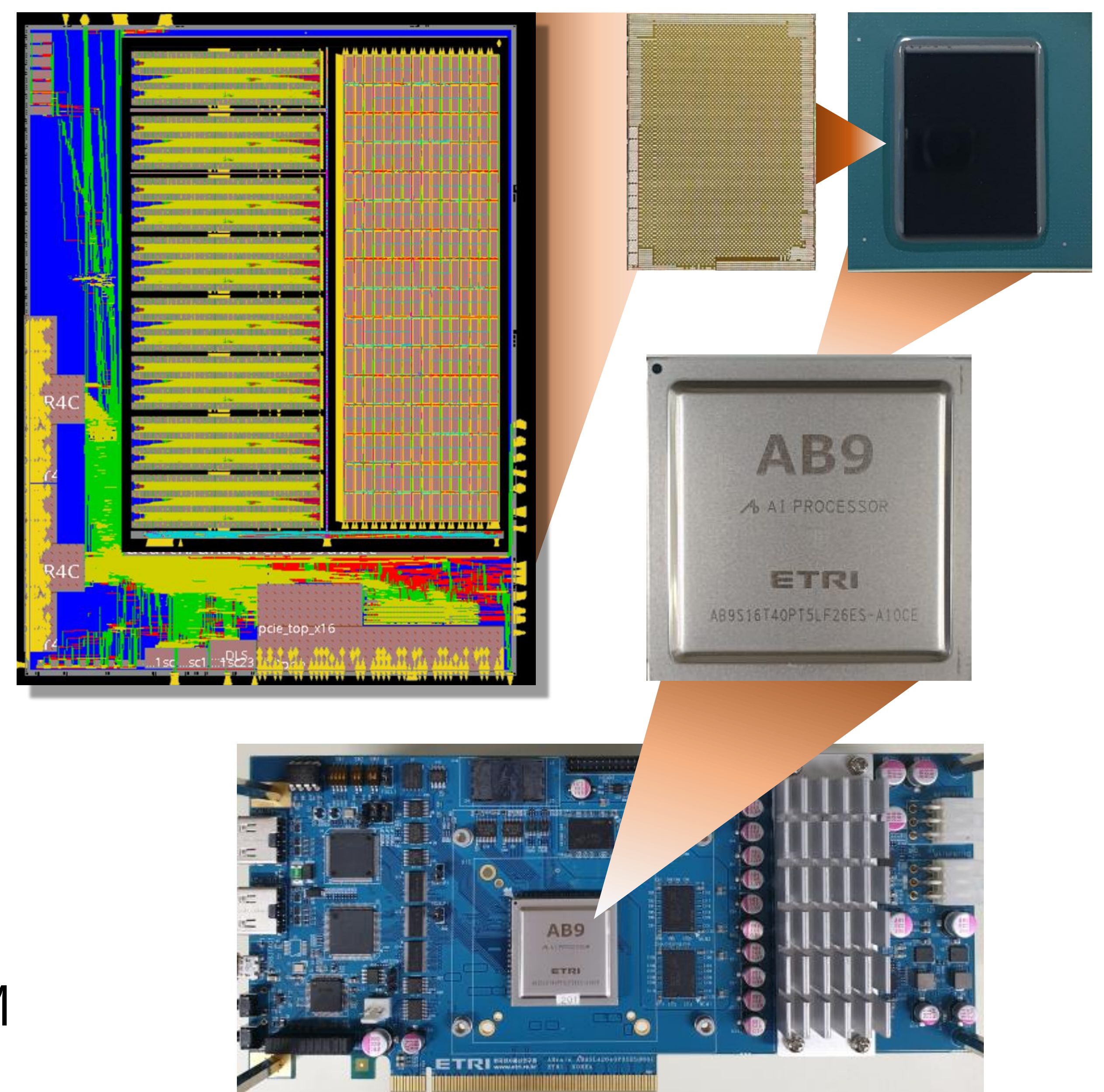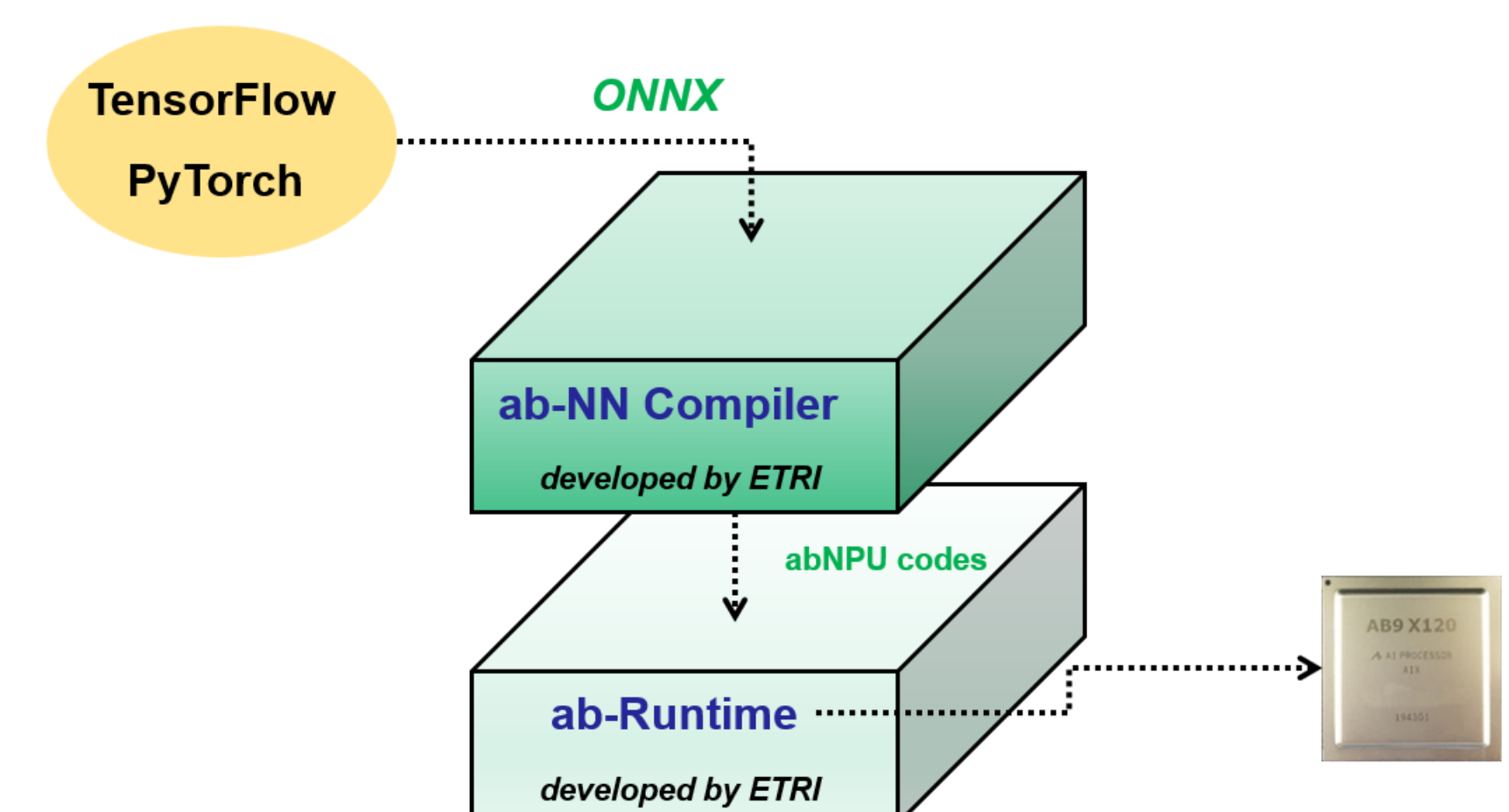# 1.25GHz 40-TFLOPS AI Processor AB9

Introduction

▶ **40-TFLOPS artificial intelligence processor implemented with 28nm CMOS process**
  - High performance deep-learning based neural network accelerator AB9
  - Effective compiler AB-NN for the high utilization and power management

Technology

▶ **AB9 Specification**
  - Performance: 40TFLOPS, 15W
  - Die Area: 17mm x 23mm ($391mm^2$)
  - Size: 1-billion transistors@28nm
  - Core frequency: 1.25GHz
  - 40 MB unified kernel/feature memory
  - 32,768 arithmetic units
  - Floating-point 16-bit (IEEE-754)
  - 16 lanes PCIe Gen3 interface
  - 2 channels LPDDR4 controller

▶ **AI-Ware**
  - Effective deep-learning compiler AB-NN
  - Real-time power/temp.-manager AB-PTM

▶ **High efficiency achievement**
  - 1.3 TOPS/W, 65.8 GOPS/mm^2
  - YOLO v2, YOLO v3, MLP, LSTM support

< AB9 Layout, Die, Chip, Chip Board >

< AI-Ware Hierarchy >

Applications

▶ **Large-scale AI Server, Autonomous Driving Car, Drone**

ETRI 한국전자통신연구원 www.etri.re.kr

기술
문의

AI SoC RESEARCH DIVISION / AI PROCESSOR RESEARCH SECTION
Jinho Han / (042) 860-6558 / E-Mail : soc@etri.re.kr
Youngsu Kwon/ (042) 860-5244 / E-Mail : yskwon@etri.re.kr