



Design + Engineering

BIDI: BUILDING THE BBC'S MEDIA CDN

2ND NOVEMBER 2020



Alistair Wooldrige

Senior product manager, OTG Media Distribution.
alistair.wooldrige@bbc.co.uk

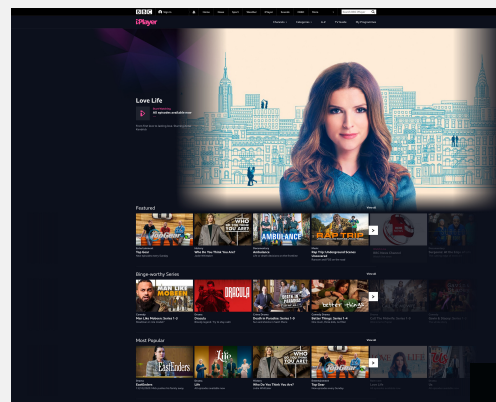
AGENDA

1. Introduction
2. BIDI growth so far
3. Future plans

INTRODUCTION

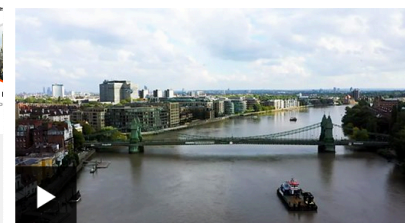
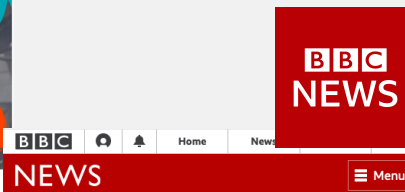
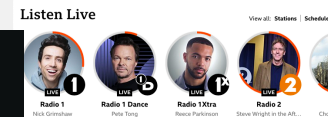
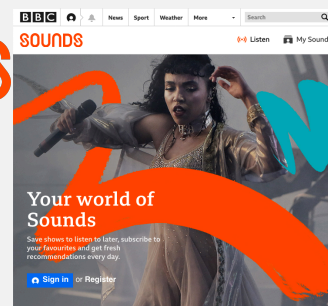
INTRODUCTION

LOTS OF MEDIA TO DISTRIBUTE OVER IP



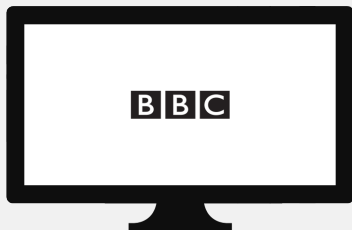
BBC iPlayer

BBC
SOUNDS



INTRODUCTION

VIDEO CLIENTS



TV

66%

PC

10%



Tablet

15%



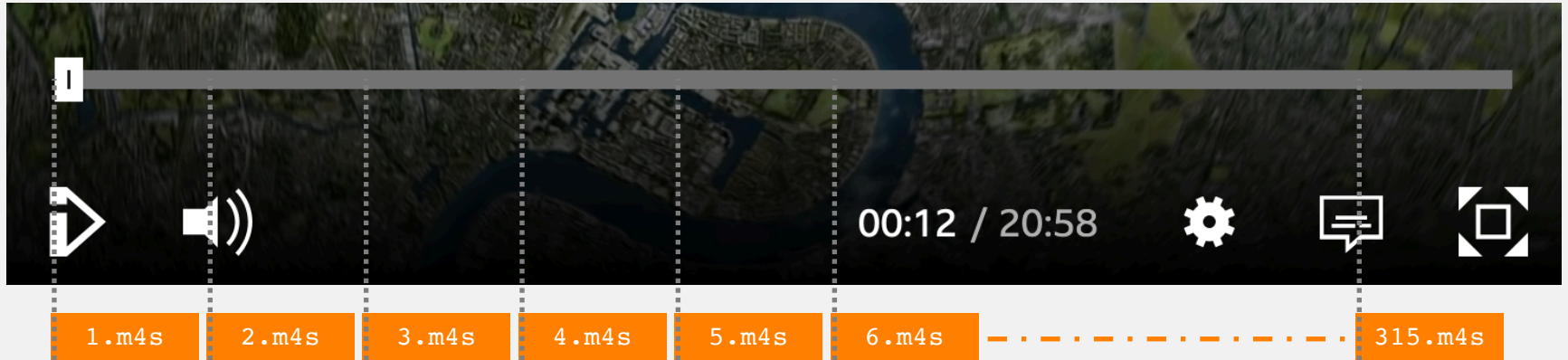
Mobile

9%



INTRODUCTION

SEGMENTED MEDIA

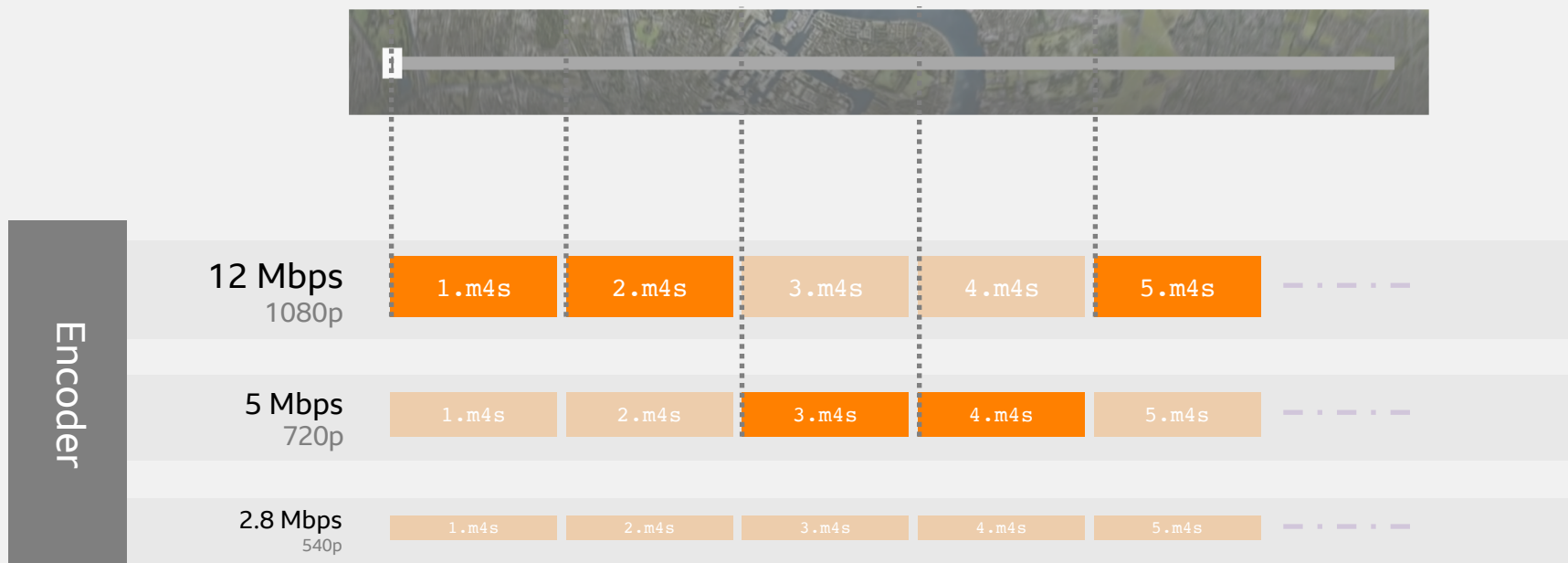


Split into ~4 second segments.

Client downloads segments in advance, over HTTP.

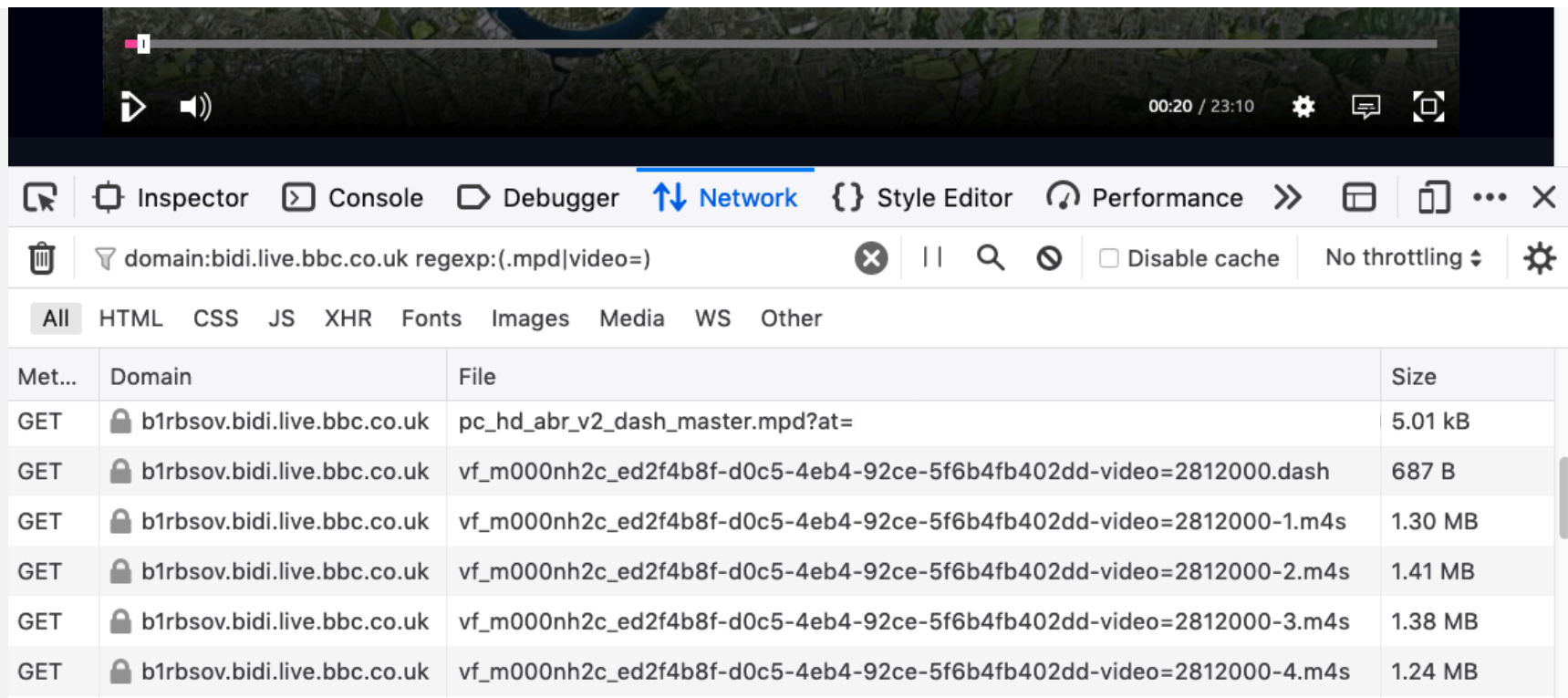
INTRODUCTION

SEGMENTED MEDIA: ADAPTIVE BIT RATE (ABR)



INTRODUCTION

SEGMENTED MEDIA: IN ACTION

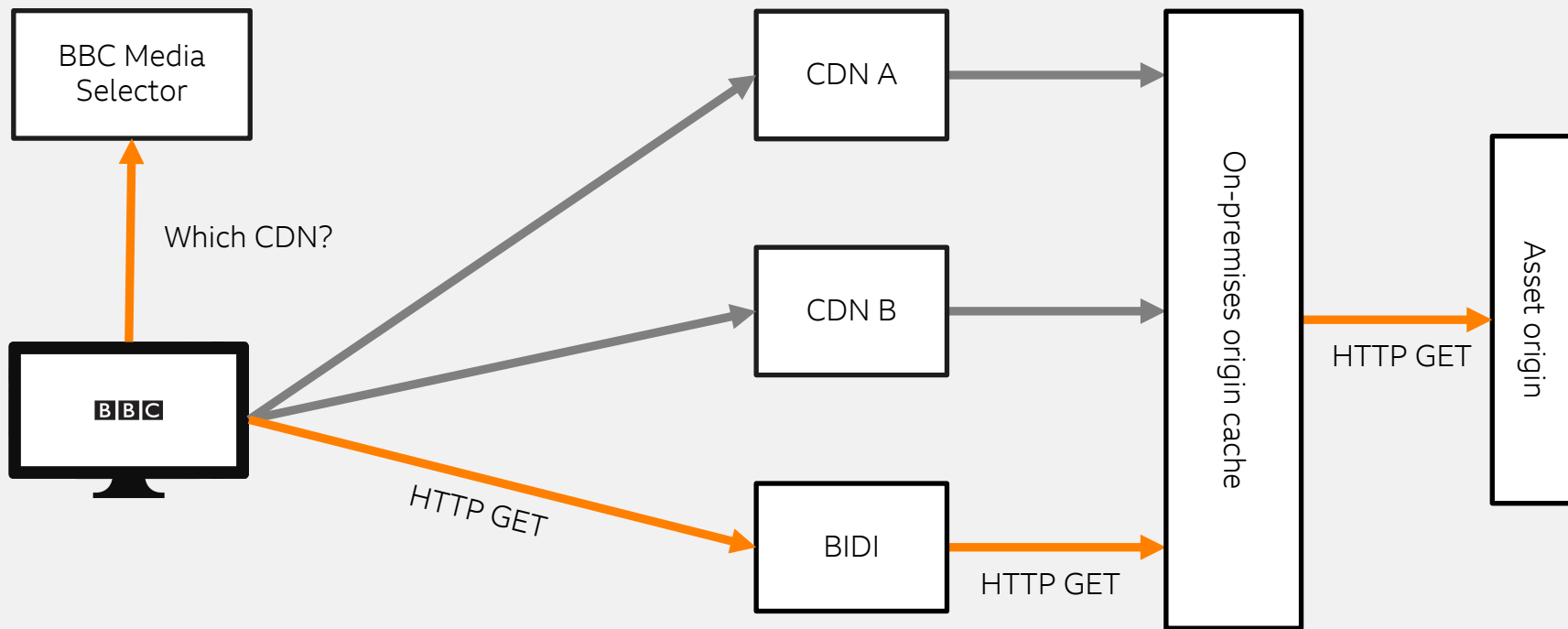


The screenshot displays a video player interface at the top, showing a progress bar and playback controls. Below the player is the Chrome DevTools Network tab, which is filtered to show requests from the domain `bidi.live.bbc.co.uk` with a regex filter `regexp:(.mpd|video=)`. The network log shows several requests for segmented media files, including a manifest file and four video segments.

Met...	Domain	File	Size
GET	b1rbsov.bidi.live.bbc.co.uk	pc_hd_abr_v2_dash_master.mpd?at=	5.01 kB
GET	b1rbsov.bidi.live.bbc.co.uk	vf_m000nh2c_ed2f4b8f-d0c5-4eb4-92ce-5f6b4fb402dd-video=2812000.dash	687 B
GET	b1rbsov.bidi.live.bbc.co.uk	vf_m000nh2c_ed2f4b8f-d0c5-4eb4-92ce-5f6b4fb402dd-video=2812000-1.m4s	1.30 MB
GET	b1rbsov.bidi.live.bbc.co.uk	vf_m000nh2c_ed2f4b8f-d0c5-4eb4-92ce-5f6b4fb402dd-video=2812000-2.m4s	1.41 MB
GET	b1rbsov.bidi.live.bbc.co.uk	vf_m000nh2c_ed2f4b8f-d0c5-4eb4-92ce-5f6b4fb402dd-video=2812000-3.m4s	1.38 MB
GET	b1rbsov.bidi.live.bbc.co.uk	vf_m000nh2c_ed2f4b8f-d0c5-4eb4-92ce-5f6b4fb402dd-video=2812000-4.m4s	1.24 MB

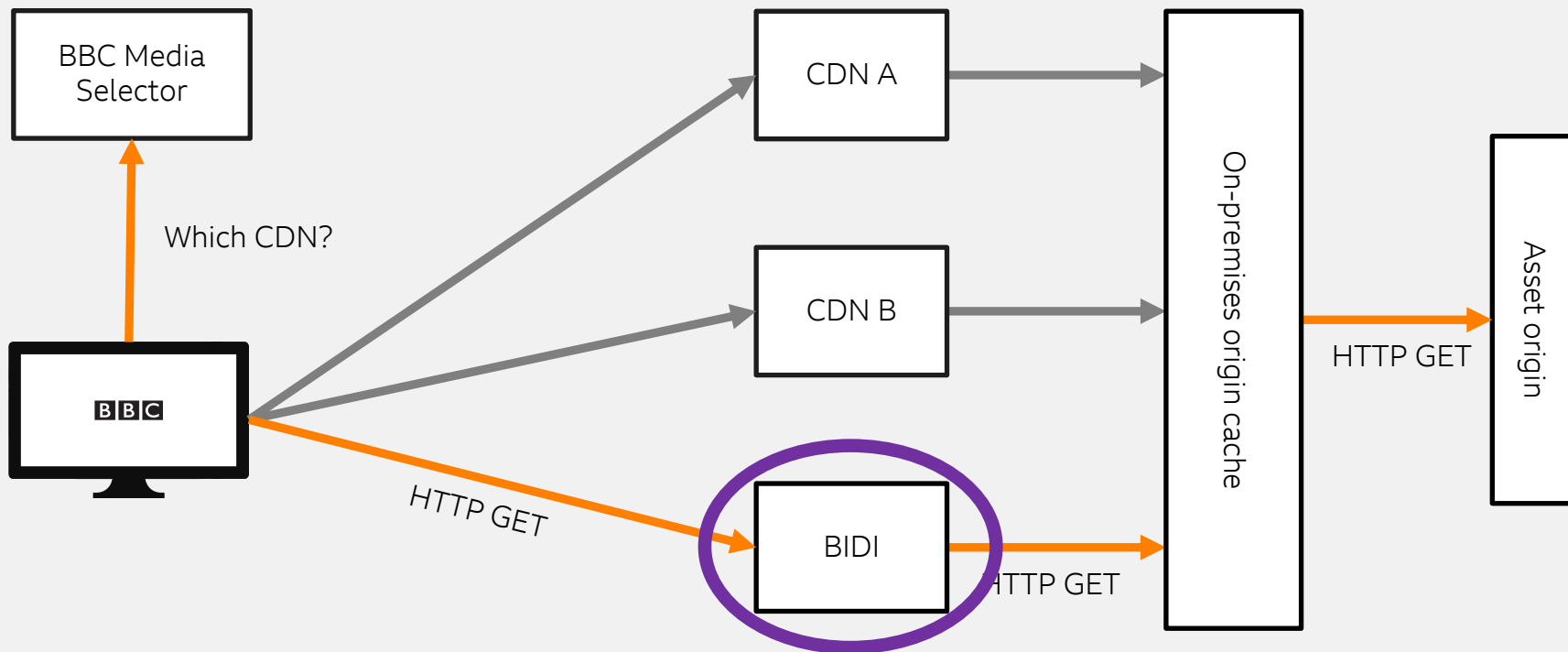
INTRODUCTION

PULL THROUGH CACHING



INTRODUCTION

PULL THROUGH CACHING



INTRODUCTION

DEFINITION

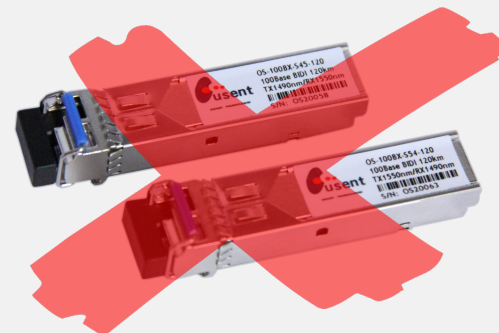
BIDI:

BBC Internet Distribution Infrastructure



Image: <https://www.flickr.com/photos/vintagehalloweencollector/421733590>

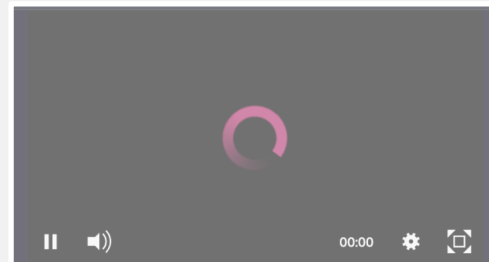
Buck Rogers (maybe)



BiDi transceivers (no)

INTRODUCTION

WHY BUILD OUR OWN CDN?



1) IP distribution future

Consumption is transitioning from broadcast (digital terrestrial, satellite, cable) to over-the-top IP distribution.

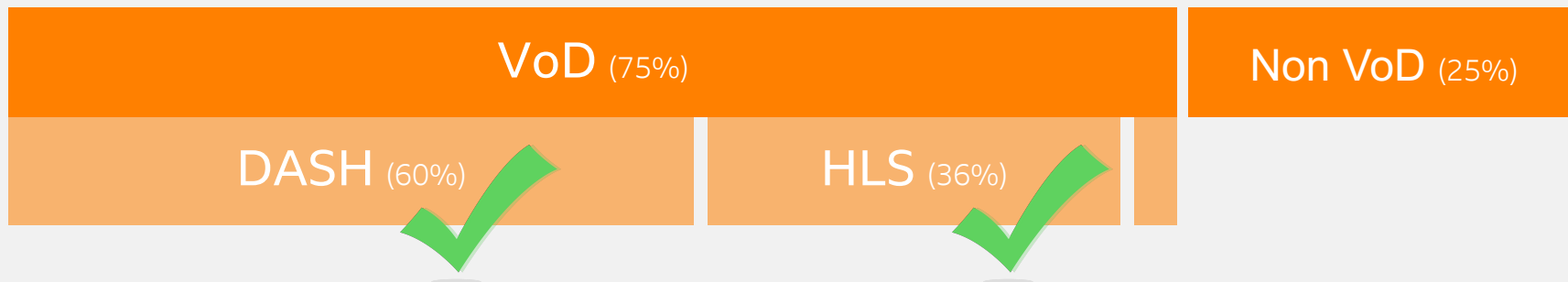
2) Quality of Service

Optimise hardware and software for high quality of service. Reduce latency for remote UK audiences.

INTRODUCTION

CONTENT SERVED BY BIDI

BIDI serves UK public service only, focussing on the high volume video on demand content sets:

BBC UK CDN traffic breakdown in September 2020:

DASH: Dynamic Adaptive Streaming over HTTP.

HLS: HTTP Live Streaming.

BIDI GROWTH SO FAR

BIDI GROWTH SO FAR

TIMELINE: R&D PROJECT



- Proof of concept built.
- Small amount of production traffic.
- Handover and joint trial.

BIDI GROWTH SO FAR

TIMELINE: PRODUCTION BUILD



- Handover complete.
- Production build & rewrite.

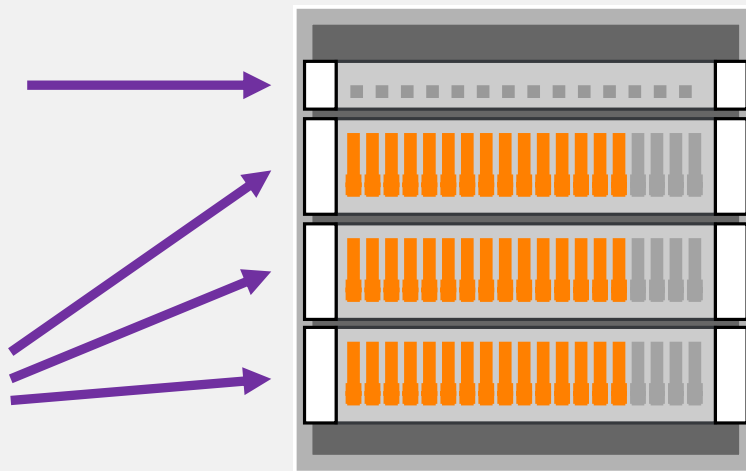
BIDI GROWTH SO FAR

BIDI BATTERY

A BIDI battery is the smallest unit of deployment within a PoP. Built with commodity hardware.

Layer-3 switch, for
aggregation and high
availability.

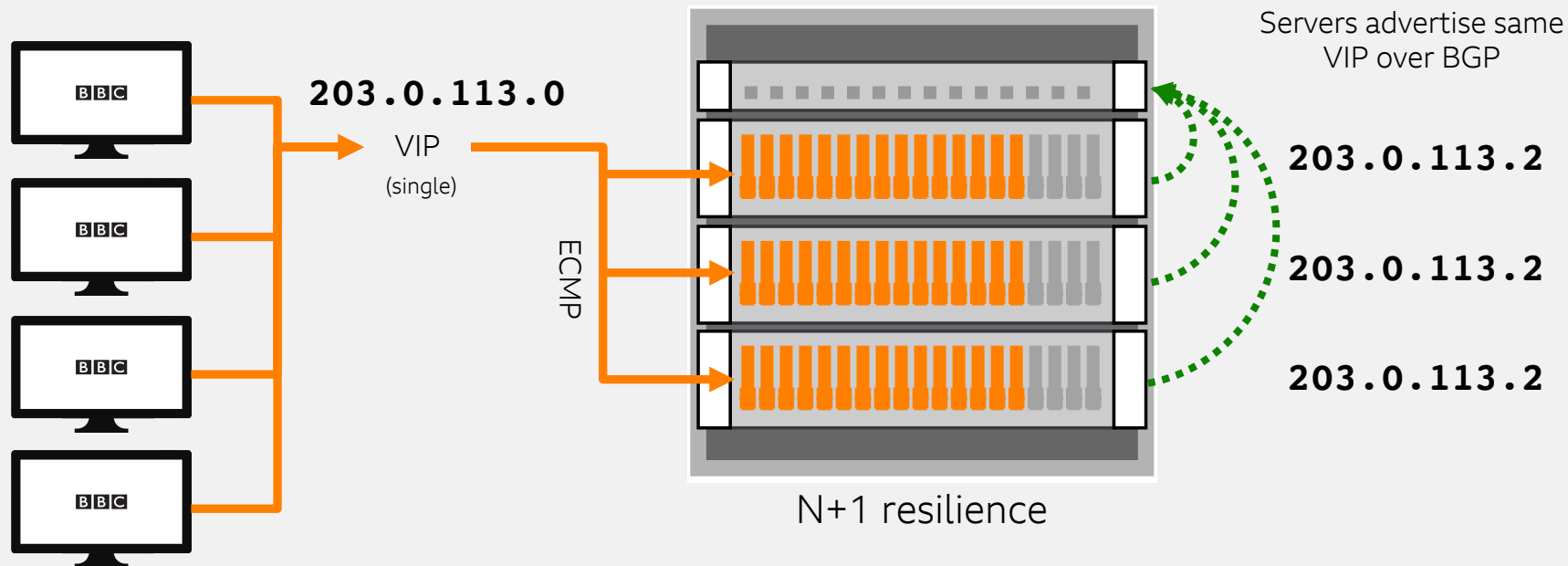
2 to 4 caching cells
(commodity servers)



BIDI battery

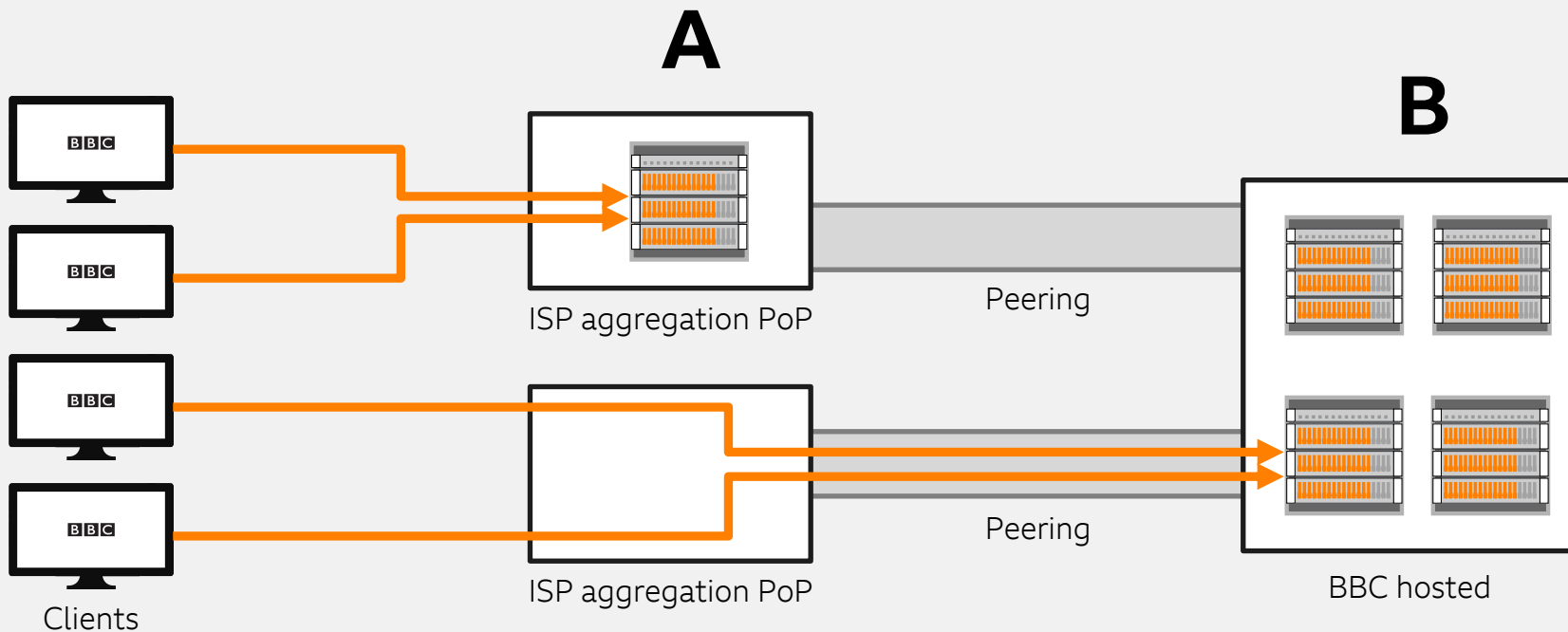
BIDI GROWTH SO FAR

BIDI BATTERY: HIGH AVAILABILITY



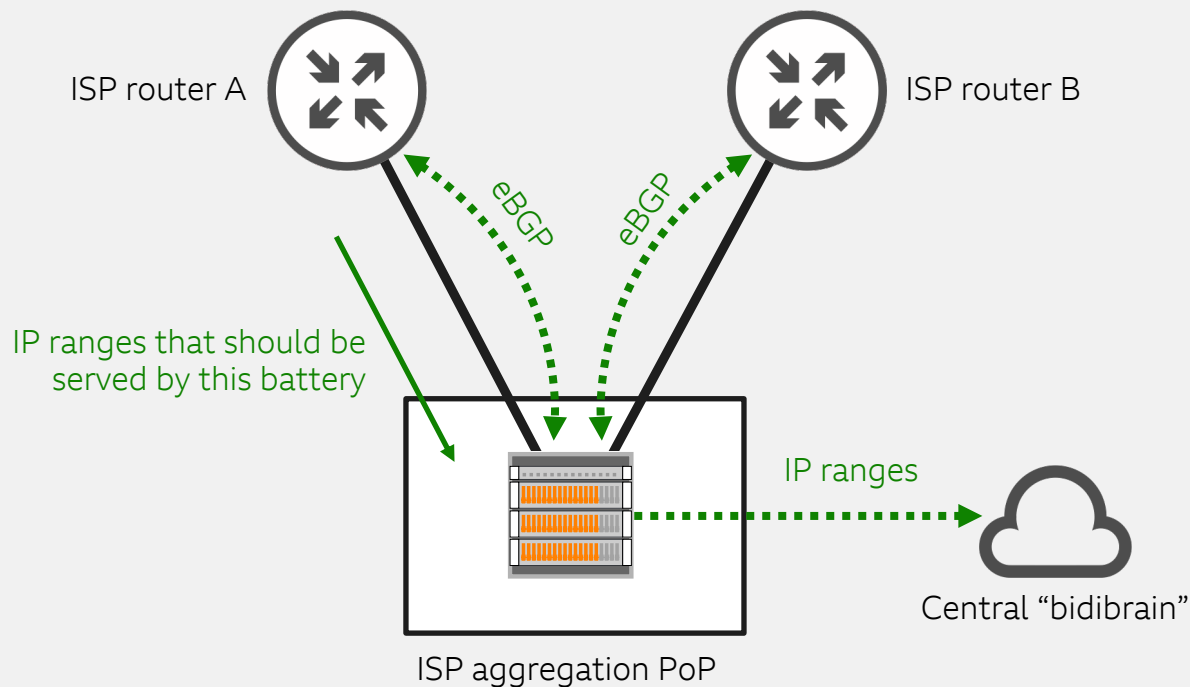
BIDI GROWTH SO FAR

BIDI BATTERY: DEPLOYMENT OPTIONS



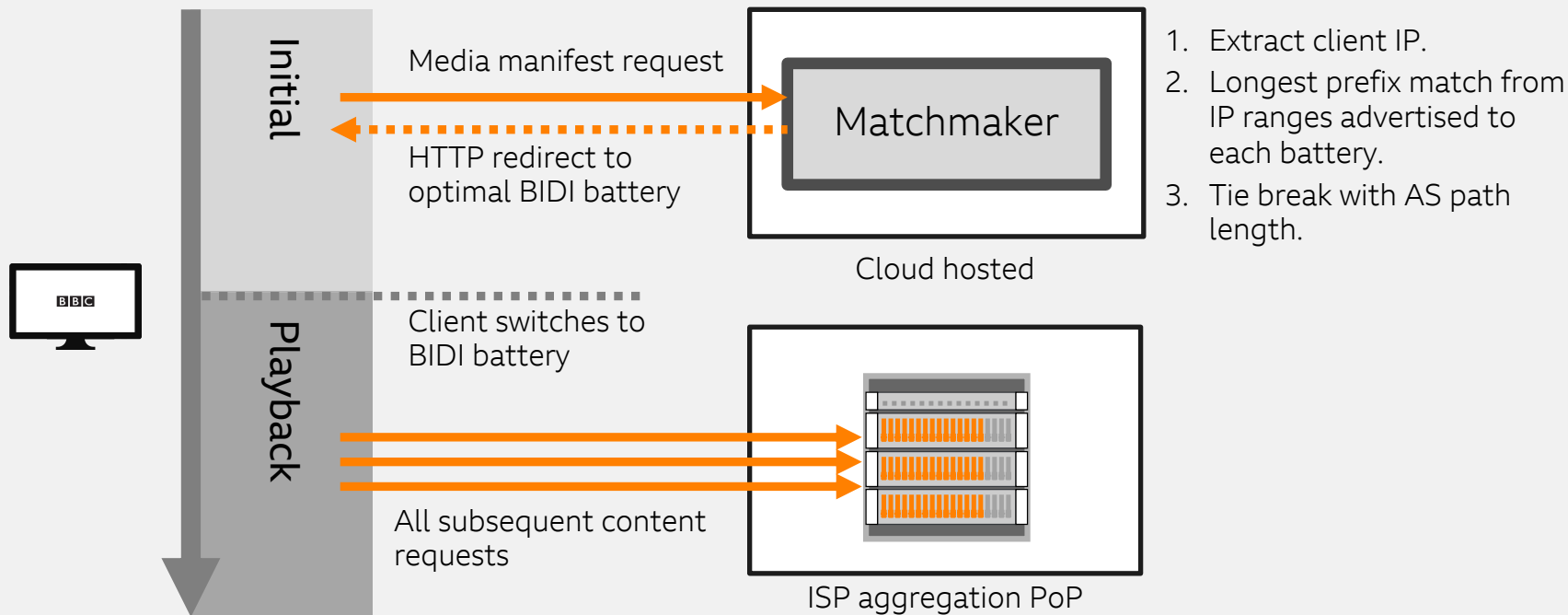
BIDI GROWTH SO FAR

BIDI BATTERY: ROUTE LEARNING – WHICH CLIENTS TO SERVE



BIDI GROWTH SO FAR

GETTING CLIENTS TO A BIDI BATTERY



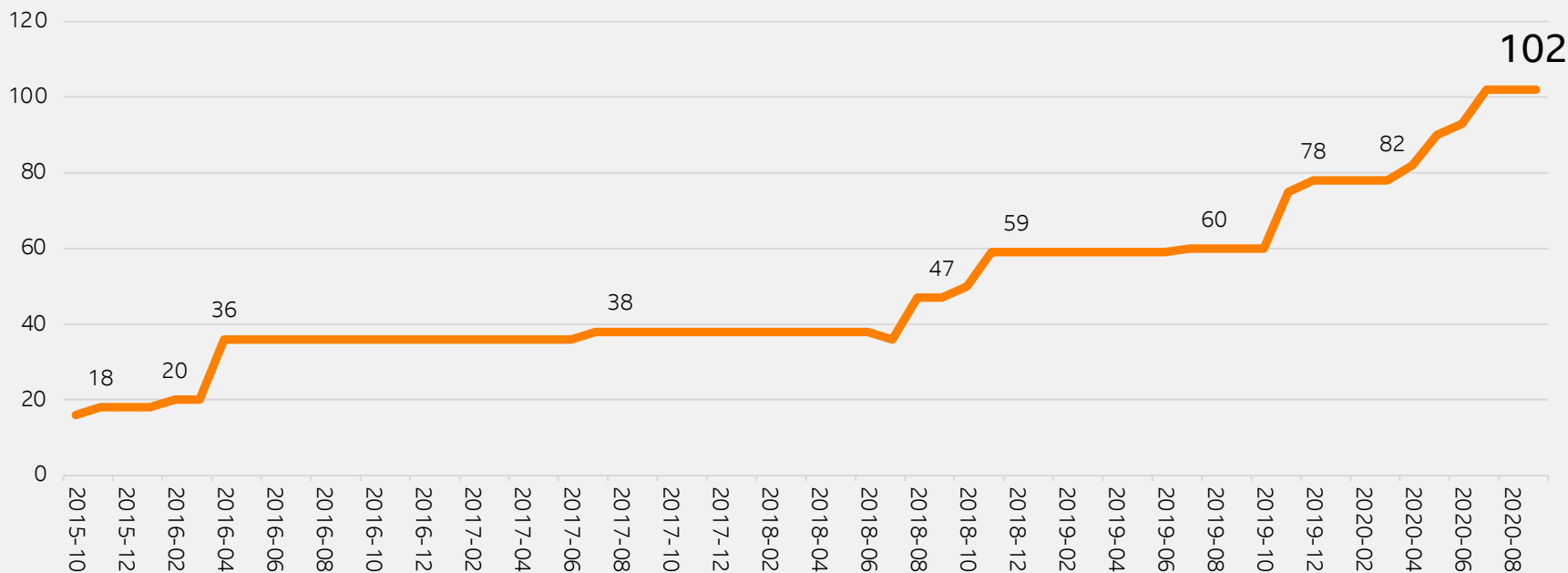
BIDI GROWTH SO FAR

TIMELINE: SCALE OUT



BIDI GROWTH SO FAR

BIDI CACHING CELLS (SERVERS) IN PRODUCTION



BIDI GROWTH SO FAR

BIDI ESTATE SIZE (NOVEMBER 2020)

102 caching servers

26 switches

24 datacentres

3 ISPs



BIDI GROWTH SO FAR

PEERING LOCATIONS

Public peering:

- LINX LON1 + LON2.
- IXManchester.
- IXScotland (from Manchester).
- LONAP.

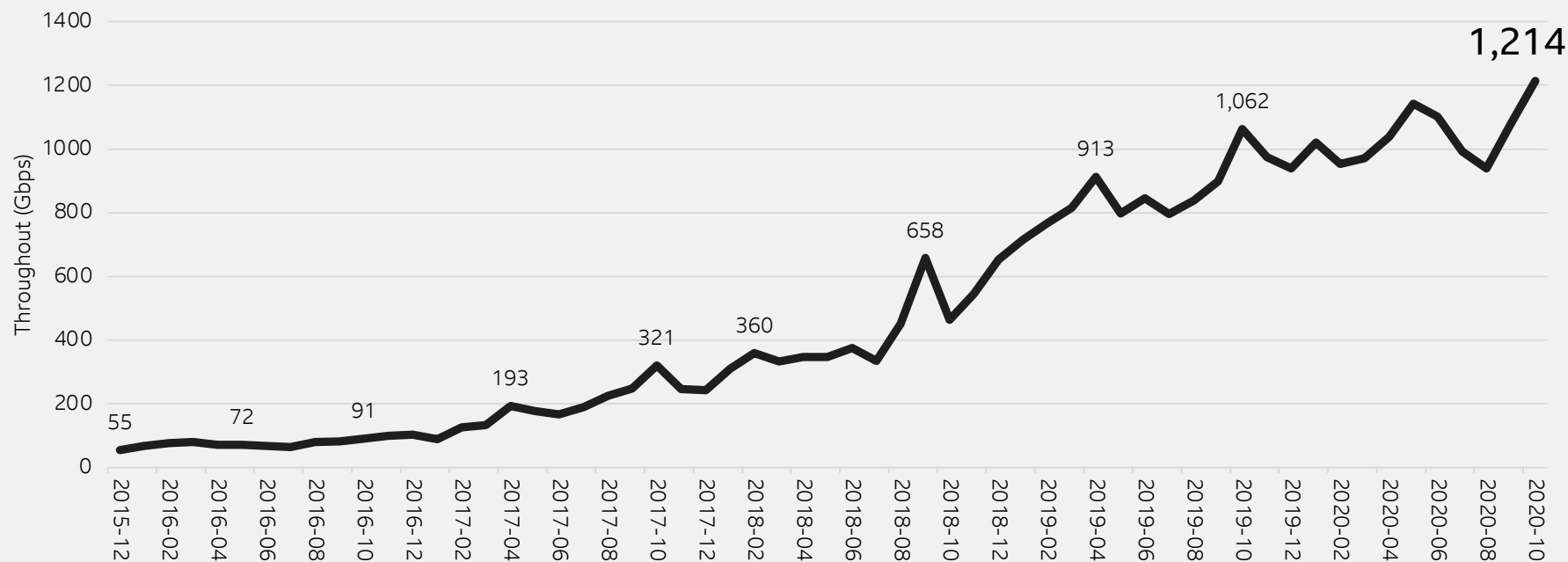
Private peering:

1. Telehouse Docklands West, London.
2. Digital Realty Sovereign House, London.
3. Equinix MA1, Manchester.
4. Equinix LD7, Slough (coming soon).



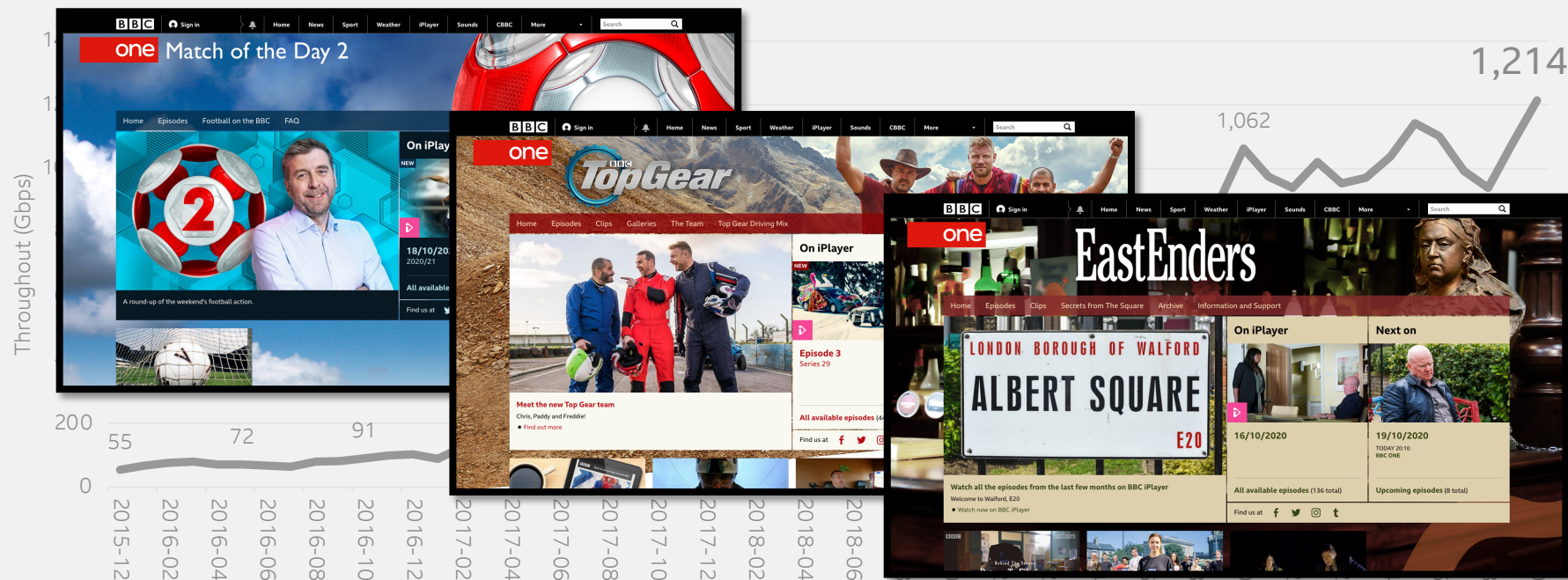
BIDI GROWTH SO FAR

BIDI MONTHLY TRAFFIC PEAK

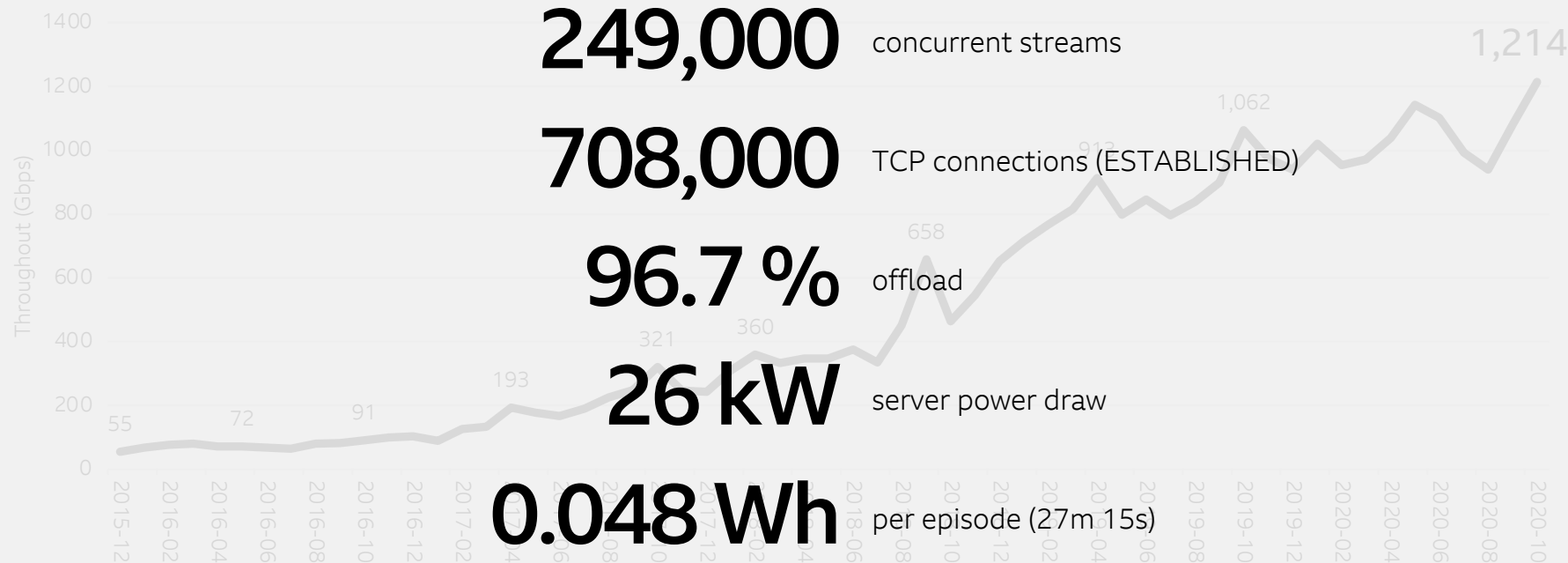


BIDI GROWTH SO FAR

BIDI MONTHLY TRAFFIC PEAK

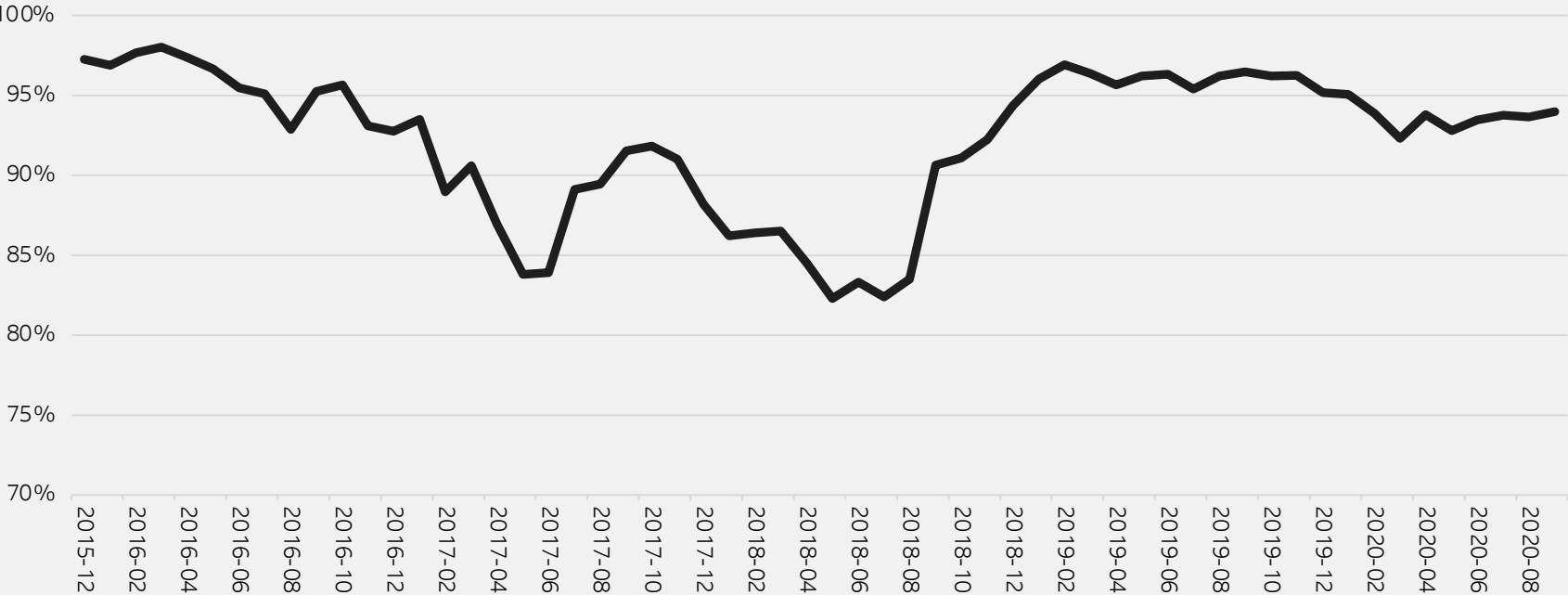


BIDI GROWTH SO FAR

BIDI MONTHLY TRAFFIC PEAK

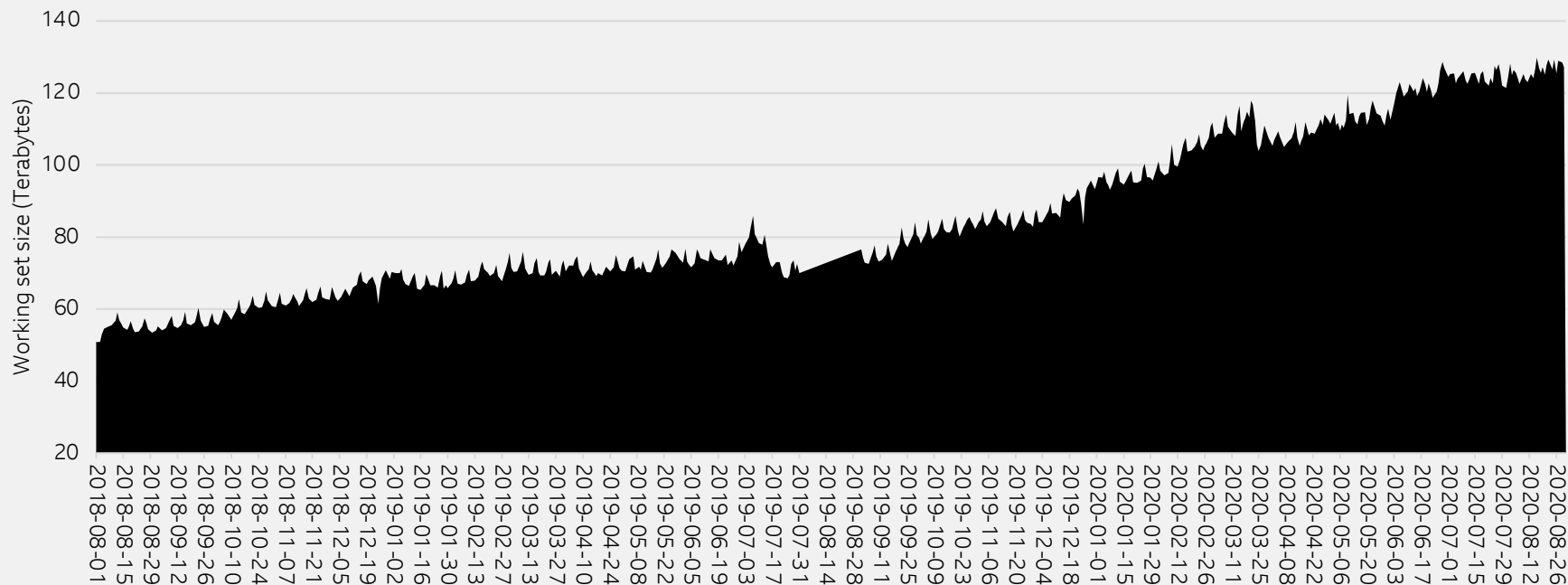
BIDI GROWTH SO FAR

BIDI OFFLOAD



BIDI GROWTH SO FAR

VOD 24 HOUR WORKING SET SIZE



BIDI GROWTH SO FAR

CACHE HARDWARE EVOLUTION

**Processor****RAM****Storage****Network****4** Gbps**20 cores**Dual socket
2.20 GHz / 190 W**128 GiB****1.3 TB SATA**6 x 480 GB SSDs
1 x HW RAID 1+0**20 Gbps**

2 x 10 G NICs

20 Gbps**36 cores**Dual socket
2.30 GHz / 280 W**256 GiB****19.2 TB SATA**10 x 1.92 TB SSDs
10 x HW RAID 0**40 Gbps**

2 x dual 10 G NICs

>70 Gbps**36 cores**Dual socket
2.60 GHz / 300 W**384 GiB****80 TB NVMe**10 x 8 TB SSDs
JBOD**100 Gbps**

1 x 100 G NIC

BIDI GROWTH SO FAR

SSD PERFORMANCE – LEARNING THE HARD WAY

- Magic controller compression
- Pesky RAID controllers
- Performance cliff edges
- Fast burnout
- Write-back caching
- High latencies / TRIM.
- Disappearing acts.

Testing disks: Lessons from our odyssey selecting replacement SSDs:

<https://www.bbc.co.uk/blogs/internet/entries/ce3eff16-228f-49a0-8d4e-d1a013e4895f>



BIDI GROWTH SO FAR

CACHE SOFTWARE EVOLUTION

**4** Gbps**Web server**

Varnish

BGP

ExaBGP

Services

Java

20 Gbpshttpd (front)
nginx
httpd (back)

ExaBGP

Node.js

>70 GbpsOpenResty
(nginx + LuaJIT)

BIRD

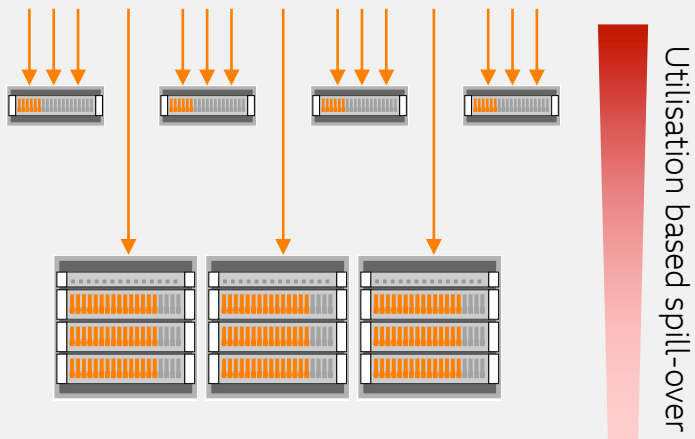
Python

FUTURE PLANS

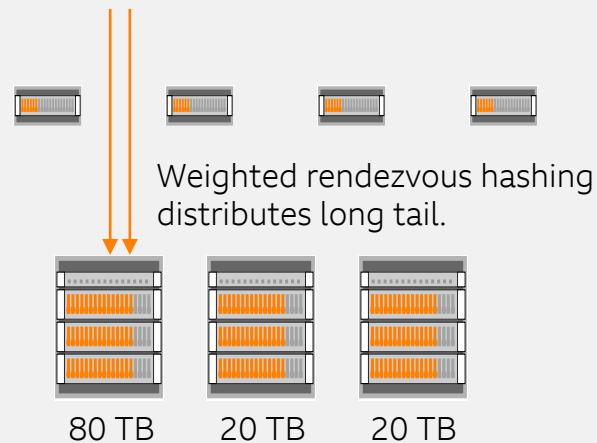
FUTURE PLANS

SMARTSHARD**Popular programmes/streams**

Served by all batteries, prioritising the edge.

**Unpopular programmes/streams**

Served by reduced number of core batteries.



FUTURE PLANS

IPV6 AND ISP PORTAL

IPv6

Still to do:

- IPv6 route ingest.
- Dual stack addressing.
- Actually doing it.

ISP portal

Web portal making BIDI information available for embedded ISPs:

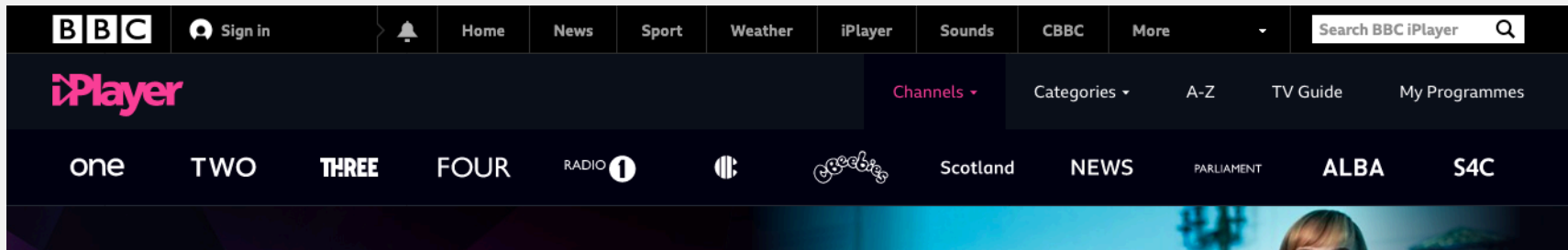
- Monitoring.
- Emergency + scheduled traffic drains.
- Route inspector.

FUTURE PLANS

LINEAR STREAMS

Onboard linear streams

- Next largest content set.
- IP distribution expected to increase.
- Less predictable traffic patterns and a high peak/mean ratio.



FUTURE PLANS

MBIDI

Next generation caching cell,
targeting deeper levels of the
network:

- Standalone operation.
- Smaller unit size.
- Increased power efficiency.

Single socket CPU

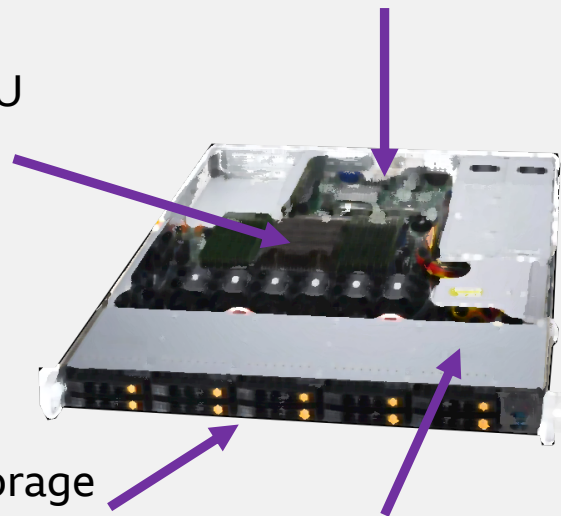
96 (logical) cores,
128 x PCIe 4 lanes.

Dual-port 100 GbE

91.8 TB cache storage

6 x 15.3 TB NVMe, PCIe 4.

Smaller 1U chassis



FUTURE PLANS

NEW SLOUGH POP

New BIDI PoP at Equinix LD7, Slough:

- Network infrastructure ready.
- On-premises origin cache operational.
- BIDI planned operational in March 2021.





Design + Engineering