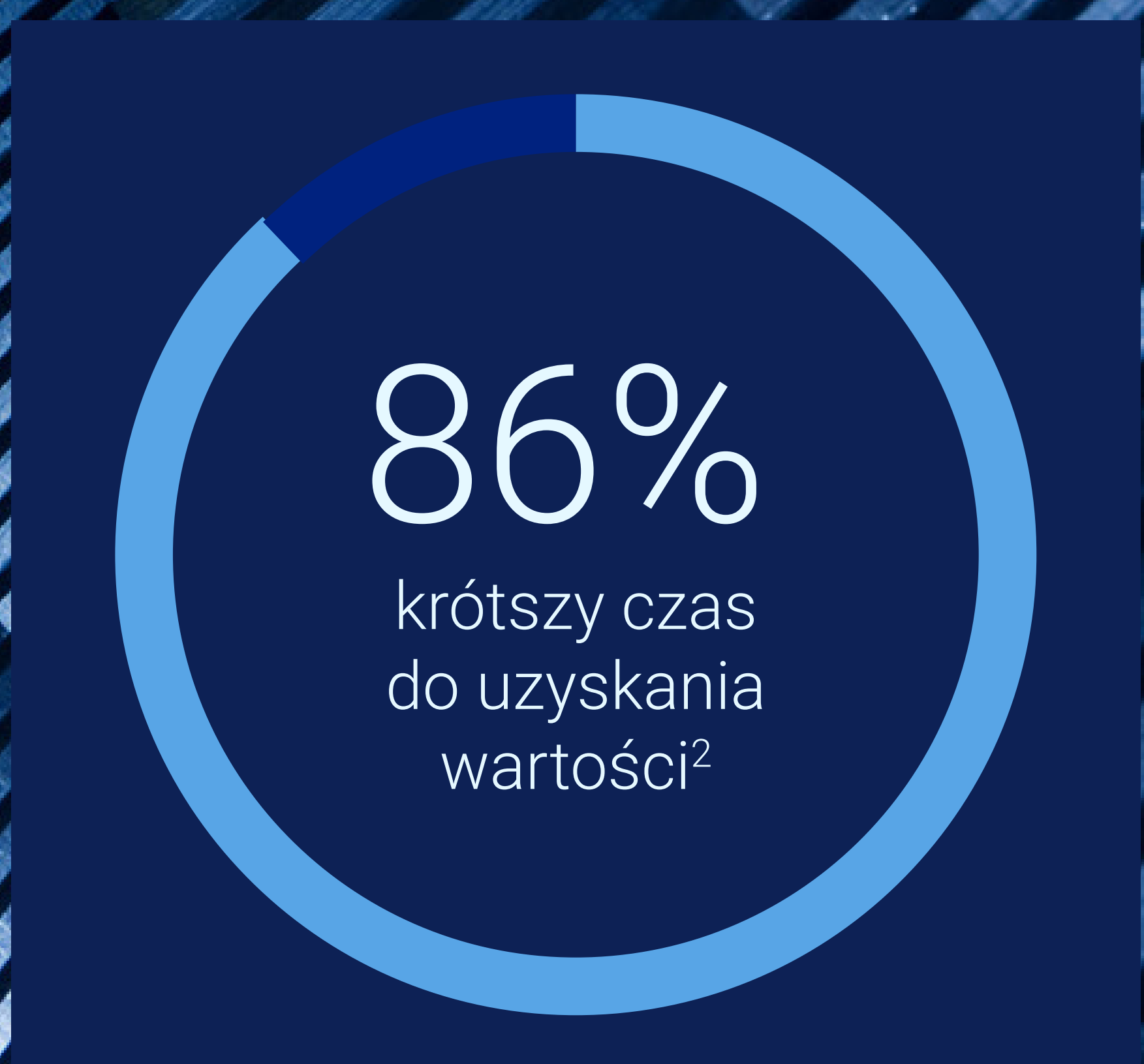
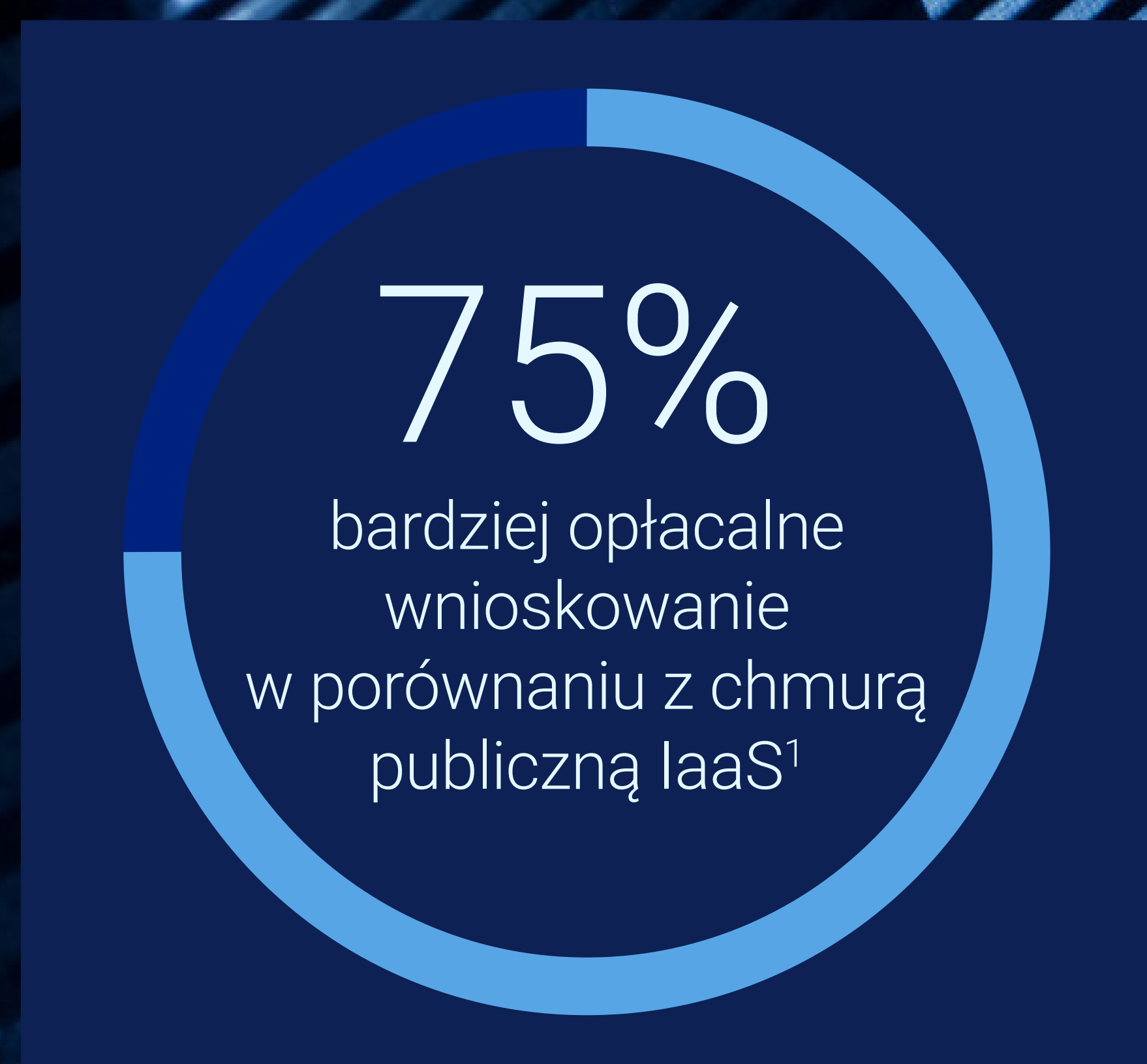


# Dell AI Factory

Rozwiązania firmy Dell w zakresie generatywnej sztucznej inteligencji dla procesorów AMD

Przyspiesz innowacje, zmniejsz koszty i chroń dane dzięki skalowalnej, modułowej architekturze na potrzeby złożonej generatywnej sztucznej inteligencji.



Kluczowe zastosowania wymagają mocy, elastyczności i skalowalności



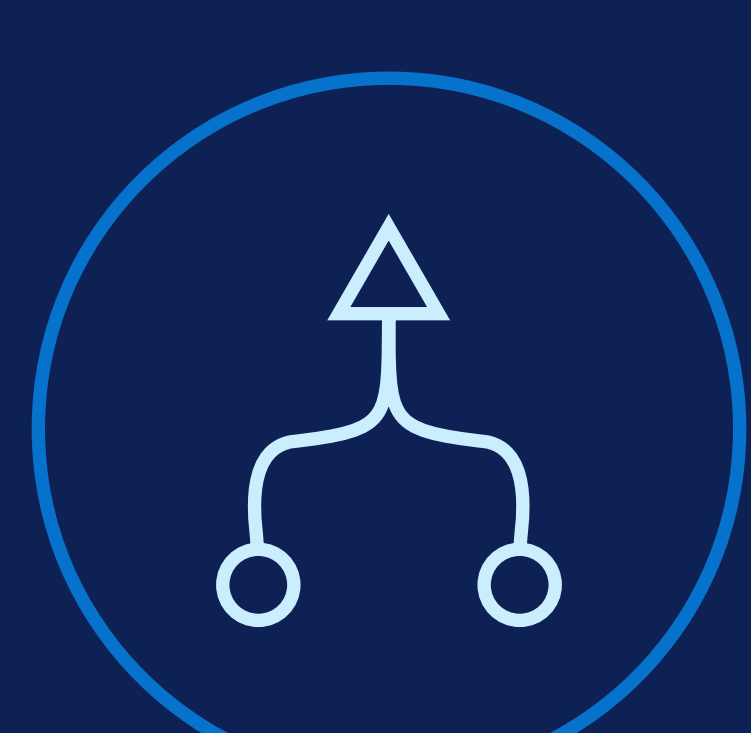
Asystenci, chatboty i tworzenie treści



Akcelerator jako usługa



Multimodalne generowanie rozszerzone wyszukiwania (RAG)



## Uproszczenie

Usprawnij wdrażanie generatywnej sztucznej inteligencji dzięki sprawdzonym, zweryfikowanym rozwiązaniom popartym ponad 340 tys. godzin pracy inżynierów.

### Optimalizacja wydajności

Akcelerator o wysokiej wydajności, otwarta architektura i sieci zoptymalizowane pod kątem SI

### SI w każdym miejscu

Dostęp do danych z każdego miejsca dzięki elastycznej pamięci masowej w wielu chmurach

### Gotowe do użycia rozwiązanie wielowęzłowe

Sprawdzone, kompleksowe fundamenty SI zapewniające szybsze wyniki



## Personalizacja

Oprogramowanie open source AMD ROCm™ i otwarte ekosystemy przyspieszają rozwój i działanie SI.

### Szybsze innowacje

Korzystaj z oprogramowania i ekosystemów open source do tworzenia unikalnych aplikacji.

### Przyspieszony rozwój

Wykorzystaj standardowe struktury branżowe z elastycznymi stosami technologii.

### Aktywacja danych

Wydajnie uruchamiaj wiele przykładów zastosowania SI jednocześnie.



## Zaufanie

82% osób decyzyjnych w kwestiach IT preferuje model lokalny lub hybrydowy<sup>3</sup>. Jakie dane, takie wyniki. Chroń je.

### Szybki początek

Lokalne fundamenty z zaufanymi zabezpieczeniami i pełną kontrolą

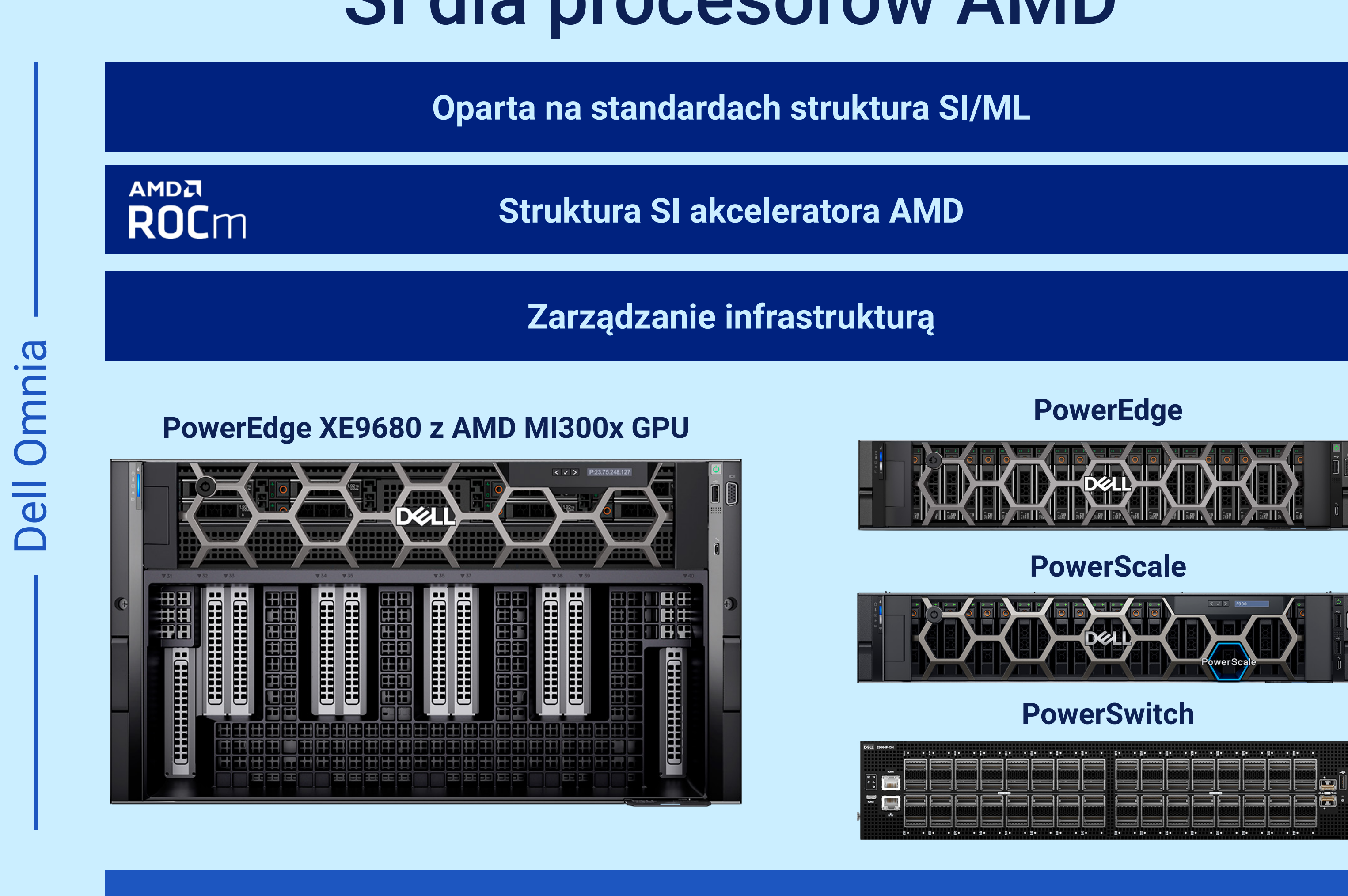
### Usprawniona łączność

Bezpieczne, bogate w funkcje sieci ze skalowalnością i zoptymalizowanymi przepływaniami ruchu

### Zautomatyzowane inicjowanie

Fundament open source do wdrażania klastrów o wysokiej wydajności i zarządzania nimi

## Rozwiązania firmy Dell w zakresie generatywnej SI dla procesorów AMD



### Wnioskowanie

Uruchom model z 70 mld parametrów na jednym akceleratorze AMD Instinct™ MI300X<sup>4</sup>.

### Dostosowanie

Wdróż i dostrój osiem równoległych modeli z 70 mld parametrów na jednym serwerze Dell PowerEdge XE9680<sup>4</sup>.

### Rozszerzenie

Włącz swoje dane do procesu generatywnego.

Zyskaj przewagę konkurencyjną dzięki sprawdzonemu, otwartemu rozwiązaniu zapewniającemu bezpieczne, lokalne aplikacje SI na dużą skalę.

Więcej informacji

<sup>1</sup> Enterprise Strategy Group, maksymalizacja zwrotu z inwestycji w sztuczną inteligencję: wdrożenie lokalne dzięki Dell Technologies może być o 75% bardziej opłacalne niż chmura publiczna, kwiecień 2024 r.

<sup>2</sup> Wartości szacunkowe na podstawie analizy przeprowadzonej przez firmę Dell w maju 2024 r., w ramach której porównano czas konfiguracji 2-węzłowego klastra Kubernetes dla LLM ogólnego przeznaczenia uzyskany przy użyciu zautomatyzowanych skryptów i poprzez ręczne wdrożenie powszechnie

stosowanego wzorca projektowego. Czas konfiguracji obejmuje tylko instalację podstawową. Rzeczywisty czas konfiguracji może się różnić w zależności od konfiguracji rozwiązania.

<sup>3</sup> Dell Technologies, Ankieta Generative AI Pulse Survey, sierpień i wrzesień 2023 r.

<sup>4</sup> Blog Dell Technologies, Silicon Diversity: wdrożenie generatywnej sztucznej inteligencji na serwerach PowerEdge XE9680 z akceleratorami AMD Instinct MI300X, maj 2024 r.