Frugal computing

On the need for low carbon and sustainable computing

Euro-CSDMS, 29 April 2024

Wim Vanderbauwhede

University of Glasgow

School of Computing Science

Global context

IPCC Sixth Assessment Reports:

- Greenhouse gas emissions must be cut drastically to keep global warming below 1.5°C (IPCC WGI AR6 report).
- They must be cut *now*, we can't afford to wait anymore (IPCC WGI, "Physical Science Basis", AR6 report).
- This is incompatible with unlimited economic growth (IPCC WG III, "Mitigation", AR6 report).

The need for frugal computing

Computational resources are finite

- Since the 1970s, we have been using increasing amounts of computational resources.
- Growth of performance per Watt has been exponential (Koomey's law)
- As a result, increasing use of computational resources has become pervasive in today's society.
- Computational resources have until recently effectively been treated as infinite.

Computational resources are finite

- Growth in demand cannot be offset by increased power efficiency.
  - Moore and Koomey's laws can’t save us (“*the free lunch is over*”).
  - With business as usual, the carbon footprint from the use of computing will become a major contributor to the world total.
- The carbon footprint of device production is also huge.
  - Moore's Law has led to very short lifetimes of compute hardware.
  - The current rate of obsolescence is entirely unsustainable.

We need *frugal* computing

- As a society, we need to start treating computational resources as finite and precious, to be utilised only when necessary, and as frugally as possible.
- Computing scientists, developers and engineers need to ensure that computing has the lowest possible energy consumption.
- And because the lifetimes of compute devices needs to be extended dramatically, this must be achieved with the currently available technologies.

The scale of the challenge

Meeting the climate targets

Limits and current emissions

- To limit global warming to below 2°C by 2040, a global reduction from 55 to 13 gigatonnes CO_2 equivalent per year ($GtCO_2e/y$) is needed.
- That means we need to cut 5%-7% per year.
- At present, emissions are **rising** between 1%-2% a year.

Limits and current emissions

- Emissions from electricity are currently about 10 $GtCO_2e$.
- But electricity consumption is rising steeply.
- And most electricity is still generated by burning fossil fuel.

Renewables and nuclear won’t save us

- Deployment is too slow.
- It takes 20 years to build a new nuclear power plant.
- Old ones are being shut down.
- Renewables+nuclear will provide only 30% of electricity by 2040.

Nor will Carbon Capture & Storage

- The energy required in the capture process will be greater than the energy made available during the release of the CO_2.
- Many scenarios assume that large areas of land will be available -- not clear if this is realistic, scalable or compatible with sustainability goals.
- There are poorly quantified risks of re-release and no credible standards or compliance procedures.
- From an ethical and intergenerational justice perspective … it looks like green wash.

(taken from the UK Research and Innovation DRI Net Zero project presentation, M. Juckes, 2022)

Nor Carbon offsetting

- The earth's land ecosystems can hold enough additional vegetation to absorb 40 - 100 $GtCO_2e$ from the atmosphere.
- Once this additional growth is achieved (takes decades), there is no capacity for additional carbon storage on land.
- The world emits 50 $GtCO_2e$ into the atmosphere per year. So all we can offset is 2 year's emissions at most.

(B. Waring, "There aren't enough trees in the world to offset society's carbon emissions – and there never will be", The Conversation 2021-04-23)

Reducing emissions is imperative

- In other words, to reduce atmospheric CO_2 to 2°C levels by 2040, the only way is
  - to reduce energy consumption
  - to reduce the amount of goods produced
- This is largely an economic problem, but there is an important role to play for technology

The carbon cost of computing

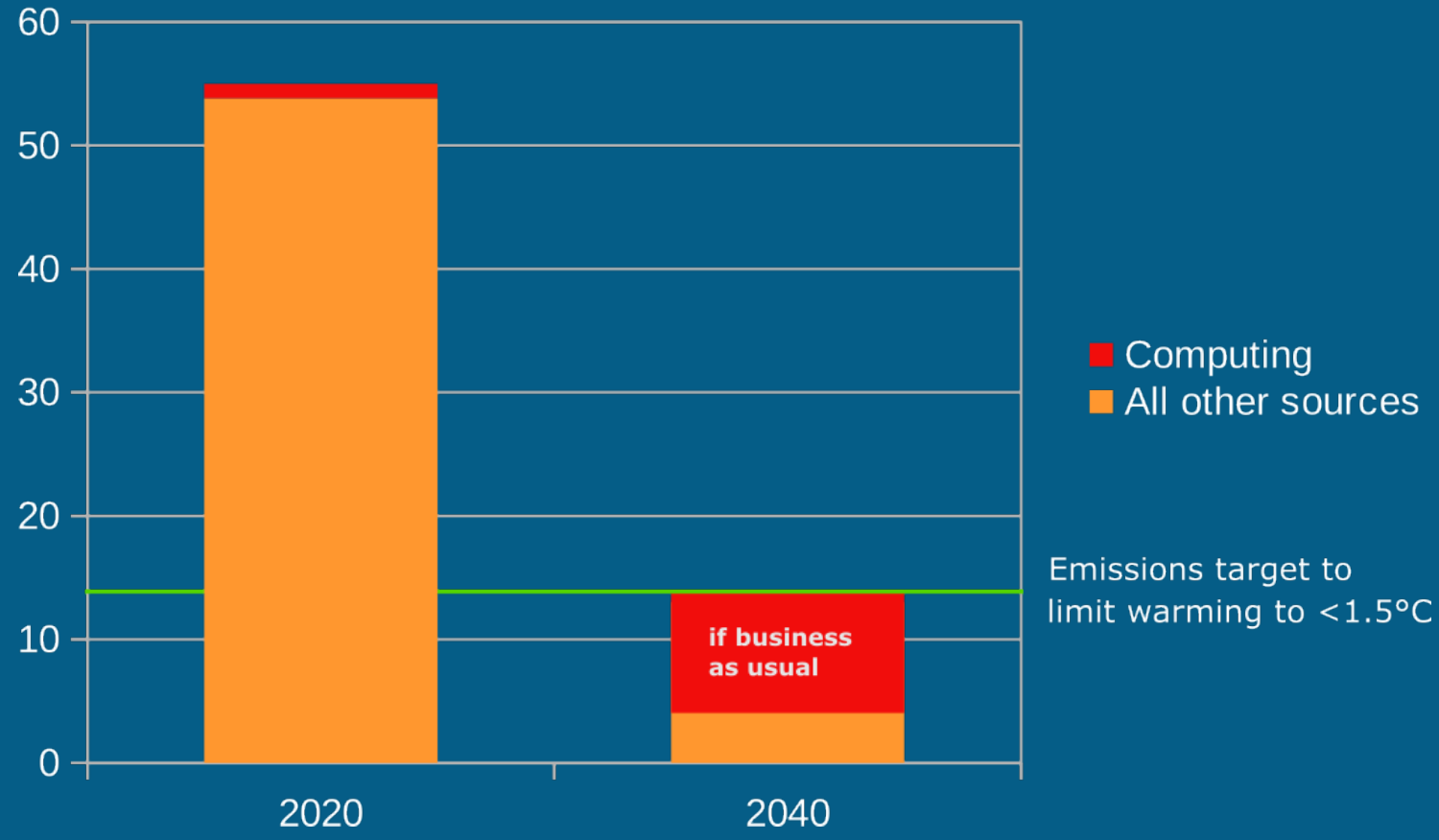
Emissions from using computing

- In 2020, emissions from using computing were between 3.0% and 3.5% of the total.
- This is already more than the airline industry
- By 2040 this will grow to 14% (4x).
- By 2040, energy consumption of compute devices would be responsible for 5 gigatonnes of CO_2

Emissions from production

- Emissions from the production of computing devices exceed those incurred during operation.
- Taking into account this carbon cost of production, computing would be responsible for 10 Gtonnes of CO_2 by 2040,
- This is almost 80% of the acceptable CO_2 emissions budget of 13 gigatonnes of CO_2.

Emissions from computing and other sources



Emission growth drivers

- High-Definition Video/VR/AR
  - Both growth in demand and in resolution
  - VR/AR encode 3-D so even higher BW needed
- Internet of Things (IoT)
  - “Internet” means every small device has a huge network+cloud footprint
  - But manufacturing of devices too has large footprint

Emission growth drivers

- “AI”, in particular Generative AI
  - Queries to e.g. ChatGPT consume 60x more energy than conventional search
  - Due to hype, growth is very steep
  - Most governments have bought into this hype
  - “AI” everywhere would lead to huge increase in emissions

Emission growth drivers

- Mobile devices
  - Still high growth in demand
  - Main driver is short replacement cycles
  - As with IoT, in practice mobile device has a huge network+cloud footprint
  - And manufacturing of devices too has large footprint
  - Network infrastructure (5G, 6G, …) growth

Achieving frugal computing

Transforming computing

- We need a radical transformation of the global use of computational resources to meet the climate targets.
- We must dramatically reduce the carbon cost of both production and operation of computing.
- We must start now, there is no time to wait

Extending the useful life

- We can't rely on next-generation hardware technologies to save energy: the production of this next generation of devices will create more emissions than any operational gains can offset.
- This does not mean research into more efficient technologies should stop.
- But their deployment cycles should be much slower.
- Extending the useful life of compute technologies to several decades must become our priority.

Extending the useful life

- We need a change in business models as well as consumer attitudes. This requires:
  - raising awareness and education;
  - providing incentives for behavioural change;
  - provide economic incentives and policies;
  - infrastructure and training for repair and maintenance.

What can we do?

Don’t believe the hype

- Renewables, nuclear, CCS and offsetting won’t save us
- Reducing electricity usage is the only feasible solution
- Reducing emissions from manufacturing is equally important
- What all this means is that growth in production and use of computing resources must slow down drastically.

Specifically on ICT for research

- Net zero computing: scoping UKRI's journey to sustainable digital research infrastructure https://zenodo.org/records/8203117
- Carbon Lifecycle Analysis for Scientific Computing

  https://indico.cern.ch/event/1377701/contributions/5894041/attachments/2839432/4962697/20240418-lca.pdf

Thank you!

For more details and references

https://limited.systems/frugal-computing/

Spread the word

- Without awareness, there is no action.
- So awareness raising is key.
- We have power to influence
  - Scientists have an important example function
  - Companies sell what we buy
  - Politicians need our votes
  - Many changes started at grassroots level

Adapt your habits

- The drivers for emissions in computing are mainly video/VR/AR, AI, IoT and mobile devices
- So your personal actions should focus on these areas
- Get as few devices as possible
- Use them for as long as possible
- Minimise your internet bandwidth usage
- Think critically about your use of AI