



Introduction:

In this challenge sponsored by Advanced Micro Devices, Inc. (“AMD”), participants are invited to develop and optimize low-level kernels and deliver significant performance gains in large reasoning models, and to push the boundaries of inference performance on AMD Instinct™ GPUs. The challenge is open for registration beginning on **April 9th, 2025**, with submissions accepted beginning on **April 15th until June 8th, 2025**. Participants are eligible to win prizes up to \$100,000 and may be selected to travel to San Jose for a special awards ceremony. Attending the award ceremony on June 12th is not a requirement for winning.

The Challenge Overview: This challenge involves optimization and performance testing of the following key kernels:

- 1. FP8 GEMM
- 2. MLA with RoPE
- 3. Fused MOE

Kernel	Submission Window
FP8 GEMM	15-Apr to 27-May

MLA with ROPE	29-Apr to 27-May
Fused MOE	13-May to 27-May

The following reference information will be provided during the GPU Mode submission window:

- PyTorch reference implementation (baseline reference)
- Triton/AMD specific optimization reference (current best performance)
- Target input shapes and configurations
- Kernel level roofline performance (theoretical maximum)

Prizes

- **Grand Prize:** \$100,000
- **First Place:** \$25,000
- **Second Place:** \$15,000
- **Third Place:** \$10,000

In addition to a chance to win a prize, select participants may be invited to San Jose, California, USA, to participate in the AMD Advancing AI Day on June 12, 2025. Details will be provided to select participants in a reasonable amount of time before the award ceremony, including information regarding travel, hotel accommodations, and any applicable restrictions.

Important Note:

- Only kernels from registered participants within the registration deadline will be considered for scoring.
- Reference kernels and associated information will be provided at the start of the particular schedule.
- Teams or individuals deemed winners are eligible to receive one (1) prize only, to be awarded at AMD's sole discretion.

- Winning a prize depends on participants' ability to meet specific performance targets. If no team meets the performance targets, AMD reserves the right to refrain from awarding a prize.
- A participant or team's place on the leaderboard is not necessarily reflective of being determined to be a winner, nor is leaderboard placement a guarantee of winning a prize.
Four (4) winning submissions will be selected by AMD in its sole discretion after reproduction and verification of results.
- The challenge will be hosted on the **GPU Mode Discord server**, operating on their KernelBot automation platform to:•Execute kernel submissions, evaluate and debug performance and populate the results on the leaderboar
- The server also enables communication among participants, sharing learning resources, and viewing competition statistics.

Challenge Schedule

- **Pre-registration Opens:** April 9, 2025 | 12:00 PM PST
- **Registration Deadline:** April 30, 2025 | 11:59 PM PST
- **Kernel Submission Window Opens:** April 15, 2025 | 12:00 PM PST
- **Kernel Submission Deadline:** June 8, 2025 | 11:59 PM PST
- **Award Ceremony:** June 12, 2025 | Details to follow

Registration and Eligibility

- Registration is mandatory to participate.
- To register, fill out the registration form below. Registration is subject to verification and approval by AMD.
- The challenge is open to individuals and teams of up to three (3) members.
- All team members must register using their legal name and contact details and provide the same team name.
- Only one person per team should upload the code in the GPU Mode Discord channel.
- Participants who were part of the early registration process, please look out for further communication.
- Participants must be **18 years or older** or of the age of majority in their country as of April 9, 2025.
- Participants under 18 must present a **parent-signed document** if they win.

- All participants must: Have a valid **Discord ID** and a valid **GitHub ID**
- For questions, contact: **prithvim@amd.com**

Not eligible:

- Any national of Belarus, Burma, Cuba, Iran, North Korea, Russia, Syria, Sudan, Venezuela, and/or any country prohibited by law or export restrictions or otherwise, regardless of legal residency, will be disqualified. Residents of the regions of Crimea, Donetsk, and Luhansk are not eligible. Employees of Sponsor, its subsidiaries, affiliates and agents, and the immediate family (defined as parents, children, siblings, domestic partner, husband or wife) of each such employee, together with those with whom such employees are domiciled, are ineligible.

Kernel Submission

- All kernels must be original, developed specifically for this challenge, and must not have been displayed publicly before submission.
- Kernel submissions must be posted on Discord.
- Participants must submit all source code, scripts, and artifacts as specified in the challenge

Scoring and Judging Criteria

Kernel submissions for the challenge will be assessed based on the following criteria.

Submissions will be compared against their relative runtime/ speed and scores assigned according to the rubric below. For the Grand Prize, any kernel delivering the highest performance above the best reference we have and closest to the published roofline for the kernel shall win the Grand Prize. All prizes will be awarded only after the organizers have successfully reproduced the results independently

- KPIs for measurements:
 - Absolute speed of user-written kernels
 - Scoring is based on relative speedup over the base reference- Pytorch reference + relative speed over other user-written kernels
- Submissions will have Weighted scoring as below.
 - MLA with ROPE:
 - Max Points = 1500
 - Fused MOE:
 - Max Points = 1250
 - FP8 GEMM:
 - Max Points = 1000

The scoring metric for the leaderboard is the absolute runtime / speed of the participant's kernel averaged over a large set of test cases – we will provide shape information for these test cases, but the input data itself will be sampled from a random distribution.

Participant Score per kernel = Max Points * [1 - (rank / 10)]

- Only top ten performant kernels will be considered for scoring.
- Submissions who do not beat the Pytorch baseline will get zero points
- Geomean of 100 runs of the kernel rounded to a particular decimal value (will be specified along with the kernel problem release) will be considered for scoring
- In case of a tie, the earliest submitted kernel will be considered

Final Scoring: Competition rankings are determined by the highest number of total (weighted) points earned from all three kernels. Final rankings and scoring for 1st, 2nd, and 3rd place prizes will be determined by accepting one submission per team on each kernel (the fastest valid submission of each team will be taken) at the end of the competition. Each team will be considered as one participant in determining scores.

Grand Prize Criteria:

1. To qualify for the grand prize, the contestant has to beat the highest performance published reference- ALTER/ Triton reference kernel.
2. Winning team will be selected on the basis of $(\text{KERNEL_TIME} - \text{FASTEST_BASELINE}) / (\text{ROOFLINE} - \text{FASTEST_BASELINE})$
3. Roofline and baseline performance of each kernel will be published at the start of the submission window for each kernel at GPU Mode.
4. In case of a tie, following precedence will be followed:
 1. Highest performance for MLA with ROPE
 2. Highest performance for Fused MOE
 3. Highest performance for FP8 GEMM

If the tie is not resolved, then the earliest submitted kernel from among the tied kernels will be considered.

How are prizes awarded?- Coming Soon!

Code of Conduct

Entries may NOT contain ANY of the following content:

- Content that is sexually explicit, profane, or pornographic.
- Content that is unnecessarily violent or derogatory of any ethnic, racial, gender, sexual orientation, gender identity, religious, professional, or age group.

- Content that promotes illegal drugs, firearms/weapons (or the use of any of the foregoing) or a particular political agenda.
- Content that defames, misrepresents or contains disparaging remarks about any third-party, including individuals or organizations.
- Content that communicates messages or images inconsistent with the positive images and/or goodwill to which we wish to associate.
- Content that violates any federal, state, or local law.
- Harassment, discrimination, or inappropriate behavior will result in disqualification.
- AMD reserves the right to disqualify any participant or team at its sole discretion.

Legal

For all entries:

Any language or information included in a participant's registration or submission is deemed to be part of the participant's entry, and participant gives AMD, its designees, successors, assigns, and licensees a royalty-free, irrevocable, non-exclusive worldwide license to use, reproduce, modify, publish, create derivative works from, and display the entry and all elements embodied therein, along with the participant's name and/or social media account handle(s), in any manner, in whole or in part, on a worldwide basis, and to incorporate it into other works, in any form, media or technology now known or later developed, including for advertising, promotional, marketing and other purposes, without further payment or consideration, notification or permission. All Entries become the property of AMD, and none will be returned. If requested, participant will sign any documentation required for Sponsor or its designees, successors, assignees, and licensees to make use of the non-exclusive rights participant is granting to AMD. Released Parties (as defined below) are not responsible for lost, late, stolen, incomplete, inaccurate, invalid, un-intelligible, garbled, delayed, or misdirected posts, all of which will be void.

Release:

By participating, Participant agrees to release and hold harmless AMD, and each of its respective subsidiaries, affiliates, suppliers, distributors, advertising/promotion agencies, and each of their respective parent companies and each such company's officers, directors, employees and agents (collectively, the "Released Parties") from and against any claim or cause of action, including, but not limited to, damage to or loss of property, arising out of participation in the Challenge or receipt or use or misuse of any prize.

Privacy:

Participants acknowledge and understand that all personal information submitted as part of the challenge will be collected and processed by AMD for the purpose of managing the challenge in accordance with its Privacy Notice, Participant can read more about their rights, how AMD handles participants' personal information, and how to contact AMD in its Privacy Notice.

Publicity:

Except where prohibited by applicable law, participation in the developer challenge constitutes each winner's consent to AMD's use of the winner's name, city, state, province or county, and country, likeness, photograph, statements made by the winner about the challenge, about AMD, and/or prize information for the challenge in any media without further payment or consideration, including, but not limited to, posting winner lists online. All submissions become the property of AMD and none will be returned.

General Conditions:

Sponsor reserves the right to terminate, amend, suspend, or modify the challenge in whole or in part, at any time and without notice or obligation, if in AMD's sole discretion, any factor interferes with its proper conduct as contemplated herein. Without limiting the generality of the foregoing, if, for any reason, the challenge is not capable of running as planned, including infection by computer virus, bugs, tampering, unauthorized intervention, fraud, technical failures, or any other causes beyond the control of AMD, which corrupt or affect the administration, security, fairness, integrity or proper conduct of the challenge, AMD reserves the right, in its sole discretion, to disqualify any individual or team who tampers with the entry process. Any attempt by any person to deliberately undermine the legitimate operation of the challenge may be a violation of criminal and civil law, and should such an attempt be made, AMD reserves the right to fully seek damages from any such person permitted by law. AMD's failure to enforce any term of these rules shall not constitute a waiver of that provision or of any other provision of these rules. The invalidity or unenforceability of any provision of these rules shall not affect the validity or enforceability of any other provision. If any provision of the rules is determined to be invalid or otherwise unenforceable, then the rules shall be construed in accordance with the terms as if the invalid or unenforceable provision was not contained therein.

Limitations of Liability:

The Released Parties are not responsible, to the extent permitted by law, for: (1) any incorrect or inaccurate information, whether caused by participant, printing errors, or omission or by any of the equipment or programming associated with or utilized in the challenge; (2) technical failures of any kind, including, but not limited to malfunctions, interruptions, or disconnections in phone lines or network hardware or software; (3)

unauthorized human intervention in any part of the entry process or the challenge; (4) technical or human error which may occur in the administration of the challenge or the processing of entries; or (5) any injury or damage to person or property which may be caused, directly or indirectly, in whole or in part, from participation in the challenge or receipt or use or misuse of any prize. If for any reason an entry is confirmed to have been erroneously deleted, lost, or otherwise destroyed or corrupted, participant or the team's sole remedy is another entry in the challenge, provided that if it is not possible to submit another entry due to discontinuance of the challenge, or any part of it, for any reason, AMD, in its sole discretion, may elect to hold a random drawing from among all eligible entries or, as the case may be, from among eligible entries received up to the date of discontinuance for any or all of the prizes offered herein. No more than the stated number of prizes will be awarded.

NOTHING IN THESE RULES SHALL DISCLAIM, LIMIT, OR EXCLUDE LIABILITY FOR ANY LIABILITY THAT MAY NOT BE DISCLAIMED, LIMITED, OR EXCLUDED PURSUANT TO APPLICABLE LAW.

Applies to participants residing in Germany only:

AMD will be liable for any culpable breach of essential contractual obligations. Essential contractual obligations are contractual obligations that need to be fulfilled to permit proper execution of these Rules and that may regularly be relied upon by the participant. AMD's liability will otherwise be limited to gross negligence and willful misconduct. In the event of any liability on the part of AMD due to a slightly negligent breach of essential contractual obligations or slightly negligent misconduct on the part of simple vicarious agents, such as the Administrator, the AMD's and the Administrator's respective subsidiaries, affiliates, suppliers, distributors, advertising/promotion agencies, and prize suppliers, and each of their respective parent companies and each such company's officers, directors, employees and agents, AMD's liability will be limited to typically foreseeable damages. The above limitations of liability will not affect any mandatory statutory liability, in particular AMD's liability in connection with the loss of life, bodily injury or illness.

Applies to participants residing in the United Kingdom only:

NOTWITHSTANDING THIS SECTION (LIMITATION OF LIABILITY), NOTHING IN THESE RULES SHALL BE CONSTRUED TO LIMIT OR EXCLUDE ANY LIABILITY OF THE AMD FOR FRAUD, DEATH, OR PERSONAL INJURY CAUSED BY AMD OR PARTICIPANTS' NEGLIGENCE. No term herein is enforceable by any person who is not a party under the Contracts (Rights of Third Parties) Act 1999 or otherwise, excluding AMD.

Disputes:

Except where prohibited, participants agree that: (1) any and all disputes, claims and causes of action arising out of or connected with this challenge or any prize awarded shall be resolved individually, without resort to any form of class action, and exclusively by the United States District Court for the Western District of Texas or the appropriate Texas State Court located in Travis County, Texas; (2) any and all claims, judgments and awards shall be limited to actual out-of-pocket costs incurred, including costs associated with entering this challenge, but in no event attorneys' fees; and (3) under no circumstances will participant be permitted to obtain awards for, and participant hereby waives all rights to claim, indirect, punitive, incidental and consequential damages and any other damages, other than for actual out-of-pocket expenses, and any and all rights to have damages multiplied or otherwise increased.

FOR PARTICIPANTS RESIDING IN THE EUROPEAN UNION:

THE ABOVE CHOICE OF LAW MAY NOT RESULT IN DEPRIVING THE PARTICIPANTS OF THE PROTECTION UNDER MANDATORY STATUTORY PROVISIONS THAT CANNOT BE WAIVED UNDER THE LAW WHICH WOULD HAVE BEEN APPLICABLE IN THE ABSENCE OF THIS CHOICE OF LAW. FOR PARTICIPANTS NOT RESIDING IN THE EUROPEAN UNION, any and all claims, judgments and awards shall be limited to actual out-of-pocket costs incurred, including costs associated with entering this Challenge, but in no event attorneys' fees, and under no circumstances will Participant be permitted to obtain awards for, and Participant hereby waives all rights to claim, indirect, punitive, incidental, and consequential damages and any other damages, other than for actual out-of-pocket expenses, and any and all rights to have damages multiplied or otherwise increased. ANY DEMAND FOR OUT-OF-POCKET COMPENSATION MUST BE FILED WITHIN ONE (1) YEAR FROM THE END OF THE CHALLENGE PERIOD, OR IN ACCORDANCE WITH THE LAW OF LIMITATIONS AS APPLICABLE LOCALLY, OR THE CAUSE OF ACTION SHALL BE FOREVER BARRED. All issues and questions concerning the construction, validity, interpretation and enforceability of these Rules, or the rights and obligations of the Participant and AMD in connection with the Challenge, shall be governed by, and construed in accordance with, the laws of the State of Texas without giving effect to any choice of law or conflict of law rules (whether of the State of Texas or any other jurisdiction), which would cause the application of the laws of any jurisdiction other than the State of Texas. Some jurisdictions do not allow for limitations of certain remedies or damages and so this provision may not apply to you.

FOR UK RESIDENTS ONLY:

TO THE EXTENT PERMITTED BY LAW, ANY CLAIMS OR DISPUTES RELATING TO THIS CHALLENGE, THE PRIZE OR THESE RULES MUST BE BROUGHT WITHIN ONE (1) YEAR OF THE TIME THE CAUSE OF ACTION OCCURRED.

Winner's List:

The Winner's List will be available in English at Data Monsters's website and on the CSDn Website in Mainland China, after the close of the challenge. Requests for the winners list must be received within fourteen (14) days of the close of the challenge. Requests received after this time will not be honored.