Assignment 3

Grasping Exa Scale

Onur Onel – 0865803
Dr. Sabine McConnell

Trent University
BSc. Software Engineering
COIS 1400

Submitted March 11, 2025

Grasping Exa Scale

In the world of Big Data and supercomputing, prefixes like Exa and Zetta illustrate the remarkable advances in technology. These measurements drive the development of supercomputers engineered to handle the most demanding scientific and industrial challenges. As supercomputing continues to evolve, it promises to transform our understanding of the world. In the following sections, using concrete examples, I will explore how extensive these measurements are and discuss their implications for modern science and technology.

Exaflop

Energy is all around us. Even though we can't see or touch it directly, we notice its effects in every moving car, every blazing fire, and every lightning strike. Einstein revealed that mass can be converted into energy, which is a cornerstone of modern physics and shows that even ordinary objects hold enormous amounts of energy. An exaflop of computing power is like a massive, untapped energy source. In the same way, an exaflop holds vast computing potential, much like the hidden energy within matter itself. The matter-energy conversion primarily occurs in nuclear power plants through equivalence. In theory, if we consider a human body weighing about 70 kilograms, using $E = mc^2$, converting all that mass into energy would produce roughly 6.29×10^{18} joules. To put that into perspective, this amount of energy is roughly 100 million times greater than the energy released by the atomic bomb dropped on Hiroshima. In a more detailed example, the energy equivalent of a human's mass could theoretically provide the global energy needs for about four days. Given that the global daily energy consumption is roughly 1.6×10^{18} J/day (Ritchie, Rosado, & Roser, 2020), a 70-kg human would have an energy equivalent calculated as $E=mc^2 = 70 \text{ kg} \times (3.0 \times 10^8 \text{ m/s})^2 = 6.3 \times 10^{18}$ Jules. Dividing the total energy by the daily consumption, 6.3×10^{18}J / 1.6×10^{18} J/day, yields roughly 3.94 days. In conclusion, these analogies show that both the physical and digital worlds are driven by massive forces. Just as energy fuels our real world, extraordinary computing powers accelerate innovation and keep us connected.

Exabyte

In computing, an exabyte is a massive unit of digital storage equal to 10^{18} bytes, which is equivalent to one quintillion bytes. To grasp this enormous data storage capacity, we can consider a straightforward example. My iPhone uses about 50 GB of iCloud storage, which is more than enough to store my details and all the associated app metadata, such as contacts, messages, social media photos, and music. In simpler terms, this 50 GB represents the entirety of my digital life, containing memories, vital documents, and daily communications. When we imagine scaling this storage requirement to the global level with approximately 8 billion people on Earth, assuming each person uses 50 GB of storage, the total storage requirement would be about 400 exabytes.

1. 50 GB × 8,000,000,000 = 400,000,000,000 GB
2. Since 1 exabyte (EB) = 1,000,000,000 GB,
3. 400,000,000,000 GB / 1,000,000,000 GB/ = 400 EB

In today's technology landscape, achieving a storage capacity of 400 EB requires building large-scale data centers. For example, Solidigm released a new SSD in the first quarter of 2025 with a capacity of 122.88 TB (Enterprise Storage Forum). Data center SSDs cost around $14,000 each, meaning that constructing an entire 400 exabyte data center requires roughly 3.5 million SSDs, which would cost approximately $50 billion. On the other hand, HDDs remain the traditional workhorse for bulk data storage, offering a balance of cost and performance. Enterprise HDDs typically cost between $15 and $25 per TB in raw hardware terms. For instance, 18–20 TB enterprise drives often sell for $400–$700 each (Athow, 2025). Therefore, achieving 400 EB of storage using HDDs would cost approximately $6–$10 billion just for the hard drives. In conclusion, the fact that 400 exabytes can encapsulate the digital lives of 8 billion people is a powerful reminder of just how vast an exabyte is. It underscores the incredible scale of our digital universe and the tremendous challenge of managing such immense data.

References

1. Hannah Ritchie, Pablo Rosado and Max Roser (2020) - “Energy Production and Consumption” Published online at OurWorldinData.org. https://ourworldindata.org/energy-production-consumption

2. Enterprise Storage Forum. *Buying guide for enterprise hard drives*. Enterprise Storage Forum.https://www.enterprisestorageforum.com/management/buying-guide-for-enterprise-hard-drives/#:~:text=Pricing

3. Athow, D. (2025, January 21). *The largest SSD and hard drives of 2025*. TechRadar. https://www.techradar.com/best/large-hard-drives-and-ssds

Assignment 3

Assumptions on Estimation Accuracy

Onur Onel – 0865803
Dr. Sabine McConnell

Trent University
BSc. Software Engineering
COIS 1400

Submitted March 12, 2025

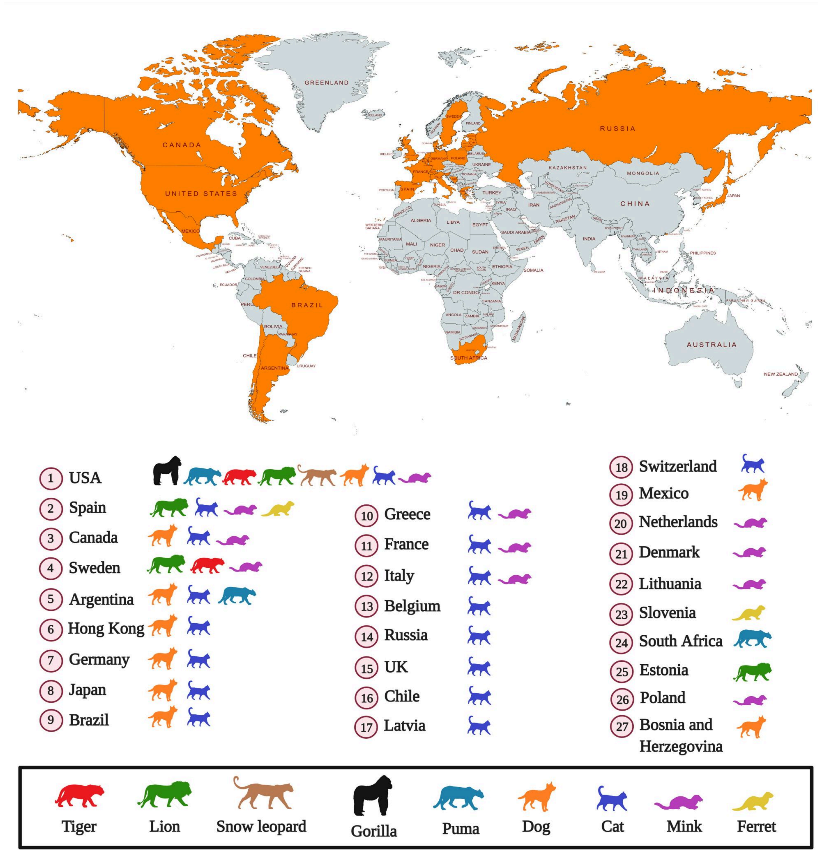
Assumptions on Estimation Accuracy

According to Professor of Mathematical Biology Dr. Christian Yates, the estimation of the total number of SARS-CoV-2 virus particles worldwide involves several critical assumptions. Daily infections are estimated at a steady 3 million, accounting for underreporting and limited testing. Each infected person is estimated to carry a peak viral load of approximately 10 billion particles, leading to a global total of roughly 2×10^{17} virus particles at any given time (Yates, 2021). Considering the virus's spherical shape, with an average diameter of 100 nanometers, the volume of a single particle is around 523,000 nm^3 (Yates, 2021). Multiplying this by the estimated particle count yields a total volume of about 120 milliliters. Accounting for inefficiencies in how spheres pack together increases this volume slightly to approximately 160 milliliters, which is equivalent to slightly less than a Coke can (Yates, 2021). However, each assumption made during this calculation introduces potential errors, which in the following sections will be carefully examined.

Potential Inaccuracies

1. Animal-Environmental Reservoirs

Dr. Christian Yates's calculations assume humans are the primary and most significant reservoir for SARS-CoV-2, ignoring the role animals and environmental sources might play in the overall viral load. On the contrary, animals and the environment can indeed serve as reservoirs, holding, replicating, and shedding virus particles into their surroundings, potentially contributing significantly to the global viral count (Sharun, 2021). Documented cases of infections in various species, such as wild deer populations, highlight substantial reservoirs beyond humans. For instance, outbreaks in mink farms resulted in billions of additional viral particles entering the environment, thereby increasing the global viral load considerably (Caserta, 2023). On the other hand, environmental surfaces, wastewater, and contaminated materials can temporarily sustain high numbers of infectious virus particles. These environmental reservoirs can persistently harbor viral particles, sometimes for days, potentially adding tens of billions of particles globally (Dargahi, 2021). Therefore, by excluding animal and environmental reservoirs from calculations, Dr. Yates's estimate could significantly underestimate the total number of SARS-CoV-2 particles.



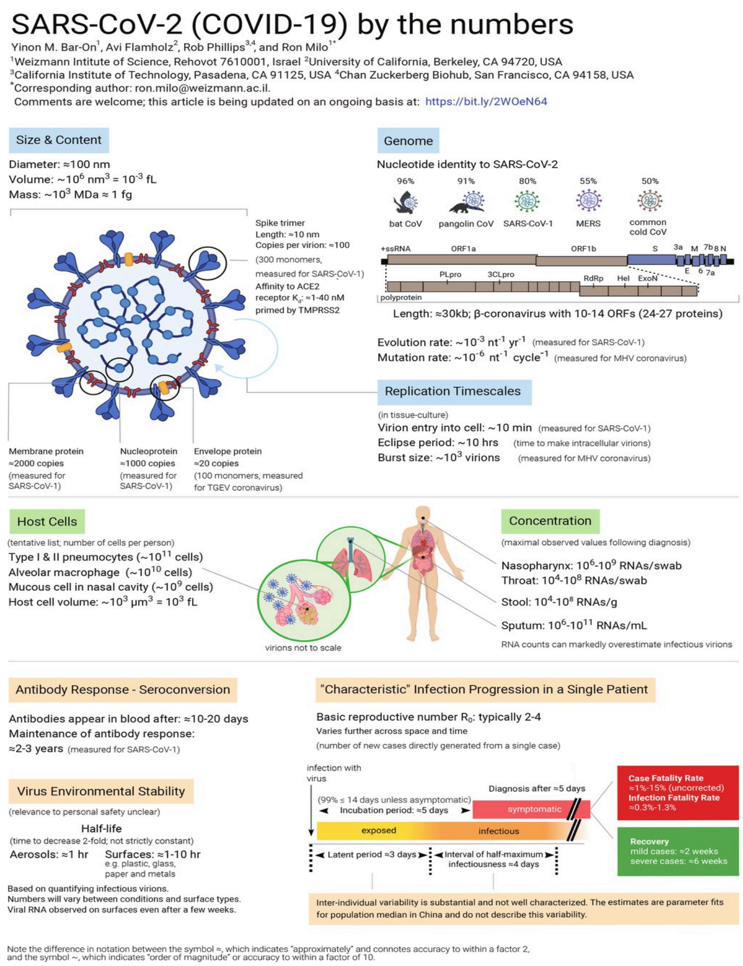
[(World Organisation for Animal Health, COVID-19 Portal, Events in animals)]

2. Human Viral Loads Based on Monkeys

Dr. Christian Yates's calculations use data from infected monkeys to estimate peak human viral loads. While animals provide valuable estimates, the University of Chicago Medical Center (2021) highlights that monkeys' immune responses and infection patterns can be significantly different from humans and apes, which potentially leads to misleading conclusions from Dr. Yates's calculations. Humans may clear infections at different rates than monkeys, or they may concentrate viruses in different tissues, affecting overall viral loads (University of Chicago Medical Center, 2021). If human viral loads were only half of those observed in monkeys, the calculation could be off by a factor of two or more. Likewise, if humans typically carry significantly higher viral concentrations than monkeys, the error would be even greater, which increases potential inaccuracies in the predictions. These discrepancies are critical and need for the careful selection of animal models when estimating human biological responses. While monkeys provide valuable insights, their immune systems are not the perfect model choice for humans; differences in antiviral responses, disease progression, and viral distribution must be accounted for in experiment and data interpretation (Haigwood & Walker). Dr. Christian Yates's estimation fails to consider these factors, which may lead to inaccurate conclusions about human infections.

3. Sphere Shape and Virus Particle Size

The coronavirus SARS-CoV-2 has been extensively studied since its emergence, but certain simplifications in its dimensions may lead to significant inaccuracies in research and modeling. While sphere-like virus particals are considered to be precisely 100 nanometers, the reality is considerably more complex. The assumption that each SARS-CoV-2 particle is a perfect sphere with a diameter of 100 nanometers oversimplifies reality. Virus particles vary slightly in size and shape, and actual measurements range between 80 and 120 nanometers in diameter (Bar-On, 2020), which significantly affects volume calculations. Given the cubic relationship between radius and volume, small changes in particle size yield large changes in volume. For example, increasing particle diameter from 100 to 120 nm results in a 1.73 increase in volume per particle. Conversely, reducing the diameter to 80 nm reduces the particle volume by about half. Additionally, real virus particles may be irregularly shaped or surrounded by protruding spike proteins that affect packing density and volume calculations (Yinon M Bar-On, Avi Flamholz, Rob Phillips, Ron Milo, 2020). The oversimplification of viral particles as perfect spheres fails to capture the biological reality and can lead to miscalculations.



(Yinon M Bar-On, Avi Flamholz, Rob Phillips ,Ron Milo, 2020)

4. Rate of Daily Infections

The calculation assumes a stable daily infection rate of 3 million individuals. However, this figure likely fluctuates significantly depending on public health measures, seasonal changes, population immunity, and the emergence of new variants. If infection rates vary widely, for example, doubling during a wave or dropping sharply due to increased vaccinations, the true number of active infections could differ substantially. Such variability would directly affect the accuracy of the calculation, causing either a significant overestimation or underestimation of the total virus particle count. For instance, if the true daily infections were double the assumed number during a peak wave, such as 6 million, the total particle count estimate would roughly double. Conversely, halving the daily infection rate would similarly halve the calculated particle count.

Suggested Improvements

Collect Direct Human Viral Load Samples: Instead of relying solely on animal models, directly measuring viral loads in infected humans provides a more accurate understanding of viral distribution. This approach could help refine estimates of average and peak viral loads, potentially reducing errors associated with extrapolating data from animals to humans.

Inclusion of Animal and Environmental Reservoirs: To improve global viral load estimates, it is essential to account for viral counts in animal reservoirs, such as mink farms and deer populations, as well as environmental sources like wastewater and contaminated surfaces. Systematic environmental surveillance can quantify virus persistence in non-human reservoirs, leading to a more comprehensive and realistic assessment of total SARS-CoV-2 particles worldwide.

References

1. Yates, C. (2021, February 9). *All the coronavirus in the world could fit inside a Coke can, with plenty of room to spare*. The Conversation. https://theconversation.com/all-the-coronavirus-in-the-world-could-fit-inside-a-coke-can-with-plenty-of-room-to-spare-154226

2. Sharun, K., Dhama, K., Pawde, A. M., Gortázar, C., Tiwari, R., Bonilla-Aldana, D. K., Rodriguez-Morales, A. J., de la Fuente, J., Michalak, I., & Attia, Y. A. (2021). SARS-CoV-2 in animals: Potential for unknown reservoir hosts and public health implications. Veterinary Quarterly, 41(1), 181-201. https://doi.org/10.1080/01652176.2021.1921311

3. Caserta, L. C., Martins, M., Butt, S. L., Hollingshead, N. A., Covaleda, L. M., Ahmed, S., Everts, M. R. R., Schuler, K. L., & Diel, D. G. (2023). White-tailed deer (Odocoileus virginianus) may serve as a wildlife reservoir for nearly extinct

SARS-CoV-2 variants of concern. Proceedings of the National Academy of Sciences of the United States of America, 120(6), e2215067120 https://pubmed.ncbi.nlm.nih.gov/36719912/

4. Dargahi, A., Jeddi, F., Vosoughi, M., Karami, C., Hadisi, A., Mokhtari, S. A., Ghobadi, H., Alighadri, M., Biparva Haghighi, S., & Sadeghi, H. (2021). Investigation of SARS CoV-2 virus on environmental surfaces. Environmental Research, 195, 110765. https://doi.org/10.1016/j.envres.2021.110765

5. University of Chicago Medical Center. (2021, March 29). *Apes show dramatically different early immune responses compared to monkeys*. ScienceDaily. https://www.sciencedaily.com/releases/2021/03/210329122858.htm

6. Haigwood, N. L., & Walker, C. M. (n.d.). *Comparison of immunity to pathogens in humans, chimpanzees, and macaques*. National Center for Biotechnology Information (NCBI). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK91442/

7. Bar-On, Y. M., Flamholz, A., Phillips, R., & Milo, R. (2020). SARS-CoV-2 (COVID-19) by the numbers. *9*, e57309 https://doi.org/10.7554/eLife.57309