Atlas

AI Computing Platform

Break through boundaries

Atlas 200	Atlas 500	Atlas 300	Atlas 800	Atlas 200 DK
AI Accelerator Module	AI Edge Station	AI Accelerator Card	Deep Learning System	AI Developer Kit
		Inference in DC	Training in DC	
Device	Edge	Cloud		Ecosystem

Release models
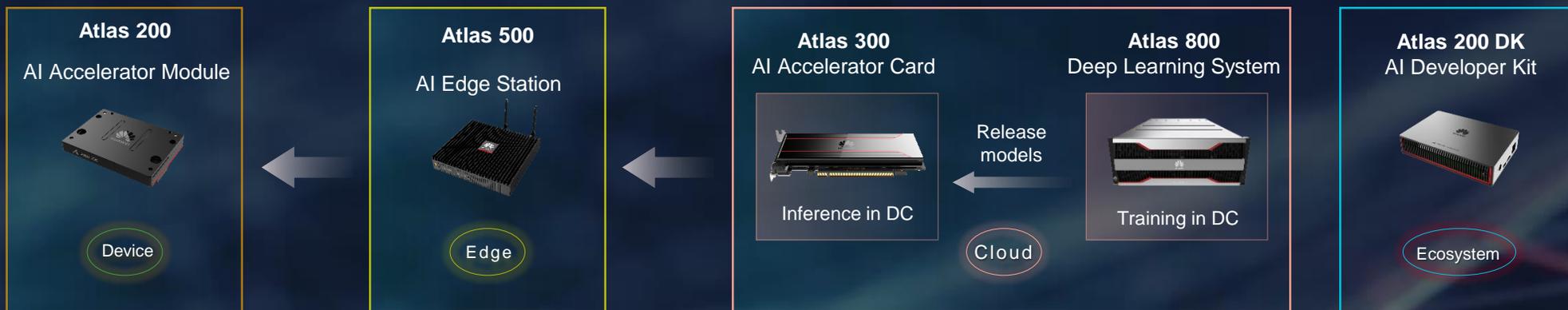
Break through compute power

Chip — Da Vinci Architecture — Ascend 310 — Ascend 910

Ascend AI Chip

Ascend 310

AI SoC with ultimate efficiency

Ascend-Mini

Da Vinci Architecture

- Half precision (FP16): 8 TFLOPS
- Integer precision (INT8): 16 TOPS
- 16-channel full-HD video decoder: H.264/265
- 1-channel full-HD video encoder: H.264/265
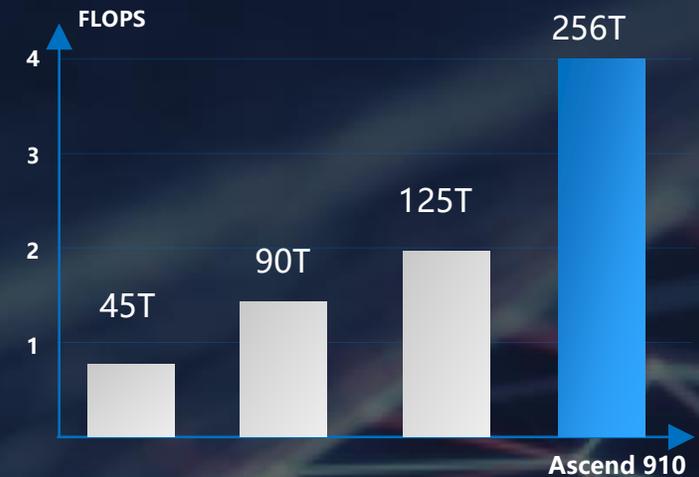- Max. power consumption: 8 W
- 12nm

Ascend 910

Highest compute density on a single chip

Ascend-Max

Da Vinci Architecture

- Half precision (FP16): 256 TFLOPS
- Integer precision (INT8): 512 TOPS
- 128-channel full-HD video decoder: H.264/265
- Max. power consumption: 350 W
- 7nm

FLOPS

256T

125T

90T

45T

4

3

2

1

Ascend 910

* Normalized to 16-bit

HUAWEI

Atlas 200 AI Accelerator Module

Superb Compute in a Compact Size

52 mm x 38 mm x 10 mm

- 16 TOPS of INT8
- 16-channel HD video real-time analytics, JPEG decoding
- 4 GB/8 GB memory, PCIe 3.0 x4 interface
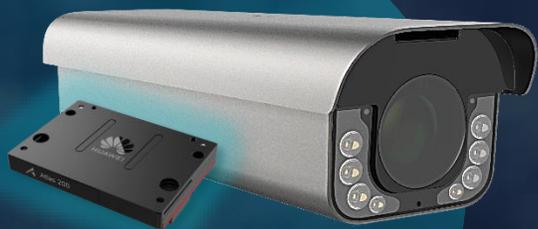- Operating temperature: -25°C to +80°C

Cameras   Drones   Robots

HUAWEI

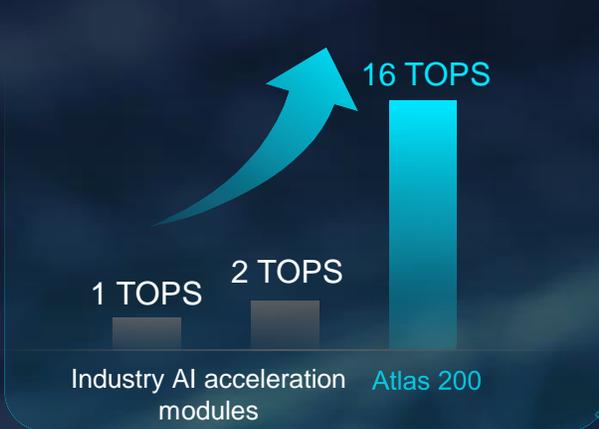# Atlas 200 Brings an 7x Leap to AI Camera Performance

Runs multiple algorithms in parallel

Pixel-Level Image Segmentation
✓ Facial recognition
✓ Human body identification
✓ Vehicle identification

Super Powerful AI Accelerator Module

16 TOPS

2 TOPS

1 TOPS

Industry AI acceleration modules

Atlas 200

Facial recognition per frame 30 ➡ 200

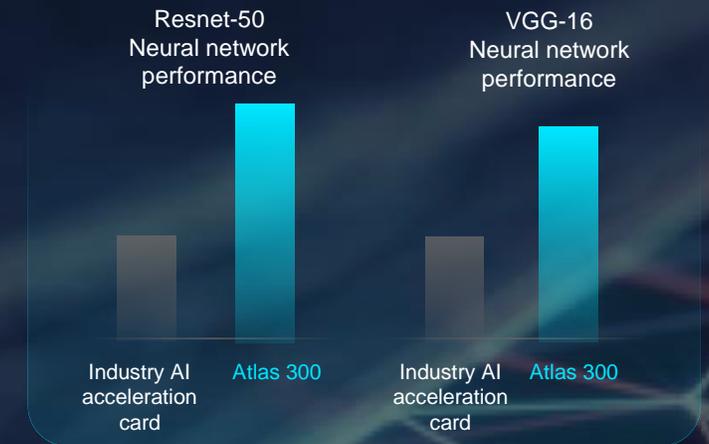● Models in the demo are Huawei employees.

HUAWEI

Atlas 300 AI Accelerator Card

Video Analytics | OCR | Speech Recognition | Precision Marketing | Medical Image Analysis

- 64 TOPS of INT8 @ 67 W

- 32 GB memory

- 64-channel HD video real-time analytics

- Standard half-height half-length PCIe card form factor, applicable to general-purpose servers

Industry-leading Performance

Resnet-50 Neural network performance

VGG-16 Neural network performance

Industry AI acceleration card

Atlas 300

Industry AI acceleration card

Atlas 300

HUAWEI

Atlas 500 AI Edge Station

Edge Intelligence

- Powerful compute
- Cloud collaboration
- Easy deployment
- Small footprint
- Low power
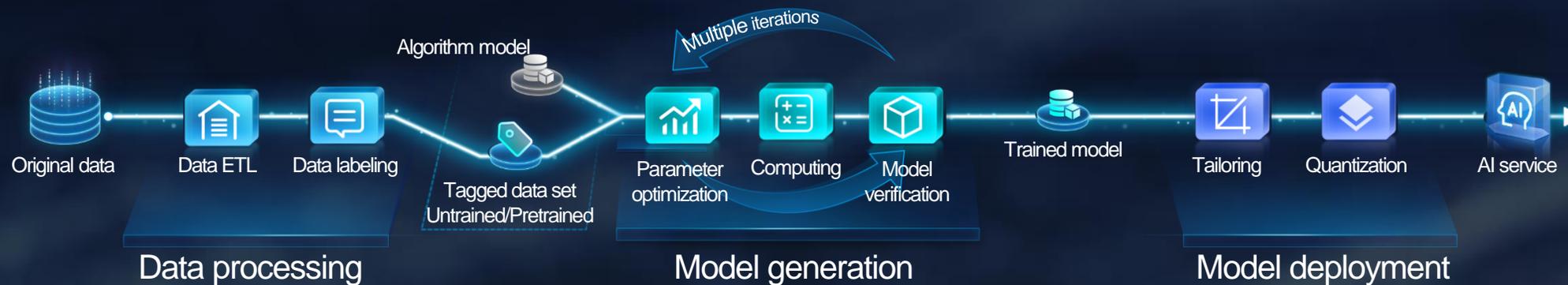
- 16 TOPS of INT8
- 25–40 W
- Wi-Fi & LTE
- 16-channel HD video real-time analytics
- Fanless design, running stably in -40°C to +70°C environments

Built for Harsh Edge Deployment Environments

Extreme cold | Extreme hot | Heavy dust

HUAWEI

Atlas 800 Deep Learning System

Algorithm model

Multiple iterations

Original data → Data ETL → Data labeling → Tagged data set Untrained/Pretrained → Parameter optimization → Computing → Model verification → Trained model → Tailoring → Quantization → AI service

Data processing   **Model generation**   **Model deployment**

2 hours   Service rollout | Out-of-the-box

HUAWEI

Atlas 200 DK AI Developer Kit

Quickly Build Development Environment in 30 Minutes

- 16 TOPS of INT8 @ 24 W
- 1 USB type-C, 2 camera interfaces, 1 GE port, 1 SD card slot
- 4 GB/8 GB memory

Quickly build development environment in 30 minutes

User-friendly GUI and a programming environment with graphical debugging capabilities

One-time development for device-edge-cloud all-scenario deployment

AI development kit with industry-leading computing capabilities

Powerful AI Developer Kit

16 TOPS

2 TOPS

0.4 TOPS

Industry AI developer boards

Atlas 200 DK AI Developer Kit