A quick guide to the Internet

David Clark
MIT CSAIL
Fall 2011

Why should you care?

- “People” say: “Why doesn’t the Internet…”
  - Protect me from spam.
  - Control porn
  - Keep terrorists from plotting
  - Etc.

- We need to translate from a cry of pain to realistic expectations.
- Must understand (in general terms) the technology to make realistic policy.

Defining the Internet

- It is not the applications:
  - Email
  - Web
  - VoIP
- It is not the technology
  - Ethernet
  - WiFi
  - Fiber optics
- So what is it?

A nice picture

A range of applications operating over a range of technologies, by means of a single interface -- the Internet Protocol IP.

Image removed due to copyright restrictions. To view the image, please see http://www.nap.edu/openbook/0309050448/gifmid/53.gif

IP: the Internet Protocol

Taken from:

Realizing the Information Future: The Internet and Beyond.

What is a network?

Or…what is the problem we are solving?

- A *shared* medium of communications.
- Why?
  - To share expensive resources
    - Cannot afford a wire between "everywhere."
  - To facilitate general communication--information sharing.

How to share?

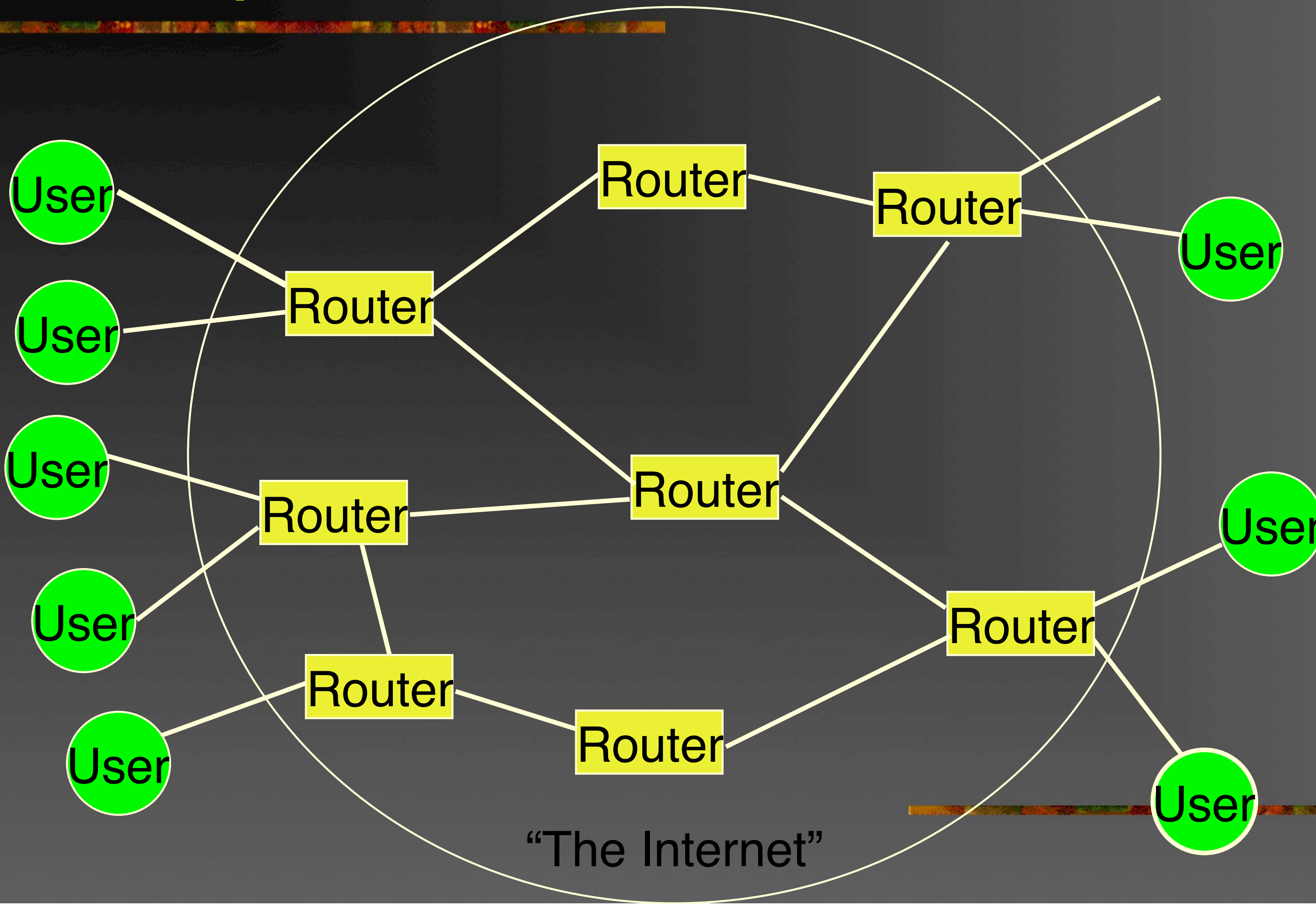
- Computer traffic is bursty.
- Older sharing method (circuit switching) was inefficient.
- About 45 years ago, the need for a new mode of sharing was felt.
- PACKETS!

What is a packet?

- Packet:
  - Some data with an address on the front.
    - Specified maximum size
  - Sent serially across a link.
  - Use a computer (a ”router” or “packet switch”) to manage the link.
  - Statistical sharing.

- A neat idea that has stood the test of time.

A simple view of the Internet



An (over) simple packet picture

Header | **Data**

A packet

Stuff…Destination/Source/Length…stuff

A packet header

Addresses: written in the form 18.26.0.166

What a router does

- When it gets a packet:
  - Check that it is not malformed.
  - Check that it is not going in circles.
  - Look at its destination address.
  - Pick the best link over which to forward it.
- In the background:
  - Computes the best routes to all destinations.

What was at the edge?

- The slide said “user”.
- It is a “host”, or a “PC”, or a “server”, or a “computer” or an “end node”.
  - The place where application code runs.
  - There might be a person there.
- Get back to this later…

Implications

- Inside the network there are only packets.
- There is no understanding of higher-level intentions.
  - The routers have a limited view of what it means to “operate correctly”.
- There is nothing like a “call”, or “placing a call”, in the router’s design.

The service model

- The other half of the Internet specification:
  - What is the commitment when I send a packet?
- Answer: very little.
- The Internet tries it best, but makes no promises.
  - It can lose, reorder, delay, or duplicate packets.
  - Usually they arrive in good order.
  - If they don't--you have no complaint.
- Called the “best effort” service.

Is this such a good idea?

- Weak expectation means Internet can run over “anything”.
- Makes the application’s job harder, but not impossible.

- So, yes, it is a good idea.
  - But now under attack.

Congestion

- More than you want to know in one slide…
- What happens if too many packets get sent?
  - In the short run, queues form in routers.
  - In the longer run, senders (are supposed to) slow down.
- Why does this work?
  - Application are expected to tolerate it.
- But if senders do not slow down?
  - Out of aggression, or because they cannot.
- A raging debate among designers.
  - And Comcast and BitTorrent…

Responses to congestion

- Four options.
  - Demand and expect them to slow down.
    - Benign socialist
  - Police them and punish them if they don’t slow down.
    - Police state
  - Let them pay to keep going fast.
    - Capitalist
  - Over-provision so net is “never” congested.
    - Pragmatic

What was at the edge?

- The slide said “user”.
- It is a “host”, or a “PC”, or a “server”, or a “computer” or an “end node”.
  - The place where application code runs.
  - There might be a person there.

What a "host" does

- Runs the application code
  - Web, email, voip, ssh, wow, etc.
- Runs software that helps cope with packets and the best effort service model.
  - Example: Transmission Control Protocol, or TCP.

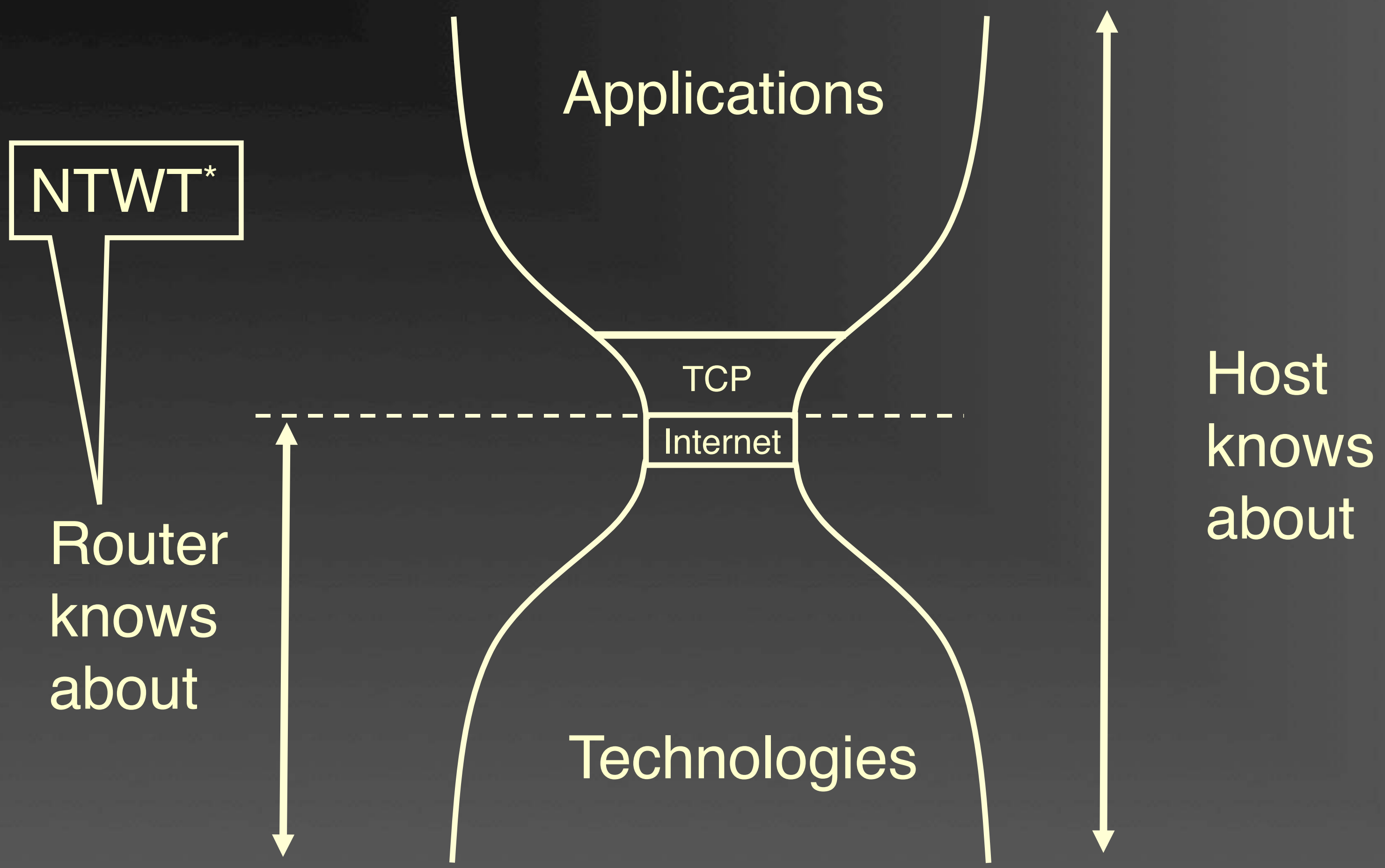
What is a protocol?

- Protocol: A specification of what can be sent when and in what format.
- A very general term, used to describe many aspects of networking.
  - The voltage to represent a 1 or 0 on a link.
  - The bit sequence to represent characters (e.g. ASCII).
  - The format of the address on the front of the packet.
  - How one reports a lost packet.
- From the Greek: "Glued on the front."

What does TCP do?

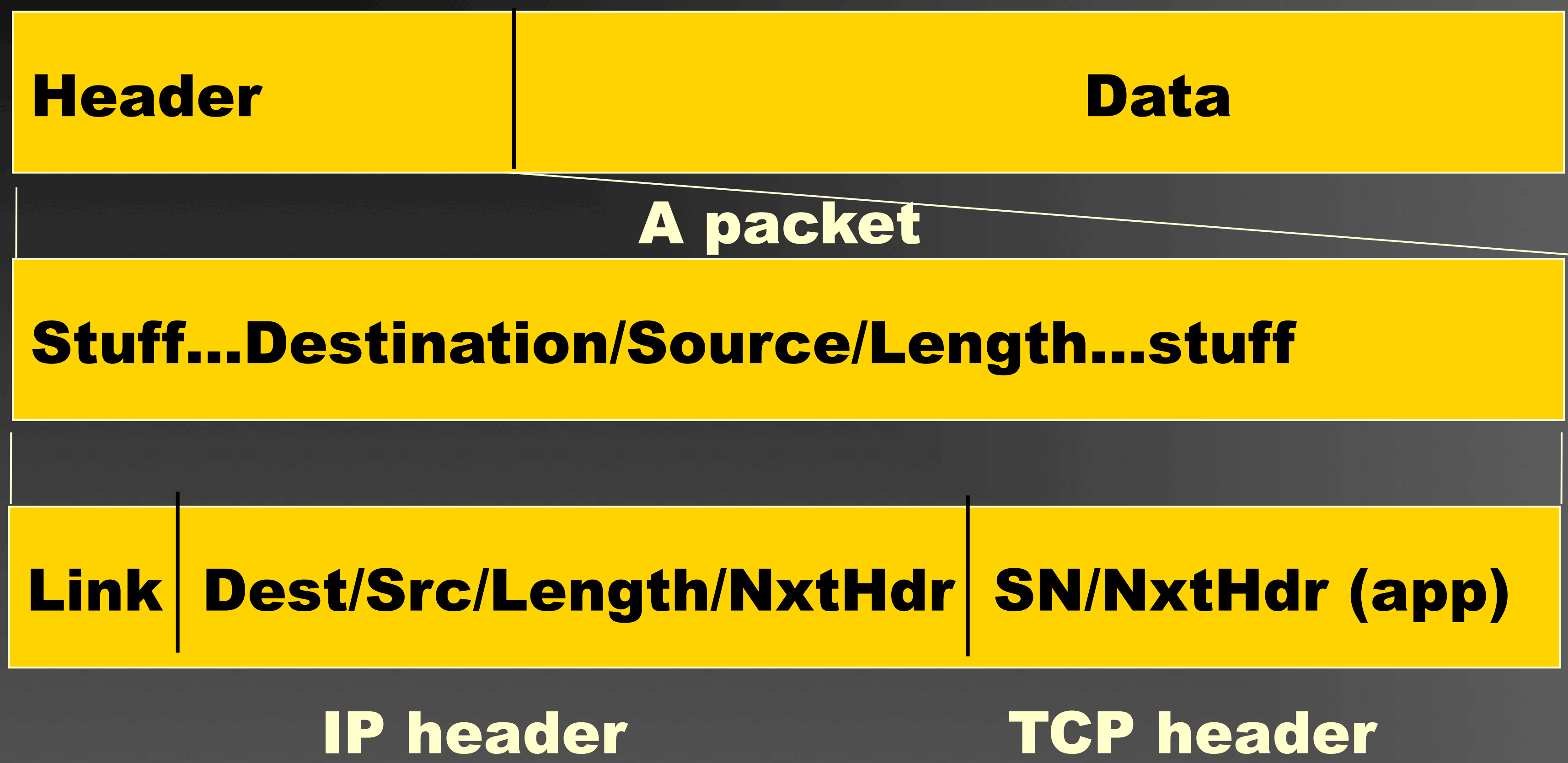
- Breaks a chunk of data (what the application wants to send) into packets at the sender.
- Numbers the packets.
- Keeps sending them until it gets an acknowledgement.
- Puts them in order at the receiver.
- Passes the date to the right application.
- Provides a very simple failure model.

Host vs. router



*Not the whole truth

A (less) simple packet picture



The *end-to-end* arguments:

The lower layers of the network are not the right place to implement application-specific functions. The lower layers of the network should implement basic and general functions, and the applications should be built “above” these functions, at the edges.

- E.g. move functions “up and out”.

- This causes function migration to the end-node.
- The network should be “as transparent as technology permits”.

Benefits of end-to-end

- User empowerment.
  - Run what you please.
- Flexibility in the face of unknown applications.
  - A network to hook computers together.
- Lower cost in core of network.
  - Eliminate special “features”.
  - Rely on edge-node equipment.
- More robust applications.
  - No unexpected failures of third-party nodes.

- An example of “getting it wrong”: make the network reliable.

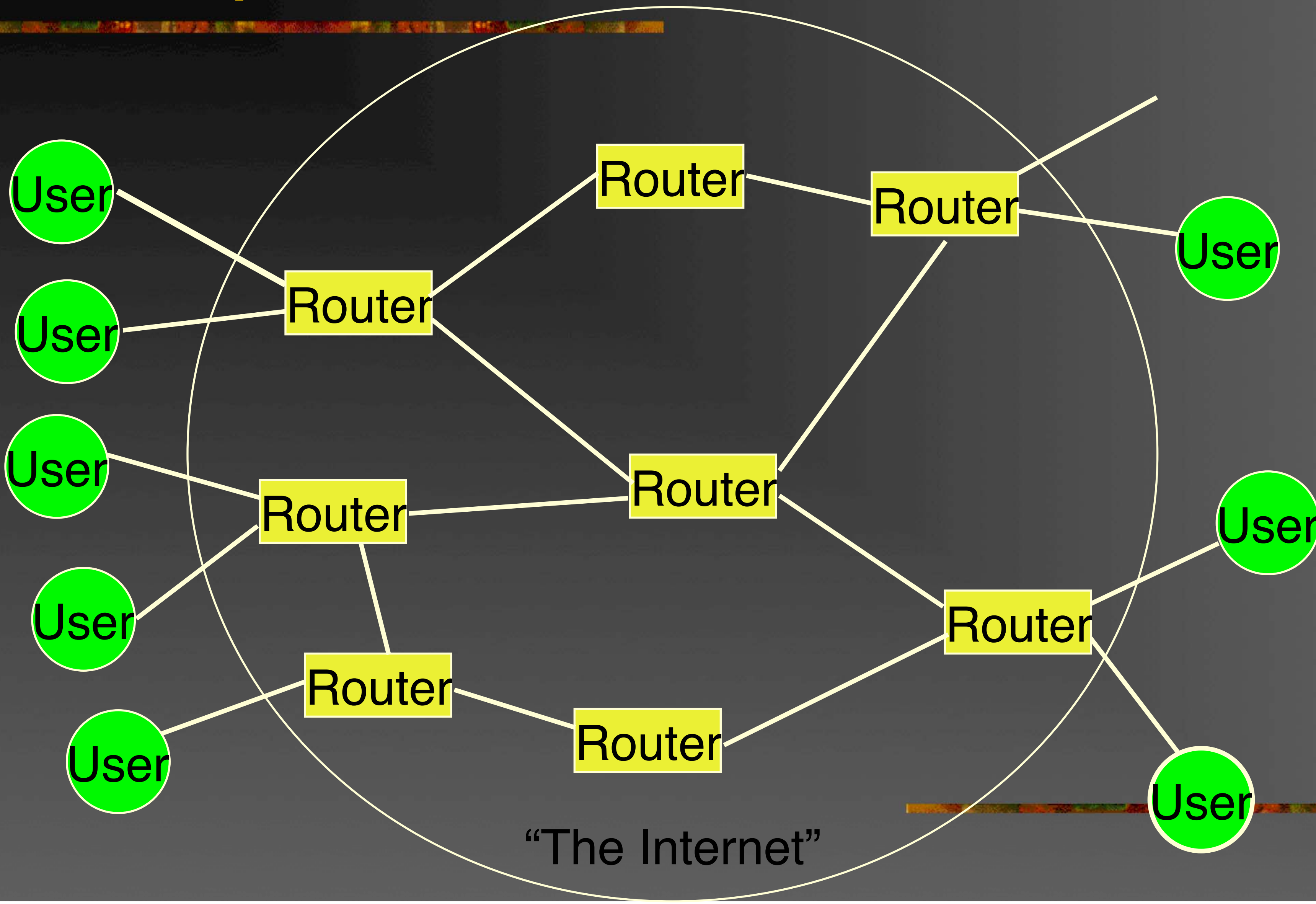
Summary

- What “the Internet” does is very simple:
  - If forwards packets.
  - It is oblivious to the purpose of the packets.
  - Packets allow effective/efficient sharing.
- Lots of applications run on the Internet
  - And there will be more tomorrow.
  - Each has its own design.
  - There is a tension/tradeoff as to where functions are placed.
- The Internet can exploit lots of technologies.

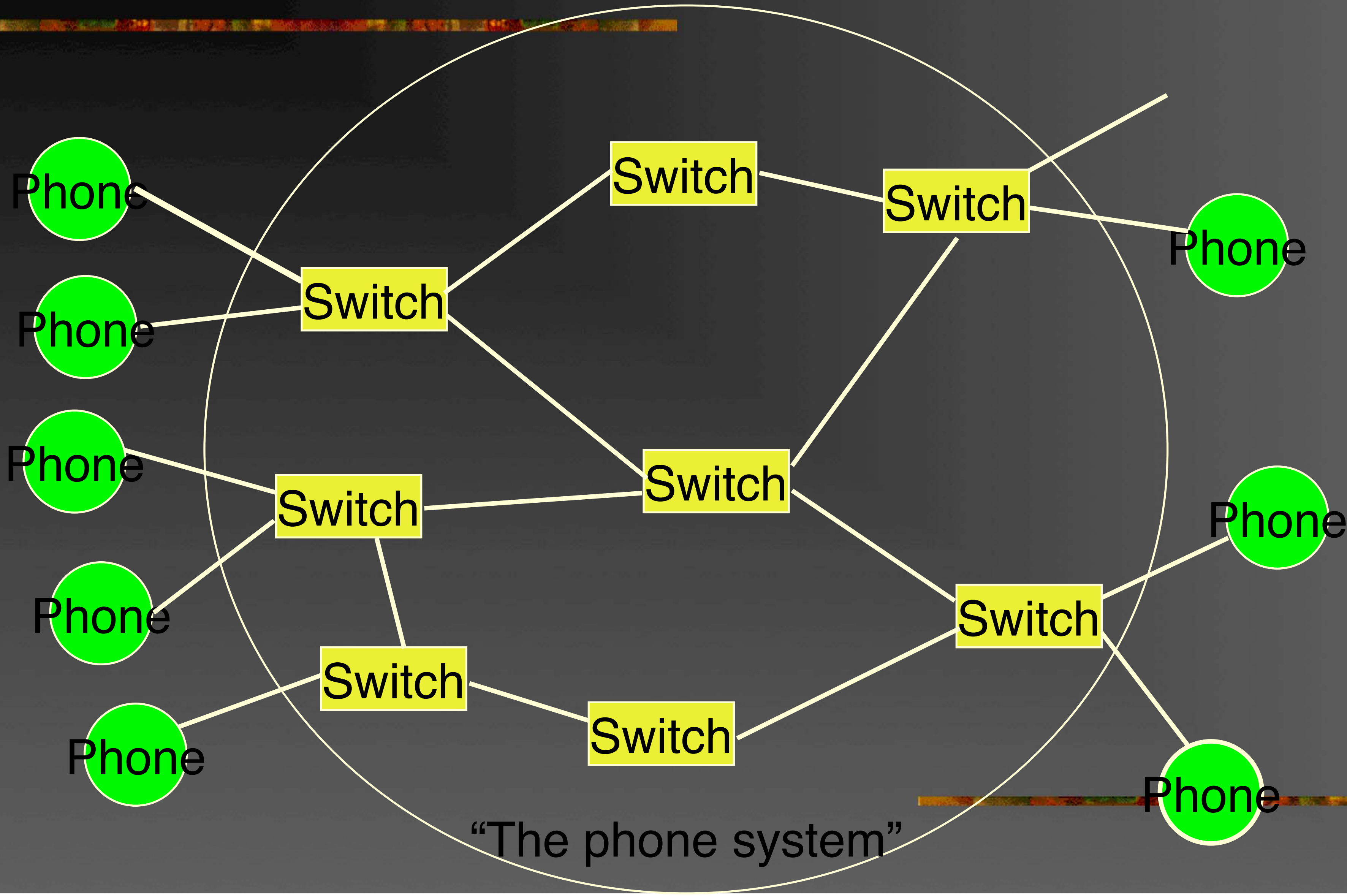
How about the phone system?

- How does it different from the Internet?
  - And *why?*
- What are the implications for policy?

A simple view of the Internet



A simple view of the phone system



The differences?

- Switches are powerful, because phones are simple.
  - The knowledge of what the phone system is for is embedded in the switches.
  - It “knows” that its purpose is to carry voice.
- Routers are simple, because end-points are powerful.
  - The knowledge of what the Internet is for is embedded in the end-points.

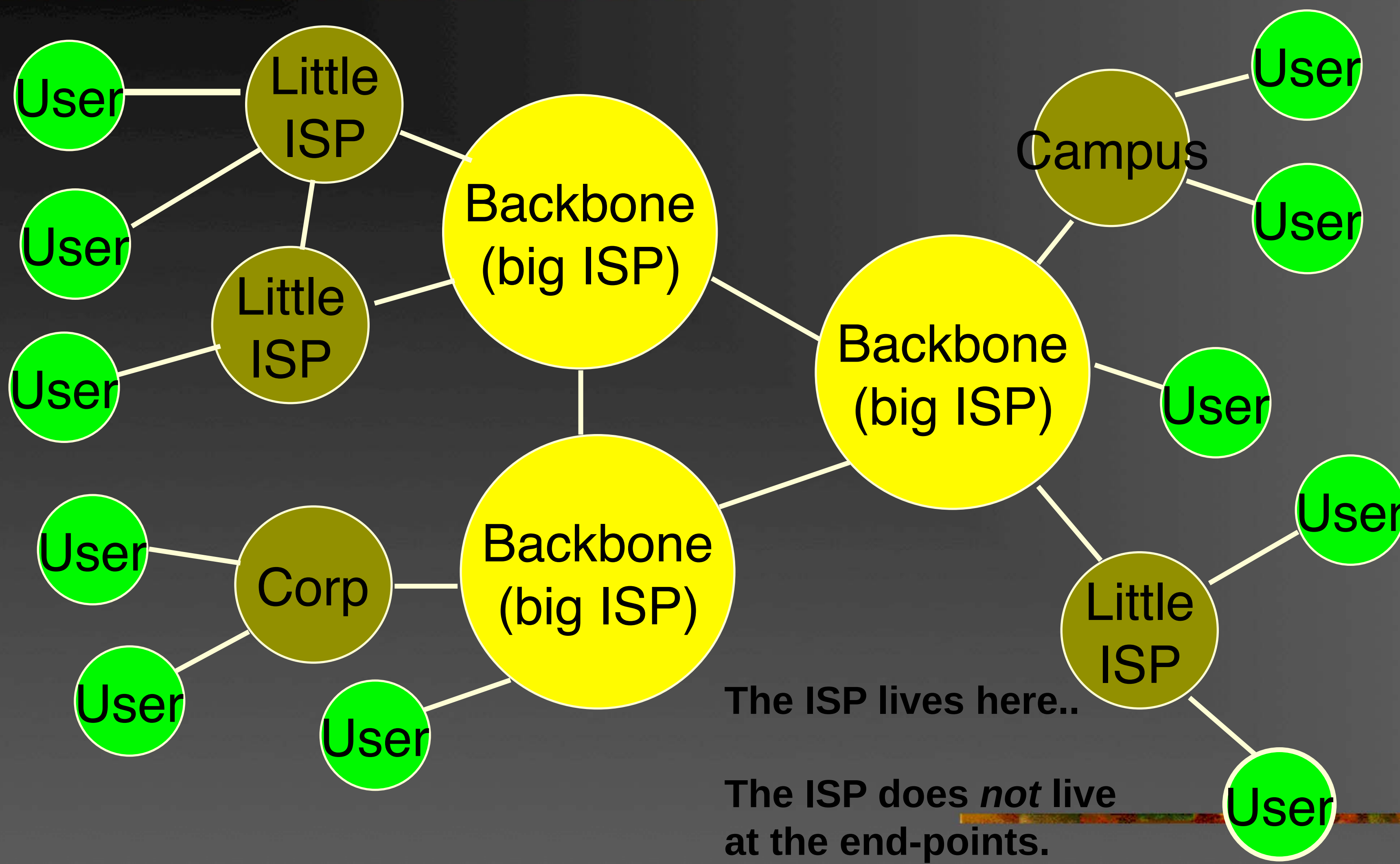
And…?

- The phone system has no open API.
  - No easy way to attach new applications.
  - Compare the generality of a voice circuit and packet carriage.
  - Very different view of layering. There is no “voice” layer.

Multiple views of system

- Topology view:
  - Routers as expression of physical distribution.
- Layered view:
  - What is the role of each "box".
  - What does this imply about limits to action?
- Administrative view.
  - Who owns/operates each part?
  - Who controls what talks to what?

A more realistic picture



The ISP lives here..

The ISP does *not* live at the end-points.

What is missing from all this?

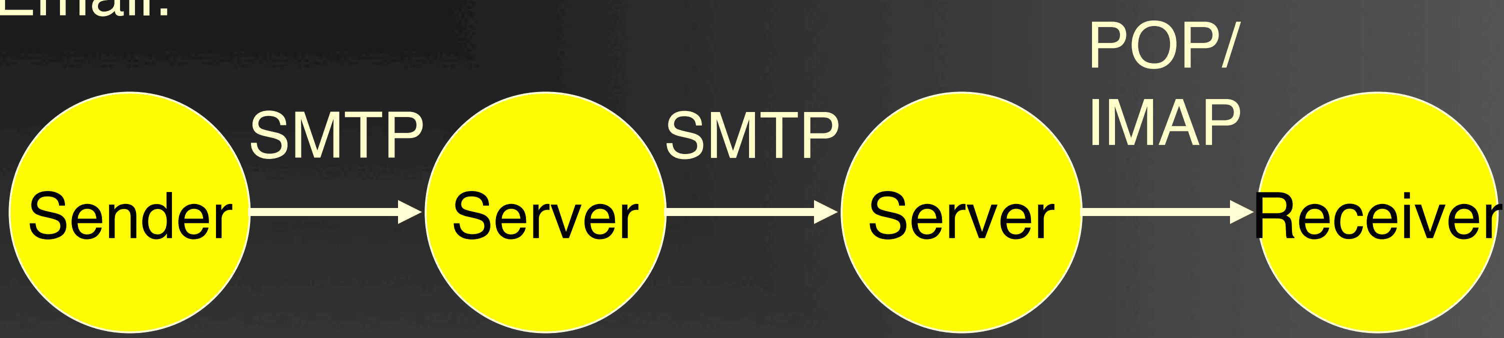
- What is it *for*?
  - How is the Internet used?
- Briefly, lets talk about applications.

Application design

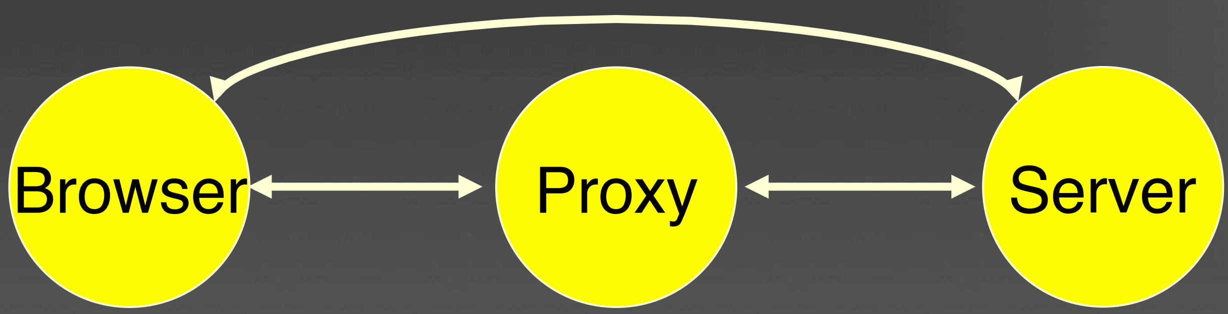
- Applications run “on” the Internet. They are not the Internet.
- Many approaches to construction
  - Patterns of communication.
  - Use of end node software and server software.
- Modern apps do not follow a simple end to end model.
  - (End to end at application level)
    - Remember the end to end argument?
- They are full of servers and services run by third parties.

Some examples:

Email:

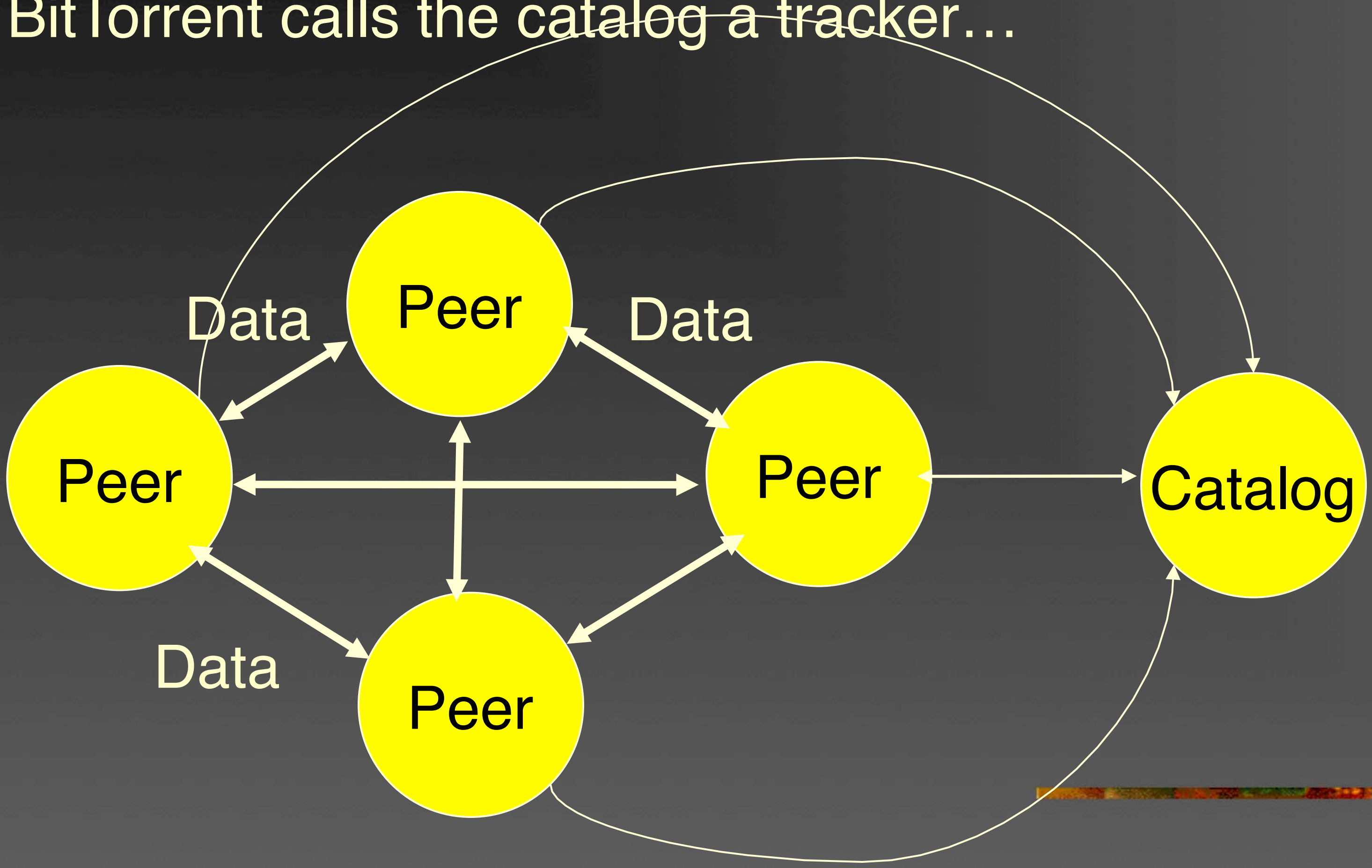


Web:



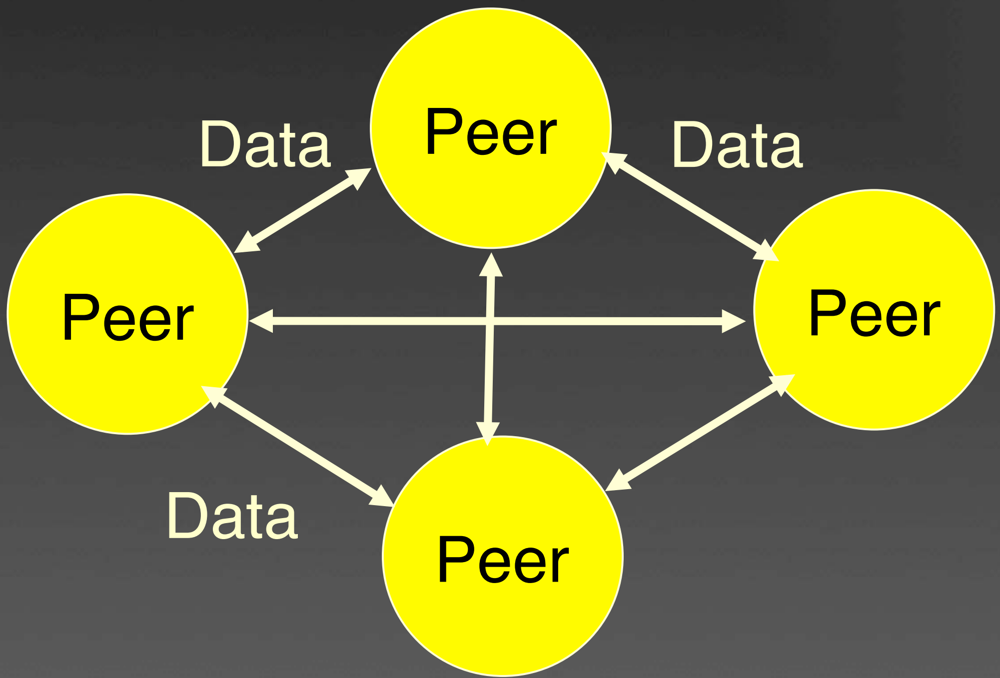
More examples:

Napster (early peer to peer)
BitTorrent calls the catalog a tracker…



More examples:

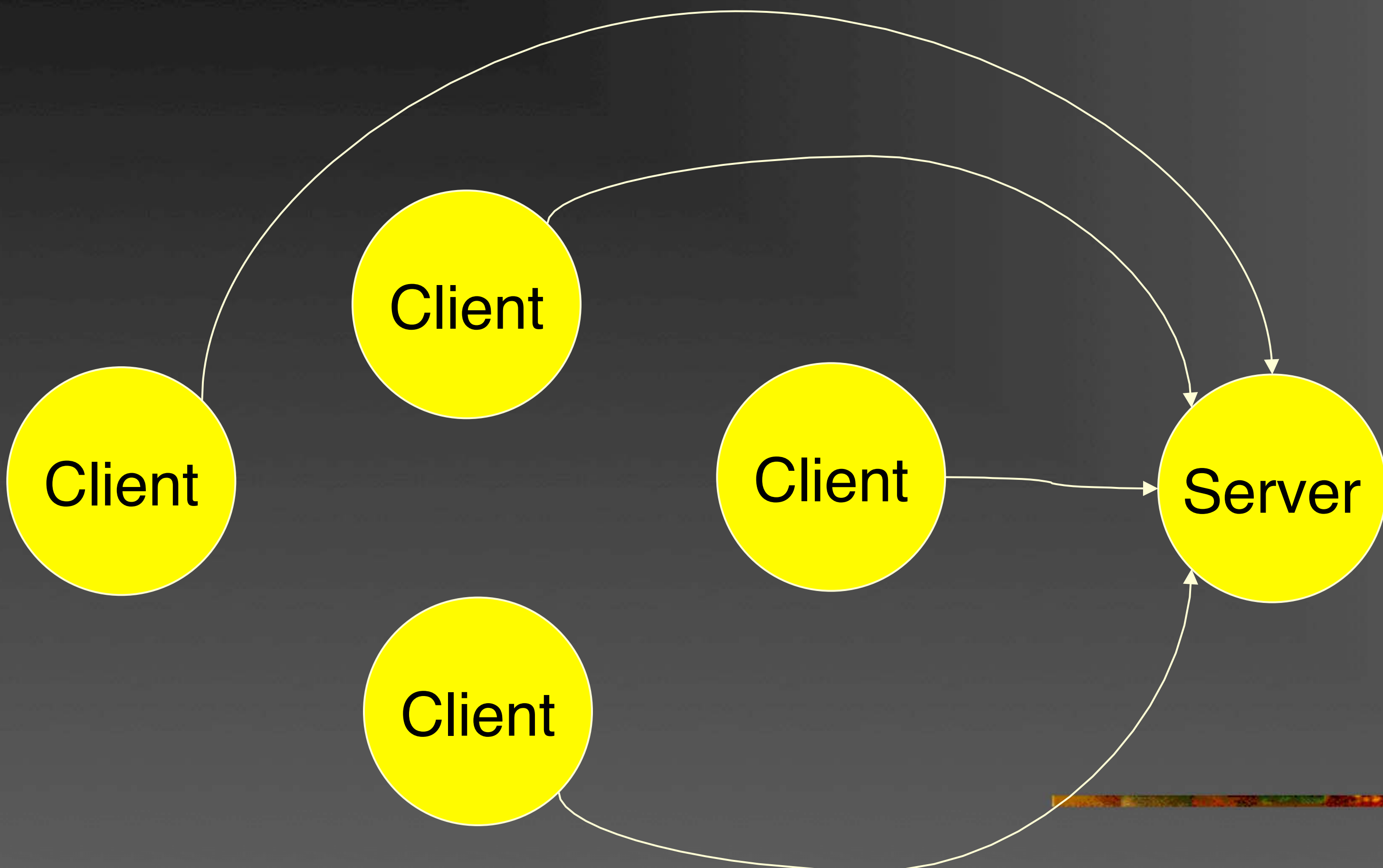
Later peer to peer
BitTorrent has a version that works without a tracker, using a distributed data base to try to find the content.



Just go feel around for the data. It might be there somewhere.

More examples:

Games (some), and IM (some)



The changing structure…

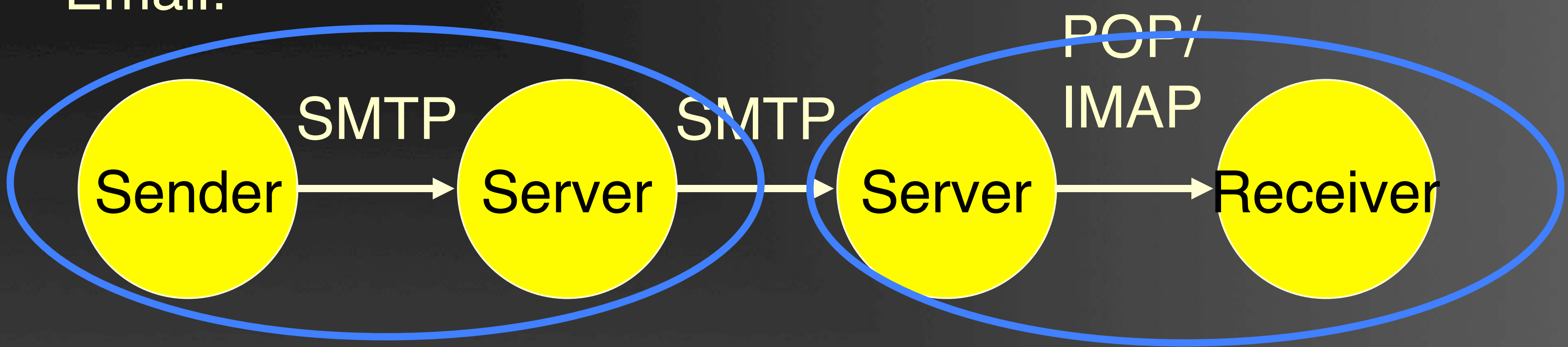
- In the old days, there were two sorts of devices:
  - Routers
  - End-node computers.
- Now:
  - Server farms
  - Cloud computing (latest buzzword…)
- So where should computing be placed?
  - And why?

What problems are we solving?

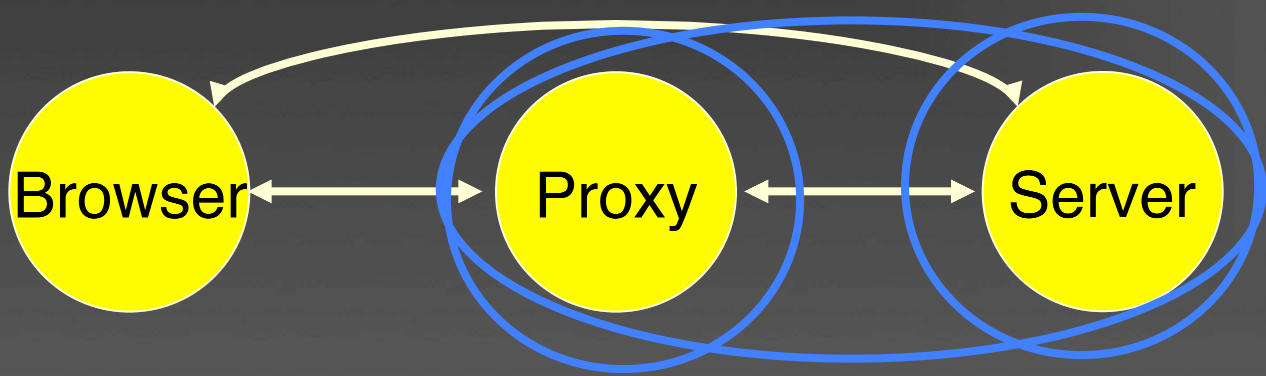
- Ease of use
- Ease of deployment
- Performance
- Economic (industry) structure
- Robustness
- Security
- Who is in control?
  - Function placement based on trust.

Trust relationships

Email:



Web:



Placement of computation

- "The Internet" is not changed by where computation is placed.
  - Except that we need some really high-capacity circuits…
- But the user and the application is strongly influenced.

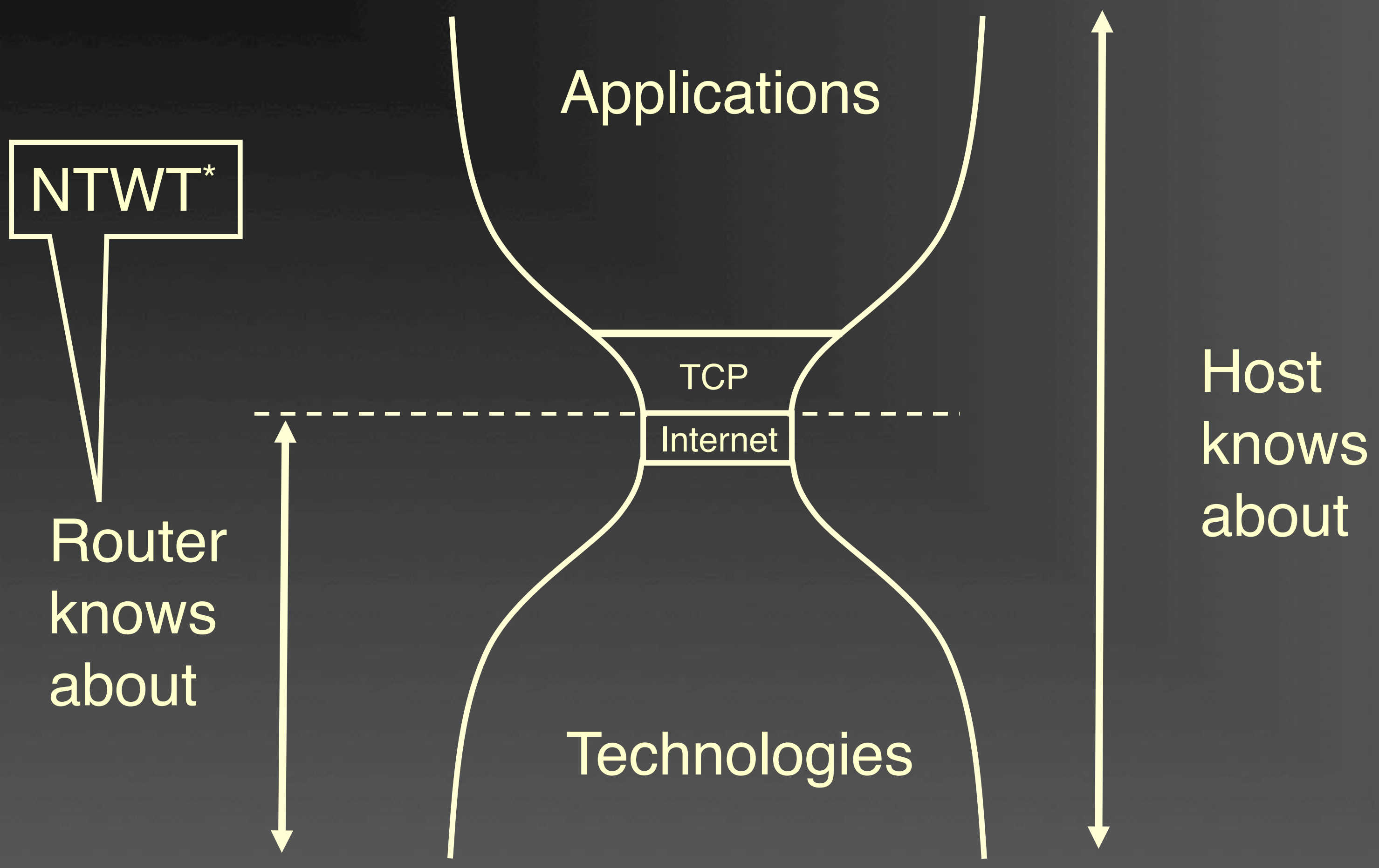
I mentioned Comcast…

- How did they disable BitTorrent?
- They “peeked” at the data part of the packet
  - Called Deep Packet Inspection, or DPI.
- When they saw a BitTorrent connection, they inserted an extra message into the flow of packets (some might say “forged”) that said “abort the connection”.

What could the app do?

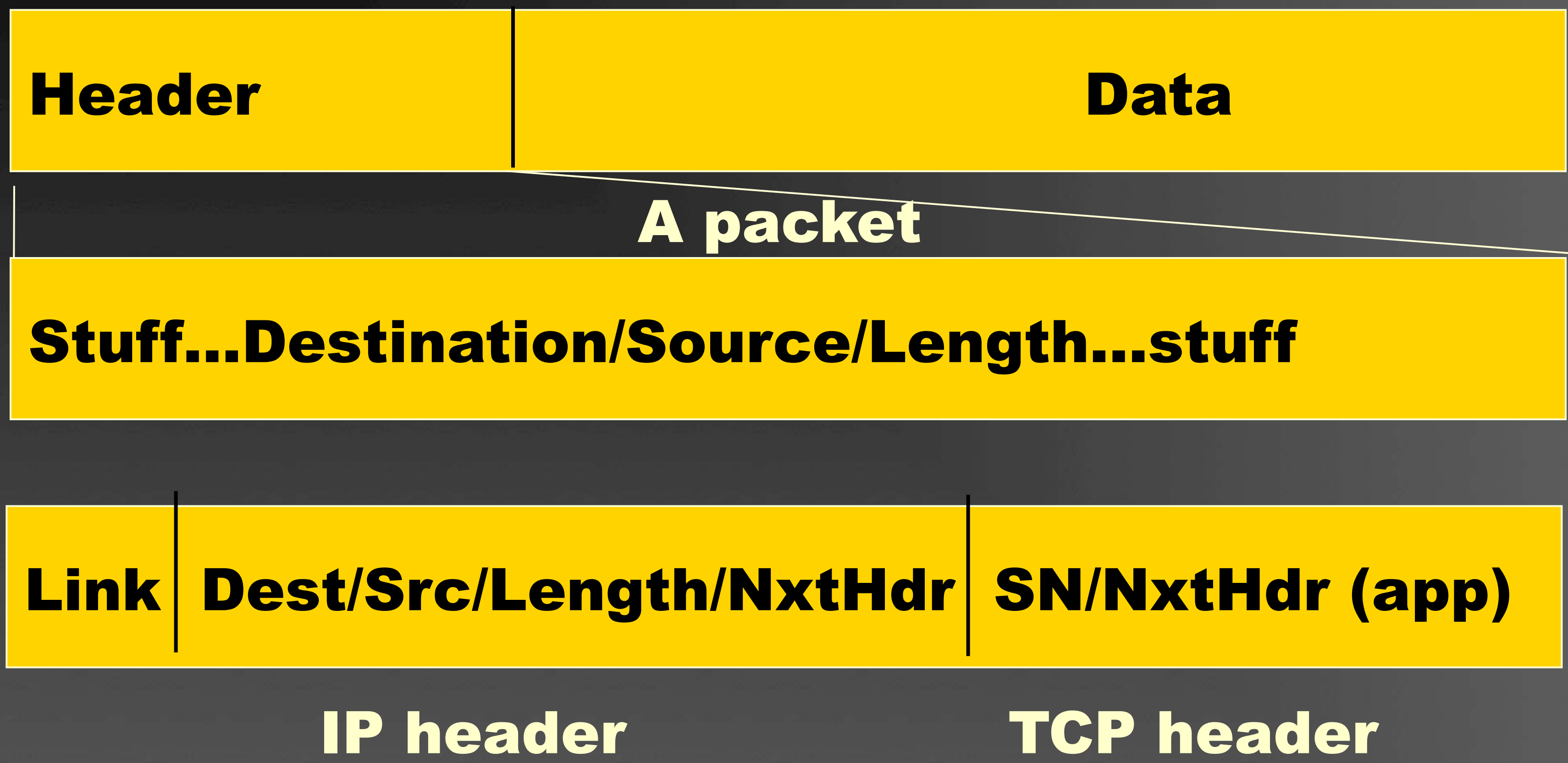
- How could an application designer prevent this sort of intervention?
- Encrypt the packets.
  - All anyone can see (unless they have the encryption key) is the header.
- Questions for later:
  - Should application designers and ISPs be in an arms race?
  - Should the user view his ISP as an enemy?

Host vs. router



*Not the whole truth

A (less) simple packet picture



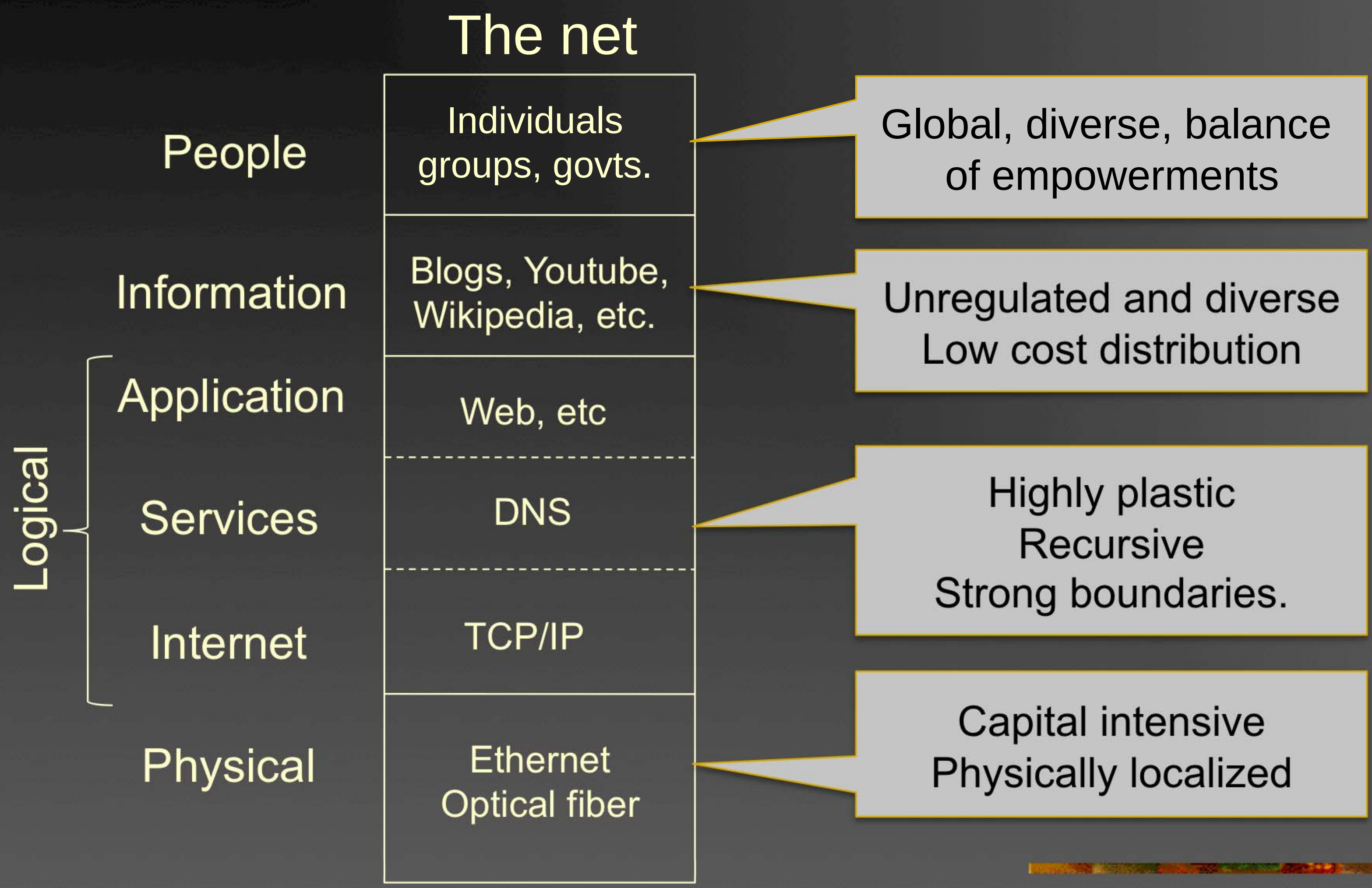
Implications

- Applications are the reason to have an Internet in the first place.
  - Only geeks send packets for the fun of it.
- Applications are where the value is generated, and money is to be made.
  - Ecommerce, advertising, etc.
- Applications greatly broaden the set of stakeholders.
  - Porn, music sharing, VoIP…
- Routers “just” forward packets.
  - Is this the right view?

A layered model of cyberspace

- The previous discussion suggests that we can describe the Internet (and cyberspace more generally) using a layered model.
  - A layered model is a classic way of Computer Science thinking.
  - Several layered models have been posed, including the formal OSI reference model.
- We will use a 4 layer model in our future discussions.

A layered model of cyber-space



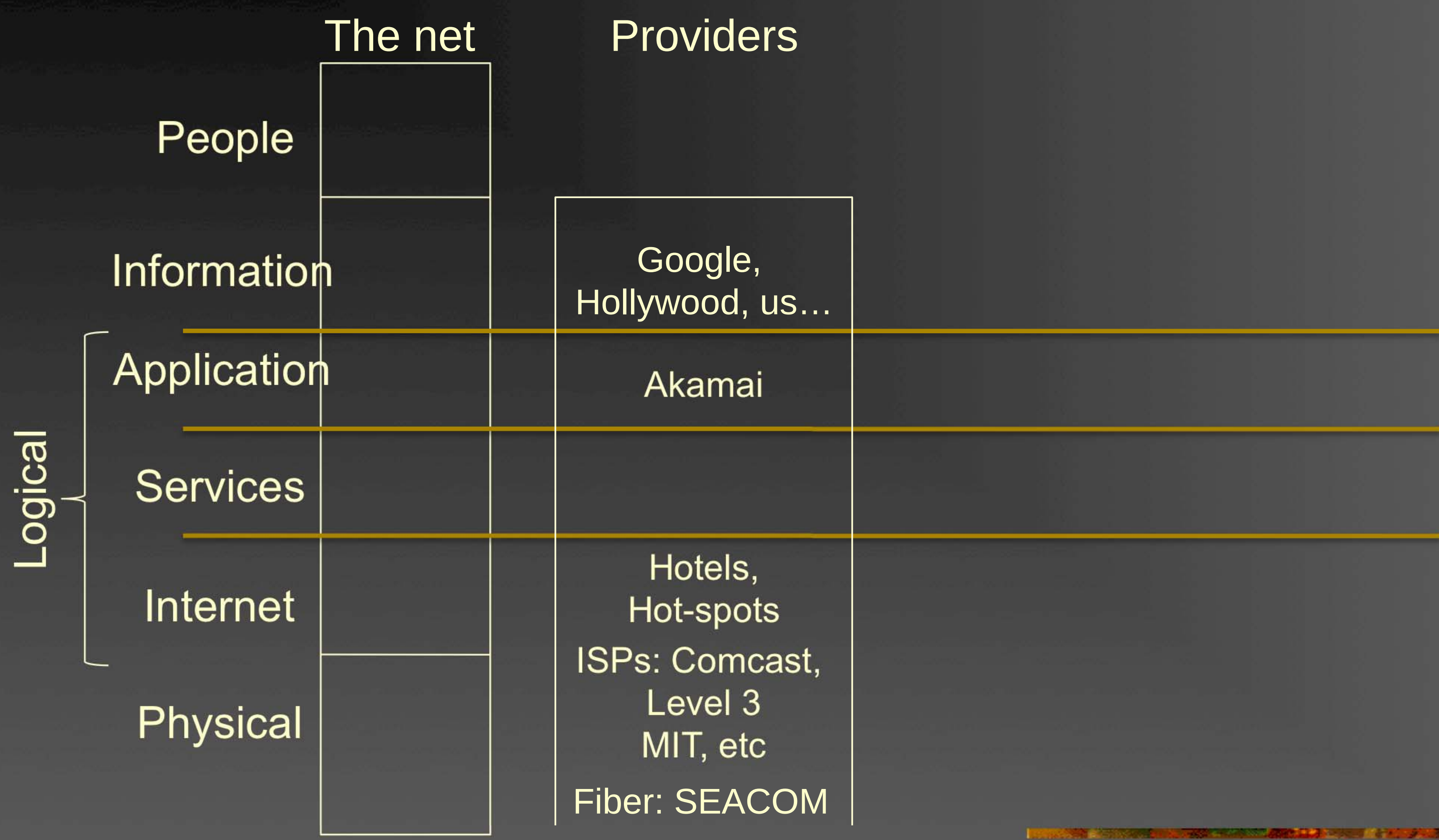
Cyberspace itself

Logical	Layer	The net
	People	Individuals groups, govts.
	Information	Blogs, Youtube, Wikipedia, etc.
Logical	Application	Web, etc
Logical	Services	DNS
Logical	Internet	IP
	Physical	Ethernet Optical fiber

Now add in the relevant actors.

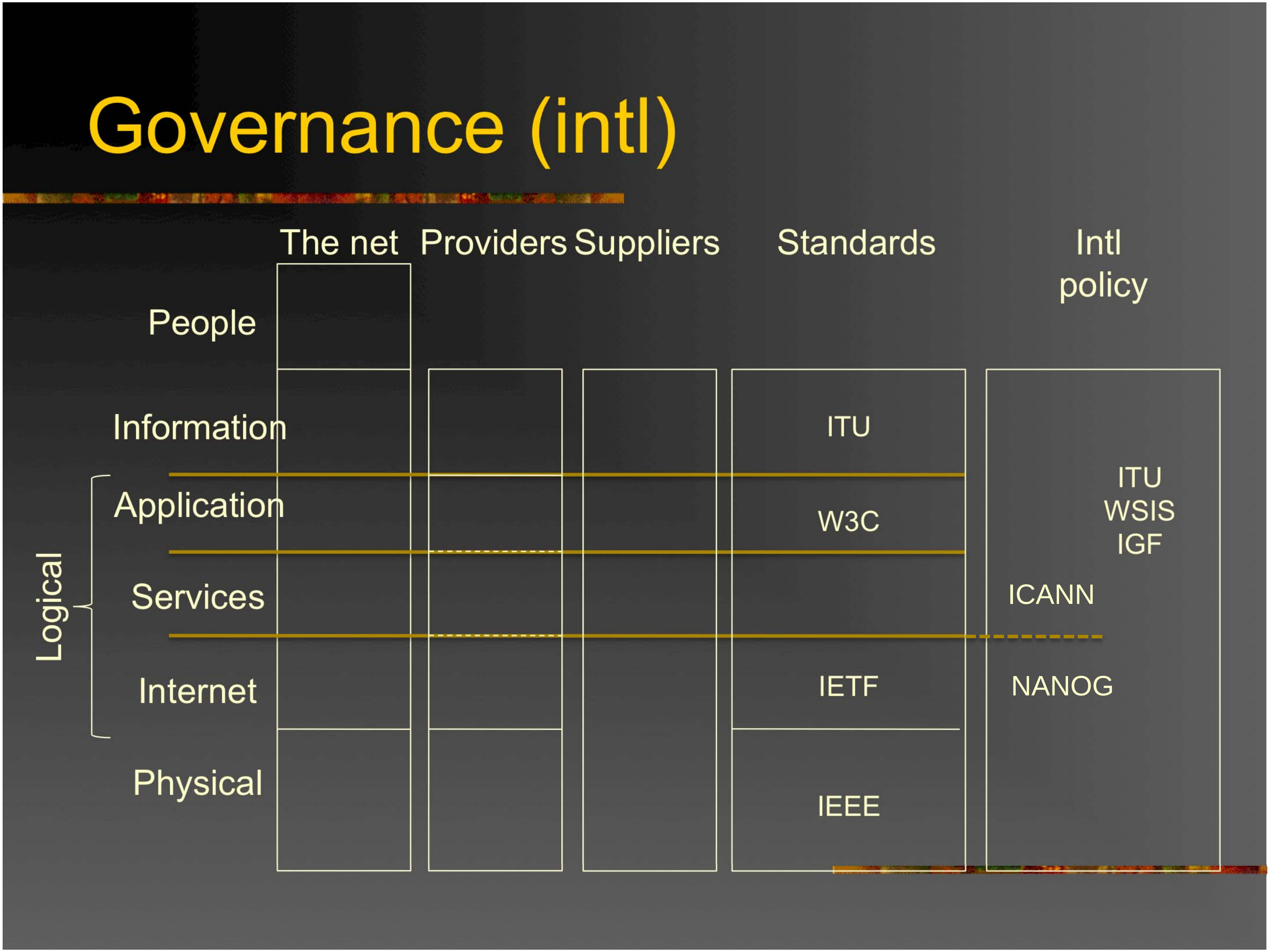
- We will emphasize the importance of cataloging actors, their objectives, their interactions, the tools of interaction, and the outcomes.
  - Which actors are successful in shaping cyberspace and its context, and by what means?
  - To test our approach: case studies of actors and their interactions.

Providers of service and content

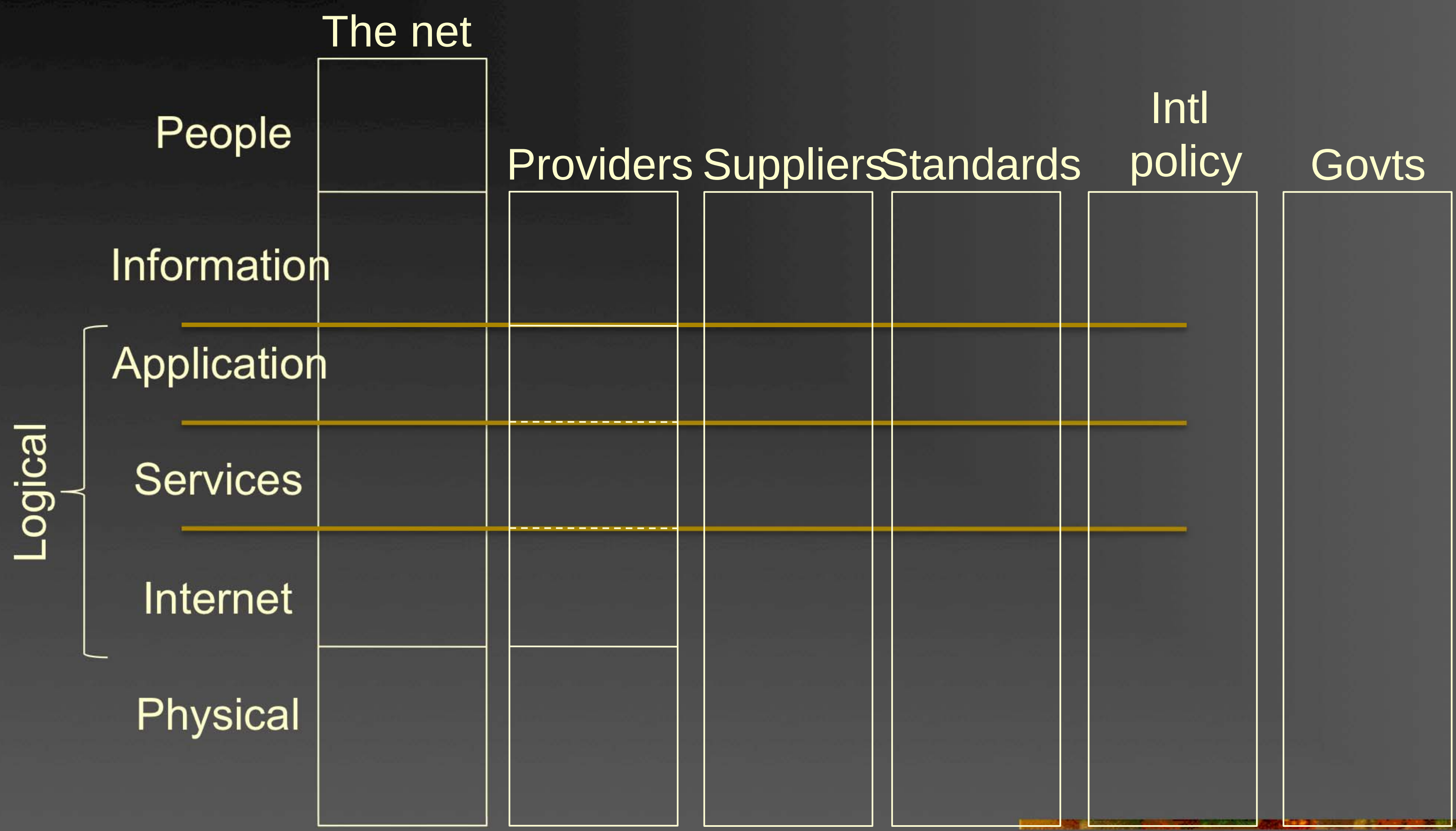


Equipment/technology

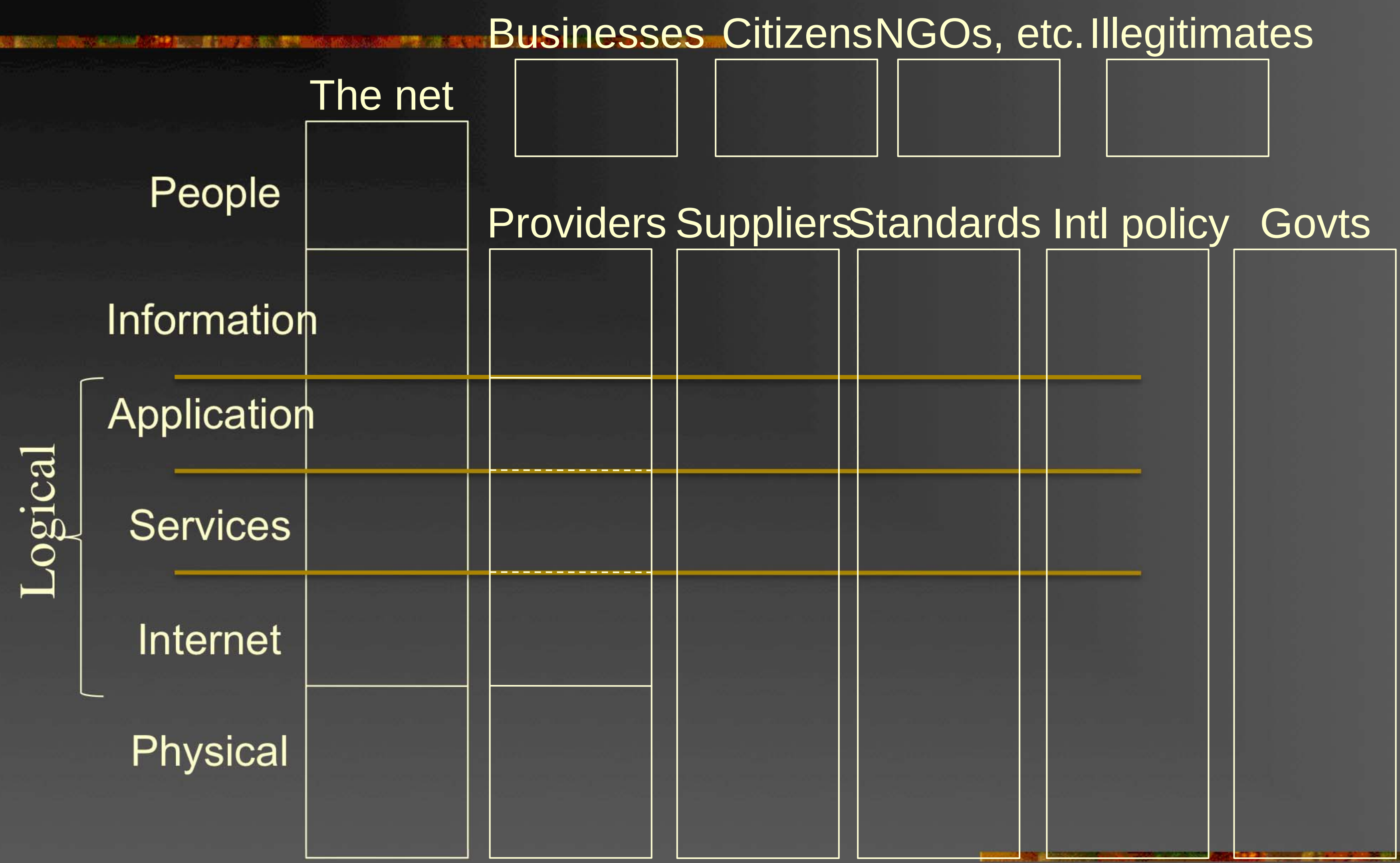
		The net	Providers	Suppliers
	People			
	Information			Data centers, Cloud providers
Logical	Application			Server hardware
Logical	Services			
Logical	Internet			Routers: Cisco, Huawei
	Physical			Fiber: Corning


Governance (intl)
The net
Providers
Suppliers
Standards
Intl policy
People
Information
Application
Services
Internet
Physical
Logical
ITU
W3C
IETF
IEEE
ITU
WSIS
IGF
ICANN
NANOG

Governments

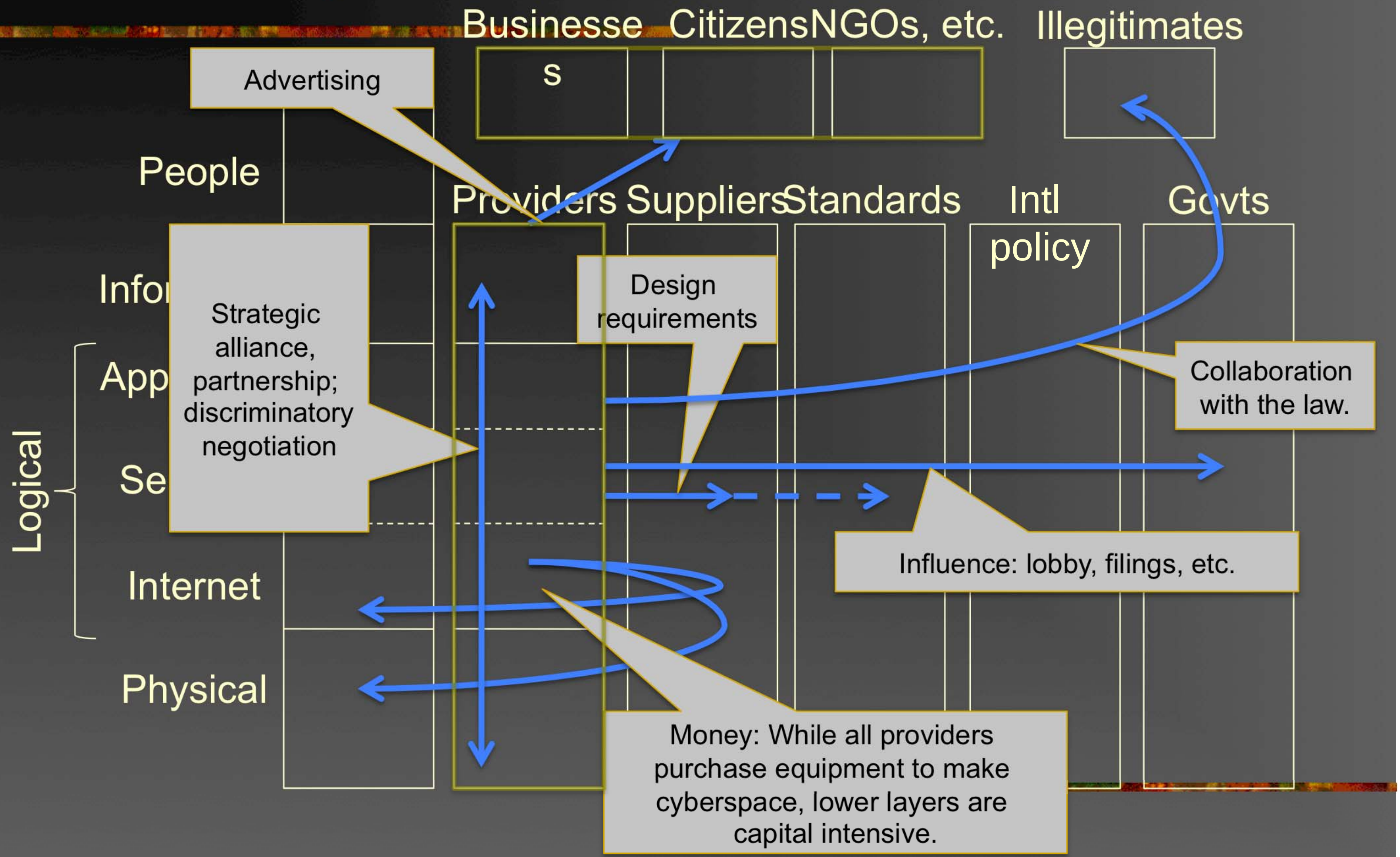


Users

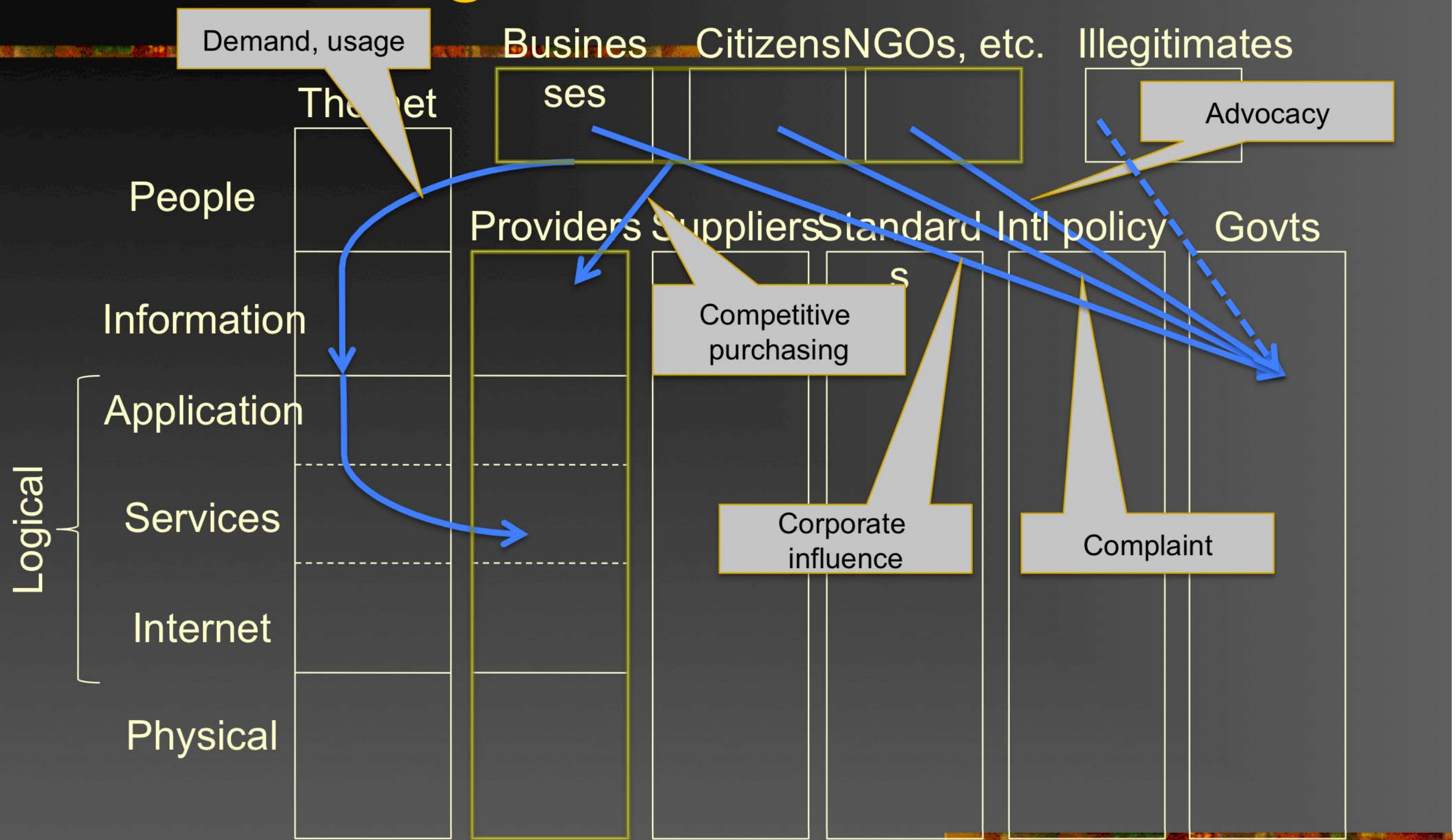


Businesses
Citizens
NGOs, etc.
Illegitimates
The net
People
Providers
Suppliers
Standards
Intl policy
Govts
Information
Application
Logical
Services
Internet
Physical

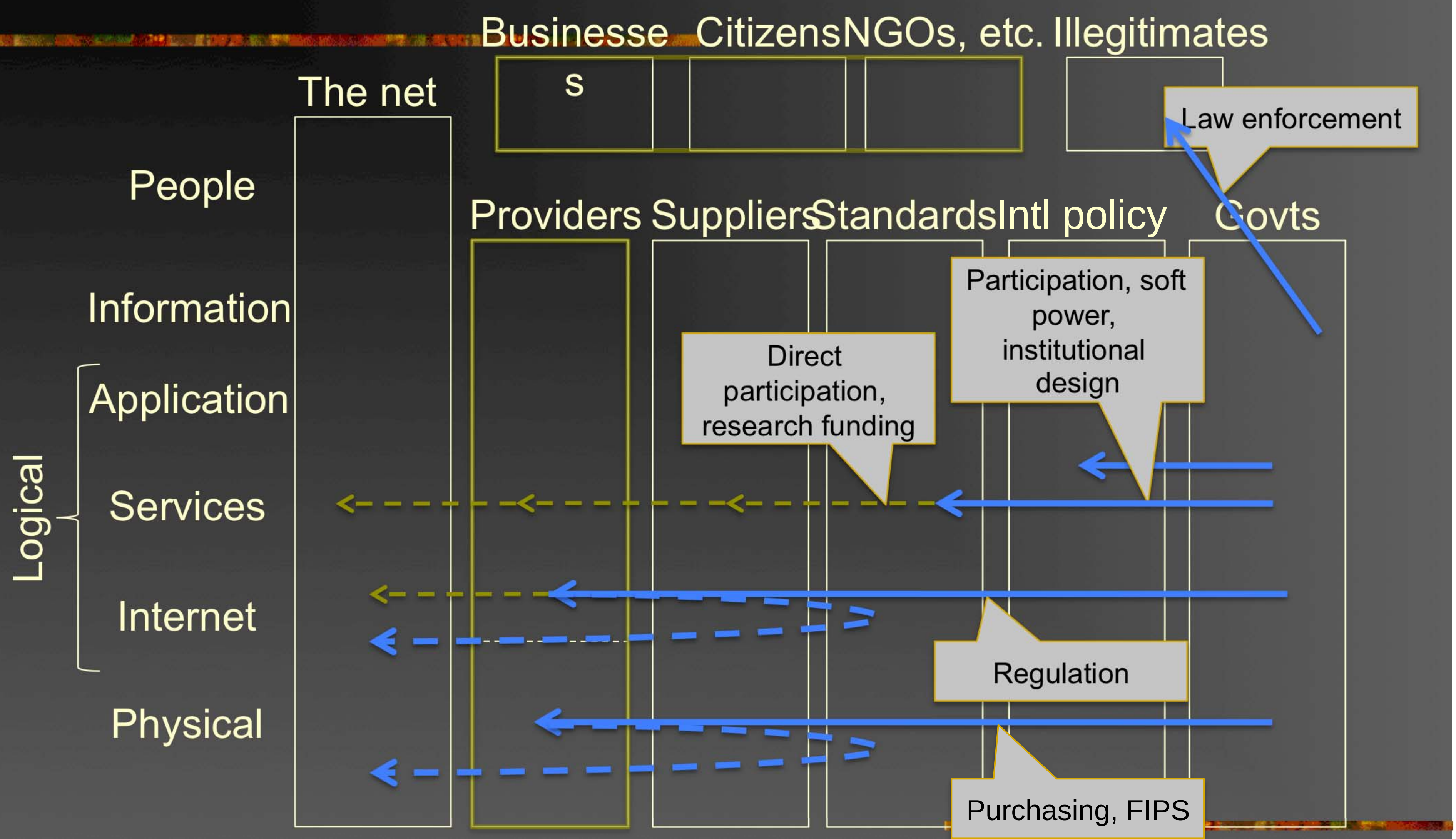
Action: providers



Action: legitimate users



Action: government



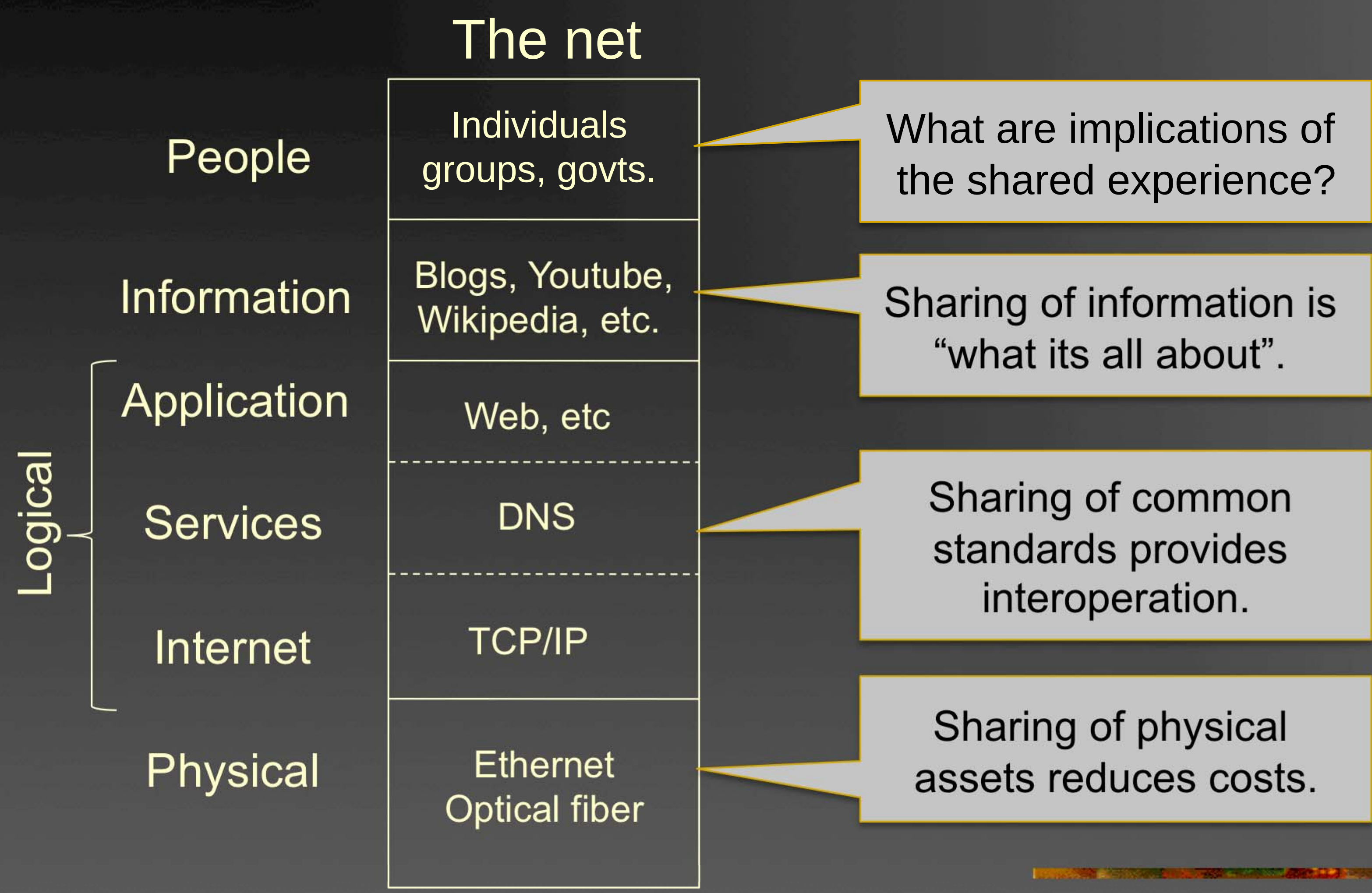
Case studies

- These three “influence pictures” are somewhat anecdotal.
- As part of research, need to provide robust grounding.
  - What tools are used?
  - Which are effective?
  - What are the range of motivations.

Interactions

- If we drew all the arrows from all the case studies:
  - The picture would be impossible to understand.
  - It would emphasize the dynamic nature of the interactions.
    - Many cycles among the actors.

Sharing at the different levels



17.447 / 17.448 Cyberpolitics in International Relations: Theory, Methods, Policy
Fall 2011