Methods for Optimizing Data Capacity

Dr. Yair Shapira
yair@divinetworks.com

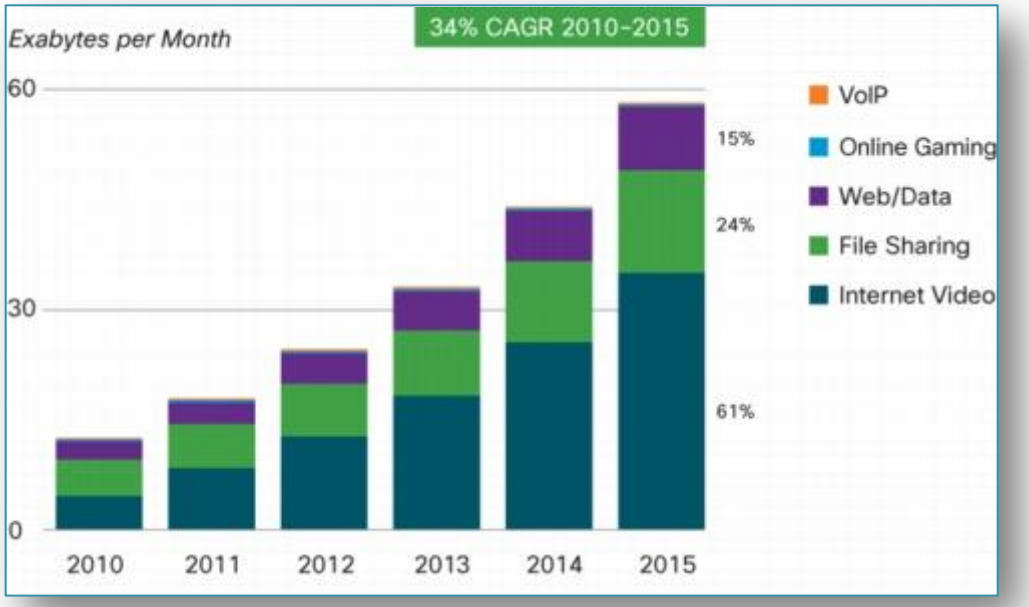
Agenda

- Methods for optimizing data capacity
  - Transparent caching
  - CDN
  - Traffic shaping
  - Video compression
- New method - Virtual capacity
- Repetitiveness characteristics in the Internet
- Virtual capacity effects on specific applications
- Examples from production networks

Meeting demand for data

- Demand for data is ever growing
  - Worldwide – 34% CAGR
  - LATAM – 48% CAGR
  - ME & Africa – 52% CAGR
  - Mobile – 92% CAGR
- ARPU stays flat
- The challenge:

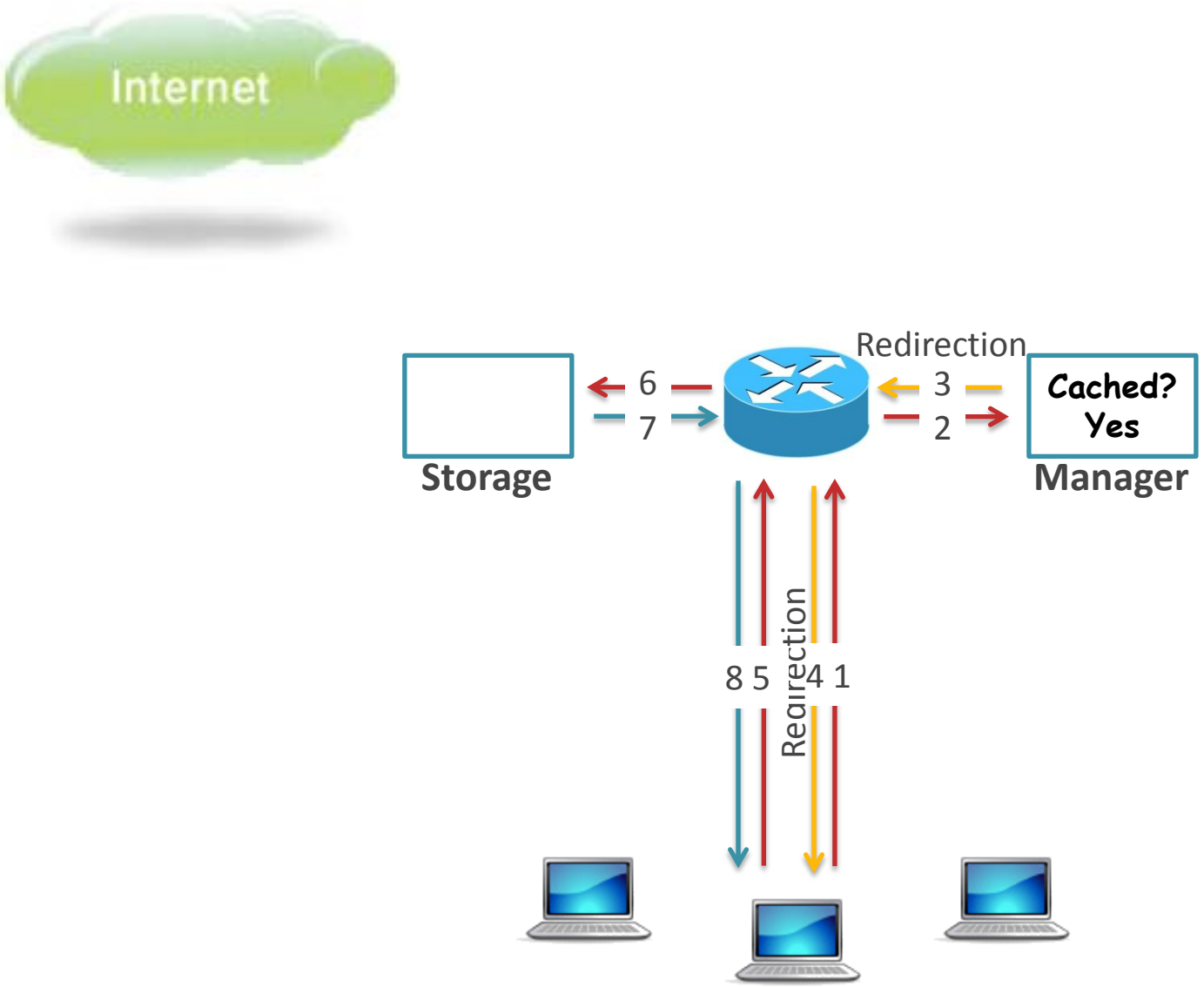
Economical Bandwidth Expansion



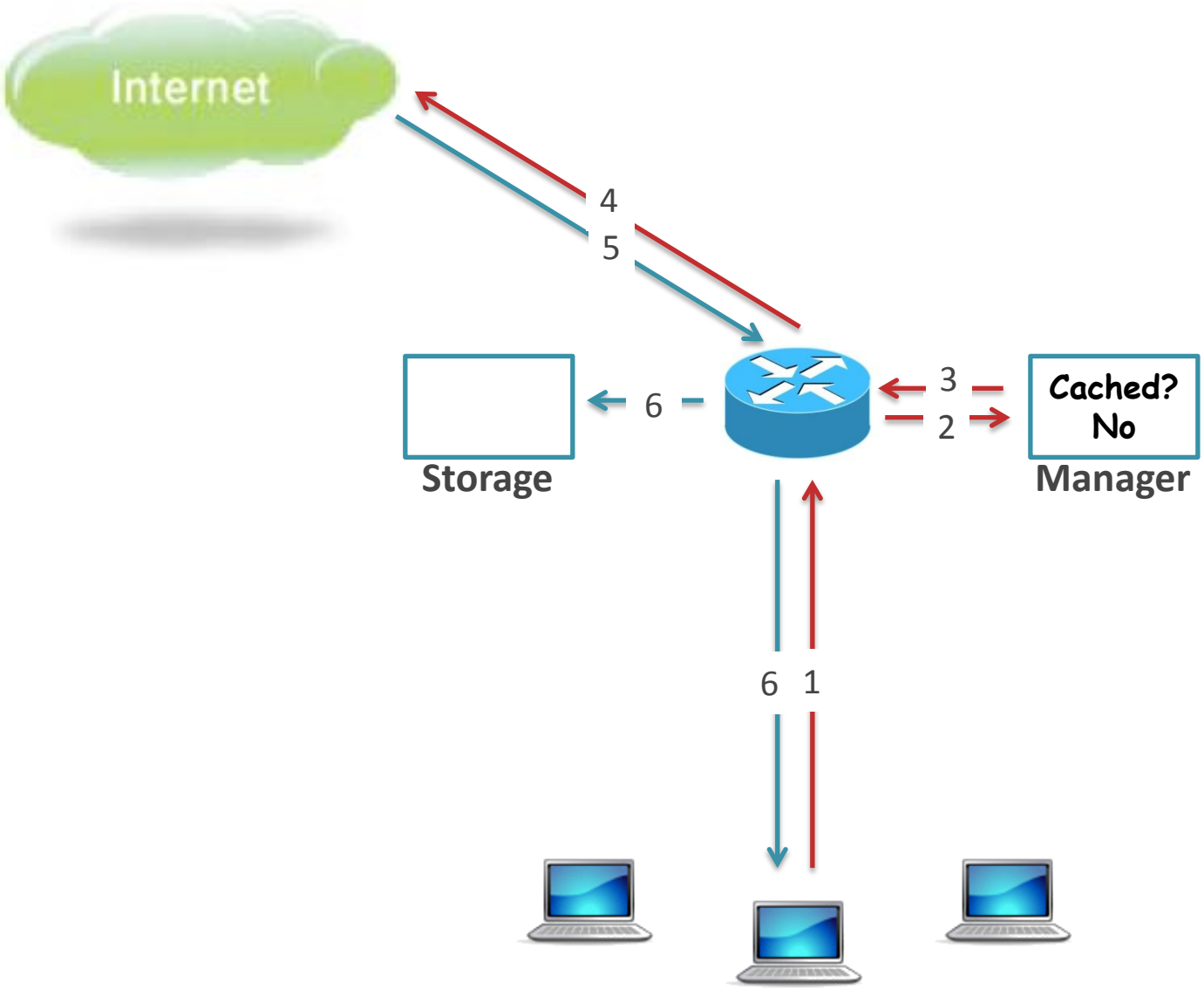
DiViNetworks

PRESENT METHODS FOR OPTIMIZING DATA CAPACITY

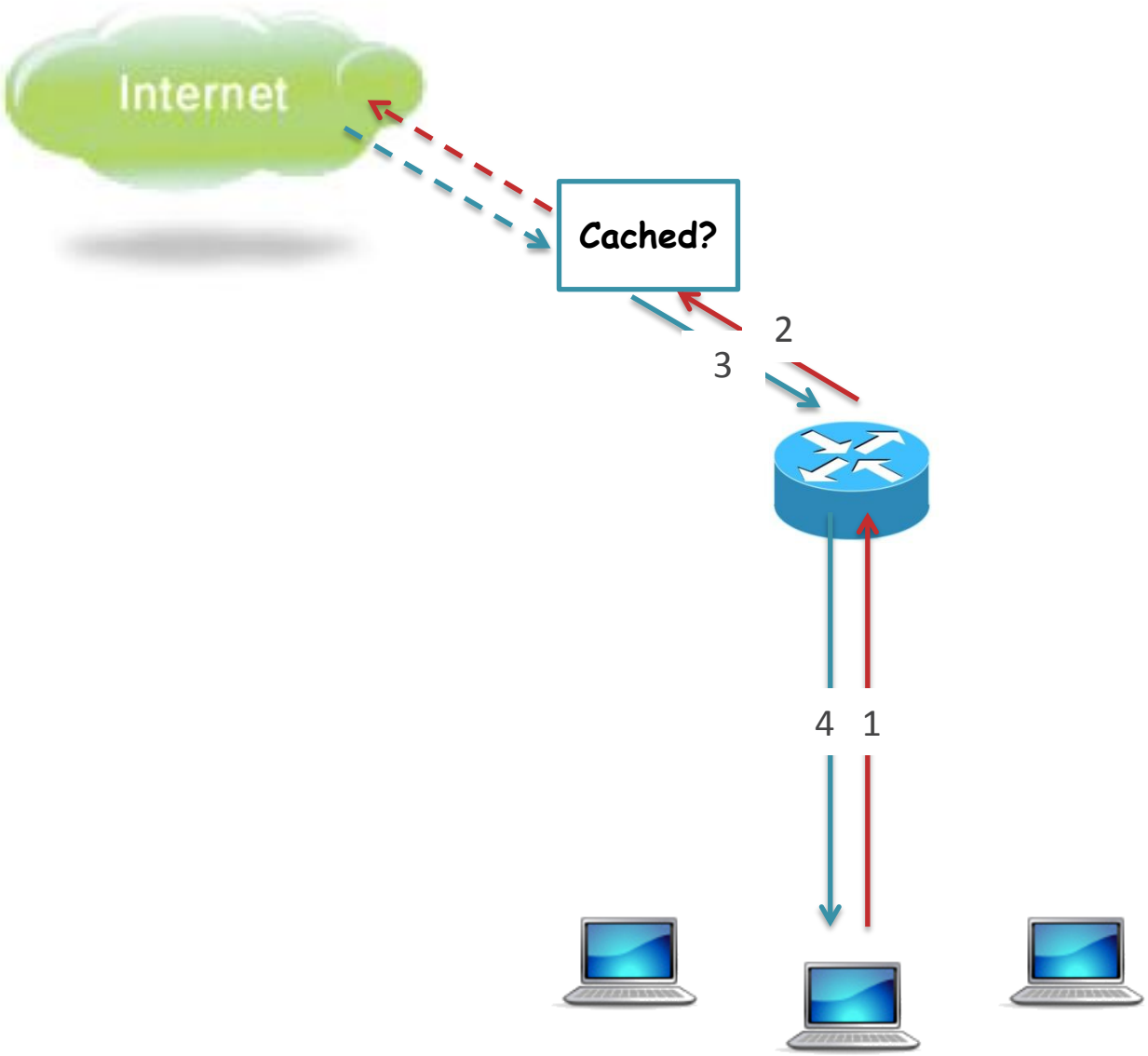
Transparent Cache – offline



Transparent Cache – offline



Transparent Cache – inline



Caching analysis from tier-1 operator

Total Traffic	1100 Gbps Peak Traffic
Part Residential	66,00%
Part HTTP	58,00%
Part Cache relevant	40,00%
Part Cachable	30,00%
Cache Volume	**50,5296 Gbps**
PoPs	74
Avg Volume/PoP	0,68 Gbps
Core PoPs	12
Avg/Core PoP	4,21 Gbps

Transparent Cache

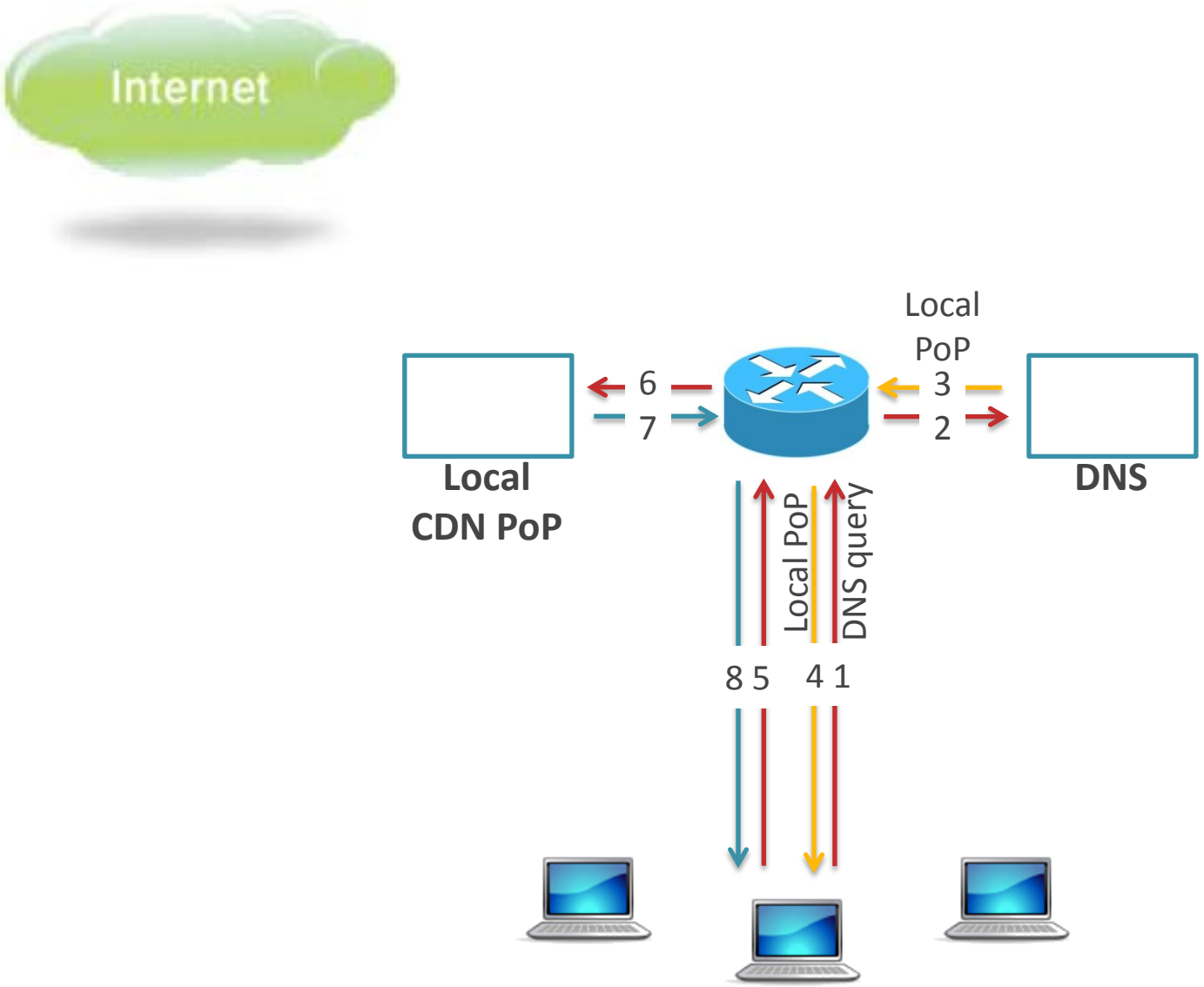
Pros

- Good QoE for supported applications and content
- Bandwidth savings
- Long track record

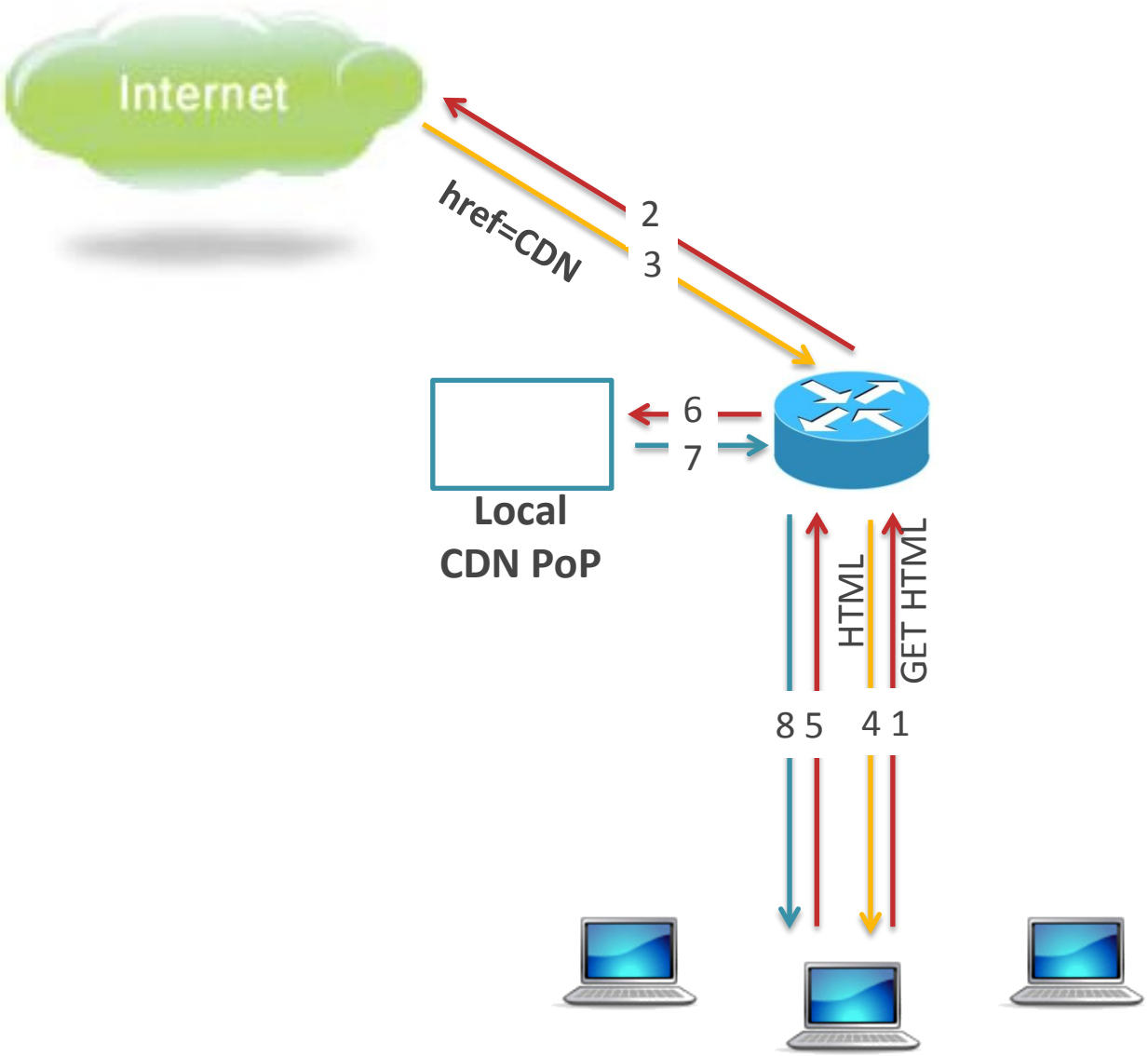
Cons

- Application & content specific
- Not effective below 0.5-1Gbps
- Constraints on routing
- High maintenance – Needs to follow Internet evolution (e.g. protocols, formats)
- Legal concerns

CDN – DNS-based



CDN – HTML-based



CDN

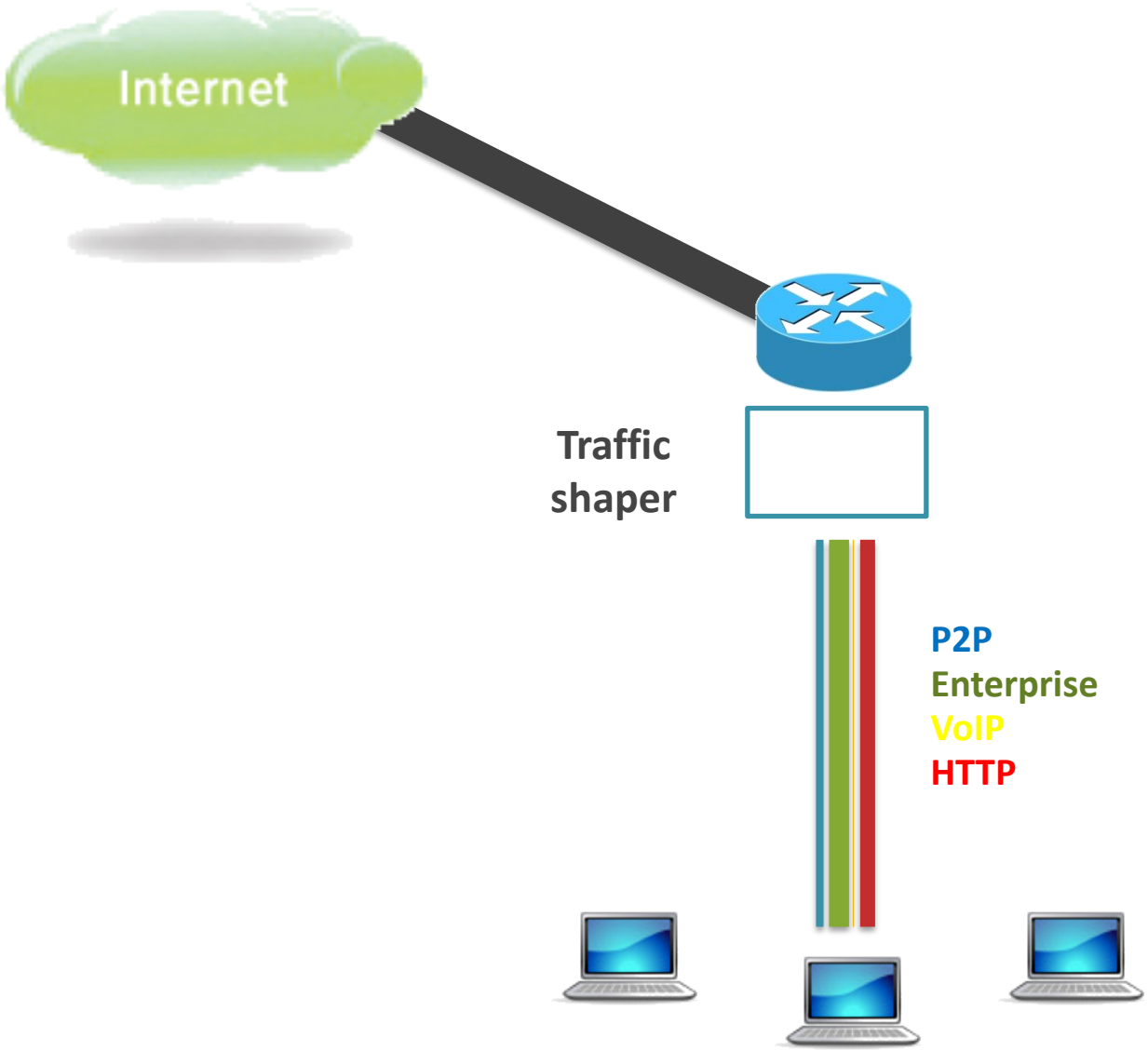
Pros

- Good QoE for managed content
- Bandwidth savings
- CDN's investment & maintenance

Cons

- Only for managed content
- CDNs won't locate PoPs where it's not economical for their customers
- Controlled by CDN, not by ISP

Traffic shaping



Traffic shaping

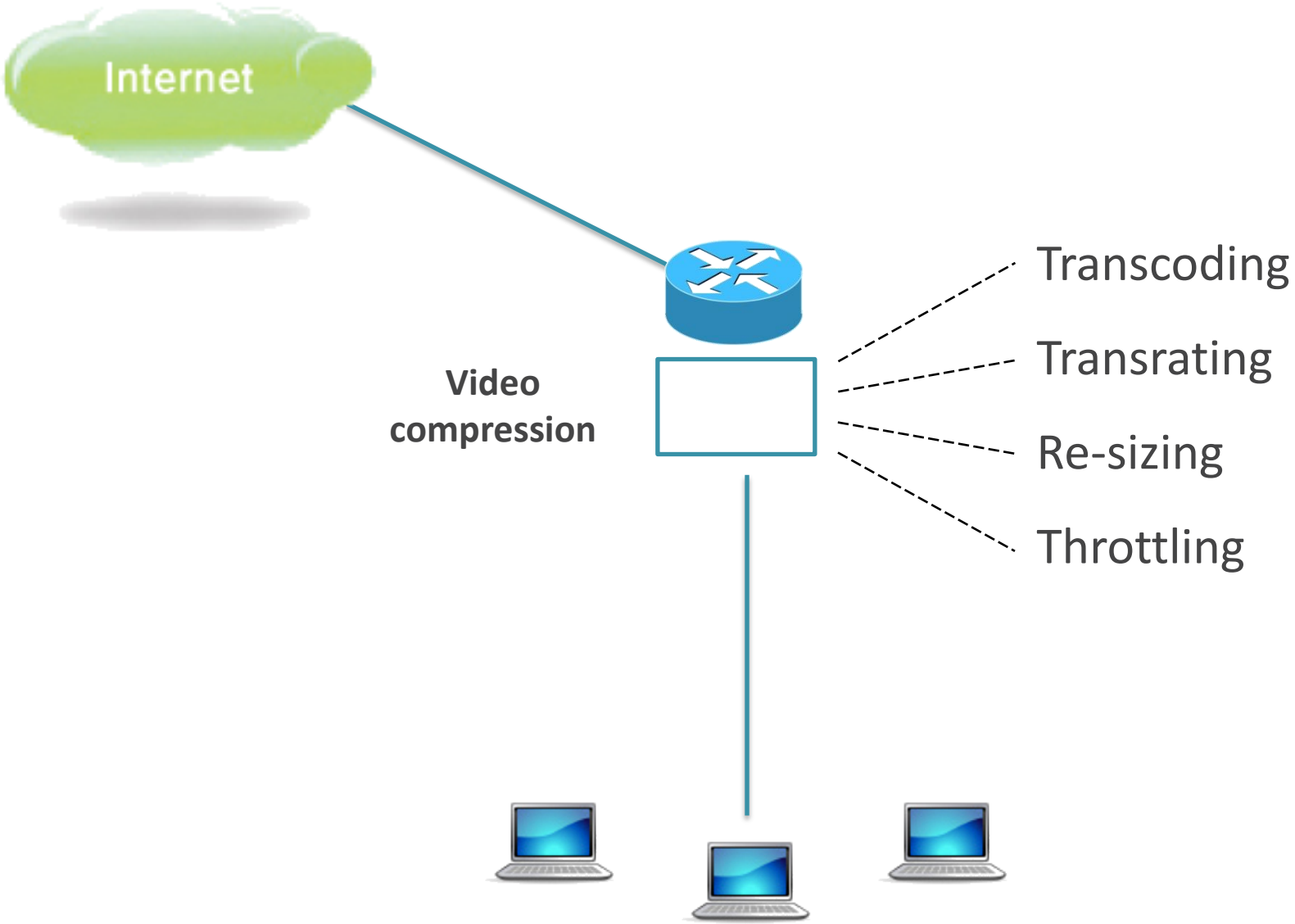
Pros

- Explicit control (non-statistical)
- Revenue generator

Cons

- Doesn’t generate traffic
- High maintenance
- Legally sensitive

Video compression



Video compression

Pros

- Limitless compression
- Selective video compression is not sensed

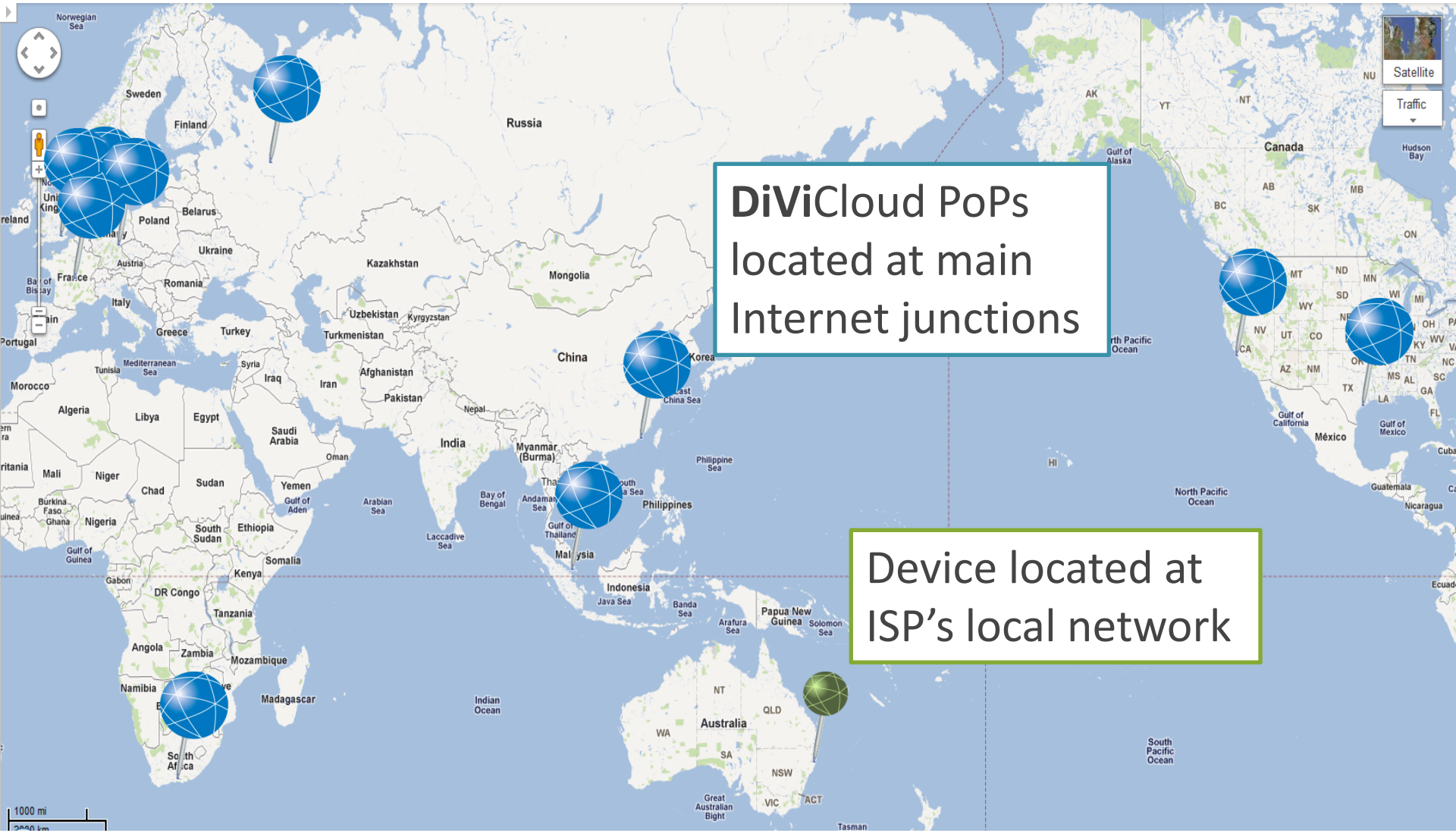
Cons

- Negative impact on QoE
- Requires high CPU power
- Redundant with ABR
- Redundant with server-based throttling

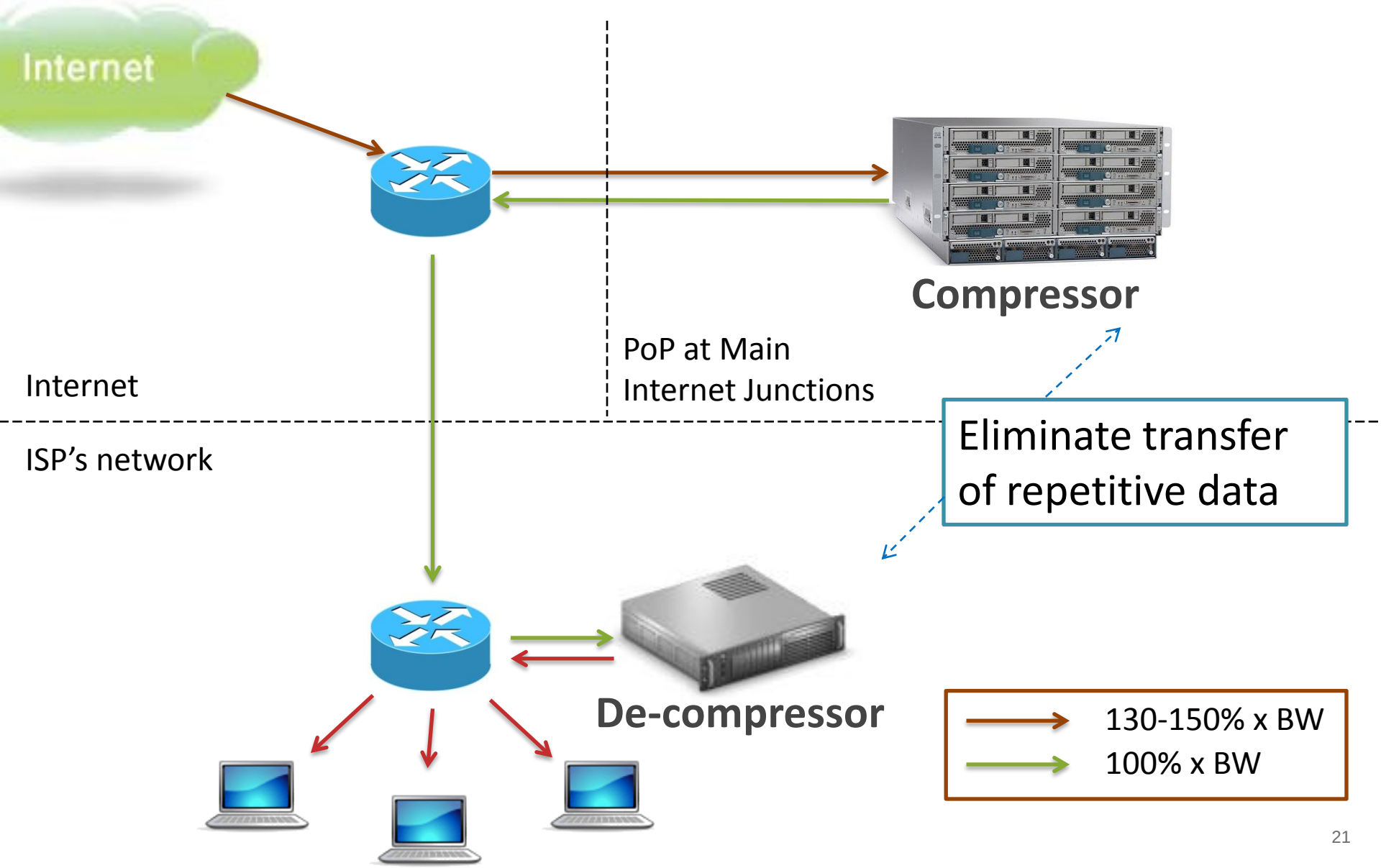
DiViNetworks

NEW METHOD FOR OPTIMIZING DATA CAPACITY – VIRTUAL CAPACITY

DiViCloud Network



Traffic Flow



Algorithm – learning

1001010101110
1010101010001
0100100111010
1011010101101
0101010110110

Response

Request

IP payload analyzed.
If pattern not identified
then learned.

1001010101110
1010101010001
0100100111010
1011010101101
0101010110110

≡ A1D

Compressor

De-compressor

Algorithm – identification

1001010101110
1010101010001
0100100111010
1011010101101
0101010110110

(2) Response

IP payload analyzed.
Pattern identified!

1001010101110
1010101010001
0100100111010

Start signing session
Next signatures B6U, GG9, K4W

Compressor

De-compressor

Algorithm – capacity generation

1001010101110
1010101010001
0100100111010
1011010101101
0101010110110

Response

Original

K4W

Signed

1001010101110
1010101010001
0100100111010
1011010101101
0101010110110

A1D

Compressor

De-compressor

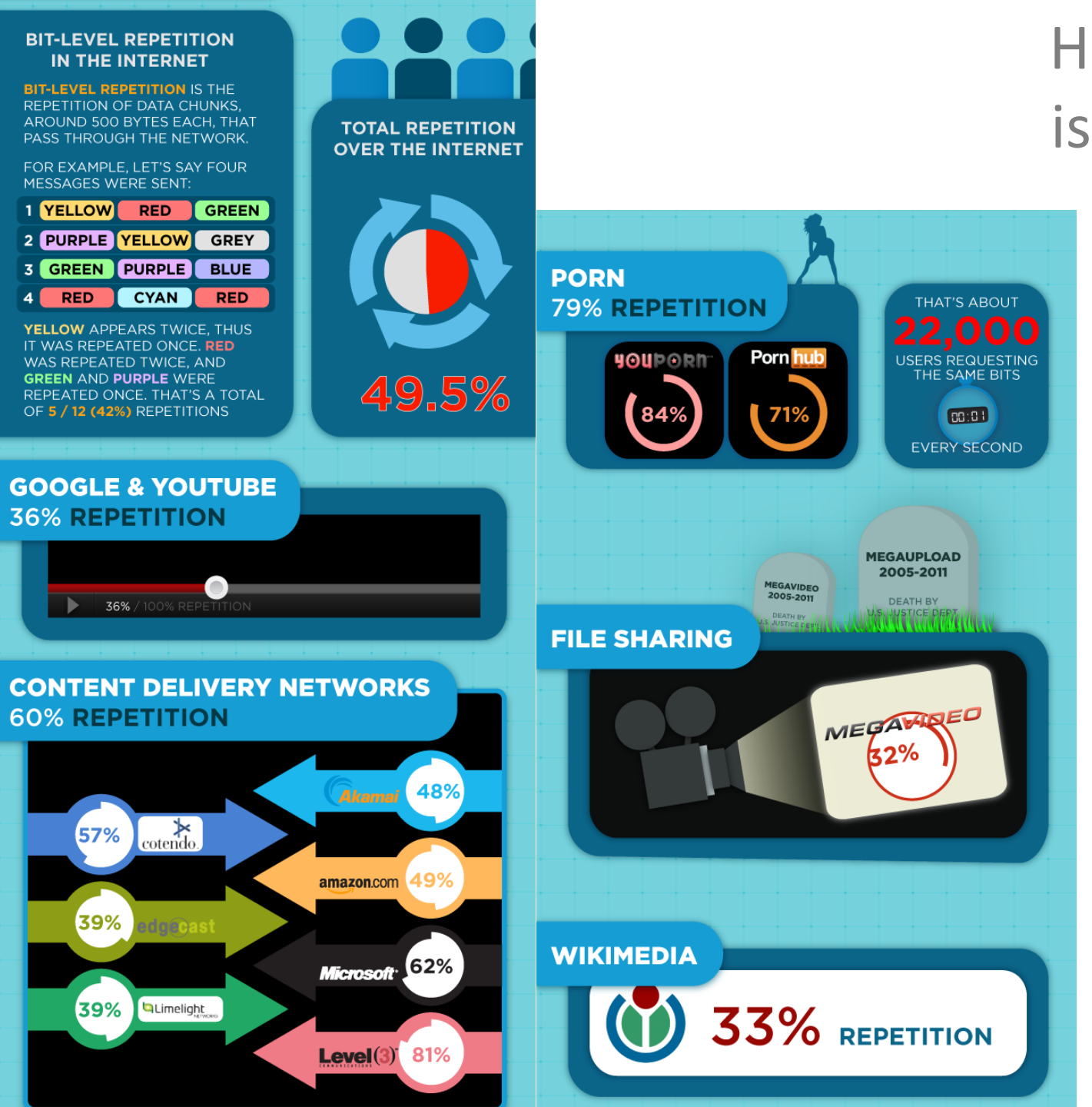
Algorithm Characteristics

- Operates at bit-stream level.
- Agnostic to protocol and content.
- No spoofing or termination applied.
- Network functions maintained.
- Future proof.
- Legally safe.

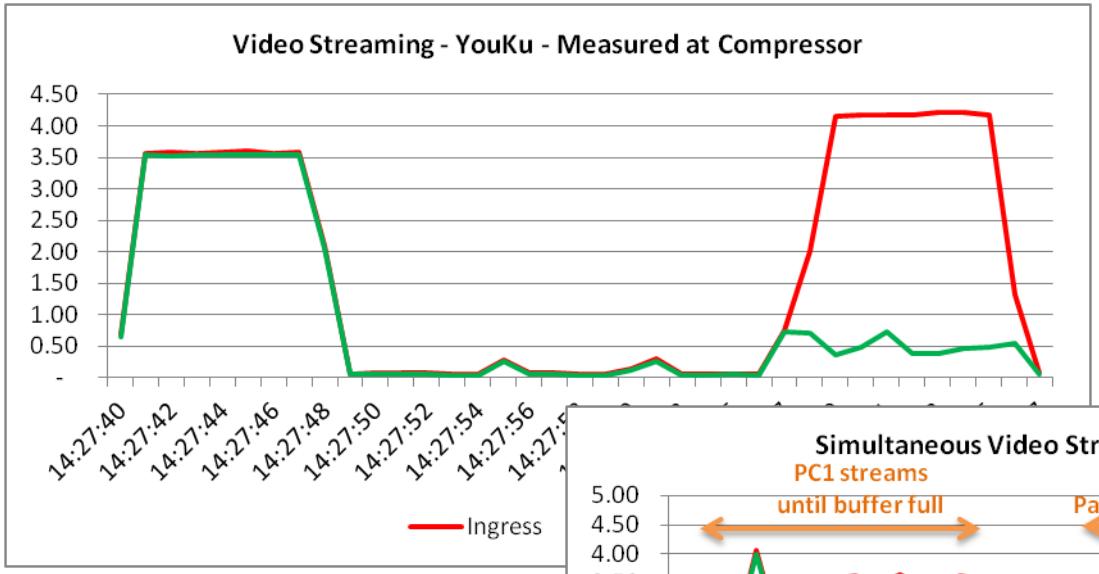
How Redundant is the Internet?

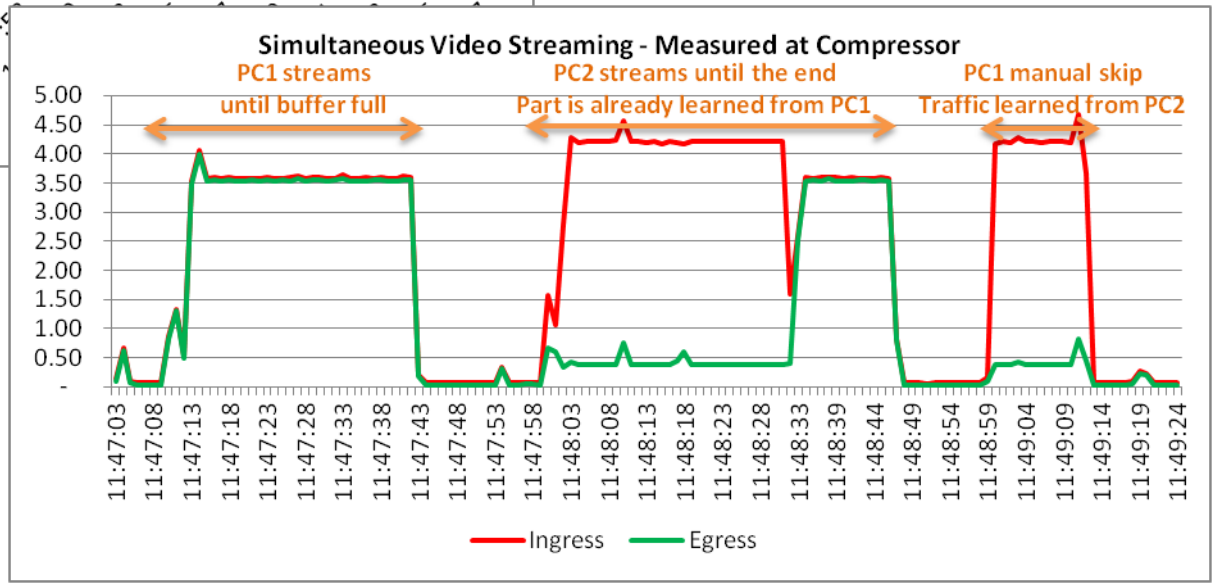
Chunk length for this analysis	*500 Bytes = 4,000 Bits*
Estimated number of clips in YouTube	*10,000,000,000 = 10^{10}*
Average files size in YouTube	*5MByte*
Chunks in average file in YouTube	*5 x 10^6 / 500 = 10^4*
Number of chunks in YouTube	*10^{10} x 10^4 = 10^{14}*
Estimated factor to cover the whole Internet	*1,000*
Number of chunks in the whole Internet	*10^{14} x 10^3 = 10^{17}*
Bits required to map all chunks in the Internet	*Log_2 10^{17} = 56 bits = 7 Byte*
Redundancy ratio	*500 Byte / 7 Byte = 71*

How Repetitive is the Internet?

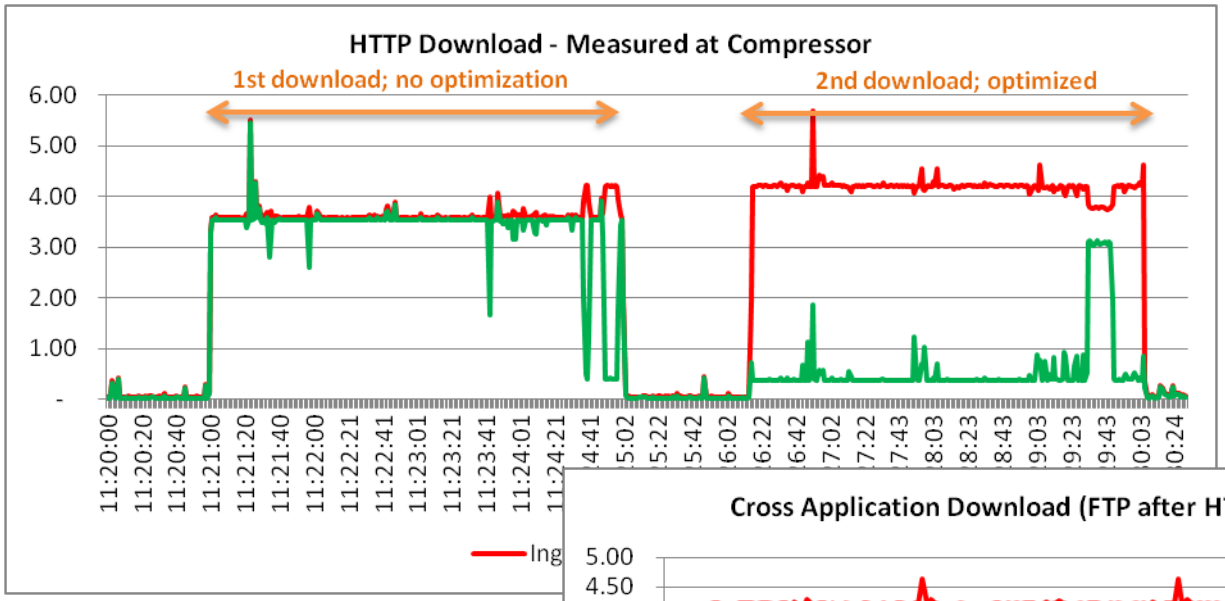


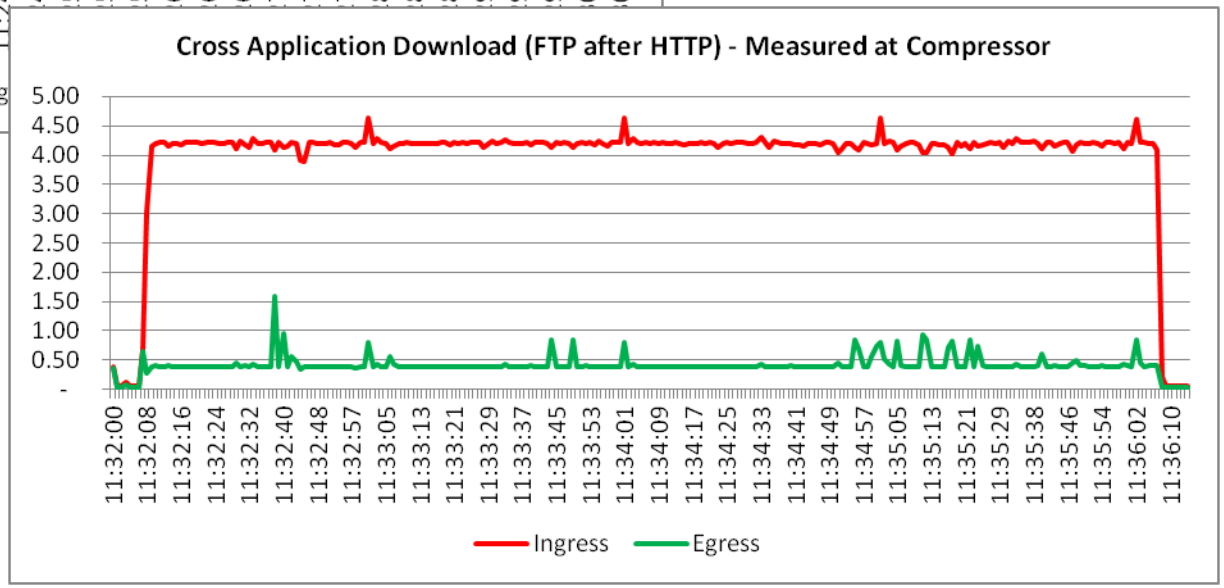
Video Progressive Download



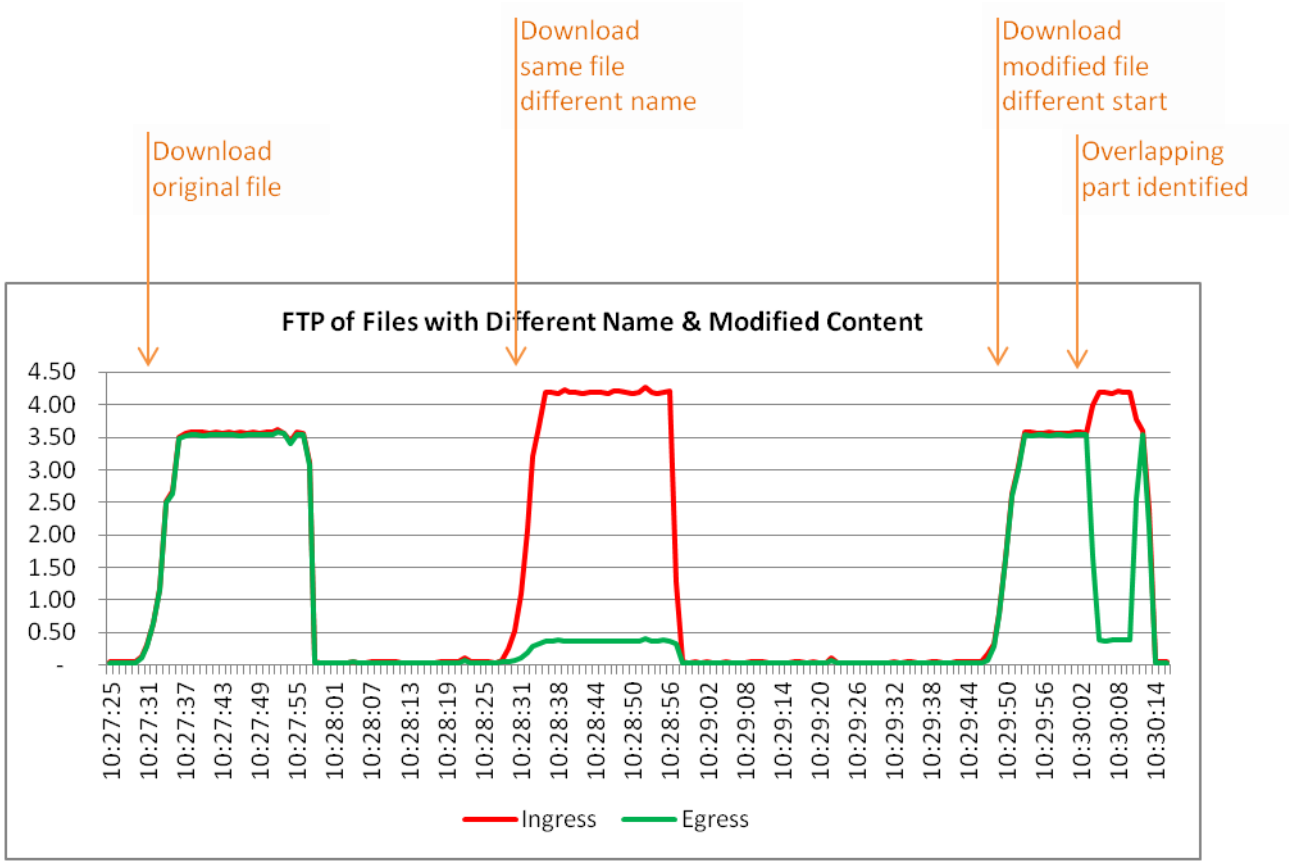


HTTP & FTP Download

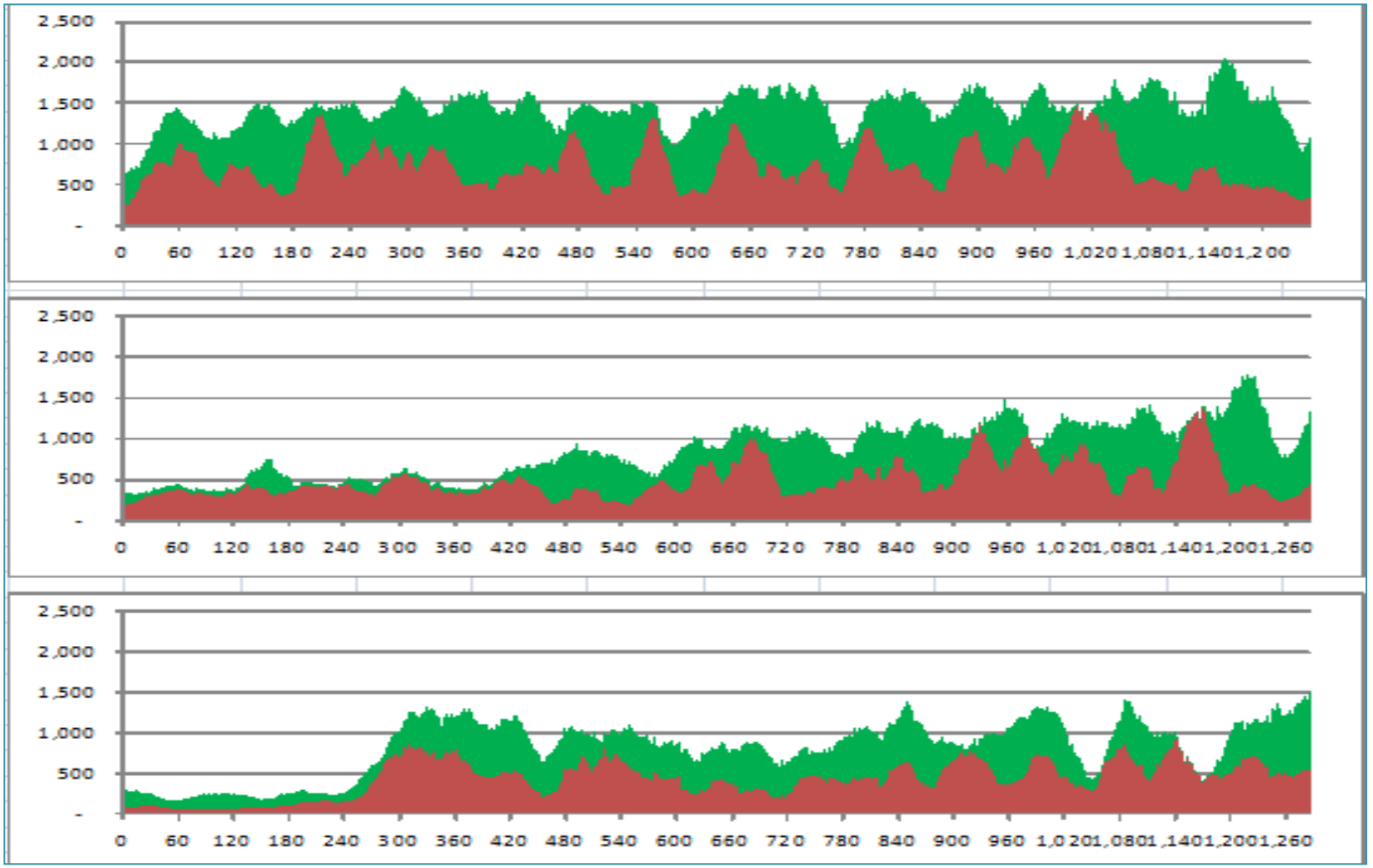




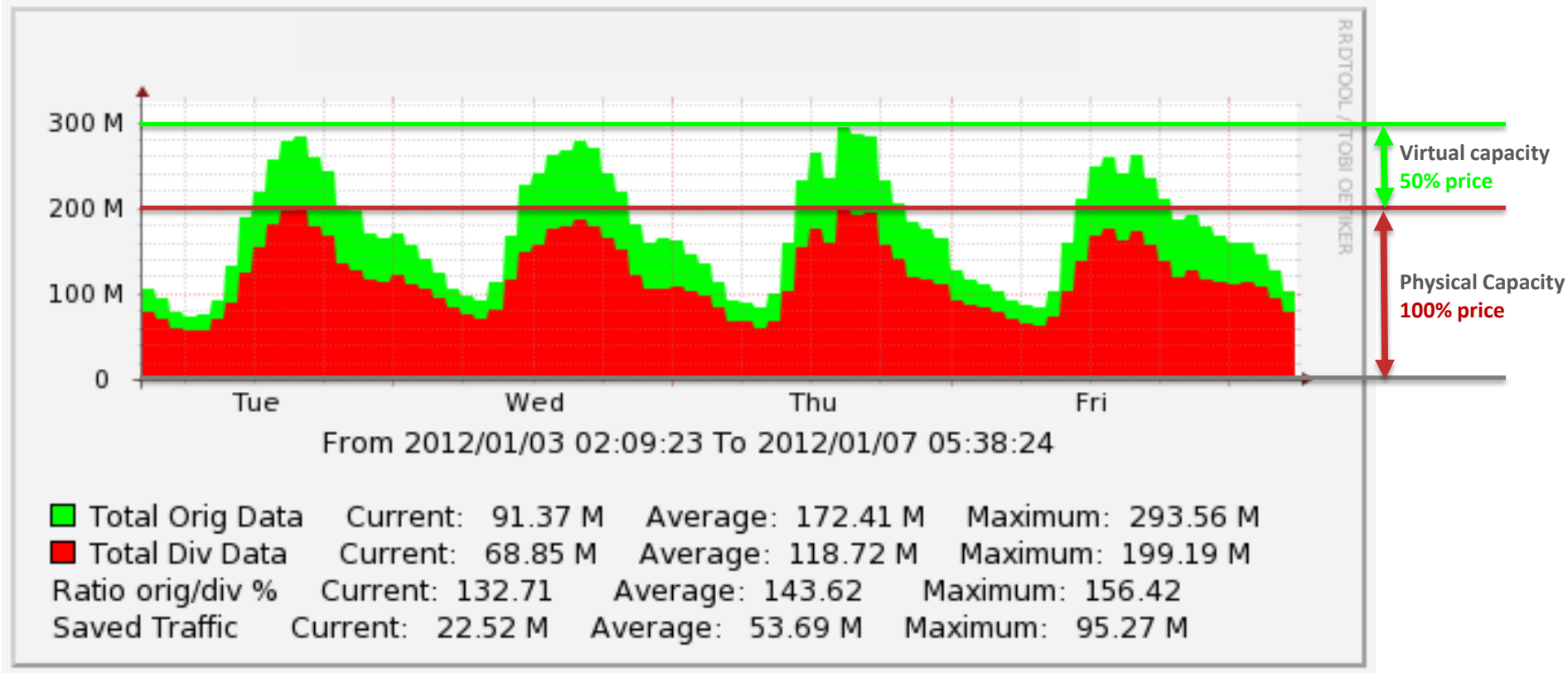
Partial Content Overlap



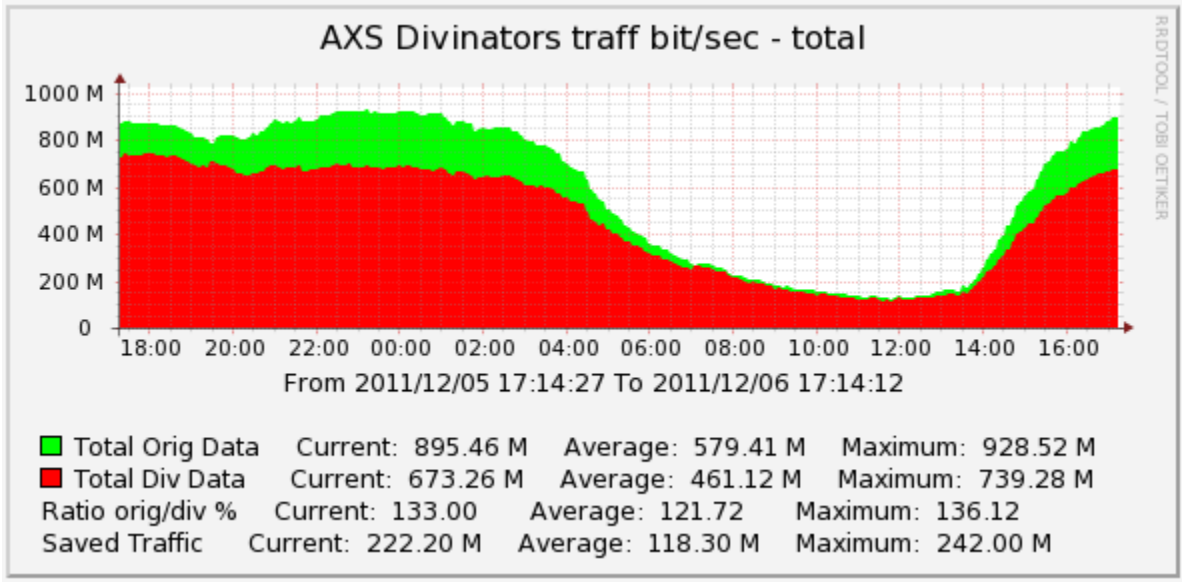
Netflix (ABR + Encryption)



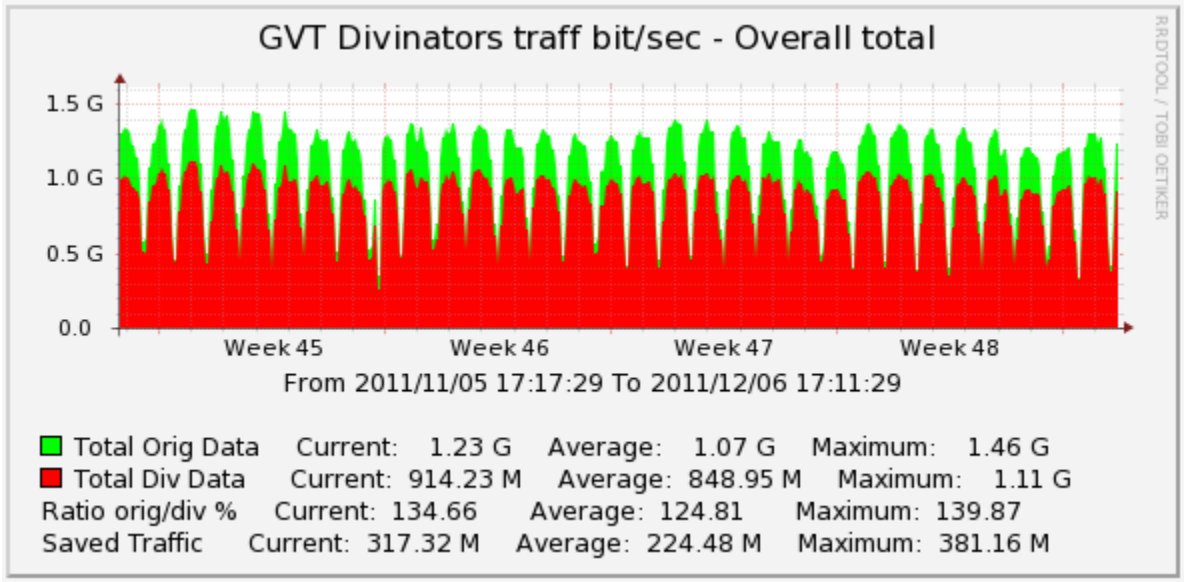
Case 1


300 M
200 M
100 M
0
Tue
Wed
Thu
Fri
From 2012/01/03 02:09:23 To 2012/01/07 05:38:24
Total Orig Data Current: 91.37 M Average: 172.41 M Maximum: 293.56 M
Total Div Data Current: 68.85 M Average: 118.72 M Maximum: 199.19 M
Ratio orig/div % Current: 132.71 Average: 143.62 Maximum: 156.42
Saved Traffic Current: 22.52 M Average: 53.69 M Maximum: 95.27 M
RRDTOOL / TOBI OETIKER
Virtual capacity
50% price
Physical Capacity
100% price

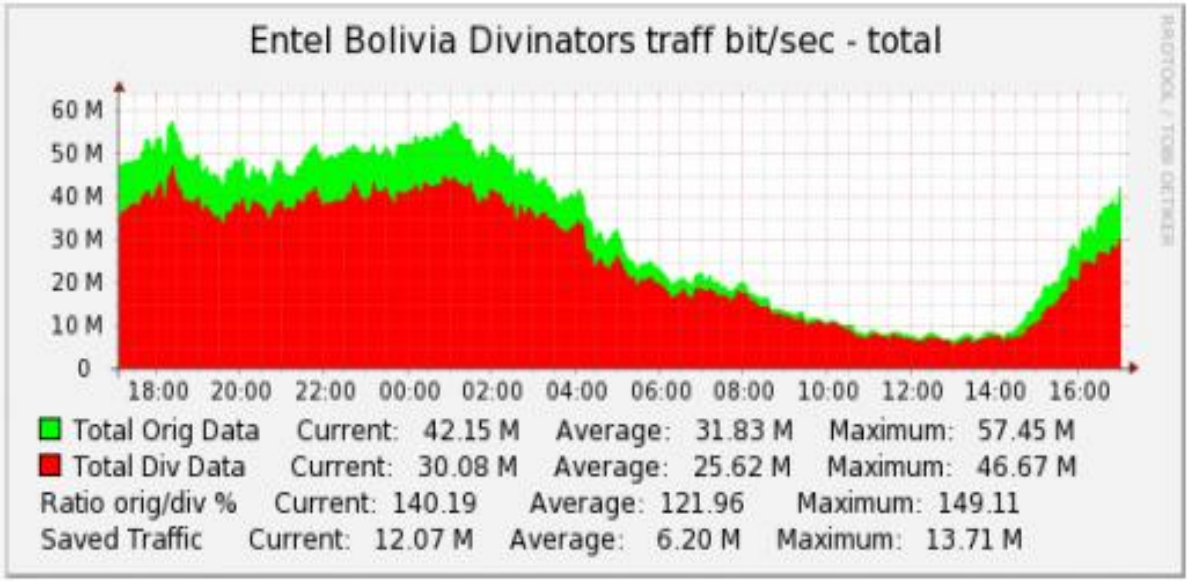
Case 2


AXS Divinators traff bit/sec - total
1000 M
800 M
600 M
400 M
200 M
0
18:00 20:00 22:00 00:00 02:00 04:00 06:00 08:00 10:00 12:00 14:00 16:00
From 2011/12/05 17:14:27 To 2011/12/06 17:14:12
Total Orig Data Current: 895.46 M Average: 579.41 M Maximum: 928.52 M
Total Div Data Current: 673.26 M Average: 461.12 M Maximum: 739.28 M
Ratio orig/div % Current: 133.00 Average: 121.72 Maximum: 136.12
Saved Traffic Current: 222.20 M Average: 118.30 M Maximum: 242.00 M
RRDTOOL / TOBI OETIKER

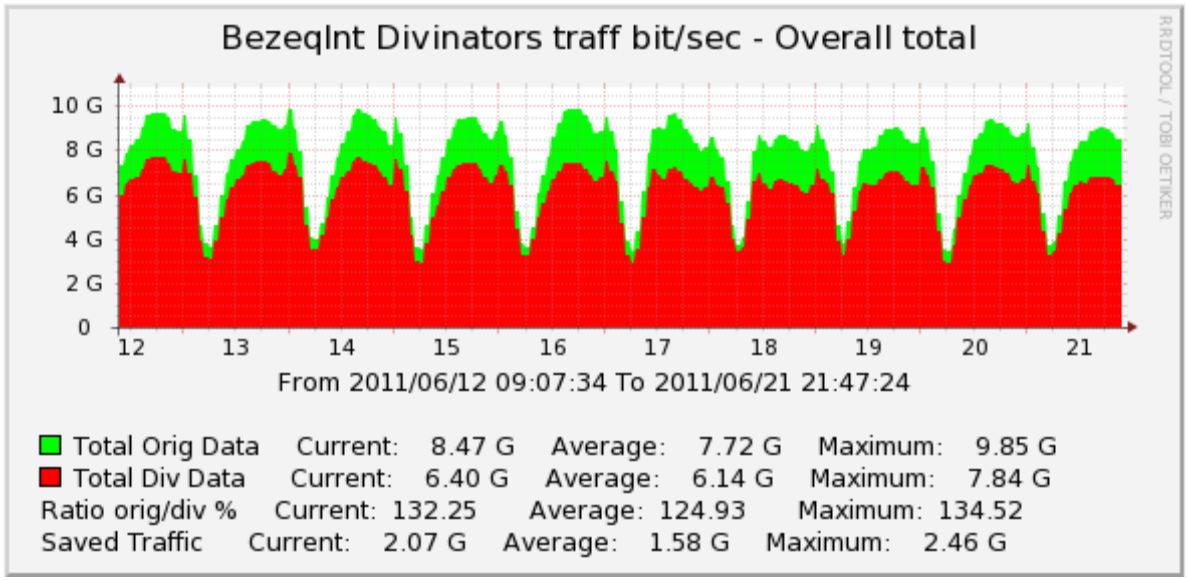
Case 3


GVT Divinators traff bit/sec - Overall total
1.5 G
1.0 G
0.5 G
0.0
Week 45
Week 46
Week 47
Week 48
From 2011/11/05 17:17:29 To 2011/12/06 17:11:29
Total Orig Data Current: 1.23 G Average: 1.07 G Maximum: 1.46 G
Total Div Data Current: 914.23 M Average: 848.95 M Maximum: 1.11 G
Ratio orig/div % Current: 134.66 Average: 124.81 Maximum: 139.87
Saved Traffic Current: 317.32 M Average: 224.48 M Maximum: 381.16 M
RRDTOOL / TOBI OETIKER

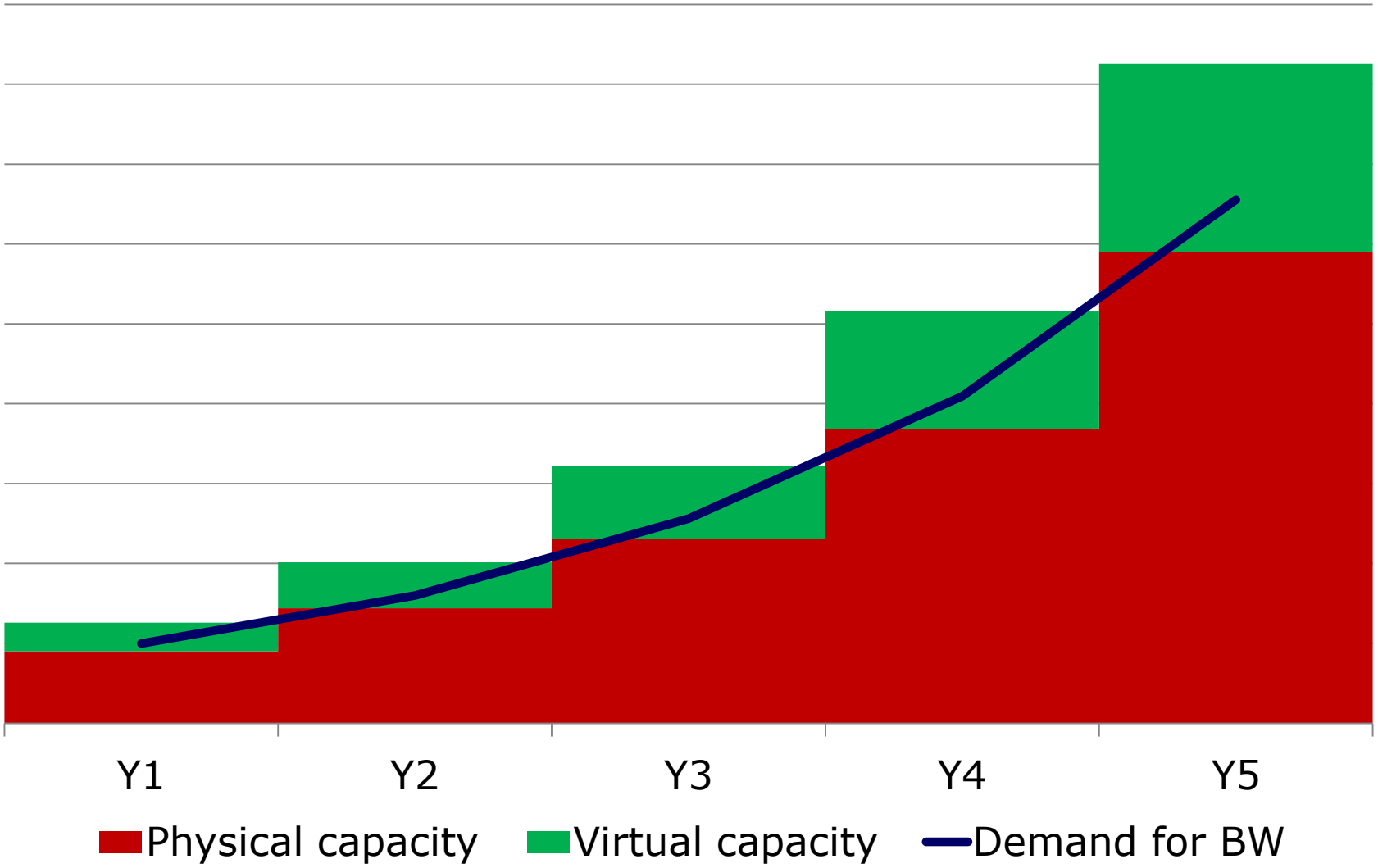
Case 4


Entel Bolivia Divinators traff bit/sec - total
60 M
50 M
40 M
30 M
20 M
10 M
0
18:00 20:00 22:00 00:00 02:00 04:00 06:00 08:00 10:00 12:00 14:00 16:00
Total Orig Data Current: 42.15 M Average: 31.83 M Maximum: 57.45 M
Total Div Data Current: 30.08 M Average: 25.62 M Maximum: 46.67 M
Ratio orig/div % Current: 140.19 Average: 121.96 Maximum: 149.11
Saved Traffic Current: 12.07 M Average: 6.20 M Maximum: 13.71 M

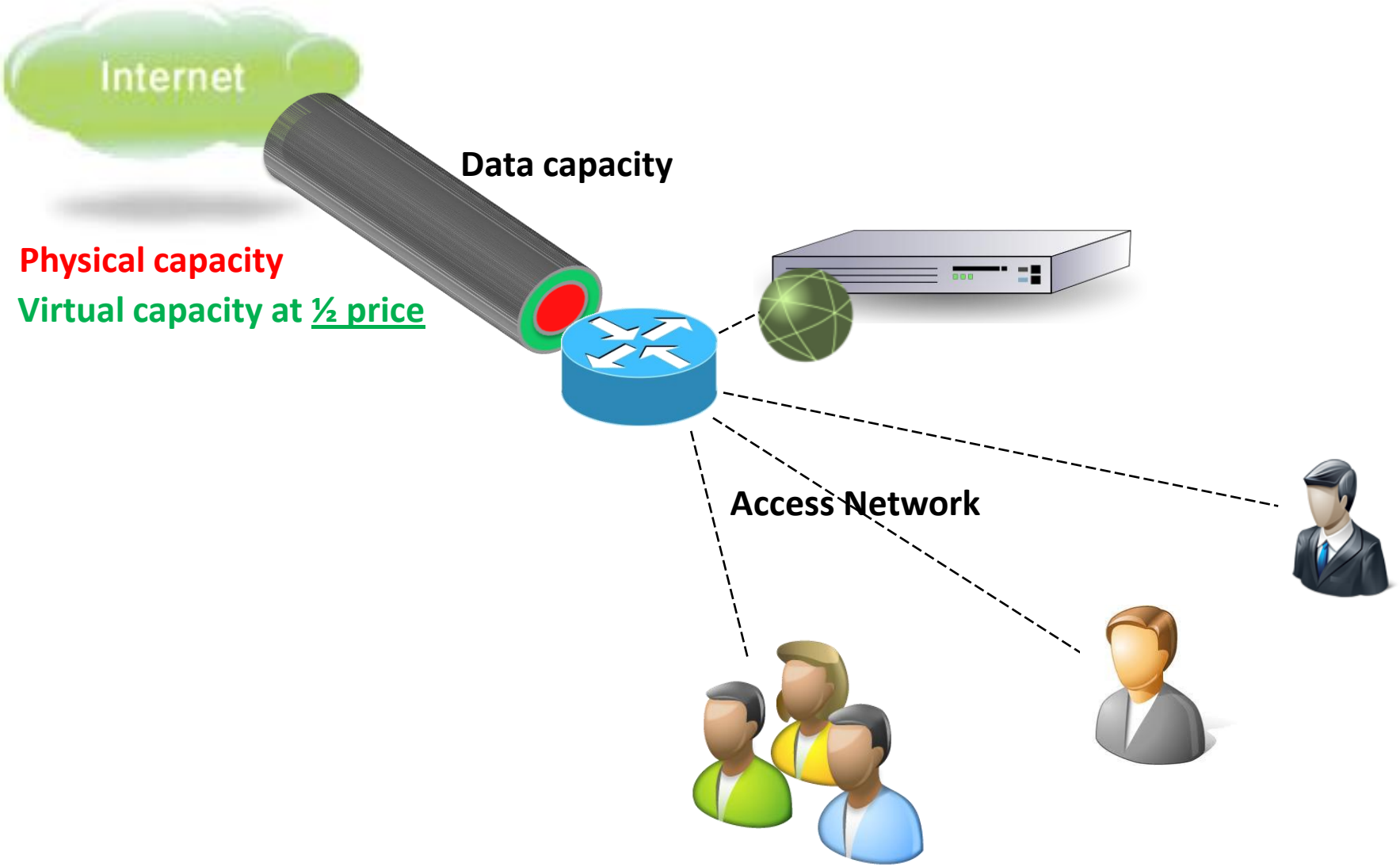
Case 5


BezeqInt Divinators traff bit/sec - Overall total
10 G
8 G
6 G
4 G
2 G
0
12
13
14
15
16
17
18
19
20
21
From 2011/06/12 09:07:34 To 2011/06/21 21:47:24
Total Orig Data Current: 8.47 G Average: 7.72 G Maximum: 9.85 G
Total Div Data Current: 6.40 G Average: 6.14 G Maximum: 7.84 G
Ratio orig/div % Current: 132.25 Average: 124.93 Maximum: 134.52
Saved Traffic Current: 2.07 G Average: 1.58 G Maximum: 2.46 G
RRDTOOL / TOBI OETIKER

Capacity planning with cloud-based capacity



DiViCloud virtually adds bandwidth



Thank You

Dr. Yair Shapira
yair@divinetworks.com