Storage and Memory Infrastructure to Support 5G Applications

Tom Coughlin
President, Coughlin Associates
www.tomcoughlin.com
Outline

• 5G and its Implementation
• Storage and Memory Technologies
• Emerging Non-Volatile Memory Technologies
• New Network Storage for the Edge and Data Center
• Conclusions
5G Technology

• Stands for 5\textsuperscript{th} generation mobile technology and will allow cell phone technology with very high bandwidth

• Whereas 4G provides data rates of 5—100 Mbps, 5G will provide up to 10 Gbps

• In addition to increasing the data transfer speeds of mobile networks 5G technology is supposed to be more scalable, allow greater connectivity and provide greater energy efficiency

• 5G is aimed at enabling applications from IoT to vehicular telematics to faster video downloads and streaming services

From Huawei Slide on 5G Vision, 2015
5G Implementation

• In early 2018 Qualcomm demonstration 5G technology in San Francisco, CA and Frankfurt, Germany. In Frankfurt browsing jumped from 56 Mbps for a median 4G user connection to >490 Mbps for a median 5G user connection. In San Francisco these speeds increased from 71 Mbps to 1.4 Gbps.

• Significant investments in cell tower infrastructure will be required and telecom companies must agree upon standards for 5G networks.

• Will some cell towers be “data centers?”
The cloud and the fog

• The Cloud refers to compute resources, including storage, located in large data centers

• The Fog refers to local networks that connect thing (e.g. IoT) together

• Local fog networks may connect to the Internet
The Role of the Fog

Connected Cars and Smart Cities
From: Tao Zhang, Cisco Distinguished Engineer, Co-Founder and Board Director of OpenFog Consortium

© 2018 Coughlin Associates
There may be many “edges” in the “fog”

- 5G is intimately tied to the Internet and other communications
- More layers of connectivity are envisioned with IoT devices resulting in multiple “edge” connectivity and possibly data storage and processing
- Eluvio introduced the concept of a content fabric that uses blockchain authenticated peer to peer delivery of content versus today’s highly centralized Content Delivery Networks (CDNs).
EXAMPLE - CDN VS CONTENT FABRIC

**LIVE STREAM**
- 8 versions / edits
- 3 formats
- 5 bitrates (4 sec seg) = 19.85 MB / seg

<table>
<thead>
<tr>
<th>Bitrate</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master</td>
<td>12 MB</td>
</tr>
<tr>
<td>20 Mbps</td>
<td>10 MB</td>
</tr>
<tr>
<td>12 Mbps</td>
<td>6 MB</td>
</tr>
<tr>
<td>6 Mbps</td>
<td>3 MB</td>
</tr>
<tr>
<td>1.3 Mbps</td>
<td>650 KB</td>
</tr>
<tr>
<td>400 Kbps</td>
<td>200 KB</td>
</tr>
</tbody>
</table>

ALL BITRATES = 19.85 MB

**MANY BITRATES IN REGION** (example: Japan)
- only one version/edit
- only push master

- 1 x 3 x 12 MB
- 36 MB per segment
- 32.4 GB per stream

13x reduction

**TYPICAL CDN PUSH**
- 8 x 3 x 19.85 MB = 476.4 MB per segment
- 900 segments per 1 h stream
- 428.76 GB per 1h stream

**SINGLE BIT RATE REGION** (example: India mobile)
- only two versions/edits
- only push low bw

- 2 x 3 x 850 KB
- 5.1 MB per segment
- 4.59 GB per stream

93x reduction

Michelle Munson, Eluvio presentation at the 2018 Creative Storage Conference (www.creativestorage.org)
Investments in infrastructure

• Faster data delivery speeds need faster equipment behind the radio delivery system

• Lower latency requirements will drive requirements for communication, processing and storage at the ”edge.”

• Changes to new distribution models, such as Eluvio’s Content Fabric could change the nature of local connection nodes, including storage, processing and communication requirements

• In particular local content rendering, content caching and buffering will be key components in providing good content QoS

• This will lead to new investments in storage/memory architectures to support new ways to access, process and capture content and data
Digital Storage and Memory Tiering

Intel Optane DIMM Announcement, June 2018


© 2018 Coughlin Associates
Touch rate versus response time indicating various types of uses

Digital storage technologies regions overlaid on the Touch Rate/Response Time chart

Hard Disk Drives

• Currently shipping up to 14 TB with 20 + TB expected by 2019, 40 + TB by next decade
• Current HDD areal densities exceed 1 Gb per square inch
• He-filled HDDs provide greater energy efficiency and less cooling requirements
• Fastest growth in HDDs is for bulk cold storage in data centers
• Besides magnetic tape and optical discs, HDDs are the most cost effective storage medium for colder data
Flash Memory

- Flash memory is increasing in storage capacity (density) and decreasing in $/GB pricing but still more expensive than HDDs.
- Flash memory is winning more applications as its price ($/GB) drops.
- Development of NVMe and NVMe-oF has enabled better access to the performance capabilities of flash memory.
- In many data center applications, flash memory is now the primary storage.
- Flash Memory can also handle more rugged environments, making this a favored storage media for remote location—such as for edge storage.
Flash memory roadmap
Emerging Non-Volatile Memories

- There is intense effort to commercialize several non-volatile memories that could replace current volatile memories, such as DRAM and SRAM
- These technologies can be applied to stand along memory chips as well as in embedded memory
- This could reduce energy expenditure in battery and low power devices and also create more efficient data centers
- These NV memories will enable both IoT devices as well as data centers at the edge or in the cloud
- The memory technologies under consideration include magnetic random access memory (MRAM), resistive RAM (RRAM or ReRAM), phase change RAM (PRAM) and ferroelectric RAM (FRAM or FeRAM)
# Memory Technology Comparison

<table>
<thead>
<tr>
<th>Technology</th>
<th>FeRAM</th>
<th>MRAM</th>
<th>ReRAM</th>
<th>PCM</th>
<th>DRAM</th>
<th>NAND Flash</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonvolatile</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Endurance</td>
<td>$10^{12}$</td>
<td>$10^{12}$</td>
<td>$10^6$</td>
<td>$10^8$</td>
<td>$10^{15}$</td>
<td>$10^3$</td>
</tr>
<tr>
<td>Write Time</td>
<td>100ns</td>
<td>~10ns</td>
<td>~50ns</td>
<td>~75ns</td>
<td>10ns</td>
<td>10μs</td>
</tr>
<tr>
<td>Read Time</td>
<td>70ns</td>
<td>10ns</td>
<td>10ns</td>
<td>20ns</td>
<td>10ns</td>
<td>25μs</td>
</tr>
<tr>
<td>Power Consumption</td>
<td>Low</td>
<td>Medium/Low</td>
<td>Low</td>
<td>Medium</td>
<td>Very High</td>
<td>Very High</td>
</tr>
<tr>
<td>Cell Size (f²)</td>
<td>15-20</td>
<td>6-12</td>
<td>6-12</td>
<td>1-4</td>
<td>6-10</td>
<td>4</td>
</tr>
<tr>
<td>Cost ($/Gb)</td>
<td>$10/Gb</td>
<td>$30-70/Gb</td>
<td>Currently High</td>
<td>$0.16/Gb</td>
<td>$0.6/Gb</td>
<td>$0.03/Gb</td>
</tr>
</tbody>
</table>

How Persistent Memory will Success, Tom Coughlin and Jim Handy, 2018 SNIA Persistent Memory Summit, January 2018

© 2018 Coughlin Associates
Storage Systems for Edge and Data Center

• PCIe based NVMe storage interfaces will be the basis of future storage systems architectures using flash-based solid state drives

• This includes network storage capabilities including Remote Direct Memory Access (RDMA) and fabrics running the NVMe protocol, NVMe-oF
NVMe Architectures and DIMMS
PERSISTENT MEMORY OVER FABRICS (PMoF)

FOR DATA REPLICATION WITH DIRECT LOAD/STORE ACCESS
Storage Systems for Edge and Data Center (2)

- Solid state storage using server memory channels (DIMMs) will be the basis of non-volatile memory using flash memory and emerging technologies, such as Intel’s Optane.
- HDD object storage will be the backbone to mass storage for colder storage (behind solid state primary storage).
- Tape (or optical discs) may be used for longer term colder data.
Conclusions

• 5G will be a factor in the growth of digital content over the next decade and will increase requirements for more digital storage and memory solutions throughout the 5G network

• This could include real time rendering, buffering and caching solutions to improve content delivery QoS, perhaps using a peer to peer network

• The demand for lower latency will drive the use of solid state storage as primary storage in data centers and at the network edge

• New storage technologies and architectures as well as traditional storage technologies will enable 5G deployments
References


• Michelle Munson, Eluvio presentation at the 2018 Creative Storage Conference (www.creativestorage.org)


• How Persistent Memory will Success, Tom Coughlin and Jim Handy, 2018 SNIA Persistent Memory Summit, January 2018

Thanks!