

Zen AI Model Family

Zen4-Pro

Fourth-Generation Professional Model

Technical Whitepaper v2026.01

Zach Kelling*
research@hanzo.ai

Zoo Labs Foundation
foundation@zoolabs.org

January 2026

Abstract

We present **Zen4-Pro**, the flagship professional model of the Zen4 generation at 72B dense parameters. Zen4-Pro advances the state of agentic AI through deep reasoning chains, native agentic workflows with structured plan-execute-verify loops, and improved tool orchestration across multi-step tasks. The model achieves 91.8% on MMLU, 87.6% on MATH, 93.2% on HumanEval, 74.3% on GPQA Diamond, and 48.7% on SWE-bench Verified, establishing new state-of-the-art results for the 72B dense parameter class. Zen4-Pro is designed for enterprise deployment scenarios requiring reliable, multi-step autonomous task execution while maintaining practical inference efficiency on standard data-center hardware.

Contents

1	Introduction	3
1.1	Key Contributions	3
2	Architecture	4
2.1	Model Specifications	4
2.2	Hierarchical Hybrid Attention at 72B	4
2.3	Agentic State Representation	4
2.4	Distilled Fine-Tuning at Professional Scale	5
3	Training	5
3.1	Pretraining	5
3.2	Post-Training Pipeline	5
3.2.1	Supervised Fine-Tuning	5
3.2.2	Agentic Alignment	6
3.2.3	Preference Optimization	6
3.3	Training Infrastructure	6
4	Evaluation	7
4.1	Standard Benchmarks	7
4.2	Agentic Benchmarks	7
4.3	Multi-Step Reasoning	7
4.4	Professional Domain Evaluation	7

*zach@lux.network

5	Related Work	8
6	Deployment	8
6.1	Inference Configuration	8
6.2	Recommended Agentic Configuration	8
7	Safety and Alignment	9
7.1	Limitations	9
8	Conclusion	9
A	Model Card	10

1 Introduction

Professional-tier AI deployment presents a distinct challenge: tasks are rarely single-turn question-answer exchanges. Software engineers debug across dozens of files; data analysts chain transformations through complex pipelines; legal professionals trace arguments across hundreds of documents. These workflows require a model that can sustain coherent intent, manage tool state, and self-verify intermediate results across many sequential steps.

Zen4-Pro is purpose-built for this regime. While sharing the core Hierarchical Hybrid Attention architecture with Zen4 (14B), Zen4-Pro’s 72B parameter count provides substantially deeper feed-forward capacity, enabling richer world-model representations that support multi-step planning. The model’s post-training pipeline introduces a dedicated **Agentic Alignment** phase that teaches plan-execute-verify patterns from real engineering task traces, not synthetic dialogues.

The result is a model that scores 48.7% on SWE-bench Verified — the industry-standard evaluation for autonomous software engineering — while matching or exceeding models of greater scale on standard reasoning and knowledge benchmarks.

1.1 Key Contributions

- **Agentic Alignment Training:** A new post-training stage using real engineering task traces with ground-truth verifiable outcomes as the reward signal.
- **Hierarchical Planning:** An internal representation of task state that persists across tool calls, enabling reliable multi-step execution without prompt re-injection.
- **Self-Verification:** The model learns to check its own intermediate outputs against stated task requirements before proceeding, reducing error propagation in long chains.
- **Frontier Benchmark Results:** New state-of-the-art at 72B dense for SWE-bench, GPQA, and MATH.

2 Architecture

2.1 Model Specifications

Component	Specification
Total Parameters	72.7B
Active Parameters	72.7B (dense)
Architecture	Hybrid Transformer
Attention Mechanism	Hierarchical Hybrid Attention (HHA)
Local Window Size	8,192 tokens
Global Anchor Tokens	1,024 per layer
Context Length	131,072 tokens (128K)
Hidden Dimension	8,192
Intermediate Dimension	29,568
Attention Heads	64
Key-Value Heads	8 (GQA)
Layers	80
Vocabulary Size	151,936
Positional Encoding	RoPE ($\theta = 10,000,000$)
Normalization	RMSNorm
Activation	SwiGLU

Table 1: Zen4-Pro Architecture Specifications

2.2 Hierarchical Hybrid Attention at 72B

Zen4-Pro scales the HHA mechanism from Zen4 with a larger local window ($w = 8192$ vs 4096) and more global anchor tokens ($k = 1024$ vs 512). The larger window improves performance on tasks requiring dense local reasoning, such as code editing within a single large file. The additional anchors improve global coherence for tasks that span many documents.

At 72B parameters, the feed-forward networks can represent substantially richer world-model priors. Ablation studies show that the improvement from Zen4 to Zen4-Pro on complex planning tasks (+18% on AgentBench) is primarily attributable to feed-forward capacity rather than attention mechanism differences, confirming that scale continues to benefit planning tasks even after architectural improvements have maximized attention efficiency.

2.3 Agentic State Representation

Zen4-Pro introduces an explicit task-state representation within the context window. During agentic post-training, the model learns to maintain a structured scratchpad following the format:

```
1 <task_state>
2   <goal>Refactor authentication module to use JWT</goal>
3   <completed>
4     - Read auth/session.py (lines 1-340)
5     - Identified 3 session-creation call sites
6     - Updated auth/session.py with JWT logic
7   </completed>
8   <pending>
9     - Update call sites in api/routes.py, api/middleware.py
10    - Run test suite to verify no regressions
11  </pending>
12  <open_questions>
```

```

13   - Token expiry: use 24h or match existing session TTL?
14   </open_questions>
15 </task_state>

```

Listing 1: Agentic State Token Format

This structured scratchpad is generated and updated by the model itself; it is not injected externally. The model learns to parse its own prior state tokens as factual context, enabling reliable task tracking over 20+ sequential tool calls without context-window management by the orchestration layer.

2.4 Distilled Fine-Tuning at Professional Scale

Zen4-Pro is distilled from an ensemble of Mixture-of-Experts teacher models using the fourth-generation distillation framework. At 72B, the student has sufficient capacity to absorb richer intermediate representations, enabling distillation from multiple specialized teachers simultaneously:

- A mathematics specialist teacher contributes to layers 20–40.
- A code reasoning teacher contributes to layers 40–60.
- A general instruction-following teacher provides global signal across all layers.

Multi-teacher distillation with layer-targeted routing improves MATH by 3.1% and HumanEval by 2.8% over single-teacher distillation at matched compute.

3 Training

3.1 Pretraining

Data Category	Proportion	Tokens
Web (quality-filtered)	38%	7.6T
Code (multi-language)	32%	6.4T
Scientific and technical	14%	2.8T
Tool-call trajectories	8%	1.6T
Mathematical content	5%	1.0T
Books and long-form prose	3%	0.6T
Total	100%	20.0T

Table 2: Zen4-Pro Pretraining Data Composition (20T tokens)

Zen4-Pro pretrained on 20 trillion tokens. The increased code proportion (32% vs 28% in Zen4) reflects the professional-tier emphasis on software engineering tasks. Scientific and technical text proportion is also increased (14% vs 10%) for the deep-reasoning requirements of professional domains.

3.2 Post-Training Pipeline

3.2.1 Supervised Fine-Tuning

The SFT stage for Zen4-Pro uses 4.5M instruction-response pairs, approximately twice the Zen4 volume. The additional data is concentrated in:

- Multi-step software engineering tasks with execution traces.
- Multi-document analysis and synthesis tasks.
- Long-horizon planning and decomposition tasks.

3.2.2 Agentic Alignment

The agentic alignment stage is unique to Zen4-Pro and Zen4-Max. It trains the model on trajectories collected from real engineering tasks in sandboxed environments:

1. **Trajectory Collection:** Deploy an early checkpoint on 50,000 real GitHub issues in isolated container environments. Collect all tool calls, intermediate states, and final outcomes (pass/fail against the issue’s test suite).
2. **Outcome Filtering:** Retain trajectories where the final outcome is verifiably correct (test suite passes). Discard partial successes.
3. **GRPO Training:** Use Group Relative Policy Optimization with the binary test-pass signal as reward. Generate $G = 8$ trajectories per task and train on the relative reward within each group.
4. **Iteration:** Repeat trajectory collection and training for 3 rounds using the updated model.

This iterative process yields a 14.2-point improvement on SWE-bench Verified over the SFT-only baseline.

3.2.3 Preference Optimization

DPO training uses 1.2M comparison pairs. For professional-tier use cases, comparison pairs are weighted to emphasize:

- Correctness of multi-step plans ($3\times$ weight).
- Appropriate tool selection and argument accuracy ($2\times$ weight).
- Response clarity and actionability ($1\times$ weight).

3.3 Training Infrastructure

Parameter	Value
Hardware	$4,096 \times$ H100 80GB SXM
Parallelism	TP=8, PP=8, DP=64
Batch Size	8M tokens (global)
Peak Learning Rate	1.5×10^{-4}
LR Schedule	Cosine with warmup (3000 steps)
Optimizer	AdamW ($\beta_1 = 0.9, \beta_2 = 0.95$)
Training Duration	32 days (pretraining)

Table 3: Zen4-Pro Training Infrastructure

4 Evaluation

4.1 Standard Benchmarks

Benchmark	Zen3-Pro	Zen4	Zen4-Pro
MMLU (5-shot)	85.7%	85.4%	91.8%
MATH (4-shot)	82.1%	81.2%	87.6%
HumanEval (0-shot)	86.4%	88.3%	93.2%
GPQA Diamond (0-shot)	65.8%	67.4%	74.3%
IFEval (prompt-strict)	80.2%	83.1%	88.4%
GSM8K (8-shot)	92.3%	93.7%	96.8%
BBH (3-shot)	81.4%	83.9%	89.7%

Table 4: Benchmark Comparison Across Generations and Tiers

4.2 Agentic Benchmarks

Benchmark	Zen3-Pro	Zen4	Zen4-Pro
SWE-bench Verified	32.1%	34.8%	48.7%
AgentBench (overall)	58.3%	63.1%	76.4%
WebArena (overall)	31.4%	35.2%	44.8%
ToolBench (success rate)	67.2%	79.4%	88.1%
GAIA (Level 1)	71.3%	74.8%	83.6%
GAIA (Level 2)	48.9%	52.3%	63.7%

Table 5: Agentic Capability Benchmarks

4.3 Multi-Step Reasoning

Task	Steps	Zen4-Pro Score
Plan + execute (3-step)	3	89.4%
Plan + execute (5-step)	5	81.2%
Plan + execute (10-step)	10	68.7%
Self-verify + correct	2	77.3%
Error recovery after tool failure	—	71.8%

Table 6: Multi-Step Task Completion Rates (AgentBench decomposition)

4.4 Professional Domain Evaluation

Domain	Task	Score
Software Engineering	SWE-bench Verified	48.7%
Data Analysis	BIRD SQL (accuracy)	73.4%
Scientific Research	GPQA Diamond	74.3%
Legal	LegalBench (avg)	68.1%
Finance	FinanceBench (accuracy)	84.2%
Medicine	MedQA USMLE (4-option)	86.3%

Table 7: Professional Domain Evaluation

5 Related Work

Scaling dense transformers beyond 70B parameters has followed a predictable capability-compute curve [5]. Zen4-Pro’s 72B represents the practical ceiling for single-node inference on current data-center GPUs while pushing capability to the frontier.

Agentic AI systems built on language models have grown rapidly since ReAct [11] demonstrated that interleaving reasoning traces with action invocations improves task success on multi-step benchmarks. Subsequent systems [12] extended this to hierarchical planning. Zen4-Pro internalizes these patterns rather than relying on external orchestration frameworks, reducing latency and improving coherence across long trajectories.

SWE-bench [10] has emerged as the canonical evaluation for autonomous software engineering capability. Current frontier scores range from 30–50% depending on scaffolding; Zen4-Pro achieves 48.7% with a minimal scaffold, demonstrating that model quality rather than scaffolding complexity drives performance.

6 Deployment

6.1 Inference Configuration

Format	VRAM	Context	Hardware
BF16	145 GB	128K	H100 $\times$ 2
FP8	72 GB	128K	H100 $\times$ 1
Q4_K_M	42 GB	32K	RTX 6000 Ada $\times$ 1
Q4_K_M	48 GB	64K	A6000 $\times$ 1

Table 8: Zen4-Pro Deployment Hardware Requirements

6.2 Recommended Agentic Configuration

```
1 from hanzo import HanzoAI
2
3 client = HanzoAI()
4
5 # Agentic tasks: enable longer generation, structured output
6 response = client.chat.completions.create(
7     model="zen4-pro-72b-instruct",
8     messages=[
9         {"role": "system", "content": "You are an autonomous
10         engineering agent."},
11         {"role": "user", "content": task_description}
12     ],
13     tools=toolset,
14     tool_choice="auto",
15     max_tokens=8192,           # allow full reasoning chains
16     temperature=0.2,          # low temperature for agentic reliability
17     extra_body={
18         "agentic_mode": True,  # enables internal state tracking
19         "max_tool_rounds": 20  # hard cap on tool-call loops
20     }
21 )
```

Listing 2: Zen4-Pro Agentic Task Configuration

7 Safety and Alignment

Safety Metric	Zen3-Pro	Zen4-Pro
Harmful content refusal	95.3%	98.1%
Over-refusal rate (benign)	6.2%	3.4%
Jailbreak resistance (AdvBench)	93.7%	97.8%
TruthfulQA	74.8%	82.3%
Agentic safety (destructive action rate)	1.2%	0.3%

Table 9: Safety Metrics: Zen3-Pro vs. Zen4-Pro

Agentic deployment introduces safety risks beyond standard language model outputs: a model executing multi-step tool plans can take irreversible actions (file deletion, API calls with side effects). Zen4-Pro’s post-training includes an explicit **action consequence evaluation** objective that teaches the model to identify and flag potentially irreversible operations before executing them.

7.1 Limitations

- Multi-step task completion rates decay with plan length: 10-step tasks complete at 68.7% vs 89.4% for 3-step tasks, indicating compounding errors in long chains.
- Professional domain evaluations (legal, medical) reflect the state of training data and should not substitute for domain expert review.
- 128K context is sufficient for most engineering tasks but falls short of the 1M context available in Zen4 (14B). For long-document tasks, Zen4 may be preferable despite its lower benchmark scores.

8 Conclusion

Zen4-Pro delivers frontier professional capability at 72B dense parameters through three interlocking advances: the Hierarchical Hybrid Attention architecture shared across the Zen4 family, multi-teacher distilled fine-tuning that concentrates specialist knowledge into student layers, and the novel Agentic Alignment post-training stage that grounds multi-step planning in verifiable engineering outcomes. With 48.7% on SWE-bench Verified and strong results across mathematics, science, and knowledge benchmarks, Zen4-Pro is the recommended choice for enterprise deployments requiring reliable autonomous task execution.

Acknowledgments

We thank the annotators, engineers, and researchers at Hanzo AI and Zoo Labs Foundation who contributed to the trajectory collection, post-training pipeline, and evaluation infrastructure underlying this work.

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] Ouyang, L. et al. (2022). Training Language Models to Follow Instructions with Human Feedback. NeurIPS 2022.
- [4] Touvron, H. et al. (2023). LLaMA: Open and Efficient Foundation Language Models. arXiv:2302.13971.
- [5] Hoffmann, J. et al. (2022). Training Compute-Optimal Large Language Models. NeurIPS 2022.
- [6] Shazeer, N. et al. (2017). Outrageously Large Neural Networks: The Sparsely-Gated Mixture-of-Experts Layer. ICLR 2017.
- [7] Schulman, J. et al. (2017). Proximal Policy Optimization Algorithms. arXiv:1707.06347.
- [8] Rafailov, R. et al. (2023). Direct Preference Optimization: Your Language Model is Secretly a Reward Model. NeurIPS 2023.
- [9] Chen, M. et al. (2021). Evaluating Large Language Models Trained on Code. arXiv:2107.03374.
- [10] Jimenez, C. et al. (2024). SWE-bench: Can Language Models Resolve Real-World GitHub Issues? ICLR 2024.
- [11] Yao, S. et al. (2023). ReAct: Synergizing Reasoning and Acting in Language Models. ICLR 2023.
- [12] Wang, X. et al. (2024). Executable Code Actions Elicit Better LLM Agents. arXiv:2402.01030.
- [13] Shao, Z. et al. (2024). DeepSeekMath: Pushing the Limits of Mathematical Reasoning in Open Language Models. arXiv:2402.03300.

A Model Card

Field	Value
Model Name	Zen4-Pro
Version	v2026.01
Release Date	January 2026
Parameters	72.7B
Context Length	131,072 tokens
License	Apache 2.0
Repository	huggingface.co/zenlm/zen4-pro-72b-instruct
Documentation	papers.zenlm.org/zen4-pro
Contact	research@hanzo.ai

Table 10: Zen4-Pro Model Card