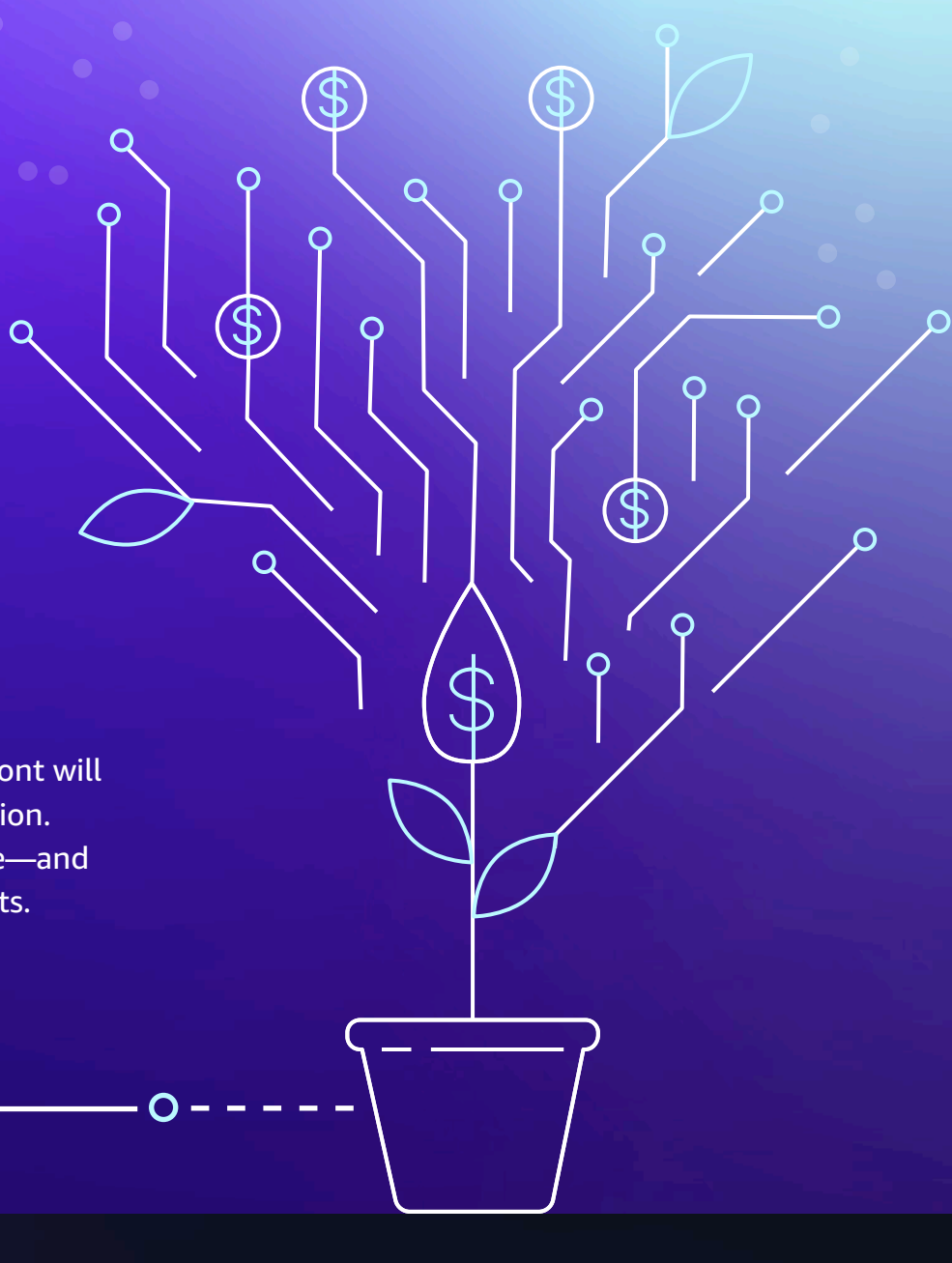




Understanding the costs of generative AI

When building with generative AI, the choices you make upfront will significantly impact the overall costs of your product or solution. Read on to discover which factors have the greatest influence—and how you can optimize them to control your generative AI costs.

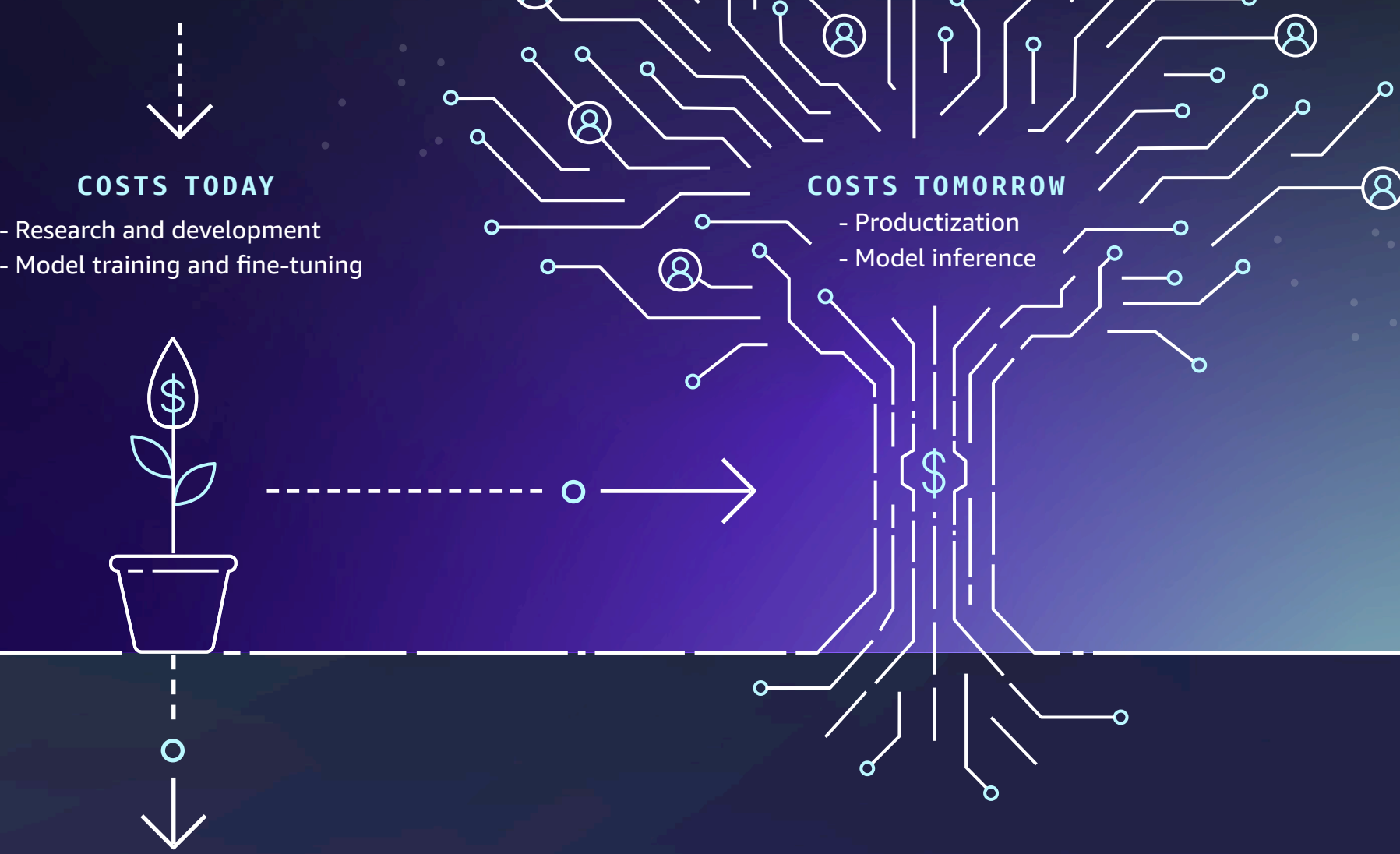


What is generative AI?

Generative AI is a type of artificial intelligence (AI) that can create new content and ideas, including conversations, stories, images, videos, and music. It is powered by large models that are pretrained on vast amounts of data, commonly referred to as foundation models (FMs).

Generative AI costs expand over time

As you plan your generative AI projects, it's important to consider not just the upfront costs of building and training the model—but also the ongoing inference expenses that will expand as your user base grows and customer demand increases.



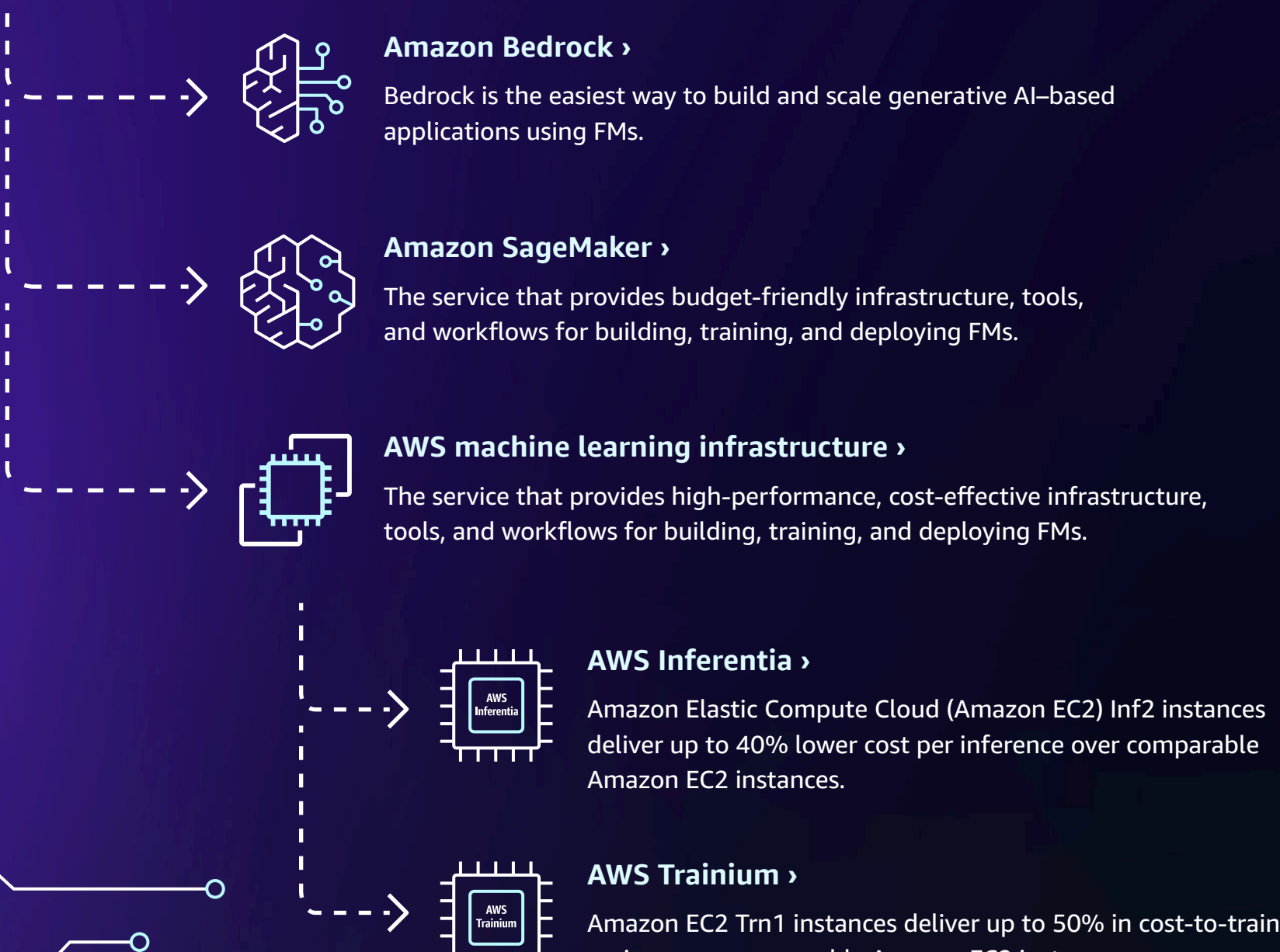
4 steps to optimizing generative AI price performance

By making the right choices at the onset of your generative AI effort, you can better control upfront and downstream costs.



How AWS can help

Amazon Web Services (AWS) offers solutions that can help you achieve the four steps outlined above—all with minimal burden on your resources and maximum impact on your generative AI investments.



Transform your business with generative AI on AWS

Understanding the costs of FM training and inference can help make generative AI more accessible—and more affordable. Learn how AWS is democratizing the technology for organizations of every size.

[Explore generative AI on AWS ›](#)

