Amazon Web Services:
Deep analysis of the organization, data centers, global network and technology

Kenneth White, MEd
Principal Scientist

Social & Scientific Systems Inc
and
Cartesian Labs LLC

Who's the professor dude?

- My career has been spent in clinical research.
- Clinical trials for new drugs, specifically
- I work at the intersection of medicine, genomics, pharmacology and complex information systems.

[queue up the *I-got-your-complexity-right-here* comments]

The drug industry?

So, you're just some corporate suit, capitalist pharma sellout?

The drug industry?

So, you're just some corporate suit, capitalist pharma sellout?

Um, no.

Problems we're working on

- New vaccines for HIV for pregnant women and treatments for infected infants & children
- Eliminating pandemic influenza (bird/swine H1N1 flu), antibiotic-resistant tuberculosis
- Developing "unprofitable" treatments for crucial diseases like malaria and dengue fever
- Pancreatic and breast cancer
- 95% of this work is government funded grant (non-profit) research

But what was that "complex" systems reference?

- Over the past 30 years, our team has grown to support the Operations Center for the largest clinical trial research network in the world
  - Over 500 field sites, most active in "resource limited" regions around the world
  - Presence in 50+ countries (Africa and SE Asia w/ fastest growth)
  - Mange 1,500+ subcontractors, 18 different currencies
  - Over 2,000 protocols (studies) under management, Phase I to Phase IV

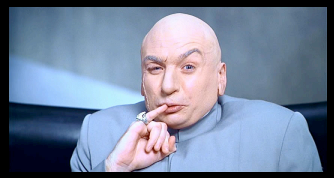
What does that mean *“Phase I to Phase IV”*?

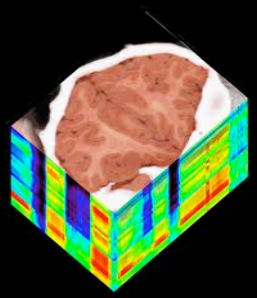
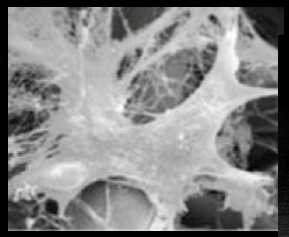
Clinical Drug Development 101

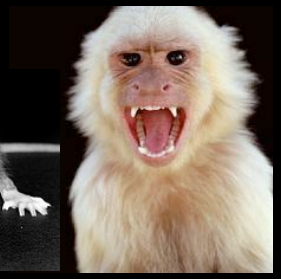
- Phase 0 ("Pre-clinical")
  1. An idea, or alternatively, pure Big Data autopilot
  2. Lit review: Is this compound likely to kill anyone?
  3. More computer models/simulations
  4. *In vitro* studies: Is compound toxic/beneficial to living tissue?
  5. *In vivo* studies: Is the compound toxic/beneficial to a real animal?
- Phase I-IV ("Clinical")
  1. *(Repeat step 5, but with people, just more rules)*

Put a different way... tl; dr:

- Phase 0





Idea | MapReduce → Lit Review → Computer Models → *In Vivo* (tissue prep) → **Successive** *In Vitro* → More Computer Models

- Phase I: First-in-human, safety, MTD, toxicity

- Phase II: Efficacy in patients

- Phase III: Large scale, multi-center, often intl.

And all of this matters because…?

Drug development and clinical trials are *highly* regulated

- Implementing the software and systems that capture and store clinical research data is:
  - a formal process
  - *routinely* audited for adherence to self-created engineering/development policies (often monthly by 3rd parties / "sponsors" / health authorities)
  - a great example of the Ellis' *Security Poverty Line*
  - created in the context of "validated computer systems" (FDA 21 CFR Part 11, Part 820, GCLP, ICH, etc)
  - increasingly (crucially) *risk-based*

- Think *chain-of-custody*

The lazy illustration: *One* System

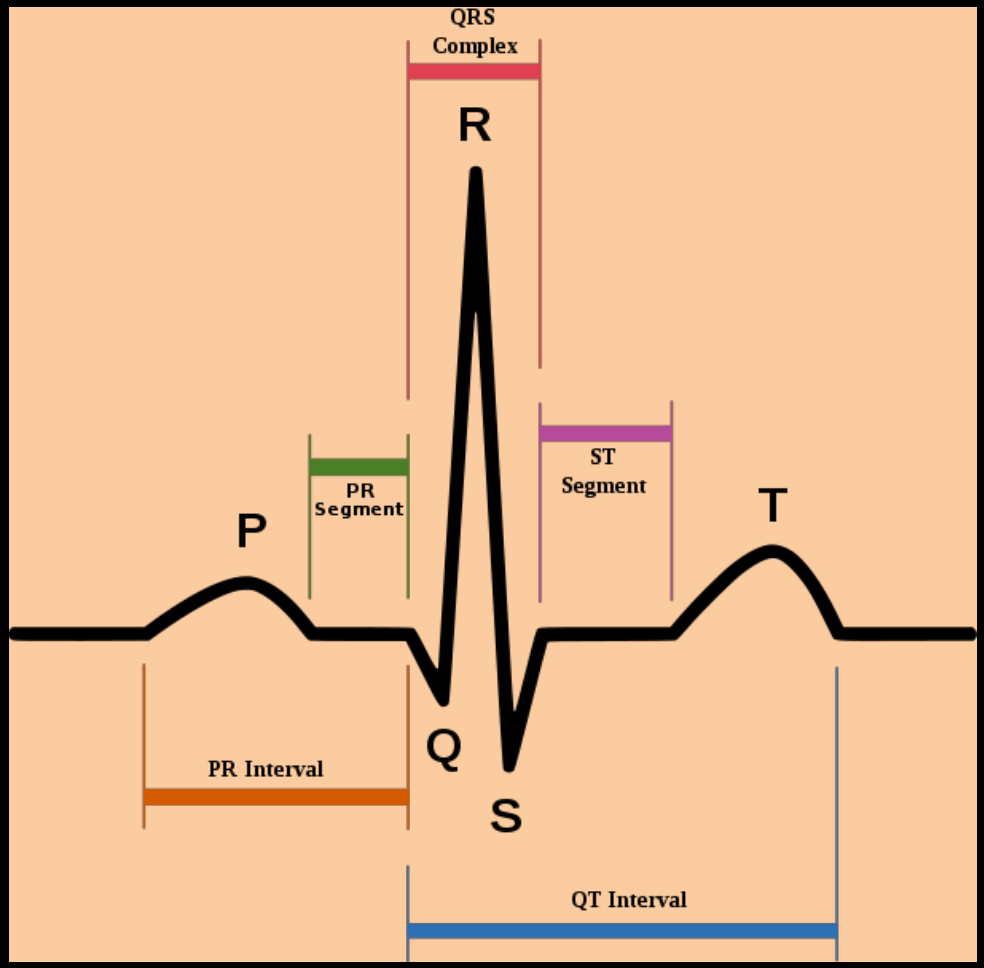
But this probably misses the more important point:

The regulatory burden is high precisely because the consequences here can be, literally, life and death.

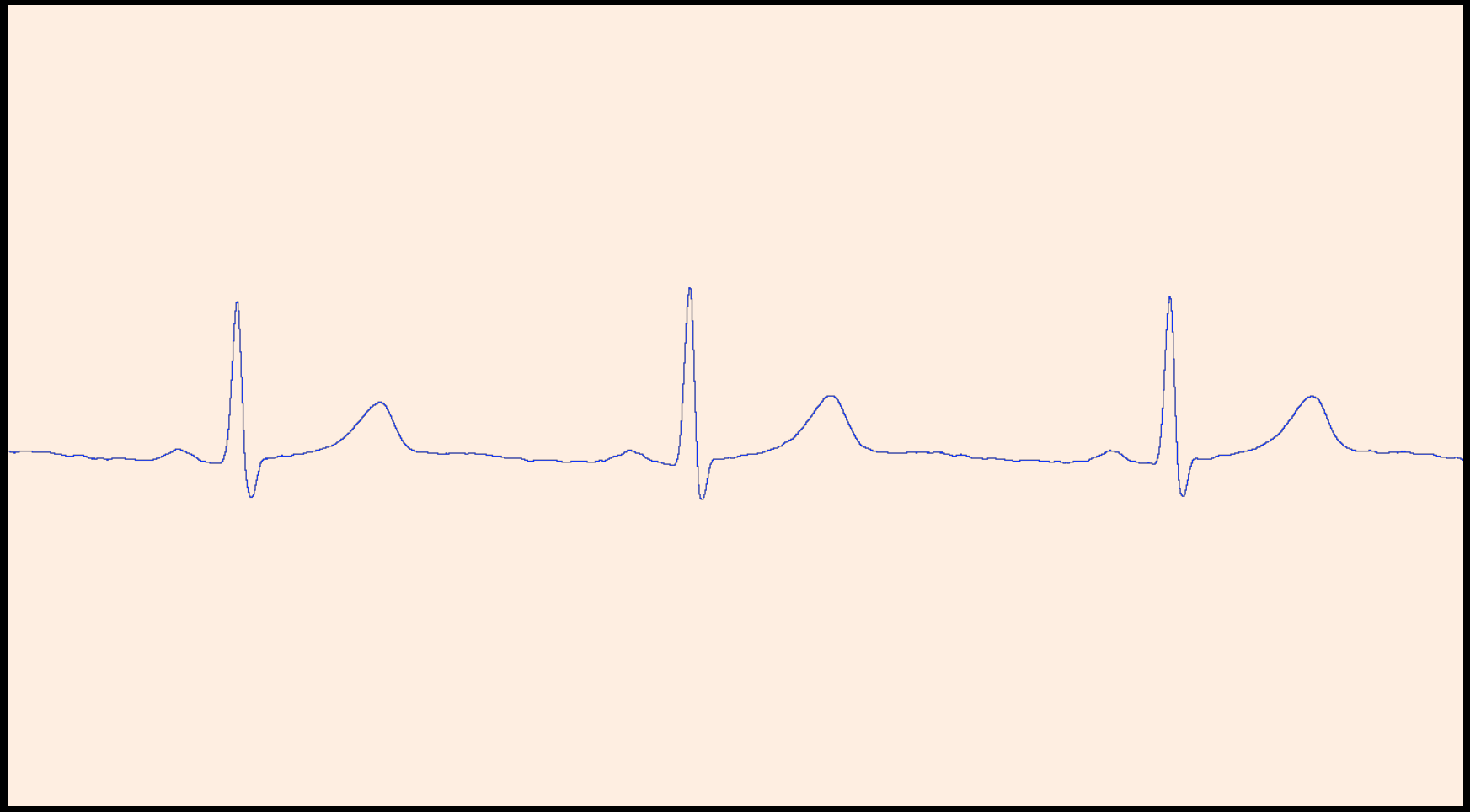
A better illustration

(aka, why science is hard)

EKG 101

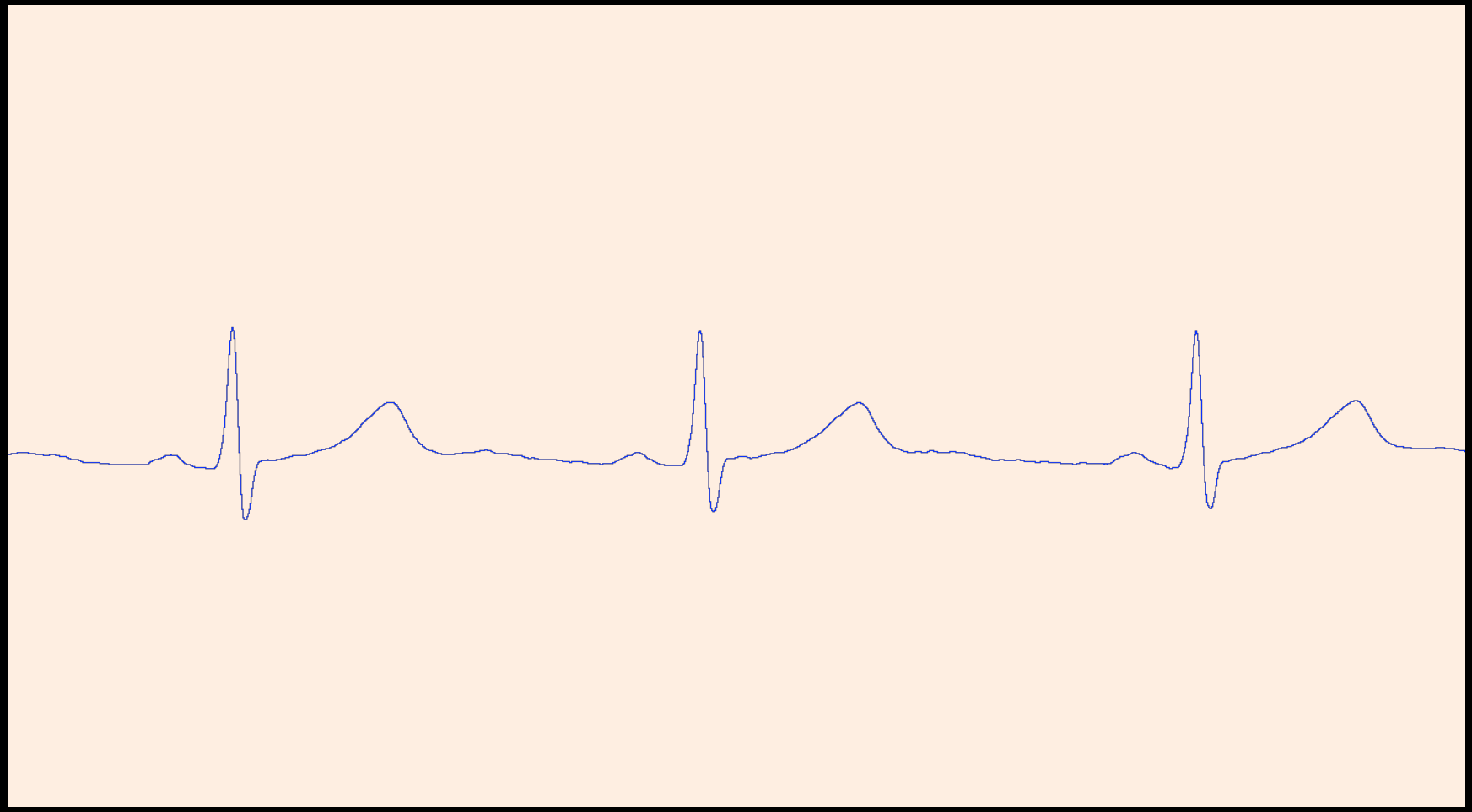

QRS
Complex
R
PR
Segment
ST
Segment
P
T
PR Interval
Q
S
QT Interval

Placebo, 1.5 hrs post-dose



Single precordial chest lead, telemetry-extracted ECG (V3 1000Hz)

Moxifloxacin, 1.5 hrs post-dose



Crtical changes are *highly* subtle to the eye

Original continuous rhythm data (moxi top, placebo bottom)

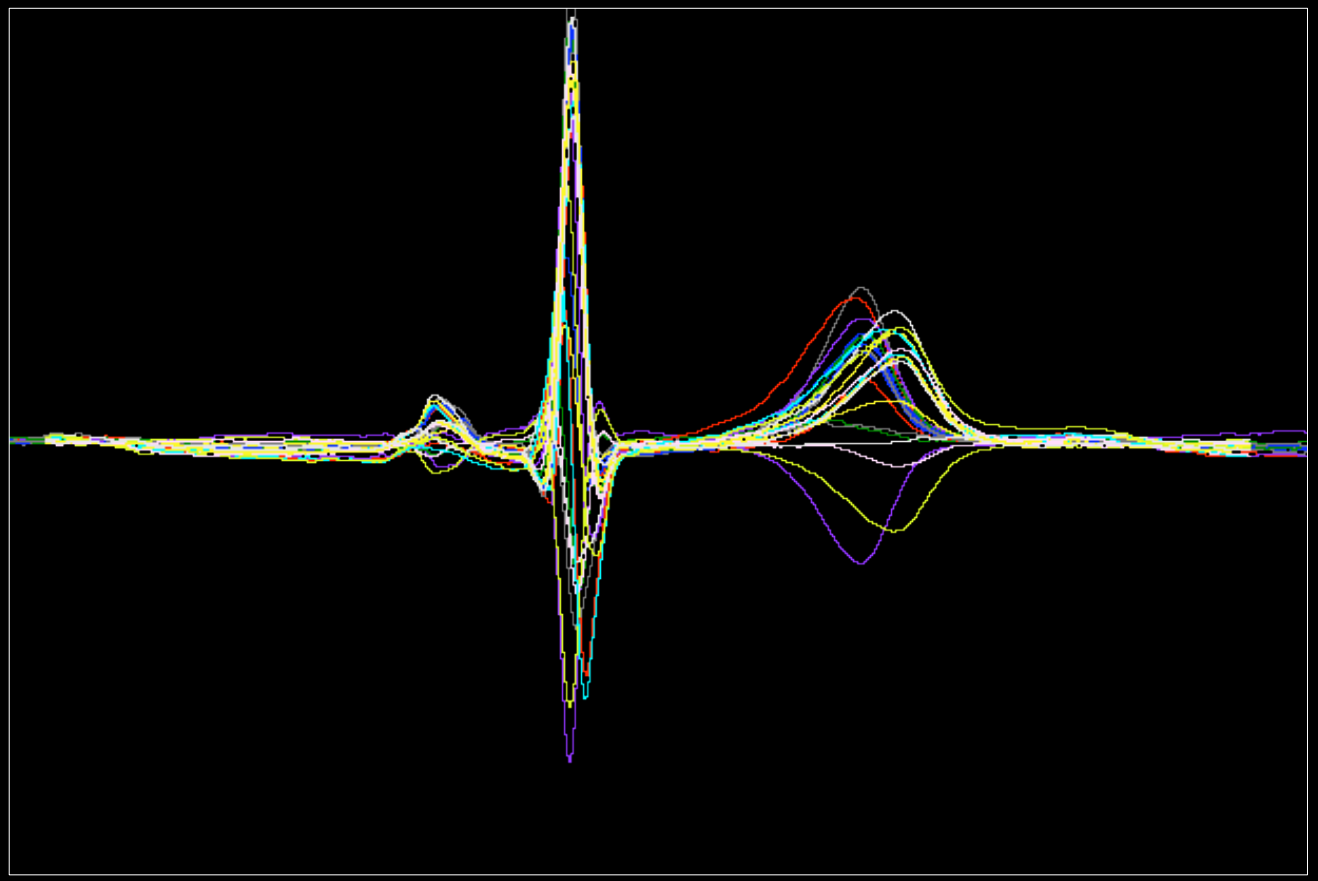
T-wave morphology & QT interval effect

Rate-corrected overlay (complex-by-complex) delta

T-wave morphology & QT interval effect

Single complex delta

T-wave morphology & QT interval effect



Single complex 12 lead "butterfly" deltas (Moxi arm shown as yellow gradient bands)

In other words

- This isn't simply a paperwork burden
- These systems aren't "just" databases
- Web apps are increasingly sophisticated and critical
- As R&D moves from animal models to human trials, the regulatory obligations are *orders of magnitude* more complex
- Rapid development is… not so rapid.

In a world where spinning up a new protocol/trial/system requires significant time, Quality Assurance and regulatory compliance, cloud-based technology has been a major disruption in how the science is executed.

Beware of the tortoise

Good turtles and bad turtles

- The Bad:
  - Institutional sloths
  - Traditional “Big Iron” IT (we have VMWare)
  - FUD
- The Good:
  - Rapid provision
  - Local / regional / global scale
  - Enlightenment: Really smart Quality people engaging in honest partnership
  - Configuration vs. customization
  - Nirvana: Pre-qualification!

The Promise of Amazon Web Services (AWS)

- Scale
- “Boring” part of infrastructure
- On-demand
  - Major cluetrain: “Hey, no one has logged in to these dev /qa/sandboxes in 2 months!”
- Pain reduction
- Zero procurement headaches
- Cost

The Promise of Amazon Web Services (AWS)

- Zero procurement headaches ←

Opportunity Cost is the "Killer Feature" of public cloud

- If you're comparing the cost of hard drives/ NAS/SAN in your server room to EC2, you're missing the point
- If you think VMs = cloud, you're missing the point
- If you really buy most of the "private cloud" vendor hype, you're *completely* missing the point
- Can you *really* fend off a large scale DDOS?
- Is that how you want to spend your time?

Home > Storage

News

Report: Iron Mountain to shutter cloud storage service

Iron Mountain to assist customers in migrating to other storage platforms

By Lucas Mearian

April 10, 2011 09:48 PM ET

2 Comments Like 16 +1 0

News

EMC shuts down online cloud storage service

Atmos offering could have pitted EMC against its service provider customers, analysts said

By Lucas Mearian

July 1, 2010 05:53 PM ET

So let’s look at AWS, shall we?

Strategic & Operations Leaders

- CEO: Jeff Bezos
  - CTO: Dr. Werner Vogels (Cornell, gossip net)
  - VP/Distinguished Engineer: James Hamilton
  - CISO: Stephen Schmidt (USDOJ/FBI Cyber, Sctn Chief)
    - VP Security, Dep. CISO: CJ Moses (USDOJ/FBI Cybercrime)
  - SVP AWS: Andy Jassy
    - VP Storage: Alyssa Henry (Microsoft)
    - GM (Elastic) MapReduce/Hadoop: Peter Sirota
    - HPC/Genomics/Sci Computing: Dr. Deepak Singh
    - Bioinformatics/Life Sciences: Dr. Matt Wood
    - Business Dev: Terry Wise
    - Solutions: Tom Stickle
    - Platform: Jacob Levanthon

Growth

- According to Feb 2012 SEC 8K filing, Amazon has hired 22,500 new employees YOY (67% increase)
- Werner Vogels (GovCloud Nov 2011):
  - *"Each day AWS adds the equivalent server capacity to power Amazon when it was a global, $2.76B enterprise (circa 2000)"*
- Currently advertising more than 800,000 public IP addresses

  https://forums.aws.amazon.com/forum.jspa?forumID=30

  http://huanliu.wordpress.com/2012/03/13/amazon-data-center-size/

Global IPs (May 2, 2012)

US East (Northern Virginia):

72.44.32.0/19 (72.44.32.0 - 72.44.63.255)
67.202.0.0/18 (67.202.0.0 - 67.202.63.255)
75.101.128.0/17 (75.101.128.0 - 75.101.255.255)
174.129.0.0/16 (174.129.0.0 - 174.129.255.255)
204.236.192.0/18 (204.236.192.0 - 204.236.255.255)
184.73.0.0/16 (184.73.0.0 – 184.73.255.255)
184.72.128.0/17 (184.72.128.0 - 184.72.255.255)
184.72.64.0/18 (184.72.64.0 - 184.72.127.255)
50.16.0.0/15 (50.16.0.0 - 50.17.255.255)
50.19.0.0/16 (50.19.0.0 - 50.19.255.255)
107.20.0.0/14 (107.20.0.0 - 107.23.255.255)
23.20.0.0/14 (23.20.0.0 – 23.23.255.255)
54.242.0.0/15 (54.242.0.0 – 54.243.255.255) NEW

US West (Oregon):

50.112.0.0/16 (50.112.0.0 - 50.112.255.255)
54.245.0.0/16 (54.245.0.0 – 54.245.255.255) NEW

US West (Northern California):

204.236.128.0/18 (204.236.128.0 - 204.236.191.255)
184.72.0.0/18 (184.72.0.0 – 184.72.63.255)
50.18.0.0/16 (50.18.0.0 - 50.18.255.255)
184.169.128.0/17 (184.169.128.0 - 184.169.255.255)
54.241.0.0/16 (54.241.0.0 – 54.241.255.255) NEW

Global IPs (cont.)

EU (Ireland):

79.125.0.0/17 (79.125.0.0 - 79.125.127.255)
46.51.128.0/18 (46.51.128.0 - 46.51.191.255)
46.51.192.0/20 (46.51.192.0 - 46.51.207.255)
46.137.0.0/17 (46.137.0.0 - 46.137.127.255)
46.137.128.0/18 (46.137.128.0 - 46.137.191.255)
176.34.128.0/17 (176.34.128.0 - 176.34.255.255)
176.34.64.0/18 (176.34.64.0 – 176.34.127.255)
54.247.0.0/16 (54.247.0.0 – 54.247.255.255)

Asia Pacific (Singapore)

175.41.128.0/18 (175.41.128.0 - 175.41.191.255)
122.248.192.0/18 (122.248.192.0 - 122.248.255.255)
46.137.192.0/18 (46.137.192.0 - 46.137.255.255)
46.51.216.0/21 (46.51.216.0 - 46.51.223.255)
54.251.0.0/16 (54.251.0.0 – 54.251.255.255)

Asia Pacific (Tokyo)

175.41.192.0/18 (175.41.192.0 - 175.41.255.255)
46.51.224.0/19 (46.51.224.0 - 46.51.255.255)
176.32.64.0/19 (176.32.64.0 - 176.32.95.255)
103.4.8.0/21 (103.4.8.0 - 103.4.15.255)
176.34.0.0/18 (176.34.0.0 - 176.34.63.255)
54.248.0.0/15 (54.248.0.0 - 54.249.255.255)

Global IPs (cont.)

US West (Northern California):

204.236.128.0/18 (204.236.128.0 - 204.236.191.255)
184.72.0.0/18 (184.72.0.0 – 184.72.63.255)
50.18.0.0/16 (50.18.0.0 - 50.18.255.255)
184.169.128.0/17 (184.169.128.0 - 184.169.255.255)
54.241.0.0/16 (54.241.0.0 – 54.241.255.255) NEW

EU (Ireland):

79.125.0.0/17 (79.125.0.0 - 79.125.127.255)
46.51.128.0/18 (46.51.128.0 - 46.51.191.255)
46.51.192.0/20 (46.51.192.0 - 46.51.207.255)
46.137.0.0/17 (46.137.0.0 - 46.137.127.255)
46.137.128.0/18 (46.137.128.0 - 46.137.191.255)
176.34.128.0/17 (176.34.128.0 - 176.34.255.255)
176.34.64.0/18 (176.34.64.0 – 176.34.127.255)
54.247.0.0/16 (54.247.0.0 – 54.247.255.255)

Asia Pacific (Singapore)

175.41.128.0/18 (175.41.128.0 - 175.41.191.255)
122.248.192.0/18 (122.248.192.0 - 122.248.255.255)
46.137.192.0/18 (46.137.192.0 - 46.137.255.255)
46.51.216.0/21 (46.51.216.0 - 46.51.223.255)
54.251.0.0/16 (54.251.0.0 – 54.251.255.255)

Asia Pacific (Tokyo)

175.41.192.0/18 (175.41.192.0 - 175.41.255.255)
46.51.224.0/19 (46.51.224.0 - 46.51.255.255)
176.32.64.0/19 (176.32.64.0 - 176.32.95.255)
103.4.8.0/21 (103.4.8.0 - 103.4.15.255)
176.34.0.0/18 (176.34.0.0 - 176.34.63.255)
54.248.0.0/15 (54.248.0.0 - 54.249.255.255)

South America (Sao Paulo)

177.71.128.0/17 (177.71.128.0 - 177.71.255.255)

Data Center Infrastructure

- Facility Build-outs: The Silent Construction Partner
- Availability Zones & Regions
- Data center locations – with pictures
- Power Capacity & Redundancy
- Backup Generators
- Acronym Soup: SAE16, SOC1, FISMA, ISO27001, Part11
- Physical, process & technical controls

Data Center Infrastructure

- Facility Build-outs: The Silent Construction Partner
  - One of the most powerful companies you've never heard of: Digital Reality Trust

Data Center Infrastructure

- Availability Zones & Regions

Direct Connect Physical Ingress

AWS Direct Connect Locations

AWS Direct Connect is available at six locations around the world. The table below shows connectivity options to different AWS Regions.

AWS Direct Connect Location	AWS Region
CoreSite One Wilshire	US West (Northern California)
Equinix DC1-DC6	US East (Virginia)
Equinix SV1 & SV5	US West (Northern California)
Equinix SG2	Asia Pacific (Singapore)
Equinix TY2	Asia Pacific (Tokyo)
TelecityGroup, London Docklands'	EU (Ireland)

Additional AWS Direct Connect locations are planned worldwide in 2012.

Data Center Infrastructure

- Facility Build-outs: The Silent Construction Partner
- Availability Zones & Regions
- Data center locations – with pictures
- Power Capacity & Redundancy
- Backup Generators
- Acronym Soup: SAE16, SOC1, FISMA, ISO27001, Part11
- Physical, process & technical controls

Data Center Infrastructure

- Data center locations
  - Where *are* these guys?
- Interesting details:
  - 30 servers per rack
  - $1,450 per server
  - AWS Uses Low-Density Containers To Cut Cooling
  - Perdix
  - 1.5 to 20MW

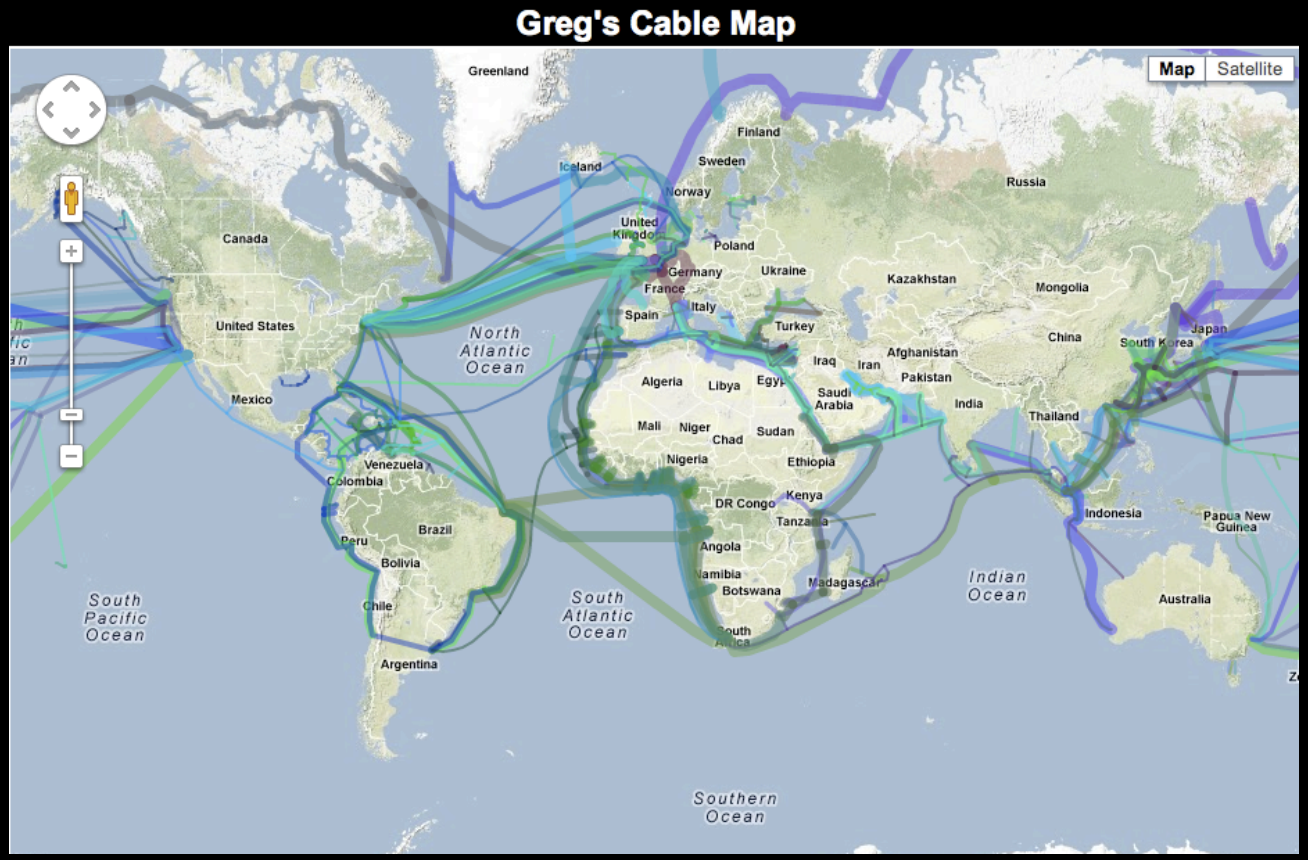
Welcome to EC2 East

DATA CENTER ADDRESS

43915 Devin Shafron Dr
Ashburn, VA 20147
United States

Data Center Infrastructure

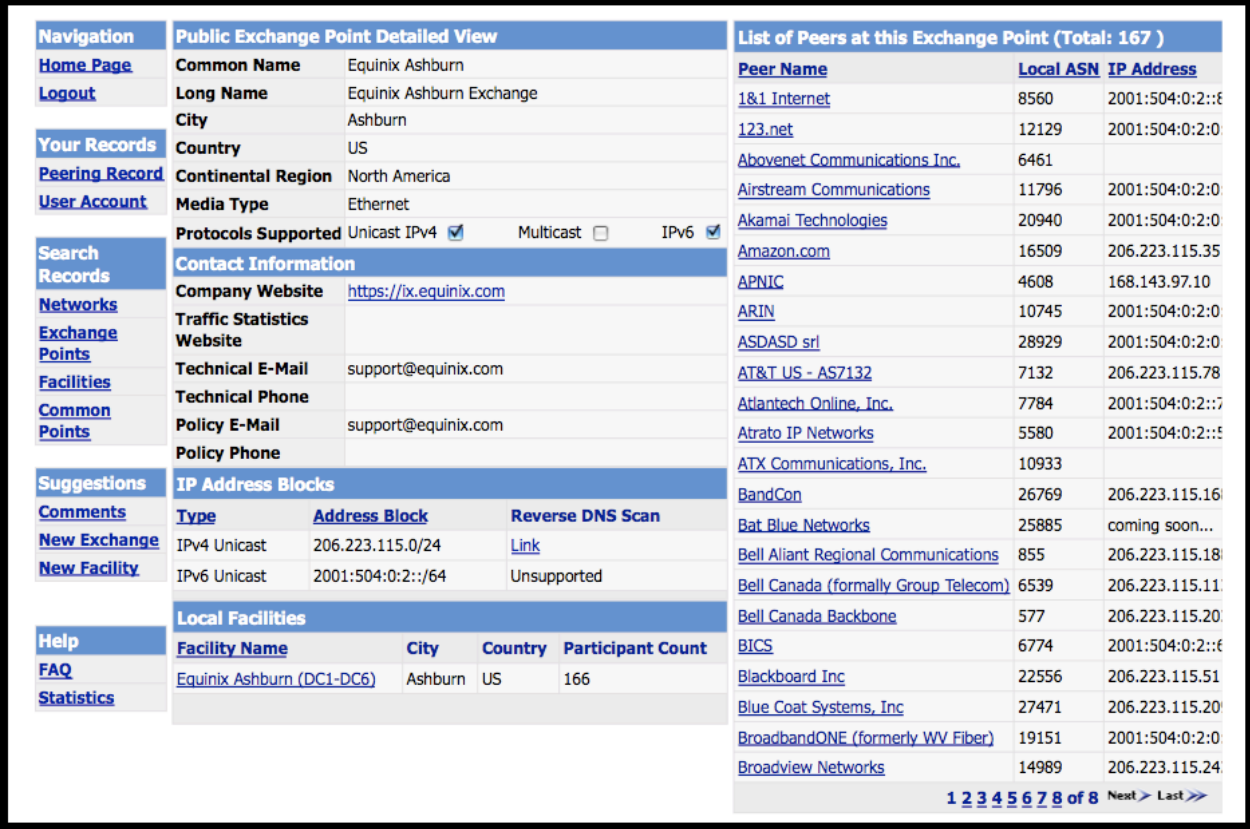
- Global network



www.cablemap.info

Data Center Infrastructure

- Public & Private Peering: The bits are out there, you just have to look



Navigation
Home Page
Logout

Your Records
Peering Record
User Account

Search Records
Networks
Exchange Points
Facilities
Common Points

Suggestions
Comments
New Exchange
New Facility

Help
FAQ
Statistics

Public Exchange Point Detailed View

Common Name	Equinix Ashburn
Long Name	Equinix Ashburn Exchange
City	Ashburn
Country	US
Continental Region	North America
Media Type	Ethernet
Protocols Supported	Unicast IPv4 ☑ Multicast ☐ IPv6 ☑

Contact Information

Company Website	https://ix.equinix.com
Traffic Statistics Website	
Technical E-Mail	support@equinix.com
Technical Phone	
Policy E-Mail	support@equinix.com
Policy Phone	

IP Address Blocks

Type	Address Block	Reverse DNS Scan
IPv4 Unicast	206.223.115.0/24	Link
IPv6 Unicast	2001:504:0:2::/64	Unsupported

Local Facilities

Facility Name	City	Country	Participant Count
Equinix Ashburn (DC1-DC6)	Ashburn	US	166

List of Peers at this Exchange Point (Total: 167)

Peer Name	Local ASN	IP Address
1&1 Internet	8560	2001:504:0:2::8
123.net	12129	2001:504:0:2:0
Abovenet Communications Inc.	6461	
Airstream Communications	11796	2001:504:0:2:0
Akamai Technologies	20940	2001:504:0:2:0
Amazon.com	16509	206.223.115.35
APNIC	4608	168.143.97.10
ARIN	10745	2001:504:0:2:0
ASDASD srl	28929	2001:504:0:2:0
AT&T US - AS7132	7132	206.223.115.78
Atlantech Online, Inc.	7784	2001:504:0:2::7
Atrato IP Networks	5580	2001:504:0:2::5
ATX Communications, Inc.	10933	
BandCon	26769	206.223.115.16
Bat Blue Networks	25885	coming soon...
Bell Aliant Regional Communications	855	206.223.115.18
Bell Canada (formally Group Telecom)	6539	206.223.115.11
Bell Canada Backbone	577	206.223.115.20
BICS	6774	2001:504:0:2::6
Blackboard Inc	22556	206.223.115.51
Blue Coat Systems, Inc	27471	206.223.115.20
BroadbandONE (formerly WV Fiber)	19151	2001:504:0:2:0
Broadview Networks	14989	206.223.115.24

1 2 3 4 5 6 7 8 of 8 Next > Last >>

Amazon Web Services: Risk and Compliance

May 2011

(Please consult http://aws.amazon.com/security **for the latest version of this paper)**