

# High Performance Computer Architecture

## Elec/Comp 526

### Spring 2019

- **Instructor**

Peter Varman

DH 2022 (Duncan Hall)

Email: pjv @ rice.edu

Tel: x3990

- **Office Hours**

Tue/Thu 2:00 – 4:00 pm or

By Appointment

- **Web Page**

[www.canvas.rice.edu](http://www.canvas.rice.edu) --- COMP/ELEC 526 Spring 2019

# Course Organization

- **Grading**
  - 5-6 Assignments (C-based Simulation Projects + HWK sets): 60%
  - Individual Project 15%
  - 2 Test . (Evening or In class): 25%
- **Course Material**
  - **Fundamentals of Parallel Multicore Architecture, Yan Solihin**
  - Parallel Computer Organization and Design, M. Dubois, M. Annavaram, P. Stenström
  - Computer Architecture: A Quantitative Approach, (5th edition), J. Hennessy and D. Patterson
  - Survey and Research Papers (Will be posted as needed)

**\*\* Will schedule 3-4 evening classes to replace in-class meetings.**

## Honor Code

All work is given under the (letter and spirit of the) Rice University Honor Code system.

## Accommodations

Any student with a disability requiring accommodations in this course is encouraged to contact me after class or during office hours. All discussions will remain as confidential as possible. Additionally, students should contact Disability Support Services.

# Goals

Develop an advanced understanding of modern parallel computer architectures and systems

Principles underlying high-performance computers

- Single Threaded and Multicore processors
- High-performance memory subsystems
- Storage systems
- Networking



## Dell Multi Core Servers based on Intel Xeon E7 (4/2)

- High degree of **implicit parallelism**
- Small degree of **explicit parallelism**



## Dell Multi Core Servers based on Intel Xeon E7 (4/2)

- High degree of **implicit parallelism**
- Small degree of **explicit parallelism**

Processor: 2x Intel Xeon Silver 4110 Processors 2.1GHz (3.0GHz Turbo Boost) 85W 11MB Cache 16 Cores (32 Threads Total)

The Aurora supercomputer at Argonne National Labs will use over 50,000 computing nodes based on Intel's Xeon Phi processors in Cray XC chassis for over 180 petaflops. (Source: Oak Ridge National Labs)

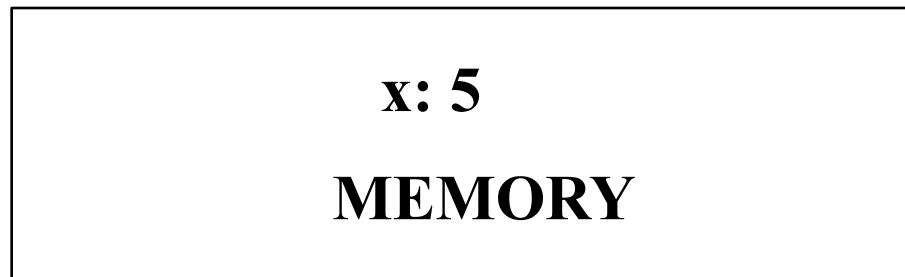
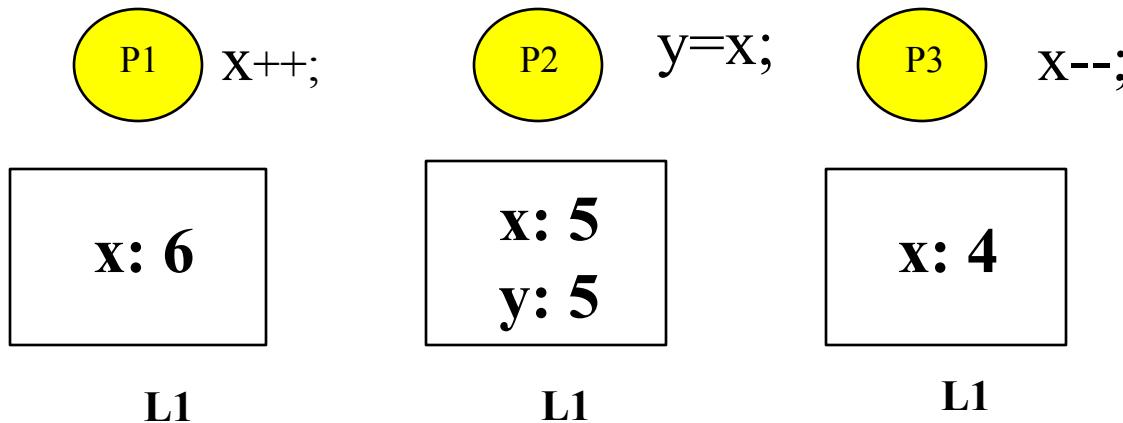


Featuring up to 72 powerful and efficient cores (72/4) with ultra-wide vector capabilities (Intel® Advanced Vector Extensions or AVX-512), the Intel® Xeon Phi™ processor raises the bar for highly parallel computing.

# High Performance Processors

- Xeon E7-48xx v3 and E7-88xx v3 series also contain functional bug-free support for Transactional Synchronization Extensions (TSX)
- TSX was disabled via a microcode update in August 2014 for Haswell-E, Haswell-WS (E3-12xx v3) and Haswell-EP (E5-16xx/26xx v3) models, due to a bug that was discovered in the TSX implementation.

# Memory Coherence and Consistency



- **What is the “true” value of x?**
- **What do we mean by the true value?**

# Storage Systems



ST4000DM005 4TB  
64MB Cache  
SATA 6.0Gb/s  
3.5"

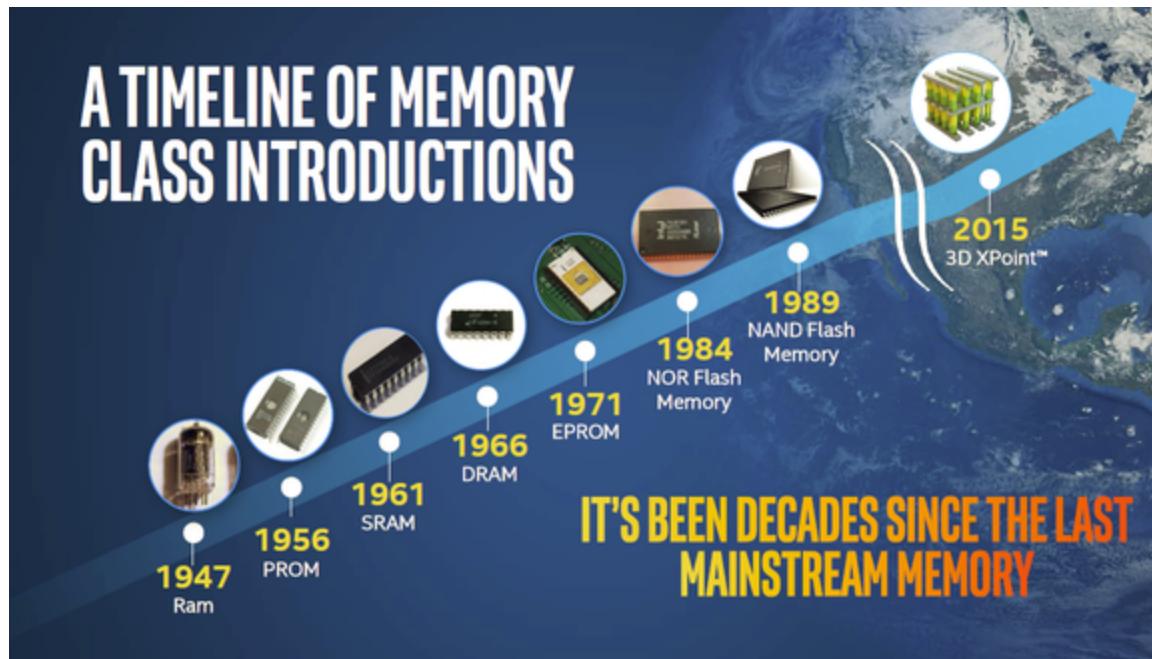
**Seagate Barracuda**

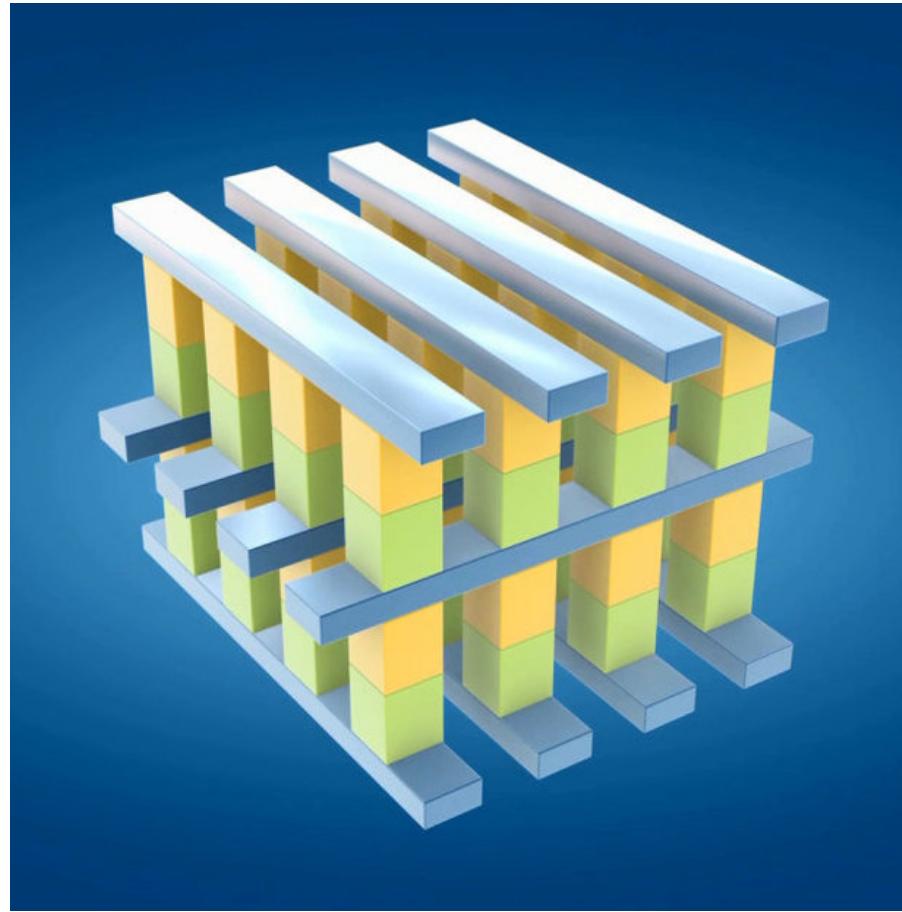


- 4 million IOPS
- 150 GB per second
- 4 PB flash capacity
- 99.9999% availability

**Dell EMC VMAX 850F  
All-Flash Storage**

# New Memory Technologies





**Intel Micron 3D XPoint**

**Byte Addressable Persistent Memory**

**Direct processor access to non-volatile storage at  
cache-line granularity**

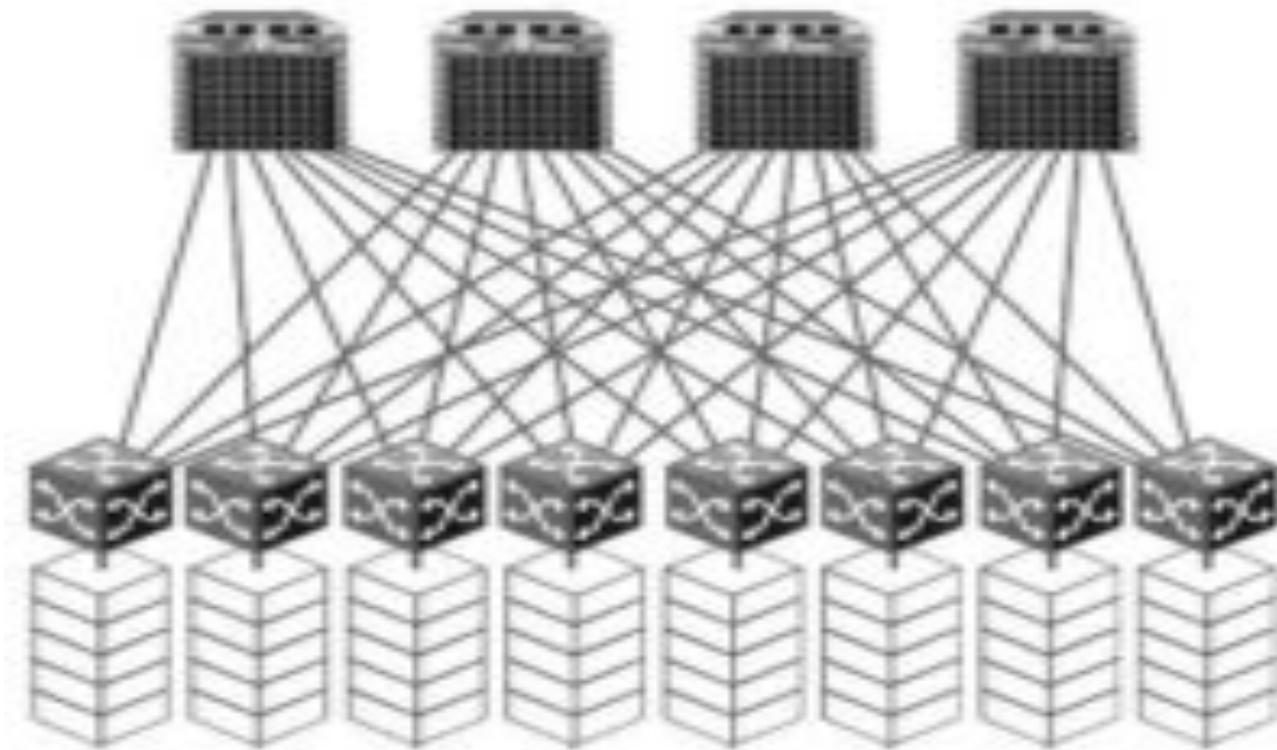
# Bye-Addressable NVM

- Today at Intel's Data Center Memory Summit, the new ‘Apache Pass’ Optane memory DIMMs were announced, with capacities from 128 GB to 512 GB.
- DIMM form factor
- Connected to the memory bus
- Directly Accessed like DRAM using LOADs and STOREs
- Non Volatile!!

# Interconnect Technologies

- RDMA – Remote Direct Memory Access
- Direct read and write of remote memory locations
- Protocol supported in NIC (Mellanox)
- Was restricted to proprietary interconnect like InfiniBand
- Software drivers (RoCE) : RDMA over Converged Ethernet
- Growing use in datacenters

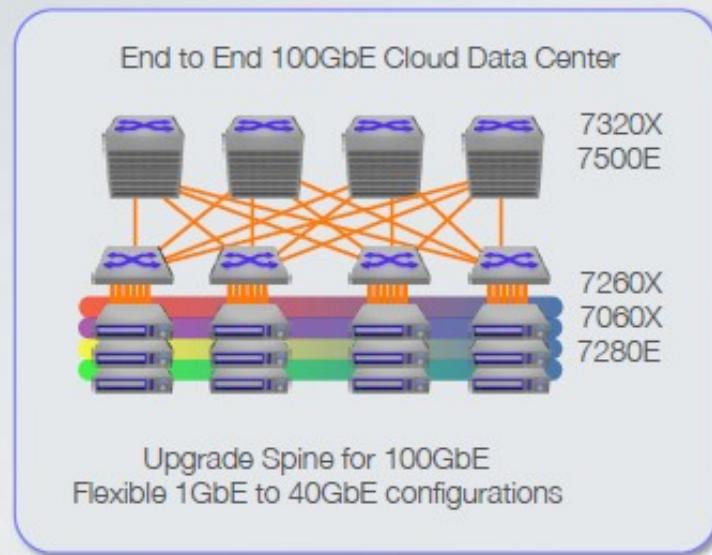
## Traditional Datacenter Networking



- Arista Gives Tomahawk 25G Ethernet Some Xpliant Competition

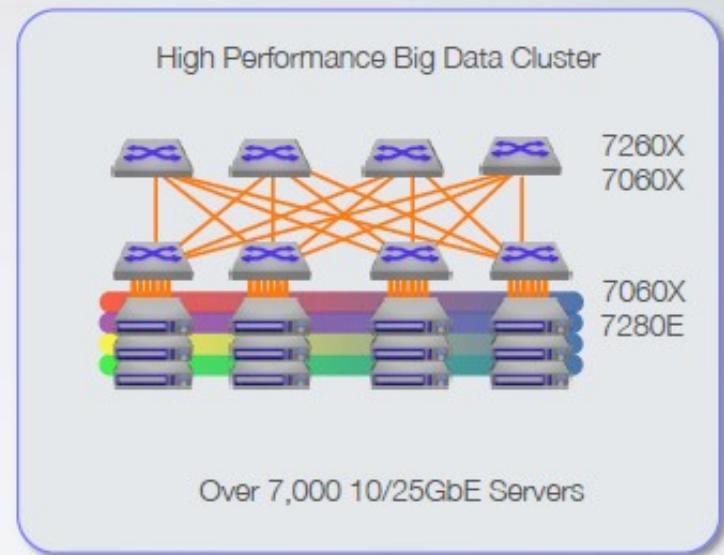
# Virtualization in the Datacenter

## Universal Cloud Scale Out with 100GbE Uplinks and 25G Servers



Scale the Leaf with 7060X and 7260X

- 100GbE Interconnect to Spine
- Choice of ToR for 40/100G uplinks



Scale the Spine with 7060X and 7260X

- Higher Speed Servers 25G to 100G
- Support for multi-rate uplinks

# Processor

- Instruction Level Parallelism
  - Pipelined, Superscalar, VLIW
- Data Parallelism
  - Identical operations on elements of large arrays (vectors)
    - Vector and Stream Processors, GGPUs
- Thread Level Parallelism
  - Application must be written as multiple cooperating threads
    - Multithreaded Processors
    - Multi Core, Multiprocessor
    - NUMA processors
    - Cluster

## ■ Memory

- Uniprocessor Caches
- Multiprocessor Caches
- Cache Coherence
- Memory Consistency models
- Virtual Memory

## ■ Synchronization

- Spin Locks
- Sleep Locks
- Lock-Free Synchronization
- Transactional Memory

- Storage Systems
  - DAS to Networked Storage: SAN, NAS
  - Hybrid Storage Arrays: Disks, SSDS, NVM
  - Distributed Storage: Reliability, RAID, Erasure Codes
  - Byte Addressable NVRAM
- Interconnect and Networking
  - Bandwidth, Latency, Topology
  - Routing
  - Scalable Networks : Close, Butterfly, Fat Trees
  - RDMA
- System Resource Virtualization (Software Defined Everything (SDX))
  - Sharing vs Isolation
  - Resource Virtualization Techniques
  - Resource Scheduling