CONTENT

Vol. 8. No. 4-2022

Published bi-monthly since 2015.

ISSN 2664-066X (Print)
ISSN 2664-0678 (Online)

Publisher
Institute of Radio and
Information Systems (IRIS),
Vienna, Austria

Address:
1010 Wien, Austria,
Ebendorferstrasse 10/6b
media-publisher.eu/
synchroinfo-journal

DRM DIGITAL BROADCASTING SYSTEM AUDIO PATH QUALITATIVE CHARACTERISTICS

Oleg V. Varlamov, *Senior Member, IEEE,*
Institute of Radio and Information Systems (IRIS), Vienna, Austria,
o.varlamov@ieee.org

DOI: 10.36724/2664-066X-2022-8-4-2-8

ABSTRACT

The Digital Radio Mondiale (DRM) digital broadcasting system is the only ITU-approved digital broadcasting system for the LF, MF and HF bands. Numerous theoretical and field studies, as well as operating experience over the past 20 years, have allowed a good study of digital broadcasting networks organization according to this standard. However, some of the issues related to the audio path formal characteristics were not considered. The article discusses the audio coders used in the DRM system and the available data rates in various immunity modes. A set of test signals for instrumental measurements and listening has been developed. The results of experimental measurements for various data rates are presented. It is shown that the range of modulating frequencies in 7 modes of immunity is not narrower than for FM transmitters. In another 5 modes of noise immunity, the range of modulating frequencies is narrower than for FM transmitters, but wider than for AM transmitters. It is confirmed that at data rates of more than 16.5 kbps, the DRM transmitter signal quality becomes no worse than AM signal transmitter quality. Above 24 kbps, the DRM transmitter signal quality becomes comparable to that of an FM transmitter.

KEYWORDS: *digital broadcasting, DRM, transmitter, sound quality, instrumental measurements.*

I. INTRODUCTION

The Digital Radio Mondiale (DRM) [1] digital broadcasting system is the only ITU-approved digital broadcasting system for the LF, MF and HF bands. The DRM also has an option for the VHF band called DRM+. In the well-known scientific literature, there are many studies related to the requirements for DRM receivers [2], transmitters [3, 4], methods for their construction [5-8], measuring equipment [9], antenna systems [10], results of on-air measurements [11- 12], planning of coverage areas [13-15] and territories [16-20], including Simulcast [21] and single-frequency networks [22-23] modes.

The issues of the DRM system sound path qualitative characteristics were studied by subjective listening tests [24], but were not confirmed by formal instrumental measurements. The article discusses the DRM digital broadcasting system sound path qualitative characteristics in various immunity modes. A comparison is made with analogue broadcasting systems with amplitude and frequency modulation. The results of the conducted instrumental measurements allow broadcasters and operators to make an initial selection of the DRM immunity modes they desire, refining it through subsequent subjective listening tests.

The article is organized as follows. The second section discusses the audio coders used in the DRM system and the available bit rates in various immunity modes. In the third section, a set of test signals is developed and the results of experimental measurements are presented. Finally, the conclusions are collected in Section 4.

II. MATERIALS AND METHODS

Unlike analog broadcasting transmitters with amplitude (AM) or frequency modulation (FM), for the quality parameters of broadcasting (i.e. audio) paths of which there are corresponding regulatory documents, for transmitters operating in the DRM standard, there are currently no such documents. This circumstance is quite natural for digital technology in general, since distortions (frequency, non-linear, etc.) that occur in the transmitter do not affect the quality of the decoded signal - until the threshold allowed for decoding is exceeded. That is why the standards for DRM transmitters are mainly set only for the parameters that determine electromagnetic compatibility [25]. The only normalized parameter that determines the transmitter "quality" is the allowable Modulation Error Ratio (MER) value, which defined in [25] as:

$$MER = 10 \times \log_{10} \left\{ \frac{\sum_{j=1}^{N} (I_j^2 + Q_j^2)}{\sum_{j=1}^{N} (\delta I_j^2 + \delta Q_j^2)} \right\}, dB$$

where Ij, Qj are the values of the received symbol coordinates, δIj, δQj are the distance from the ideal position of the chosen symbol (the centre of the decision box) to the actual position of the received symbol. The value of MER in the transmitter output signal must be at least 30 dB [25]. As a rule, when fulfilling the Out-of-band emission regulations, the requirements for the allowable MER value are met automatically.

Thus, the broadcasting transmitter quality parameters (in the usual sense, these are the range of transmitted audio frequencies, harmonic distortion, intermodulation distortion, crosstalk between stereo channels, etc.) in the DRM system are determined solely by the audio coding parameters. The possibility of using one or another type of audio coders provided in the DRM system and their operation modes are determined by the available transmitted digital stream bit rate (Fig. 1 [26]), which, in turn, is determined by the occupied frequencies bandwidth and immunity modes.

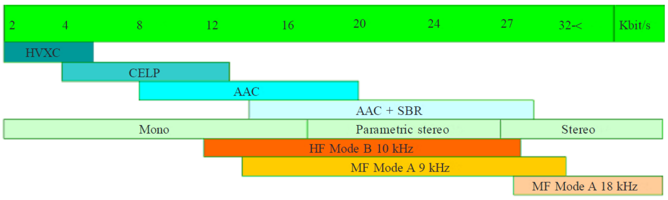


Fig. 1. Bit rate requirements for the different audio coders used in the DRM system

The main characteristics of these audio coders:

For HVXC – Harmonic Vector Excitation Coding (speech only) – bit rate 2000-6560 bps, SBR technology can be used to extend the audio frequency range from 4 kHz to 8 kHz. It can be used when organizing additional speech channels, for example, when transmitting multilingual news;

For CELP – Coded Excited Linear Prediction coder (speech, music at higher speeds possible) - data rate 3860-14000 bps, SBR technology can be used to extend the audio frequency range to 14 kHz;

For AAC (speech, music) coder, as the data rate increases from 8 to 20 kbps, the audio frequency range expands from 4 to 6 kHz. Universal audio coder, designed mainly for "complex" HF channels with low data rates;

The AAC+SBR coder has three audio bandwidth limits:

– 0 875 Hz – at data rates of 14 000-18 460 bps;
– 13 125 Hz – at data rates of 18 480-22 460 bit/s;
– 5375 Hz – at data rates of 22480-28460 bps.

The main application is in standard HF channels with speeds of 17-21 kbps;

Parametric Stereo – minimum bit rate 16480 bps, three audio frequency ranges – same as AAC+SBR mono. It can be used in the LF and MF bands or in "good" single-hop HF channels;

AAC+SBR Stereo – minimum bit rate 26480 bps, two audio frequency ranges:

– 13 125 Hz - at data rates of 26 480-28 480 bit/s;
– 5375 Hz – at data rates of 28480 bps.

It can be used in the LF and MF bands and in dual (18 kHz) MF channels.

The bit rates available in the main service channel (MSC) of the DRM system for the immunity modes used in the LF and MF bands are given in Tables 1 and 2 [26].

Table 1

Data rate (bit/s) in standard mode, Mode A (ground wave)

Parameters ⇓	Bandwidth (kHz)					
	4.5	5	9	10	18	20
64-QAM, rall = 0.5	9 392.5	10 620	19 695	22 142.5	40 935	45 840
64-QAM, rall = 0.6	11 272.5	12 740	23 625	26 570	49 115	54 995
64-QAM, rall = 0.71	13 305	15 045	27 892.5	31 367.5	57 982.5	64 940
64-QAM, rall = 0.78	14 745	16 660	30 910	34 770	64 260	71 970
16-QAM, rall = 0.5	6 262.5	7 080	13 125	14 760	27 285	30 555
16 QAM, rall = 0.62	7 827.5	8 850	16 412.5	18 452.5	34 112.5	38 200

Table 2

Data rate in standard mode, Mode B (skywave)

Parameters ⇓	Bandwidth (kHz)					
	4.5	5	9	10	18	20
64-QAM, rall = 0.5	7 200	8 280	15 332.5	17 477.5	31 817.5	35 760
64-QAM, rall = 0.6	8 640	9 930	18 402.5	20 975	38 180	42 905
64-QAM, rall = 0.71	10 200	11 730	21 720	24 750	45 065	50 660
64-QAM, rall = 0.78	11 300	12 990	24 075	27 450	49 950	56 140
16-QAM, rall = 0.5	4 800	5 520	10 222.5	11 655	21 210	23 835
16-QAM, rall = 0.62	6 000	6 900	12 777.5	14 565	26 515	29 800

It should be noted that options with a minimum data rate (i.e., with maximum immunity) do not suit either broadcasters or listeners in terms of the audio signal quality and can only be used to transmit voice information. “Acceptable quality” of audio content is achieved at bit rates of at least 14 kbit/s (preferably more than 20…22 kbit/s), for which, when using a bandwidth of 9 or 10 kHz, SNR at the receiving point is required from 11 dB to 13 dB (and 15 dB to 20 dB, respectively) depending on propagation conditions.

The ability to use higher bit rates (and get better audio quality) requires higher SNR, and therefore higher transmitter power while maintaining coverage, which broadcasters are usually reluctant to do.

At the same time, the concept of "acceptable quality" has not yet been standardized and formalized. In the DRM consortium promotional materials and in numerous publications, the concept of "quality comparable to VHF broadcasting" appears, based on subjective listening tests (for example, [24]) without specifying technical characteristics.

It is known that the quality of lossy audio encoders can be determined only on the subjective listening tests basis, and instrumental methods for measuring their characteristics do not give an adequate idea of the correctness of their work.

However, the large number of immunity modes provided by the DRM standard, and the corresponding number of different data rates with fine steps between them, make subjective listening tests task for all possible combinations difficult and prohibitively expensive. At the same time, broadcasters and operators who are used to focusing on standardized “quality classes” constantly have questions about what audio bandwidth will be available in a particular mode.

III. Instrumental Measurements

To answer these questions and in order to formalize the audio coding qualitative parameters in various DRM system immunity modes (i.e., at various data rates), a series of instrumental measurements was carried out - with the full understanding that they will not give an adequate idea of the actual audio quality.

When carrying out instrumental measurements, all parameters were checked that are normalized when testing both AM and FM transmitters (bandwidth of modulating frequencies, frequency response permissible deviation, harmonic distortion, intermodulation distortion, immunity from integral noise, crosstalk between stereo channels). The results of these measurements allow interested structures (broadcasters and operators) to compare the DRM transmitters qualitative characteristics with the AM and FM broadcast transmitters parameters and make an initial selection of the DRM immunity modes they desire, refining it through subsequent subjective listening tests.

A summary list of parameters characterizing the broadcast path of AM and FM transmitters quality, is given in Table 3.

Table 3

A summary list of parameters characterizing the broadcast path of AM and FM transmitters quality

Parameter	AM	FM stereo
1. Nominal modulating frequencies range, Hz	50-10000	30-15000
2. Frequency response permissible deviation, dB, not more than: FM AM: Up to 75 Hz and over 6600 Hz AM: 75 Hz to 6600 Hz	+0,7;-1,3 ±0,7	±0,8
3. Harmonic distortion, %, no more than: FM (m=100%, up to 7000 Hz) AM: Up to 100Hz (m=10%, 90%/m=50%) AM: 100Hz to 4000Hz (m=10%, 90%/m=50%) AM: Over 4000Hz (m=10%, 90%/m=50%)	3,8/1,5 2,0/1,0 4,0/2,0	0,5
4. Intermodulation distortion, %, no more than: AM: m=90% AM: m=50% FM, 3/5 orders, dB, no more	10 6	-50/-55
5. Unweighted (integral) noise level, dB, no more	-58	-62
6. Weighted (psophometric) noise level, dB, no more	-60	-65
7. Crosstalk attenuation between stereo channels, dB: 1000 Hz 120 Hz, 400 Hz, 5000 Hz, 10000 Hz		50 40

Instrumental measurements of the DRM transmitter broadcasting path parameters were carried out with TRAM-100 transmitter, and DRM coder-modulator DMOD3. A DT700 measuring and control receiver from Fraunhofer was used as a measuring demodulator.

The prepared set of test signals included all the necessary signals to determine AM and FM transmitters parameters in accordance with Table 3, as well as a set of musical fragments for subsequent subjective listening tests.

To reduce measurements time, if possible, common measurement series of frequencies are used, which are a combination of measurement series provided for by the standard measurement methods. The single test signal duration is chosen equal to 2 seconds, which allows to carry out Fourier analysis with sufficient accuracy to determine the harmonic distortion and measure the signal amplitude without taking into account the transient process. Single test signals are separated by 1 second pauses to facilitate their identification. In addition to frequency response measurements were introduced both on a noise signal with a uniform spectrum in the frequency band from 20 Hz to 22 kHz, and using a swept frequency source (sweep generator from 20 Hz to 20 kHz). The nominal range of modulating frequencies is determined by the tolerance for uneven frequency response.

In view of the above, the following set of test signals is defined:

– to measure the frequency response and harmonic distortion according to AM standards, 50% modulation: 1000 Hz, 30, 50, 63, 125, 500, 2000, 4000, 5000, 6000, 7000, 8000, 10000, 15000, white noise 20 seconds;

– for measuring the harmonic distortion according to AM standards, 10% modulation: 1000 Hz, 30, 50, 63, 125, 500, 2000, 4000, 5000, 6000, 7000, 8000, 10000, 15000;

– to measure the harmonic distortion according to AM standards, 90% modulation: 1000 Hz, 30, 50, 63, 125, 500, 2000, 4000, 5000, 6000, 7000, 8000, 10000, 15000;

– to measure frequency response and harmonics distortion according to FM standards, 100% deviation: 1000 Hz, 30, 50, 63, 125, 500, 2000, 4000, 5000, 6000, 7000, 8000, 10000, 15000;

– a ditional pause 1 sec;

– to measure intermodulation distortion according to AM standards m=90%: 80 Hz – 72%, 8000 Hz – 18% (products are calculated relative to 8000 Hz);

– a ditional pause 2 sec;

– to measure intermodulation distortion according to AM standards m=50%: 80 Hz – 40%, 8000 Hz – 10% (products are calculated relative to 8000 Hz);

– to measure intermodulation distortion according to FM standards: equal tones 5000 Hz and 7000 Hz, up to 100% deviation;

– a ditional pause 2 sec;

– to measure the levels of weighted and unweighted noise: 1000 Hz, 100%, 2 sec, then 2 sec pause;

– to measure crosstalk between stereo channels: 120Hz – 4 sec (2 sec in one channel, then 2 sec in another channel), 1s pause, then similarly at frequencies of 400, 1000, 5000, 10000 Hz;

– to measure the frequency response, a sweep frequency from 20 Hz to 20 kHz is additionally used, the duration is 4 seconds (100%).

A pause at the beginning of a set of test signals is 20 seconds, at the end - 10 seconds. The total duration of a set of test signals is 4 minutes 25 seconds. This set of test signals was generated in the Cool Edit Pro program and recorded on a CD that was used for instrumental measurements. The timing diagram of the signals recorded on the measuring disk is shown in Figure 2.

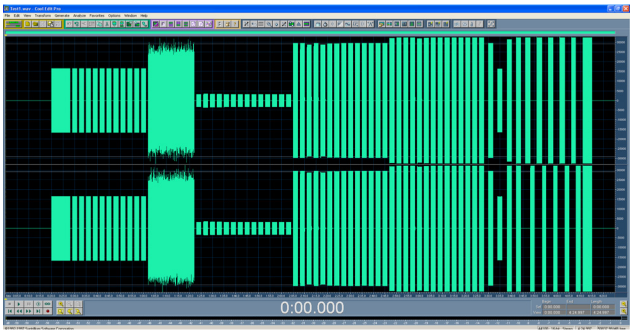


Fig. 2. The timing diagram of the signals recorded on the measuring disk

For all available bandwidths (4.5; 5; 9 and 10 kHz), for immunity mode "A" was set for 16 QAM modulation with immunity levels "0" and "1" and for 64 QAM modulation with noise immunity levels "0", "1", "2", "3". At sufficient transmission rates, the SBR (AAC +) mode was turned on. For some immunity modes, in addition to the "mono" mode, parameters were also measured in the "stereo" and "parametric stereo" modes. The output decoded signal recorded files with test signals were saved for further processing, and with musical fragments - for subjective listening tests. As an example, Fig. 3 shows the timing diagram of the output demodulated signal for mode "A", 9 kHz, 64 QAM, immunity level "1". The spectrum envelope of the swept frequency fragment, illustrating the range of reproducible frequencies for this mode, is shown in Figure 4.

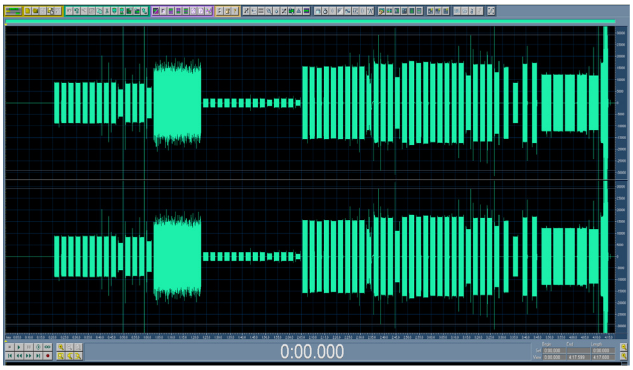


Fig. 3. Timing diagram of the output test signal for mode "A", 9 kHz, 64 QAM, immunity level "1"

The instrumental measurements results showed that the levels of unweighted (integral) noise and weighted (psophometric) noise in all modes of operation were less than -75 dB with the norm for AM no more than -58/-60 dB and the norm for FM no more than 62/-65 dB respectively.

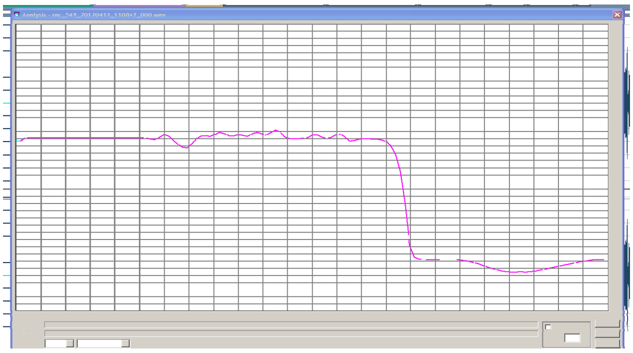


Fig. 4. Sweep spectrum envelope illustrating the reproducible frequency range for mode "A", 9 kHz, 64 QAM, immunity level "1"

Crosstalk attenuation between stereo channels in stereo modes was -77/-78 dB, while the norm in FM mode was -40/-50 dB. The crosstalk between stereo channels in the "Parametric stereo" mode, in accordance with its operation algorithm, depends on the frequency and is -9 dB at a frequency of 120 Hz; -11.6 dB at 400 Hz; -14 dB at 1000 Hz; -21 dB at 5000 Hz and -35 dB at 10000 Hz. The remaining results of instrumental measurements are given in Table 4.

Table 4

Instrumental measurements results

Parameter ***Mode***	**Modulating frequency range, Hz**	**Harmonic distortion, %**	**Intermodulation distortion, dB**
DRM A10/64/3	16700	0,1	-75
DRM A10/64/2	16700	0,05	-75
DRM A9/64/3	16700	0,05	-75
DRM A10/64/3 Stereo	15000	0,1	-75
DRM A10/64/3 Param. stereo	15000	0,1	-75
DRM A9/64/3 Stereo	15000	0,2	-75
DRM A9/64/1	15000	0,1	-75
FM Stereo Norm	30-15000	0,5	-50/-55
DRM A9/64/2	13800	0,1	-75
DRM A10/64/1	13700	0,2	-75
DRM A9/64/0	12700	0,05	-60
DRM A10/64/0	12000	0,05	-75
DRM A10/16/1	10800	0,05	-60
DRM A10/16/0	10700	0,05	-60
DRM A9/16/1	10700	0,1	-60
DRM A5/64/3	10700	0,2	-60
DRM A5/64/2	10700	0,05	-60
DRM A4,5/64/3	10700	0,05	-60
AM Norm	50-10000	3,8/1,5 2,0/1,0 4,0/2,0	-20/-24
DRM A4,5/64/2	3700	0,05	-
DRM A5/64/1	3100	0,05	-
DRM A4,5/64/1	2800	0,1	-
DRM A4,5/64/0	2600	0,1	-
DRM A9/16/0	2500	0,05	-
DRM A5/64/0	2500	0,05	-

The harmonic distortion given values correspond to the worst case in the entire frequency range. Intermodulation distortion is defined only for those modes, the bandwidth of which allows them to be measured.

As can be seen from Table 4, the transmitter in all studied DRM immunity modes has significantly lower intermodulation distortion and harmonic distortion than required by the standards for AM and FM broadcast transmitters. The range of modulating frequencies in 7 modes of immunity is not narrower than for FM transmitters. In another 10 modes of immunity, the baseband frequency range is narrower than for FM transmitters, but wider than for AM transmitters. A graphical representation of the modulating frequencies range dependence on the immunity mode and the corresponding data rate is shown in Figure 5.

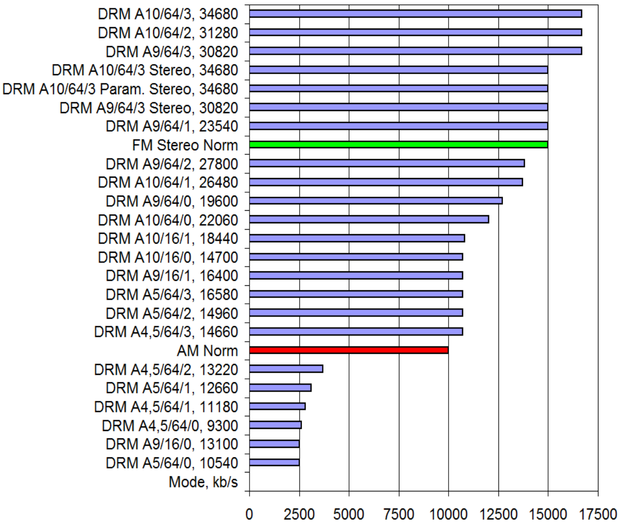


Fig. 5. The range of modulating frequencies vs. immunity mode and the corresponding data rate

If you measure the broadcast transmitter parameters in the DRM mode using traditional analog measuring instruments as a "black box", i.e. without a priori knowledge that it uses digital coding, you can get very high performance, unattainable for AM and FM transmitters. The applied method of instrumental measurements, using spectral analysis, made it possible to reveal some well-known features of low-speed digital coding. In particular, individual spectral components in the high-frequency part of the spectrum can be shifted from their original position.

So, in Fig. 6, the spectrum of the input test signal fragment is shown in blue, and the spectrum of the output signal corresponding to it in mode “A”, 10 kHz, 64QAM, immunity level “3” is shown in pink. The figure shows that the input frequencies of 7, 8, 10 and 15 kHz at the output are represented by slightly different frequency values. This is a well-known feature of this type of audio encoders, due to the algorithm of their work, which cannot be detected when measuring the frequency response with a voltmeter.

The noted circumstance once again confirms the above statement that it is expedient to determine the lossy audio encoders operation quality must based on the results of subjective listening tests, and the instrumental measurements results of their characteristics can only be used for preliminary selection of immunity modes and their comparison.

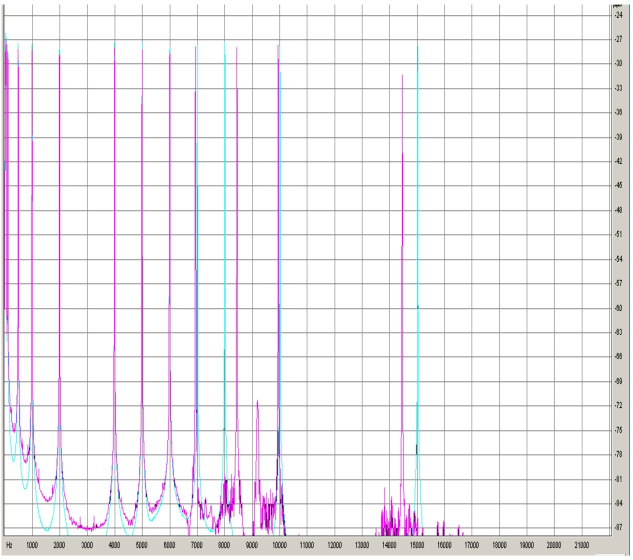


Fig. 6. Comparison the input (blue color) and output (pink color) spectra of the test signal fragments in the "A" mode, 10 kHz, 64QAM, immunity level "3"

The data obtained using instrumental measurements confirm the results of many studies that at data rates of more than 16.5 kbps, the signal quality of a DRM transmitter becomes no worse than AM transmitter signal quality. Above 24 kbps, the signal quality of a DRM transmitter becomes comparable to that of an FM transmitter.

When listening to recorded speech and music fragments encoded at rates of more than 20...24 kbit/s, digital processing artifacts were hardly noticed by ear, and the quality of the sound content could be called "comparable to FM". New xHE-AAC (Extended High Efficiency Advanced Audio Coding) codec allows you to get even higher quality audio content at the same encoding speeds.

IV. CONCLUSION

To study the parameters of the DRM broadcast transmitter, the following work was performed:

– a technique for instrumental measurements of the parameters of the DRM broadcast transmitter has been developed;

– a set of test signals for instrumental measurements and listening was prepared;

– instrumental measurements of the DRM broadcast transmitter parameters with different data rates in the MSC (in various immunity modes) were performed and fragments for instrumental analysis and listening were recorded.

Processing and measurement results analysis showed:

– the measured levels of unweighted (integral) noise and weighted (psophometric) noise in all modes of operation were less than -75 dB with the norm for AM not more than -58 / -60 dB, and the norm for FM not more than -62 / -65 dB, respectively;

– the transmitter in all studied immunity modes has significantly lower intermodulation distortion and harmonic distortion than the norms for AM and FM broadcast transmitters;

– the range of modulating frequencies in 7 modes of immunity is not narrower than for FM transmitters. In another 5 modes of noise immunity, the range of modulating frequencies is narrower than for FM transmitters, but wider than for AM transmitters;

– crosstalk between stereo channels in stereo modes was -77/-78 dB, while the norm in FM mode was -40/-50 dB. The crosstalk between stereo channels in the "Parametric stereo" mode, in accordance with its operation algorithm, depends on the frequency and is -9 dB at a frequency of 120 Hz; -11.6 dB at 400 Hz; -14 dB at 1000 Hz; -21 dB at 5000 Hz and -35 dB at 10000 Hz.

The data obtained using instrumental measurements confirm the results of many studies that at data rates of more than 16.5 kbps, the DRM transmitter signal quality becomes no worse than AM transmitter signal quality. Above 24 kbps, the DRM transmitter signal quality becomes comparable to that of an FM transmitter.

When listening to recorded speech and music fragments encoded at speeds of more than 20...24 kbit/s, digital processing artifacts were practically not noticed by ear, and the quality of the audio content could be called "comparable to FM".

REFERENCES

[1] ETSI ES201 980 V4.1.1 (2014–01) Digital Radio Mondiale (DRM); System Specification

[2] O.V. Varlamov, "Development of requirements for receiving equipment of digital broadcasting networks of the DRM standard", *T-Comm*, vol. 7, no. 9, pp. 39-42, 2013.

[3] O.V. Varlamov, V.G. Lavrushenkov "The quality criteria for the DRM standard transmitting device and the measuring equipment," *Broadcasting. Television and radio broadcasting*. No. 3. Pp. 44-48, 2004.

[4] O. V. Varlamov, "Theoretical Approach to Calculating Reverse Intermodulation Distortion in Voltage Mode Class D RF Power Amplifiers," *2022 Systems of Signals Generating and Processing in the Field of on Board Communications*, 2022, pp. 1-6, doi: 10.1109/IEEECONF53456.2022.9744320.

[5] N. Gromorushkin, O. V. Varlamov, A. V. Dolgopyatova and A. A. Voronkov, "Operation Problems of the EER Transmitter with Narrowband Antenna," *2019 Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, Russia, 2019, pp. 1-5. DOI: 10.1109/SOSG.2019.8706736

[6] O. V. Varlamov, D. C. Nguyen and S. E. Grychkin, "Simultaneous Application of Several Synthetic Methods for High Efficiency Radiofrequency Amplification," *2021 Systems of Signals Generating and Processing in the Field of on Board Communications,* 2021, pp. 1-5, doi: 10.1109/IEEECONF51389.2021.9416126.

[7] Varlamov O.V., Nguyen D.C., Grychkin S.E., "Combination of synthetic high-performance RF amplification techniques," *T-Comm*, vol. 15, no.9, pp. 11-16, 2021. DOI: 10.36724/2072-8735-2021-15-9-11-16

[8] O. V. Varlamov, "Multiphase PWM characteristics in the EER transmitter envelope path," *2021 International Conference on Engineering Management of Communication and Technology (EMCTECH)*, 2021, pp. 1-5, doi: 10.1109/EMCTECH53459.2021.9619166.

[9] O.V. Varlamov, V.N. Gromorushkin, V.G. Lavrushenkov and I.V. Chugunov, "Generator of test signals for measuring characteristics of EER SSB switching power amplifiers," *T-Comm*, vol. 5, no. 9, pp. 47-49, 2011.

[10] O. V. Varlamov and E. P. Stroganova, "Frequency extension circuit for EER transmitters operating with electrically short antennas," *2018 Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, 2018, pp. 1-5. DOI: 10.1109/SOSG.2018.8350577

[11] O.V. Varlamov, "Using the extraordinary wave for digital DRM NVIS broadcasting," *T-Comm*, vol. 9, no. 1, pp. 32-38, 2015.

[12] O.V. Varlamov, "Study of DRM digital broadcasting in the MF fading zone," *T-Comm*, vol. 9, no. 2, pp. 41-45, 2015.

[13] O.V. Varlamov, "The radio noise effect on the coverage area of DRM broadcast transmitter in different regions," *T-Comm*, vol. 9, no. 2, pp. 90-93, 2015.

[14] O. V. Varlamov and V. O. Varlamov, "Distribution of maximum levels of atmospheric radio noise in LF and MF ranges in the territory of the Earth," *H&ES Research*, vol. 9, no. 5, pp. 42-51, 2017.

[15] V. M. J. D. Santos and Y. A. Kovagin, "Building digital broadcasting networking in the low and midium frequencies," *T-Comm*, vol. 13, no. 4, pp. 55-63, 2019.

[16] O.V. Varlamov, "The technology of creating a digital broadcasting network of the DRM standard for the Russian Federation", D.Sc. Thesis, Moscow, MTUCI, 2017.

[17] O. Varlamov, V. Varlamov and A. Dolgopyatova, "Digital Radio Broadcasting Network in the Arctic Region," *2019 24th Conference of Open Innovations Association (FRUCT)*, Moscow, Russia, 2019, pp. 457-462. DOI: 10.23919/FRUCT.2019.8711933

[18] Varlamov O.V., Varlamov V.O., Dolgopyatova A.V., "DRM broadcasting international network to create an information field in the Arctic region," *T-Comm*, vol. 13, no.9, pp. 9-16, 2019.

[19] O. V. Varlamov and A. A. Bychkova, "Basis of Technical Design and Development a Single-Frequency DRM Digital Broadcasting Network for Venezuela," *2021 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO,* 2021, pp. 1-7, doi: 10.1109/SYNCHROINFO51390.2021.9488396.

[20] O. V. Varlamov and Abi Assali Bychkova, "Development of a DRM standard digital simultaneous radio broadcasting network for Venezuela," *REDS: Telecommunication devices and systems.* Vol. 10. No. 2, pp. 23-27, 2020.

[21] O.V. Varlamov, "Analog to digital signal power ratio in simulcast DRM transmission," *T-Comm*, vol. 10, no. 12, pp. 81-84, 2016.

[22] O. V. Varlamov, "Organization of single frequency DRM digital radio broadcasting networks. Features and results of practical tests," *2018 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)*, Minsk, 2018, pp. 1-8. DOI: 10.1109/SYNCHROINFO.2018.8456925

[23] Varlamov O.V., "Organization of single frequency DRM digital radio broadcasting networks. Features and results of practical tests," *T-Comm,* vol. 12, no.11, pp. 4-20, 2018.

[24] S. Meltzer and G. Moser, "MPEG-4 HE-AAC v2 - audio coding for today's digital media world," *EBU TECHNICAL REVIEW*, – January 2006, 1 / 12.

[25] ETSI EN 302245 V2.1.1 (2018-06) "Transmitting equipment for the Digital Radio Mondiale (DRM) sound broadcasting service; Harmonised Standard for access to radio spectrum". ETSI, 2018.

[26] Report ITU-R BS.2144, "Planning parameters and coverage for Digital Radio Mondiale (DRM) broadcasting at frequencies below 30 MHz," 2009.

IMPACT OF TUNNELING ON NETWORK CAPACITY

Knaj Nouma Khalil,
Master student of the Tartous University, Tartous, Syrian Arab Republic
knajnouma@gmail.com

DOI: 10.36724/2664-066X-2022-8-4-9-13

ABSTRACT

For many companies, setting up a VPN for secure, en-crypted communication is a cost-effective alternative to pur-chasing, operating, and managing a separate physical net-work. Many institutions, corporations, government agencies and non-profit organizations want to have their own private IP network for secure and reliable connectivity between of-fices across multiple geographies. A virtual private network (VPN) is a secure, encrypted connection over a public public network. Creating a separate network requires the purchase of equipment, its installation and maintenance. A VPN-based solution using the public Internet is becoming a cost-effective solution for many corporations. As a research task, the au-thors define an assessment of the impact of the tunneling technology used to solve problems arising from the lack of network support from existing data transfer protocols. The possibility of using the IPSec protocol and MPLS technology to implement tunneling is considered and compared. The results of the comparison and evaluation of the impact of the choice of protocol on the required bandwidth are presented.

KEYWORDS: *Packets, Bandwidth, Virtual Private Networks, Internet Protocol Security (IPsec), Multi Protocol Label Switching (MPLS), Label Switching Routers (LSR), Tunneling*

INTRODUCTION

Many institutions, corporations, government agencies and non-profit organizations want to have their own private IP network for secure and reliable connectivity between offices across multiple geographies. A virtual private network (VPN) is a secure, encrypted connection over a public public network. Creating a separate network requires the purchase of equipment, its installation and maintenance. A VPN-based solution using the public Internet is becoming a cost-effective solution for many corporations.

On the way to the destination address, data packets pass through many different networks. Packets use network protocols, but not all networks support the necessary protocols. To solve this problem, you can use the approach of placing the package inside another package using the protocols that are supported on the given network. This process of packet encapsulation is called tunneling. Tunneling allows VPN packets to reach their destination, which is usually a private network. The tunneling process provides a secure, encrypted connection between networks.

The aim of the work is to study the effect of tunneling on network throughput. Let's look at the MPLS protocol and the IPsec protocol, comparing the impact of each on network throughput for a Voice of IP (VoIP) application.

1. INTERNET PROTOCOL SECURITY (IP-SEC) CAPABILITIES

Tunneling is the technique of putting a packet of data into another packet (containing routing information) and sending it over the Internet. The packets travel along a path called a tunnel.

Internet Protocol Security, known as IPSec, is used to secure Internet communications over an IP network. Many VPNs use the IPsec protocol suite, which runs on top of an existing IP network. IPSec secures communications over the Internet by verifying the session. It encrypts every data packet during the connection. IPSec operates in 2 modes – transport and tunneling.

Transport mode encrypts the message in the data packet, while tunneling mode encrypts the entire data packet. IPSec can be used with other security protocols.

For example, suppose a company uses IPv6 to connect one of its offices (office A) to another (office B). The network between offices A and B only supports IPv4. It is necessary to pack (encapsulate) IPv6 packets into IPv4 packets in order to successfully transfer data between offices.

2. CAPABILITIES OF MULTI PROTOCOL LABEL SWITCHING TECHNOLOGY

A number of efforts have been made to increase the forwarding rate of packets in IP routers by introducing the concept of fixed-length labels. These efforts have been consolidated by the IETF (Internet Engineering Task Force) into the MPLS technology [RFC 3031, RFC 3032].

MPLS is a label lookup packet forwarding technology that does not affect the packet's IP header. An MPLS network consists of several routers called LSRs (Label Switching Routers). Routers that connect to IP routers are called LERs (Label Edge Routers).

The RFC defines the MPLS header format between MPLS capable devices (routers), between the second and third level headers. The MPLS Label header, as shown in Figure 1, consists of 20 bits (220 values or labels). The label value can be used by the LSR to find the next hop.

The EXP field consists of three bits of information and is used to implement functions related to QOS quality control. The next field is one bit called bottom-of-stack. It is used as a flag when more than one label is assigned to a packet, as in the case of MPLS VPN or MPLS TE.

The next byte is an eight-bit MPLS TTL (time to live) field that serves the same purpose as the IP TTL byte in the IP header. Each time an LSR forwards a packet, it decrements the TTL field in the packet's header, and if the value reaches zero, the packet is dropped.

An MPLS router (LSR, LER) forwards MPLS packets by looking up the MPLS label in its forwarding table and then immediately forwards the datagram to the appropriate egress interface. There is no need to process the header of every packet on every hop to determine the destination IP address and then look up the longest prefix match in the forwarding table based on the destination IP address.

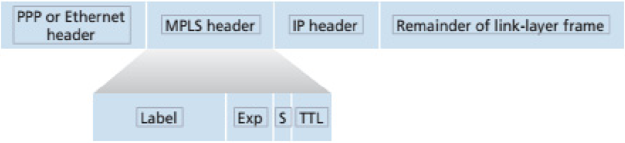
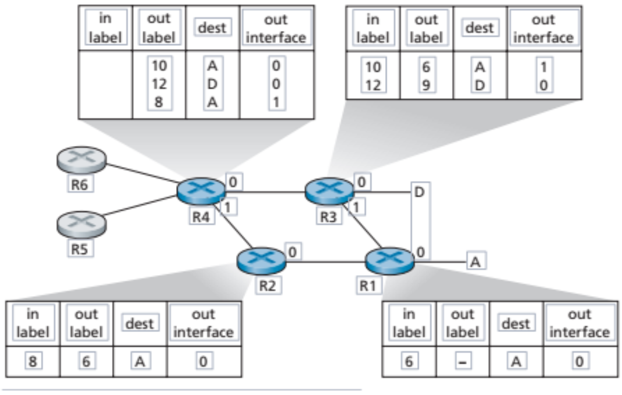


Figure 1. MPLS header

Figure 2 uses the following designations: R1-R4 – MPLS routers; R5 and R6 are standard IP routers.



in label	out label	dest	out interface
	10 12 8	A D A	0 0 1

in label	out label	dest	out interface
10 12	6 9	A D	1 0

in label	out label	dest	out interface
8	6	A	0

in label	out label	dest	out interface
6	–	A	0

Figure 2. MPLS transfers

According to the MPLS concept, router R1 announces to R2 and R3 that destination A is reached by any MPLS frame with label 6.

Router R3 announces to R4 that received frames with MPLS tags 10, 12 will be forwarded to destinations A, D.

Router R2 announces to R4 that the received frame with MPLS label 8 will be switched in direction A.

Router R4 can reach A with outgoing MPLS label 10 on interface 0 and outgoing MPLS label 8 on interface 1.

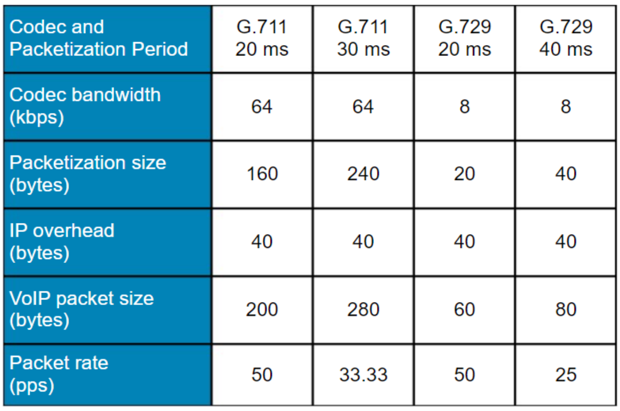
3. CODEC TYPE SELECTION

Various types of codecs based on different technologies are used to process audio and video data. To assess the quality of codecs, the Mean Opinion Score (MOS) indicator is used. MOS values are in the range from 1 to 5 (1 – "bad", 5 – "excellent"). The rating is given taking into account the P.800 and P.830 ITU-T standards.

G.711 is the default standard in any IP device and is the most widely used. The data transfer rate in one direction is 64 kbps, with a logarithmic compression ratio of 1:2 (16-bit sampling is reduced to 8-bit). On the MOS scale, it has a value of 4 (Fig. 3). G.711 can be used for LAN VoIP applications where there is a margin of available bandwidth.

G.729 is a codec that provides a drastic reduction in bit rate and payload size, but a slight reduction in voice quality.

G.729 and G.711 are most commonly used in VoIP applications. However, G.729 is more suitable for networks with limited bandwidth and low latency.



Codec and Packetization Period	G.711 20 ms	G.711 30 ms	G.729 20 ms	G.729 40 ms
Codec bandwidth (kbps)	64	64	8	8
Packetization size (bytes)	160	240	20	40
IP overhead (bytes)	40	40	40	40
VoIP packet size (bytes)	200	280	60	80
Packet rate (pps)	50	33.33	50	25

Figure 3. Codec characteristics

4. PACKET SIZE ESTIMATION FOR VOIP

We determine required network bandwidth for the implementation of VoIP according to formula:

B = (Pt x 8 /1000) x N,

where Pt = Payload – total packet size, bytes:

Pt= RP + O,

where RP is the packetization size defined as RP = (Tpack / 1000) × (codec bandwidth × 1000 / 8), bytes per packet; Tpack – packetization period; O – Overhead, bytes; N – the number of packets per second.

The packets per second (PPS) rate can be calculated by taking the reciprocal of the packetization period. Packet size is the amount of digitized and encapsulated voice in each IP packet, and it depends on the codec used. In G.711, a 20ms voice duration is digitized and encapsulated in one IP packet, so the packet rate is 1/20ms = 1/0.02 = 50pps. G.711 uses Pulse Code Modulation (PCM) 8000 samples with 8 bits for each sample with a total bandwidth of 8000*8=64Kbps. These 64 kbps contain 50 packets, which means the size of each packet

64 *1000 / 8*50 = 160 bytes.

На выходе кодека голосовые кадры помещаются в real-time protocol (RTP) packets to give them the necessary information for real-time end-to-end transmission, such as sequence numbers, timestamps.

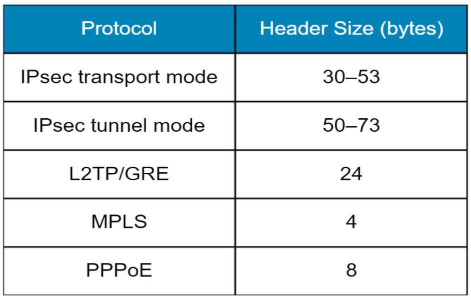
The overhead added by RTP to the payload is 12 bytes. In turn, the User Datagram Protocol (UDP) transport layer protocol adds an overhead of 8 bytes. The overhead added by IP is 20 bytes, and Ethernet adds 18 bytes. We will not add the size of the Ethernet header now, because this will be done in the last step after the tunneling process.

Considering the addition (excluding the overhead of tunneling and without adding the size of the Ethernet header), we get 160 + 12 + 8 + 20 = 200 bytes.

5. ESTIMATING OVERHEAD COSTS WHEN USING SECURITY AND TUNNELING PROTOCOLS

The implementation of tunneling leads to the appearance of so-called "overhead", that is, to an increase in share of service traffic in the transmitted information. There are many non-security protocols for tunneling IP packets and frames. Data for some common tunneling protocols with overhead for each packet is shown in Figure 4.

IPSec adds an Encapsulating Security Payload (ESP) header to the IP header to provide authentication and encryption. In transport mode, only the payload of the IP packet is encrypted. While in tunnel mode, the entire IP packet is encrypted, including the header. Encrypting the IP header eliminates the ability of routers to know the next hop of the packet and therefore route it, so an encrypted IP packet needs a different header to use in the routing process.



Protocol	Header Size (bytes)
IPsec transport mode	30–53
IPsec tunnel mode	50–73
L2TP/GRE	24
MPLS	4
PPPoE	8

Figure 4. Sizes of additional headers in security and tunneling protocols

ESP supports the use of a list of encryption and integrity protection algorithms such as HMAC. The operation of HMAC is based on the use of a given block size of 64 bytes. 8 bytes of message length are added to each packet (including 1 bit of the padding procedure identifier). After adding 8 bytes to a packet, if its size is not a multiple of 64, it must be padded so that it can be handled.

8 bytes of the padding HMAC procedure ID will be added to the 200 byte packet.

200 + 8 = 208 bytes.

ESP also adds some additional data and overhead: ESP header = 8 bytes; ESP initialization vector = 16 bytes; ESP trailer = 16 bytes; tunnel mode header = 20 bytes.

The package size is:

208 + 8 + 16 + 16 + 20 = 268 bytes.

Now you can add an 18 byte overhead for the link layer Ethernet header. It turns out:

Pt = 268+18= 286 bytes.

Let's calculate the bandwidth per voice connection when using the IPSec protocol using the formula:

$B_{1\ IPSec}$ = (Pt X 8 / 1000) X N,

where t= packetization size + O = 286 bytes; N – packet rate, N = 50 pps.

We get B1ip-sec = 114.4 kbps.

When using the MPLS protocol with a header size of 4 bytes, to create a tunnel, we calculate the overhead:

O= Eth + MPLS+ IPv4 + UDP +RTP =
= 18+ 4+ 20 +8 +12= 62 bytes.

Pt = packetization size + O = 160 + 62 = 222 bytes.

N is the packet rate, N = 50 packets per second.

Let's calculate the throughput per voice connection when using MPLS

B_1mpls = 88,8 kbit/s.

Thus, when using the security protocol and tunneling technology MPLS, the required network bandwidth per connection is reduced by 114.4 / 88.8 = 1.29 times compared to the IPSec protocol.

6. PLANNING AND CALCULATING BANDWIDTH FOR A CORPORATE NETWORK

The process of planning a telecommunications network requires carrying out the necessary measurements or traffic calculations. The results are used to determine the network topology, the bandwidth of the backbone group, the necessary lines to provide communication between network divisions. The most delay-sensitive, growing in demand, and revenue-generating applications are VoIP.

After calculating the throughput for a single voice call, we can calculate the total throughput based on the estimated traffic for all calls in Erlang multiplied by the throughput for a single call.

As an example, consider a corporate network that has a central office in Moscow (200 employees), two branches in St. Petersburg (100 employees) and Kazan (150 employees). They need to be able to communicate with each other through the central office.

Table 1

The results of the calculation of telephone load intensity in the CNN

Branch	Moscow	St. Petersburg	Kazan
Number of employees	200	150	100
Traffic volume (minutes/day)	9600	7200	4800
HNN (minutes)	1632	1224	816
A (Earl)	27.2	20.4	13.6

We need traffic delivery capabilities to implement VoIP (in terms of number of channels and required bandwidth), so we need to calculate the traffic during busy hour (HHH) in Erlang.

Traffic in HNN can be calculated by taking a percentage of the total number of daily minutes of calls in HNN (Table 1). The telecom industry defaults to a multiplier in HTN (17%) with a percentage of blocked calls or a Blocking Target (1%).

The results of calculating the required bandwidth in kilobytes for each branch of the company are presented in Table 2.

Table 2

The results of required throughput calculation

kbit/s	Moscow	St. Petersburg	Kazan
IPSec	3112	2333	1556
MPLS	2415	1811	1208

Comparing the required bandwidth when using MPLS and IPSec for a specific corporate network structure, we see that the use of MPLS technology can significantly save network bandwidth.

CONCLUSION

1. Many institutions, corporations, government agencies and non-profit organizations want to have their own private IP network for a secure and reliable connection between offices in several geographic regions. Instead of creating a separate physical network that is expensive to purchase, install, and maintain, a VPN solution over the existing public Internet is becoming a viable solution for many corporations.

2. Packets travel through many different networks on their way to their final destination. Communication protocols in networks differ. To overcome this problem, tunneling is performed, that is, packing packets inside other packets using protocols that are supported in a particular section of the network. The tunneling process allows a secure, encrypted connection to be established between users on different networks.

3. The performed analysis and calculation showed that when using the security protocol and tunneling technology MPLS, the required network bandwidth per connection is reduced by 1.29 times compared to the IPSec protocol.

4. Creating a tunnel with guaranteed QoS requirements for a highly sensitive application such as VoIP using MPLS saves significant bandwidth compared to IPSec, the secure tunneling protocol used in VoIP. This circumstance is especially important for the organization of corporate communications using the resources of the public Internet.

REFERENCES

[1] V. Olifer, N. Olifer. Computer networks. Principles, technologies, protocols: A textbook for universities. 5th edition. St. Petersburg: Piter, 2018. 992 p.

[2] A. B. Goldstein, B. S. Goldstein. MPLS technology and protocols. St. Petersburg: BHV-Peterburg, 2014. 304 p.

[3] V. P. Koryachko, D. A. Perepelkin. Analysis and design of data transmission routes in corporate networks. Moscow: Hotline – Telecom, 2020. 235 p.

[4] B.Ya. Lichtsinder. Traffic of multiservice access networks (interval analysis and design). Moscow: Hotline – Telecom, 2018. 290 p.

[5] S.N. Stepanov. Theory of teletraffic: concepts, models, applications. Moscow: Hotline – Telecom, 2015. 868 p.

[6] I. V. Stepanova, M. O. A. Abdulvasea, N. Zhuven. Analysis of promising approaches to improving the reliability of convergent corporate communication networks. *T-Comm*. 2015. Vol. 9, no. 12, pp. 44-51.

FULLY OPTICAL NETWORKS FOR QUANTUM ENCRYPTION KEY TRANSMISSION SYSTEMS

Maharramov Vagif Ali
Azerbaijan Technical University, Baku, Azerbaijan
mvg476@mail.ru

Mansurov Tofig Magomed ogly
Azerbaijan Technical University, Baku, Azerbaijan
tofiq-mansurov@rambler.ru

DOI: 10.36724/2664-066X-2022-8-4-14-17

ABSTRACT

The development of experimental quantum physics in recent years has led to the fact that the abstract ideas of quantum mechanics began to find practical application for the protection of information in a rapidly developing field such as fiber-optic communication lines (FOCL). On the basis of (smart) translucent SMART mirrors, optical splitters, switches, schemes of unidirectional, counter, bidirectional and universal multiplexers and demultiplexers, and the principle of distributing quantum encryption keys between authorized users have been developed. It is shown that it is possible to improve the speed of the process of switching and multiplexing of information flows in comparison with traditional mechanical switches. The paper proposes a circuit solution for an all-optical network (All-Optical Networks - AON) with the possibility of generating and distributing quantum encryption keys between authorized users. The aim of this work is to develop a principle for constructing quantum systems for the secure distribution of encryption keys on all-optical networks between selected (4 or more) users, as well as methods for generating, transmitting and receiving these keys in real time. In this regard, there is a need to develop a switch and a multiplexer of information flows.

KEYWORDS: *Quantum entanglement, switching, multiplexing, mirror, optical splitter, commutator, multiplexer, demultiplexer, information flow.*

Information about authors:
Maharramov Vagif Ali
Azerbaijan Technical University, Professor, Doctor of Physical and Mathematical Sciences, Baku, Azerbaijan

Mansurov Tofig Magomed ogly
Azerbaijan Technical University, Professor, Doctor of Technical Sciences, Baku, Azerbaijan

INTRODUCTION

In 2022, Frenchman Alain Aspe, American John Clauser and Austrian Anton Zeilinger were announced as winners of the Nobel Prize in Physics. The wording of the Nobel Committee states that these scientists are noted "for experiments with entangled photons, which demonstrated the violation of Bell's inequalities and gave rise to quantum computer science." Scientists have described the effect of "quantum entanglement", when the particles that were part of the same system continue to "feel" each other's state changes even at a distance of several kilometers.

The history of these studies began back in the mid-1930s with an article by Albert Einstein, Boris Podolsky and Nathan Rosen, in which a paradox was formulated by which the authors tried to show the incompleteness of quantum mechanics. Attempts to comprehend this paradox, to which the laureates made an important contribution, ultimately made it possible to better understand the quantum basis of our world.

The idea of protecting information using quantum objects was first proposed by Stefan Weisner in 1970. Decades later, C. Bennett and J. Brassard, having familiarized themselves with the work of S. Weissner, proposed to transmit a secret key using quantum objects and in 1984 proposed the possibility of creating a fundamentally secure channel using quantum states [1]. A detailed analysis of theoretical and experimental works in this direction was made in [2].

The development of experimental quantum physics in recent years has led to the fact that the abstract ideas of quantum mechanics began to find practical application for the protection of information in a rapidly developing field such as fiber-optic communication lines (FOCL). Experimental research in the field of quantum cryptography and currently proposed protocols for data transmission and encryption key can only involve two authorized users. With a larger number of users, it is very difficult to ensure the integrity of the encryption keys or its confidentiality. This problem was first considered in [3] and a new approach was proposed, the essence of which is that when transmitting confidential information over FOCL, the encryption key is transmitted after authorized users make sure that there are no unauthorized connections to FOCL. At the same time, the detection of unauthorized users is carried out by controlling the parameters of optical noise with a given photon statistics, known only to authorized users.

The main task of quantum cryptography is the search for efficient algorithms and the development of schemes for the practical implementation of the transfer of confidential information using quantum objects, i.e. single photons [2, 4]. The modern data coding system in telecommunication systems (classical cryptography) is based on the use of ciphers (keys), for deciphering which it is necessary to be able to factorize (decompose into prime factors) large numbers. Since there are no fast algorithms for factorization of large numbers for modern computers (although they have already been developed for quantum computers), which makes it possible to ensure secrecy. However, it can be expected that in the near future such algorithms will be found and the entire security system may be destroyed.

Therefore, to fully protect the transmitted data, it is necessary to use absolutely random sequences of numbers as encryption keys (used only once, to transmit one message from the sender Alice to the recipient Bob), which cannot be reliably determined by the spy Eve. According to Shannon's mathematically proven statement [5], a data transmission cannot be decrypted if the message is encrypted with a one-time random key, the length of the key is equal to the length of the message, and this key is known only to authorized users.

The aim of this work is to develop a principle for constructing quantum systems for the secure distribution of encryption keys on all-optical networks between selected (4 or more) users, as well as methods for generating, transmitting and receiving these keys in real time. In this regard, there is a need to develop a switch and a multiplexer of information flows.

A NEW APPROACH TO SWITCHING AND MULTIPLEXING INFORMATION FLOWS

In the process of developing fiber-optic networks, one has to face a number of complex scientific and technical problems. One of them is the creation of high-speed splitters of information optical streams that perform the functions of both an optical switch and a multiplexer, providing the required accuracy of spatial modulation or spectral selection of the stream and thereby stabilizing the position of the optical stream in the focal surface of the receiver or transmitter of optical information [6, 7].

The main advantage of all-optical networks is their virtually unlimited bandwidth. The practical value of this property lies in the possibility of multiplying the speed of information transmission over fiber optic communication channels on a global scale. This task of research in the field of all-optical networks is a very relevant and promising task of a theoretical and practical nature.

In addition to the important task of improving the parameters and designs of backbone optical cables, the issue of creating reliable and affordable optical signal switches is no less acute, without which it is impossible to build branched optical networks [8].

Switches are one of the most important units of optical information transmission systems built on the basis of standard hierarchical structures. Without them, it is practically impossible to carry out automatic control of the movement of data flows and monitoring issues over an extensive network. A huge variety of devices used in technology that perform the function of switching optical signals is determined by the particular features of their use in a particular type of network [9].

One of such technologies proposed by us, as the principle of key distribution, based on bidirectional multiplexing of information flows, is considered in [7, 8].

As you know, the basis for building all-optical networks is the creation of fundamentally new circuit solutions for switching and multiplexing information flows. On fig. 1 shows variants of circuit solutions for switching and multiplexing information flows based on moving Smart mirrors.

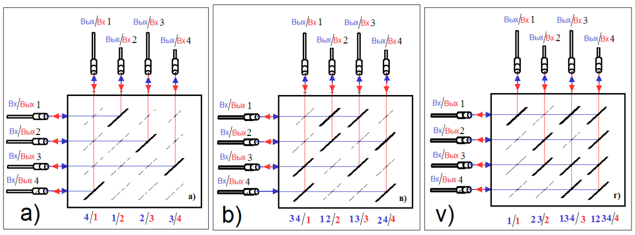


Figure 1. Options for switching and multiplexing information flows based on moving Smart mirrors

Similar switches and multiplexers can be created from a set of optical lenses in the form shown in Figure 2.

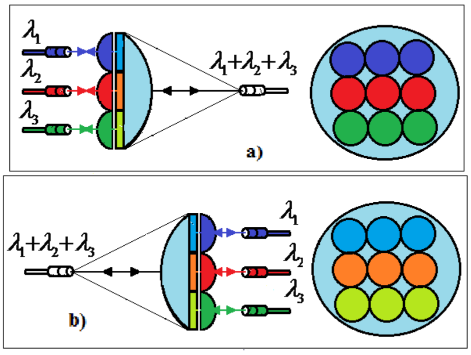


Figure 2. Technology of switching and multiplexing optical flows based on focusing lenses

All optical filters are designed as semitransparent dichroic mirrors operating at a selected wavelength [8, 9]. A dichroic mirror is an optical element that reflects radiation of only one wavelength and freely transmits other wavelengths. Such mirrors are a glass substrate with a deposited multilayer dielectric structure, which reflects only one wavelength due to the interference effect.

The principle of quantum keys distribution built on the basis of translucent mirrors

One of the problems of cryptography has always been the problem of key distribution, which is currently successfully solved using asymmetric encryption algorithms with a private key that does not leave its owner.

Currently, there are several protocols used in quantum cryptography to distribute encryption keys. Historically, the first implementation of a quantum key distribution system was a polarization coding scheme operating according to the BB84 protocol [10].

However, the strength of this and many other encryption algorithms is ensured by the current lack of computing power in the world for the possibility of successful cryptanalysis, so it is worth looking for new methods and technologies for key distribution.

One of such technologies with polarization coding was proposed in [6], where the principle of quantum key distribution was created on the basis of bidirectional multiplexing of information flows using the features of an optical splitter and a semitransparent mirror. The proposed scheme of a quantum cryptographic installation with polarization coding is shown in Figure 3.

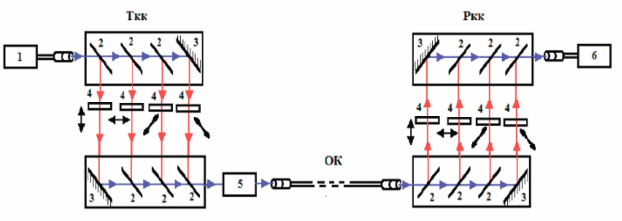


Figure 3. The proposed principle of quantum key distribution based on bidirectional multiplexing of information flows:
Tkk – quantum key transmitter or station Alice; Rkk – quantum key receiver or station Bob; OK - optical cable; 1 – semiconductor laser; 2 – translucent mirrors operating in the optical splitter (OR) mode; 3 – reflective mirrors; 4 – polarizers; 5 – absorbing filter; 6 – single-photon detector of photons from an avalanche photodiode

A semiconductor laser quantum encryption key transmitter emits short light pulses (eg, 1.0 ns duration). In principle, a semiconductor laser can operate in a continuous mode, and the formation of a short pulse (for example, 1.0 ns duration) can be formed by the OR, depending on the required polarization plane, turning the CCD into an active position. The polarization planes of photons can be -45°, 0°, +45°, and 90°. In other words, in order to transmit one bit from the required plane of polarization of a photon, it is activated according to this degree of polarization, chosen by smart translucent mirrors (SMI) in Tkk. In principle, all four types of polarization can be transmitted simultaneously. Then the pulses are attenuated by absorbing filters 5 to ensure the single-photon condition, i.e. the average number of photons per pulse is chosen to be less than one. After that, the photon is emitted in the direction of the quantum key receiver or station Bob. As is known, an important condition for the correct detection of information by station Bob is the preservation of the polarization of photons in the optical fiber.

The pulses arriving at the input of the receiver of the quantum encryption key or the Bob station pass through the CPL set and the initial polarization state is automatically determined. After that, coming from the avalanche photodiode to the input of a single-photon detector, photons are detected in the corresponding status codes. It is very important to note here that if an external intruder on the way from Alice to Bob is noticed by the transmitter or receiver of information, then they, for their part, can automatically change the direction of information transfer to the optical splitter with smart translucent mirrors, noted in [6, 7, 11, 12].

The proposed scheme of a quantum cryptographic setup with polarization coding is characterized by ease of implementation, fast detection, and high reliability. After creating SMART optical separators, multiplexing (MUX), demultiplexing (DMUX) and the principle of quantum key distribution (QKKD), we will start creating an optical network with the possibility of generating and distributing quantum encryption keys.

ALL-OPTICAL NETWORK WITH THE ABILITY TO GENERATE AND DISTRIBUTE QUANTUM ENCRYPTION KEYS BETWEEN AUTHORIZED SUBSCRIBERS

As noted in [3], the most effective ways to protect information transmitted over FOCL are quantum cryptography methods, however, according to this scheme, it is impossible to create systems for generating and distributing encryption keys for more than two authorized users. Therefore, for the first time in [9], a new approach was proposed for creating a system for generating and distributing encryption keys between several authorized users for transmitting confidential information over FOCL.

Subsequently, expanding this idea, Figure 4 proposes the principle of constructed AON networks, where the formation and distribution of quantum encryption keys between authorized subscribers is carried out [13].

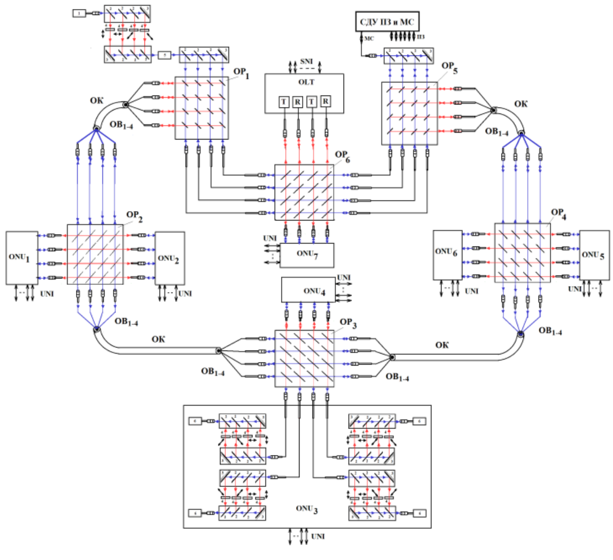


Figure 4. Scheme of an all-optical network based on optical splitters with CCD:
SNI – trunk connection interfaces; OLT – central node; ONU1-ONU7 – subscriber nodes, UNI – subscriber connection interfaces; m - main and backup transmitters; R – main and backup receivers; SDU PZ and MS – remote control system for translucent mirrors and network monitoring; OK – optical cable; OV1÷4 – optical fibers; OP1÷OP7 – optical splitters; 1 – semiconductor laser; 2 – translucent mirror; 3 – reflective mirror; 4 – polarizer (Glan prism); 5 – absorbing filter; 6 – avalanche photodiode

Here, the AON network is based on the principle of building AON networks based on OR with VMI, considered in [7].

The operation of the network is similar to the principle of operation described in Figure 3, as well as the operation of the networks described in [3, 9, 13]. On the other hand, if unauthorized users are detected, depending on the location of Eva penetration, the subscriber of this subscriber node can turn off the information receipt network or change the route using the capabilities of a universal optical splitter [6, 7].

CONCLUSION

Taking into account the proposed principle of constructing translucent and smart translucent mirrors, optical splitters, the scheme of unidirectional, counter, bidirectional and universal multiplexers (MUX) and demultiplexers (DMUX), the option of building an all-optical network based on an OR with an SPL, the principle of distributing quantum keys based on bidirectional multiplexing information flows made it possible to create an AON network with the ability to generate and distribute quantum encryption keys between authorized subscribers. On the other hand, the advantage of the network lies in the fact that, firstly, all ORs and CCDs are remotely controlled and, secondly, all optical fibers are under continuous monitoring at a wavelength that does not interfere with the normal functioning of the network built for its intended purpose.

REFERENCES

[1] Bennet C.H., Brassard G. Proc. *IEEE Intern. Conf. on Comput. Sys. and Sign. Proces*., Bangalore (India), 1984, pp. 175-179.

[2] Gisin N., Ribordy G., Titlel W. et al. Quantum cryptography. *Rev. Mod. Phys*. 2002. Vol. 4, pp. 145-175.

[3] A.O. Zenevich. Quantum systems for transmitting encryption keys over fiber-optic communication lines. *International conference "Innovative technologies in telecommunications"*. Baku, December 4-6, 2019, pp. 13-15.

[4] I.I. Ryabtsev, I.I. Beterov, D.B. Tretyakov et al. Experimental quantum informatics with single atoms and photons. *Bulletin of the Russian Academy of Sciences*. 2013. Vol. 83. No. 7. P. 606.

[5] Shannon, C.E. Communication Theory of Secret Systems. *Bell Syst. Tech. Jour*. 1949. Vol. 28. P. 658.

[6] V.A. Maharramov. The principle of an optical splitter. *International conference "Innovative technologies in the shopping mall"*. Baku December 4-6, 2019. P. 155-158.

[7] V.A. Maharramov. Fully optical networks based on Smart mirrors. *Problems of infocommunication*. No. 1(11), Minsk, Belarus, 2020, pp. 19-26.

[8] V.A. Maharramov, T.M. Mansurov. About one technology of switching and multiplexing of information flows. *XXVII International Conference "Modern means of communication"*. Minsk, October 27-28, 2022, pp. 184-186.

[9] V.A. Maharramov. All optical networks based on translucent mirrors. *"Machine-building and Energy: New Concepts and Technologies" International Scientific-practical Conference*, December 2-3, 2021, AzTU, Baku, Azerbaijan.

[10] V.L. Kurochkin, I.I. Ryabtsev, I.G. Neizvestnyy. Experimental setup for quantum cryptography with single polarized photons. *JTF*. 2005. Vol. 75. no. 6, pp. 54-58.

[11] I.R. Gulakov, A.O. Zenevich, T.M. Mansurov. Components of fiber-optic communication lines. Minsk, BGAS, 2020. 336 p.

[12] T.M. Mansurov, A.O. Zenevich, I.A. Mammadov. Fiber optic coupler/optical power switch. *REDS: Telecommunication devices and systems*. 2021. No.2, pp. 29-36.

[13] V.A. Maharramov, T.M. Mansurov. A new approach to building all-optical networks for a quantum encryption key transmission system. *XXVI International Conference "Modern means of communication"*. October 21-22, 2021, pp. 128-131. Minsk, Belarus.

DEVELOPMENT OF AN AUTOMATED ROAD CONTROL SYSTEM "ALA-ARCHA"

Valery V. Mamrega,
"Darbaza-Avtomatik" LLC, Bishkek, Kyrgyzstan
mamregavv@gmail.com

DOI: 10.36724/2664-066X-2022-8-4-18-21

ABSTRACT

This article explores the subject of computer vision systems – a technology that allows vehicles to identify, track, and also classify objects on the roadway. The objectives of the study are to consider the principle of operation of these automated systems, their advantages in comparison with modern road regulation, as well as the problems of implementation and development of these systems. The research was carried out on the basis of the analysis of information from open information resources. The statistics of accidents at work are presented, the high rates of which are due to large volumes of production and an outdated system for monitoring compliance with safety rules and the availability of personal protective equipment for employees. The scheme of interaction of the components of a computer vision system is considered, which will allow monitoring of events occurring in production during operation, monitoring the situation at the enterprise for the occurrence of a potentially dangerous situation for personnel and equipment, and, accordingly, this system will be able to prevent an emergency, as well as avoid personal injury by reacting even to minor deviations from operating parameters. The research was carried out on the basis of the study and analysis of materials published in open information sources.

KEYWORDS: *machine vision, traffic control system, computer vision, transport, neural networks, production, computer vision, safety compliance, monitoring, production process, safety.*

Information about author:

Valery V. Mamrega, *Private researcher of machine vision. General Director of LLC "Darbaza-Avtomatik", Bishkek, Kyrgyzstan*

INTRODUCTION

Every year around the world there is a tendency to increase the number of vehicles, respectively, density of the flow and its intensity, which, in turn, causes the need for traffic management [1-5]. One of the priority vectors of development in this direction is the development of automated traffic control systems. The main task of automated system is to increase the vehicle traffic regulation efficiency, reduce intersection delays of their traffic trajectories, as well as increase the level of comfort and safety of traffic participants [6-11].

One example of an automated control system is the use of a special cabin with a visit control point at the entrance to the parking area or airport area, where you can stay for a limited amount of time [12-15]. The sequence of entry and exit consists in the approach of the vehicle to the barrier, next to which there is a point for issuing "visitor cards". The driver, having received the card, takes it to his salon, waits for the opening of the barrier and drives into the territory. When leaving, the driver returns card back, barrier opens and vehicle leaves the territory [16-18]. These systems are manufactured by companies such as Mallenom systems (Automarshal barriers).

RESEARCH METHODS AND RESULTS

From the material presented above, we can conclude that one of the priority areas in the development of automated control systems is the development and improvement of software, namely the identification of a vehicle on the road and the highest quality data collection about it. Also, in addition to the technical aspect, it should not be overlooked that it is necessary to develop a reliable system, coupled with ease of maintenance and affordable cost of the finished complex, so the development of a traffic control system is one of the priority areas in the field of traffic management [19-21].

As part of the implementation of the project "Development of the automated road regulation system "Ala-Archa"", the goal was determined, which is to develop a control system, its logic and control algorithms, as well as the analysis and selection of equipment necessary for the implementation of a checkpoint that controls the entry of vehicles funds to the reserve "Ala-Archa", located on the territory of Kyrgyzstan, as well as departure from it.

In the course of the project, an analysis of modern systems for regulating the movement of vehicles at the entrance to a closed area, as well as leaving it, was carried out. A significant disadvantage of this system is the need for the driver to perform the following actions – stop in front of the barrier, open the window, reach for the button to receive the visitor's card, take the card into the salon, close the window, wait for the barrier to open, and only after all of the above actions the driver will be able to enter the territory. The driver will have to carry out a similar sequence of operations when leaving the territory [22-23].

Accordingly, the imperfection of the existing traffic regulation system is due to inappropriate software that does not provide autonomous operation of the checkpoint, as a result of which drivers must perform a certain sequence of actions in order to be able to enter and exit the territory, which, in turn, takes a large number of time, and, accordingly, entails the formation of a queue in front of the checkpoint. In addition, if the driver was in the closed area for more than the set time, he must go to the payment point in front of the barrier at the exit, pay the amount that is calculated depending on the time that the driver spent in the territory in excess of the established limit, and only after payment does he get the opportunity to leave.

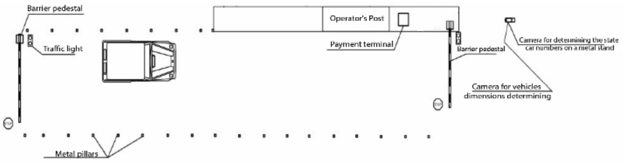

Barrier pedestal
Traffic light
Operator's Post
Payment terminal
Barrier pedestal
Camera for determining the state car numbers on a metal stand
Camera for vehicles dimensions determining
Metal pillars

Figure 1. Car movement scheme at the entrance to park territory

After analyzing the existing traffic control system, a proposal was put forward to organize an automated collection of payment from vehicles based on determining the dimensions of vehicles at the entrance to the Ala-Archa Natural Park, as well as at the exit from it.

Figure 1 shows a diagram of the movement of a vehicle at the entrance to the park. The vehicle drives up to the first barrier, after which the system for determining the dimensions of the car is located, and, provided that there is no vehicle in the dimension determination zone, the first barrier lets the car into the dimension determination zone, and closes immediately after the passage. If there is another vehicle in dimension determination zone, the first barrier remains in closed position. Then the driver activates fare payment button, while determining car dimensions using CCTV cameras, which makes it possible to identify vehicle dimensions with an accuracy of several centimeters, in parallel, the car number is read and data about vehicle is sent to the system. After making the payment, a second barrier opens to enter the park. After payment this vehicle has access to enter the park within 24 hours.

Figure 2 shows a vehicle movement diagram (when leaving the park). The vehicle approaches the barrier at the territory exit, video surveillance system identifies vehicle presence, after identifying vehicle presence, open barrier, vehicle leaves the park. This system logic allows five times to reduce the time for car to leave the closed area.

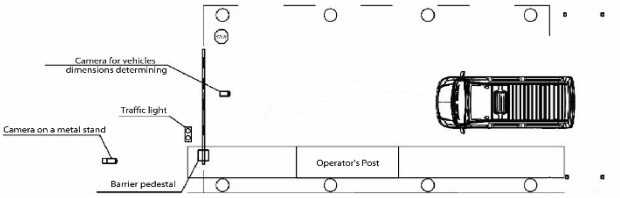


Figure 2. Car movement scheme (leaving the park)

To implement this logic, the following resources are required: Server, barrier, payment terminal, image recognition cameras.

The project is based on the Open CV computer vision library and implemented on the Python programming language platform. During the implementation of the project, the accuracy of recognition of dimensions and car numbers was increased to 92%, and during the optimization of the system in 2021, this figure was already brought up to 98%. Such high performance was achieved by training the neural network with more than 10 million frames of images of vehicle numbers.

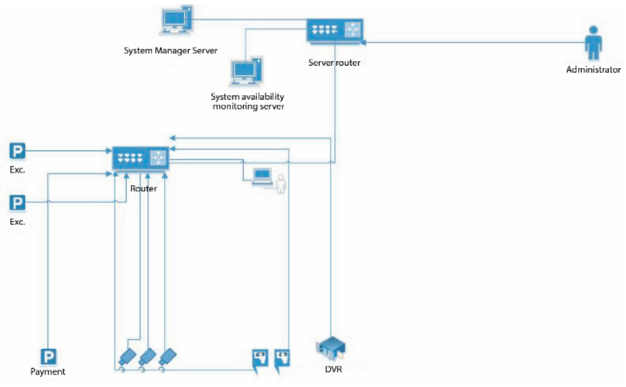


Figure 3. Scheme of interaction between the components of system logic for automated payment collection from vehicles based on determining the dimensions of vehicles at the entrance to natural park "Ala-Archa"

DEPARTURE LOGIC

1. At the entrance there are two barriers, for each passage – a vestibule.

2. The first barrier works automatically after magnetic loop is triggered, the second one opens only after payment, or checks list of free cars (white list) against billing, or the list of cars whose owners have paid fare or more than 24 hours have not passed since the payment.

3. Exposure at the first barrier – when the loop is approached, a signal is sent to open (it does not matter in what position the barrier boom is now - if it is open and the vehicle enters the loop, the barrier will close) after the boom has taken a vertical position (open) , which withstands a passage waiting pause of 120 seconds, if the car does not cross photo elements on the barrier, then auto-closing occurs, if it crosses photo cells, barrier closes immediately after crossing the photo cells.

4. Each entrance barrier has two magnetic loops (one in front of the barrier, second behind it).

5. When the car is on the second barrier (loop 2 of the second barrier), the car that drove up to the first barrier, barrier will not open, opening occurs only after passing through the second car barrier, which was in the magnetic loop 2, in this case, first barrier will open automatically.

6) After payment and opening of the second barrier, there is a waiting at the intersection of photocells for 120 seconds, after which, if no action occurs, the barrier will close automatically; if there is an intersection of photocells, barrier will close immediately after intersection.

7. In order to correctly determine the dimensions and make payment, there should be only one car in the vestibule.

8. If there are two vehicles in the vestibule, the dimensions and numbers of vehicles may be considered incorrectly.

9. The vehicle must move along the vestibule no more than 5 km / h for the correct reading of numbers. If the car number was not considered for any reason, it must be entered manually. In no case should the vehicle be allowed to drive back and forth along the vestibule, this can lead to damage to the fence, as well as barriers if the driver is inattentive.

ENTRY LOGIC

1. When the car approaches exit barrier and runs into the magnetic loop, there is an exposure for 10 seconds (for license plate recognition), and car is searched in the system to transfer it to passage history.

2. After which the exit barrier opens and waits for 40 seconds to pass through the photocells, if there is no passage, barrier closes automatically.

CONCLUSION

Comparing the proposed system for organizing entry and exit from the park with existing analogues, we can conclude that thanks to the improved logic of the system, drivers only need to activate pressing the ticket purchase button and pay for it, the system performs all other operations autonomously, significantly increasing the speed of vehicles passing by territory of the park and reducing the time for the vehicle to leave the closed area, as well as unloading the work of the cash desk operator and administrator. These advantages of the developed system are provided by the implemented logic based on the computer vision library "Open CV" and the programming language "Python", as well as a set of special equipment described in this article.

REFERENCES

[1] Automated traffic control system (ASUDD). URL: https://korda-group.ru/products/asudd.

[2] Barriers and barriers. URL: https://clck.ru/X9XLJ.

[3] Barrier with a security post. URL: https://rem-stroytrest.ru/shl.html.

[4] Automatic traffic control system (ASKRDD). URL: https://www.vzglyad.biz/ru/asrpdd.html.

[5] Optimization of the work of adaptive traffic lights based on the use of machine vision. URL: https://clck.ru/XFAX4.

[6] Vizilter Yu.V., Zheltov S.Yu. Technical vision in mobile object control systems. *Proceedings of the scientific and technical conference-seminar*. 2010, pp. 11-44.

[7] Optimization of the work of adaptive traffic lights based on the use of machine vision. URL: https://avt.global/video/computervision.

[8] Machine vision: demand and prospects. URL: https://www.tbforum.ru/blog/mashinnoe-zrenie-vostrebovannost-i-perspektivy.

[9] Machine vision. URL: https://www.iksmedia.ru/articles/5685849-Mashinnoe-zrenie-kak-nauchit-lokomo.html.

[10] Machine vision for traffic analytics. URL: https://www.secuteck.ru/articles/kompyuternoe-zrenie-dlya-analitiki-dorozhnogo-dvizheniya-3-uspeshnyh-kejsa.

[11] Machine vision. What is it and how to use it? Processing of images of an optical source. URL: https://habr.com/en/post/350918.

[12] Mastering computer vision - 8 basic steps. URL: https://habr.com/ru/post/461365.

[13] Computer vision. URL: https://exponenta.ru/comp-vision.

[14] Video monitoring of industrial safety using computer vision. URL: https://digdes.ru/nlab/mashinnoe-zrenie-v-promyshlennosti.

[15] Researchers recorded the brain activity of rats and used it to improve machine vision. URL: https://habr.com/en/search/?q=%D0%BC%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5%20%D0%B7%D1%80%D0%B5%D0%BD%D0%B8%D0%B5&target_type=posts&order=relevance.

[16] Computer vision. URL: https://yandex.ru/company/technologies/computer_vision.

[17] How traffic was regulated: from Caesar to the present day. URL: https://www.prostranstvo.media/kak-regulirovali-dorozhnoe-dvizhenie-ot-cezarja-do-nashih-dnej.

[18] What role do computer vision systems play in the fourth industrial revolution? URL: https://www.baslerweb.com/ru/vision-campus/otrasli-i-zadachi/rol-kompyuternogo-zreniya-v-ehpohu-industriya-4-0.

[19] Computer vision. Tasks, scopes, prospects. URL: https://vc.ru/ml/166105-kompyuternoe-zrenie-zadachi-oblasti-primeneniya-perspektivy.

[20] Just about the complex. How does computer vision work. URL: https://vc.ru/u/536956-infosistemy-jet/162112-prosto-o-slozhnom-kak-rabotaet-kompyuternoe-zrenie.

[21] How to choose an automatic barrier: what to look for, the best models and reviews. URL: https://vidsyst.ru/sistemy-kontrolya-dostupa/shlagbaum/avtomaticheskij.html22.

[23] Automarshal. URL: https://www.mallenom.ru/products/videokontrol-i-uchet-avtotransporta/avtomarshal/?utm_source=yandex&utm_medium=cpc&utm_campaign=cid|30020285|search&utm_content=gid|2929484070|aid|4772959164|11695495130_&utm_term=%D0%B0%D0%B2%D1%82%D0%BE%D0%BC%D0%B0%D1%82%20%D1%88%D0%BB%D0%B0%D0%B3%D0%B1%D0%B0%D1%83%D0%BC%D1%8B&yclid=5200370865234540402.

GLOBAL CONNECTIVTY REPORT 2022

CHAPTERS 3-4. ACCELERATING PROGRESS TOWARDS UNIVERSAL AND MEANINGFUL CONNECTIVITY & THE CRITICAL ROLE OF MIDDLE-MILE CONNECTIVITY

Michael Kende, Sonia Livingstone, Scott Minehane, Michael Minges, Simon Molloy, and George Sciadas,
ICT Data and Analytics Division of the ITU Telecommunication Development Bureau, Geneva, Switzerland

DOI: 10.36724/2664-066X-2022-8-4-22-32

ABSTRACT

The Global Connectivity Report 2022 takes stock of the progress in digital connectivity over the past three decades. It provides a detailed assessment of the current state of connectivity and how close the world is to achieving universal and meaningful connectivity, using a unique analytical framework. It goes on to showcase solutions and good practices to accelerate progress. The second part of the report consists of seven thematic deep dives on infrastructure, affordability, financing, the pandemic, regulation, youth, and data. *Chapter 3* explores options to accelerate progress towards universal and meaningful connectivity. Expanding broadband networks is needed to eliminate the remaining blind spots and improve the quality of connectivity. Measures include reducing constraints on foreign direct investment to attract capital for upgrading and expanding digital infrastructure; ensuring sound ICT sector regulation to help build competitive markets and enhance predictability; promoting infrastructure sharing to reduce costs; ensuring the supply of adequate, inexpensive spectrum to help reduce coverage gaps; and ensuring sufficient capacity and a shift to new generations of mobile broadband. Solutions to ensure an adequate energy provision to power ICT infrastructure include policy incentives, reducing duties and taxes on green power equipment and allowing independent power producers. *Chapter 4* explores the importance of middle-mile connectivity as countries develop digital economies – achieving better quality, lower costs and greater redundancy. The "middle mile" consists of infrastructure responsible for storing and exchanging data. It is an overlooked yet critical link in the connectivity chain, between international connectivity – or "first-mile" connectivity – and "last-mile" connectivity, made of the infrastructure that connects the users, which is hence more visible and tangible. The three key components of a domestic data infrastructure ecosystem are Internet exchange points (IXPs), data centres and cloud computing.

KEYWORDS: *ITU, universal and meaningful connectivity, critical role of middle-mile connectivity, digital onnectivity, Internet of Things (IoT).*

INTRODUCTION

This chapter looks at potential solutions to accelerate progress towards universal and meaningful connectivity and mitigate the dangers of online threats to user security and safety. Consistent with the universal and meaningful connectivity framework introduced in Chapter 2, solutions are organized around these enablers: infrastructure, affordability, device, skills, and security and safety. The chapter also examines specific policy options to address the needs of disadvantaged groups and aspects of environmental risk.

Meaningful connectivity implies safety of use. Threats include a breach of data privacy, misinformation and harmful content, and overuse of digital technology. It is important to know how to mitigate risks to preserve trust in the use of the Internet. Countries need to enact better data protection laws to safeguard privacy, social media companies need to moderate content to detect false and inciteful content, and media literacy must be part of any digital skills training [2].

To achieve universal connectivity, disadvantaged groups such as persons with disabilities, older persons, those with low incomes and people living in remote areas, require special attention. Greater collaboration is needed across governments, agencies, advocacy organizations and digital companies to accelerate the acquisition of digital skills. To reduce the gender gap, non-governmental organizations should be supported in providing mentoring and digital skills training for women and girls. Technology companies, too, can play a role, not only by supporting skills initiatives but also by setting their own gender equity targets. Digital products and services should be customized to the needs of women in terms of design, safety and security. Training of older persons is necessary if they are to access online public services. Measures to reduce the digital disability gap include raising awareness, enacting laws that require online public services to be accessible to persons with disabilities, adapting products by adhering to international design guidelines, and supporting entrepreneurs in the development of contextually relevant digital assistive technologies. Since data are often lacking, there is a need to ensure that the scope of ICT surveys addresses disadvantaged groups as well.

Among the challenges posed by increased digital connectivity, e-waste continues to grow, and what happens to over four-fifths of e-waste is unknown. As a minimum, the recycling process should be made easier for consumers. Connectivity will help reduce carbon emissions across the economy, for example video conferencing for work and education will reduce travel while the greater use of sensors will generate energy efficiencies across many sectors. Furthermore, there is considerable untapped renewable potential from solar, wind, hydro and geothermal sources in many low- and middle-income countries. As major energy users, ICT companies can provide the scale of investment to make renewable energy economically feasible. Governments can help enormously by creating climate friendly energy strategies and liberalizing markets, particularly by welcoming independent renewable power producers.

INFRASTRUCTURE

This section outlines areas where government measures can expand high-speed telecommunication network coverage to achieve meaningful connectivity.

• Reducing constraints on foreign direct investment (FDI) can be effective in attracting capital to upgrade and expand digital infrastructure. Of the 43 low- and middle-income economies included in the OECD FDI restrictiveness indicator, only six were fully open to foreign investment in their telecommunication sector (OECD 2022). Such restrictiveness limits investment by large international telecommunication groups and the expertise and technology transfer they represent. Some countries profess to have a liberalized sector but often impose restrictions, particularly when governments retain a stake in telecommunication operators.

• Ensuring sound ICT sector regulation will help build competitive markets and enhance predictability, attracting investment. The ITU ICT Regulatory Tracker measures regulatory performance among countries with a framework that identifies how far countries have travelled on their regulatory journey and which ‘generation of regulation’ they fit into: G1 indicates regulated public monopolies with a command and control approach; G2 indicates basic reform with partial liberalization and privatization; G3 enables investment for innovation and access, has dual focus on stimulating competition in service and content delivery, and provides for consumer protection; and G4 indicates integrated regulation, led by economic and social policy goals. A fifth stage of regulation (G5) is a collaborative generation of regulation where digital transformation is promoted across all sectors of the economy. Almost 40% of countries are at the G1 or G2 generation of regulation, hampering their ability to expand connectivity (Figure 1).

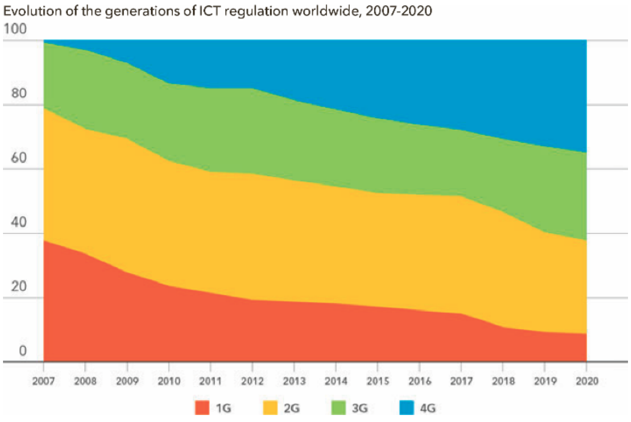


Figure 1. Performance in the ITU ICT Regulatory Tracker

Source: ITU

• Regulation can improve investment but can also introduce additional rules and costs. A light regulatory touch can result in competitive markets with higher adoption and cheaper prices, particularly in low-income countries. Seventy per cent of the population use mobile money services in Somalia for example, ahead of most countries in Africa. In Cambodia, light touch regulation has stimulated competition and foreign investment. Cambodia also leads LDCs in data usage and mobile-phone ownership, is second for mobile-broadband affordability, and is one of few countries where ownership is higher among women than men.

• Promoting the sharing of infrastructure can reduce costs. Operators could, for example, share mobile towers and underground ducts. Network deployment investment is reduced by laying fibre-optic cable along railway lines, power transmission grids and pipelines. Estimates suggest that sharing antenna sites can save operators up to 40% on both capital expenditure and 5G deployment.

• Ensuring the supply of adequate, inexpensive spectrum can help reduce coverage gaps, ensure sufficient capacity and support the shift to new generations of mobile broadband. Low frequency spectrum is needed for rural areas, as it provides wide coverage, requires fewer sites and reduces investment costs. Challenges are delaying progress in this area. In some countries migration from analogue to digital television has been delayed, reducing availability of low frequency spectrum for mobile broadband use. Some countries also auction frequencies with high reserve prices, thereby raising investment costs, which results in higher prices for users. Some governments charge recurring fees for spectrum, raising the cost of deploying infrastructure in rural areas.

• Ensuring that energy provision is adequate to power ICT infrastructure is essential. This is a challenge in some low- and middle-income countries especially in remote rural locations. Diesel is often used but this is expensive and unkind to the environment (GSMA 2020b). Renewable solutions are not always feasible or price competitive, for instance because of a lack of sunlight, infrequent wind, or the need for expensive battery storage. Solutions to these challenges include reduced duty, tax incentives on green power equipment, and allowing independent power producers.

• Recalibrating universal service funds (USFs) can help deployment of infrastructure in unserved areas. Many funds have been unsuccessful, suffering from challenges such as poor design, mismatches in funds collected and disbursed, political interference, lack of training and education, and maintenance and energy supply.

Niche technologies for expanding telecommunication infrastructure

Although a universal solution has yet to be found, a variety of technological solutions to cheap access for people living in rural and remote regions have been available for many years. Examples of such technological solutions include:

• TV white space (TVWS) utilizes buffer frequencies between TV channels to provide broadband Internet access. In remote parts of Colombia cellular coverage has not been feasible due to frequency bands being congested, high licensing costs and limited communication range. TVWS is being used as an alternative to connect rural schools and coffee plantations in geographically challenging locations such as mountainous rainforests.

• High altitude platform service (HAPS) such as Loon (operated by Alphabet, the parent company of Google) uses a network of hot air balloons to provide connectivity to unserved locations. Loon was used during floods in Peru in 2017 as well as in Kenya to provide Internet access to a region covering 50 000 sq km. Loon stopped operating in January 2021 as it could not be made commercially viable.

• Networked tethered flying platforms (NTFP) are tethered gas balloons. Due to their altitude, an NTFP can replace numerous regular cell towers, lowering costs. NTFPs are being proposed for use in Australia where 70% of the land mass has no cellular coverage.

• Satellites provide backbone transmission services as well as direct to consumer television and broadband access. Low earth orbiting (LEO) satellites blanketing the Earth delivering affordable service to handheld devices have been promoted as a solution for remote areas but remain unaffordable for many low- and middle-income countries. LEOs are providing important backhaul transmission services to the Internet in landlocked or remote islands. They can be a useful backup when terrestrial systems are damaged, for example if an undersea volcano were to damage a submarine cable, as was the case in Tonga, or other disasters disrupt the Internet network.

In addition to the niche technologies above, improvements in wireless cellular technologies are lowering the cost of deploying last-mile access. The OpenRAN project is promoting the use of inter-operable open source software and hardware to reduce the cost of proprietary products.9 Moving to a cloud-based, software-driven environment can lower the cost of cellular networks. In Japan, Rakuten launched the world’s first cloud-based mobile network, claiming 40% lower costs than those of traditional cellular networks.

Many people are learning digital skills without formal training, resulting in shortcomings in acquiring further skills. They use social media acquiring basic skills from family and friends People with limited literacy in Africa have used a simple customized version of the Internet with audio and icon-based interfaces. These applications often mean people are ‘unconscious Internet users’, not knowing what the Internet is or that they are actually using it, and therefore unaware of the variety of uses, benefits and risks it can bring. Informal training often omits important security skills such as protecting privacy, for example, minimizing the digital trail left on social media and elsewhere. Nor does it teach how to distinguish between fact and misinformation. The result is an urgent need to train millions of people formally in using the Internet to ensure they have safe and meaningful connectivity [1].

COVID-19 has seriously hampered the provision of face-to-face digital literacy training. Although programmes have moved online, they are not practical for those who have never used the Internet. If there is no other option, courses should be provided in a webinar format with instructors able to interact with students.

DIGITAL SKILLS

Overcoming digital illiteracy is critical to shrinking the usage gap. Effective and large-scale programmes are needed to address the challenge. Providing digital literacy as part of the school curriculum is a solution for those at school. Recent data on how many countries include digital skills training in the curriculum is not available. Data compiled a decade ago indicate that 55% of countries included basic computer skills training for primary schools and 74% for upper secondary schools [3-4].

Worldwide only 40% of primary, 51% of lower secondary and 66 per cent of upper secondary schools had Internet access in 2020. Giga, a partnership between UNICEF, ITU and the private sector, seeks to connect every school to the Internet. The programme has shown that schools can be "anchor tenants" in a community, extending access and digital skills to those living close by. Funding school connectivity remains a challenge however, with many low- and middle-income countries struggling to build schools with electricity let alone Internet access. Increasingly, the private sector is helping to support digital literacy in schools.

Ensuring school connectivity and digital skills

The private sector plays a key role connecting schools in certain countries (World Benchmarking Alliance 2020). Safaricom's 47-in-1 Initiative is installing a computer lab in one primary school in every county in Kenya. Mobile operator Millicom has committed to the Organization of American States (OAS) goal of connecting every public school in Latin America and the Caribbean to the Internet by 2030, providing Internet access to 2 000 schools throughout the region. Vodafone's Instant Network Schools provides Internet to schools with refugee students. Launched in 2013 in partnership with UNHCR, it has provided school connectivity to 36 schools in five African countries reaching over 86 500 refugees.

Those not in school and without digital skills also need to be reached. The Rwanda Government launched the Digital Ambassador Program (DAP) with the target of training 5 000 youth and sending them all over the country to provide digital skills training to 5 million people. By December 2019, DAP had reached nearly 50 000 people. An evaluation of DAP made specific, practical recommendations to further enhance its impact: i) greater community outreach to increase participation; ii) minimizing technical aspects; and iii) linkages to programmes such as mobile money, device and service charge affordability and national content.

The private sector is providing digital literacy training to adults. The Mobile Internet Skills Training Toolkit (MISTT) was developed by GSMA for mobile operators. Available in Bengali, English, French, Hindi, and Kinyarwanda, MISTT uses a 'train the trainers' approach, whereby staff from the mobile operator train sales agents who then teach customers. MISTT has been used in countries throughout South Asia and Africa. South Africa mobile group MTN offers MISTT in eight African countries and, as of April 2021, has trained over 18 million people, finding that incentives (commissions for trainers and free data for trainees) had a real impact.

AFFORDABILITY

The cost of devices and Internet use represents a major barrier to connectivity. This section sets out recommendations on what can be done to overcome this challenge both in regard to devices and services.

Device affordability

The price of a device is a significant barrier that stops many people developing digital skills. Price reduction has its challenges, however. Very few countries manufacture and therefore control pricing of these products, and importing countries have no say in how the pricing is arrived at. Three approaches set out below offer promise.

Governments do, however, affect device price (and therefore affordability) through imposing import duties and sales taxes. The World Trade Organization (WTO) Information Technology Agreement (ITA) calls for countries to eliminate duties on information technology products. Despite the initiative having 82 signatories, many of the world's poorest countries, particularly in Africa where the impact could be greatest, have not signed. Although sales taxes serve a purpose, taxes on devices should be kept relatively low and certainly not higher than for other products.

A4AI has carried out research on smartphone pricing. They found that the average world price in 2021 was around one-quarter of monthly income, that in South Asia the figure rises to 40%, and in the LDCs it is 53%. Among its recommendations for lowering device prices, A4AI calls for using USF funding to subsidize the cost, highlighting the examples of Malaysia and Costa Rica (A4AI 2020).

Some operators are playing their part to lower the costs of handsets. Mobile group MTN which operates in 21 countries throughout sub-Saharan Africa and the Middle East has launched several initiatives (MTN 2021). Working with Chinese manufacturers, MTN introduced a handset that costs less than USD 40 across its markets. In Zambia it is subsiding handsets and in Uganda it offers customers an installment plan amounting to USD 0.17 per day.

There is a market opportunity for low-cost manufacturers. TECNO, the brand of the Chinese mobile phone manufacturer Transsion, has the highest mobile phone sales in Africa because it sells affordable handsets.

Service affordability

In over half of countries worldwide, ITU analysis suggests broadband services remain unaffordable. Governments can however take action to remedy this in these three areas:

1. Reduce taxes on services to make them more affordable. In 2017, of total payments made by mobile operators to governments, almost a third was specific to the mobile sector (mobile consumption taxes, spectrum and licence

fees, etc.). This was in addition to other, economy-wide, general taxes paid by telecommunication operators and consumers. Reductions in sector-specific taxes enhances affordability and increases demand, with spillover effects on other industries. GSMA studies find that increased demand from lowering taxes and indirect impacts across the economy raise the tax base, off-setting the loss of sector-specific taxes. Uganda for example has a range of taxes that negatively impact affordability (Stork and Esselaar 2018). In addition to value-added tax, the government levies a mobile services excise tax and an Internet data tax that has replaced a social media tax. Almost half of what is spent on mobile airtime in Uganda consists of taxes.

2. Governments should encourage operators to offer plans that reflect different income levels and circumstances and that offer a minimum of 2 GB data a month for the cheapest plans. In almost all low- and middle-income countries, prepaid and data-limited mobile offers dominate Internet access packages. In Zambia, for example, mobile operators offered 17 plans ranging from a one-hour plan featuring 5 MB of usage, to weekly bundles offering unlimited access to popular social media services such as Facebook and WhatsApp. An ITU report found such bundles successfully enabled access to mobile Internet for lower-income users at low cost. This illustrates that while affordability need not be a barrier to Internet use, it limits how much is consumed and when it is consumed, a less than perfect solution when measured against the aspiration of universal and meaningful connectivity.

3. Make mobile data more affordable in a world where 6 GB a month is reasonable. COVID-19 has made users look at data consumption, one hour of Zoom for example consumes between 0.5 GB and 2.5 GB.26 Data consumption patterns vary widely across the world and generally relate to income levels. ITU data for 2020 show that an individual in Finland and Kuwait, for example, consumed 30 GB of mobile data a month in contrast to less than 1 GB for those living in 21 low- and middle-income countries. The volume of monthly data that a person would need to access key online activities was recently estimated at 660 MB per user per month and included access to public services, health information, shopping, learning, and news. When recreational activities were included, the estimated volume of data rose to 6 GB per month (an extra 5.2 GB). Such a monthly data package in the six low- and middle-income countries included in the study costs more than 2% of income for the bottom 40% of the population.

However, there are concrete measures that can make data more affordable in low- and middle-income countries. Governments can:

- Ensure provision of unlimited broadband access to community centres and schools, with access to those in the surrounding community who cannot afford it at home.
- Ensure that the temporary COVID-19 concessions that were put in place by operators in many countries (higher data allowances or providing free Wi-Fi) are maintained for the poorest segment of the population , those needing medical support and for students.
- Subsidize data use for the poorest segment of the population through social tariffs similar to those for food allowances.
- Apply zero ratings for critical services such as e-government, education and health services.
- Create haritable data donation schemes. In Australia for example, users can donate their unused monthly data to those in need.

SECURITY AND SAFETY

To be sustainable, meaningful connectivity must equate to having limited or no risks associated with connecting to the Internet. This section explores the nature of online threats to user security and safety, and considers personal data, misinformation, overuse of digital technology, and vulnerability of children.

According to a global survey carried out in 2019, eight out of ten Internet users are concerned about their online privacy and one in four do not trust the Internet. Over a third of Internet users in the European Union experienced a security incident of some description in 2019. Personal data breaches, online harassment, children accessing inappropriate websites, hacking, viruses, pharming and phishing, and the spread of misinformation are just some of the negative consequences of going online.

Protecting personal data is a critical issue but only 23% of countries around the world have adequate data protection laws on a par with the EU General Data Protection Regulation (GDPR). One source reports that 69% of countries have data protection laws, however, many are not implemented; do not adequately reflect present day user needs; often require no user consent for use of personal information; offer limited control mechanisms for transferring personal data abroad; and lack provisions for a data protection authority [5-9].

Countries that are falling short need to create adequate data protection laws or update their existing laws in line with best practice, and many telecommunication operators and platform companies should exceed minimum duty of care requirements and put in place a single policy that meets international best practice to ensure the security and safety of their customers.

The spread of misinformation is rising steeply, driven by social media platforms. Analysis of social media use in the United States found that 17% of information from among the top 100 news platforms came from unreliable sources, up from 8% in 2019. The World Health Organization noted that the spread of misinformation about COVID-19 is “proving to be as much a threat to global public health as the virus itself”. Top social media platforms have begun to label or take down false information, but ex post facto action is often too late. Hate speech is now a major concern and has led to documented violence against ethnic minorities.

Social media companies should take more action, too, for example by increasing moderators on the ground in all countries to detect false and inciteful content.

In situations of political conflicts they need to come to a balanced judgement on the type of content they restrict. They need to demonstrate greater transparency of how platforms use algorithms to disseminate content, or add features that discourage the sharing of harmful content or that limit the spread of viral content.

The overuse of digital technology is a now a recognized health risk with a range of dangers. Gaming addiction is estimated to affect around 5% of the population. Internet addiction is also recognized in many countries, for example, in Germany the rate has been estimated at 2%, while in Bangladesh over a quarter of young adults are Internet addicted. Efforts to limit online gaming addiction include parental controls, limited access set by some online gaming companies and, in China for example, restricted access for those under 18 (Figure 2).

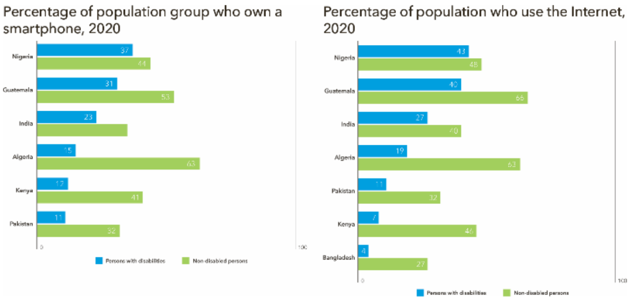


Figure 2. Disability gap for smartphone ownership and mobile Internet use

Note: Based on survey results for adults aged 18 and over. n=49-260 for persons with disabilities and n=900–1 866 for persons without disability.

Source: GSMA Consumer Survey 2020

ACCELERATING CONNECTIVITY AMONG DISADVANTAGED GROUPS

To attain universal connectivity, special attention must be paid to the needs of disadvantaged groups including persons with disabilities, older persons, women and girls in some countries, those with low-incomes, and people living in remote areas. People with one or more disadvantages are at greater risk of digital exclusion (for instance women with low-incomes and older persons with disabilities.) Other groups at risk are country specific, such as migrants, refugees or ethnic minorities [10-14].

Persons with disabilities

It is estimated that 1 billion people have a disability or about 15% of the global population. Global statistics about the connectivity status of persons with disabilities do not exist. GSMA has collected data for some middle-income countries that indicates significant gaps separating persons with disabilities and the rest of the population in smartphone ownership and Internet use.

Many leading hardware manufacturers have adapted products to be more disability friendly through features such as enabling large fonts and screen readers, and many adhere to the W3C global standard for web accessibility. Designed with disability experts, the GSMA Principles for Driving the Digital Inclusion of Persons with Disabilities offers guidelines for the mobile industry to reduce the gaps in access and use.

Innovation is an important means of empowering persons with disabilities, for example, persons with visual impairment are using smartphone technology to scan and read documents, to get accessibility ratings for public places and audio and vibration alerts for approaching obstacles.

Older persons

Available survey data indicate gaps between rates of Internet usage by age group. Young people use it most and older persons use it least. In Norway, for instance, the Internet use gap is much less pronounced at 92% for the 75+ age group compared with 99% for the 15-74 age group. However, for most economies the age gap is wide, at more than 50 percentage points in over half of economies providing data and 80 percentage points in Kazakhstan (Figure 3).

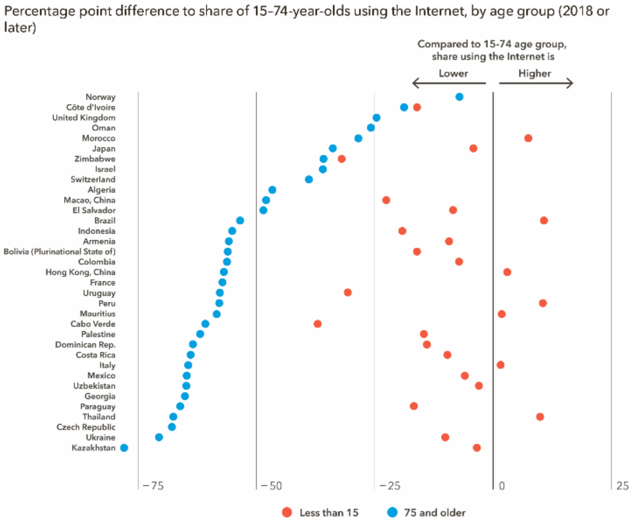


Figure 3. The inter-generational gap in Internet use

Source: ITU

Digital skills training for older persons is an imperative for governments if older persons are to access online public services.

The New Zealand Government has allocated funding to train up to 4 700 older persons in digital skills over three years. Some digital companies are providing training, for example in Singapore, Singtel is upgrading community centres with Internet access and tablets, its staff volunteer for one-on-one digital skills training, and it opens its shops early to provide training workshops.

Training should be designed for and delivered exclusively to older persons. Design should take comfort levels, learning relevance and application focus into account. Course numbers should be small and include modules on security to build trust in using online services. Training should be ongoing to reinforce learning and have a lasting effect.

CHAPTER 4. THE CRITICAL ROLE OF MIDDLE-MILE CONNECTIVITY

INTERNET EXCHANGE POINTS

As IXPs grow to handle an increasing amount of data, they are relocated to a professionally managed data centre, allowing companies that need to exchange data to be closer to the IXPs and with their servers located in the same data centres. Similarly, as demand increases for cloud computing, service providers also situate data centres to be closer to customers. The IXPs sit at the core of this ecosystem (Figure 4).

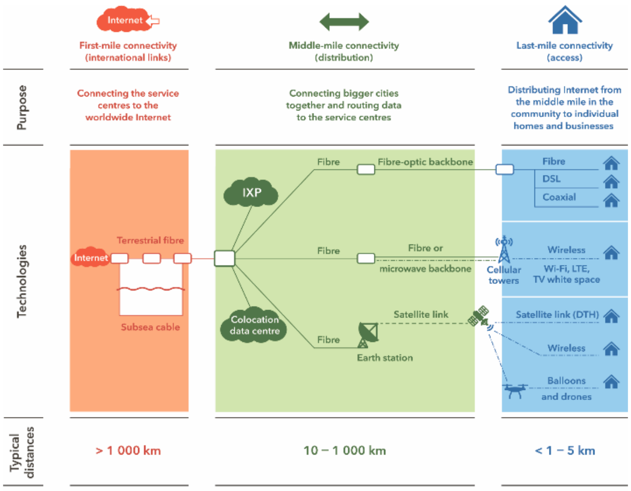

Internet
First-mile connectivity (international links)
Middle-mile connectivity (distribution)
Last-mile connectivity (access)
Purpose
Connecting the service centres to the worldwide Internet
Connecting bigger cities together and routing data to the service centres
Distributing Internet from the middle mile in the community to individual homes and businesses
Technologies
Internet
Terrestrial fibre
Subsea cable
IXP
Colocation data centre
Fibre
Fibre-optic backbone
Fibre
Fibre or microwave backbone
Fibre
Satellite link
Earth station
Cellular towers
Fibre
DSL
Coaxial
Wireless
Wi-Fi, LTE, TV white space
Satellite link (DTH)
Wireless
Balloons and drones
Typical distances
> 1 000 km
10 – 1 000 km
< 1 – 5 km

Figure 4. The connectivity chain and the "miles" of connectivity

Note: IXP = Internet exchange point; DSL = digital subscriber line; DTH = direct-to-home; LTE = Long-term Evolution.

Source: Adapted from World Bank (2021)

IXPs are a core component of data infrastructure, enabling Internet service providers (ISPs) and content providers to exchange their data traffic – known as "peering". The IXP method of data exchange offers substantial advantages, some of which are listed below:

- It is less costly than using international bandwidth, since traffic is not sent back and forth over costly overseas links. Latin America could slash by one-third the USD 2 billion a year it spent on international bandwidth through greater use of IXPs, according to one study. Studies for Kenya and Nigeria also find that IXPs reduce overseas payments and improve latency.
- ISPs do not need to make peering agreements with each potential partner.
- Redundancy is enhanced, since countries do not rely on international bandwidth if there is a disruption.
- IXPs also improve quality, since they are situated closer to the user and hence have less latency.
- IXPs reduce the time it takes to retrieve data, enhancing user engagement. In Rwanda, it is 40 times faster to access a local website (<5 milliseconds) compared with those hosted in the United States or Europe (>200 milliseconds).

IXPs grow organically – and attract big content providers

IXPs begin as locations for ISPs to exchange traffic. Initially, this may not amount to much traffic, since in many developing countries locally relevant content is limited or is hosted abroad. IXP participation has grown more diverse over time, now frequently including companies, governments, content providers and cloud operators as members. Diverse and growing participation also stimulates demand for data centres, boosting the economy. Companies also want to be closer to end users to reduce latency and enhance the Internet experience [15-20].

IXPs also reduce the need for international bandwidth due to a reversal of network routing. Instead of countries having to pay international transit fees to access content overseas, large content and cloud providers are increasingly moving to IXPs (Table 1). These companies handle the backhaul to their data centres, on occasion through their own submarine cables. Content providers have now overtaken telecommunication carriers as the largest users of international capacity. Three content providers – Google, Facebook and Netflix – account for two-thirds of all mobile application traffic, highlighting the importance of attracting content providers to IXPs.

Table 1

Top 10 companies by presence on an IXP, December 2021

Company	ASN*	Type	Number of IXPs present on
Hurricane Electric	6 939	Network service provider (NSP)	275
Cloudflare	13 335	Content delivery network (CDN)	263
Packet Clearing House	3 856	Educational/Research	212
Google	15 169	Content	207
Microsoft	8 075	Content	194
Akamai	20 940	CDN	182
Facebook	32 934	Content	168
Amazon	16 509	Enterprise	129
Subspace	32 261	CDN	116
Netflix	2 906	Content	107

**ASN = Autonomous System Number uniquely identifying organizations routing traffic over the Internet.*

Note: There were almost 24 000 organizations with an ASN in December 2021. CDNs deliver other companies' data to the IXP, whereas "Content" refers to companies that deliver their own content to the IXP.

Source: PeeringDB (www.peeringdb .com)

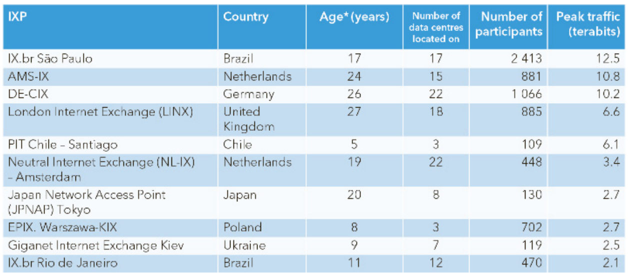
According to Packet Clearing House, there were 726 active IXPs around the world in December 2021. Despite the benefits of an IXP, 65 countries and territories do not have one. These are mainly countries where there is only one ISP or are small island States where the volume of domestic traffic may be insufficient to warrant an IXP. In contrast, a number of countries have multiple IXPs – much needed to reduce latency in large countries with dispersed populations.

Multiple IXPs also deliver redundancy and, through competition, are likely to reduce costs of use. However, introducing multiple IXPs in a country in the early stage of middle-mile connectivity risks reducing the scale of the IXP and its attractiveness to content providers.

The top 10 IXPs by the volume of traffic exchanged speak to well-developed ecosystems with high levels of participants (Table 2). While most are based in high-income nations, two entries are based in Brazil and one in Ukraine. This top 10 group boasts an average age of 17 years, reflecting the importance of experience in developing an efficient IXP. Most have hundreds of participants and are available in multiple data centres to better reach their customers. Some are expanding operations into other countries. For instance, Deutscher Commercial Internet Exchange (DE-CIX) is available in 16 other countries. Of these, nearly all are high- and upper-middle-income nations – but this model could be more widely applied through partnerships in developing nations.

Table 2

Top 10 IXPs by traffic exchanged, December 2021



IXP	Country	Age* (years)	Number of data centres located on	Number of participants	Peak traffic (terabits)
IX.br São Paulo	Brazil	17	17	2 413	12.5
AMS-IX	Netherlands	24	15	881	10.8
DE-CIX	Germany	26	22	1 066	10.2
London Internet Exchange (LINX)	United Kingdom	27	18	885	6.6
PIT Chile – Santiago	Chile	5	3	109	6.1
Neutral Internet Exchange (NL-IX) – Amsterdam	Netherlands	19	22	448	3.4
Japan Network Access Point (JPNAP) Tokyo	Japan	20	8	130	2.7
EPIX. Warszawa-KIX	Poland	8	3	702	2.7
Giganet Internet Exchange Kiev	Ukraine	9	7	119	2.5
IX.br Rio de Janeiro	Brazil	11	12	470	2.1

Age = From the year it was established.*
Sources: PeeringDB (www .peeringdb .com) and IXP websites

The existence of an IXP does not guarantee its potential benefits. Although the number of IXPs has grown in developing nations, many in low-income nations are stuck in first gear, with few participants and very little traffic. Average membership per IXP in low-income nations is 9, compared with a world average of 57.

Interestingly, upper-middle-income economies – not high-income economies – have higher membership levels and traffic per IXP. This is because large countries such as Brazil, Russian Federation and South Africa have well-developed IXP ecosystems and boast some of the largest IXPs in the world.

Regionally, there are also notable gaps. Europe generates 260 gigabits per second (Gbit/s) of traffic per IXP, the highest of any region. Most long-established IXPs are European, with many years of experience. On the other hand, Africa (excluding South Africa) has on average 14 participants per IXP, compared with a world average of 57, and generates just 9 Gbit/s per IXP, compared with a world average of 173. While South Africa accounts for just over 10 per cent of the continent's IXPs.

Stages of IXP growth

IXPs progress in stages, and each higher stage of development increases its impact. The first IXPs are typically established by universities or as non-profit associations of ISPs. They are located in small server rooms, with technical tasks carried out by volunteers. As traffic increases, and new participants join, a more sustainable technical and operational environment becomes necessary – more formal governance, staff hiring and equipment upgrades. The final stage sees many participants wishing to join without having to deploy a physical connection to the exchange. Multiple IXPs in different locations are created and the IXP is relocated to a colocation data centre (discussed below).

Developing countries are at different stages of maturity in regard to IXPs. At one end of the scale are countries with no IXPs, while at the other are countries that boast a dense fabric of multiple IXPs located in connected data centres, usually operated by the private sector and with many different participants. As countries progress through the stages, prices drop, performance improves and traffic increases (Figure 5).

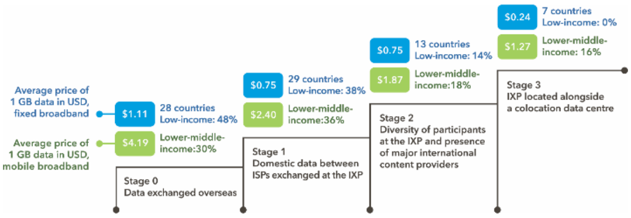

Average price of 1 GB data in USD, fixed broadband
Average price of 1 GB data in USD, mobile broadband
$1.11
28 countries Low-income: 48%
$4.19
Lower-middle-income:30%
Stage 0 Data exchanged overseas
$0.75
29 countries Low-income: 38%
$2.40
Lower-middle-income:36%
Stage 1 Domestic data between ISPs exchanged at the IXP
$0.75
13 countries Low-income: 14%
$1.87
Lower-middle-income:18%
Stage 2 Diversity of participants at the IXP and presence of major international content providers
$0.24
7 countries Low-income: 0%
$1.27
Lower-middle-income: 16%
Stage 3 IXP located alongside a colocation data centre

Figure 5. IXPs and stages of growth
Sources: ITU; Srinivasan et al. (2021)

Data centres

Data centres provide space, power and cooling for servers hosting data and network cabling. They play a fundamental role in the digital economy by the storing of data, and the local hosting of domestic content. In addition, data centres offer a significant advantage when processing large volumes of data – and in the development of big data services [21-24]. Their presence is also a measure of the digitalization of the economy, reflecting demand from the information and communication technology (ICT) sector and beyond – finance and insurance, transportation, legal and accounting activities, research and development, advertising and the public sector. IXPs also benefit from data centre hosting, attracting more diverse participants and enjoying more professional facilities [25-28].

Data centres can be classified into four broad categories:

- Enterprise data centres are single-tenant facilities owned by a company to store data. They are either located on a company's site or in a dedicated facility off-site.
- Carrier data centres are provided by telecommunication operators to host their clients' data. This has historically locked clients into exclusive use of the operator's data centre services. However, growing numbers of telecommunication operators are now providing "carrier-neutral" connectivity.
- Multi-tenant data centres (MTDCs) are operated by companies that rent out space for data storage. Leading operators usually have certifications for security and reliability.
- Hyperscale data centres belong to major content and cloud providers such as Facebook, Google, Amazon and Microsoft (who together account for over half). There were close to 600 hyperscale data centres at the end of 2020, more than doubling in five years (Synergy Research Group, 2021).

In the past, data centre connectivity was often via a direct telecommunication link between a company's office and the data centre. Today, however, more flexible communication links from the data centre to the outside world are essential – many companies may be using the centre and employees can also be at disparate sites throughout a country or the world and, working from home.

Three ways to organize data centre connectivity

Data centre connectivity can be accomplished in different ways. In some instances, telecommunication companies may operate MTDCs but require the tenant to use their services – this can result in higher prices and a lack of flexibility. A further option is for an IXP located in the data centre to handle data exchange – an attractive option, since this is often done through free peering arrangements. The benefit is magnified when content and cloud providers are also in the data centre. A third option is carrier-neutral MTDCs operated by companies that do not provide telecommunication services – tenants are free to use any telecommunication provider to handle their data transport needs. Notably, some telecommunication operators now offer open peering MTDCs and own IXPs.

Global overview of data centres shows large disparities

PeeringDB provides a global listing of companies that exchange traffic over the Internet and the data centres they are located in. Globally, there were 4 300 data centres connected to the Internet in November 2021. Large disparities exist in connected data centre penetration: 57 economies do not have a connected data centre. While connected data centre penetration is 2.7 per million inhabitants in high-income nations, it is considerably less in low- and middle-income nations. Similarly, sharp regional disparities exist – with a penetration of more than 1.5 connected data centres per 1 million inhabitants in Europe and North America, compared with less than 0.5 in other regions. While such disparities are related to income and demand for large-scale data storage, they are also caused by a lack of complementary infrastructure (particularly energy) and by policies that have inhibited private investment [29-30].

Another view of data centre dispersion is to examine where the leading carrier-neutral MTDC operators are headquartered. The MTDC big picture is dominated by United States-headquartered operators, including the two largest, Digital Realty Trust (DRT) and Equinix, with some hundreds of data centres between them. Of the 2 113 organizations with a connected data centre, 1 565 (74%) report operating just one. The top 20 MTDCs account for less than 1% of organizations offering connected data centres – but do account for over a quarter (27%) of the total data centres and 74% of the total of those operating more than five data centres.

Mapping data centre locations of the 20 largest MTDCs reveals stark geographical gaps (Figure 6). Dense concentrations occur in developed regions such as North America and Western Europe, the powerhouses of the digital economy, while in much of the rest of the world there are none.

Data centres are costly to build, and in many low- and middle-income countries, the private sector lacks the capital and necessary expertise.

Major MTDC operators such as Equinix, DRT and NTT rarely have data centres in low- and middle-income countries. However, some MTDC companies partner with local investors to build data centres. In India, EdgeConneX (2021), a large MTDC operator, is partnering with local company the Adani Group to help build six data centres. Private companies with a regional focus are also operating MTDCs in developing countries – several companies are building data centres in Africa.

Development partners are providing investment funding in government data centres, but companies often do not want to locate in State-owned facilities. In 2021, China loaned Senegal USD 18 million for a government data centre, with Chinese company Huawei providing equipment and technical support. In Togo, also in 2021, the World Bank provided USD 24 million to the Government for the country's first world class data centre, providing space for non-government tenants. The centre is built by French company APL, and managed by Africa DataCenters, which operates nine facilities throughout the continent.

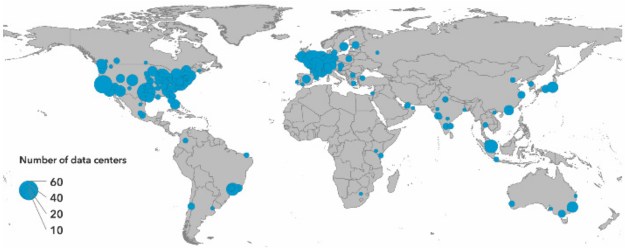


Figure 6. Data centre locations of top 20 MTDC operators

Note: The size and colour of the dots refer to how many data centres there are in that location. For example, there are 63 in London.

Source: Compiled from locations reported by top 20 MTDC operators

Cloud computing

Cloud computing has transformed data storage and analysis by allowing users to access scalable data storage and computing resources as needed. As broadband connectivity has boomed, delays associated with remote storage, processing and analysis of data have dropped significantly – and cloud use by businesses and governments has taken off. Cloud computing is especially attractive, since it helps avoid costs associated with maintaining on-site hardware, software and storage.

Microsoft Azure, Amazon Web Services and Google Cloud – large firms from the United States –dominate the cloud. They have hyperscale cloud data centres, most located in developed countries with stringent data protection and sovereignty regulations.

However, the lack of a cloud data centre in a country is overcome through "on-ramps" to cloud computing services. Customers can meet cloud providers at IXPs located in colocation data centres, avoiding costly international transit to access cloud services – and enjoying greater security and reliability, improved performance and reduced costs.

Cloud and content providers have emerged as some of the largest investors in backbone infrastructure, including submarine cables to route traffic from MTDCs to their own hyperscale data centres. Countries no longer need to spend money on international bandwidth to access popular content and the cloud, since the providers will come to a country's MTDC if conditions are favourable.

Big data analysis and sharing applications are often available only on the cloud, and while it may seem attractive to store data on the cloud, there are three factors to consider. First, it can be costly to store data on the cloud. Organizations often use a "hybrid cloud approach",12 storing on the cloud only the data needed for cloud analytics. Second, latency is a key issue for applications such as finance and gaming, if stored on the cloud. Third, national security grounds may dictate that sensitive data be stored in the country – the cloud data centre needs to be located in the country and the cloud provider needs to adhere to national data laws.

CONCLUSIONS

There is no single pathway to universal and meaningful connectivity. The scope and nature of intervention depends on where a country stands on the path from basic connectivity for the few to meaningful connectivity for all. Multiple factors are at play, including a country's institutional framework, income level, demographics, geography, and culture that require a range of options, rather than a single solution, and which can differ significantly across countries within a region.

There is a huge divide in core data infrastructure between high-income and other countries. Many low- and middle-income economies have inadequate data infrastructure that cannot support transformation to sustainable digital economies – and which function at higher cost and with poorer quality. While investment has in the past been flat because of a perceived lack of demand, many such countries have seen accelerated use of the Internet, spurred by COVID-19. Investment limitations in core data infrastructure persist, however.

Scale is critical. Private investment in data centres has not been forthcoming in countries with small populations – though possibilities are emerging. Smaller, energy-efficient facilities are increasingly viable, as are schemes involving countries working together on regional data infrastructures featuring Internet exchange points. Most low- and middle-income countries in fact increasingly have the scale to attract investment, especially in view of data infrastructure operators' need to be close to customers, to reduce latency. Often what holds back investment is the absence of an enabling environment and an immature data infrastructure ecosystem – it does take time for IXPs to achieve large scale. In short, countries need to build data ecosystem environments that attract investment.

There are five building blocks to create a more conducive environment for middle-mile connectivity.

- *Liberalization:* Liberalization of the telecommunication market fosters growth in core data infrastructure. Deregulation increases investment opportunities and provides businesses more options in their choice of providers. Introducing unhindered competition in the international transit market would benefit IXPs, making large ISPs less dominant and more likely to join an IXP. For example, South Africa attributes its leading data centre position in Africa to the early liberalization of the telecommunication market. Similarly, Equinix, one of the world's largest multi-tenant data centre (MTDC) operators, entered the Mexican market following the country's 2013 telecommunication reform.
- *Data protection:* Data protection laws are especially important for attracting investment into MTDCs and cloud computing. Such laws stimulate investment if they require certain data to be stored in the country – and offer protection to investors who are looking to limit reputational risk arising from data breaches. Europe has the highest share of countries (96%) with a data protection law, due to the 2018 introduction of the EU General Data Protection Regulation. Since then, a growing number of economies – including China, Japan, Singapore, Thailand, India, Brazil and the United States state of California – have adopted data protection regulations. Globally, two-thirds of countries have data protection laws, but a number of developing countries have yet to adopt one.
- *Energy:* Data centres consume a lot of energy in powering servers and keeping them cool – a challenge that has become more pointed in the context of the climate crisis. Investors have been more focused on a strategic path towards carbon neutrality than on price. Governments could facilitate investment in this regard by liberalizing energy markets, thereby allowing independent renewable power producers and suppliers to enter the market. With set targets for carbon neutrality, most major MTDCs prefer renewable sources to be available in countries where they invest. The largest hyperscale data centre owners are the world's leading buyers of renewable power purchase agreements.
- *Collaboration:* This is essential across the many parties involved in a country's data infrastructure – governments, IXPs, ISPs, data centre operators and investors (such as development partners, content developers and cloud providers). Governments need to grasp the vital role that IXPs play in developing a country's data ecosystem – and put in place enabling policies, strategies, laws and regulations. Developing countries should pursue partnerships with large IXPs, providing capacity-building as well as helping to establish facilities. Those developing countries with enabling data infrastructure policies need to market more robustly their advantages, thereby encouraging private sector investment. While some development agencies have supported IXPs and data centres, more work can be done.
- *Key metrics:* There are no official international sources of key metrics for IXPs and data centres at a country level – in spite of the great importance of data infrastructure. Improving the availability of key statistics on the digital

economy at country level is essential. Timely, comparable and reliable statistics on data infrastructure are essential for countries to measure their performance and better understand the relationship between international and domestic traffic exchange. Several organizations collect relevant administrative statistics related to IXPs and data centres, and many IXPs and MTDC operators also report on their activities. Groups such as the Expert Group on Telecommunication/ICT Indicators (EGTI) could partner with those already collecting relevant statistics – to identify and define key indicators, to review and harmonize existing data sets.

REFERENCES

[1] A. Alhassan, and A.A. Kilishi, "Weak economic institutions in Africa: a destiny or design?". International Journal of Social Economics. 2019. Vol. 46. No. 7, pp. 904-919. doi.org/10.1108/IJSE-12-2018 -0651

[2] Alliance for Affordable Internet. 2020. From luxury to lifeline: Reducing the cost of mobile devices to reach universal Internet access. Web Foundation. https://a4ai.org/research/from-luxury-to-lifeline-reducing-the-cost-of-mobile-devices-to-reach-universal -Internet-access/

[3] World Bank. 2021. World Development Report 2021: Data for Better Lives. Washington, DC: World Bank. Available at https://openknowledge.worldbank.org/handle/10986/35218.

[4] Broadband Commission for Sustainable Development. 2020. The Digital Transformation of Education: Connecting Schools, Empowering Learners. https://www.broadbandcommission.org/Documents/working-groups/SchoolConnectivity _report.pdf.

[5] T. Burt. 2019. "CyberPeace Institute fills a critical need for cyberattack victims". Microsoft Blog. https://blogs.microsoft.com/on-the-issues/ 2019/09/26/cyberpeace-institute-fills-a-critical-need-for-cyberattack-victims.

[6] R. Chen, and M. Minges, 2021. "Minimum Data Consumption: How Much is Needed to Support Online Activities, and Is It Affordable?". World Bank Analytical Insights. https:/www.worldbank .org/en/topic/digitaldevelopment/brief/minimum-data-consumption-how-much-is-needed-to-support-online-activities-and-is-it-affordable.

[7] Global e-Sustainability Initiative (GeSI). 2015. #SMARTer2030: ICT Solutions for 21st Century Challenges. http://smarter2030.gesi.org/downloads/Full_report.pdf.

[8] P. Gilbert, 2020. "Safaricom, Google partner on Kenyan smartphone financing". Connecting Africa. 28 July. http://www.connectingafrica.com/author.asp?section_id = 761 & doc_id = 762720 & print = yes.

[9] Enabling Rural Coverage. https://www.gsma.com/m obileforde velopment wp-content/uploads/2018/02/Enabling_Rural_Coverage_English _February_2018.pdf.

[10] Survey of Universal Funds: Key Findings. https://www .gsma.com/publicpolicy/wp-content/ uploads/2016/ 09/ GSMA2013 _Report _SurveyOfU niversalSe rviceFunds_KeyFindings.pdf.

[11] T. Hassan, M. M. Alam, A. Wahab, et al. 2020. Prevalence and associated factors of Internet addiction among young adults in Bangladesh. *J. Egypt. Public. Health*. Assoc. 95, 3. https:/doi.org/10.1186/s42506-019-0032-7.

[12] IEA. 2021. "Data Centres and Data Transmission Networks". IEA, Paris https://www.iea.org/reports data-centres-and-data-transmission-networks.

[13] ITU. 2021a. Connectivity in the Least Developed Countries. https://www.un.org/ohrlls/sites/www.un.org.ohrlls/files/21-00606_1e_ldc-digital_connectivit -rpt_e.pdf.

[14] C. Lewis, 2017. Universal Access and Service in South Africa: Policy Success, Policy Failure and Policy Impact (SSRN Scholarly Paper ID 2980803). Social Science Research Network. doi.org/10.2139/ssrn.2980803.

[15] K. Müller, H. Glaesmer, E. Brähler, K. Wölfling, and M. Beutel, 2013. "Prevalence of Internet addiction in the general population: Results from a German population-based survey". Behaviour and Information Technology. No. 33, pp. 757-766. 10.1080/0144929X.2013.810778.

[16] Optus. 2020. Sustainability Report 2020. Australia. https://www.optus.com.au/content/dam/optus/documents/about-us/sustainability/reporting/2020/Optus-Sustainability-Report-FY2020.pdf.

[17] L. Silver, A. Smith. 2019. "In Some Countries, Many Use the Internet without Realizing It". Pew Research Center Fact Tank. https://www.pewresearch.org/fact-tank/2019/05/02/in-some-countries-many-use-the-Internet-without-realizing-it.

[18] "Soaring 'SuperTowers' Aim to Bring Mobile Broadband to Rural Areas*". IEEE Spectrum: Technology, Engineering, and Science News*. https:// spectrum.ieee.org/tech-talk/telecom/ wireless/soaring-supertowers-to-bring-mobile-broadband-to-rural-areas. 2018.

[19] D. Strusani, and G.V. Houngbonon, 2020. "Accelerating Digital Connectivity Through Infrastructure Sharing". *EMCompass*. https://doi.org/10.1596/33616.

[20] M. Agudelo, R. Katz, E. Flores-Roux, M.C. Duarte Botero, F. Callorda and T. Berry. 2014. Expansión de infraestructura regional para la interconexión de tráfico de internet en América Latina. Available at http://scioteca.caf.com/handle/123456789/522.

[21] Deutscher Commercial Internet Exchange (DE-CIX). 2021. "White Paper: 10 Reasons Why You Should Peer". Available at www.de-cix .net en/ resources/white-papers/10-reasons -why-you-should-peer.

[22] C. Dietzel. 2020. "Why the Internet Holds Firm: Internet Infrastructure in Times of Covid-19". DE-CIX. June. Available at www.de-cix.net/en/ resources/articles/why-the-internet -holds -firm-internet-infrastructure-in-times-of-covid-19.

[23] DP Facilities (n.d.). "The Critical Role Data Centers Play in Today's Enterprise Networks: Part 3 – Why Cloud On-Ramps Are Key for an Enterprise Migrating to the Cloud". Four-part blog series. Available at www.dpfacilities.com/blog/cloud -onramps-are-key-to-migration.

[24] Feldmann et al. 2020. "The Lockdown Effect: Implications of the COVID-19 Pandemic on Internet Traffic*". Proceedings of the ACM Internet Measurement Conference*. 1-18 October. doi.org/10.1145/3419394.3423658.

[25] T. Katsuyasu. 2020. "The impacts of COVID-19 pandemic on the IXPs in APAC region". IX.br IX Forum. 2-4 December. Available at https://forum.ix.br/files/apresentacao/arquivo/1025/20201204-ixbr-forum-katsuyasu-for-publish.pdf.

[26] K. Olsen. 2021. "Peering in Norway: Traffic growth and shifting patterns". Netnod. 20 September. www.netnod.se/blog/peering-norway-traffic-growth-and-shifting-patterns.

[27] Oxford Business Group. 2021. "Data Centres in Africa Focus Report: How is Africa positioned as a destination for data centres?" Available at https://oxfordbusinessgroup.com/news/focus-report-how-africa-positioned-destination-data-centres.

[28] Sandvine. 2020. "Mobile Internet Phenomena Report 2020". Available at www.sandvine.com/download-report-mobile -internet-phenomena-report-2020-sandvine.

[29] S. Srinivasan, S. Comini and M. Minges. 2021. "The Importance of National Data Infrastructure for Low and Middle-Income Countries". SSRN Scholarly Paper ID 3898094. Rochester, NY: Social Science Research Network. Available at https:// doi.org/10.2139/ssrn.3898094.

[30] The Economist. 2021. "Data Centres Are Taking Root in Africa". 4 December. Available at www.economist.com/middle east-and-africa/2021/12/04/data-centres-are-taking-root-in-africa.

ERICSSON MOBILITY REPORT (REVIEW. PART II)

Peter Jonsson
Ericsson, Stockholm, Sweden
www.ericsson.com

DOI: 10.36724/2664-066X-2022-8-4-33-41

ABSTRACT

While 5G roll-outs are by no means complete, they are well under way. For many in the industry, efforts to utilize 5G to go beyond simply providing "fast connectivity" are already in focus. Our articles explore how the industry is already looking forward to what comes next, asking how to: make gains in sustainability; utilize technologies like IoT and edge to maximize efficiencies and push exciting use cases; use 5G as a springboard for innovation; truly capture the opportunities for all consumers and enterprises in all regions; and, with all this growing potential, how to keep 5G safe and secure. 5G technologies play a key role in modernization, providing multiples of capacity while becoming more energy efficient. Innovative network technologies enable service providers to introduce new services that in turn support societies and enterprises to reduce their carbon emission footprint. In this edition, we share some examples of how 4G and 5G technologies make it possible to unleash the power of IoT connectivity to enhance both enterprises' business performance and sustainability. The transition to cellular LPWA and 4G/5G technologies makes it possible to unleash the power of IoT connectivity. We explore the positive impact of these technologies in areas such as business efficiency and sustainability. Deploying edge computing is key to enabling latency-critical and bandwidth-hungry 5G use cases, and can cost less than on-premise IT resource for an enterprise. This capability represents huge untapped growth potential for service providers. As 5G grows in prominence due to advancing digitalization, networks become a more enticing target for threat actors. We explore the motivators, opportunities and capabilities of threat actors, and how to protect 5G networks.

KEYWORDS: *5G technologies, digitalization, mobile subscriptions.*

This article was written in cooperation with Telia Company, a market-leading service provider in Sweden, providing innovative services for more digitalized and sustainable societies across the Nordics and the Baltics and with MTN, Africa's largest mobile network operator.

Information about authors:
Executive Editor of Ericsson Mobility Report: Peter Jonsson
Project Manager: Anette Lundvall
Collaborators: Katja Kalliorinne (Telia), Staffan Thorsell (Telia), Amith Maharaj (MTN Group), Emmanuel Lartey (MTN Group), Farhan Khan (MTN Group)
Contributors: Harald Baur, Greger Blennerud, Fredrik Burstedt, Warren Chaisatien, Mikko Karikyto, Anna-Maria Kastedt, Per Lindberg, Michael Martinsson, Rhys Hemi Mataira, Leena Mattila, Amardeep Mehta, Frank Muller, Ravi Shekhar Pandey, Lars Sandstrom

INTRODUCTION

5G technologies play a key role in modernization, providing multiples of capacity while becoming more energy efficient. Innovative network technologies enable service providers to introduce new services that in turn support societies and enterprises to reduce their carbon emission footprint. In this edition, we share some examples of how 4G and 5G technologies make it possible to unleash the power of IoT connectivity to enhance both enterprises' business performance and sustainability.

The transition to cellular LPWA and 4G/5G technologies makes it possible to unleash the power of IoT connectivity. With Telia, we explore the positive impact of these technologies in areas such as business efficiency and sustainability.

MTN considers 5G to be an innovation platform that could completely transform society and businesses. Here's how new ways of working will allow service providers to fully capture the 5G opportunity in Sub-Saharan Africa.

Deploying edge computing is key to enabling latency-critical and bandwidth-hungry 5G use cases, and can cost less than on-premise IT resource for an enterprise. This capability represents huge untapped growth potential for service providers.

UNLEASHING THE POWER OF IOT CONNECTIVITY

Telia Company's purpose is to reinvent better, connected living, and it strives to improve business efficiency. The transition to cellular LPWA and 4G/5G technologies makes it possible to unleash the power of IoT connectivity to enhance enterprises' business performance and sustainability.

At the crossroads of change

In recent years, Telia has seen a continuous rise in the number of cellular-connected IoT devices on its networks across the Nordic and Baltic countries. 2021 saw an increase of 44 %, more than double the growth compared to 2020.

The growth is primarily fueled by large-scale smart meter deployments, based on the low-power wide-area (LPWA) IoT technologies, NB-IoT and Cat-M.

In addition, the adoption of embedded universal integrated circuit cards (eUICC)1 has simplified the global deployment of connected devices by allowing remote SIM provisioning of multiple network profiles.

NB-IoT and Cat-M technologies are ideal for connecting massive volumes of low-cost, low-complexity IoT devices with long battery life and limited data throughput demand.

These technologies, which form part of the 5G standard, are the successors to 2G and 3G networks that are being replaced as the industry moves to adopt broadband and critical IoT, powered by 4G and 5G.

IoT devices migrating to modernized networks

2G and 3G networks are being phased out globally to enable the reuse of valuable radio spectrum for 4G and 5G deployments. By modernizing the networks with the latest technology and replacing old equipment, it is possible to realize new business opportunities and create significant energy savings at the same time.

About 30 % of all cellular IoT devices are still connecting through 2G/3G networks. However, enterprises are migrating their IoT devices and services to Cat-M and NB-IoT networks, which are more energy efficient, reliable and have higher capacities.

Across Europe, the sunsetting of 3G networks is happening before 2G, but the order and the schedule varies from country to country and between service providers. Telia will decommission its 3G networks before 2G, with the 3G sunset already in motion across Telia's markets in the Nordics and Baltics.

Globally, the number of IoT devices connected via 2G and 3G has been in slow decline since 2019 (Figure 1). The combined segment of cellular LPWA, broadband and critical IoT (4G/5G) overtook 2G/3G in terms of IoT connection numbers for the first time in 2020. LPWA IoT technologies are expected to make up about 50 % of all cellular IoT connections in 2027.

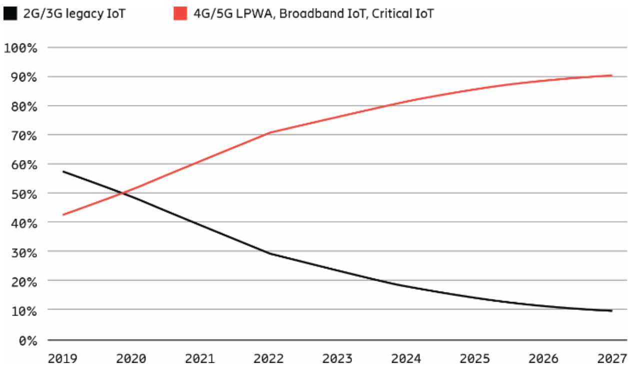


Figure 1. % age share of 2G/3G vs 4G/5G connections for cellular IoT

Source: Ericsson Mobility Visualizer.
Note: NB-IoT and Cat-M access technologies are also referred to as LPWA technologies

Extending IoT connectivity reach with cellular LPWA

LPWA IoT technology supports solutions requiring low total cost, long battery life and the ability to operate in remote locations. Its energy efficiency comes from sending smaller amounts of data at defined time intervals and then quickly powering down the transmitter in between. The two different cellular LPWA IoT network technologies, NB-IoT and Cat-M – both under the 5G standard – are inherently more secure and have longer reach than previous generations. For example, Cat-M can have a reach of up to 100km and NB-IoT up to 120km from a radio base station.

The extended reach and high-penetration capabilities make it possible to cost-efficiently connect sensors in cities, remote rural, coastal, and maritime areas, and even deep inside buildings or underground. In several tests throughout its development, Telia has shown that NB-IoT can connect devices placed as deep as 80m underground.

The transformational power of enterprise digitalization Organizations that embrace this new era of digitalization enjoy increased efficiency and cost reductions, thanks to better predictability and greater control.

Digitalization also means companies are becoming software businesses, generating proprietary data. They are no longer an isolated part in a vertical market, but a data-driven, interconnected element of a wider, digital ecosystem of services.

For example, when an agriculture machinery manufacturer equips a tractor with more than 300 IoT sensors and the ability to process more than 150,000 measurements per second, the business and its value creation changes. The tractor is now a data-generating unit, part of an ecosystem of related services such as weather forecasting, commodity pricing and crop yield predictions.

There are many more examples: A car manufacturer that harnesses IoT connectivity is no longer just selling cars, they are also enabling carpooling services and shared ownership alternatives, while gathering and handling information about the driver, roads, traveling habits and even the weather. Providers of consumer IoT services improve the health and lifestyle of consumers thanks to health monitoring, lifestyle optimization and entertainment apps.

Enterprises can be transformed and their new capabilities turned into new customer values and chargeable services. Internal processes and cost control become more effective too, as every decision can be based on real-time data.

Monitoring enables less repeating and reactive maintenance, and there is no longer a need for so many trips or manual efforts, leading to clear sustainability gains such as reduced CO2 and pollution from fossil-fueled vehicles. Smarter energy systems, smarter grids and better monitoring allow for a more efficient use of resources.

A truly data-driven, or rather data-native, company makes data the basis for all decision making, regardless of whether it relates to technology, business decisions or sustainability.

IoT connectivity goes underground for pest control

As cities expand and urbanization grows, there is typically an increase in common underground pests, such as rats.

Poison traps have traditionally been used for pest control. However, this method allows poison to enter the food chain whereby birds, foxes and domestic pets eat the poisoned rats above ground.

A pest control company in Denmark developed a new digital trap that enables an ethical, non-toxic approach. At first, the solution utilized 2G (GSM/GPRS) for connectivity, but due to the heavy steel covers below the surface of the drains, 25% of the traps could never connect to the network. By migrating from legacy connectivity to NB-IoT technology, the connection success rate rose from 75% to 100%. NB-IoT technology fulfilled the performance requirements to connect the traps deep underground, enabling performance monitoring and information gathering about the number of triggered traps, maintenance needs and sewer flooding in hard-to-reach places. This gave the company a competitive advantage.

Navigating treacherous waters with IoT Hundreds of thousands of islands making up the archipelagos of Finland and Sweden are battered by brutal storms every winter. Navigation marks are extensively deployed to support marine traffic safety, but these often break free from their anchors and float across long distances. In the past, local maritime authorities had to go out on resource-demanding and fuel-consuming runs each spring to find the marks and return them to their correct locations.

A Finland-based global provider of advanced tracking and sensor solutions took on the challenge, developing a tracker using NB-IoT and aiming to deploy them in over 20,000 navigation marks in the Finnish archipelago. NB-IoT is the ideal connectivity solution for the hard-to-reach offshore navigation marks thanks to its extensive reach and the ability to operate for up to ten years on a single battery charge. Sea routes will be digitalized by remotely tracking navigation marks, which create savings in cost and resources, reduce CO2 emissions and make the waters safer.

Remove roaming limits

Global multinational enterprises (MNEs) need to connect IoT devices across different countries and regions. For an MNE to procure local solutions from a local service provider in each market would be very challenging to implement and to operate, both technically and commercially.

Using cellular network capabilities, they can change the connectivity profile of devices through eUICC. The SIM profile is changeable over-the-air (OTA) and can be set to become a local network device to fulfill the legal requirements that exist in each market, or to have a roaming profile when allowed.

A Finland-based manufacturer of industrial and marine gearboxes, as well as drives for process industries, needed to set up easy-to-use and cost-effective mobile connections for some of its 200,000 gearboxes across 40 countries. In many critical segments of the process industry, optimized gearboxes that allow for uninterrupted operations and cost-effective maintenance are vital. Unplanned maintenance leads to production loss.

Installed sensors measure data such as oil quality, relative humidity, temperature, gearbox vibrations, pressure, and cleanliness of the equipment. Pre-installed IoT devices transmitted the relevant information to different stakeholders, such as the process control system, the operations and maintenance personnel and the equipment manufacturer. Through eUICC SIM cards, health monitoring the equipment and anticipating subscription costs became transparent and easier to manage.

Transforming tomorrow with IoT

4G and 5G networks will continue to evolve, further enhancing IoT connectivity capabilities with higher data speeds, lower latency, improved security, and extreme reliability. Supported by 4G networks, businesses can achieve better efficiencies and performance with cellular IoT technologies, and Telia's 5G network presently supports use cases such as remotely controlled high-lift wheel loaders, autonomous field robots for mechanical weed control and automated port operations.

Service providers are uniquely positioned to support the digital transformation of a wide range of industries with evolving cellular IoT technologies, as they enable industries to become truly data driven, efficient and sustainable to further contribute to a better society. As 5G and IoT transform connectivity and unlock new intelligence, the possibilities are only limited to what enterprises and service providers can imagine.

THE EVOLUTION OF MTN'S CONNECTIVITY PLATFORM

Continued investment in 4G – and the expansion of 5G – technologies are expected to play a crucial role in realizing MTN's ambitions, and will enable it to meet evolving market demands and monetize new use cases across markets in the Sub-Saharan Africa region.

MTN Group, South Africa, has defined its strategic "Ambition 2025" plan. It is built on MTN's current market position, where connectivity is the foundation, while platforms are gradually expanded to capture new growth opportunities and deliver value. In this context, 5G network deployment and evolution across markets plays an important role in enabling new services for consumers, enterprises, industries and society. For MTN, 5G is an innovation platform with the ability to transform various aspects of business and livelihoods beyond pure connectivity.

Data connectivity and usage – drivers for revenue growth

In the Sub-Saharan Africa region, connectivity is still dominated by 3G and 2G technologies, with 4G only making up around 20 % of mobile subscriptions by the end of 2021.1 However, demand for data connectivity and digital services is increasing across markets. Operating in 18 markets across the Middle East and Africa, MTN is pursuing these new growth opportunities.

Continuous network modernization and coverage build-out, supported by MTN's Rapid Rural Rollout (R3) program, has enabled it to capture strong new subscriber growth and stimulate increased data usage. This has resulted in increased data service revenues, despite price pressure in the markets. In South Africa, MTN networks experienced strong data growth as the number of customers actively using the internet grew by 12.5%, leading to a mobile data traffic growth of almost 60% in 2021. The average mobile data traffic per pre-paid subscriber was 2.3GB and 10.3GB for post-paid subscribers.

MTN considers data as a main driver of revenue growth over the medium term. Initiatives to stimulate further data adoption include data service bundling, segmented value propositions and the development and launch of freemium data propositions, supported by strategic over-the-top partnerships.

MTN's strategic priorities up to 2025

MTN continues to invest in 4G technologies and has expansive plans for 5G to realize the opportunities it has identified to evolve and expand its service offerings for the consumer, enterprise and industry segments. MTN's strategic priorities are articulated in its Ambition 2025 strategic framework, which is underpinned by 10 key technology strategic pillars intended to enable growth in connectivity and platforms businesses.

Some of the most important pillars are ensuring best-in-class, ubiquitous access across mobile and fixed networks, maintaining network leadership and efficiencies, and the monetization of infrastructure. Other priorities include investment in sustainable technologies and zero-touch, service-aware networks. 5G networks will play an essential role in delivering on the technology pillars to realize the Ambition 2025 plan.

Monetization of network infrastructure includes a network-as-a-service (NaaS) strategy, where network sharing (national roaming, MOCN and MORAN) is the starting point, followed by 5G network slicing which enables exposure of network functionality via APIs to build new enterprise services. An additional step will be the monetization of data exposed via online third parties.

Building 5G for timely monetization

MTN's 5G network build-out strategy is based on meeting evolving market demands with the timely deployment of the relevant technology enablers, in order to optimize the potential for monetizing new use cases. So far, 5G subscriber uptake has been driven by a combination of increased 5G device penetration and fixed wireless access (FWA) subscriber uptake. The average 5G subscriber mobile data consumption is approximately twice that of 4G subscribers.

Mobile broadband and FWA are currently the main 5G services marketed by MTN. It stresses better user experience as the main value, in a manner that relates to consumer needs, rather than bandwidth and latency which are not relative to the consumer. Interest in high-speed, good-quality broadband increased as working from home practices spread during the pandemic. 5G FWA will compete with fiber-to-the-home as an alternative, cost-efficient home broadband solution.

The deployment of 5G SA architecture, enabling network slicing, will be driven by consumer and enterprise use case evolution over time. In the 2023-2024 timeframe, the initial target will be consumers (enhanced mobile broadband/FWA).

This will be followed by deployments for enterprises, as ultra-reliable low-latency communications (URLLC) for critical services – which are crucial for high-end industrial applications – and 5G-era massive machine-type communications (IoT) use cases start to emerge. Over-the-top services will also be an important offering to create stickiness.

The challenge of migrating to SA architecture is not related to the technology as such, but rather how to monetize these new types of services, while also adhering to local market regulations related to net neutrality.

The enterprise opportunities

5G will enable a range of new services across different sectors, such as mining, manufacturing, utilities and agriculture. MTN is sharing information with enterprise customers and industry verticals about the value of 5G connectivity and low latency for optimizing its operations, as well as the introduction of new services. Dedicated private networks are already being deployed in proof-of-concept trials to validate the value of new services.

An AI-based face recognition system at mining sites is one example of a service being evaluated – this is currently 4G based, but will evolve to 5G. According to MTN, the main new opportunities in the African market that can be addressed with 5G technologies are related to areas such as virtual education, industrial automation, telemedicine, remote health care and smart cities.

MTN's 5G deployment strategy

5G is still in its infancy in South Africa. Within the country, MTN is a leading service provider, with around 35 million mobile subscribers. Of these, about 50 % are active mobile data users. At around USD 6.30, it has the highest blended average revenue per user (ARPU)2 of all service providers in South Africa.

MTN launched its 5G commercial services in June 2020 and reached 200,000 5G subscribers by the end of 2021. Continued 5G subscriber uptake will be strongly impacted by the availability of a wider range of low-cost 5G smartphones. In the recent spectrum auction, MTN acquired 100MHz of spectrum across three frequency bands: 40MHz in the 3.5GHz band, 40MHz in the 2.6GHz band and 2x10MHz in the 800MHz band.

MTN's initial 5G network deployment strategy focuses on high-value urban areas and hot spots, where they will deploy high-quality 5G New Radio (NR) equipment on the mid-band 3.5GHz frequency (40MHz bandwidth) as a capacity layer. Initially, hot spots being targeted include key markets, university locations, institutions and residential areas serving consumers with high data usage potential.

Long term, a coverage layer on the 700MHz band will ensure that regulatory requirements for 5G coverage are fulfilled. Deployments in high-band spectrum (mmWave) will be carried out on a more limited basis in areas with high-capacity traffic demand and in areas for deployment of FWA services. 5G is also available in some areas through dynamic spectrum sharing (DSS), a technology which allows both 4G and 5G to be deployed in the same band and on the same radio. MTN has deployed about 1,000 5G mobile sites and aims to reach 25 % 5G population coverage by the end of 2022, and 60 % by 2025.

MTN will begin decommissioning its 3G network in 2025/2026, with 4G and 5G becoming the principal technologies used to deliver telecoms services to its customers. In Sub-Saharan Africa, 5G subscriptions will represent around 10 % of all mobile subscriptions by 2027,3 with South Africa expected to lead the adoption rate in the region. Local market research forecasts that the number of 5G subscribers in South Africa is expected to reach 11 million by 2025.

Strategy execution – addressing the new opportunities

MTN knows that the traditional business models and "ways of doing things" will not be sufficient to enable it to make the most of the emerging 5G opportunities. To really benefit from 5G's capabilities, MTN will need to tie its 5G vision and roadmap closely to its digital transformation strategies. It will need to introduce network slicing if it expects to see revolutionary business models and service pricing.

Network slices will be created on demand and will be independently controlled and managed with a degree of customization that could previously only be achieved with dedicated physical networks. Network slices allow partners to integrate into network platforms in a similar way to a dedicated private network, but with far less effort. They will also enable MTN to expand its role from connectivity to other areas of the value chain – such as cloud and edge services, orchestration and applications.

ENABLING DEMANDING USE CASES WITH CSP EDGE COMPUTING

Edge computing is key to enabling latency-critical and bandwidth-hungry 5G use cases, representing significant growth potential for communications service providers (CSPs).

Demand for immersive use cases has been held back by factors in the development of a new ecosystem, including networks, devices and applications. As this ecosystem matures, we expect the value brought about by edge computing will overcome the cost advantages held by large-scale data centers. Our analysis indicates that it is clearly possible for a CSP to build-out edge computing with an annual cost base not materially higher than a data center.

Historically, enterprises could either run their application workloads on-premise, based on the company's own IT infrastructure, or hosted in centralized data centers. There are several fundamental differences between these deployment options, including cost, control, security and regulatory compliance.

With the rollout of 5G, CSP mobile networks present an attractive proposition for running demanding enterprise applications close to target customers.

A cost analysis of deployment shows that the cost to CSPs to deliver edge compute resources to enterprise customers is nearly half of what it would cost for an enterprise to build its own on-premise infrastructure with similar performance, reliability and data security.

Enter edge

Edge refers to the distribution of compute resource and applications to geographically distributed sites on the premises of an enterprise or in a CSP network. It provides compute resource and storage closer to where the data is generated and consumed. It offers significant advantages by enabling advanced data processing capabilities located close to where they are needed, reducing the latency inherent in centralized data centers. Deploying software at the edge comes with an increased cost compared with centralized deployment, but also enables a range of enhanced capabilities, including increased performance, reliability, data security and privacy, as well as reduced cost/bandwidth for the transport network.

Since data does not need to travel to remote locations for processing, analysis and rendering, enterprises can save precious milliseconds on round-trip times (RTT) while benefiting from more reliable data throughput. Enterprise on-premise edge computing can help insulate their networks from cyberattacks and distributed denial-of-service (DDOS) attacks on more centralized locations.

There is also reduced risk of data being intercepted in transit, further adding to the security and privacy features of edge computing.

Edge computing can help organizations to fully comply with jurisdictional data regulations and sovereignty laws by allowing data to be processed close to its source.

CSPs can leverage the proximity of their existing sites to end users to set up edge compute, providing low-latency and high-performance IT capabilities for enterprise workloads as a service. For example, one way enterprises can reduce on-premise IT infrastructure is by deploying “infrastructure-less” branch offices; all IT on-premise applications, from communication, image processing and analytics to specialized enterprise services, can be hosted on the network edge.

A number of considerations must be addressed while rolling out compute capabilities alongside connectivity. There can be limitations to adding resources to some sites due to constraints on space, power and/or network capacity. Another challenge could arise from low fault tolerance of the commercial off-the-shelf (COTS) hardware used at the edge sites. CSPs may also require new sites to provide both continuous coverage and compute capabilities at critical locations to enable particularly demanding use cases.

The cost of the edge

To compare the cost of deploying compute resources at different scales, we convert capital expenditure into depreciation by dividing each asset category by the number of years it will be written down, and then add the resulting depreciation to the annual opex, providing a snapshot of the yearly cost structure. For example, power and cooling systems are written down over 14 years, whereas COTS servers are typically written down over 3 years.

Capex includes:

- Server capex is mainly the cost of COTS servers and virtualization software.
- Other capex consists of the cost of components such as power distribution and cooling systems.

Opex includes:

- The electrical power required to run and cool the servers.
- Other opex, mainly the cost of operations and maintenance (O&M).

As an example, we estimate the cost of compute resources for a CSP in Sweden. Initially, edge compute rollout is expected to be on aggregation sites having power capacity installed up to 10kW, hosting an average of 8 server units, each with 4 cores. With approximately 8,000 access sites and 1 aggregation site per 10 access sites, there is a virtual processor (vCPUs) capacity of 25,600 (800 sites x 8 servers per site x 4 cores per server) for enterprise applications at CSP-owned edge sites.

Capex depends on the required capacity plus redundancy in the edge hardware components to meet the reliability requirements for edge services or applications. The geographic distribution can also be leveraged to improve the system availability by avoiding a single point of failure. We categorize the capex into server capex and other capex due to the faster cycle of server performance improvement compared to others. Servers are typically depreciated over 3 years while investments in power and cooling systems are depreciated over 14 years.

Upgrading aggregation sites with edge compute capability, with an average of 8 units of servers, can draw up to 1.6MW (800 sites x 8 servers per site x 250W per server) for running the servers. With an assumed power efficiency factor of 2, 3.2MW power is needed on average to power all the aggregation sites. The cost of compute resource at each aggregation site is estimated be around USD 20,000. Hence the USD per critical watt for an edge site is USD 20,000/(8 servers x 250W/server) = USD 10/W. This cost is very similar to USD per critical watt for building a large-scale data center.

Opex is the sum of electricity cost and O&M. For the current study, we assume it to vary in the range of USD 0.10-0.15/kWh. For O&M, the cost of full-time employees required to manage and maintain the distributed edge servers is projected.

We constructed four different scenarios to estimate and compare the compute resource cost, based on USD per vCPU-hour.

- Scenario 1 is a base case with costs assumed for a small- or medium-sized enterprise handling its compute needs with its own IT infrastructure.
- Scenario 2 is an estimation of cost for a large-scale data center to provision the same capacity as the first case.

- Scenario 3 is built around provisioning the capacity used in the first two cases by deploying edge computing on the CSP network.
- Scenario 4 is an extension of the third case, with the addition of the cost to implement a set of measures to reduce power consumption. These include using renewable energy, dynamic usage of battery/power storage at peak times and advanced cooling technologies, including a heat exchanger for the server cabinets.

Server capex is the most significant parameter for all the scenarios except the base case where O&M (other opex) dominates due to the lack of scale.

Electricity cost is the second largest factor in USD/CPU-hour for scenario 3. This leads to the significance of additional power efficiency elements in scenario 4.

With an estimate of expenditure in use cases suitable for edge deployment, the cost of edge compute resources can be just 10 % more than that of a large-scale centralized one. Capacity utilization is the most important parameter for increasing the cost efficiency of the edge resources (Fig. 2).

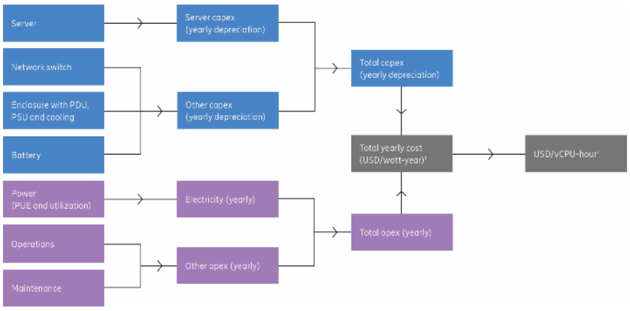

Server
Server capex (yearly depreciation)
Network switch
Enclosure with PDU, PSU and cooling
Battery
Other capex (yearly depreciation)
Total capex (yearly depreciation)
Total yearly cost (USD/watt-year)^{1}
USD/vCPU-hour^{2}
Power (PUE and utilization)
Electricity (yearly)
Operations
Maintenance
Other opex (yearly)
Total opex (yearly)

Figure 2. Annual cost estimation framework for compute resources

CSP edge infrastructure resources are marginally more expensive than those at a large-scale data center but much less than those at an enterprise on-premise compute solution. CSP edge infrastructure also provides better latency and proximity to enterprise applications.

When comparing the costs of a large-scale traditional data center and a CSP network edge, we need to consider that those alternatives enable different use cases. Positive features for a CSP-operated edge include high location sensitivity, reduced latency (in the millisecond range), and guaranteed connectivity. However, edge compute infrastructure will have limited scalability compared to large data centers.

The short- to mid-term edge opportunity for CSPs should be seen in the wider context of the enterprise opportunity, where edge computing will be an enabler for a broad range of use cases, for example offerings such as private 5G networks, IoT platforms, cloud gaming and immersive experiences with XR. In the long term, when compute is deeply integrated in mobile networks, the most demanding use cases, including closed-loop industrial control systems, industrial robotics, extended reality with real-time synchronous haptic feedback (the Internet of Senses) and negotiated automatic cooperative driving for autonomous vehicles, will open up an expanding set of opportunities.

SECURING 5G NETWORKS IN AN EVOLVING THREAT LANDSCAPE

5G is, by design, more secure than previous generations, but it is being deployed and operated in an evolving and complex threat landscape. New, demanding use cases served by telecom networks can increase attack motivations and attack vectors are multiplying. These factors are exponentially increasing the need to protect networks.

The evolving 5G threat landscape

With the introduction of 5G and billions of new devices, the threat landscape in which telecom networks operate is evolving significantly. Networks provide vital infrastructure for business-, mission- and society-critical applications, and as a result, threat actors are motivated to constantly evolve to seek out weaknesses.

Safeguarding 5G networks

As the value and volume of personal, business sensitive and public service information increases with continued digitization, security and privacy laws and regulations have been expanding. This is a reaction to decreasing risk tolerance and the deteriorating cyber security environment.

Regulators know the importance of 5G and see safeguarding these networks as vital. The threat landscape for 5G is more complex than with previous generations due to the convergence with traditional IT, enabling IT threat actors to attack telecom networks in a similar way. In addition, networks often have new functionalities, such as network slicing for service separation and isolation, along with an increased use of AI/ML for automation. While AI is widely explored for its potential in addressing security concerns in networks, it is also important to consider the security and transparency of AI. Edge computing places cloud resources closer to the access, bringing new challenges whilst enabling mission-critical, low-latency applications.

Attacks on telecom networks are rising

Threat actors are increasingly skilled and pervasive, and attacks are becoming more frequent. Research from CrowdStrike, a US cyber security company, shows which industry verticals are most frequently impacted by targeted intrusions.1 The data showed that, between July 2020 and June 2021, the telecom industry was the most targeted, attracting 40 % of attacks compared to 10 % for the next-highest industry vertical. It should be noted that the data does not distinguish between the telecom enterprise and the telecom network intrusions for the industry.

What motivates threat actors?

The main motivations to target telecom networks are surveillance/espionage, financial gain and disruption/sabotage. In recent years, the most common type of attack in the cybersecurity landscape has been the deployment of financial gain ransomware.

To achieve bigger payoffs, ransomware operators have shifted their targeting to high-profile organizations in industries such as manufacturing. Threat actors know this industry sector has a low tolerance towards downtime and is more inclined to pay out as a result.

With increased use of 5G within different industry verticals' networks, the motivation to attack 5G networks should be looked at from the perspective of the related industry sector.

Personal data is also always of high interest. One objective of espionage is to obtain call metadata, especially call detail records (CDRs). This means customer billing and customer care systems are primary targets. LightBasin was observed targeting business support systems to obtain CDRs.

Disruption is the least typical of these motivations for targeting telecom networks. These attacks often have their roots in ideology, driven by personal, group or nation-state agendas. During the first quarter of 2022, a number of these attacks occurred on European networks, including targeted attacks to prevent local gamers from participating in a tournament and network-wide disruptive cyberattacks, putting critical services at risk.

Due to a shift in the tactics used by cybercrime and nation-state threat actors, and the increasing use of common IT platforms in telecoms, the likelihood of attacks has increased.

The opportunities for threat actors

New features within 5G networks bring many advantages, enabling new use cases. However, the technical complexities can create new opportunities for threat actors.

The ongoing transformation to cloud native introduces new concepts, new deployment methods and more complex partnership structures. With this trend, deployments are becoming more complex. This requires new types of competence and skill sets, from both vendors and service providers. Consequently, the risk for misconfigurations, which expose weaknesses, is increased.

Vulnerabilities in virtualization, cloud services, or network slicing can have a considerable impact, as they may enable access to unauthorized resources (Fig. 3).

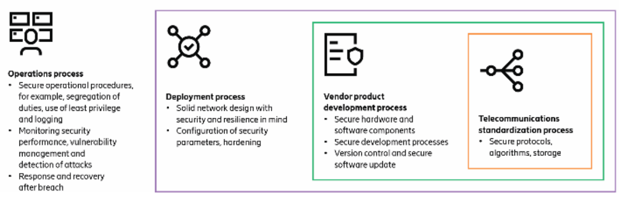

Operations process
• Secure operational procedures, for example, segregation of duties, use of least privilege and logging
• Monitoring security performance, vulnerability management and detection of attacks
• Response and recovery after breach
Deployment process
• Solid network design with security and resilience in mind
• Configuration of security parameters, hardening
Vendor product development process
• Secure hardware and software components
• Secure development processes
• Version control and secure software update
Telecommunications standardization process
• Secure protocols, algorithms, storage

Figure 3. Protecting 5G end-users requires a holistic approach including the four key layers

5G will connect billions of devices, and not all these devices have sufficient security protection. Devices used for Industrial IoT are often optimized for a specific task, with design driven by cost efficiency.

Vulnerabilities in these devices can be used to target the 5G network, or the industry vertical. This requires protection of devices to be provided from the network side. In general, any exposed interface provides an initial entry point for a threat actor. LightBasin accessed target networks via incorrectly exposed interfaces on the GPRS roaming exchange (GRX), a closed inter-service provider network.

Threat actors are increasingly using valid credentials for accessing targets. In addition to the traditional social engineering techniques for obtaining human identities, threat actors are looking for weaknesses presented by the surge of machine identities that are needed in cloud-native deployments. Strong multi-factor authentication, with management and monitoring of privileged accounts, is essential to prevent and detect account misuse. It will also limit the impact of credential theft and the exploitation of vulnerabilities.

What are the capabilities of threat actors?

Threat actors have shown the capability to build targeted and context-specific malware. Nation state threat actors routinely exhibit good operational security and use various defense evasion techniques to hide their activities, making it possible for them to move laterally in the target organization before being noticed. For instance, LightBasin carefully deleted traces in log files after their activities.

Threat actors try to blend their communication into normal traffic and use legitimate protocols, such as ICMP and HTTP. In addition to these, LightBasin used telecom-specific protocols to bypass firewalls and stay under the radar.

As the industry moves away from proprietary protocols and dedicated infrastructure, intrusion of telecom networks does not necessarily depend on extensive knowledge of these networks and their protocols. Threat actors targeting telecommunications networks will increasingly resort to routine vulnerability exploitation, supported by public availability of exploit code.

Even though 5G interconnects are more secure, older network generations will be used for several years, and attacks via interconnected interfaces will continue and will be more complex and difficult to detect as threat actors increasingly focus on defense evasion.

Trust in mobile networks is paramount

Trust in mobile networks, especially 5G, is the foundation for digitalization. To enhance trust, the GSMA Network Equipment Security Assurance Scheme (NESAS), jointly defined by 3GPP and GSMA, provides an industry-wide security assurance framework to facilitate improvements in security levels.

NESAS defines security requirements and an assessment framework for secure product development and product lifecycle processes, and uses 3GPP-defined security test cases for the security evaluation of network equipment. NESAS is intended to be used alongside other mechanisms to ensure a network is secure and, in particular, to ensure an appropriate set of security policies covering the entire lifecycle of a network is in place.

3GPP standardization made major improvements in terms of security and privacy compared to 4G. 5G has been designed with new functionality that is intended to make it more resilient towards various existing frauds, subscriber privacy and eavesdropping issues, than earlier generations.

For instance, the industry is putting considerable effort into protecting the interconnect networks between the service providers, encrypting, and otherwise hiding subscriber identifiers, and preventing the modification of the user data sent between user equipment and radio base stations.

5G also provides a standardized and well-defined way to deploy zero-trust functions like authentication and authorization of API usage, and protected communication between and to the 5G network functions.

It's time for the active defense of telecom networks

With networks being used in new contexts, connecting a greater variety of mission-critical processes, it is no longer enough to rely solely on standardized and regulatory-based security controls. Now the active defense of telecom networks is also required.

The entire industry is currently accelerating the journey from passive defense to active defense strategies. The embedded security inside network products is critical but still not enough. The telecom networks of today are built to evolve, and security must do the same.

Securing 5G networks

Telecom networks' availability and performance are more valuable than ever, which makes them attractive targets for malicious actors. Powerful security monitoring and automation, identity management, effective incident response handling and solid business continuity planning are critical to securing networks. Building a secure 5G network requires a holistic approach, rather than a focus on individual technical parts in isolation, to protect end users. Network operations is one of four key layers enabling the holistic approach, alongside standards, product development processes and network deployments.

Methodology

Forecast methodology. The forecast time in the Mobility Report is six years. The subscription and traffic forecast baseline is established using historical data from various sources, validated with Ericsson internal data, including measurements in customer networks. Future developments are estimated based on macroeconomic trends, user trends, market maturity and technological advances. Other sources include industry analyst reports, together with internal assumptions and analyses. Historical data may be revised if the underlying data changes – for example, if service providers report updated subscription figures.

Mobile subscriptions. Mobile subscriptions include all mobile technologies. Subscriptions are defined by the most advanced technology that the mobile phone and network are capable of. LTE (4G) subscriptions, in most cases, also include the possibility for the subscription to access 3G (WCDMA/HSPA) and 2G (GSM or CDMA in some markets) networks. A 5G subscription is counted as such when associated with a device that supports New Radio as specified in 3GPP Release 15, and connected to a 5G-enabled network. Mobile broadband includes radio access technologies HSPA (3G), LTE (4G), 5G, CDMA2000 EV-DO, TD-SCDMA and Mobile WiMAX. WCDMA without HSPA and GPRS/EDGE are not in-cluded. FWA is defined as a connection that provides broadband access through mobile network enabled customer premises equipment (CPE).

Subscribers. There is a large difference between the numbers of subscriptions and subscribers. This is because many subscribers have several subscriptions. Reasons for this could include users lowering traffic costs by using optimized subscriptions for different types of calls, maximizing coverage and having different subscriptions for mobile PCs/tablets and mobile phones. In addition, it takes time before inactive subscriptions are removed from service provider databases. Consequently, subscription penetration can be above 100%, which is the case in many countries today. However, in some developing regions, it is common for several people to share one subscription, for example via a family- or community-shared phone.

Mobile network traffic. Ericsson regularly performs traffic measurements in over 100 live networks covering all major regions of the world. These measurements form a representative base for calculating worldwide total mobile network traffic. Mobile network data traffic also includes traffic generated by FWA services. More detailed measurements are made in a select number of commercial networks with the purpose of understanding how mobile data traffic evolves. No subscriber data is included in these measurements.

Population coverage. Population coverage is estimated using a database of regional population and territory distribution, based on population density. This is then combined with proprietary data on the installed base of radio base stations (RBS), together with estimated coverage per RBS for each of six population density categories (from metro to wilderness). Based on this, the portion of each area that is covered by a certain technology can be estimated, as well as the % age of the population it represents. By aggregating these areas, world population coverage per technology can be calculated.

REFERENCES

[1] https://www.ericsson.com/en/reports-and-papers/mobility-report.

[2] Ericsson ConsumerLab, 5 ways for a better 5G. 2021.

[3] Source for network and device statistics: GSA and GSMA. 2022.

[4] https://www.ericsson.com/en/reports-and-papers/mobility-report/mobility-visualizer.

[5] https://www.ericsson.com/en/internet-of-things/platform.

[6] ericsson.com/mobility-visualizer.

ADVISORY BOARD ANNOUNCED FOR GEO WEEK 2023

Geo Week Advisory Board set to help craft programming, recommend speakers, and deliver critical insights to geospatial and built world professionals

Organizers of Geo Week, the premier event that champions the coming together of geospatial technologies and the built world, have announced an impressive list of influential leaders within the geospatial and built world industries who will be participating on the 2023 event's Advisory Board.

The 2023 event will take place February 13-15, 2023 in Denver, Colorado.

The Advisory Board will assist in developing conference programming comprised of both general sessions and breakout sessions that delve into the full spectrum of data needs, work processes, software integration and standards in both the geospatial and BIM worlds. Specific vertical industries include architecture, engineering, & construction; asset & facility management; disaster & emergency response; earth observation & satellite applications; energy & utilities; infrastructure & transportation; land & natural resource management; mining & aggregates; surveying & mapping; and urban planning / smart cities.

Members of the 2023 Advisory Board include:

- *Dr. Qassim Abdullah, Woolpert, Inc.*
- *Ashley Chappell, NOAA*
- *Kelly Cone, ClearEdge3D*
- *Kevin Dowling, Kaarta*
- *Martin Flood, GeoCue Group*
- *Birgitta Foster, Sandia National Laboratories*
- *Thomas Haun, Turner Staffing Group*
- *Kourosh Langari, Caltrans*
- *Amar Nayegandhi, Dewberry*
- *Lindsay Prichard-Fox, TiverBuilt*
- *Barbara Ryan, World Geospatial Industry Council (WGIC)*
- *Scott Simmons, Open Geospatial Consortium*
- *Dr. Jason Stoker, U.S. Geological Survey (USGS)*
- *Daniel Stonecipher, Schneider Electric*
- *Dr. Stewart Walker, LIDAR Magazine*
- *Jennifer Wozencraft, US Army Corps of Engineers*
- *Geoff Zeiss, Between the Poles*

The Advisory Board is responsible for recommending conference topics and speakers, reviewing submitted abstracts, consulting on the program, and acting as a resource to develop different aspects of the event. In 2023, the Geo Week conference program will showcase real-world use cases and highlight emerging trends in technology and processes. In addition to the extensive conference program, Geo Week offers a vendor-neutral show floor featuring the newest geospatial and built world products and solutions to qualify and compare.

Geo Week is part of a network of events and media for the global geospatial and built markets organized by Diversified Communications, a leading organizer of conferences, trade shows, and online media with 16 years in the technology arena. Geo Week, taking place February 13-15, 2023, reflects the increased integration between the built environment, advanced airborne/terrestrial technologies, and commercial 3D technologies. Powerful partnership events will also take place at Geo Week, including ASPRS (American Society for Photogrammetry and Remote Sensing) Annual Conference.
Diversified Communications also produces Lidar & Geospatial Newsletter, 3D Technology Newsletter, AEC Innovations Newsletter, GeoBusiness Show (UK), Digital Construction Week (UK), Commercial UAV Expo and Commercial UAV News.

www.geo-week.com