Toward Real-Time GPU-Accelerated Lattice Boltzmann LES for Molten Salt Reactor Thermal Hydraulics: Benchmark Validation and Demonstration

Suhyeon Kim[a], Jiho Kim[a], and Eung Soo Kim[a*]
[a] *Nuclear Engineering, Seoul National University, Seoul 08826, Korea*
*Corresponding author: kes77@snu.ac.kr

**Keywords :* molten salt fast reactor, lattice Boltzmann method, large-eddy simulation, thermal hydraulics, GPU computing, benchmark validation

1. Introduction

In a liquid-fuel molten salt reactor, the circulating salt transports heat while passing through the heat-generation region. A three-dimensional calculation must therefore represent recirculating flow, volumetric heating, heat removal, curved walls, and temperature-dependent properties over useful physical time intervals. These requirements make repeated design, uncertainty, and future digital-twin calculations expensive. Published Molten Salt Fast Reactor (MSFR) studies also show that mixing and the hottest-salt location depend strongly on the core and loop representation [1,2].

The lattice Boltzmann method (LBM) is attractive for this problem because its local collision and regular streaming operations map efficiently to graphics processing units (GPUs) [3]. This study evaluates a GPU-accelerated solver implementation that couples a three-dimensional 27-velocity (D3Q27) fluid LBM with a three-dimensional 15-velocity (D3Q15) thermal LBM. The word validation in the title refers to agreement within the stated benchmark conditions. The word demonstration refers to execution of the coupled GPU solver on an MSFR-scale quarter-domain reconstruction. It does not denote experimental reactor validation or a safety-grade prediction.

Real-time performance is defined here by the ratio of simulated physical time, denoted by t_sim, to elapsed wall-clock time, denoted by t_wall. A ratio of one or greater is required for a real-time claim. Kernel timing and end-to-end timing are reported separately because hydraulic preconditioning, compilation, diagnostics, and output materially change the comparison. Results measured directly, quantities derived from those measurements, and future hypotheses are kept separate.

The evidence is organized from separable tests to the final demonstration. A fixed-flux natural-convection cavity tests thermal advection, diffusion, and source reconstruction. A periodic Taylor--Green vortex tests the D3Q27 cumulant path in wall-free decaying turbulence. The completed CNRS Step 0.1 and Step 0.3 cases test laminar flow and frozen-flow thermal transport against independent participant data. Pump and heat-exchanger surrogates then exercise the coupled component path and provide the single-GPU reference performance case. Finally, a completed quarter-domain calculation demonstrates coupled MSFR-scale field generation while exposing the remaining curved-wall thermal conservation error.

2. Numerical Framework

2.1. Fluid and Thermal LBM

For fluid direction i, let f_i be the particle distribution and c_i its D3Q27 lattice velocity. Let ρ be density, u velocity, F applied force density, and Δ t the physical timestep. The macroscopic fields are reconstructed as

$$\rho = \sum_i f_i \ , \rho u = \sum_i c_i f_i + \left(\frac{\Delta t}{2}\right) F \qquad (1)$$

The flow solver relaxes cumulants on the D3Q27 lattice [4]. It permits a Smagorinsky coefficient C_s. A predefined Taylor--Green parameter study chose C_s=0, which removes the explicit Smagorinsky viscosity increment. The Taylor--Green and quarter-domain results are therefore described as an implicit-LES-like baseline, not as validation of a transferable positive Smagorinsky closure. Here, LES denotes large-eddy simulation.

For thermal direction i, let h_i be the D3Q15 thermal distribution. The transported scalar is sensible enthalpy per unit volume H. With reference density ρ_0, temperature-dependent specific heat c_p, reference temperature T_ref, and temperature T,

$$H(T) = \rho_0 \int_{T_{ref}}^{T} c_p(\theta) d\theta \qquad (2)$$

The collision operator is a modified multiple-relaxation-time formulation based on Huang and Wu [5]. It evaluates local thermal diffusivity, reconstructs temperature from enthalpy, and centers volumetric heating or a linear heat sink. If \bar h_i is the source-shifted thermal distribution and R is the net volumetric enthalpy source, physical enthalpy is reconstructed as

$$H = \sum_i \bar{h}_i + \left(\frac{\Delta t}{2}\right) R \qquad (2)$$

Fluid velocity is passed to the thermal solver, and reconstructed temperature is passed to the fluid-property evaluation during each coupled step.

2.2. Boundaries and GPU Execution

Curved stationary walls use the Bouzidi--Firdaouss--Lallemand (BFL) interpolated bounce-back rule [6]. Fluid BFL imposes no slip. The same stationary reflection applied to thermal distributions represents an

adiabatic, zero-normal-flux wall; it is not a prescribed wall temperature or conjugate heat-transfer model. The quarter domain uses post-collision half-link specular reflection on its two symmetry planes so that diagonal populations reach the correct tangentially adjacent cell.

Distributions use in-place Esoteric Pull streaming [10]. Construction-time cell roles and boundary links are lowered to generated Warp GPU kernels and invoked from JAX through its foreign-function interface. All reported calculations exercise this production path rather than a separate Python reference solver.

2.3. Evidence and Performance Metrics

The real-time factor R_RT is

$$R_{RT} = \frac{t_{sim}}{t_{wall}} \tag{3}$$

Throughput is reported in billions of lattice-site updates per second (GLUPS). One fluid-plus-thermal grid step is counted once as a coupled update; counting the two lattices separately gives twice that value. Thermal, periodic turbulent, laminar boundary-condition, component-surrogate, and reactor-geometry evidence are interpreted separately rather than combined into one validation claim.

3. Benchmark Validation

3.1. Fixed-Flux Natural-Convection Cavity

The thermal path was assessed in a cavity with uniformly heated and cooled vertical sides, adiabatic horizontal walls, Rayleigh number Ra=10^7, and Prandtl number Pr=0.7. Equal and opposite wall powers were represented in adjacent fluid cells. Published finite-volume profiles and an analytical infinite-slot solution supplied comparison data [7].

In 64-bit calculations near nondimensional time 0.2, refinement from N=32 to N=128 reduced the temperature-profile L2 error from 0.00214952 to 0.00072836. At N=128, the computed Nusselt number was 12.50465776 versus the analytical value 12.22606424, a +2.28% difference. Stratification and mid-height maximum velocity differed by -2.28% and +5.95%, respectively. Mean enthalpy drift was -1.80×10^-11 J/m3, and recorded spanwise temperature and velocity variations were zero. The compressed 16-bit-storage path lost accuracy with refinement and long runtime, so it is not treated as evidence equivalent to the 64-bit result.

3.2. Taylor--Green Bulk-Flow LES Evidence

The D3Q27 cumulant path was tested on the triply periodic Taylor--Green vortex at Reynolds number 1600 and initial Mach number 0.05 using the HiOCFD reference data [8]. Forty production calculations covered grids from 64^3 to 160^3 and ten Smagorinsky coefficients. The predefined scoring metric selected C_s=0, after which 216^3 and 264^3 confirmation calculations were completed.

Over nondimensional times 1 through 10, the 264^3 result had normalized root-mean-square errors of 0.435% for kinetic energy and 0.783% for total energy decay. The reference decay peak occurred at t^*=8.97 and the LBM peak at t^*=8.92. A decay history formed from resolved molecular dissipation alone had 7.521% normalized error, while the measured total-decay error was 0.783%. The non-molecular decay residual was positive at 94.12% of comparison times and decreased with refinement.

The observation is consistent with a grid-dependent numerical energy sink and supports an implicit-LES-like interpretation for this periodic bulk-flow case [9]. The result does not uniquely identify the numerical dissipation mechanism. It also does not validate wall turbulence, a positive-C_s model, heated-scalar transport, or transfer to MSFR Reynolds numbers.

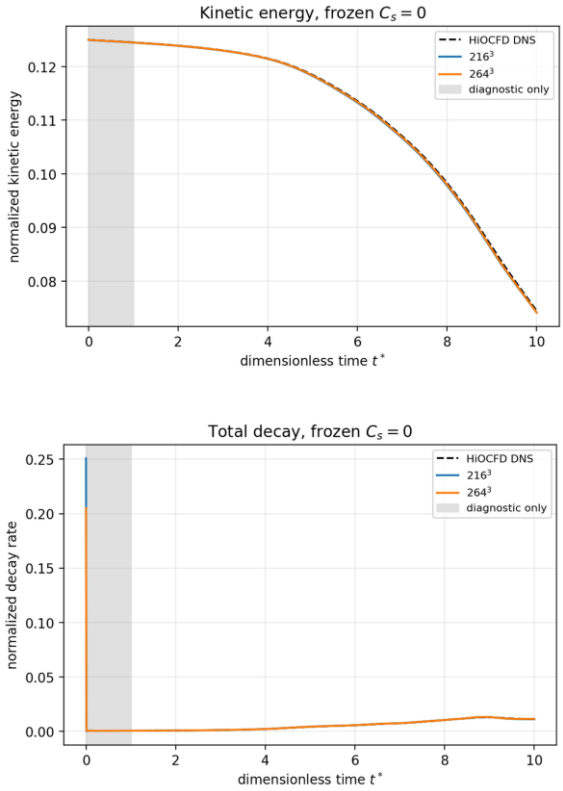


Fig. 1. Taylor--Green confirmation histories at Re=1600 with frozen Cs=0: (a) normalized kinetic energy and (b) total decay for the 216^3 and 264^3 grids against HiOCFD DNS. The shaded t*<1 interval is diagnostic only.

3.3. CNRS Step 0.1 and Step 0.3

Steps 0.1 and 0.3 of the CNRS MSFR numerical benchmark reported by Tiberga et al. [2] CNRS Step 0.1 is a two-dimensional 2 m by 2 m lid-driven cavity at Reynolds number 40 with a lid speed of 0.5 m/s [2]. Five FP32 D3Q27 grids from N=16 to N=64 reached the velocity-convergence and spanwise-invariance criteria. From N=48 to N=64, the full-field velocity change had a mean absolute value of 3.806×10^-4 m/s and a root-mean-square value of 2.181×10^-3 m/s. Extending the accepted N=16 field from 23.4 s to 64.0 s changed the full field by 3.298×10^-5 m/s.

Step 0.3 held the prescribed velocity field obtained from Step 0.1 fixed and advanced the FP64 D3Q15 thermal solver with the prescribed 1 GW volumetric source field, linear heat removal -10^6(T-900\ K) W/m3, and adiabatic half-way walls. All five grids met the frozen temperature and heating-removal criteria. Against the the TUD-S2 participant result reported in the benchmark [2], centreline mean absolute error decreased from 20.323 K to 3.933 K on AA' and from 27.671 K to 7.569 K on BB' between N=16 and N=64. At N=64, the final mean temperature was 1150.000000 K and normalized heating-removal imbalance was 3.392×10^-13.

A source-free 200-step diagnostic gave relative scalar drift of approximately 1.02×10^-5 in FP32 and 3.7×10^-14 in FP64. These results validate the implemented frozen-flow thermal benchmark within the stated numerical configuration; they are code-to-code evidence, not experimental MSFR validation.

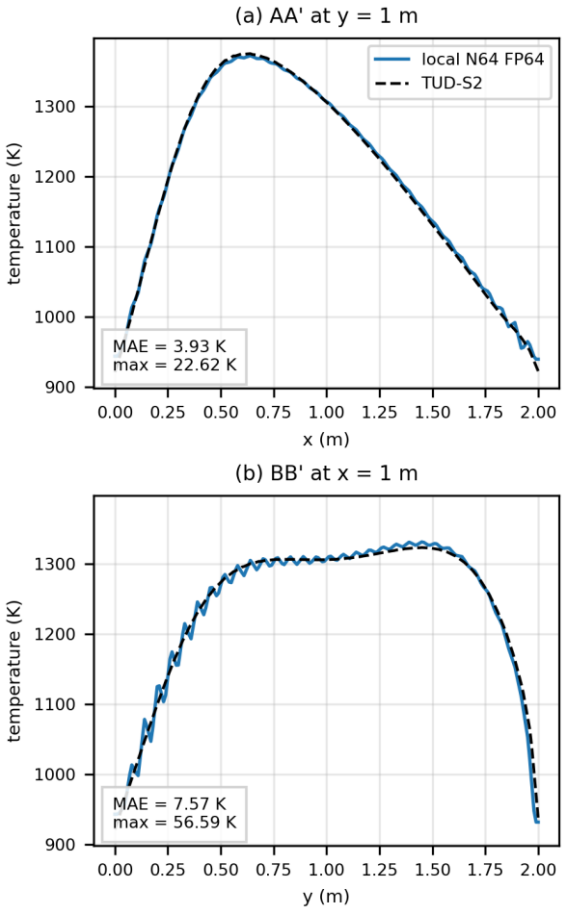


Fig. 2. CNRS Step 0.3 FP64 N=64 temperature profiles on (a) AA' and (b) BB' versus TUD-S2. Mean absolute errors are 3.93 K and 7.57 K, respectively.

4. MSFR-Scale Demonstration

4.1. Pump, Heat Exchanger, and Performance Controls

A three-dimensional square-loop surrogate first exercised separated pump, volumetric-heating, and heat-exchanger regions with D3Q27 fluid and D3Q15 thermal LBM. Its 24,000-step field sequence produced coherent circulation, 0.07884% section-flow spread, and finite velocity, gauge-pressure, and temperature fields. It is retained as a component-integration visualization, not as a calibrated MSFR loop or a performance benchmark.

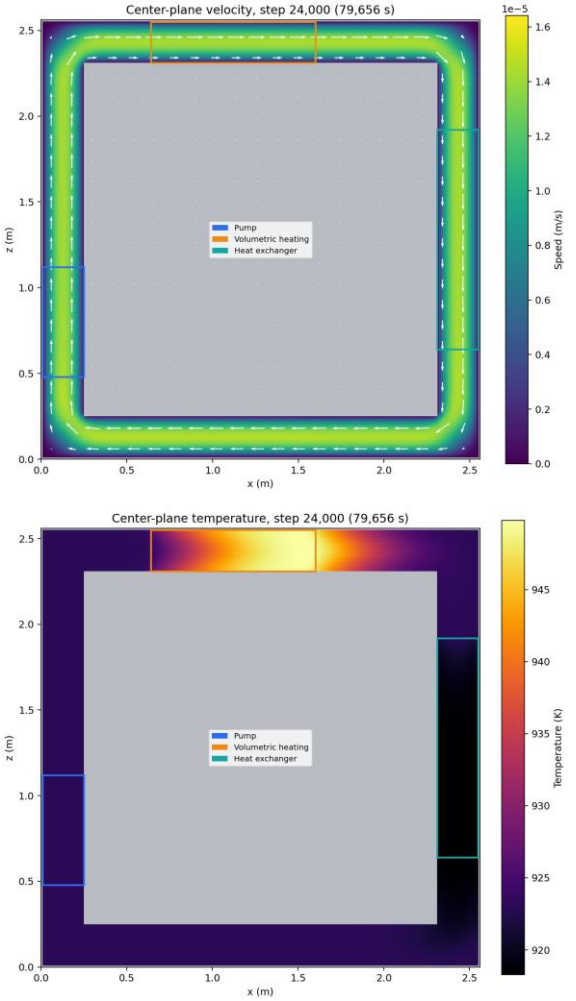


Fig. 3. Square-loop component surrogate at step 24,000: (a) velocity and (b) temperature. Colored outlines identify the pump, volumetric-heating, and heat-exchanger regions.

The reference performance case used a rounded-loop control on a 256×26×256 grid at 0.01 m spacing. It contained 494,304 fluid cells, 362,320 active D3Q27 BFL links, and 197,616 active D3Q15 BFL links. The accepted FP32 calculation used one NVIDIA GeForce RTX 3090, a 27 kPa pump input, uniform initial temperature, no volumetric heating, no heat removal, and no explicit drag. After 150,000 hydraulic preconditioning steps, 49,120 coupled steps represented 5.011052 s of physical time. Mean flow was 505.321 kg/s with 0.1672% spread among monitored sections.

Measured steady throughput was 2.661940 coupled GLUPS, or 5.323880 GLUPS when the two lattices were counted separately. Compile-inclusive coupled throughput was 2.644396 GLUPS. The measured and derived comparison is summarized in Table I. Both real-time factors are below one; the calculation was not real time.

Table I. Rounded-loop performance on one RTX 3090 for 5.011052 s of simulated time. Throughput and wall times were measured. Real-time factors and gaps were derived using R_RT = t_sim / t_wall.

Metric	Timing scope	Value
Coupled throughput	Steady kernel	2.66 GLUPS
Coupled throughput	Compile-inclusive	2.64 GLUPS
Wall time	Kernel interval	31.650832 s
Wall time	End to end	177.028944 s
Real-time factor	Kernel interval	0.1583 (6.32x gap)
Real-time factor	End to end	0.0283 (35.3x gap)

A separate rounded-loop source/sink control exercised volumetric heating and heat-exchanger removal on the same coupled LBM production path. This control is used only as component-integration and energy-accounting evidence; the reactor-scale thermal discussion is based on the completed quarter-domain calculation.

4.2. Quarter-Domain Setup and Uniform Control

The approximate Rouch Geometry-II quarter-domain model contains (160×160×192) cells at 0.015625 m spacing, corresponding to 4,915,200 cells. It was generated from analytic surfaces, and the discrete volumetric heat source was normalized to 698.975 MW over the quarter domain. The model combines published Rouch geometry and operating data with explicitly documented auxiliary assumptions; it is not an exact reconstruction of unavailable Geometry-II CAD data.

The production calculation used FP32 storage, weakly compressible D3Q27 cumulant collision with Smagorinsky LES modeled flow with C_s=0, D3Q15 thermal LBM, physical curved-wall BFL links, and half-link specular symmetry. The D3Q27 topology contained 951,848 physical BFL and 194,348 specular links. The D3Q15 topology contained 509,870 physical BFL and 107,802 specular links. Pump pressure rise was 902 kPa.

The uniform-temperature, zero-source, zero-heat-removal control completed 16,000 steps, or 3.333333 s. Quarter flow was 4734.076 kg/s, +0.0217% from the 4733.05 kg/s target. Temperature remained within 948.149963--948.150146 K, and the average domain-integrated thermal-balance residual was +6.98 W. This control supports the final run's hydraulic mapping and preservation of a uniform temperature field under the same quarter topology.

4.3. Heated Quarter-Domain Result

The heated and heat-removal calculation completed the same 16,000 steps without numerical divergence. Quarter flow was 4766.239 kg/s, +0.701% from the target. The final temperature range was 868.320--1170.437 K. Core inlet and outlet temperatures were 910.493 K and 999.540 K, a rise of 89.047 K. Heat exchanger inlet and outlet temperatures were 1012.416 K and 912.249 K, a drop of 100.167 K. Final heat removal was 739.239 MW from the volumetric sink and 716.339 MW from the section enthalpy difference.

With half-link specular symmetry, the symmetry-layer mean absolute vertical speed was 2.061 m/s, and the hottest cell was located within the fluid domain at (16,50,142), where the temperature was 1170.437 K. The final field therefore retained active tangential transport along the symmetry planes and placed the reported maximum away from the symmetry boundaries.

The final run-integrated thermal-balance residual was -251.998 MJ, or -10.82% of generated energy. This residual is the principal remaining numerical limitation of the quarter-domain thermal result and is associated with the current curved-wall thermal treatment. The quarter calculation consequently demonstrates production-path feasibility and MSFR-scale field generation, but it does not establish closed thermal conservation or validate peak temperature, turbulent heat flux, wall shear, pressure loss, or a reactor safety margin.

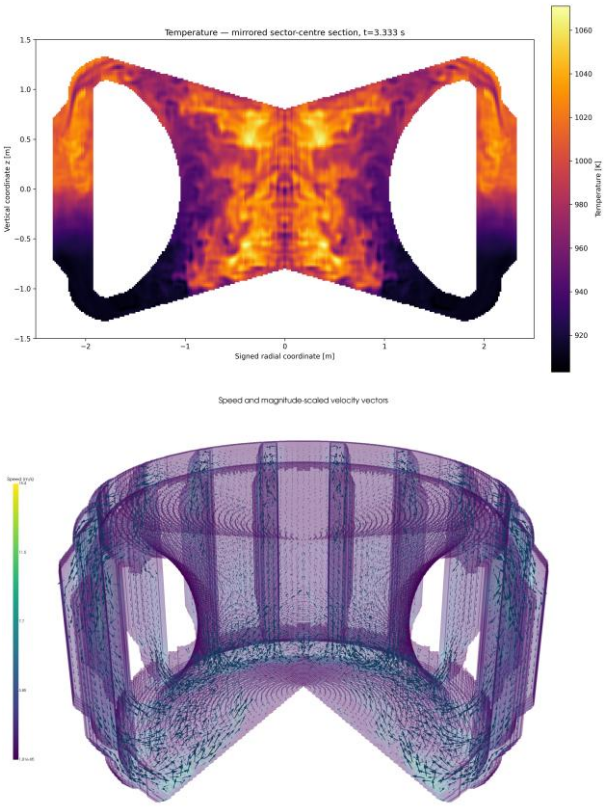


Fig. 4. Final mirrored sector-centre fields after 3.333333 s in the completed half-link specular-symmetry quarter calculation: (a) temperature and (b) speed. Mirroring is a display operation; only one quarter was solved.

5. Discussion

5.1. Supported Interpretation

The measured evidence supports four separate conclusions. First, the D3Q15 thermal LBM

implementation reproduces the fixed-flux cavity refinement trend and the completed CNRS frozen-flow thermal benchmark in FP64. Second, the D3Q27 cumulant path with C_s=0 reproduces the selected wall-free Taylor--Green energy and decay histories with sub-percent normalized errors at 264^3. Third, the rounded-loop control establishes high single-GPU coupled throughput but remains below real time. Fourth, the quarter result demonstrates that the coupled GPU solver can advance an MSFR-scale, curved-wall, heated and cooled domain for 3.333333 s while producing three-dimensional fields and target-scale flow.

These conclusions do not combine into a validated MSFR LES. The Taylor--Green case is isothermal, periodic, and wall free. CNRS Step 0.3 holds a laminar velocity field fixed. The pump/heat-exchanger visual surrogate and the rounded-loop performance control have different physical purposes. The quarter geometry is approximate, uses C_s=0, lacks a validated turbulent thermal diffusivity and wall heat-transfer model, and retains a -10.82% thermal-energy residual.

5.2. Path to Higher Fidelity and Real-Time Execution

The next development stage targets higher physical fidelity while preserving the local GPU execution structure. Priorities are a conservative curved-wall thermal treatment; pressure-based incompressible-LBM collision variants where they provide a verified accuracy advantage; wall-aware subgrid-scale closures, including the wall-adapting local eddy-viscosity (WALE) model [11]; and a separately validated turbulent thermal-diffusivity or heat-flux closure. These methods must be assessed in wall-bounded-flow and heated-scalar benchmarks before they are used for MSFR wall shear, pressure loss, turbulent heat flux, or peak-temperature claims.

Because LBM collision and many closure evaluations are cell-local or compact-stencil operations, these refinements are structurally compatible with GPU execution. In a separate matched rounded-loop control, the weakly compressible and pressure-based collision variants sustained 2.154 and 2.162 coupled GLUPS, respectively. This indicative single-run comparison was not based on a single run and not accompanied by repeatability statistics and did not improve thermal conservation, but it shows that at least some collision-level physics changes can be introduced without a large throughput penalty. For WALE, wall models, and thermal subgrid-scale transport, limited overhead remains a hypothesis to be measured on the complete GPU solver workflow.

The only current real-time factors are derived from the rounded-loop performance control: 0.1583 for its compile-inclusive kernel interval and 0.0283 end to end. Obtained on a single consumer-class RTX 3090, the kernel result is within one order of magnitude of real time and quantifies the kernel-level speedup required for real time as 6.32-fold. Newer high-end accelerators and multi-GPU domain decomposition therefore provide credible engineering paths toward real time simulation. This is a scaling hypothesis rather than a measured multi-GPU result. A complete real-time claim still requires matched-accuracy measurement of the quarter-domain workload, including preconditioning, compilation, diagnostics, output, load balance, and inter-GPU communication.

6. Conclusion

On one RTX 3090, the rounded-loop control sustained 2.661940 coupled GLUPS and advanced 5.011052 s of physical time with a compile-inclusive kernel real-time factor of 0.1583, corresponding to a quantified 6.32-fold kernel gap. The quarter-domain calculation advanced 4,915,200 cells for 3.333333 s, matched the target flow within 0.0217% in the uniform control, and completed the heated case with a final temperature range of 868.320--1170.437 K.

The -10.82% quarter thermal-energy residual prevents a validated temperature or safety claim. Nevertheless, reaching a kernel real-time factor within one order of magnitude on a single consumer GPU provides a quantitative basis for pursuing newer high-end accelerators and multi-GPU execution. Conservative curved-wall thermal transport, incompressible or low-Mach collision formulations, wall-aware LES closures such as WALE, and turbulent thermal transport can be developed within the same GPU-oriented framework and may retain high throughput, but their accuracy and performance must be measured before they are promoted from future targets to demonstrated results. The present work therefore establishes benchmark-specific validation and an MSFR-scale coupled-solver demonstration while positioning higher-fidelity real-time analysis as a credible next milestone.

REFERENCES

[1] H. Rouch, O. Geoffroy, P. Rubiolo, A. Laureau, M. Brovchenko, D. Heuer, and E. Merle-Lucotte, "Preliminary thermal-hydraulic core design of the Molten Salt Fast Reactor (MSFR)," Annals of Nuclear Energy, Vol. 64, pp. 449-456, 2014. https://doi.org/10.1016/j.anucene.2013.09.012.
[2] M. Tiberga et al., "Results from a multi-physics numerical benchmark for codes dedicated to molten salt fast reactors," Annals of Nuclear Energy, Vol. 142, 107428, 2020. https://doi.org/10.1016/j.anucene.2020.107428.
[3] S. Chen and G. D. Doolen, "Lattice Boltzmann method for fluid flows," Annual Review of Fluid Mechanics, Vol. 30, pp. 329-364, 1998. https://doi.org/10.1146/annurev.fluid.30.1.329.
[4] M. Geier, M. Schönherr, A. Pasquali, and M. Krafczyk, "The cumulant lattice Boltzmann equation in three dimensions: Theory and validation," Computers & Mathematics with Applications, Vol. 70, No. 4, pp. 507-547, 2015. https://doi.org/10.1016/j.camwa.2015.05.001.
[5] R. Huang and H. Wu, "A modified multiple-relaxation-time lattice Boltzmann model for convection-diffusion

equation," Journal of Computational Physics, Vol. 274, pp. 50-63, 2014. https://doi.org/10.1016/j.jcp.2014.05.041.
[6] M. Bouzidi, M. Firdaouss, and P. Lallemand, "Momentum transfer of a Boltzmann-lattice fluid with boundaries," Physics of Fluids, Vol. 13, pp. 3452-3459, 2001. https://doi.org/10.1063/1.1399290.
[7] S. Jiracheewanun, G. D. McBain, S. W. Armfield, and M. Behnia, "Natural convection in the cavity with the differentially heated isoflux boundaries," ANZIAM Journal, Vol. 48, pp. C977-C990, 2008. https://doi.org/10.21914/anziamj.v48i0.113.
[8] HiOCFD5, "WS1 DNS of the Taylor-Green vortex at Re = 1600," 5th International Workshop on High-Order CFD Methods, accessed August 24, 2026. https://how5.cenaero.be/content/ws1-dns-taylor-green-vortex-re1600.
[9] T. Dairay, E. Lamballais, S. Laizet, and J. C. Vassilicos, "Numerical dissipation vs. subgrid-scale modelling for large eddy simulation," Journal of Computational Physics, Vol. 337, pp. 252-274, 2017. https://doi.org/10.1016/j.jcp.2017.02.035.
[10] M. Lehmann, "Esoteric Pull and Esoteric Push: Two simple in-place streaming schemes for the lattice Boltzmann method on GPUs," Computation, Vol. 10, No. 6, 92, 2022. https://doi.org/10.3390/computation10060092.
[11] F. Nicoud and F. Ducros, "Subgrid-scale stress modelling based on the
square of the velocity gradient tensor," Flow, Turbulence and Combustion,
Vol. 62, No. 3, pp. 183-200, 1999.
https://doi.org/10.1023/A:1009995426001.