COEN-4730/EECE-5730 Computer Architecture

Lecture 11
Servers, Reliability, and Power

Cris Ababei
Dept. of Electrical and Computer Engineering
MARQUETTE UNIVERSITY
BE THE DIFFERENCE.

Credits: Slides adapted from presentations of Sudeep Pasricha and others: Kubiatowicz, Patterson, Mutlu, Elsevier

1

1

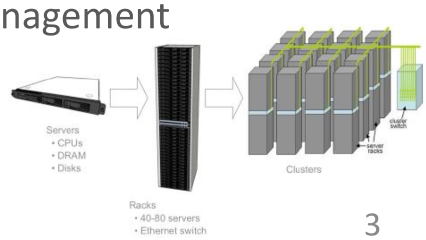
Outline

- Servers
- Availability, Reliability
- Power

2

What is a Server?

- A computer specialized for business users
  - File server, data server, application server
  - Database, file and printer sharing, email server
  - Web server, DNS server, firewall server, ftp server
  - Business applications: payroll, enterprise resource planning, customer relationship management
  - Small business
  - Big enterprise



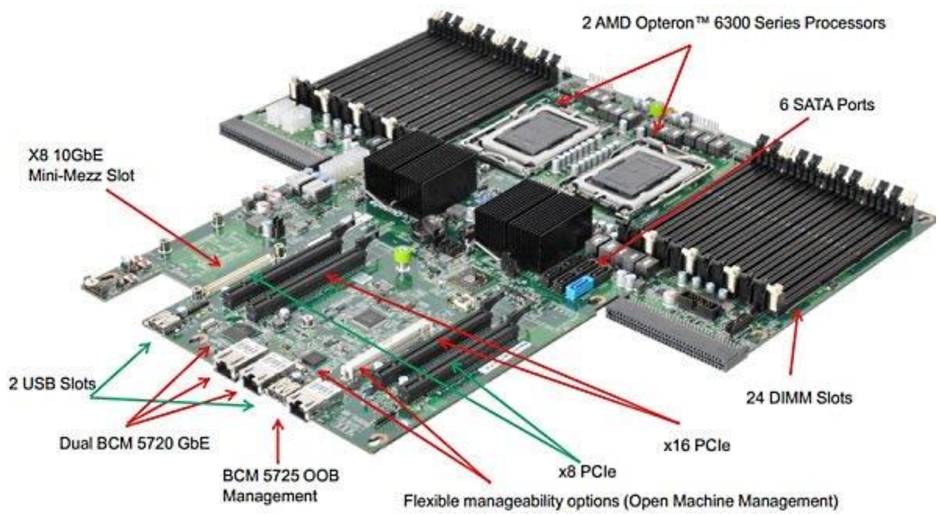
3

3

Example: FB Datacenter Racks

4

Example: AMD Open 3.0 Server Hardware



AMD Open 3.0 motherboard

5

Example: ASRock 1U12LW-C2750 Server

6

Desktop vs. Server

Desktop	Server
1-2 Desktop CPUs	Up to 64 server CPUs
192GB memory max	2 TB memory max
7 PCI/PCIe slots	Up to 192 PCIe slots
Fast high-res video	Basic video
Typically SATA disks	SAS, SATA, SSD, SCSI disks
Single user applications	Multi-user applications
Sound and multi-media	No sound systems
Monitor, keyboard, mouse	Shared/remote KVM
Designed for 9x5 operations	Designed for 24x7 operations
Little to no high-availability features	High availability and redundancy
Little to no manageability features	Support for manageability

7

Key Server Requirements

- Metrics tied to business value
  - Reliability – error-free operation as per-specifications
  - Availability – uptime of system including fault-tolerant operation
  - Serviceability – maintain server (install, upgrade, debug)
  - Scalability – handle increasing amounts of workload
  - Security – avoid vulnerabilities; protect data
  - Performance
  - Costs

8

8

Server Components

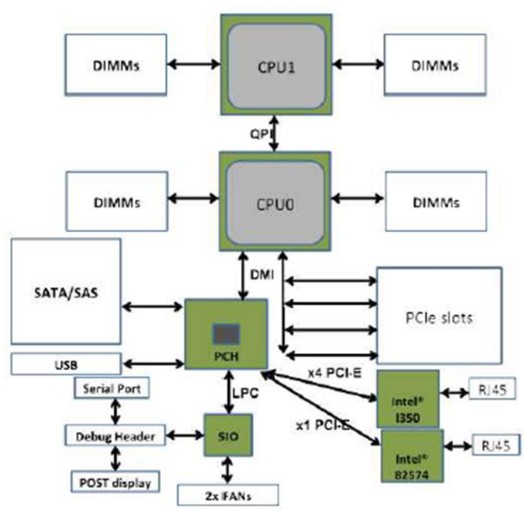
- CPU – for processing
  - Intel Xeon, Itanium, AMD Opteron, IBM POwer7
  - 1-64 CPUs in multiple sockets
- Memory and storage – of data/OS, etc.
  - DDR, DDR2, DDR3, …
  - Serial ATA (SATA), SCSI, Serial Attached SCSI (SAS), Fibre Channel
  - Direct Attached Storage (DAS), Network Attached Storage (NAS), Storage Area Network (SAN)
- I/O Bus and network interface – for communication
  - Ethernet, PCIExpress, …
- Operating Systems
  - Windows server, Unix, Linux, Solaris, …

9

9

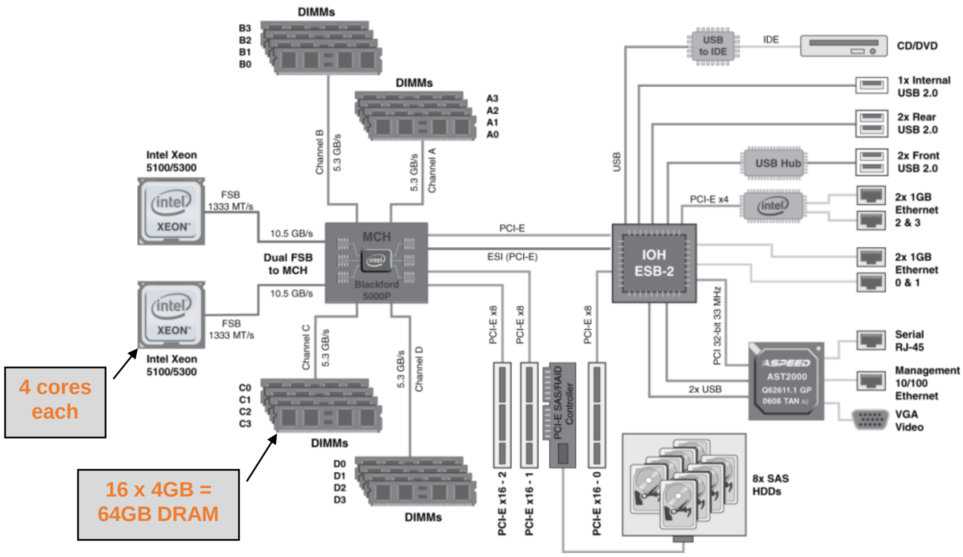
2-socket Server Basic Architecture

- 1-2 multicore chips
- 8-16 DIMMS
- 1-2 Ethernet ports
- 2-6 internal SATA/SAS disks
- External storage expansion
- Configuration/size vary
  - Depends on tier role
  - 1U-2U (1U = 1.75 inches)



10

Example: Sun Fire x4150 1U Server



11

Example Configurations

- Facebook server configurations for different services

Standard Systems	I Web	III Database	IV Hadoop	V Haystack	VI Feed
CPU	High 2 x E5-2670	Med 2 x E5-2660	Med 2 x X5650	Low 1 x L5630	High 2 x E5-2660
Memory	Low 16GB	High 144GB	Medium 48GB	Low 18GB	High 144GB
Disk	Low 250GB	High IOPS 3.2 TB Flash	High 12 x 3TB SATA	High 12 x 3TB SATA	Medium 2TB SATA
Services	Web, Chat	Database	Hadoop	Photos, Video	Multifeed, Search, Ads

12

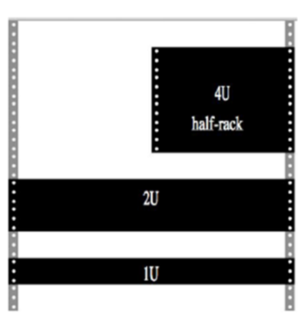
Server Form Factors

- Tower chassis servers
  - Upright free standing units + full systems
  - Affordable, entry-level server for small/remote offices
- Rackmount servers
  - Complete server optimized for ultra-compact vertical arrangement within a standard 19-inch mounting rack/cabinet
  - Flexible, located in computer rooms or datacenters
- Blade servers
  - Small form-factor servers housed in blade enclosures designed for modularity and high-density footprints
  - Very efficient use of space, amortized sharing of power supplies, fans, networking. Used in datacenters. Growing segment.
- Micro-slice servers
  - Multiple small server boards share an enclosure
  - Amortize cost of enclosure, disks, switch, power supply,...

13

Rack-mounted Servers

- Typically, 19 or 23 inches wide
- Typically, 42 U
  - U is a rack unit, 1.75 inches
- Slots:

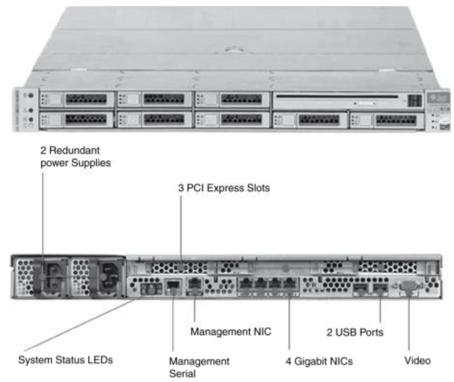


14

Rack-mounted Servers: Sun

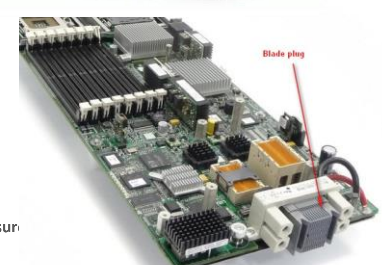
Standard 19" rack with 42 1U server

Sun Fire x4150 1U server



15

Blade Servers: HP C7000



- 10U enclosure for standard racks
- 16 half-height blades or 8 full-height blades
- Hot-pluggable; small form-factor SAS/SATA drives
- Power supplies
  - 6x 2250 power supplies, or 2400 W power supplies
  - 12V DC supply, no-redundancy, N+N redundancy, N+1 redundancy
  - AC power = 3-phase or single phase 48V DC
- 10 ActiveCool fans
  - Side ducts for interconnect modules
  - Separate fans for power supplies
- 8 Interconnect bays – single-wide or double-wide
  - VC Eth, VC FC, Eth, IB, storage switches
  - Gig Eth, 10Gig Eth, 4GB/8Gb FC, SAS, 4x DDR (20GB)
- Passive shared power backplane and active signal midplane
  - 5Tb/s aggregate BW
- Two bays for on-board administrator module
  - "Dynamic power saver", for subset of power supplies, dynamic power capping, fan management, enclosur troubleshooting, iLO access, DVD media sharing, …
  - Sensors, thermal conditions, power conditions, system configuration, management network
  - Systeme status display, HP insight manage

16

Enclosure-level Density Optimization

- Objective functions
  - Minimum costs – min blade costs (max blades per enclosure to amortize costs) and min switch costs (number of internal and external ports in switches)
  - Constrained by volume space within enclosure, minimum space required for server-class components, max power budget for server blade
  - Maximum flexibility – maximize switches for various network protocols, maximize performance of blades (highest power budget and volume) and switches (highest network speed protocols and highest external network connectors)
- Multi-objective optimization across power envelope, per server volume space, switch bandwidth oversubscription ratio, network protocols, …

17

Platform (HW) Management

- Management tasks
  - Turn on/off, recovery from failure (reboot after system crash), system events and alerts log, console (keyboard, video, and mouse (KVM)), monitoring (health), power management, installation (boot OS image)
- Platform management system
  - Automates all these operations
  - Out-of-Band (OOB), secure (privileged access point to the system), low-power (always on), flexible and low-cost

18

Management Processors

- An embedded computer on each server
  - Custom processors: e.g., HP iLO (Integrated Lights-Out)
  - Small processor core, memory controller, dedicated NIC, specialized devices (Digital Video Redirection, USB emulation)
  - E.g., IBM remote supervisor adapter (RSA), Dell remote assistant card (DRAC)
- Some iLO functions
  - Video redirection (textual console, graphic console)
  - Power management (monitoring, regulator, capping)
  - Security (authentication, authorization, directory services, data encryption, …)
- Standards: Intelligent Platform Management Interface (IPMI)
  - Baseboard management controller (simpler interfaces/functionality)

19

Outline

- Servers
- Availability, Reliability
- Power

20

Availability Classifications

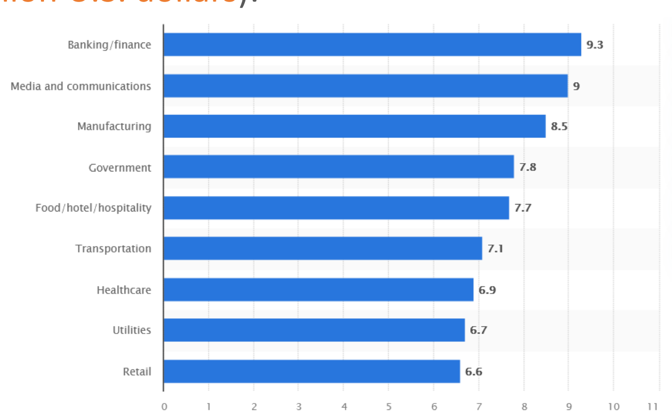
- Availability quoted in “9s”
  - E.g., Telephone system has five 9s availability
  - 99.999% availability of 5 minutes downtime per year

Uptime	Downtime in one year
99% (two 9’s)	87.6 hours
99.9% (three 9’s)	8.76 hours
99.99% (four 9’s)	53 min
99.999% (five 9’s)	5 min
99.9999% (six9’s)	32 sec
99.99999% (seven 9’s)	3 sec

21

Why is Availability Important?

- Mission-critical (100% uptime), business-critical (minimal interruptions)
- Average cost per hour of server downtime worldwide, by vertical industry (in million U.S. dollars):



Source: https://www.statista.com/statistics/780699/worldwide-server-hourly-downtime-cost-vertical-industry/

22

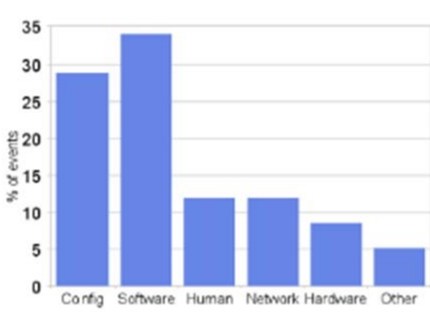
Types of Faults

- Permanent
  - Defects, bugs, out-of-range parameters, wearout, …
- Transient (temporary)
  - Radiation issues, power supply noise, EMI, …
- Intermittent (temporary)
  - Oscillate between faulty and non-faulty operations
  - Operation margin, weak ports, …

Real-world Service Disruptions

- Large number of techniques on hardware fault-tolerance
- Software, operator, maintenance-induced faults
  - Affect multiple systems at once

Source of "disruptions events" at Google



Source of enterprise "disruption events"

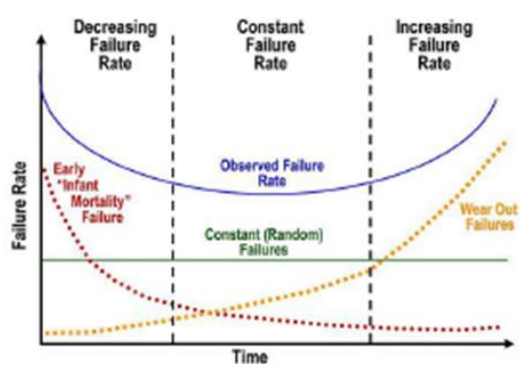
Disruption event = service degradation that triggered operations team scrutiny

Improving MTTF & MTTR

- Two issues
  - Error detection
  - Error correction
- Observations
  - Both are useful (e.g., fail-stop operation after detection)
  - Both add to cost; so, use carefully
  - Can be done at multiple levels (HW/SW)
    - General, chip, disks, memories, networks, system, DC
- Some terminology
  - Fail-fast – either function correctly or stop when error detected
  - Fail-silent – system crashes on failure
  - Fail-stop – system stops on failure
  - Fail-safe- automatically counteracting a failure

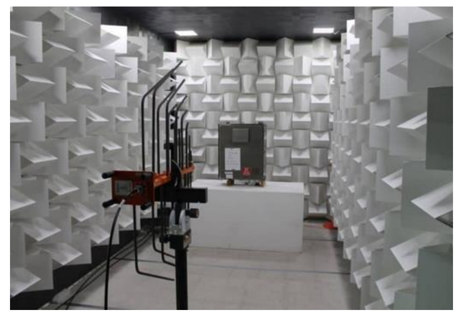
General: "Infant Mortality"

- Many failures happen in early stages of use
  - Marginal components, design/SW bugs, etc.
- Use "burn-in" testing to screen such issues
  - E.g., Stress test HW and SW before deployment



Extensive Validation

- High-level steps
  - Units built in a way that simulates factory methods
  - All components evaluated: electrical, mechanical, software bundles, firmware, system interoperability
  - Failure diagnostics and interaction with design team
  - Potential beta customer testing
- Extensive testing
  - Accelerated thermal lifetime testing (-60C to 90C)
  - Accelerated vibration testing
  - Manufacturing verification
  - Reliability of user interface and full rack configuration
  - Static discharge, repetitive mechanical joints, etc.
  - Dust chamber: simulate dust buildup
  - Environmental testing: model shipping stresses
  - Acoustic emissions and EMI standards
  - FCC approval (US), CE approval (EU)
  - Power fluctuations and noise: semi-anechoic chamber
  - On-site datacenter testing: TPC benchmarking



27

RAID: Dealing with Faults in Storage Systems

- Redundant arrays of inexpensive disks (RAID)
  - A collection of disks that behaves like a single disk with: High capacity, high bandwidth, high reliability
  - Key idea in RAID: error correcting information across disks
  - Many organizations; two distinguishing features:
    - Granularity of the interleaving (bit, byte, block)
    - Amount and distribution of redundant information
  - Patterson's classification – RAID levels 0 to 6:

Level	Description
RAID0	Block-level striping without parity mirroring
RAID 1	Mirroring without parity striping
RAID 2	Bit-level striping with dedicated parity
RAID 3	Byte-level striping with dedicated parity
RAID 4	Block-level striping with dedicated parity
RAID 5	Block-level striping with distributed parity
RAID 6	Block-level striping with double-distributed parity
RAID 1+0	Disk mirroring and data striping without parity

28

Dealing with Faults in Memories

- Permanent faults (Stuck at 0/1 bits)
  - Address with redundant rows/columns; i.e., spares
  - Built-in-self-testing (BIST) and fuses to program decoders
- Transient faults
  - Bits flip 0->1 or 1->0
  - Parity
    - Add a 9th bit
    - E.g., Even parity: make 9th bit 1 if number of ones in byte is odd

29

Dealing with Network Faults

- Use error detecting codes and retransmissions
  - CRC: cyclic redundancy code
  - Receiver detects error and requests retransmission
    - Requires buffering at the sender side
  - An Ack/Nack protocol is typically used
    - To indicate when receiver received correct data or not
  - Timeouts to deal with situations of lost messages
    - Error in control signals or with acknowledgements
- Permanent faults
  - Use network with path diversity

30

Dealing with Faults in Logic

- Triple modular redundancy (TMR)
  - Three copies of compute unit + majority voter
  - Issues: synchronization & common mode errors
- Dual modular redundancy (DMR)
  - Two copies of compute unit + comparator
  - Can use simpler 2^{nd} copy (e.g., parity detector)
- Checkpoint & restore
  - Periodic checkpoints of state
  - On error detection, rollback & re-execute from checkpoint
  - Issues: checkpoint interval, detection speed, number of checkpoints, recovery time, …

31

Datacenter Availability

- Mostly system-level, SW-based techniques
  - Using clusters for high availability
    - Active/standby; active/active
    - Shared-nothing/shared-disk/shared-everything
- Reasons
  - High cost of server-level techniques
    - Cost of failures vs. cost of more reliable servers
  - Cannot rely on all servers working reliably anyway
    - Example: with 10K servers rated at 30 years of MTBF, you should expect to have 1 failure/day
- But, components must be reliable enough…
  - ECC based memory used – detection is important!

32

Outline

- Servers
- Availability, Reliability
- Power

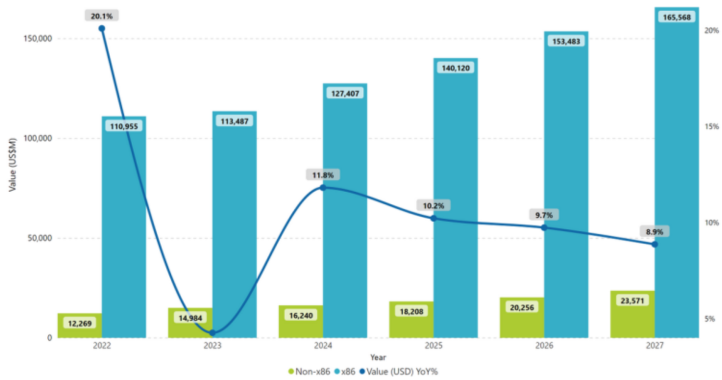
33

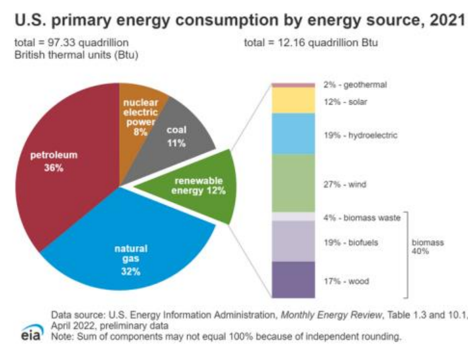
Why is Power Important?

- Desire to reduce electricity use
  - For mobile devices, impacts battery life
  - For tethered devices, impacts electricity costs
    - Delivery of power to buildings
    - Gets worse with large datacenters ($7M for 100 racks)
- Environmental friendliness
  - Compute equipment energy use has been increasing (e.g., training LLM such as ChatGPT and others)
  - Need to reduce amount of emitted CO_2
- Power delivery, packaging, cooling costs
  - At high-end 1W cooling for 1W of power!
- Compaction, density, reliability
  - Thermal failures
    - 50% server reliability degradation for +10C
    - 50% decrease in hard disk lifetime for +15C

34

Why is Power Important?



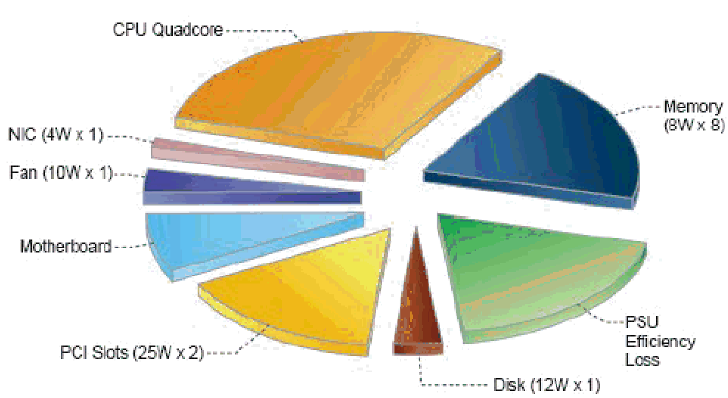


35

35

Total Power Consumption

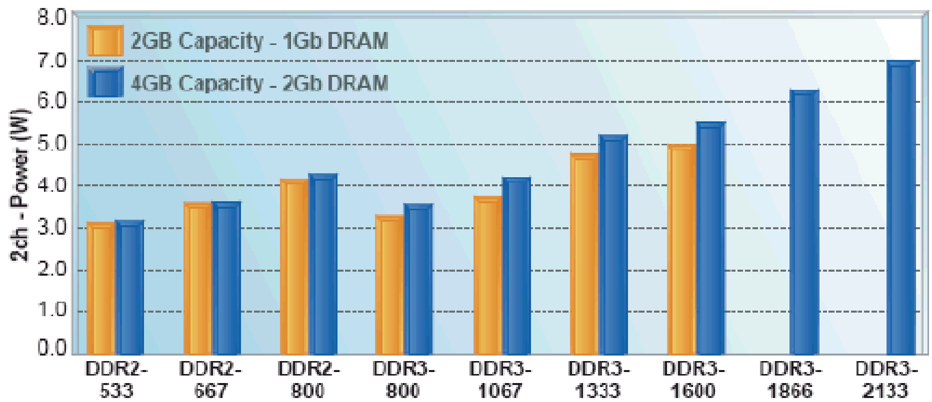
- Power consumption of a Quadcore Intel Xeon server



Server Power Consumption (Source: Intel Labs)

36

Memory Power Consumption



RDIMM Memory Power Comparison (Source: Intel Platform Memory Operation)

37

Power Consumption in ICs

$$P = C{*}Vdd^2{*}F_{0\rightarrow 1} + Tsc{*}Vdd{*}Ipeak{*}\,F_{0\rightarrow 1} + Vdd{*}I_{leakage}$$

- Dynamic (active) power consumption
  - Charging/discharging capacitors
  - Depends on switching activity
- Short circuit currents
  - Short circuit path between power rails during switching
  - Depends on size of transistors
- Leakage current or static power consumption
  - Leaking transistors, diodes
  - Gets worse with technology downscaling and lower Vdd
  - Gets worse with higher temperatures

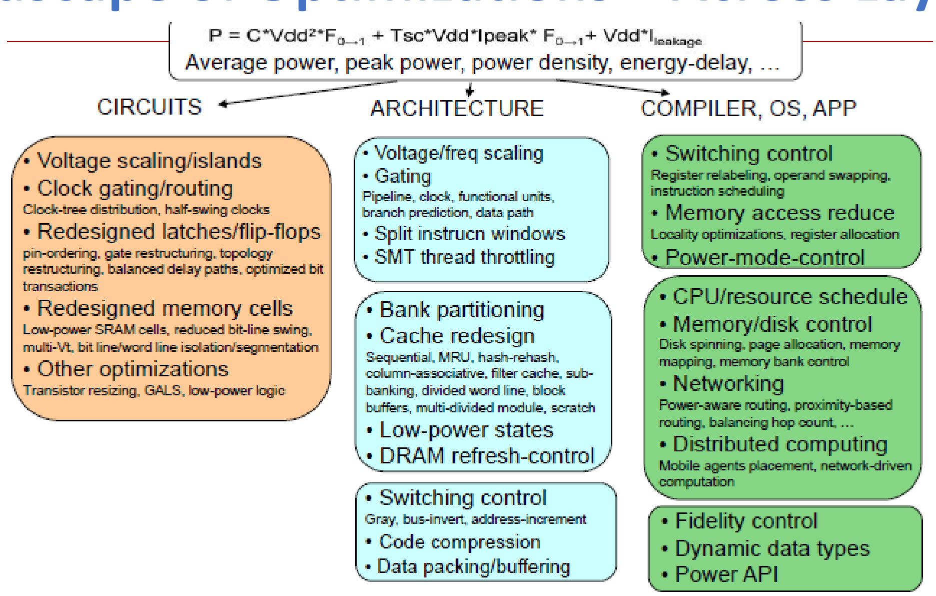
38

Metrics

- Energy (Joules) = Power (Watts) * Time (sec)
  - Power limited by infrastructure (power supply)
- Power density = power/area
  - The major metric for system cooling
- Combined metrics
  - How to trade off performance for power savings
  - Energy-Delay-Product (EDP), …

39

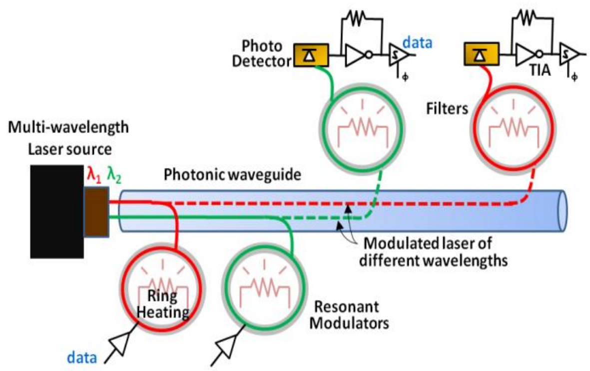
Landscape of Optimizations – Across Layers

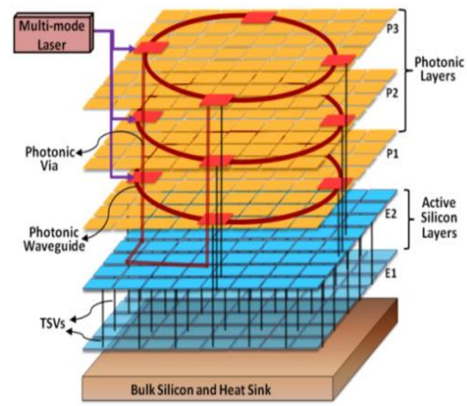


40

Replace Copper Wires with Optics

- Networks-on-chip (NoCs) have high latency and power dissipation
- What if we used photonic interconnects on chip?

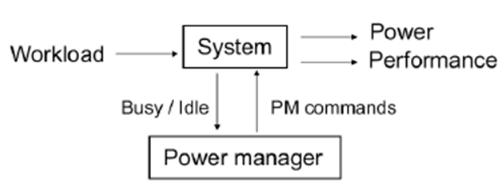




41

Power Management

- Components with multiple power modes/states
  - Active: different levels of performance/power consumption
  - Idle: different power consumption/wake-up time
- Select power states to match constraints
  - Exploit fluctuations in use
  - Done in HW/SW and/or by user
  - Tradeoffs: power saving Vs. QoS Vs. speed of resuming



42

Advanced Configuration and Power Interface (ACPI)

- Standard for power management of systems
  - Describes power stages for system, cores, devices,…
  - Interface for SW to query and manage power states
- Global system states
  - G0: working – system in responsive, user application run
  - G1: sleeping – appears to be off. Within G1:
    - S1 (caches flushed, CPU halted)
    - S2 (CPU power off)
    - S3 (suspend to RAM)
    - S4 (hibernate to storage)
  - G2: soft off (wakeup on LAN)
  - G3: hard off (mechanical)

43

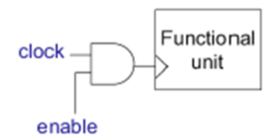
Advanced Configuration and Power Interface (ACPI)

- Device states
  - D0 fully on operating state
  - D1 and D2 are intermediate states (vary by design)
  - D3 is powered off state (device unresponsive)
- Processor states
  - C0 is fully on
  - With P states related to DVFS stages
  - C1 to C3 are idle modes
  - Clock may be stopped, but, state is maintained
  - C4 and beyond are various power off state
  - First the cache, then cores, and finally the whole chip

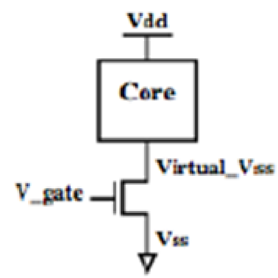
44

Power Management in Processors

- **Clock gating of idle units**
  - Clock is major power contributor
  - Done automatically in most designs
  - Near instantaneous on/off behavior



- **Power gating (C4 and beyond)**
  - Turn off power to unused cores/caches
  - Large delay for on/off
    - Saving SW state, flushing dirty cache lines, turn off clock tree
    - Carefully done to avoid voltage spikes or memory bottlenecks
  - Area & power consumption of gate
  - Opportunity: use thermal headroom for other cores



45

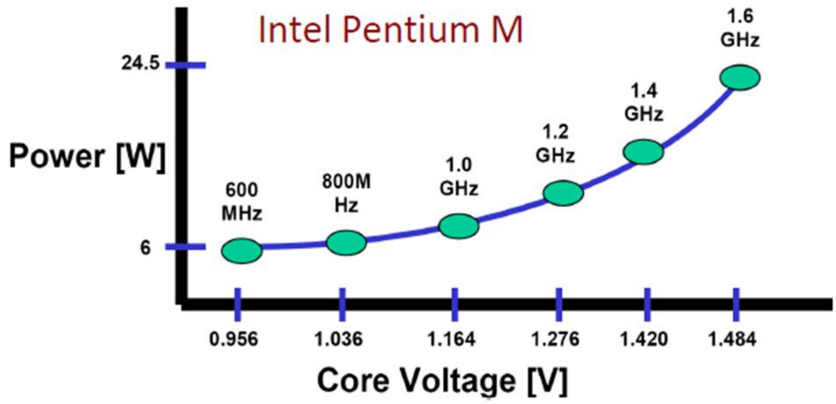
DVFS

- **Set frequency to lowest needed**
- **Scale back Vdd to lowest required by that frequency**
  - Lower voltage -> slower transistors
  - Power = C * Vdd^2 * f
- **Provides P states for power management**
  - Heavy load: frequency, voltage, power high
  - Light load: frequency, voltage, power low
  - Tradeoff: power savings Vs. overhead of scaling
  - Effectiveness limited by voltage range

46

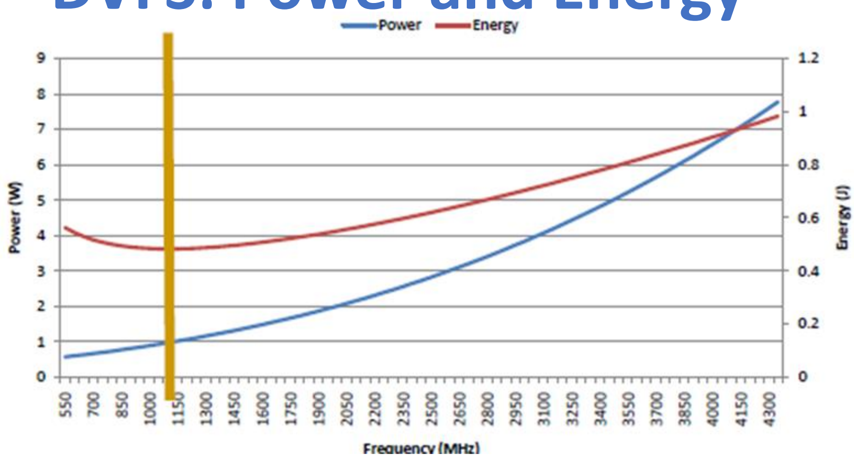
Example DVFS Implementation

- Transitions typically take a few microseconds



47

DVFS: Power and Energy



- Assumptions
  - Running time is linear w/ frequency
  - Vdd = 0.78V to 1.62V, f = 550 MHz to 4350 MHz

48

DRAM Power States

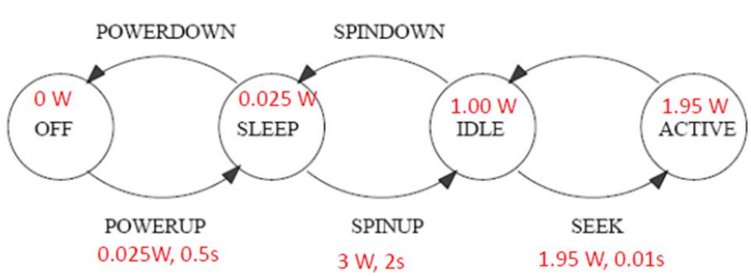
Power State	Operating Mode	Resync -time	% Active power
Active	All modules ready	0 cycles	100%
Standby	Column multiplexers disabled	2 cycles	60%
Napping	Row decoders turned off	30 cycles	10%
Power Down	Clock sync to Controller interface turned off	9000 cycles	1%
Disabled	No refresh; data lost	Reboot	0%

- Example: 5 states in DR-DRAM
- Tradeoff: power savings Vs. resync penalty

49

Disk Drive Power Modes

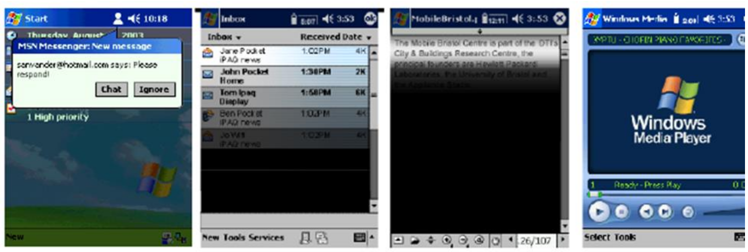
- Common optimization
  - Stop spinning disk when it is unused for a certain period of time
  - Example: Toshiba notebook drive



50

Display Power Management

- Turn-off displays, use smaller displays
- Energy-aware user-interface
  - Spatial – focus on informational content
  - Temporal – focus on content of interest at given time
  - Reduced energy (2-10X) and better ease-of-use
- Leverage usability-friendly energy-reducers
  - E.g., Contrast, personalization, visibility of surrounding text



Global savings of 8.3 Megawatt-hours per day if Google switched to black background!

51

Per-server Power Management: e.g., HP Power Regulator

- Monitor & manage individual and groups of servers by physical or logical location (power domain)
- Monitor vital power information
  - Power consumption in Watts
  - BTU/hr output
    - British Thermal Unit (BTU) per Hour: is a measurement of heat energy.
    - One BTU is amount of heat required to raise one pound of water by one degree Fahrenheit.
  - Ambient air temperature
- Policy based power management
  - Power cap policy: Set maximum BTUs/hr or Wattage threshold (capped on a server by server basis)
  - Temporary conservation policy: Set time of day to drop to lower selected priority servers into lower power state
  - Severe facility issue: Drop lower priority servers into lower power state when sever facility issues occur
  - Energy efficiency policy: Set all servers in power domain to dynamic power regulating

52

Cluster-level Power Management

- Power-aware load distribution to a server cluster
  - Try to create idle resources to send to low-power/off states
  - Sophisticated policies (predictions, economy-based, batching)
  - Interactions between intra-server DVS and inter-server load balancing
  - Impact of heterogeneity
  - Interactions with performance and more broadly service-level agreements (SLAs)

53

Readings

- Luiz André Barroso, Jimmy Clidaras, and Urs Hölzle, The Datacenter as a Computer, An Introduction to the Design of Warehouse-Scale Machines, Second Edition, 2013 (Ch.3-6):
  - https://link.springer.com/book/10.1007/978-3-031-01761-2
- Hot Chips: A Symposium on High Performance Chips
  - https://www.hotchips.org/archives/
- Open Compute: www.opencompute.org
- Google: https://www.google.com/about/datacenters/
- Top 500: https://www.top500.org/lists/top500/

54

Assignment

- Search online about how AI is used & impacting design and management of servers and datacenters/WSCs
- Write report to summarize your findings
- Upload report to D2L

55

55