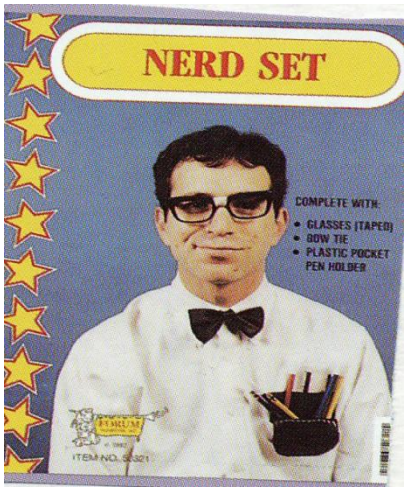
“Shock and Awe:" Network Science as an Interdisciplinary Networking Course

Walter Willinger
AT&T Labs-Research

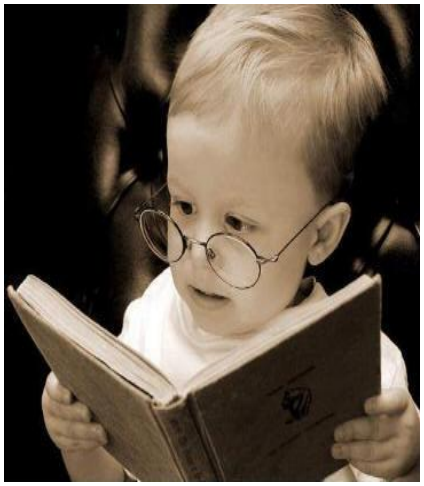
walter@research.att.com

As Networking Researchers …

- … we are at our best when …
  - we think like engineers
  - we argue like engineers
  - we are engineers

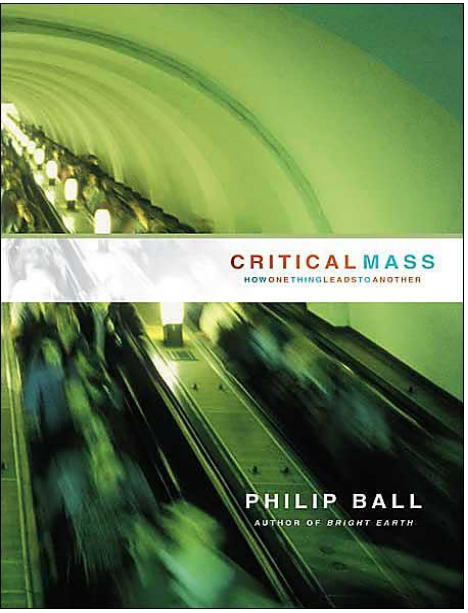


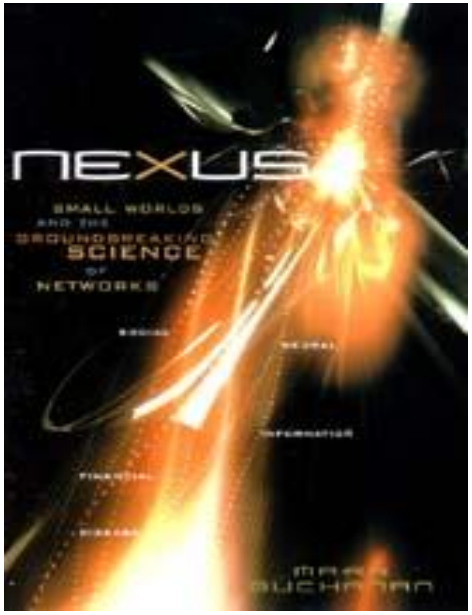
- … and we are at our worst when …
  - we forget about our engineering background
  - we are embarrassed by our engineering thinking
  - we pretend to be scientists (in the sense of "network scientists" …)

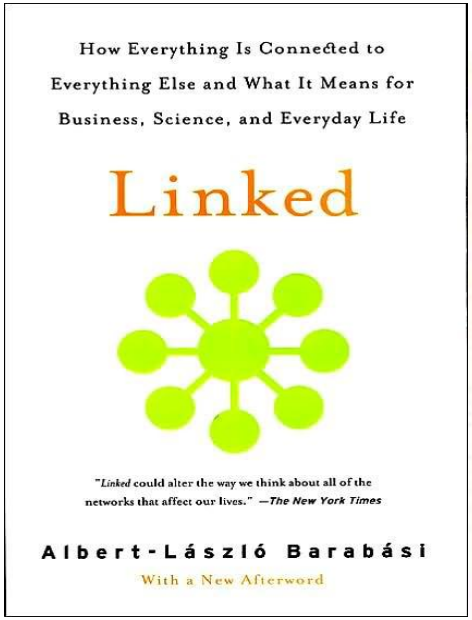


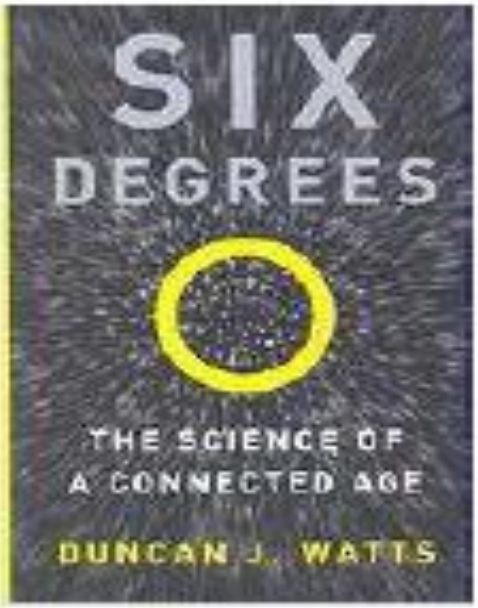
Why criticizing "Network Science"?

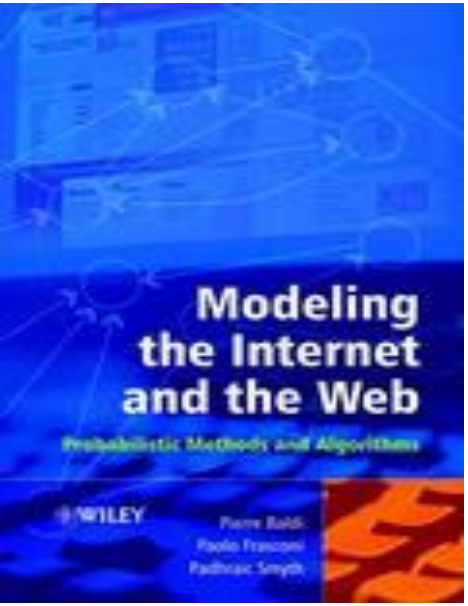
- To be honest, there is an element of "envy" …
  - Very good at marketing "Network Science"
    - Lots of books, within a short time, for diverse audiences
  - Well-oiled publication machine
    - Secret sauce for publishing in the leading scientific journals
  - Portraying itself as truly "inter/cross/multi-disciplinary"
    - Physics, Internet, Biology, Social Science, Economics, …

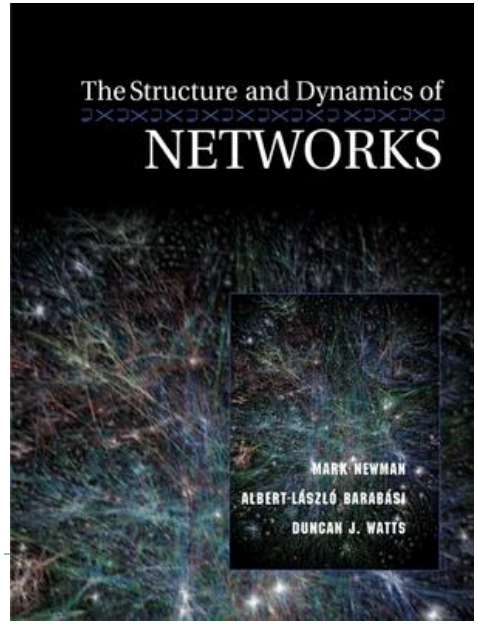

CRITICAL MASS
HOW ONE THING LEADS TO ANOTHER
PHILIP BALL
AUTHOR OF BRIGHT EARTH

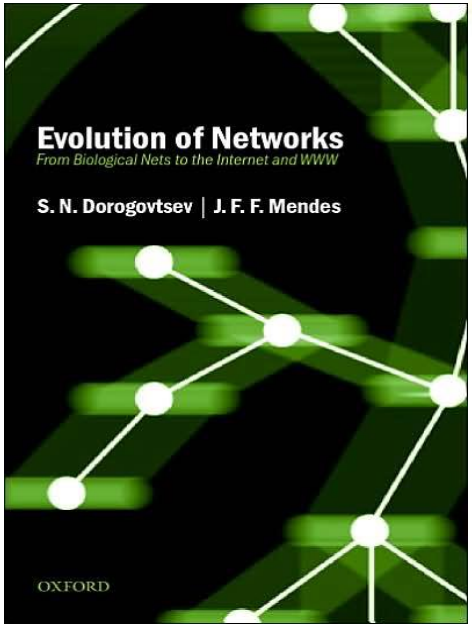

NEXUS
SMALL WORLDS AND THE GROUNDBREAKING SCIENCE OF NETWORKS

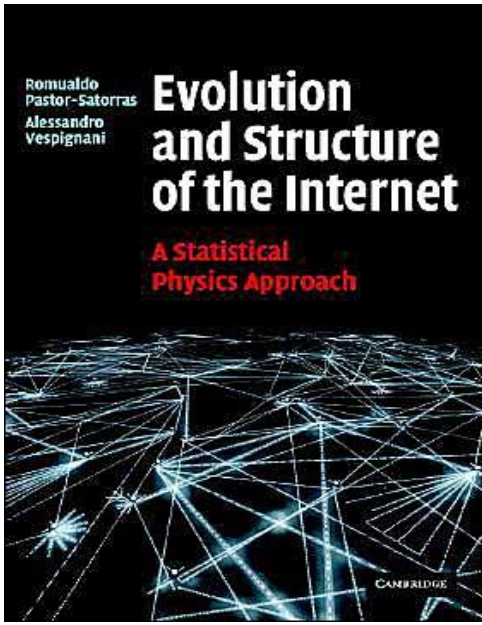

How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life
Linked
"Linked could alter the way we think about all of the networks that affect our lives." —The New York Times
Albert-László Barabási
With a New Afterword


SIX DEGREES
THE SCIENCE OF A CONNECTED AGE
DUNCAN J. WATTS


Modeling the Internet and the Web
Probabilistic Methods and Algorithms
WILEY


The Structure and Dynamics of NETWORKS
MARK NEWMAN
ALBERT-LÁSZLÓ BARABÁSI
DUNCAN J. WATTS


Evolution of Networks
From Biological Nets to the Internet and WWW
S. N. Dorogovtsev | J. F. F. Mendes
OXFORD

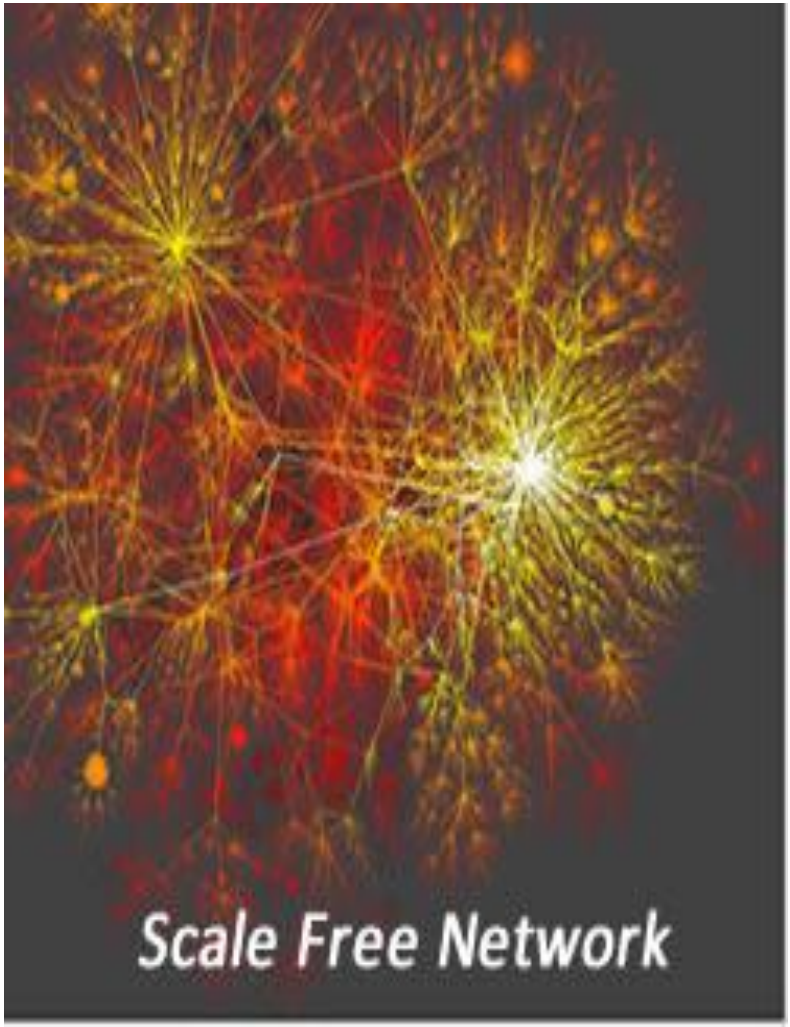

Romualdo Pastor-Satorras
Alessandro Vespignani
Evolution and Structure of the Internet
A Statistical Physics Approach
CAMBRIDGE

Why criticizing "Network Science"?

- To be honest, there is an element of "envy" …
  - Very good at marketing "Network Science"
    - Lots of books, within a short time, for diverse audience
  - Well-oiled publication machine
    - Secret sauce for publishing in the leading scientific journals
  - Portraying itself as truly "inter/cross/multi-disciplinary"
    - Physics, Internet, biology, social science, economics, …

- More seriously, there is an element of "disgust" …
  - Soft to the point of being ignorant about data quality
  - Soft to the point of being dishonest about model validation
  - Soft to the point of being disparaging about domain knowledge

Textbook Example:
"The Internet is Scale-free" (2000-today)



The "Secret Sauce" for Writing such Stuff

- Step 1: Get access to high-volume Internet datasets
  - Traceroute- or BGP-based measurements from your favorite source(s)
- Step 2: Take the (high) quality of the datasets as given
  - traceroute- or BGP-based datasets as basis for mapping Internet connectivity at the router- or AS-level
- Step 3: Focus on a single "intriguing/universal" feature of the data
  - Power-law-type node degree distributions
- Step 4: Propose a new class of "network" models
  - Scale-free networks of the preferential attachment type
- Step 5: Argue for the validity of the proposed model class
  - Can reproduce the original "intriguing/universal" feature of the data
- Step 6: Use analysis and/or simulation to obtain model-based predictions about the Internet and its properties
  - "Achilles' heel" of the Internet

"Network Science" Principles at work …

- Never ever question the available measurements!
  - More data is always better
  - The fact that data can often be garbage is a foreign concept
- Avoid specificity at all cost (it pays to remain vague …)!
  - Networks are nothing but abstract generic graphs
  - The rest is just details that don't matter
- Under any circumstances, avoid discussing model validation!
  - Strictly an exercise in data-fitting; not more, not less
- Important Corollary
  - Know enough to appear knowledgeable
  - Argue that everything else are "details" that don't matter
  - Decry "detailed-oriented thinking" as "not seeing the forest for the trees"

Networking Research on a Slippery Slope …

- There is a current trend to imitate "Network Science"
  - Give me data, will model (what do you mean by "data quality"?)
  - Give me models, will simulate (what do mean by "proof"?)
  - Give me models, will use them (what do you mean by "model validation"?)

- There is a current trend to make "doing science" look easy
  - Collecting data is easy …
    - when in reality, checking their quality is hard (when done carefully)
  - Modeling is easy …
    - when in reality, model validation is hard (when done with care)
  - Using models is easy …
    - when in reality, using them to gain insight is hard (if insight means advancing science)

Recent Example (October 2010) Scale-Free Networks & Data Center Design

- **Scafida: A Scale-Free Network Inspired Data Center Architecture,** by L. Gyarmati, T. A. Trinh
- Appeared in: ACM SIGCOMM Computer Communication Review, Volume 40, Number 5, October 2010

- From the public review: *"All the reviewers agree that the paper asks a compelling and timely question. What if we were to design datacenter networks to be scale-free? The answers provided are not perfect, but certainly shed light on the problem. Surprisingly such topologies provide properties quite similar to carefully designed topologies."*

Even more recent example (Sigcomm'11)

- **The Evolution of Layered Protocol Stacks Leads to an Hourglass-Shaped Architecture,** by S. Akhshabi and C. Dovrolis
- Appeared in**:** Proceedings of the ACM SIGCOMM 2011

- From the abstract: *"We propose EvoArch, an abstract model for studying protocol stacks and their evolution. EvoArch is based on a few principles about layered network architectures and their evolution in a competitive environment … EvoArch produces an hourglass structure that is similar to the Internet architecture from general initial conditions in a robust manner."*

“Engineering” Principles to the Rescue …

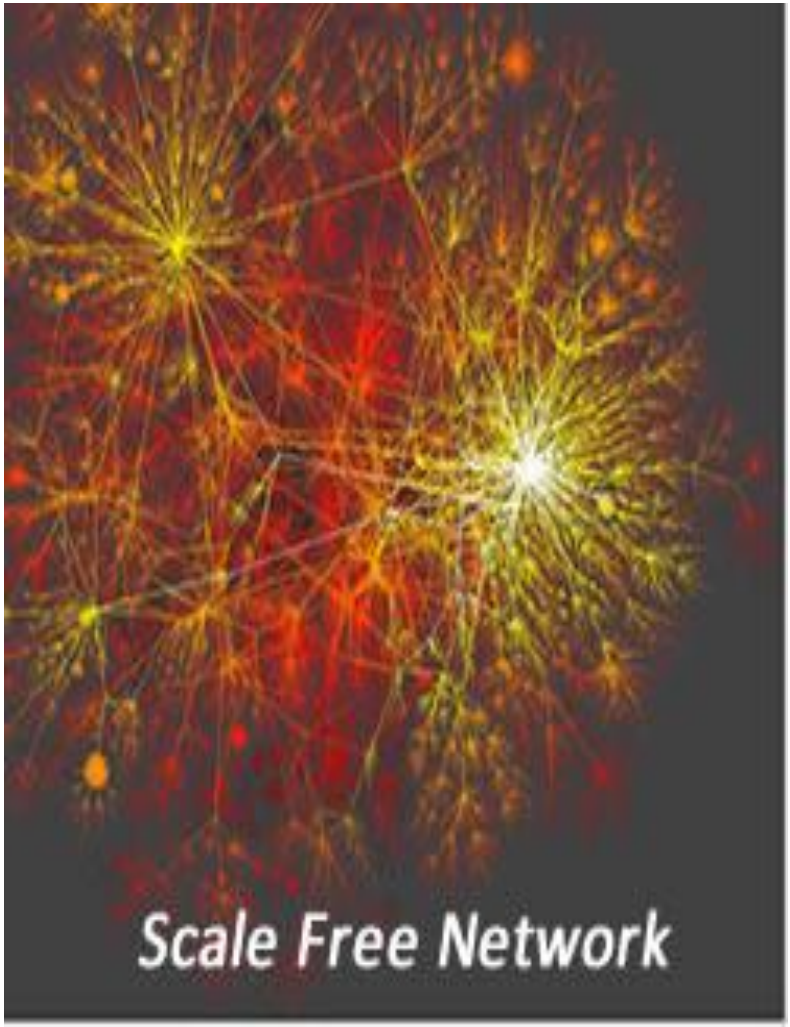
- Details do matter!
  - Insist on proper domain knowledge
- Know your data!
  - Premise: (Internet) measurement is hard
  - General Rule: What we want to measure is typically not what we can (or think we) measure
- Take model validation serious!
  - Modeling/model validation has to be more than an exercise in data-fitting
  - Modeling/model validation has to become an exercise in reverse-engineering
- Important Corollary
  - Engineering is all about “details matter”, “know your data”, “not data-fitting” but “reverse-engineering”, ….
  - **“[Engineering has] nothing to offer but blood, toil, tears and sweat.”**

Asking for the Impossible? An Engineering-inspired "Network Science"

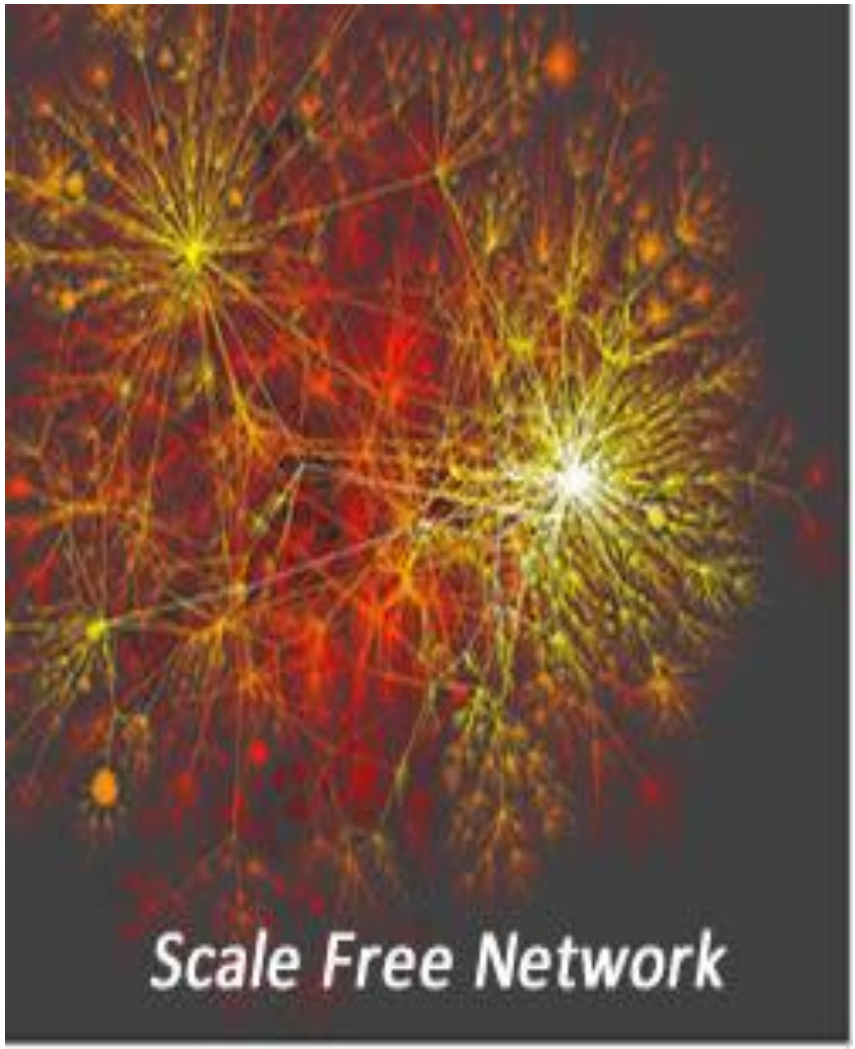
- Requires a different mindset
  - Past
    - "All models are wrong, but some are useful" (Box)
    - Not very helpful, often used as "last defense" by Network Science
  - Future
    - "When exactitude is elusive, it is better to be approximately right than certifiably wrong" (Mandelbrot)
    - Poses a challenge that has the potential of advancing science

- Requires some paradigm shifts
  - Making "data hygiene " part of any measurement-driven scientific endeavor in the peta/exa/zetta-byte age
  - Completely abandoning the existing data-fitting mentality
  - Combining the current physics-centric approach with an engineering-based perspective for the benefits of science as a whole

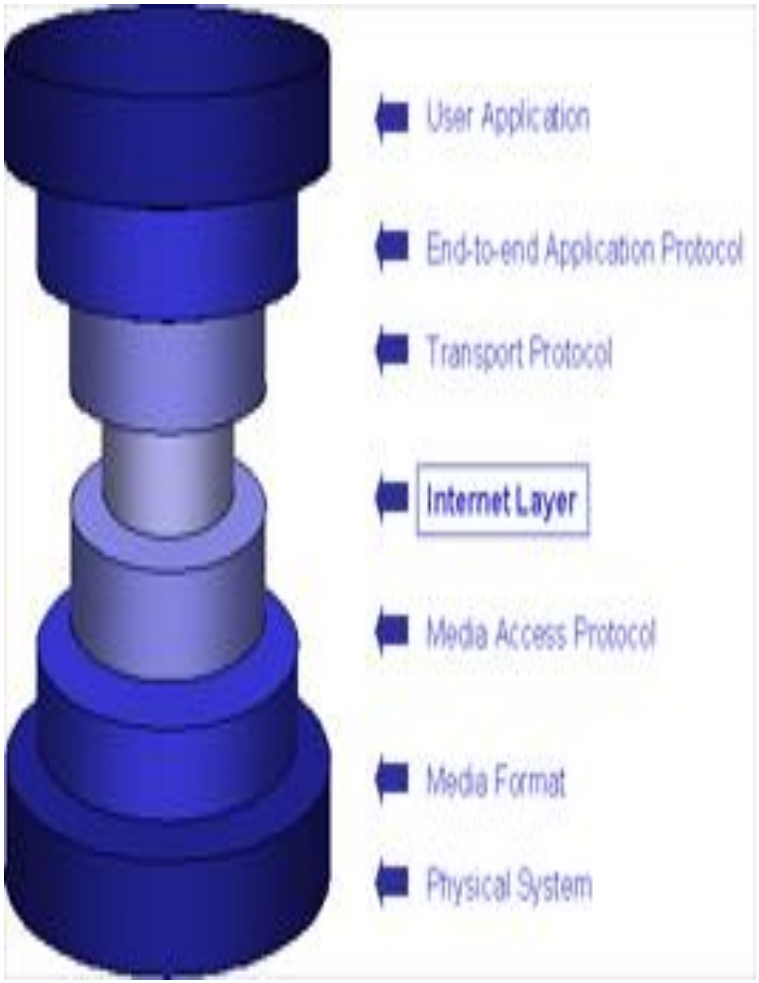
The Internet is scale-free …



Scale-free Networks for Data Center Design

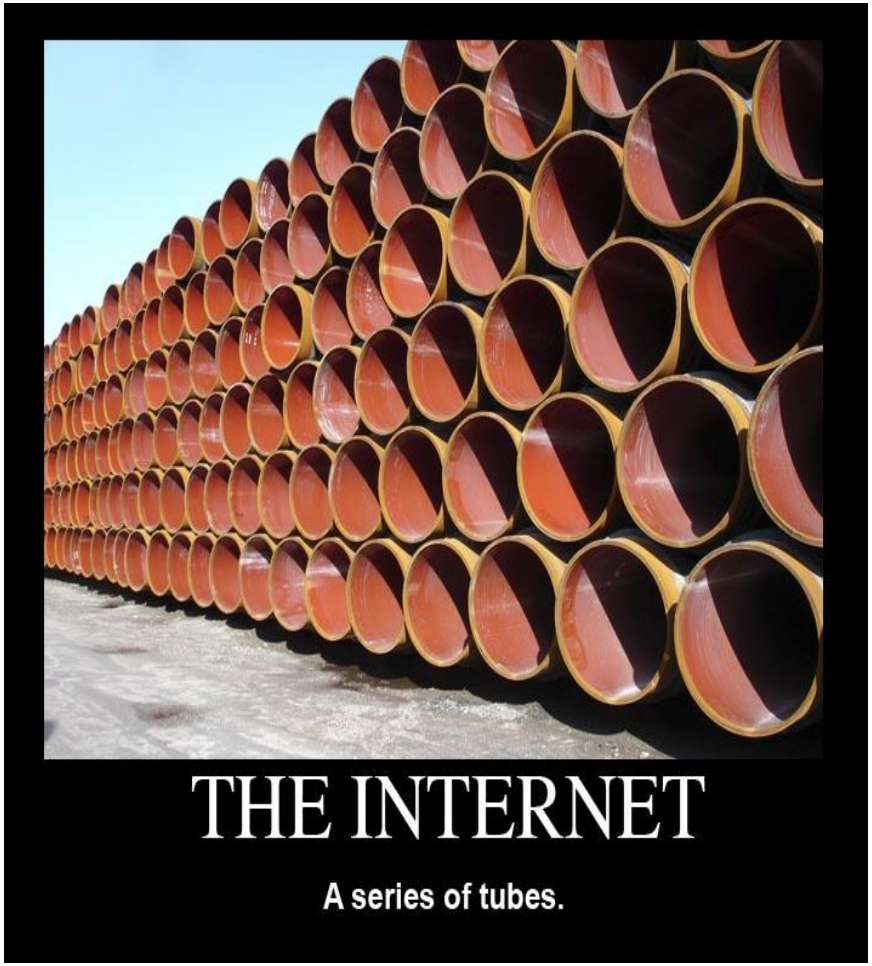


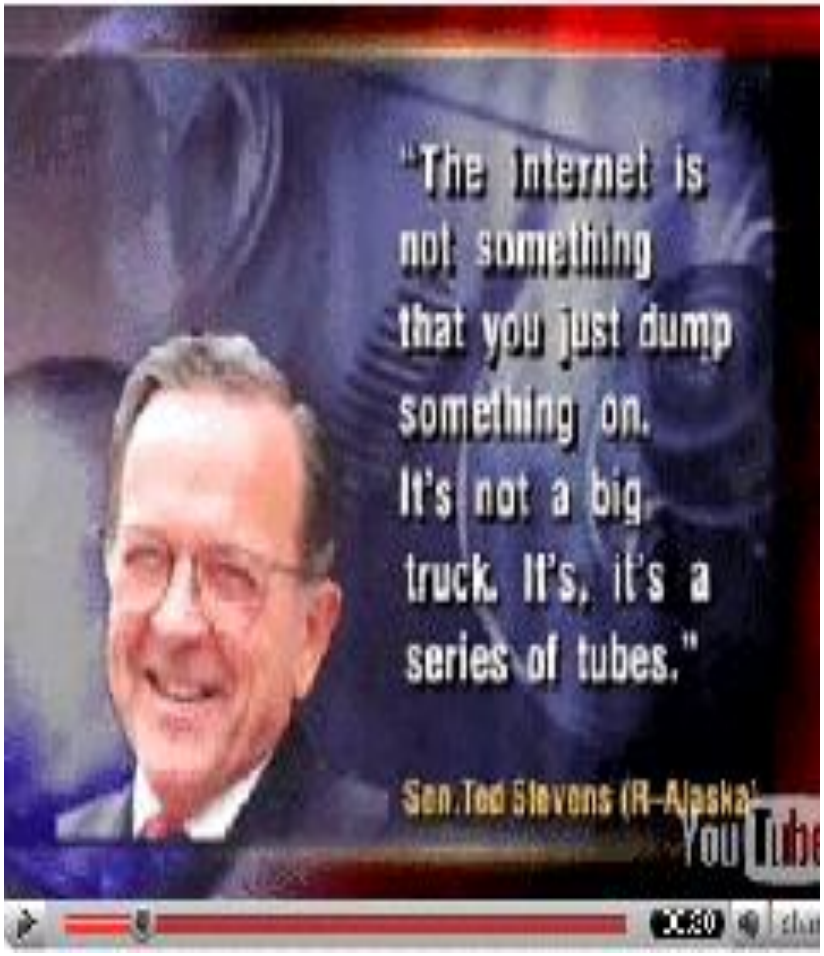
… the Evolution of Layered Protocol Stacks Leads to an Hourglass-Shaped Architecture …”



Geoff Huston, 2003, http://isoc.org/wp/ispcolumn/?p=49

“The Internet is a series of tubes …”





Which is the most ridiculous statement?

- General scientific literature (Nature 2000)
  - "The Internet is Scale-free …"
- Recent networking literature (CCR 2010)
  - "A Scale-Free Network Inspired Data Center Architecture"
- SIGCOMM 2011 paper:
  - "The Evolution of Layered Protocol Stacks Leads to an Hourglass-Shaped Architecture …"
- General public (Senator T. Stevens, 2006)
  - "The Internet is a Series of Tubes …"

After undergoing 15 Minutes of “Shock and Awe” Therapy …

Which Statement is NOT entirely wrong?

- General scientific literature (Nature 2000)
  - "The Internet is Scale-free …"
- Recent networking literature (CCR 2010)
  - "A Scale-Free Network Inspired Data Center Architecture"
- SIGCOMM 2011 paper:
  - "The Evolution of Layered Protocol Stacks Leads to an Hourglass-Shaped Architecture …"
- General public (Senator T. Stevens, 2006)
  - "The Internet is a Series of Tubes …"