

Zen AI Model Family

Zen4-Max

Fourth-Generation High-Performance Research Model

Technical Whitepaper v2026.01

Zach Kelling*
research@hanzo.ai

Zoo Labs Foundation
foundation@zoolabs.org

January 2026

Abstract

We present **Zen4-Max**, a 480B total parameter Mixture-of-Experts (MoE) model activating 48B parameters per token, optimized for research-grade scientific reasoning, mathematical proof verification, and complex long-horizon analysis. Zen4-Max achieves 93.4% on MMLU, 90.1% on MATH, 79.8% on GPQA Diamond, 94.7% on HumanEval, and 61.2% on FrontierMath, establishing it as the highest-performing model for formal reasoning tasks at the 48B active parameter inference cost. Through a novel **Proof-Augmented Training** (PAT) methodology that incorporates formal verification signals into the pretraining and post-training pipeline, Zen4-Max demonstrates qualitatively superior performance on tasks requiring logical completeness and self-consistency.

Contents

1	Introduction	3
1.1	Key Contributions	3
2	Architecture	4
2.1	Model Specifications	4
2.2	Expert Specialization for Research Domains	4
2.3	Proof-Augmented Training Architecture	4
2.3.1	Corpus-Level PAT	5
2.3.2	Reward-Level PAT	5
3	Training	5
3.1	Pretraining Data	5
3.2	Training Infrastructure	6
3.3	Post-Training Stages	6
4	Evaluation	6
4.1	Standard Benchmarks	6
4.2	Research and Formal Reasoning	7
4.3	Scientific Domain Evaluation	7
4.4	Long-Output Coherence	7

*zach@lux.network

5	Related Work	7
6	Deployment	8
6.1	Hardware Requirements	8
6.2	Research Workflow Integration	8
7	Safety and Alignment	9
7.1	Limitations	9
8	Conclusion	9
A	Model Card	10
B	PAT Ablation Study	10

1 Introduction

Scientific progress increasingly depends on AI systems capable of reliably assisting with tasks that require sustained, logically consistent reasoning: deriving mathematical proofs, designing experiments, interpreting empirical results, and synthesizing findings across hundreds of papers. These tasks share a common requirement that distinguishes them from most benchmarks: correctness cannot be approximated by fluency. A mathematical proof is either valid or it is not; a synthesized research hypothesis is either consistent with its supporting evidence or it contradicts it.

Zen4-Max is designed for this regime. Its primary architectural innovation, Proof-Augmented Training, introduces formal verification signals into both the pretraining corpus and the post-training reward function, teaching the model to distinguish correct chains of reasoning from plausible-but-incorrect ones. The result is measurable improvement on FrontierMath (61.2%), competition-level mathematics benchmarks, and scientific reasoning tasks that require multi-step argument validation.

Positioned between Zen4-Pro (72B dense) and Zen4-Ultra (1T MoE), Zen4-Max provides a practical inference cost (48B active parameters, achievable on 4 H100s) while approaching the capability ceiling of the Zen4 family on reasoning-intensive tasks.

1.1 Key Contributions

- **Proof-Augmented Training (PAT):** A training methodology incorporating formal verification signals (lean4, isabelle, and symbolic computation results) into the pretraining corpus and using verifier outputs as post-training rewards.
- **480B/48B MoE:** 480B total parameter MoE with 48B active, using 64 experts per layer with top-6 routing, optimized for research-domain specialization.
- **FrontierMath Performance:** 61.2% on FrontierMath, representing a significant advance on research-mathematics benchmarks that are resistant to pattern matching.
- **Long-Horizon Coherence:** Maintained logical consistency over outputs exceeding 8,000 tokens, enabling full proof derivations and experimental design documents.

2 Architecture

2.1 Model Specifications

Component	Specification
Total Parameters	480B
Active Parameters	48B per token
Architecture	Mixture of Experts (MoE)
Number of Layers	80
Dense Base Layers	24
MoE Upper Layers	56
Attention Mechanism	Hierarchical Hybrid Attention (HHA)
Local Window Size	16,384 tokens
Global Anchor Tokens	2,048 per layer
Context Length	131,072 tokens (128K)
Hidden Dimension	7,168
Experts per MoE Layer	64
Active Experts per Token	6 (top-6 routing)
Expert Intermediate Dim	2,816
Attention Heads	56
Key-Value Heads	8 (GQA)
Vocabulary Size	151,936
Positional Encoding	RoPE ($\theta = 10,000,000$)
Normalization	RMSNorm
Activation	SwiGLU

Table 1: Zen4-Max Architecture Specifications

2.2 Expert Specialization for Research Domains

Zen4-Max uses 64 experts per MoE layer organized into four research-domain clusters:

Expert Cluster	Count	Primary Data Sources
Mathematics and formal proof	16	ArXiv math, Lean4 libraries, OEIS
Physical sciences	14	ArXiv physics, NIST databases
Life sciences and chemistry	12	PubMed, ChEMBL, protein databases
Computer science and systems	14	ArXiv CS, GitHub, formal specs
General reasoning	8	All domains
Total	64	—

Table 2: Zen4-Max Expert Cluster Configuration

Top-6 routing across 64 experts activates 9.4% of expert parameters per token. The routing mechanism uses a sigmoid-normalized gating function rather than softmax, which reduces routing sensitivity to outlier logit values and improves stability on long contexts where token distributions shift significantly.

2.3 Proof-Augmented Training Architecture

PAT operates at two levels: the pretraining corpus and the post-training reward function.

2.3.1 Corpus-Level PAT

The pretraining corpus includes formal proof corpora that provide ground-truth logical structures:

- Mathlib4 (Lean4): 250,000+ theorems with complete formal proofs.
- Isabelle Archive of Formal Proofs: 720+ theory developments.
- HOL Light standard library: 12,000+ theorems.
- Symbolic computation results from SymPy and Mathematica (open datasets).

These corpora are presented in a unified format that interleaves the formal statement, the tactic proof steps, and the verified status of each step. The model learns to distinguish verified from unverified reasoning by observing this structure.

2.3.2 Reward-Level PAT

During post-training, mathematical generation tasks use a hybrid reward:

$$R_{\text{PAT}} = \alpha \cdot R_{\text{outcome}} + \beta \cdot R_{\text{verify}} + \gamma \cdot R_{\text{style}} \quad (1)$$

where R_{outcome} is the binary correctness of the final answer, R_{verify} is the output of a learned step-level verification model that scores each reasoning step’s logical validity, and R_{style} penalizes informal hedging language in mathematical contexts (phrases like “approximately” or “should be” in proof steps). Empirically, $\alpha = 0.6, \beta = 0.3, \gamma = 0.1$ produces optimal FrontierMath performance.

3 Training

3.1 Pretraining Data

Data Category	Proportion	Tokens
Web (quality-filtered)	30%	7.5T
Code and formal specifications	25%	6.25T
Scientific literature (ArXiv)	20%	5.0T
Mathematical content	12%	3.0T
Formal proofs (Lean4, Isabelle)	5%	1.25T
Tool-call trajectories	5%	1.25T
Books and long-form prose	3%	0.75T
Total	100%	25.0T

Table 3: Zen4-Max Pretraining Data Composition (25T tokens)

Formal proof data (5%) is over-represented relative to its natural occurrence in the web corpus by approximately $50\times$. This over-representation is critical for PAT: without sufficient formal proof exposure, the verification reward signal cannot be internalized.

3.2 Training Infrastructure

Parameter	Value
Hardware	$8,192 \times$ H100 80GB SXM
Expert Parallelism	64 (one expert per device)
Tensor Parallelism	TP=4
Pipeline Parallelism	PP=10
Data Parallelism	DP=32
Global Batch Size	12M tokens
Peak Learning Rate	1×10^{-4}
LR Schedule	Cosine with warmup (2500 steps)
Optimizer	AdamW ($\beta_1 = 0.9, \beta_2 = 0.95$)
Training Duration	52 days

Table 4: Zen4-Max Training Infrastructure

3.3 Post-Training Stages

1. **SFT** (5M pairs): Heavy emphasis on mathematical derivation and scientific analysis tasks. Includes 200K formal proof demonstrations in natural language augmented with lean4 verification certificates.
2. **PAT-GRPO**: Proof-Augmented GRPO over 80K mathematics tasks with SymPy-verified answers. $G = 16$ rollouts per task; the step-level verifier provides intermediate rewards at each reasoning step.
3. **DPO**: 900K comparison pairs, weighted 3:1 toward research tasks.
4. **Safety Alignment**: Standard constitutional AI pass.

4 Evaluation

4.1 Standard Benchmarks

Benchmark	Zen3-Max	Zen4-Pro	Zen4-Max	Δ vs Zen3
MMLU (5-shot)	87.3%	91.8%	93.4%	+6.1%
MATH (4-shot)	84.6%	87.6%	90.1%	+5.5%
HumanEval (0-shot)	89.7%	93.2%	94.7%	+5.0%
GPQA Diamond	72.1%	74.3%	79.8%	+7.7%
IFEval	84.3%	88.4%	89.1%	+4.8%
BBH (3-shot)	84.9%	89.7%	91.4%	+6.5%

Table 5: Standard Benchmark Results: Zen4-Max vs. Prior Art

4.2 Research and Formal Reasoning

Benchmark	Zen3-Max	Zen4-Max
FrontierMath (overall)	48.3%	61.2%
FrontierMath (Level 1)	71.4%	82.7%
FrontierMath (Level 2)	45.8%	59.3%
FrontierMath (Level 3)	27.7%	41.6%
AIME 2024 (avg 2 sets)	71.3%	81.4%
OlympiadBench	59.2%	68.7%
ProofWriter (deductive)	83.4%	91.2%
FOLIO (first-order logic)	79.8%	87.6%

Table 6: Research and Formal Reasoning Benchmarks

4.3 Scientific Domain Evaluation

Domain Benchmark	Metric	Zen4-Max
GPQA Diamond (physics)	Accuracy	81.3%
GPQA Diamond (chemistry)	Accuracy	78.4%
GPQA Diamond (biology)	Accuracy	79.7%
SciBench (university-level)	Accuracy	86.2%
MedQA USMLE (4-opt)	Accuracy	88.7%
ClimateBench	Accuracy	74.3%

Table 7: Scientific Domain Evaluation

4.4 Long-Output Coherence

A key capability for research assistance is maintaining logical consistency across long generated outputs. We evaluate this using a custom **Proof Coherence Score** that measures the rate at which self-contained mathematical arguments in long outputs are free of logical contradictions, verified by a symbolic checker.

Output Length	Zen4-Pro	Zen4-Max	Zen4-Thinking
1K tokens	94.1%	96.8%	97.3%
4K tokens	87.3%	92.4%	95.1%
8K tokens	74.2%	86.7%	91.8%
16K tokens	61.8%	78.3%	85.4%

Table 8: Proof Coherence Score by Output Length

5 Related Work

FrontierMath [7] provides a benchmark of research-mathematics problems designed to be resistant to pattern matching and memorization, representing a meaningful signal of genuine mathematical reasoning capability. Prior published scores on this benchmark range from single digits for standard instruction-tuned models to the mid-50s for extended-thinking models; Zen4-Max’s 61.2% represents the current best result in our evaluation.

Formal verification integration into language model training has been explored in the context of Lean-based theorem provers [8] and RLHF with verifier reward models [9]. PAT extends this

by incorporating verification signals at the pretraining stage, not only during post-training, which we find is critical for deep internalization of proof structure.

Expert specialization in MoE models through domain-partitioned data mixtures was explored in prior work on mixture-of-experts for multilingual models [10]. Zen4-Max applies this principle to scientific domains with more granular expert routing.

6 Deployment

6.1 Hardware Requirements

Format	VRAM	Context	Hardware
FP8	4× H100 80GB	128K	Recommended
FP8	2× H100 80GB	64K	Reduced context
INT4 (AWQ)	4× A100 80GB	32K	Budget option
INT4 (AWQ)	2× A100 80GB	16K	Minimal config

Table 9: Zen4-Max Deployment Configurations

6.2 Research Workflow Integration

```

1 from hanzo import HanzoAI
2
3 client = HanzoAI()
4
5 system_prompt = """You are a mathematical research assistant.
6 When proving theorems, proceed step by step.
7 Each step should follow logically from prior steps or stated axioms.
8 Flag any step requiring an unproved assumption.
9 """
10
11 response = client.chat.completions.create(
12     model="zen4-max",
13     messages=[
14         {"role": "system", "content": system_prompt},
15         {"role": "user", "content": "Prove that sqrt(2) is irrational."}
16     ],
17     max_tokens=4096,
18     temperature=0.0,          # deterministic for proof generation
19 )

```

Listing 1: Zen4-Max for Mathematical Proof Assistance

7 Safety and Alignment

Metric	Zen3-Max	Zen4-Max
Harmful content refusal	96.1%	98.4%
Over-refusal rate	5.3%	3.1%
Jailbreak resistance	94.8%	97.9%
TruthfulQA	76.8%	84.7%
Scientific claim accuracy	83.4%	89.2%

Table 10: Safety and Alignment Metrics

Scientific claim accuracy measures the rate at which the model’s scientific assertions match peer-reviewed consensus, evaluated by domain experts on 500 sampled claims per scientific domain. The improvement from 83.4% to 89.2% reflects the increased scientific literature proportion in pretraining and the PAT reward signal penalizing unsupported assertions.

7.1 Limitations

- FrontierMath performance (61.2%) indicates significant headroom on research-level mathematics; Zen4-Thinking is recommended for tasks where extended reasoning budget is more important than knowledge breadth.
- Formal proof output is not guaranteed to be machine-verifiable without post-processing; the model learns proof style from Lean4 but does not run a lean4 kernel during inference.
- Expert routing is non-deterministic across batch sizes; results may vary slightly when comparing single-item and batched inference.

8 Conclusion

Zen4-Max establishes the highest standard for research-grade AI assistance at a practical inference cost. Through Proof-Augmented Training, it internalizes formal reasoning structure during pretraining rather than only imitating it during fine-tuning. The 480B/48B Mixture-of-Experts architecture concentrates domain expertise in research clusters that are selectively activated by the routing mechanism for each token. With 90.1% on MATH, 79.8% on GPQA Diamond, and 61.2% on FrontierMath, Zen4-Max is the recommended choice for research institutions, quantitative finance, pharmaceutical discovery, and any application where logical completeness and scientific accuracy are primary requirements.

Acknowledgments

We thank the formal methods research community whose open Lean4 and Isabelle corpora made Proof-Augmented Training possible, the Zoo Labs Foundation science team for domain evaluation design, and the Hanzo AI infrastructure team for the training cluster.

References

- [1] Vaswani, A. et al. (2017). Attention Is All You Need. NeurIPS 2017.
- [2] Brown, T. et al. (2020). Language Models are Few-Shot Learners. NeurIPS 2020.

- [3] Shazeer, N. et al. (2017). Outrageously Large Neural Networks: The Sparsely-Gated Mixture-of-Experts Layer. ICLR 2017.
- [4] Fedus, W. et al. (2022). Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity. JMLR 2022.
- [5] Hoffmann, J. et al. (2022). Training Compute-Optimal Large Language Models. NeurIPS 2022.
- [6] Hinton, G., Vinyals, O. and Dean, J. (2015). Distilling the Knowledge in a Neural Network. arXiv:1503.02531.
- [7] Glazer, E. et al. (2024). FrontierMath: A Benchmark for Evaluating Advanced Mathematical Reasoning in AI. arXiv:2411.04872.
- [8] Polu, S. et al. (2022). Formal Mathematics Statement Curriculum Learning. ICLR 2023.
- [9] Lightman, H. et al. (2023). Let’s Verify Step by Step. arXiv:2305.20050.
- [10] Artetxe, M. et al. (2021). Efficient Large Scale Language Modeling with Mixtures of Experts. EMNLP 2022.
- [11] Chowdhery, A. et al. (2023). PaLM: Scaling Language Modeling with Pathways. JMLR 2023.
- [12] Romero, A. et al. (2015). FitNets: Hints for Thin Deep Nets. ICLR 2015.

A Model Card

Field	Value
Model Name	Zen4-Max
Version	v2026.01
Release Date	January 2026
Total Parameters	480B
Active Parameters	48B per token
Context Length	131,072 tokens
License	Apache 2.0 (weights)
Repository	huggingface.co/zenlm/zen4-max
Documentation	papers.zenlm.org/zen4-max
Contact	research@hanzo.ai

Table 11: Zen4-Max Model Card

B PAT Ablation Study

Training Configuration	MATH	FrontierMath	GPQA
Base (no PAT)	85.3%	48.1%	74.8%
+ Formal corpus (pretraining)	87.4%	53.6%	76.3%
+ Step-level reward (GRPO)	89.1%	58.4%	78.7%
+ Style penalty (γ term)	90.1%	61.2%	79.8%

Table 12: PAT Ablation: Contribution of Each Component

Each PAT component contributes incrementally. The formal corpus provides the largest gain on FrontierMath (5.5 points), while the step-level reward provides the largest gain on MATH (1.7 points, where intermediate step quality is most important). The style penalty provides a smaller but consistent benefit across all three benchmarks.