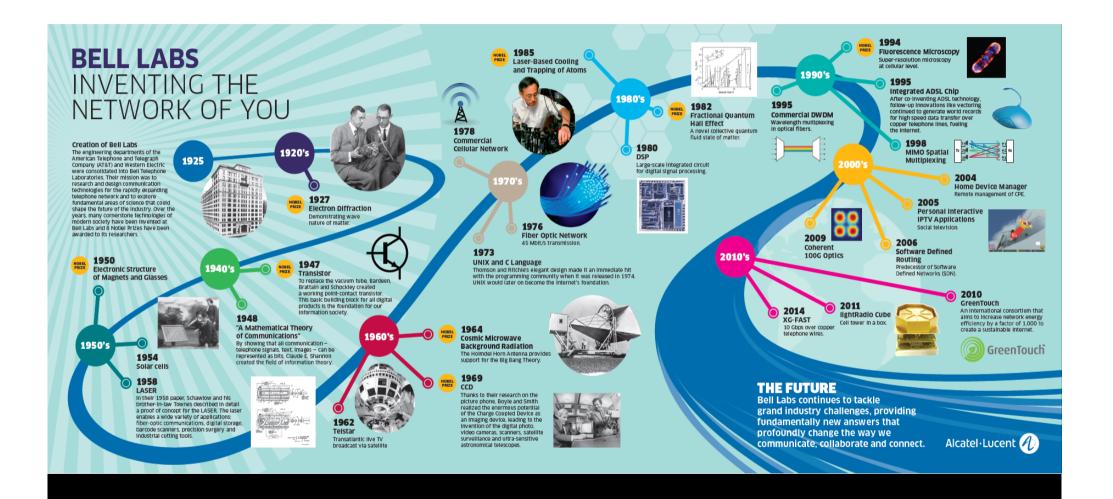


#### **FUTURE COMMUNICATION CLOUDS**

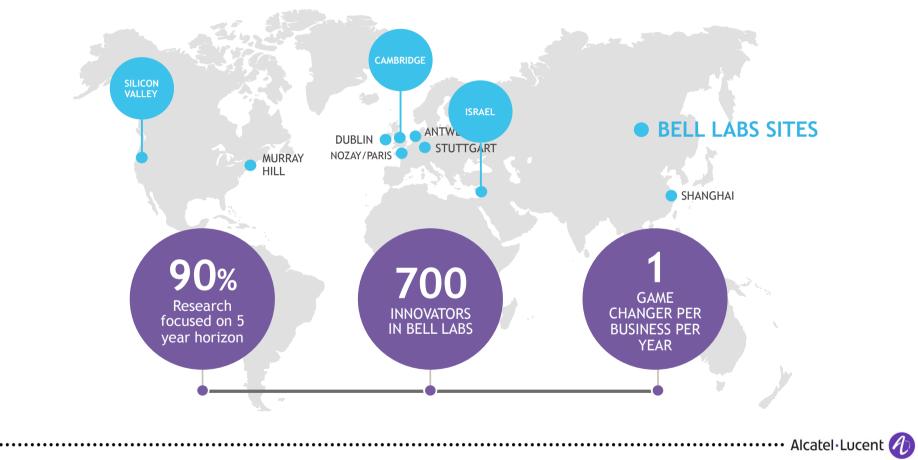
Krishan Sabnani

Bell Labs 🕖



#### Who/Where We Are

**GLOBAL INNOVATION** 



COPYRIGHT © 2015 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

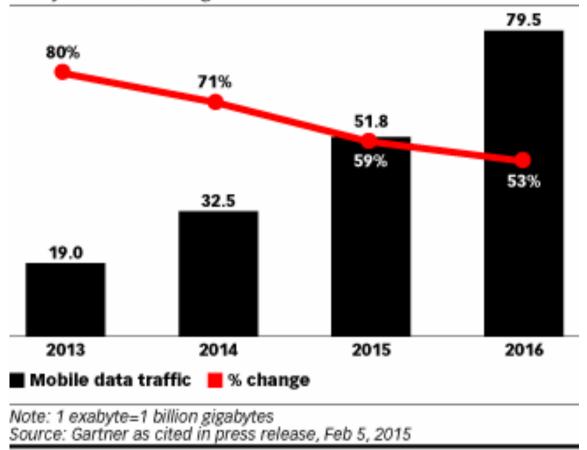
3

# What Happens in an Internet Minute?



#### Mobile Data Traffic Worldwide, 2013-2016

exabytes and % change



185120

www.eMarketer.com

#### SOME CUSTOMER STORIES



- I want a network which is elastic, that scales with my business, software definable and ondemand
  - John Donovan COO AT&T



- If Google bought Sprint, with its spare change, and converted it to a company that runs like Google, they would crush us...
- We need to break 130 years of Alexander Graham Bell's way of doing things
- I am going to build a truly elastic, programmable network
  - Ben Parker VP, Verizon Wireless

#### SOME CUSTOMER STORIES It's all about OPEX!

- Google has 1000 devices in its spare parts inventory, a Telecom Provider has 50,000
- Need trained personnel to maintain all the different boxes we have
- One network for mobile and fixed

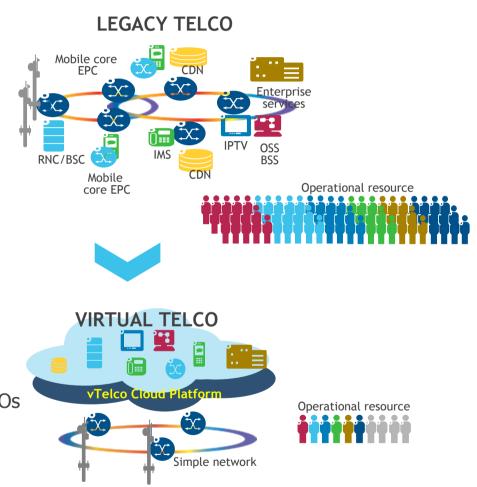


COPYRIGHT © 2012 ALCATEL-LUCENT. ALL RIGHTS RESERVED.

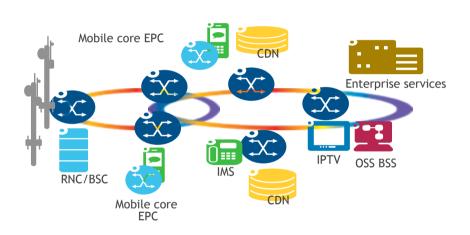
#### **Network Function Virtualization**

Network Function Virtualization aka vTelco

- Perceived benefits
  - Reduced equipment cost
  - Reduced power, space and operational expenses
  - Faster Time-to-Market
- Requirements for success
  - Portability between vendors and hypervisors
  - Full automation
  - Co-existence with legacy networks
- NFV Industry Specification Group (ISG)
  - Operator-led under the auspices of ETSI
  - Started in January 2013
  - Study NFV and drive standardization efforts in other SDOs



#### **NETWORK FUNCTION VIRTUALIZATION** MOVING NETWORK FUNCTIONALITY TO THE CLOUD



Legacy Service Provider



# Virtualized Service Provider

Operational resource

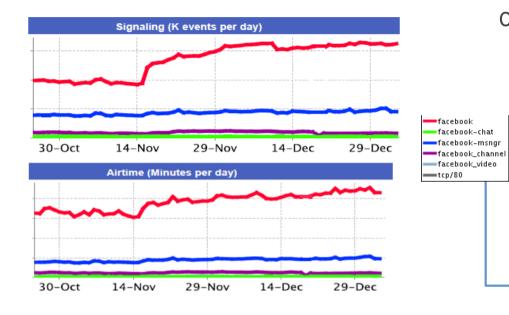


#### **Overview**

Current cellular networks face many performance issues created by modern applications

- Bandwidth limitation
- Inability to deal with short bursts of data
  - -> Sample solutions: Massive MIMO, mm waves, Aloha like-schemes in 5G
- Tight delay needs especially for a new class of resource- and interaction-intensive applications can not be met by even these enhanced networks
  - -> Solution : Create an overlay network of Cloudlets
- Efficient use of network capacity
  - -> One Solution: Preload content using Mobile Smartloading (MSL)

#### An Application Upgrade can Significantly Increase Network Load



| Application      | Airtime %<br>Change | Signaling %<br>Change | Volume %<br>Change | Subscriber %<br>Change |
|------------------|---------------------|-----------------------|--------------------|------------------------|
| facebook         | 27%                 | 60%                   | 0%                 | 4%                     |
| facebook_channel | 9%                  | 15%                   | 9%                 | -3%                    |
| facebook_video   | 390%                | 277%                  | 366%               | 154%                   |
| facebook-chat    | 16%                 | -3%                   | 11%                | -17%                   |
| facebook-msngr   | 23%                 | 18%                   | 17%                | 6%                     |

On Nov 15 of 2013, Facebook released new versions of their mobile app for Android and iOS.

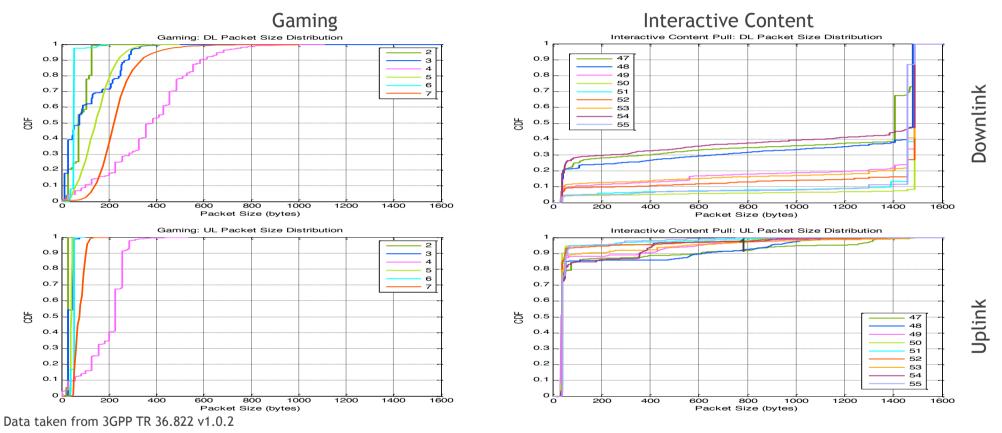
#### FACEBOOK SIGNALING INCREASED

- 60% increase in Facebook signaling
- 25% increase in Facebook airtime
- increase in the amount of signaling per user
- 5-10% increase in overall signaling load!!

#### FACEBOOK VIDEO USAGE INCREASED

- 350% increase in Facebook video volume
- more subscribers using video maybe sharing news feeds

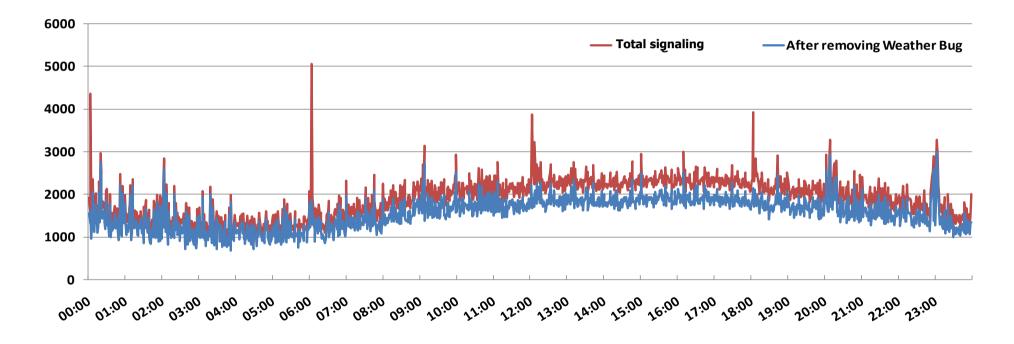
#### **Predominance of Short Bursts**



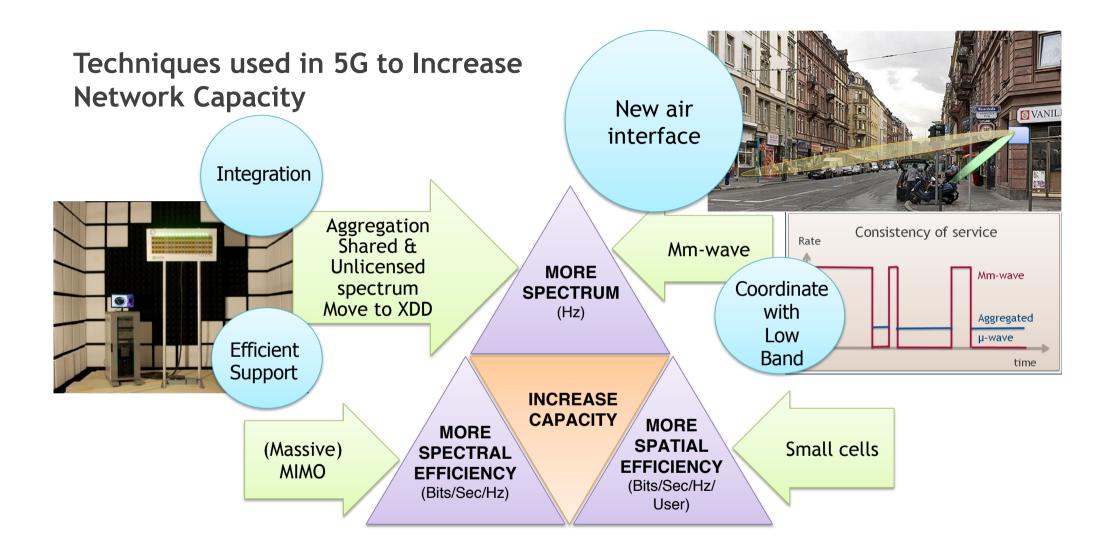
More than 50% of Today's 4G Traffic is Short Bursts of a Few 100 Bytes

12

An Example of Signaling Overload - Burst of Short Packets Wireless Ignorant Applications can Take Down Control Planes



Millions of Android phones running one weather application all contact server once every 6 hours in a synchronized way that cause signaling spikes overloading the control plane



#### Working towards a Solution for Short Bursts New Air Interface and Protocols for 5G

- Efficiently support very low latency scenarios
- Efficiently support short burst applications (Aloha-like scheme)
- Enable low energy user and machine devices
- Allow fast random access for long battery life devices
- Flexibly support both short burst traffic and high bandwidth traffic in same spectrum

## SG IS ABOUT A COMMUNICATIONS SERVICE THAT ADAPTS TO THE CONSUMER,

RATHER THAN THE CONSUMER ADAPTING TO THE COMMUNICATIONS SERVICE.

Alcatel·Lucent 🕖

### An Overlay of Cloudlets for Resource- and Interaction-Intensive Applications

#### Wireless Networks and Cloud Infrastructure will not Meet Latency Needs for Resource- and Interaction-Intensive Applications

- Many mobile and NFV applications require latency of a few 10's of msec
  - Current networks and cloud infrastructure can not provide acceptable experience
- Bring computing close to users
- Cloudlets provide a compute assist for applications near the network edge
  - Mobile devices such as Google Glass do not have the necessary processing power
- This overlay is similar to a CDN network
  - It brings cloud-computing-resources and relevant content close to the clients

Some charts from a keynote by Mahadev Satyanarayanan at ACM Mobicase 2014, "Cloudlets: *At the Leading Edge of Cloud-Mobile Convergence*" and a keynote by Victor Bahl at Microsoft Devices & Networking Summit 2015, "emergence of micro data center (cloudlets/edges) for mobile computing".

# New Resource- and Interaction-Intensive Applications (Enabled by Wearable Devices such as Google Glass)



- Entirely new genre of applications
- Combine mobile and cloud with *real-time cognitive engines* scene analysis, object/person recognition, speech recognition, language translation, planning, navigation, question-answering technology, voice synthesis, ...
- Seamlessly integrated into inner loop of human perception and cognition

#### **Crisp Interactive Response**

- Humans are amazingly fast, accurate and robust
- face detection under hostile conditions < 700 ms
  - (low lighting, distorted optics)

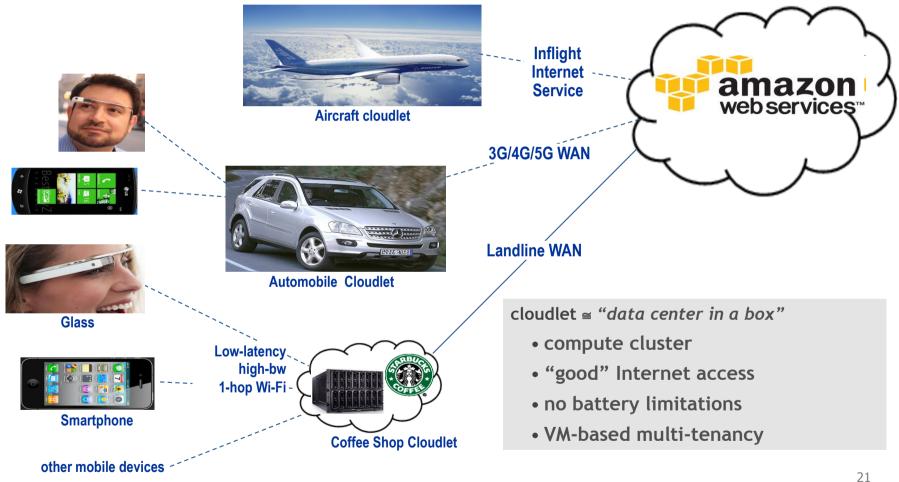
| - face recognition            | 370 ms – 620 ms                                  |
|-------------------------------|--|
| - is this sound from a human? | 4 ms   |
| - Head tracking               | < <b>16 ms</b><br>(2004 NASA study, Ellis et al) |
|                               | (2007 HAJA Study, L(IIS Ct at))                  |

- Not enough to just match humans
- we need to be "superhuman"
- allow enough time budget for additional cognitive processing
- consistent with the theme of the "Tactile Internet"

Safe goal: *E2E Latency < "few tens of ms"* 

#### **Bring the Cloud Closer**

Create a Small Cloudlet Nearby



#### **Especially Good for Mobile Devices** battery life improvement



55 50 Battery Life (in Hours) 05 25 07 25 25 75% increase in battery life 25 20 15 25 10 20 30 35 40 **#Network Transactions per Hour** -Battery Life Today -Battery Life w/ Proxy

these types of saving occur across the board for all battery types and all types of mobile devices

\* Samsung Standard LI-ION battery with rating of 1500mAh/3.7Vdc

#### calculated for a 30 msec speedup / network transaction

### Saving Bandwidth





locating objects of interest



customer queue analytics

#### current approach

• upload the captured video to the cloud for remote analysis

#### observations

- too much data captured per hour (>10GB/hour)
- bandwidth limits scale and use of system
- unable to support near real-time tracking & security

#### Micro Data Centers Exist Today



# Mobile Smartloading

#### Efficient Use of Network Capacity & Faster Content Access

#### What is Mobile Smartloading (MSL)?

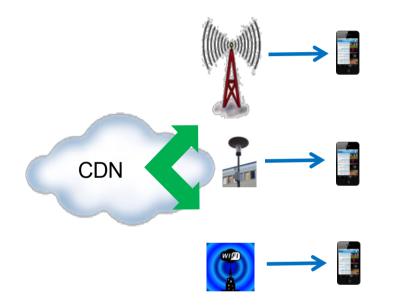
Congestion-aware automatic download and/or upload of any multimedia content to a smart-phone or tablet, selecting the most appropriate time and access technology





- HD viewing experience---stall/buffer free---even during peak periods
- Least cost delivery--using surplus RAN capacity, WiFi, Small Cells
- Policy-based opportunistic pre-loading---network, device, and user aware
- Excess capacity monetization for additional ARPU---Off-peak pricing

#### **Opportunistic Pre-placement of Content into User Devices**



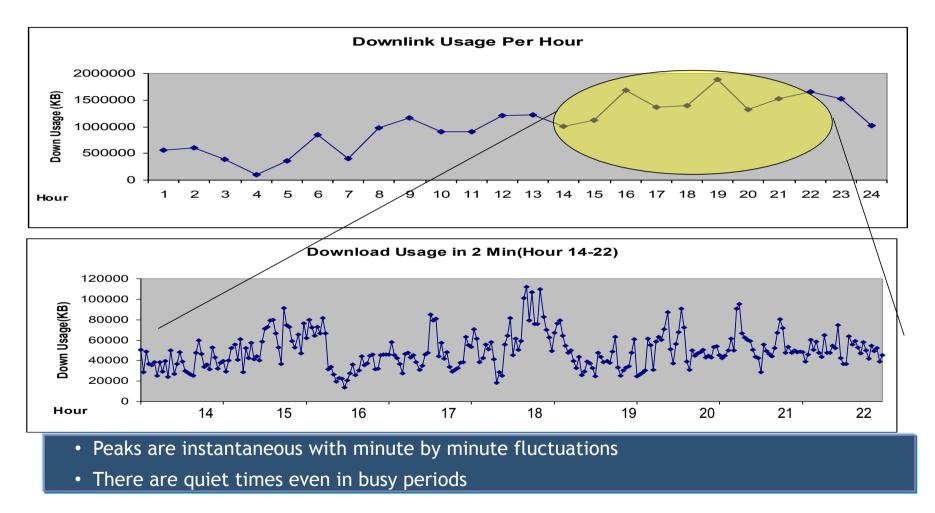
- > When the cell site is not congested
- > When in range of a high capacity metro cell

- > When at home with WiFi Access
- Cloud based, policy driven delivery scheduling
- With consideration for device battery & storage, Network congestion & quality
- Automatic suspend/resume to handle changes in network and device state

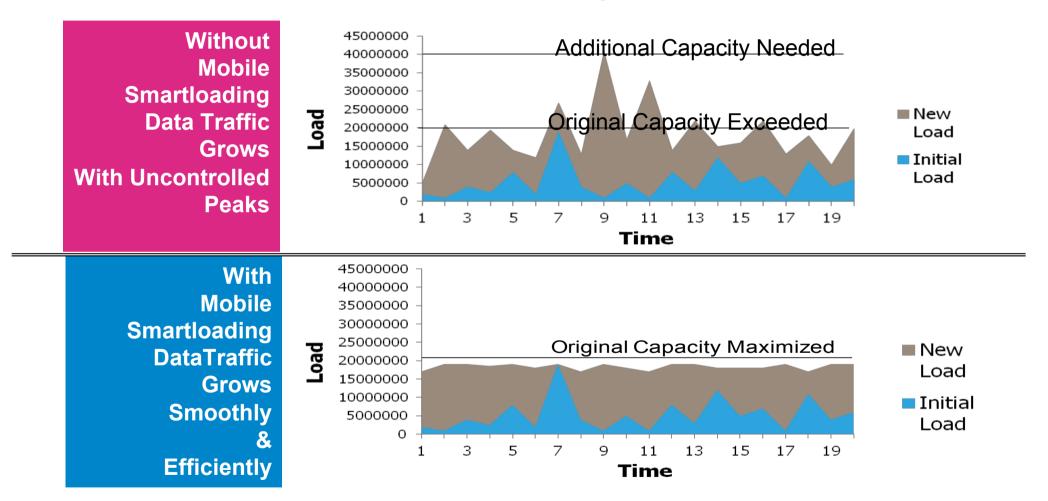
#### **Benefits for Mobile Operators**

| Network<br>Efficiency | <ul> <li>Defer expensive RAN upgrades</li> <li>Fully maximize network utilization</li> </ul> |
|-----------------------|--|
| Increase<br>ARPU      | <ul> <li>Sell Smartloaded apps &amp; pricing</li> <li>Monetize unused capacity</li> </ul>    |
| Greater<br>Relevance  | <ul> <li>Leverage network intelligence</li> <li>Tap into App ecosystem \$\$</li> </ul>       |

#### Data Traffic Trends: At The Base Station



#### **MSL Increases Utilization and Lowers Capex**



#### Conclusions

- Current wireless networks have performance issues such as limited bandwidth and inability to deal with short bursts of packets
  - They are being fixed in 5G
- Even these fixes will not deal with latency requirements for resource-and interaction-intensive applications
  - An overlay of Cloudlets
- Efficient usage of wireless networks
  - An example: Mobile Smartloading (MSL)

#### Predictions

- Future networks will be converged and network functions will be virtualized.
- CDNs will be enhanced to add computing close to users. The result will be similar to an overlay of cloudlets
- Content will be preloaded onto smartphones
- Spaces around us will be software-defined