



**Tachyum**<sup>TM</sup>

# Universal Processor Semiconductor Company



## Products

Fastest processor for  
cloud servers/HPC/AI

## Value

3x lower cost  
10x lower power

## Initially: AI/HPC



Tachyum ports Linux & GCC

## Our Market: Hyperscale



Existing binaries run via QEMU  
Recompiled applications

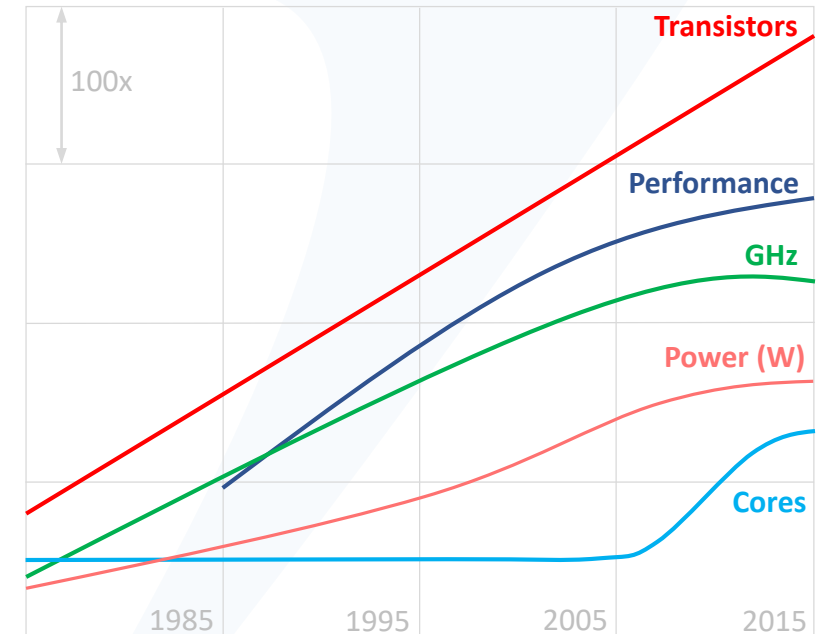
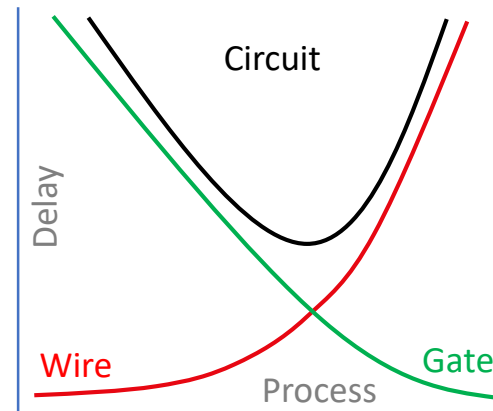
# Performance Plateau: Slow Wires

## Processor Plateau 4.5GHz today

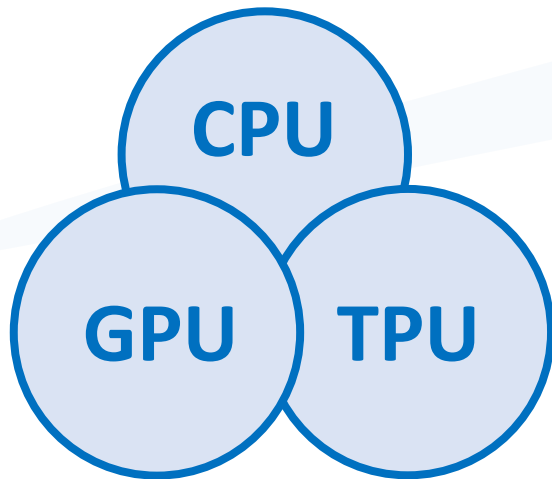
- 2005 Pentium-4 was 3.8GHz
- Transistor speedups 30% per 24 months → should be 20GHz now
- Process shrink makes wires slower

GPUs are only better in a few apps

ARM is too slow



# Universal Processor for Servers / HPC / AI



**Best of**

## Server / Supercomputer / AI Chip

For hyperscale datacenters

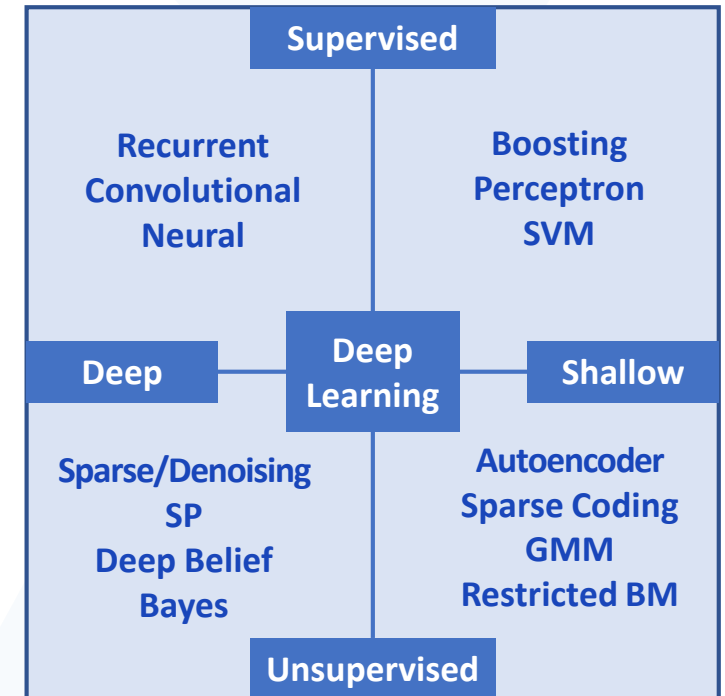
**1<sup>st</sup> time humanity can simulate human brain sized AI in real-time**

## Best of CPU, GPU and TPU

CPU: easy to program but costly

GPU: faster but harder to program

TPU: faster but very limited



# Solution: Fast & 10x More Efficient Processor



## Faster than Xeon but smaller than ARM processors

Short wires reduce wire delays

Out-of-Order execution with compiler vs. hardware

## Existing applications can run without changes

Recompile to exceed Xeon performance

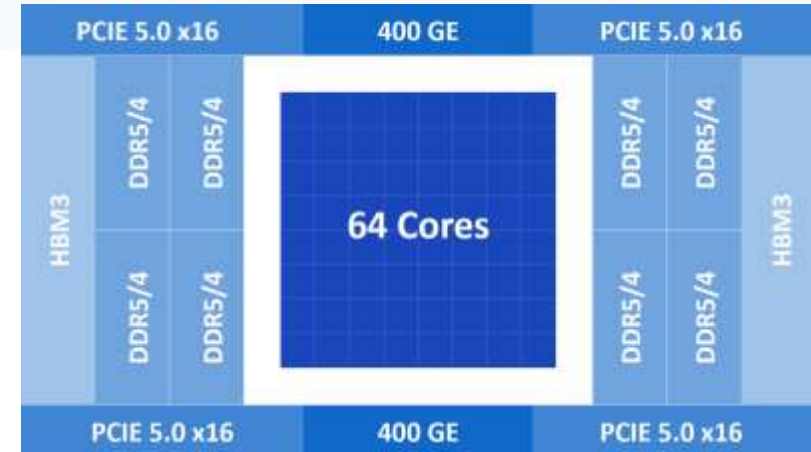
Dynamic binary translation for existing binaries

## 64 cores, each faster than Xeon Core

4GHz, 8 DDR4/5, 72 PCIE 5.0, 2x 400G Ethernet

4 HBM2e, 32MB coherent cache

7nm FinFET process with no custom design



# Human Scale AI: Capital Expenditure Free

## 50-70% of server time wasted

<5% servers have AI GPU (CPUs are too slow)

## Universal Processor / HPC / AI Chip

Enables 10x more AI using idle servers

## Rat → Human Brain Needs 1,000x Speedup

256,000 Tachyum chips deliver  $10^{19}$  FLOPS

## >100 Brain-Capable Datacenters

Facebook has 442,368 servers in datacenters

40% utilization means 256,000 idle servers



# Chips Shipping in 2020

Faster & 10x lower power processor

3x lower cost in \$24B market growing 15% annually

**Forbes** Human Brain Sized AI : Coming Soon To a Cloud Data Center Near You



Solutions to Unsustainable Energy Consumption of Data Centers and Brain Scale AI

**HPC** WIRE coming onto the HPC / hyperscale scene with intriguing & bold claims



1<sup>st</sup> real-time human brain sized AI

**MICROPROCESSOR** *report*

massive profits available in the server market, Tachyum is taking educated risk



attractive proposition for hyperscale, which could build a single architecture that ...



2018 Winner



2018 CESA