

# Ascend AI: Building a Versatile, User-friendly, All-scenario AI Platform

---

Danny  
Huawei Computing Ecosystem Development



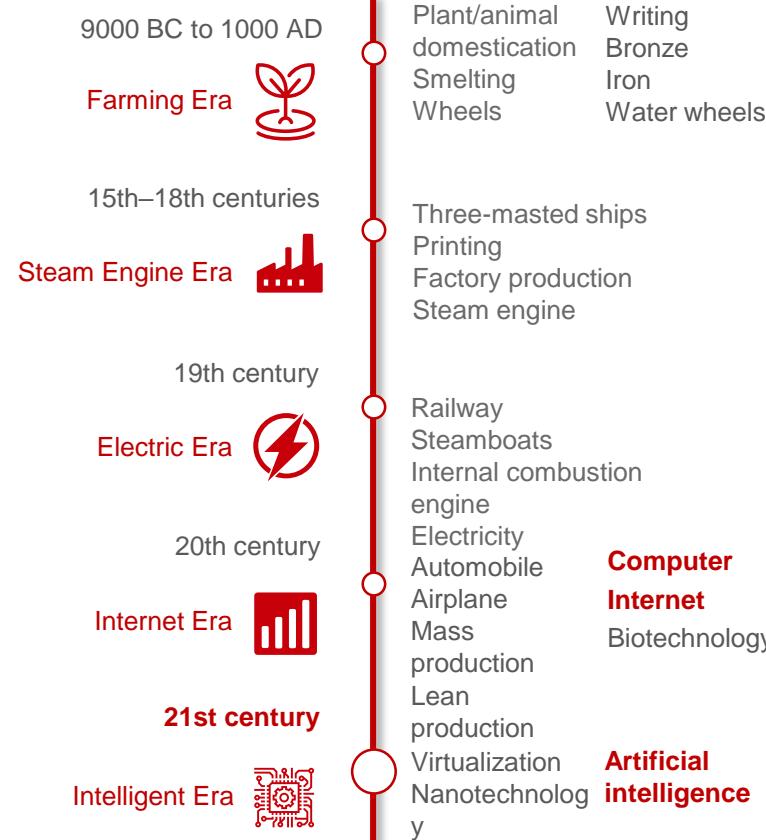
# Towards an AI Era

---

Development Trends of AI

# AI: The New Driver of Social Development

## AI is a new general purpose technology (GPT)



## A high potential market awaits



AI application and productivity curve

**Phase 1**  
Partial exploration of technologies and applications

**Phase II**  
Collision between technical development and social environment

**Phase III**  
Mutual improvement of technology and social environment

**Phase 4**  
Stable development & looking forward to a new GPT

**AI industry market scale in 2025**  
**USD7.44 trillion**

Average annual growth rate in the next 5 years: **30%+**

**AI in 2030 will enable**  
Global GDP grows by **1.2%** per year

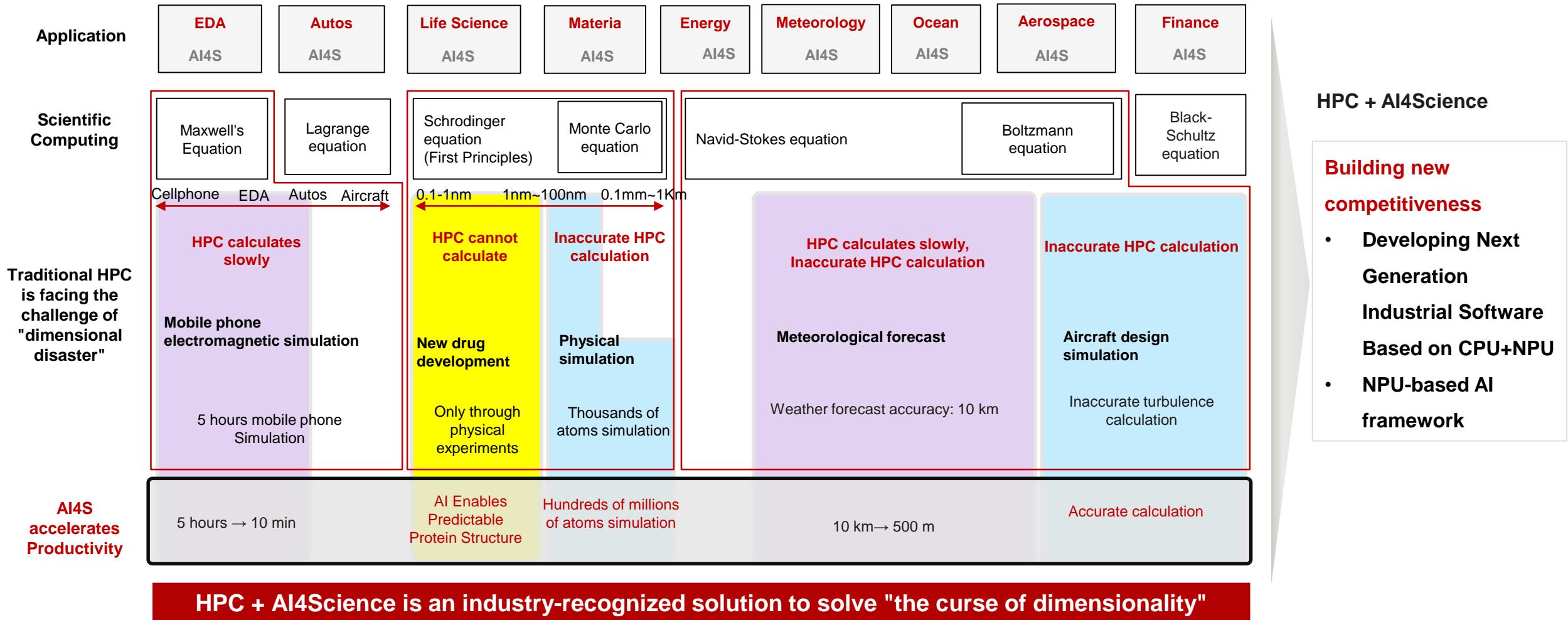


# The Adoption of AI in Different Sectors

The adoption rate of AI in mainstream industries is projected to reach 20% by 2026.



# AI4Science is becoming a new demand in the industry

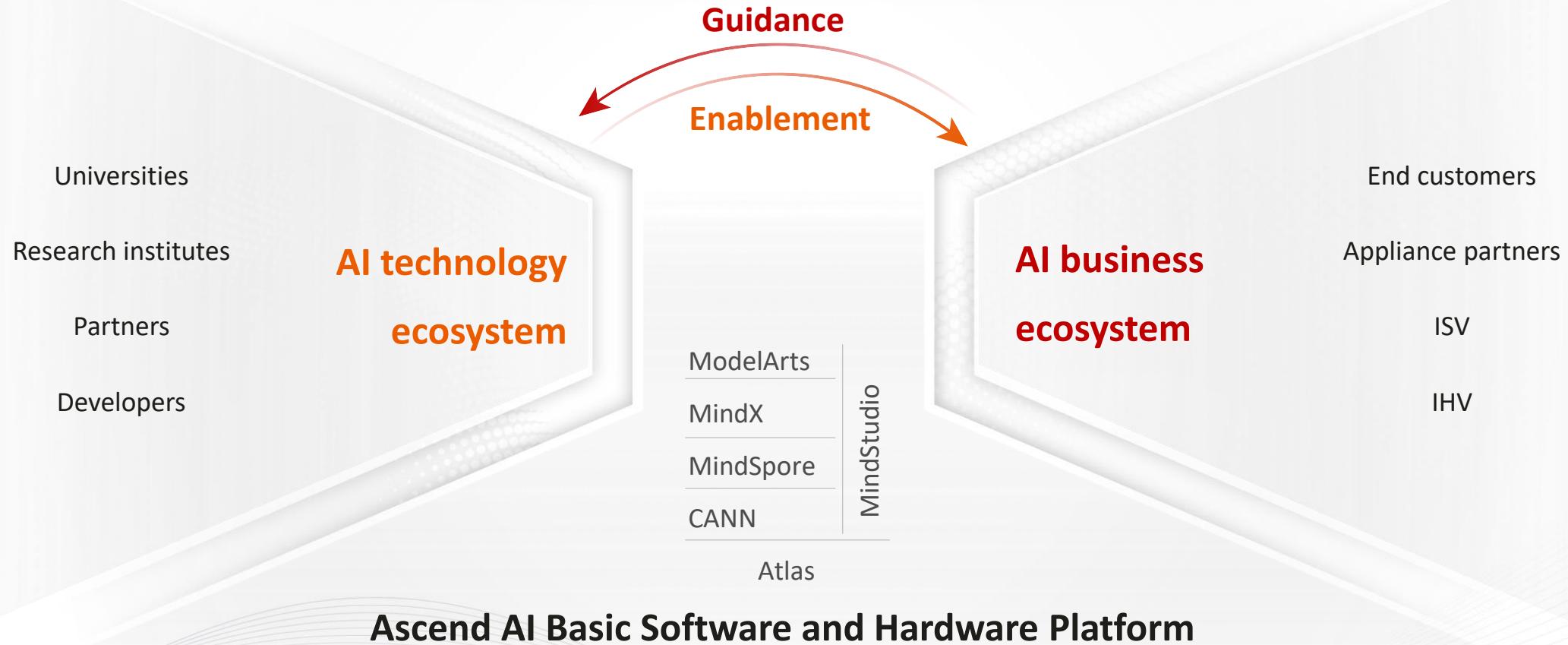




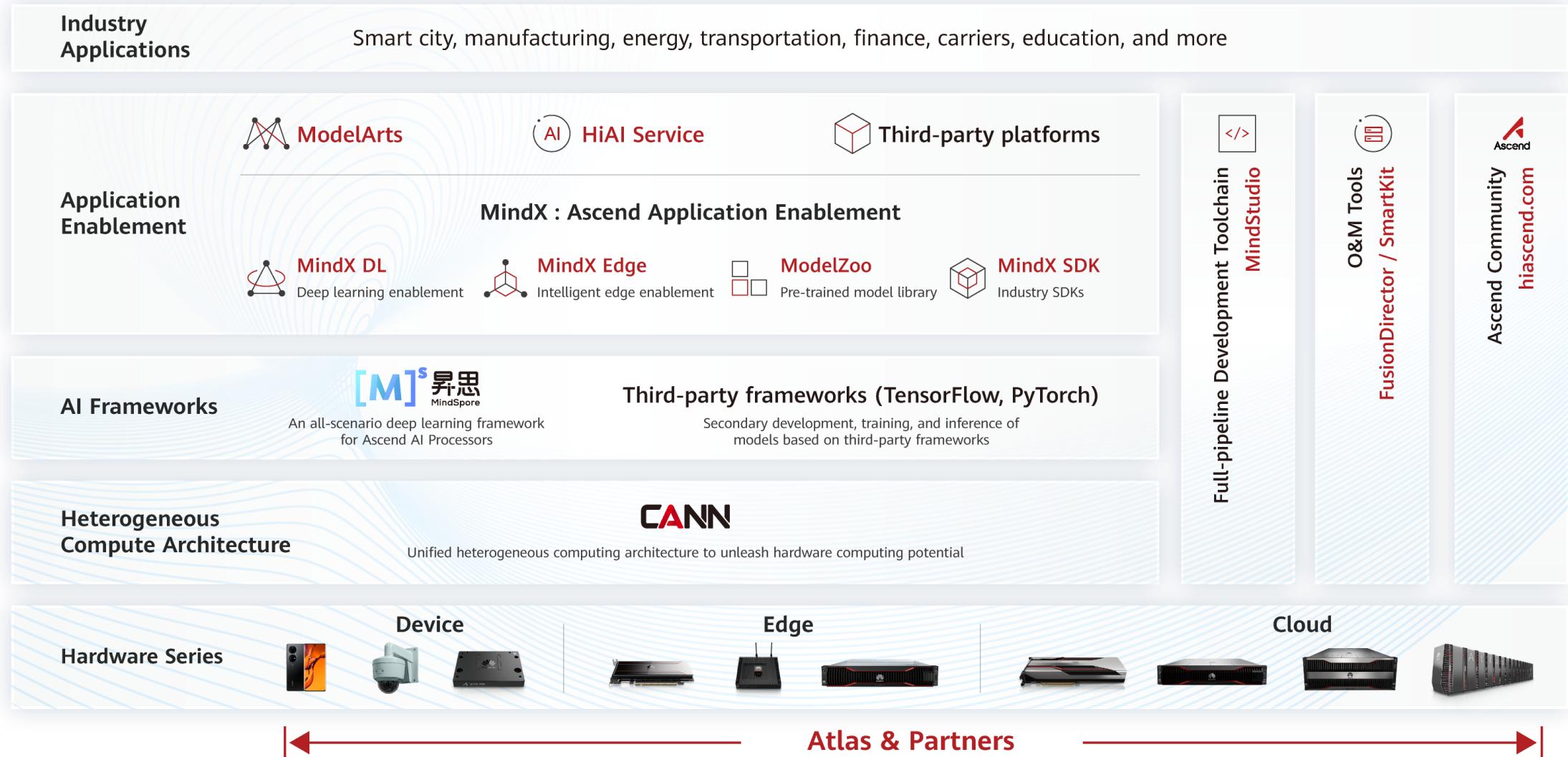
# One Platform + Dual AI Ecosystems

Developing AI foundational technologies to build a new ecosystem

# Ascend AI strategy: Developing AI foundational technologies to build a new ecosystem



# Building a Leading High-Performance AI Computing Platform



# Full Series AI Hardware based on Ascend AI Processor

## Atlas inference series



Atlas 200  
AI accelerator module  
**22 TOPS (INT8)**  
9.5w/25w power consumption



Atlas 500  
AI edge station  
**2U height**  
2 Kunpeng/Intel CPU  
8 Atlas inference card  
**MAX 1120 TOPS (INT8)**



Atlas 800  
inference server



Atlas 300I inference card  
**NEW TBD**



**140 TOPS (INT8)**  
70 TFLOPS (FP16)  
72w power consumption

88 TOPS (INT8)  
44 TFLOPS (FP16)

67w power consumption



Atlas 300T AI training card  
**Industry-leading** single-card computing power

**280 TFLOPS (FP16)**  
150w power consumption



Atlas 800  
training server



Atlas 900 PoD

4U height  
4 KunpengCPU/ 2 Intel CPU  
8\*32G HBM + 32\* DDR4 32/64GB  
8 Atlas training card integrated  
**MAX 2.24 PFLOPS (FP16)**  
5.6kw power consumption



Atlas 900 AI cluster  
**MAX Scalable 1.3 EFLOPS FP16**

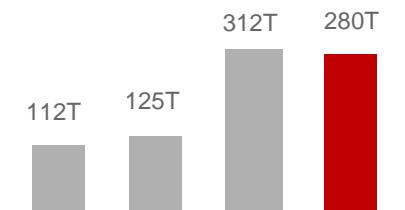
## Atlas 300I Pro: The most powerful inference card for video analysis

Computing power per card vs peers: 2.1x higher



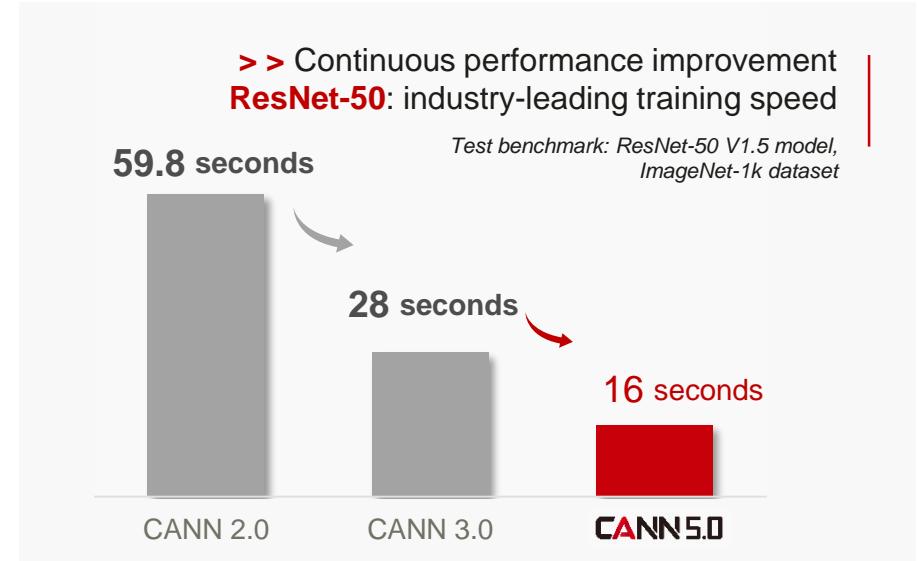
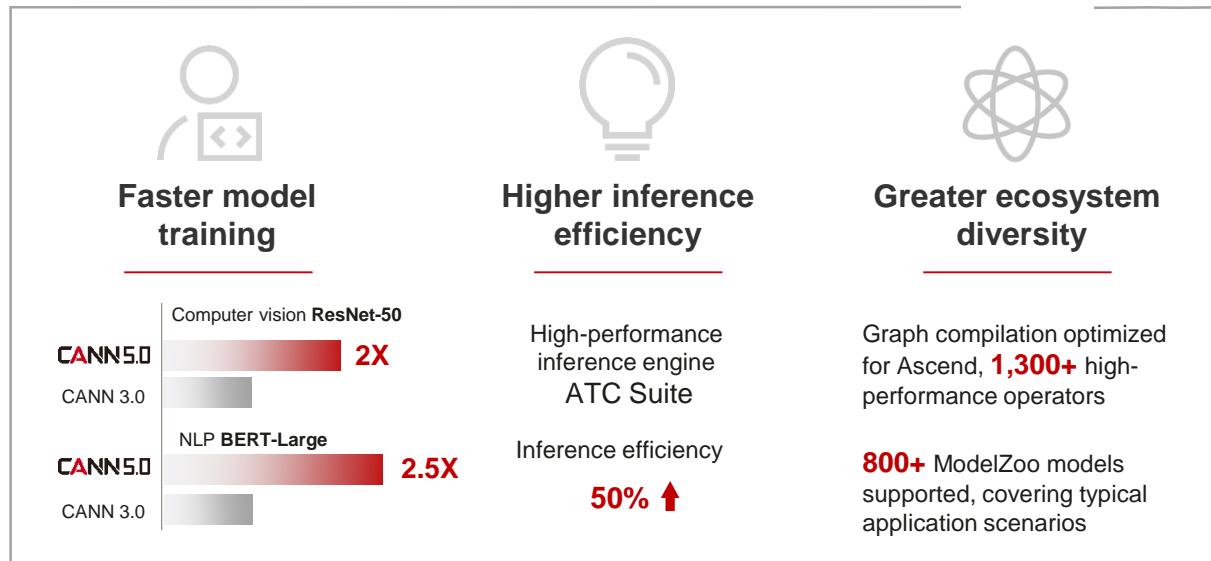
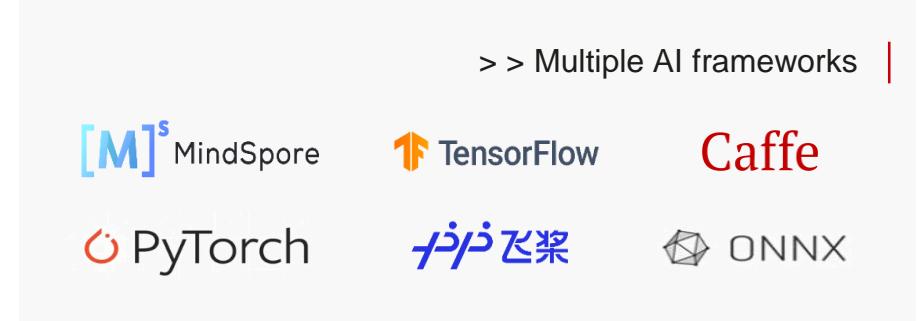
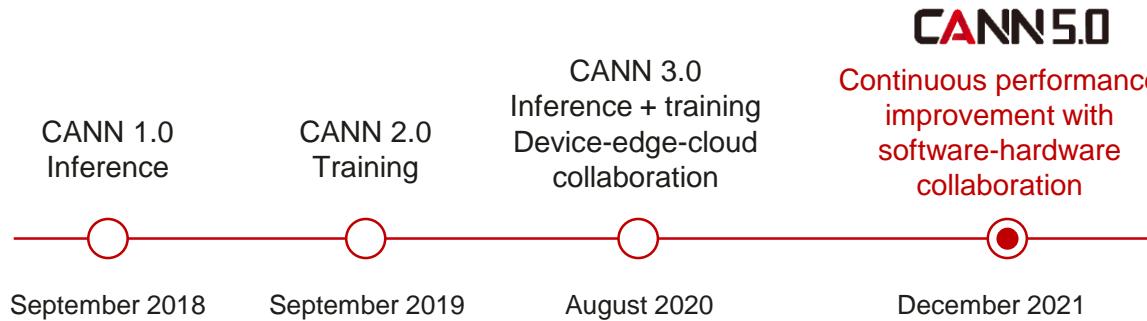
Channels of video analysis per card vs peers: 3x more

## Atlas 300T: Performance is in the first tier

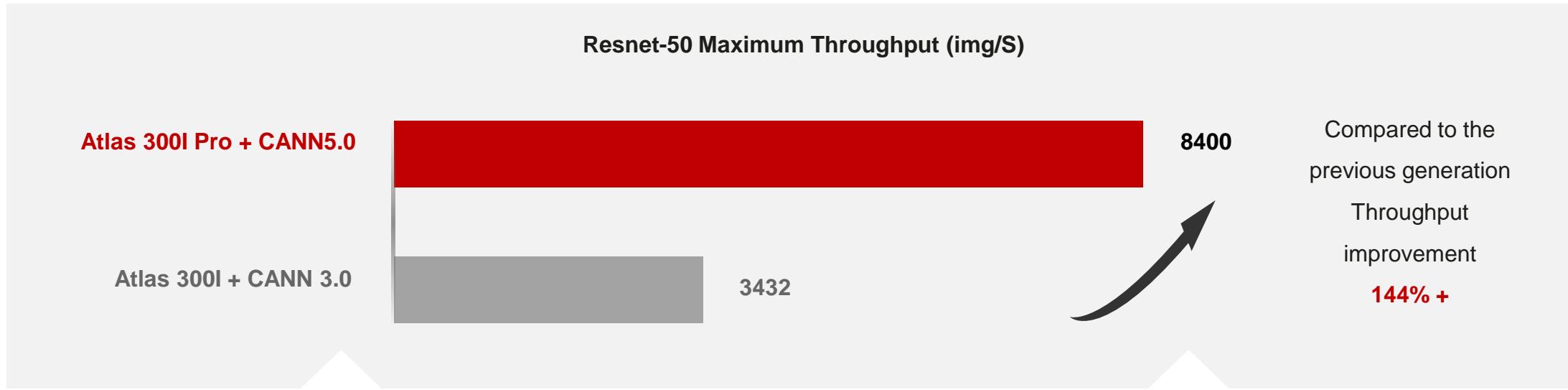


Peer AI training cards      **Atlas 300T**

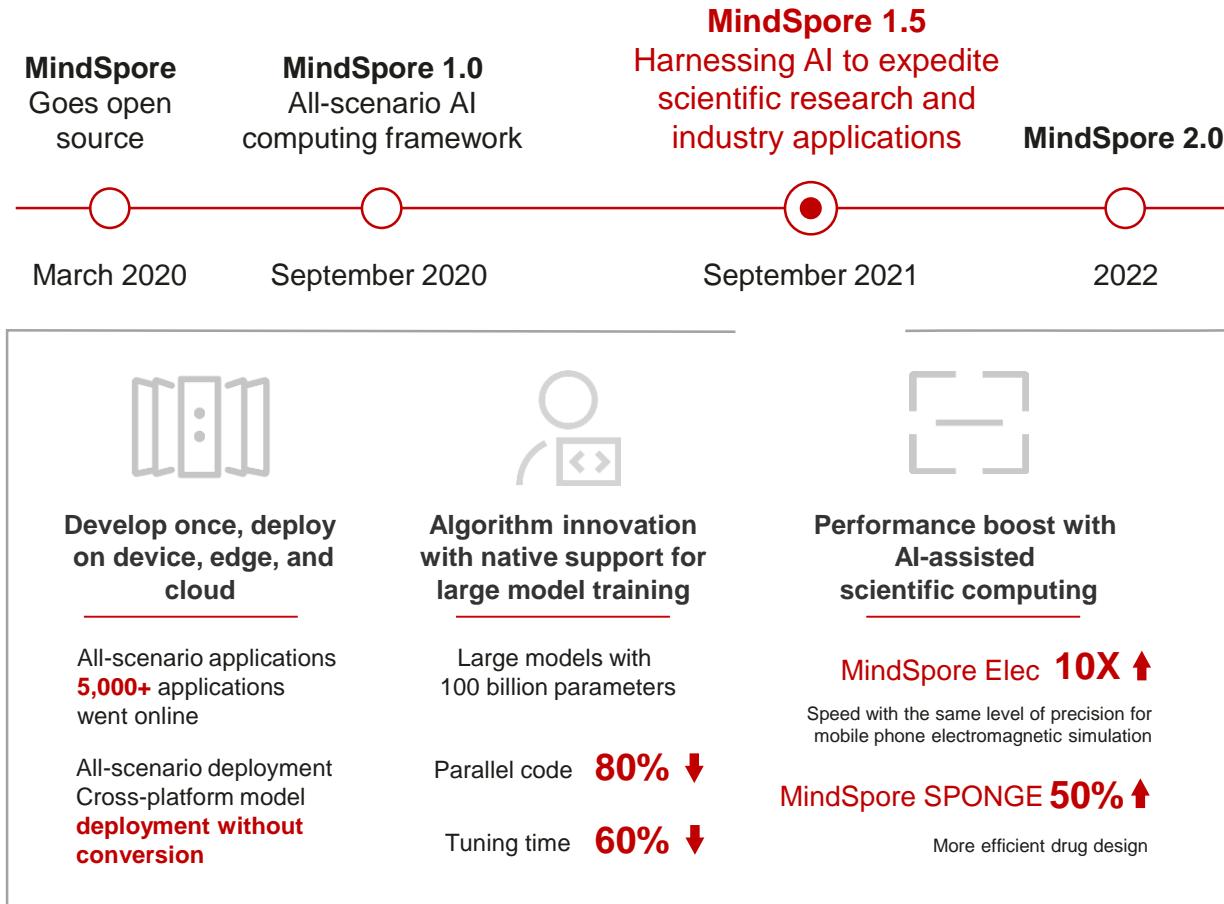
# Heterogeneous Compute Architecture CANN: Unleashing Computing Potential through Software and Hardware Collaboration



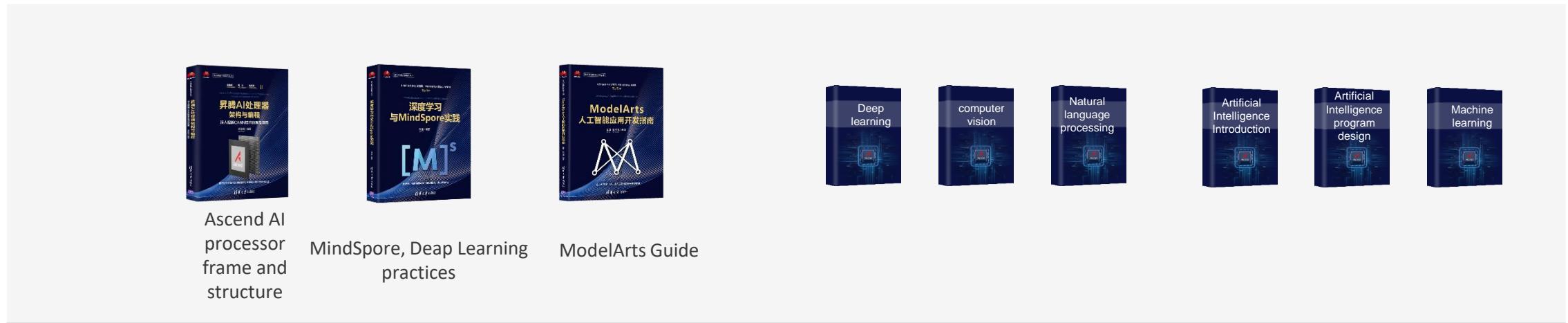
# Software and hardware synergy, Atlas 300I Pro+CANN 5.0 doubles performance



# Open Source AI Framework MindSpore: Building an Ecosystem for Device-Edge-Cloud Scenarios



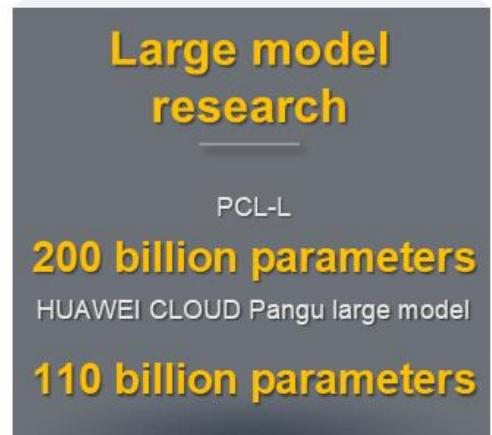
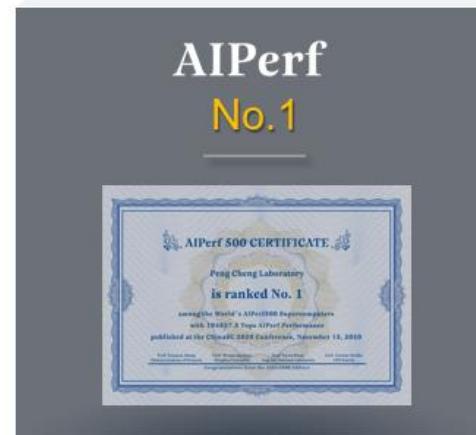
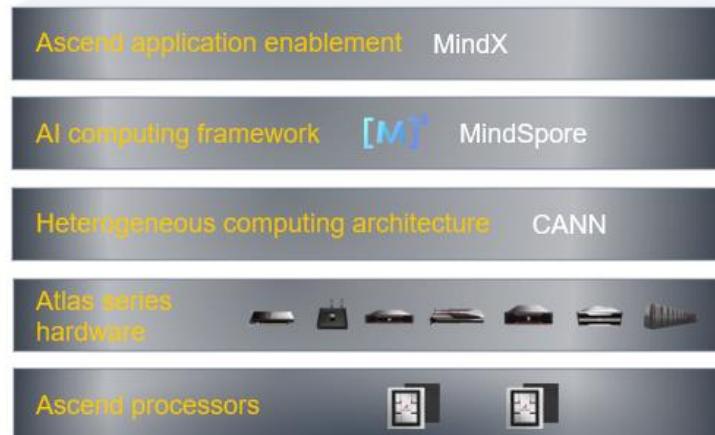
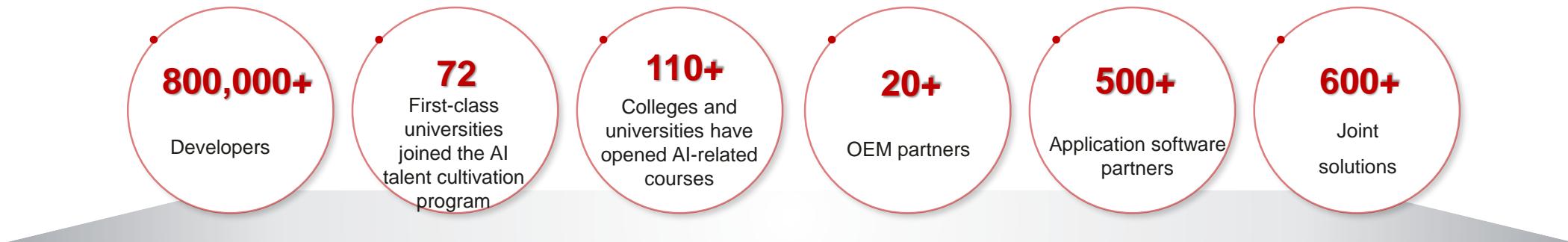
# Cooperate with universities to create education and training material to support talent cultivation and development in AI



Cooperation with Top Universities in Europe in AI



# Creating a Vibrant Ascend AI Ecosystem

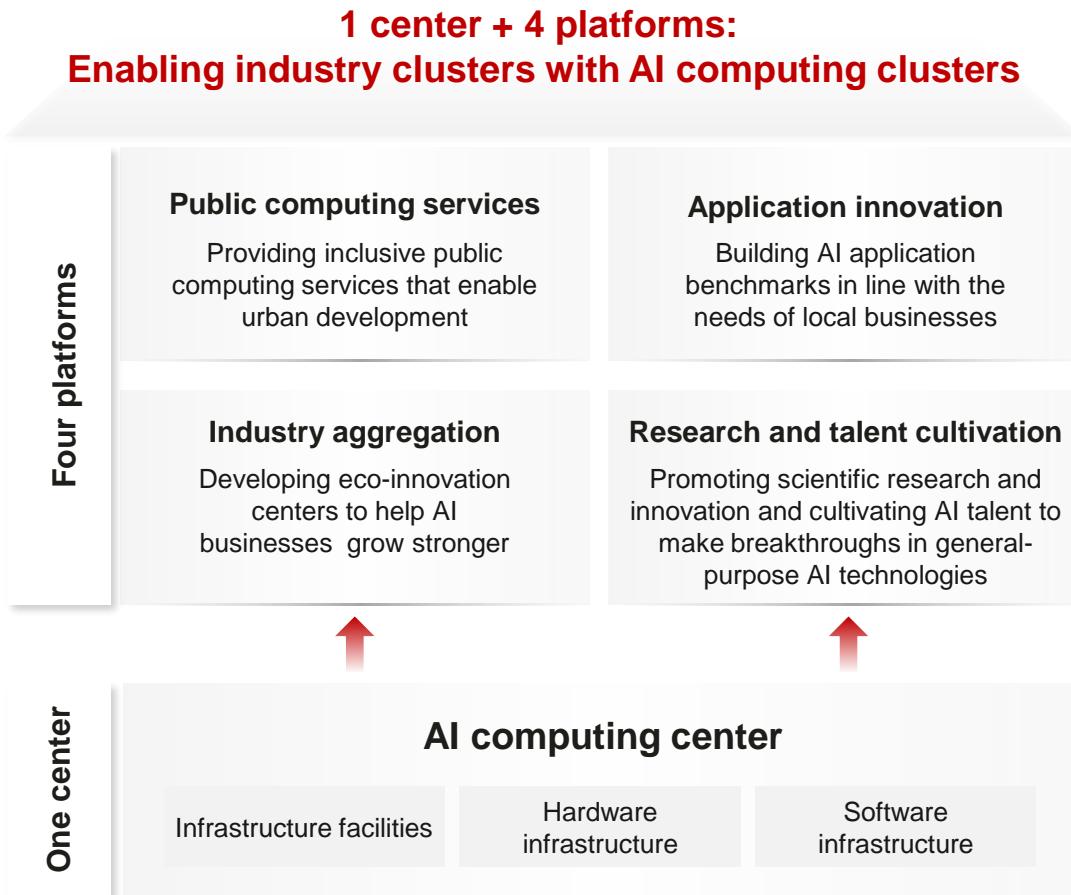


# Enable Industries

Ascend Turbocharges Intelligent Transformation

HUAWEI

# Ascend AI: Enabling the construction of AI computing centers to consolidate the foundation of the digital economy



**More than 20 cities** are actively planning and building AI computing centers



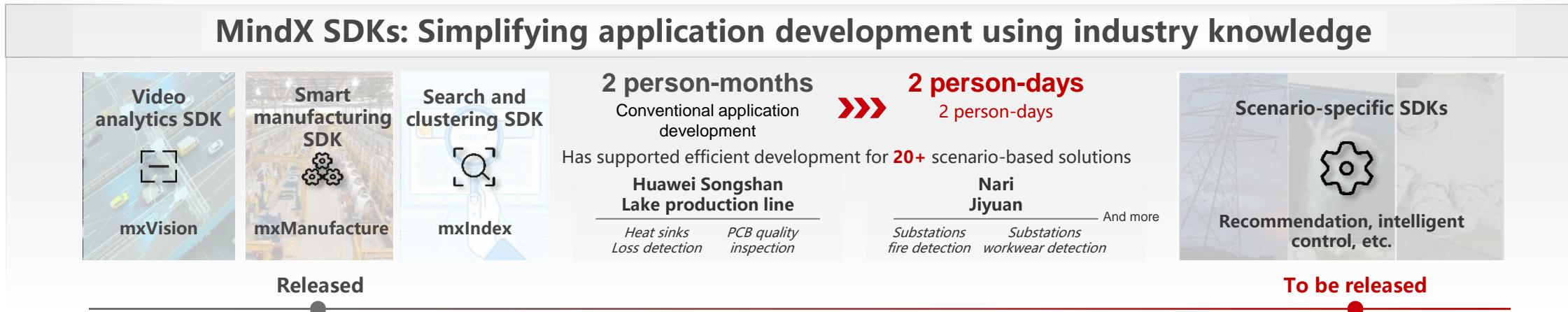
**AI computing power will become a new type of public resource in the digital economy, bringing intelligence closer**

# Empowering partners to accelerate AI application innovation in various industries

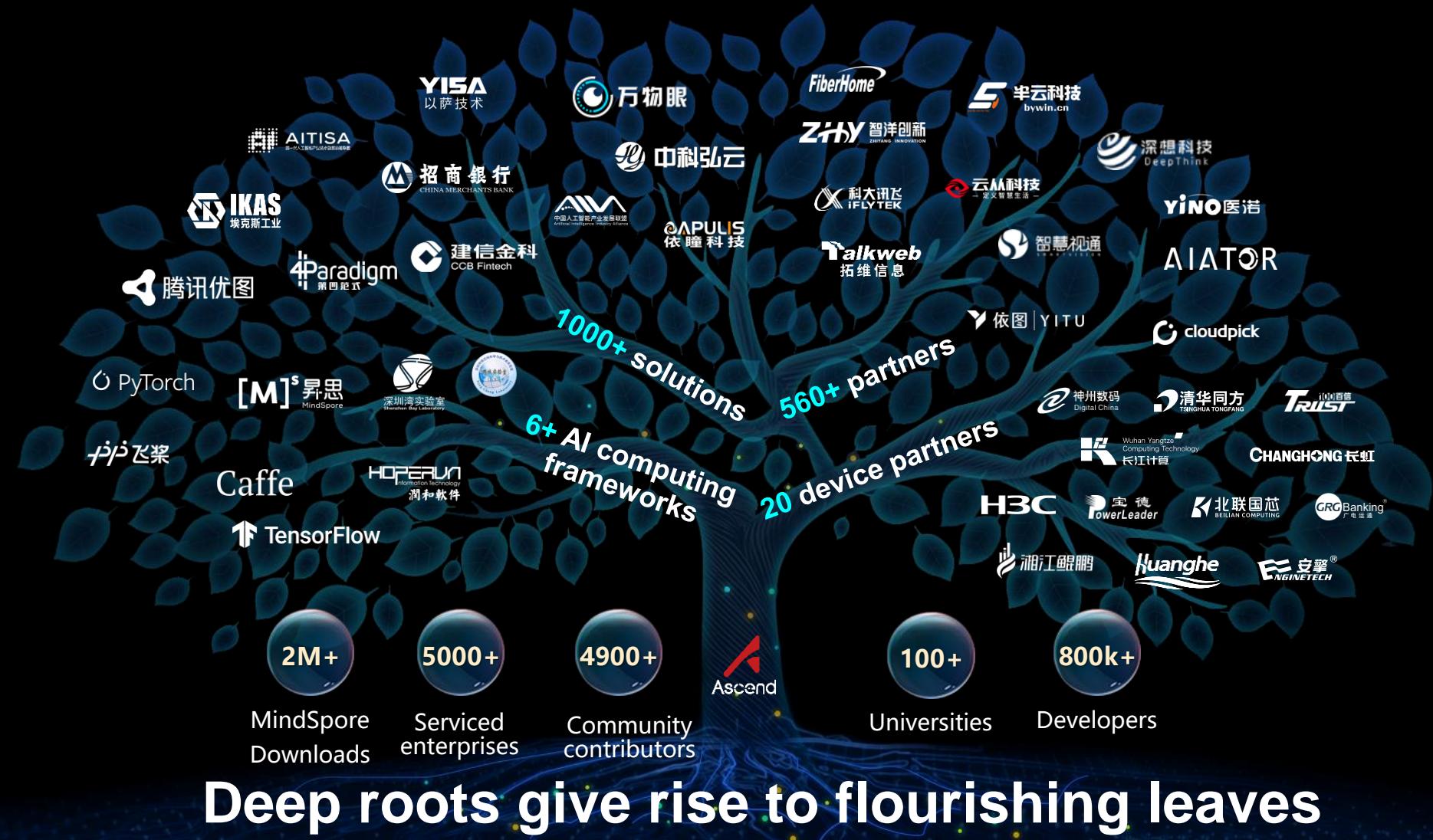
## Top customers/partners: Setting application benchmarks with Ascend-based innovation

 <b>中国移动</b> Cloud album 6,000 monthly active users for mCloud	 <b>intellifusion</b> 云天励飞 Smart Longgang in Shenzhen, China 140+ algorithms to improve service capabilities	 <b>招商银行</b> CHINA MERCHANTS BANK Smart bank branch 1,000+ branches in China	 <b>中国工商银行</b> INDUSTRIAL AND COMMERCIAL BANK OF CHINA Financial brain Explainable AI recommendation system	 <b>兴业银行</b> INDUSTRIAL BANK CO.,LTD. Financial superbrain Risk control, OCR, and audit
 <b>字节跳动</b> Ultimate training performance Typical model performance is twice that of the highest in the industry	 <b>中国南方电网</b> CHINA SOUTHERN POWER GRID AI inspection of power transmission lines Market share: 70%	 <b>ADLINK TECHNOLOGY INC.</b> 凌华科技 190+ production lines in the Huawei Songshan Lake factory Intelligent quality inspection at workstations, lines, and workshops	<b>20+ provinces in China</b> Cancellation of provincial border road toll charges Market share: 80%	<b>19 provinces in China</b> Highway video cloud Market share: 70%+

## MindX SDKs: Simplifying application development using industry knowledge



# Enabling Partners and Cultivate Talent with Open Hardware and Open Source Software



# Thank you.

把数字世界带入每个人、每个家庭、  
每个组织，构建万物互联的智能世界。

Bring digital to every person, home and  
organization for a fully connected,  
intelligent world.

Copyright©2018 Huawei Technologies Co., Ltd.  
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

