

Anthropic Announces Claude Opus 4, the World's Most Powerful Programming Model

Anthropic today announced the new generation of Claude models: Claude Opus 4 and Claude Sonnet 4.

Anthropic today announced the next generation of its Claude models: **Claude Opus 4** and **Claude Sonnet 4**. The company claims that Claude Opus 4 — the most advanced version — is **the world's most powerful programming model**, excelling at agentic workflows and handling complex, long-running tasks. Meanwhile, Claude Sonnet 4 improves programming and logic performance over the previous Claude Sonnet 3.7 version.

According to the comparison chart below, Claude Sonnet 4 scored an all-time high of **72.7%** on **SWE-bench Verified** — a popular benchmark for automated programming capabilities. With this strength, **Microsoft's GitHub** announced that Sonnet 4 will be the core model for the new programming agent in **GitHub Copilot**.

Anthropic describes the performance of the new models as follows:

These models advance customers' AI strategies across the board: Opus 4 breaks boundaries in programming, research, writing, and scientific discovery, while Sonnet 4 brings cutting-edge performance to everyday tasks as an immediate upgrade from Sonnet 3.7.

	Claude Opus 4	Claude Sonnet 4	Claude Sonnet 3.7	OpenAI o3	OpenAI GPT-4.1	Gemini 2.5 Pro Preview (05-06)
Agentic coding <small>SWE-bench Verified^{4,5}</small>	72.5% / 79.4%	72.7% / 80.2%	62.3% / 70.3%	69.1%	54.6%	63.2%
Agentic terminal coding <small>Terminal-bench^{2,3}</small>	43.2% / 50.0%	35.5% / 41.3%	35.2%	30.2%	30.3%	25.3%
Graduate-level reasoning <small>GPQA Diamond⁵</small>	79.6% / 83.3%	75.4% / 83.8%	78.2%	83.3%	66.3%	83.0%
Agentic tool use <small>TAU-bench</small>	Retail 81.4%	Retail 80.5%	Retail 81.2%	Retail 70.4%	Retail 68.0%	—
	Airline 59.6%	Airline 60.0%	Airline 58.4%	Airline 52.0%	Airline 49.4%	—
Multilingual Q&A <small>MMMLU³</small>	88.8%	86.5%	85.9%	88.8%	83.7%	—
Visual reasoning <small>MMMU (validation)</small>	76.5%	74.4%	75.0%	82.9%	74.8%	79.6%
High school math competition <small>AIME 2025^{4,5}</small>	75.5% / 90.0%	70.5% / 85.0%	54.8%	88.9%	—	83.0%

Both new Claude 4 models support two modes:

1. **Almost instant response**
2. **Extended Thinking** (for problems that require in-depth analysis)
Programmers can switch between the two modes as needed.

Claude free users only get access to the Sonnet 4 model, while **Pro, Max, Team, and Enterprise** users get access to both models and extended thinking. Anthropic also says the models are available on **the Anthropic API**, **Amazon Bedrock**, and **Google Cloud's Vertex AI**. Notably, Anthropic **isn't offering discounts** on the new models:

1. Claude Opus 4: \$15/1 million input tokens and \$75/1 million output tokens
2. Claude Sonnet 4: \$3/1 million input tokens and \$15/1 million output tokens

Along with the new models, Anthropic announced updates to **Claude Code** with support for background tasks via **GitHub Actions**, native integration with **VS Code** and **JetBrains**, and more. Additionally, **the Anthropic API** adds a code execution tool, MCP connectivity, Files API, and the ability to cache prompts for up to **1 hour**.

You finished reading the article "**Anthropic Announces Claude Opus 4, the World's Most Powerful Programming Model**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.