WHITE-BOX MODELLING OF PARALLEL COMPUTING DYNAMICS

Ayesha Afzal <ayesha.afzal@fau.de>, Georg Hager (advisor), Gerhard Wellein (advisor)

Department of Computer Science | Erlangen National High Performance Computing Center | Friedrich-Alexander University Erlangen-Nürnberg, Germany

1. MOTIVATION AND CHALLENGES

Problem

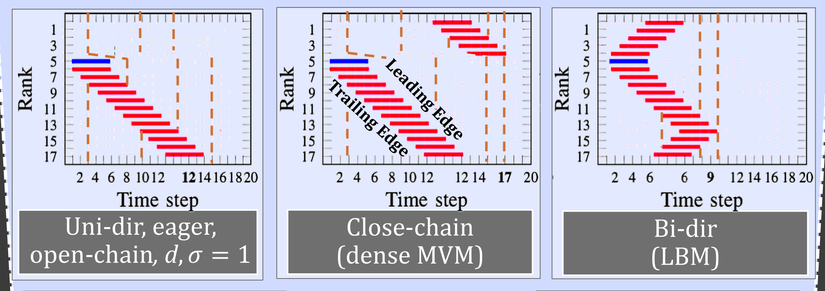
Variability can bring acceleration, speed down, or neither

Model for runtime of distributed-memory parallel applications = Model for runtime of computation (Roofline, ECM or other model) + Model for runtime of communication (Hockney, LogGP or other model) + Model for runtime driven by variability

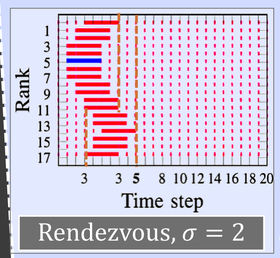
Research vision

"White-box" first-principle performance modelling of distributed-memory applications that incorporates the influence of variability

2. COMMUNICATION-DEPENDENT IDLE WAVE PROPAGATION VELOCITY v_p (NO VARIABILITY)



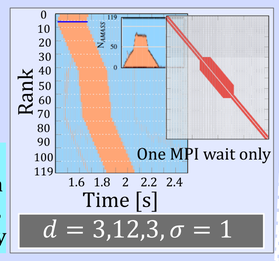
Uni-dir, eager, open-chain, $d, \sigma = 1$ | Close-chain (dense MVM) | Bi-dir (LBM)



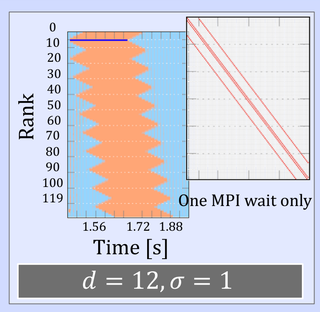
Rendezvous, $\sigma = 2$

$$\frac{v_p\left[\frac{\text{ranks}}{\text{s}}\right]}{} = \frac{\sigma.\kappa(d)}{T_{comp} + T_{comm}}$$

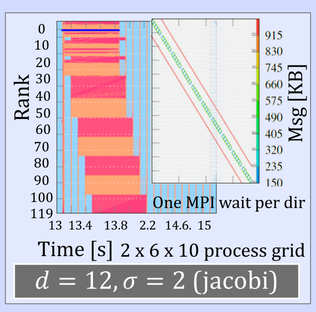
Idle period on rank 5 propagates depending on communication protocol, concurrency and topology



$d = 3,12,3, \sigma = 1$



$d = 12, \sigma = 1$

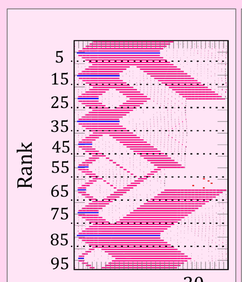


$d = 12, \sigma = 2$ (jacobi)

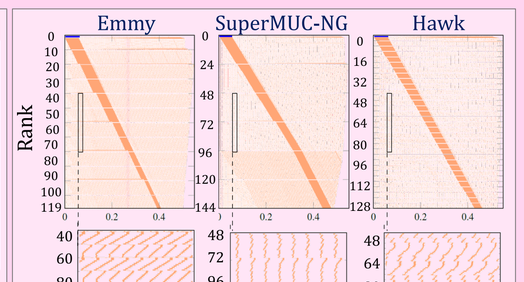
Afzal et al. DOI:10.1109/CLUSTER.2019.8890995 DOI: 10.1007/978-3-030-78713-4_19 DOI:10.13140/RG.2.2.28259.50726

Afzal et al. DOI:10.1109/CLUSTER.2019.8890995 DOI: 10.1007/978-3-030-78713-4_19 DOI:10.13140/RG.2.2.28259.50726

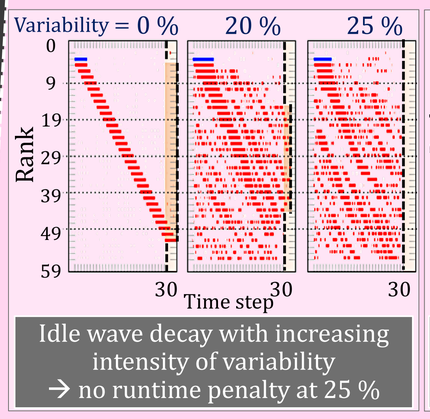
3. VARIABILITY-DRIVEN IDLE WAVE DECAY IN SCALABLE PROCESSES



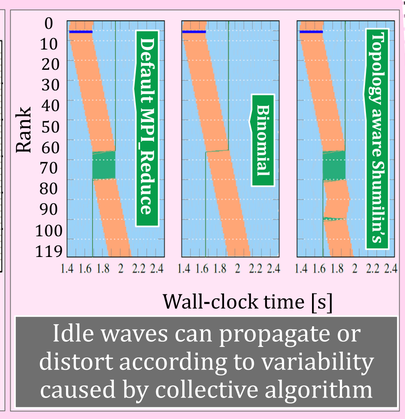
Nonlinear interaction of idle waves



Idle wave decay driven by structured variability owing to system topology

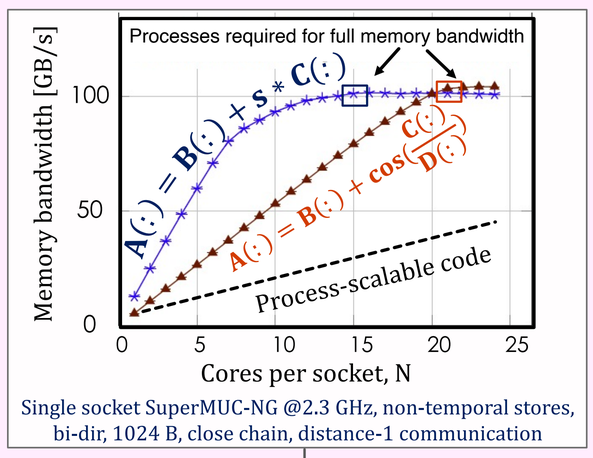


Idle wave decay with increasing intensity of variability → no runtime penalty at 25 %



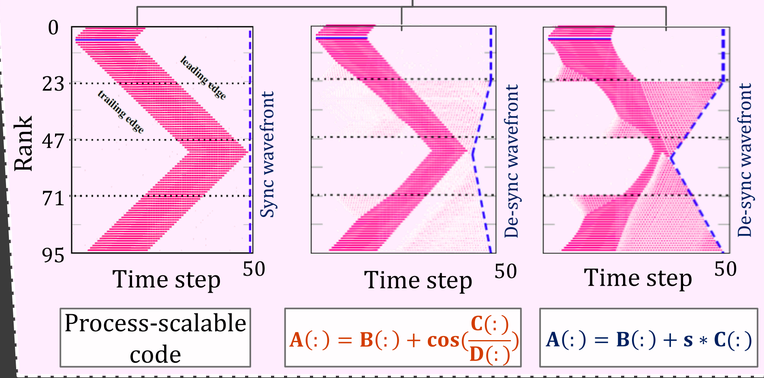
Idle waves can propagate or distort according to variability caused by collective algorithm

4. DE-SYNC WAVEFRONT FORMATION FOR ACROSS-PROCESSES BOTTLENECK



Single socket SuperMUC-NG @2.3 GHz, non-temporal stores, bi-dir, 1024 B, close chain, distance-1 communication

N=24 (4 sockets)



Process-scalable code | $A(:) = B(:) + \cos(\frac{C(:)}{D(:)})$ | $A(:) = B(:) + s * C(:)$

Afzal et al. DOI: 10.1007/978-3-030-50743-5_20 DOI:10.13140/RG.2.2.28259.50726

KEY-TAKEAWAYS

De-sync of processes causes automatic communication overlap and bottleneck evasion

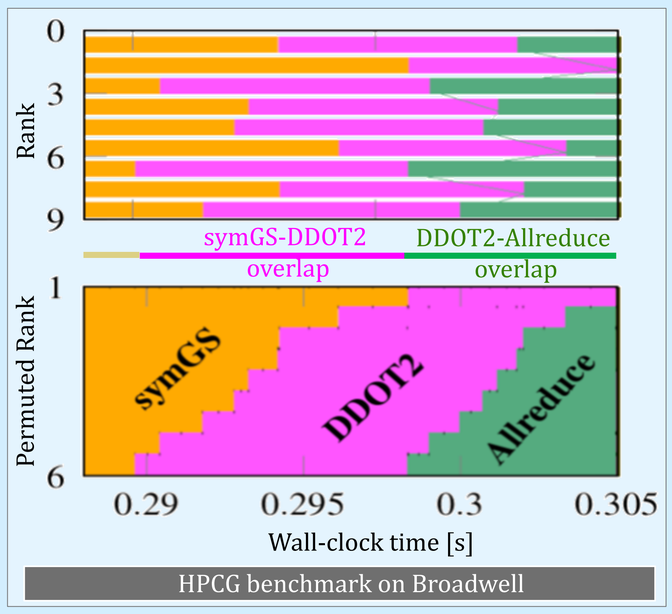
Variability is not all detrimental

Slower idle waves, strong bottleneck and variability help in amplifying the de-sync

Across context, if de-sync is desirable, apply optimization techniques amplifying the de-sync

Afzal et al. arXiv: 2011.00243 [cs.DC]

5. BANDWIDTH SHARING MODEL FOR BACK-TO-BACK DIFFERENT KERNELS TO PREDICT IF DE-SYNC IS AMPLIFIED OR DAMPED



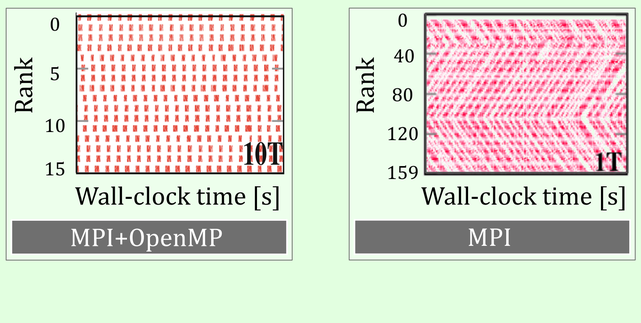
HPCG benchmark on Broadwell

Possible de-sync scenarios
1. Multi-phase applications
2. Spatial multitasking in GPUs
3. Task-parallel programs

6. USABILITY AND GENERALIZATION

Chebychev Filter Diagonalization

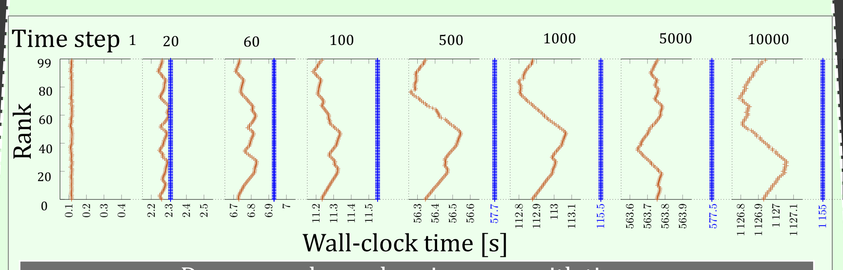
128x64x64 size, 6.7 GB data set, non-blocking MPI routines, 16 Emmy sockets



MPI - parallelized Lattice Boltzmann Fluid Solver

✚ Actual measurement
✚ Compositional model

302^3 lattice cell, 8 GB data set, non-blocking MPI routines, 10 Emmy sockets



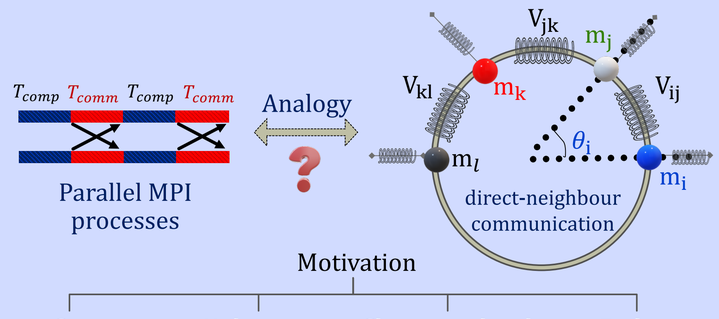
De-sync and speed-up increase with time

Afzal et al. DOI:10.1109/CLUSTER.2019.8890995 DOI: 10.1007/978-3-030-50743-5_20 http://tiny.cc/IReserachPoster-iSC21

7. PHYSICAL OSCILLATOR MODEL

We propose
A non-linear physical coupled phase oscillators model

The phenomenology of distributed applications suggests a physical interpretation of individual processes as a set of coupled oscillators whose inherent frequencies are influenced by coupling potential



Motivation: Cost effective | Unknown analogy | Cluster and code characterization tool | End-user tool

$$\dot{\theta}_i = \underbrace{\omega_i}_{\text{Inverse time step}} + \underbrace{\zeta_i(t)}_{\text{Noise}} + \underbrace{\frac{v_p}{N_p}\sum_{j=1}^{N} T_{ij}\, V_{ij}(\theta(t,\tau))}_{\text{Interaction term}}$$

N_p : System size; T_{ij} : Communication topology

$V_{ij}(\theta)$: Coupling potential reflects bottleneck structure

$\zeta_i(t)$: Local noise; τ: Interaction noise

8. OUTLOOK AND CONCLUSION

Addressing the analysis and simulation challenges of the dynamics of parallel computing

Exploration of various aspects and application scenarios of "white-box" modelling approach

Encompassing the performance-limiting bottlenecks, analysis guides appropriate code changes and insight into the hardware-software interaction

Connecting parallel computing dynamics to the physical world, oscillator model can serve as a high-level cluster characterization tool

KONWIHR OMI4papps | NHR FAU | FAU Friedrich-Alexander-Universität Erlangen-Nürnberg

POSTER ARTIFACT http://tiny.cc/PosterArtifacts