ouees-201506 Part 3: Environmentally-sustainable computing

Kenji Rikitake

23-JUN-2015
School of Engineering Science
Osaka University
Toyonaka, Osaka, Japan
@jj1bdx

Lecture notes on GitHub

- https://github.com/jj1bdx/oueees-201505-public/
- Don't forget to *check out the issues*!

Sustainability: economic feasibility, energy efficiency, scalability

Economic feasibility of computing

- Device production: can we make it?
- Complexity: can we solve it?
- Energy consumption: can we feed them?

Tackling with physics

- Speed of light = latency
- Heat dissipation
- Device density
- Radio bandwidth limitation
- Scaling by distribution

Tackling with complexity

- Addressing objects
- Routing computation
- Autocracy .vs. distribution
- Concurrency .vs. consistency
- System administration cost

Tackling with scalability

- Scalable: *handling growth*
- Scaling up: higher processing power
- Scaling out: more computer units
- Consistency .vs. scalability
- Efficiency issues: power consumption, parallelized speed gain, inconsistency allowance

Energy consumption: the final frontier

(Information) CAPITALISM

Towards information capitalism

- Mercantillism: collecting wealth, colonialization, trade barriers
- Industrial capitalism: factory, labor division, industrialization, imperialism
- Information capitalism: investment (derivatives), for-profit, commoditization

An information capitalism principle: hyper over-provisioning

- Resource extinction *instantly* kills the system
- For preventing the extinction or starvation, *keep the resources as much as you can*
- Implication: *expansionism*
- Assumption: resources are *infinite*

Question: are natural resources *infinite*?

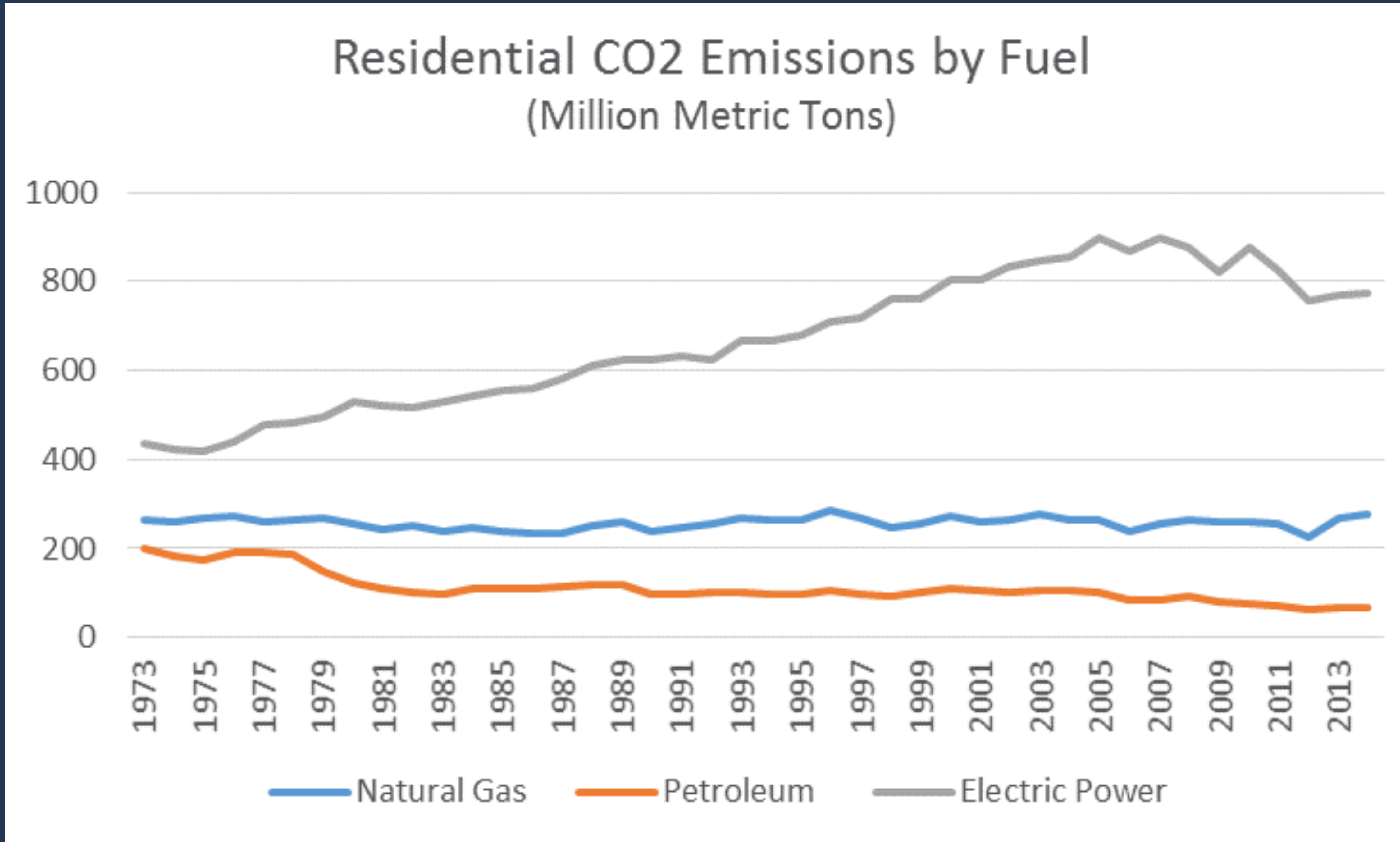
Our lives depend on electricity

Electricity as energy

- Well-established transportation technologies: high-voltage wires (with superconductivity)
- Can be saved in various forms: chemical energy (batteries), potential energy (dams), physical energy (flywheels)
- Relatively easier to control the flow
- Safer than natural gas and liquid fuels

Problems on energy consumption

- Quantity: *exponentially increasing*
- Efficiency: improvement stagnated (e.g., electricity delivery loss)
- Demand and desire: more and more people want to *modernize* their lives
- Many stakeholders of conflicting interests

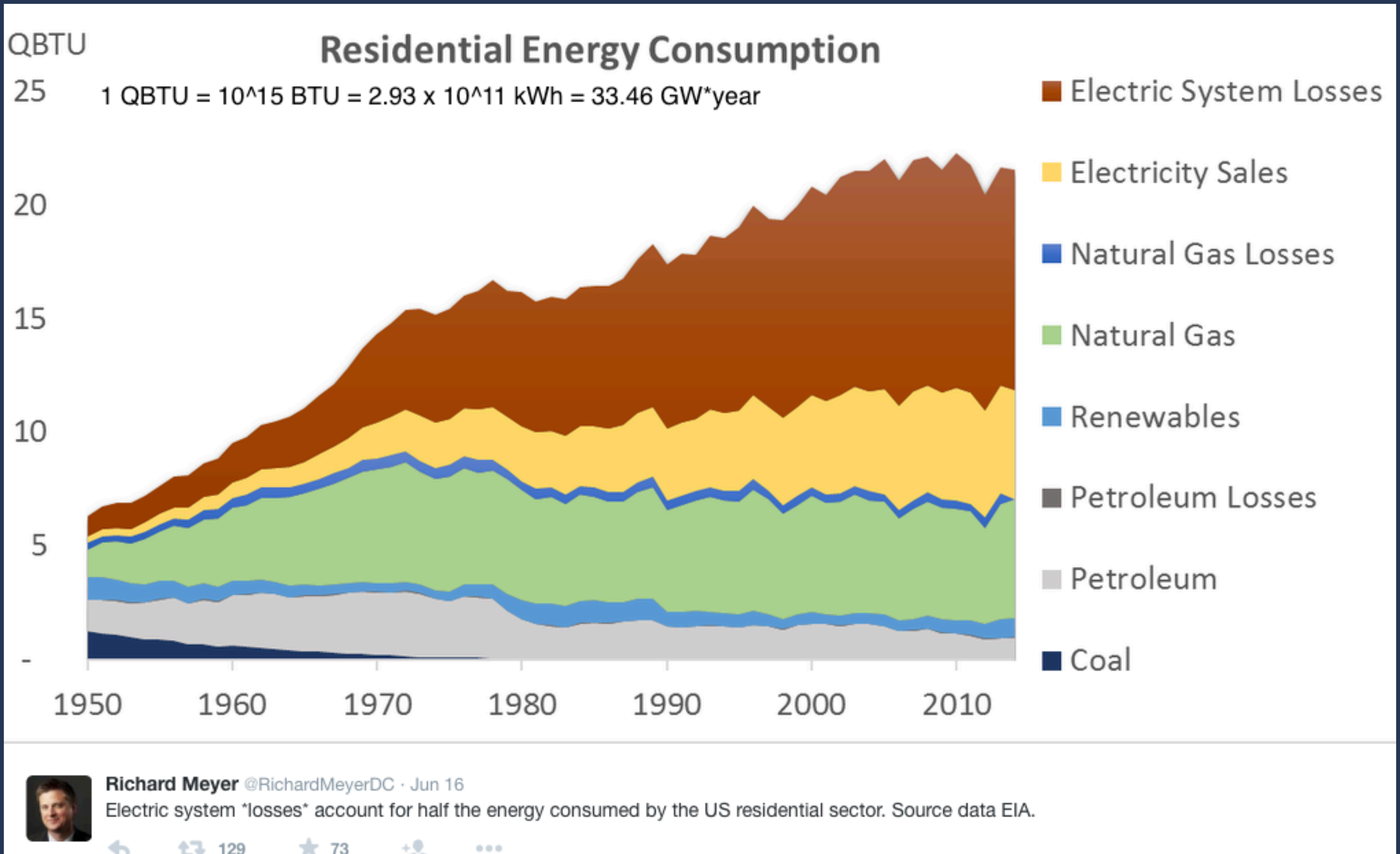


Richard Meyer @RichardMeyerDC · Jun 16

Nearly all of the recent (40 yrs) growth in US residential CO2 emissions is due to electric power consumption.

3 4

Source: https://twitter.com/RichardMeyerDC/status/610547856693399552



Source: https://twitter.com/RichardMeyerDC/status/610536366594781184

An alarming prediction: Internet may use up all electricity supply capability by 2025

Source: CEET Annual Report 2013, Center for Energy-Efficient Telecommunications, University of Melbourne Page 3, http://www.ceet.unimelb.edu.au/publications/downloads/ceet-annualreport-2013.pdf

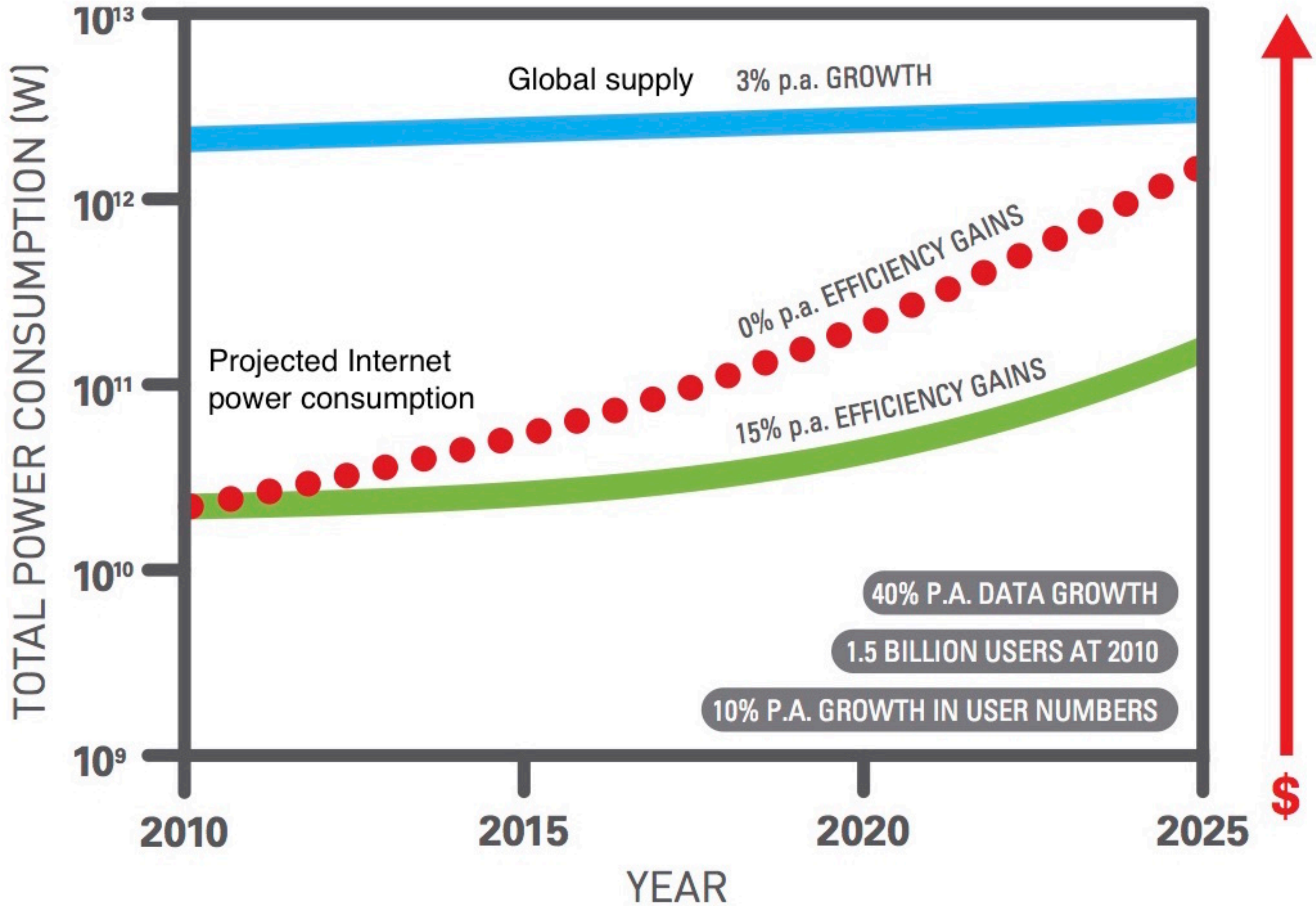


Figure 1 The growth of power consumption of the Internet over the coming years assuming current growth rates in traffic and number of users.

Data centers in the USA

- 12M servers in 3M data centers [1]
- 2013: 91TWh / 34 x 0.5GW power plants
- 2020: 140TWh / 50 x 0.5GW power plants
- 2020: 150Mt CO_2 pollution

[1] Data Center Efficienct Assessment, Natural Resource Defense Council, August 2014

Data center metrics

- Server utilization rate
- Power Usage Effectiveness (PUE)

Server utilization rate

- [processing load] / [maximum server capacity]
- 10% utilization rate server can spend 30% to 60% of power

Power Usage Effectiveness (PUE)

- Measuring *cooling* efficiency
- [total power] / [server-consumed power]
- should be <2.0, closer to 1.0 is better

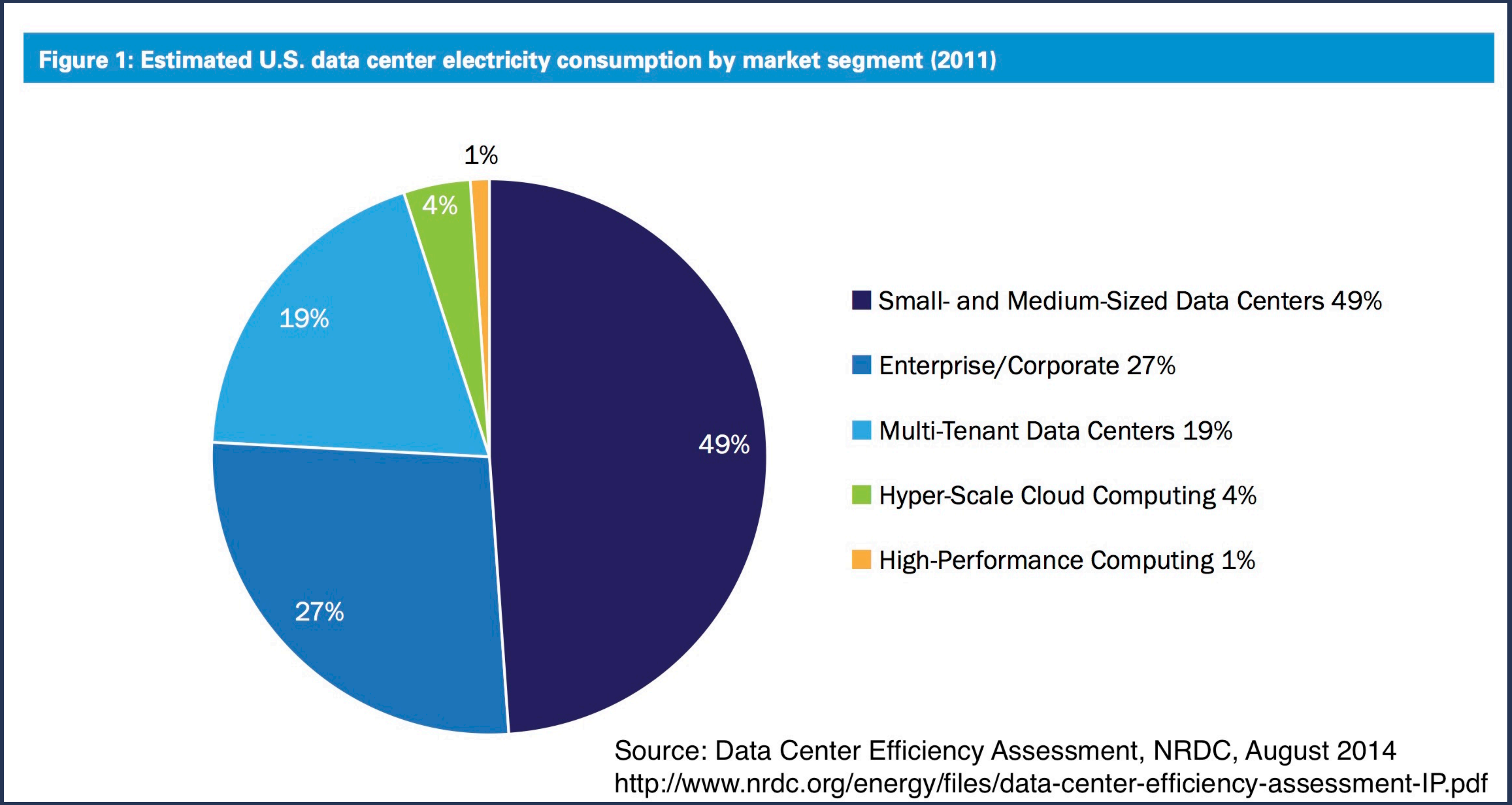
USA data centers in 2011

- Only 5% of DC-spent power is low PUE
- 40% of servers in small-to-medium DCs consume 49% of total electricity
  - PUE ~= 2.0, utilization: low as 10%
  - older servers (3 years old)

DC operating issues

- Too much over-provisioning (~ +50%)
- Low virtualization daployment rate (~30%)
- Too many *unused* servers (20~30%)
- Power management not well deployed

Small and inefficient data centers are the majority



The numbers

Source: Data Center Efficiency Assessment, NRDC, August 2014, Appendix 2
http://www.nrdc.org/energy/files/data-center-efficiency-assessment-IP.pdf

U.S. Data Center Segmentation Energy Use Methodology and Assumptions

Segment	% of stock (based on # of servers)	Average PUE	Average server utilization	Average server age (years)	2011 Electricity Use (MWh)	Server power at average utilization level (SPECpower_ssj2008) (watts)	DC market segmentation by electricity consumption
Small- to Medium-sized Data Centers	40%	2.0	10%	3	37,500,000	149	49%
Enterprise/ Corporate	30%	1.8	20%	2	20,500,000	120	27%
Multi-tenant Data Centers	22%	1.8	15%	2	14,100,000	113	19%
Hyper-scale Cloud Computing	7%	1.5	40%	1	3,300,000	101	4%
High-performance Computing	1%	1.8	50%	2	1,000,000	169	1%
	100%				**76,400,000**		**100%**

Issues of Japanese DCs

- Natural disasters (earthquakes)
- Fluctuating *industrial* power supply: not something solvable by "saving energy" in the *residential* sector
- The price of electricity is *very high*
- Japan is at an *edge* of world Internet trans-ocean links and have little direct oversea links

The future is grim

Can we sustain the level of the modern computing society?