Computing for the Future of the Planet

Andy Hopper

Computing is a <u>crucial</u> weapon in our armoury for ensuring the future of the planet.

Computing will play a key part in <u>optimising use of physical resources</u> and ultimately their <u>substitution by the digital world</u>.

Computing will be a tool for enabling <u>developing societies to improve their standard of living</u> without undue impact on the environment.

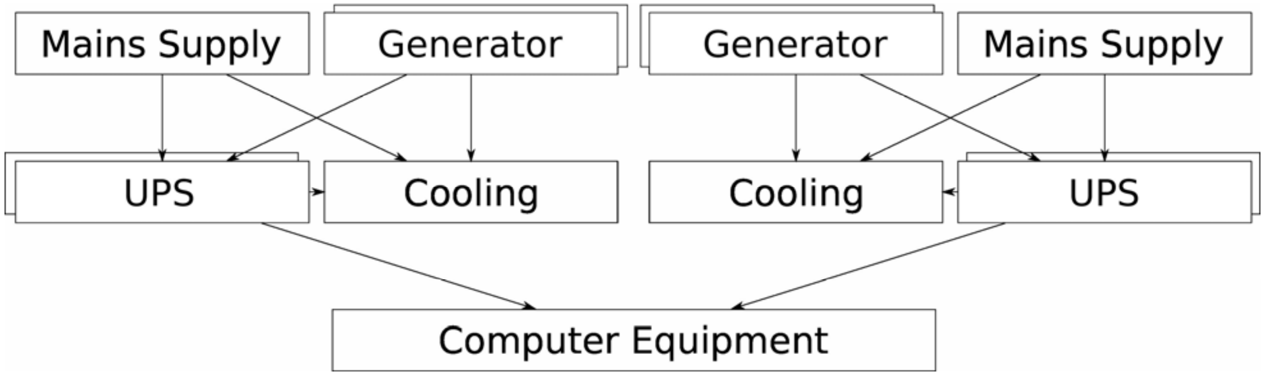
Computer Laboratory
University of Cambridge

A. Hopper and A. Rice, “Computing for the Future of the Planet”, *Phil. Trans. R. Soc. A*, Oct 2008.

Computing for the Future of the Planet

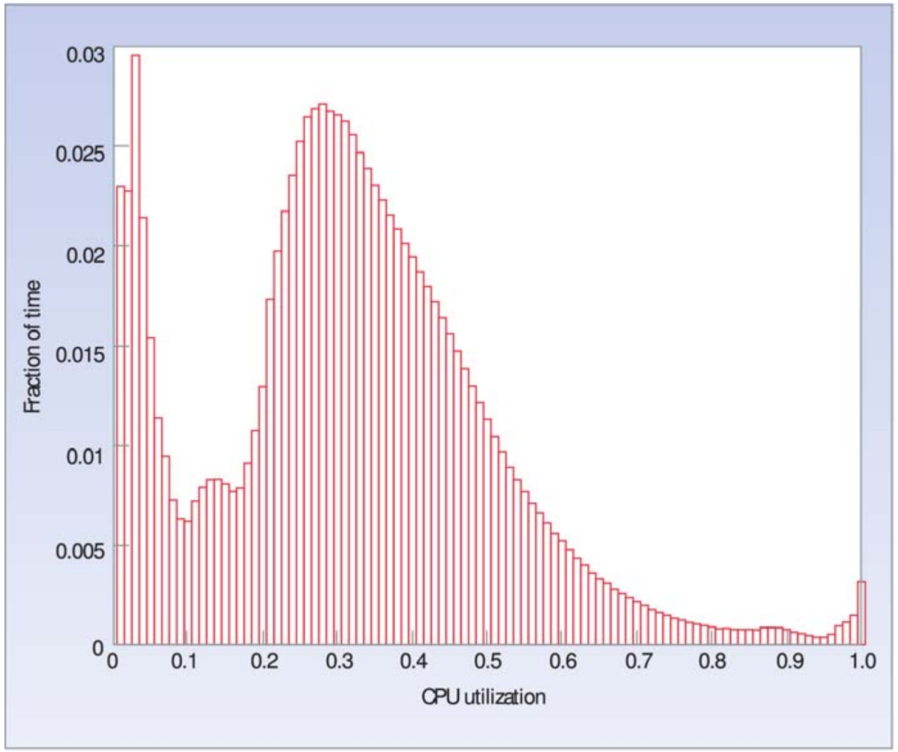
1. Optimal Digital Infrastructure
2. Sense and Optimise
3. Predict and React
4. Digital Alternatives to Physical Activities

1 - Optimal Digital Infrastructure



- Redundancy doubles (energy) cost of datacenter
- Reduce redundancy by using new tools which monitor interdependency of components
- Decrease restart times of services (from ~4 hrs)

Utilisation of Servers



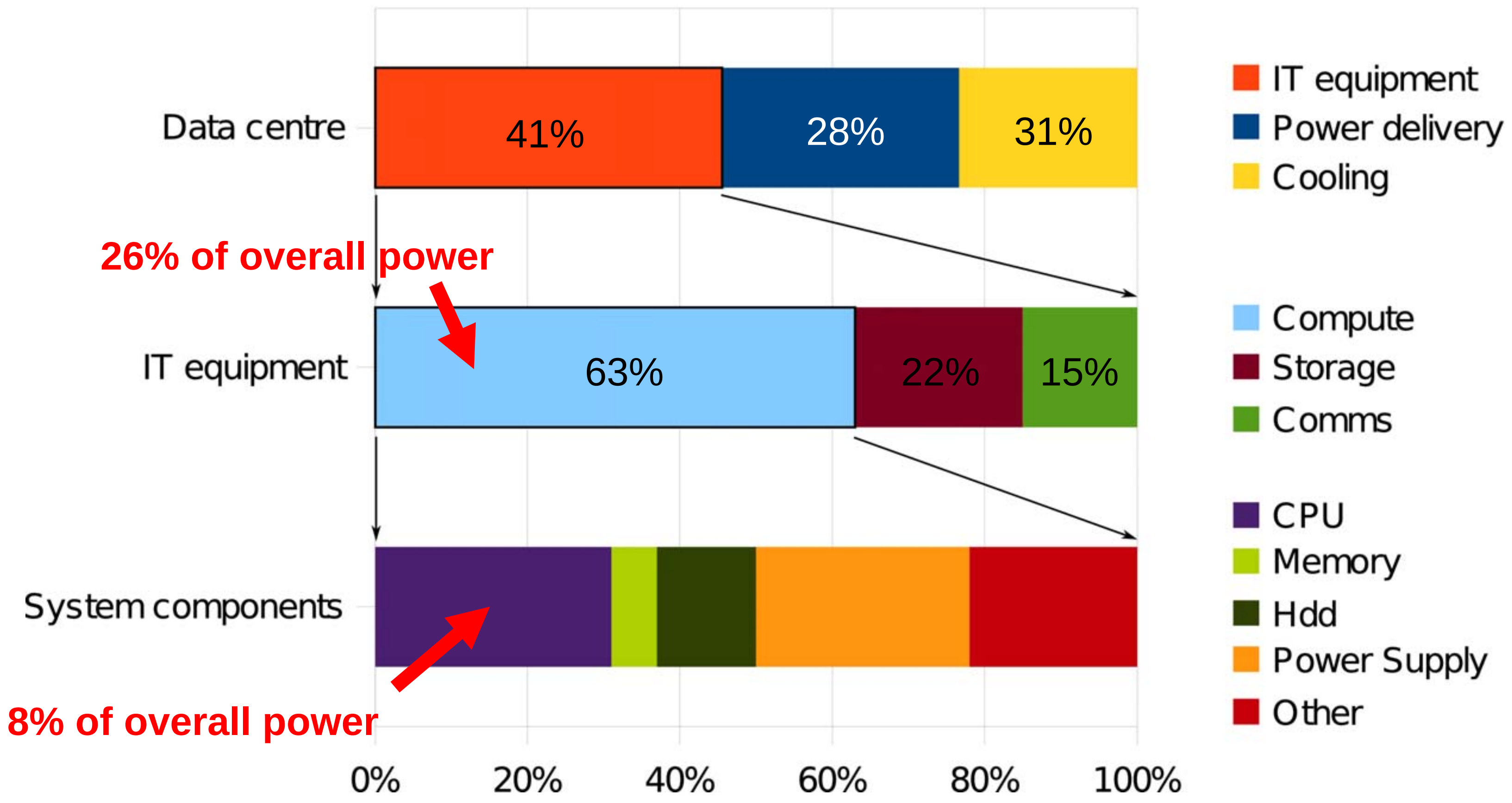
Consolidation

Load concentration

Migration

L.A. Barroso and U. Hőlzle, The Case for Energy-Proportional Computing,
IEEE Computer 40, 33–37. (DOI 10.1109/MC.2007.443.)

Use of Energy by Servers



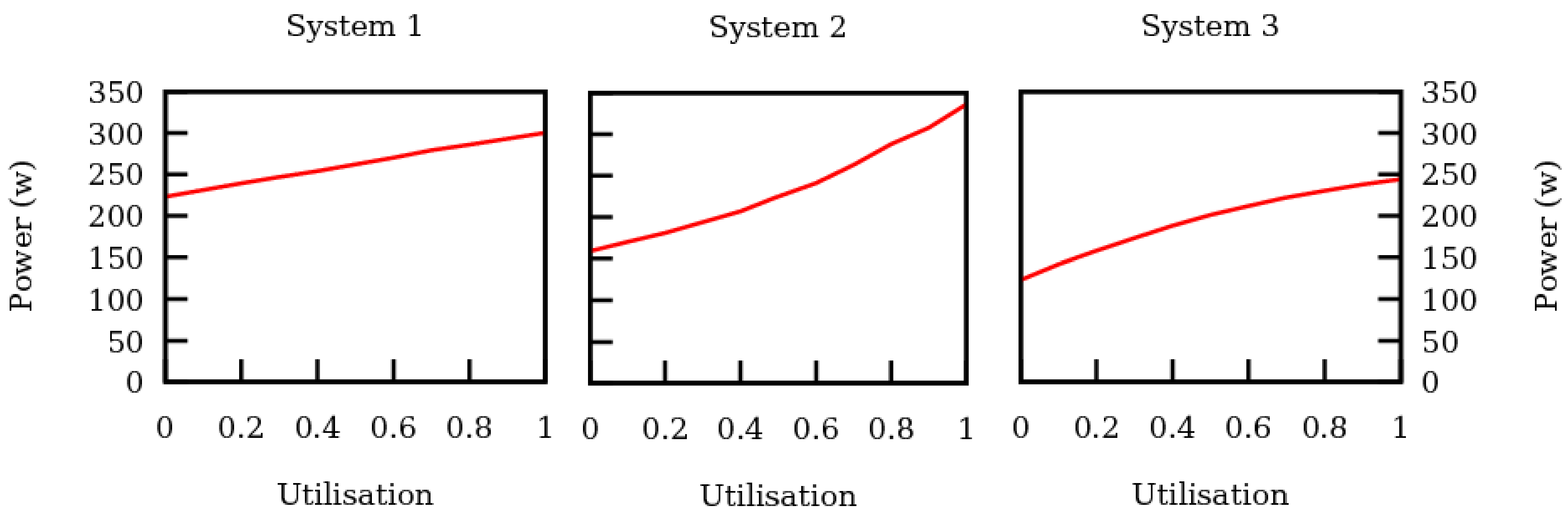
Source: Data Center Efficiency in the Scalable Enterprise,
Dell Power Solutions, Feb 2007

Energy Efficient Computing

A. Rice

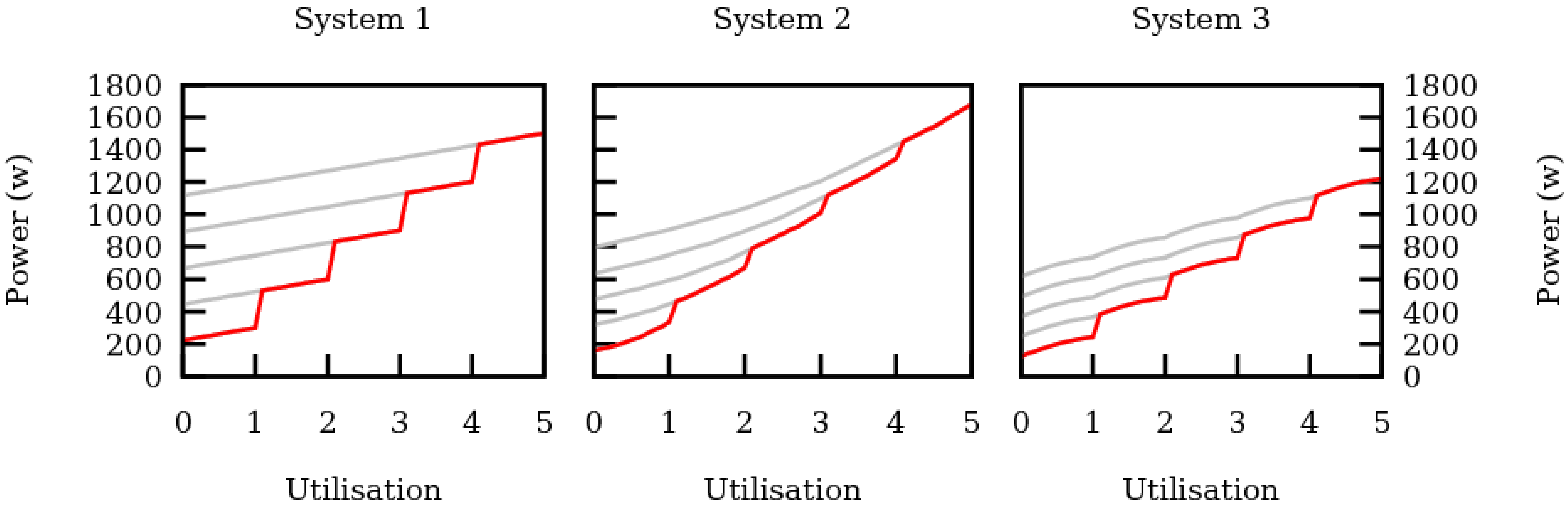
- Adaptive datacentres
  - Improve fault recovery by automatically back tracking through computational blocks
  - Use machine readable descriptions of service agreements
  - Include energy optimisation not just fault tolerance as part of adaptation strategy
  - Run “closer to the wire”
- Scale energy use with useful work done at all levels
- Develop principles
  - Switch off if not in use
  - Don’t send data if not wanted
  - Know sources of network traffic

Power and Load – Single Server



- Server consumes ~50% power when idle (power SPEC marks 2008)
- Energy efficiency is poor as server can be idle much of the time

Power and Load – Multiple Servers

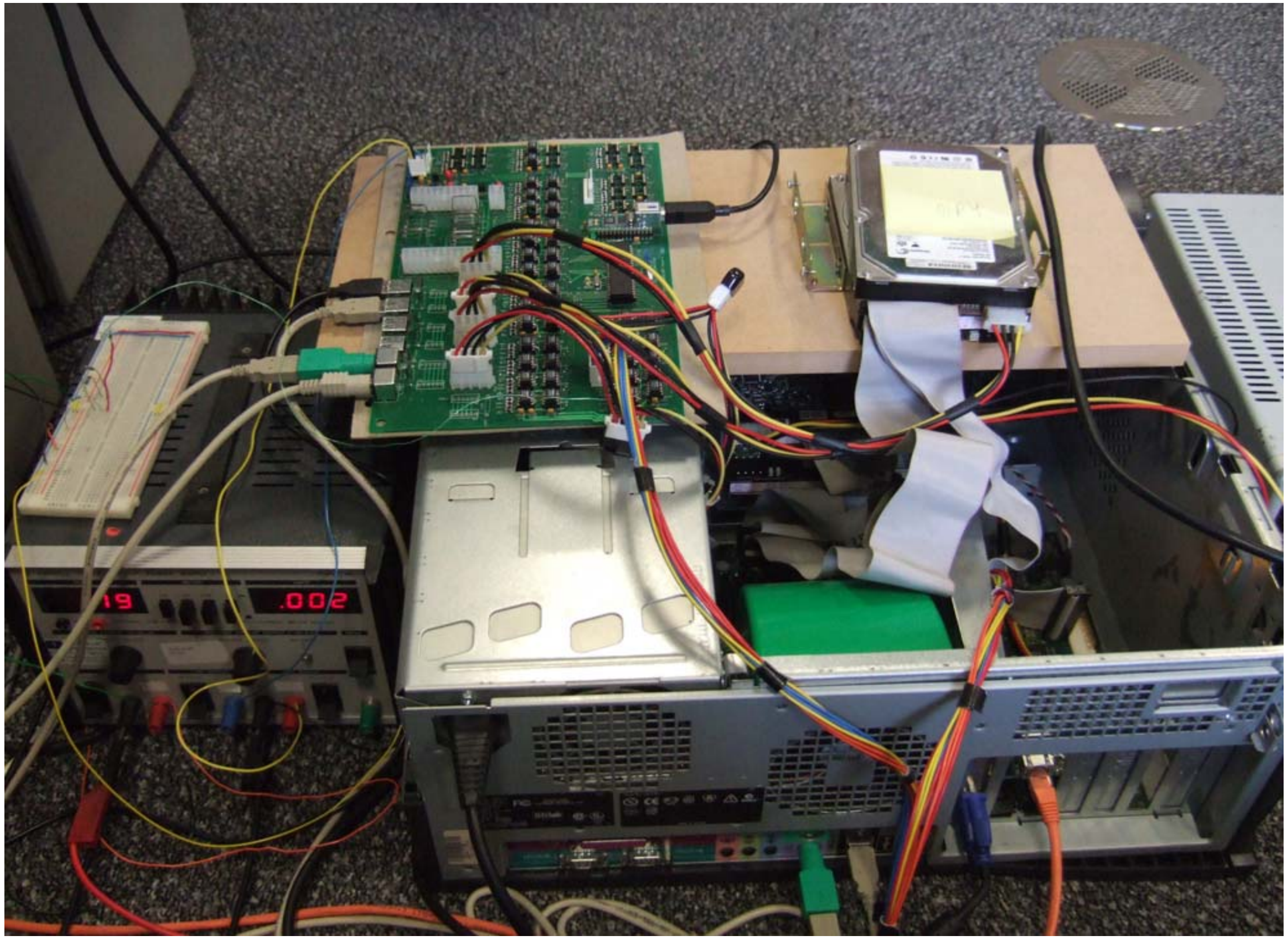


- Machines not in use are switched off
- Tasks are moved between machines
- Some tasks can be delayed
- Shape of power scaling curve less important for larger clusters

Power and Load – Multiple Servers

- Load concentration gives 30-80% improvement but has specific application requirements
- XEN virtualisation comes close to energy proportional computing in the SAN context
  - Tasks move between two servers in 250msec down time and 60sec/1Gb Ethernet,10sec/10Gb Ethernet elapsed time
- Non-interactive jobs are delay tolerant
  - Data indexing, batch simulation, climate models

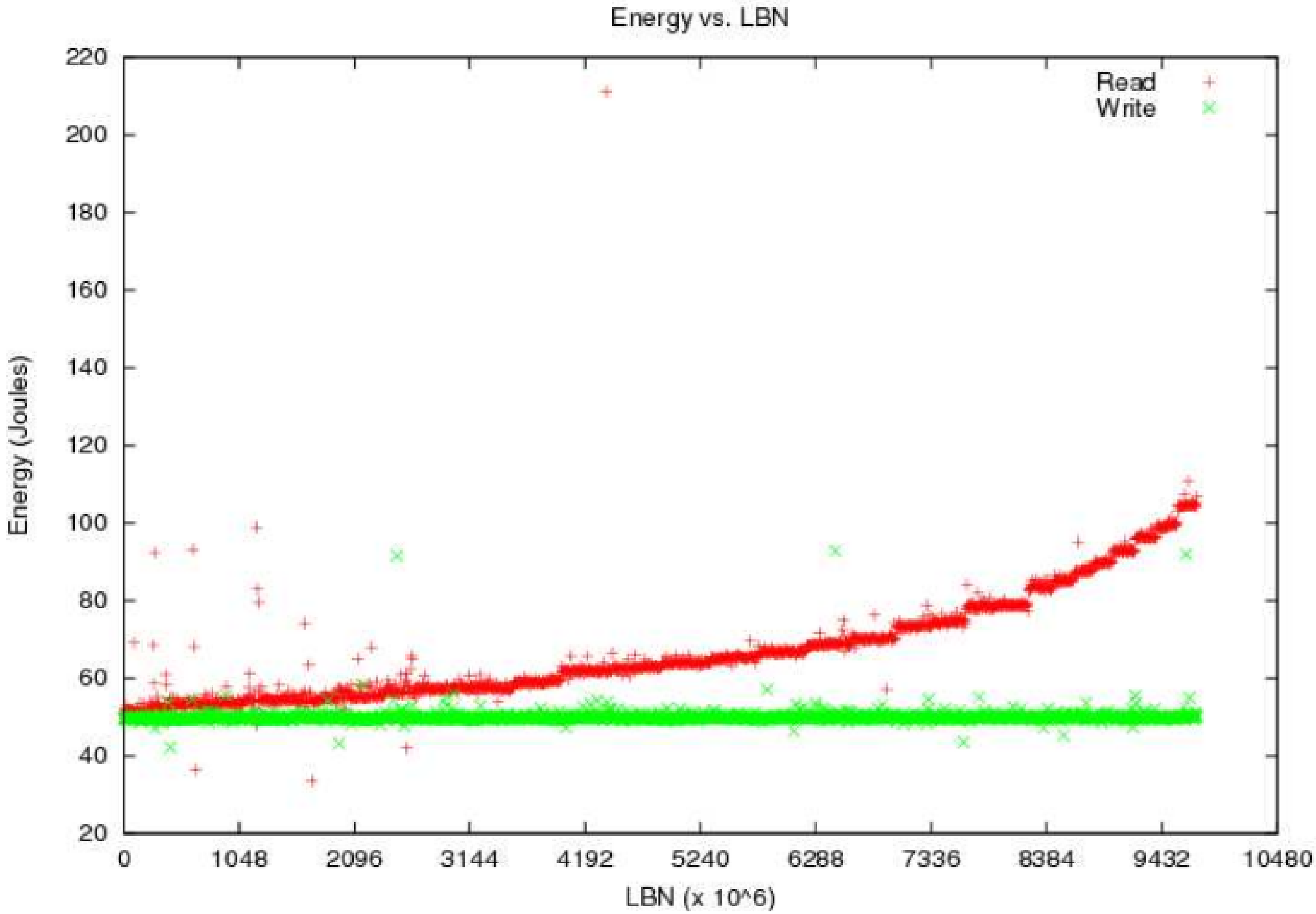
Fine Grain Power Analysis



Anthony Hylick

Use of Energy by Discs

Anthony Hylick



Use of Remote Energy Sources

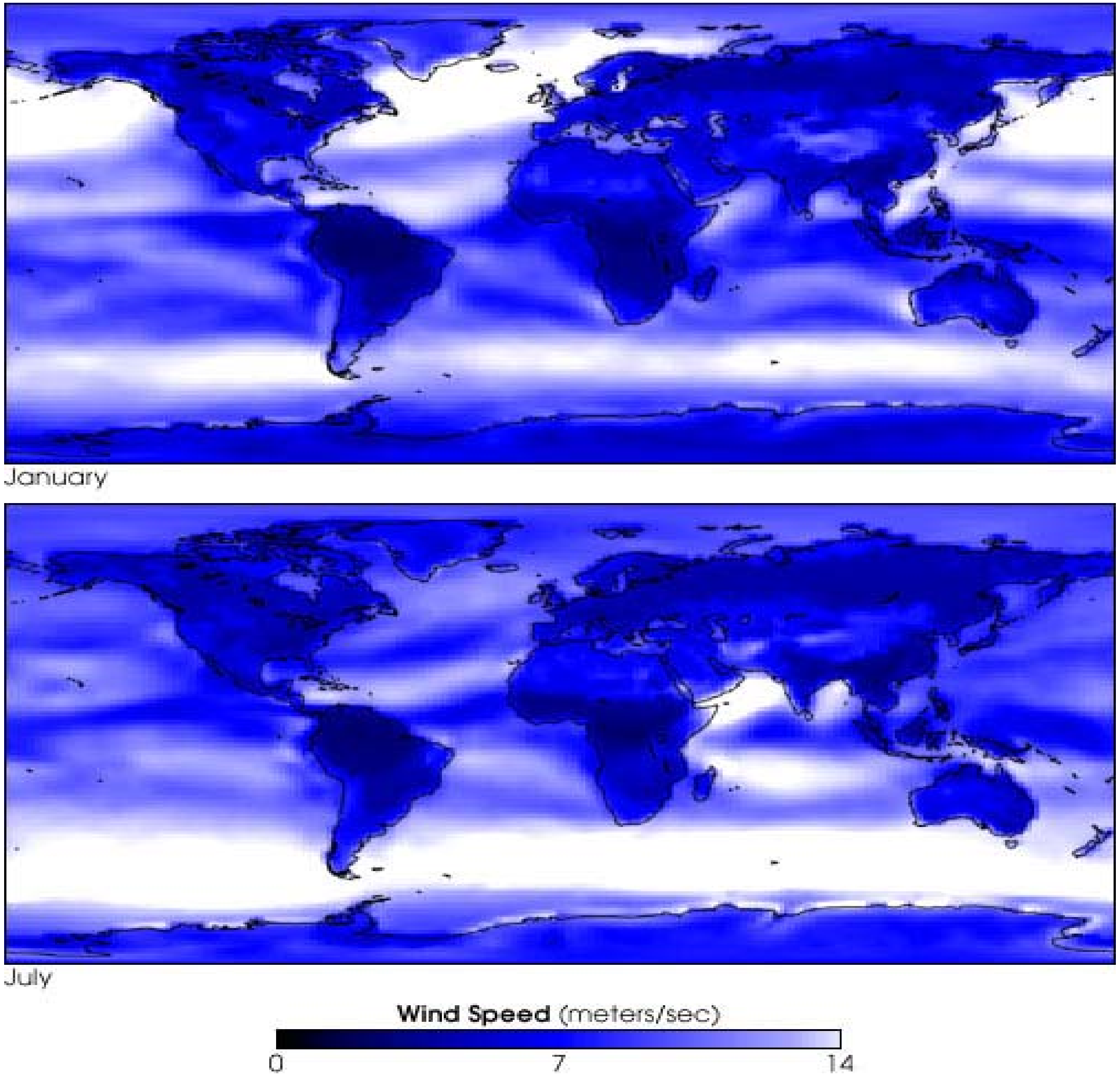
Sun

Siemens press picture

- Keep moving computing tasks to where energy is available
  - Cheaper to transmit data rather than energy
  - Use energy that cannot be used for another purpose
  - At what granularity should jobs be shipped?
  - Do we ship program, data, or both?

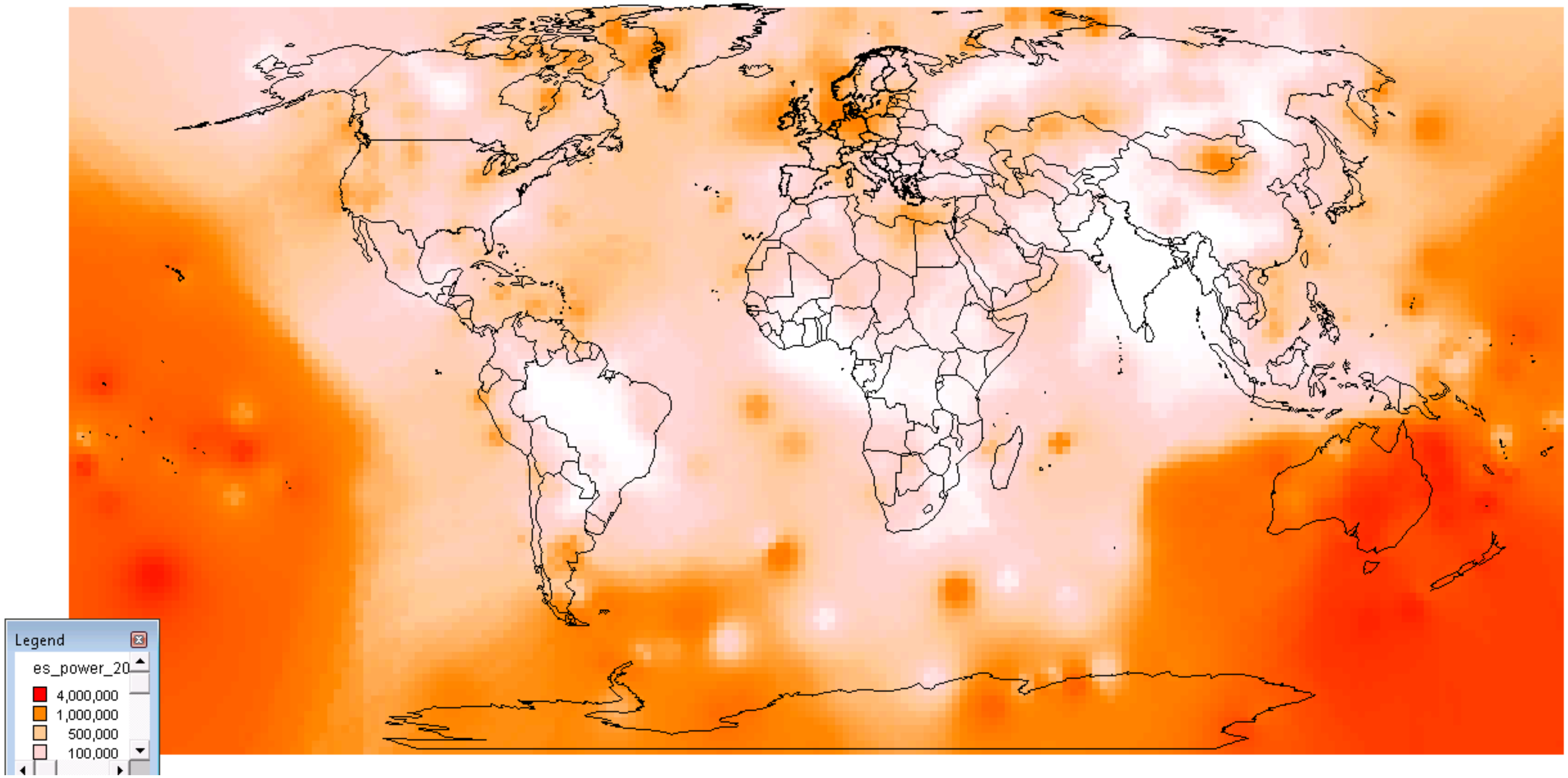
Wind Power Meteorology

visibleearth.nasa.gov



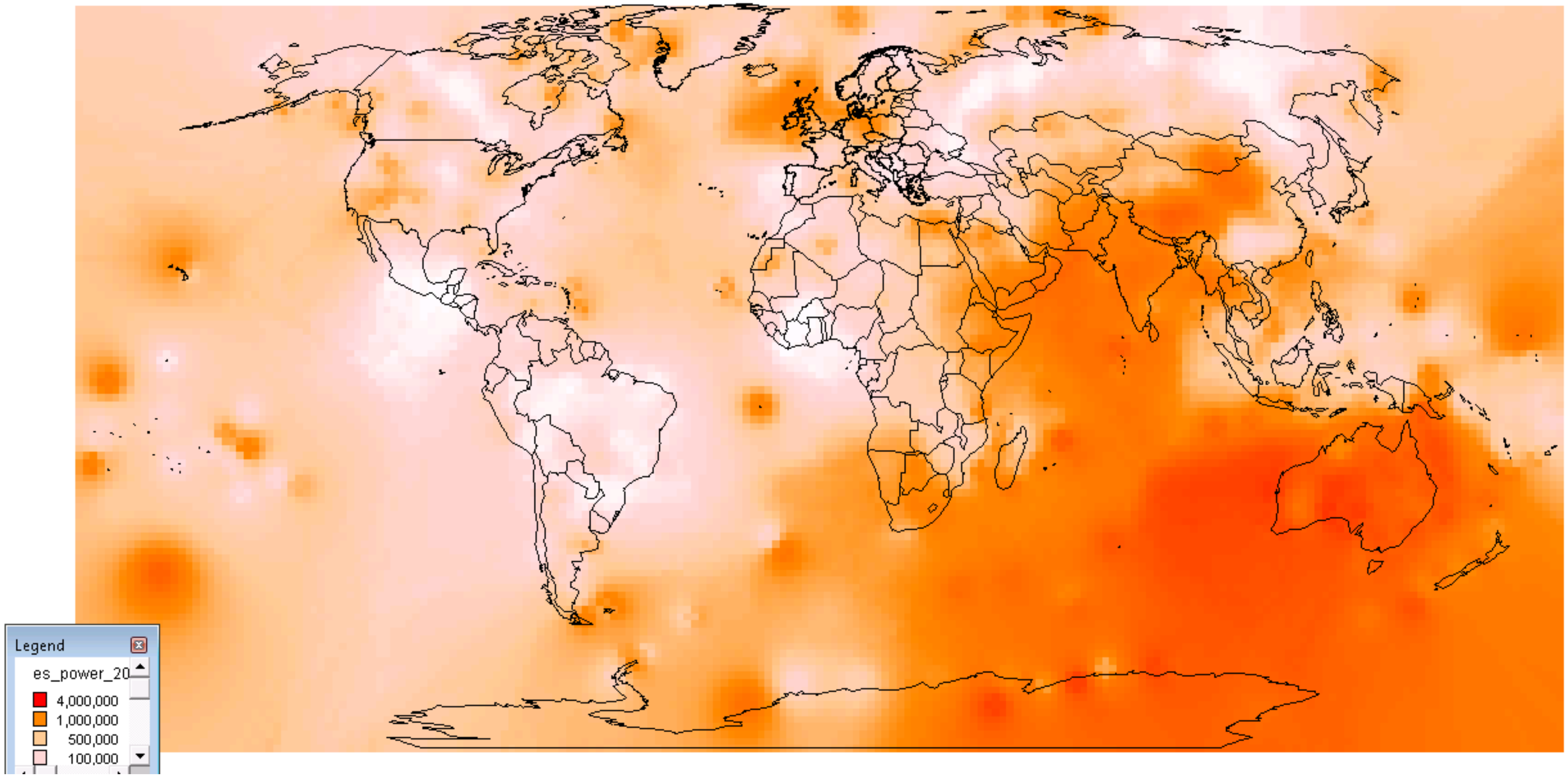
- What is the equivalent latency map?
- Where do we put the server farms?

Energy available from Wind + Solar



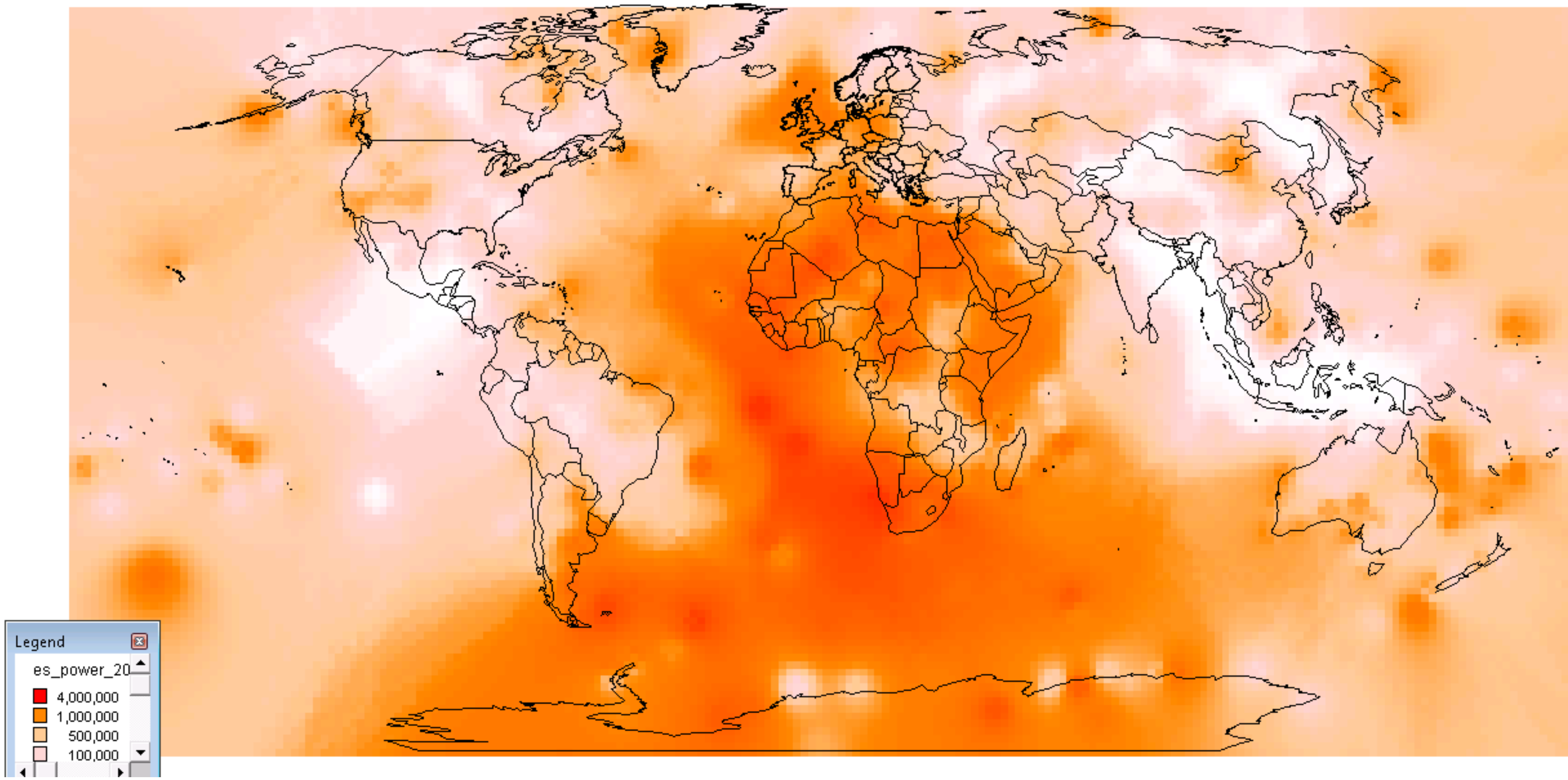
Sherif Akoush

Energy available from Wind + Solar



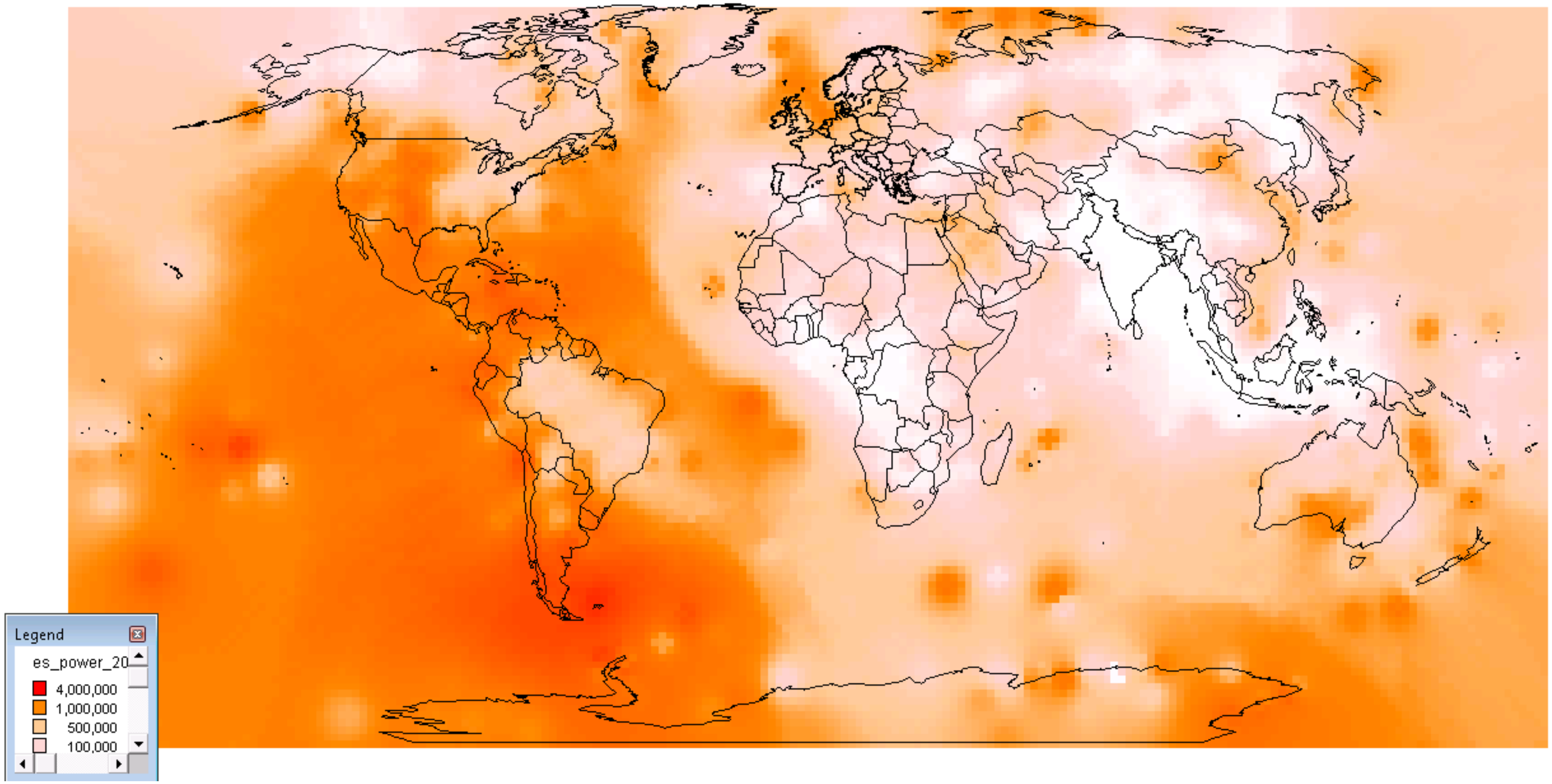
Sherif Akoush

Energy available from Wind + Solar



Sherif Akoush

Energy available from Wind + Solar



Sherif Akoush

The Overall Goal

- Optimal Digital Infrastructure
  - Components switched off if not doing useful work
  - Energy proportional computing and communications at all levels
  - Where possible use energy that would otherwise be lost (virtual battery)

- Components
  - Servers / Server Farms
  - Networks
  - Workstations
  - Terminals

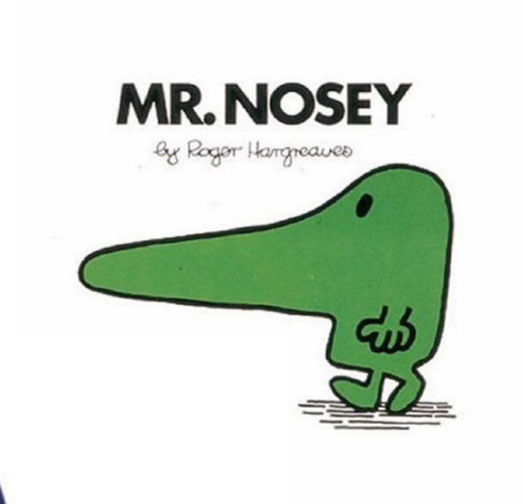
- For the first time over-provisioning and technology improvements may not save the day!

2 - Sense and Optimise

- A sensor-based digital model of the planet

- “Googling” Earth!
- “Googling” Space-Time!

- How do we do it?
  - coverage
  - fidelity
  - scalability
  - performance
  - usefulness

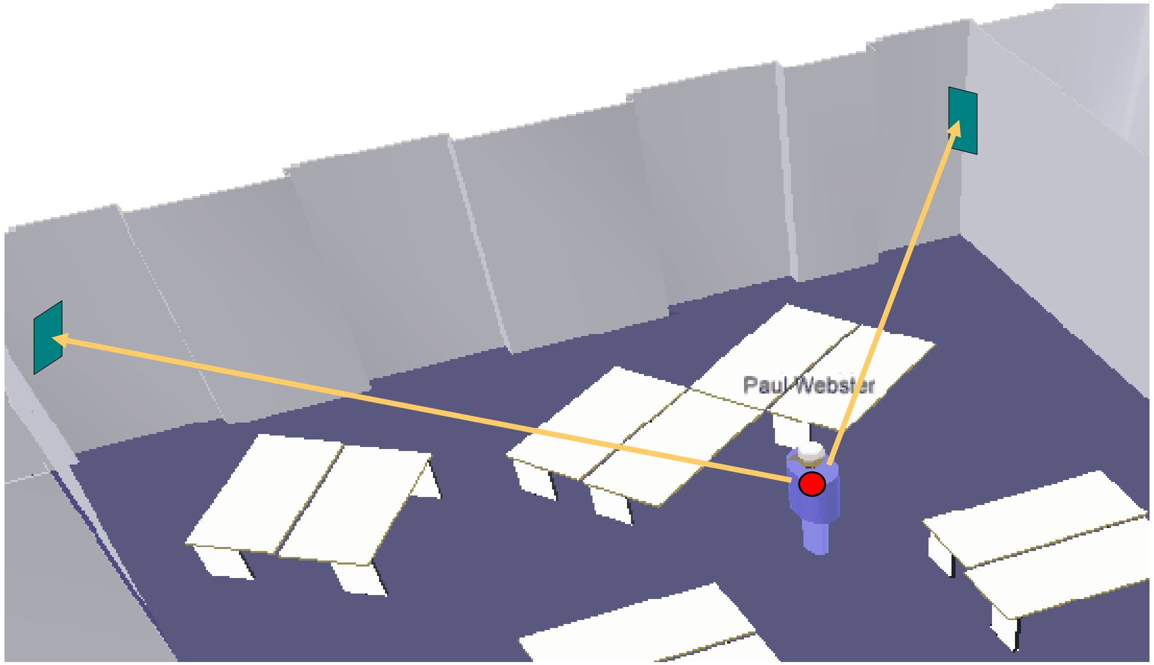


World Model

- Sensing
  - World is already full of sensors but more is to come
  - Publishing data
- Storing
  - Create a global repository
  - What are the data and computational models?
  - Consistency
- Indexing
  - Web pages
  - Sensor data
  - Shift from query-based to event-based (“where is” to “there is”)
- Interpreting
  - Observation and reaction
  - Classification
  - Optimisation
  - Prediction

Sensing Indoors

Ubisense



- Ultrawideband location system
- Measure pulse time-differences-of-arrival and angles-of-arrival

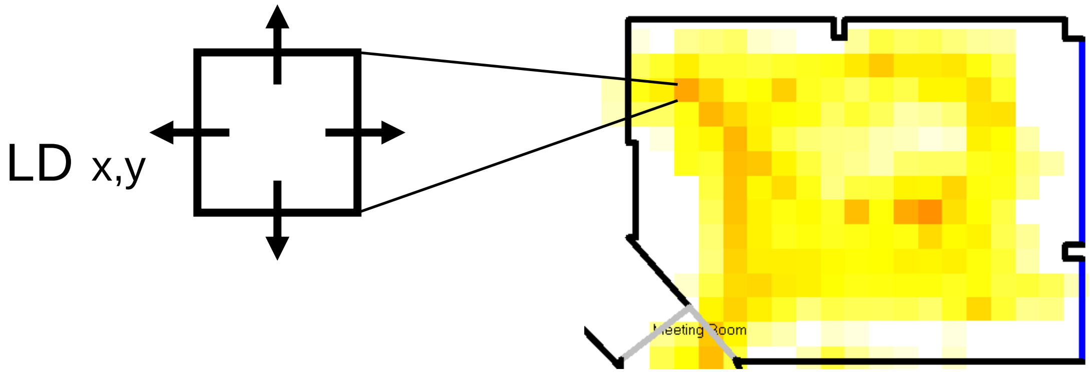
World Model Example

A real-world environment where people are wearing location tags

A 3D rendering of a **"World Model"** constructed and updated in real time using location and other systems

World Model Interpretation

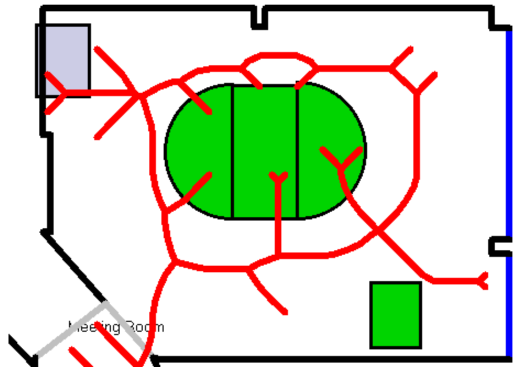
R. Harle



- 2D linkage diagram of grid transitions

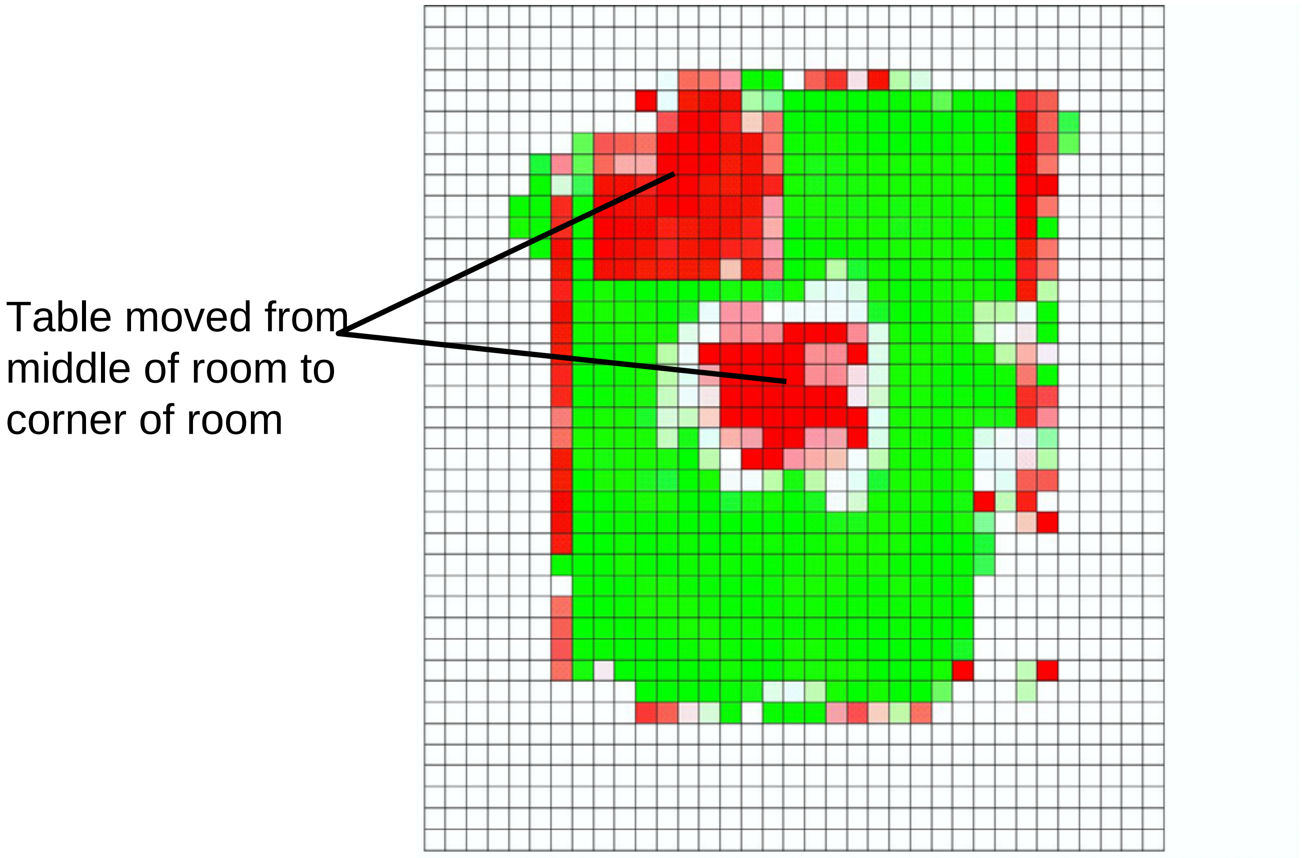
- Threshold updates using
  - maximum linkage time
  - minimum linkage length

- Various topological maps can be created

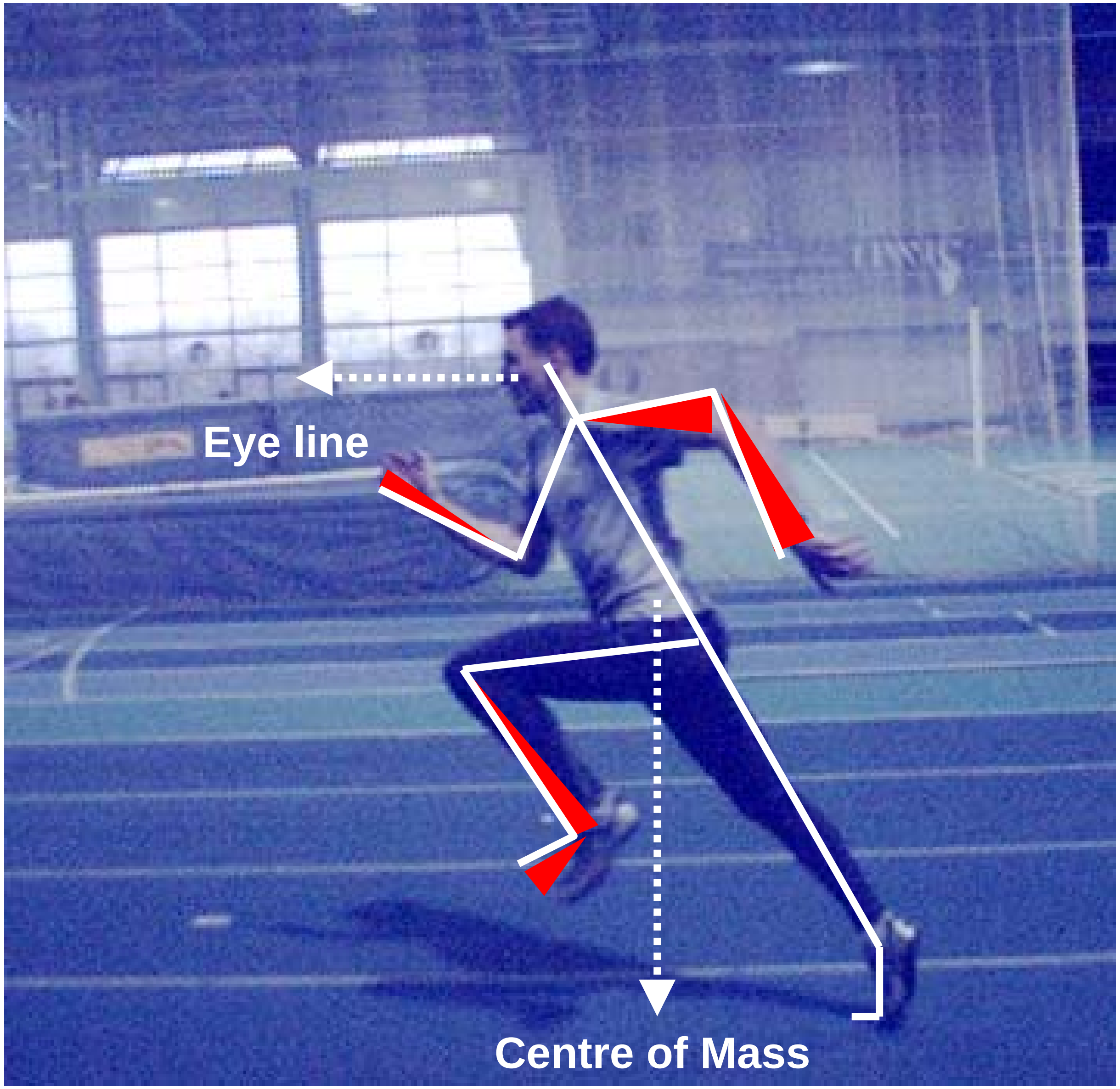


World Model Consistency

R. Harle



Sensing Athletes



Sensing Outdoors – Use of Vehicles

J. Davies, D. Cottingham, A. Beresford, B. Jones

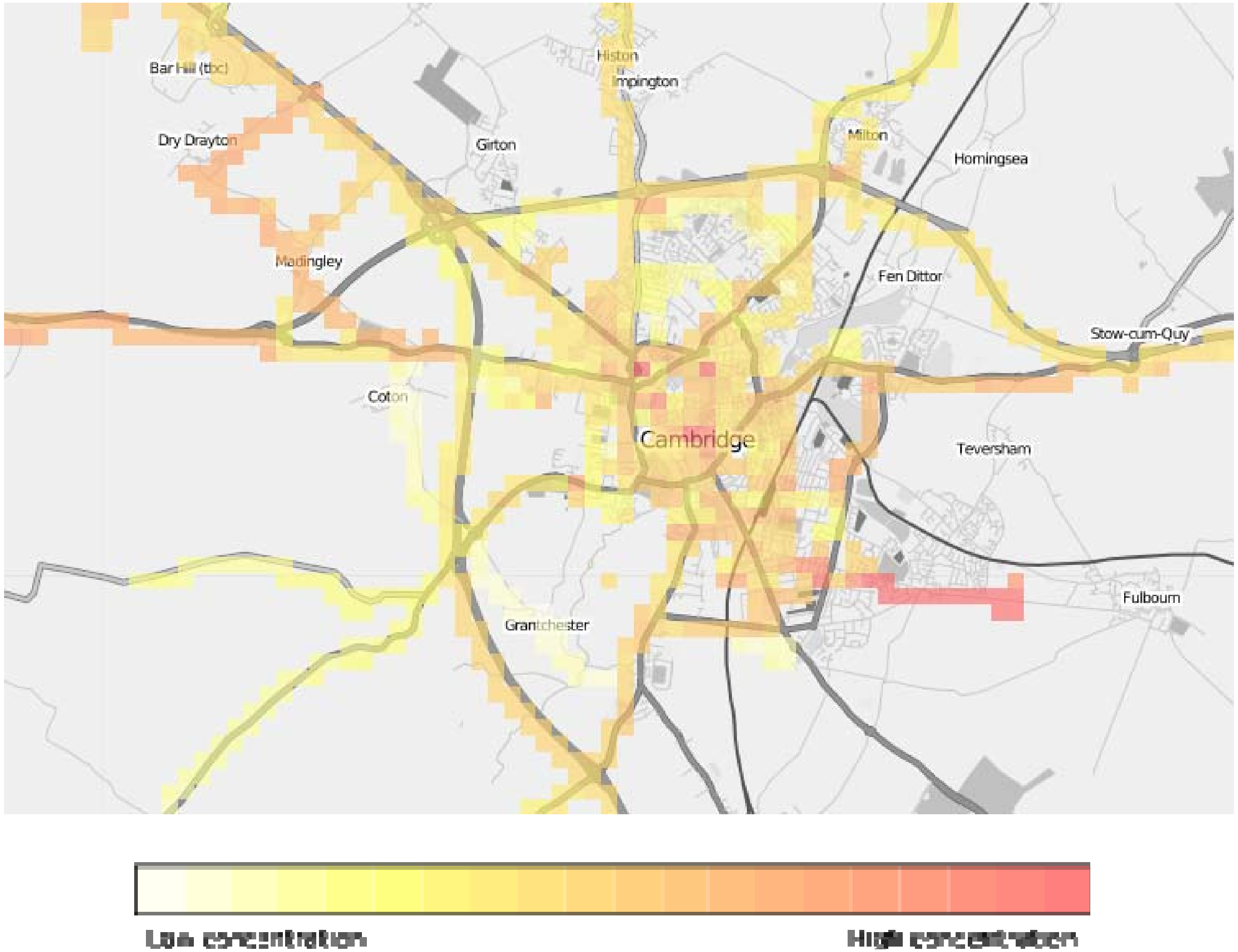
- Objective
  - Take a road vehicle
  - Embed power/processing
  - Add sensors (lots!)
  - Add storage (lots!)
  - Add networks (lots!)
  - Research platform

- Future platforms
  - Mobile “phone” as sensor?
  - Federated open Global Repository?

Concept

Reality

CO_2 in Cambridge



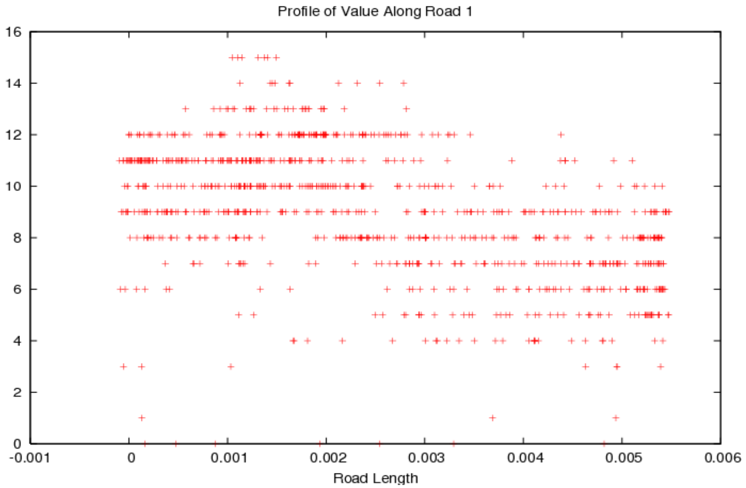
Mapping the Spectrum

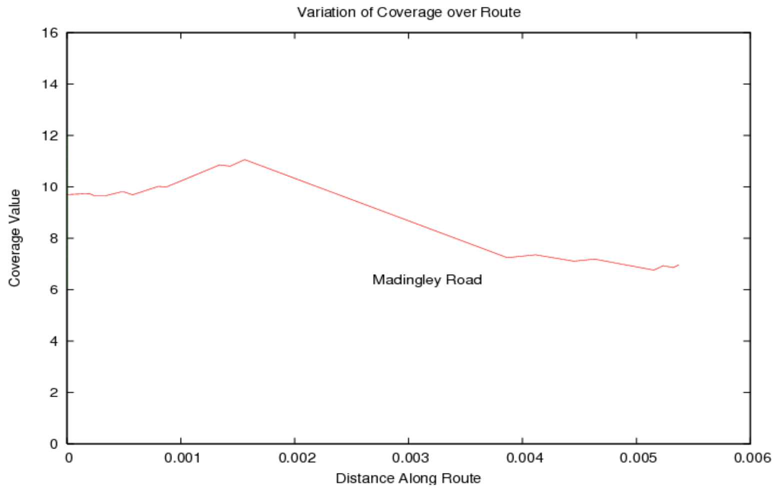
D. Cottingham

- Measured 3G signal strength
- Red is poor reception
  Blue is excellent
  Orange circles are base stations
- Results sent to the Global Repository
  - What are the standards for exchanging data?
  - How is the data marked up?
  - How does this generalise for all data?

RFeye™ DC-6GHz Node

Sensor Data Processing





- Section of Madingley Road, Cambridge
  - 1,002 input points, multi-valued at any location
  - 22 output points, single-valued function

Generating a Road Map of Cambridge

J. Davies

- GPS traces from vehicles sent to Global Repository
- Location data converted into a directed graph of the road network

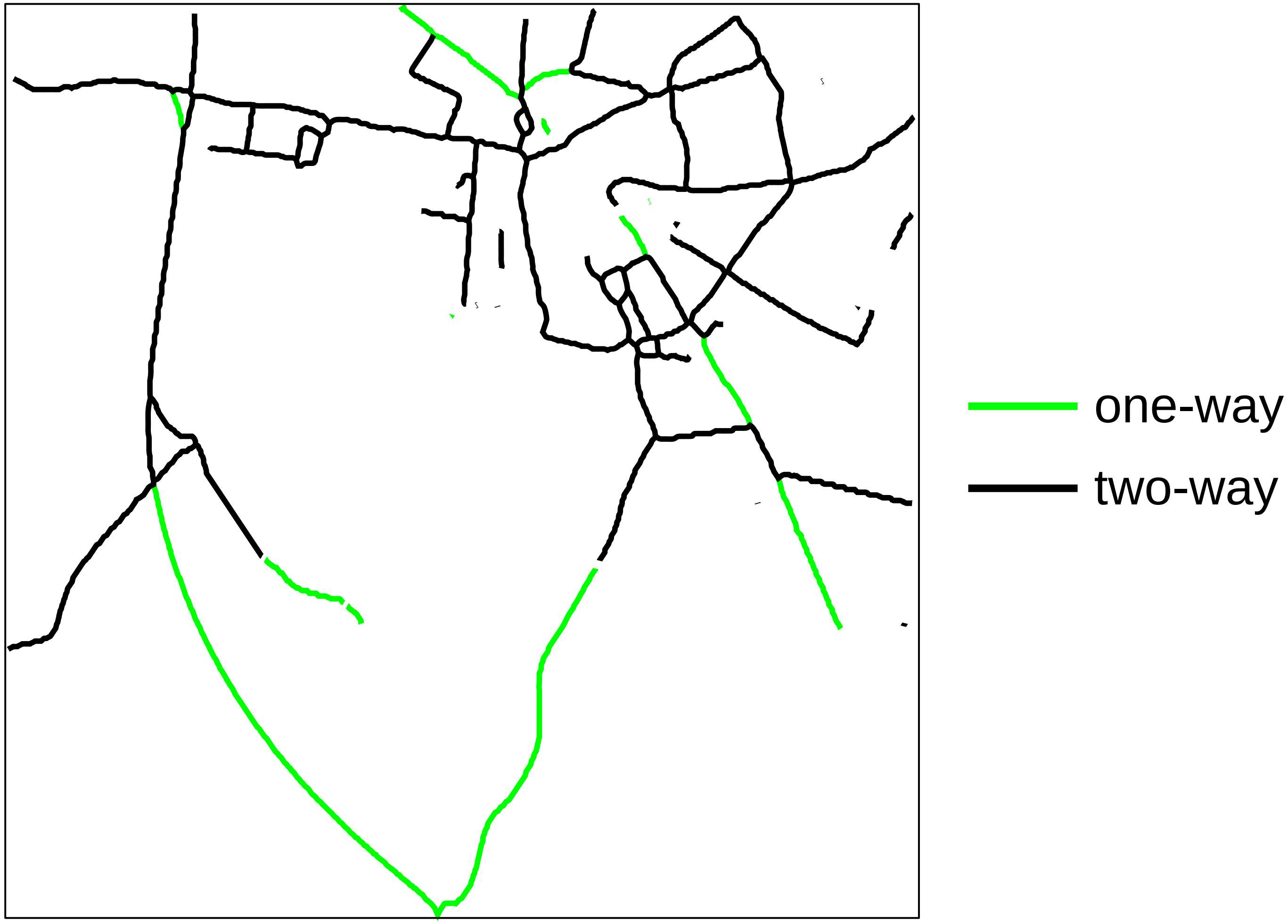
2D histogram of cells

3x3 cell blur filter

Threshold

Extract outlines

Compute centrelines



Compute directions

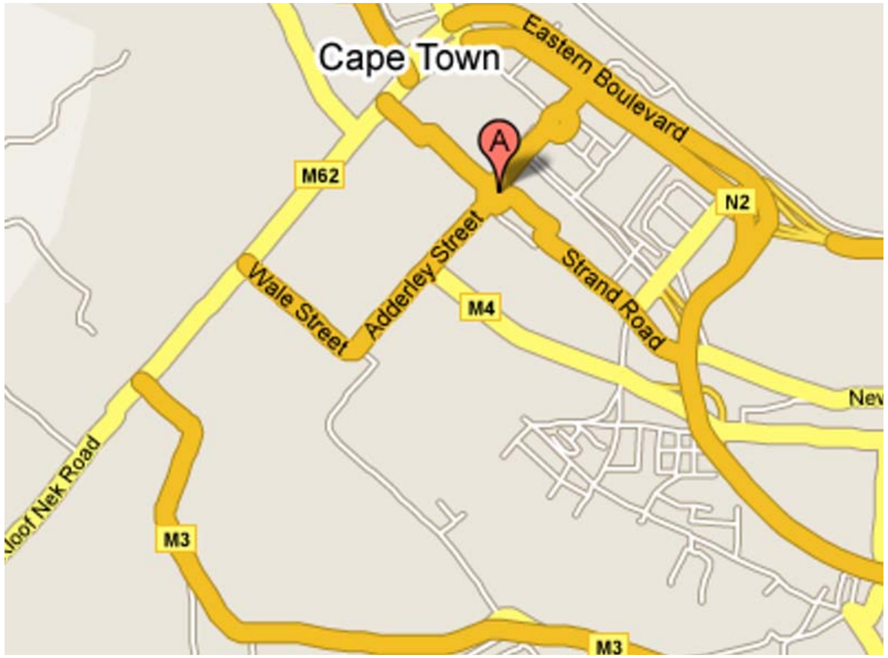
Finished map

Sensing – Humans as Sensors

www.openstreetmap.org

- Openstreet map is an example of human sensing
  - Pubs, post boxes, potholes, etc
- Reward for content creation?
- Enticing and wealth creating for developing world?

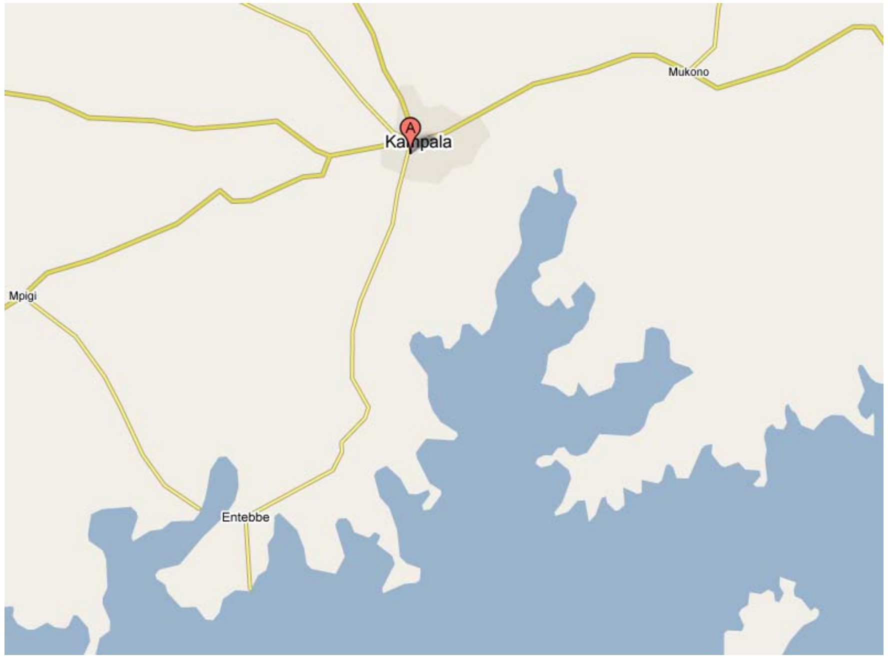
Cape Town



Data from Google

Data from OpenStreetMap

Kampala

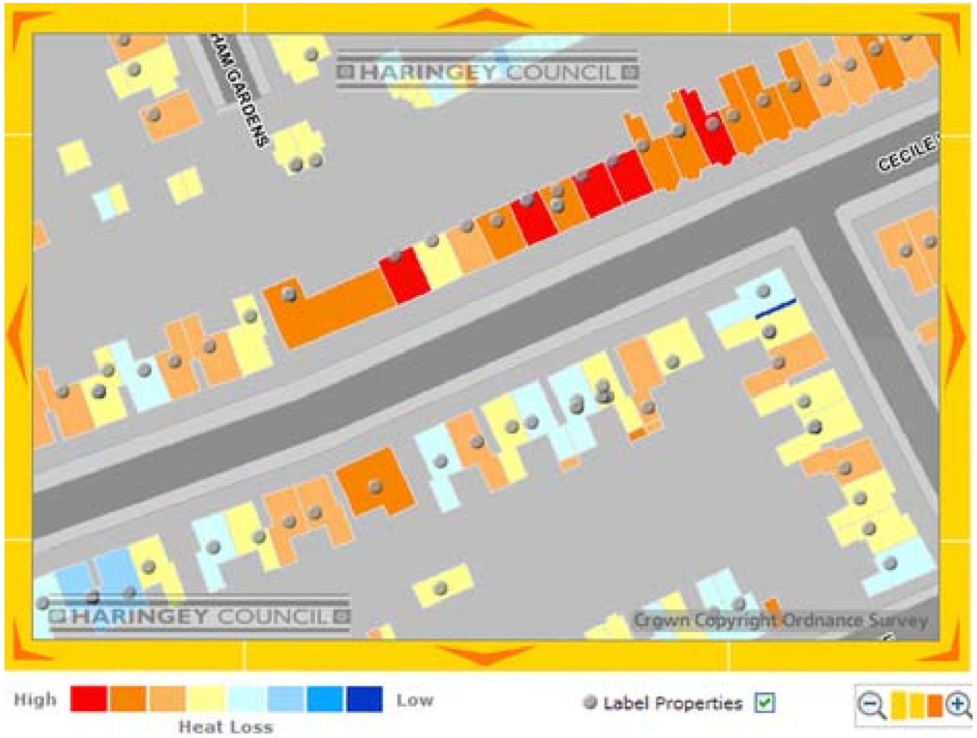


Data from Google

Data from OpenStreetMap

Thermal Maps

www.seeit.co.uk/haringey/Map.cfm



- London Borough of Haringey used aerial survey to generate thermal images
- Should this be a real-time global service like GPS?
- What applications would be written if data was free?

Personal Energy Meters

- Collect information about individual energy consumption (direct and indirect)
- Present itemised breakdown
  - travel, heating, water usage, transportation of food, etc
- Use World Model
  - upload own energy use to help digital optimisation
  - download energy profile of devices and goods
- Lots of lovely computing problems!
  - measurement, indexing, caching, event-delivery, prediction, use of social networking, security, privacy, correctness, etc

Regulation/ Incentives/ Ethics/ Privacy

www.raeng.org.uk/policy/reports/pdf/dilemmas_of_privacy_and_surveillance_report.pdf

- Generating data, changing individual behaviour

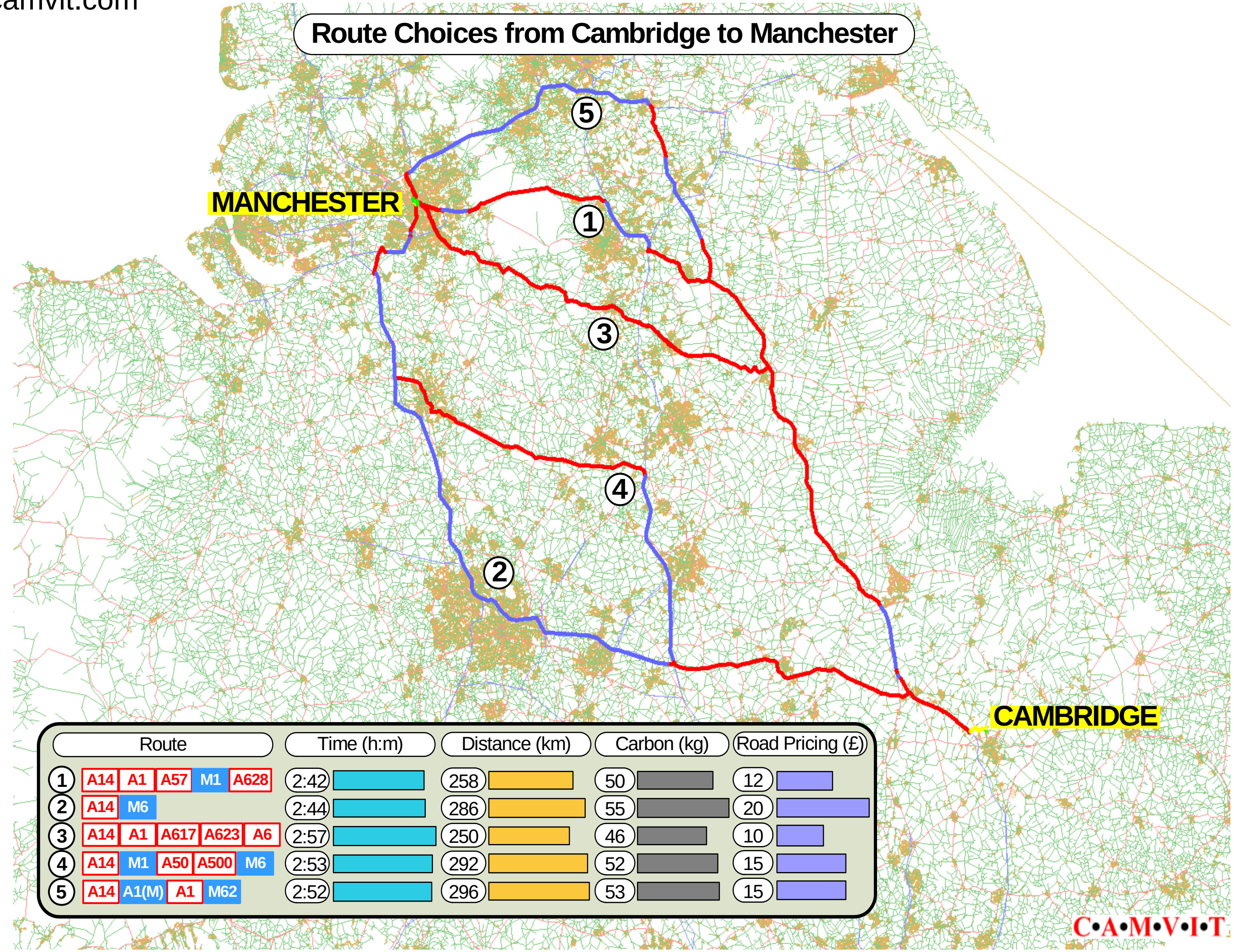
- Engineering
  - design out dangers
  - prepare for failure

- Dilemmas
  - of value, privacy, stakeholders, governance, etc
  - who to trust?

- Which surveillance scenario?
  - big brother, big mess, the citizens themselves
  - reciprocity: watching the watchers

3 - Predict and React

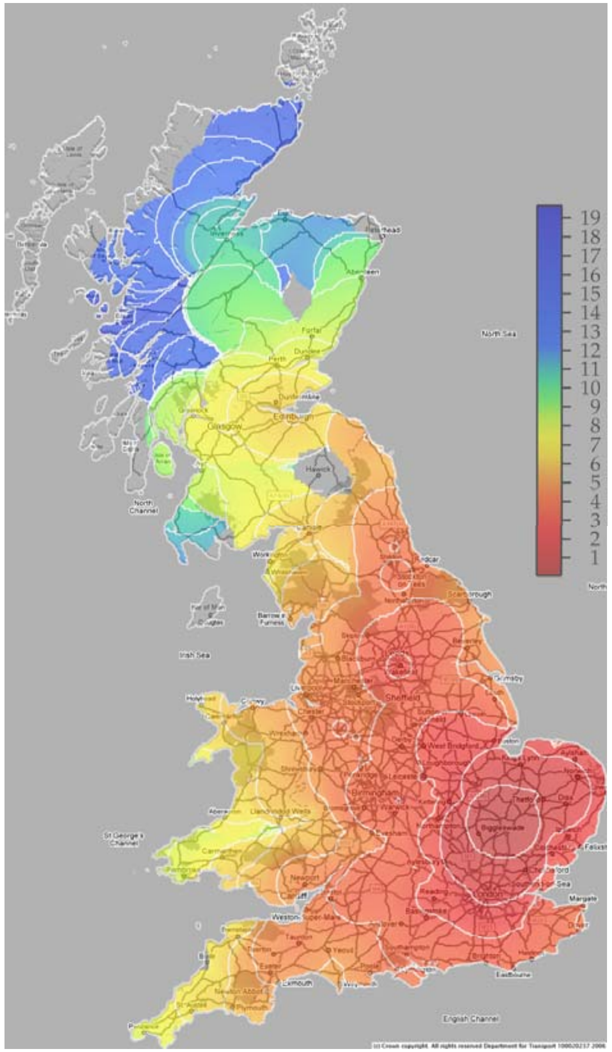
www.camvit.com



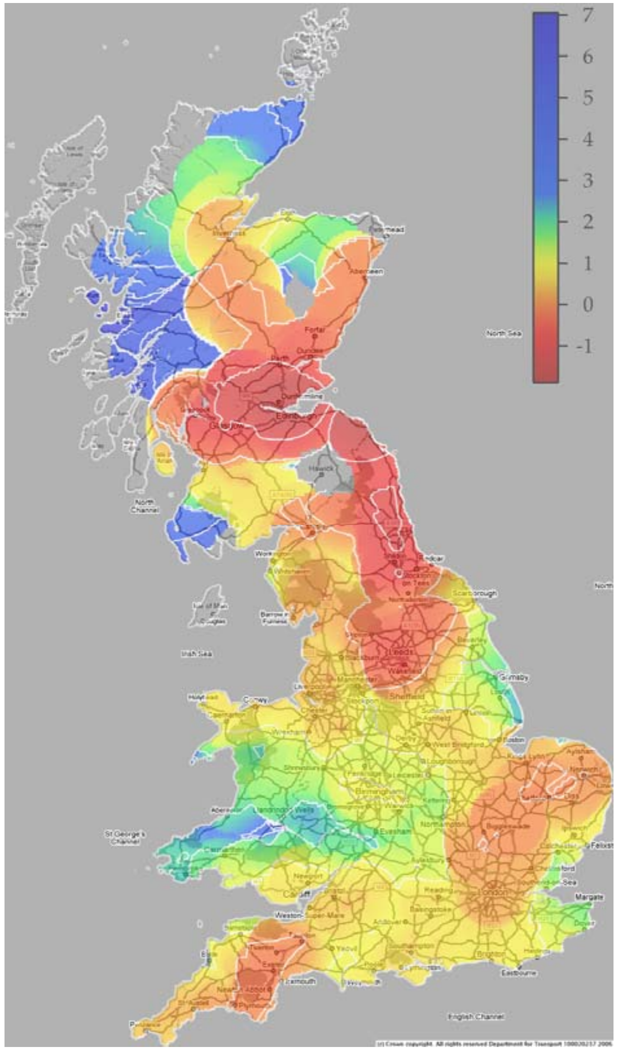
Route		Time (h:m)	Distance (km)	Carbon (kg)	Road Pricing (£)
1	A14 A1 A57 M1 A628	2:42	258	50	12
2	A14 M6	2:44	286	55	20
3	A14 A1 A617 A623 A6	2:57	250	46	10
4	A14 M1 A50 A500 M6	2:53	292	52	15
5	A14 A1(M) A1 M62	2:52	296	53	15

Information Collection / Distribution

www.mysociety.org



Travel times by Train (+1hr Taxi)
from Cambridge

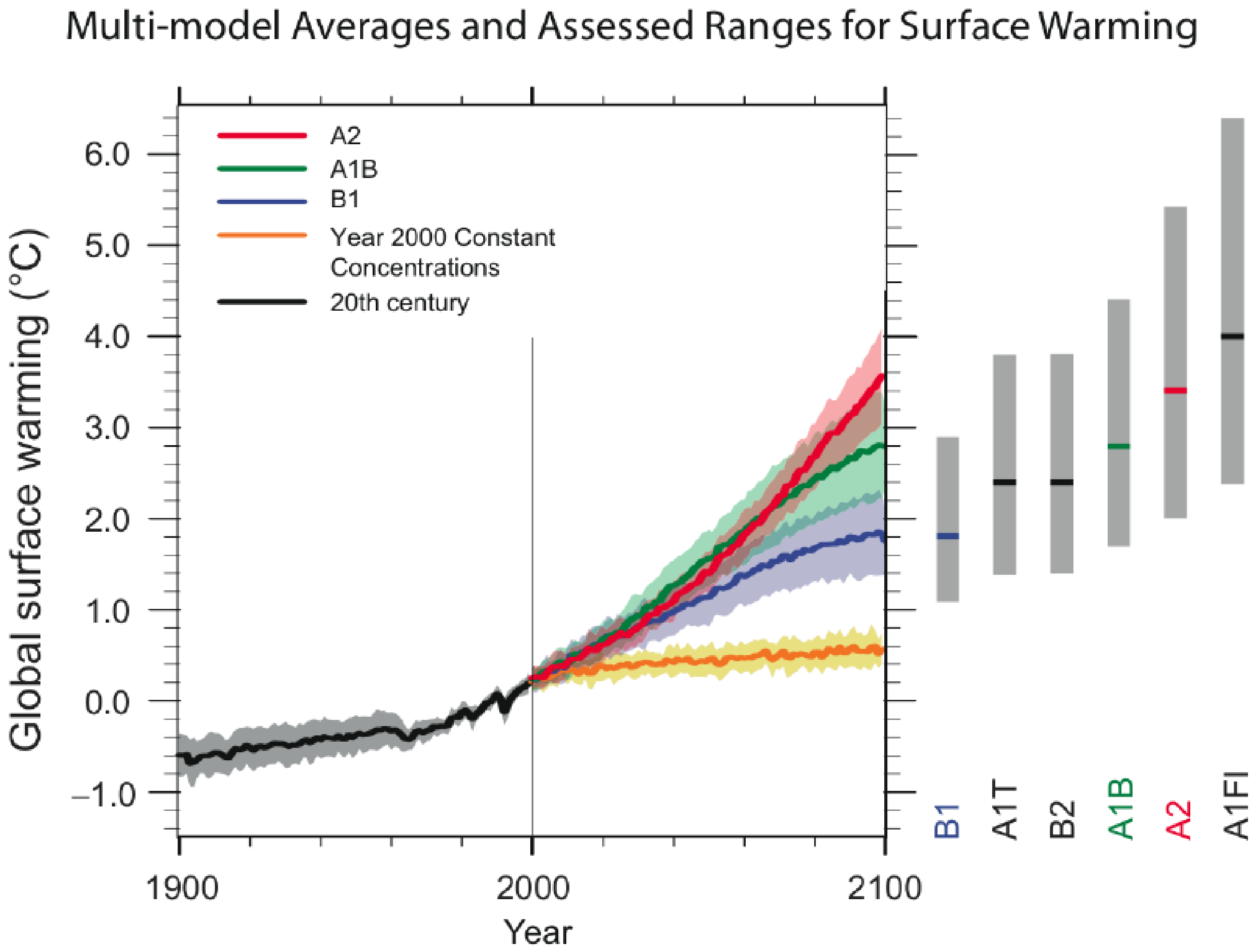


Travel times Train vs Car
from Cambridge

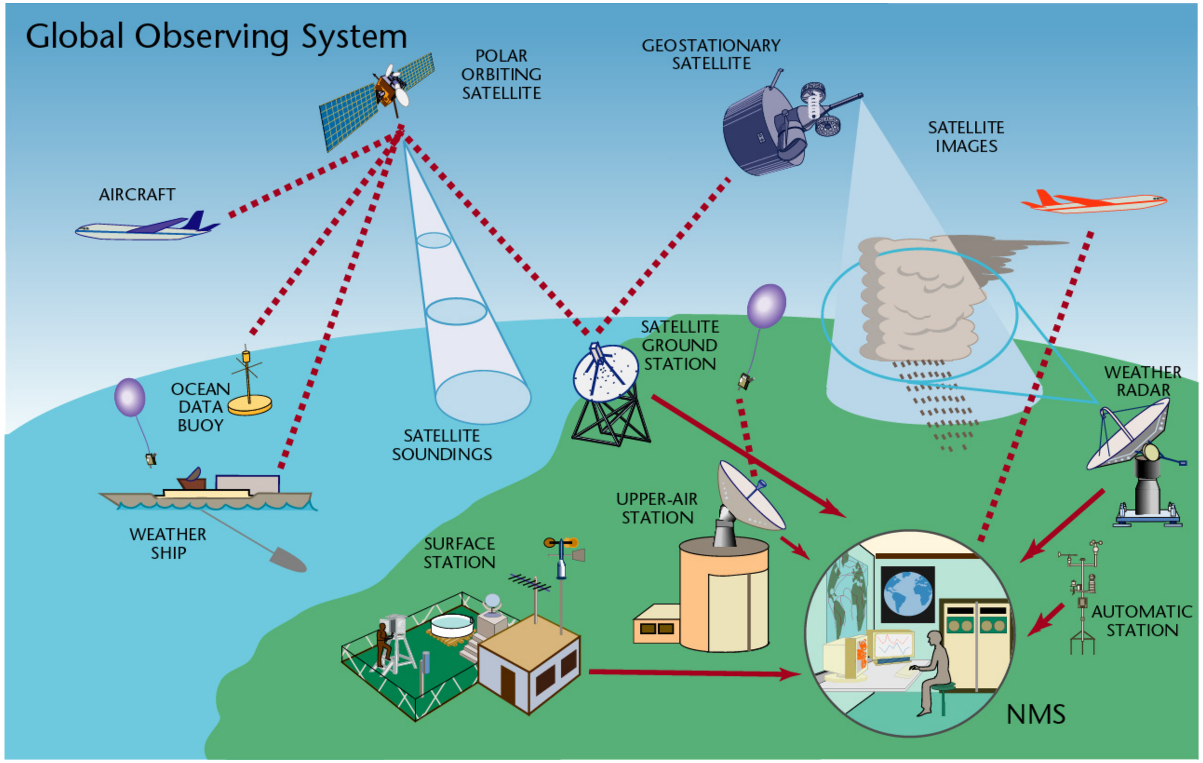
A transport network from scratch?

- Counting cars, taxis motorbikes, cycles
- Run without traffic lights?

Large Scale Models

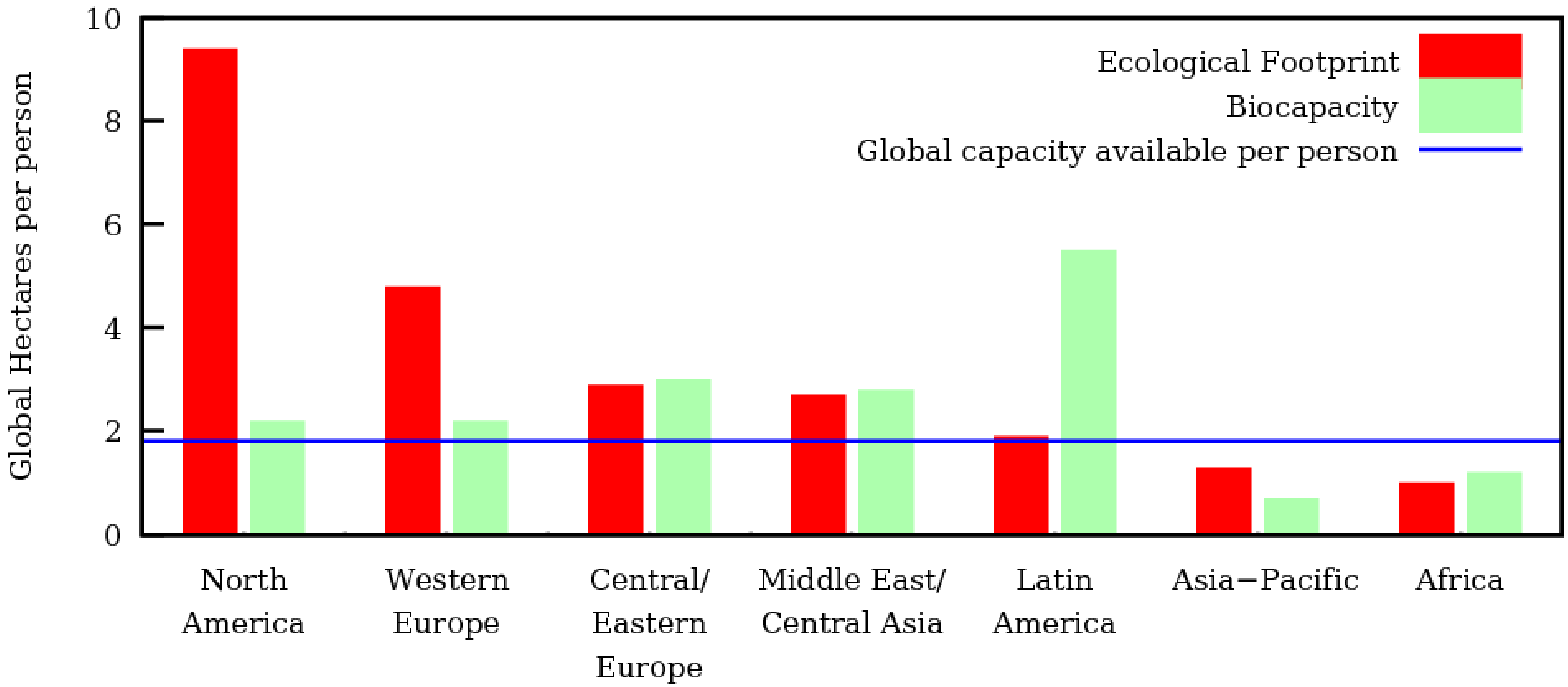


Global “Scientific” Computing



- Requirements
  - Accurate and correct model
  - Algorithm separated from implementation, verified code
  - Shared data, up to date data
  - Deadline driven computation (part of a control loop)
  - Scaleable computer power

4 - Digital Alternatives



Kitzes, J., Wackernagel, M., Loh, J., Peller, A., Goldfinger, S., Cheng, D., Tea, K. 2008 Shrink and share: humanity's present and future Ecological Footprint. Philosophical Transactions of the Royal Society 363, 467–475. (DOI 10.1098/rstb.2007.2164.)

Physical to Digital

GuardianUnlimited network

- Move bits rather than people or products
  - iTunes, Tesco Online, etc
- Good news or bad news?

Shift to Cyberspace?

- Can we construct a digital world in which we can conduct our lives?
  - on a ultra-cheap open platform
  - using miniscule power
  - fed with sensor data from the real-world
  - accessible to every human

- Scaling up virtual worlds is a challenge
- Key to wealth creation in developing world?

South African Township

- No power services to buildings
- Mobile phones are common
- Top-ups in units of 1 SMS
- A meal is an SMS or two

Computing for the Future of the Planet

www.engineeringchallenges.org

- Lists, dimensions, and quantifies computing problems
- Targets the world outside (and inside) computing
- A vision and an architecture
- Contemplates the unbounded upside of computing!

Engineers set 'grand challenges' to enhance life

14 targets

- Make solar energy affordable
- Get energy from fusion
- Develop carbon sequestration methods
- Manage the nitrogen cycle
- Expand access to clean water
- Restore and improve urban infrastructure
- Advance health informatics
- Make better medicines
- Reverse-engineer the brain
- Prevent nuclear terror
- Secure cyberspace
- Enhance virtual reality
- Advance personalised learning
- Engineer tools for scientific discovery