ENABLING AI Infrastructure
Safe Harbor Statement

This presentation contains forward-looking statements (including within the meaning of Section 21E of the United States Securities Exchange Act of 1934, as amended, and Section 27A of the United States Securities Act of 1933, as amended) that are based on current expectations and beliefs of Broadcom’s management, current information available to Broadcom’s management, and current market trends and market conditions and involve risks and uncertainties, that may cause actual results to differ materially from those contained in forward-looking statements. Accordingly, undue reliance should not be placed on such statements.

Our filings with the SEC, which are available without charge at the SEC’s website at https://www.sec.gov, discuss some of the important risk factors that may affect our business, results of operations and financial condition. Actual results may vary from the estimates provided. We undertake no intent or obligation to publicly update or revise any of the estimates and other forward-looking statements made in this presentation, whether as a result of new information, future events or otherwise, except as required by law.
Agenda

- **Enabling AI Infrastructure:** Charlie Kawwas
- **Scalable AI Networks:** Ram Velaga
- **AI Server Interconnect:** Jas Tremblay
- **Optical Interconnects:** Near Margalit
- **Foundational Technology for AI Interconnect:** Vijay Janapaty
- **Custom AI Accelerators:** Frank Ostojic
- **Closing Remarks:** Charlie Kawwas
ENABLING AI Infrastructure

Charlie Kawwas, Ph. D.
President, Broadcom

OPEN // SCALABLE // POWER EFFICIENT
Heritage of **Technology Innovation**

**Semiconductors**
- AT&T Bell Laboratories
  - Est. 1960
- LSI Logic
  - Est. 1981
- Broadcom
  - Est. 1991

**Infrastructure Software**
- HP
  - 1961
- Avago Technologies
  - 2005
- Broadcom
  - 2009
- Brocade
  - Est. 1995
- VMware
  - Est. 1998
- CA Technologies
  - Est. 1976
- Symantec
  - Est. 1982

**Open // Scalable // Power Efficient**

Broadcom Proprietary and Confidential. Copyright © 2024 Broadcom. All Rights Reserved. The term “Broadcom” refers to Broadcom Inc. and/or its subsidiaries.
Semiconductors

<table>
<thead>
<tr>
<th></th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
</tr>
</thead>
<tbody>
<tr>
<td>FY23 Revenue</td>
<td>$17B</td>
<td>$17B</td>
<td>$20B</td>
<td>$26B</td>
<td>$28B</td>
</tr>
<tr>
<td>2019-2023 CAGR</td>
<td>13%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual R&amp;D spend</td>
<td>$3B+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Technology Leadership → Organic Growth
Strategy

Market
10 Year Horizon

Technology
Leadership

Execution
Excellence

OPEN // SCALABLE // POWER EFFICIENT
Enabling AI infrastructure

17 Franchises

Networking
Industrial
Broadband
Wireless
Storage

OPEN // SCALABLE // POWER EFFICIENT
Two AI Markets

Consumer AI
Search, Social Networking, ...

Enterprise AI
Public Cloud, On-Prem, ...
Consumer AI
Search, Social Networking, …

Enterprise AI
Public Cloud, On-Prem, …

AI Accelerators (XPU)
Custom

AI Connectivity
Merchant

Merchant
Broadcom’s AI Journey…

AI Revenue as % of Semiconductor Solutions Revenue

2019: <5%
2020: <5%
2021: <5%
2022: ~10%
2023: ~15%

FY2024 Target: 25%
... is Accelerating

<table>
<thead>
<tr>
<th></th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
<th>2024</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI Revenue as % of Semiconductor Solutions Revenue</td>
<td>&lt;5%</td>
<td>&lt;5%</td>
<td>&lt;5%</td>
<td>~10%</td>
<td>~15%</td>
<td>35%</td>
</tr>
</tbody>
</table>

$10B+

Revised FY2024 Target
# Multi-Generational strategic engagements

<table>
<thead>
<tr>
<th>Customer #1</th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Customer #2</th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Customer #3</th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ramp → Production</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
The Build-Out is On…

# of XPU Nodes per Cluster

- 4K+ (2022)
- 10K+ (2023)
- 30K+ (2024)
- 1M? (2027+)

OPEN // SCALABLE // POWER EFFICIENT
Scale Up

CPU

CPU

XPU

XPU

XPU

XPU

Switch
Interconnect

Scale Up
Interconnect

Open // Scalable // Power Efficient

AI Server0

CPU

NIC

PCIe Switch

XPU

NIC

PCIe Switch

Switch

XPU

CPU

NIC

NIC
Scale Out

OPEN // SCALABLE // POWER EFFICIENT
Scale Up + Scale Out = Back End
Interconnect
1. Power  
2. Complexity  
3. Ecosystem
Open  Scalable  Power Efficient

Sustainable Technology Leadership

Ethernet  PCIe  Optics  SerDes  DSP  Custom
ENABLING AI Infrastructure

Scalable AI Networks

Ram Velaga
General Manager
Core Switching Group
Broadcom

OPEN // SCALABLE // POWER EFFICIENT
New AI Models → 100X+ Scale ➔ Distributed Computing

1 Million
AI Accelerators
THE NETWORK IS THE COMPUTER
“Network I/O is Key for Recommendation Workloads.”

Meta, OCP 2022 Global Summit
What Makes AI Networks Unique?

- Very high bandwidth
- RDMA traffic → Bulk data transfers
- Intermittent data surges
- Straggler data significantly impacts job completion time
- Training jobs run over long durations (hours, days)
Broadcom’s AI Network Solutions

Tomahawk 5 (Leaf) Tomahawk 5 Tomahawk 5 Tomahawk 5

Ramon (Spine) Ramon Ramon Tomahawk 5

GPU

Endpoint Scheduled

Tomahawk 5 (Leaf) Tomahawk 5 Tomahawk 5 Tomahawk 5

Ramon (Spine) Ramon Ramon Jericho3-AI

GPU

Switch Scheduled
Jericho3-AI Fabric: Switch Scheduled Ethernet Network

10% Performance improvement = network more than pays for itself

32,000 AI Accelerators at 800Gbps each
Lowest time spent in networking
Tomahawk: Endpoint Scheduled Ethernet Network

RELENTLESS, UNMATCHED ADVANCEMENT

80x Bandwidth Increase
90+% Energy Consumption Reduction

* G=Gbps, T=Tbps
Introducing Thor2: AI Optimized NIC

Thor2

- 400G high-performance NIC
- High-scale RDMA
- Industry’s lowest power
- Longest reach 100G Serdes
Thor2 Consumption Models

Board

Chiplet

IP
End-to-End High Performance Ethernet AI Network

- Perfectly load balanced fabric
- Programmable E2E Congestion Control
- Zero Impact Failover (ZIF)
- Secure & Multi-tenant fabric
- Efficient Interconnects

Largest Cluster Scale
Ethernet Beats InfiniBand: 10+% Improvement in Job Completion Time

*Source: Broadcom Internal Estimates*
Ethernet Provides 30x Faster Failover than InfiniBand

**Optics Annual Failover Rate**

- 2%

**Failures per Month**

- 15

**30x Reduction**

Fast recovery reduces job completion time

* Typical industry failure rate. ** Assuming 4K node cluster using 9.2K optic modules

Source: Broadcom Internal Estimates
Reducing AI Interconnect Cost and Power

- Extended Reach for Copper Cables
  - 4m+ DAC (2x IEEE spec)
- Linear Pluggable Optics
  - 33% Lower Power
- Co-Packaged Optics
  - Lowest Power & Cost
Ethernet is the De-facto AI Network
## Large Ethernet AI Clusters

<table>
<thead>
<tr>
<th>Company</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>60,000+</td>
</tr>
<tr>
<td>Oracle</td>
<td>30,000+</td>
</tr>
<tr>
<td>Meta</td>
<td>20,000+</td>
</tr>
<tr>
<td>ByteDance</td>
<td>10,000+</td>
</tr>
</tbody>
</table>
Ultra Ethernet: AI at Scale

Incredibly Strong Industry Reception: 55+ companies
Modernizing RDMA…

**Classic RDMA**

- In-order packet delivery
- Go-back-n → inefficient
- No multipathing
- DCQCN → hard to tune

**UltraEthernet**

- Out-of-order placement, in-order message completion
- Selective Ack and retransmit
- Packet-level multipathing
- Configuration-free congestion control
Ethernet Network for AI

- Amazon: Trainium, Inferentia
- NVIDIA: GPU
- AMD: MI
- Microsoft: MAIA
- Meta: MTIA
- Intel: Gaudi
- Google: XPU
- Tesla: Dojo

Ethernet AI Network

Open // Scalable // Power Efficient
Comprehensive Ethernet Portfolio and Ecosystem

Network Control & Automation

Operation System

Hardware Platform

OPEN // SCALABLE // POWER EFFICIENT
ENABLING AI Infrastructure

AI Server Interconnect

Jas Tremblay
General Manager
Data Center Solutions Group
Broadcom

OPEN // SCALABLE // POWER EFFICIENT
Server Interconnect History

PCI Standards Group: Founded in 1992 with 900+ Companies
Server Evolution

Compute Centric

High-Performance Interconnect
AI Server: Open → Options

Accelerators
- AMD
- Broadcom
- Intel
- G
- Microsoft
- Meta
- NVIDIA

Central Processing Unit
- Amazon
- AMD
- Intel

NIC
- Amazon
- Broadcom
- Intel
- NVIDIA

Storage & Memory
- KIOXIA
- Samsung
- Seagate
- Solidigm
- Western Digital
AI Server: Interconnect

- PCIe Switch
  - 144 lanes
  - PCIe Gen 5
  - 120 nanosec latency

- CPU
- NIC
- AI Accelerators
- Storage

Most Trusted PCIe Device
AI Server: PCIe Interconnect

End-to-End AI PCIe Connectivity
**AI Server: PCIe Retimer**

<table>
<thead>
<tr>
<th>Retimer</th>
<th>Competition</th>
<th>Broadcom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>7 / 16nm</td>
<td>5nm</td>
</tr>
<tr>
<td>SerDes</td>
<td>Third Party</td>
<td>In House</td>
</tr>
<tr>
<td>Power</td>
<td>1x</td>
<td>35% Less</td>
</tr>
<tr>
<td>Reach</td>
<td>1x</td>
<td>40% More</td>
</tr>
</tbody>
</table>

Source: Broadcom Internal Estimates

**End-to-End AI PCIe Connectivity**
First to Market for 20 Years … 5 Successful Generations

1B+ PCIe ports shipped
• PCIe Gen 5 Leadership
• Open: PCIe Golden Node
• Scale: 38% more lanes
• Power Efficient: 50% Less

Switch + Retimer
PCle Switching Interconnect Leadership

Accelerating PCle Switch + Retimer Introduction
AI Server: Scale Up

PCle Gen 7 Switch
- Accelerator Fabric Link
  Latency: ~120 nanosecond
- Scale: Hundreds of Nodes

AMD/Broadcom Collaborate on Scale Up Fabrics
AI Server Connectivity

Key Points

- PCIe is key to AI Server OPEN interconnect
- Decades of market leadership in SCALABLE PCIe Switches and Retimers
- Broadcom end to end PCIe delivers 40% greater reach with 50% lower POWER

Source: Broadcom Internal Estimates
ENABLING AI Infrastructure

Optical Interconnects

Near Margalit, Ph. D.
General Manager
Optical Systems Division
Broadcom
The Optical Interconnect for AI
Optical Interconnect for AI

- VCSEL: Multi-mode, short reach
- EML: Single-mode, long reach
- CPO (SiPho): Power and cost leadership
Broadcom Optics Leadership

50M+
High-Speed Communication Lasers
Shipped Annually

10 Trillion
Device Service Hours in the field
with < 1 FIT

1G VCSEL
2004

10G VCSEL
2013

25G VCSEL
2019

50G VCSEL
2023

100G EML
2025

AI Data Center

2.5G EML and DML
AT&T

10G EML and DML
AT&T

50G EML and DML
Avago Technologies

100G EML

1998

OPEN // SCALABLE // POWER EFFICIENT
Broadcom Optics Leadership

- **50M+** High-Speed Communication Lasers Shipped Annually
- **10 Trillion** Device Service Hours in the field with < 1 FIT

Timeline:
- 1G VCSEL
- 2004
- 10G VCSEL
- 2013
- 25G VCSEL
- 2019
- 50G VCSEL
- 2023
- 100G VCSEL
- 2025
- 200G VCSEL

Next Gen CPO
- Bailly CPO
- Next Gen EML
- 200G EML
- 2023
- 50G VCSEL
- 2019
- 10G VCSEL
- 2013
- 1G VCSEL
- 1998

OPEN // SCALABLE // POWER EFFICIENT
Why Co-Packaged Optics (CPO)?

- **Lowest Cost/Bit**: Massive reduction in components and interconnects
- **Power/Performance Leadership**: Eliminate electrical interconnect power
- **Increased Reliability**: Silicon integration with pluggable Laser

### POWER COMPARISON

#### 100G/lane
- **Pluggable**: 14W/800G
- **CPO**: 5W/800G
- **LPO**: 9W/800G

#### 200G/lane
- **Pluggable**: 25W/1.6T
- **TBD**: 8W/1.6T
- **CPO**: 14W/800G

**OPEN // SCALABLE // POWER EFFICIENT**
The Journey to Integration at Scale

Conventional Module Design

Integrated SiPh Module Design

Co-Packaged Optics

Engineering and manufacturing limits to scale

Module Integration = First step to improved scale

Highly Integrated Optical Engines (up to 6.4T)

FIRST to Productize High-Density Photonic Integration
CPO Applications: Ethernet Switch, XPU Interconnect

- **Ethernet Switch**: More than 50Tbps optical connectivity co-packaged with an Ethernet Switch.
- **XPU Interconnect**: Up to 50Tbps optical connectivity co-packaged with a XPU.
51.2T TH5-Bailly CPO available today

- 64 Channel FR4 PIC with Integrated Mux/Demux
- 64 Channel EIC with CMOS Driver and TIA
- Advanced Packaging
- ODM System Integration
Optical Interconnects

Key Points

- Industry leadership on 100G → 200G VCSEL and EML
- First commercial CPO with pluggable lasers
- Unmatched >70% power reduction* and >30% cost savings

*Source: Broadcom Internal Estimates
ENABLING AI Infrastructure

Foundational Technology for AI Interconnect

Vijay Janapaty
General Manager
Physical Layer Products
Broadcom
High Speed Links: 2\textsuperscript{nd} Highest Power & Cost in AI Cluster

- 1M+ XPU/Cluster
- 10M+ High Speed Links
- 2x B/W Increase Every 2 Years
10M+
High Speed Links

Backplane/DAC
Reach: < 5m
Lowest Power
Lowest Cost
SerDes

Optics
Reach: 10m–10 Km
Highest Power
Highest Cost
SerDes
DSPs

Source: Broadcom Internal Estimates
• High-Speed Analog-to-Digital Conversions

• High-Speed Custom Digital-Signal Processing

• Leading-Edge Process Technology Nodes

• Deployment in 100s of Systems

#1 SerDes for 4+ Generations
Industry-Leading 5 nm 112 Gb/s SerDes: Peregrine

- 45+ dB Backplanes
  Higher Fanout

- 4+m DAC Cables
  Lowest Power and Cost for Intra-Rack & Inter-Rack

- Drives CPO Natively
  50+% Power Reduction

- Drives LDO Natively
  33+% Power Reduction

Same SerDes Benefits
Across All Broadcom Products

OPEN // SCALABLE // POWER EFFICIENT
Industry’s Best 200 Gb/s SerDes in 3 nm
Industry’s First: CMOS DSP + Driver + TIA for 100 G/λ PAM4 Optics

Promise from 4-19-2022 Teach-in: CMOS DSP + Driver + TIA

Broadcom iDSPs ⇒ Lowest Power/Cost MM/SM Pluggable Modules
Best-Performing 200 G/λ DSPs in the Industry

1.6T DR8 Modules
- 16x100G to 8x200G
- 212 Gb/s FEC Bypass
- 226 Gb/s w/ Inner FEC
- Integrated Driver

1.6T DR8 Modules
- 8x200G to 8x200G
- 212 Gb/s FEC Bypass
- 226 Gb/s w/ Inner FEC
- Integrated Driver

Integrated DSPs for 1.6T Generation

Performance Result from Partner 1.6T Module

200G iDSPs ⇒ Performance, Power and Latency Leadership
Foundation Technology

Key Points

- 100 Gb/s **BEST-IN-CLASS** Innovative Connectivity
- 200 Gb/s **LEADS** Open Ecosystem
- **DIFFERENTIATION** from Analog + DSP Expertise, R&D Scale
ENABLING AI Infrastructure

Custom Accelerators (XPUs)

Frank Ostojic
General Manager
Custom AI Product Division
Broadcom
#1 in custom silicon for 10 years
Consumer AI Requires Custom Accelerators (XPUs)

- Architecture optimized for internal workloads
- Larger volumes $\rightarrow$ lower power & cost
- Memory & I/O architecture customized for AI workloads
Industry-Leading IP Portfolio for Differentiated XPUss

Optimized Architecture + Lowest Power → Best Performance/TCO

$3B+$ R&D Investment

- SerDes IP
- AI-Optimized NICs IP
- Advanced Pkg
- Jericho3-AI
- Switching IP Cores
- Co-Packaged Optics
- Buffer Memory IP

OPEN // SCALABLE // POWER EFFICIENT
Broadcom IP Enables XPU Critical Components

1. Compute
   - Processing Unit Architecture (Customer Owned)
   - Design Flow & Performance Optimization (Broadcom Owned)

2. Memory
   - HBM PHY, Integration & Performance (Broadcom Owned)

3. Network IO
   - Architecture & Implementation (Broadcom Owned)
   - Full Solution Chiplets (Hardware, Firmware, & Software)

4. Package
   - 2.5D, 3D, & Silicon Photonics Architecture & Implementation (Broadcom Owned); Vertical Integration Advantage
Delivering Multi-Generational Custom AI Accelerators over the Past Decade
Deep, Strategic, Multi-Year Partnership with Consumer AI Customers

10 Years & 10 XPU Co-Development

2014–2023

4 Years & 4 XPU Co-Development

2024–2025

2026 & Beyond

CUSTOMER #1

CUSTOMER #2

OPEN // SCALABLE // POWER EFFICIENT
Industry’s Fastest Time-to-Production with Annual Cadence

**XPU Platform Engineering**

Years prior to Design Start

- Automated Design Flow
  - AI Design Flow Optimized over 10 Years of XPU experience

- AI IP
  - Plug & Play Ready
    - Silicon
    - Firmware
    - Software-API

- SOC Package
  - Pre-Qualified
    - Thermal
    - Mechanical
    - Electrical

**Development Co-Development with Consumer AI Customer**

- 7–9 Month Development
- FAB
- 3-Month Production Ramp

**OPEN // SCALABLE // POWER EFFICIENT**

Broadcom Proprietary and Confidential. Copyright © 2024 Broadcom. All Rights Reserved. The term "Broadcom" refers to Broadcom Inc. and/or its subsidiaries.
Deep & Strategic Collaboration During Architectural Phase

**INVESTMENT PERIOD**
2023-2025

- Early IP investments years before XPU design kick-off across critical components

**ARCHITECTURE PHASE**
2026 & Beyond

- Concurrent technology, IP & packaging investments for next two generations of XPU
Execution Excellence in Delivering to Consumer-AI Requirements

XPU Complexity
Size of the bubble = Complexity of the XPU

F (Complexity) = \{Compute Performance, Network Bandwidth, Memory Bandwidth, Power Delivery, Thermal Integrity, Mechanical Reliability\}
Custom AI Accelerators

Key Points

- Focused on Consumer-AI Custom Accelerators
- $3B+ Investment Scale in Differentiation XPU IP
- 10-year Engineering Excellence in High-Complexity XPUs
- 3-year Deep Architectural Advantage in Co-Development
ENABLING AI Infrastructure

Charlie Kawwas, Ph. D.
President, Broadcom
## XPUs → Multi-Generational Strategic Engagements

<table>
<thead>
<tr>
<th></th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Customer #1</strong></td>
<td>Production</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Customer #2</strong></td>
<td>Production</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Customer #3</strong></td>
<td>Ramp → Production</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Broadest & Industry-Leading AI Networking Portfolio

**Ethernet**

- Broadcom GEMINI 3D Brings 256GB/s Data Rates to Backplane
- Broadcom's First 400G Phy in Production

**PCIe**

- Industry’s First End-to-End PCIe/CXL Connectivity Portfolio
- Broadcom Upgrades Co-Packaged Optics Switch to 512Tbps

**Optics**

- World Leader in AI Infrastructure Optics
- Broadcom Introduces Industry’s First 5nm 100G/400G Optical

**Foundational Technology**

- Broadcom Delivers 200G/400G CXL 7.0 for Next-Generation Switches and AI Networks

**OPEN // SCALABLE // POWER EFFICIENT**
... Enabling AI Infrastructure
Open  Scalable  Power Efficient

Sustainable Technology Leadership
Ethernet  PCIe  Optics  SerDes  DSP  Custom

Investment Scale  Innovation Leadership  Execution Excellence
Thank you
ENABLING AI Infrastructure