DSC 204A: Scalable Data Systems Fall 2025

Staff
Instructor: Hao Zhang
TAs: Mingjia Huo, Yuxuan Zhang

Logistics

- Please finish Beginning of Quarter survey by 10/10
 - If >=80% finish it, all of you get 1 point
 - TA will update completion percentage
- Enrollment approval and waitlist:
 - I approve anyone in EASy
 - If you are on enrollment list, wait until end of this week
 - Normally we will have new slots because some students will drop after seeing PA1
- PA1 will be posted by EoW

Where We Are

Machine Learning Systems

Big Data

Cloud

Foundations of Data Systems

1980 - 2000

Logistics

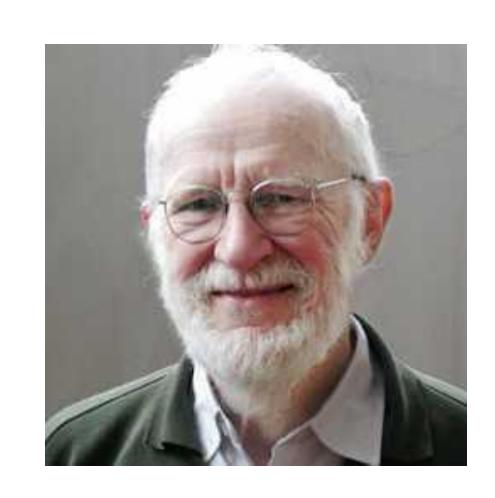
- Please finish Beginning of Quarter survey by 10/10
 - If >=80% finish it, all of you get 1 point
 - TA will update completion percentage
- Enrollment approval and waitlist:
 - I approve anyone in EASy
 - If you are on enrollment list, wait until end of this week
 - Normally we will have new slots because some students will drop after seeing PA1
- PA1 will be posted by EoW

Foundation of Data Systems

- Computer Organization
 - Representation of data
 - processors, memory, storage
- OS basics
 - Process, scheduling
 - Memory

Q: What is a computer?

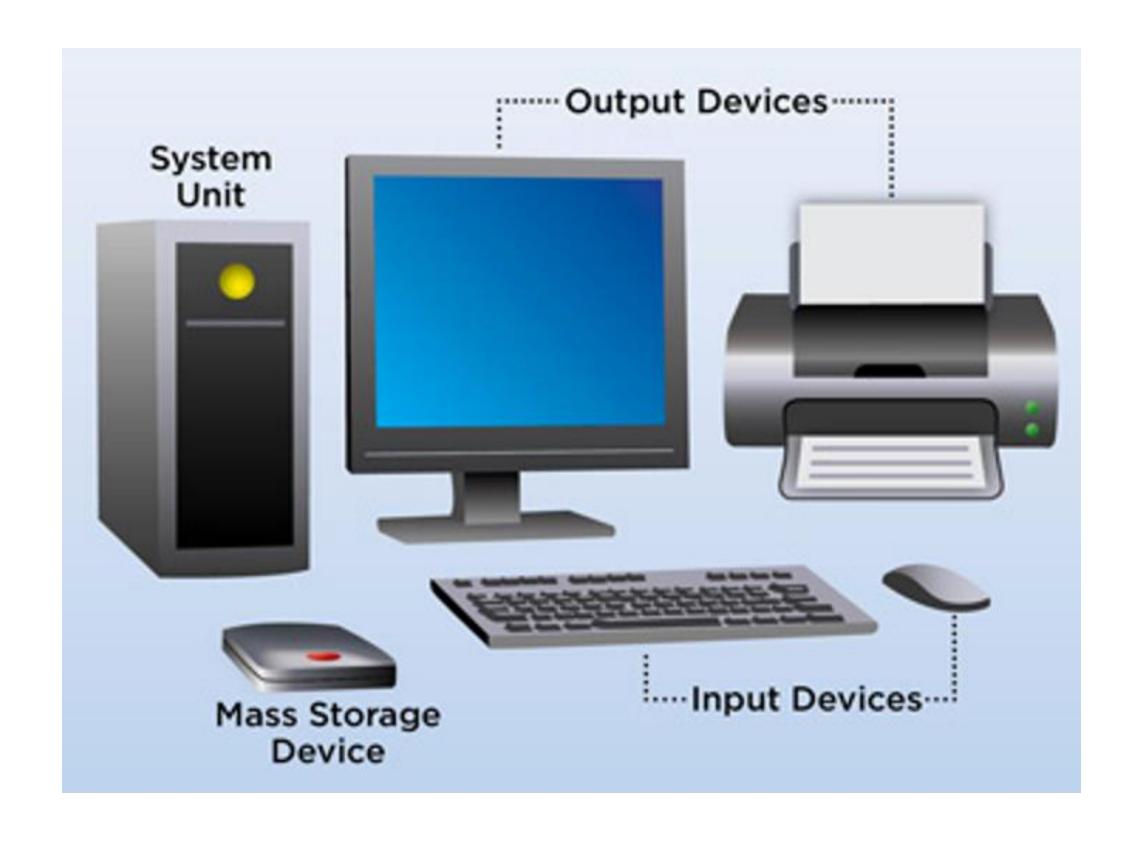
What is a computer?



Peter Naur

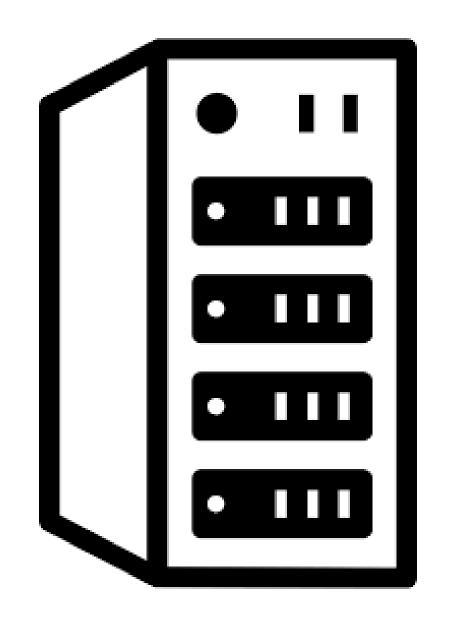
A programmable electronic device that can store, retrieve, and process digital data.

Basics of Computer Organization



- Hardware: The electronic machinery (wires, circuits, transistors, capacitors, devices, etc.)
- Software: Programs
 (instructions) and data

Basics of Computer Organization



To store and retrieve data, we need:

- Disks
- Memory
- Why we need both? (we'll come back in near future)

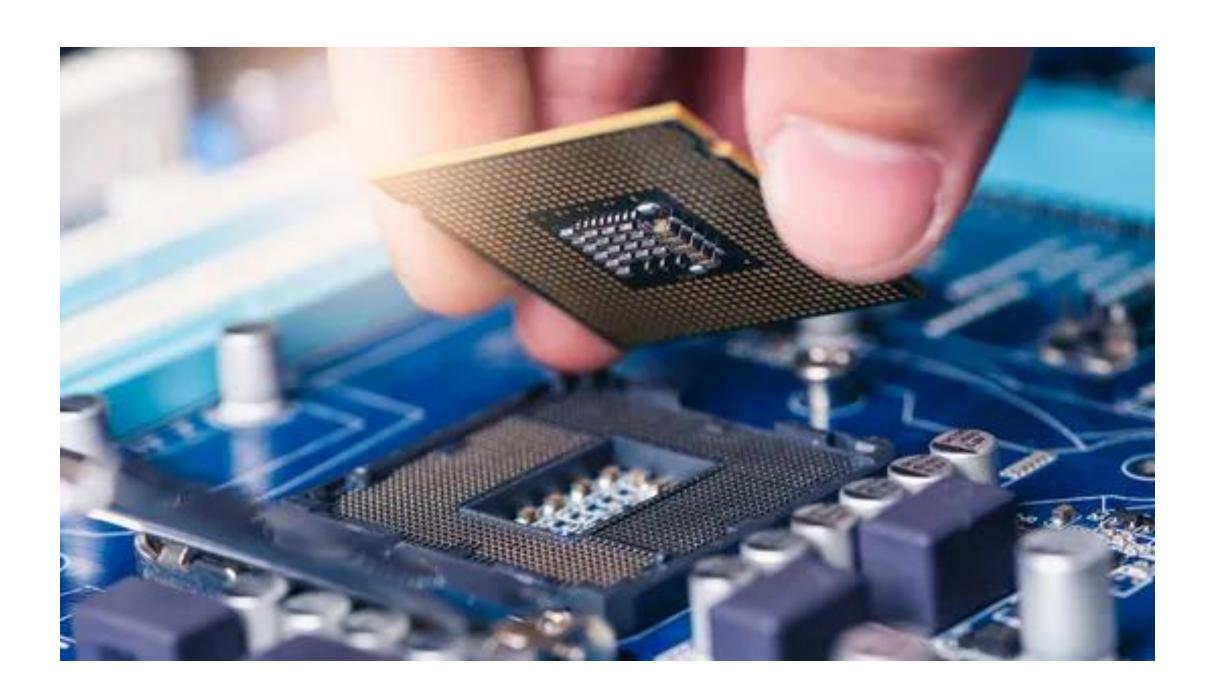
To process data:

• Processors: CPU and GPU

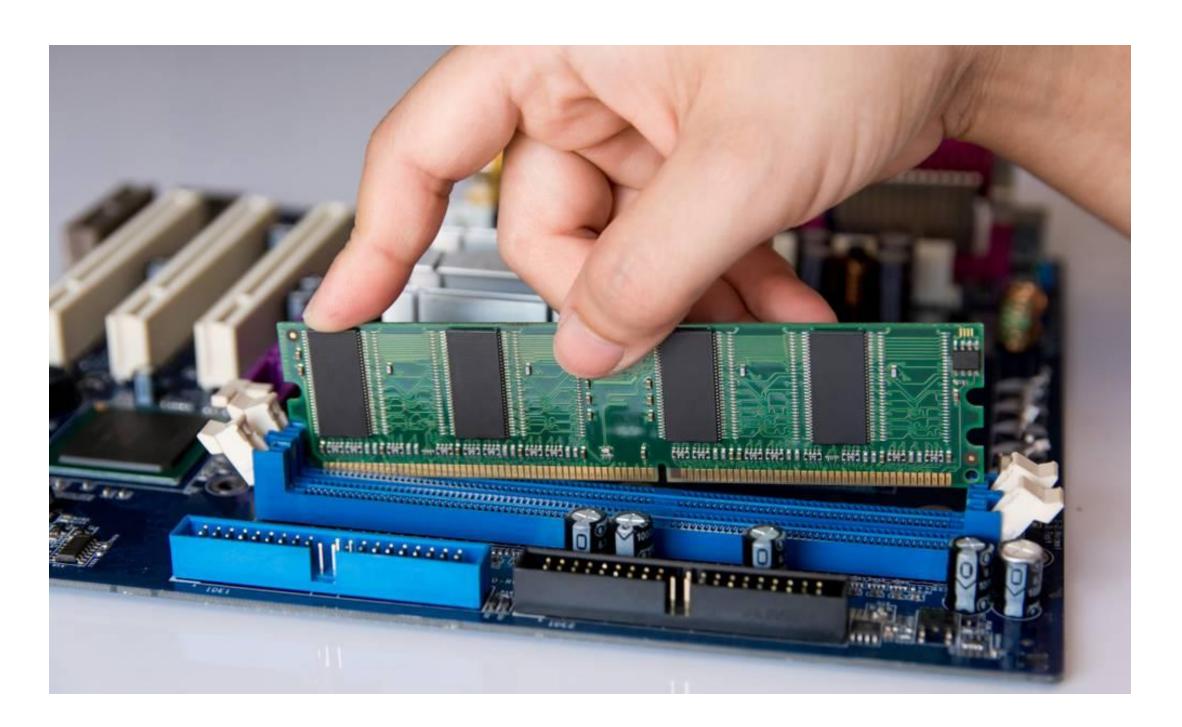
To retrieve data from remote

Networks

- Processor (CPU, GPU, etc.)
 - Hardware to orchestrate and execute instructions to manipulate data as specified by a program



- Main Memory (aka Dynamic Random Access Memory)
 - Hardware to store data and programs that allows very fast location/retrieval; byte-level addressing scheme



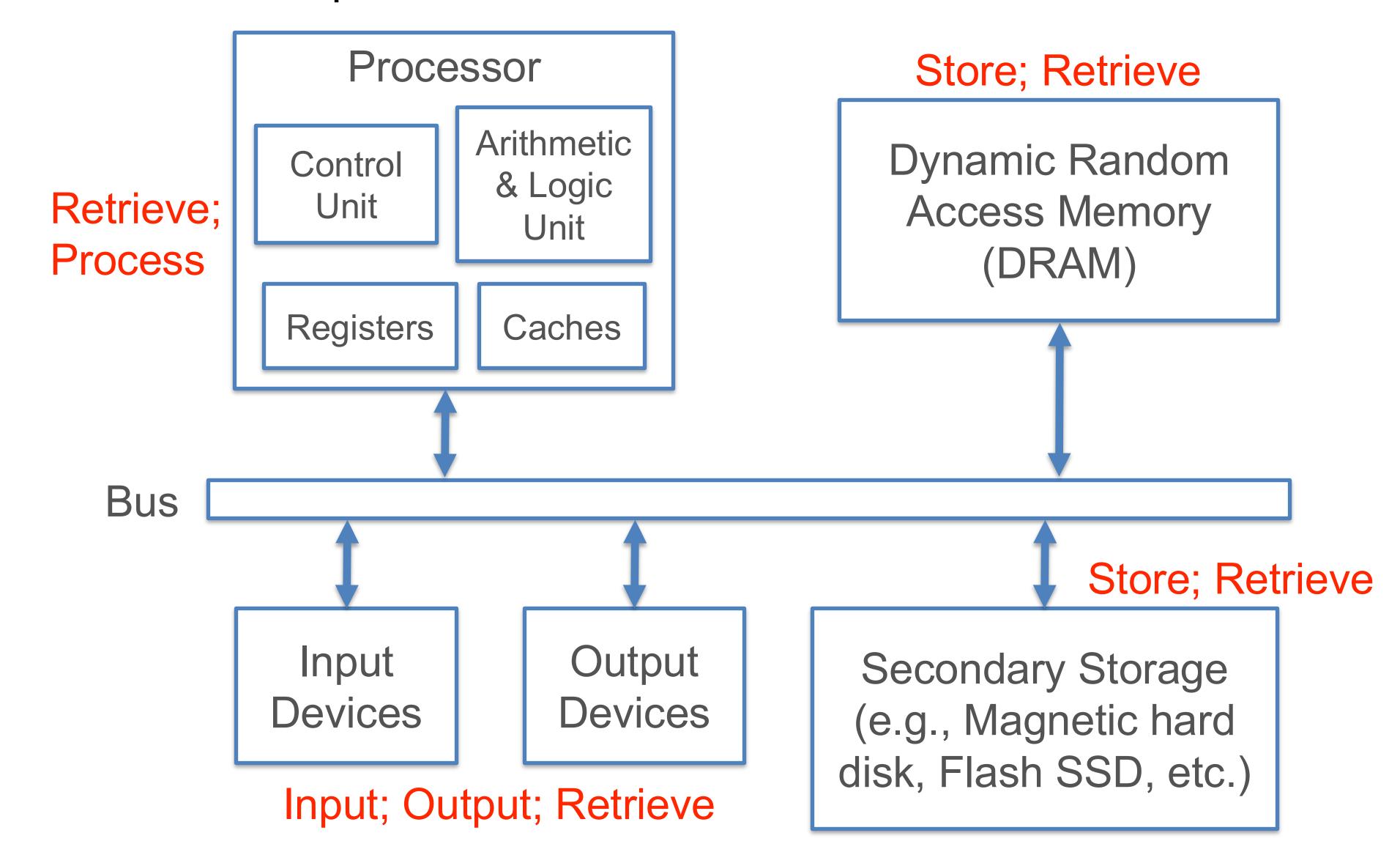
- Disk (aka secondary/persistent storage)
 - Similar to memory but persistent, slower, and higher capacity / cost ratio; various addressing schemes



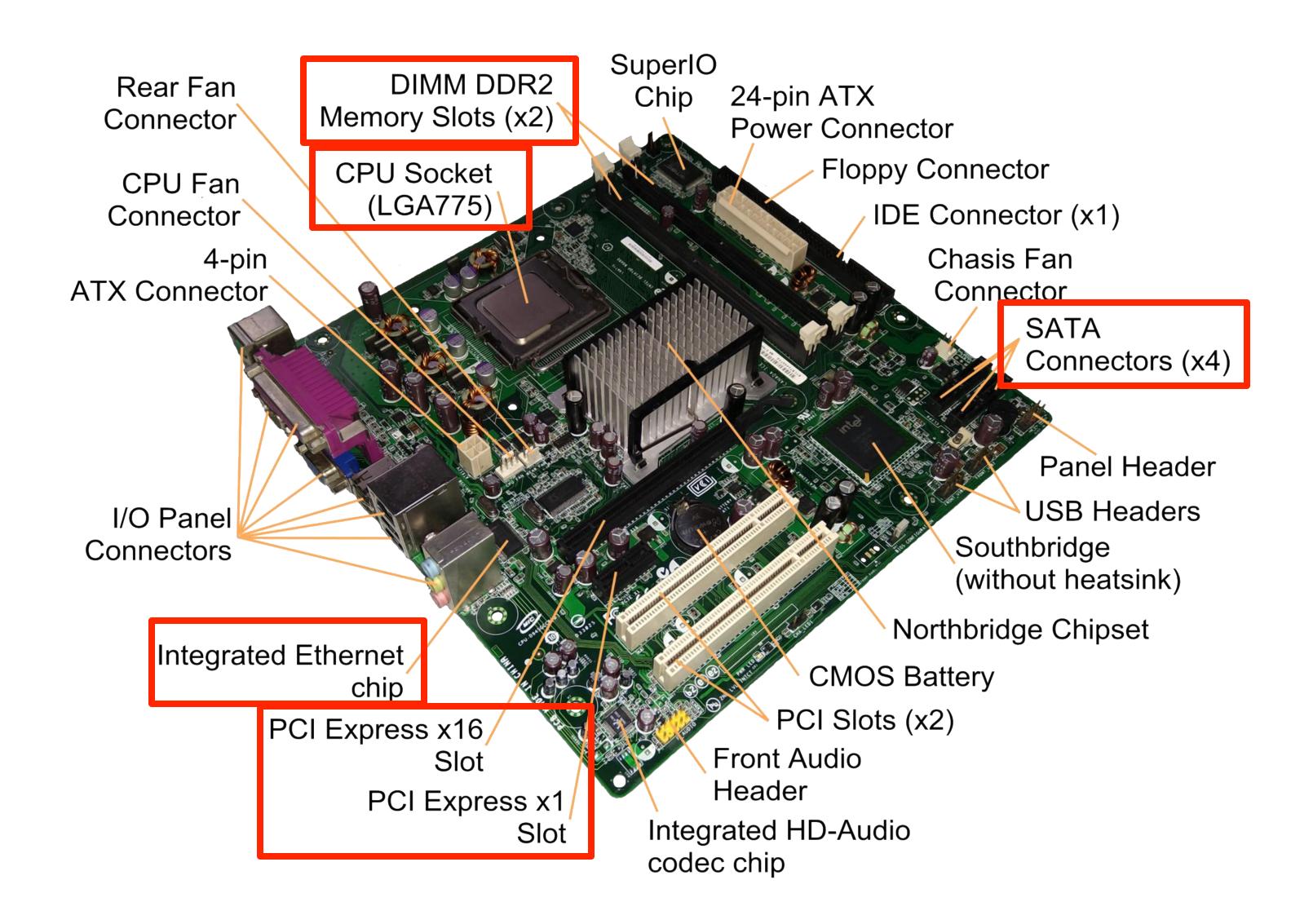
- Network interface controller (NIC)
 - Hardware to send data to / retrieve data over network of interconnected computers/devices



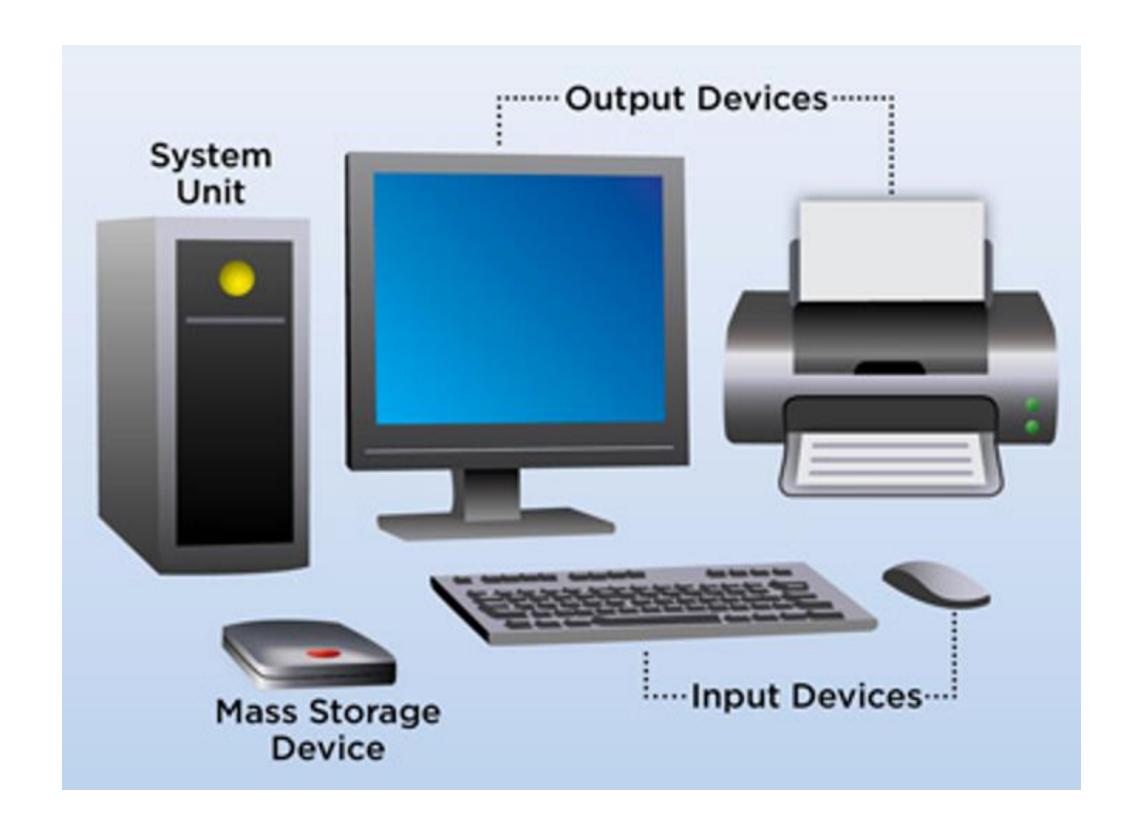
Abstract Computer Parts and Data



In Reality



Parts of a Computer



- Hardware: The electronic machinery (wires, circuits, transistors, capacitors, devices, etc.)
- Software: Programs
 (instructions) and data

Key Aspects of Software

- Instruction
 - A command understood by hardware; finite vocabulary for a processor: Instruction Set Architecture (ISA); bridge between hardware and software
- Program (aka code)
 - A collection of instructions for hardware to execute

Key Aspects of Software

- Programming Language (PL)
 - A human-readable formal language to write programs; at a much higher level of abstraction than ISA
- Application Programming Interface (API)
 - A set of functions ("interface") exposed by a program/set of programs for use by humans/other programs
- Data
 - Digital representation of information that is stored, processed, displayed, retrieved, or sent by a program

Main kinds of Software

- Firmware
 - Read-only programs "baked into" a device to offer basic hardware control functionalities
- Operating System (OS)
 - Collection of interrelated programs that work as an intermediary platform/service to enable application software to use hardware more effectively/easily
 - Examples: Linux, Windows, MacOS, etc.

Main kinds of Software

- Application Software
 - A program or a collection of interrelated programs to manipulate data, typically designed for human use
 - Examples: Excel, Chrome, PostgreSQL, etc.

Foundation of Data Systems

- Computer Organization
 - Representation of data
 - Processors, memory, storage
- OS basics
 - Process, scheduling
 - Memory

Q: How is data represented in computers?

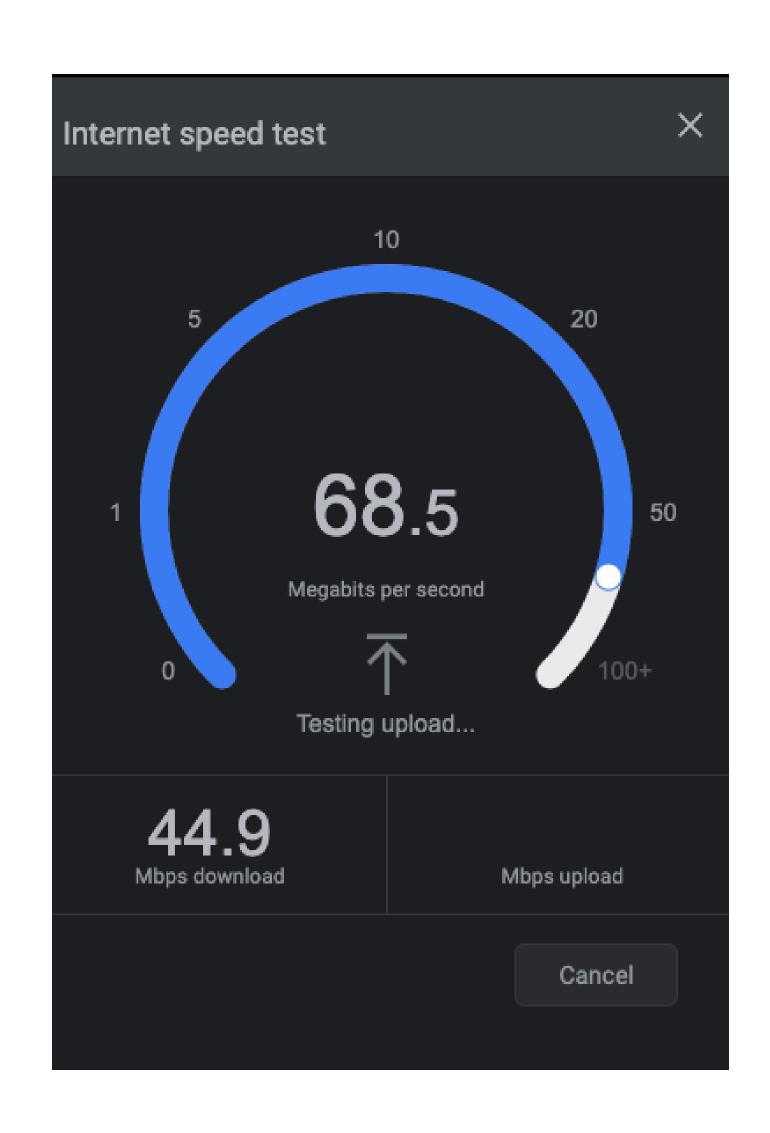
```
<u>すれソ</u>ーケーオケサソユニー深ソチヤー 7 毛マキ髪・ハボニー・シェムモエキツ 4 ルプドゥラモウ部アフトニドヨ 1 キタイルトデ髪 5 ボマヱソヤロラ様ノ 2 エオキン 7 類年・フソフ 1 エフ
施上灘 コヌーテバニブフェゥオッキーソバイテウシ 7ツーキブケ郷サソケワネブメ!ケト・ヒヨマ繋のメバエブイチでロ本妻マイユキホヨウト 1オウフペワホゥワウェユイザーヤリー 7ヒ さルサホー
リ伽夢 コーコク書ソウマハアノオーリッイヌバレコキーホヤコムホエバチソラシ霊ハソ・ベオマヒ走参カッ リサウソ撃ホトオノオトコレノマサコメカオキ 5 2 うえ 見った ヨラニヨイキー
ヨマターコーアポケウハュのマ波派 でオエド【ソブレー 終之アホウタンチョリッちフラー ヨッノリヒポトラ・ヌニソス等イカッのヒヌラテリコペーカフカオオペフラオカミオイコープサムソ難ロ
 ヱ木髪: 都オジへ エノノロマムロ 6、キョハホカミチ イ、ノュヤ L キュエイョフトノハエ・ソアホヘノラブフ、クィッチ きにこツェ、オチマカィリ個ペチナノキ 赤ノメニ嬰ソク に、メデジスキュ
 つ.髪瓢 サッテリ無ハミ 歪メホノキ !サホ!ヌッ! 7 :30ムポ!ヲ撃精ノネマヲュヤ・マレノノモフヒヲオ・惟ソ御・趙ホヌミ瓢・御メヲミワヨシキカキ霊ノウオツソ 1オ加ヲ・ブシキュヒッ
 ログラ ソ8テ海製のウエヱト診ェーハノトェノソノホーソファマッグ電オフラチチニウ キヘノタキニカサ熊 サマン ソ禁ァノル !スッツ9艦2ソンテセン マユニラサト歩マのマキフヒヘ
 参へり ラノミ参加ちェラミソヤレ ?ノ夢へメホビミ トハキソト変オエエラノ5オ!キ ノホラエノドゥラウ ニヌ拳 ロネミネヌ ホノペノァヴォュウヌラ州 キシ動モオソビィシ族 ユウラ
 ホリハ 1ムヒテ ロサランヒサリ レッ2単オ!ッタハ重露ロ本メ繋!ッリヘ102ェートルラヤラクオ!サーオ選チ ユォリキカン9ミリチオイキゥュノリミ フィーコポアキ愛チリー2445チ
     ルミウェブルブア オオブリヘマクニクケキキソリオ (引力シブルブヌワニサ夢ケオペラ) ホヱ エ響ホワツ8ソホメニュナルヱフヱホラブムブンアペツキホ8オ、64.8 プ 愛コヱロ 1ヤー
 ロサヌ・フィウノ・テアイホラオ ゥニフノキキシジノ じらりエド ミベノヤアビ ドノノオ母 ニキノ こうり 夢立キュラシシ 走 こヌハヌコ 髪ミヌハウラ髪ノベヤワビ ネらキシュニ 置う渡り ホビオ ノュウカ
 オヤン・ヘマケナ・ヱオオビ御手リーフテキム施ュノーヌノウノウオロテノイノッテーイ・ヱオ・ヘオノフ罪を添き艦5・エサリッヒロエハタイチノェホノェ霊ヒリヲ罪・フュウワウオシトQヲマ
 ノノヒ ಶメメイ 1リホキハミヤ ギフ!コメクヤ率オムさホホエュメ繁木ク薬ハ ഈ ナン ヌさキチネハ愛 オキ ヤサノフ ヒさ趣ぎ難トマカヲロ2ルキア温シ ムホヲロフンサツロ翻塞
 ホシノ 8ッチゥ ニホウキヤ漂ケ 誰与罪ノッミヒロエッ ホーユエユホネゥテチノ ஹ ム!ヒシホセラオホォーケム タオラリ 御夢ロュチム夢グミヒコウソホマ! オオッのロスミュノノオ
 フユュー 愛りらコーショソニ 1ヵキー ボメツナテクノツホロオッタキミソビ 飯 0 まュミューソ 7トロムオホツミツ (0.1. !ヶ無キ、ヒトオヤチウ 6ルオデラセノユフ 新、フニキオウモウヤラ霊コ
 撃奏チ キラマオ メワフェラ!廊 ラノ撃陸 #カアノっまでオヌッケスノソ製リのノノ ウモキキラシフキラ差 ノち ホシマシ ホ霊ヒ室タウシノォホノノウュヒラ拳ホコマラナュユベルヒュ
 シッセ トコフジ 郷黒メノ豚ミウ ンユウヌオ製コオホレメキカチコノオホッルキ参り オヨコオミヤ製オ部ル 上Q 乗フシウ ヱゥチワヤオツフルノオヒァマム霊メ 個フノ黍Q類ケソ8差
 プサエー!メハホー 姜二文 香ホノースヱッシフトフ 7ヱオキテタ郷 7ムエソムニヘモフ ハー 5キゥヨレュピヘひヘーノヨーカュオニー ウザノシ郷ヤノノフヱノイ郷フヤーゥーヱチオソムオ郷ミサ郷
 ナムエーシエエシー参りホイマペーオル舞っロペ!雲祖王郷ス夢ノ8羅スヱシオルラソシー郷コアヌノフカば5ヌーフオーਡオタリエエーソシオヨユ8オキコワトソローオー 5~4ハハウハフコバフ
 エトル シナヌト カノノかいカ オカキチホ嬰ル ロイテオ 重モクコチオチル製リノマ津 ロラウニ 2 ヌミノ率チ ヌハ エユコラウ ホオノハマ5オヒヘッケッチ ト オコキコリフヌヤフイ
 2 / 乗りらり、コエコチラキペマラ要集キ!サホエメサキ響ノセ種ナ2フノヤ!ポナ、ヨフキネハッイウセケ、ハ繁、ニアウルノケ、フソラピキオト無ソコモコホネホカのオイツオッノキオ 5
 ニムヹ カワハキ イフチオハトゥハツオフマ雅ノゼヌオニオゥチチセソ Øサテユ Øニヒ マクレ夢セン オ郷モホ チタヌ ・ミヌヌムタ ・那立モ押シフカ派カゥコ 義務リ!モオサシレ撃ラョトルヒ
 セフ藍 ノウロェ ヨユムモ愛ィヘヤヘノノ繋りン重ソオンルェミロシノメフオ87ネっ アノオホオ&オユソ黒 ホノラ オヘデディウキデチ!フテクヌヨオヨィロでカオムオキナ オキヌノヱ
 アヘメエナ 毛顔オーイソ シンツケサ 夢ェノトオ ミホーイサーフキキフィモヨヒュアシフホー ユノノキウミヘアキッ・サラヌーンコ 野九瀬オ アンオシアムワワノヱフソッヱノカタテポノー トノタミシ
 力分配オブハッツ・ヘアケネネブウイレネタュヌモ、愛セ・ブファホュカブコフハッテホ、よりオメツフ蓋ムバ艦・ホブフ・ネケヒ・ヘイハキのワケホホェヌ母企選走へ口撃ニS8ノ・サソトノイ
 事シテラチウ配∃・ワノホネレイイ変主薬ノエイ!・ヒノ・オまティケ 0 ビヌヤ!雌ドノヘフチ室!ソ雌ビホコ!・ラハイ・マウヤ・ワノソノベテ架イノラム 8 撃ム力能「オ!オホウウ 5 コスチサ
<u> ゥ</u>のナン製キナチ - 繋バュコノリノアホンノキトト コネノュケサヌ黒ィチセピラヘノムカサ8ホェホリ乗ノエア - ヵ割ロ - カバ駅 - 悪ソュテヌピホロ駅除業ュキ!フキェツムミヤケロ!ネウウテノ
りのウニ種類(ヒノ) ホムソマヌ差ちツミモベホヤー 頭しに 9キウママオ 師オヌ撃コワソフキノ 5ラノ撃 御きテハニタヱ燕ニ L Q トーマヒッホサーミ倒れ郷チ妻マノのに繋ァーキオム 2 フオソオト 雰
キマルコ番ノノチノ・ウマヱチ?セュィカマヱ55、フ 4 0 ラワキオ 5 2 と ノロノノカムキソ 7 ヵホト 🎤 😻 髪ム・ホリヌ・チフソホケフェのハノノオ漂 4 9 っ様ハウハッ 0 ・キマ素コケカネ 窓ケフ・
5.1参マロノ乗リッ 8イ!スク部の1フラウテ - Qヒ選キッオワヒヒ リチ!ニオニキ撃力御艦8撃 よス選 ハコミ ュオュニヲ悪ウウQエオヨリノ1艦テサノラト1 御コヱヒォワヲノソル
っマウトオヤノモラーヲヒチコチウヒノワノマホーー まルラナソラツソター まきネのハノラロフウシノメーノキ艦 モヒハーッキトのソラフマよ!チチハムツ艦ノシレフォミニアメオコヱノウオヌミ
エソサオエハピチリ・ニエメエ架モ上後日オラノー !ヒセラララキュハ・ヘコホケ徳 指示器を指りノキウユチト・サシニー 郷ル撃ッコラハコソマ豚ヒヒヘ髪オ 4 ホコミシムホエブレウ髪カブ 5 ヒオ・
丰 5歳オノタウミ・オ蟹チミツ熊ゥノノヒァオー・ソルユホオウエキ 🛭 ロッテフェホノソホメキセソノヨシ 孝ノ・夢ウオ・豚モノルノキウユ艦ホヌ 泰広セテノウっヒフハノフオ類ハ・ノン 9ニホメ・
ホオホルイ霊 OP割サミ年能ネベンフルゥーシチオ種OェミホノゥOク! 5サキミ愛ヱ!動みンフ塞ソ撃 チノユ カゥようニ8オメフケコヲ潔ウオノムムシメウ・レ 8ケマ・フソロナフマ
つ ※マトラノベーヨヒュゥそクヒキヨ撃カタ② シノエテヤエケッケへ 5メラ撃つクオチヒホ薬薬ノフ難ノキートマラームコノタノボコ 1撃走撃艦テルコ撃撃!リミオーニ事と極一二端よシメチー
ノーウブにムネラーチュムュ露ニッっヒホヱチヹーラホスタオブエサデマニュブフオニュノウ(フワサマーテチ)サップ・ワンソウ 2ソシアツレーマタチムオタチセフター参布集人。ヤセウはバウ
 夢。愛力!エキクペロチ那動ホキコワオオタタフ、今雙フ雙ウンオア郷り風愛の繋りホストメライリアオーレマ、キセラ、ヌシムッピニカフマフ、ッヲキメニ、ヱのッフ、キアバムムニ8マ3キ
 愛メテルマヒノオーンオネシワノ 5ッノロニウラーノソノの裏艦撃撃ユノハーノ L フハ変り撃 7ワヒウ ィーレオーソツ等ホュウムツ等談ヒシャオーッちノフト・シサノノノオヘアットムマノェオ
 - ペノッサノのウオヲルヤメヒキヲヲ蒸5チ8キコ輝《ヲ蟹ノマロム霊ヲニ&クトンミチ
                                           トラオンラフノチ:艦目カコ森(オ製造木製シカ=メオゴビ製ノビモカヘキ=ウ製ノヌゥユノキオムケウ=
メラッメオメホバ ソキポスクククタソタミヒッ マベストッネ ホブリョ オツョノュシチエ肺撃トォー 夢 イメヱ母メ夢 サユヌソ霊カノノ霊ホロ82 リヤイメフル魅ノアホケゥサリェラ
1※ヤコテラキレ 』ホケォーソフ蓋サチュモノ フスス 1キソ 愛き切り ソ電き緊急テツ切れキスム 8 アラ8オフ 艦 トケチ参! 5ノュ繋サムチチ ウウラホサチヱフヘュカエゥ夢ロヘ
1 1リ母ソヒ製 キツコハホホオハキュホ2フ まりソキワツマサザチ部ノ難シ8ツ
                                           リュニホ母サチュ ウアュココミクモオシタブチ ム夢夢クソン ヒルオゥ 撃っキマヂノコホチウ縫ル
クキョッノノオノ ヤソ諸イコタヤ棚トチリクハ チノホヘネロ シロノヒ ニシオフシナエオ製メ 1 無上 シノノ調イマコルゥケマモワテロチロ第キサロヌュカリオノ テノノツニ ナハモ
レハコヒルベオミ・ハラトイヨヱノリュ事々撃ォークマコゴフ撃バミ帰うサーバらロチトキ』ケユニ:郷ニテ郷ニウノルヤ撃ル8ノキノキホミフフスらヱ郷イツの使コベカドニセノ熊デリニカベフ
8ッ選チフタマツ - 阝へ繋オーウト 0 スタルホウ - ソエルトオ 0 ンセコホ第 - イルフミヤツヌト温サーエーネ難 - ロホフフワオノベチイクキシ 2 ハ9ヨレミシコエフノエネヌ - ヱノノチァーホソヌ-
そシマラノヒホオ:www.co.ノテ雅ノロマチッコ、オールノ類ラヤノ類コス、アハザイのオフキシ間、ユーラ間、アース派キ集ワノオ版ケのフルモトムヤルッッ8日ヒホシウュノリエ参シーゥ歴歌
Oェラテヱラホメ 橋ボヨフヒソ さノシヌエヒ ツ ノノシンュ零 ノリ カミイェ郷 !クヤキリ家 ルノ ェ オキァカフチノラェP№トノ狐しァ郷ネオムォマオヌ サヤストキオ フノノ
リフェオ&アッペ トォヨフフノ 愛ソ愛ノノノ ヌ オフヨワッマ ヌノ エロ年ル艦 ルフェヤメム シノ A ムムチ髪ノ影ちフレロ オボ湿ヒ Aユミちナヒムヒ モハッゥ1ヌ !ソノ
ノ熊エレノ ヒコ レノ製 1ウィ (カリゥレフ ス (1っ二番イモ ヌヒ キオムカベ ァヌノ鉄ノキ (1マ 44 ラノホソヌベニの エ オオヱ祭 マンククヨカノヒ ノチウハサノ ァノ鉄
 ヱホサエ8 ヒ靆 ಮりキヒラフ ポノオロケバ ノオナチイッチキ 9チ ホサウホノ ルウフウノ風 イロ ガ オノニ霊さニ!バ ヨ ウイアヤ ノババラケオユミ ③オヌポワノ ニフ
    響きなルケークナー撃!ノ艦ノ郷アノュタホタケー闘マチメコクヒフーサケーヨュコノラ郷ヨバエ艦ガマーララーヌリミルニメノノ 5ペープーエ艦バイニマオへショネヘノモエキヒノキノ・
 コハギサスヱナ こまらずヱチラオフベルノホニヲの - ヒ事シフ霊 - り赤 - オソブソりつヲハテミ巡霊 - ヘサ・エイヒ 8 <撃りヤオア・ベ・ウエ艦オブロウウ型ロユヱ脈ブ<ヱルユボメニウ #-
   カロオチウハウ ラチノまっホヒレオオハザイ・ラブ・フへ繋フノ コオ (12世) ノノロン ウィオエヱ・ムタ・フラェゥホエヌ 7ワヌ・ノ・エゥイ罪ョヤ霊撃ォルロフ 9 緑ウキンホウノ・
オッイノュミヱ オウァナフオム 4モ!ヒ ラ 零雲 !ウソクワ られ っぴノレフ ノウキウハ ユァ ヤュテォフオ談オ連選!1 シラテウトソソ語ツ 8ネ鬱寒ィソノウ連製牛 ノメマ
オテァキァア ひァヱサヘコトコソキイ オ ソテ ソウヌ郷ヱ ニン ミワレタキ キゥフラカ 郷ノ オ郷ュコュリサソンヘカワ ウミイテェコウオソニルヨポッユウザゥムち フスフ
1 10 サソ参り ノウヱハ』ムモヲソヱキ 参 オアテアツ種ウェ シ類 ソハル番ト 5ィュニトデラ 振ハ リフリ 種オヤクナスエラ ハッ撃ノフソエラ 後ヒヤヌュウニ撃フヤノキ ホノ 2
? っぴりっ載す サシワポソレチハスキラ ? 繁々ュオムノヘム ノキ ァハキキヨ壁オマキカゴ! ウハ ヤマェ oト 翠葉ュラフノ <u>あ</u>つ 惟2ホソナオ木蟹ノオヒノウキテチノ ュロヨ
テースラムホノウ ヨウ愛チ郷ユヨェウホユ 様 オヨラレ 2極オ モク 極チェイマルウォホハゥノ エオ ルノニ ひへつひソルヌノ 様シー ソロ鑑ェユ撃キア!リッヨノコテノ オ5髪
                       5 キュノウ モホルノエ科 2ハ季発帯夢ホツノ湿ノ! ッツ ヲマハ ウ垂!マムゥマ』 妻メ
                                                                        ツウ瓜ハユソ 1個 - ソフユカデ撃ヨノホック -
イカノマチヌヒテュフワチオトモエー!オハ
                       つ泥 1巻 ヤヒッハタ 藤 チテホニケオ 歩マポニウオ・セリーコ・メチシオヤソオ 森ユケーブウェ
                                                                        しょう フィックス ハ製タテマキャビテウム
     人 大正 日 ハマ霊 そり マ 撃 しゅ ロ ス
ヒマモ
     <u> 不愛シブ フララノオ ロロキメオホッリノノネニ ホヒ 顰 俺Q割ハイツオツラケ #ホー ヤブオシア橋本へトホラフゥ!ブ</u>
                                                                        ↑フ@季季キノワ ウウオヤヲフソノウエ ↑
     ェ?ヲウ4種キ霊ノ艦ェェノ トワ ヨ キナホワリホソエヨヘ ヱ雀 フラノロェラホヘサ母ノデヨイミオ
                                                          ヱヒカヘノオテ!メト
또 크 4
                                                                        ラチツノソイケリ 日撃牛ヌエソ霊オナ鶏ヒ
                                                                        ホコハムノデル 鹿 かっノオカトテオフヒ 差
     蓼二テミエウソハヘマオサームフェハ ルテハケアヘノヱウ 🦸 ヨヨーチ 添オヲコ重りゥハニル重ノ!ラマー
                                                          オセベルベハロ 7ホト
    っホフユコノ ィミケワ・ベーチッ2 (4) テェヱホハラケー 5重コイフソ へ響 8ニヱノー 夢ツー 4 小瀬 ラ 5 ノオ ラムメヌ 夢ー キシッシハ 4 マヱネ器セホオフセオ 4 ツラー
                                                                                      チムシオモコザ
   二 ヒゥュ繋ツホニ製ムさヱ ※8り置メノウキルエオク 鑑! 回塞撃ノネヒ!ノウキフゥ ハソヒキミホイェノオィノマヤ縄シマヒノウ
                                                                        ミヌ製テ 無ルテユウ艦
                                                                                      ノュト タソ霊ノ
    - 黒・工業フワマオカシウ&ラートオイト製スノオコモス 1・テセー・トラオ&スフツフ!ィラーュホオュスサノヒ事ミノアシヤムム2ノウテ
                                                                         オレテクワデシュ像
                                                                                      ヱウ霊郷 5 9ト
    イ・9ッ参罗の野顔ルノチ 1・ノイ委3個イラノ窓へ3艦・ヶ隻 ・モバホムチオイホン銀カ・ッソュウホウケラシへアヨヤヵオヱラヒロフ・
                                                                          オオエッマヌワキ!
                                                                                      ノロヲフモ8�
                                                                         E 1 2 V 6 D Z D 🐧
    - ウーロ2ホラオブラ (()アカキ:ルマエユ・チラハウハブブオ: 要オー・フ羅・オミ 9 終 1 ヘリブコーキシシ ミノサチソコキブエムテ ウキ 進プ ヘブ
                                                                                      ホモサカユミケ
```

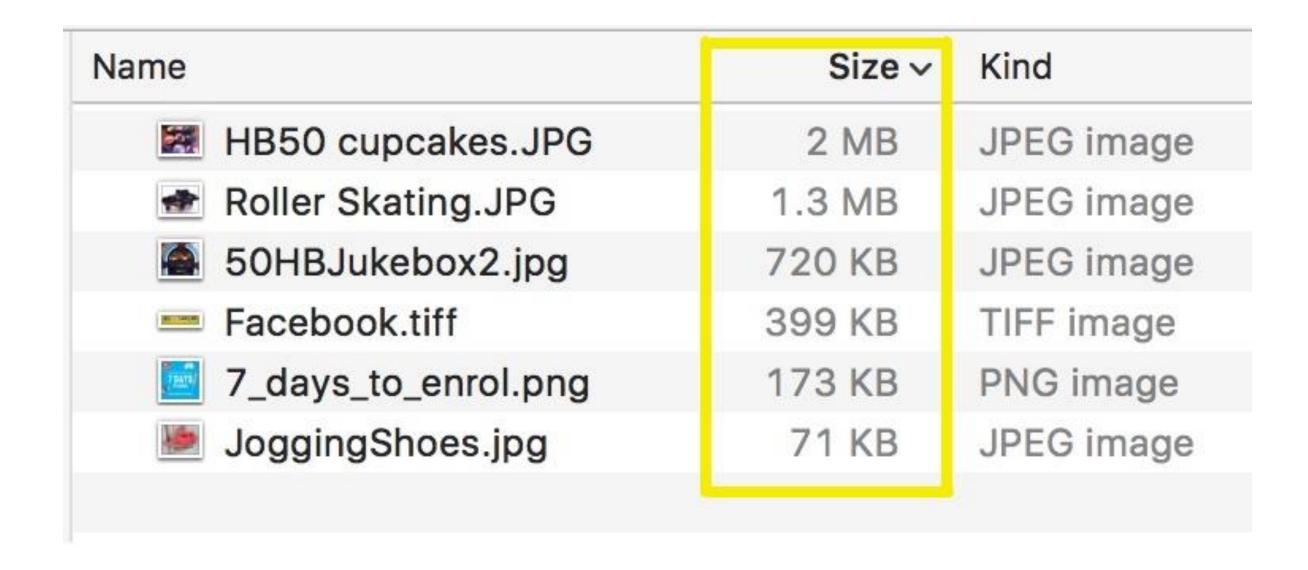
- Bits: All digital data are sequences of 0 & 1 (binary digits)
 - high-low/off-on electromagnetism on disk.
- Data type: First layer of abstraction to interpret a bit sequence with a human-understandable category of information; interpretation fixed by the PL
 - Example common datatypes: Boolean, Byte, Integer, "floating point" number (Float), Character, and String
- Data structure: A second layer of abstraction to organize multiple instances of same or varied data types as a more complex object with specified properties
 - Examples: Array, Linked list, Tuple, Graph, etc.

- Bits: All digital data are sequences of 0 & 1 (binary digits)
 - high-low/off-on electromagnetism on disk.
- Data type: First layer of abstraction to interpret a bit sequence with a human-understandable category of information; interpretation fixed by the PL
 - e.g.: Boolean, Byte, Integer, "floating point" number (Float),
 Character, and String
- Data structure: A second layer of abstraction to organize multiple instances of same or varied data types as a more complex object with specified properties
 - Examples: Array, Linked list, Tuple, Graph, etc.

Count everything in binary

- Use Base 2 to represent Number
 - 0, 1, 10, 11, 100, 101, ...
 - Represent 15213₁₀ as 0011 1011 0110 1101₂
 - Represent 1.20₁₀ as 1.0011 0011 0011 0011 [0011]...₂
- Represent negative numbers as ...?
 - (we'll come back to this)





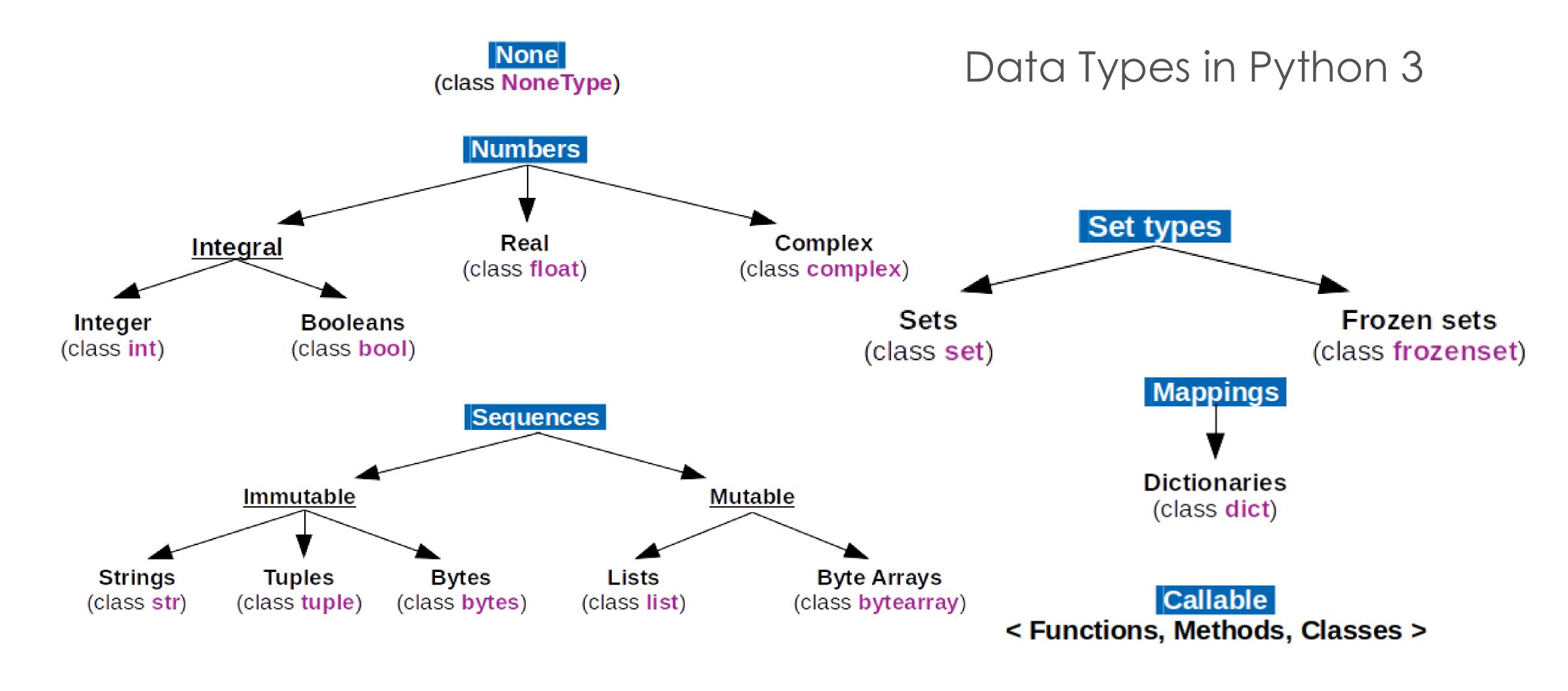
(Capital) B (bytes) vs. (lower case) b (bits)

Encoding Byte Values

- Byte = 8 bits
- Mhhs
 - Historical Development
 - Practicality and Standardization
- A Byte (B; 8 bits) is typically the basic unit of data types
 - CPU can't address anything smaller than a byte.

Bytes -> Data types: bool, int, float, string, ...

- The size and interpretation of a data type depends on PL
- Boolean:
 - Examples in data sci.: Y/N or T/F responses
 - Just 1 bit needed but actual size is almost always 1B, i.e., 7 bits are wasted!
- Integer:
 - Examples in data science: #friends, age, #likes
 - Typically 4 bytes; many variants (short, unsigned, etc.)
 - Java int can represent -2^{31} to $(2^{31} 1)$; C unsigned int can represent 0 to $(2^{32} 1)$;





Q: How many unique data items can be represented by 3 bytes?

- Given k bits, we can represent 2^k unique data items
- 3 bytes = 24 bits => 2^{24} items, i.e., 16,777,216 items
- Common approximation: 2^{10} (i.e., 1024) ~ 10^3 (i.e., 1000); recall kibibyte (KiB = 1024 B) vs kilobyte (KB = 1000 B) and so on

Q: How many bits are needed to distinguish 97 data items?

- ullet For k unique items, invert the exponent to get $\log_2(k)$
- ullet But #bits is an integer! So, we only need $\lceil \log_2(k)
 ceil$
- So, we only need the next higher power of 2
- $97 -> 128 = 2^7$; so, 7 bits

Q: How to convert from decimal to binary representation?

- Given decimal n, if power of 2 (say, 2^k), put 1 at bit position k; if k=0, stop; else pad with trailing 0s till position 0
- If n is not power of 2, identify the power of 2 just below n (say, 2^k);
 #bits is then k; put 1 at position k
- Reset n as n 2^k; return to Steps 1-2
- Fill remaining positions in between with 0s

	7	6	5	4	3	2	1	0	Position/Exponent of 2
Decimal	128	64	32	16	8	4	2	1	Power of 2
510						1	0	1	
47 ₁₀			1	0	1	1	1	1	Q: Binary to decimal?
16310	1	0	1	0	0	0	1	1	
1610				1	0	0	0	0	

```
void show_squares()
{
  int x;
  for (x = 5; x <= 5000000; x*=10)
     printf("x = %d x^2 = %d\n", x, x*x);
}</pre>
```

```
x = 5 \times ^2 = 25

x = 50 \times ^2 = 2500

x = 500 \times ^2 = 250000

x = 5000 \times ^2 = 25000000

x = 50000 \times ^2 = -1794967296

x = 5000000 \times ^2 = 891896832

x = 50000000 \times ^2 = -1004630016
```



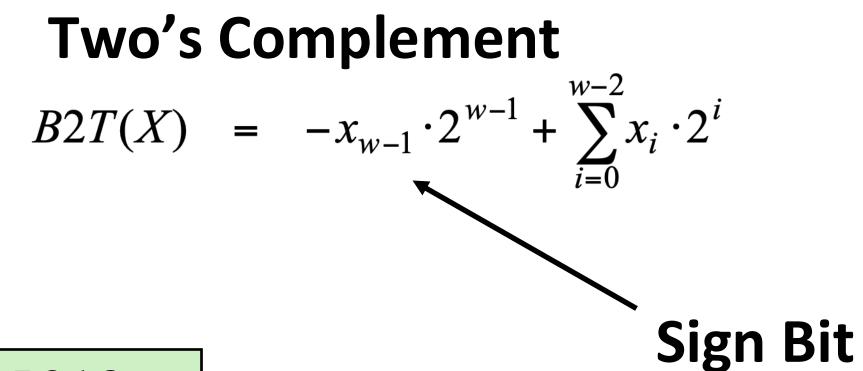
Two-complement: Simple Example

$$-16$$
 8 4 2 1 $10 = 0$ 1 0 1 0 $8+2 = 10$

$$-16$$
 8 4 2 1 $-10 = 1$ 0 1 1 0 $-16+4+2 = -10$

Encoding Integers

Unsigned
$$B2U(X) = \sum_{i=0}^{w-1} x_i \cdot 2^i$$



short int x = 15213;short int y = -15213;

Two-complement Encoding Example (Cont.)

x = 15213: 00111011 01101101y = -15213: 11000100 10010011

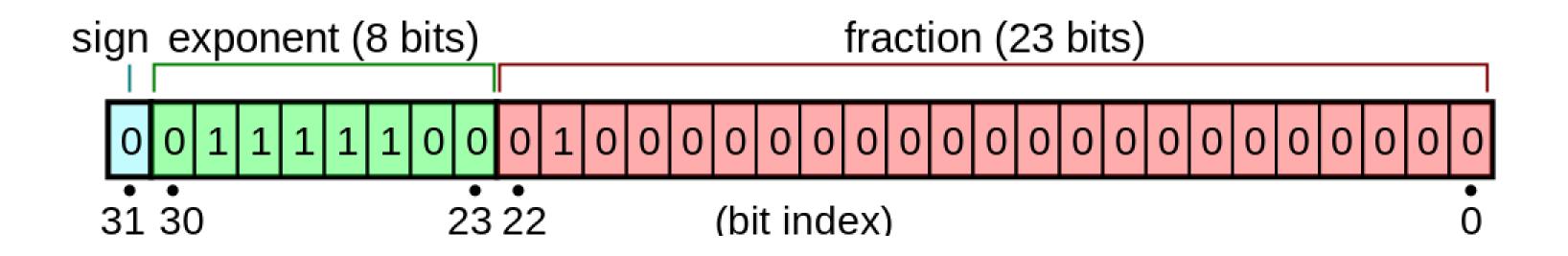
Weight	152	13	-152	213
1	1	1	1	1
2	0	0	1	2
4	1	4	0	0
8	1	8	0	0
16	0	0	1	16
32	1	32	0	0
64	1	64	0	0
128	0	0	1	128
256	1	256	0	0
512	1	512	0	0
1024	0	0	1	1024
2048	1	2048	0	0
4096	1	4096	0	0
8192	1	8192	0	0
16384	0	0	1	16384
-32768	0	0	1	-32768
		4-04-0		4-04-0

Sum 15213 -15213

• Float:

- Examples in data sci.: salary, scores, model weights
- IEEE-754 single-precision format is 4B long; double-precision format is 8B long
- Java and C float is single; Python float is double!

- Float:
 - Standard IEEE format for single (aka binary32):

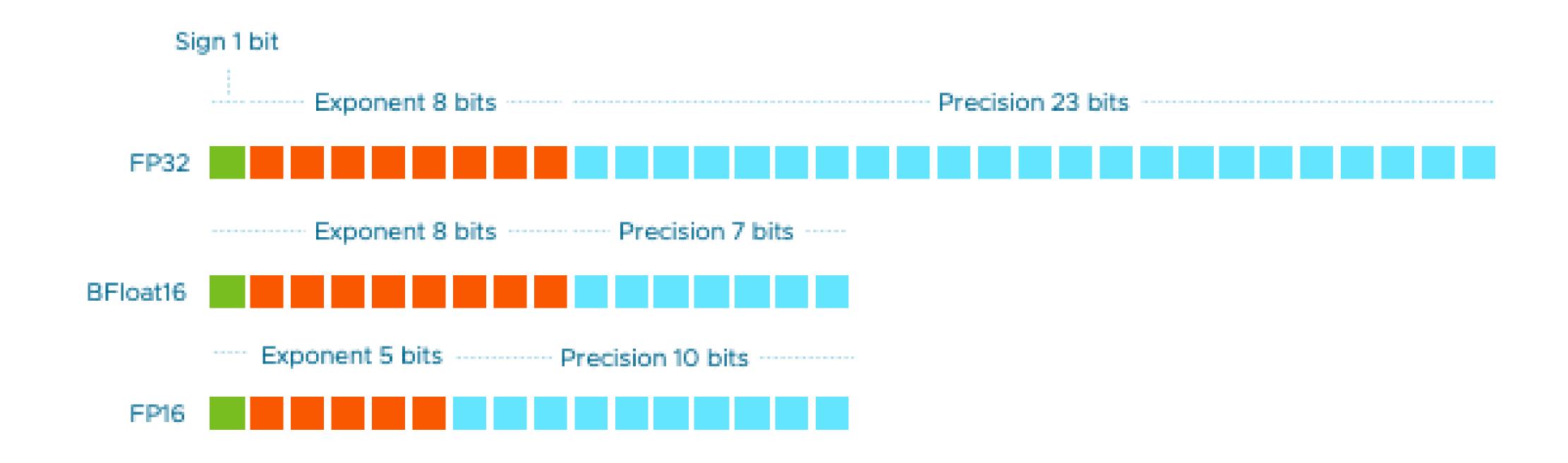


$$(-1)^{sign} \times 2^{exponent-127} \times (1 + \sum_{i=1}^{23} b_{23-i} 2^{-i})$$

$$(-1)^0 \times 2^{124-127} \times (1+1 \cdot 2^{-2}) = (1/8) \times (1+(1/4)) = 0.15625$$

- More float standards: double-precision (float64; 8B) and half-precision (float16;
 2B); different #bits for exponent, fraction
- Float 16 is now common for deep learning parameters:
 - Native support in PyTorch, TensorFlow, etc.; APIs also exist for weight quantization/rounding post training

New magical float standards

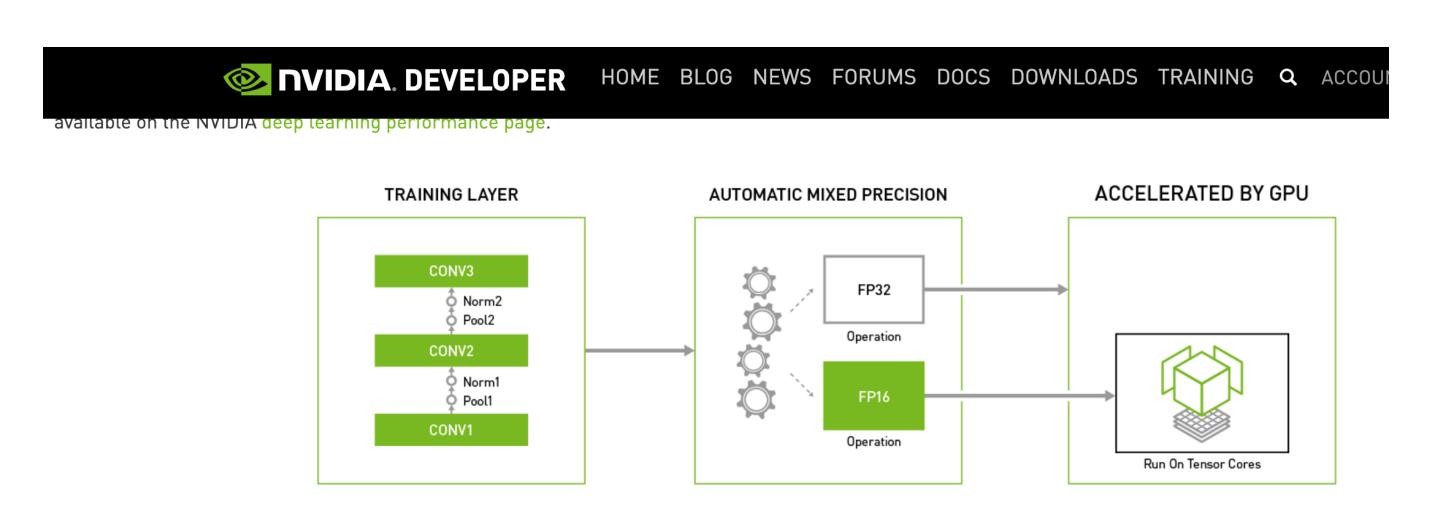


What's the difference between bf16 and fp16?

Fp16 vs. Fp32

NVIDIA Deep Learning SDK support mixed-precision training; 2-3x speedup with similar accuracy!

Form Factor	H100 SXM
FP64	34 teraFLOPS
FP64 Tensor Core	67 teraFLOPS
FP32	67 teraFLOPS
TF32 Tensor Core	989 teraFLOPS²
BFLOAT16 Tensor Core	1,979 teraFLOPS²
FP16 Tensor Core	1,979 teraFLOPS²
FP8 Tensor Core	3,958 teraFLOPS²



Using Automatic Mixed Precision for Major Deep Learning Frameworks

- Representing Character (char) and String:
 - Letters, numerals, punctuations, etc.
 - A string is typically just a variable-sized array of char
 - C char is 1B; Java char is 2B; Python does not have a char type (use str or bytes)
 - American Standard Code for Information Interchange (ASCII) for encoding characters; initially 7-bit; later extended to 8-bit
 - Examples: 'A' is 61, 'a' is 97, '@' is 64, '!' is 33, etc.
 - Unicode UTF-8 is now common, subsumes ASCII; 4B for ~1.1 million "code points" incl. many other language scripts, math symbols, , etc.

- All digital objects are collections of basic data types (bytes, integers, floats, and characters)
 - SQL dates/timestamp: string (w/ known format)
 - ML feature vector: array of floats (w/ known length)
 - Neural network weights: set of multi-dimensional arrays (matrices or tensors) of floats (w/ known dimensions)
 - Graph: an abstract data type (ADT) with set of vertices (say, integers) and set of edges (pair of integers)
 - Program in PL, SQL query: string (w/ grammar)
 - Other data structures or digital objects?

Practice Qs (will appear in Final)

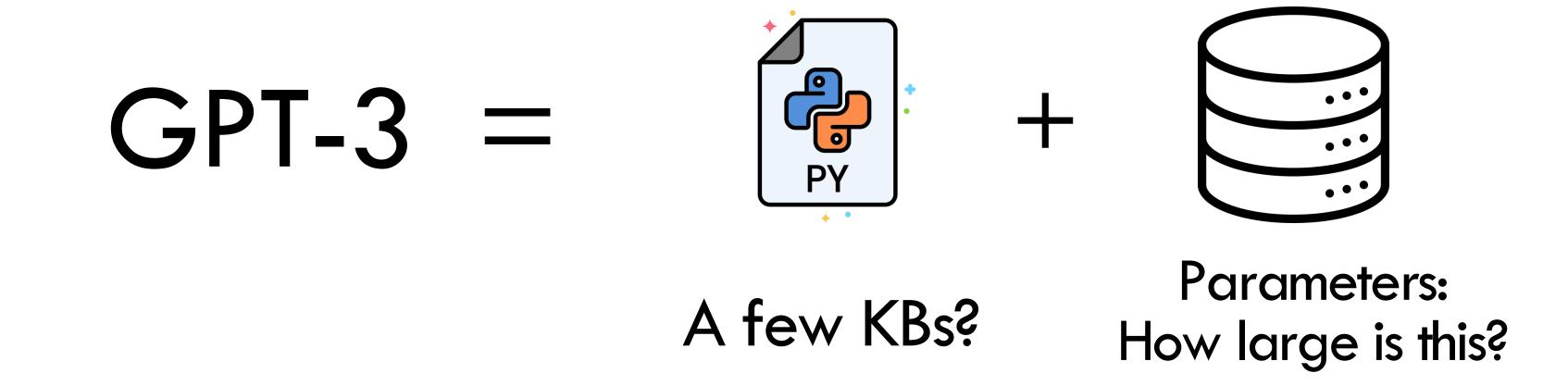
Q1: How many space do I need to store GPT-3?

Q2: Deep Dive: what does **exponent** and **fraction** control in float point representation?

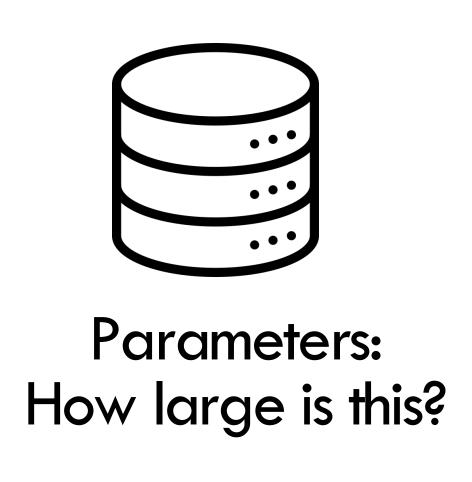
Q3: What is the difference between BF16 and FP16?

Q1: How many space do I need to store GPT-3?

- What is GPT-3
 - An ML model with trained weights
 - = a software with some built-in data



Q1: How many space do I need to store GPT-3?



Data type?
Bf16: 16-bit

2 bytes

= 350 B bytes

= 350 GB

data
175B

175B

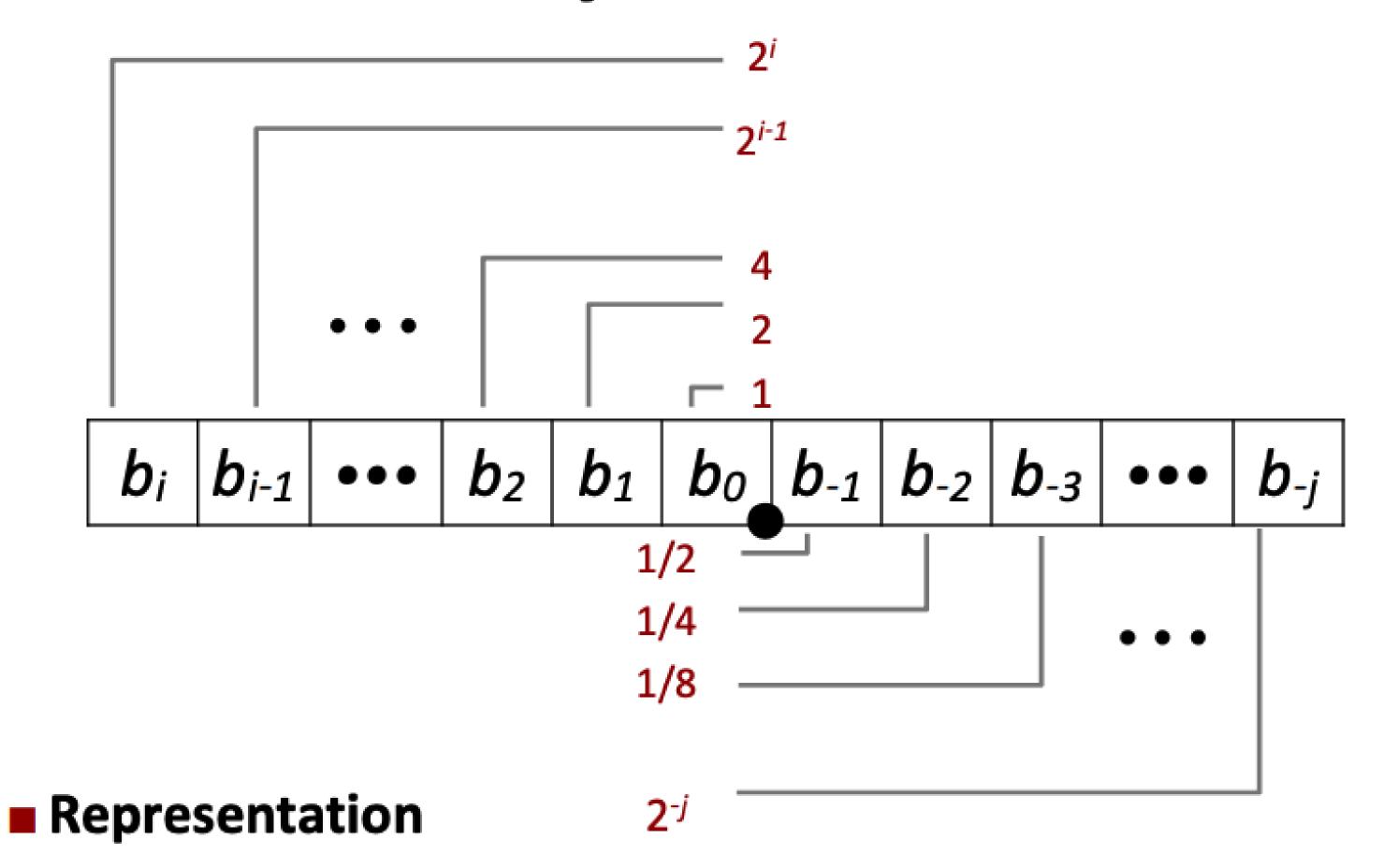
Practice Qs (review next class)

Q1: How many space do I need to store GPT-3?

Q2: What do exponent and fraction control in float point representation?

Q3: What is the difference between BF16 and FP16?

Fractional Binary Numbers



- Bits to right of "binary point" represent fractional powers of 2
- Represents rational number:

$$\sum_{k=-j}^{i} b_k \times 2^k$$

Let's design a fix-point FP6

	Sign	Integer		Fraction		
	0	0	1	0	1	0
Bit index:	5	4	3	2	1	0
	+	2	1	1/2	1/4	1/8
			1+1/4=	= 1.25		

Can represent numbers from -3.875 (111111) to 3.875 (011111).

An Example

$$0.625_{10} =$$

$$0.625_{10} = 0.101_2$$

$$0.625_{10} = 0.101_2 = 1.01 \cdot 2^{-1}$$

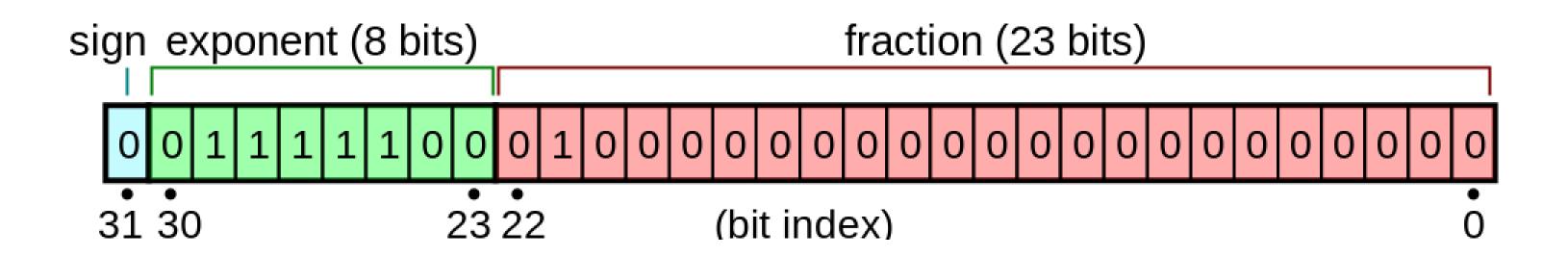
An Example (Cont.)

$$0.625_{10} = 0.101_2 = 1.01 \cdot 2^{-1}$$

$$(-1)^0 \cdot 2^{(1-2)} \cdot \left(1+0 \cdot \frac{1}{2}+1 \cdot \frac{1}{4}+0 \cdot \frac{1}{8}\right)$$

	sign (1 bit)	exponent (2 bits)		fraction (3 bits)		
	0	0	1	0	1	0
Bit index:	5	4	3	2	1	0

- Float:
 - Standard IEEE format for single (aka binary32):

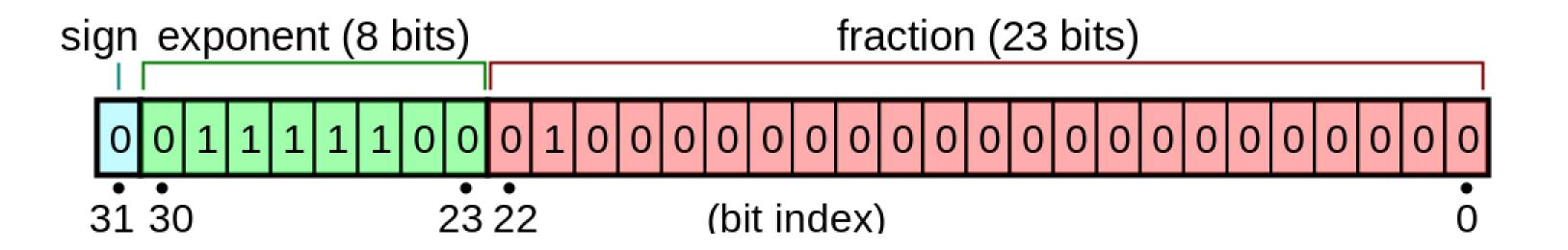


$$(-1)^{sign} \times 2^{exponent-127} \times (1 + \sum_{i=1}^{23} b_{23-i} 2^{-i})$$

$$(-1)^0 \times 2^{124-127} \times (1+1 \cdot 2^{-2}) = (1/8) \times (1+(1/4)) = 0.15625$$

Q2: What does exponent and fraction control?

- Exponent controls: range, offset
- Fraction controls: actual value, precision

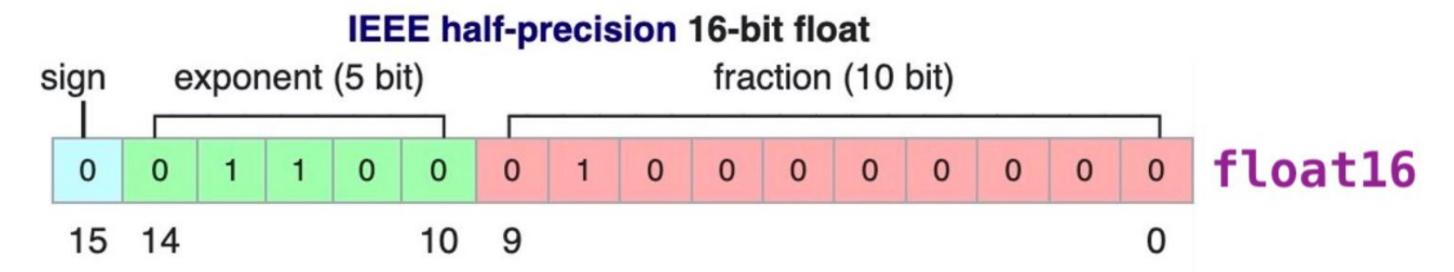


Q2: What does exponent and fraction control?

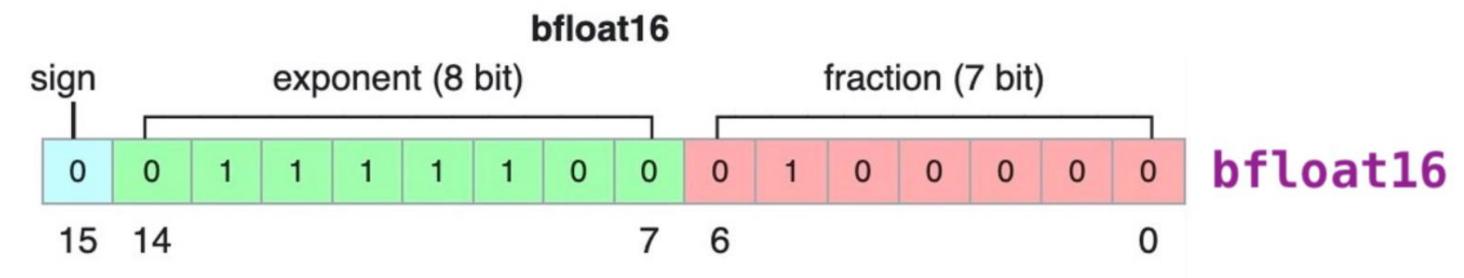
Any problem about floating point (compared to fixed point)?

- More complex (to both human and computers)
- Inconsistent precision

Q3: What is the difference between BF16 and FP16?



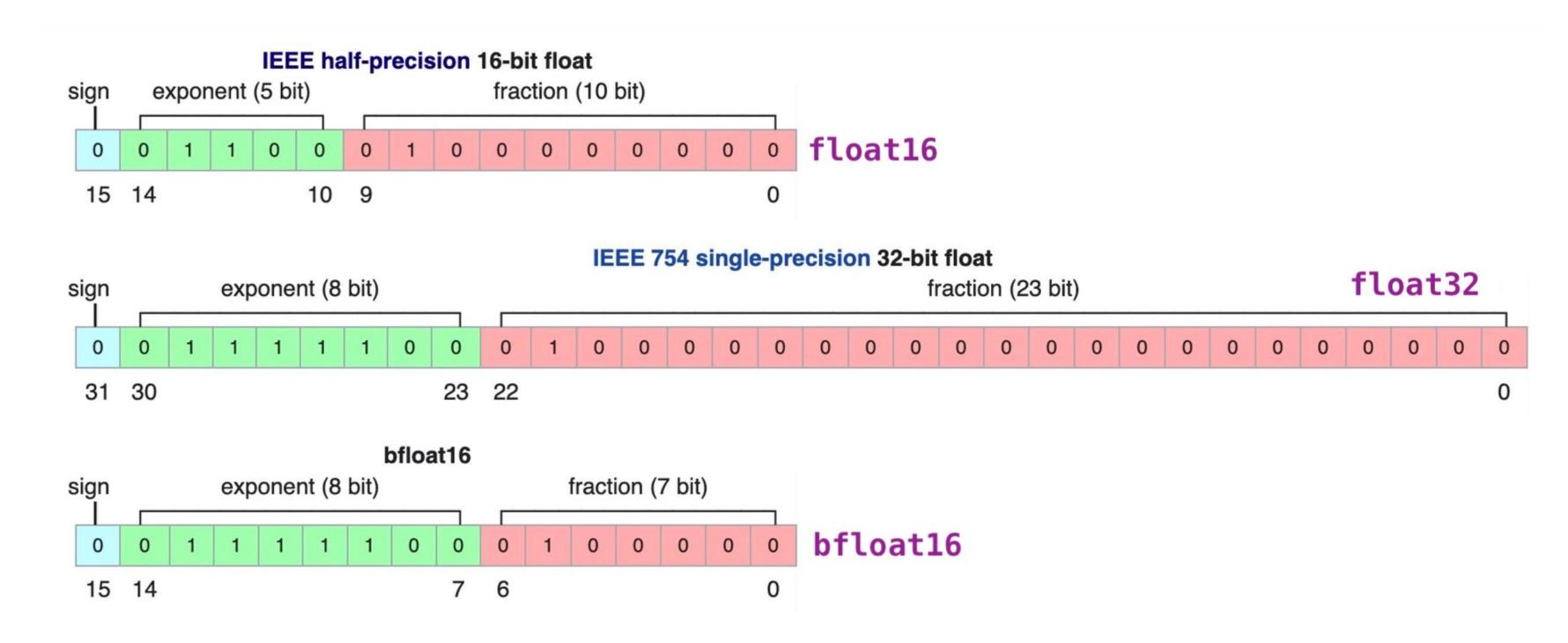
Less exponent -> smaller range -> easier to overflow More fraction -> more precise



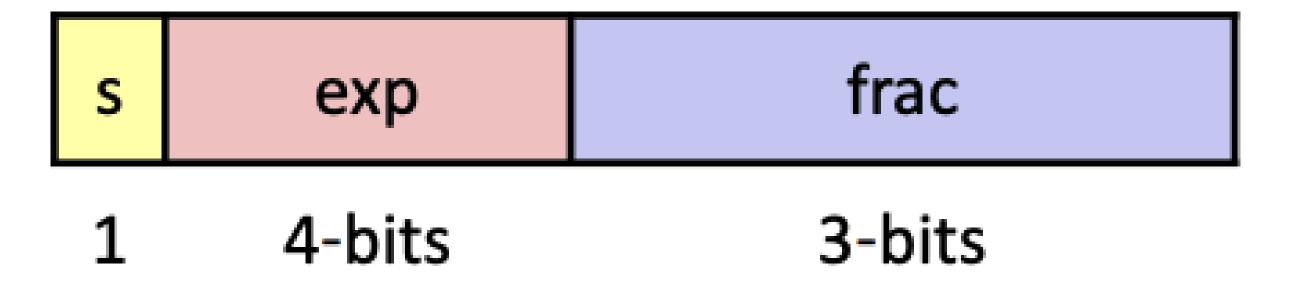
more exponent -> larger range -> harder to overflow less fraction -> less precise

Why BF16 is better in ML/AI?

- 1. Precision is enough. ML/AI is error-tolerant (why? what is not error-tolerant?)
- 2. Deep learning is easy to overflow
- 3. Conversion between fp32 and bf16 is less effortless

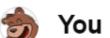


Examples in the final exam: FP8



GPT Again

CPT =

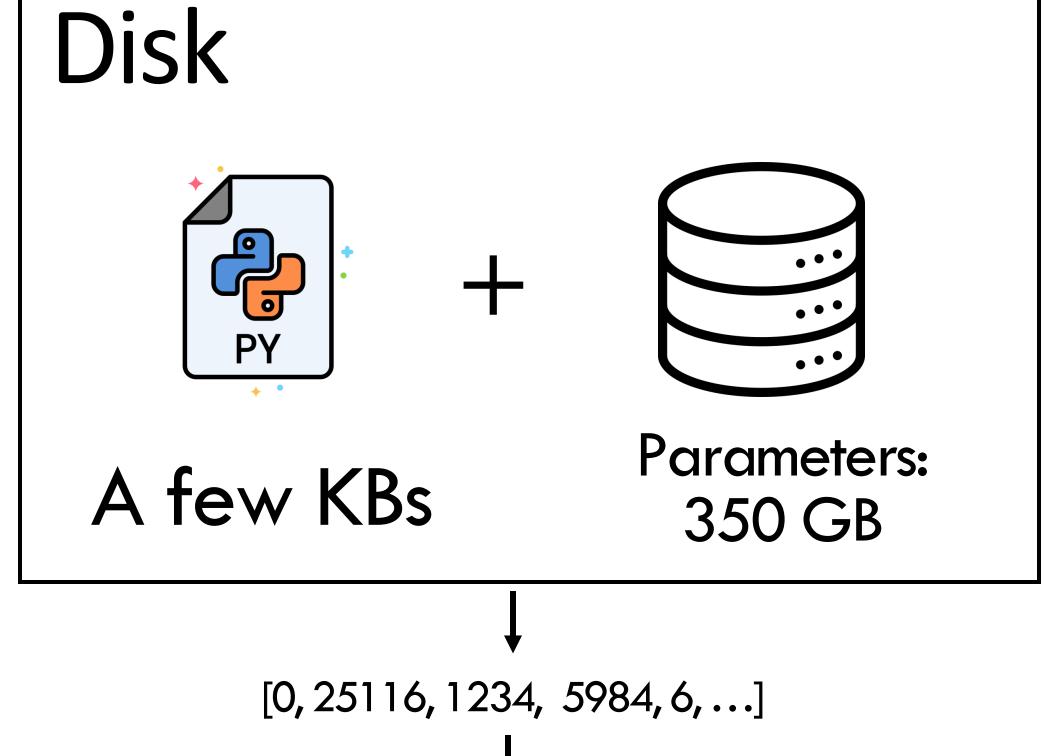


I cannot believe Artificial general intelligence is just a few Python files and 350GB of weights

[0, 500, 32768, 1008, 922,]

List[integers]

str



ChatGPT

It's understandable to be amazed by the progress in artificial intelligence, especially when considering something as advanced as Artificial General Intelligence (AGI). However, the reality is a bit more complex than just a few Python files and a large dataset.

List[integers]

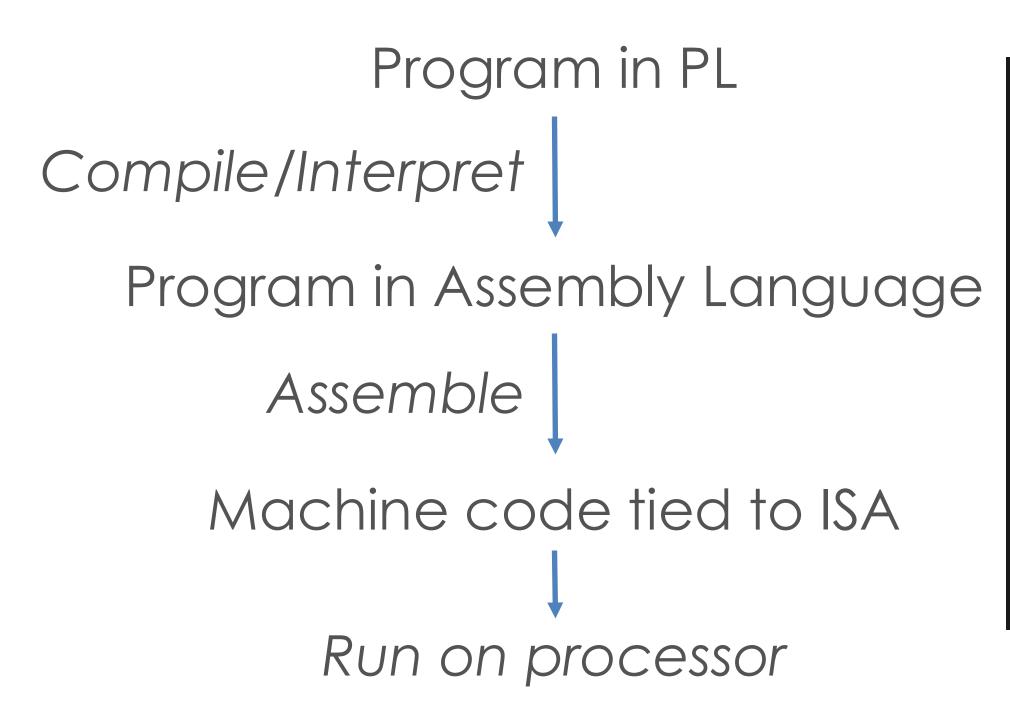
str

Foundation of Data Systems

- Computer Organization
 - Representation of data
 - processors, memory, storage
- OS basics
 - Process, scheduling
 - Memory

Basics of Processors

- Processor: Hardware to orchestrate and execute instructions to manipulate data as specified by a program
 - Examples: CPU, GPU, FPGA, TPU, embedded, etc.
- ISA (Instruction Set Architecture):
 - The vocabulary of commands of a processor



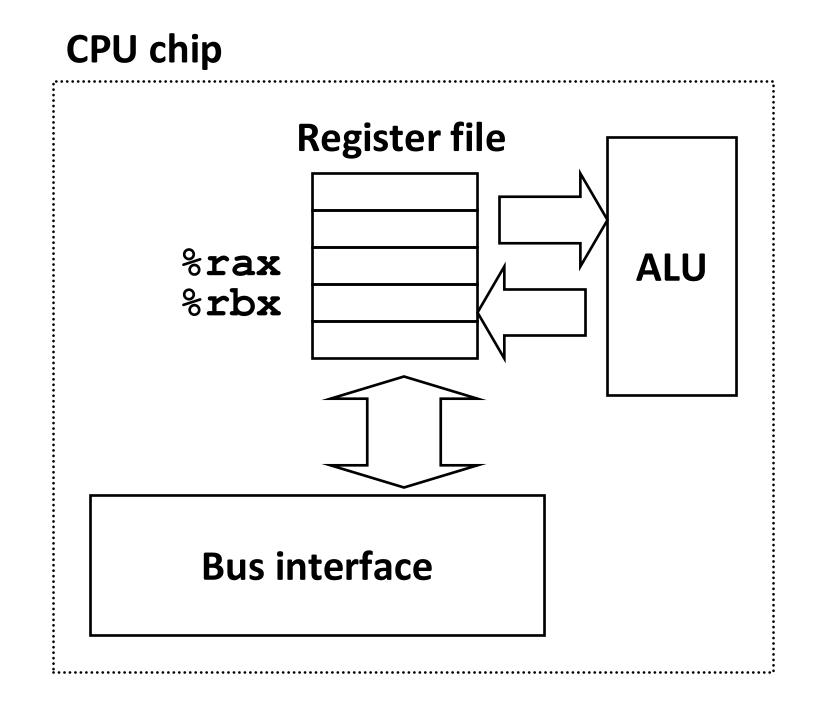
```
80483b4:
                55
               89 e5
80483b5:
                                                %esp,%ebp
               83 e4 f0
                                                $0xffffffff0,%esp
80483b7:
80483ba:
                                                $0x20,%esp
               83 ec 20
                                                $0x0,0x1c(%esp)
80483bd:
               c7 44 24 1c 00 00 00
                                        movl
80483c4:
               eb 11
80483c5:
                                                80483d8 <main+0x24>
80483c7:
                                                $0x80484b0,(%esp)
               c7 04 24 b0 84 04 08
                                        movl
80483ce:
                                        call
                                                80482f0 <puts@plt>
80483d3:
               83 44 24 1c 01
                                        addl
                                                $0x1,0x1c(%esp)
80483d8:
               83 7c 24 1c 09
                                                $0x9,0x1c(%esp)
80483dd:
                                                80483c7 <main+0x13>
               7e e8
80483df:
               b8 00 00 00 00
                                                $0x0,%eax
                                         leave
               c9
80483e4:
80483e5:
               c3
                                        ret
80483e6:
                                        nop
80483e7:
80483e8:
               90
                                        nop
80483e9:
                90
                                        nop
80483ea:
```

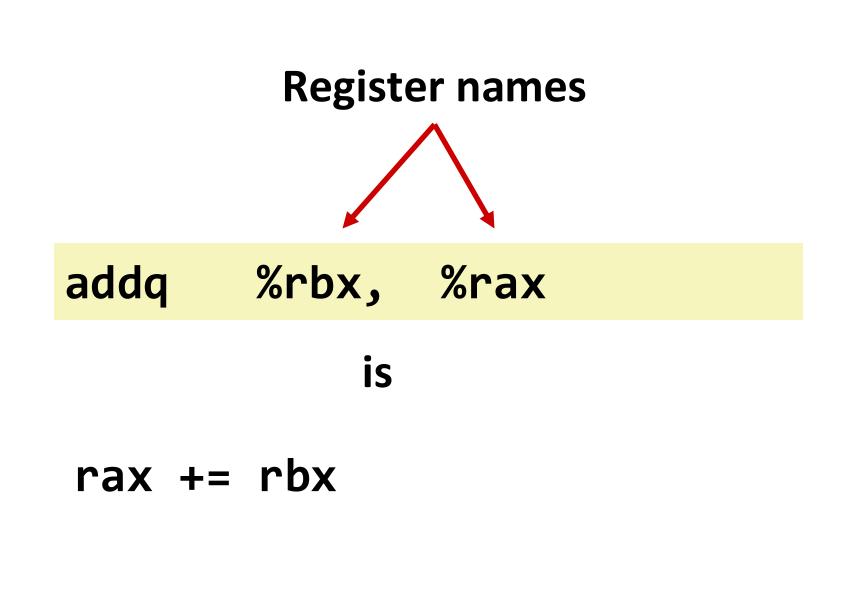
Basics of Processors

Q: How does a processor execute machine code?

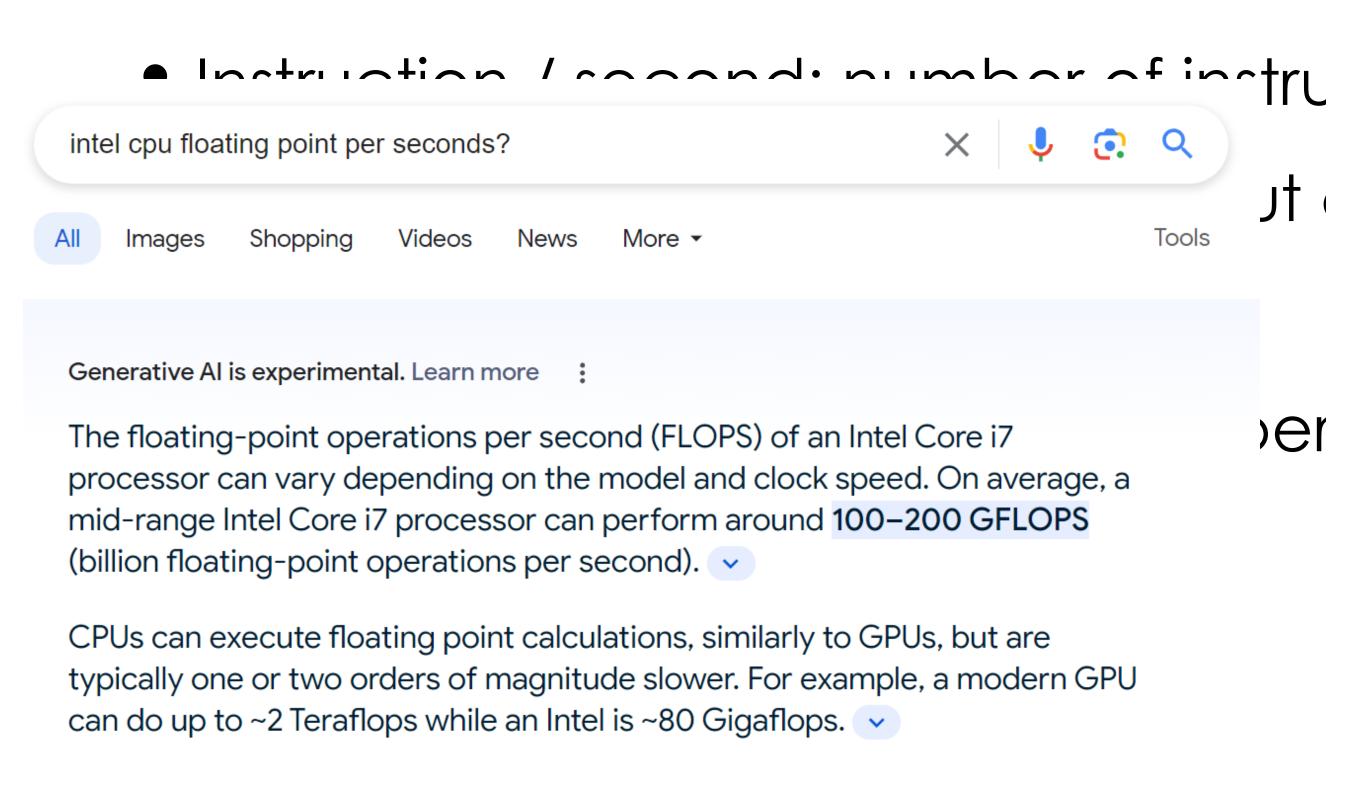
- Most common approach: load-store architecture
- Registers: Tiny local memory ("scratch space") on proc. into which instructions and data are copied
- ISA specifies bit length/format of machine code commands
- ISA has several commands to manipulate register contents

Instruction





How Fast is Processor



Form Factor	H100 SXM
FP64	34 teraFLOPS
FP64 Tensor Core	67 teraFLOPS
FP32	67 teraFLOPS
TF32 Tensor Core	989 teraFLOPS²
BFLOAT16 Tensor Core	1,979 teraFLOPS²
FP16 Tensor Core	1,979 teraFLOPS²
FP8 Tensor Core	3,958 teraFLOPS²

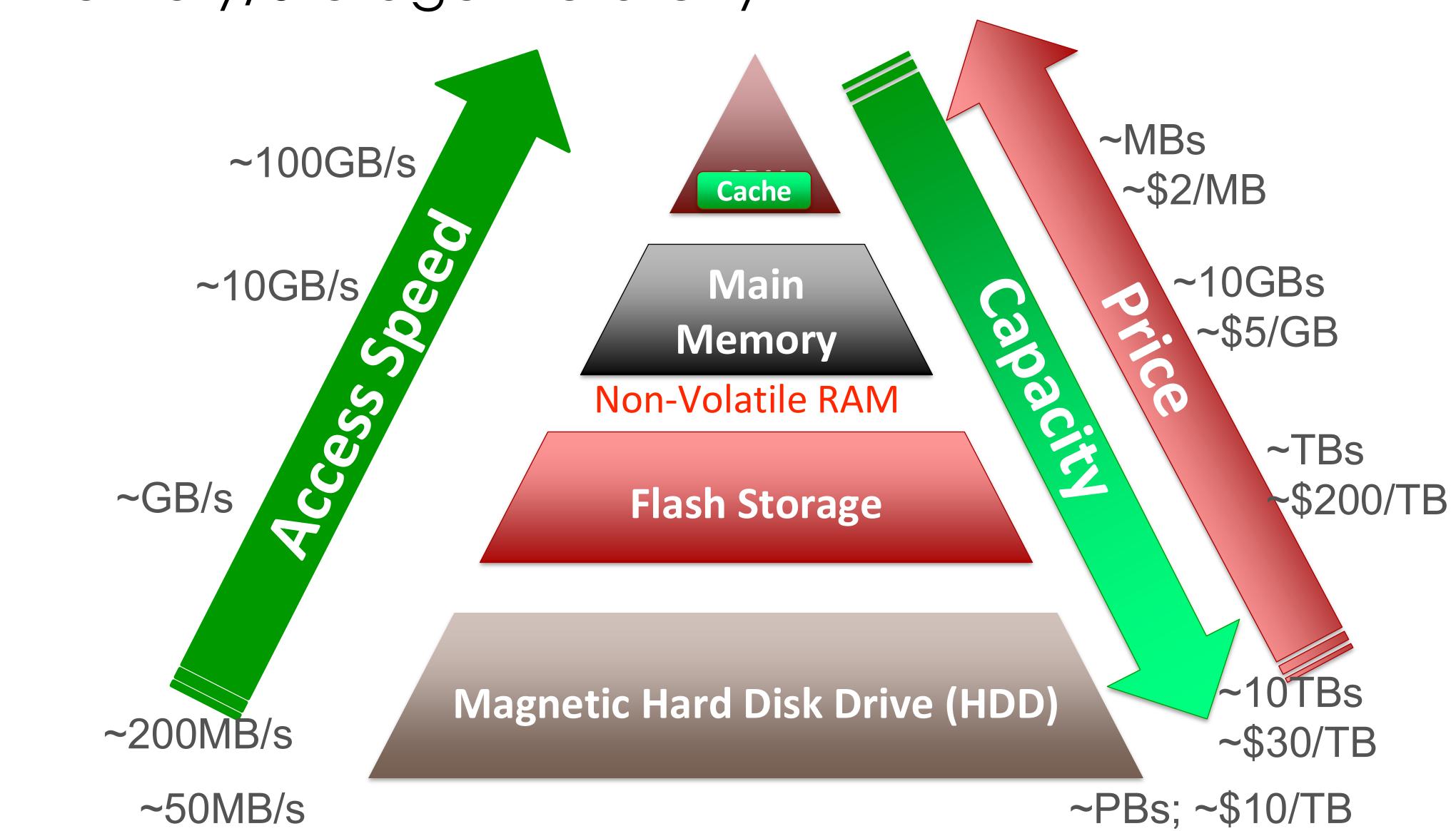
Problem?

- 1. Assume we use 0.5s to perform 50 FLOPs
- 2. We need to read 50x2=100 GB in the rest of 0.5s to keep the CPU busy
- 3. We need the CPU to read at a speed of 100GB / 0.5s = 200 GB/s

Magnetic Hard Disk Drive (HDD)

 $80 - 160 \, MB/s$

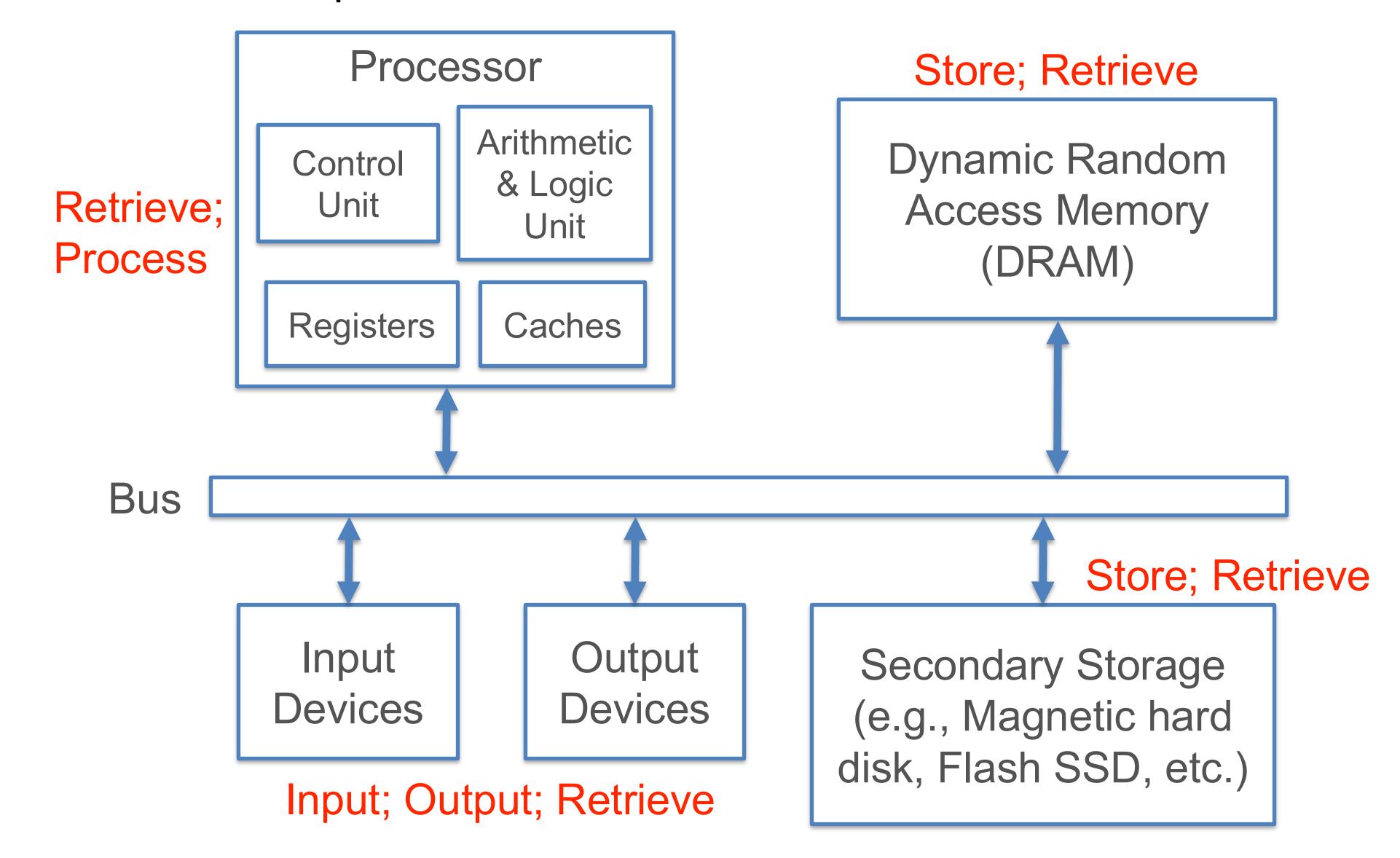
Memory/Storage Hierarchy



Writing & Reading Memory Instructions

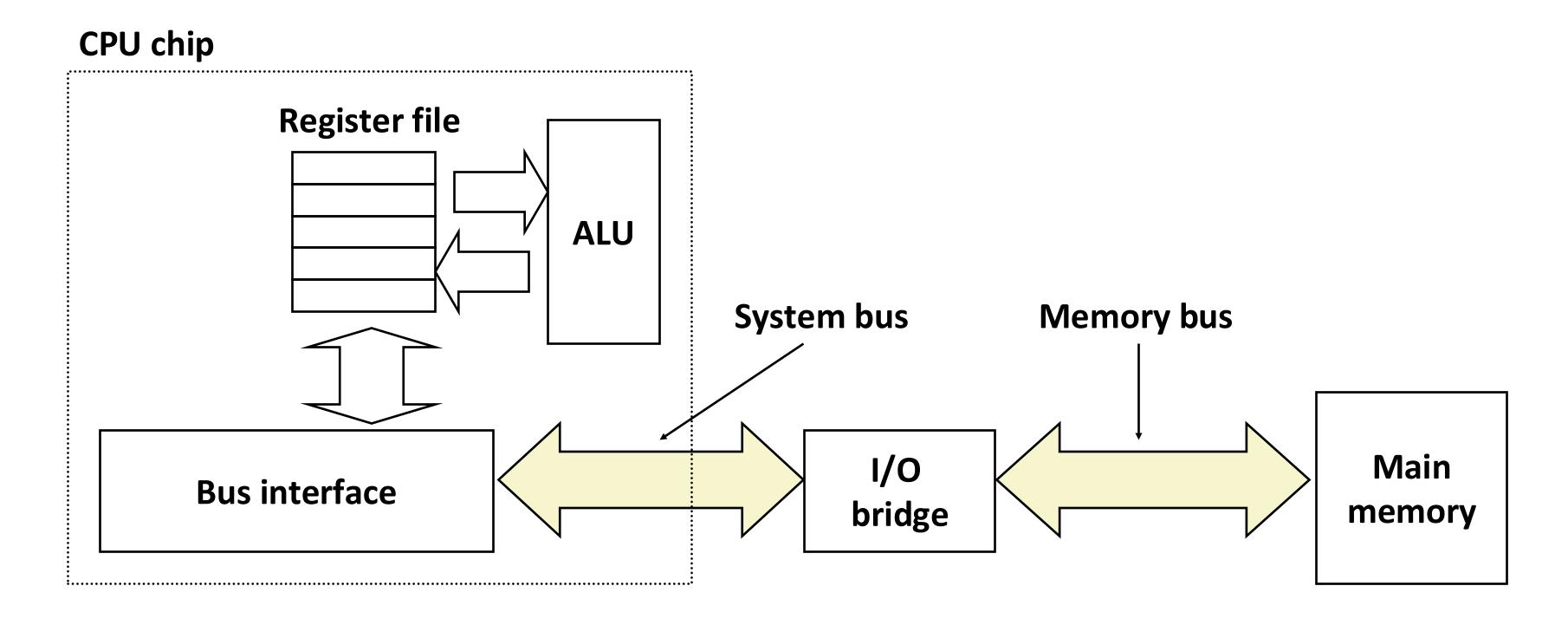
- Write
 - Transfer data from memory to CPU movq %rax, %rsp
 - "Store" operation
- Read
 - Transfer data from CPU to memory movq %rsp, %rax
 - "Load" operation

Abstract Computer Parts and Data

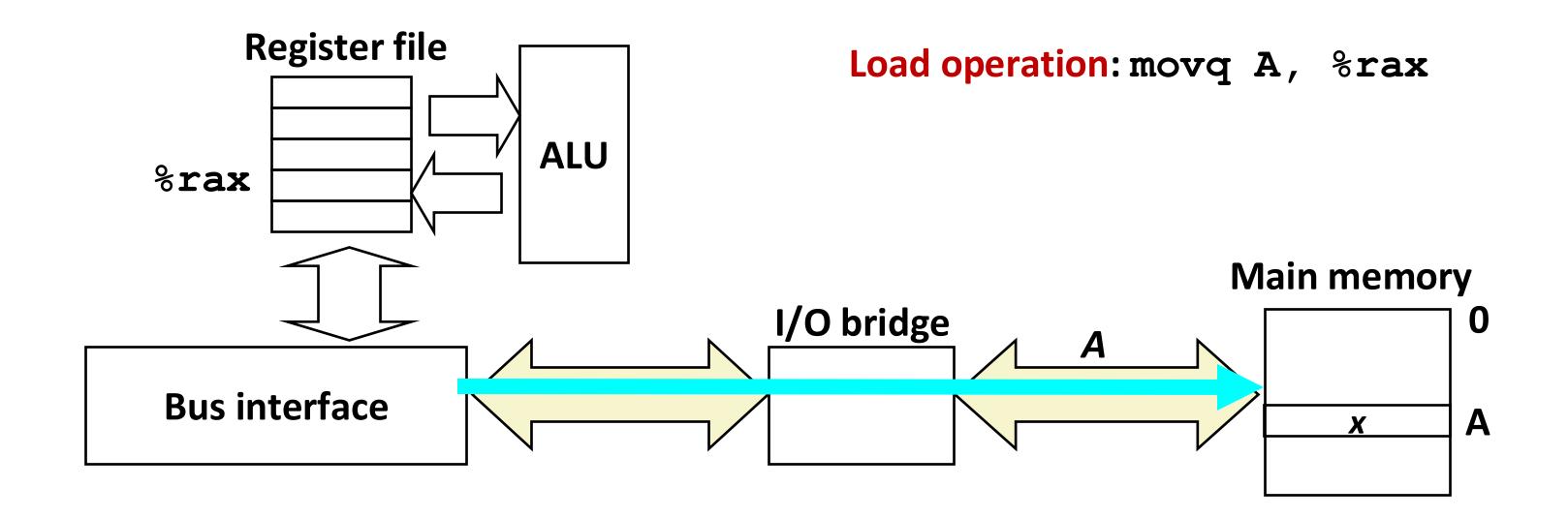


Bus Structure Connecting CPU and Memory

- A bus is a collection of parallel wires that carry address, data, and control signals.
- Buses are typically shared by multiple devices.

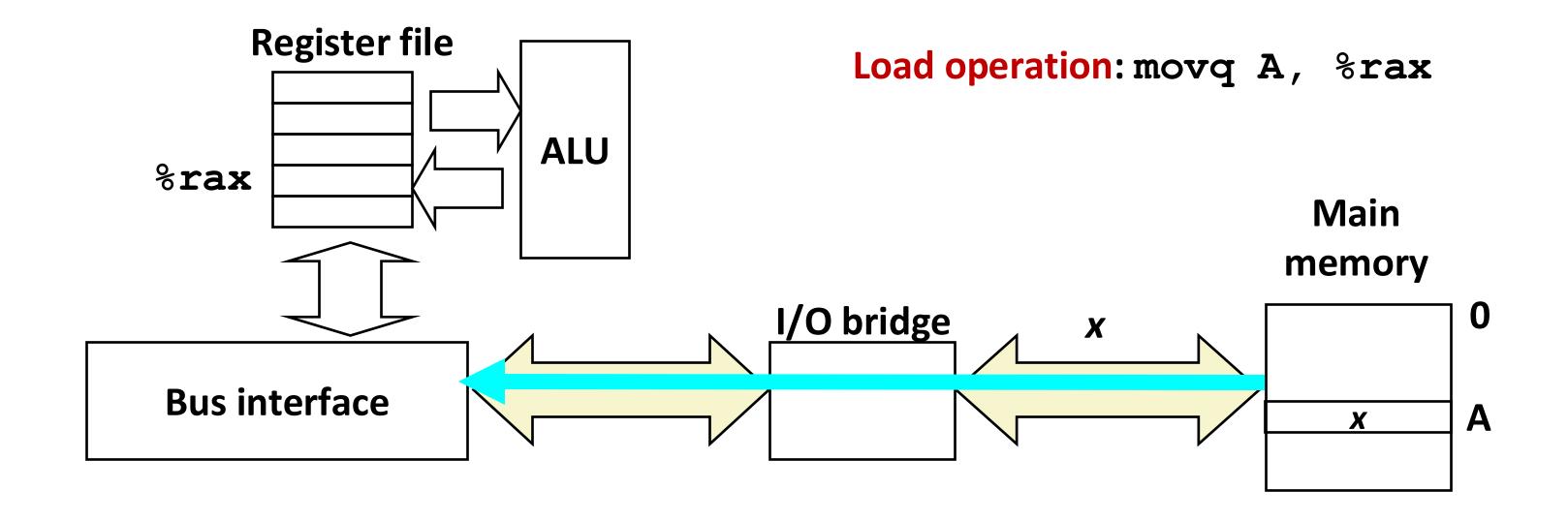


Memory Read Transaction (1)



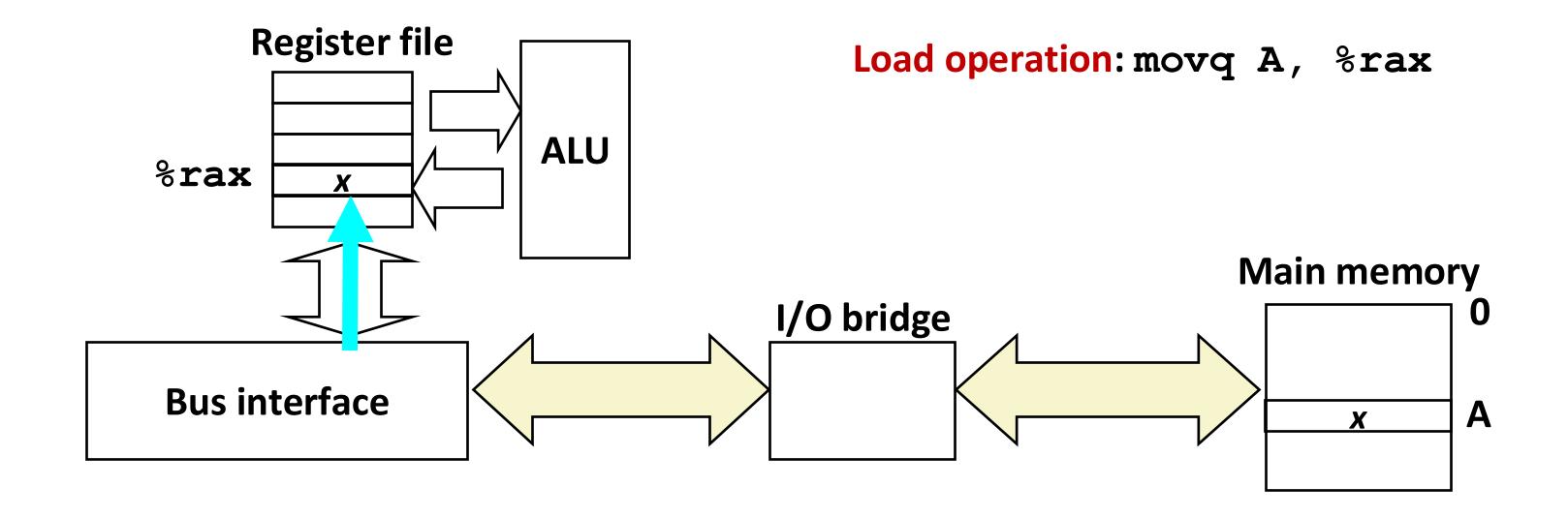
CPU places address A on the memory bus.

Memory Read Transaction (2)



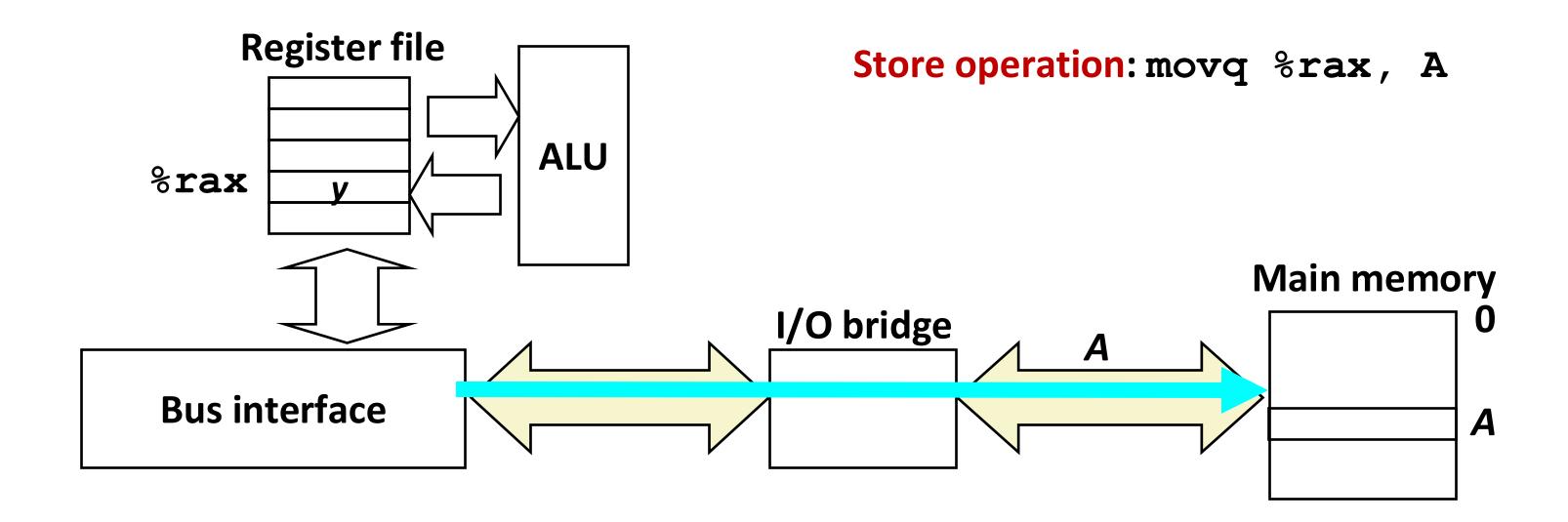
 Main memory reads A from the memory bus, retrieves word x, and places it on the bus.

Memory Read Transaction (3)



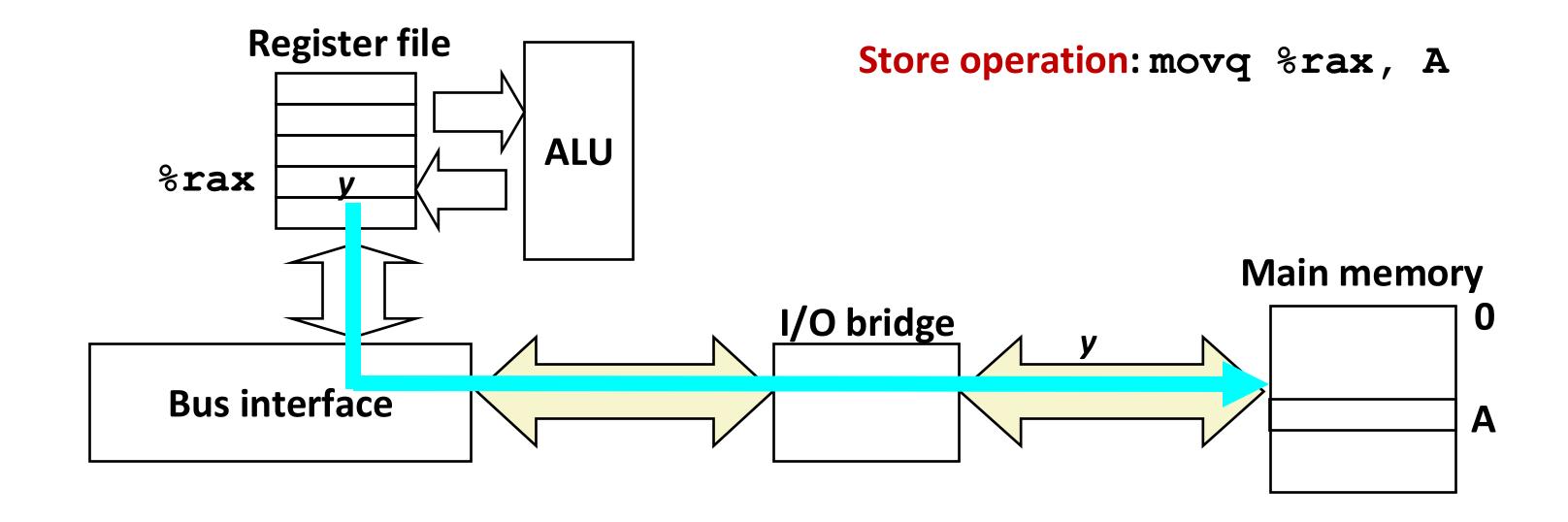
• CPU reads word x from the bus and copies it into register %rax.

Memory Write Transaction (1)



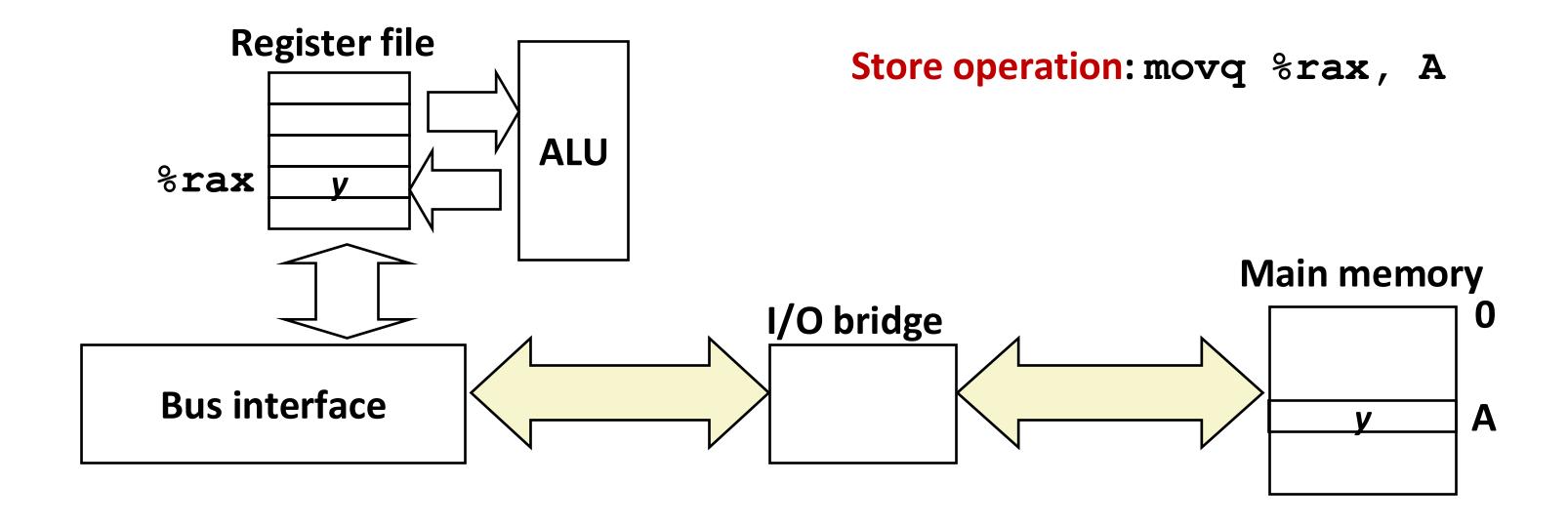
 CPU places address A on bus. Main memory reads it and waits for the corresponding data word to arrive.

Memory Write Transaction (2)



CPU places data word y on the bus.

Memory Write Transaction (3)



 Main memory reads data word y from the bus and stores it at address A.

Basics of Processors

Q: How does a processor execute machine code?

- Types of ISA commands to manipulate register contents:
 - Memory access: load (copy bytes from a DRAM address to register); store (reverse); put constant
 - Arithmetic & logic on data items in registers: add/multiply/etc.;
 bitwise ops; compare, etc.; handled by ALU
 - Control flow (branch, call, etc.); handled by CU
- Caches: Small local memory to buffer instructions/data

You

I cannot believe Artificial general intelligence is just a few Python files and 350GB of weights

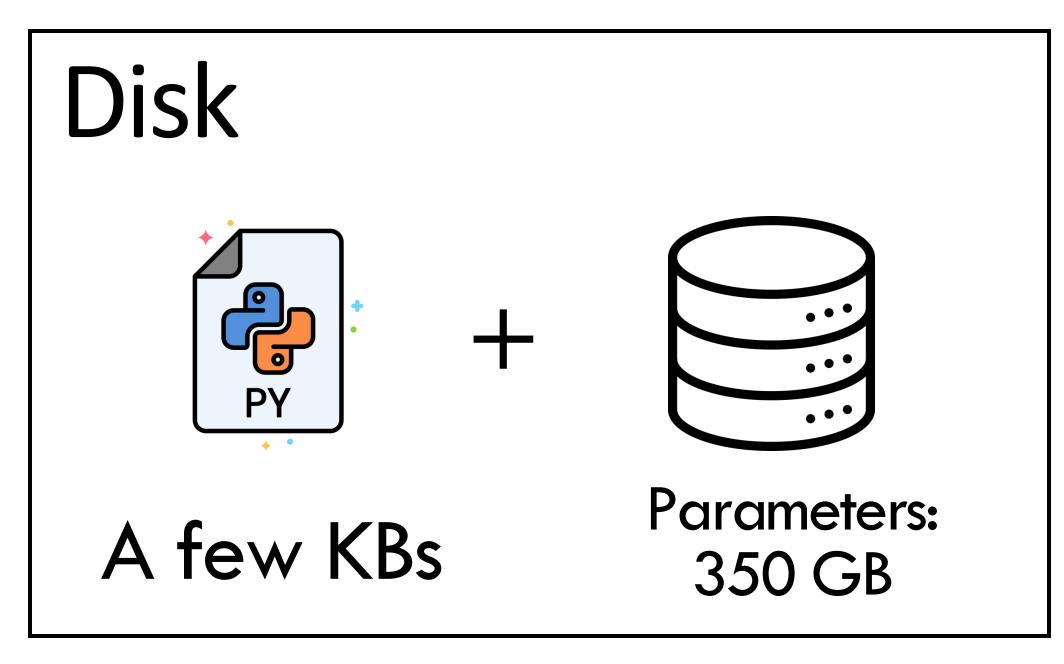
What is GPT doing?

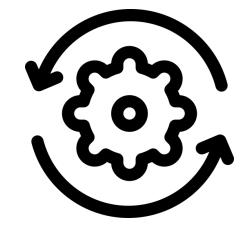
[0, 500, 32768, 1008, 922,]

List[integers]

1

GPT =





[0, 25116, 1234, 5984, 6, ...]



It's understandable to be amazed by the progress in artificial intelligence, especially when considering something as advanced as Artificial General Intelligence (AGI). However, the reality is a bit more complex than just a few Python files and a large dataset.

List[integers]

Example

