

A note on power efficiency and income in data center networks

A. F. J. Levi

University of Southern California

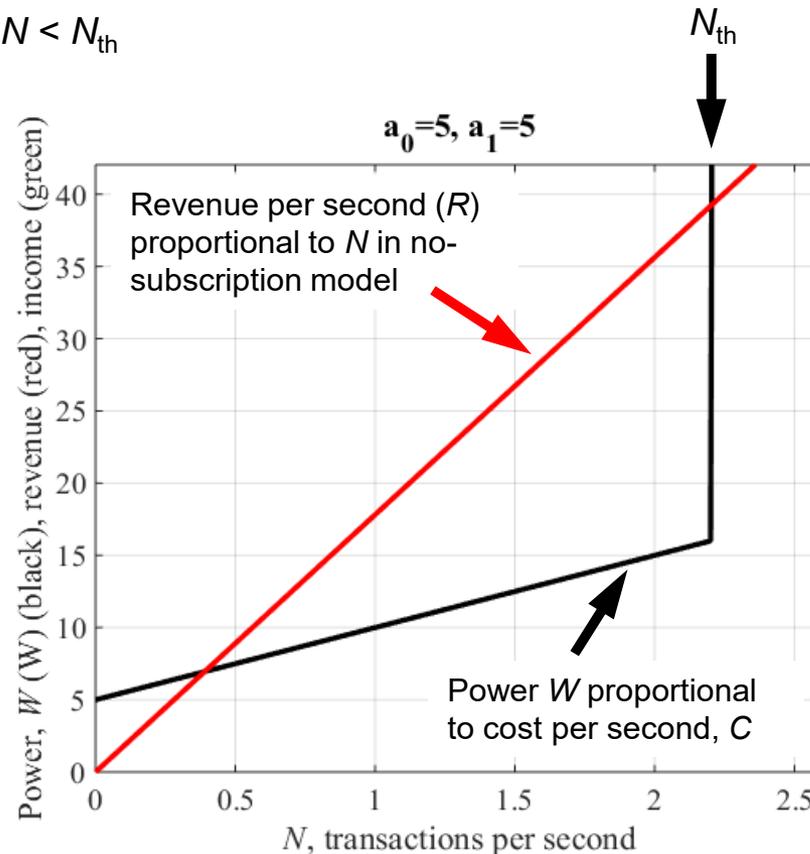
ARPA-E meeting, Seattle, October 23-24, 2018

A single performance metric may not capture opportunity to improve system efficiency

- Let n = number of transactions
- There is ambiguity in n as it may refer to:
 - part of, or average over, a specific application
 - part of, or average over, multiple applications
 - a specific server
 - an average over a rack
 - an average over a cluster of racks
 - an average over the entire data-center.
 - ...
- Number of transactions per second, $N = n/s$
- Datacenter consumes W power for given N with maximum energy efficiency at $\min(W/N)$
- At peak energy-efficiency in the system, single metric is average number of transactions per Joule, $\varepsilon = n / J = N / W$
- This may not correspond to the operating point that maximizes income
- A commercial business is likely to respond to financial incentives
 - Is there a way to tie energy efficiency to income?

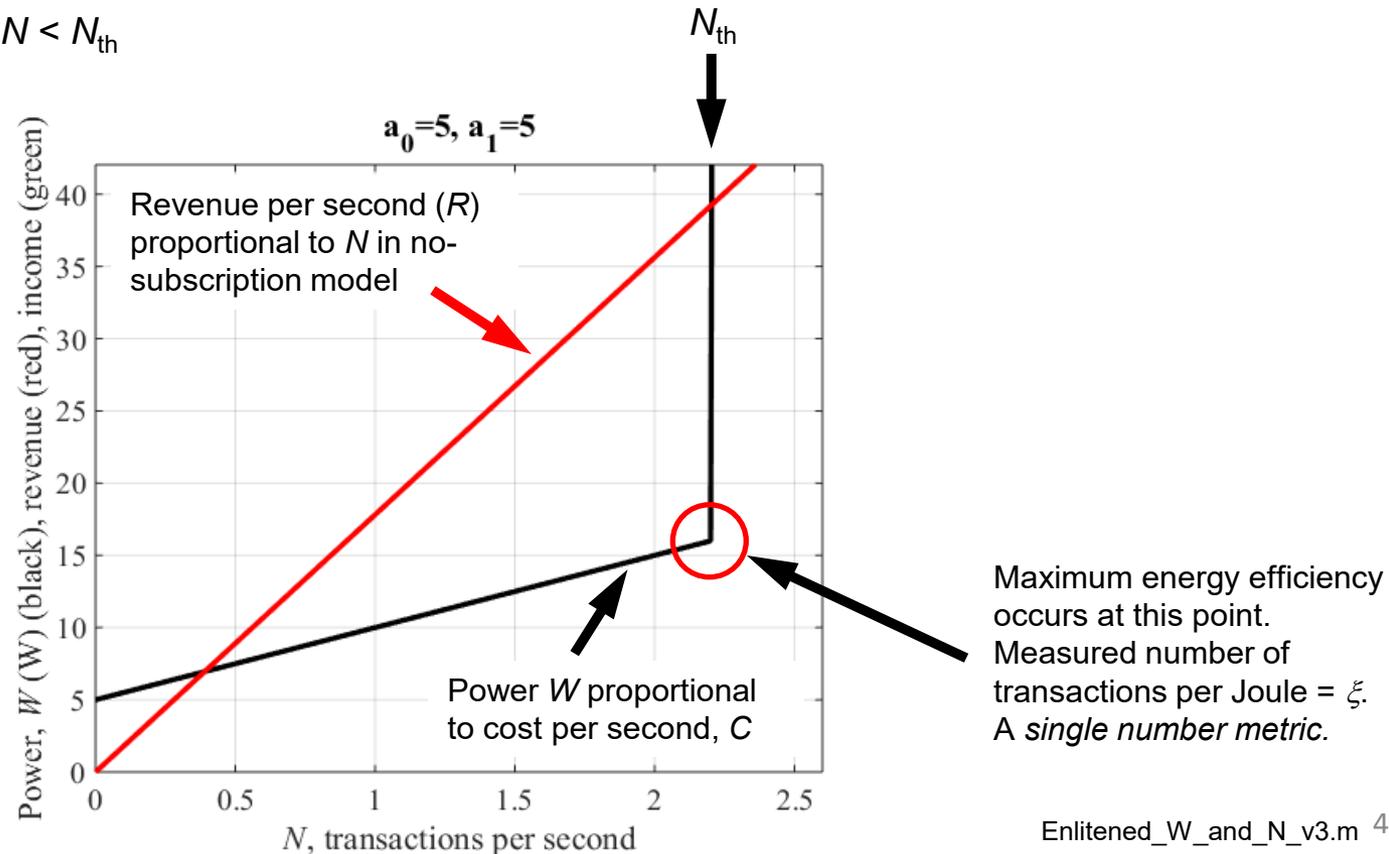
An extreme business model with commodity pricing of transactions per second

- Assume gross revenue per second (R , red curve) is proportional to number of transactions per second ($n/s = N$)
 - Transactions are treated as a **commodity** and there is no offset due to customer subscription
- Assume only significant cost is electricity for running the datacenter
 - This is only for illustrative purposes and it is assumed electricity is priced as a commodity
- Assume simplest *non-linear* relationship between power (W) consumed by the datacenter and average number of transactions per second (N) is characterized by an offset, slope and cut-off at N_{th}
 - $W = a_0 + a_1 N$, for $0 < N < N_{th}$



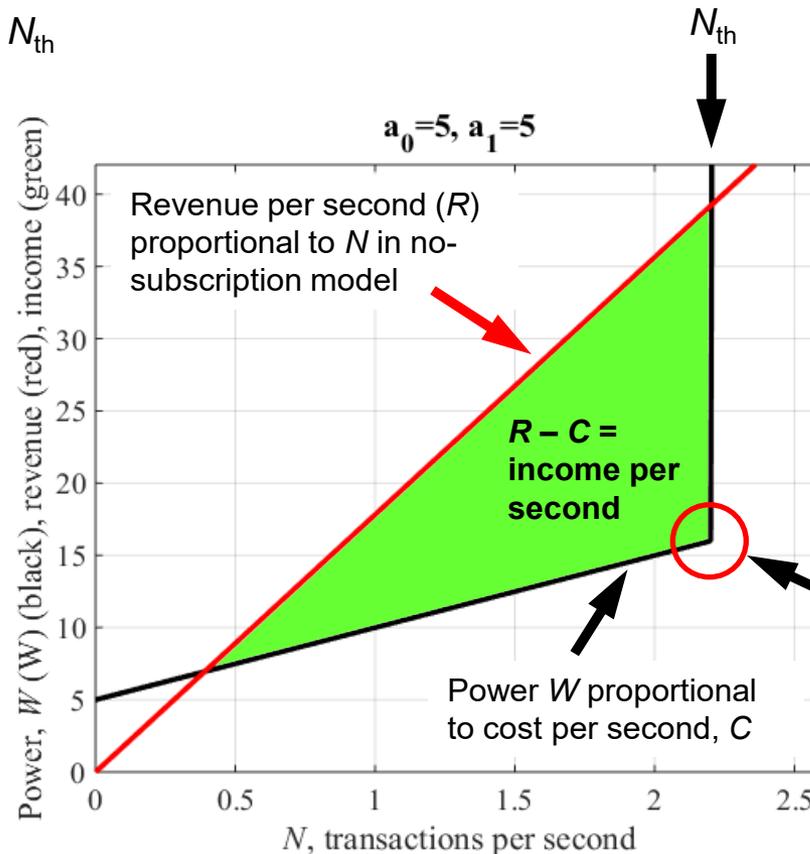
An extreme business model with commodity pricing of transactions per second

- Assume gross revenue per second (R , red curve) is proportional to number of transactions per second (N)
 - Transactions are treated as a **commodity** and there is no offset due to customer subscription
- Assume only significant cost is electricity for running the datacenter
 - This is only for illustrative purposes and it is assumed electricity is priced as a commodity
- Assume simplest *non-linear* relationship between power (W) consumed by the datacenter and average number of transactions per second (N) is characterized by an offset, slope and cut-off at N_{th}
 - $W = a_0 + a_1 N$, for $0 < N < N_{th}$



An extreme business model with commodity pricing of transactions per second

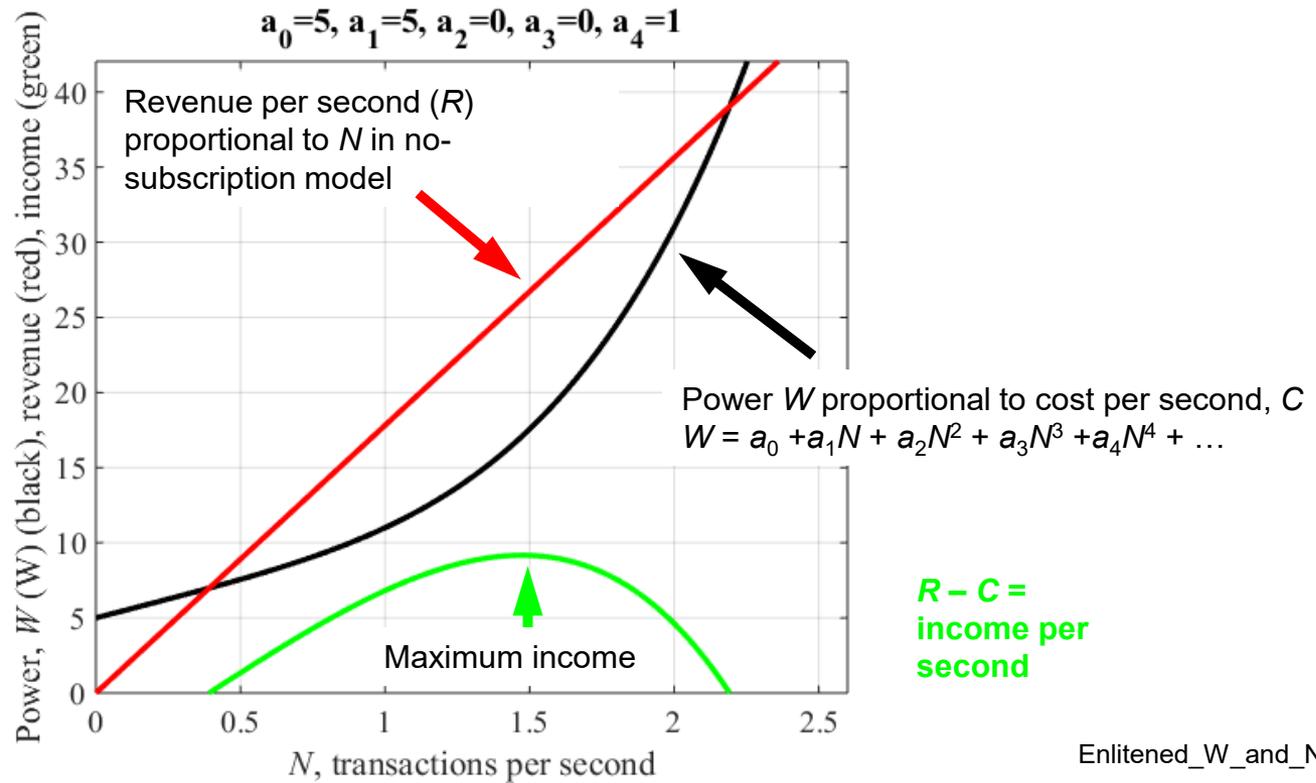
- Assume gross revenue per second (R , red curve) is proportional to number of transactions per second (N)
 - Transactions are treated as a **commodity** and there is no offset due to customer subscription
- Assume only significant cost is electricity for running the datacenter
 - This is only for illustrative purposes and it is assumed electricity is priced as a commodity
- Assume simplest *non-linear* relationship between power (W) consumed by the datacenter and average number of transactions per second (N) is characterized by an offset, slope and cut-off at N_{th}
 - $W = a_0 + a_1 N$, for $N < N_{th}$



Maximum energy efficiency also maximizes income. Measured number of transactions per Joule = ξ . A single number metric.

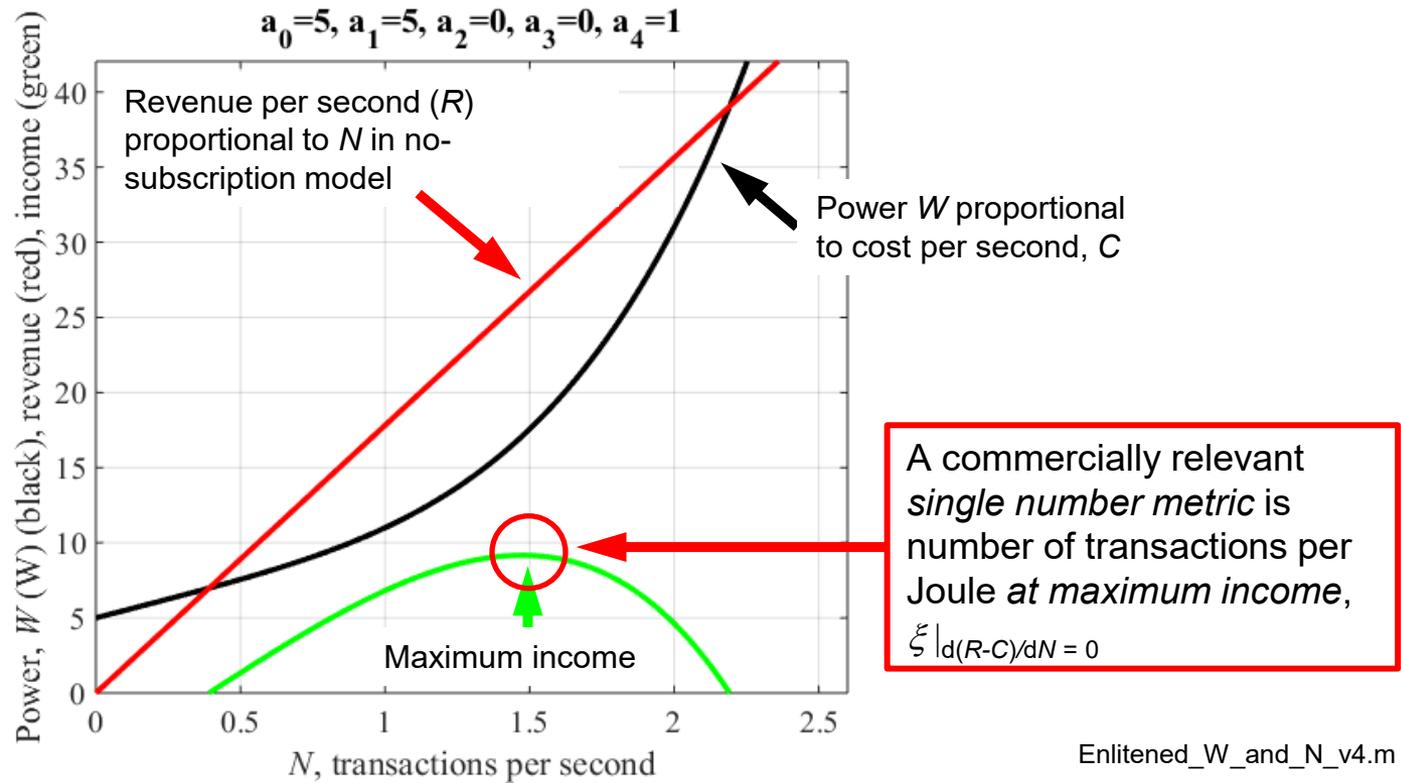
A business model with commodity pricing of transactions per second

- Assume gross revenue per second (R , red curve) is proportional to number of transactions per second (N)
 - Transactions are treated as a **commodity** and there is no offset due to customer subscription
- Assume only significant cost is electricity for running the datacenter
- Assume *non-linear* relationship between power (W) consumed by the datacenter and average number of transactions per second (N) is *characterized by a polynomial*
 - $W = a_0 + a_1N + a_2N^2 + a_3N^3 + a_4N^4 + \dots$
 - $R - C =$ income per second
 - ENLITENED-enabled system reconfiguration can change value of polynomial coefficients, $a_n = a_n(t)$



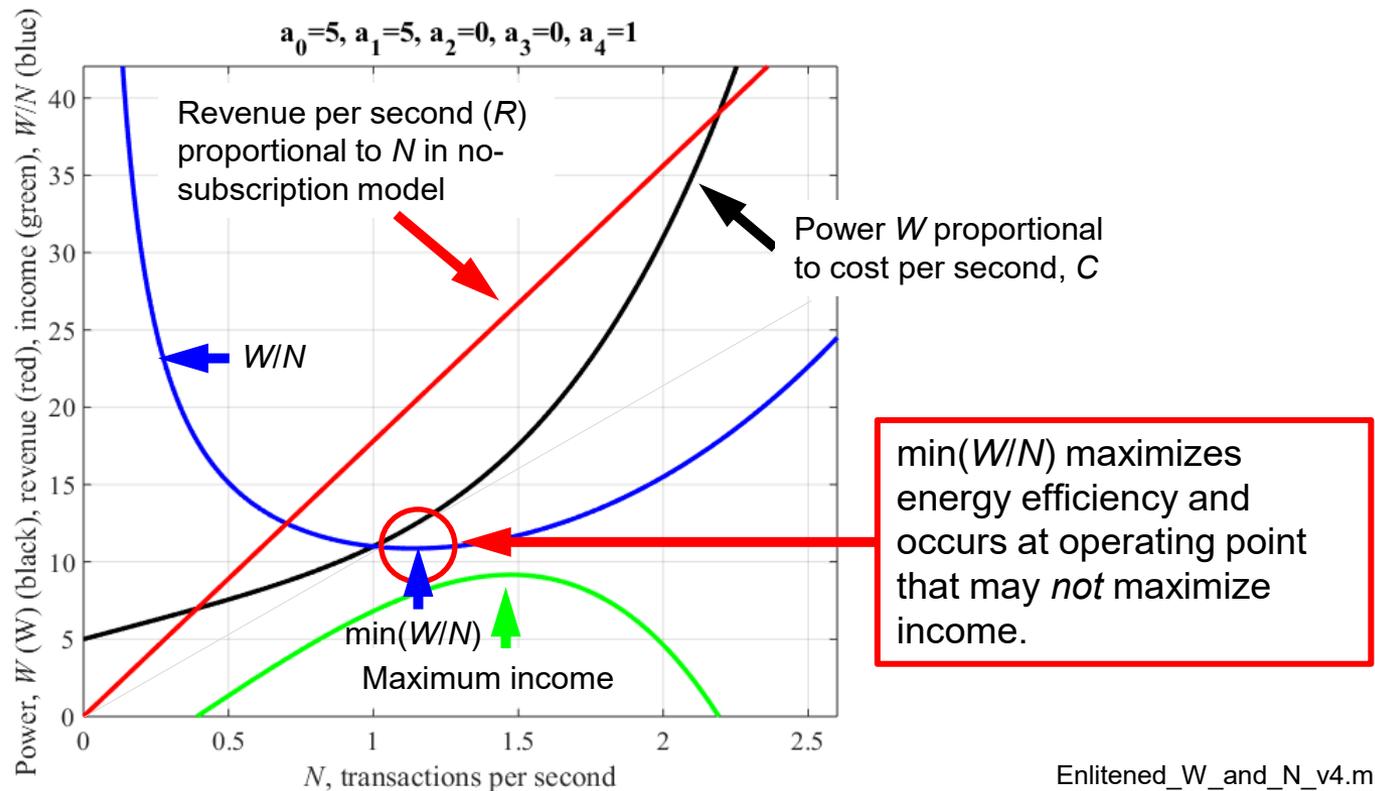
A business model with commodity pricing of transactions per second

- Assume gross revenue per second (R , red curve) is proportional to number of transactions per second (N)
 - Transactions are treated as a **commodity** and there is no offset due to customer subscription
- Assume only significant cost is electricity for running the datacenter
- Assume *non-linear* relationship between power (W) consumed by the datacenter and average number of transactions per second (N) is *characterized by a polynomial*
 - $W = a_0 + a_1N + a_2N^2 + a_3N^3 + a_4N^4 + \dots$
 - Maximum income when $d(R-C)/dN = 0$



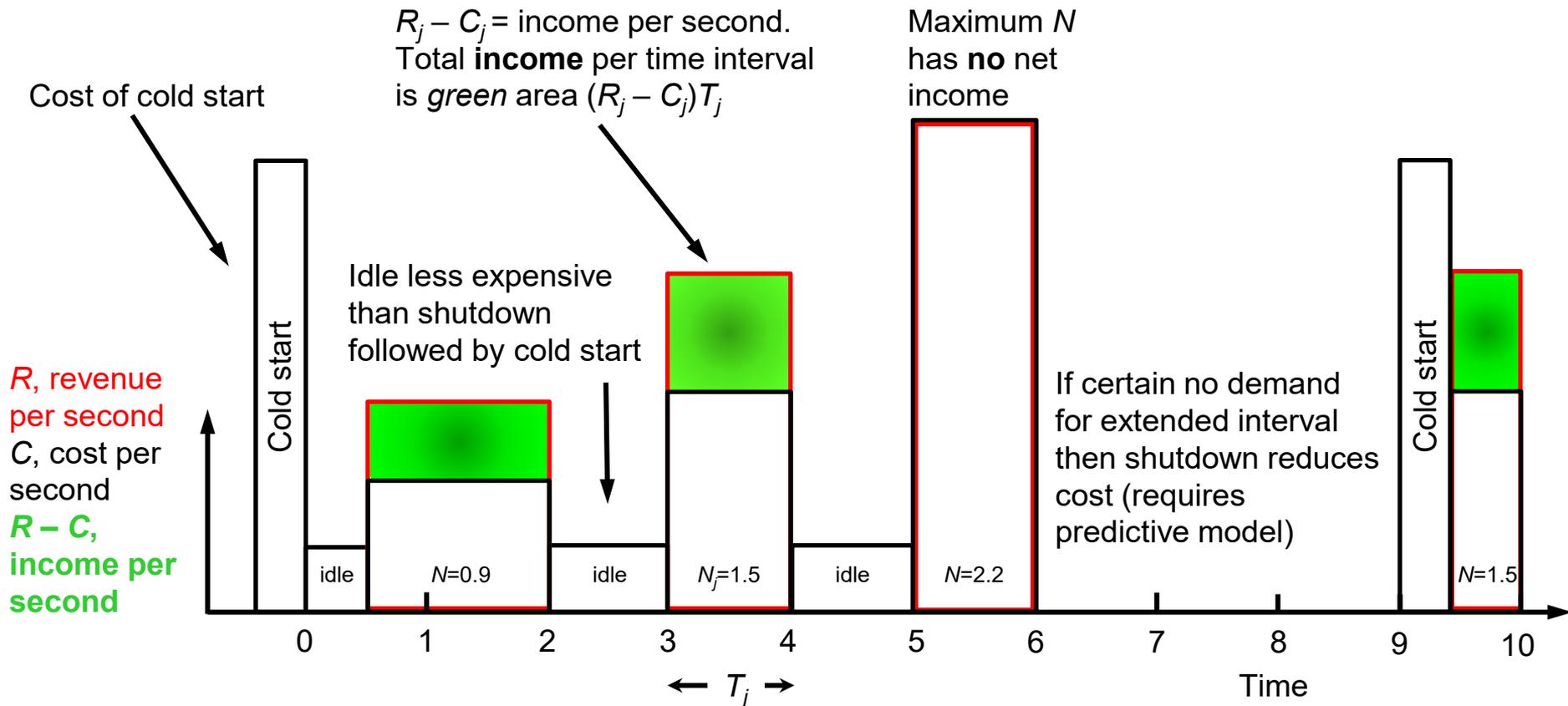
A business model with commodity pricing of transactions per second

- Assume gross revenue per second (R , red curve) is proportional to number of transactions per second (N)
- Assume only significant cost is electricity for running the datacenter
- Assume *non-linear* relationship between power (W) consumed by the datacenter and average number of transactions per second (N) is *characterized by a polynomial*
 - $W = a_0 + a_1N + a_2N^2 + a_3N^3 + a_4N^4 + \dots$
 - In general for $\{a_n > 0\}$ maximum income ($d(R-C)/dN = 0$) occurs at operating point not coincident with maximum energy efficiency ($\min(W/N)$)
 - There is a *gap* between maximizing *income* and maximizing *energy-efficiency*



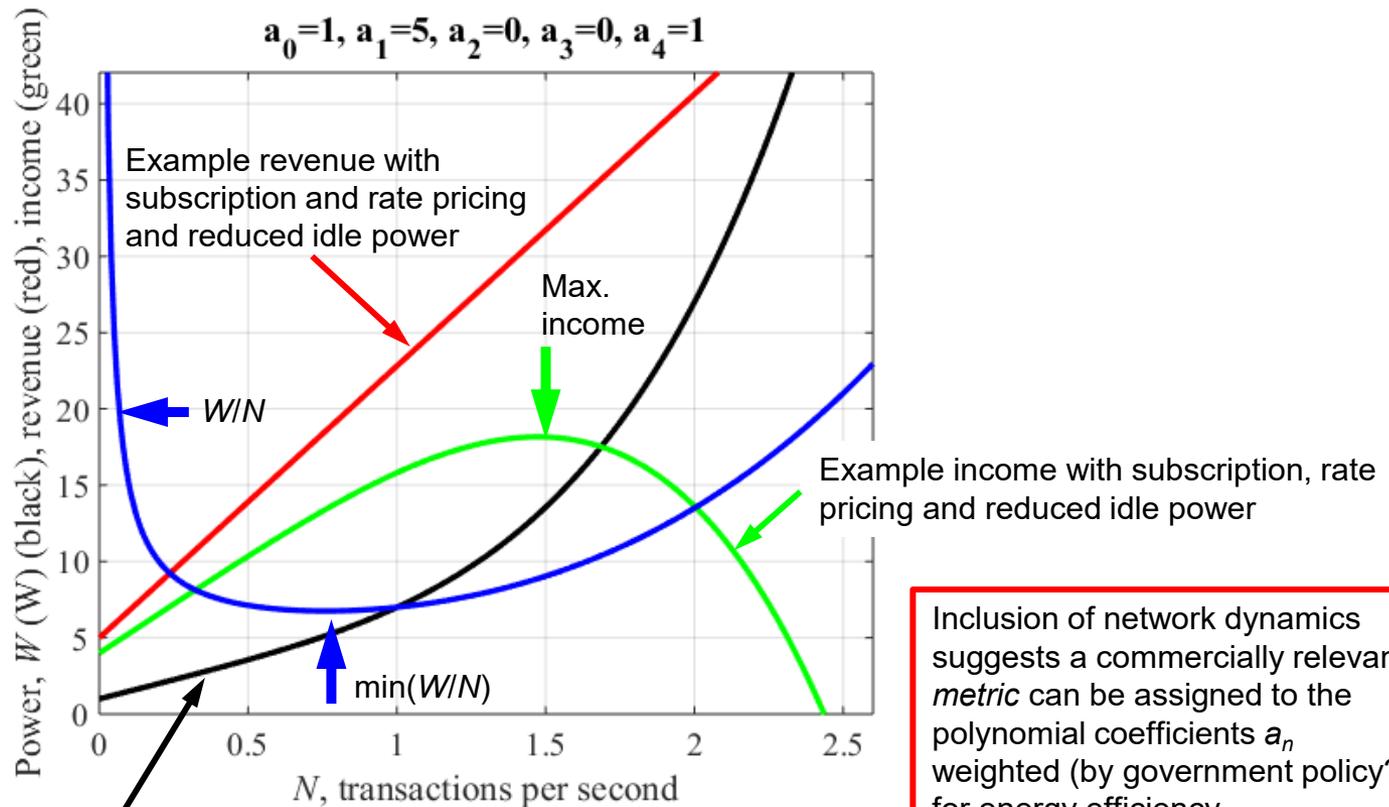
Variation in revenue, cost, and income per unit time

- Total number of transactions in j -th piecewise constant interval is number of transactions per second, N_j , multiplied by number of seconds in interval, T_j
 - Transactions are treated as a commodity (*no* quality-of-service, *no* branding)
 - Temporal dependence of workload changes operating income per second
- Overhead associated with cold start
 - Idle may be more efficient (cost less) than turning machines off
- Architectures that reduce idle power are beneficial and can increase total income



Reduction in power dissipation increases *income*

- Reduced idle power increases income at low number of transactions per second N
 - Operating point for maximum income per second little changed if there is strong non-linear relationship between power (W) consumed by the datacenter and average number of transactions per second (N)
 - There is an increased *gap* between maximizing *income* and maximizing *energy-efficiency*
- Additional potential increase in profitable operating range with subscription plus rate pricing

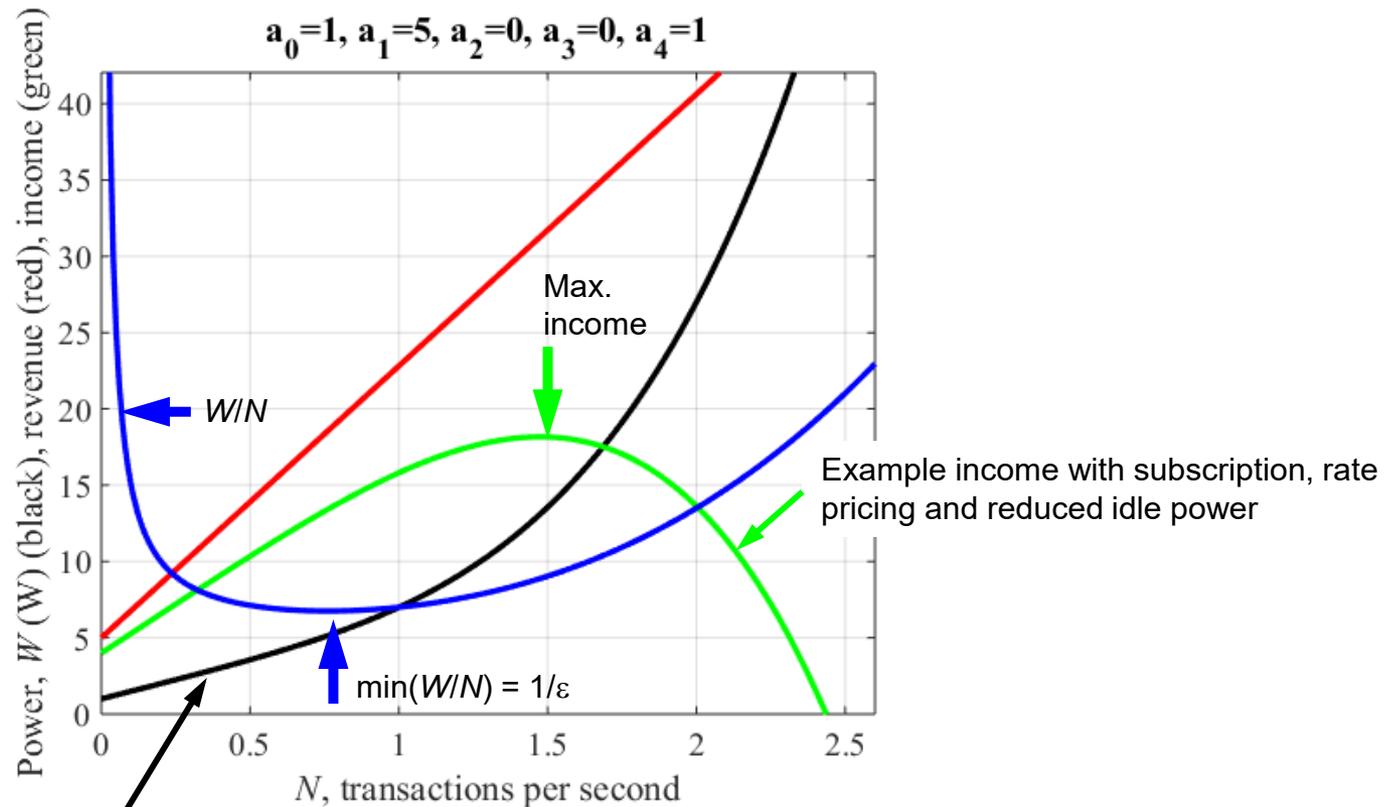


Reduced idle power has benefit of increasing income at low N

Inclusion of network dynamics suggests a commercially relevant *metric* can be assigned to the polynomial coefficients a_n weighted (by government policy?) for energy efficiency

The increasing gap between max. income and max. energy-efficiency

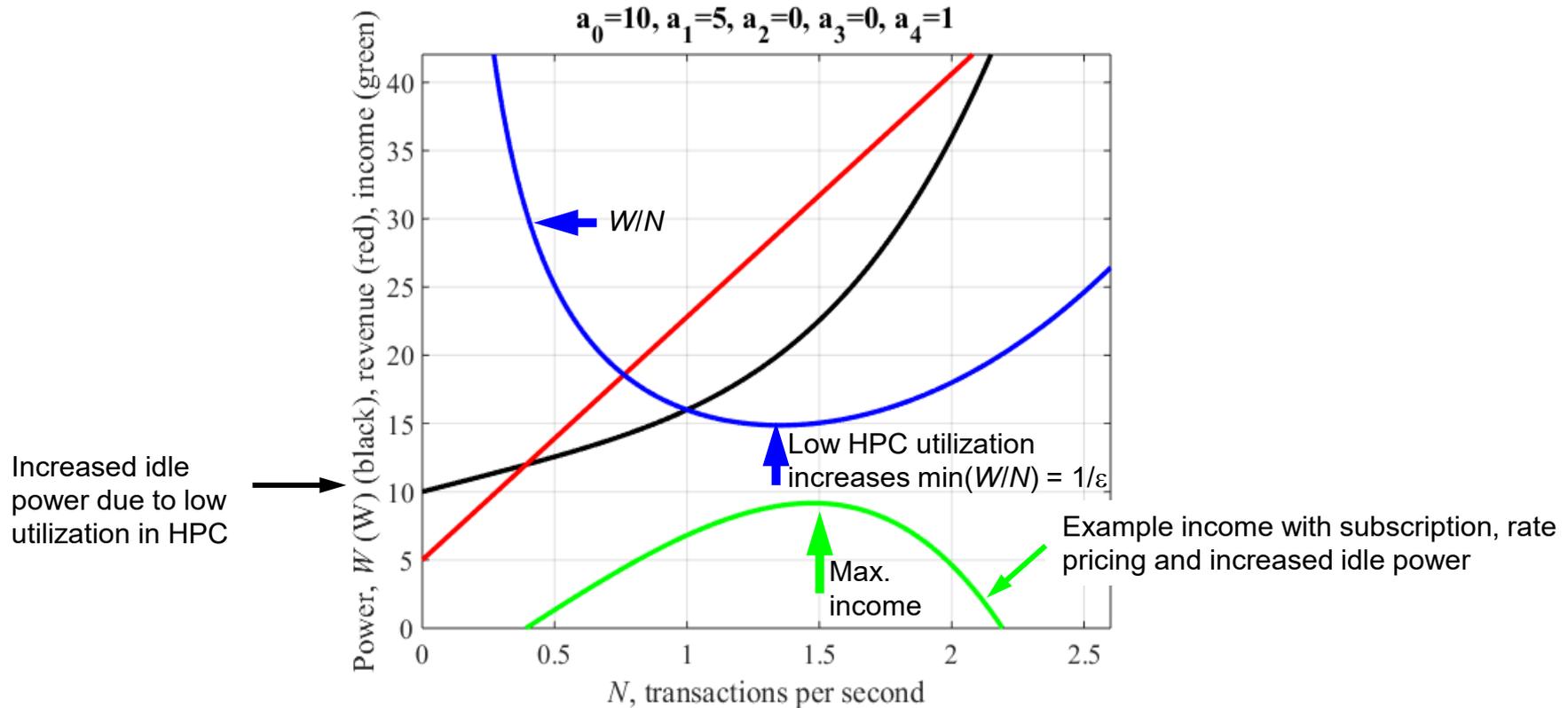
- For the case considered, maximum energy efficiency $\min(W/N)$ occurs at $N_{\min} = (a_0/3a_4)^{1/4} = 0.76$ (power 5.1) with $\varepsilon = 0.65$ transactions per J and maximum income occurs at $N = 1.5$ (power 13.6) giving power ratio of $13.6/5.1 = 2.7$



Reduced idle power has benefit of increasing income at low N

The increasing gap between max. income and max. energy-efficiency

- For the case considered, the maximum energy efficiency $\min(W/N)$ occurs at $N_{\min} = (a_0/3a_4)^{1/4} = 1.35$ (power 20) with $\varepsilon = 0.067$ transactions per J
- Low HPC utilization increases a_0 coefficient and is fundamentally inefficient



Big data may need big buffer switches that optics does not provide

<https://www.arista.com/assets/data/pdf/Whitepapers/BigDataBigBuffers-WP.pdf>

- Optical switches can provide high-bandwidth resources suggesting big data applications in high-performance storage or streaming video
- However, optical switches have no significant optical buffer resource
- Increasing buffer size in contemporary electronic switches enhances performance of some bandwidth intensive applications
 - Example: Arista 7500R modular universal spine switches with non-blocking 150Tbps of system capacity, 10/25/40/50/100GbE optimized for large virtualized and cloud networks use ultra deep buffers of up to 24GB *per line card*
 - Example: Arista 7280E, 1RU 100GbE TOR switch, has 9GB buffers
- In an optical switch architecture the benefits of buffers, including deep buffers, must migrate to the electrical domain and contribute to energy consumption

