Scaling in Biology and Computation

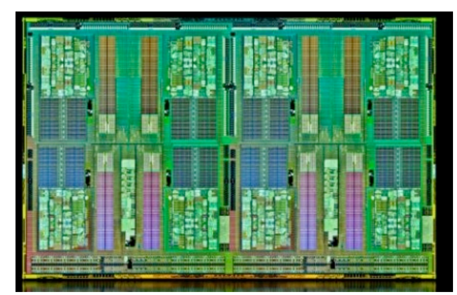
Melanie Moses
Department of Computer Science
University of New Mexico

SFI Complex Systems Summer School
June 20, 2017

How do complex systems get big?

- Today
  - Scaling in Biology
  - Scaling in Computation
  - Can time-energy minimization explain ubiquitous patterns in both domains?
- Tomorrow
  - How to build a scalable biologically-inspired swarm of robots

Intel 4004



AMD Opteron

Scaling in Biology

A whale is

100 000 000 000 000 000 000

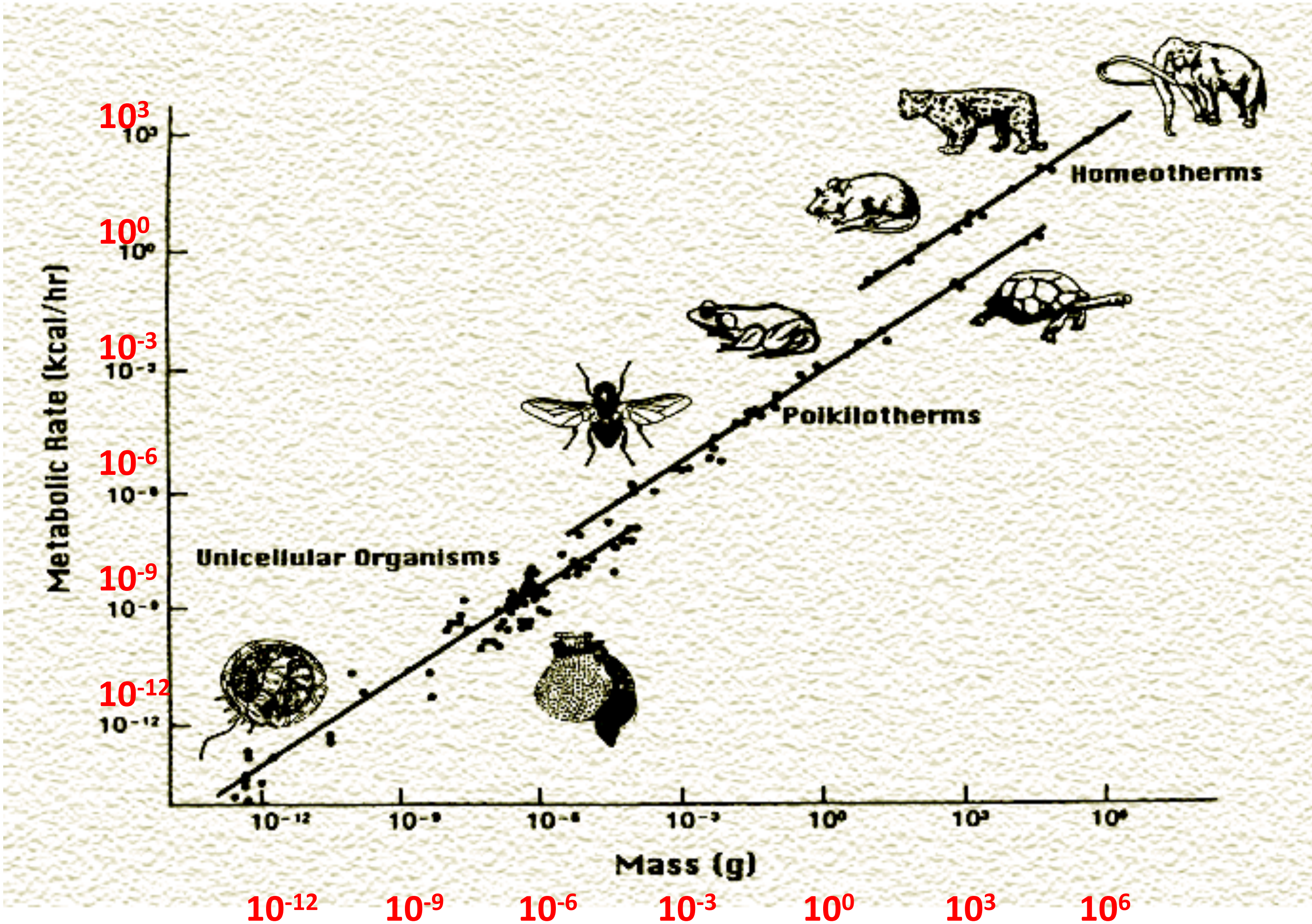
times bigger than an E. coli

10^{-12} g

10^{8} g

Glaw et al PLoS ONE 2012

Metabolic Scaling
A striking universal(?) pattern



Hemmingson 1960

Analyzing Scaling Relationships

The scaling exponent is the slope on log-log plot

$\log(B) = ¾ \log(M) + \log(c)$

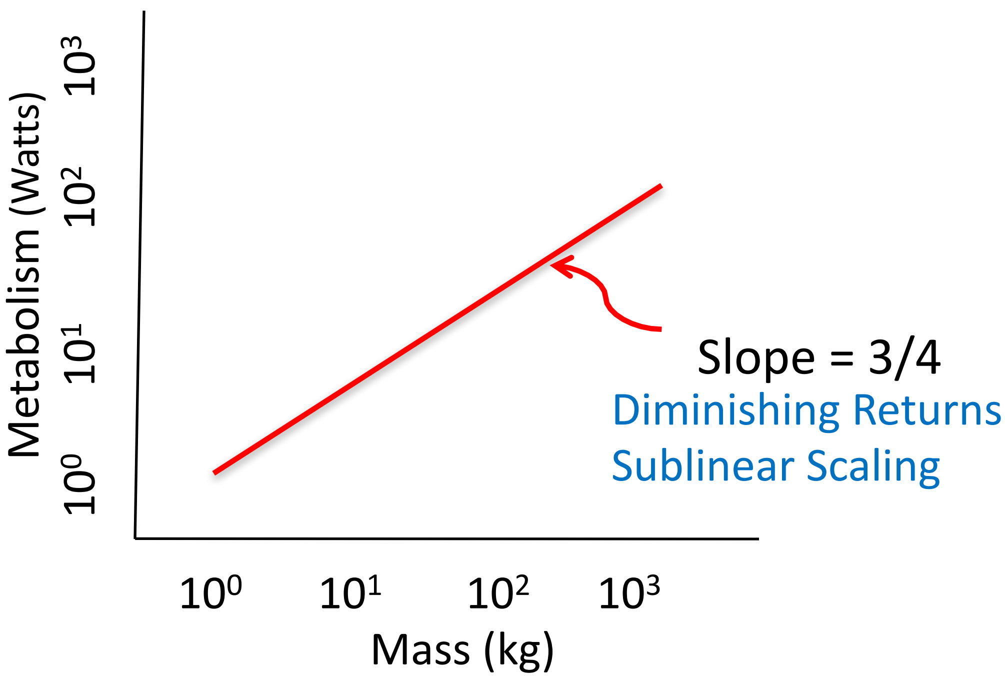
$y = mx + b$

The intercept is log(c)

The slope is ¾

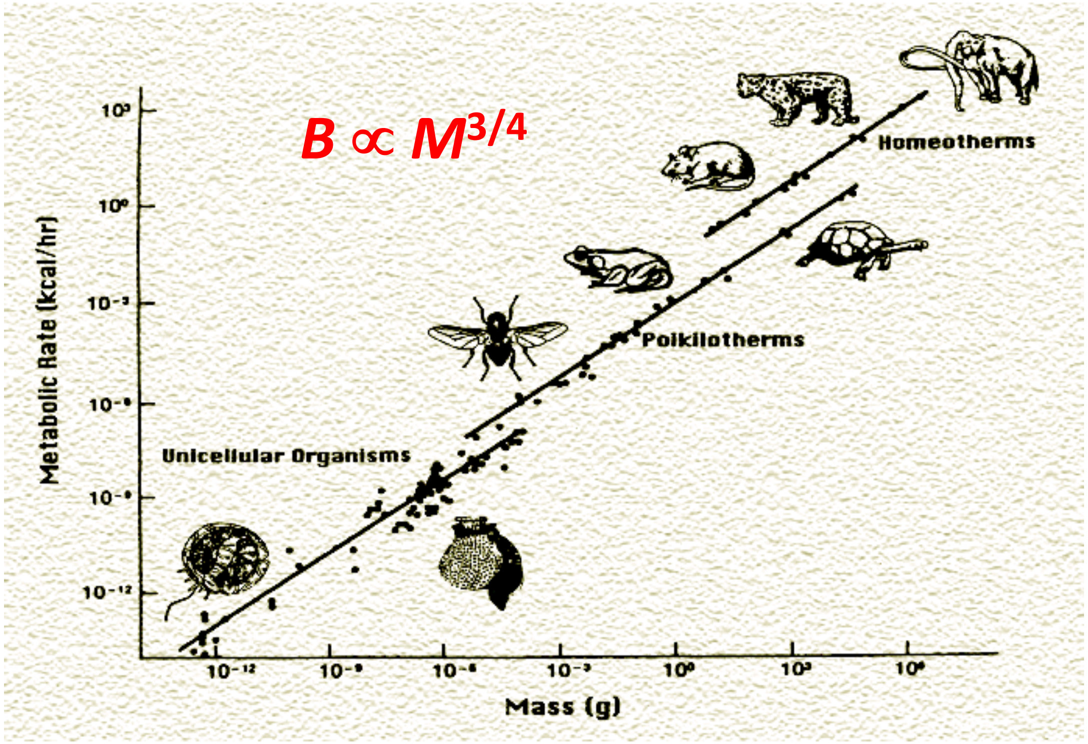
$B = cM^{3/4}$

$B \propto M^{3/4}$



Metabolic Scaling
A striking universal(?) pattern

Metabolic rate scales sub-linearly with mass



Hemmingson 1960

Metabolism is rate of energy use

Metabolism is measured as

- J/s or kcal/day
- Rate of O_2 in, or CO_2 out
- Rate of Food consumption*

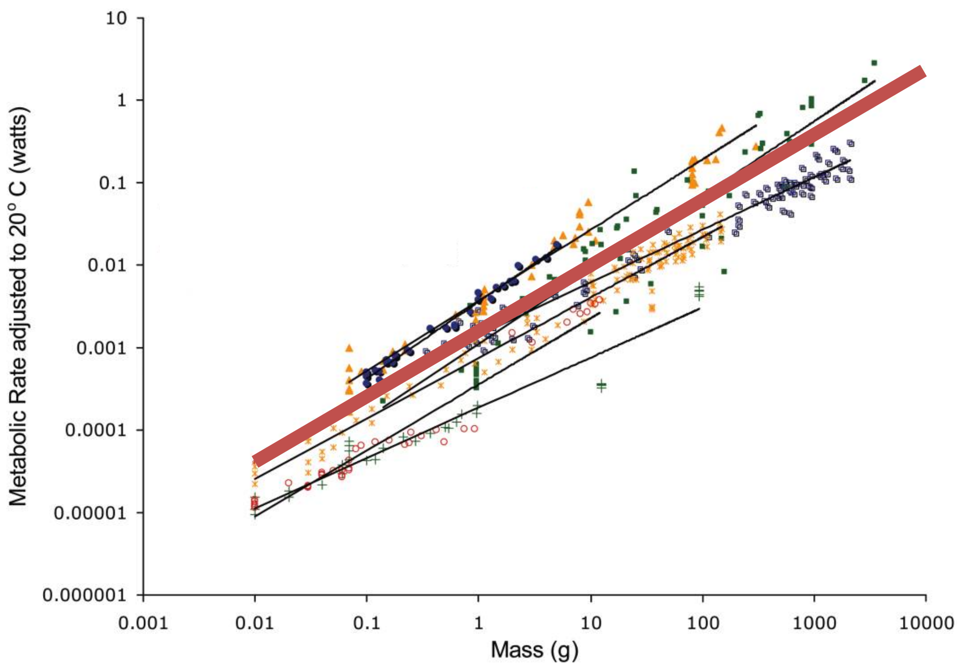
Metabolism governs the pace of life

- Physiology
- Growth
- Reproduction
- Lifespan
- Photosynthesis & carbon flux
- Ecosystem dynamics
- …

(*in non-growing animals)

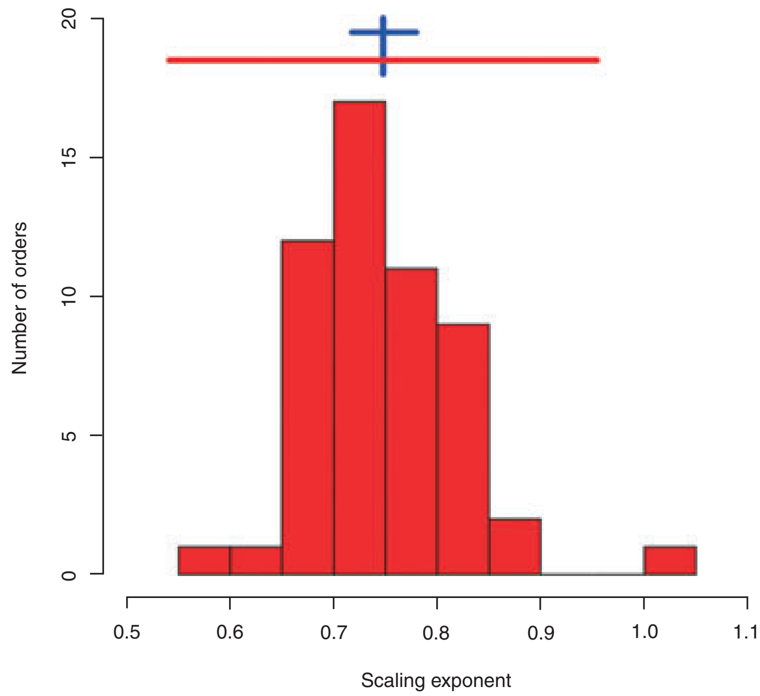
Meaningful(?) variation around a mean of ¾

Metabolic Rate in Growing Fish
$B \propto M^{3/4}$ (0.6 – 0.9)



Moses et al AmNat 2008

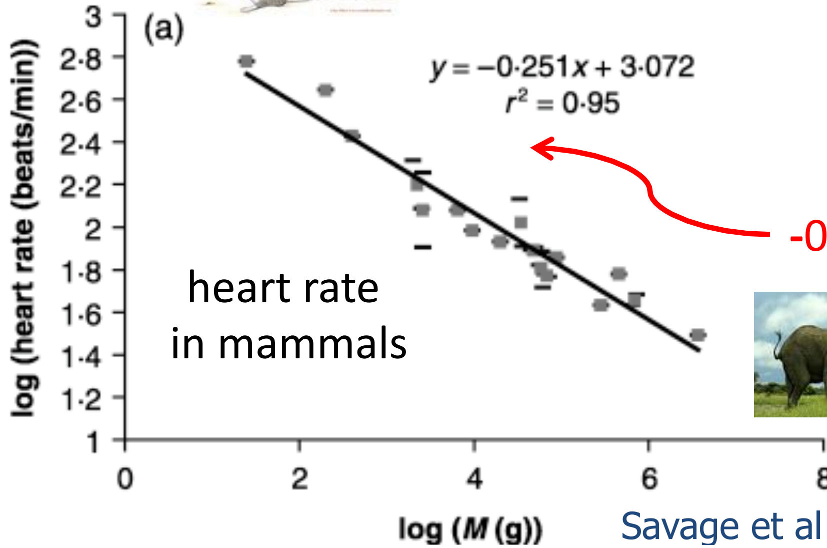
Metabolic Rate in Mammalian Orders
Mean $B \propto M^{0.749}$ (0.7 – 0.8)



Isaac & Carbone Ecol Lett 2010

Physiological Rates ~ $M^{-1/4}$

Physiological Times ~ $M^{1/4}$



Savage et al 2004

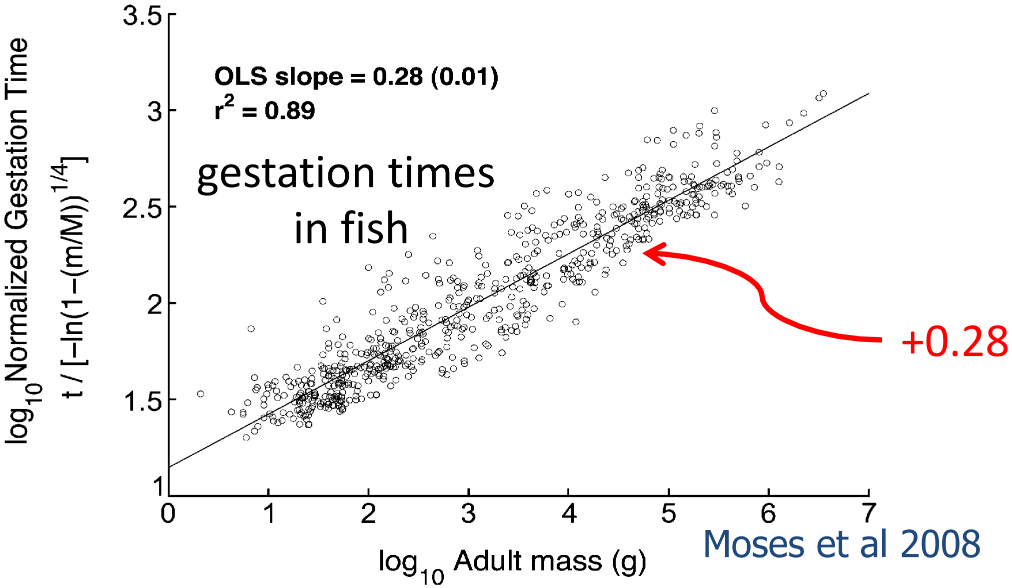
Whole animal: $B \sim M^{3/4}$

Mass-specific Scaling

$$\frac{B}{M} \propto M^{-1/4}$$

mass specific rates $\propto M^{-1/4}$

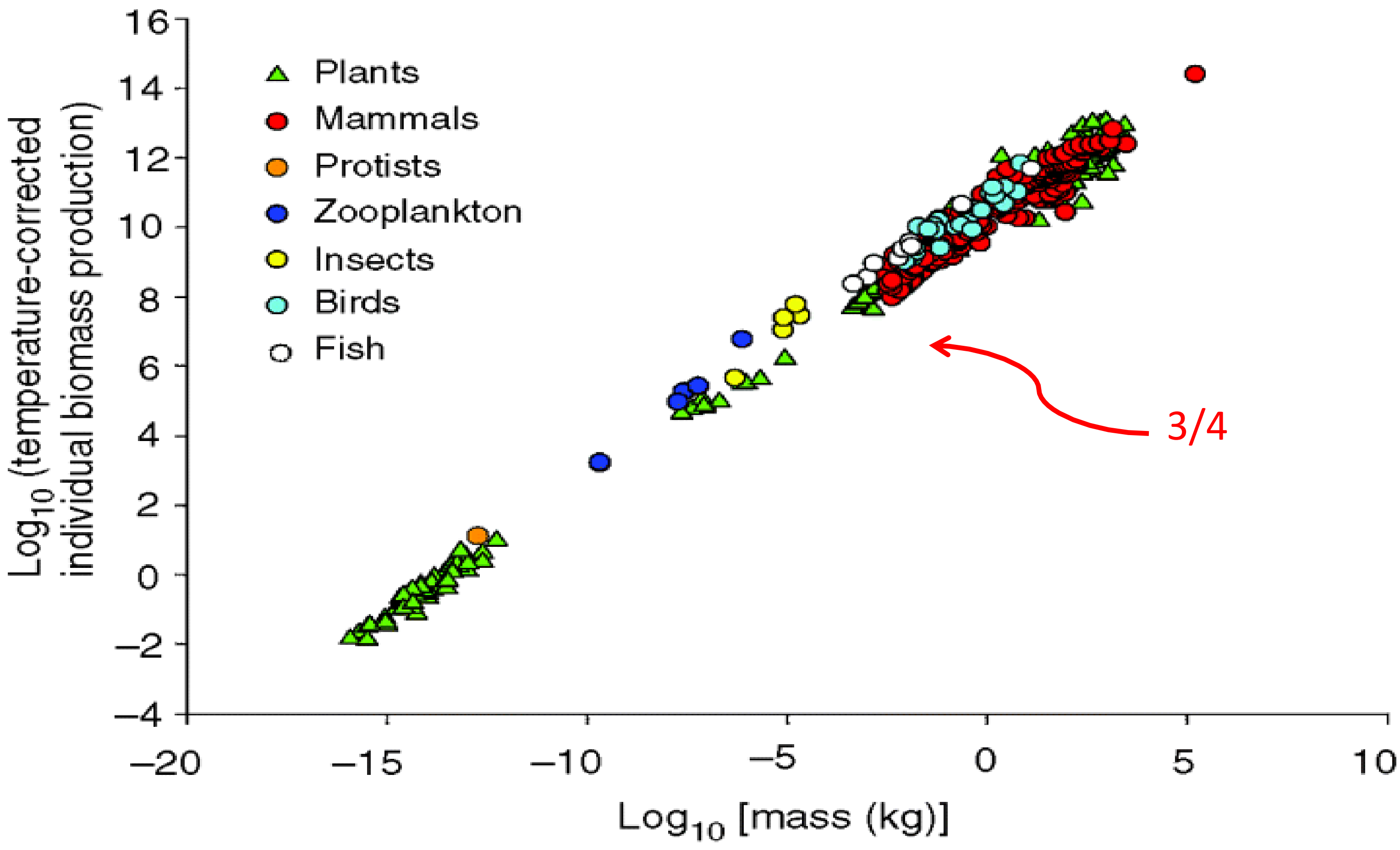
times $\propto M^{1/4}$



Moses et al 2008

Mice live fast and die young

Biomass Production: $P \propto M^{3/4}$



Ernest et al 2003

Universal Growth Curve

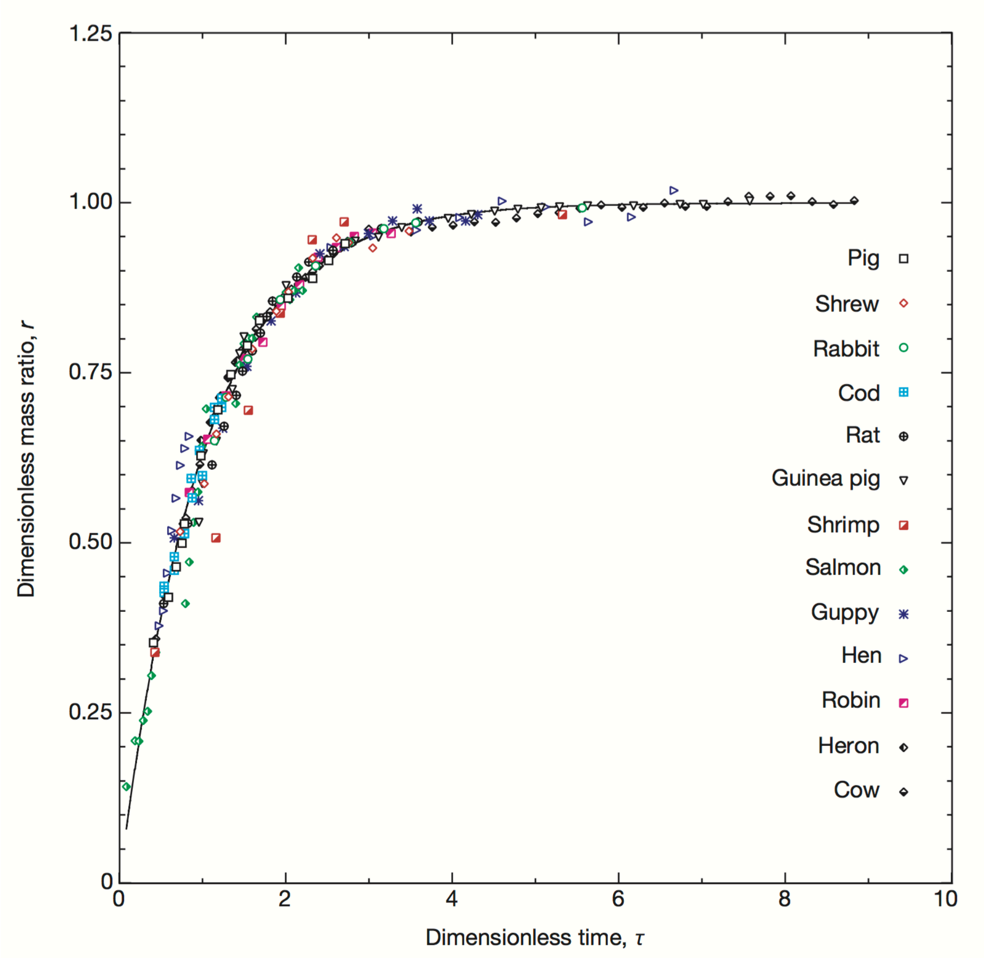
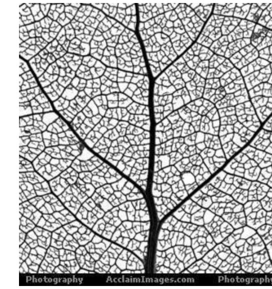


Figure 2 Universal growth curve. A plot of the dimensionless mass ratio, $r = 1 - R \equiv (m/M)^{1/4}$, versus the dimensionless time variable, $\tau = (at/4M^{1/4}) - \ln[1 - (m_0/M)^{1/4}]$, for a wide variety of determinate and indeterminate species. When plotted in this way, our model predicts that growth curves for all organisms should fall on the same universal parameterless curve $1 - e^{-\tau}$ (shown as a solid line). The model identifies r as the proportion of total lifetime metabolic power used for maintenance and other activities.

WBE, Nature 2001

Fractal Networks Generate 3/4 powers

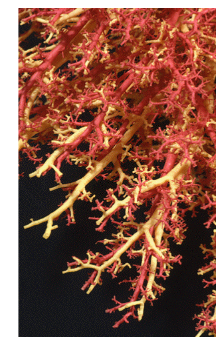


Centralized hierarchical, fractal branching

1. Constant branching ratio,

2. Area preserving

3. Space filling



4. Invariant terminal units
   - -Capillaries same length, radius & delivery capacity
   - -Metabolism proportional to # of capillaries

5. Network volume proportional to plant or animal mass

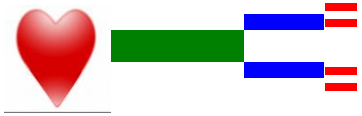
West, Brown & Enquist
Science 1997

Fractal Networks Generate 3/4 powers

Centralized hierarchical, fractal branching

1. Constant branching ratio, b

2. Area preserving $N_k A_k = c$



3. Space filling $\frac{l_{k+1}}{l_k} = b^{1/3}$

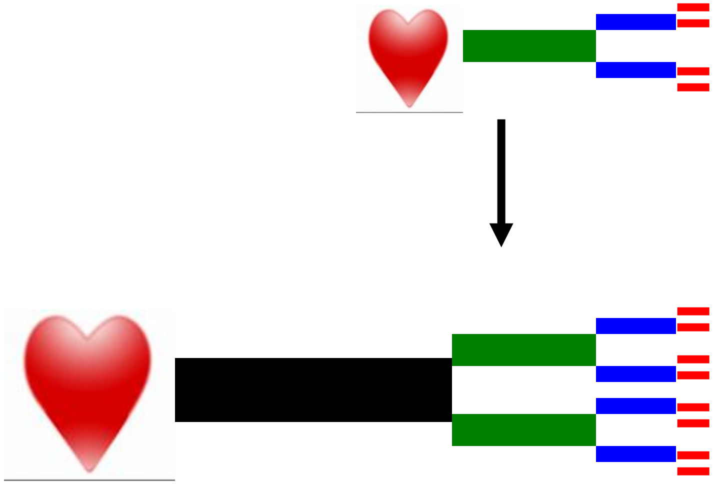
4. Invariant terminal units
   - -Capillaries same length, radius & delivery capacity
   - -Metabolism proportional to # of capillaries

5. Network volume proportional to plant or animal mass

Metabolic Rate is proportional to the number of capillaries

To double metabolic rate, double the number of capillaries

Additional network (black)
is needed to connect the 2 smaller networks



$$V_{net} = \pi b^k A_{cap} l_{cap} \sum_{i=0}^{k} b^{i/3}$$

$V_{net} \propto (b^k)^{(4/3)}$

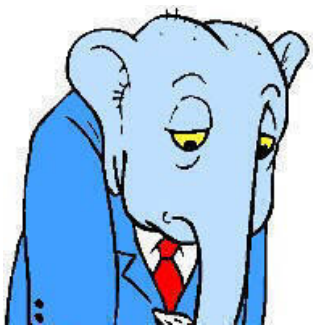
$V_{net} \propto N_{cap}^{(4/3)} \propto B^{(4/3)}$

$B \propto V_{net}^{3/4}$

Increasing Volume <u>100 times</u> increases metabolic delivery <u>30 times</u>

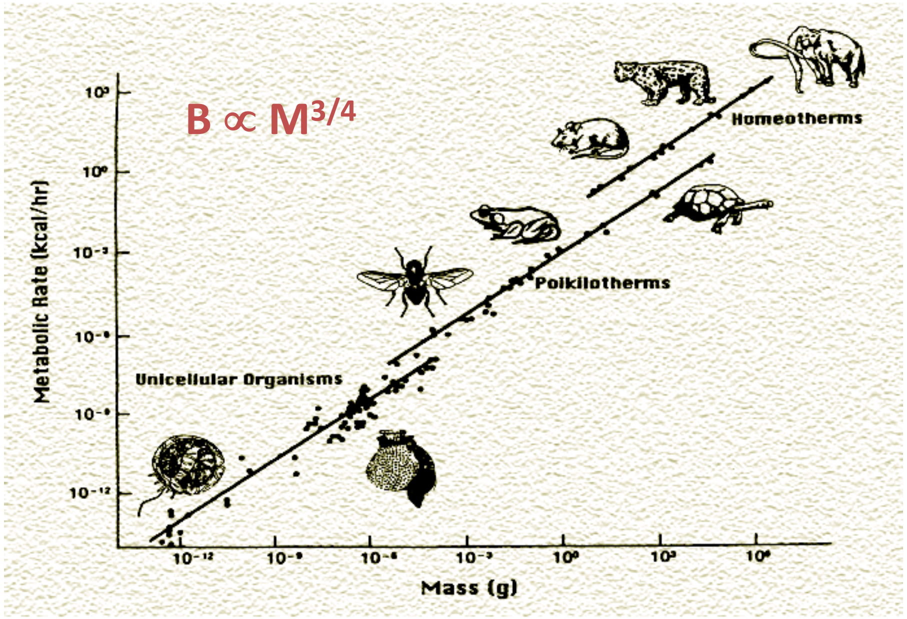
Diminishing returns: Network size grows faster than network delivery rate

Assuming* network volume is proportional to organism volume, then $B \propto V_{net}^{3/4} \propto M^{3/4}$



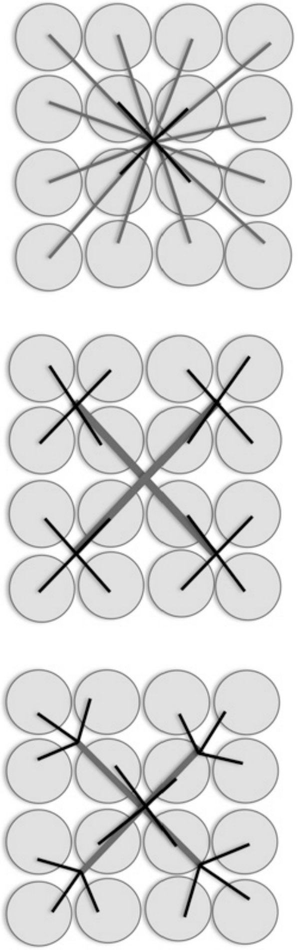
Revised Scaling Theory:
3D centralized transportation networks generate ¾ power scaling
No fractals required

Rates of Metabolism, Physiology, Growth, Reproduction & evolution show ¼ power scaling with mass



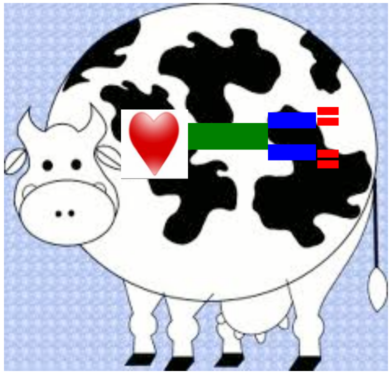
1 million heartbeats: Lifetime investment in growth, reproduction, lifespan are invariant wrt mass

West, Brown & Enquist *Science* 1997

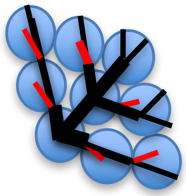


Banavar, Moses et al PNAS 2010

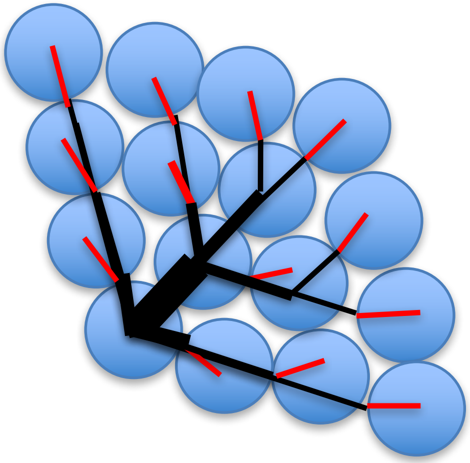
In fractal networks: delivery rate ~ $V_{net}^{3/4}$
Embedded in Euclidean animals: delivery rate ~ $V_{animal}^{2/3}$



- The 4th linear dimension is 'the last mile' the length of the service volume (capillaries) $l_s \sim M^{1/12}$



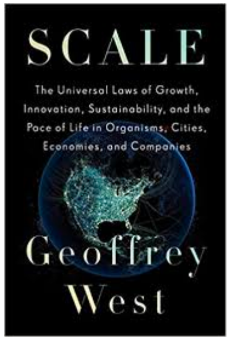
- When matching supply & demand *maximum velocity* ~ $l_s \sim M^{1/12}$
- Mean path length ~ $M^{1/3}$; transport times ~ $M^{1/4}$
- 3D centralized transport network has at most ¾ power scaling of delivery rate vs volume

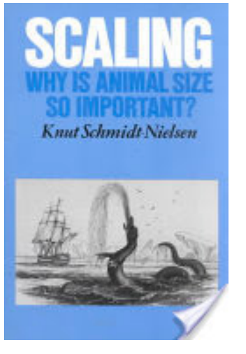


- Velocity through the last mile generates ¾ powers
- Prediction: ¾ (D/(D+1)) scaling is a more general characteristic of resource distribution networks

Recap: Metabolic Scaling in Plants and Animals

- $B \sim M^{3/4}$
  - Approximately, in plants and animals
  - Physiological rates & times show ¼ power scaling
  - Variation in exponents reflects interesting life history tradeoffs
  - Some evidence of curvilinearity (Kolokotrones & Savage *Nature* 2010)
- ¾ powers explained by the properties of 3D centralized, hierarchical fractal branching networks





Scaling in Social Systems

Colonies of social snimals also vary enormously in size

~20 Ants

Ants are
Abundant, Diverse, Dominant

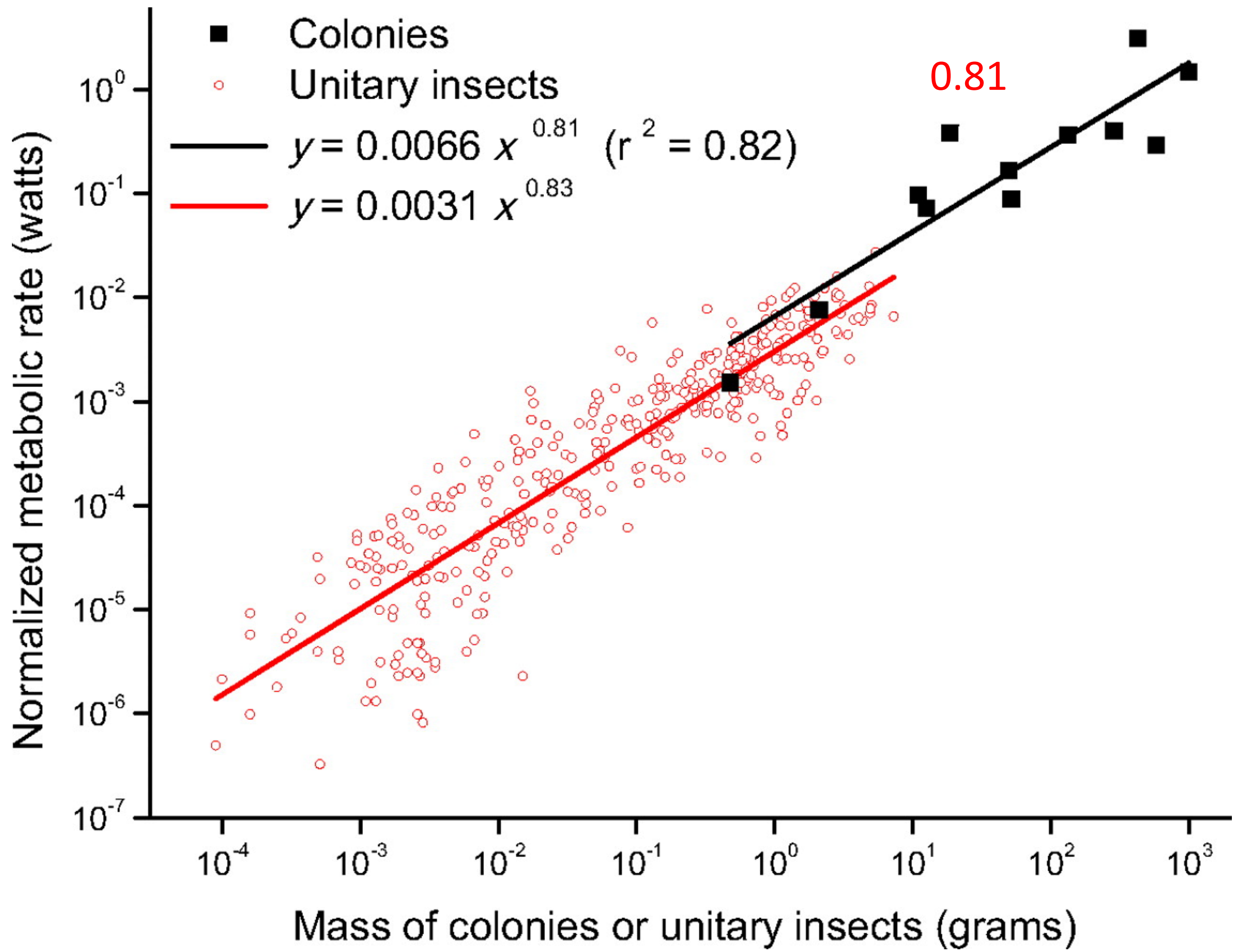
14,000 species

10^{19} ants

15% of terrestrial animal biomass

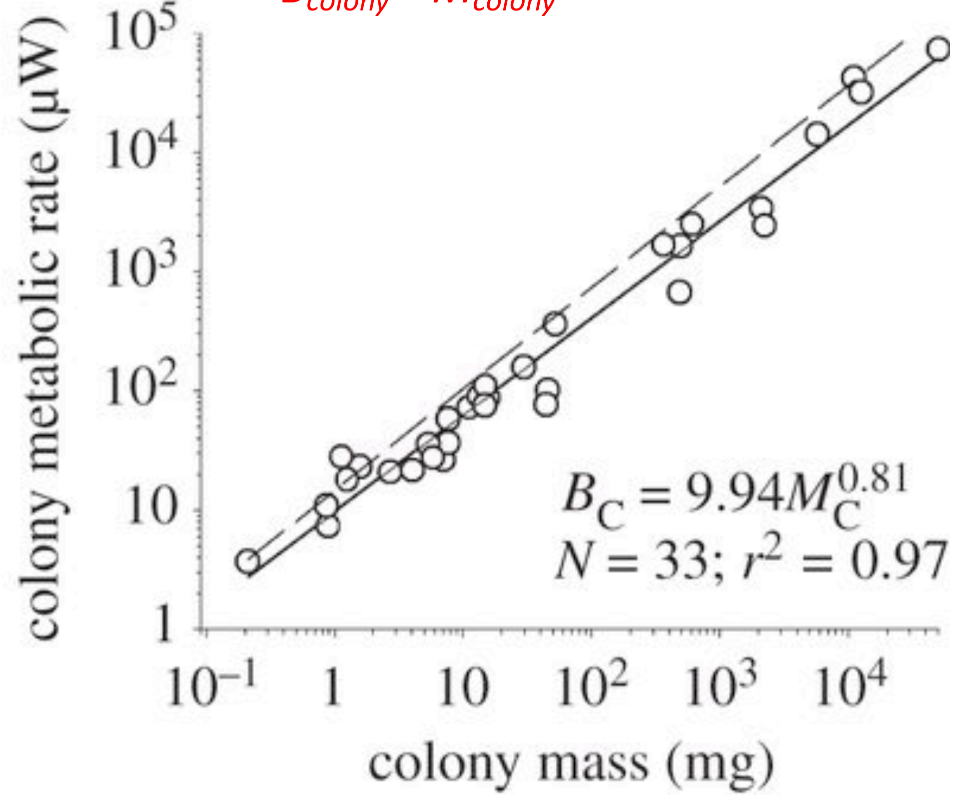
Foraging strategies
adapt to a variety of environments
from simple behaviors
with no central control

Metabolic rate and body mass for resting unitary insects and whole colonies.

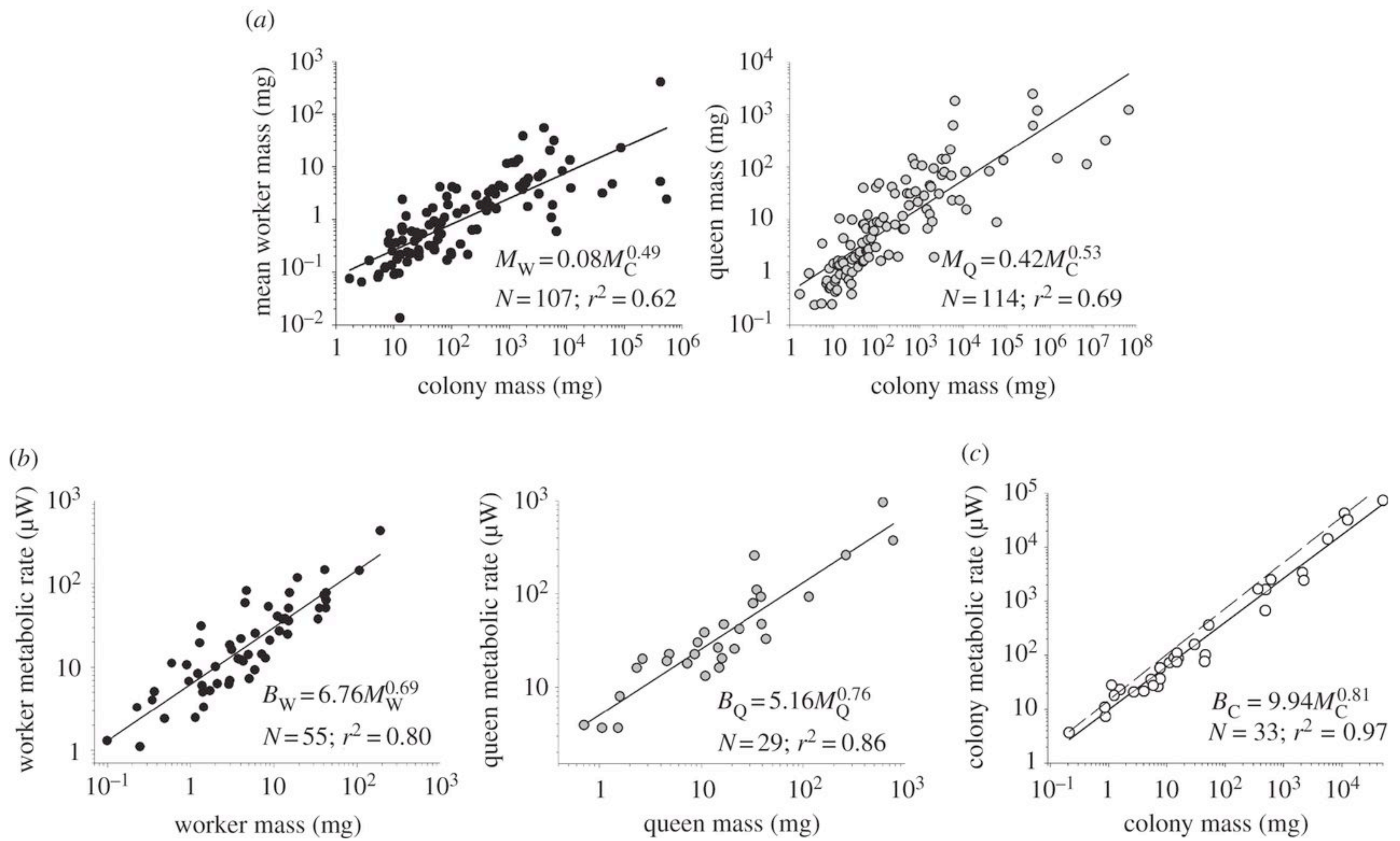


Hou C et al. PNAS 2010;107:3634-3638

$B_{colony} \sim M_{colony}^{\ 0.81}$



Shik et al Biol. Letters 2012



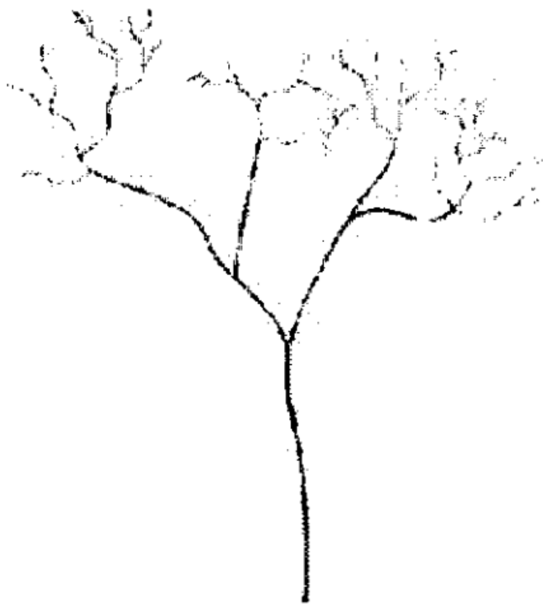
Why is colony metabolic rate more constrained than ant metabolic rate?

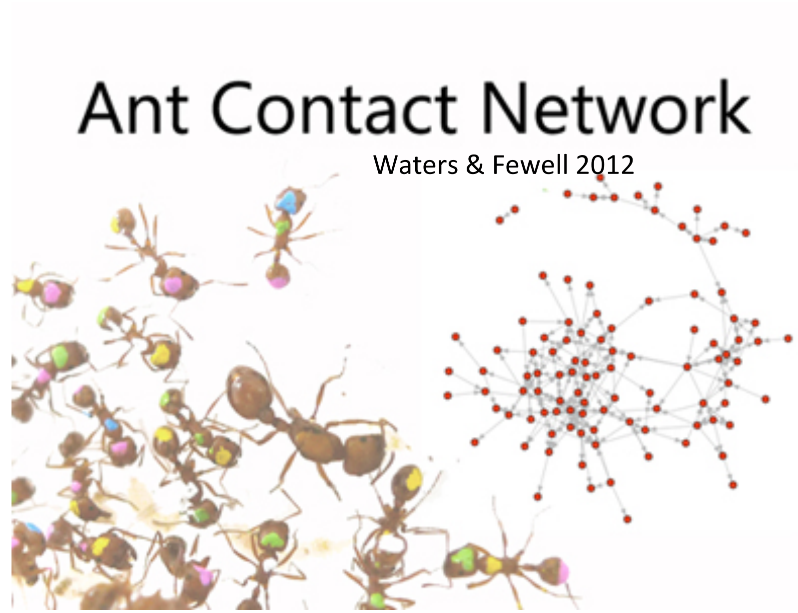
Note: the metabolic rate of disorganized, unrelated ants is linear with the number of ants.

Shik et al Biol. Letters 2012

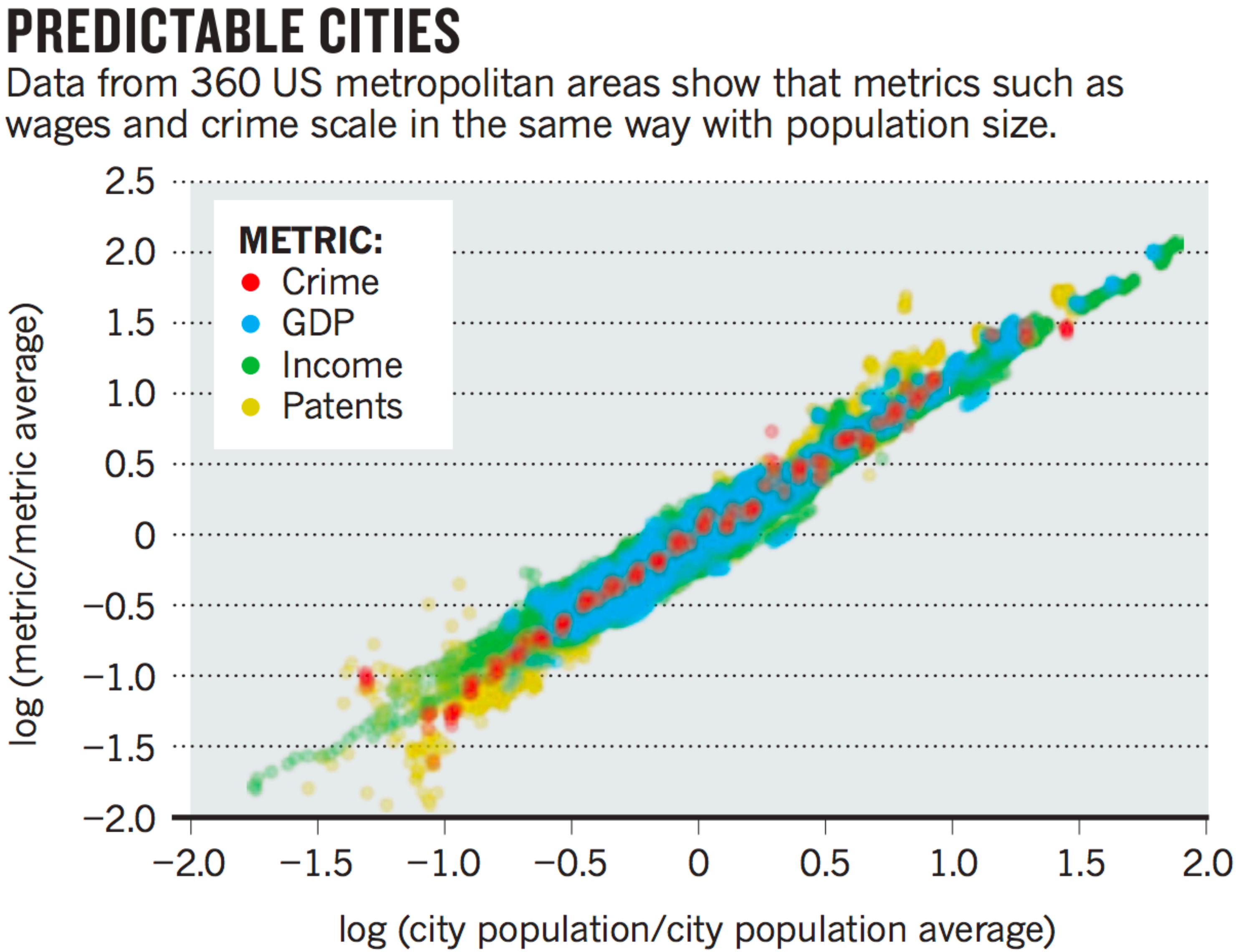
Does network scaling explain colony metabolic scaling?

Foraging trail network of Pheidole militicida
Jun et al 2003





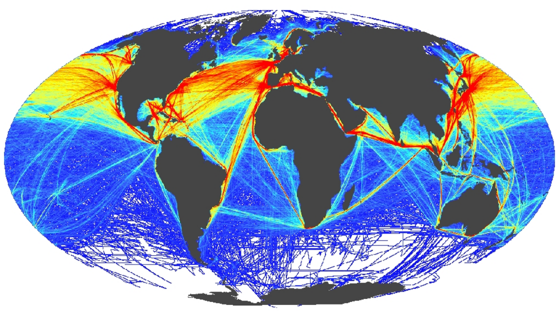
Waters & Fewell 2012



Bettencourt & West *Nature* 2010

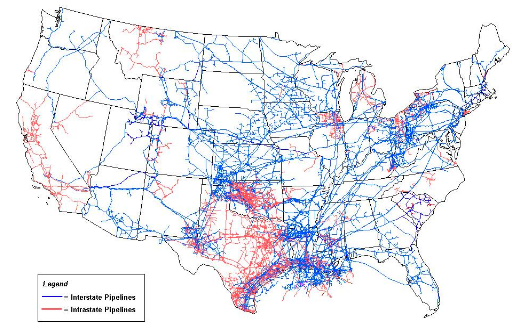
Scaling in human societies: Industrial Metabolism

Biological Metabolism
100 W from food

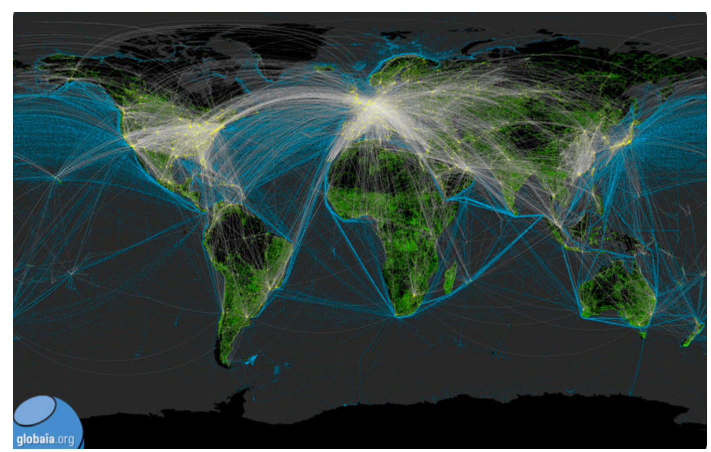


Halpern et al *Science* 2008

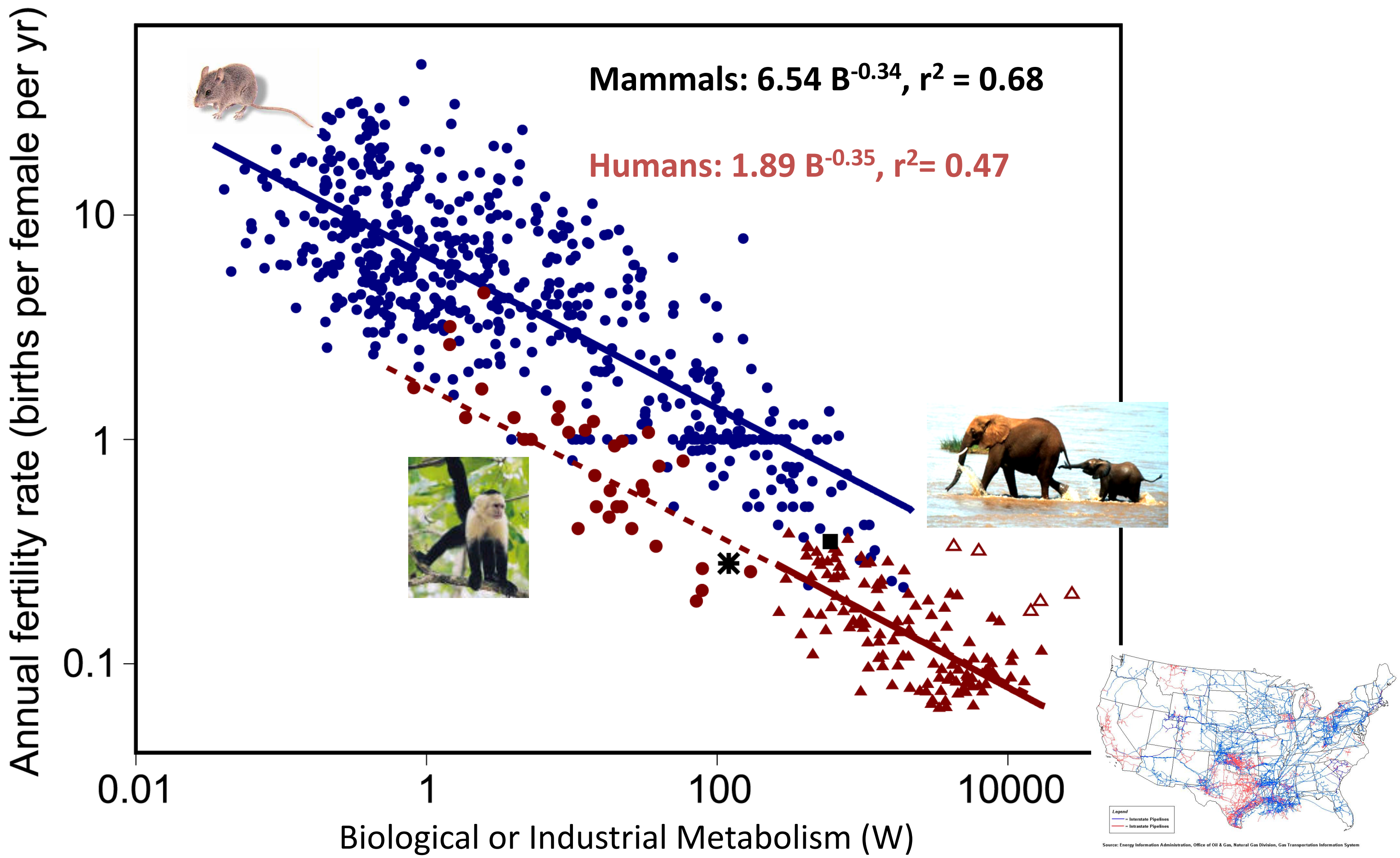
Per Capita IM
300 - 11,000 W
Fossil Fuels delivered by networks



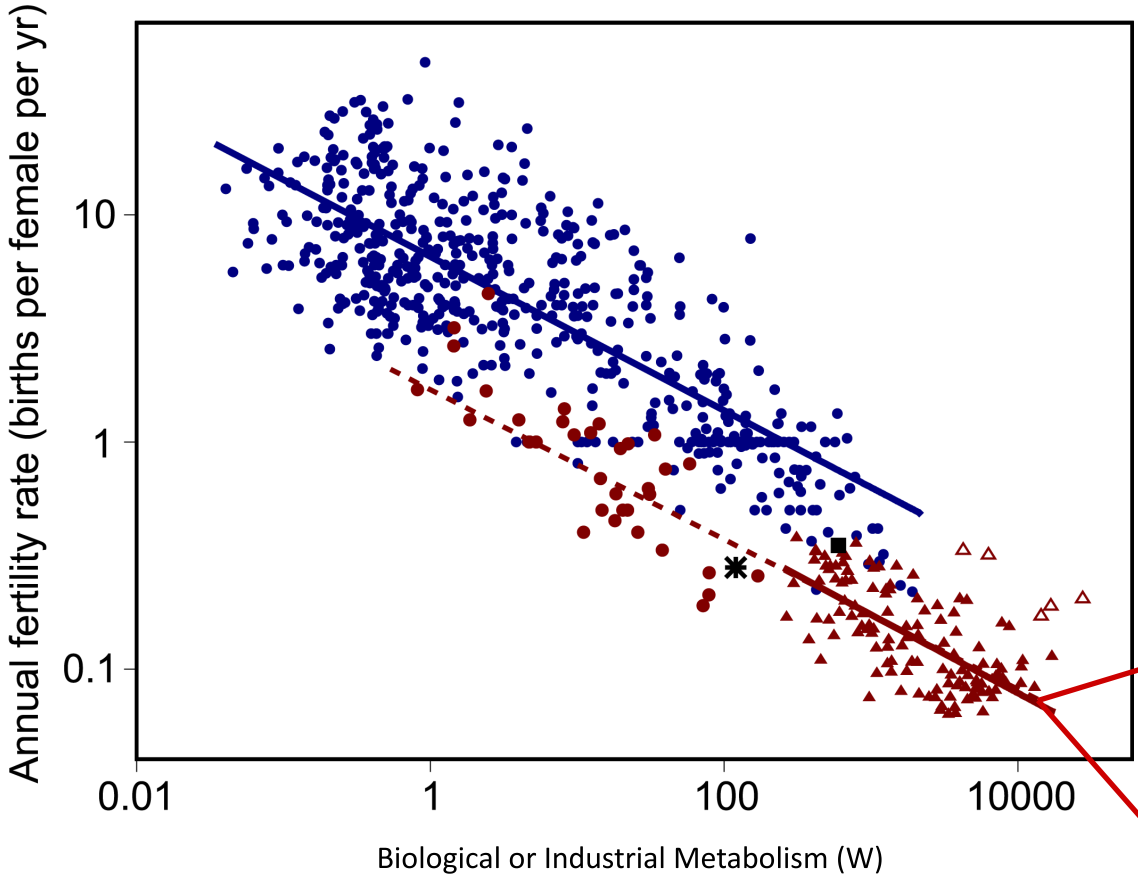
Source: Energy Information Administration, Office of Oil & Gas, Natural Gas Division, Gas Transportation Information System



Reproductive Rate vs. Metabolism: Humans and other mammals

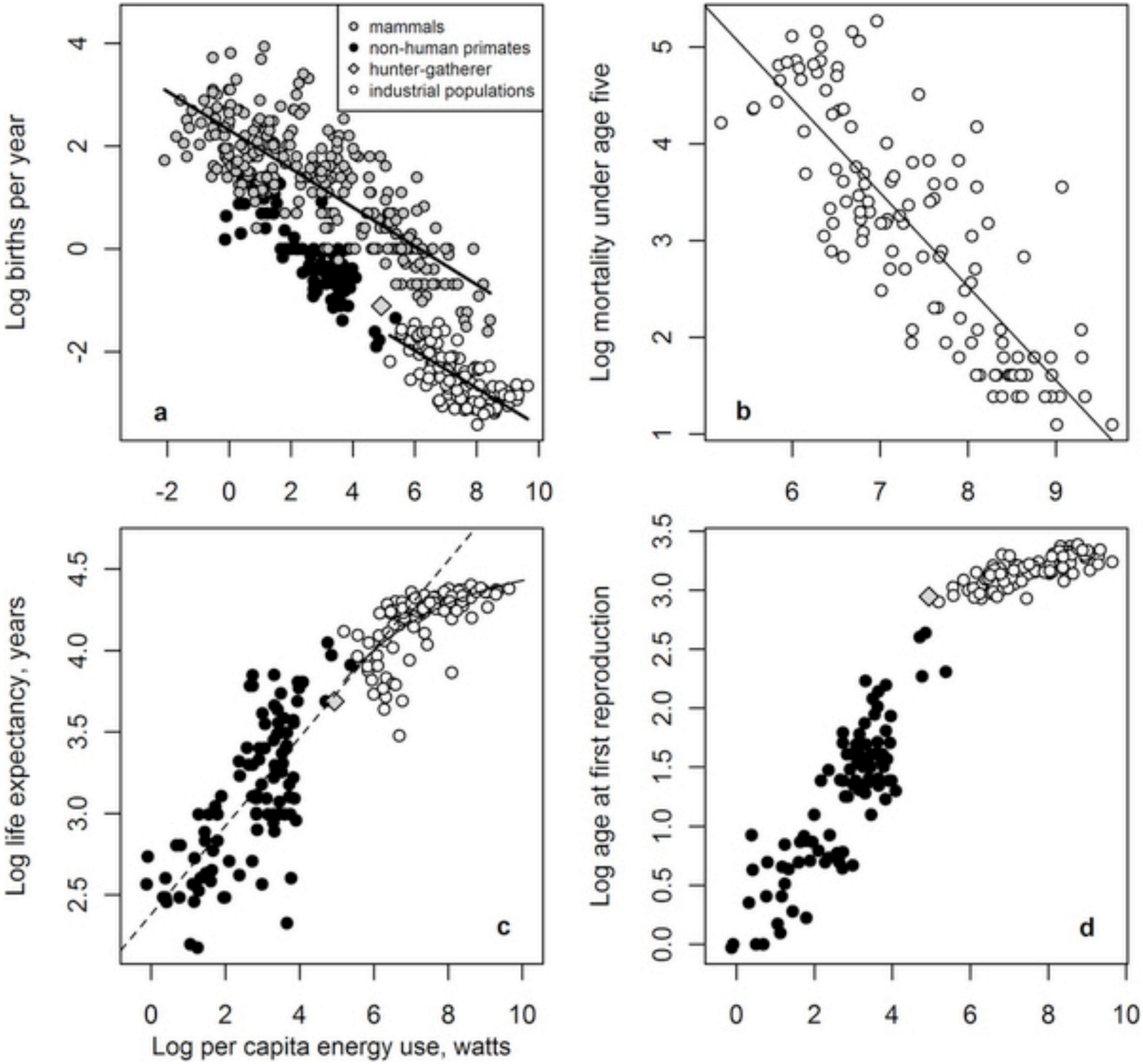


Moses & Brown 2003

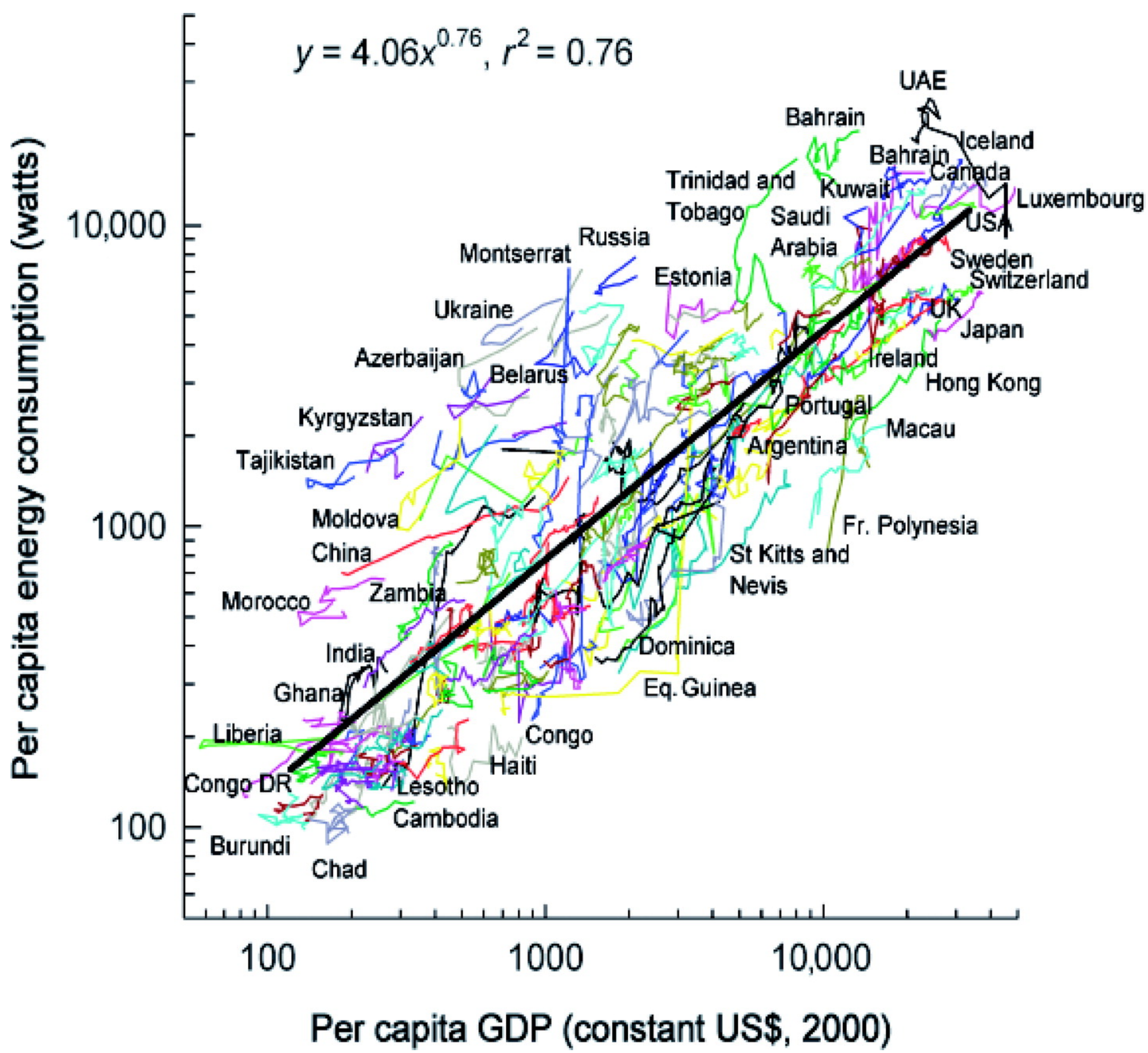


Each North American consumes the energy of a 30,000 kg primate

Reproductive rates have dropped accordingly

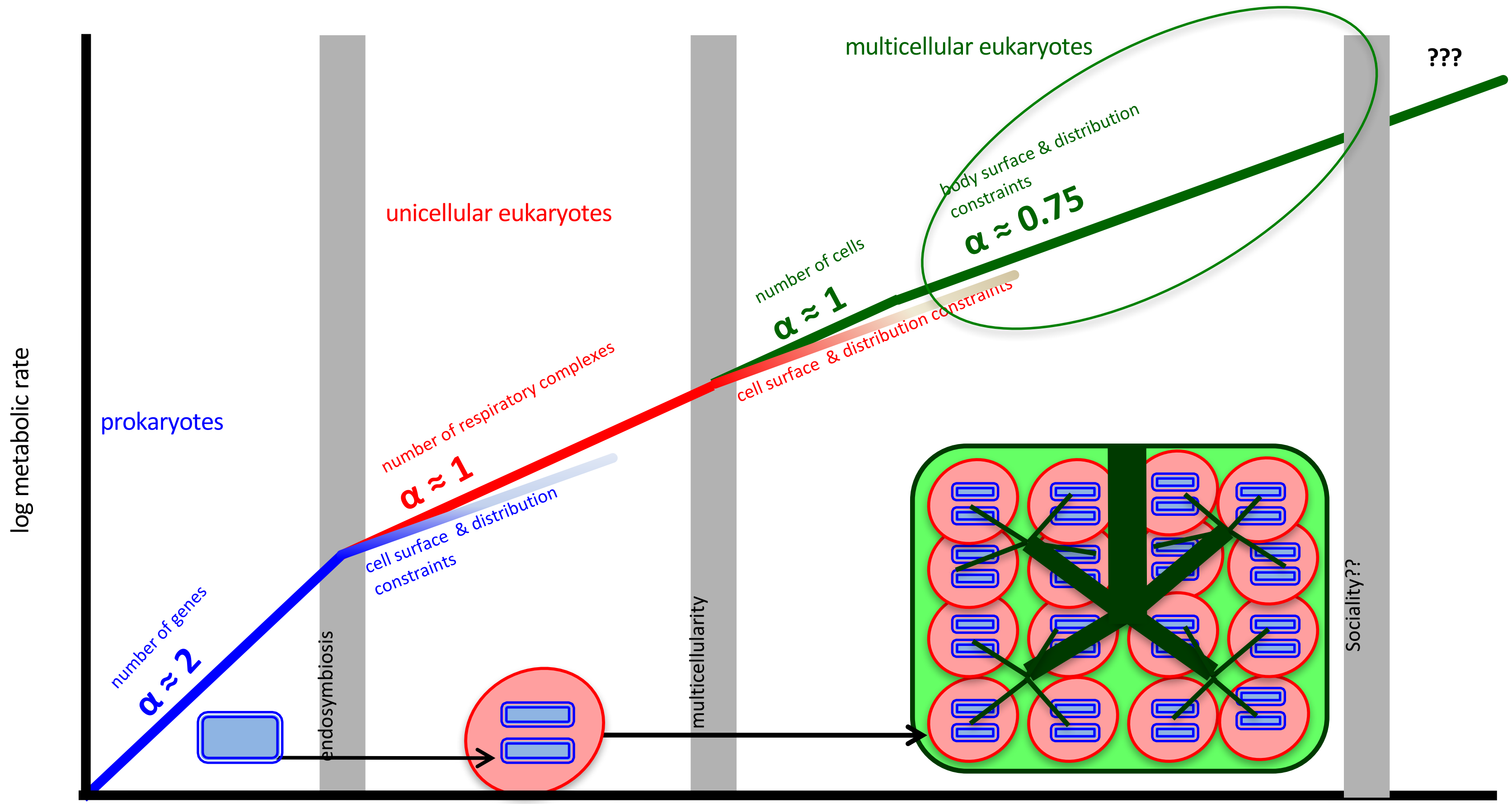


Berger et al 2011



Brown et al Bioscience 2011

Scaling intercepts and slopes shift after evolutionary innovations



Delong et al PNAS 2010

Recap: Scaling in social systems

- Scaling in ant colonies
  - Scaling of metabolic rate with colony size is less variable than with ant size
  - $B_{colony} \sim M_{colony}^{\ 0.81}$
- “Scaling” in human societies at the level of countries is highly variable but sublinear
- Resource distribution networks are not clearly hierarchical, fractal or centralized

Scaling in Computation

Computational complexity

Moore's Law

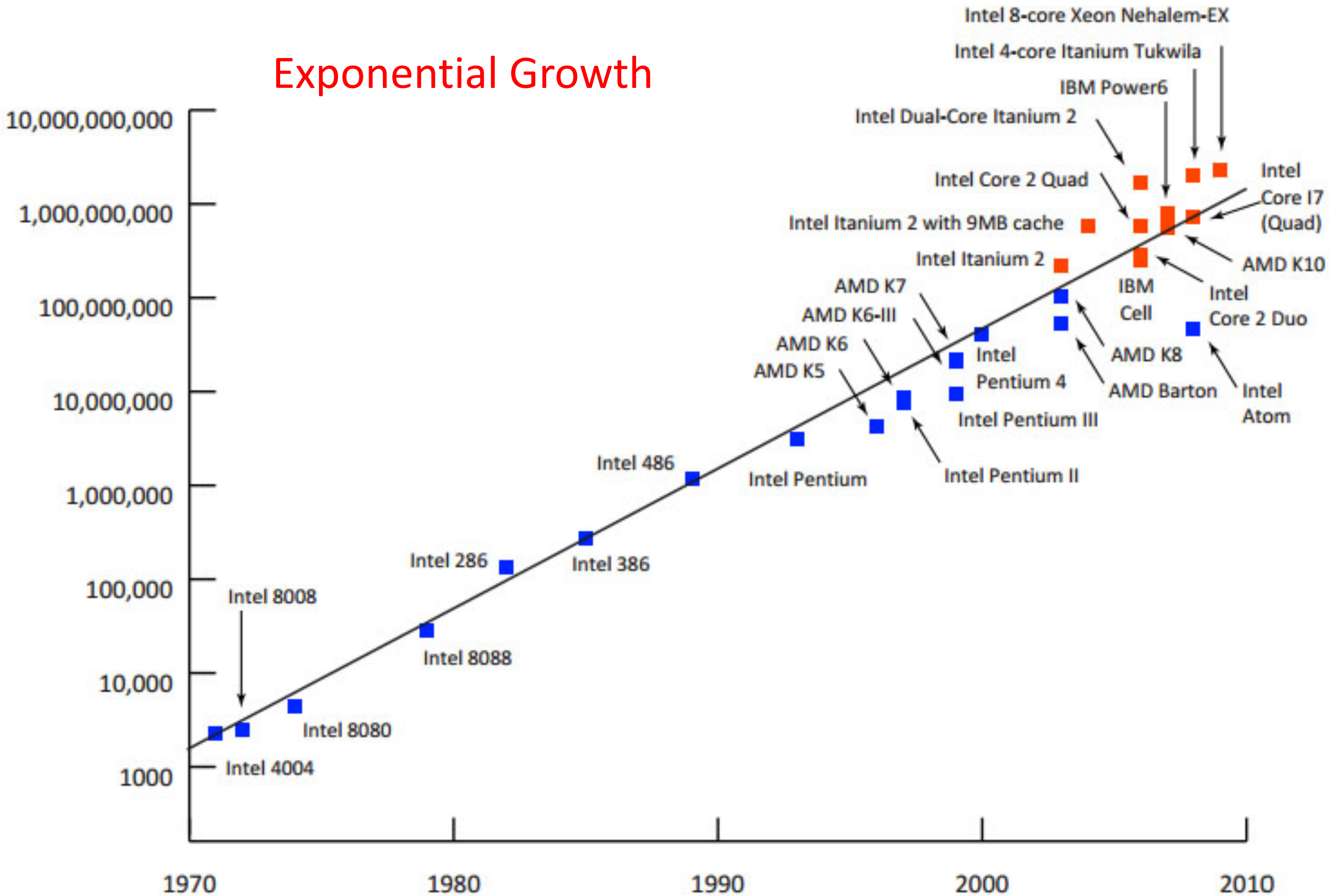
Scaling in CS

- How does computation time scale with input size?
- Example: Sorting Algorithms
  - https://en.wikipedia.org/wiki/Sorting_algorithm
  - https://visualgo.net/en/sorting
  - https://www.toptal.com/developers/sorting-algorithms
  - Bubble Sort $O(n^2)$
  - Quick Sort O(n log n)

Description	O-notation
constant	O(1)
logarithmic	O(log n)
linear	O(n)
n log n	O(n log n)
quadratic	$O(n^2)$
cubic	$O(n^3)$
polynomial	$O(n^k)$, $k \geq 1$
exponential	$O(a^n)$, $a > 1$

"scaling" in computing power: Moore's Law

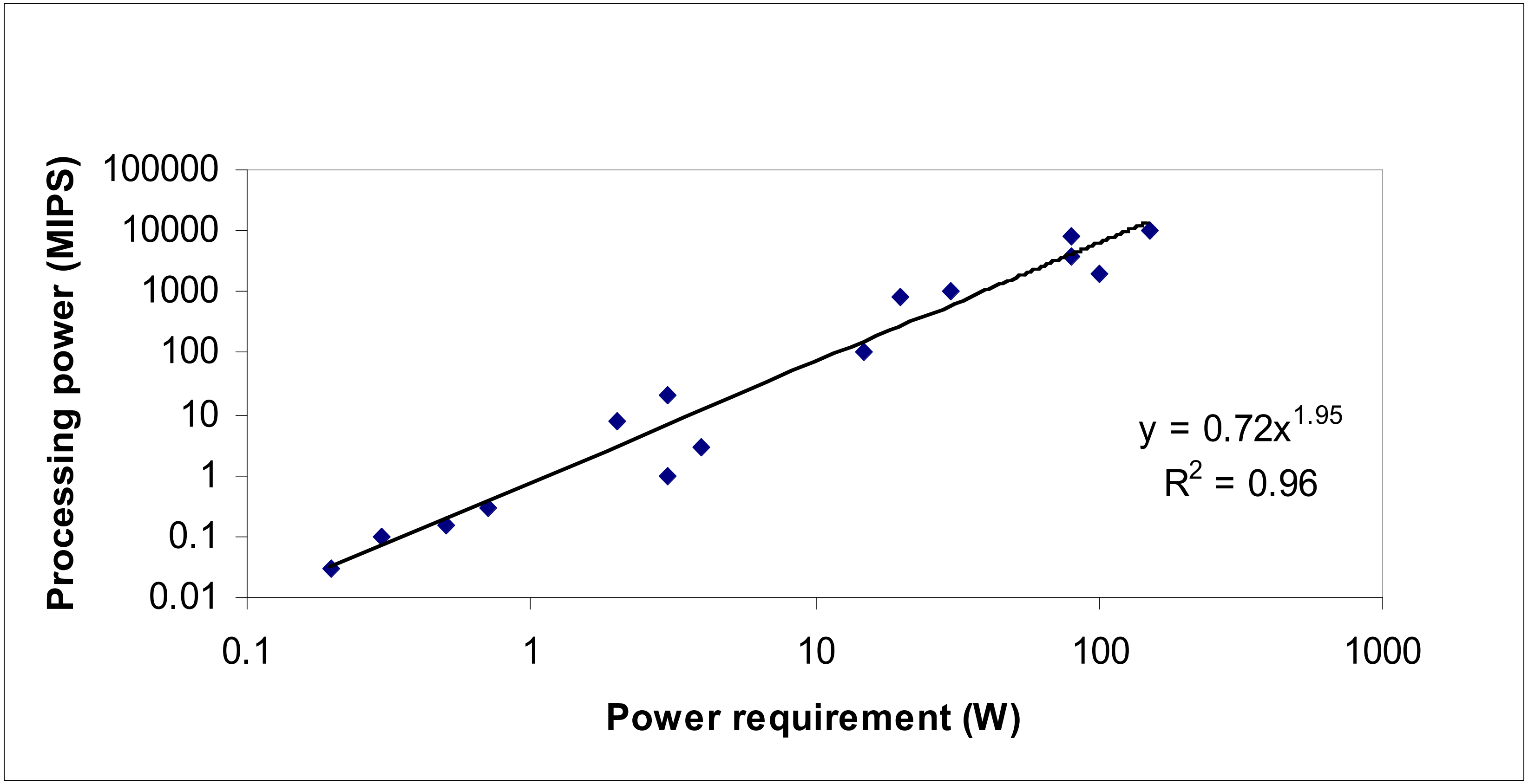
The number of transistors per chip doubles every 2 years



http://www.embedded.com/design/programming-languages-and-tools/4375996/Using-Java-for-multicore-programming-complexity--Part-1

Power scaling: Computing Power vs Power Consumption

Thousand-fold increase in power,
Million-fold increase in MIPS
(Million Instructions Per Second)



1970
100 Watts powered 15 MIPS

2005
100 Watts powered 6700 MIPS

Analyzing Scaling Relationships

The scaling exponent is the slope on log-log plot

$\log(MIPS) = 2\log(P) + \log(c)$

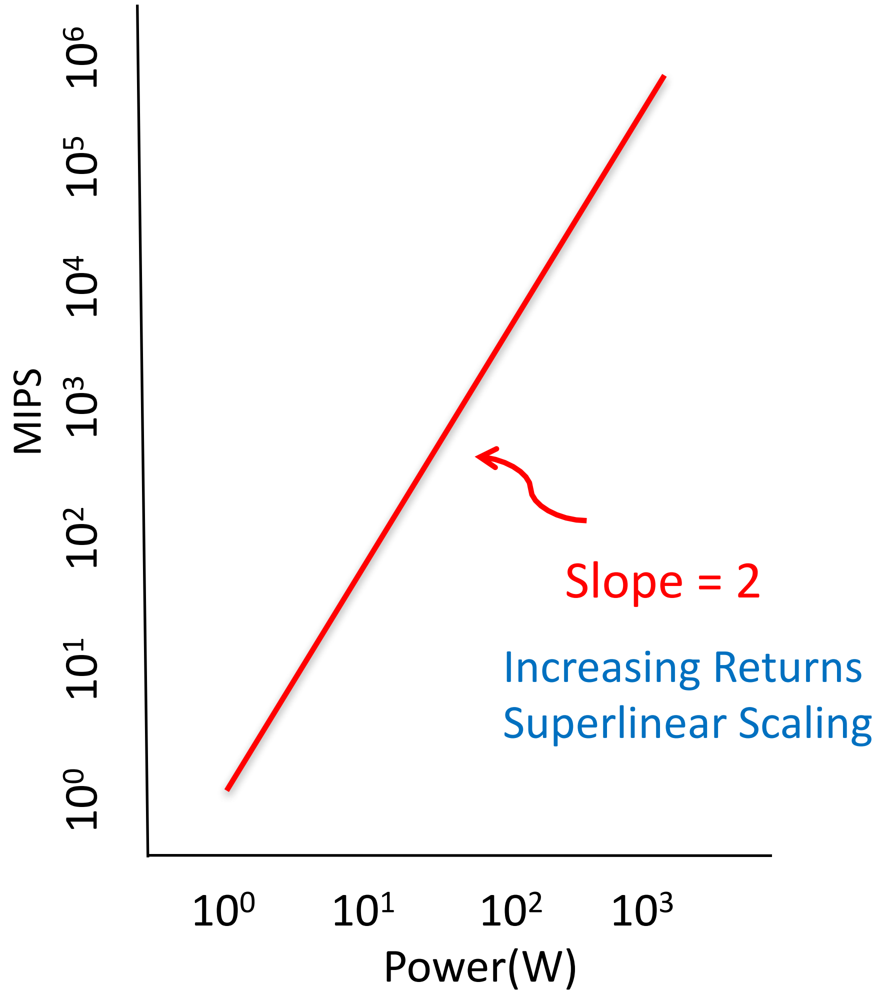
$y = mx + b$

The intercept is log(c)

The slope is 2

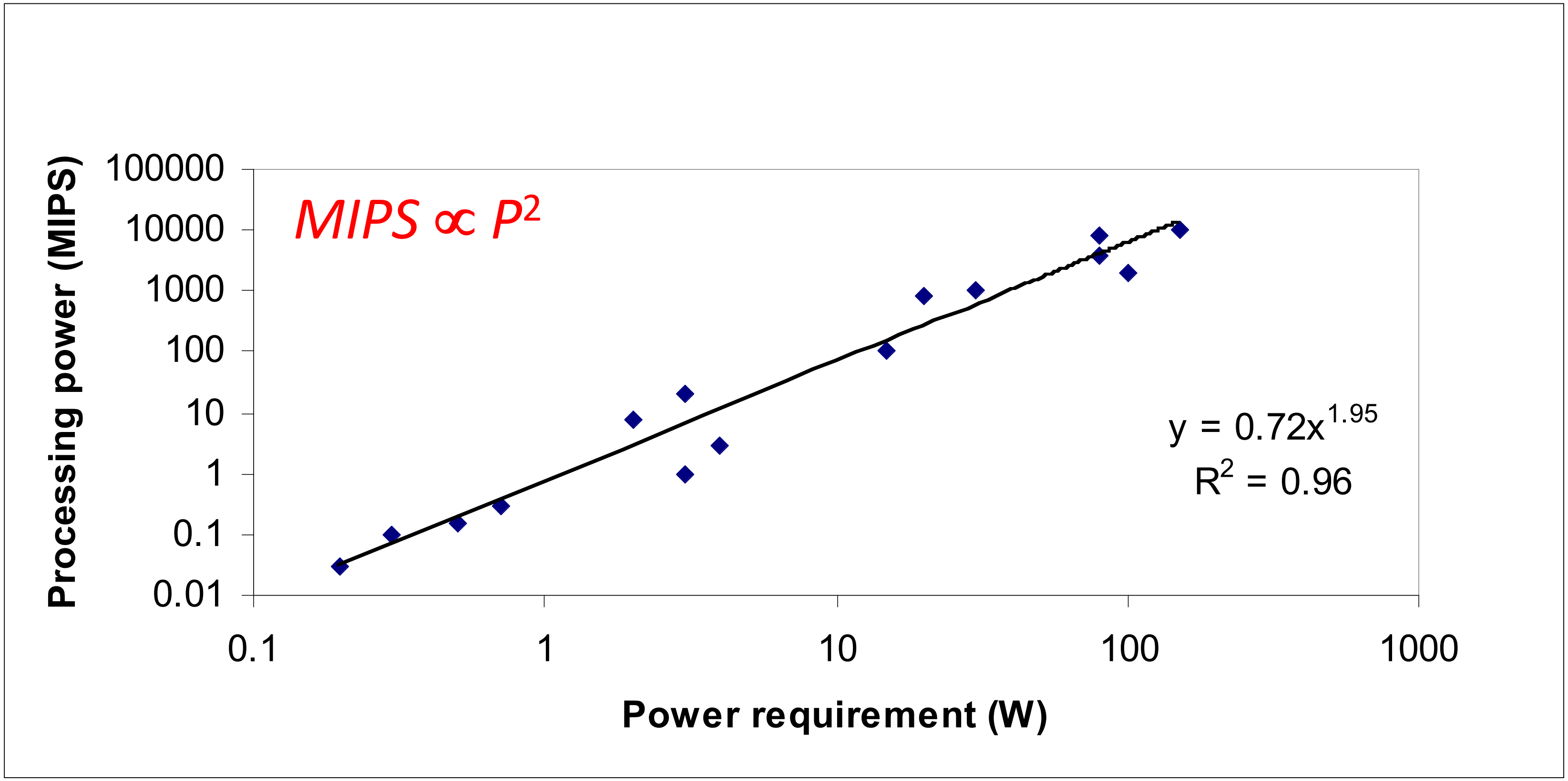
$MIPS = cP^2$

$MIPS \propto P^2$



Power scaling: Computing Power vs Power Consumption

Thousand-fold increase in power,
Million-fold increase in MIPS
(Million Instructions Per Second)

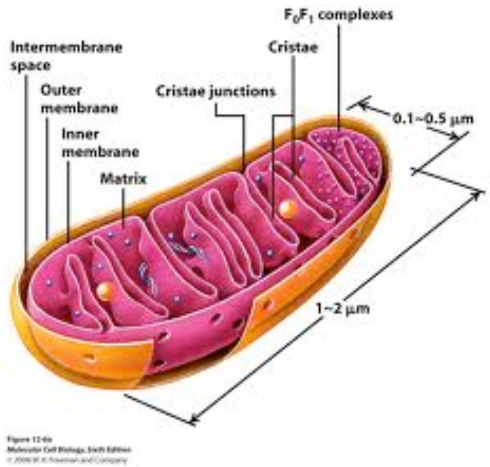


1970
100 Watts powered 15 MIPS

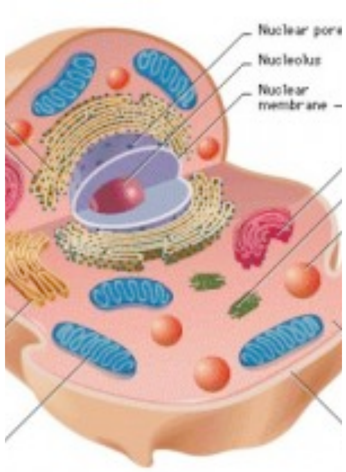
2005
100 Watts powered 6700 MIPS

Scaling in Biology & Computation

Toward a Unified Framework

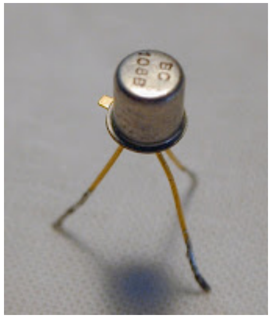


Mitochondrion

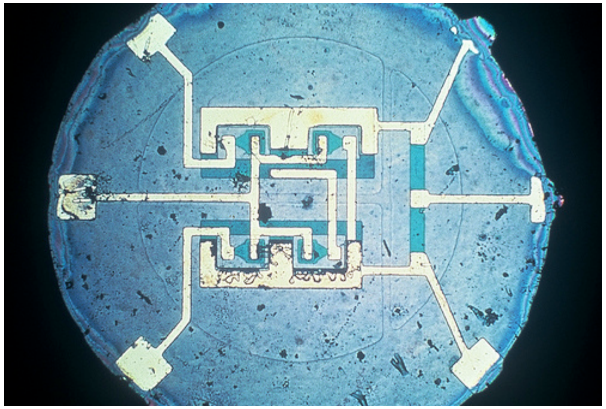


1 to 1000s of Mitochondria per cell

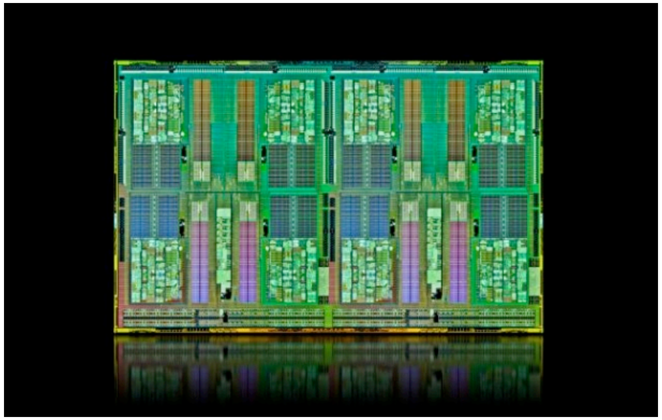
Trillions of mitochondria



Transistor

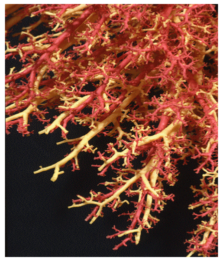
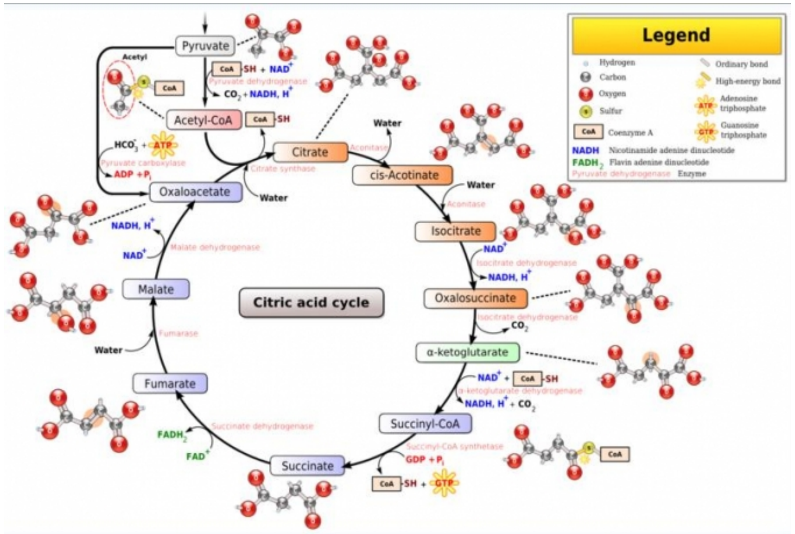


Integrated circuit



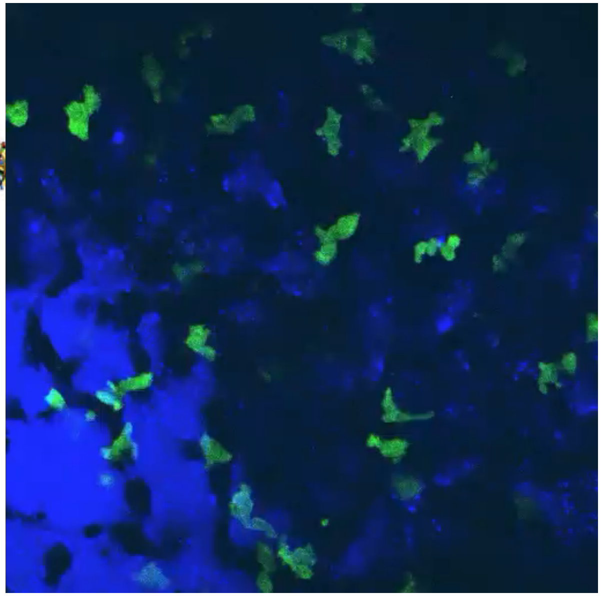
AMD Opteron multi-core chip billions of transistors



Living systems acquire and transform

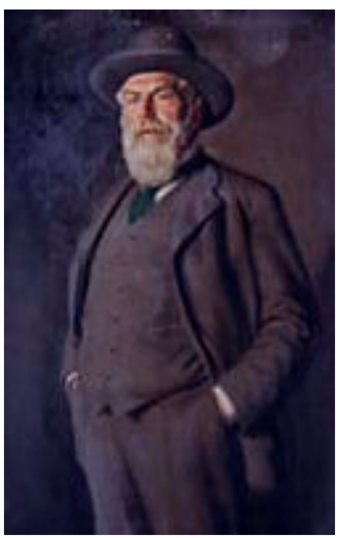
energy and information

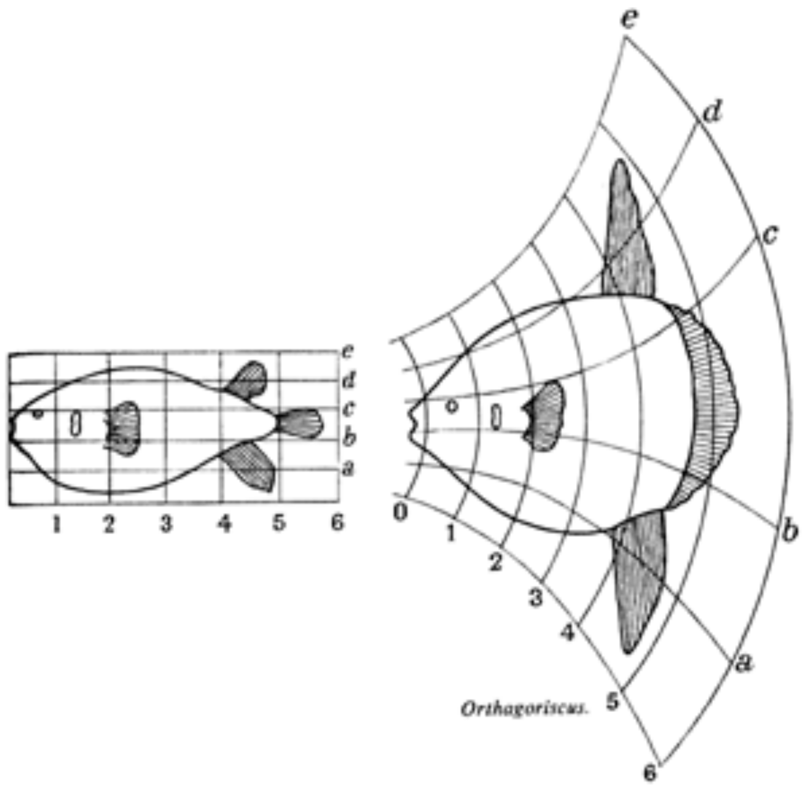




"We must envisage a living organism as a special kind of system to which the general laws of physics and chemistry apply.

And because of the prevalence of homologies of organization, we may well suppose, as D'Arcy Thompson has done, that certain physical processes are of very general occurrence…"

attributed to Alan Turing by Evelyn Fox Keller in Making Sense of Life





Energy and time determine scaling in biological and computer designs

Melanie Moses[1,2,3], George Bezerra[1], Benjamin Edwards[1], James Brown[2,3] and Stephanie Forrest[1,2,3]

Energy (oxygen) transported through cardiovascular network to mitochondria

Bits delivered by “interconnect” on microprocessors to transistors

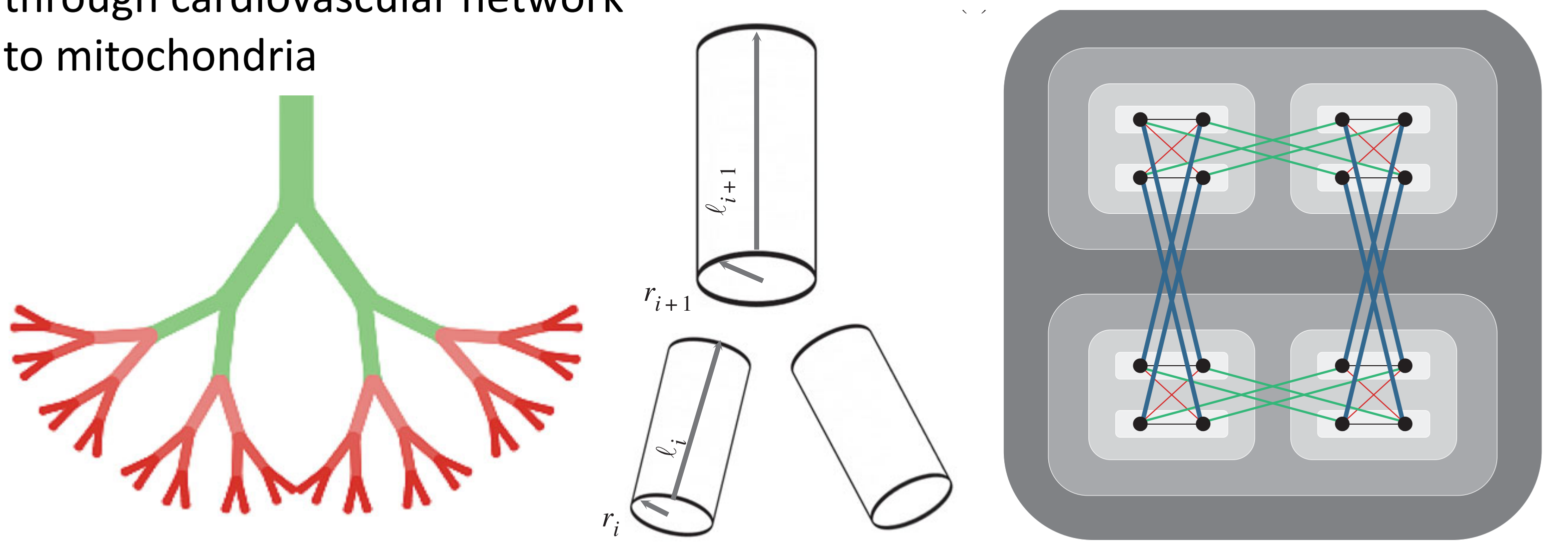


Figure 1. Idealized branching models in biology (*a*) and computers (*c*). (*a*) A cardiovascular tree with branching factor $\lambda = 2$, $H = 5$ hierarchical branchings and $N = 32$ terminal branches at level 0 that represent capillaries. (*b*) The radius and length of successive branches: D_r defines the relative radius and D_l defines the relative length of pipe or wire between successive hierarchical levels (i and $i + 1$) in both biology (*a*) and computers (*c*). (*c*) The semi-hierarchical branching of logic wires on a computer chip. Each module within a hierarchical level is shaded the same colour. The purple, red, green and blue (thinnest to thickest) wires cross 0, 1, 2 and 3 modules, respectively. The wire lengths and widths increase as they cross more levels according to D_l and D_r. D_w defines the number of wires, determined by the ratio of internal (intra-module) communication per node to external (inter-module) communication per node. Here $D_w = 2$ so that a node is connected to all nodes within a module (in this case only 1) by a purple wire, 1/2 of the nodes in the next hierarchical level by red wires, 1/4 of the nodes in the next level by green wires, and 1/8 of the nodes in the next level by blue wires.

Fractal Networks Generate 3/4 powers

Centralized hierarchical, fractal branching

1. Constant branching ratio,

2. Area preserving

3. Space filling

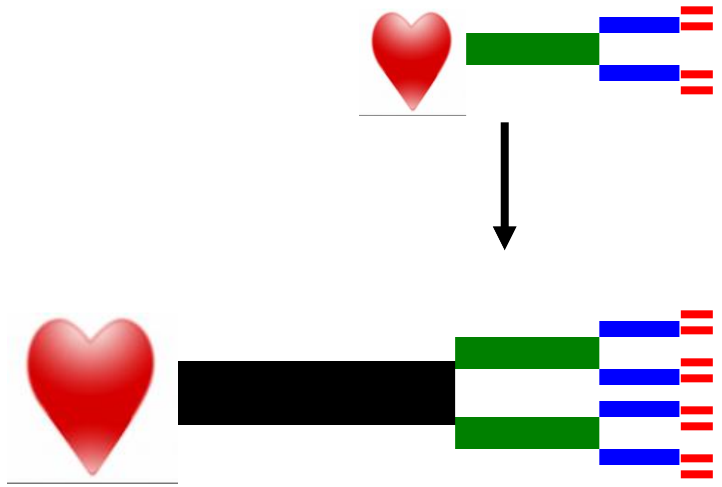
4. Invariant terminal units
   - -Capillaries same length, radius & delivery capacity
   - -Metabolism proportional to # of capillaries

5. Network volume proportional to mass

Metabolic Rate is proportional to the number of capillaries

To double metabolic rate, double the number of capillaries

Additional network (black)
is needed to connect the 2 smaller networks



$$V_{net} = \pi b^k A_{cap} l_{cap} \sum_{i=0}^{k} b^{i/3}$$

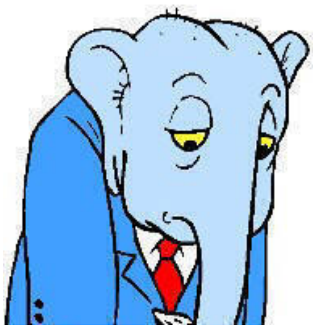
$$V_{net} \propto (b^k)^{(4/3)}$$

$$V_{net} \propto N_{cap}^{(4/3)} \propto B^{(4/3)}$$

$$B \propto V_{net}^{3/4}$$

Increasing Volume <u>100 times</u> increases metabolic delivery <u>30 times</u>

Diminishing returns: Network size grows faster than network delivery rate



Dec Alpha H-tree (1994), a 2D WBE fractal network

A centralized network that delivered a timing signal

Wire lengths and radii follow WBE predictions in 2D

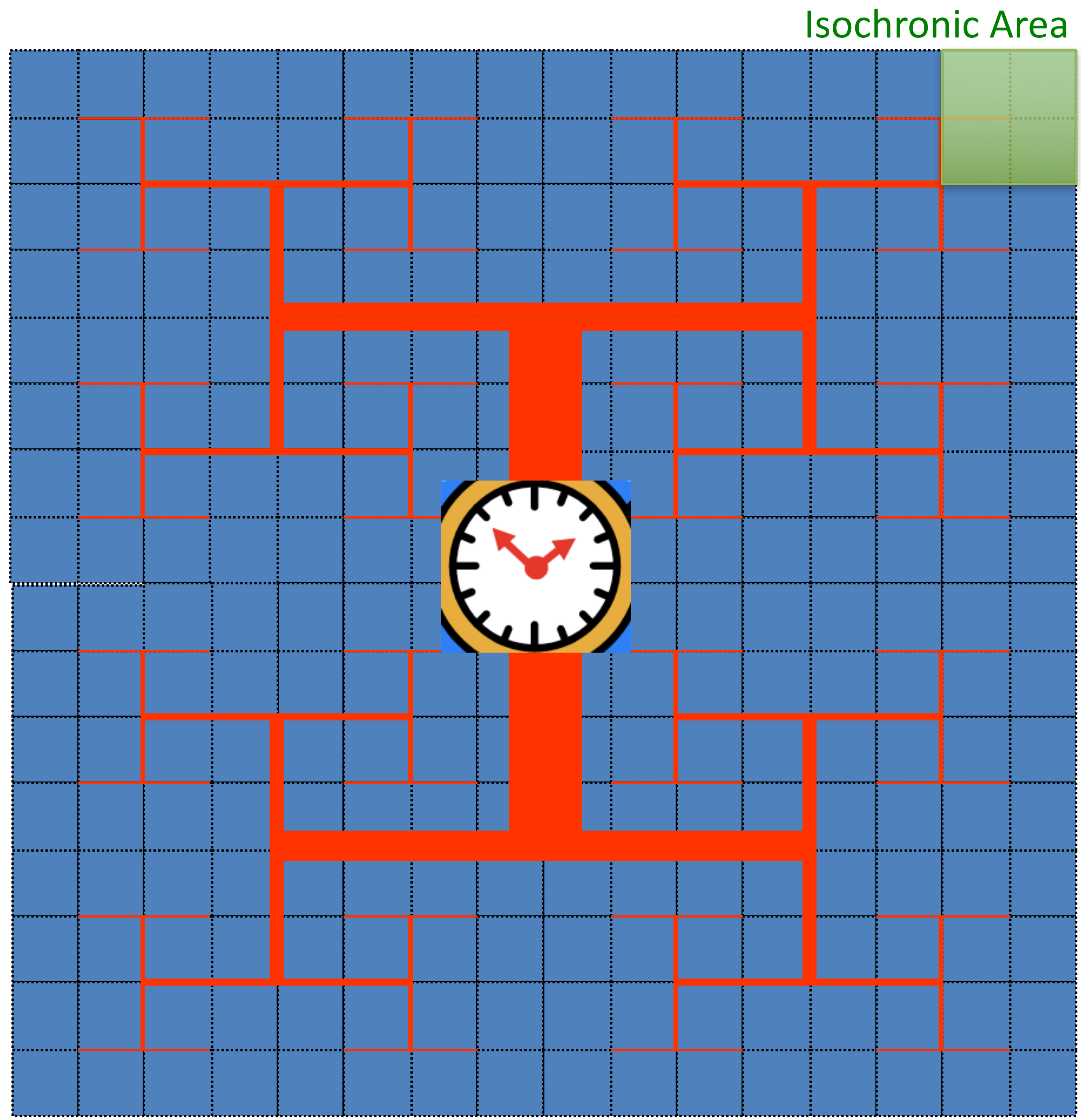
Allowed unprecedented speed (300 MHz)

Clock speed is limited by the isochronic area (last mile)

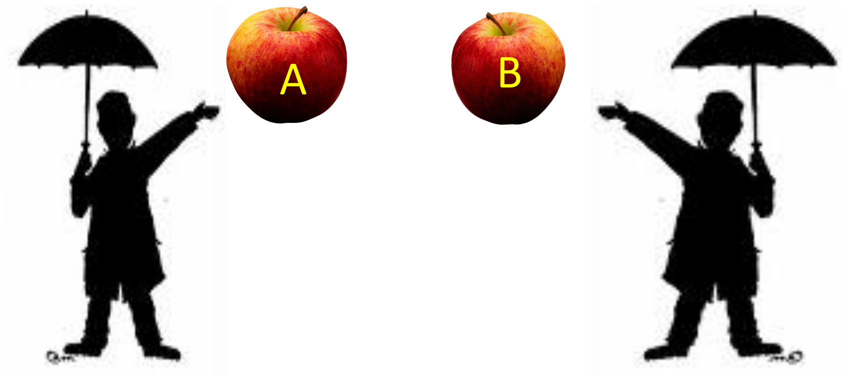
Clock tree area ~ $A_{chip}^{3/2}$

The clock consumed 40% of the chip's power

Diminishing Returns

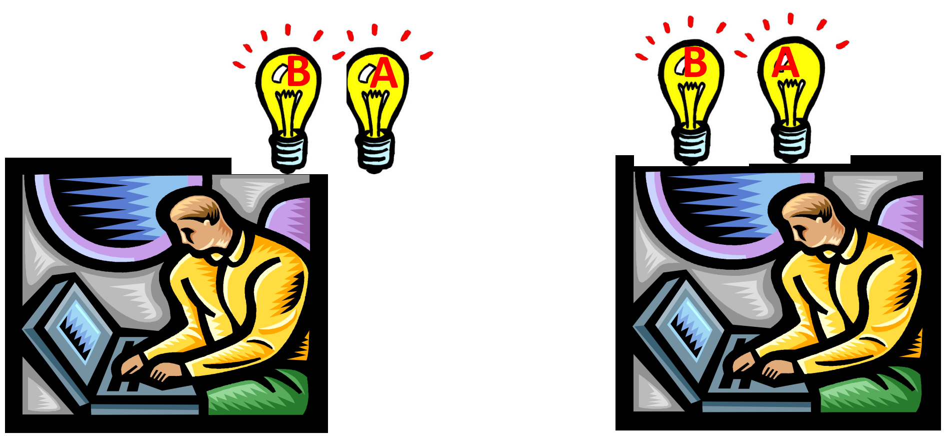


Scaling in Information Networks: Increasing Returns in Information Exchange



If you have an apple and I have an apple and we exchange apples
then you and I will still each have one apple.
But if you have an idea and I have an idea and we exchange these ideas,
then each of us will have two ideas.

--George Bernard Shaw



Important Scaling Differences

1) Information can be copied

2) Information can be communicated locally (Rent's Rule)

3) Technology is (barely) still improving

Transistors are getting smaller

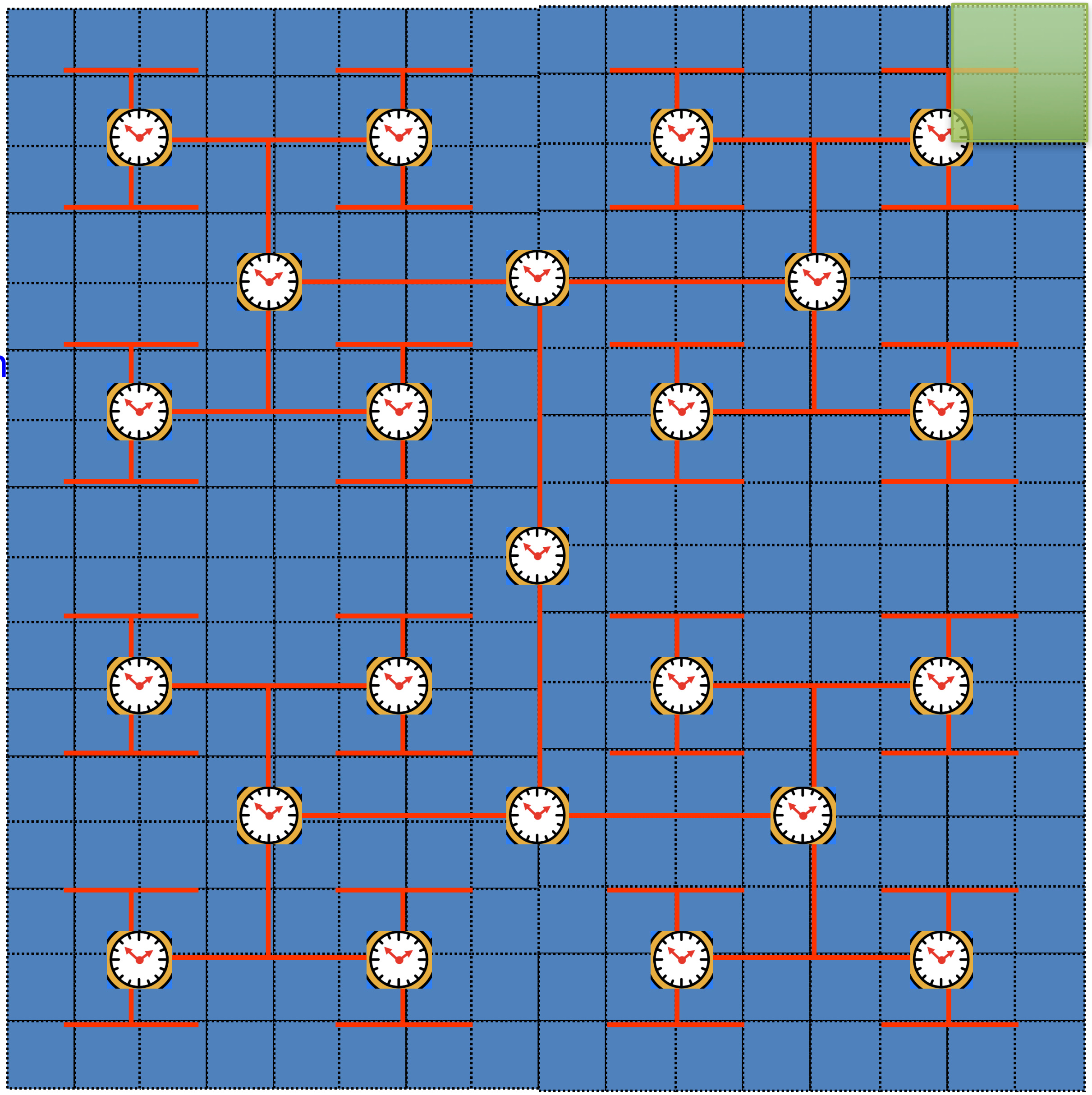
Partially Decentralized Intel Itanium 2 H-tree (2004)

Information can be copied

Amplifiers regenerate clock signal at each branch

Decentralized communication generates linear scaling of clock power & area with chip area

Synchronize more transistors with less power

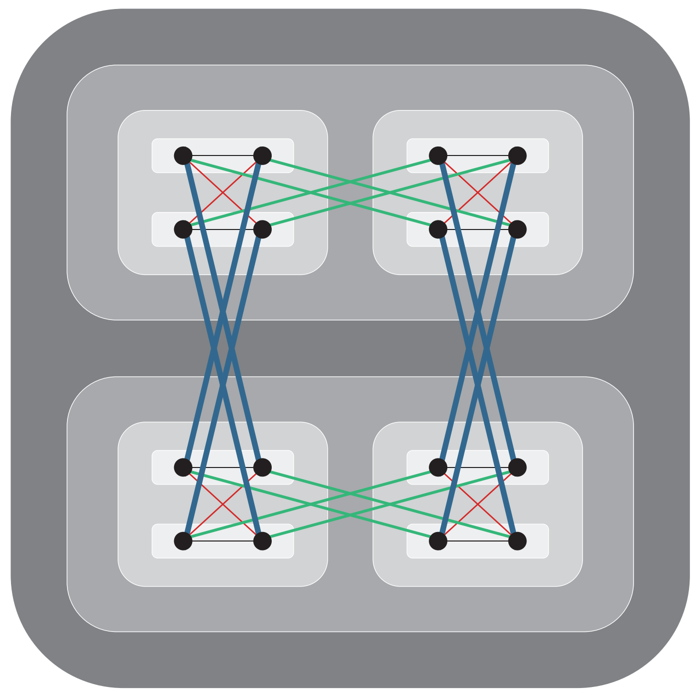


Important Scaling Differences

2) Information can be communicated locally (Rent's Rule)

The probability of communicating is proportional to the distance between the nodes

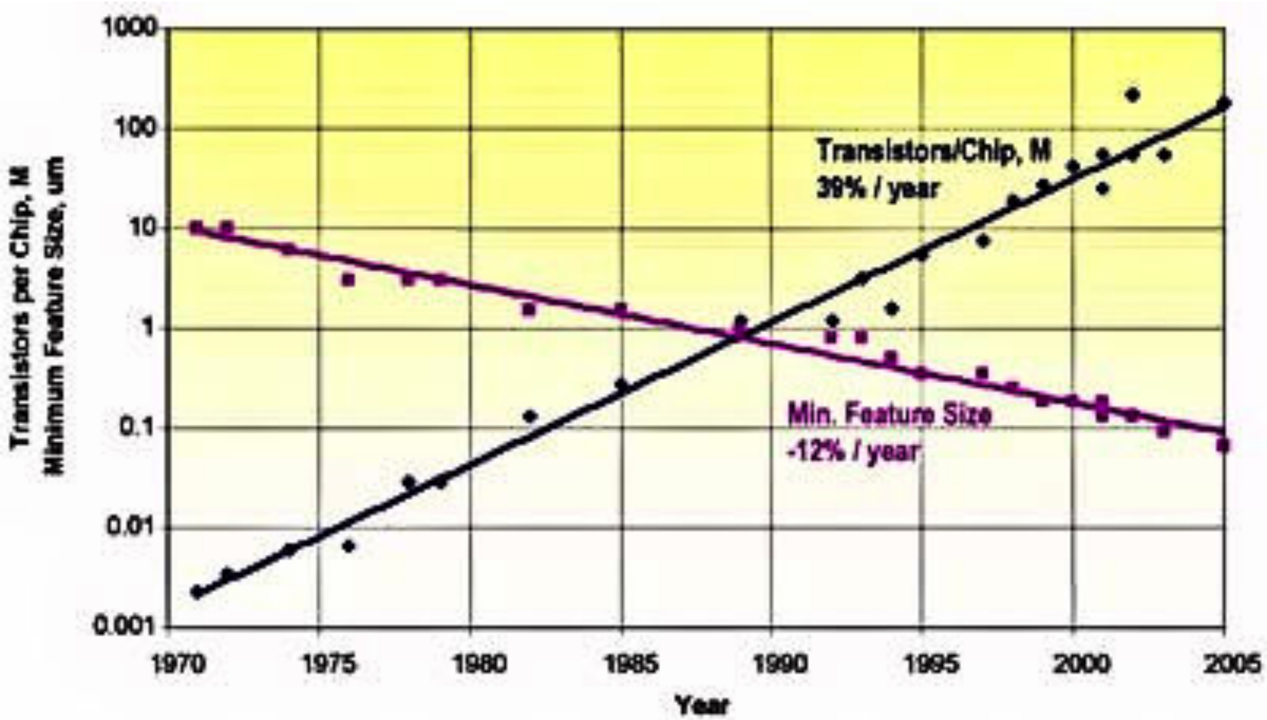
The wire lengths and widths increase as they cross more levels according to D_l and D_r. D_w defines the number of wires, determined by the ratio of internal (intra-module) communication per node to external (inter-module) communication per node. Here D_w 1/4 2 so that a node is connected to all nodes within a module (in this case only 1) by a purple wire, 1/2 of the nodes in the next hierarchical level by red wires, 1/4 of the nodes in the next level by green wires, and 1/8 of the nodes in the next level by blue wires.



Important Scaling Differences

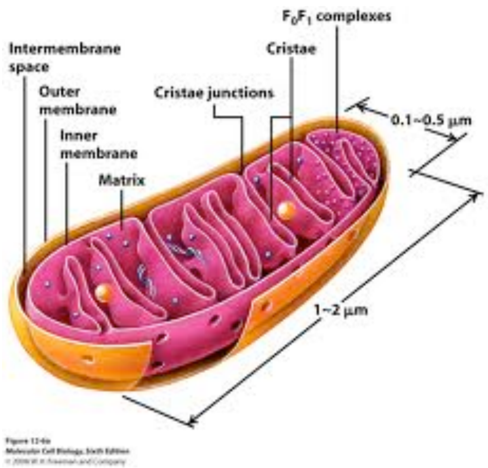
3) Technology is (barely) still improving

– Transistors are getting smaller

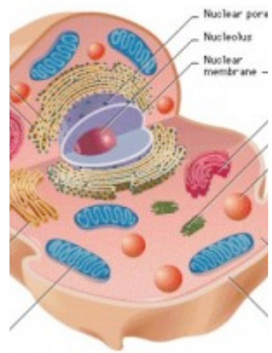


http://www.embedded.com/design/mcus-processors-and-socs/4217276/Reduce-SoC-device-package-leakage-power-with-improved-power-management-protocols

Wednesday, June 21

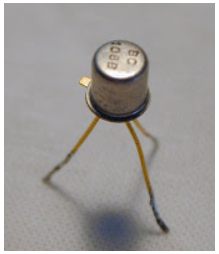


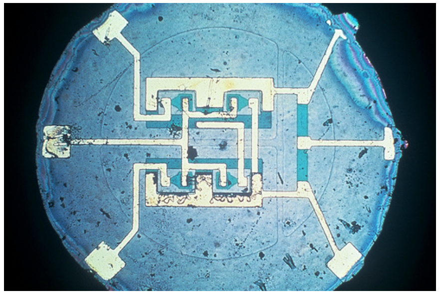
Mitochondrion

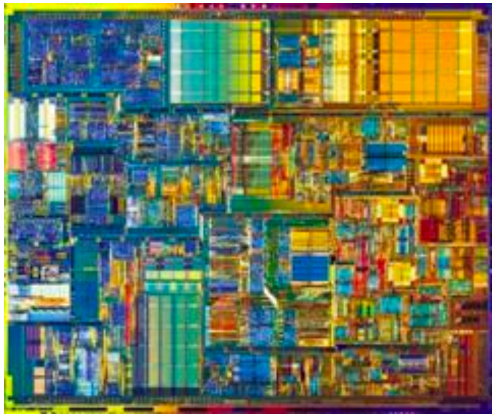


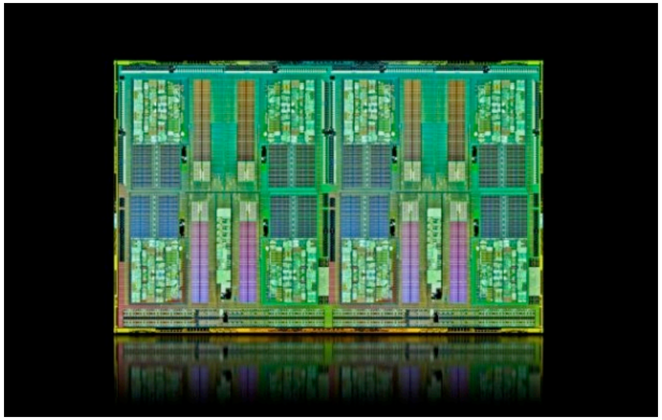
1 to 1000s of
Mitochondria per cell

Trillions of
mitochondria









Transistor

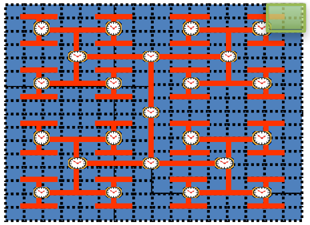
Integrated circuit

Pentium, millions of
transistors

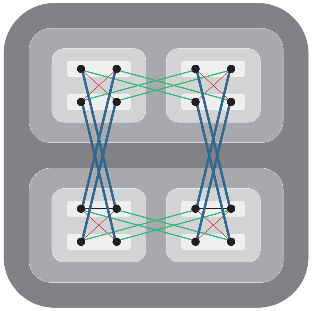
AMD Opteron multi-core chip
billions of transistors

Important Scaling Differences

1) Information can be copied

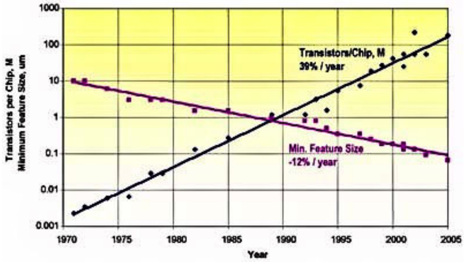


2) Information can be communicated locally (Rent's Rule)

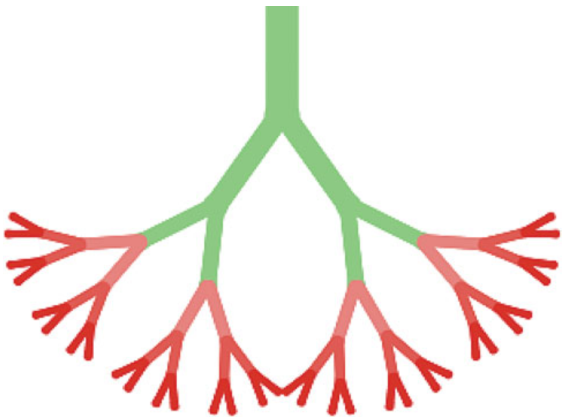


3) Technology is (barely) still improving

Transistors are getting smaller



Fractal Network Differences



Animals

D_r = ? > 2 (slowing)

D_l = 3 (volume)

$$r_i = r_0 \lambda^{i/D_r}$$

$$l_i = l_0 \lambda^{i/D_l}$$

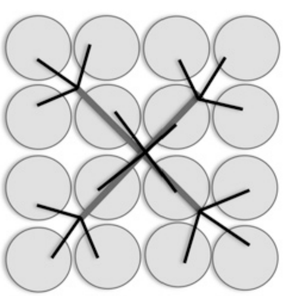
$$w_i = w_0 \lambda^{i/D_w}$$

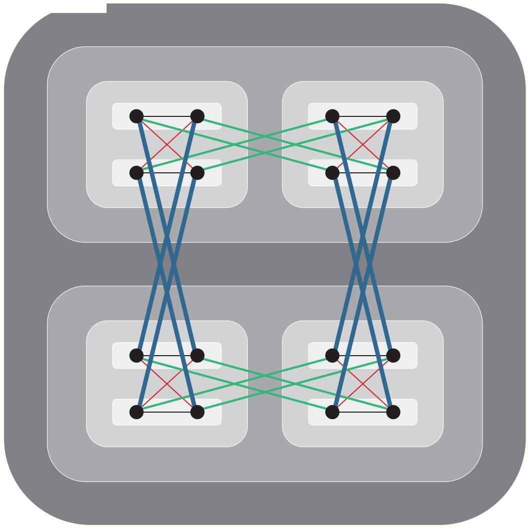
Chips

D_r = ?

D_l = 2 (area)

D_w = ? (Rents Rule)





Scaling Model Assumptions

$$\min_{D_r, D_w, D_l} (E_{sys} \times T_{sys})$$

- Living systems and computer chips are designed to maximize the rate at which resources are delivered to terminal nodes of a network and to minimize the energy dissipated as it is delivered and processed.
  - Minimize Energy dissipation & Delivery Time
    (Minimize the energy-time product)
  - Explicitly consider energy & time in the network AND nodes
  - match supply and demand (pipelining)

- Biology: minimize energy dissipated in the network & maximize metabolic rate
- Computers: minimize total energy consumption on the chip and maximize rate that bits are processed (MIPS)

$$\min_{D_r, D_w, D_l} (E_{sys} \times T_{sys})$$

$$E_{net} \propto N^{1-1/D_l} \sum_{i=0}^{H} \lambda^{i(1/D_l + D_w - 1)}$$

$$T_{net} \propto R_0 C_0 \propto \frac{l_0^2}{r_0^2} \propto N^0$$

$$E_{net} \propto N^{1-1/D_l} \sum_{i=0}^{H} \lambda^{i(1/D_l + D_w - 1)}$$

$$T_{node} \propto N^{-1/D_l}$$

D_l = 2 (area-filling in 2D chips)

To minimize the E_{sys} x T_{sys}:

D_r = 2

D_w <= 2

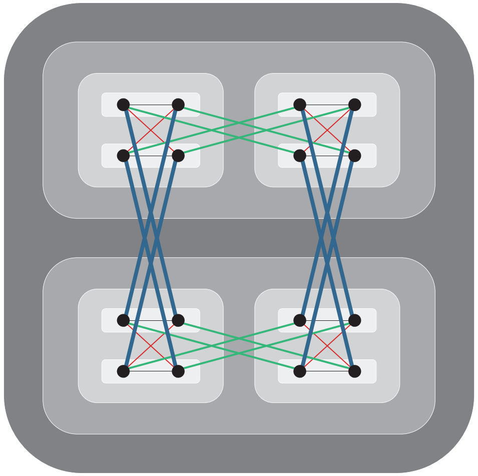
	general	energy – time minimization
mammals		
E_{net}	$l_0 u_0 N^{2/D_r - 1}$	$N^{1/12}$
E_{node}	N	N
T_{net}	$u_0^{-1} N^{1-2/D_r}$	N^0
T_{node}	$u_0^{-1} N^{1-2/D_r}$	N^0
$E_{sys} \times T_{sys}$	$l_0 + u_0^{-1} N^{2-2/D_r}$	$N^{1/12} + N$
computers		
E_{net}	N^{1-1/D_l}	$N^{1/2}$
E_{node}	N^{1-1/D_l}	$N^{1/2}$
T_{net}	N^0	N^0
T_{node}	N^{-1/D_l}	$N^{-1/2}$
$E_{sys} \times T_{sys}$	$N^{1-1/D_l} + N^{1-1/D_l}$	$N^{1/2} + N^{1/2}$

Power

$$P = \frac{E_{sys}}{T_{sys}} \propto N^{1/2}$$

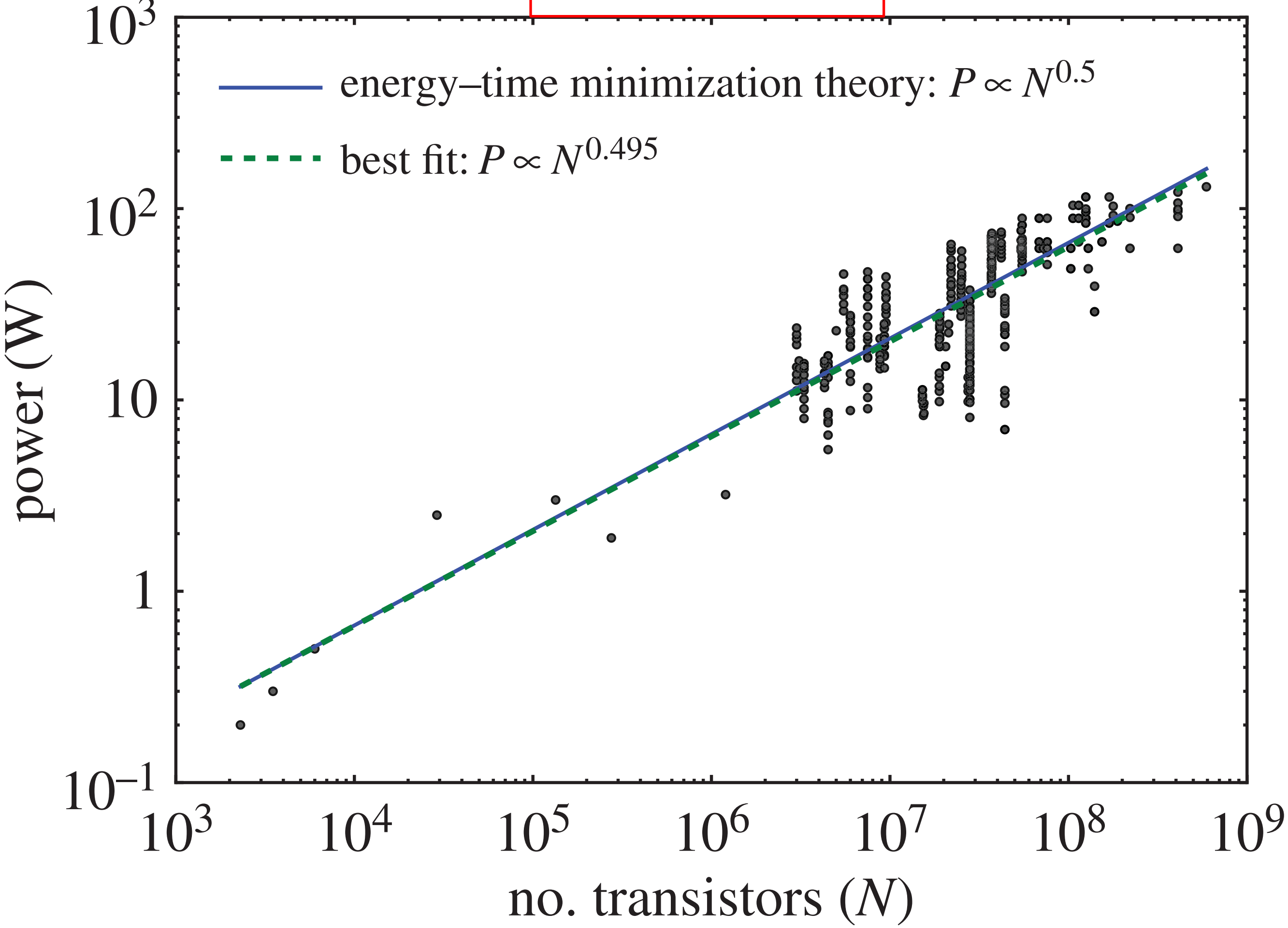
Throughput

$$Tp \propto \frac{N}{T_{sys}} \propto N$$



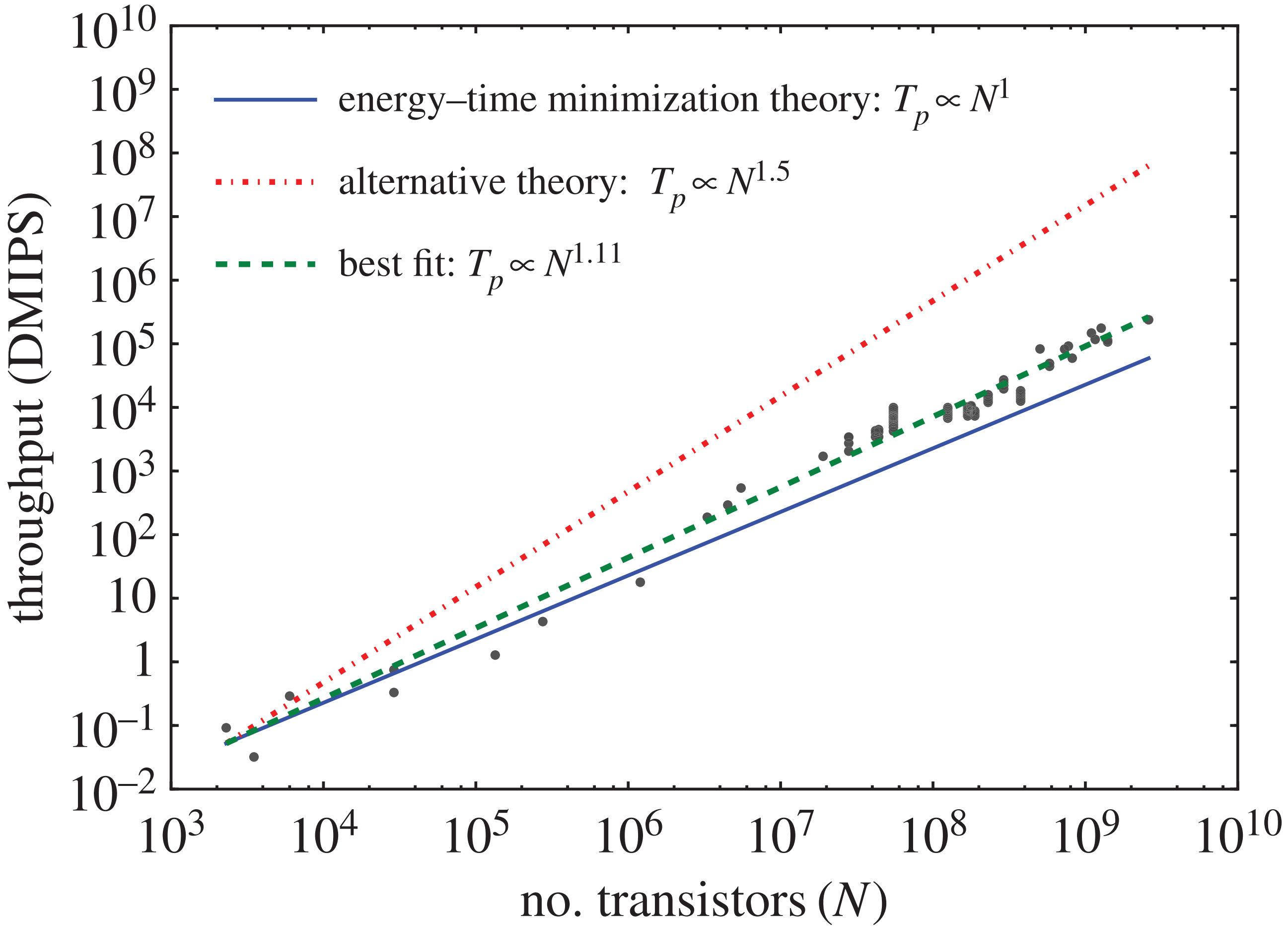
Power vs Size of Microprocessors

$$P = \frac{E_{sys}}{T_{sys}} \propto N^{1/2}$$



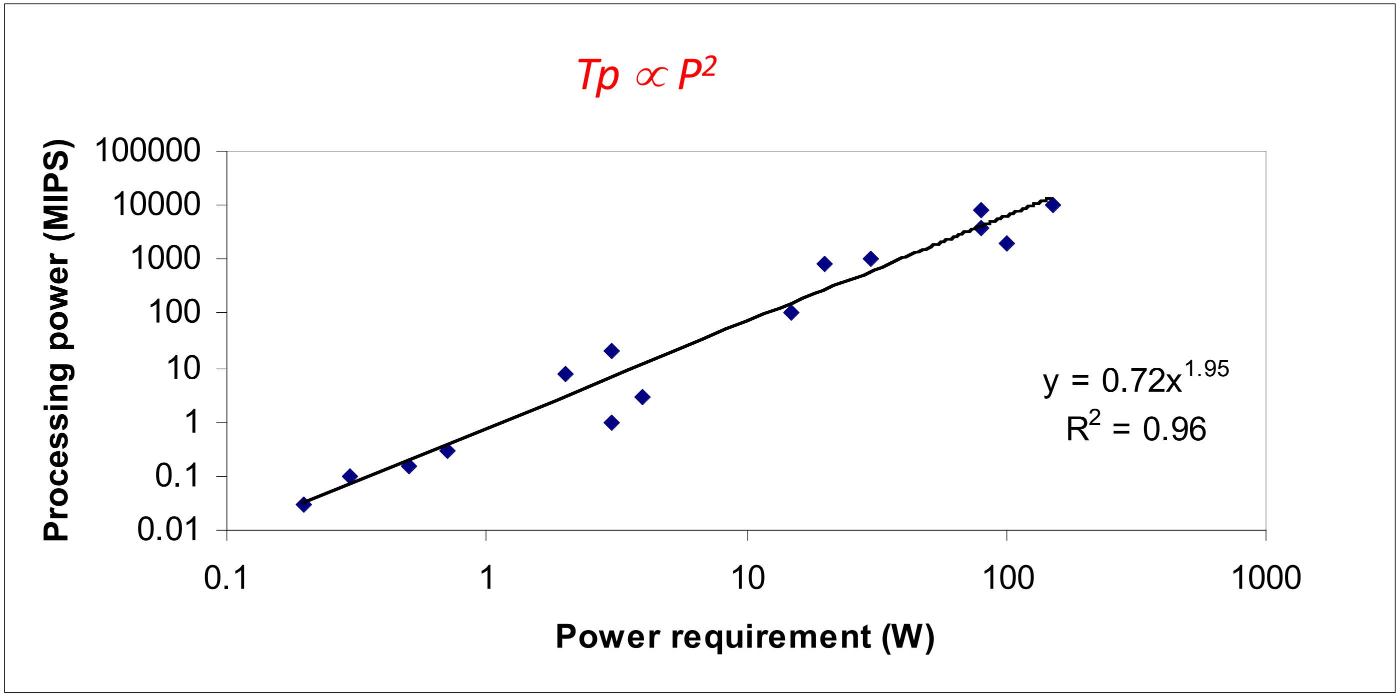
“Hegemony of the network”
Linear scaling between throughput and # of transistors

$$Tp \propto \frac{N}{T_{\text{sys}}} \propto N.$$



Power scaling: Increasing returns

Thousand-fold increase in power, Million-fold increase in MIPS



1970 **2005**

In 1970, 100 Watts powered 15 MIPS. In 2005, 6700 MIPS

Transistors perform computations
Power consumption is dominated by wires

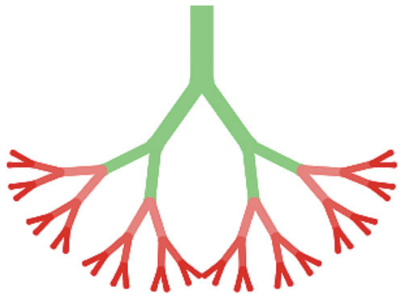
Energy Time minimization in Biological Scaling

- Minimizing time and energy dissipation in network & nodes
- Match the delivery rate by network to consumption rate in nodes
- Allow blood velocity to slow

$$\min(E_{\mathrm{sys}} \times T_{\mathrm{sys}}) = \min_{D_{\mathrm{r}}, D_{\mathrm{w}}, D_{\mathrm{l}}} \left(RN + \frac{N^2}{Q} \right).$$

	general	energy – time minimization
mammals		
E_{net}	$l_0 u_0 N^{2/D_r - 1}$	$N^{1/12}$
E_{node}	N	N
T_{net}	$u_0^{-1} N^{1-2/D_r}$	N^0
T_{node}	$u_0^{-1} N^{1-2/D_r}$	N^0
$E_{\mathrm{sys}} \times T_{\mathrm{sys}}$	$l_0 + u_0^{-1} N^{2-2/D_r}$	$N^{1/12} + N$
computers		
E_{net}	N^{1-1/D_l}	$N^{1/2}$
E_{node}	N^{1-1/D_l}	$N^{1/2}$
T_{net}	N^0	N^0
T_{node}	N^{-1/D_l}	$N^{-1/2}$
$E_{\mathrm{sys}} \times T_{\mathrm{sys}}$	$N^{1-1/D_l} + N^{1-1/D_l}$	$N^{1/2} + N^{1/2}$

Energy dissipation in the network is minimized when D_r = 2 (area preserving branching)
Energy * time is minimized when $D_{\mathrm{r}} = 24/11 = 2.18$.

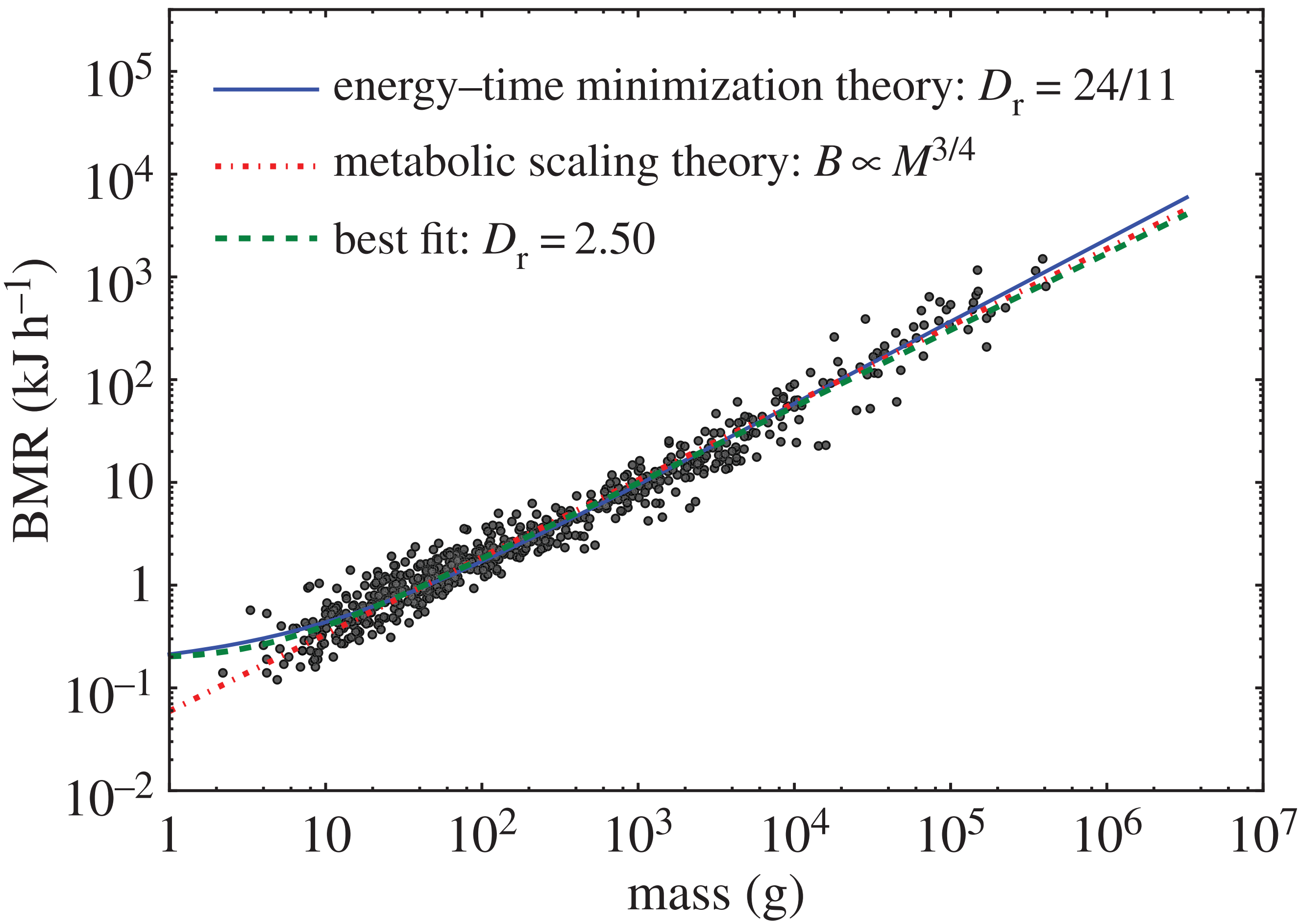


We predict the optimal D_r given that

- blood must slow (D_r > 2)

Metabolic Scaling Prediction that accounts for blood slowing

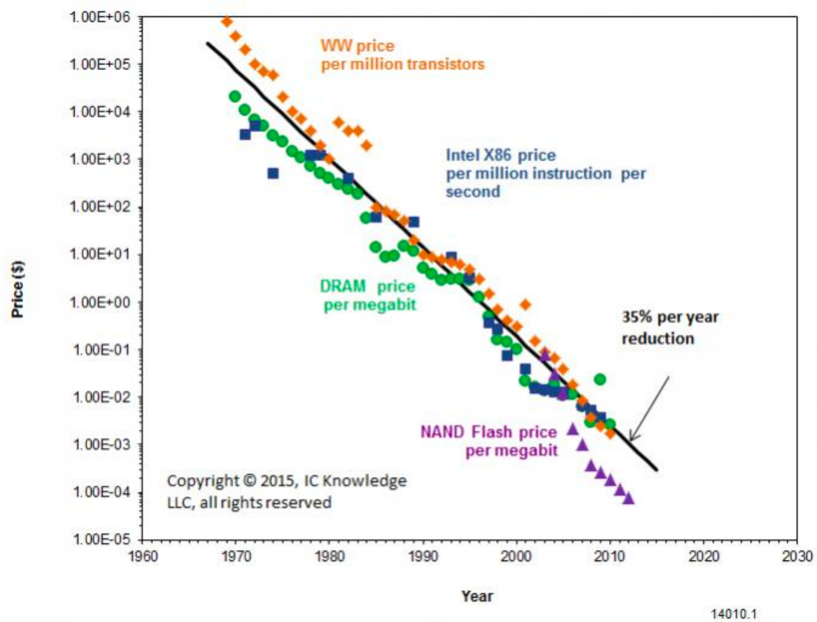
$$B \propto M^{(18-8D_r)/(6+D_r)} + M^{(24-2D_r)/(18-3D_r)}.$$



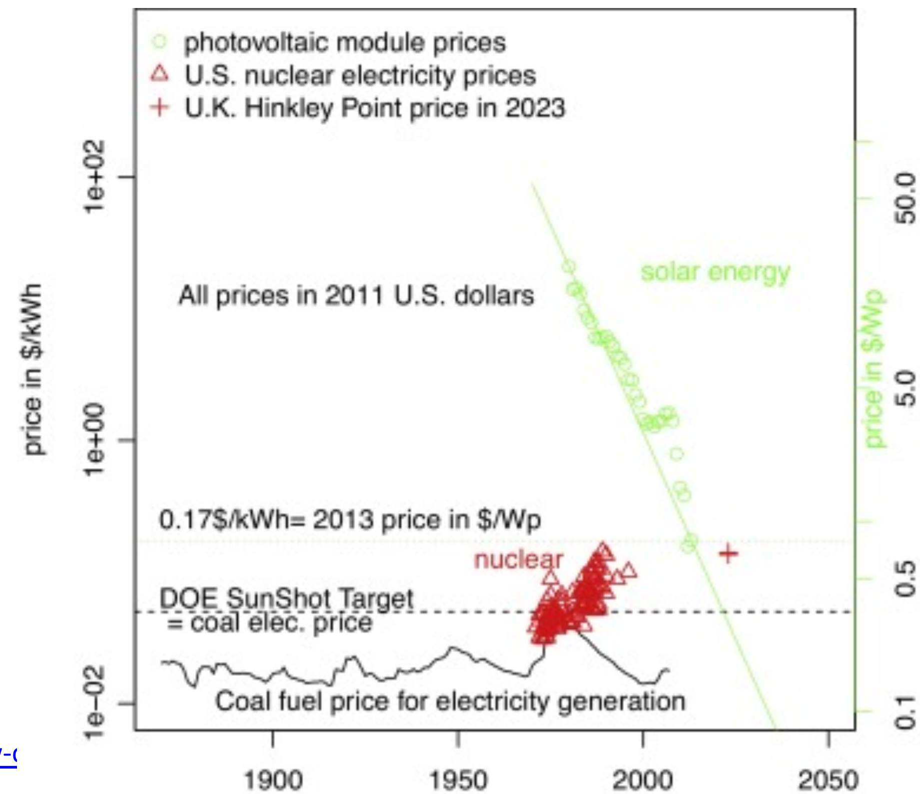
Recap: Scaling theory for Biology and Computation

- Networks that deliver resources are optimized to reduce energy dissipation and increase flow rates: minimizing the energy–time product
- Differences CS & Biology
  - Information can be copied
  - Communication locality (Rent's Rule)
  - Transistor size decrease
- By accounting for slowing and minimizing the energy time product, we explain curvilinearity in metabolic scaling of mammals
- We use the network scaling framework to explain scaling patterns in microprocessors. This result corresponds to ideal scaling, as suggested by Dennard, where the linear dimensions of transistors and wires scale at the same rate, wire delay is constant, and Rent's exponent is 1/2.
- Show that Rent's exponent is ½ in the sense that any further decrease in communication locality has minimal impact on minimizing the energy time product.

Moore's Law is not just for computers
Cost of Photovoltaics Scales



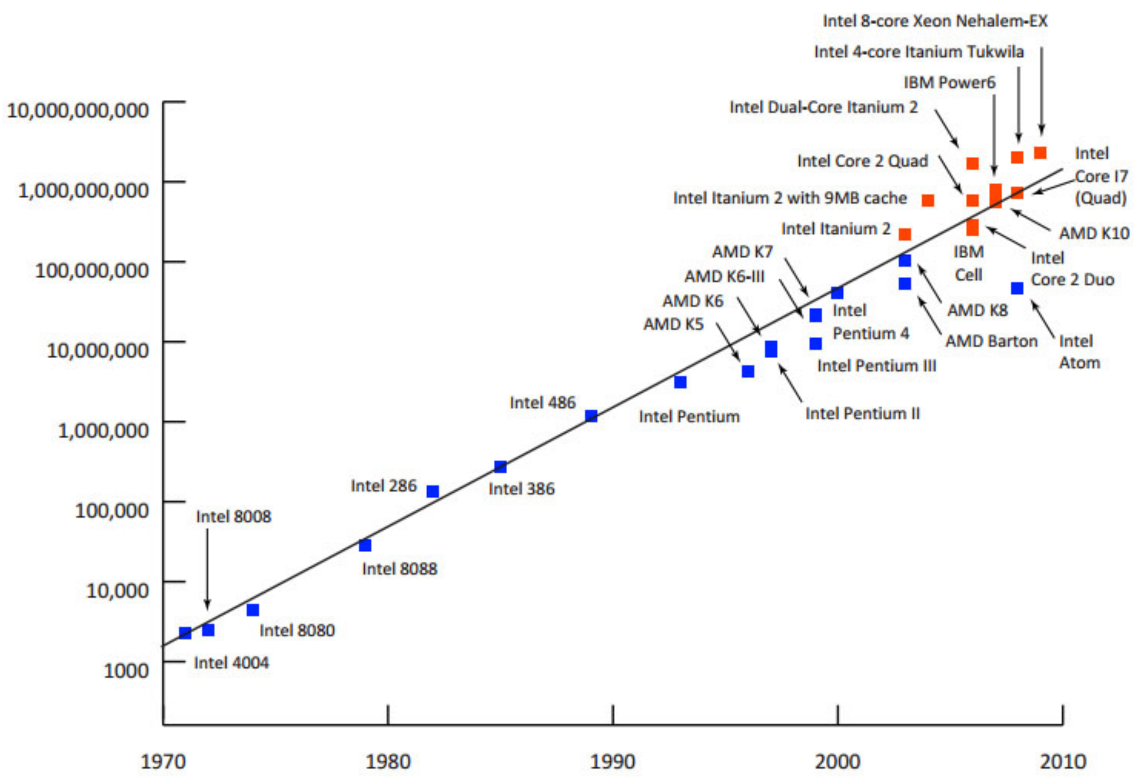
https://www.semiwiki.com/forum/content/4522-moore%C2%92s-law-long-live-moore%C2%92s-law-%C2%96-part-1-a.html

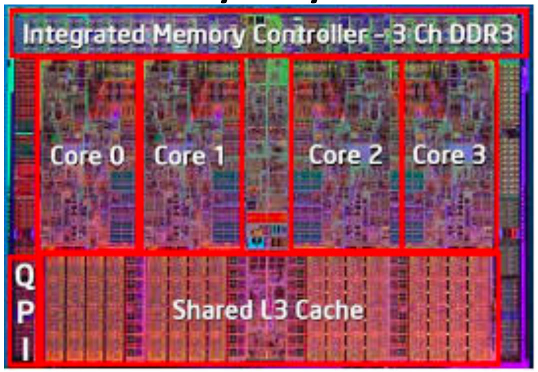


J. Doyne Farmer, François Lafond 2015

“scaling” in computing power: Moore’s Law

The number of transistors per chip doubles every 2 years



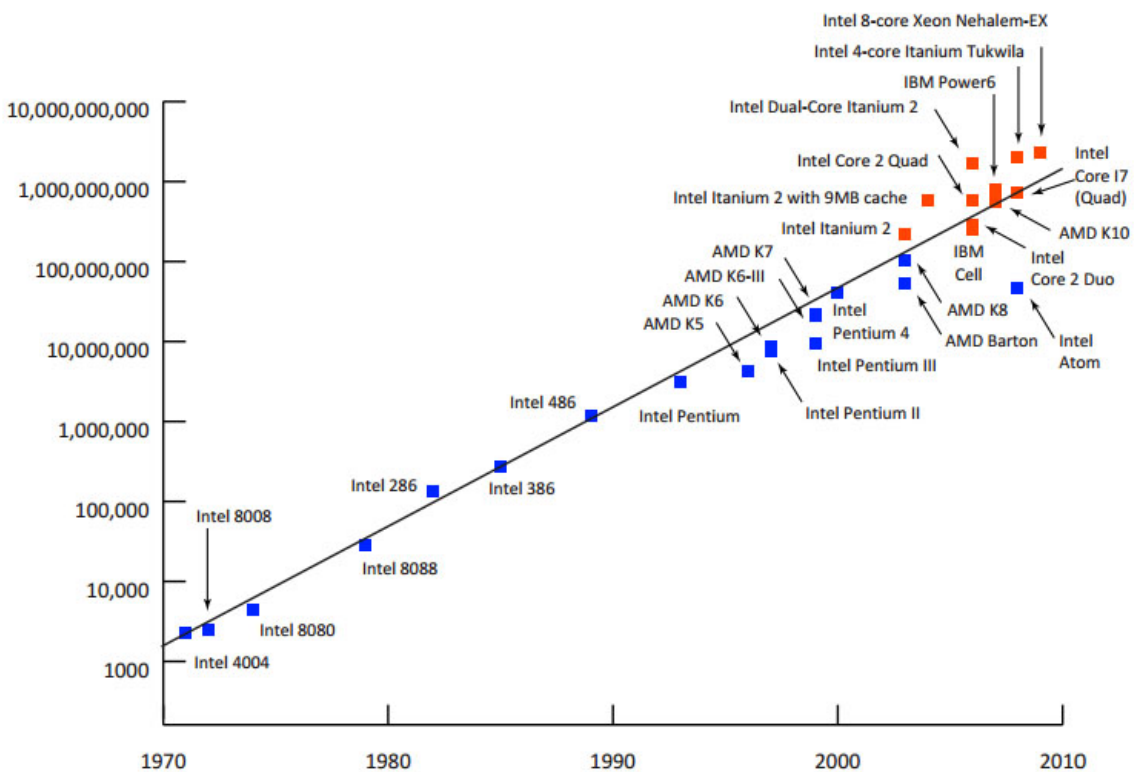


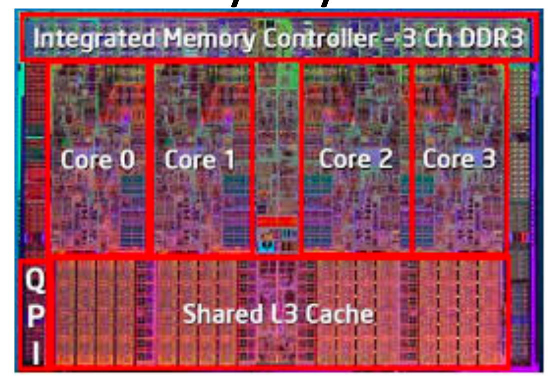
Transition to Multicore

http://www.embedded.com/design/programming-languages-and-tools/4375996/Using-Java-for-multicore-programming-complexity--Part-1

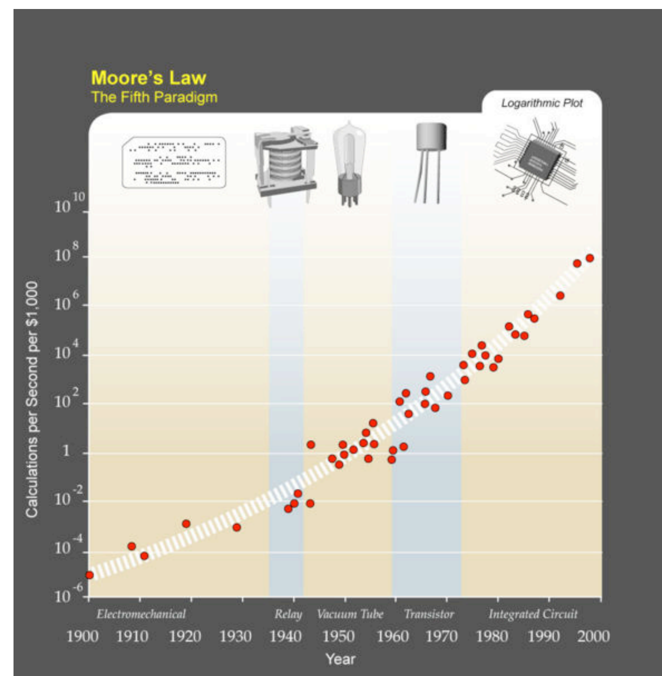
“scaling” in computing power: Moore’s Law

The number of transistors per chip doubles every 2 years





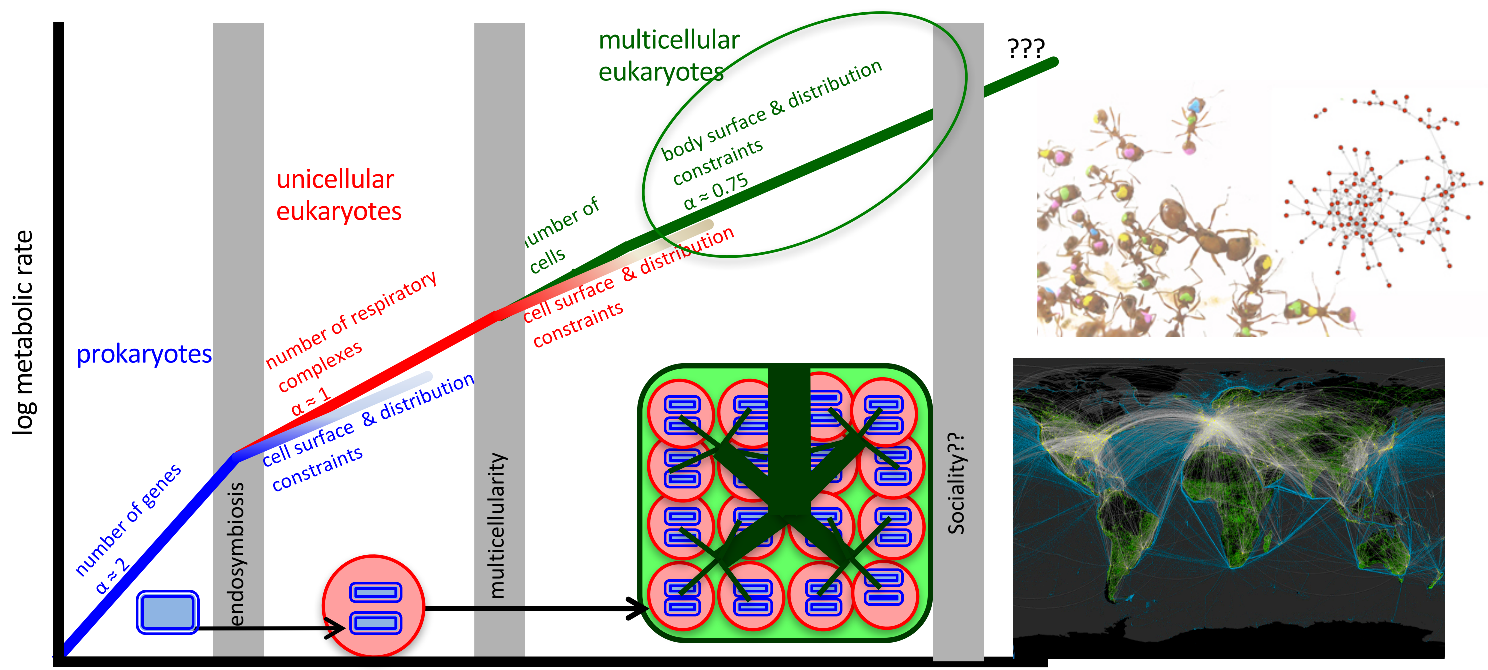
Transition to Multicore



http://www.embedded.com/design/programming-languages-and-tools/4375996/Using-Java-for-multicore-programming-complexity--Part-1

Scaling intercepts and slopes shift after evolutionary & technological innovations

- Innovations in chip components mimic innovation in the evolution of bacteria
- Single-core chip scaling mimics unicell scaling
- Multi-core chips echo the transition to multicellularity
- Multi-agent computation as a model for scaling in social systems?
- Computer scaling deviates from animal metabolic scaling in part due to decentralization
    - Decentralized designs dominate in the transition to sociality



Delong et al PNAS 2010