

# High-performance compute

**Purpose-built compute infrastructure for Enterprise AI to help accelerate development and speed to market of autonomous technology with the leading experience of on-prem power and cloud flexibility**

## THE CHALLENGE

- High-performance computing hardware, specifically GPUs, is essential for AV development but scarcely available in the public cloud, potentially making it difficult for AV organizations to implement, operate, and scale
- Complex AV workflows can be necessary for research and development, but some companies utilize a tailor-made infrastructure toolchain that can be difficult and expensive to manage
- The complexity and scalability of public cloud can make it difficult to accelerate development while also controlling costs

## SOLUTION HIGHLIGHTS

- Dedicated capacity you can rely on, without constraints, that can help empower effortless scalability from one system to thousands and beyond
- Achieve dependable operations with regular updates and maintenance provided by specialty teams from NVIDIA and Deloitte, freeing you to focus less on infrastructure and more on AV development
- Has the potential to unlock maximum value with flexible deployment options, tailored to your workflows and delivered at predictable prices

## DEPLOYMENT OPTIONS

### On premise

Have Deloitte stand up, integrate, and manage your own dedicated AI supercomputer

### NVIDIA DGX Cloud

Have Deloitte safely migrate your AV development workflows to a high-performance AI-optimized cloud

### Hybrid

Have Deloitte deliver the maximum value of dedicated on prem paired with the accessibility of DGX Cloud

## VALUE PROPOSITION

### On prem

- Up to 50% reduction in TCO
- Greater than 1.4X improved performance
- Up to 2X data science productivity

### DGX Cloud

- On prem performance, cloud convenience
- No egress fees
- Up to 2X data science productivity

### Hybrid

- Up to 30% reduction in TCO
- Up to 1.3X improved performance

## DIFFERENTIATION

