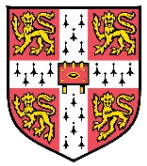


UNIVERSITY OF CAMBRIDGE

Isaac Newton Trust

LEVERHULME TRUST

Can ISPs become CSPs?

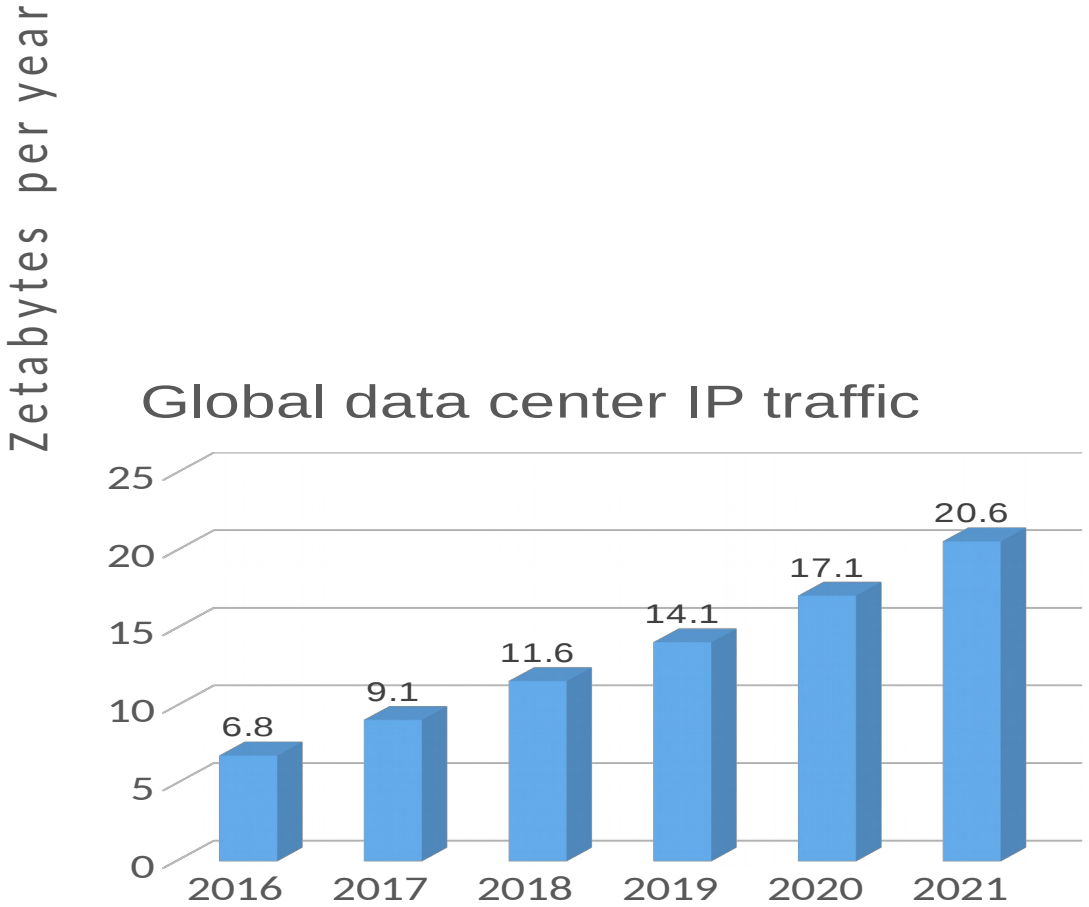
A vision of Computing Service Providers

Dr Noa Zilberman

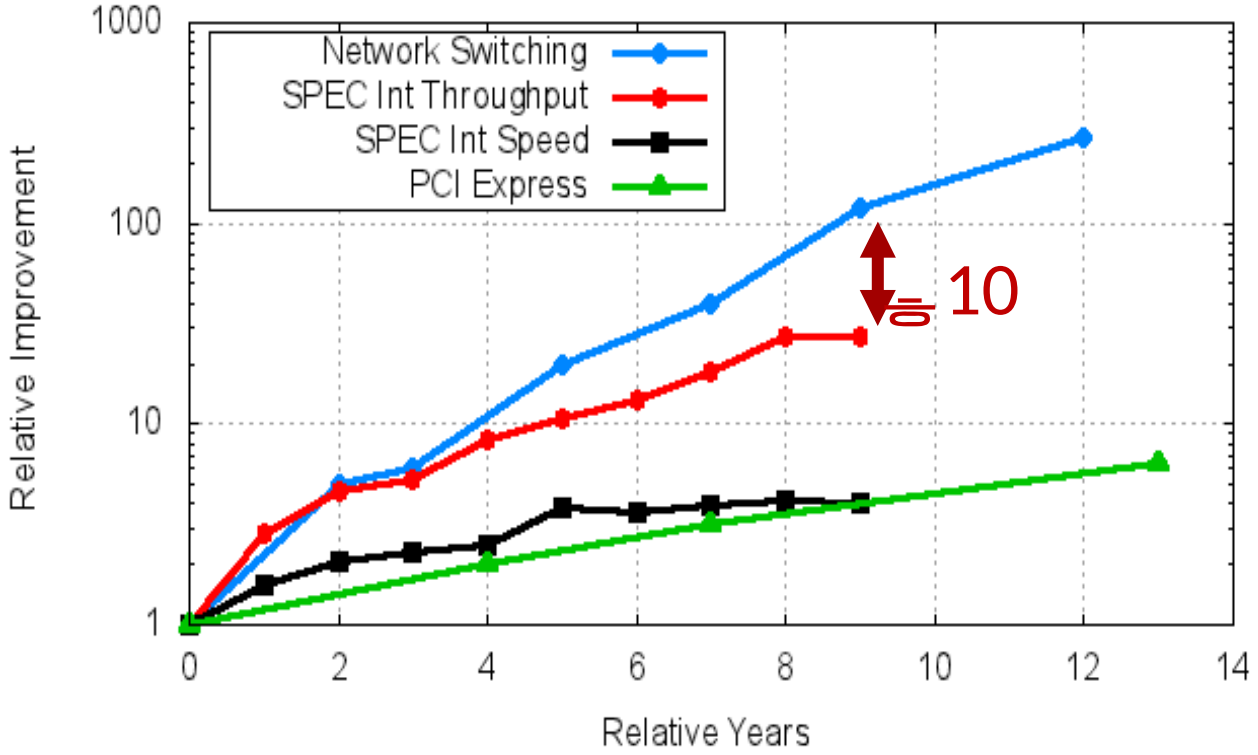
University of Cambridge

March 2019

The performance gap between networking and computing grows



Source: Cisco global cloud index, 2016-2021

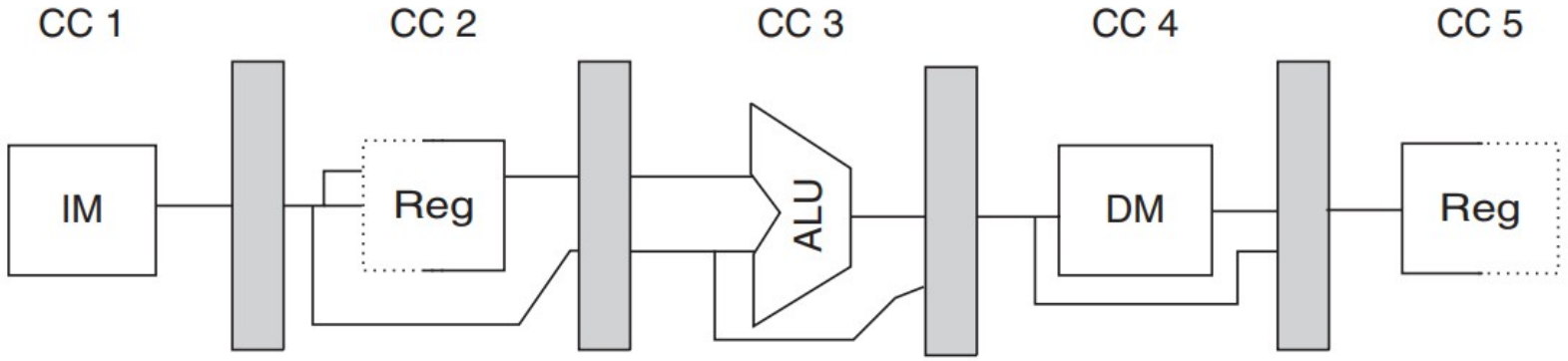


Zilberman et al., "From photons to big data applications: Terminating terabits", RSTA 2016

CPU vs switch architecture

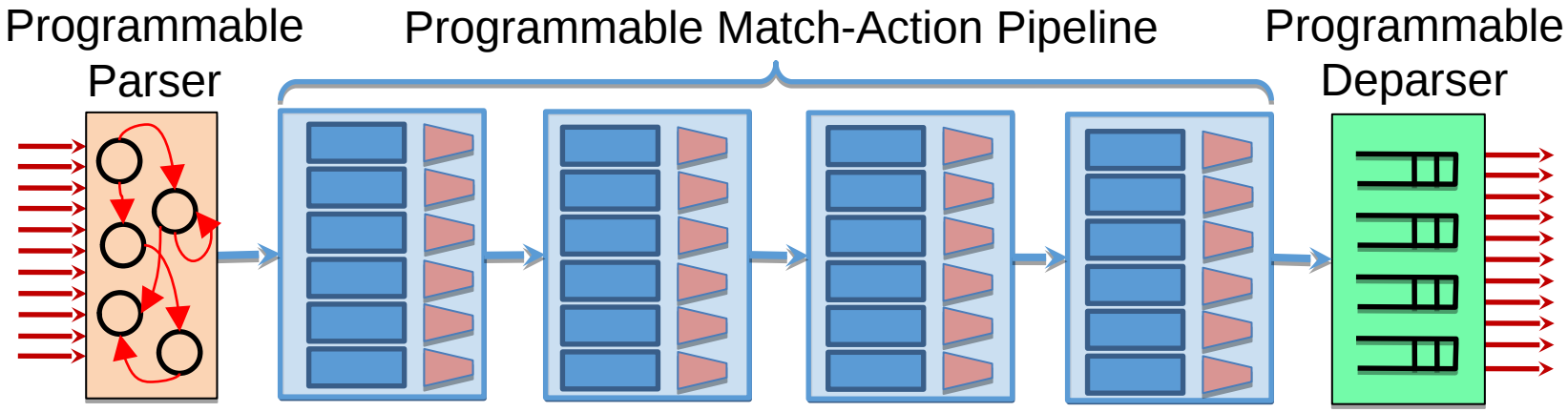
CPU

Moving **instructions**
Bus width 64bit



Switch

Moving **data**
Bus width ≥256Byte

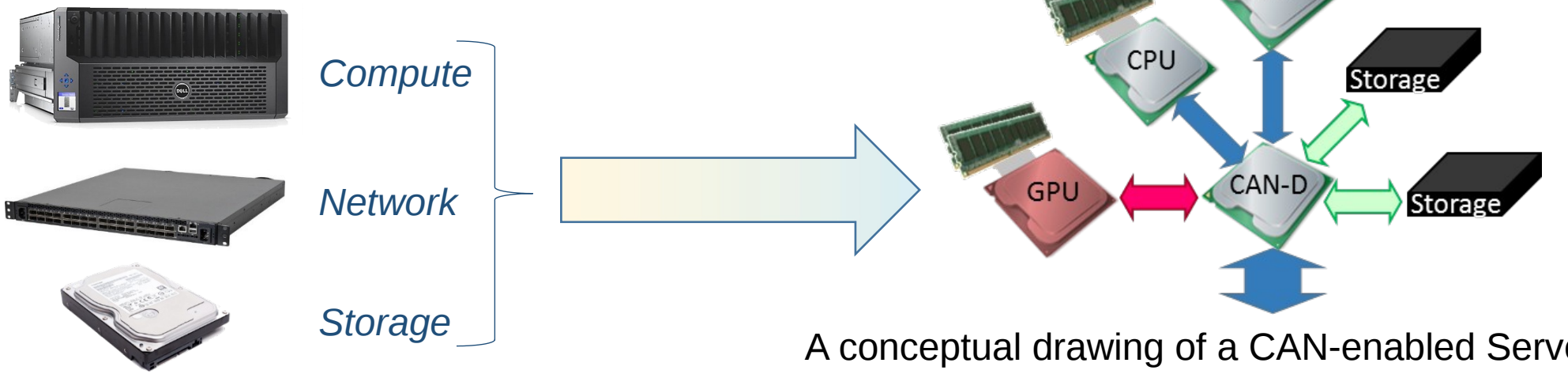


Sources: Hennessey & Patterson - “Computer Architecture, A Quantitative Approach”, p4.org

CAN: Computing As Network

Create a “Tiny Terabit Data Centre”

- “Pull” all data centre components into a single computer
- A network-fabric device at the core
- Other components are “peripherals”
- Optimise for high data rates



A conceptual drawing of a CAN-enabled Server

Zilberman et al., “From photons to big data applications: Terminating terabits”, RSTA 2016

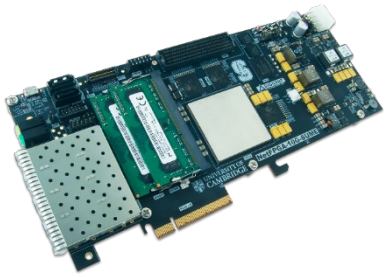
In-network computing

In the context of *this* talk:

In-network computing:

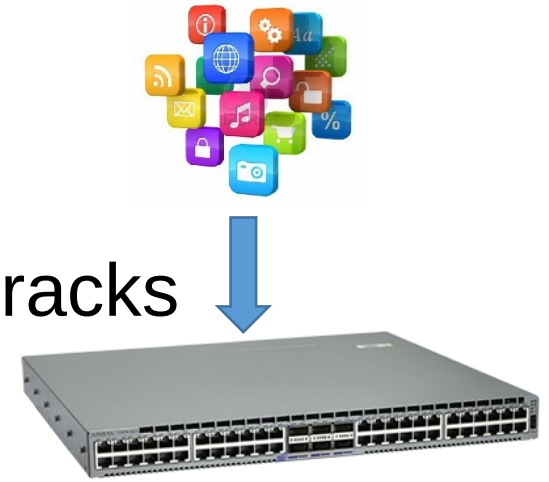
The execution of native host applications within the network using standard network devices:

network interface cards (NIC), switches



In-network computing: potential benefits

- Every op done in the network frees CPU cycles
- Superior application throughput
- Significant latency reduction.
- A single network switch can “replace” multiple server-racks
  - saving ~1GWh/year.
- The cost is almost zero:
  - The devices are already within the network!



We are not there yet

Tokusahi et al., “The case for in-network computing on demand”, Eurosys 2019

Computing as an Infrastructure

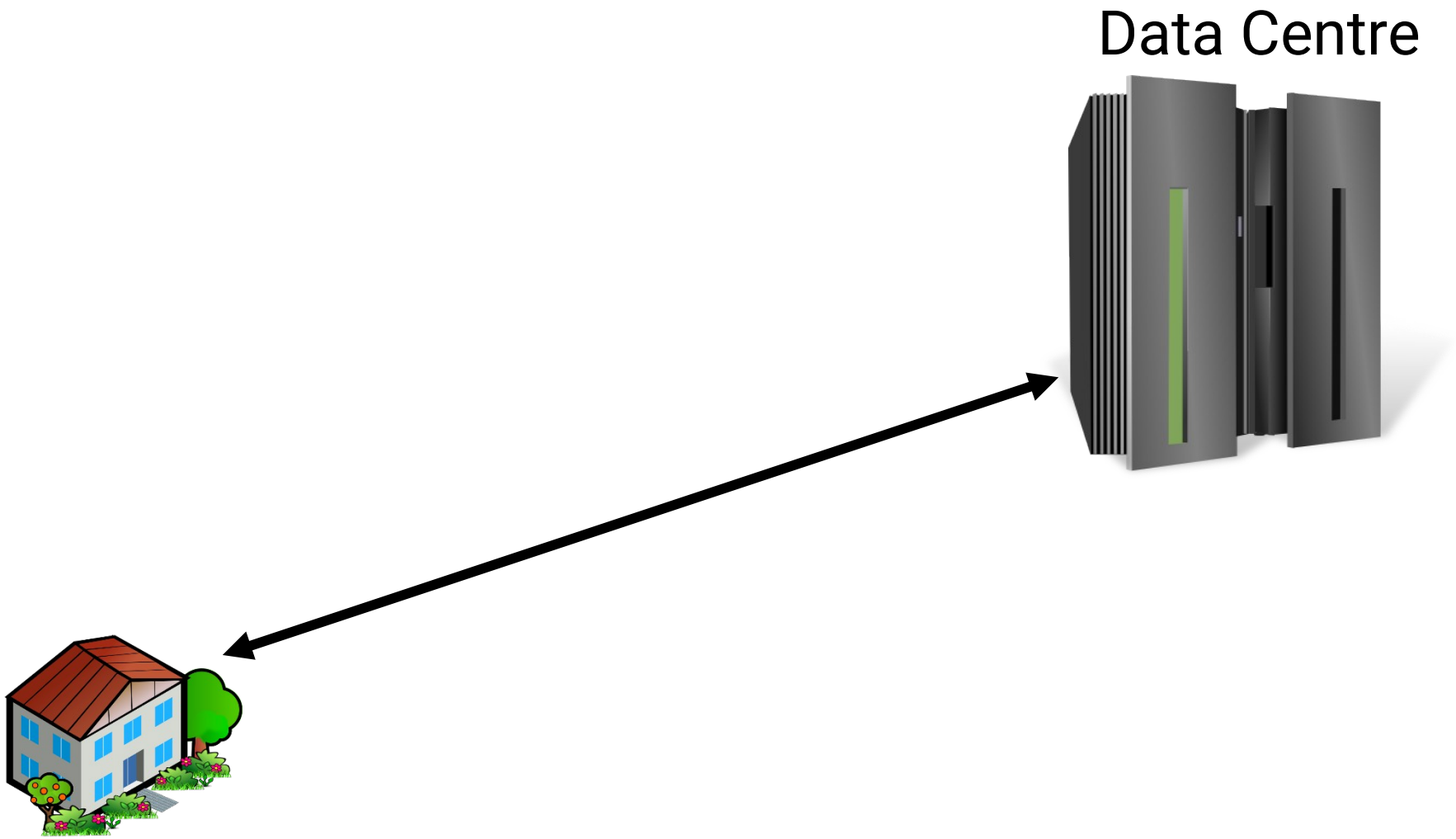
Infrastructure

Infrastructure is deployed:

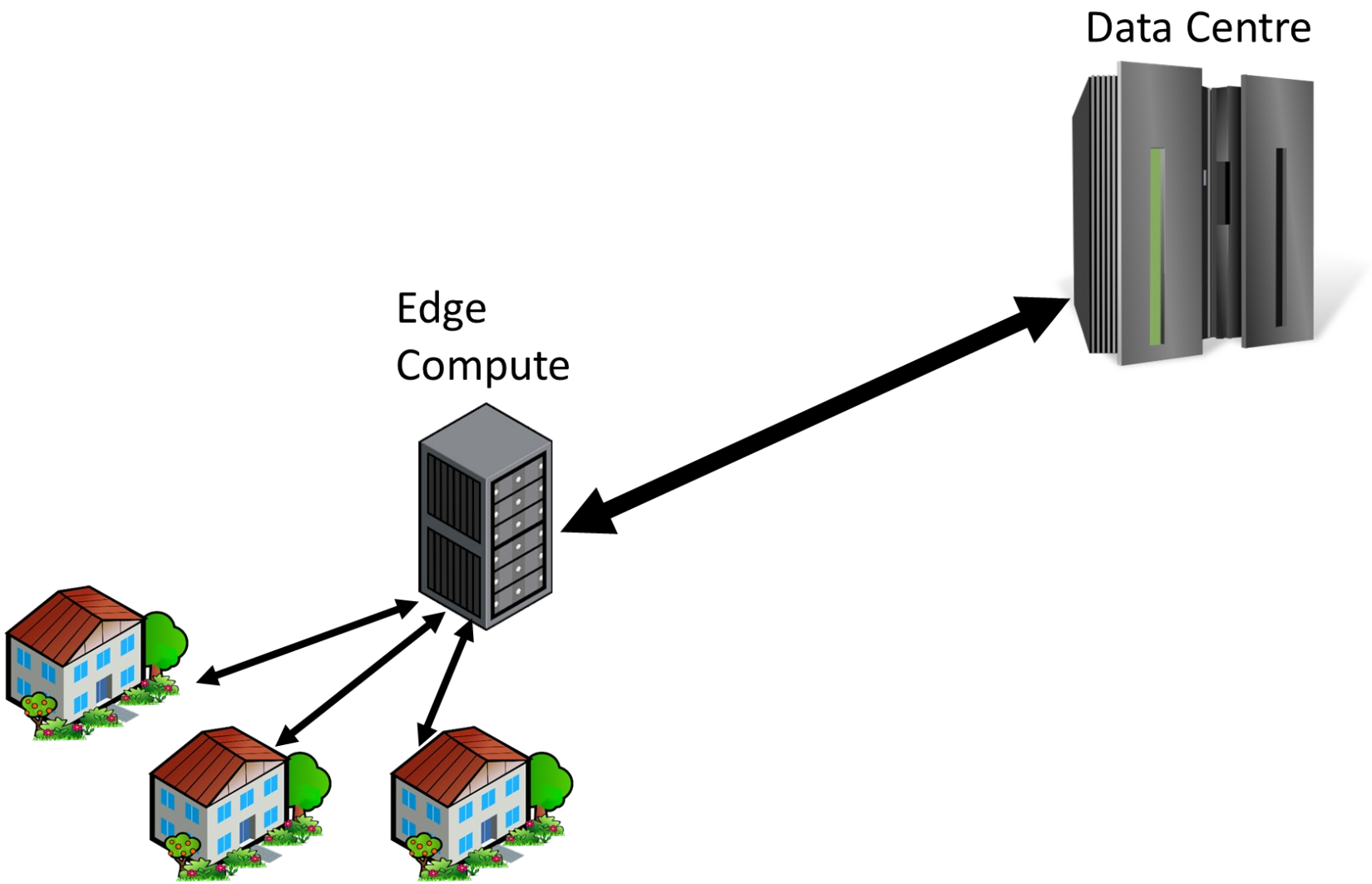
- At varying scales
- For varying needs

Computing becomes akin to Infrastructure

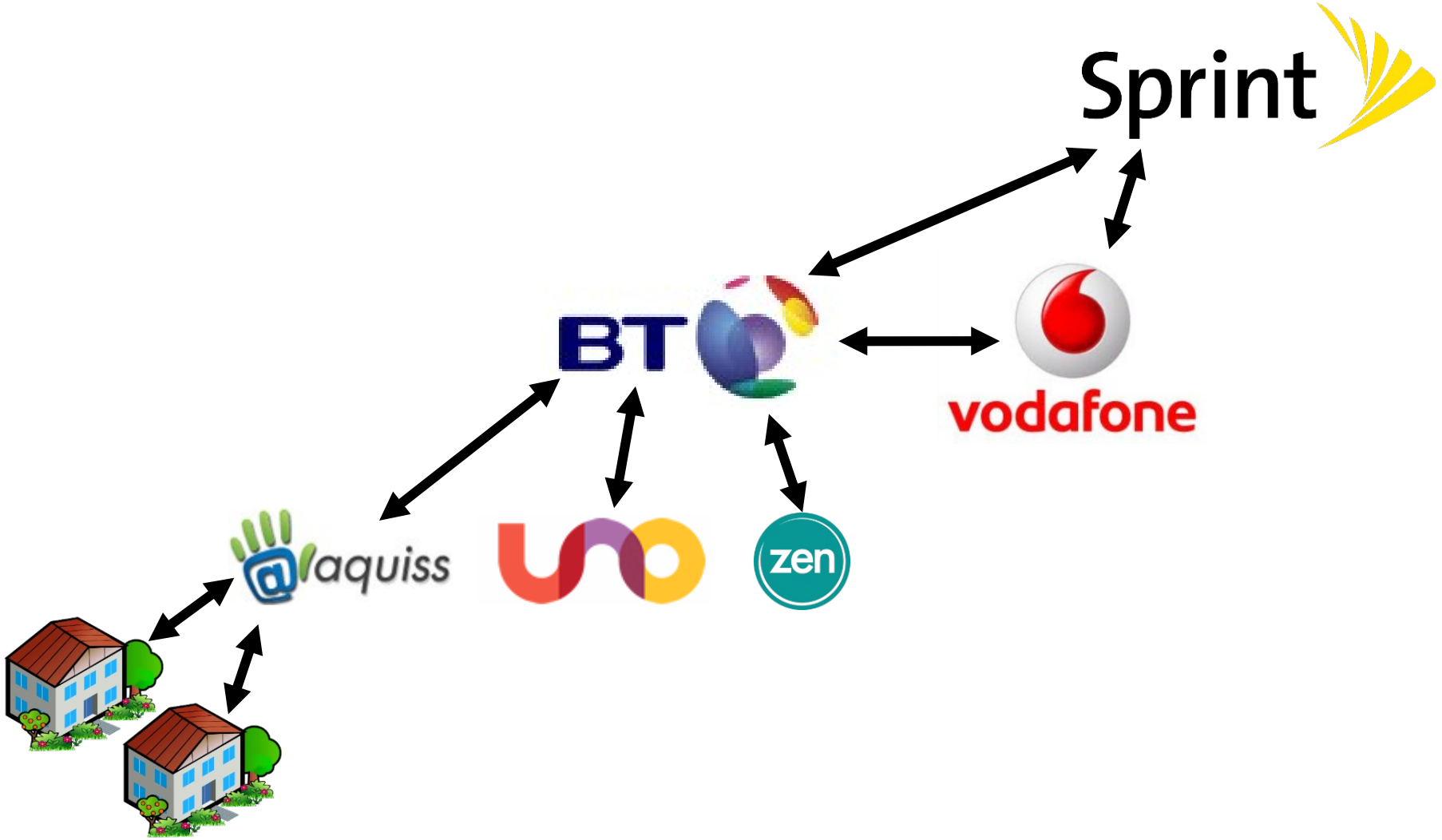
Computing infrastructure today



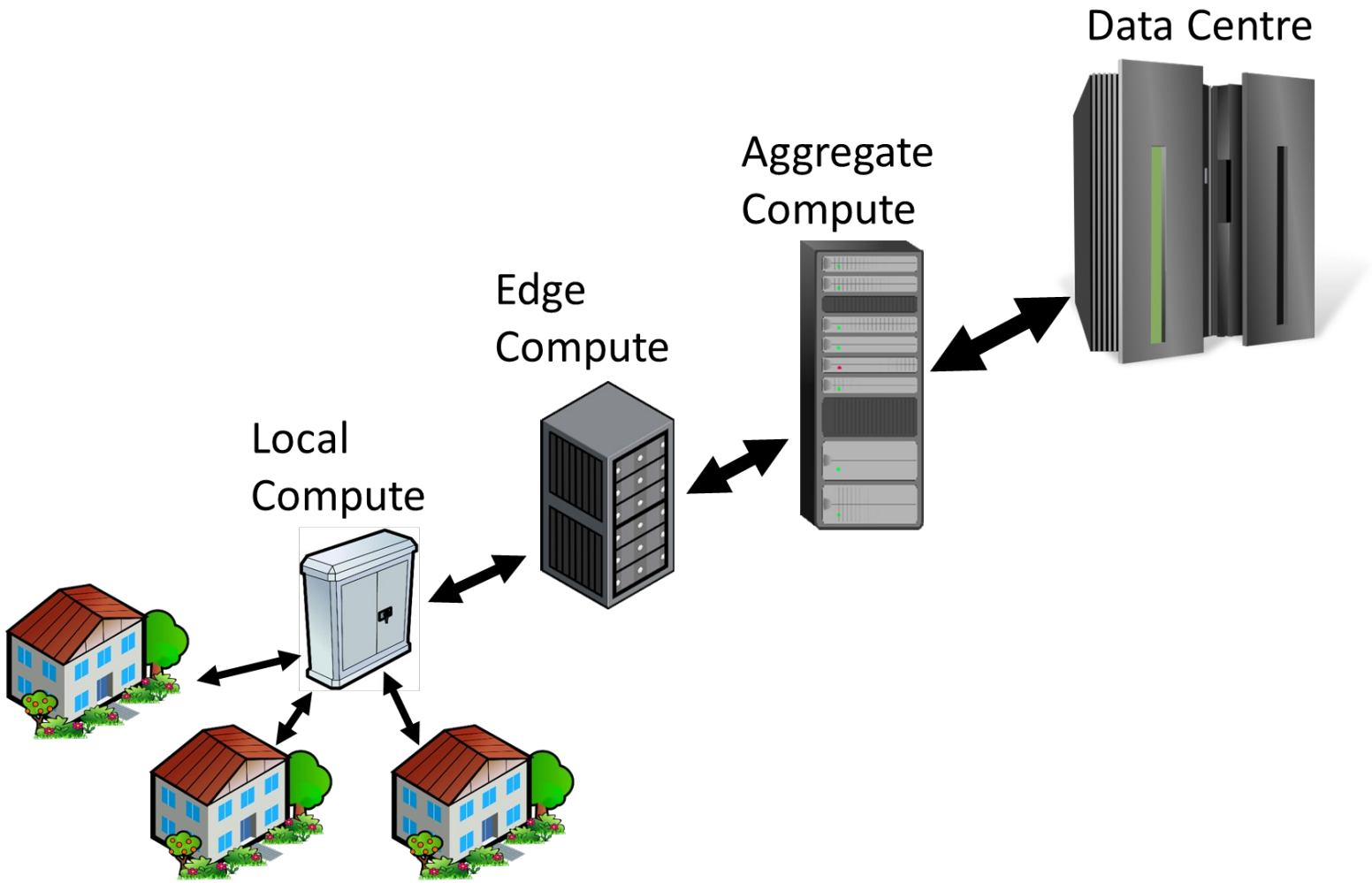
Computing infrastructure - emerging



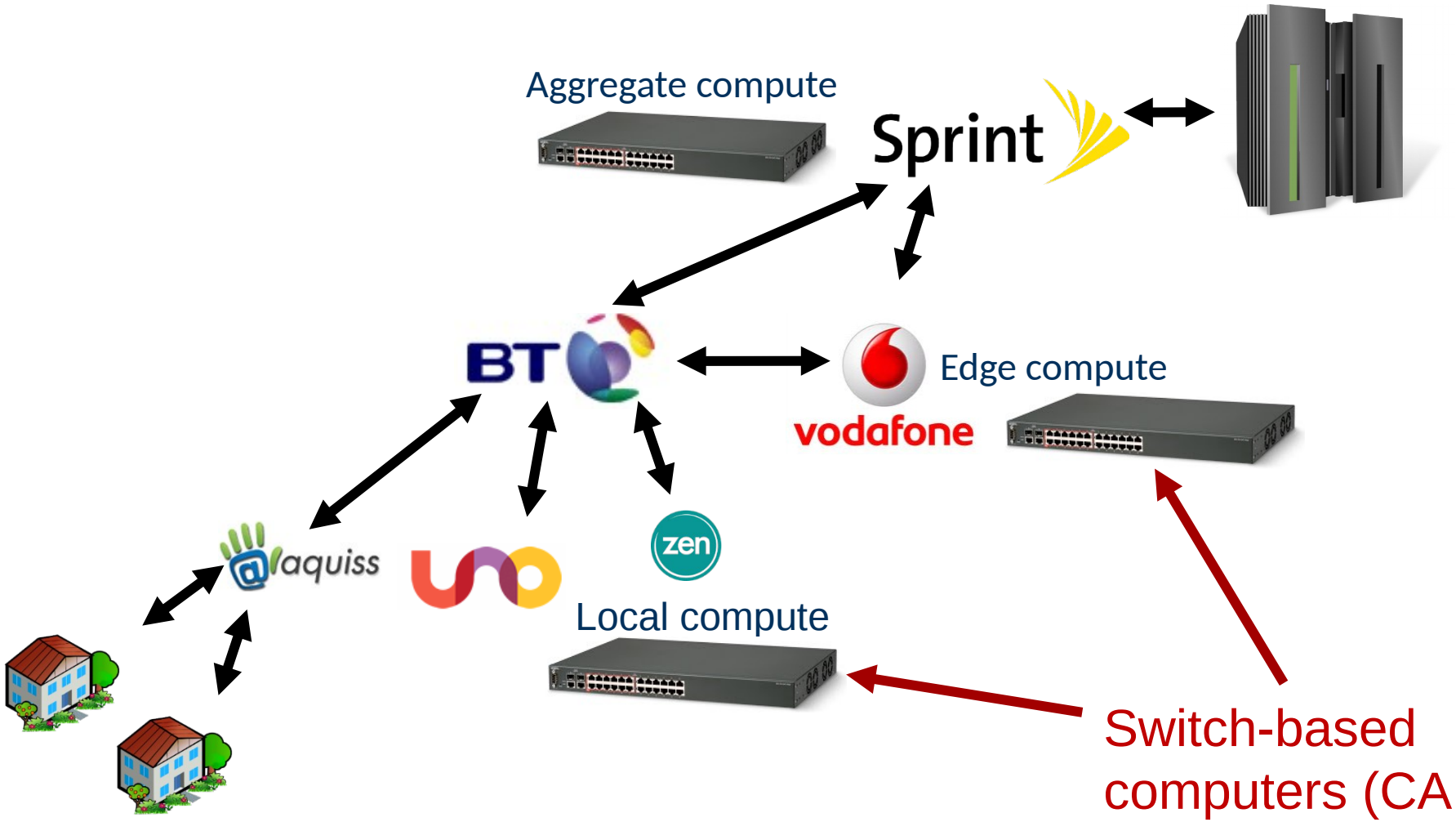
Communication infrastructure



Scaling computing infrastructure

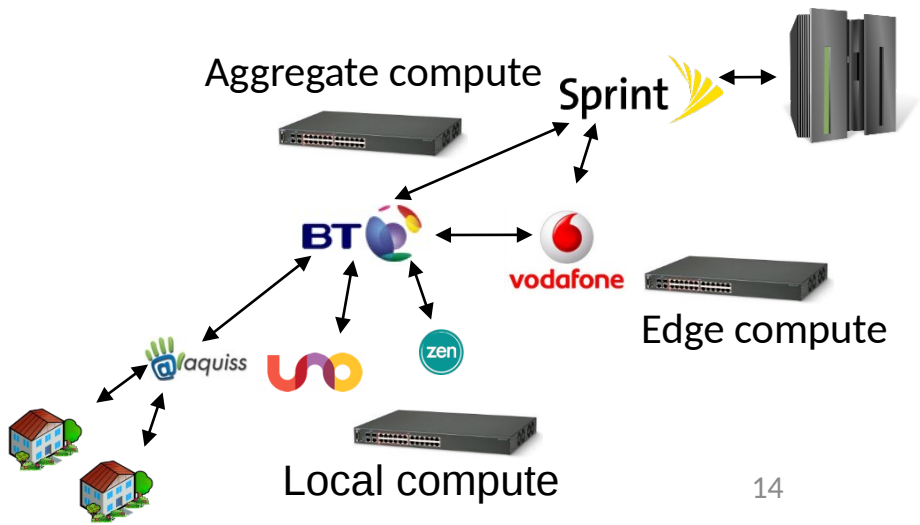


Unified communication and computing infrastructure



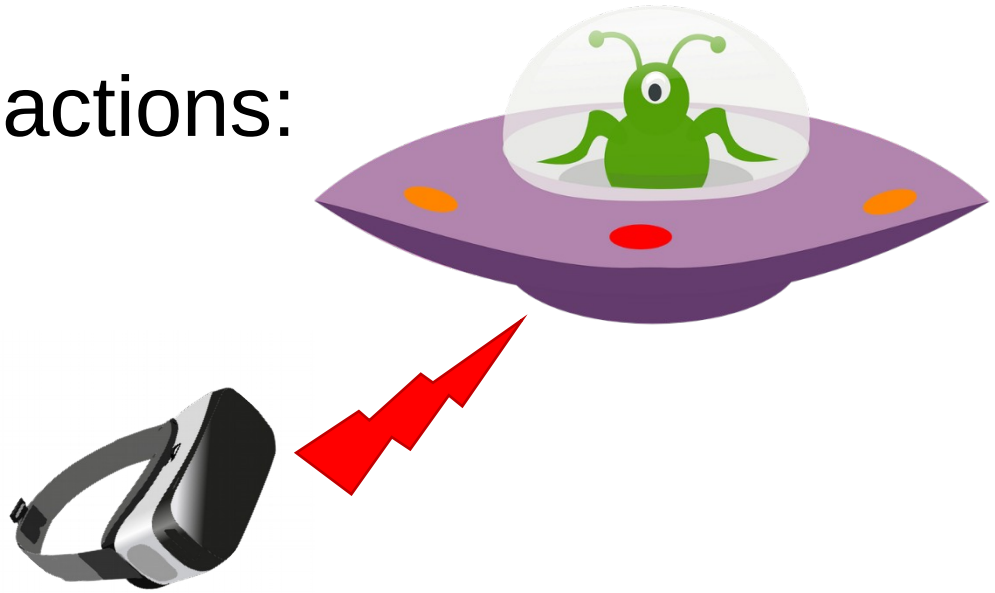
Unified communication and computing infrastructure

- Terminate data at the edge
  - Before it gets to the data centre
- Reduce the complexity of every stage
- Increase scalability and longevity
- Reduce power consumption
- Improve privacy and data control

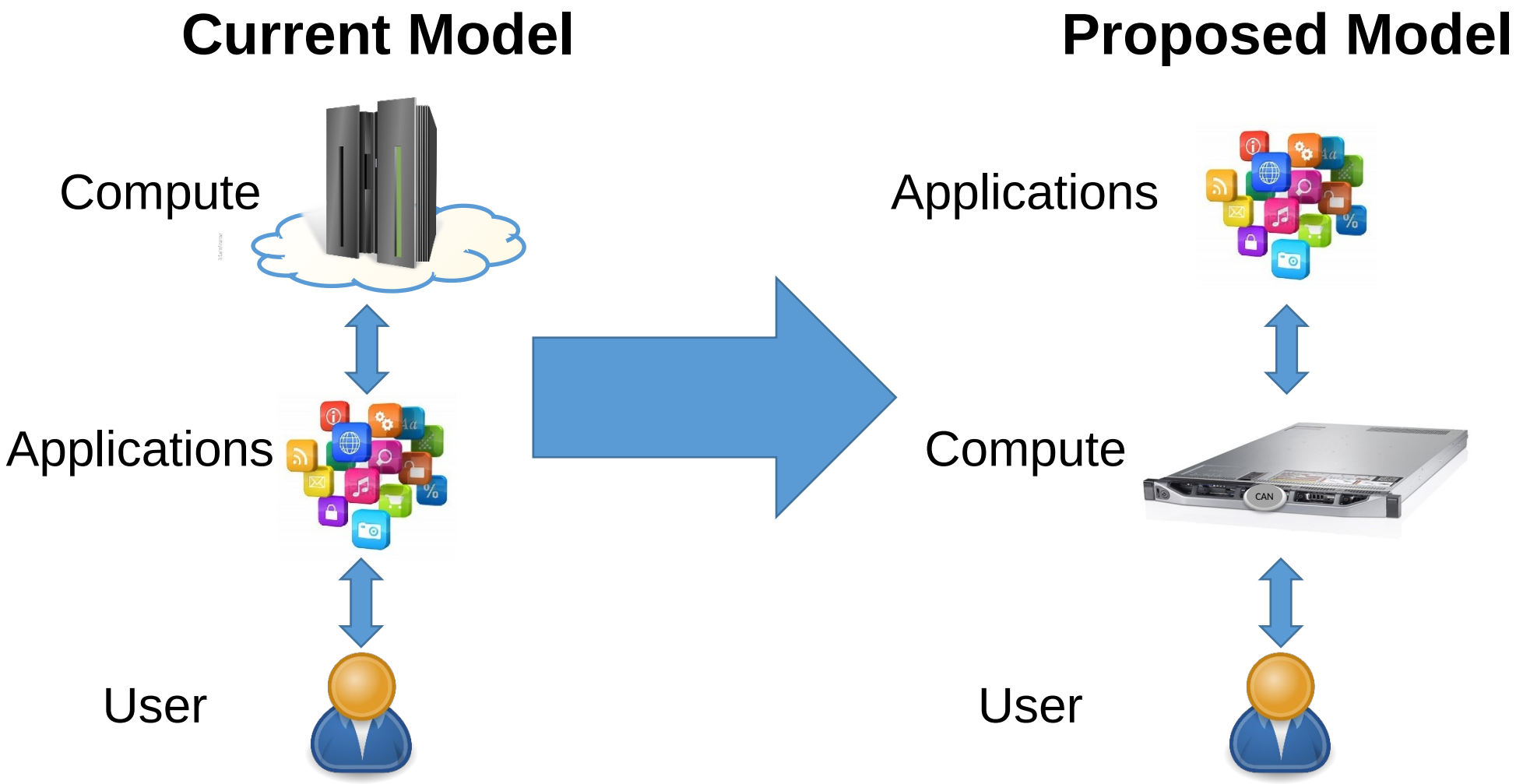


Scaling Computing Infrastructure – Cloud Providers

- Cloud providers have incentives to move computing closer to users:
  - Reduce the load on data centre infrastructure
  - Many applications require quick reactions:
    - Online gaming
    - Face recognition
    - Hand-writing recognition
    - Better smart-assistant experience

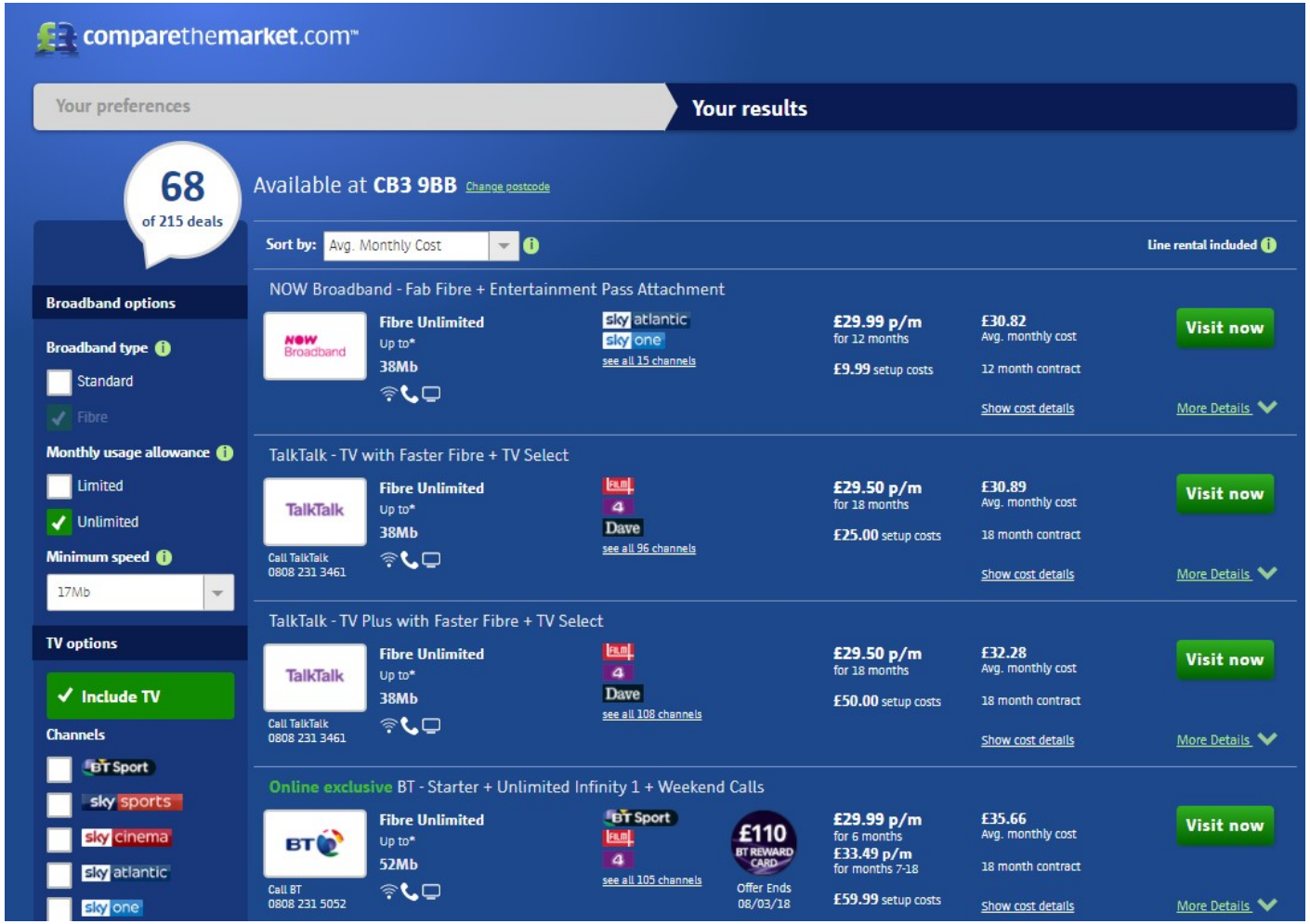


Rethinking application deployment model

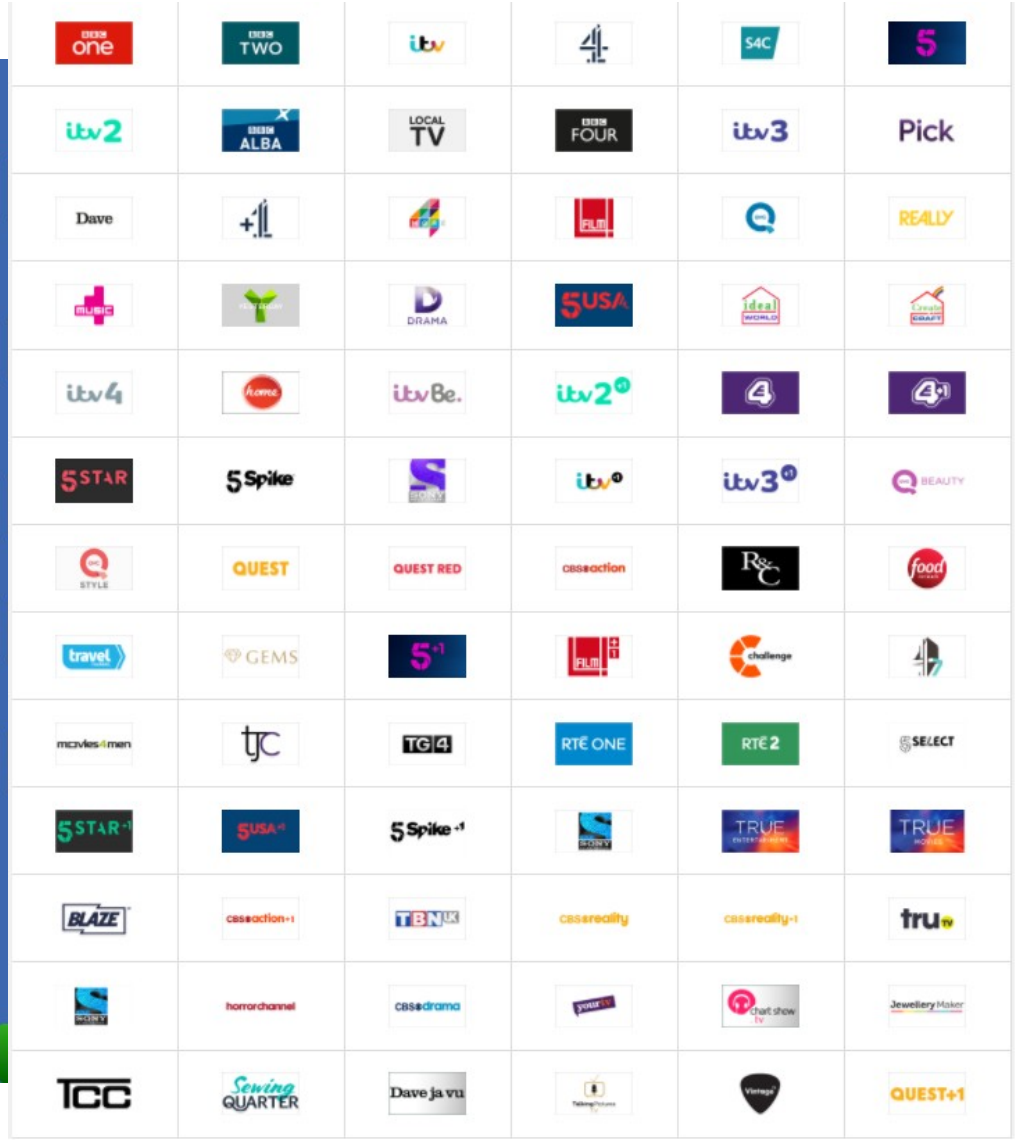


Zilberman, "Revolutionising Computing Infrastructure For Citizen Empowerment"*, Data for policy 2017*

Give us the right to choose our provider!



Give us the right to choose our apps pack!



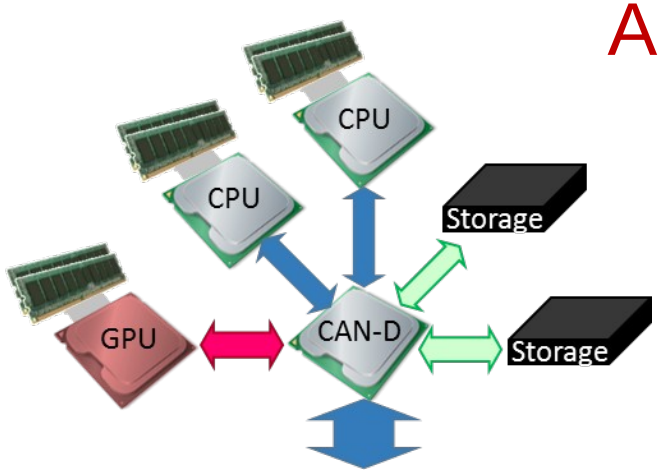
Taking back control of data

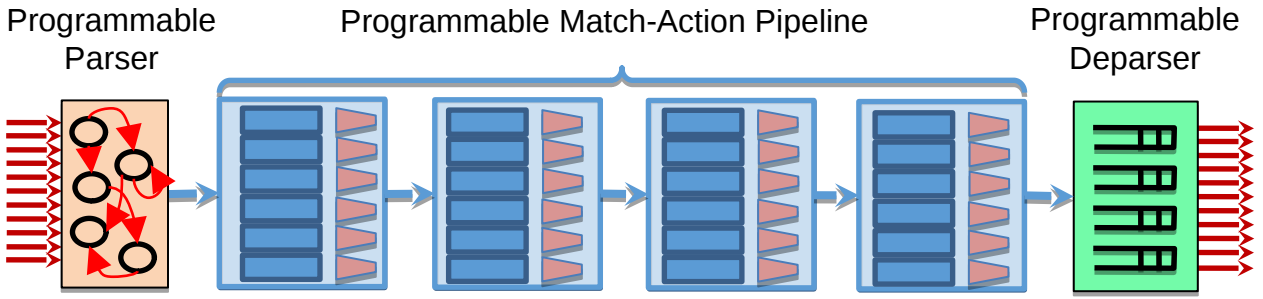
- **You** choose your compute service provider
  - Like choosing a television service provider
- Privacy and data control as a service
- Competitive market
- Improved resilience

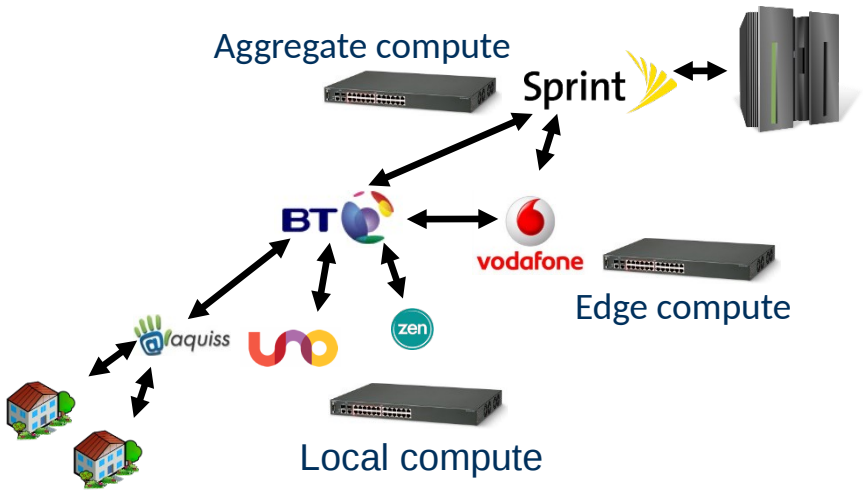
Thinking Ahead

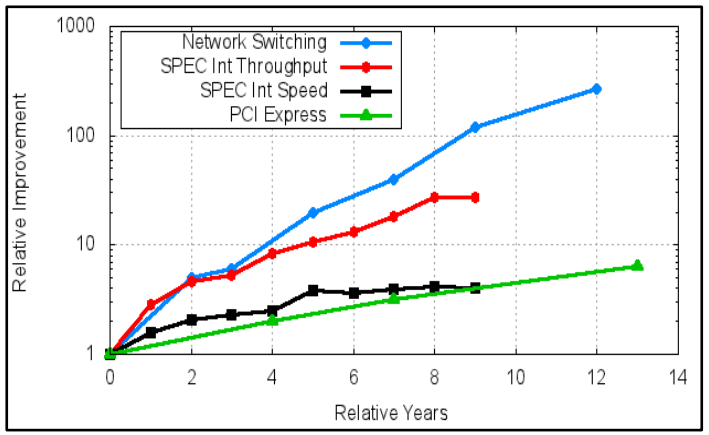
- Can new types of computing entities emerge?
  - E.g. neighbourhood clouds
  - Can be led by community or non-profit organisations

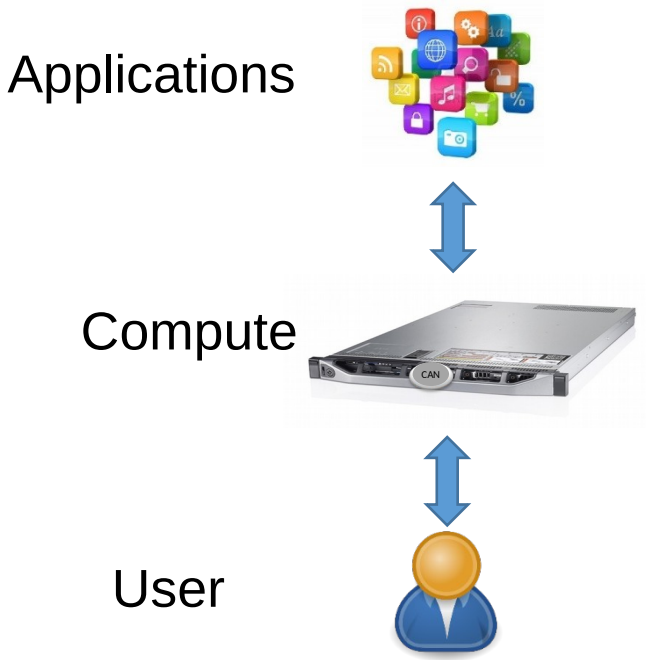
All papers are available at: www.cl.cam.ac.uk/~nz247











Questions?