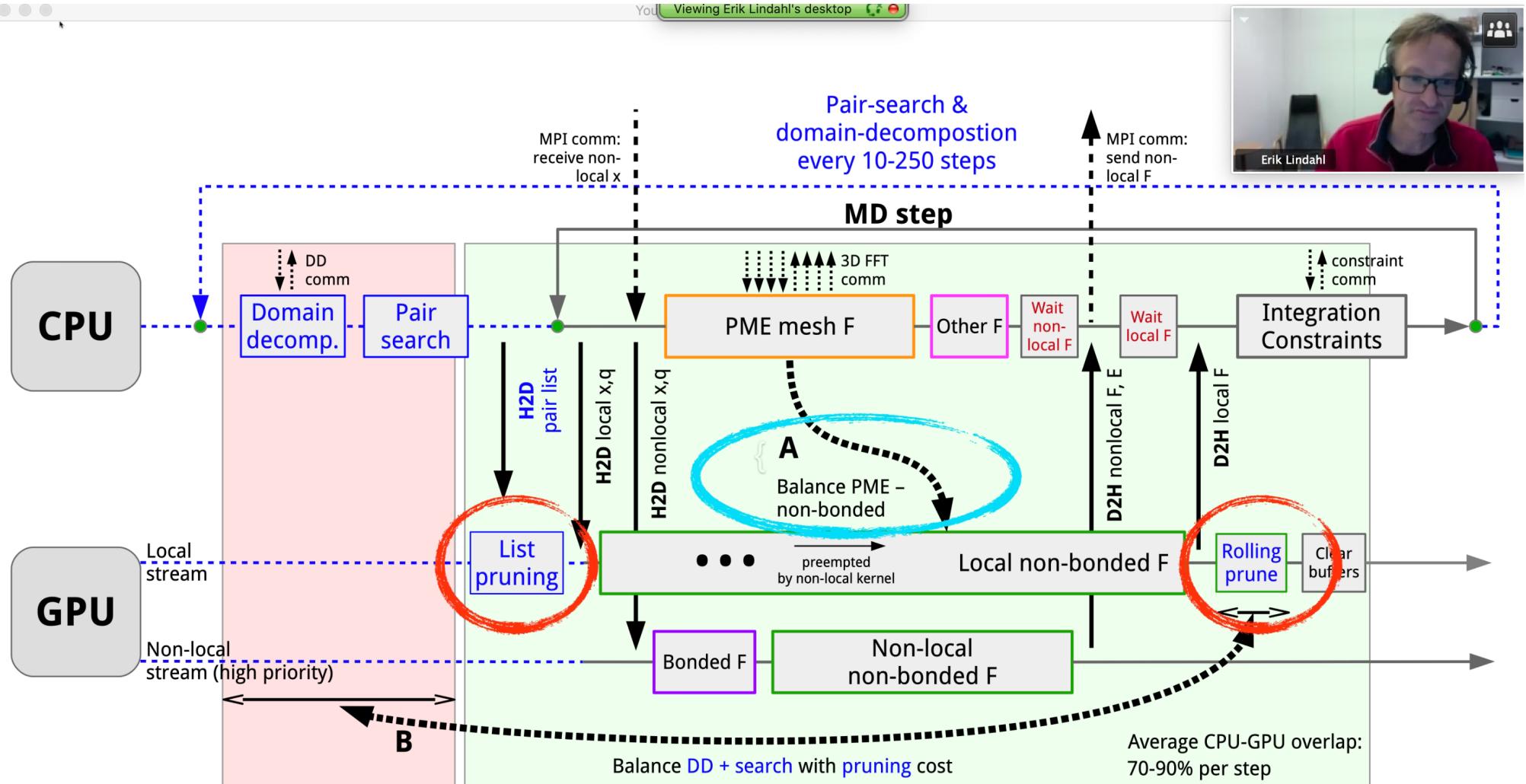
4x4 tiles not sufficient to keep modern GPUs busy, but 8x8 much less efficient

Don't fall in trap of maximising FLOPS;
our goal is runtime performance

Introduce more complex neighborlists:
superclusters with 4-16 4x4 clusters

Generate GPU-optimal data structure on the **CPU** using 128-bit AVX SIMD for bounding box operations (only three elements in cartesian coordinates). Re-use data for 10-250 steps on the GPU. Works in parallel with data from other nodes.





Data structure is already on device, so we can prune the list a bit (partial is fine!) any time the device is idle

cf. OpenGP idle event handler

You are viewing: Russ Beutler's desktop

Jeopardy! - The Training Arcade

intelsoftware.thetrainingarcade.com/games/jeopardy/dpcc

Apps FIRE! Real Estate Dropbox NETGEAR Five things you nee... Happiness Education Career Intel Marketing 'One API' Project -... Secret Flying | Che... My LastPass Vault My LastPass

\$1,600 II

DPC++ ONEAPI PARALLEL

\$200 \$200 \$200

\$400 \$400

\$600

Russ Beutler

Department or Event Name

Intel Confidential

Intel 3