

RECENT CAPABILITY AND PERFORMANCE ENHANCEMENTS OF THE WHOLE-CORE TRANSPORT CODE nTRACER

Namjae Choi¹, Hansol Park², Han Gyu Lee¹, Seungug Jae¹, Sori Jeon¹, Han Gyu Joo¹

¹Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea

²University of Michigan
2355 Bonisteel Boulevard, Ann Arbor, Michigan 48109, USA

njchoi94@gmail.com, hansolp@umich.edu, hglee531@snu.ac.kr, jsuk9417@snu.ac.kr,
sori226@snu.ac.kr, joohan@snu.ac.kr

ABSTRACT

The whole-core transport code nTRACER has made many advances in recent years. Several innovative cross section treatment methods were developed, a new axial transport solver was introduced for stabilizing the 2D/1D scheme, and substantial computational enhancements were achieved using NVIDIA CUDA and Intel Math Kernel Library (MKL). In addition, gamma transport solver was implemented to predict the power distributions more physically, and the flexibility of the restart calculation was improved using an offline processing code nTIG (nTRACER Input Generator). This paper is the compilation of the recent progresses in nTRACER developments.

KEYWORDS: nTRACER; cross section; stabilization; CUDA; gamma transport

1. INTRODUCTION

nTRACER [1] is a 2D/1D direct whole-core transport code being developed at Seoul National University. It claims to support “the practical numerical reactor,” and it has shown successes in various applications ranging from commercial reactors – OPR1000 [1], APR1400 [2], AP1000 [3], VERA [4], and BEAVRS [5] – to critical experiments – KRITZ and B&W [6] so far. Recently, nTRACER had further extended its capabilities to push the direct whole-core transport calculation towards a more feasible level. First of all, methods to generate multi-group cross sections were improved. It was demonstrated that using asymptotic scattering kernel of hydrogen in epithermal energy range and neglecting angle dependence of multi-group cross sections in resonance energy range introduce huge biases in spectra and thus in reaction rates. These problems were resolved by using free-gas scattering kernel of hydrogen in epithermal energy range and by using region-wise spectral SPH factors, respectively. Additionally, the methodology of the subgroup method was improved with the introduction of Macro Level Grid (MLG) scheme and Equivalent Dancoff factor Cell (EDC) method to significantly reduce the time consumed by the resonance treatment without loss of accuracy [7]. Second, an axial transport solution scheme using Method of Characteristics (MOC) was developed to enhance the stability of the 2D/1D solver. Third, computational enhancements were achieved with cutting-edge high performance computing (HPC) technologies like NVIDIA CUDA and Intel Math Kernel Library (MKL). Lastly, capabilities such as gamma transport and nTIG (nTRACER Input Generator) were developed to extend the applicability.

This paper compiles recent advances of nTRACER, some of which were or will be published as separate papers, in terms of three aspects: methodological improvements, computational enhancements, and other uncategorized capability extensions. The methods, capabilities, and algorithms will be briefly introduced, and the results will be presented comprehensively. The readers are encouraged to refer to the reference papers for more detailed information.

2. METHODOLOGICAL IMPROVEMENTS

This section describes the improved features in multi-group cross section processing and the 1D MOC axial solver. In the first subsection, improvements made to the multi-group library and to the subgroup method are introduced. In the second subsection, the 1D MOC axial solution scheme and the strategies to mitigate the 2D/1D instability while retaining accuracy are presented.

2.1. Multi-group Cross Section Processing [7]

There have been several improvements in multi-group cross section processing. Among the factors related to the multi-group cross sections, following two were shown to be of great significance on the accuracy of the neutron spectrum solution: the temperature dependence of hydrogen scattering kernel above the thermal energy cut-off (2 ~ 4 eV) and the angle dependence of multi-group cross sections in resolved energy range. nTRACER had neglected these factors and thus spectrum solutions were largely biased in a consistent manner for all pressured water reactor (PWR) problems under normal conditions.

Scattering kernel of hydrogen has a unique feature that the thermal agitation effect is not at all negligible even at high incident neutron energies. It is because a neutron at any energy can stop with a one head-on collision with hydrogen and thus the probability of down-scattering below the thermal average energy of $3/2 k_B T$ drastically decreases. Since NJOY is used to generate the multi-group library of nTRACER and NJOY uses temperature independent scattering kernel (asymptotic kernel) above a thermal energy cut-off in the THERMR module, neutrons having energies above the thermal cut-off are over-thermalized. This results in less capture reactions by the broad resonances of U-238 and plutonium isotopes near and above the thermal energy cut-off (~ 4 eV), which results in positive reactivity biases of about +70 to +250 pcm. This problem was resolved by using the hydrogen scattering matrix tallied from a Monte Carlo code that considers the hydrogen velocity sampled from Maxwell-Boltzmann distribution in all energy range [8].

In heterogeneous geometry for lattice calculation, spectrum in resonance energy range strongly depends on directions of neutrons at the periphery of fuel rod. Due to this different self-shielding condition for different directions, multi-group cross section at the periphery of a fuel rod becomes significantly angle-dependent in resonance energy groups. However, there is no practical way to use angle-dependent total cross sections in the multi-group neutron transport solver. It was demonstrated that the neglect of this angle dependency results in large condensation error (reactivity error of -200 to -300 pcm) even though exact region-wise effective cross sections are used. Ref. [9] demonstrated that the spectral SPH factor can resolve this problem and that region-wise SPH factors can be functionalized over pin-cell parameters. This was supported by the observation that more than 95% of the condensation error comes from the intra-cell effect. Thus, we constructed a library of region-wise spectral SPH factors functionalized over various parameters of the PWR pin-cell types. In the application, region-wise spectral SPH factors are obtained for each pin-cell type by table look-up and applied to modify the multi-group transport operator to have the correct leakage into the fuel rod from the moderator region.

The subgroup method used to determine effective cross section in the resolved energy range was also optimized. In order to obtain escape cross sections for each resonant isotope in all regions, a fixed source problem (FSP) with the intermediate resonance (IR) approximation should be solved for each subgroup level of each isotope. These FSPs are called the subgroup FSPs (SGFSPs). The conventional way to

reduce the computational cost of the SGFSP is the categorization approach. It categorizes isotopes which have similar resonances within each group, sets a representative isotope for each category, and solves SGFSPs with reduced number (usually four) of subgroup levels of the representative isotope. Then, the escape cross section of the representative isotope obtained at the reduced number of subgroup levels is converted and interpolated to the escape cross sections of the isotopes of interest at exact subgroup levels (mostly seven levels). However, it was shown that errors in effective cross sections induced from the interpolation errors of escape cross sections result in consistent reactivity errors of about +60 to +150 pcm and that the categorization scheme fails for regions having similar amounts of different resonant isotopes with overlapped resonances like the AIC control rod regions.

In order to resolve this problem, the MLG scheme was developed [10]. This scheme determines the optimum macroscopic subgroup levels that minimize the error of the pellet averaged effective cross section induced from the interpolation error of escape cross section. The escape cross section for each isotope is obtained by interpolation. It turned out that eight levels are the optimum number for cylindrical fuel rod, and the performance of this scheme was shown to be much superior to that of the conventional approach in accuracy and computational time. The pellet averaged effective cross sections obtained by the MLG scheme agreed very well with those obtained by the conventional scheme without categorization and interpolation of escape cross sections. On top of that, the number of the SGFSP calculations is fixed regardless of the number of resonance isotopes in a core and is even smaller than that required for the conventional scheme. In order to reduce the subgroup calculation time further, the EDC method [11] was adopted to the subgroup method. Each two-dimensional pin-cell is approximated by one dimensional pin-cell of which dimension is determined by preserving the Dancoff factor obtained by the Enhanced Neutron Current Method [12]. This scheme significantly reduced the computing time without noticeable loss of accuracy.

2.2. Augmented Axial MOC [13]

The numerical instability of the 2D/1D method has long been an issue. There are several causes of the instabilities: 1. transverse leakage causing the total source to become negative, 2. negative self-scattering of hydrogen due to transport correction, and 3. unphysical flux shapes induced by polynomial expansion in axial nodal kernel. All the causes lead to the occurrence of negative fluxes, which spoils the CMFD relations. And often the axial solver is the origin of instability.

Therefore, after elaborate investigations of the causes of the instability, we developed an axial transport solver that suppresses the occurrence of negative fluxes while minimizing the impact on the solutions. It employs flat source method of characteristics augmented by several techniques such as subgrid, flux form function, leakage splitting, P_L scattering expansion, and limited transport correction.

The axial MOC solver works on homogenized pin basis. Coarse CMFD meshes are imbedded by subgrids on which the axial MOC sweeps are performed. The flux distribution in the subgrid is reconstructed using current coarse mesh solution and flux form function evaluated at previous axial sweep. The fine mesh fission source distribution is then evaluated using the fine mesh flux and the coarse mesh homogenized fission cross section, and fixed source iteration is performed. The transverse leakage is isotropized and expanded with quadratic function, as is done in conventional nodal methods. The expanded function is then piece-wise integrated in each fine mesh. **Figure 1** illustrates the subgrid scheme and the treatment of transverse leakage in the axial MOC solver.

Several schemes are augmented to prevent the occurrence of negative flux. To minimize the impact on the global solution, the schemes are selectively applied depending on the characteristics of the regions. First, to prevent the source term to become negative by the transverse leakage, *isotropic leakage splitting* is

employed in non-fuel regions; in fuel regions, there are often enough fission sources to make total source positive, while it is not true for non-fuel regions:

$$\begin{cases} \frac{d\phi_g}{ds} + \left(\Sigma_{tr,g} + \frac{L_g}{\phi_g} \right) \phi_g = Q_g & (L_g > 0) \\ \frac{d\phi_g}{ds} + \Sigma_{tr,g} \phi_g = Q_g - L_g & (L_g \leq 0) \end{cases} \quad (1)$$

Second, in homogeneous non-fuel cells, especially in water blocks, P_L scattering expansion is directly applied to avoid negative self-scattering cross section that occurs in hydrogen during transport correction. However, for heterogeneous non-fuel cells in which water is dominant such as shroud and empty guide tube cells, the P_L method cannot be applied in that homogenization of high-order scattering matrices is not possible, while the negative self-scattering terms still persist. In such cells, *limited transport correction* scheme is employed, which migrates the negative self-scattering term to the left hand side:

$$\begin{cases} \frac{d\phi_g}{ds} + (\Sigma_{tr} - \Sigma_{gg})\phi_g = Q_{f,g} + \frac{1}{2} \sum_{g' \neq g} \Sigma_{g'g} \phi_{g'} & (\Sigma_{gg} < 0) \\ \frac{d\phi_g}{ds} + \Sigma_{tr} \phi_g = Q_{f,g} + \frac{1}{2} \sum_{g'} \Sigma_{g'g} \phi_{g'} & (\Sigma_{gg} \geq 0) \end{cases} \quad (2)$$

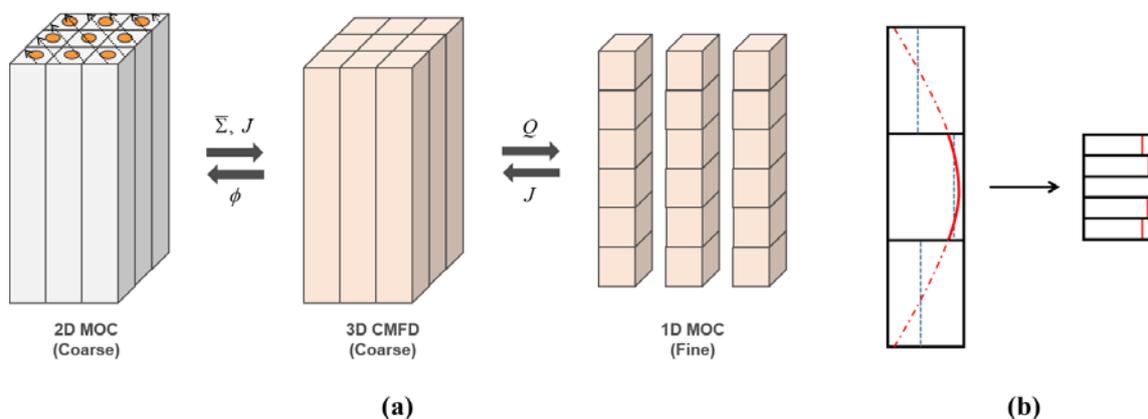


Figure 1. (a) Illustration of subgrid scheme, and (b) treatment of transverse leakage in fine meshes.

3. COMPUTATIONAL ENHANCEMENTS

This section mainly explains the GPU acceleration algorithms of nTRACER. Also, a brief explanation on how the Intel MKL library routines are utilized in nTRACER is given.

3.1. GPU Acceleration with CUDA [14]

The improvement of CPU processing power is facing limitations due to power constraints. However, the 2D/1D method is still not fast enough to be utilized in the industries. Thus, many codes are nowadays focusing on developing massive parallelization techniques employing thousands of CPU cores to achieve time-wise feasibility. However, the practicality of such approaches is questionable, because computing platforms with such size are not readily affordable. Thus, nTRACER has adopted an alternative approach

of using GPUs which are increasingly gaining attentions with the rise of artificial intelligence (AI) and big data computing.

Especially, we focus on the computing potential of cheap consumer-grade GPUs that are used for gaming. Even though the gaming GPUs lack double precision computing capability, they present significant single precision computing performances. Thus, to obtain maximal performance on the gaming GPUs, extensive use of single precision arithmetic is indispensable. It is widely admitted that the blinded use of double precision is unnecessary, and there have been many researches across a variety of fields on the use of mixed precision arithmetic in scientific computing. We also employ mixed precision techniques to exploit the gaming GPUs while preserving the accuracy.

Our target computing platform consists of less than a dozen of computing nodes, each equipped with a few gaming GPUs. Clusters with such size is expected to be easily affordable by academia and industries. Thus, we keep sticking to the plane-wise domain decomposition instead of shifting towards assembly-wise domain decomposition. Each GPU takes one or more planar MOC planes and communicates with neighbors when performing axially coupled calculations such as CMFD acceleration and axial sweep.

Ray tracing calculation is the most favorable calculation to GPUs. The potential of GPU acceleration in ray tracing calculation was first explored by researchers in MIT [15] and further extended by nTRACER [16] to anisotropic scattering treatments along with the consideration of CPU – GPU concurrency. To fully exploit the vectorization capability of GPUs, the ray tracing algorithm should be changed to a Jacobi style in terms of group sweep, as illustrated in **Algorithm 1**. Fine-grain thread-level parallelism is applied to group sweeps, and coarse-grain warp-level parallelism is applied to rotational ray sweeps, as illustrated in **Figure 2** where rotational rays are distinguished by colors. *Warp* is a computational unit in CUDA that consists of 32 work items. The efficiency is maximized when threads in each warp perform SIMD (Single Instruction Multiple Data) operations, and the Jacobi sweep algorithm is well-suited for the condition.

Algorithm 1. Parallel ray tracing algorithm.

```

FOR  $G$  DO Group Block Sweep
  FOR  $r$  PARALLEL DO Rotational Ray Sweep
    FOR  $s \in r$  DO Ray Segment Sweep
      FOR  $g \in G$  PARALLEL DO Group Sweep
        FOR  $p$  DO Polar Angle Sweep
          Calculate outgoing track angular flux
          Locally accumulate angular flux change
        END DO
        Atomically accumulate region scalar flux
      END PARALLEL DO
    END DO
  END PARALLEL DO
END DO
    
```

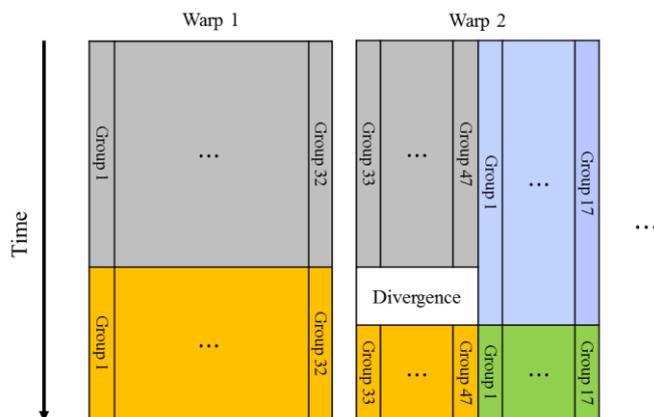


Figure 2. Ray and energy group threading scheme.

Note that the energy groups are divided into a number of blocks, and each block is solved sequentially. It is to have auxiliary CPUs interact with GPU through task parallelism. It is also a means to implement mixed precision arithmetic in the ray tracing calculation. While GPU is performing ray tracing calculation for a certain block with single precision arithmetic, CPUs proceed to subsequent blocks, calculate sources with double precision arithmetic, and copy them to the single precision buffers on GPU. The CPUs and GPU runs asynchronously to each other. **Figure 3** illustrates the concurrent execution diagram of CPUs and GPU. In this manner, auxiliary CPU computing power can be effectively utilized with additional benefits of preserving the accuracy and hiding the source calculation overhead.

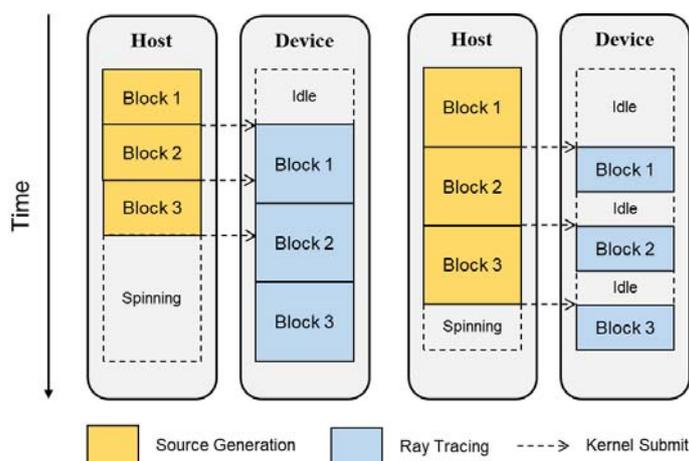


Figure 3. Concurrent execution diagram.

Mixed precision in CMFD power iteration is achieved by *iterative refinement* technique, as described in **Algorithm 2**, where the subscripts indicate the floating point precision in bytes. Parallel iterative linear system solvers cannot reduce errors below certain decimal level due to the randomly occurring round-off errors at the last significant digit, and for single precision arithmetic, the lower-bound is not sufficient. However, by incorporating the idea of iterative refinement technique into the fission source iteration framework, convergence level equivalent to full double precision arithmetic can be achieved with single precision iterative linear system solver, as illustrated in **Figure 4**. The linear system solver currently used in the GPU solver is BiCGSTAB with modified Sparse Approximate Inverse (SPAI) preconditioner that neglects the out-scattering terms for computational efficiency [17].

Algorithm 2. Power iteration with iterative refinement.

DO WHILE Not Converged

Compute total source: $\vec{Q}_{(8)}^{(n)} = \frac{1}{k^{(n)}} \chi_{(8)} \vec{\psi}_{(8)}^{(n)} + \mathbf{S}_{(8)} \vec{\phi}_{(8)}^{(n)}$

Compute residual: $\vec{r}_{(4)} = \vec{Q}_{(8)}^{(n)} - \mathbf{M}_{(8)} \vec{\phi}_{(8)}^{(n)}$

Solve (iteratively) for correction: $\mathbf{M}_{(4)} \vec{d}_{(4)} = \vec{r}_{(4)}$

Add correction: $\vec{\phi}_{(8)}^{(n+1)} = \vec{\phi}_{(8)}^{(n)} + \vec{d}_{(4)}$

Update fission source: $\vec{\psi}_{(8)}^{(n+1)} = \mathbf{F}_{(8)} \vec{\phi}_{(8)}^{(n+1)}$

Update eigenvalue

Check convergence

END DO

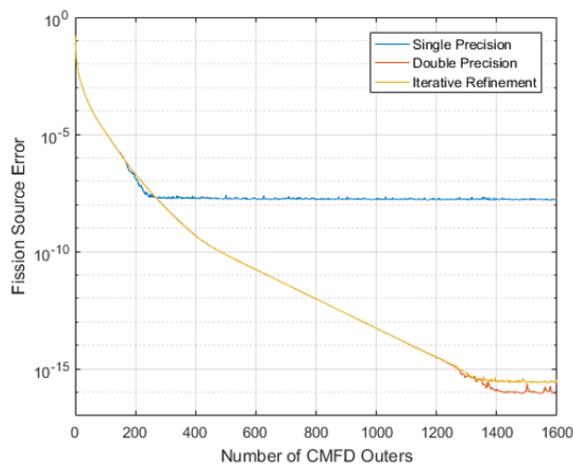


Figure 4. Power iteration residual behavior.

3.2. Application of Intel MKL

nTRACER is a legacy code with more than a decade of development period, and many parts of the code became obsolete. Hence, for conventional CPU applications, old built-in CMFD and depletion modules were rewritten such that they can utilize the linear algebra packages in Intel MKL extensively. In case of CMFD calculation, the BiCGSTAB kernel was rewritten using the Basic Linear Algebra Subprograms (BLAS) and the Sparse BLAS routines in MKL, and the matrices were reformatted to the Compressed Sparse Row (CSR) format. The Block Incomplete LU (BILU) preconditioner was also replaced by the parallel ILU preconditioner of MKL. Scattering matrices are stored in vector form and the source update

subroutine utilizes Vector Mathematical Library (VML) of MKL. In case of depletion calculation, Krylov subspace method is currently employed and the BLAS routines are intensively utilized for the Arnoldi process, scaling and squaring, and the Taylor expansion for computing the matrix exponential. The matrix of each region is rather small, so the parallelization is applied on the region loop and the single-threaded BLAS routines are called inside the parallel loop for increasing upper-level cache hit.

4. ADDITIONAL FEATURES

4.1. Gamma Transport [18]

To explicitly evaluate the gamma heating effect, nTRACER has recently implemented a gamma transport solver. It employs MOC as the high-order transport solver, and it is accelerated by fixed source CMFD, assuming that the CMFD formulation is valid for the photon physics. The multi-group photoatomic cross sections are generated from the GAMINR module of NJOY, and an 18-group structure is used. Following **Figure 5** shows the effect of gamma transport for VERA problem 5A-2D [19]. Fission power is the power calculated using the fission Q-value, which assumes that all the released energy is deposited on the spot.

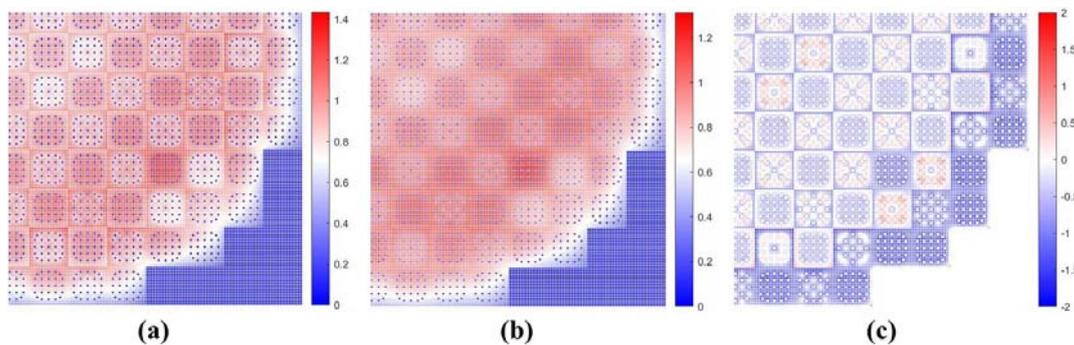


Figure 5. (a) Normalized fission power distribution, (b) normalized gamma heating power distribution, and (c) relative difference (%) between fission power and explicitly calculated power.

Currently, the gamma transport solver is one-way coupled with the neutronics solver. Once the neutron flux distributions are fully converged, the gamma transport solver starts fixed source iterations using the converged neutron flux as the source term. Consideration of T/H feedback remains as a future work.

4.2. nTRACER Input Generator (nTIG)

For improving the flexibility of restart calculation function in nTRACER, an external pre-processing code nTIG was developed. **Figure 6** schematically describes the working mechanism of the nTIG code under the interaction with nTRACER.

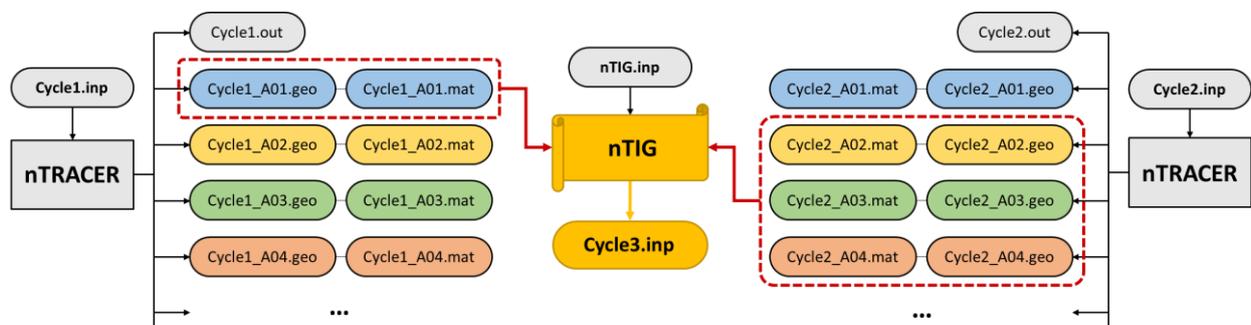


Figure 6. Schematic diagram of nTIG working mechanism.

Previously, nTRACER generated the restart file of each core state as a large monolithic binary file, which is difficult to maintain. It could be decoded only by nTRACER at runtime, so the manipulations of the core conditions were limited. However, in nTIG restart framework, nTRACER generates assembly-wise geometry and material information files. nTIG can assemble arbitrary assemblies of any core states using the information files to generate an nTRACER input file which describes a new core configuration. Since the restart file exists as a readable input file, adjusting ‘microscopic’ core conditions becomes especially easier. For example, manipulating a single pin, which can represent control rod movement, replacement of ruptured fuel, or inserting irradiation samples inside a burned fuel assembly, can be readily done by using nTIG functions or even with hand by manually modifying the information files of the target assembly or the final restart input file.

5. COMPREHENSIVE RESULTS

This section demonstrates 2D and 3D quarter-core calculation results in a comprehensive manner. The problem solved throughout this section is the APR1400 SKN3 initial core [2].

5.1. 2D Core Analysis

5.1.1. Effectiveness of advanced cross section treatments

Table 1 and **Figure 7** compare the computing time and accuracy of various combinations of cross section treatment methods for the 2D HZP case. ISO indicates the reference subgroup scheme which explicitly considers all the isotopes, and CAT is the conventional categorization scheme with 4-level approximation. MCH is the case in which the hydrogen scattering matrix is tallied from MC. The reference solution was obtained by the continuous energy Monte Carlo code PRAGMA [20] using 20 billion histories in total and the reference eigenvalue is 1.00294 with standard deviation of less than 1 pcm. A single computing node containing 20 cores of Xeon E5-2630 v4 CPU was used for calculation.

Table 1. Summary of core calculation results with various options.

Case	ISO	CAT	MLG	MLG + EDC	MLG + EDC + SPH	MLG + EDC + SPH + MCH
Subgroup Time (s)	2089.5	593.9	160.2	52.3	53.6	53.4
Total Time (s)	2911.5	1409.6	985.1	877.5	877.0	873.8
Reactivity Error (pcm)	-165	-130	-155	-115	121	38
Assembly Power RMS Error (%)	0.78	0.75	0.75	0.74	0.59	0.38
Assembly Power Max Error (%)	2.41	2.43	2.41	2.43	2.02	1.39

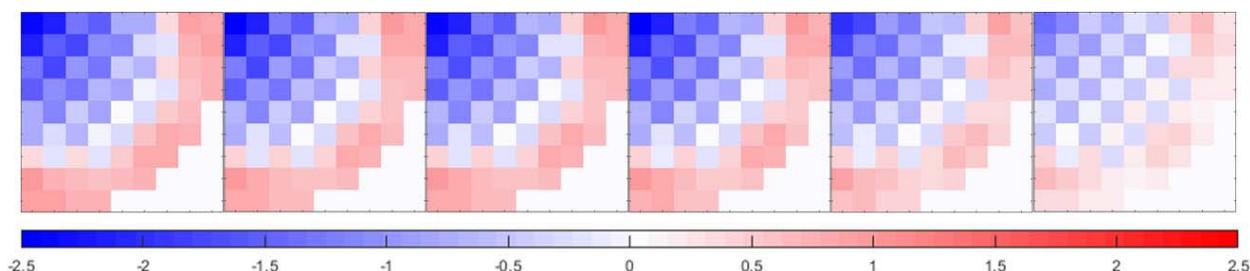


Figure 7. Assembly power relative error distributions of different cases.

It can be noticed that the MLG scheme yields almost identical result with the reference case, with more than tenfold reduction in subgroup calculation time. It can be further reduced to one-third by employing

the EDC scheme without noticeable loss of accuracy. In addition, it can be observed that the reactivity bias can be effectively captured without adding overhead with the application of the spectral SPH factors. It also improves the global power distribution to some degree.

5.1.2. Computational enhancements

Table 2 compares the computing time of the CPU parallel code and the GPU accelerated code spent for full convergence (relative l_2 -norm of the MOC fission source change being less than 5×10^{-7}). A single compute node with 20 cores of Xeon E5-2630 v4 CPU and one GeForce RTX 2080 Ti GPU was used. The EDC scheme has not been ported to GPU yet, so the comparison was made with the MLG scheme.

Substantial speedup is achieved with the use of a gaming GPU. Especially, the speedup in the ray tracing calculation is significant. The overall speedup factor, which includes time spent on CPU side, of the MLG subgroup calculation is around 14, while it reaches almost 70 as far as the ray tracing time is considered. In the same manner, the overall speedup ratio of the planar MOC is around 24, but the ray tracing speedup ratio will be much higher than that, though it cannot be directly measured due to the concurrency. The least efficient calculation appeared to be the CMFD acceleration. However, most of the time are spent for cell homogenization and linear system setup which are performed on CPU side, and the speedup of the power iteration is still meaningful. In the meantime, it is proved that the mixed precision strategy is valid; the degree of the pin power relative difference illustrated in **Figure 8** is completely negligible.

Table 2. Computing time comparison.

Calculation	CPU	CPU + GPU	Speedup
Subgroup (Ray Tracing)	165.0s (157.8s)	12.1s (2.3s)	13.6 (68.6)
MOC	1196.3s	49.8s	24.0
CMFD (Power Iteration)	46.2s (33.9s)	20.0s (3.0s)	2.3 (11.3)
Total	1407.5s	81.9s	17.2

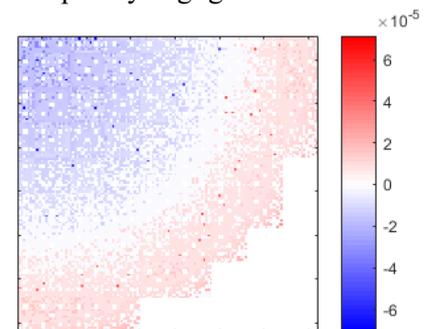


Figure 8. Pin power relative error.

Table 3 shows the performance of the renewed CMFD and depletion modules with Intel MKL. 2D HFP depletion calculation was carried out on a single compute node with 16 cores of Xeon E5-2640 v3 CPU. Note that the CMFD and depletion calculation time, which were originally under-optimal, are remarkably improved to have acceptable portions in the total core follow time. **Figure 9** illustrates the k_{eff} change with burnup and the difference of the two solvers, with the original solver as the reference. It can be noted that the renewed solvers can closely reproduce the result of original solvers, with subtle differences due to the changes in the iteration schemes and the treatments of the gadolinium residues.

Table 3. Computing time comparison.

Calculation	Built-in	MKL	Speedup
Subgroup	1:00:41	1:00:28	-
MOC	9:40:52	9:18:04	-
CMFD	2:28:02	0:19:09	7.7
Depletion	11:07:06	1:10:48	9.4
Other	1:38:19	1:25:10	-
Total	25:55:00	13:13:39	1.7

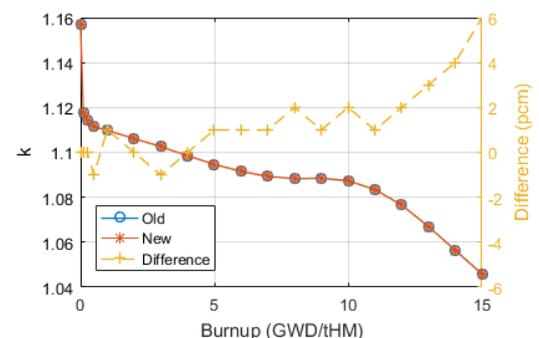


Figure 9. k_{eff} with burnup and difference.

5.2. 3D Core Analysis

5.2.1. Performance of the augmented axial MOC solver

3D HZP steady-state calculations were carried out using three different axial solution methods: two-node diffusion Source Expansion Nodal Method (SENM), whole-node SP₃ SENM, and axial MOC. Only a single case is presented here due to the limitation of space, and more verification results can be found in Ref. [13]. As can be seen in **Table 4** which summarizes the convergence results, the axial nodal kernels are only conditionally stable with the presence of damping while the axial MOC kernel does not require damping. Note that damping is an under-relaxation scheme that suppresses the drastic updates of the CMFD correction factors. Since the selection of the damping factor is rather artificial and case-dependent, the axial MOC solver, which provides methodological stability, is considered advantageous. Meanwhile, the accuracy of the axial MOC solver was comparable with that of the SP₃ SENM solver, as illustrated in **Figure 10** and summarized in **Table 5**. The reference solution was obtained from PRAGMA with total histories of 19.2 billion, which gives negligible uncertainties for the axial power.

Table 4. Summary of convergence status and k_{eff} .

Axial Solver	No Damping	Damping
Diffusion SENM	Div.	1.00090
SP ₃ SENM	Div.	1.00092
MOC	1.00090	1.00090

Table 5. Summary of axial power errors.

Axial Solver	Max Error (%)	RMS Error (%)
Diffusion SENM	2.42	0.30
SP ₃ SENM	1.94	0.18
MOC	2.33	0.11

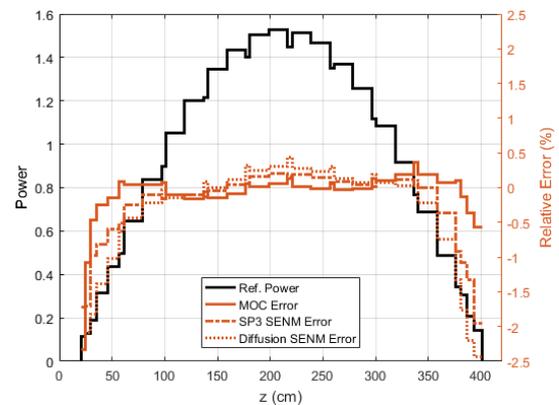


Figure 10. Axial power error distributions.

5.2.2. Computational enhancements

A 3D HZP steady-state calculation using the GPU acceleration module is presented. The 3D core consists of 36 axial planes. In-house *Soochiro* clusters were used for computing, whose specifications are listed in **Table 6**. For the CPU-based calculation, 18 compute nodes of Soochiro 3 were deployed, which results in total 288 CPU cores and 2 planes per node. For the GPU-accelerated calculation, 9 compute nodes of Soochiro 4 were deployed, which is equivalent to using 36 GPUs with each GPU solving one plane.

Table 6. Computing cluster specifications.

Cluster	Soochiro 3	Soochiro 4
# of Nodes Used	18	9
CPUs per Node	16 Cores of Intel Xeon E5-2640 v3	20 Cores of Intel Xeon E5-2630 v4
GPUs per Node	-	4 NVIDIA GeForce GTX 1080 (6 Nodes) 4 NVIDIA GeForce RTX 2080 Ti (3 Nodes)
Interconnect	Intel Omni-path Infiniband EDR (58Gbps)	Mellanox Infiniband FDR (56Gbps)

Table 7 compares the computing time of the CPU version and the GPU accelerated version. Axial MOC is used in both versions. As in the 2D core case, calculations employing MOC show substantial speedups, though the ratios had decreased. In the 2D core case, all 20 CPU cores were assigned to support a single GPU, while they are shared by four GPUs in this case. Also, note that the GPU model of the former and the latter nodes are different. Since nTRACER employs plane-wise decomposition, the computing time will be bound to the old GTX GPUs. To our experiments, the ray tracing performance of the RTX GPUs

was approximately three times faster than the GTX GPUs. The CMFD calculation being the least efficient has the same reason with the 2D case: homogenization and linear system setup overheads are dominant. In addition, the communication cost between GPUs is more expensive because additional latencies are introduced due to the data transfers through PCI-e channels. Still, the overall effectiveness of exploiting GPUs is significant, and the inefficiency of CMFD acceleration can be tolerated.

Table 7. Computing time comparison.

Calculation	CPU	CPU + GPU	Speedup
Subgroup	338.6s	23.6s	14.3
MOC	1672.7s	78.8s	21.2
CMFD (Power Iteration)	57.9s (46.8s)	44.6s (16.4s)	1.3 (2.9)
Axial	168.8s	23.7s	7.1
Total	2238.0s	170.7s	13.1

6. CONCLUSIONS

Recent advancements in the direct whole-core transport code nTRACER were briefly introduced and the newest results were demonstrated comprehensively. Introduction of SPH factors practically eliminated the reactivity biases caused by neglecting the angle dependence of multi-group cross sections. Paired with the improvement of hydrogen scattering kernel, nTRACER achieved less than 0.4% assembly power RMS difference with continuous energy Monte Carlo for a commercial core. MLG and EDC schemes enabled significantly faster subgroup calculations while preserving accuracy. The computational speed was further enhanced with the introduction of high performance computing techniques. Efficient GPU acceleration algorithms were developed using CUDA and a single gaming GPU was capable of giving the throughput equivalent to hundreds of CPU cores. By employing Intel MKL, outdated CMFD and depletion modules were rewritten to take advantage of efficient multithreading on CPU. Additionally, axial MOC solver was developed with augmentations to eliminate the negative flux in axial solution, and it showed enhanced stability and accuracy compared to the nodal kernels. Finally, gamma transport solver was implemented to enable explicit calculation of gamma heat, and a flexible restart input generator nTIG was developed.

Optimization of nTRACER is still ongoing. There are still many under-optimal routines, especially in the cross section treatments, which serve as global bottlenecks. Legacy data structures that are unfriendly in terms of computation should be restructured as well. Other than the ‘ground’ code optimization, GPU acceleration of sub-channel T/H and depletion calculations are being performed for a complete offloading of core follow calculations, and the transient calculation module is also being offloaded to GPUs to carry out the whole-core transient calculations in a feasible time span. In addition, investigations of an optimal energy group structure for nTRACER multi-group library are being performed to replace current HELIOS 47-group structure and improve the accuracy while retaining efficiency. Eventually, all the legacies of nTRACER will be removed and it will be reformed to a next generation direct whole-core transport code.

ACKNOWLEDGMENTS

This research is supported by National Research Foundation of Korea (NRF) Grant No. 2016M3C4A7952631 (Realization of Massive Parallel High Fidelity Virtual Reactor).

REFERENCES

1. Y. S. Jung, C. B. Shim, C. H. Lim, H. G. Joo, “Practical Numerical Reactor Employing Direct Whole Core Neutron Transport and Subchannel Thermal/Hydraulics Solvers,” *Annals of Nuclear Energy* **62**, pp. 357-374 (2013).

2. H. Hong and H. G. Joo, "Analysis of the APR1400 PWR Initial Core with the nTRACER Direct Whole Core Calculation Code and the McCARD Monte Carlo Code," *Transactions of the Korean Nuclear Society Spring Meeting*, Jeju, Korea, May 18-19 (2017).
3. H. Hong, Y. S. Ban, H. G. Joo, D. A. Salazar, F. Franceschini, "Analyses of AP1000® First Core with Direct Whole Calculation Codes," *Proceedings of the PHYSOR 2016*, Sun Valley, Idaho, USA, May 1-5 (2016).
4. M. Ryu, H. H. Cho, C. H. Lim, Y. S. Jung, H. G. Joo, "Preliminary Assessment of nTRACER and McCARD Direct Whole Core Transport Solutions to VERA Core Physics Benchmark Problems," *Transactions of the American Nuclear Society Winter Meeting*, Anaheim, California, USA, Nov. 9-13 (2014).
5. M. Ryu, Y. S. Jung, H. H. Cho, H. G. Joo, "Solution of the BEAVRS Benchmark Using the nTRACER Direct Whole Core Calculation Code," *Journal of Nuclear Science and Technology* **52**, pp. 961-969 (2014).
6. S. Jeon and H. G. Joo, "Analyses of the B&W-1810 and KRITZ-2 Critical Experiments with nTRACER," *Transactions of the Korean Nuclear Society Spring Meeting*, Jeju, Korea, May 17-18 (2018).
7. H. Park, "Resonance Treatment Innovations for Efficiency and Accuracy Enhancement in Direct Whole Core Calculations of Water-Cooled Power Reactors," *Ph.D. Thesis*, Seoul National University (2018).
8. H. Park and H. G. Joo, "The Impact of Hydrogen Asymptotic Scattering Approximations on Thermal Reactor Analysis in Deterministic Multigroup Approach," *Transactions of the American Nuclear Society Winter Meeting*, Orlando, Florida, USA, Nov. 11-18 (2018).
9. H. Park and H. G. Joo, "Practical Resolution of Angle Dependency of Multigroup Resonance Cross Sections Using Parametrized Spectral Superhomogenization Factors," *Nuclear Engineering and Technology* **49**(6), pp. 1287-1300 (2017).
10. H. Park and H. G. Joo, "Effective Subgroup Method Employing Macro Level Grid Optimization for LWR Applications," *Annals of Nuclear Energy* **129**, pp. 461-471 (2019).
11. K. Yamaji, H. Koike, Y. Kamiyama, K. Kirimura and S. Kosaka, "Ultrafine-group Resonance Treatment Using Equivalent Dancoff Factor Cell Method in Lattice Physics Code, GALAXY," *Journal of Nuclear Science and Technology* **55**, pp. 756-780 (2018).
12. A. Yamamoto, "Evaluation of Background Cross Section for Heterogeneous and Complicated Geometry by the Enhanced Neutron Current Method," *Journal of Nuclear Science and Technology* **45**, pp. 1287-1292 (2008).
13. N. Choi and H. G. Joo, "Stability Enhancement of Planar Transport Solution Based Whole-core Calculation Employing Augmented Axial Method of Characteristics," *Annals of Nuclear Energy*. (Submitted for publication.)
14. N. Choi, J. Kang, H. G. Joo, "Preliminary Performance Assessment of GPU Acceleration Module in nTRACER," *Transactions of the Korean Nuclear Society Autumn Meeting*, Yeosu, Korea, Oct 25-26 (2018).
15. W. Boyd, K. Smith, B. Forget, "A Massively Parallel Method of Characteristics Neutral Particle Transport Code for GPUs," *Proceedings of the M&C 2013*, Sun Valley, Idaho, USA, May 5-9 (2013).
16. N. Choi, J. Kang, H. G. Joo, "Massively Parallel Method of Characteristics Neutron Transport Calculation with Anisotropic Scattering Treatment on GPUs," *Proceedings of the HPC Asia 2018*, Tokyo, Japan, Jan. 29-31 (2018).
17. J. Kang and H. G. Joo, "GPU-based Parallel Krylov Linear System Solvers for CMFD Calculation in nTRACER," *Transactions of the Korean Nuclear Society Spring Meeting*, Jeju, Korea, May 17-18 (2018).
18. S. Jae, N. Choi, H. G. Joo, "Implementation and Verification of an Explicit Neutron/Photon Heating in nTRACER," *Proceedings of the ICAPP 2020*, Abu Dhabi, UAE, Mar. 15-19 (2020).
19. "VERA Core Physics Benchmark Progression Problem Specifications," CASL-U-2012-0131-004 (2014).
20. N. Choi, K. M. Kim, H. G. Joo, "Initial Development of PRAGMA – A GPU-Based Continuous Energy Monte Carlo Code for Practical Applications," *Transactions of the Korean Nuclear Society Autumn Meeting*, Goyang, Korea, Oct. 24-25 (2019).