

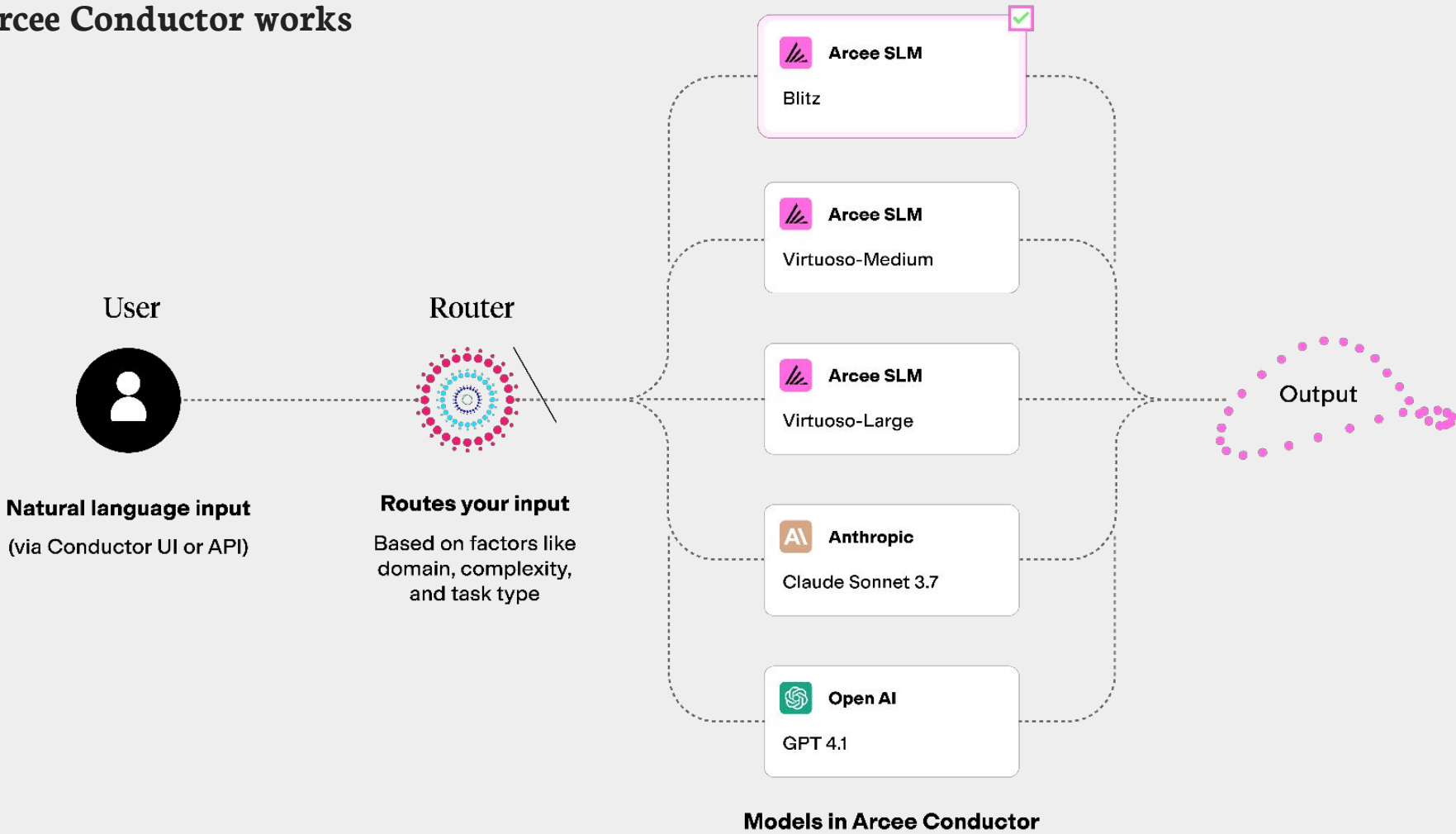
Arcee Conductor

Intelligent model routing

The right model for every prompt

Arcee Conductor is an intelligent model routing platform that directs each input to its ideal AI model based on domain, task type, and complexity. By dynamically routing between large language models (LLMs) and our own specialized small language models (Arcee SLMs), Conductor maximizes cost efficiency without compromising performance.

How Arcee Conductor works



Why choose Arcee Conductor vs. a single AI model?

Dramatic Cost Savings

- **Average savings of 64%** vs. traditional model usage
- Cost per million tokens reduced by **\$11** vs. single-model approach (* For enterprises processing **500M** tokens per month, potential savings of **\$281,562** per year)

Ease of Use

- **Seamlessly routes to either SLMs or LLMs**
- Arcee Conductor's API connects to your existing language model stack with a single line of code and no interruption to the end-user experience

Performance without Compromise

- Precise, intelligence routing that **matches or exceeds performance** vs. using only a leading LLM
- Results show Arcee Conductor consistently performs better than a single premium LLM across standard evaluation metrics including MMLU, Math-500, GPQA-D and HumanEval

Full, Transparent Reporting

- Each output delivered with **detailed analytics on cost** and on **the rationale for model selection**
- “Compare” feature enables you to easily compare outputs of different AI models

About the AI models in Arcee Conductor

AI Model	Description	Excels at...
Blitz <i>Arcee SLM</i>	Our most affordable small language model (SLM), Blitz combines 24B speed with advanced DeepSeek knowledge. A practical workhorse model that can tackle a range of tasks without the overhead of larger architectures	<ul style="list-style-type: none">● Real-time conversation and customer support● Rapid information retrieval and summarization● Mobile applications where speed and efficiency are critical
Virtuoso-Medium <i>Arcee SLM</i>	A versatile model capable of handling diverse input with precision and adaptability across multiple domains. Ideal for dynamic use cases requiring significant computational power	<ul style="list-style-type: none">● Data interpretation, content and report generation● Knowledge retrieval● Advanced language understanding
Virtuoso-Large <i>Arcee SLM</i>	Our most powerful general-purpose model, designed to excel at handling complex prompts and queries across domains. It offers unparalleled capability for nuanced understanding, contextual adaptability, and high accuracy	<ul style="list-style-type: none">● Advanced content creation, such as technical writing and creative storytelling● Data summarization, report generation for cross-functional domains● Detailed knowledge synthesis and deep-dive insights from diverse datasets
Claude 3.7 Sonnet	A highly intelligent model optimized for complex reasoning input while maintaining exceptional language capabilities, designed to provide nuanced responses	<ul style="list-style-type: none">● Complex problem solving and multi-step reasoning● In-depth analysis of nuanced topics and scenarios● Creative content generation with logical coherence and depth
GPT-4.1	A sophisticated multimodal model that seamlessly processes and generates content across text, image, and audio inputs	<ul style="list-style-type: none">● Complex multimodal analysis combining images and text● Creative content generation with visual understanding● Advanced reasoning across multiple input formats

Optimize output across AI models, reduce usage costs, and maximize performance with intelligent routing. [Try Arcee Conductor](#) today.

