Web Requests, Down To The Atom

Chris Zacharias

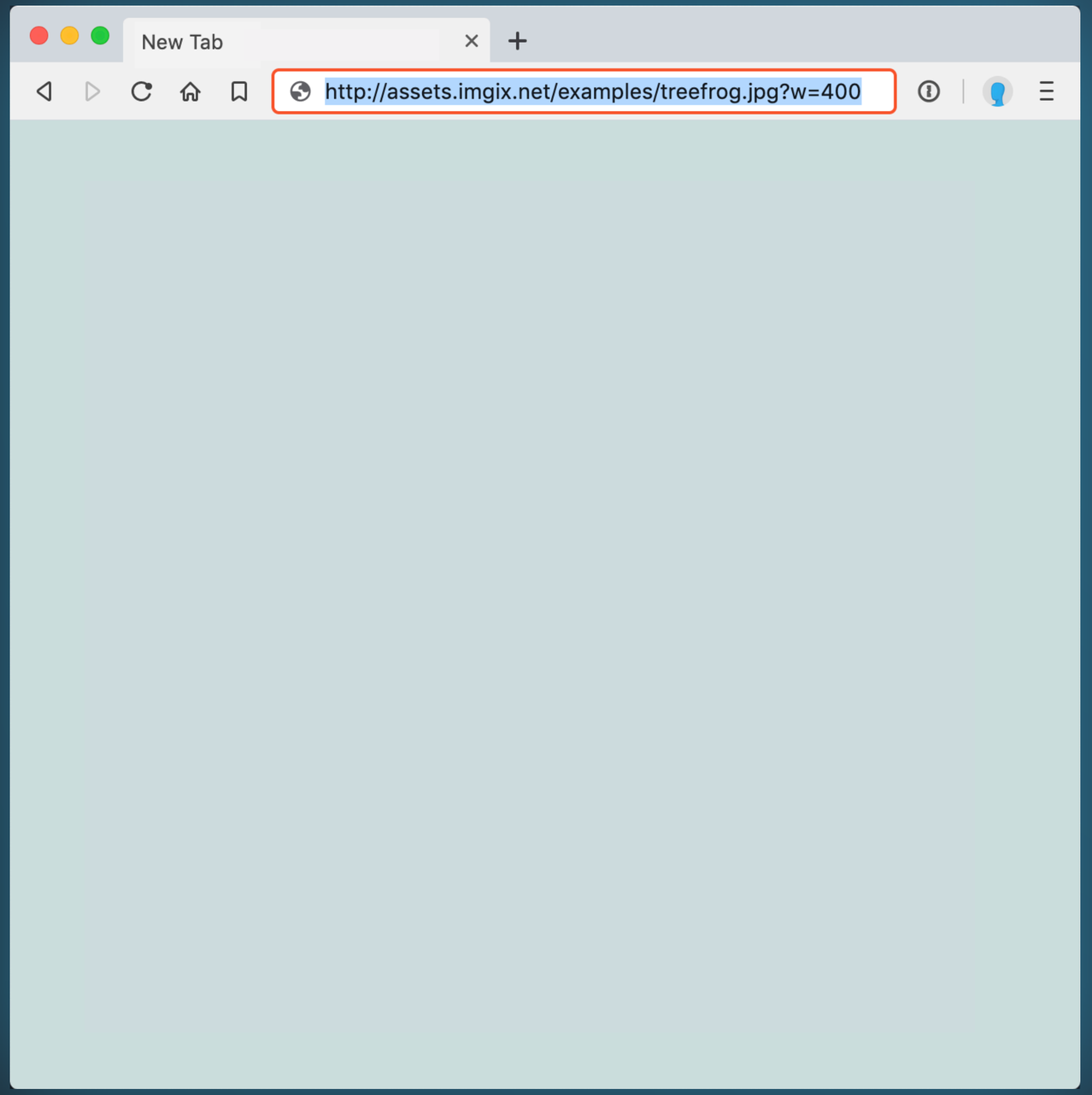
Web Unleashed 2019

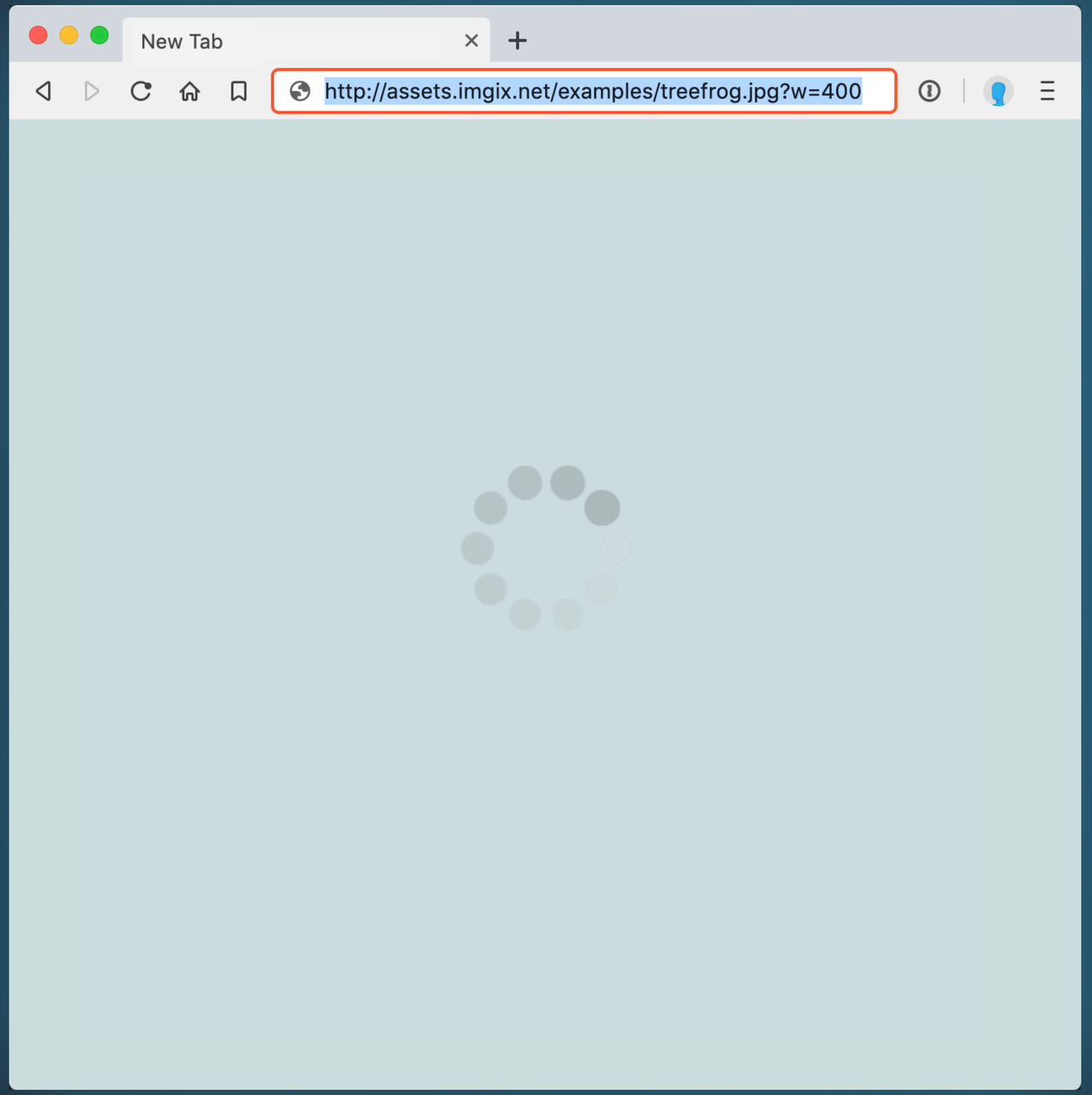
Toronto

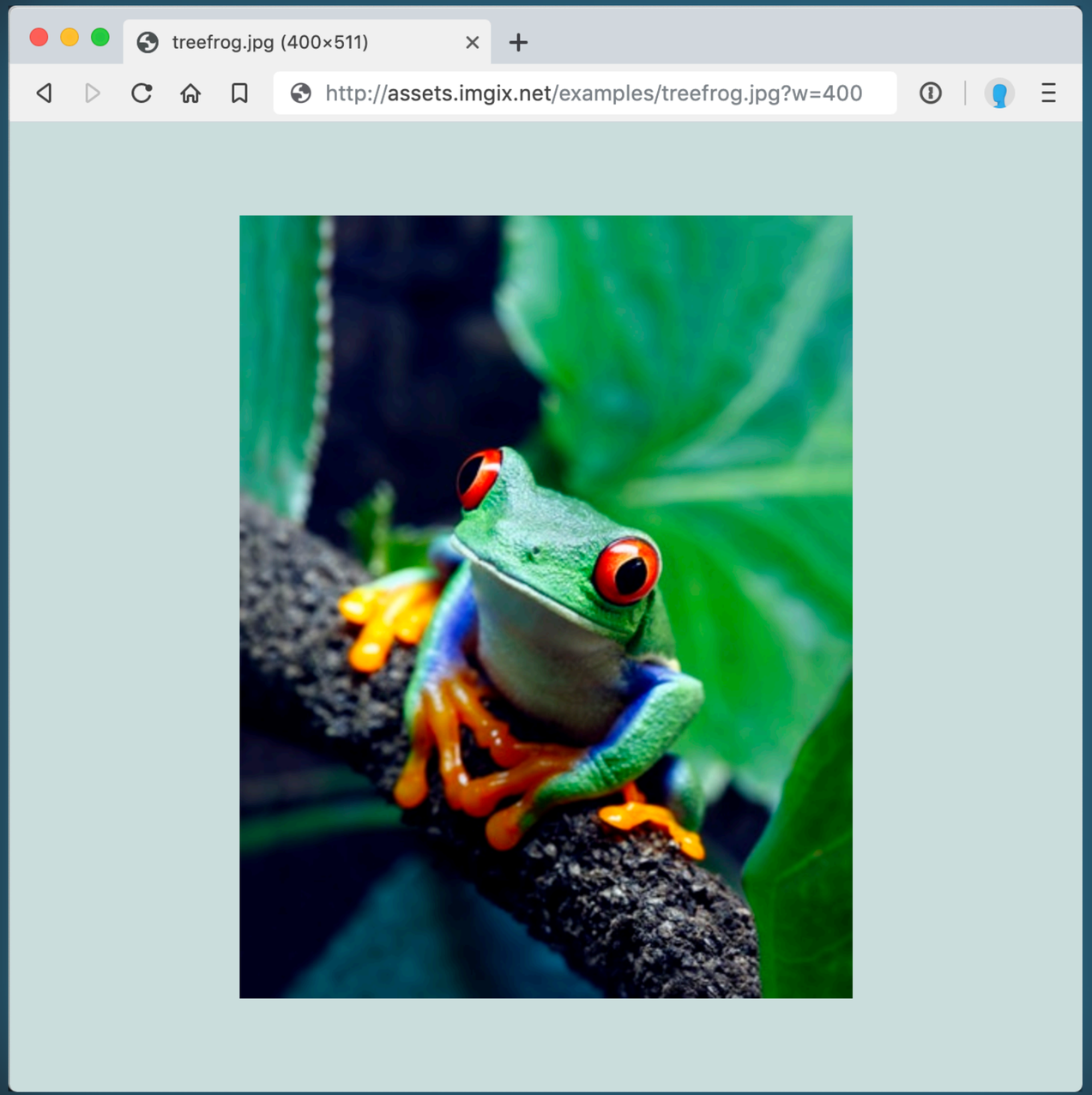
imgIX

Understanding how the Internet fundamentally works is important.

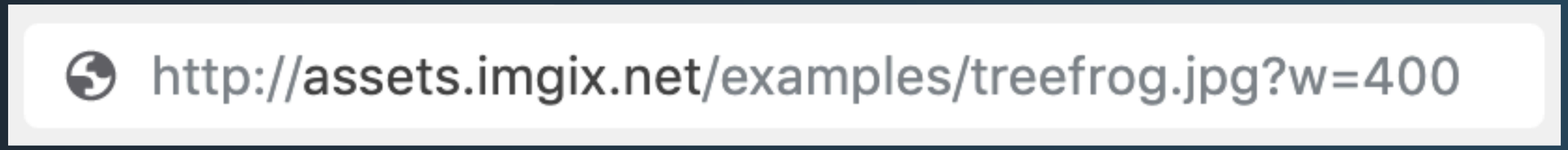
- Identify performance opportunities at every layer.
- Make smarter decisions around compression and optimization.
- Become better at triaging and debugging network-related issues.
- Recognize potential attack vectors and vulnerabilities.
- Understand when and where costs occur.
- Communicate better with other engineers.


New Tab
http://assets.imgix.net/examples/treefrog.jpg?w=400

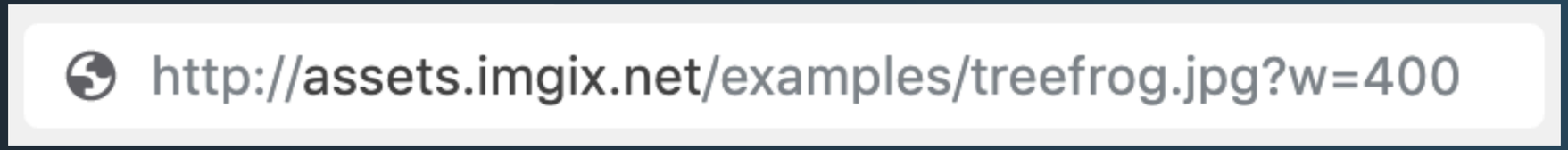

New Tab
http://assets.imgix.net/examples/treefrog.jpg?w=400


treefrog.jpg (400×511)
http://assets.imgix.net/examples/treefrog.jpg?w=400

A web request begins with a URL.



The URL answers three simple questions.



What kind of request is it?

Where do I send the request?

What is it that I want to request?

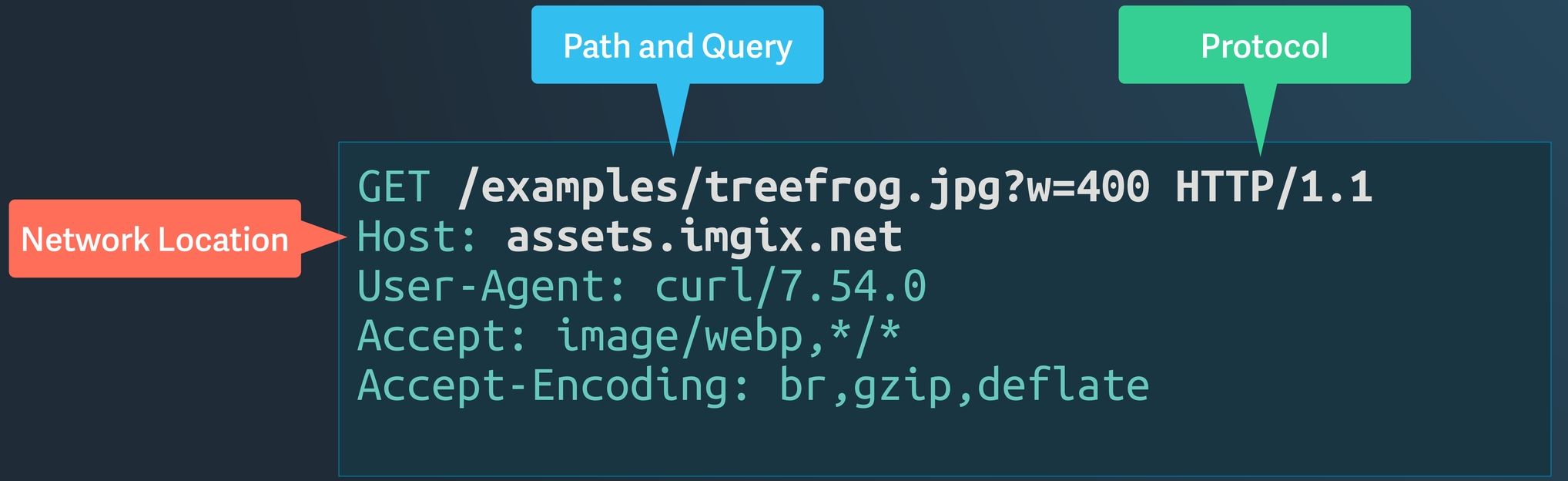
The URL answers three simple questions.

What kind of request is it?

Where do I send the request?

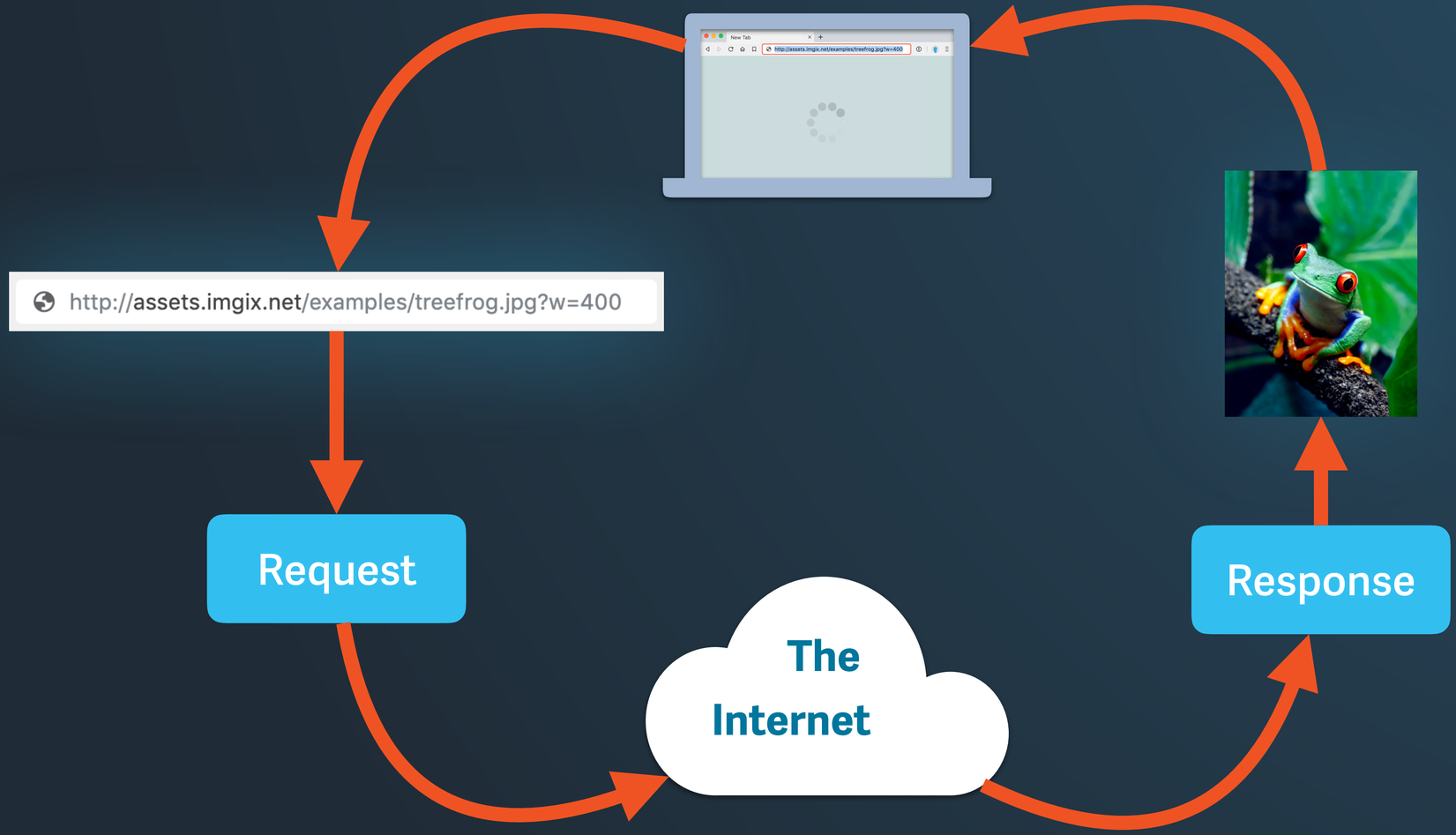
What is it that I want to request?

The browser builds an HTTP request out of the URL we provide.

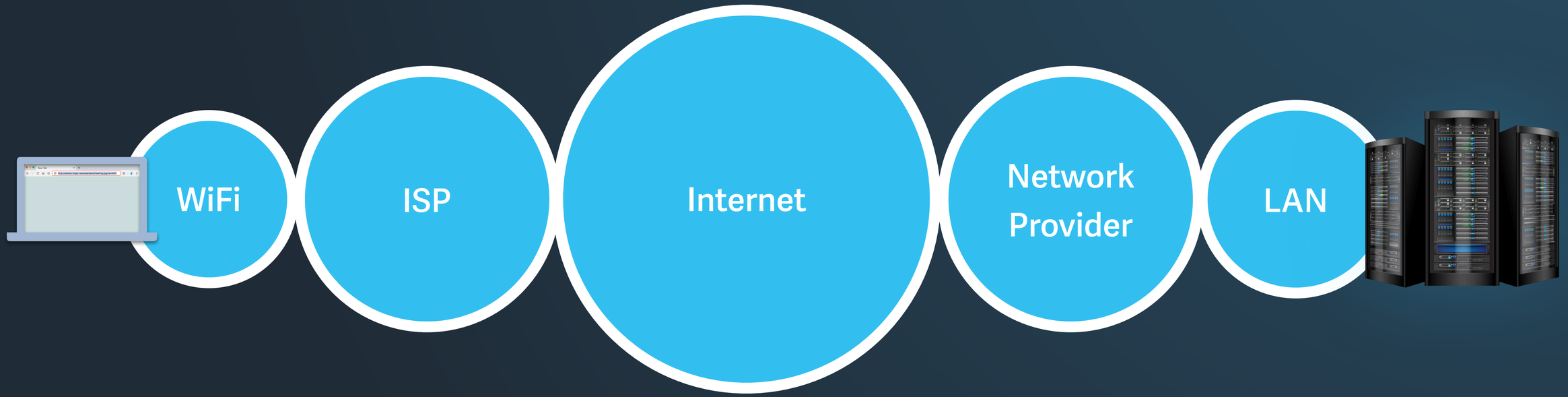


An HTTP GET request.

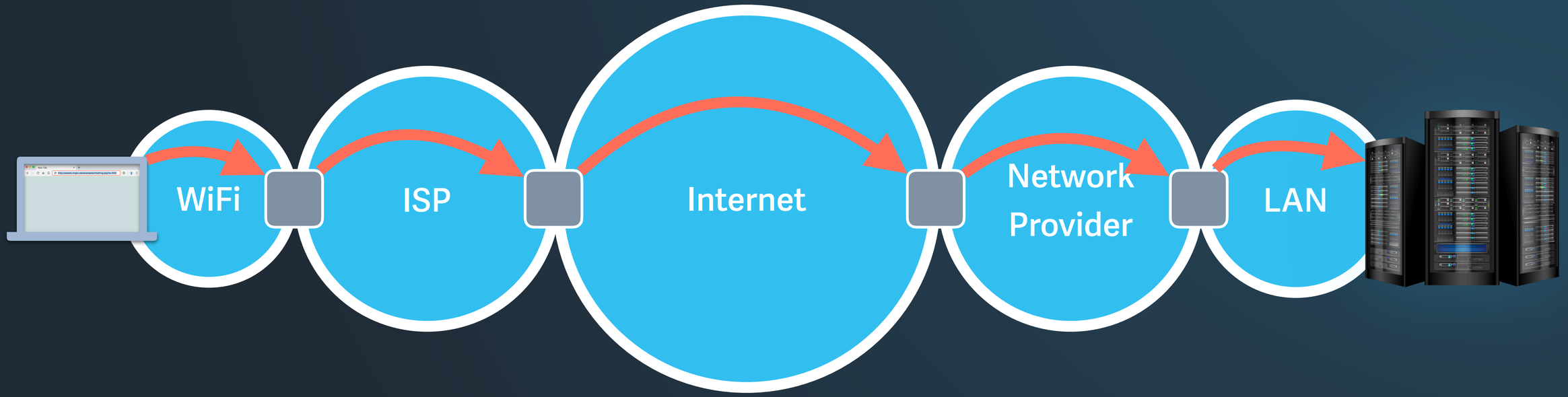
The request is sent out onto the Internet and (hopefully) we get a response.



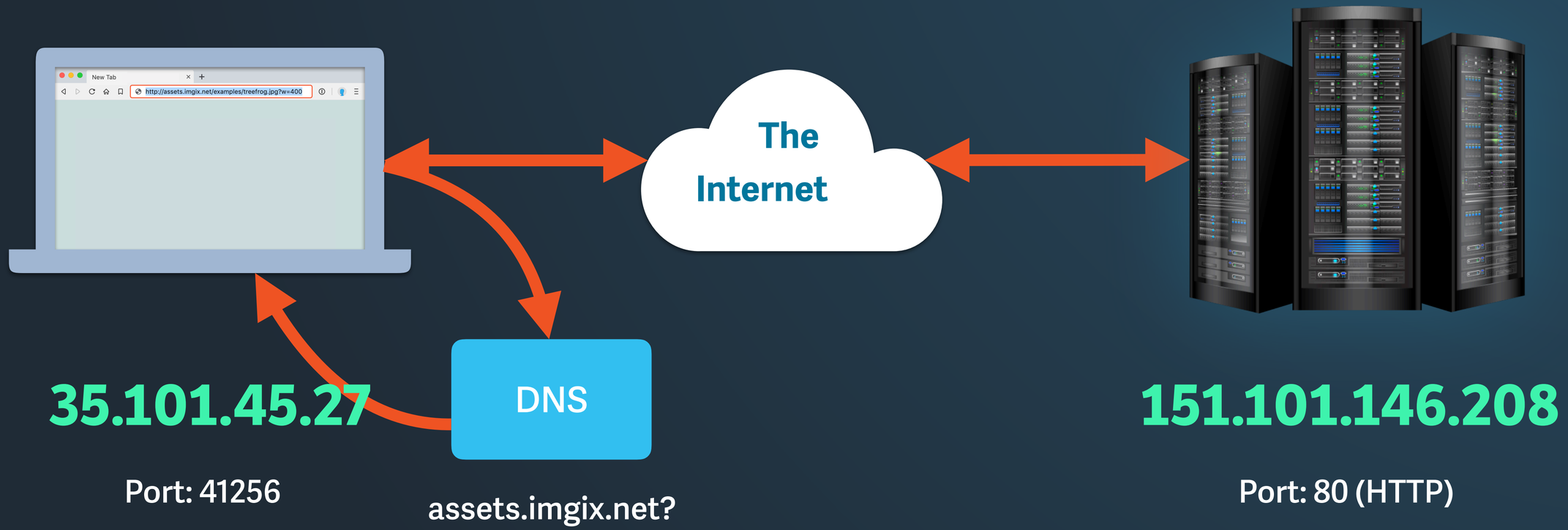
The Internet is just a collection of networks that connect clients to servers.



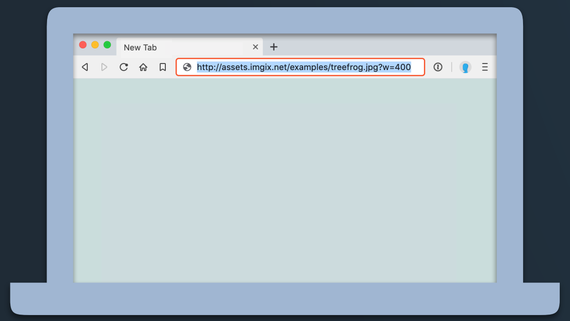
By “hopping" between networks, we can find our way to the right server.



To perform a request, we need a network connection, an IP address and a port.



Once we have those things, we can establish a TCP connection.



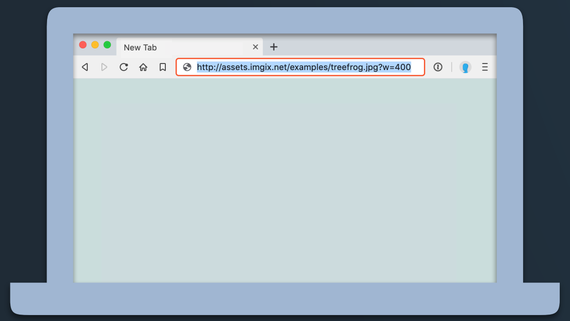
35.101.45.27

Port: 41256

151.101.146.208

Port: 80 (HTTP)

Packets are sent between client and server to transfer data.



35.101.45.27

Port: 41256

TCP Handshake

DATA (GET /file.jpg)

DATA (200 OK)

DATA (Response)

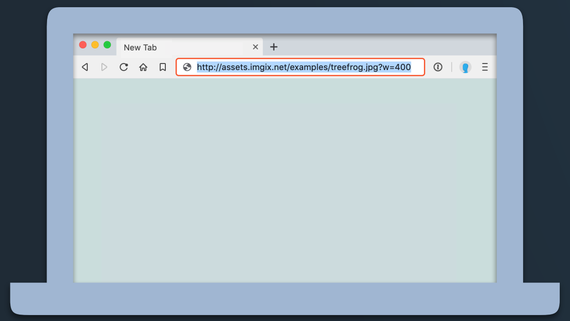
ACK

TCP Teardown

151.101.146.208

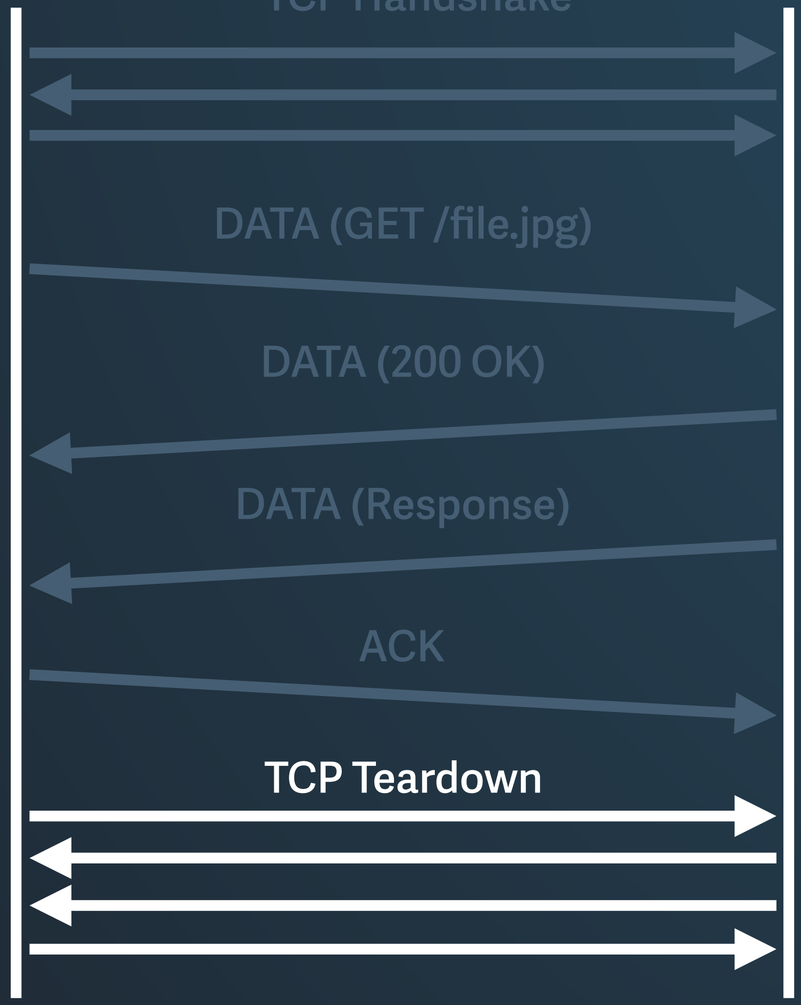
Port: 80 (HTTP)

The connection is shut down when it is no longer needed.



35.101.45.27

Port: 41256

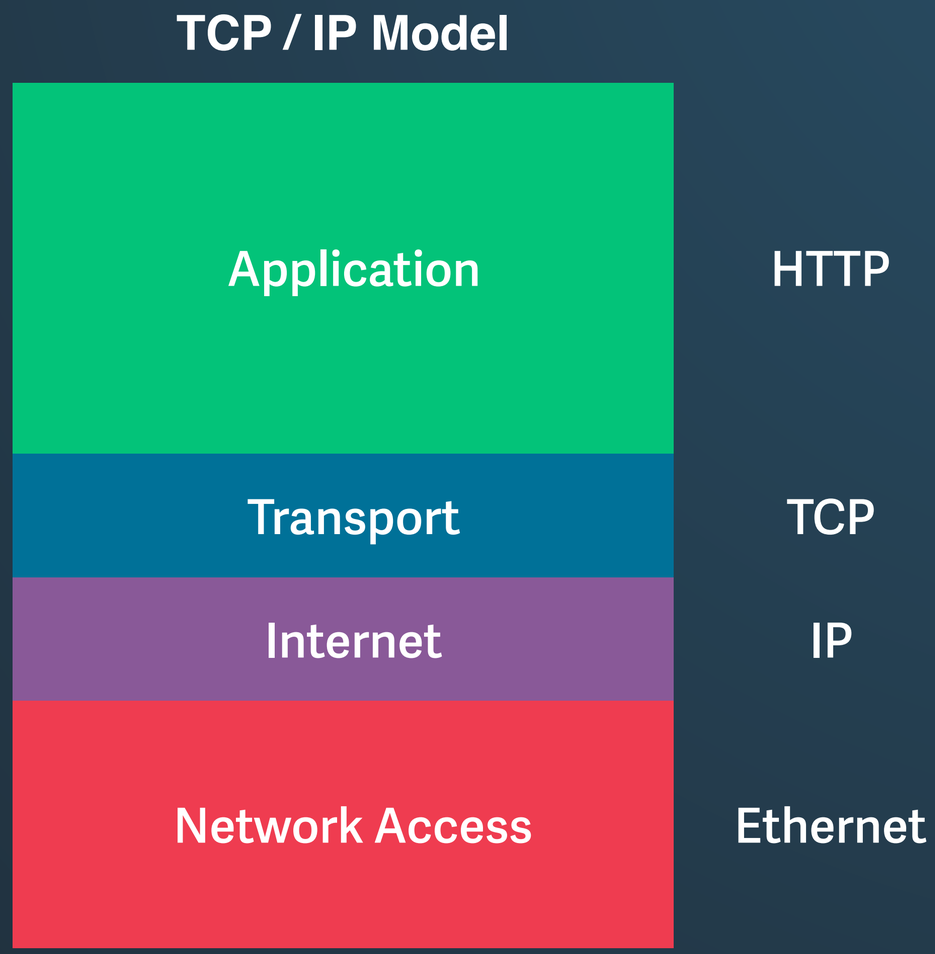


151.101.146.208

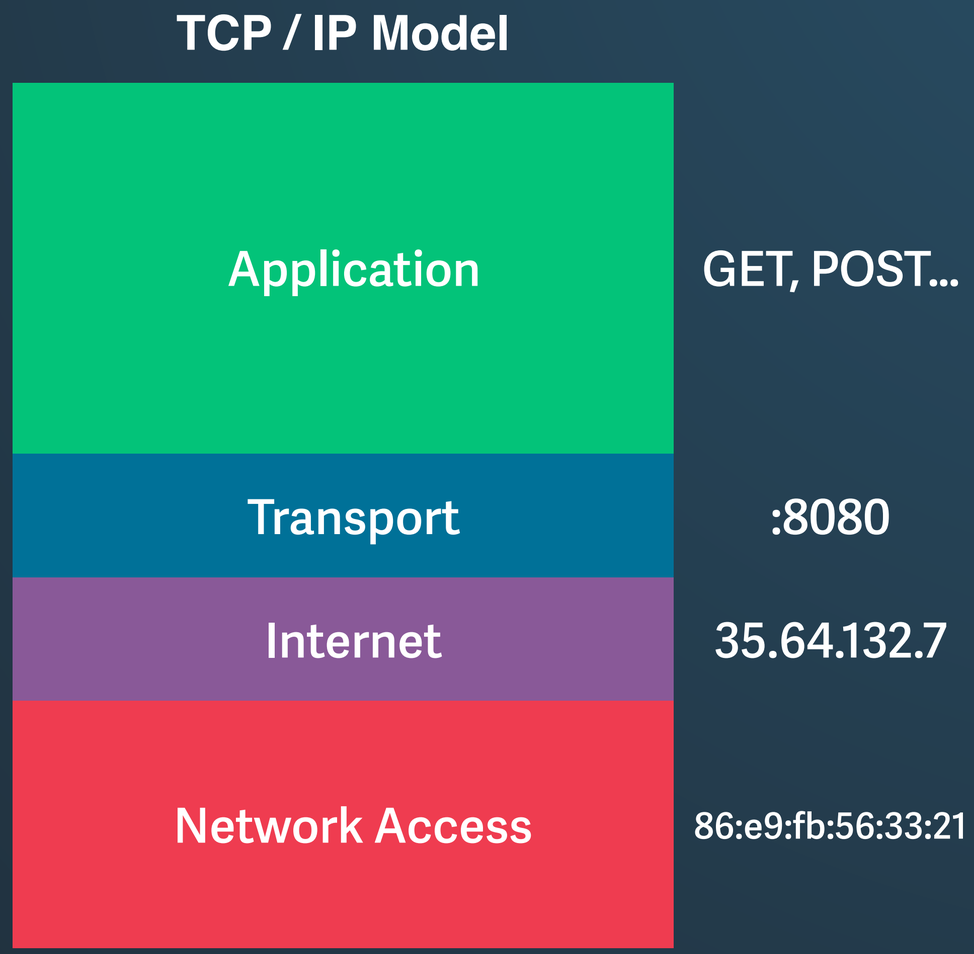
Port: 80 (HTTP)

To describe the networking stack,
we make use of abstract models.

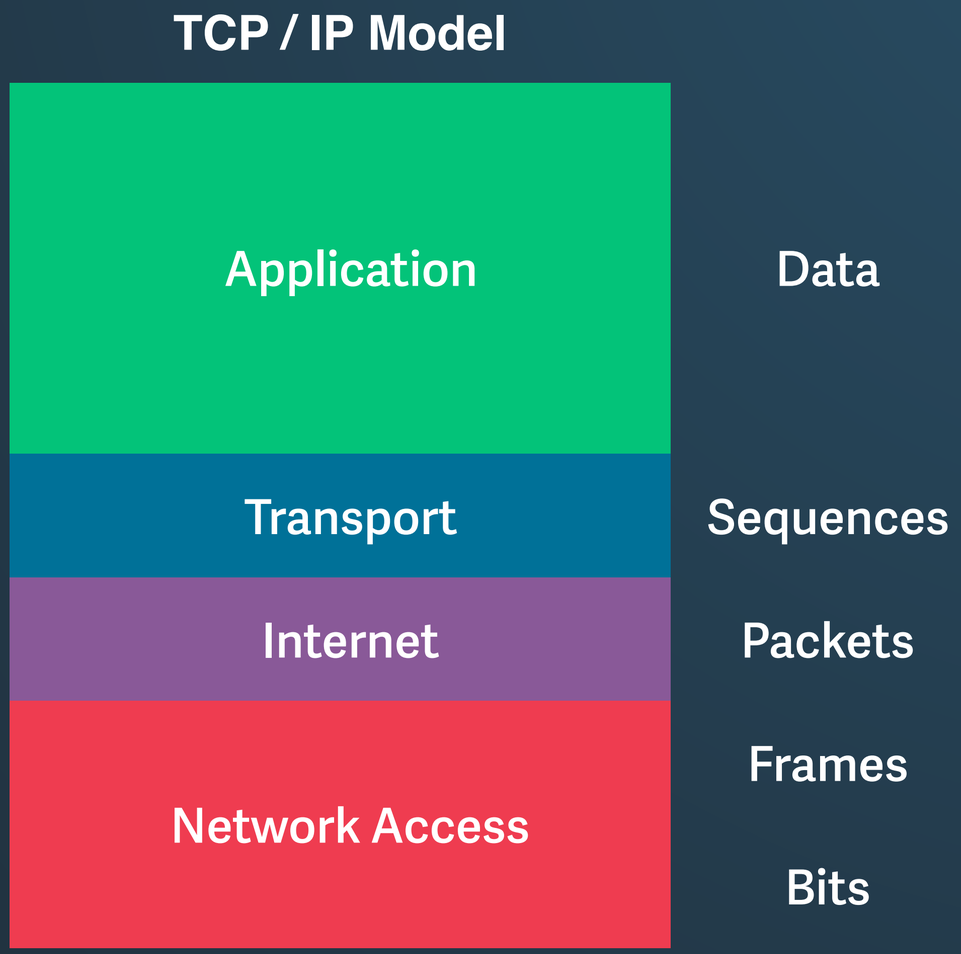
This is the most common understanding of the network stack.



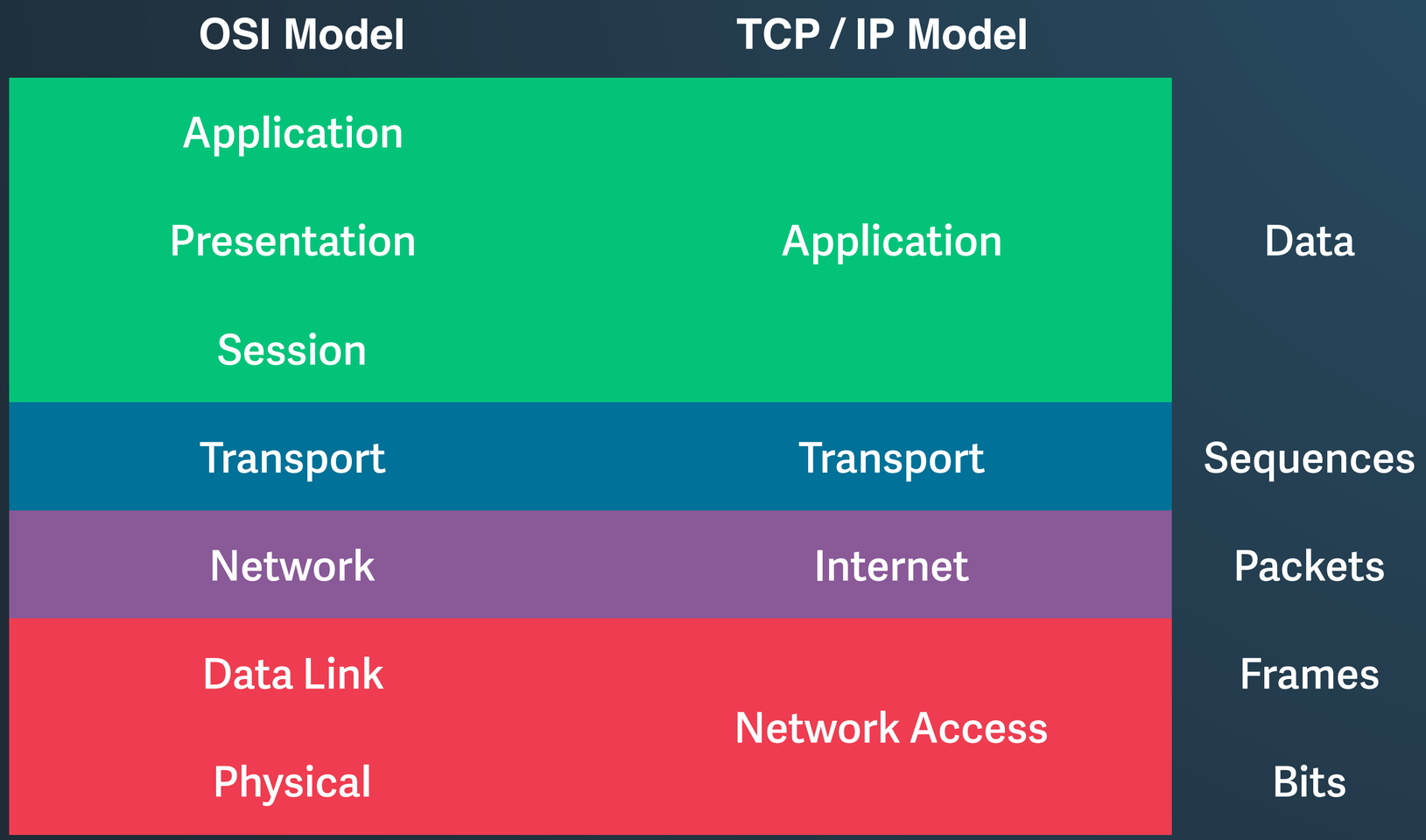
This is the most common understanding of the network stack.



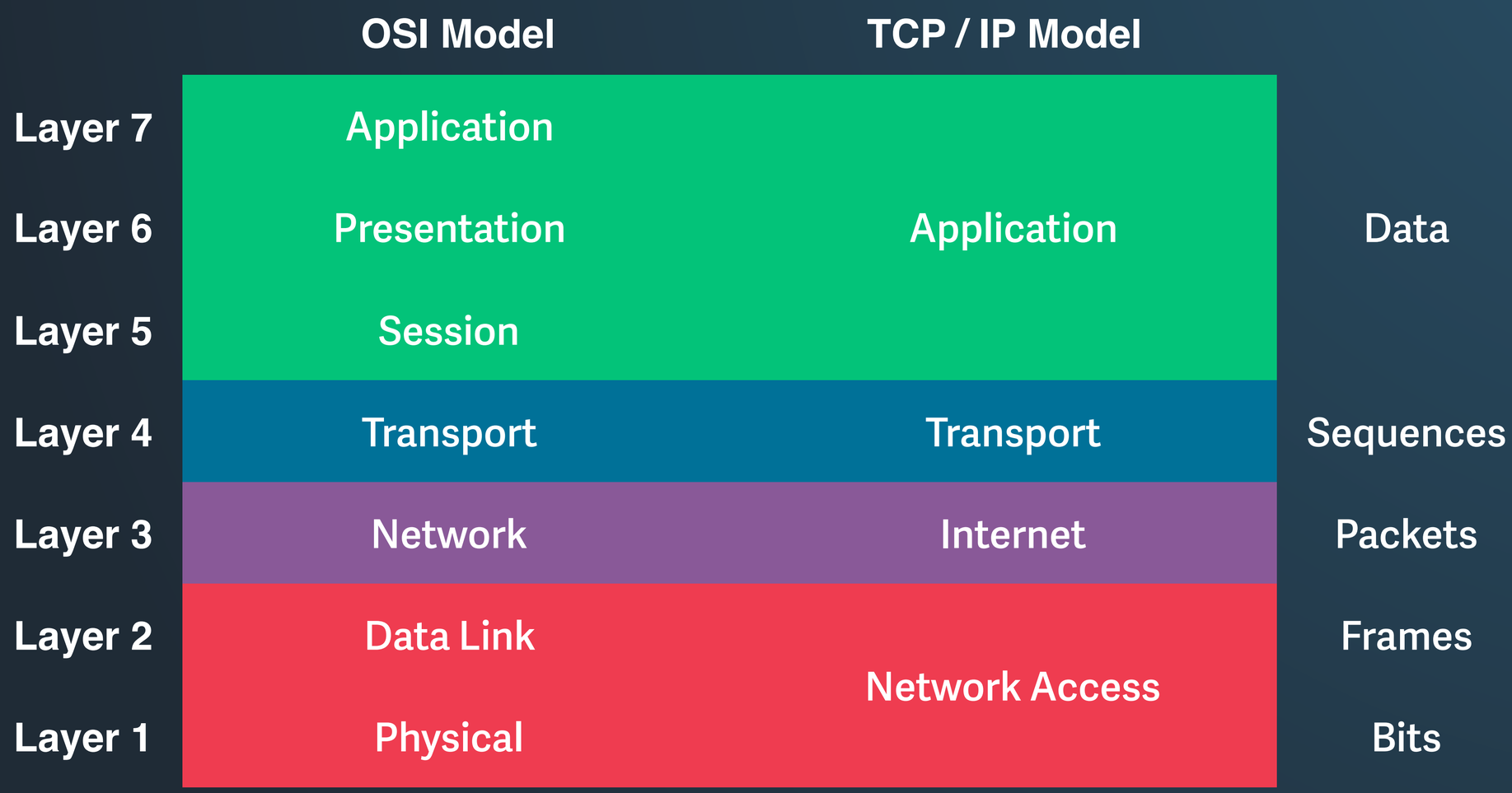
This is the most common understanding of the network stack.



The OSI model is often used to talk about networking in the abstract.



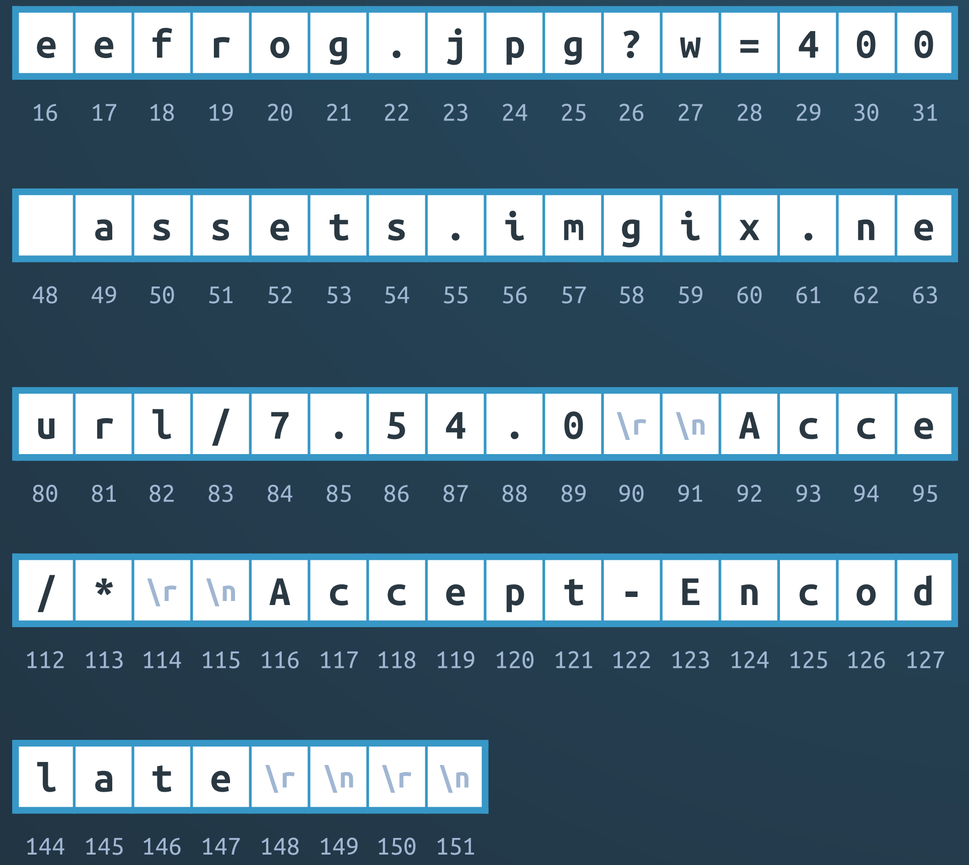
You may frequently hear about features being “L7”, “L4”, etc.



Given that the Internet is stateless, we need to provide a lot of extra information with our request.

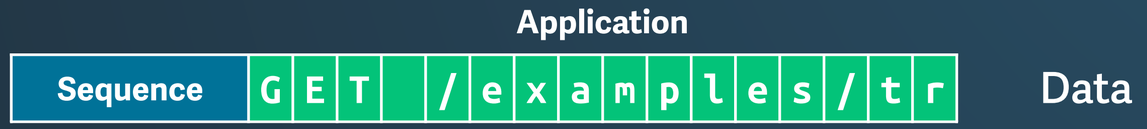
The data in a web request is encoded as a stream of bytes.

The stream is broken apart into chunks.

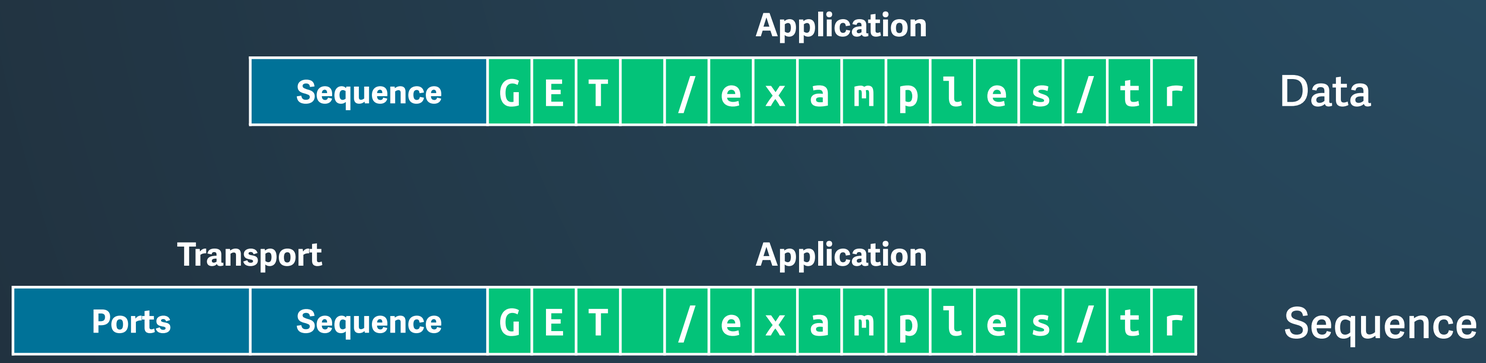


Those chunks are given IDs to keep them in order.

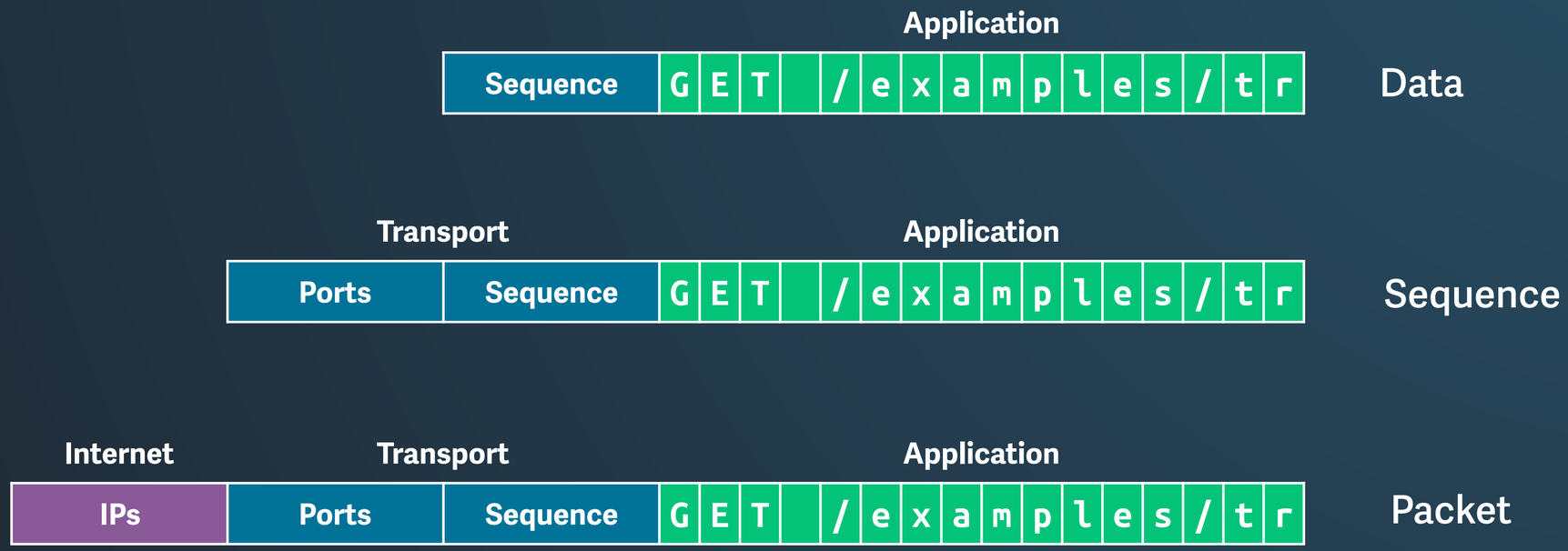
The chunks need to be packaged for sending across the Internet.



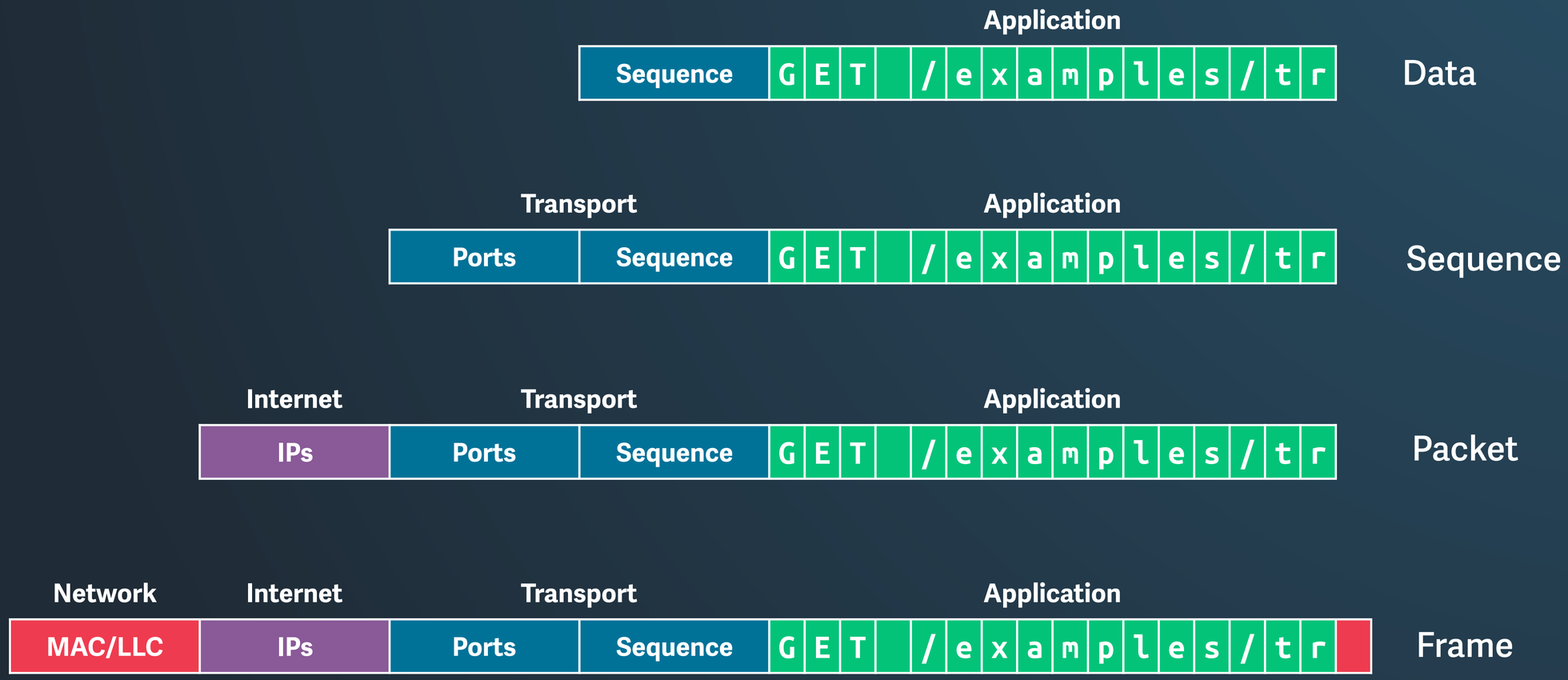
The TCP headers are added to handle the connection between two hosts.



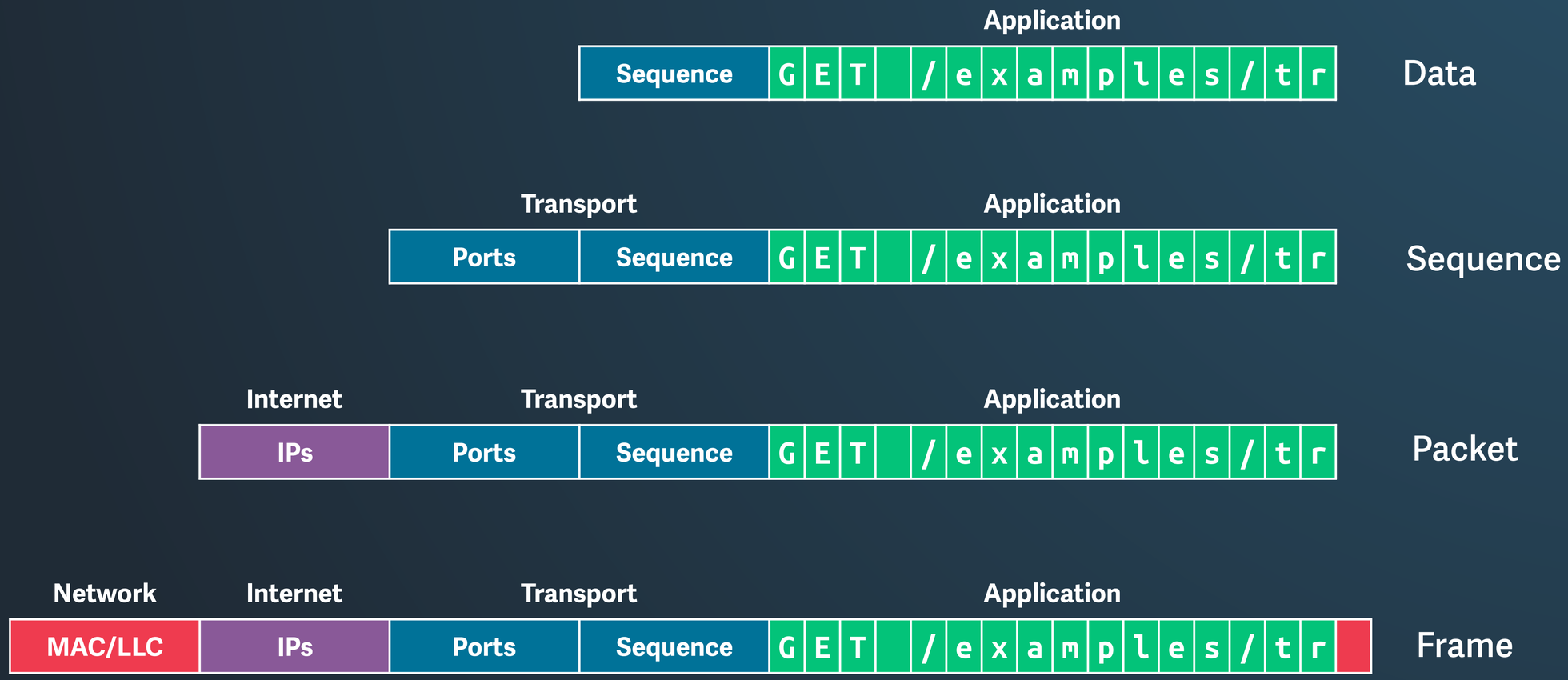
The IP headers provide the necessary context to navigate the Internet.



The MAC/LLC headers manage how packets navigate the local network.



The MAC/LLC headers manage how packets navigate the local network.

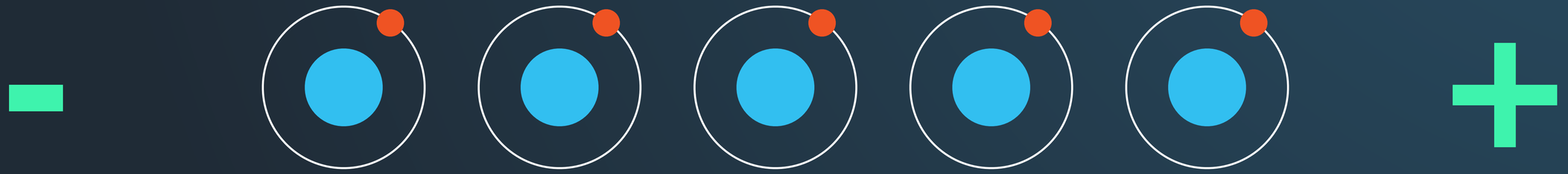


The standard “MTU” or Maximum Transmission Unit is 1526 bytes.

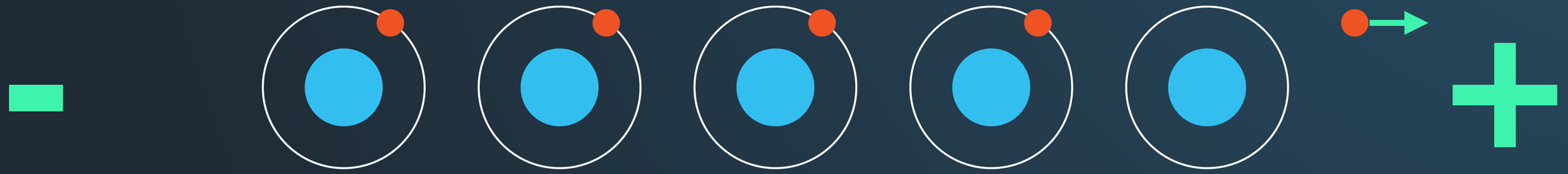
To understand how data is sent over a physical wire, we are going to work backwards from the atom.

Electrically conductive atoms have a single electron in their outermost shell.

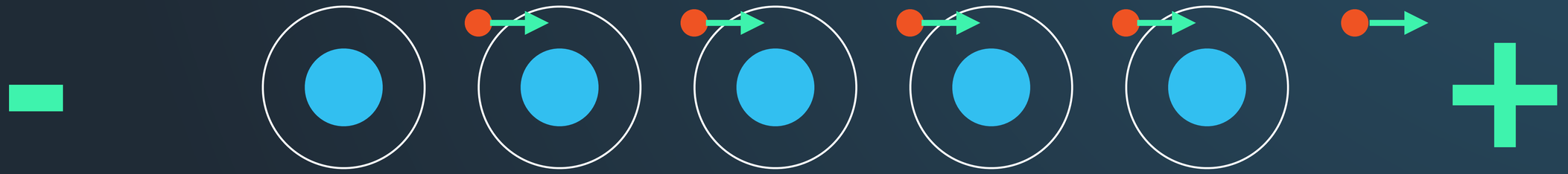
By adding poles with positive and negative charges...



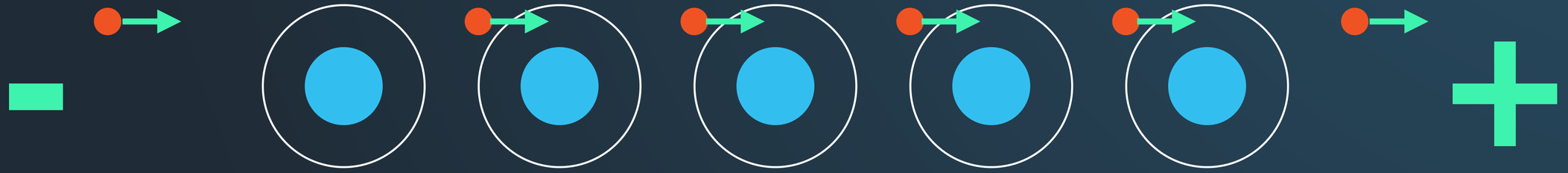
... we cause the atoms nearest to the positive pole to lose an electron.



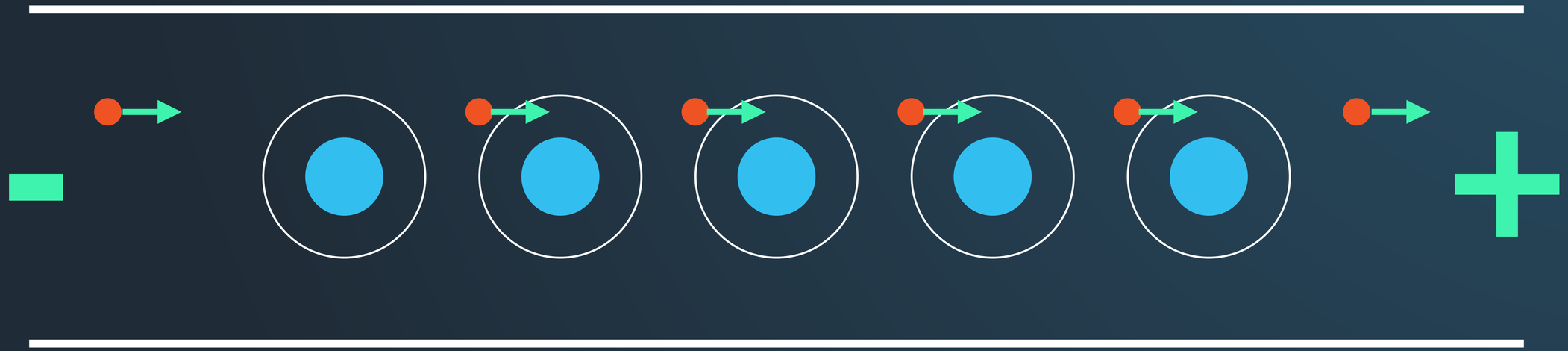
Atoms missing an electron pull electrons from other atoms.



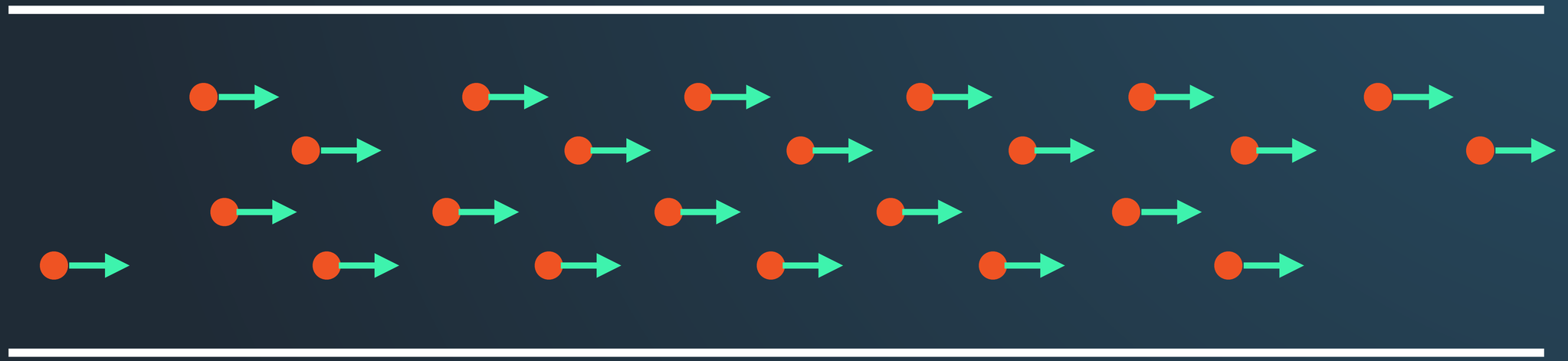
The negative pole provides electrons to atoms that end up without one.



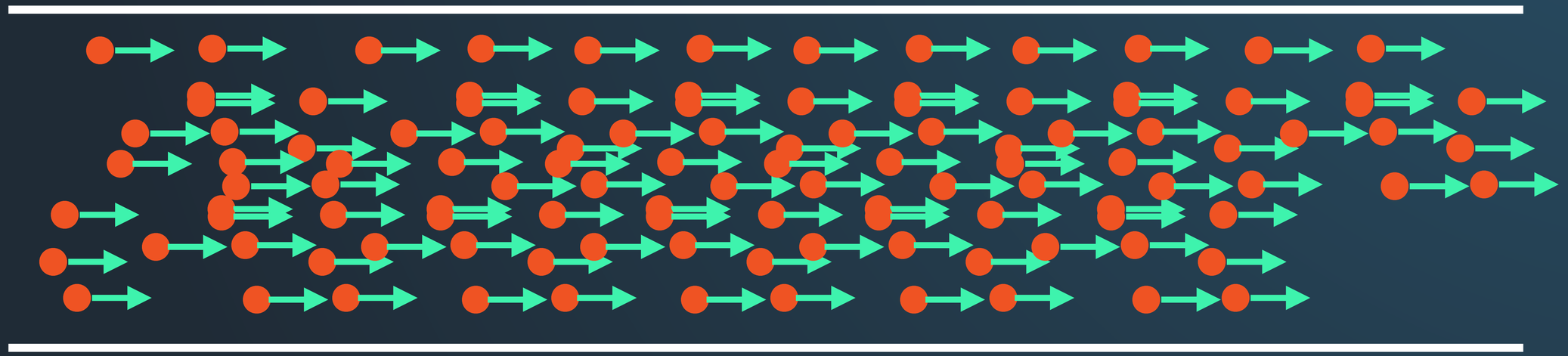
This creates an electrical current.



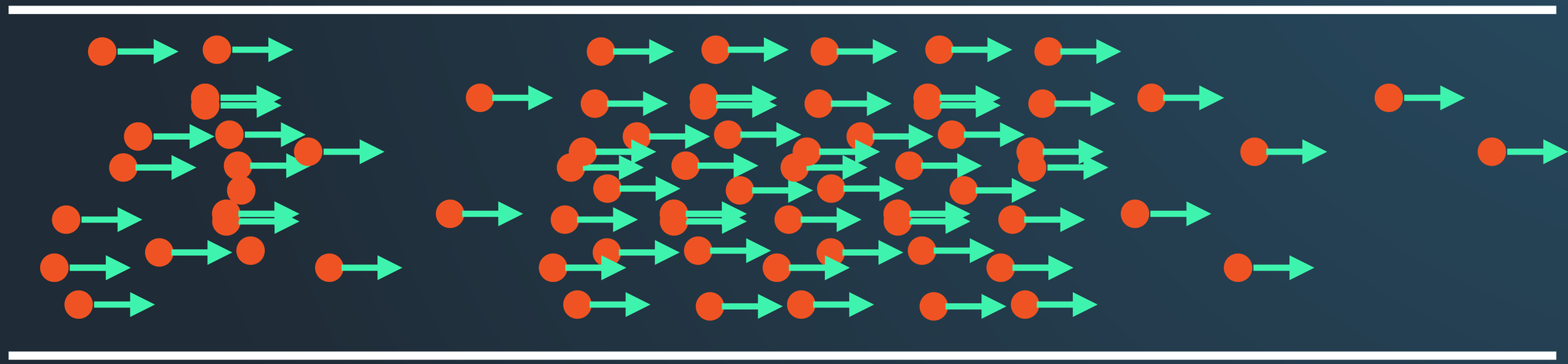
In direct current, electrons flow in one direction at a rate known as voltage.



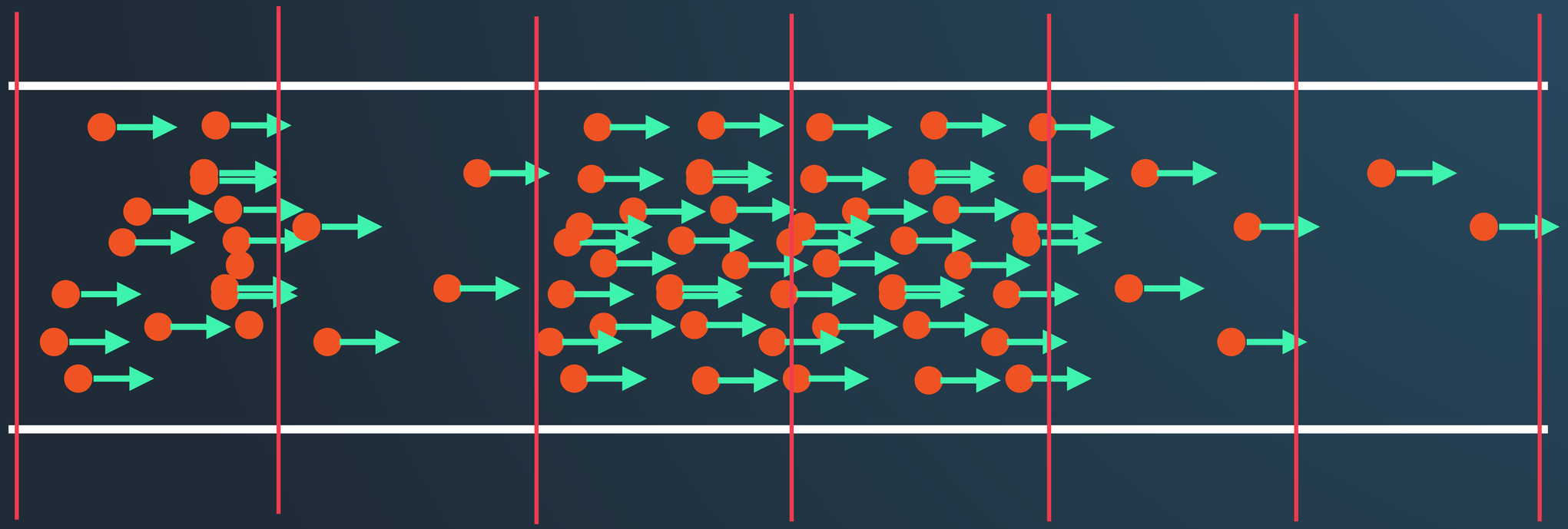
A higher voltage means more electrons are flowing in the current.



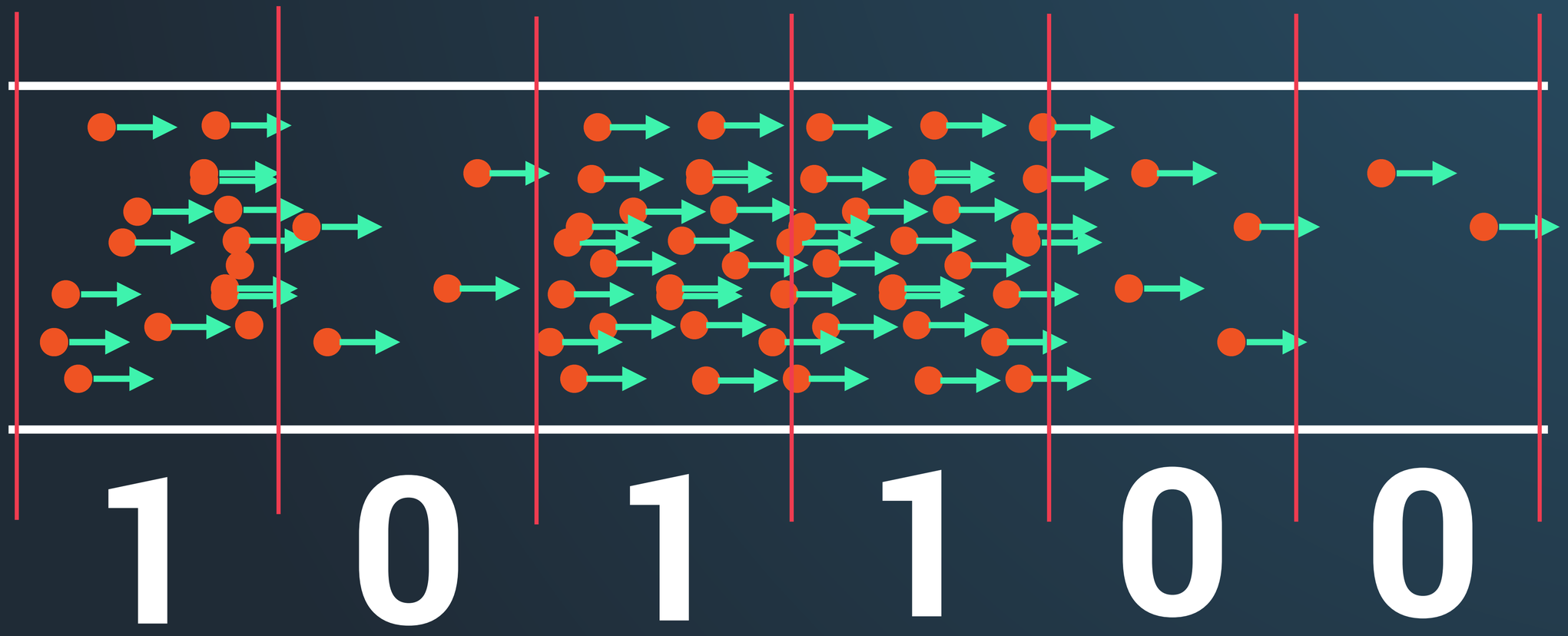
The voltage can be made to vary over time.



A measurement of the voltage is taken regularly on a clock cycle.



By thresholding the voltages to discrete values, we can derive bits.



Numbers, text and other data can be translated into and out of bits.

01011010 = 90

0	1	0	1	1	0	1	0
2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0

$2^0 * 0 = 0$

$2^1 * 1 = 2$

$2^2 * 0 = 0$

$2^3 * 1 = 8$

$2^4 * 1 = 16$

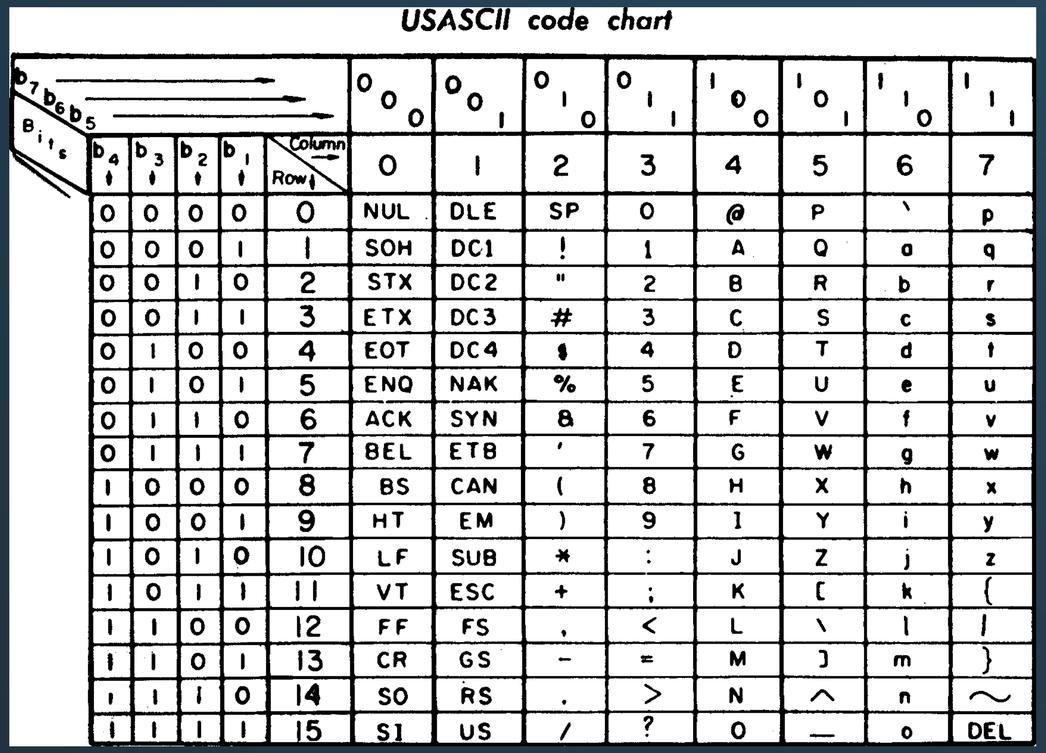
$2^5 * 0 = 0$

$2^6 * 1 = 64$

$2^7 * 0 = 0$

\+ 90

01011010 = “Z”

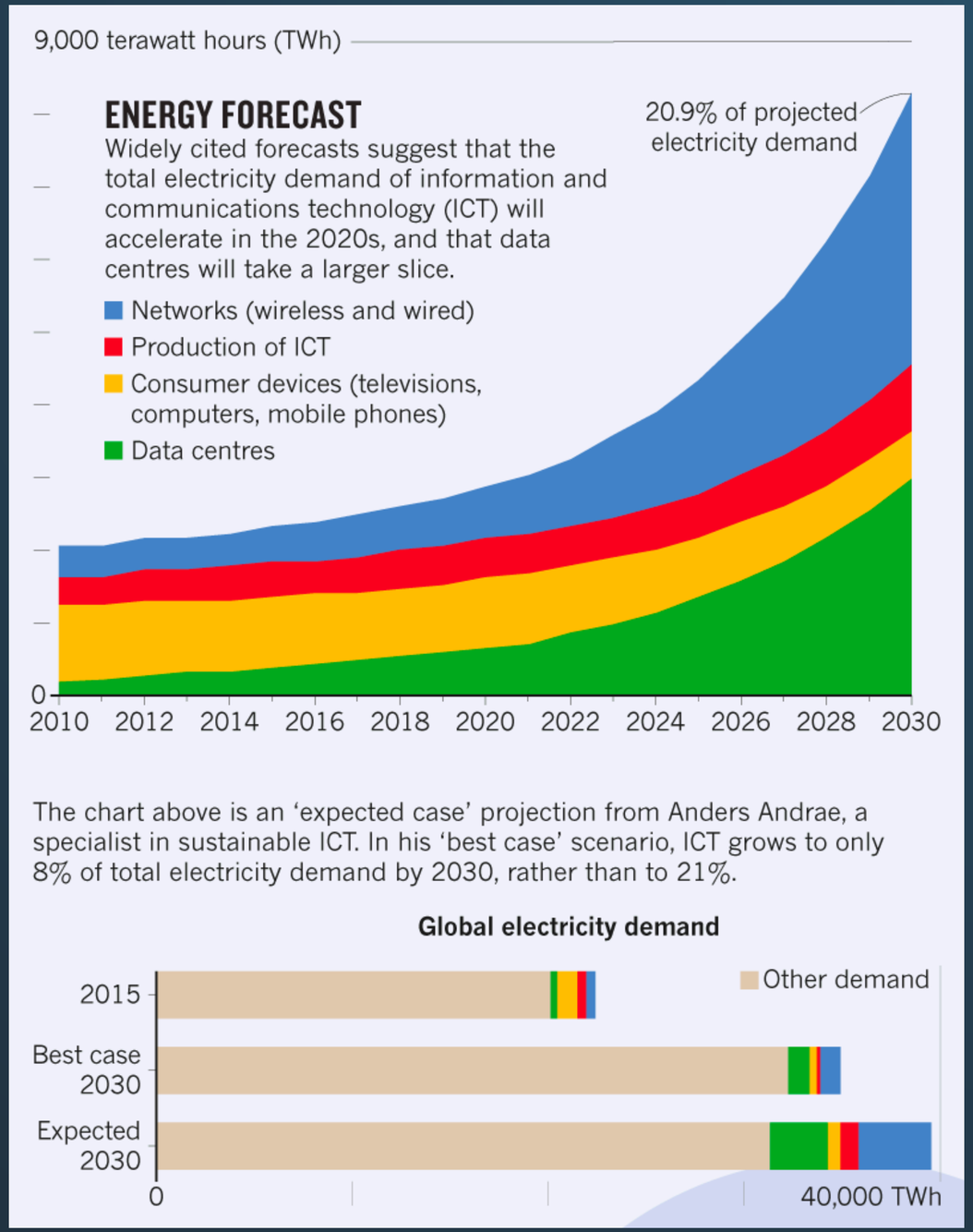


USASCII code chart

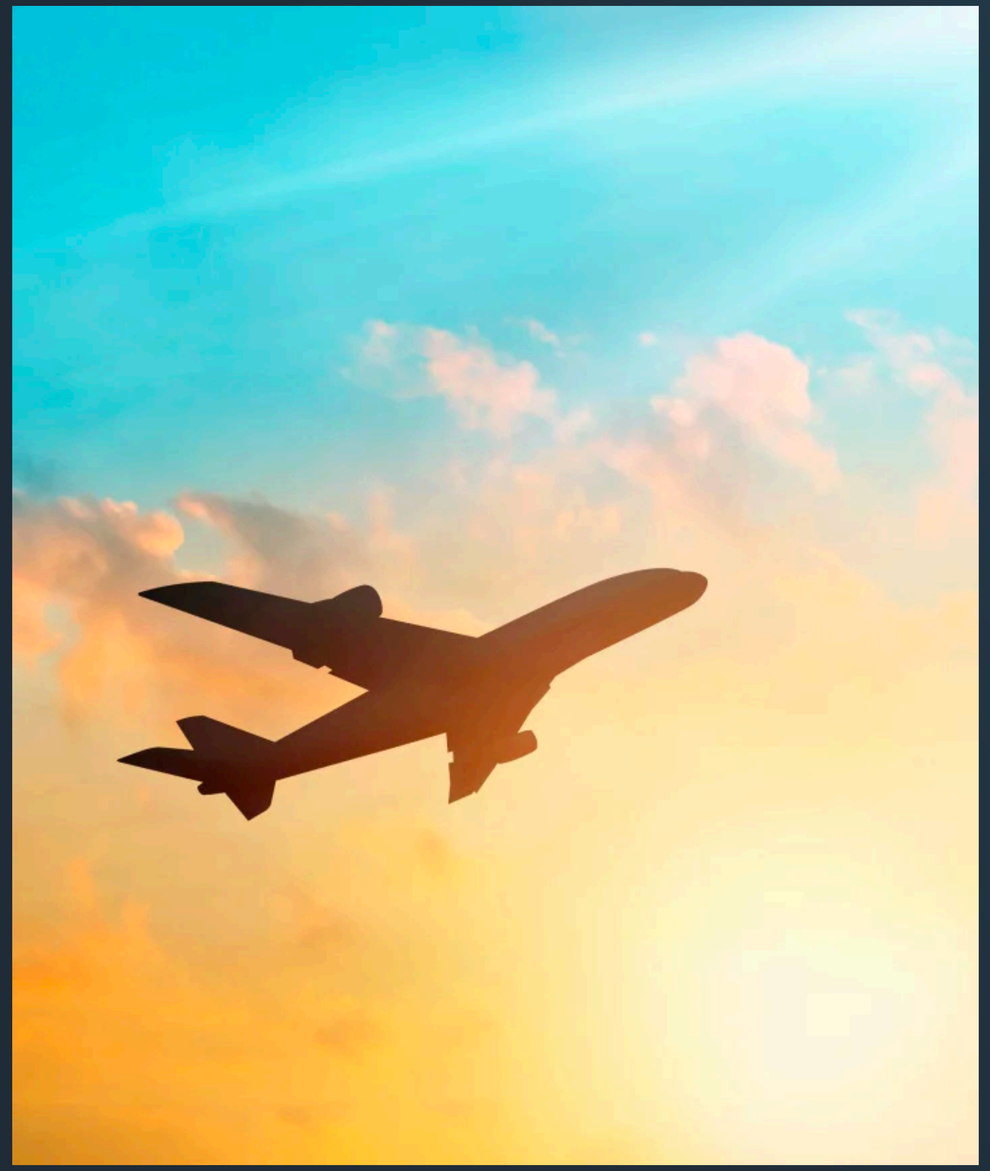
b7 b6 b5 / Bits	0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
b4 b3 b2 b1 / Row ↓ Column →	0	1	2	3	4	5	6	7
0 0 0 0	0	NUL	DLE	SP	0	@	P	`	p
0 0 0 1	1	SOH	DC1	!	1	A	Q	a	q
0 0 1 0	2	STX	DC2	"	2	B	R	b	r
0 0 1 1	3	ETX	DC3	#	3	C	S	c	s
0 1 0 0	4	EOT	DC4	$	4	D	T	d	t
0 1 0 1	5	ENQ	NAK	%	5	E	U	e	u
0 1 1 0	6	ACK	SYN	&	6	F	V	f	v
0 1 1 1	7	BEL	ETB	'	7	G	W	g	w
1 0 0 0	8	BS	CAN	(	8	H	X	h	x
1 0 0 1	9	HT	EM	)	9	I	Y	i	y
1 0 1 0	10	LF	SUB	*	:	J	Z	j	z
1 0 1 1	11	VT	ESC	+	;	K	[	k	{
1 1 0 0	12	FF	FS	,	<	L	\	l	|
1 1 0 1	13	CR	GS	-	=	M	]	m	}
1 1 1 0	14	SO	RS	.	>	N	^	n	~
1 1 1 1	15	SI	US	/	?	O	_	o	DEL

Over 2 TRILLION GIGABYTES of data will be transmitted over the Internet in 2020.

It will take a lot of electricity to transit that much data.



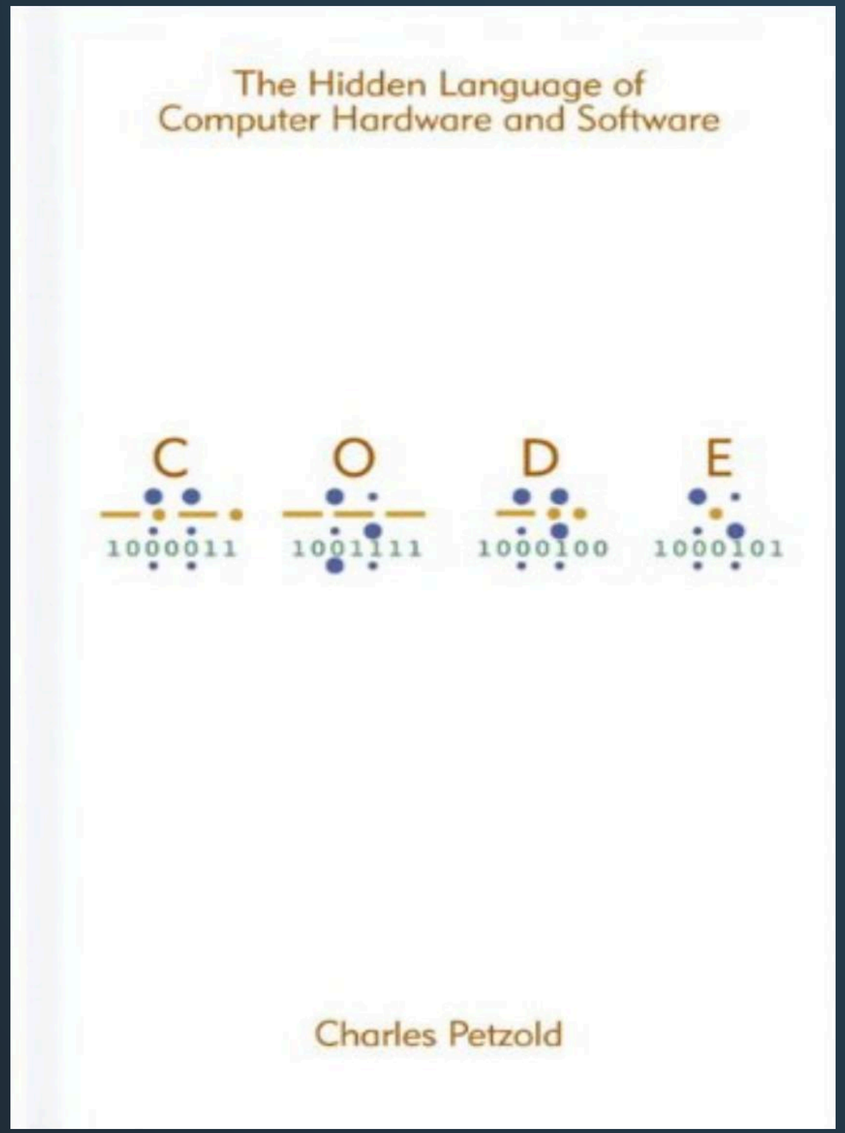
https://www.nature.com/articles/d41586-018-06610-y



Carbon emissions due to the Internet surpassed those of the airline industry in 2015.

https://www.theguardian.com/environment/2015/sep/25/server-data-centre-emissions-air-travel-web-google-facebook-greenhouse-gas

Do your part. Use fewer bytes.


The Hidden Language of
Computer Hardware and Software
C
O
D
E
1000011
1001111
1000100
1000101
Charles Petzold

Thank you.

Chris Zacharias

chris@imgix.com

@zacman85

imgIX