

Совет по ИТ ПИЯФ
9 февраля 2016 года

Взгляд в будущее ИТ инфраструктуры

в 2016 году и далее

А. Е. Шевель

План презентации

- Исторические даты
- Мировые тенденции
 - Большие данные
 - Конфигурируемые архитектуры
- Наша реальность и перспективы
 - Облачное хранение данных + комьютинг
 - Локальная проводная сеть и WiFi сеть
 - Электронная почта

Октябрь 1995

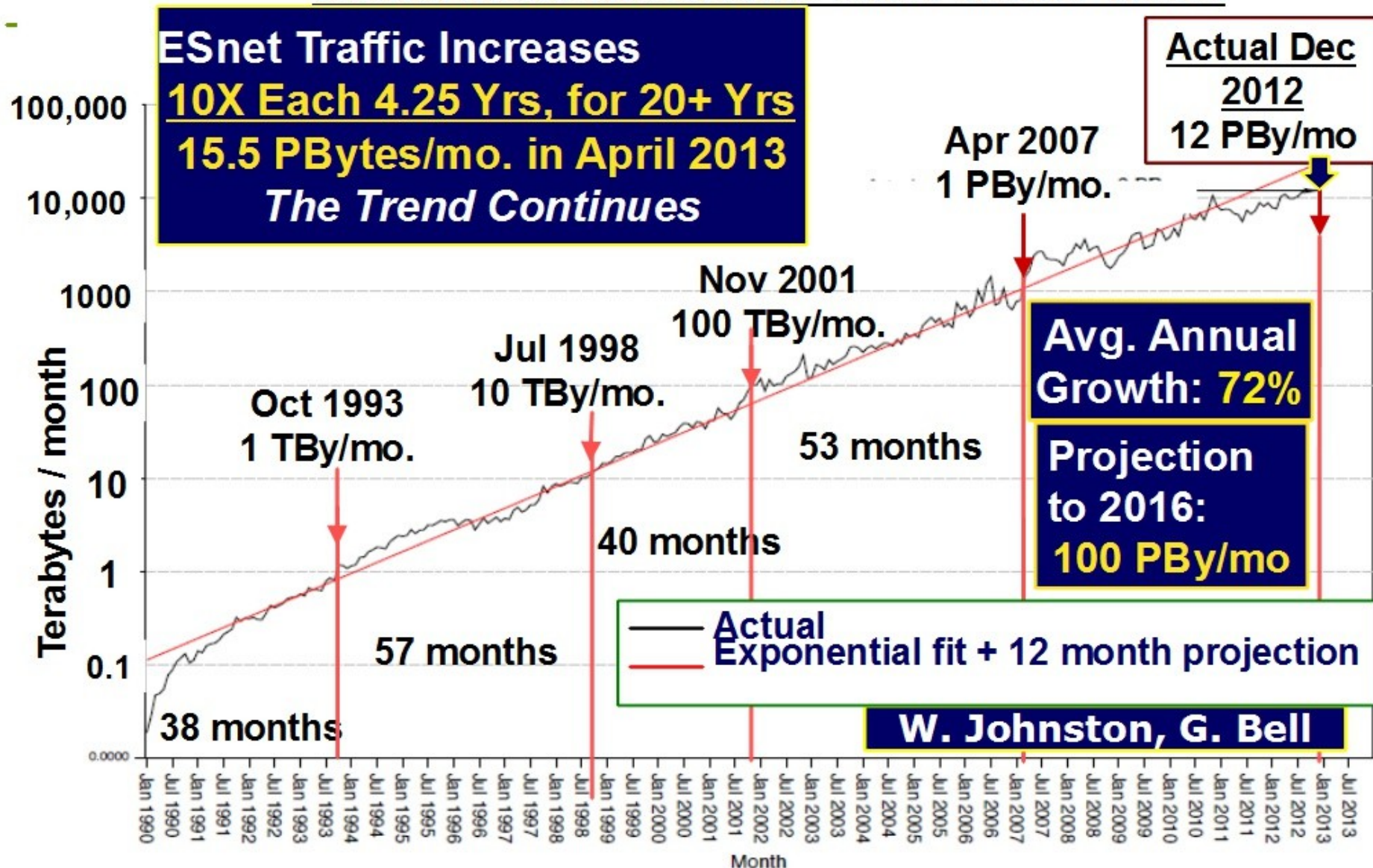




Исторические даты компьютеризации и коммуникации

- **20 лет** со времени старта публичного в ОФВЭ и ПИЯФ доступа в Интернет через спутниковый канал (старт = октябрь 1995, 64 Kbit)
 - Рост за 20 лет в 16 раз (во многих других местах рост более, чем в тысячу раз)

Remarkable Historical ESnet Traffic Trend

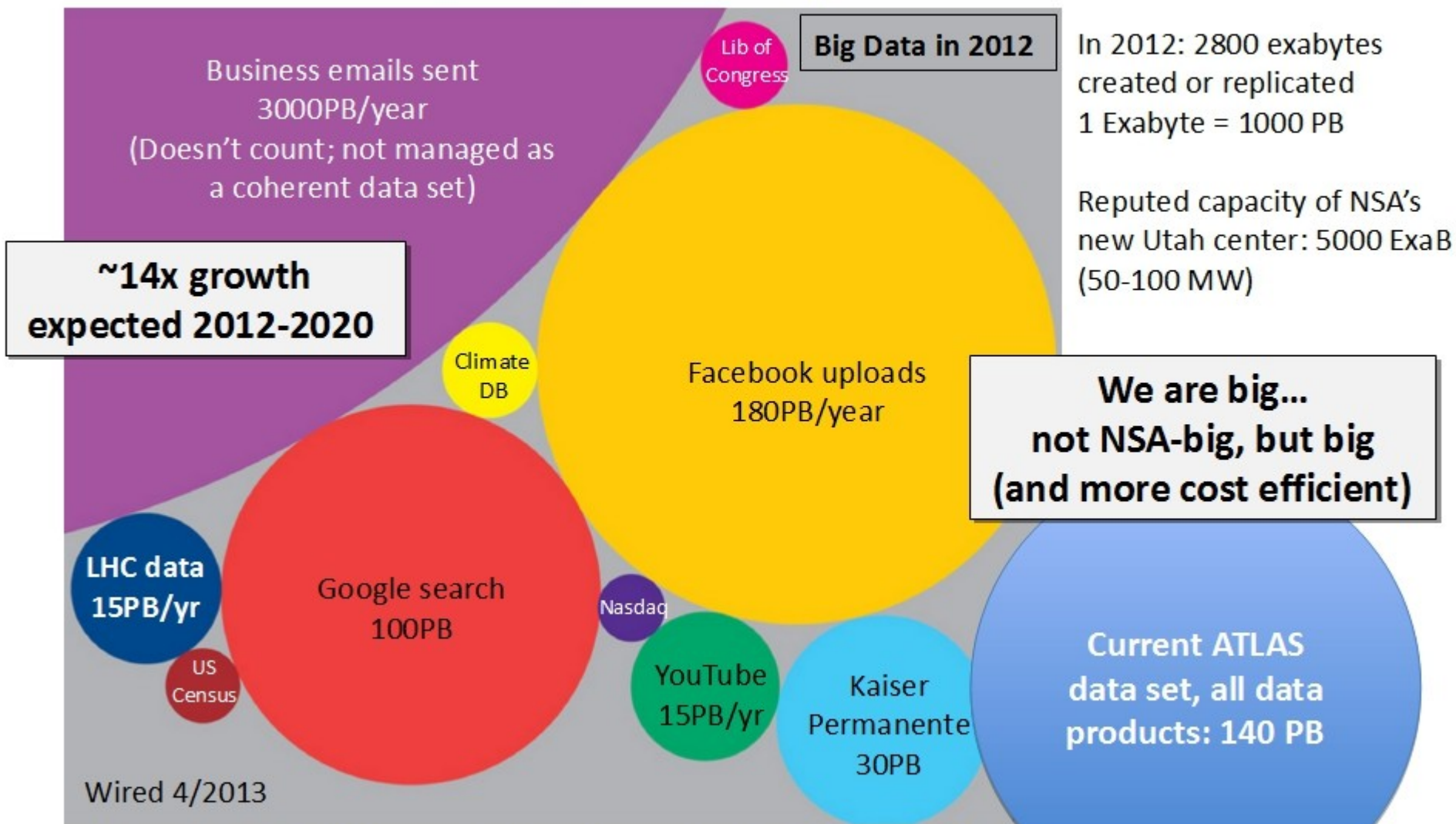


Log Plot of ESnet Monthly Accepted Traffic, January 1990 – December 2012

Продуктивность научных групп в значительной степени обусловлена удобством и скоростью обмена данными как внутри организации, так и далеко за её пределами.

Data Management

Where is LHC in Big Data Terms?



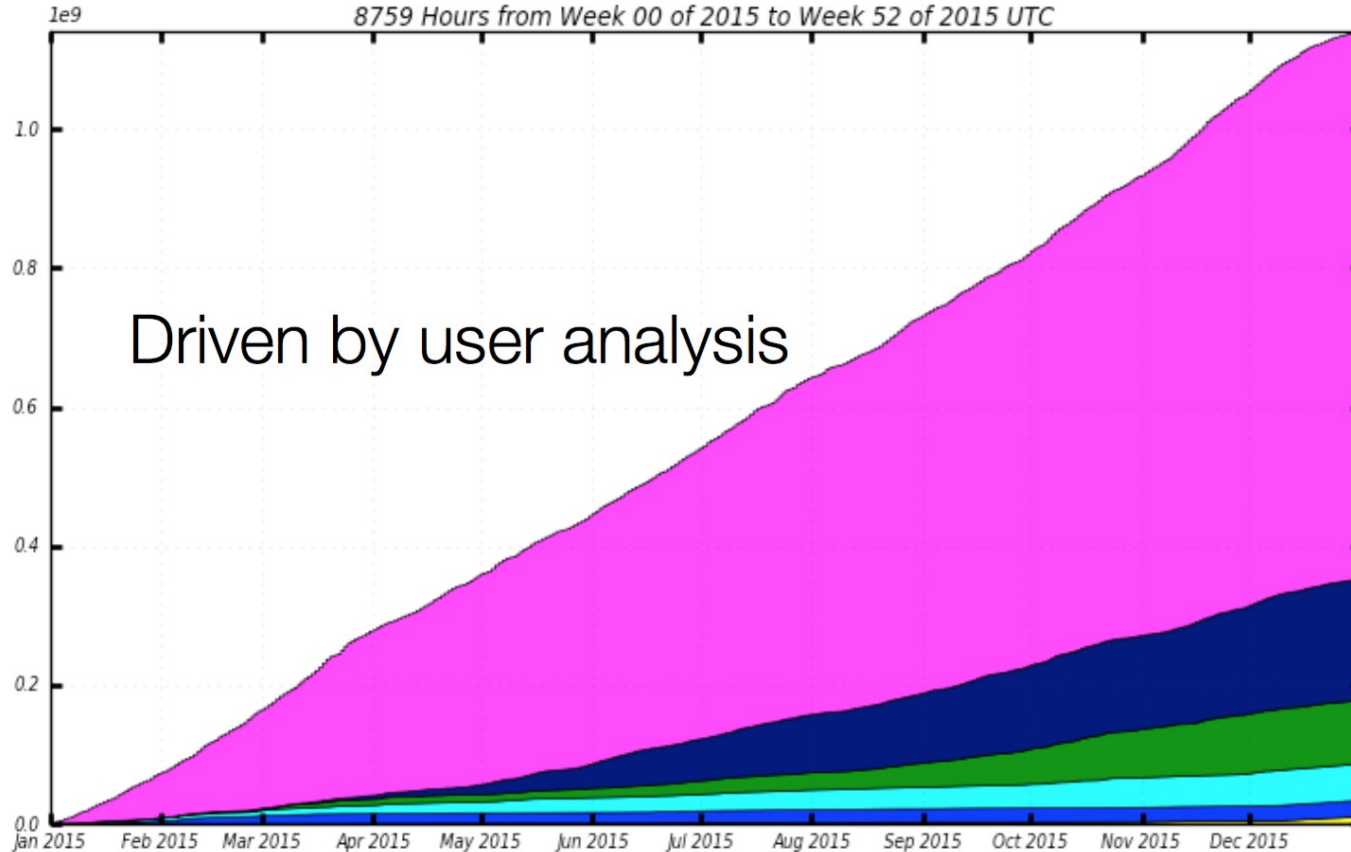
<http://www.wired.com/magazine/2013/04/bigdata/>

Data volume processed : the Exabyte area!



NBytes Processed in GBs

8759 Hours from Week 00 of 2015 to Week 52 of 2015 UTC



1 Exabyte

Total: 1,141,100,582 , Average Rate: 36.18 /s

40 ZETTABYTES
[43 TRILLION GIGABYTES]
of data will be created by 2020, an increase of 300 times from 2005



Volume
SCALE OF DATA



It's estimated that **2.5 QUINTILLION BYTES** [2.3 TRILLION GIGABYTES] of data are created each day



Most companies in the U.S. have at least **100 TERABYTES** [100,000 GIGABYTES] of data stored

The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015 **4.4 MILLION IT JOBS** will be created globally to support big data, with 1.9 million in the United States



As of 2011, the global size of data in healthcare was estimated to be **150 EXABYTES** [161 BILLION GIGABYTES]



30 BILLION PIECES OF CONTENT are shared on Facebook every month



By 2014, it's anticipated there will be **420 MILLION WEARABLE, WIRELESS HEALTH MONITORS**

4 BILLION+ HOURS OF VIDEO are watched on YouTube each month



400 MILLION TWEETS are sent per day by about 200 million monthly active users



The New York Stock Exchange captures **1 TB OF TRADE INFORMATION** during each trading session



Velocity
ANALYSIS OF STREAMING DATA



Modern cars have close to **100 SENSORS** that monitor items such as fuel level and tire pressure

By 2016, it is projected there will be **18.9 BILLION NETWORK CONNECTIONS** - almost 2.5 connections per person on earth



1 IN 3 BUSINESS LEADERS don't trust the information they use to make decisions



in one survey were unsure of how much of their data was inaccurate

Veracity
UNCERTAINTY OF DATA

Poor data quality costs the US economy around **\$3.1 TRILLION A YEAR**



Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTec, QAS



Облачная архитектура

- Пять важных характеристик
 - Доступ по требованию в режиме самообслуживания; Быстрые каналы связи; Большой пул ресурсов; Гибкость; Измеряемость потребления ресурсов
- Три сервисные модели
 - SaaS, PaaS, IaaS
- Четыре модели развёртывания
 - Частное, коммунальное, публичное, гибридное
- Ref: csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

Программно-конфигурируемые ресурсы

- Программно-конфигурируемые сети (ПКС – software defined networks - SDN)
- Программно-конфигурируемые Системы Хранения Данных (ПКСХД – software defined storage - SDS)
- Программно-конфигурируемые Центр Обработки Данных (ПКЦОД – software defined data center - SDDC)

Наши реалии

- Предполагается, что приток специалистов в Институт в среднесрочной перспективе уменьшится или станет отрицательным:
 - следует планировать автоматизацию рутинных операций, чтоб уменьшить необходимость ручного вмешательства;

Виртуальная реальность в ОФВЭ

- Виртуализированы: мейл сервер, кластер ОФВЭ и все прочие сервисы (перевод завершился в 2015).
- Виртуальная компьютерная инфраструктура:
 - Openstack (в экспериментальном тестировании);
 - Программно-конфигурируемая Система Хранения Данных CEPH (в экспериментальном тестировании).

Перспективные предложения (на несколько лет)

- СХД для документов ПИЯФ (чтоб не посылать мейлами сотни МБ);
- Мейл сервер: отображение в облака.
- WiFi – международный роуминг по авторизации.
- Виртуальная компьютерная инфраструктура (включая все сервисы вида серверов доменных имён, подобного).
- Внутренние каналы связи – 10 Гбит.
- Сменные консультанты по ИТ в ПИЯФ

Благодарю за внимание!

Add Ref: http://hepd.pnpi.spb.ru/hepd/events/meetings/sessiya_2015/dec23/shevel_2015-12-23.pdf