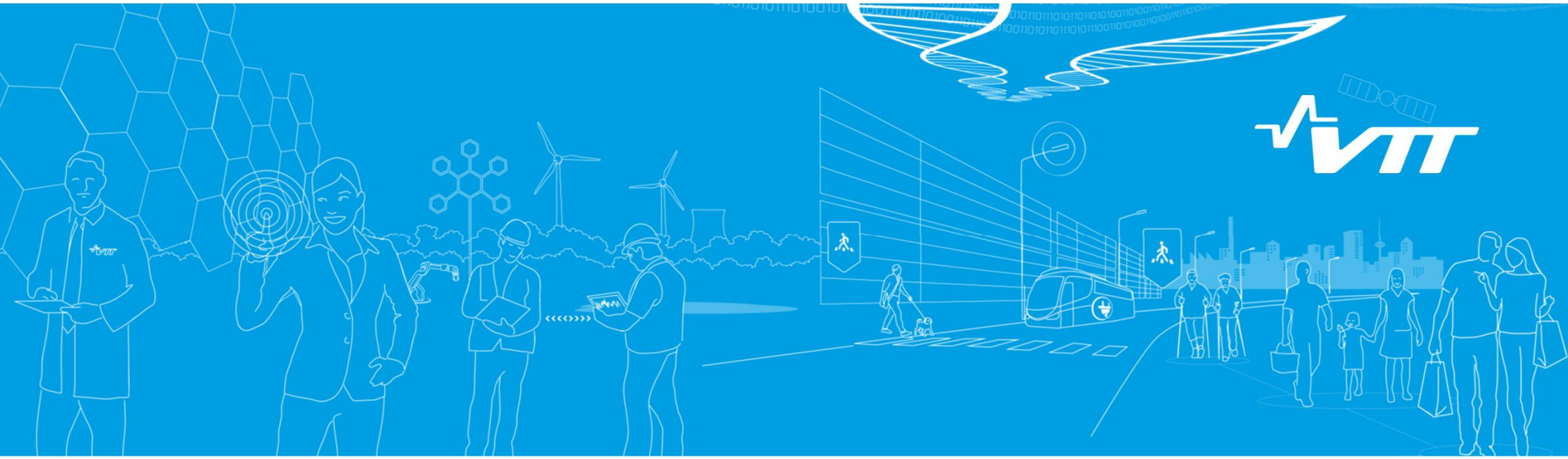
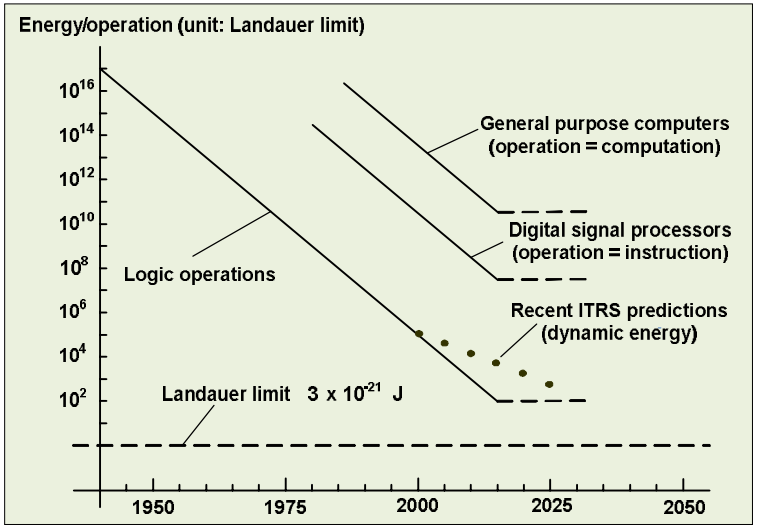


Energy efficiency in 5G networks

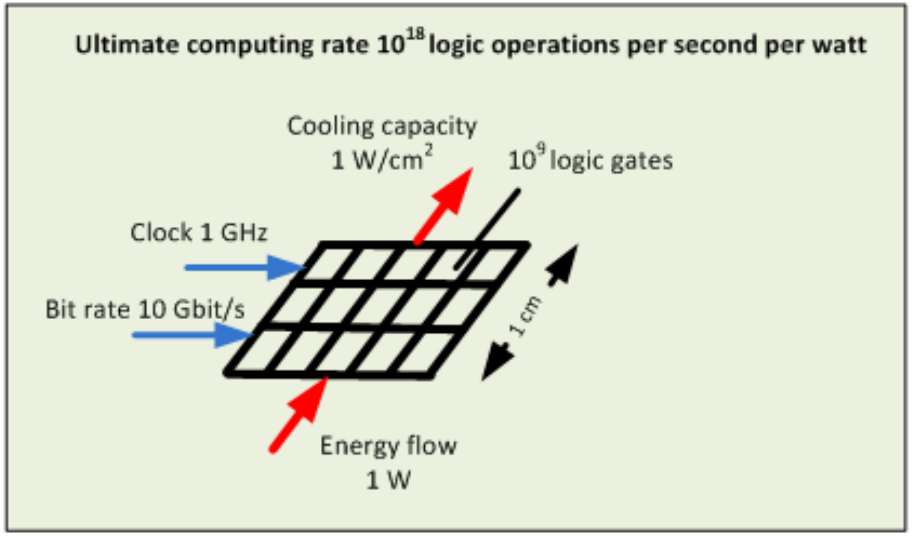


Aarne Mämmelä, aarne.mammela@vtt.fi
VTT Technical Research Centre of Finland

IFIP Networking 2015, Toulouse, France, 20.5.2015

Outline

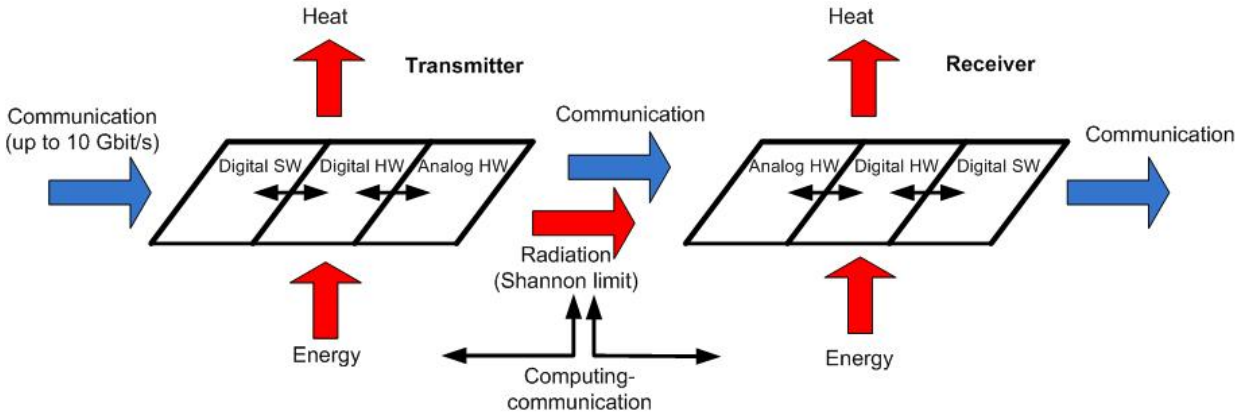
- Introduction
- KPIs and basic resources
- Trends in mobile data traffic
- Trends in power consumption
- Trends of energy efficiency
- Fundamental limits
- Solutions are trade-offs
- Summary
- Bibliography



Introduction

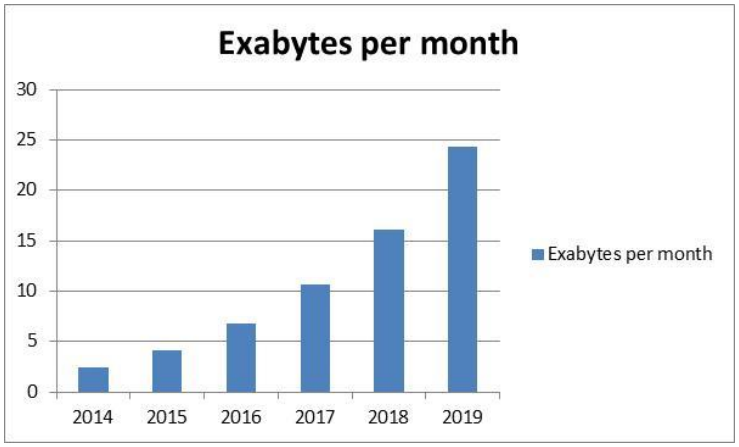
Introduction

- 5G systems will be deployed in about 2020 (Andrews 2014, Boccardi 2014).
- Technology evolution and fundamental limits in digital electronics (Keyes 2005, Zhirnov 2013, Markov 2014) will have a direct effect on all layers.
- A trend is towards small cells (Cox 2008, Kishiyama 2013, Liu 2014) where circuit power (computation) will become important in addition to transmission power (communication) leading to computation-communication trade-off (Lettieri 1999).
- The speaker's background is in the physical layer algorithms.



Importance of energy (ITU-R 2015)

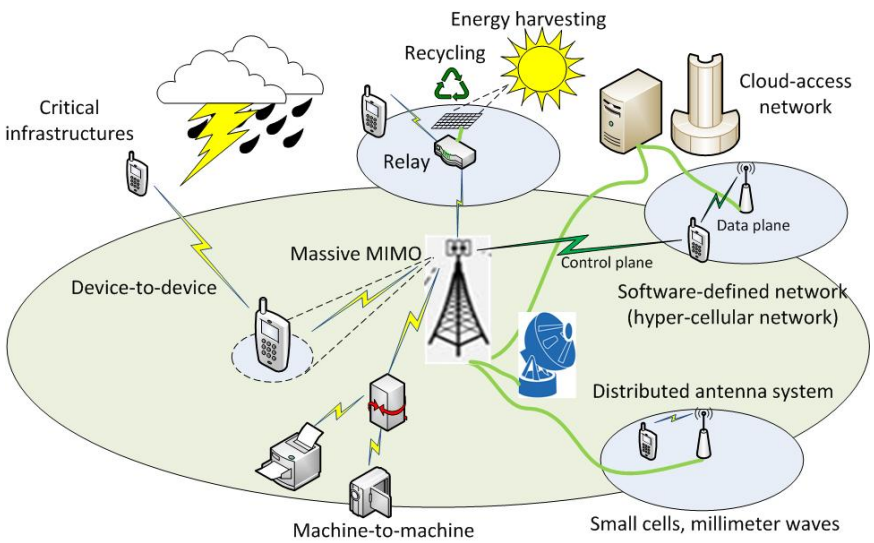
- Energy is expensive for operators and users and its production has environmental effects.
- Battery capacity is increasing only 1.5x/decade or 4 %/year.
- Mobile traffic is increasing exponentially (up to 1000x/decade) and power consumption should be adapted to the traffic load (Blume 2010).
- Link throughput requirements are increasing up to 10 Gbit/s.
- Moore's law (energy efficiency 100x/decade) is slowing down and will have thermal noise death by 2020-2030 (now 10x/decade).
- Cooling efficiency will not improve significantly, and active cooling is using energy.



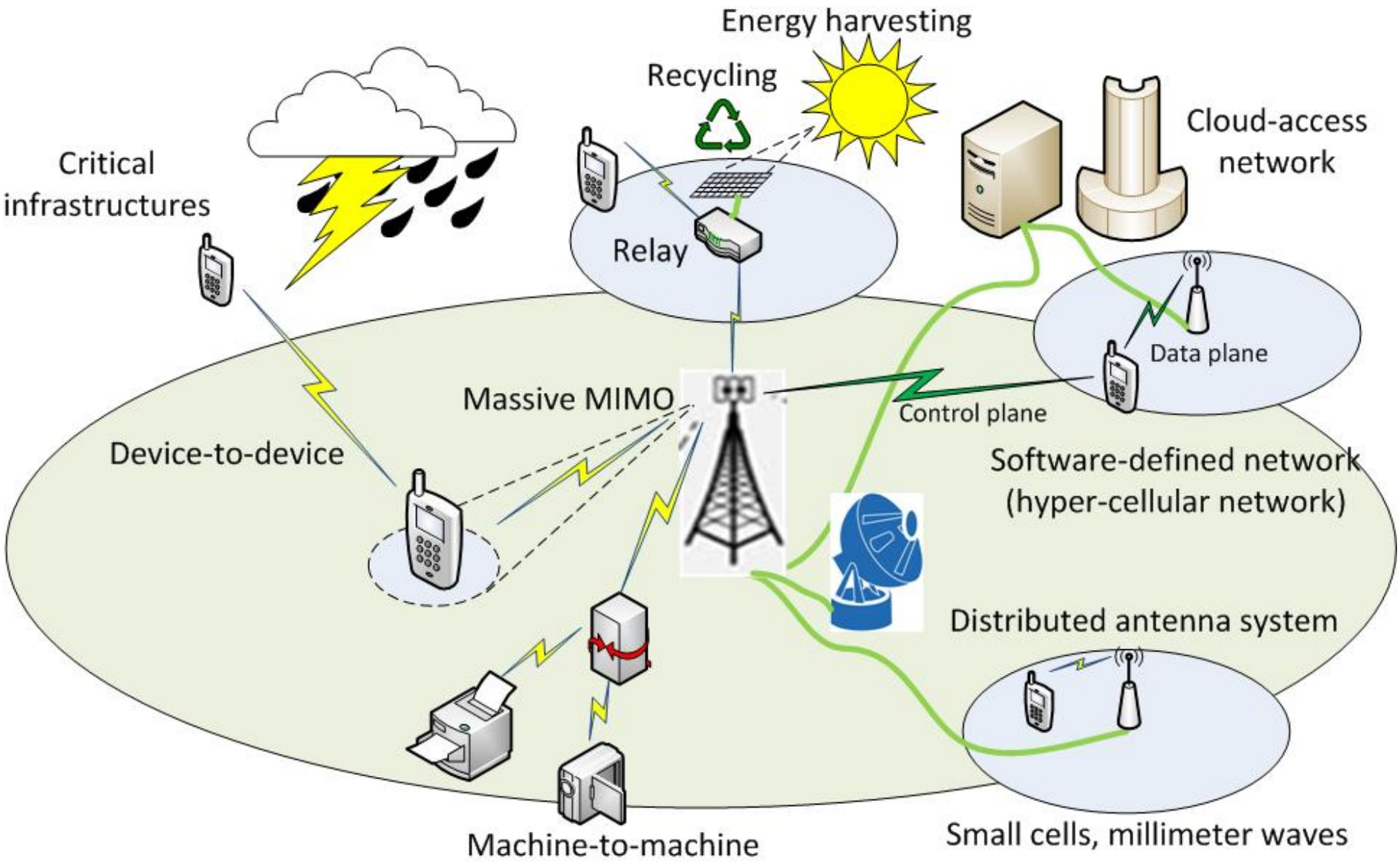
Global mobile traffic (Cisco 2015)

Revolutionary technologies for 5G (Andrews 2014, Boccardi 2014)

1. Small cells (increased area spectral efficiency) (Cooper's prediction, Chandrasekhar 2008)
   - Heterogeneous networks, microcells < 1 km, picocells < 100 m, femtocells < 10 m, energy harvesting networks
   - Software-defined network, separate data and control planes (phantom cell, hyper-cellular network, macro-assisted small cell)
   - Cloud radio access networks (centralized baseband), distributed antenna systems
   - Cooperative communication (coordinated multipoint, relaying)
2. Massive MIMO (increased area spectral efficiency)
   - Number of antennas much larger than the number of devices
3. Millimeter waves (larger bandwidth)
   - 30-300 GHz, wavelength 1-10 mm
4. Smart devices and device-centric connectivity
   - Distributed services
   - Device-to-device (D2D) connectivity
   - Local caching
   - Advanced interference rejection
5. Machine-to-machine (M2M) communications
   - Critical (e.g. vehicles) and massive (e.g. sensors) deployments
   - Minimal data rate with very high link reliability virtually all the time
   - Very low latency and real-time operation
   - Long battery life > 10-15 years



Revolutionary technologies for 5G



KPIs and basic resources

Key performance indicators (KPIs)

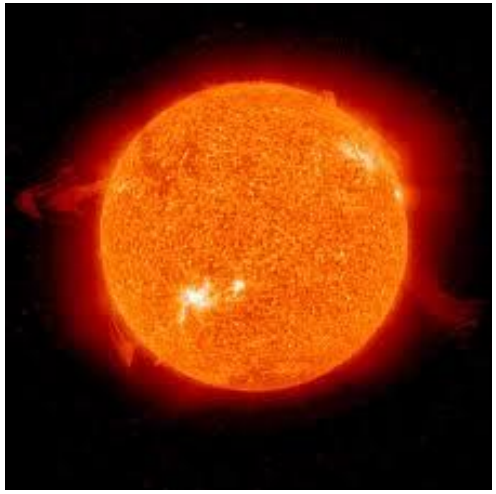
- Peak data rate, user experienced data rate
- Traffic volume density (sum bit rate/km^2)
- Number of connections/km^2
- End-to-end latency (ms)
- Battery life (3 days for smart phones, up to 10-15 years for machine-to-machine systems, NGMN 2015)
- Mobility (km/h)
- Reliability

ENERGY

Basic resources

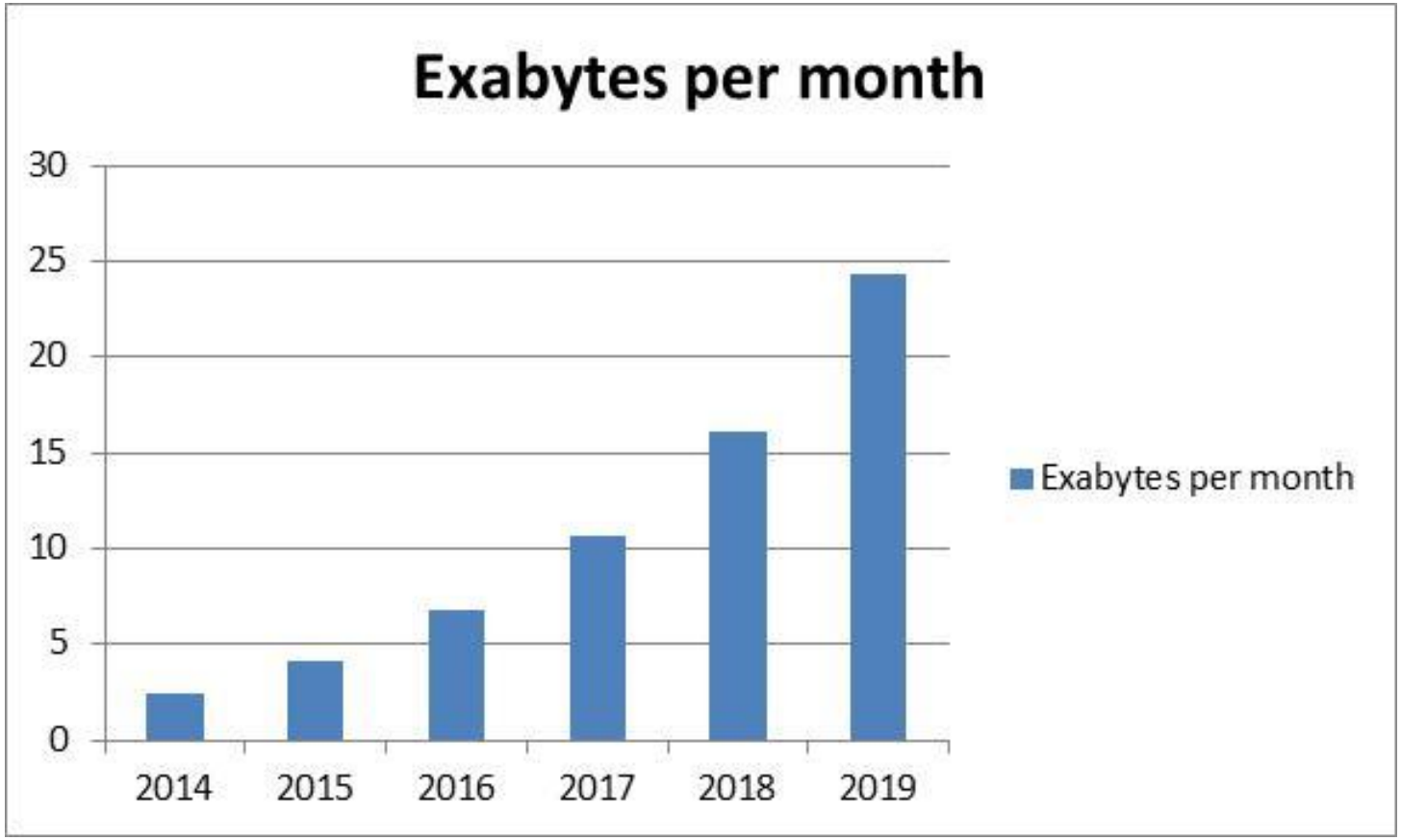
- Materials (silicon, copper, etc.)
- Energy (power) – computation-communication trade-off
- Time (delay) – diversity (redundancy), multiplexing
- Frequency (bandwidth) – diversity (redundancy), multiplexing
- Space (area, volume) – diversity (reduncancy), multiplexing
- Capital (cost) – human work

Materials Energy
Time Frequency Space
Capital



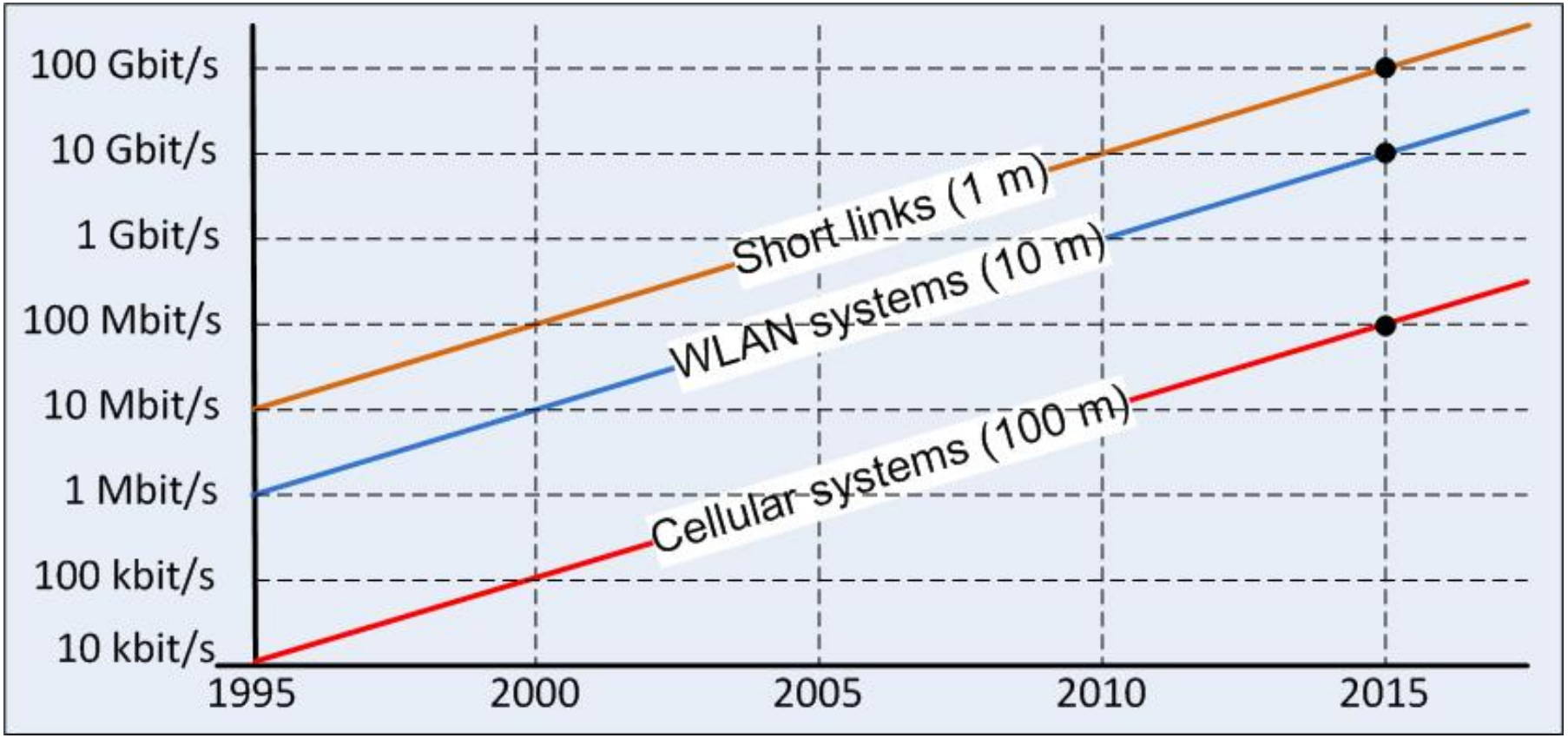
Trends in mobile data traffic

Global mobile traffic (Cisco 2015)



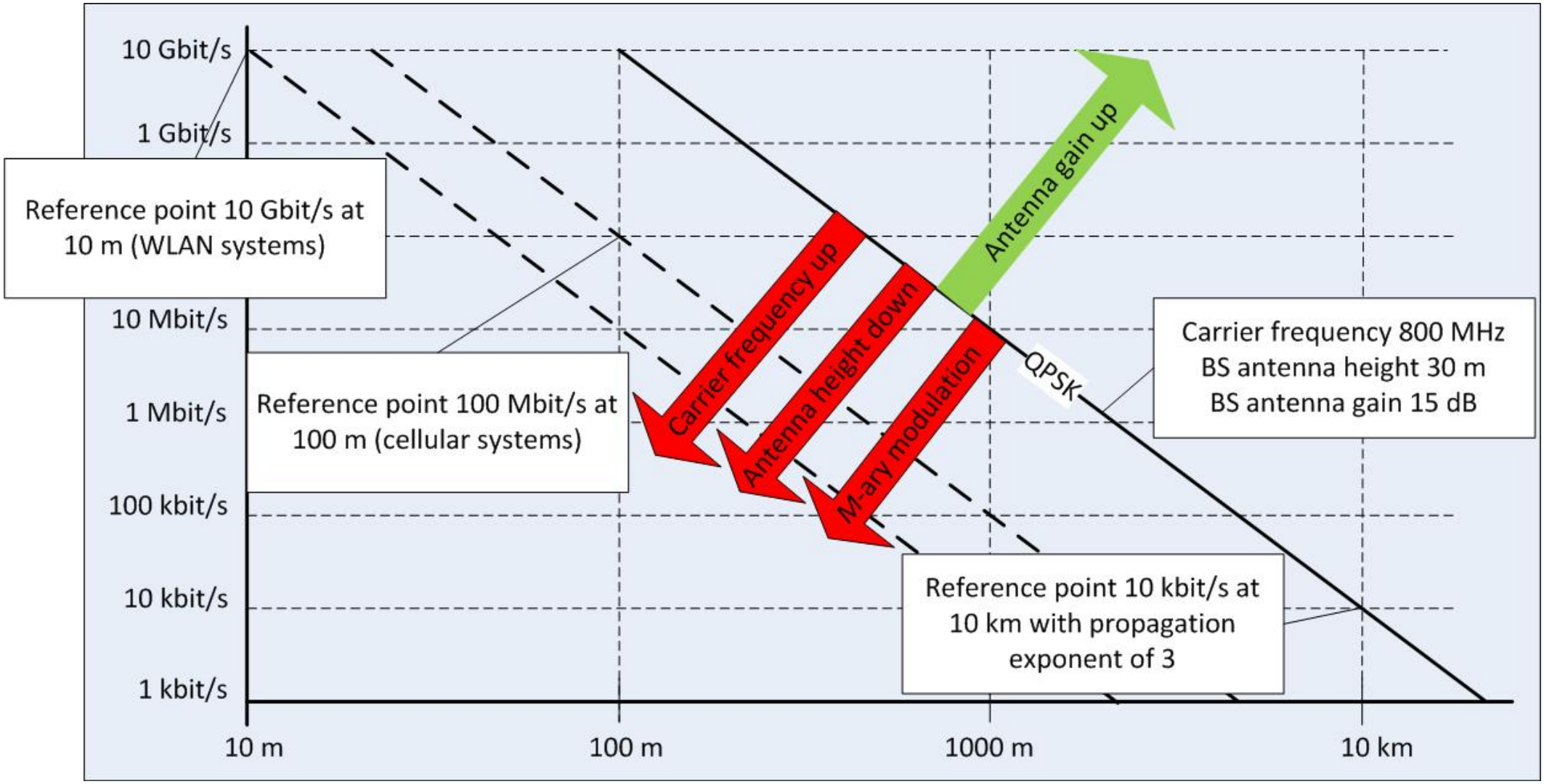
- Annual growth is 57 %, implying 100-fold growth in ten years (some companies expect 1000-fold growth).
- Majority of the traffic will be video traffic (Ericsson 2014)

Bit rates in wireless systems (Fettweis 2012)



- Bit rates at the same link distance have been increasing 100x/decade
- We use the year 2015 as a reference point (after this some saturation expected for constant power consumption)
  - 100 Mbit/s at 100 m
  - 10 Gbit/s at 10 m
  - 100 Gbit/s at 1 m

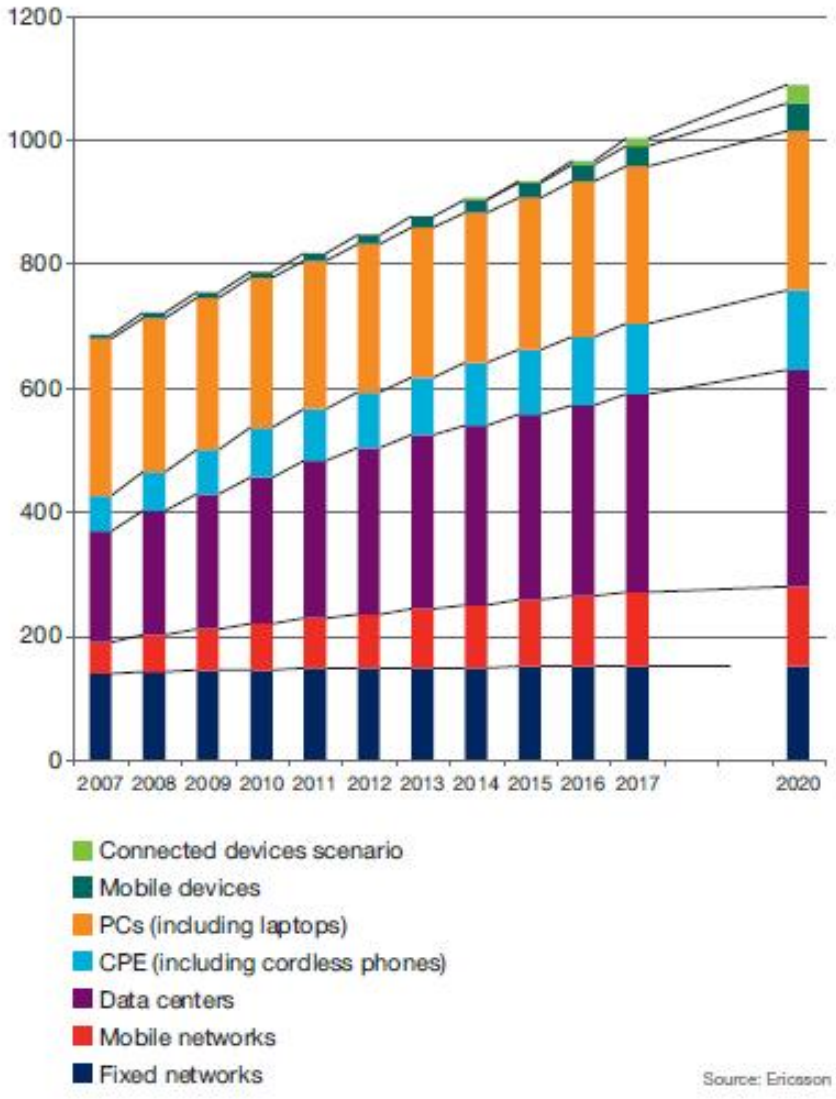
The bit rates vs. link distance (Cox 2008)



- Bandwidth is decreased with *M*-ary modulation when increasing *M*

Trends in power consumption

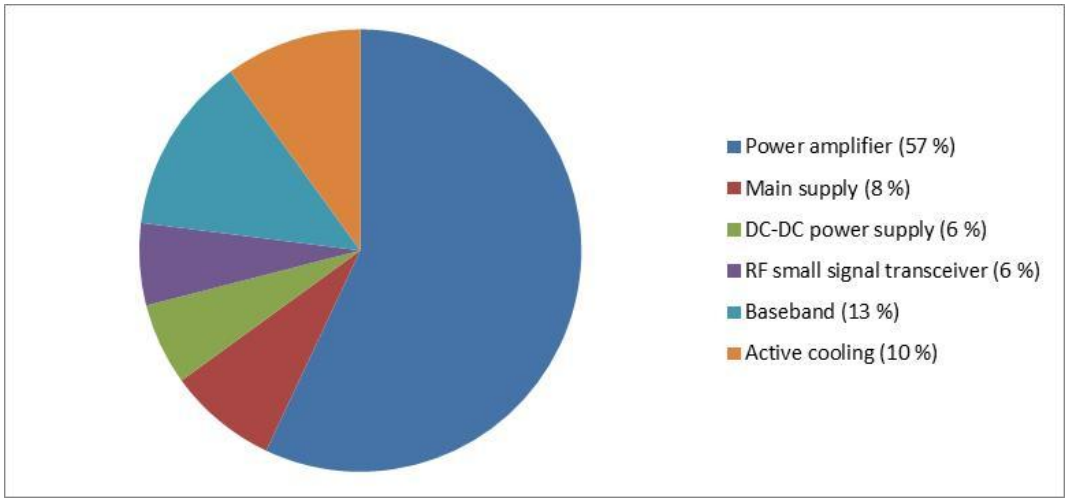
Trend in power consumption (TWh) (Ericsson 2013, used with permission of Ericsson)



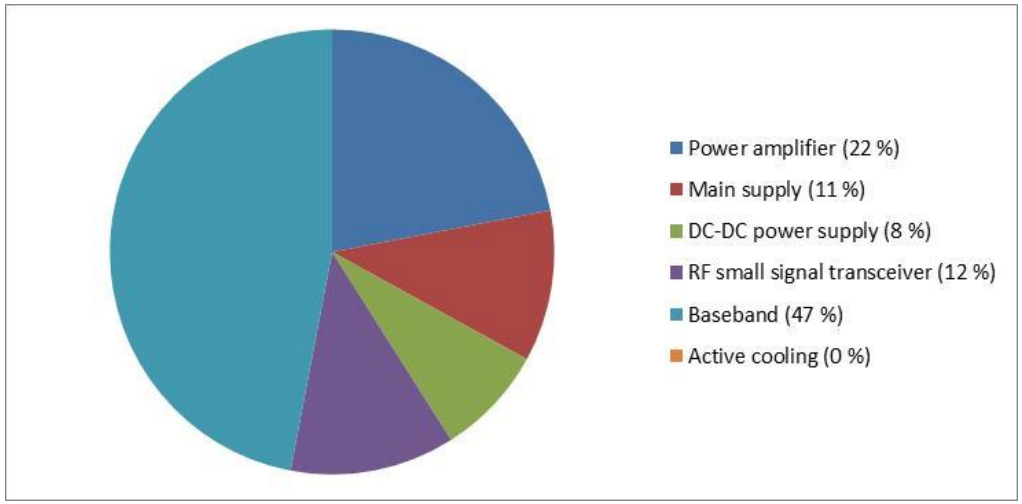
- Data centers will dominate the power consumption in the network due to cloud access networks
- PC, personal computer, CPE, customer premises equipment

Power consumption of various LTE BS types (%) (EARTH D2.3)

Macrocells > 1 km

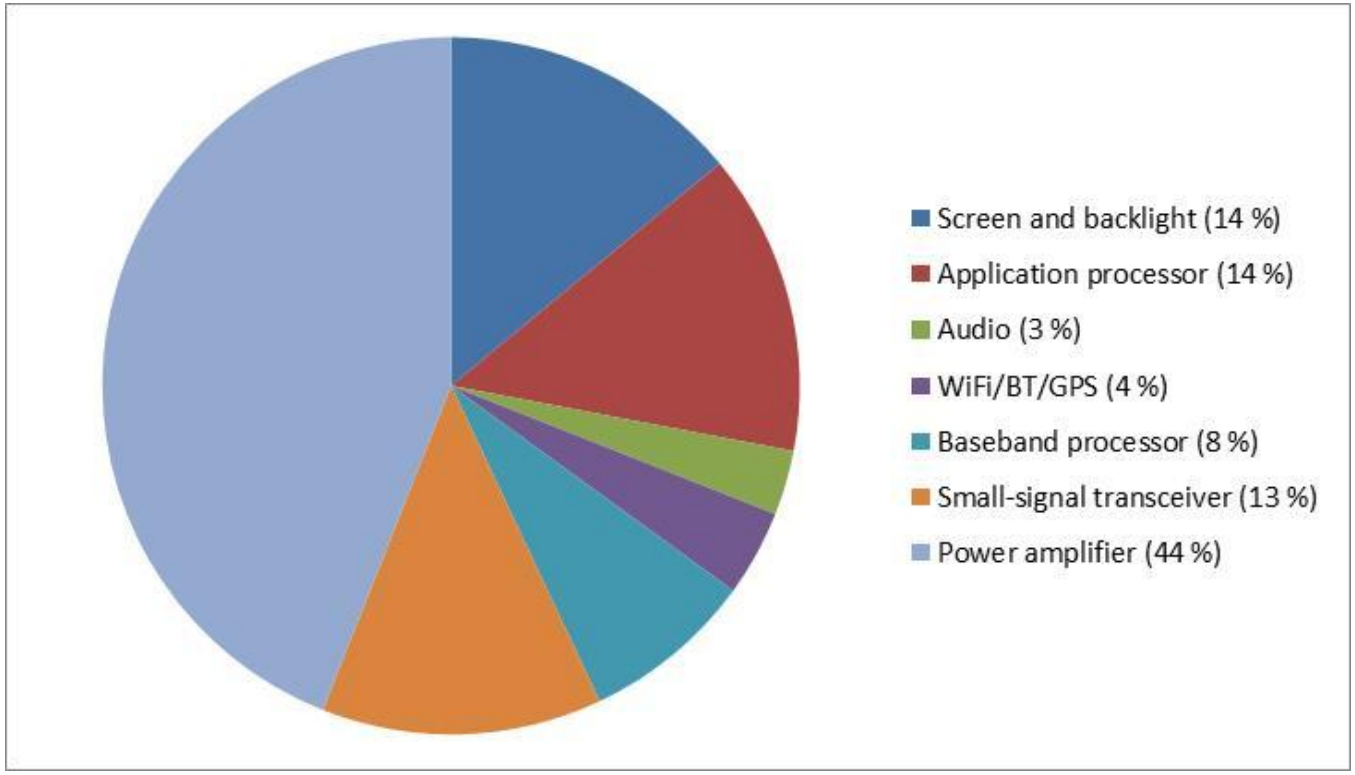


Femtocells < 10 m



- Baseband dominates in small cells because of computation-communication trade-off
  - In small cells the transmission power is reduced and circuit power starts to dominate
- Increasing carrier frequency will increase path loss and transmission power
- Increasing number of antennas will increase antenna gain and decrease transmission power but increase computation (circuit) power

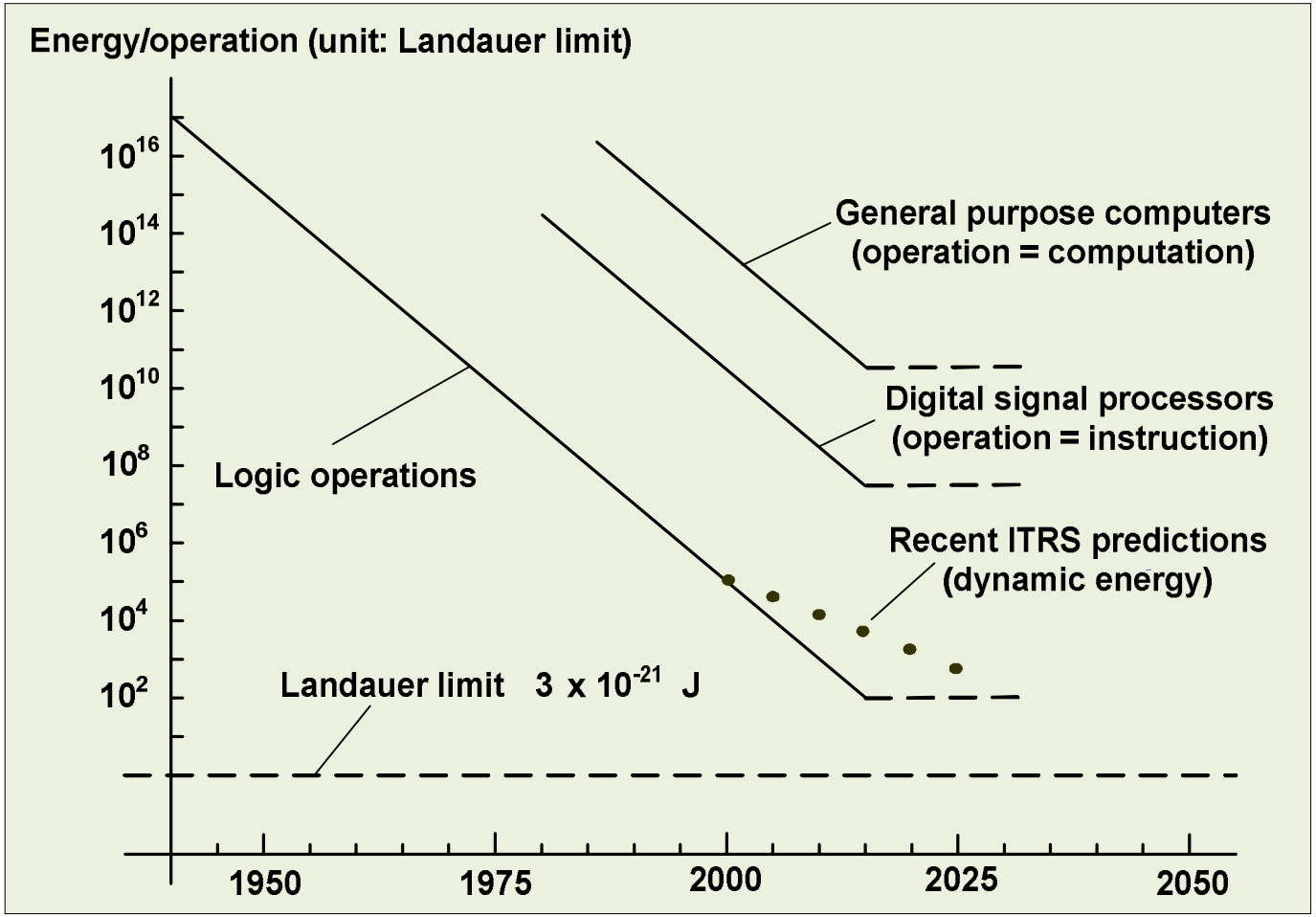
Power consumption in an LTE smart phone (%) (Wang 2014)



- Percentage of power amplifier is reduced at short links (max. power used here)
- Power control also reduces the average power amplifier power
- BT, Bluetooth, GPS, Global Positioning System

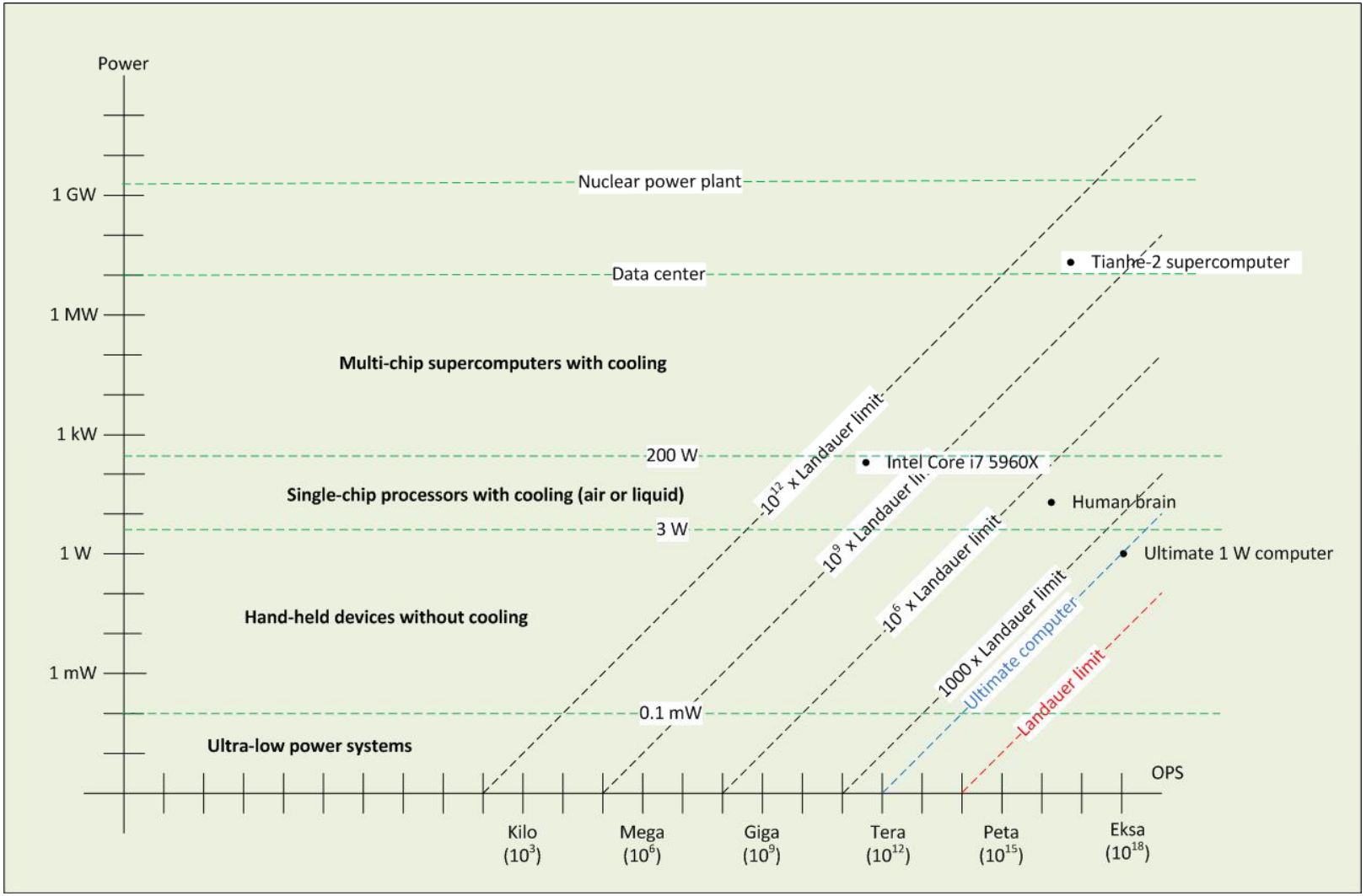
Trends in energy efficiency

Technology trends

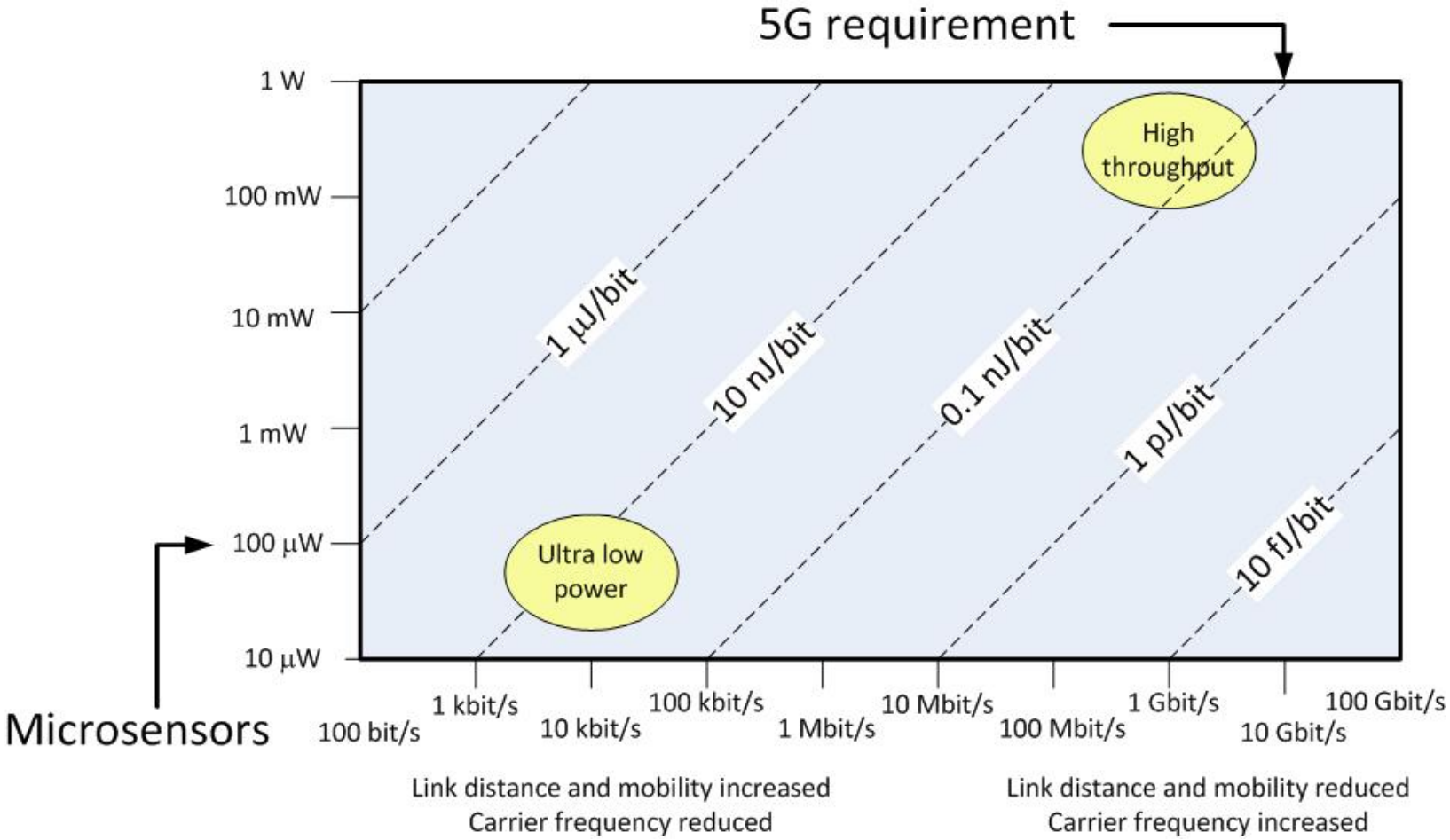


- Saturation of technology has started in about 2000 (Keyes 2005, ITRS 2012, Frantz 2012, Koomey 2011). Landauer limit is a noise limit for the switching energy identical to Shannon limit.

Power consumption vs. computing power



Power consumption vs. bit rate



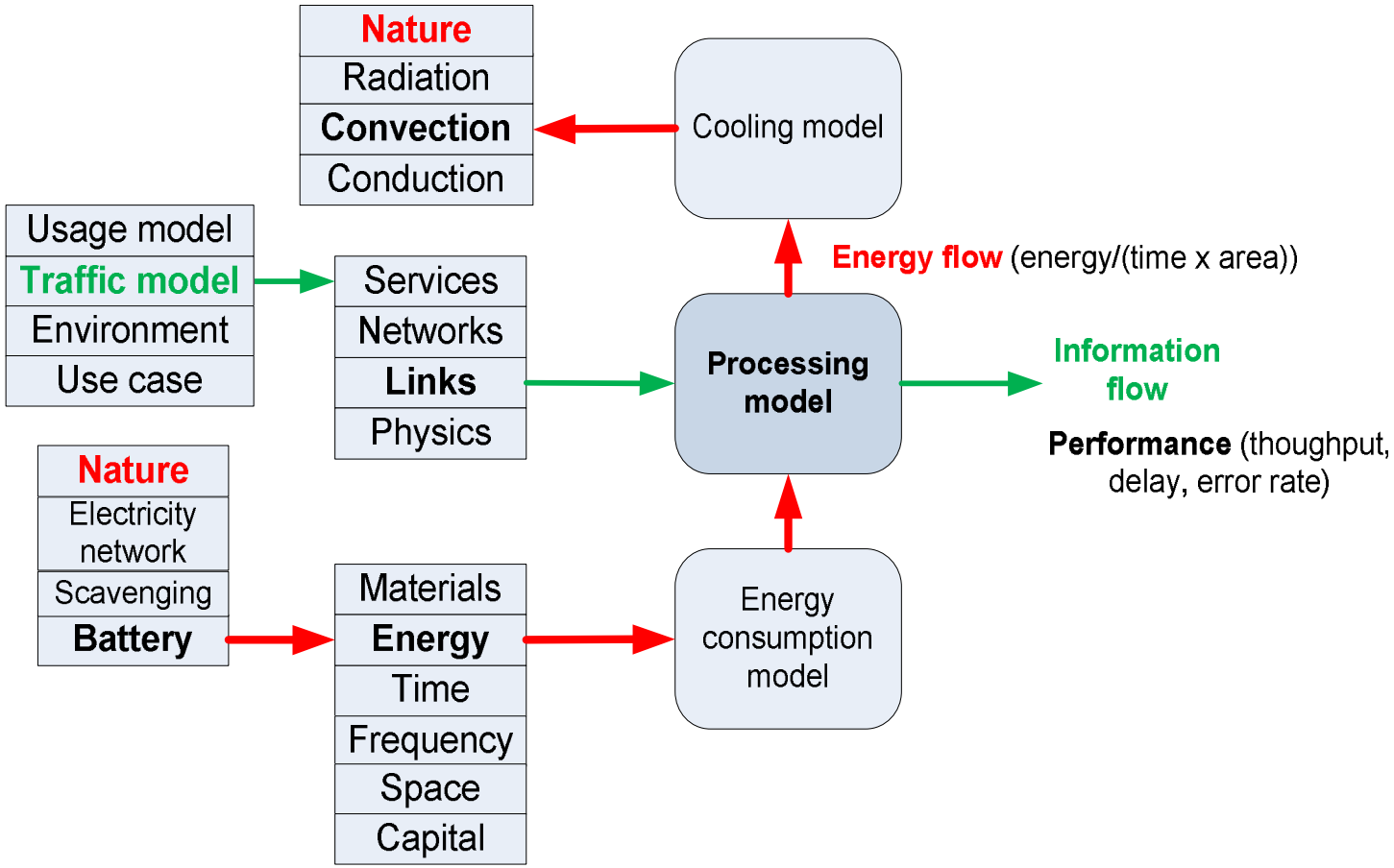
- One 16-bit multiplication with hardware using the technology of the 2020's
  - 3000 logic gates = 3 x 10^5 x Landauer limit = 1 fJ

Fundamental limits

Fundamental limits (Markov 2014)

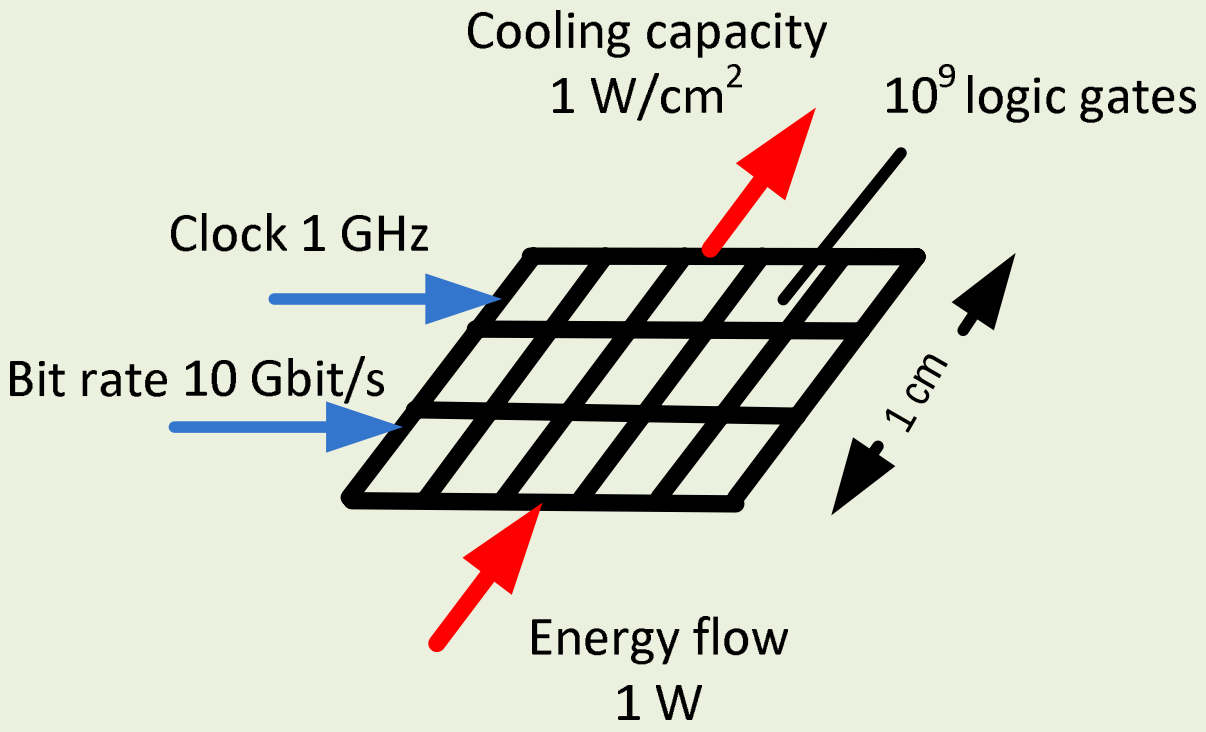
- Conservation of energy – cooling problems
- Entropy law – available energy reduced or kept constant, new energy needed
- Absolute zero – power density of thermal noise depends on absolute temperature (Zhirnov 2003)
- Upper velocity limit – propagation delays, delays of wires on a chip (Ho 2001)
- Uncertainty principle – lowest line width 4 nm (Zhirnov 2013)
- Channel capacity – maximum link spectral efficiency, Shannon limit for received energy/bit
- Energy limit for computation (circuit power) – Landauer limit for energy/irreversible logic operation
- Unprovable theorems – Gödel incompleteness theorem
- Unsolvable problems – Church-Turing conjecture, Turing machine never stops
- Intractable problems – Cook-Levin theorem, exponential complexity, Turing machine stops after an extremely long time
- Unpredictable deterministic systems – chaotic systems
- Maximum speed-up in parallel processing – not linear, Amdahl's and Gustafson's laws, energy efficiency in general reduced in parallel processing
- Resolution of optical imaging instruments – Abbe diffraction limit

Systems approach, emphasizing the information and energy flows



Ultimate computer

Ultimate computing rate 10^{18} logic operations per second per watt

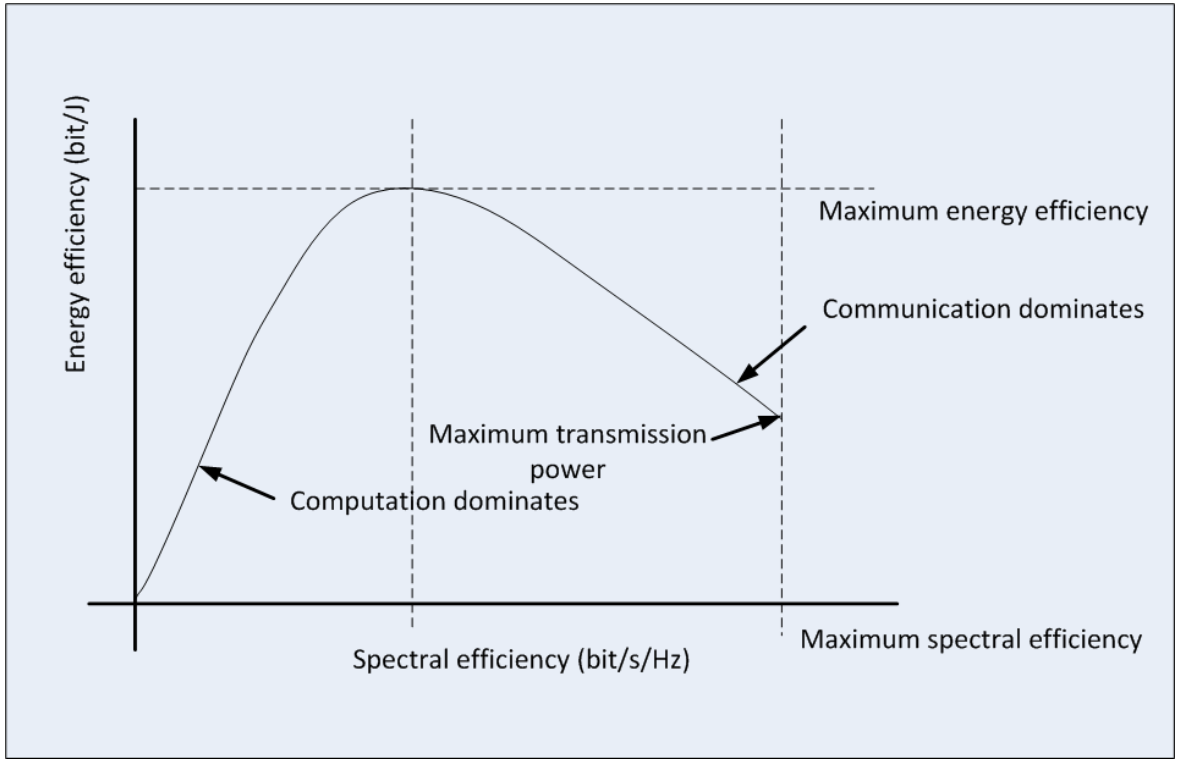


- Ultimate computer has an energy efficiency (energy/logic operation) of 100 x Landauer limit

Solutions are trade-offs

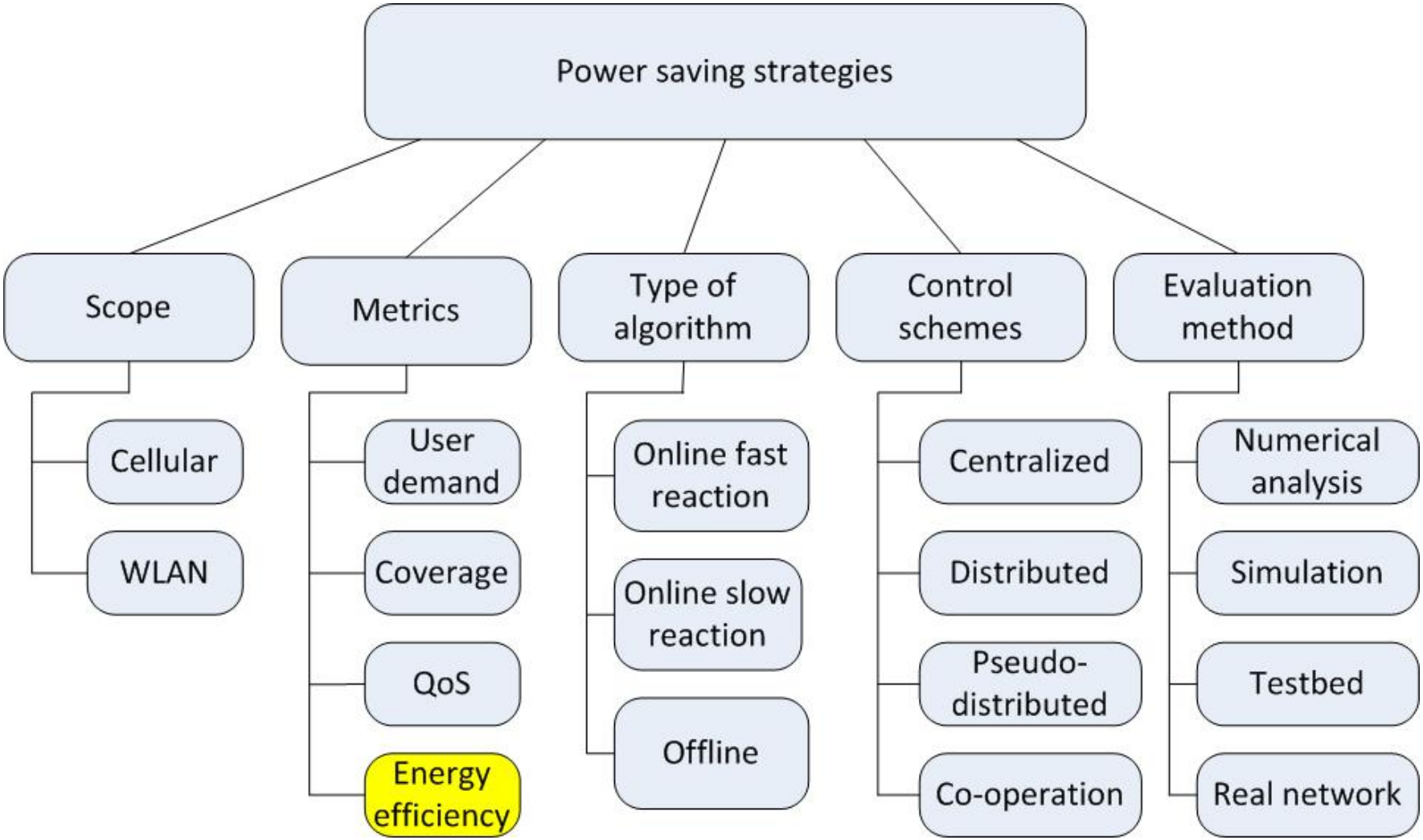
Why trade-offs?

- We are approaching fundamental and practical limits. We must make a trade-off, for example either spectral efficiency or energy efficiency is maximized but not both (Chen 2011, Deng 2013). Valid for uplink (terminals) and downlink (base stations).



- Energy efficiency (bit/J) is reduced at low and high spectral efficiency (bit/s/Hz)

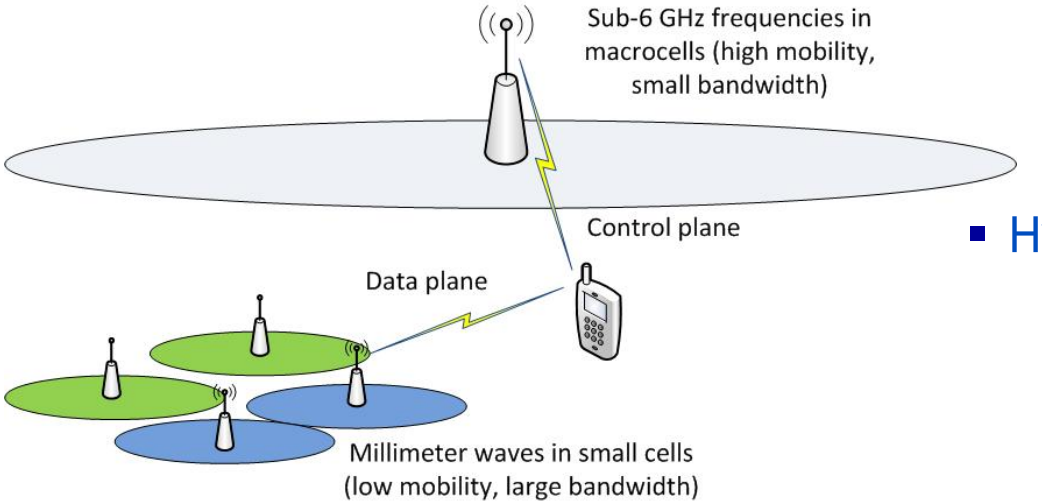
Taxonomy of power saving strategies (Budzisz 2014)



- Effect to the whole system must be considered, several methods to be combined (Alagoz 2011, Mittag 2011, Calhoun 2012)

Optimization methods

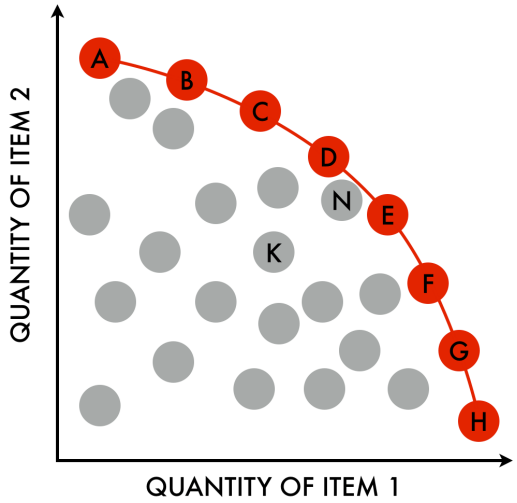
- Single-objective (single-criterion) optimization: either spectral efficiency (bit/s/Hz) or energy efficiency (bits/J) is maximimized to obtain better average performance
  - Hyper-cellular networks (phantom cells) (Kishiyama 2013, Liu 2014): data plane and control plane are separated (sleep modes in small cells used, Blume 2010)
- Multi-objective (multi-criteria) optimization (Marler 2004, Chen 2011, Deng 2013)
  - Near fundamental limits approaching peak performance needs joint optimization of both spectral efficiency and energy efficiency implying the use of trade-offs since the optimum is not unique, for example cross-layer design



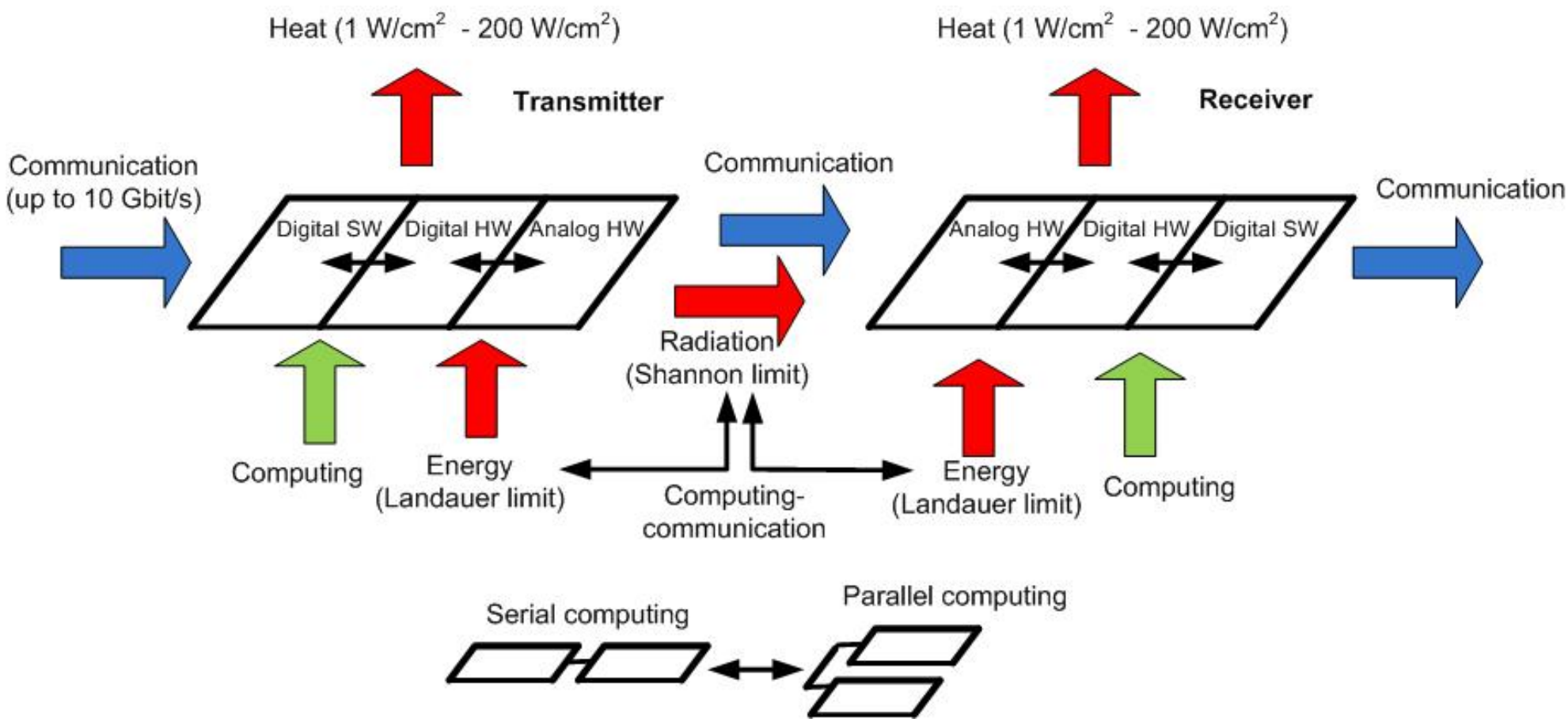
- Hyper-cellular network (phantom cells)

Trade-offs

- Performance-energy trade-off (limited battery capacity)
  - Energy efficiency – spectral efficiency trade-off (Shannon capacity)
  - Energy-space-time trade-off (cooling problems)
  - Communication-computing trade-off (Shannon and Landauer limits)
  - Precision-bandwidth trade-off (limited word length)
  - Analog-digital trade-off (part of precision-bandwidth trade-off)
  - Software-hardware trade-off (performance-flexibility trade-off)
  - Serial processing – parallel processing trade-off (limited clock rate)
- Performance-delay trade-off (spatially distributed systems)
- Communication-control trade-off
  - Centralized control – distributed control (cognitive radios)
  - Centralized computing – distributed computing (clouds)
- Order-chaos trade-off (interacting network elements, transmitter power control)

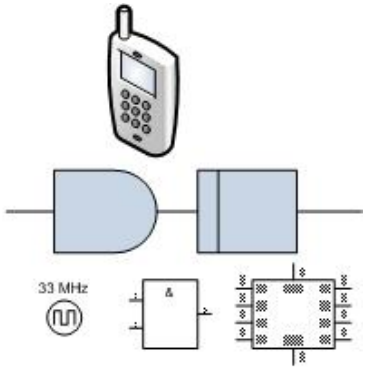


Trade-offs


Heat (1 W/cm^2 - 200 W/cm^2)
Transmitter
Heat (1 W/cm^2 - 200 W/cm^2)
Receiver
Communication
(up to 10 Gbit/s)
Digital SW
Digital HW
Analog HW
Communication
Radiation
(Shannon limit)
Analog HW
Digital HW
Digital SW
Communication
Computing
Energy
(Landauer limit)
Computing-
communication
Energy
(Landauer limit)
Computing
Serial computing
Parallel computing

Physical level exists at all layers

Service layer	Functional level
	Structural level
	Physical level
Network layer	Functional level
	Structural level
	Physical level
Physical layer	Functional level
	Structural level
	Physical level



- Each OSI layer includes three decription levels
  - Functional level (describes what the layer should do)
  - Structural (architectural) level (describes what parts are needed and how they are connected)
  - Physical level (physical implementation of the parts, inc. processors, memory, etc.)

Summary

Materials Energy
Time Frequency Space
Capital

Summary

- Future designs will be resource-limited and trade-offs are needed for peak performance
- Moore's exponential law has been slowing down since about 2000 and it will eventually undergo a thermal noise death in about 2020-2030.
- The switching energy of CMOS electronics will approach the level of 100 x thermal noise spectral density.
- Future networks are expected to carry 1000 times more mobile data in ten years, but the energy efficiency is improving only 10 times in ten years.
- The technology trends and fundamental limits in the digital electronics will have a direct effect on all layers.
- Networks do not only consume energy in transmission in power amplifiers but also in computation (circuit power) in algorithms and protocols.
- Computation energy will be more important in small cells that are expected to be used in places where user density is high.
- Although peak energy efficiency will saturate, the average energy efficiency may still grow (consumed energy should be proportional to the number of transmitted bits using sleep modes).

Bibliography

Bibliography

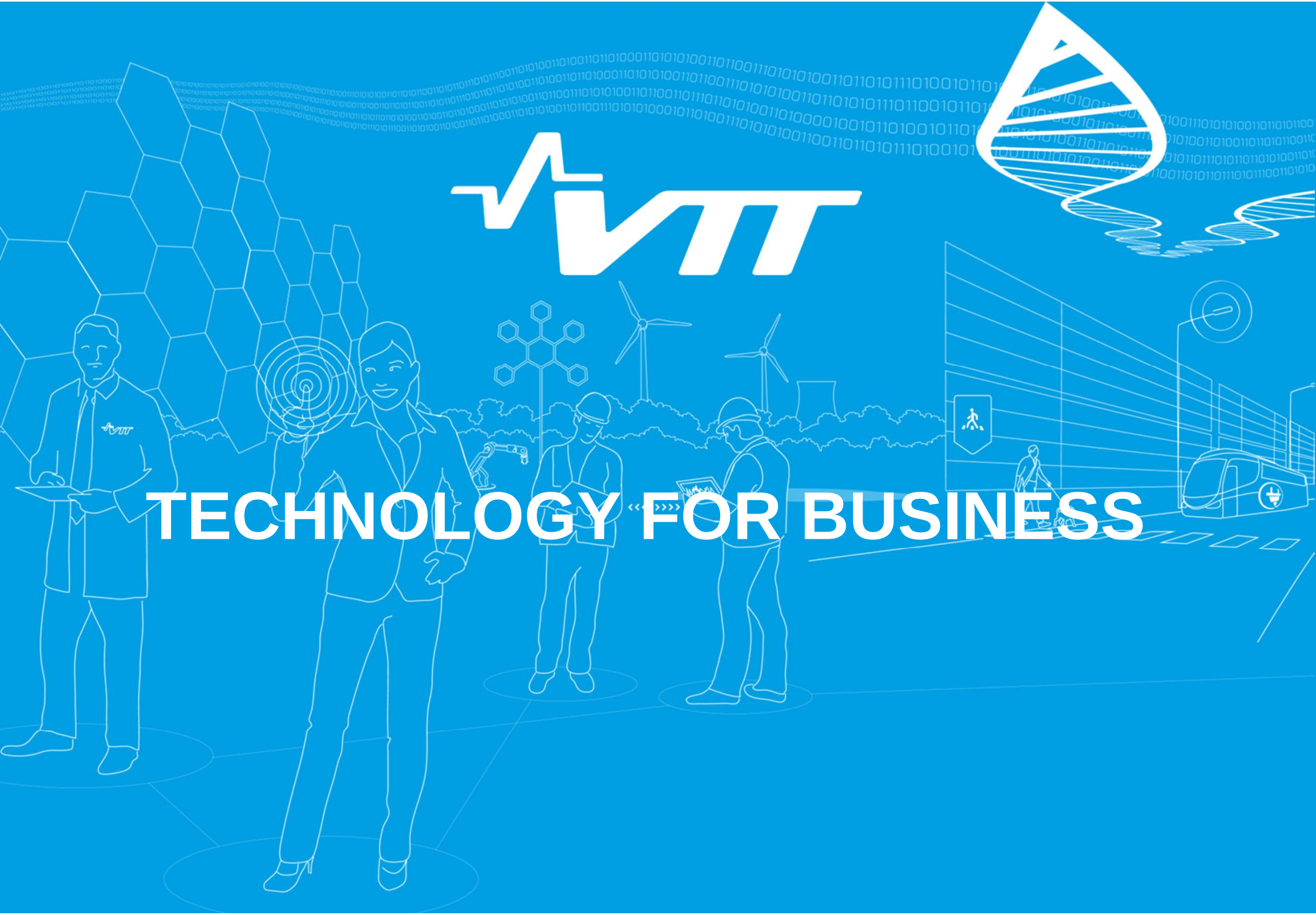
- 5G PPP, 5G Vision: The 5G Infrastructure Public Private Partnership: The next generation of communication networks and services, February 2015.
- F. Alagoz and G. Gur, “Energy efficiency and satellite networking: A holistic overview,” Proceedings of the IEEE, vol. 99, no. 11, 2011, pp. 1954 – 1979.
- J. G. Andrews et al. “What will 5G be?” IEEE Journal on Selected Areas in Communications, vol. 32, pp. 1065 – 1082, no. 6, 2014.
- E. Björnson et al., “Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?” IEEE Transactions on Wireless Communications, to be published.
- O. Blume et al., “Energy savings in mobile networks based on adaptation to traffic statistics,” Bell Labs Technical Journal, vol. 52, pp. 77-94, no. 2, 2010.
- B. Boccardi et al., “Five disruptive technology directions for 5G,” IEEE Communications Magazine, pp. 74-80, February 2014.
- China IMT-2020, IMT Vision towards 2020 and Beyond, IMT-2020 (5G) Promotion Group, February 2014.
- L. Budzisz et al., “Dynamic resource provisioning for energy efficiency in wireless access networks: A survey and an outlook,” IEEE Communications Surveys & Tutorials, vol. 16, no. 4, 2014 , pp. 2259 – 2285.
- B. H. Calhoun et al., “Body sensor networks: A holistic approach from silicon to users,” Proceedings of the IEEE, vol. 100, pp. 91- 106, no. 1, 2012.
- V. Chandrasekhar et al., “Femtocell networks: A survey,” IEEE Communications Magazine, pp. 59-67, September 2008.
- Y. Chen et al., “Fundamental trade-offs on green wireless networks,” IEEE Communications Magazine, pp. 30-37, June 2011.
- Cisco , Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014–2019, White Paper, February 2015.
- D. C. Cox and H. Lee, “Physical relationships,” IEEE Microwave Magazine, vol. 9, no. 4, August 2008, pp. 89 – 94.
- L. Deng et al., “A unified energy efficiency and spectral efficiency tradeoff metric in wireless networks,“ IEEE Communications Letters, vol. 17, pp. 55 – 58, no. 1, 2013.
- EARTH, Deliverable D2.3, Energy efficiency analysis of the reference systems, areas of improvements and target breakdown, version 2.00, January 31, 2012, http://www.ict-earth.eu.

Bibliography

- Ericsson Energy and Carbon Report, June 2013.
- Ericsson Mobility Report, June 2014.
- G. Fettweis, “A 5G wireless communications vision,” Microwave Journal, vol. 55, pp. 24-36, December 2012.
- G. Frantz, “Signal core: A short history of the digital signal processor,” IEEE Solid-State Circuits Magazine, vol. 4, pp. 16–20, Spring 2012.
- Z. Hasan et al., “Green cellular networks: A survey, some research issues and challenges,” IEEE Communications Surveys & Tutorials, vol. 13, pp. 524-540, no. 4, Fourth Quarter 2011.
- R. Ho et al., “The future of wires,” Proceedings of the IEEE, vol. 89, pp. 490-504, April 2001.
- ITU-R, Future Technology trends of terrestrial IMT systems, Report ITU-R M.2320-0, Geneva 2015.
- ITRS 2012, www.itrs.net.
- R. W. Keyes, “Physical limits of silicon transistors and circuits,” Reports on Progress in Physics, vol. 68, pp. 2701–2746, Dec. 2005.
- Y. Kishiyama et al., “Future steps of LTE-A: Evolution towards integration of local area and wide area systems,” IEEE Wireless Commun., vol. 20, pp. 12–18, February 2013.
- J. G. Koomey et al., “Implications of historical trends in the electrical efficiency of computing,” IEEE Annals of the History of Computing, vol. 33, pp. 46–54, 2011.
- P. Lettieri and M. B. Srivastava, “Advances in wireless terminals,” IEEE Personal Communications, vol. 6, pp. 6 – 19, no. 1, 1999.
- J. Liu et al., “CONCERT: A cloud-based architecture for next-generation cellular systems,” IEEE Wireless Communications, pp. 14-22, December 2014.
- I. L. Markov, “Limits on fundamental limits to computation,” Nature, vol. 512, pp. 147–154, 14 August 2014.
- R. T. Marler ad J. S. Arora, “Survey of multi-objective optimization methods for engineering,” Struct. Multidisc. Optim., vol. 26, pp. 369-395, 2004.
- R. Min et al., “Energy-centric enabling technologies for wireless sensor networks,” IEEE Wireless Communications, August 2002, pp. 28-39.
- J. Mittag, “Enabling accurate cross-layer PHY/MAC/NET simulation studies of vehicular communication networks,” Proceedings of the IEEE, vol. 99, pp. 1311 – 1326, no. 7, 2011.
- NGMN, 5G White Paper, NGMN Alliance, 17 February 2015.

Bibliography

- Nokia Networks, 5G radio access: System design aspects, White Paper, Nokia Networks, 2015.
- NSN, Flatten network energy consumption, Technology Vision 2020, White Paper, December 2013.
- A. Osseiran et al., “Scenarios for 5G mobile and wireless communications: The vision of the METIS project,” IEEE Communications Magazine, vol. 52, pp. 26 – 35, no. 5, 2014.
- E. Perahia, “IEEE 802.11n development: History, process, and technology,” IEEE Communications Magazine, pp. 48-55, July 2008.
- G. Piro et al., “NetNets powered by renewable energy sources,” IEEE Internet Computing , vol. 17, pp. 32-39, January/February 2013.
- J. M. Rabaey et al., “PicoRadio supports ad hoc ultra-low power wireless networking,” Computer, vol. 33, no. 7, pp. 42 – 48, 2000.
- M. Z. Shafiq et al., “Large-scale measurement and characterization of cellular machine-to-machine traffic,” IEEE/ACM Transactions on Networking,, vol. 21, pp. 1960 – 1973, no. 6, 2013.
- L. Suarez et al., “An overview and classification of research approaches in green wireless networks,” EURASIP Journal on Wireless Communication and Networking, 2012:142, pp. 1-18.
- S. Sudevalayam and P. Kulkarni, “Energy harvesting sensor nodes: Survey and implications,” IEEE Communications Surveys & Tutorials, vol. 13, pp. 443-461, Third Quarter 2011.
- T. Wang and L. Vandendorpe, “On the optimum energy efficiency for flat-fading channels with rate-dependent circuit power,” IEEE Transactions on Communications, vol. 61, pp. 4910-4921, December 2013.
- B. Wang et al., “End-to-end power estimation for heterogeneous cellular LTE SoCs in early design phases,” in Proc. PATMOS’14.
- V. V. Zhirnov et al., “Limits to binary logic switch scaling-A Gedanken model,” Proc. IEEE, vol. 91, pp. 1934–1939, Nov. 2003.
- V. V. Zhirnov and R. K. Cavin, “Future microsystems for information processing: Limits and lessons from living systems,” IEEE Journal of the Electron Device Society, vol. 1, pp. 29-47, February 2013.


VTT
TECHNOLOGY FOR BUSINESS