NETWORK LINK PERFORMANCE PREDICTION

@ W3C WEB AND NETWORK IG

Jonas Svennebring
Principal Engineer

Jon Devlin
Director of
Business Development

LEGAL NOTICES/DISCLAIMER

- Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.
- No computer system can be absolutely secure.
- Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.
- Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.
- Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.
- Cost reduction scenarios described are intended as examples of how a given Intel- based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.
- Results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.
- Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.
- Intel, the Intel logo and others are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others.
-

BBB COPYRIGHT NOTICE

Pictures / video taken from the Big Buck Bunny movie:
(c) copyright 2008, Blender Foundation
www.bigbuckbunny.org

AGENDA

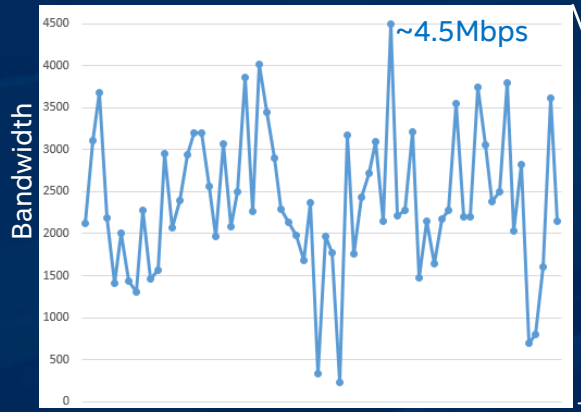
- Wireless Network Challenge
- Intel LPP Technology
- Application Usage of link prediction
- Who Benefits?
- API considerations
- Wednesday Demo: Remote gaming with/without LPP

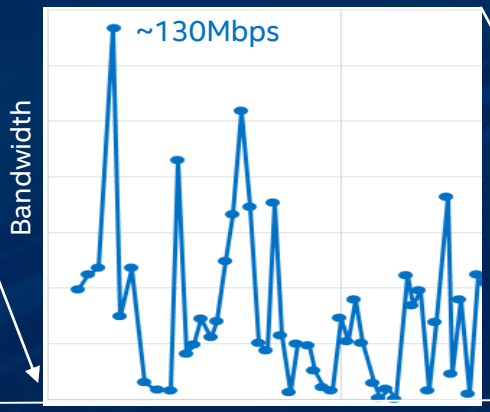
1. Wireless Network Challenge

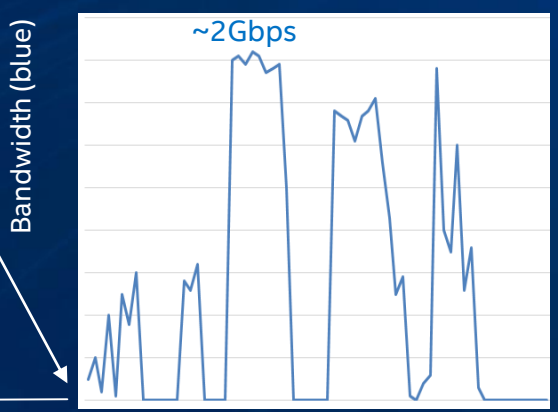
WIRELESS NETWORK CHALLANGES

Networks are better, but variations are larger

A) Large variations in quality _between_ networks
B) Large variations _within_ networks







WIRELESS NETWORK CHALLANGES

Networks are better, but variations are larger

A) Large variations in quality _between_ networks
B) Large variations _within_ networks
C) Big difference between Edge and Cloud
D) Edge can be many things, very different behavior

Networks are ”best effort” today
- limits the type of services allowed

Can we make it more deterministic?

INTEL LINK PERFORMANCE PREDICTION - LPP

Bring network awareness to the application

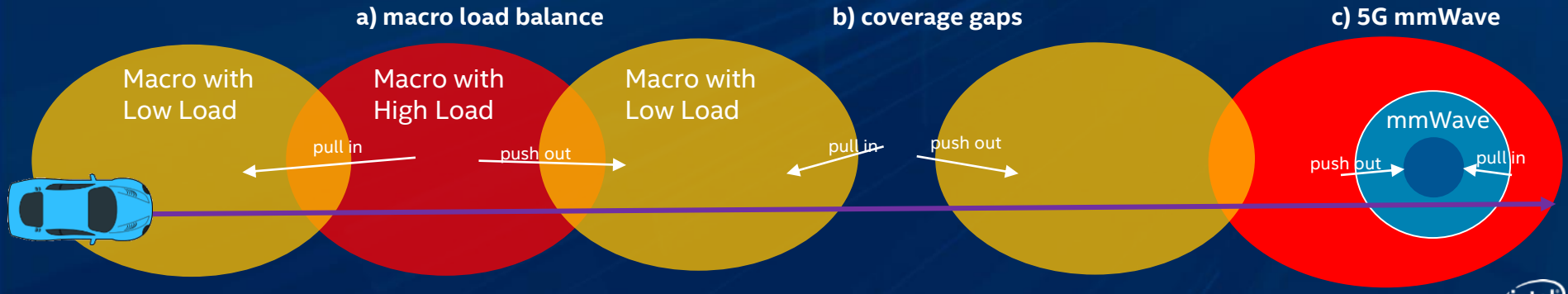
- Provided as ”hints” - application is still in control
- Current and near future link performance
- Multiple parameters: bandwidth, latency, cell load...

INTEL LINK PERFORMANCE PREDICTION - LPP

Bring network awareness to the application

- Provided as ”hints” - application is still in control
- Current and near future link performance
- Multiple parameters: bandwidth, latency, cell load…

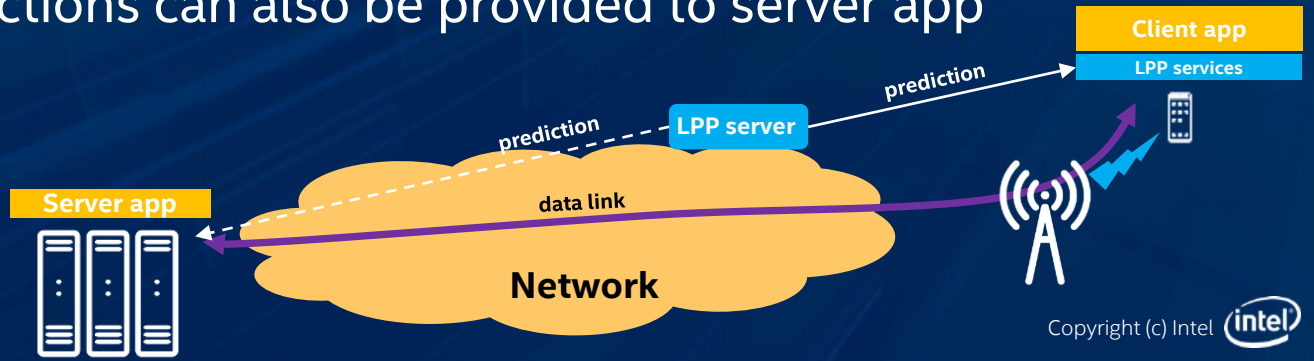
Examples:



2. Intel LPP Technology

INTEL LINK QUALITY PREDICTION (LPP) TECHNOLOGY

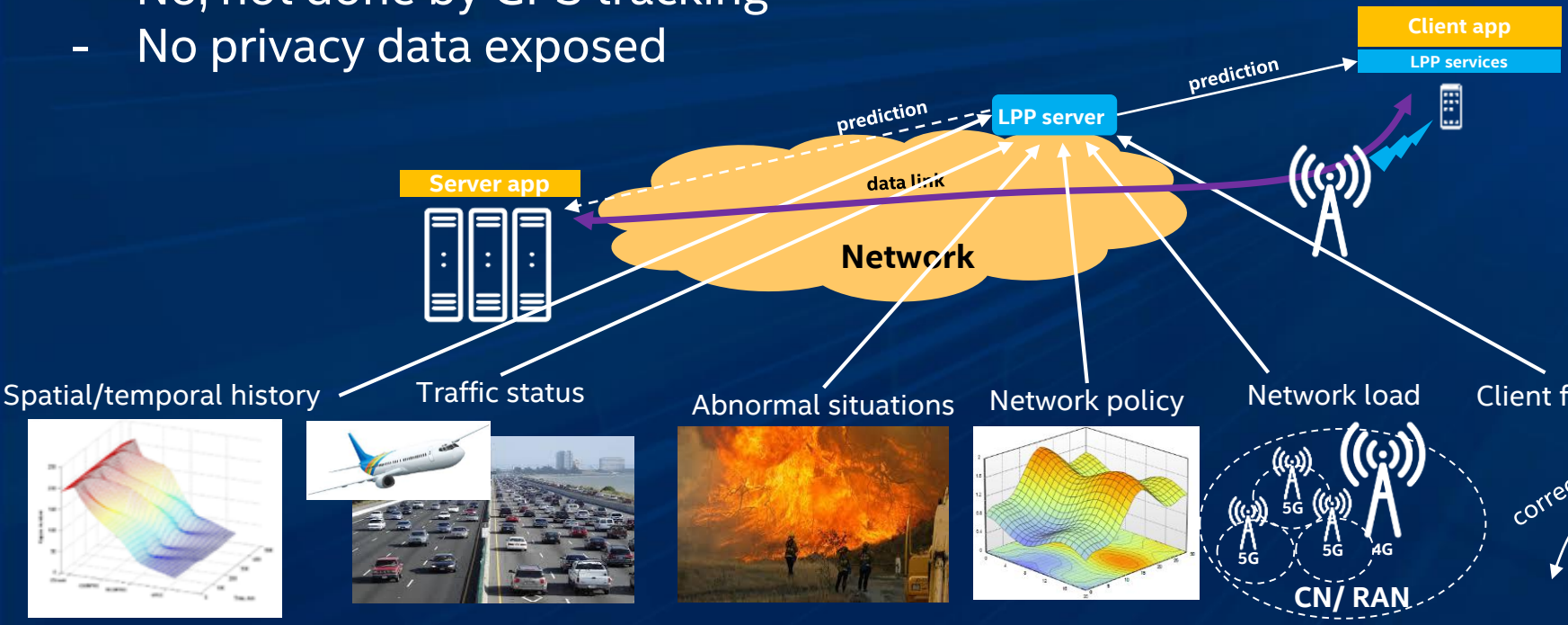
- Client/server connection as normal
  - Agnostic to Cloud, Edge etc.
  - No data is touched or routed through LPP server
- LPP server added to give link performance hints
  - LPP server resides in Operator network
  - Easy to use client service library to enable
  - Optionally predictions can also be provided to server app



INTEL LINK QUALITY PREDICTION (LPP) TECHNOLOGY

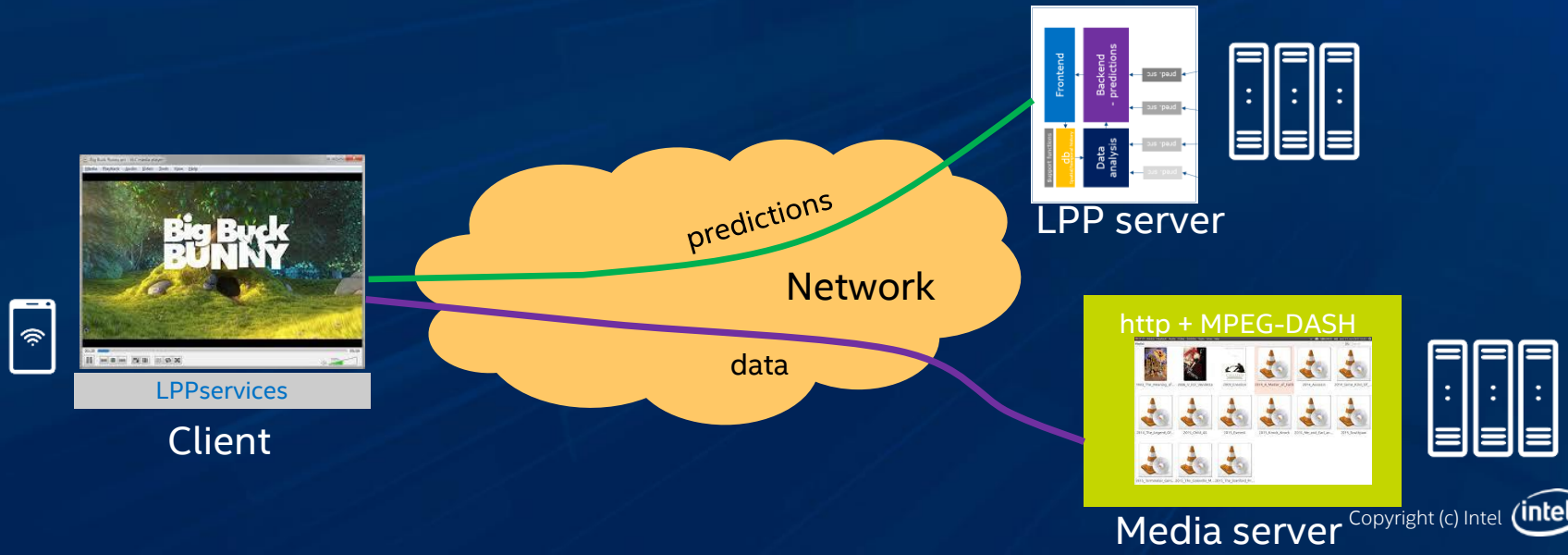
Predictions generated from a number of inputs…

- No, not done by GPS tracking
- No privacy data exposed



EXAMPLE USE-CASE: MEDIA STREAMING WITH LPP

- LPP server at location A
- Media played over standard MPEG-DASH media server at location B
- Client running web-browser with js-media player



STANDARD MPEG DASH – QUICK INTRO

- DASH – Dynamic Adaptive Streaming over HTTP, standard since 2012.
- Streaming players such as Youtube, Netflix, Amazon, BBC… are using DASH.

- Example: A movie is cut into small pieces, each 5 seconds long, and encoded in multiple different quality levels. Eg. movie_piece[0…1080]_quality[0..5].mov

- When download a piece the time is measured and used to select next quality:
- <2s download → increase quality next piece
- <5s download → keep quality
- >5s download → stall movie, lower quality next piece

- DASH is **reactive**, it assumes bandwidth is the same as it has been recently.

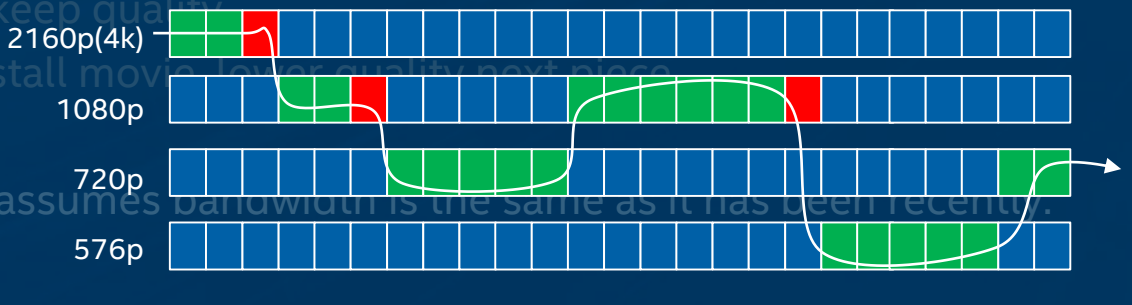
STANDARD MPEG DASH – QUICK

- DASH – Dynamic Adaptive Streaming over HT
- Streaming players such as Youtube, Netflix, A

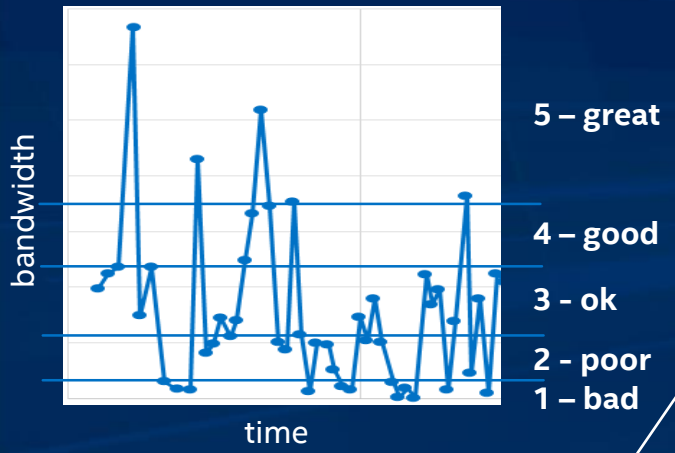
- Example: A movie is cut into small pieces, eac
  different quality levels. Eg. movie_piece[0...1

- When download a piece the time is measured and used to select next quality:
- <2s download → increase quality next piece
- <5s download → keep quality
- >5s download → stall movie, lower quality next piece

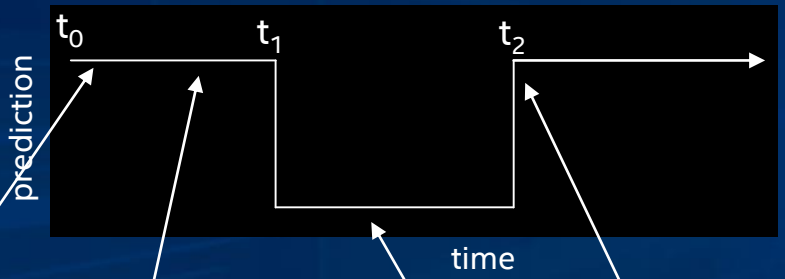
- DASH is **reactive**, it assumes bandwidth is the same as it has been recently.



DEMO – EXAMPLE OF LPP PRE-FETCH STRATEGY



Coverage gap scenario



t_1–n: Media player prefetches data to cover for the upcomming coverage gap.

t_0: Media player knows connection is good and can keep a lower buffer level – reducing cost / bandwidth when user ”jumps”

t_1: Media player consumes buffer, poor connection.

t_2: Media player rebuffers as needed.

DEMO – DASH REFERENCE JAVASCRIPT CLIENT + LPP

Target buffer level
- based on prediction

Actual buffer level

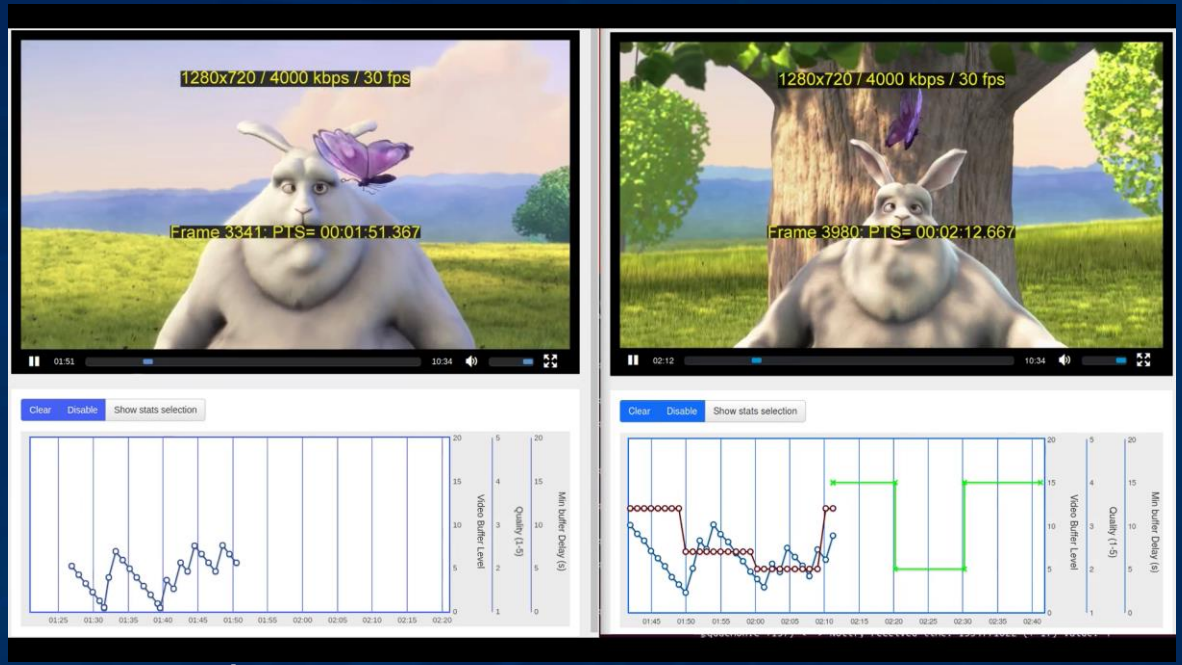
t_1–n: start prefetch

t_1: consume buffer

LPP prediction

t_2: rebuffer

SHOW DEMO...



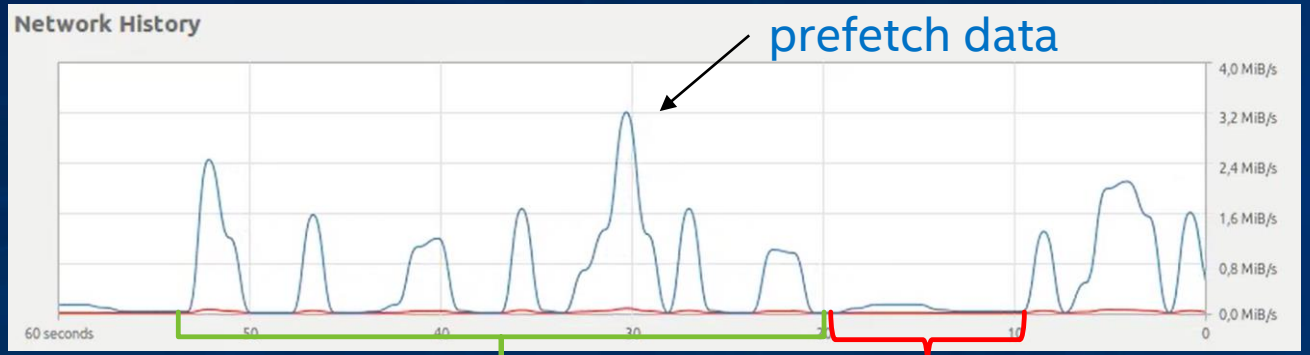
Normal DASH

DASH + LPP

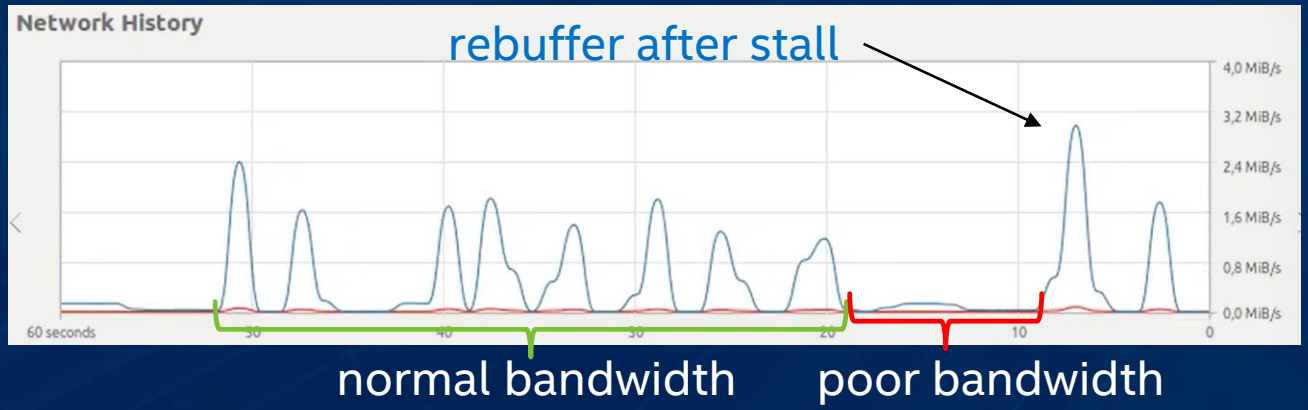
Note: Quality level is locked in this demo, when picture freezes this indicates its running out of video buffer.

EXAMPLE STREAM REQUIRES ~4MBPS PER PLAYER

DASH + LPP



DASH only



3. Application Usage of link prediction

EXAMPLE USAGE

- Pre-buffer when network is going to be poor
  - Improved user-experience
- Minimize buffer when network is good
  - Less buffer delay for real-time media
  - Less data transmission
- Remote Gaming – frame handling
  - Improved user-experience
- Network policy usage
  - New / improved usage scenarios for e.g. 3GPP based wireless networks

EXAMPLE USAGE: DYNAMIC BUFFER CONTROL WHEN STREAMING

Target: Maximizing user-experience / avoid stalls in media stream

Strategy: When bandwidth will be less than required by stream then prefetch data to cover the glitch

Measured results: Significant reduction in stalls for scenarios tested

LPP not active:

```
Stream d2cd85e125c308ae started
Got one stall @ 92.0 -- 6564 ms
Got one stall @ 96.0 -- 8166 ms
Got one stall @ 204.0 -- 6796 ms
Got one stall @ 208.0 -- 9508 ms
Got one stall @ 212.0 -- 1618 ms
Got one stall @ 316.0 -- 3873 ms
Got one stall @ 320.1 -- 8879 ms
Got one stall @ 324.0 -- 4289 ms
Got one stall @ 432.0 -- 4852 ms
Got one stall @ 436.0 -- 7476 ms
Got one stall @ 440.0 -- 539 ms
Got one stall @ 548.0 -- 880 ms
Got one stall @ 552.0 -- 9105 ms
Got one stall @ 556.0 -- 4284 ms
Stream d2cd85e125c308ae paused @ 634.566666
##############################
Stream: d2cd85e125c308ae Date UTC: Mon, 27 May 2019 05:37:15 GMT
LQP disabled DefaultBufferLevel 30
14 stalls occurred for 76.8 s of 712 s (10.8%)
Total time: 711.5 s, Play time: 634.707 s
Downloaded data: 912043 kB, Played data: 912043 kB
```

Stalls while playing

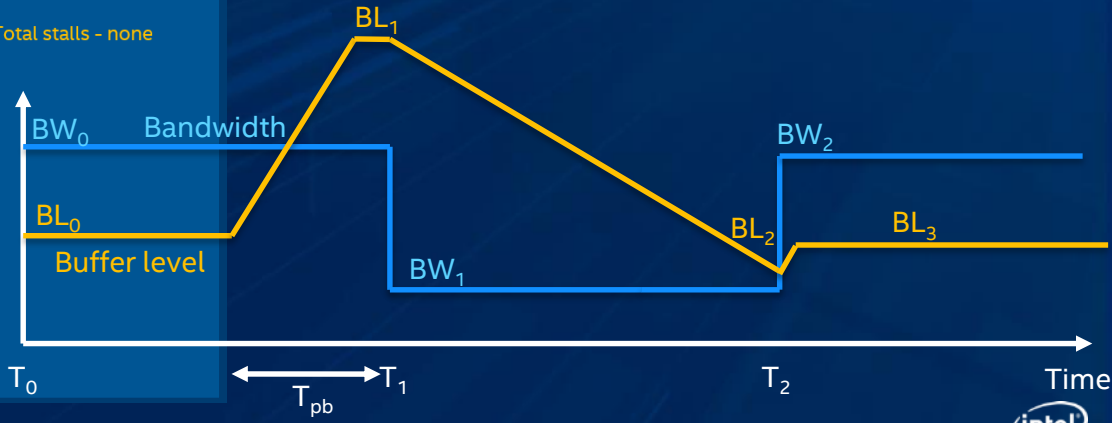
Total stalls

LPP activated:

```
Stream ce9262c4fbf1b353 started
Stream ce9262c4fbf1b353 paused @ 634.566666
##############################
Stream: ce9262c4fbf1b353 Date UTC: Mon, 27 May 2019 05:17:30 GMT
LQP enabled
0 stalls occurred for 0.0 s of 635 s (0.0%)
Total time: 634.6 s, Play time: 634.58 s
Downloaded data: 912043 kB, Played data: 912043 kB
```

Stalls while playing - none

Total stalls - none

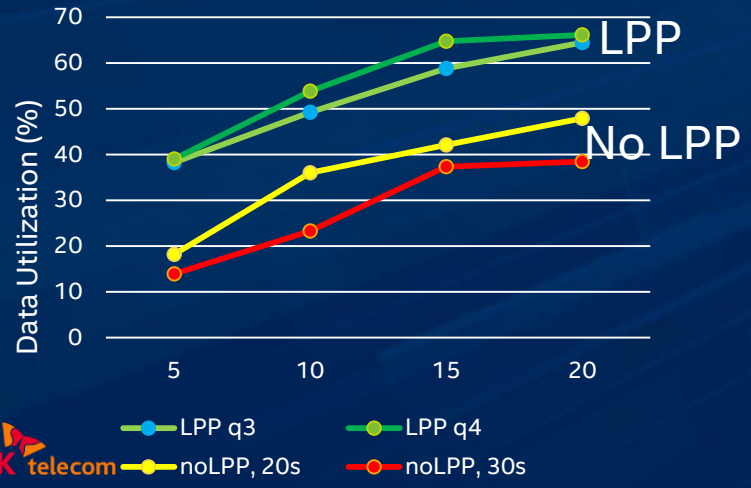


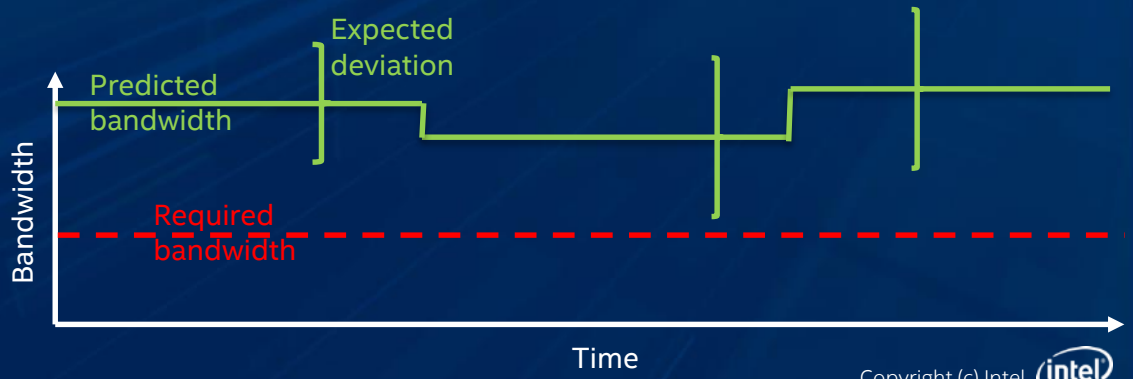
EXAMPLE USAGE: MIN BUFFER LEVEL WHEN STREAMING

Target: Minimize load on network when link connection is good

Strategy: Minimize buffer level, less data discarded when users jump/change media stream

Measured results: Significant improved data utilization rate, i.e. Less load on network





EXAMPLE USAGE: REMOTE GAMING FRAME HANDLING WITH LPP

network

video frames

user input

monitor

FB

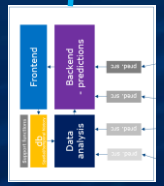
GPU

input

CPU

NIC

client



LPP server

FB

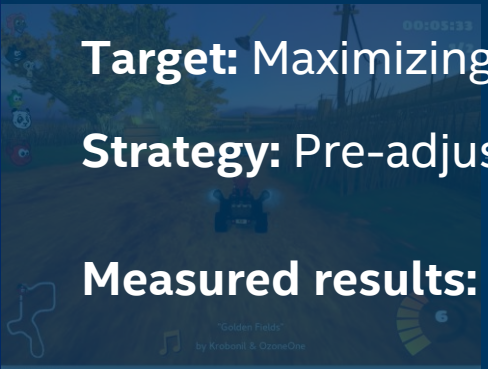
GPU

CPU

NIC

server

EXAMPLE USAGE: REMOTE GAMING FRAME HANDLING WITH LPP



Target: Maximizing user-experience / avoid stalls/lag when playing

Strategy: Pre-adjust channel bandwidth usage to coming availible bandwidth

Measured results: So far so good… more coming

Gaming vs. media streaming:
- No opportunity for buffering
- Server endpoint drives data transfer
  - I.e. server needs to get the predictions

Similar concept can be used for AR/VR with some adjustments

EXAMPLE USAGE: NETWORK POLICY USAGE

- Non-time critical traffic accounts for ~10% of network traffic in 3GPP networks
- Communicating ”network load” or ”network transfer policy” can push these loads to low-usage areas/times
- Enterprice coverage (5G indoor), free wifi etc. can leverage same solution

What’s the incentive given?

-”be nice” is not a good motivation

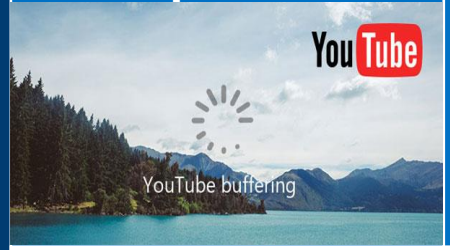
4. Who Benefits?

Use Case 1 – Bundled Video Mobile Plan

Description	**Bundle prediction capability with specific OTT applications (e.g. Oksusu, YouTube Red) without using user's data allowance. An example is T-Mobile/Netflix. LPP provides extra value for the user and/or to optimize the use of the SP network for users that are heavy video users.**
	
Value Proposition	Premium consumer experience, and efficient use of SP network resources. The combination of premium content and premium quality, unlocks additional subscriber WTP. This could be applied for a defined period of time (sports match, major event) or at all times.

"For a guaranteed quality of experience we will be able to charge a premium" Telefonica COO

Use Case 2 – OTT Video Optimized Slice

Description	**Take advantage of intelligent prediction of Network conditions to maximize Quality of Experience for Subscribers**
	
Value Proposition	Allows OTT applications (e.g. YouTube) to take advantage of dynamic network capabilities vs always assuming poor conditions (i.e. default buffering of 30 seconds of video content) = Increased subscriber stickiness for OTT

"5G Consumers expect to be able to stream videos seamlessly wherever they are"

Use Case 3 – Mobile Cloud Gaming Slice

Description	**Online games download parts of interactive games such as maps and textures, when prediction, based on the expected gameplay of the user and the connection quality (note gamers favor 4K content at 120 FPS with URLLC)**
	
Value Proposition	Premium consumer experience for mobile gamers (closer to console/PC experience when travelling). This could be applied for a defined period of time (particularly competitive event) or at all times.

"In certain services, lower latency is a must" EVP, Elisa

MONETIZATION OPTIONS – PREDICTION AS A SERVICE

Description	Offer 'prediction as a service' through an API. This includes user level predictions or predictions aggregated across a number of users. The insight drawn that prompts a particular movement of traffic within a network can also be provided to businesses or other third parties (e.g. enroute to sports event)
	
Value Proposition	Network optimization is improved by controlling timing of ad's with locations where excellent downlink expected = increased % impressions vs skips vs full views

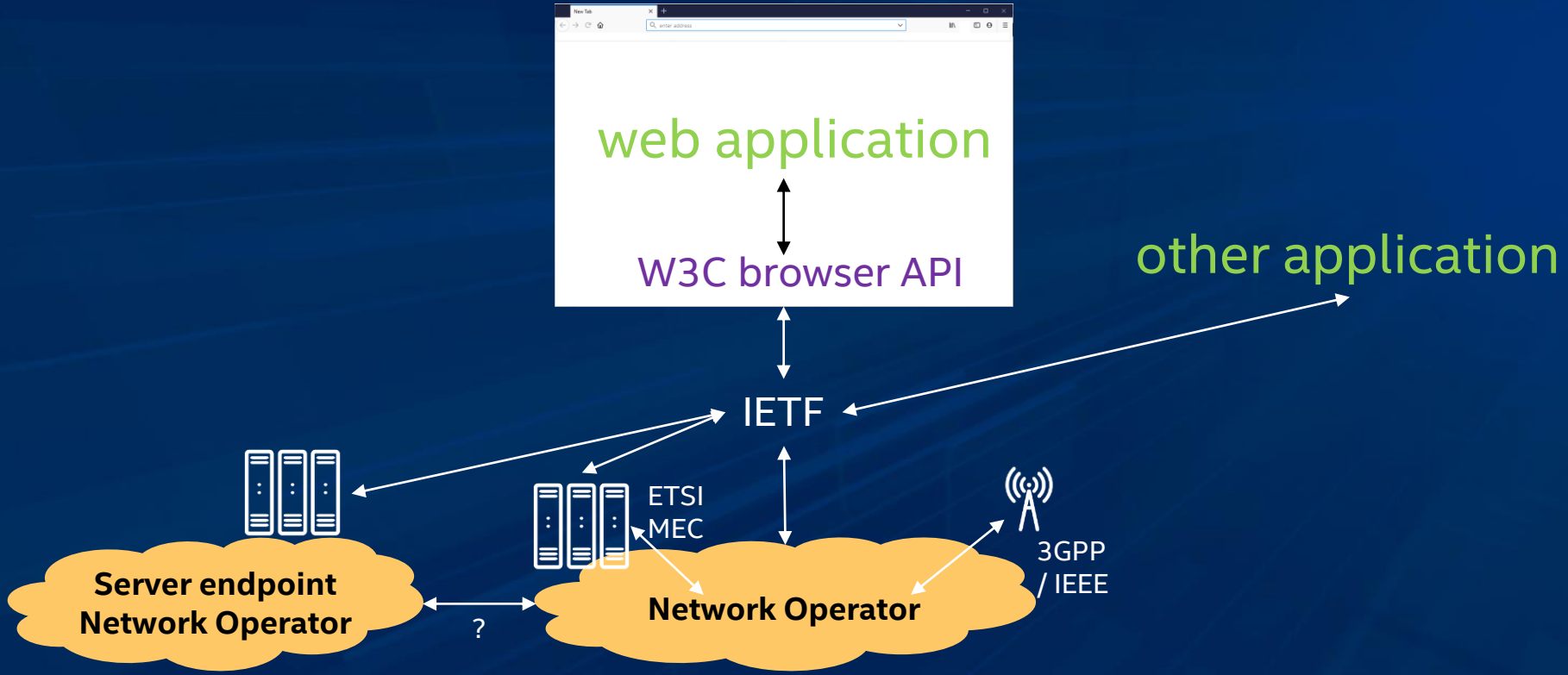
"5G interactive content like live-streaming, 4K videos and AR will become more powerful" MarTech

5. API considerations

NETWORK STATUS / PREDICTION - CONSIDERATIONS

- Get status/prediction
  - Information types
  - Value
  - Deviation
  - Probability
- Timeframe
  - Current conditions
  - Forward looking: short/medium/longterm
- Per client vs. per application status
- Who can get the link predictions?
  - Sharing handler
- What are we optimizing for?
  - Most likely
  - Worst case
  - Best case
  - Application defined specifiction
- Accuracy
  - Red / Yellow / Green
  - Descrete: e.g 1-5, 1-10, %...
  - Relative current: ++, +, 0, -, --
  - Magnitute
  - Actual

W3C VS. OTHER LAYERS

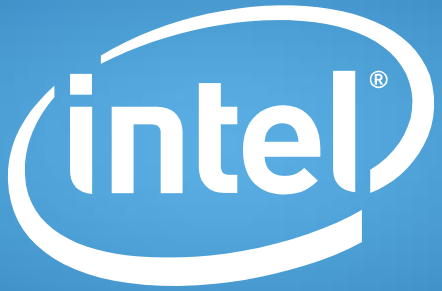


GOING GLOBAL...

How do you know where to get the status/prediction from?

- Service connection point
  - ”DNS”, network id -> server IP
  - Extending DHCP
  - Other?
- Security Certificates
  - Encryption and Authentication

6. Demo: Remote gaming with/without LPP


intel