

REGIONAL PERSPECTIVE & ION/DYNES CAPABILITIES FOR RESEARCHERS IN DELAWARE

Gregory D. Palmer
MAGPI Executive Director
University of Pennsylvania



Data Volumes Exploding

"Every two days now we create as much information as we did from the dawn of civilization up until 2003."

-Eric Schmidt, CEO Google

Examples of Data Volumes

Unit	Value	Example
Kilobytes (KB)	1,000 bytes	a paragraph of a text document
Megabytes (MB)	1,000 Kilobytes	a small novel
Gigabytes (GB)	1,000 Megabytes	Beethoven's 5th Symphony
Terabytes (TB)	1,000 Gigabytes	all the X-rays in a large hospital
Petabytes (PB)	1,000 Terabytes	half the contents of all US academic research libraries (the ASDC holds roughly this much data)
Exabytes (EB)	1,000 Petabytes	about one fifth of the words people have ever spoken
Zettabytes (ZB)	1,000 Exabytes	as much information as there are grains of sand on all the world's beaches
Yottabytes (YB)	1,000 Zettabytes	as much information as there are atoms in 7,000 human bodies



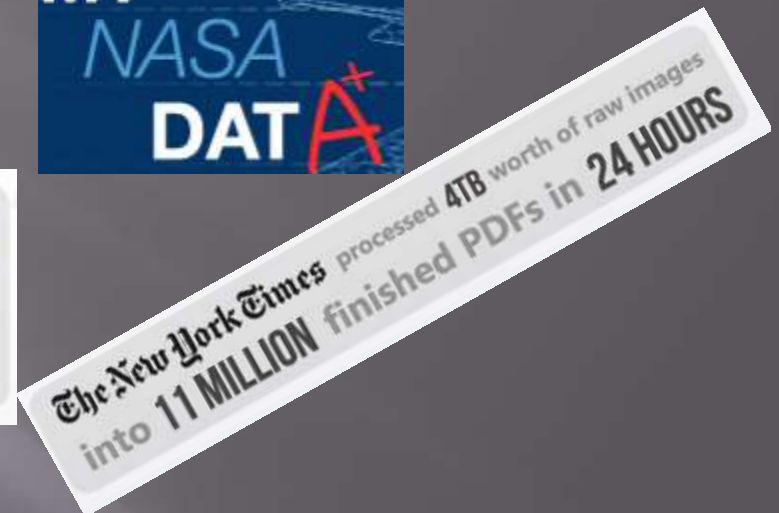
What Do you Mean by Data?

- ▣ It's personal stuff!
 - My pictures
 - My tax returns
 - My contacts list
 - My Social Media
 - My Government

2.7 Zetabytes of data exist in the digital universe today.

235 Terabytes of data was collected by the U.S. Library of Congress in April 2011

Thanks to
Wikibon

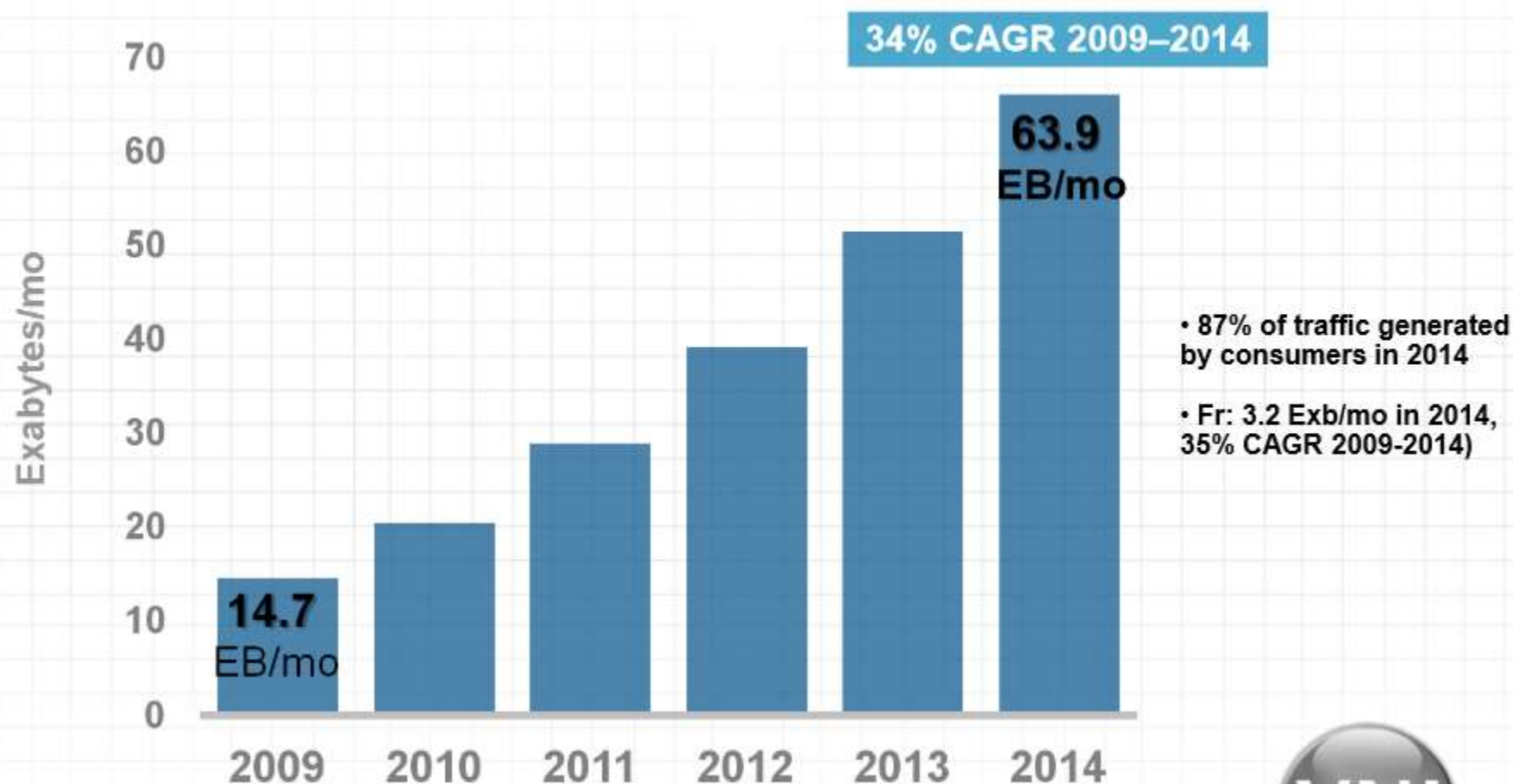


Global IP Traffic Growth

Global IP Traffic

IP traffic will increase 4.3 fold from 2009–2014

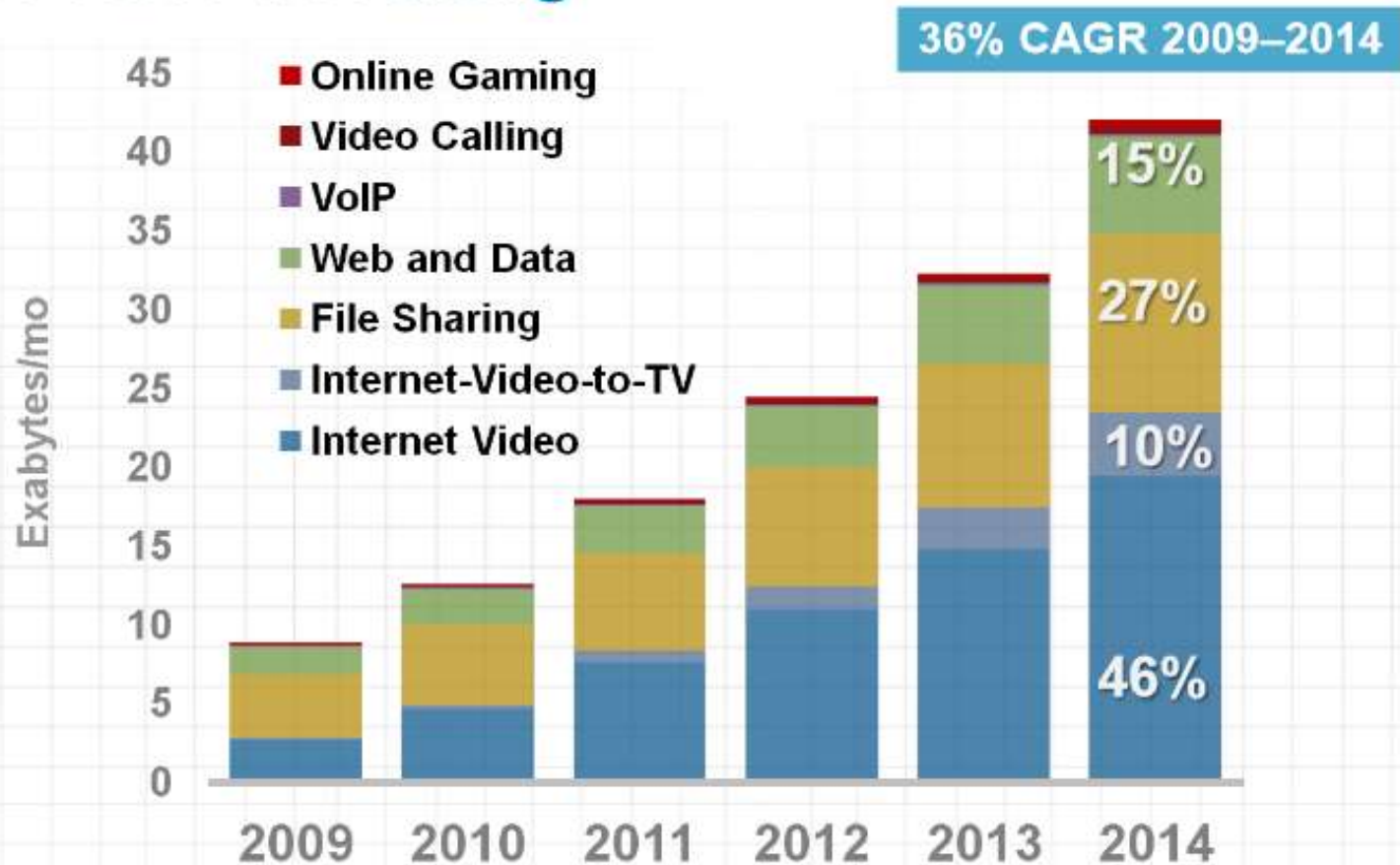
In 2014, global IP traffic will reach 3/4 of a zettabyte



Source: Cisco Visual Networking Index (VNI) Global Forecast, 2009–2014

Global Consumer Internet Traffic Growth:

Video to exceed 91% of global traffic by 2014; 56% will be video streaming



October 26, 2010

Netflix Delivers One-Two Punch to Video/Data Providers

There was one tidbit in Sandvine's recently released "Fall 2010 Global Internet Phenomena Report" that caused some double takes: Netflix represents more than 20 percent of downstream traffic during peak hours on fixed-access networks. The heaviest usage is from 8 p.m. to 10 p.m.

What do you mean by Data?

- ▣ It's Scientific Stuff!
- ▣ Sloan Digital Sky Survey - 2000, amassed more in its first few weeks than all data collected in the history of astronomy.
- ▣ LHC - Processes about one petabyte of data every day The center hosts 10,000 servers with 90,000 processor cores. Some 6000 changes in the database are performed every second.
- ▣ And now...Bioinformatics





How Big is a Sequencing Project

Whole Genome Samples	Exome Samples	Space
0	200	1.6 TB
0	1000	8.0 TB
100	0	1.5 TB
1000	0	15 TB
100	1000	23 TB

Methods of Moving Research Data

- ▣ You can get someone to carry it to your collaborator.
- ▣ Estimated time across campus: 1 hour, 15 minutes
- ▣ Requirements: 1 post-doc, (with muscles) Maybe two...



Methods of Moving Research Data

- ▣ By Ground
- ▣ Shipping Hard Drives
 - ▣ 3 days, weather permitting
- ▣ Analysis: 3 days
- ▣ Response: 3 days
- ▣ OOPS! One drive damaged
- ▣ Start Over



OR



Methods of Moving Research data

- ▣ By Air
- ▣ Shipping Hard Drives
 - Faster, but more expensive
 - Analysis feedback still delayed
 - 2 days each way



Methods of Moving Research Data

- ▣ By Commodity Internet
 - 1 TB ~ 12 hours
 - 2 TB ~ 1 day
 - 10 TB ~ 7 days
- ▣ Corrupt Data?
 - Start Over!
- ▣ Retransmits?
 - Start Over!

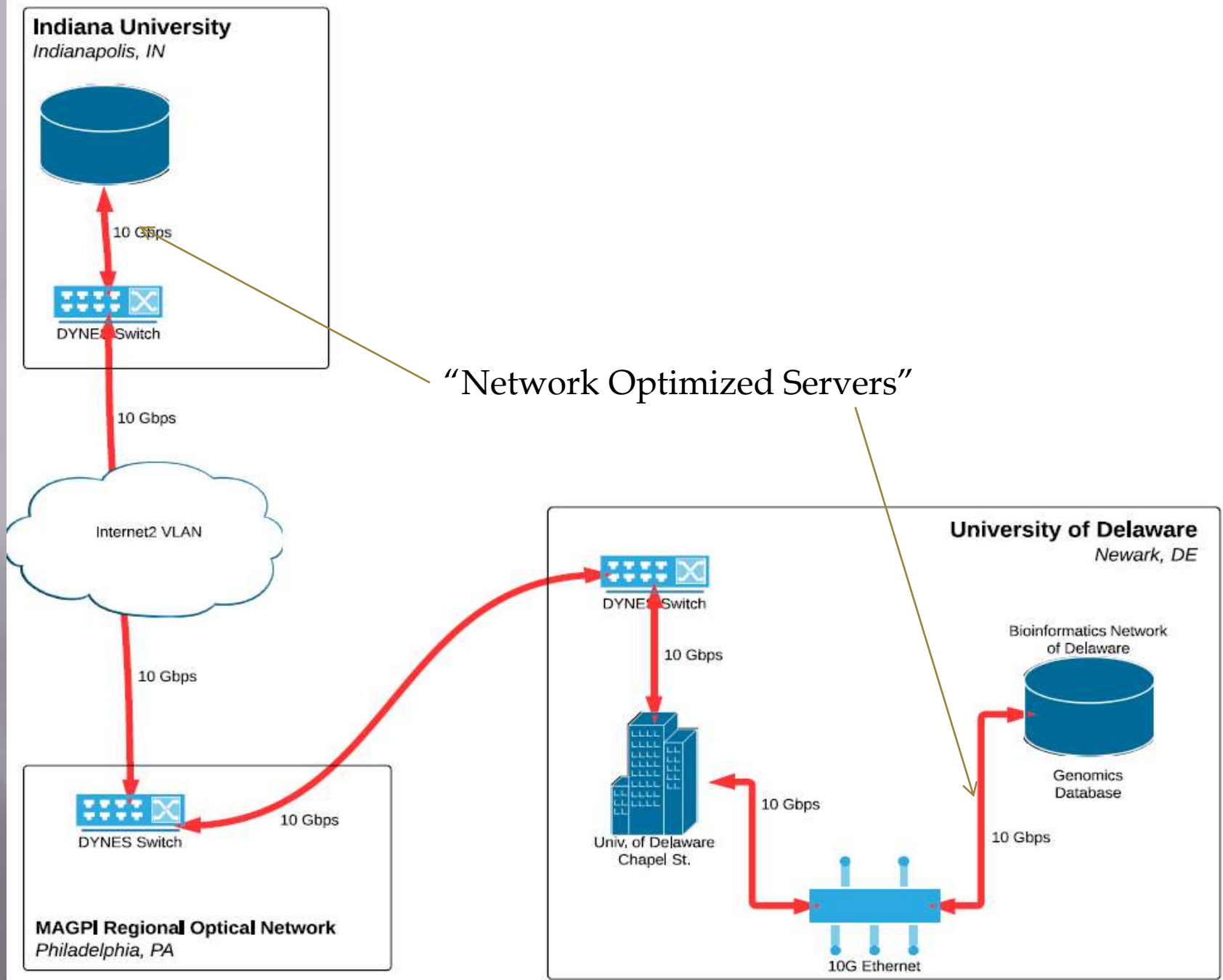


Advanced Networks

- ▣ Thank you Internet2!
 - 8.8 Terabytes per second of total capacity
 - Layer 2 services
 - ▣ Software Defined Networking & Openflow
 - ▣ VLAN Tagging
 - Layer 3 services
 - ▣ IP routing to 70,000 connected sites in the U.S.
 - ▣ 84 international peering relationships
- ▣ The University of Delaware – Member since 2000

UDel/Indiana University Pilot Program

- ▣ Utilizing the Internet2 Advanced Layer 2 Service
- ▣ Via the Inter-Operable Network Platform, (ION)
- ▣ With “Network Optimized Servers”
 - Developed by Dr. Martin Swamy, Indiana Univ.
- ▣ Uses VLAN tagging between sites
- ▣ Tested at line speed in the lab, (10 Gbps)
- ▣ Get the baseline via traditional IP
- ▣ Then test with the optimized servers.



Simple Questions, Please!