Before the
FEDERAL COMMUNICATIONS COMMISSION
Washington, D.C. 20554

In The Matter of	)	
	)	
Use of Spectrum Bands Above 24 GHz For Mobile Radio Services	)	GN Docket No. 14-177
	)	
Establishing a More Flexible Framework to Facilitate Satellite Operations in the 27.5-28.35 GHz and 37.5-40 GHz Bands	)	IB Docket No. 15-256
	)	
Petition for Rulemaking of the Fixed Wireless Communications Coalition to Create Service Rules for the 42-43.5 GHz Band	)	RM-11664
	)	
Amendment of Parts 1, 22, 24, 27, 74, 80 90, 95 And 101 to Establish Uniform License Renewal, Discontinuance of Operation, and Geographic Partitioning and Spectrum Disaggregation Rules And Policies for Certain Wireless Radio Services	)	WT Docket No. 10-112
	)	
Allocation and Designation of Spectrum Fixed-Satellite Services in the 37.5-38.5 GHz, 40.5-41.5 GHz and 48.2-50.2 GHz Frequency Bands; Allocation of Spectrum to Upgrade Fixed and Mobile Allocations in the 40.5-42.5 GHz Frequency Band; Allocation of Spectrum in the 46.9-47.0 GHz Frequency Band for Wireless Services; and Allocation of Spectrum in the 37.0-38.0 GHz and 40.0-40.5 GHz for Government Operations	)	IB Docket No. 97-95

COMMENTS OF 4G AMERICAS

Table of Contents

I. Introduction 1

II. To Encourage 5G Development, the Commission Should Support Flexible Rules That Encourage Spectrum Use For The 28 GHz and 39 GHz Bands. 3

III. The FCC Should Adopt Broad Geographic Licensing Areas. 5

IV. Any Performance Metrics Should Reflect the Realities of Millimeter Waves and 5G Technology. 9

V. Flexible-Use Rules And Unpaired Blocks Will Best Support TDD and FDD Deployment. 11

VI. Licensing and Band Plans Should Promote Investment in 5G Network Deployment. 14

VII. The Commission Should Continue Exploring Use Of Other Bands Not Included In the NPRM. 15

VIII. Security is a Fundamental Design Criterion. 17

IX. Conclusion 18

COMMENTS OF 4G AMERICAS

I. INTRODUCTION

4G Americas[1] responds to the *Notice of Proposed Rulemaking* (NPRM) in the above-referenced proceedings related to the use of spectrum bands above 24 GHz for mobile and other uses.[2] 4G Americas commends the Federal Communications Commission ("FCC") on seeking to develop "a regulatory framework that will help facilitate so-called Fifth Generation (5G) mobile services."[3]

4G Americas, the leading voice for 5G in the Americas, is an industry trade organization comprised of our region's national operators and manufacturers. The organization's mission is to foster—throughout the ecosystem—the advancement of the Americas' LTE mobile broadband technology and its evolution beyond into 5G network functionality.

4G Americas regularly works with government agencies, regulatory bodies, technical standards organizations, and other global wireless organizations in our efforts to promote seamless interoperability and convergence. This includes 4G Americas' role as a Market Representation Partner in the 3rd Generation Partnership Project (3GPP), its membership in the International Telecommunication Union (ITU) and the Inter-American Telecommunication Commission (CITEL) of the Organization of American States, and its collaborative working agreements with other agencies throughout the Western Hemisphere.

---

[1] 4G Americas' Board of Governors includes representatives from such telecommunications pioneers as América Móvil, AT&T, Cable & Wireless, Cisco, CommScope, Entel, Ericsson, HP, Intel, Mitel, Nokia, Qualcomm, Sprint, T-Mobile US, Inc., and Telefónica.

[2] *Use of Spectrum Bands Above 24 GHz For Mobile Radio Services*, GN Docket No. 14-177, et al., Notice of Proposed Rulemaking, FCC 15-138, ¶ 1 (rel. Oct. 23, 2015) ("NPRM").

[3] *Id.*

To further its mission, 4G Americas plays a major role in researching and educating the industry on technological advancements. As an example of that leadership, 4G Americas recently published a White Paper regarding 5G—*5G Technology Evolution Recommendations* (*5G Evolution White Paper*)[4] —which expands on the White Paper that 4G Americas submitted in response to the Commission's *Notice of Inquiry* (NOI) in this same proceeding.[5] The *5G Evolution White Paper*, attached, further explores technological considerations for future 5G deployment, and provides technical information in response to many of the questions contained in the Commission's NPRM.

In response to the NPRM's inquiries, 4G Americas also provides specific responses to several points. *First*, the Commission should adopt flexible service rules that encourage spectrum use for the 28 GHz and 39 GHz bands in support of 5G development. *Second*, the Commission should adopt broad geographic licensing areas—rather than county-sized licensing areas—to better reflect the reality of 5G technology. *Third*, the Commission should avoid using a population-based performance metric to measure spectrum usage. *Fourth*, flexible use rules and unpaired spectrum blocks are the best means to support multiple duplexing methods. *Fifth*, the Commission should adopt a licensing scheme that provides exclusive use and a minimum of 200 MHz blocks to incentivize investment. *Sixth*, because 5G technology development will attempt to manage the increasing consumer and industrial demands on spectrum, the

---

4 4G Americas, *5G Technology Evolution Recommendations*, White Paper (October 2015), *attached, and also available at* http://www.ghzericas.org/documents/4G%20Americas%20Recommendations%20on%205G%20Requirements%20and%20Solutions_10%2014%202014-FINALx.pdf.

5 *See Comments of 4G Americas to NOI*, citing 4G Americas, *4G Americas' Recommendations on 5G Requirements and Solutions*, White Paper (October 2014), *also available at* http://www.4gamericas.org/documents/4G%20Americas%20Recommendations%20on%205G%20Requirements%20and%20Solutions_10%2014%202014-FINALx.pdf.

Commission should continue to consider proposing rules in the future for flexible centimeter wave use below 24 GHz and the remaining bands initially considered in its *Notice of Inquiry* which were not proposed in this rulemaking. *Seventh*, 4G Americas agrees that security is a fundamental design criterion for upper microwave flexible use services.

II. To Encourage 5G Development, the Commission Should Support Flexible Rules That Encourage Spectrum Use For The 28 GHz and 39 GHz Bands.

4G Americas commends the Commission on its efforts to make the 28 GHz and 39 GHz bands available.[6] 4G Americas believes it would benefit marketplace developments if the FCC were to expeditiously make these proposed frequencies available for 5G, and—importantly—to do so under flexible rules which incentivize incumbent license holders to put the spectrum to its highest-value use.[7] Moving to a *Report and Order* in the near term will reduce investment risk and uncertainty in the U.S. A compelling public interest rationale exists for the FCC to act quickly for these bands. Making them available will foster added investment, innovation, and competition in the U.S., and allow the U.S. to be a leader internationally. Based on the outcome of the World Radiocommunication Conference (WRC-15) and announced plans in certain countries, it is clear other countries are considering these bands for 5G, or as the International Telecommunication Union calls it, IMT 2020.[8] Early action by the U.S. will help expedite global economies for 5G applications and devices.

To further that public interest, 4G Americas supports the Commission's proposal to provide flexibility to existing fixed licensees by authorizing mobile operations in the 28 GHz and

---

6 NPRM ¶¶ 28, 39.

7 NPRM ¶¶ 93, 182.

8 *See* World Radiocommunication Conference, Resolution Com6/20 (WRC-15).

39 GHz bands.[9] 4G Americas believes this course of action provides for expeditious development, implementation, and deployment of 5G systems in line with mobile industry expectations and similar activities worldwide.

4G Americas also supports the Commission's efforts to make sure this valuable spectrum is used, but cautions against complex systems such as an overlay rights auction that will not give licensees exclusive rights. The Commission seeks comment on alternative approaches for geographic areas and bands with existing Local Multipoint Distribution Service (LMDS) and 39 GHz licensees.[10] As one alternative, the Commission considers establishing an overlay right that would allow new licensees flexibility in use, subject to noninterference with the incumbent licensees, and asks whether the Commission should award overlay rights through auction.[11] Such an approach, however, would be problematic because it would not provide the full panoply of rights in the spectrum; a preferable approach would be to auction any licenses that are returned to the FCC or dormant licenses in the 28 and 39 GHz band. 4G Americas believes these licenses should be auctioned in a timely manner to create equal opportunity for various regions across the country to enable 5G deployments.

The Commission raises equipment power measurement techniques and radiofrequency (RF) exposure compliance.[12] Like flexible service rules, equipment operating rules should also be flexible, once basic power and RF exposure parameters are established. The industry is examining new device-to-device (D2D) protocols and expanded mesh applications. The use of multi-hop D2D communication can be a more efficient mode of transmission when nearby

---

9 NPRM ¶ 96.

10 NPRM ¶ 97.

11 *Id.*

12 *Id.* at ¶ 318.

devices have end-user data to convey between each other.[13] The use of direct D2D communications can extend coverage beyond the reach of conventional infrastructure. Cooperative devices can also provide joint transmission and/or reception between multiple devices—a kind of CoMP (coordinated transmission/reception) but on the device side. The Commission has *no more than one-hop* rules developed in earlier rule Parts that were designed to protect government assets.[14] While protection of adjacent government assets are concerns in the 39 GHz band, the Commission could consider more innovative ways to ensure compliance with power limits and RF compliance than restrictions on multi-hop communications in proposed bands.

The Commission also seeks comment on "the technical characteristics for the mobile applications envisioned" for the 39 GHz band.[15] 4G Americas notes that 3GPP has developed a timeline for determining "technical characteristics," consistent with the ITU-R's mandate of their submission of such characteristics to concerned ITU-R Study Groups by March 2017. 4G Americas' *5G Evolution White Paper* provides an early look at considerations for mobile technical characteristics in higher-frequency bands.[16]

III. THE FCC SHOULD ADOPT BROAD GEOGRAPHIC LICENSING AREAS.

4G Americas urges the Commission to consider larger geographic licensing areas that will better reflect the realities of 5G technology development, such as the Internet of Things and smart cities. The FCC has proposed using counties as the base geographic-area unit for licenses

---

13 *See 5G Evolution White Paper* at 33-34, § 6.8.

14 *See, e.g.*, Part 90.

15 NPRM ¶ 46.

16 *See, e.g.*, *5G Evolution White Paper* at 16-26, §§ 4.1-4.3 (discussing potential technical requirements for 5G).

in the 28 GHz, 37 GHz, and 39 GHz bands.[17] As the Commission notes, counties are significantly smaller than traditional license areas such as Basic Trading Areas (BTAs) and Economic Areas (EAs).[18] The Commission invites comment on its proposal, explaining that "the characteristics of millimeter wave spectrum must be taken into account" in determining the appropriate geographic scope of licenses.[19] Establishing the optimal geographic licensing area will be essential to successfully leveraging spectrum bands above 24 GHz to support high-capacity networks and future 5G services.

To achieve its goals of facilitating 5G mobile services and promoting efficient access to spectrum,[20] the Commission should select a large geographic area licensing unit like BTAs or EAs or Partial Economic Areas (PEAs). Indeed, adopting broader geographic licensing areas in the 28 GHz, 37 GHz, and 39 GHz bands will best fit the types of services expected to flourish in these mmW bands.[21] In particular, larger geographic licensing areas will help ensure a thriving mmW spectrum ecosystem by incentivizing investment in the spectrum and supporting innovative 5G use cases.

Many of the use cases for 5G applications contemplate large licensing tracts covering densely-populated areas. The automotive sector, for example, is expected to be one of the drivers for 5G services, with a number of 5G use cases in development for vehicle-to-vehicle and

---

[17] NPRM ¶ 110.

[18] *Id.*

[19] *Id.* at ¶ 109.

[20] *See id.* at ¶¶ 1, 3.

[21] *See id.* at ¶ 111.

vehicle-to-infrastructure communications.[22] Broad geographic licensing areas will best support these automotive use cases by providing the continuous coverage necessary to ensure continuity of service for drivers traveling throughout different counties. Other 5G anticipated use cases would likewise benefit from larger licensing areas. Indeed, self-backhaul, the Internet of Things (IoT), smart grids, smart cities, earthquake and tsunami detection networks, and telemedicine are all 5G use cases that may span several counties.[23] Support of mobile service will require contiguous coverage to provide an attractive and effective service. So implementing a broad geographic licensing approach that is tailored to the mmW use cases, including enhanced mobile broadband services, will help draw investment to these new services.

Larger geographic licensing areas are also necessary because of the technical characteristics of the mmW spectrum. The radio signals in bands above 24 GHz have unique technical characteristics, propagating over short distances with atmospheric absorption further restricting coverage.[24] With these characteristics, mmW bands will be best suited to deliver ultra-high data rates and expand capacity in densely-populated urban areas.[25] But dense urban

---

[22] Nokia Solutions and Networks, 5*G Use Cases And Requirements* at 8 (2014), *available at* http://networks.nokia.com/sites/default/files/document/5g_requirements_white_paper.pdf; *see also 5G Evolution White Paper* at 5, § 2.1.4.

[23] *See* Ericsson, 5G Use Cases, *available at* http://www.ericsson.com/res/docs/2015/5g-use-cases.pdf; *5G Evolution White Paper* at 2-8, 18, §§ 2, 4.1.6.

[24] *Use of Spectrum Bands Above 24 GHz For Mobile Radio Services, Amendment of the Commission's Rules Regarding the 37.0-38.6 GHz and 38.6-40.0 GHz Bands, Implementation of Section 309(j) of the Communications Act – Competitive Bidding, 37.0-38.6 GHz and 38.6-40.0 GHz Bands, Petition For Rulemaking of the Fixed Wireless Communications Coalition to Create Service Rules for the 42-43.5 GHz Band*, Notice of Inquiry, FCC 14-154, ¶ 35 (Oct. 17, 2014) ("NOI").

[25] *See* Letter from Scott K. Bergmann, Vice President, Regulatory Affairs, CTIA – The Wireless Association® to Marlene H. Dortch, Secretary, Federal Communications Commission, GN Docket No. 14-177, at 1 (Oct. 19, 2015) ("CTIA *Ex Parte*") (advising the

areas like Chicago and Washington, D.C. often sprawl across several counties, putting county-based licensing at odds with mmW spectrum bands' most valuable potential.

Moreover, county-based licenses would require extensive operator coordination to minimize interference issues. By employing geographic area licensing on the basis of BTAs, EAs, or PEAs, the Commission would reduce such burdensome measures by ensuring that interference would need to be managed only along the perimeters of the larger service areas. With such a framework, contiguous licenses for larger geographic areas would bear distinct operational advantages while also attracting investment and innovation in the mmW bands. Similarly, larger geographic licensing areas would minimize complexities and overhead associated with time division duplexing (TDD) implementation, which could require careful synchronization among operators.

Finally, as the Commission recognizes, a county-based licensing regime would create administrative complexities, including the need to verify buildout.[26] Rather than shouldering a burdensome geographic licensing framework that does not fit well with the characteristics of mmW spectrum or leading 5G use cases, the Commission should adopt BTAs, EAs, or PEAs as the geographic licensing unit for these bands. To the extent market forces or consumer demand do not support services throughout the BTAs, EAs, or PEAs, a licensee could partition, disaggregate, or lease the portion of its license that may go unused, as the Commission has proposed.[27] In short, the Commission should minimize administrative burdens and encourage

FCC that the benefits of spectrum bands above 24 GHz may be limited to "densely populated areas of the country" due to their inherent propagation limits).

[26] NPRM ¶ 107.

[27] *Id.* at ¶¶ 232, 238.

providers to invest in mmW spectrum during this nascent development stage of 5G services by adopting geographic licensing areas that are larger than the proposed county-based scheme.

IV. Any Performance Metrics Should Reflect the Realities of Millimeter Waves and 5G Technology.

4G Americas notes that there are better metrics to ensure use of the new spectrum bands than the proposed 40% population coverage performance metric and associated milestones. In its NPRM, the FCC proposes specific performance requirements to encourage licensees to provide service to customers in a timely manner.[28] 4G Americas commends the Commission's underlying goal of ensuring that these beneficial services are provided to the public and that spectrum does not lie fallow, while simultaneously trying to strike an appropriate balance regarding operational flexibility.[29]

Deployments in the current bands will differ, however, from traditional lower-frequency deployments. Some of the specific criteria that the FCC suggests hark back to prior outdoor macrocell scenarios, which have a more ubiquitous large-area nature and are supported by a mature, well-established economic model for operators. 5G mmW deployments, on the other hand, are likely to differ from lower frequency macrocell deployments in quite significant ways, in both physical as well as economic aspects. Indeed, the FCC seems to recognize this in its comments, as it openly requests industry guidance and suggestions on how to measure spectrum use.[30]

Millimeter wave propagation is inherently shorter range and quasi-optic in character, hence deployments are expected to be primarily small cells, likely with less ubiquity due to the

---

[28] NPRM ¶ 213.

[29] NPRM ¶ 197.

[30] *See, e.g.*, NPRM ¶¶ 207-208.

stronger shadowing behavior of higher frequency cells deployed amongst the clutter. Consequently, in comparison to macrocells, small cells may be more heavily utilized for capacity augmentation in higher-density locations – i.e., on top of existing macrocell coverage. Millimeter wave is less applicable for augmenting large-area wireless coverage.

Millimeter wave small cells are more likely to be deployed in locations possessing a higher-density of users and/or devices, i.e., where additional capacity is needed. Examples of this may include locations with significant commerce such as public plazas, public transportation hubs, malls, stadiums, corporate campuses, and downtown business centers. Furthermore, with the emergence of IoT, 4G Americas expects that some mmW deployments may also focus on a completely different target audience—machine-type applications—which involve "things" rather than people as their target audience, and could be located in either public or private (e.g., industrial sites) locations.[31]

In summary, we expect the potential range of mmW deployment scenarios to differ significantly from prior conventional FCC criteria that were predicated upon residential census measurements aggregated over large geographic areas. We believe that an entirely new and more appropriate metric will need to be developed for mmW deployments to better capture the level of spectrum utilization.

We believe that it is more appropriate to base performance metrics on usage and/or service levels, rather than census data. The chosen metrics need to be flexible, reflecting the considerable diversity of 5G applications presently being discussed in the relevant standards bodies and industry fora.[32] For example, a more appropriate measure of adequate spectrum use

---

[31] As the Commission notes, NPRM at ¶ 318, massive Machine Type Communications will be one of the drivers of 5G; *see also 5G Evolution White Paper* at 3, § 2.1 (IoT).

[32] *See 5G Evolution White Paper* at 22, § 4.2.10.

might be whether it serves a meaningful quantity of connected devices, carried traffic, or session count. We encourage the FCC to consider one or more of these innovative approaches, recognizing that reconsideration would be possible in the future should the approach warrant adjustment or prove to be inappropriate.

Another important point is that the economic realities of small cell deployments are still nascent, as the Commission recognizes,[33] and the supporting business models are just beginning to emerge. Any performance criteria the Commission imposes needs to account for the elevated amount of operator risk involved relative to macrocell deployments.

Deployment costs still loom large, including outdated, per-site local zoning procedures developed for macro cells. These costs extend to issues such as backhaul. In addition, important dependencies also exist with regard to the development of a viable ecosystem of supporting chipsets and devices and infrastructure, capable of serving consumer mobile broadband as well as IoT applications in these bands. Hence we believe that all dates for measuring performance criterion must be predicated upon the availability of supporting devices and infrastructure.

V. FLEXIBLE-USE RULES AND UNPAIRED BLOCKS WILL BEST SUPPORT TDD AND FDD DEPLOYMENT.

4G Americas supports the Commission's proposal to adopt flexible use in the proposed bands "by allowing TDD and FDD deployment."[34] Recognizing that TDD offers multiple advantages and may eventually be the chosen duplexing method, we support the Commission's proposal regarding flexible duplexing rules and recommend that at this stage of millimeter wave technology development, the Commission refrain from mandating a specific duplexing method.

---

33 NPRM ¶ 212.

34 NPRM ¶ 269.

If the FCC decides on the necessity to specify paired or unpaired blocks,[35] we recommend unpaired blocks, as this arrangement can support either TDD or frequency-division duplexing (FDD) (pairing the spectrum with other bands as determined by the licensee). A technology neutral approach is warranted because TDD may be the optimal duplexing method for mobility at the millimeter wave frequencies for a host of reasons, such as manufacturability, smart antenna efficiency, smart antenna simplicity, and spectral efficiency.[36]

Manufacturability. With the current state of the technology, building a mobile device duplexer which would be required for FDD to operate at these frequencies does not seem economically viable. The isolation necessary for full FDD duplex operations is far greater than the ~20 dB that can be achieved with known duplex technologies. Furthermore, those technologies would have a 3 to 5 dB insertion loss. TDD, on the other hand, needs no duplexer.

Smart Antenna Efficiency. Advanced antennas systems (AAS), such as those with next generation beamforming, operate at peak efficiency when the transmit and receive operations are on the same frequency and can transmit and receive as soon as the channel conditions are known. TDD inherently provides that capability.

Smart Antenna Simplicity. TDD and AAS in combination permit massive MIMO, which enables higher spectral efficiency and supports asymmetry ratios of 10:1. With continued growth in video and other high-resolution content applications expected for 5G, download traffic demand has been and is expected to continue to be a major factor driving the need for efficiency gains.

---

[35] *See, e.g.*, NPRM ¶ 270.

[36] *See* NPRM ¶¶ 268-270.

Spectral Efficiency. TDD spectral efficiency is likely to be far greater because the uplink:downlink ratio can be adjusted to match the uplink:downlink traffic demand.

While it is true that there are incumbent FDD operations in the 27.5-28.35 GHz and 38.6-40.0 GHz bands, the 27.5-28.35 GHz band was developed without a channelized band plan. The 39 GHz service (38.6-40.0 GHz) does have 14 pairs of 50 MHz channels. For both bands, existing operations can easily be protected by coordination among users. While that approach will lead to small protection zones representative of the antenna pattern, network equipment cannot be deployed within that protection zone, so a device may not transmit. Eventually moving to a synchronized TDD plan for backhaul and eliminating the protection zones will benefit the incumbent licensee too. This in turn will benefit the market and allow for deployments that meet the needs of all current and future users, ensuring a successful expansion of broadband services. Sound engineering design and deployment will mitigate interference between backhaul and mobility.

The mobile industry understands that traditional backhaul solutions must continue to be supported. The extreme densification of 5G radio nodes will require new interference avoidance solutions, including possibly the exchange of information between schedulers on the network side. Upgraded backhaul will be critical to support the necessary exchange of information with very small delays.[37] Therefore, a new backhaul model will evolve and appear on street "furniture" such as traffic lights, light standards, telephone poles, sides of buildings—closer to the device. Backhaul cannot be an afterthought when developing new technologies and spectrum

[37] *See 5G Evolution White Paper* at 31, § 6.5 (Advanced Inter-Node Coordination).

must support these new backhaul deployment models.[38] Accordingly, 4G Americas supports the Commission's proposal to adopt flexible use rules and unpaired spectrum.

VI. LICENSING AND BAND PLANS SHOULD PROMOTE INVESTMENT IN 5G NETWORK DEPLOYMENT.

4G Americas reiterates the importance of licensing spectrum on an exclusive basis to provide certainty for investment in 5G network deployment. In general, the Commission proposes to license 5G services in the 28, 37 and 39 GHz bands through competitive bidding and allow licensees full flexibility to aggregate and disaggregate spectrum to provide the coverage and capacity needed in their particular licensed markets. Global harmony in spectrum bands, as well as the same licensing structure, is vital to benefit the U.S. ecosystem with economies of scale.[39] New novel hybrid licensing schemes are still in the experimental phase; therefore, we encourage the FCC to hold off on expanded implementation of any innovative hybrid approaches for the bands above 24 GHz[40] until such time as the viability and effectiveness of hybrid schemes have been demonstrated.

With regard to band plans,[41] 4G Americas recommends a minimum block size of 200 MHz based on the 5G requirement of high data rates everywhere in the order of Gbps in dense environments to several 100 Mbps in urban and suburban environments. Coupled with industry's plans for 5G requirements for high reliability and availability and low energy consumption for devices, a 200 MHz channelization scheme will provide for efficient use of spectrum, consistent with the Commission's goals.

---

38 *See id.* at 35-36, § 6.10 (discussing wireless backhaul and access integration).

39 *See id.* at 49, § 8 ("5G development should provide global harmonization under a single framework").

40 NPRM ¶¶ 100, 294 (proposing hybrid authorization scheme for the 39 GHz band).

41 NPRM ¶ 116.

Therefore, 4G Americas proposes the following:

28 GHz (27.5-28.35 GHz). This band is currently allocated for LMDS, for A1 sub blocks on a Basic Trading Area (BTA) basis. The FCC proposes to expand the LMDS authorization for Fixed Service to include Mobile Service. Currently, licenses are assigned in single 850 MHz license blocks.

4G Americas proposes that all remaining unassigned licenses in this band be awarded as three 200 MHz and one 250 MHz block.

37 GHz (37.0-38.6 GHz) and 39 GHz (38.6-40.0 GHz). 4G Americas proposes that the FCC combine the 37 and 39 GHz bands to make a single contiguous band of 3 GHz. 4G Americas proposes that the Commission consider any combinations of 500/200 MHz blocks for this band.

VII. THE COMMISSION SHOULD CONTINUE EXPLORING USE OF OTHER BANDS NOT INCLUDED IN THE NPRM.

4G Americas urges the Commission to move forward to make additional bands of spectrum available so that the U.S. can maintain a leading role in 5G development. The Commission deferred on proposing service rules in the NPRM at this time for the other bands that had been introduced and explored in the NOI.[42] The Commission stated that it may revisit the other bands at a later date after it "develop[s] a further record in this proceeding, as technology develops, and as [the FCC] develop[s] a further record on compatibility issues with other allocated . . . services."[43] 4G Americas notes, however, the importance of early spectrum allocation to stay ahead of development. The time it takes for proposed spectrum reallocation

---

[42] NPRM ¶ 60.

[43] *Id.*

and rulemaking necessitates expedition by the Commission, proposing flexible use rules for as much centimeter and mmW spectrum as possible. Indeed, both Commissioners Pai and O'Rielly commented that the Commission should move forward with the additional bands explored in the 2014 NOI, seeing "no persuasive reason for leaving these bands on the cutting room floor."[44]

The Commission mentions the desire to wait for the outcome of WRC-15,[45] recognizing that the conference could influence the Commission to address some of these other bands at a later date.[46] As the Commission knows, since the NPRM was issued, many of the bands "left on the cutting room floor" in the NPRM have been teed-up as bands that will be studied for possible identification for IMT 2020 at WRC-19. WRC-15 adopted a WRC-19 agenda item to consider new IMT spectrum at the WRC-19, and specifically "to conduct and complete in time for WRC-19 the appropriate sharing and compatibility studies, taking into account the protection of services to which the band is allocated on a primary basis" for a list of frequency bands.[47] Given the overlap in bands, the importance of innovation, and the goal of maintaining U.S. leadership in 5G development, we urge the Commission to reevaluate in future proceedings the bands from the 2014 NOI that have been removed from the 2015 NPRM in light of the WRC-15 outcome.

---

[44] *See* NPRM at 134, Statement of Commissioner Ajit Pai; *see also* NPRM at 137, Statement of Michael O'Rielly.

[45] The World Radiocommunication Conference (WRC) is a conference typically held every four years where ITU (United Nations International Telecommunications Union) member states ("administrations") consider allocation proposals towards the goal of harmonizing spectrum use internationally. Administrations at WRCs make a number of decisions that impact the future of the mobile industry and other participants like vendors, operators and industry groups can attend as members of administrations' delegations or as observers.

[46] *See* NPRM ¶¶ 13-14.

[47] *See* World Radiocommunication Conference, Resolution Com6/20 (WRC-15) (identifying spectrum bands for study in advance of WRC-19).

In addition, 4G Americas is of the view that spectrum below 24 GHz will play an important role in the next generation of mobile radio because of the difference in propagation characteristics that could be relevant for different 5G applications.[48]

As the Commission is aware, "besides just technology advances and system architecture evolution, it is clear that additional spectrum allocated for mobile broadband will be required to meet the projected demand."[49] Accordingly, the Commission should press forward in making additional spectrum available for flexible use to allow America's innovators to meet the demands of consumers of both mobile broadband and industrial/commercial IoT for more connected devices.

VIII. SECURITY IS A FUNDAMENTAL DESIGN CRITERION.

The Commission seeks comment on how to ensure that effective security features are built into key design principles for all mmW band communications devices and networks.[50] 4G Americas agrees with the Commission that support for security will be a fundamental component in the design of any new network architecture and protocols developed for upper microwave services.[51] Mission-critical applications of mobile services such as smart grids, telemedicine, industrial control, public safety and automotive, have security requirements to defend against intrusion and to ensure uninterrupted operations.[52] Industry's development of 5G can potentially address the core security objectives of confidentiality, integrity and availability.[53] To ensure

---

[48] *5G Evolution White Paper* at 8, 22, and 30, §§ 3.1, 4.2.9, and 6.2.

[49] *See id.* at 47.

[50] *See* NPRM ¶ 261.

[51] *5G Evolution White Paper* at 23, § 4.3.2 and at 40, § 6.14.

[52] *Id.* at 22, § 4.2.8 (Security).

[53] *See* NPRM ¶ 261.

integrity of user information, 5G standards could ensure that information is not tampered with accidentally or deliberately during transit, through authentication at the source of the information and at its reception. 4G Americas assures the Commission that its members are exploring means to ensure the integrity of each data object in 5G networks.[54] Likewise, standards development could address the ability to keep confidential information away from unauthorized users, both through proper use of authentication and data protection through encryption and other means. 5G networks will likely have to be able to defend against Denial of Service and other security attacks.

IX. CONCLUSION

4G Americas thanks the Commission for its efforts to make high frequency spectrum available for mobile and other uses. The Commission's adopted rules should reflect the needs of 5G development, including flexible use rights, exclusive spectrum rights, and rules that will allow 5G use cases to succeed in the market. 5G use cases will have very diverse requirements. For instance, sensor applications will require low data rates while remote surgery and autonomous vehicles require high data rates and low latencies. The mmW and upper cmW service rules should be flexible enough to support existing and future use case requirements, thus avoiding the need to introduce a dedicated regulatory framework for each emerging use case.[55]

4G Americas likewise continues to urge the Commission to make more spectrum available for 5G. Additional spectrum is absolutely crucial as the U.S. telecommunications industry moves forward to 5G capability, since such additional spectrum is the only means by which providers, like those represented by 4G Americas, can successfully and efficiently

---

[54] *Id*. at ¶ 263.

[55] *See 5G Evolution White Paper* at 22, §4.2.10.

accommodate the tremendous surge in mobile wireless data demands that 5G in part seeks to address.

Respectfully submitted,

Chris Pearson

President, *4G Americas*
1750 112th Avenue N.E.
Suite B220
Bellevue, WA 98004

January 26, 2016

Attachment: 5G Technology Evolution Recommendations White Paper

The Voice of 5G for the Americas

5G Technology Evolution Recommendations

October 2015

TABLE OF CONTENTS

Executive Summary ... 1

1 Introduction ... 2

2 Market Drivers and Use Cases for 5G ... 2

2.1 Internet of Things (IoT) ... 3

2.1.1 Smart Grid and Critical Infrastructure Monitoring ... 3

2.1.2 Smart Cities ... 4

2.1.3 m-Health and Telemedicine ... 4

2.1.4 Automotive ... 5

2.1.5 Sports and Fitness ... 6

2.2 Extreme Video, Virtual Reality and Gaming Applications ... 6

2.3 Explosive Increase in Density of Data Usage ... 6

2.4 Public Safety ... 7

2.5 PSTN Sunset ... 7

2.6 Context-Aware Services ... 8

3 Co-Existence of LTE End-to-End Ecosystem with 5G ... 8

3.1 Relation Between 5G Radio Access and LTE ... 8

3.1.1 New Radio Access Technology Tight Interworking with LTE ... 9

3.1.2 New Radio Access Technology Loosely Interworking with LTE ... 10

3.2 Alternatives for the Evolution Path of LTE ... 10

3.3 Co-Existence of LTE with New 5G Radio Access and Core ... 11

3.3.1 5G Network and Interworking Objectives ... 11

3.3.1.1 Option 1 – 5G Packet Core Interfacing with Multiple Access Technologies ... 13

3.3.1.2 Option 2 – Introducing an Interworking Function to Enable Interworking ... 14

3.4 4G Americas 5G Performance Requirements ... 15

4 Requirements for 5G ... 16

4.1 User-Driven Requirements ... 16

4.1.1 Battery Life ... 16

4.1.2 Per-User Data Rate and Latency ... 17

4.1.3 Robustness and Resiliency 17
4.1.4 Mobility 17
4.1.5 Seamless User Experience 17
4.1.6 Context-Aware Network 18
4.2 Network-Driven Requirements 18
4.2.1 Scalability 18
4.2.2 Network Capacity 19
4.2.3 Cost Efficiency 19
4.2.4 Automated System Management and Configuration 19
4.2.4.1 Status of SON Technology in 4G 19
4.2.4.2 General Direction of SON Evolution Towards 5G 19
4.2.4.3 How Evolution to 5G Affects SON 20
4.2.4.4 SON Evolution to Evolved LTE SON in Non-Virtualized Network 20
4.2.4.5 SON Evolution to Evolved LTE SON as Driven by Virtualization 20
4.2.5 Network Flexibility 21
4.2.6 Energy Efficiency 21
4.2.7 Coverage 22
4.2.8 Security 22
4.2.9 Diverse Spectrum Operation 22
4.2.10 Unified System Framework 22
4.3 Aspects of 4G Network Architecture That Can Be Enhanced By 5G 22
4.3.1 Enhancement of Networking Flexibility 23
4.3.2 Additional Support for Essential Functions as Fundamental Attributes of Networking Layer . 23
4.3.3 Providing More Flexible Mobility Solutions 23
4.3.4 Expanded Form of Multi-RAT Integration and Management 24
4.3.5 Enhanced Efficiency for Short-Burst or Small-Data Communication 25
4.3.6 Expanding Context Information Known to the Network 25
4.3.7 Self-Organization 25
5 Regulatory Considerations 26

6 Potential Technologies for 5G 28
6.1 Massive MIMO 28
6.2 RAN Transmission at Centimeter and Millimeter Waves 29
6.3 New Waveforms 30
6.3.1 Advanced Multi-Carrier Transmission 30
6.3.2 Non-orthogonal Transmission 30
6.4 Shared Spectrum Access 31
6.5 Advanced Inter-Node Coordination 31
6.6 Simultaneous Transmission Reception 32
6.7 Multi-RAT Integration and Management 33
6.8 Device-to-Device Communication 33
6.9 Efficient Small Data Transmission 34
6.10 Wireless Backhaul/Access Integration 35
6.11 Flexible Networks 36
6.12 Flexible Mobility 37
6.13 Context-Aware Networking 39
6.14 Information Centric Networking (ICN) 40
6.14.1 Introduction 40
6.14.2 ICN Architecture 41
6.14.2.1 Naming Data 41
6.14.2.2 Mobility Support 42
6.14.2.3 Data-Centric Security 42
6.14.2.4 Name-Based Routing and Forwarding 43
6.14.2.4.1 ICN Routing Operations 43
6.14.2.5 Pull-Based Connectionless Transport 44
6.14.2.6 In-Network Storage 44
6.14.3 Examples OF ICN Benefits and Use-Cases 45
6.14.3.1 Simplified and Enhanced Network Substrate 45
6.14.3.2 Examples of Application Areas 46

6.15 Moving Networks....47
7 Spectrum....47
8 Conclusions and Recommendations....48
Acronym List....50
Acknowledgements....51

EXECUTIVE SUMMARY

This white paper expands the foundation of 4G Americas' view of 5G recommendations that was published in the October 2014 white paper titled, *4G Americas' Recommendations on 5G Requirements and Solutions*.

In the wake of worldwide deployments of 4G technologies, the industry has already started laying the foundation for 5G; and until recently, much of this discussion has been outside North and South America.

The International Telecommunication Union (ITU) has defined requirements and approved standards for earlier mobile communication generations: IMT-2000 ("3G") and IMT-Advanced ("4G"). It will play the same role in defining requirements for 5G and eventually approve standards based on technologies that meet these requirements.

Today's conversation is centered on predicting what the mobile industry will look like in 2020 and beyond. Some countries have indicated that they would like to see 5G deployed even before 2020. The consensus emerging is that the industry must give enough time for technology breakthroughs that deserve the moniker 5G. Also, there is a consensus that because LTE and LTE-Advanced are presently being deployed, there is considerable life left in 4G. In fact, the LTE family of technologies should remain viable through at least 2020 because they will continue to evolve and advance in terms of higher speeds and greater capacity. Carrier aggregation (CA), small cell enhancements and device-to-device signaling are just some of the examples of how LTE is advancing.

When will 5G be deployed? What will characterize networks in 2020 and beyond? What are the likely solutions and technologies that will come in to play? These are some questions currently being discussed among operators, the supplier community, research institutions, standards bodies, trade organizations and governments. Examination of 5G requirements and solutions is basically an exercise in planning a network evolution plan that spans six to seven years. While past generations have been identified by a major new technology step, such as the definition of a new air interface, the expectation is that 5G will be approached from an end-to-end system perspective and include major technology steps both in the radio access network and core network. These steps can be evolutionary or revolutionary by introducing a completely new concept.

This paper examines the 5G market drivers, use cases, requirements, regulatory considerations and technology elements. The 5G market drivers and use cases described in Section 2 include the Internet of Things (IoT), extreme video, Public Switched Telephone Network (PSTN) sunset, public safety and context-aware services. Based upon these 5G market drivers and use cases, Section 3 describes the co-existence of the LTE end-to-end ecosystem with 5G, Sections 4 and 5 describe the requirements and regulatory aspects such as low latency, high throughput, mobility on demand, high reliability and resiliency and network flexibility. Section 6 describes potential technologies for 5G, including packet core, RAN and aspects applicable to end-to-end 5G systems. Section 7 discusses the spectrum aspects associated with 5G networks including both licensed and unlicensed spectrum options.

These 5G requirements and recommendations have been identified by 4G Americas and developed for the purpose of being considered for the further development of the end-to-end 5G system.

1 INTRODUCTION

Wireless has evolved since its inception and seen several generational changes to the service offerings. The generational term is meant to help delineate differences between technologies and discussions are ramping up about fifth-generation (5G) wireless access. 5G is associated with the next step of IMT and IMT-2020 and initial planning is currently under way in the ITU. Additionally, a number of other changes in the end-to-end system will be a part of 5G evolution both in the Radio Access Network (RAN) and core network.

There seems to be a broad consensus that 5G will be introduced around 2020. However, the work on 5G is still in its early stages in terms of examining use cases, requirements and component technologies. A primary question is whether or not 5G will include another new air interface or a collection of air interfaces, each for a different scenario and use case.

Because the industry is working to define the 5G requirements for deployments in the 2020 timeframe without having actual technology developed, one should be cautious that some requirements may not be feasible in the expected timeframe.

The aim of this paper is to address the following areas from a North American perspective:

- What are the key use cases and corresponding key challenges and requirements for wireless beyond 2020?
- What are the key new technology components and solutions that can be used in combination to address these challenges and requirements?

It should be noted that the future of 5G wireless access as referred to above is much more than just radio-interface technology. 5G wireless access should be seen as the overall future solution to providing wireless access to people and devices.

A clear definition of 5G or 5G requirements is not yet available. However, requirements such as support of large number of connected devices, "always online," energy efficiency and support of flexible air interfaces may not be achieved by just an evolution of current systems; instead, those requirements may require 5G to have new protocols and access technologies.

2 MARKET DRIVERS AND USE CASES FOR 5G

3G and 4G technologies have mainly focused on mobile broadband use cases, providing enhanced system capacity and offering higher data rates. This focus will clearly continue in the future 5G era, with capacity and data rates being driven by services such as video.

However, the future will also be more than just enhancements to the "conventional" mobile broadband use case. Future wireless networks should offer wireless access to anyone and anything. Thus, in the future, wireless access will go beyond humans and expand to serve any entity that may benefit from being connected. This vision often is referred to as the "Internet of Things (IoT)", "Networked Society," "Machine-to-Machine communications (M2M)" or "machine-centric communications." North American operators' best customers are no longer humans; they are increasingly machines such as smart utility meters, digital signage and vehicle infotainment systems.

As users begin to use more interconnected devices to play games and collaborate, the ability of the devices themselves will need to expand to create personal networks. Even machines that need to talk to wide-area networks for alarm monitoring, home health, fleet management and many other applications will continue to grow. It is not inconceivable for machine customers to outnumber human customers.

The industry foresees a future of wireless in an increasingly interconnected world where voice, video, medical, entertainment and other applications and services will be served by a highly integrated and automatically configurable network. Users will simply request the information they need and the information will be delivered to their desired location and device.

Issues that need to be addressed are the ways in which the user community will need to interact with information in an environment where on-demand and high-speed mobile data have become a reality. That means there will be some significant changes in the way that user interfaces are designed and higher-layer features are developed. The Internet model known today may not be what best serves 5G users.

In summary, 5G is about enabling new services and devices, connecting new industries and empowering new user experiences. This will entail connecting people and things across a diverse set of scenarios.

This section describes several envisioned 5G use cases across multiple industry verticals. It should be noted that what is described is only a subset of the use cases that can be envisioned. Also, new, yet unknown use cases will most likely emerge, and 5G should have the flexibility to adapt to them.

2.1 INTERNET OF THINGS (IoT)

A wide variety of cellular-enabled IoT applications will be prevalent by 2020, ranging from smart utility grids to earthquake/tsunami detection sensors that warn the public. All such applications can and are beginning to get deployed even on today's cellular networks. However, IoT applications are predicted to grow at a much faster pace than what existing networks and cellular technologies can optimally handle.

To support possibly billions of IoT devices, a wireless network infrastructure is needed which is not only highly scalable in terms of its capacity, but can also optimally handle differing service needs of various IoT verticals. Examples of differing service needs include diverse requirements for mobility, latency, network reliability and resiliency. These diverse requirements may require re-architecting key components of the cellular network, such as to support mobility on-demand only for those devices and services that need it. The following example use cases involving Machine Type Communication (MTC)[1] will become the norm in the 2020 timeframe.

2.1.1 SMART GRID AND CRITICAL INFRASTRUCTURE MONITORING

Today's societies depend on a wide array of critical infrastructure to function properly. Malfunction or damage to this infrastructure could result in huge financial impact, quality of living degradation and even loss of life. The 2003 power blackout in the Northeastern U.S. is an example of how infrastructure failure can bring an entire region and its economy to a halt. Other examples of such disruption include bridge and building structural failure leading to collapse, or water and sewer systems malfunctioning. It is therefore important to monitor the "health" of critical infrastructure reliably and cost effectively.

Critical infrastructure monitoring is an expensive undertaking, often requiring service levels achievable only by dedicated wire-line connectivity. For instance, in order to detect a fault in a high-voltage

---

[1] Within the industry several different terms are used to describe machine to machine communications, these terms include IoT, M2M, MTC, etc. This white paper uses these terms interchangeably.

transmission line and be able to take corrective action to prevent cascading failures, the required communication latency is beyond what current wireless networks can achieve. Similarly, structural monitoring requires the provisioning of a large number of low-data-rate, battery-powered wireless sensors, which today's wireless networks are not optimized to support both in terms of battery life and cost efficiency. 5G will be designed to support reliable low-latency communications among densely deployed devices that are subject to power constraints and wide-ranging data-rate requirements.

2.1.2 SMART CITIES

Massive urbanization is an ongoing trend around the world that's severely straining city services, resources and infrastructure. According to the World Health Organization, by 2030, six out of every ten people will live in a city. Smart City initiatives aim at improving cost, resource and process efficiency of cities, while maintaining a high living quality for their rising populations. The following are three potential examples of 5G-enabled Smart City use cases:

Smart Transportation: Traffic congestion is becoming a major issue in many urban areas and is leading to productivity loss, environmental pollution and degradation of quality of life. 5G will enable a real-time collection of massive amounts of data from vehicles, drivers, pedestrians, road sensors and cameras to help streamline traffic flow. For example, it can help optimize traffic lights and road usage, direct public transportation to where it is needed most, navigate vehicles to avoid congestion and raise tolls to limit traffic entering a congestion zone.

Smart Building: Urban buildings are major consumers of energy and resources. Streamlining building operations will lead to increased productivity and energy efficiency. For example, 5G-connected sensors/actuators can help optimize building temperature, humidity and lighting based on current activities inside them. They will also enable buildings to detect when hidden pipes and cables need repair, unauthorized access takes place, office supplies are running low, and even when garbage bins are full. This information allows building management to take appropriate action in a cost-effective and timely manner.

Smart Home: Home security and automation applications constitute another M2M service area that is expected to grow significantly in the future. Examples include the transmission of home security alarms and surveillance video data to commercial monitoring stations.

2.1.3 M-HEALTH AND TELEMEDICINE

Telemedicine is a major tool for improving healthcare access both in remote rural and urban areas. It can help reduce health care costs while improving health outcomes. A major enabler in this field is the comprehensive use of cloud-based electronic medical records that would serve as repositories of medical information about individual patients. With 5G, these records, which contain high-resolution medical images and video, can be made available to physicians and medical professionals anytime and anywhere. Remote, real-time general physician and specialist consultations would also contribute to cost savings, convenience and better and timelier medical outcomes.

A major hurdle in the realization of this scenario is the lack of a wireless infrastructure that would need to handle the voluminous nature of medical images and video with sizes ranging from hundreds of Megabytes to Gigabytes per instance. The increasing use of diagnostic tools such as 3D and 4D ultrasounds[2], CAT scans and MRIs, and the miniaturization of this equipment to a portable/hand-held form factor, will lead to even higher demands being placed on wireless networks. In addition, massive improvements in the quality of low-cost mobile displays and application software now available to medical professionals make this field ripe for massive adoption. Bio-connectivity, which is the continuous and automatic medical telemetry (e.g., temperature, blood pressure, heart-rate, blood glucose) collection via wearable sensors, is another strong emerging trend that will add to the wireless communications requirements. 5G will enable these and other future medical applications through significant improvements to wireless data throughput and network capacity.

2.1.4 AUTOMOTIVE

Advanced Driver Assistance Systems (ADAS) and Autonomous Vehicles are emerging trends in the automotive space. Together, they bring a number of benefits, including better safety, fewer collisions, less congestion, better fuel economy and even higher productivity for the drivers. 5G wireless technologies supporting high-speed, low-latency vehicle-to-vehicle and vehicle-to-infrastructure communications are key enablers of ADAS and Autonomous Vehicles. In addition, today's drivers and passengers are demanding richer infotainment options which are adding to the strain on wireless networks. The following section describes several potential automotive use cases for 5G.

Vehicular Internet/Infotainment: Increasing content consumption by vehicle occupants will greatly contribute to the need for wireless bandwidth and mobile network capacity. Typical infotainment options include video, audio, Internet access and upcoming applications such as augmented reality and heads-up displays. For these applications, vehicle occupants will expect a user experience comparable to those offered by their home and office networks. The vehicles themselves form another group of Internet users for map, traffic data and high-resolution picture download, as well as sensor data and image upload.

Pre-Crash Sensing and Mitigation: Collisions lead to injury and property damage, as well as time and productivity loss due to traffic congestion. Pre-crash sensing enables vehicles to sense imminent collisions and exchange relevant data among vehicles involved, allowing vehicles and drivers to take counter-measures to mitigate the impact of the collision. Pre-crash sensing requires highly reliable and extremely low latency vehicle-to-vehicle communications.

Cooperative Vehicles: Limited highway capacity in many cities often results in severe traffic congestion. Cooperative vehicles use Vehicle-to-Vehicle and Vehicle-to-Infrastructure communications to safely operate vehicles as a self-driving car train on a highway in order to improve highway capacity, reduce occurrence of driver error and achieve better fuel economy. To ensure safety and reliability while operating as a self-driving car train, reliable and very low-latency communications among vehicles and with the infrastructure are needed.

Inter-Vehicle Information Exchange: Peer-to-peer inter-vehicular communication using D2D cellular technology under the guidance of the operator policies can allow vehicles to communicate information related to road safety and traffic congestion directly in a mesh fashion, thus offloading data from the traditional RAN infrastructure. This is just one possible example of the type of information that can be exchanged.

---

[2] 4D Ultrasound are ultrasound videos with 3D images where the 4th dimension is time. Most 4D ultrasound videos today are less than 5 frames per second. For greater diagnostic utility, they would need to be much higher resolution and about 30 frames per second. For cardiac applications, greater than 100 frames per second would be desirable.

2.1.5 SPORTS AND FITNESS

Fitness-related applications, such as activity and body monitoring applications that track walking, running, and biking activities, metabolic rate, cardiovascular fitness, sleep quality, etc. will constitute a significant vertical market in M2M services. Some of these applications will utilize body or personal area networks to collect biometric information and then use cellular networks to transmit it back to centralized data acquisition sites.

2.2 EXTREME VIDEO, VIRTUAL REALITY AND GAMING APPLICATIONS

Future wireless communication systems will support extreme video and gaming applications that use features such as augmented and virtual reality. Such immersive multimedia services would require the use of technologies such as 3D audio, 3D video and ultra-high-definition formats and codec(s). Examples of such services include:

- Mobile telepresence with 3D rendering capabilities that will extend well beyond the traditional wired office environment.

- Internet gaming, including wirelessly delivered gaming control with high-resolution graphics and dynamic management of feedback mechanisms via smartphone to ensure an enhanced, augmented reality gaming environment.

- Adoption of higher resolution devices, head-mounted displays and wearables in fields such as emergency services, public safety, telemedicine, smart cities, professional services and retail is expected to place further demands on mobile networks.

This type of interactive experience will require the network to support much lower latencies and much higher bandwidths than what are possible today.

2.3 EXPLOSIVE INCREASE IN DENSITY OF DATA USAGE

The use cases outlined so far in this section identify some general trends in the industry:

- The number of devices using cellular networks is expected to increase significantly in the coming years. In other words, the density of cellular devices (devices:area) will increase. A large part of this increase will be coming from M2M services.

- Some future services will require much higher data rates compared to what is typically achievable today. Examples of such services have been provided in the previous use case on extreme video and Internet gaming.

Concentration of devices using Ultra High Definition (4K and 8K) video and high-resolution picture and video-sharing applications occur at event venues such as stadiums. In addition, significant variations in UL:DL traffic ratios imply a need for an air-interface design that can more flexibly assign traffic capacity to the different transmission directions. The effect of these two factors on network traffic will be multiplicative, resulting in an explosive increase in data traffic demand per square mile of the coverage area, especially in urban environments. Some of this is validated by looking at existing traffic trends where data traffic density in urban environments such as stadiums, financial businesses, hospitals, universities and major transportation corridors has increased dramatically. To handle this surge of data traffic demand by 2020, network capacity would have to be increased by orders of magnitude.

2.4 PUBLIC SAFETY

The U.S. is planning to deploy an LTE broadband network for public safety at 700 MHz to leverage pricing of standardized commercial equipment. Canada is currently evaluating the use of an LTE broadband network for public safety at 700 MHz. It seems natural that future wireless broadband networks will also need to consider public safety in their fundamental design. Some of the “special” public safety needs include:

Mission-Critical Voice: This allows a public safety responder to push a button (push-to-talk) to communicate with other public safety responders. This needs to be extremely reliable, working both on and off network without any delay for dialing phone numbers. The feature needs to allow communication with one or more groups (e.g., local police, regional police and local public safety) in real time. Public safety users must be able to monitor multiple groups simultaneously (scanning communication on different groups) and allow additional users to join an on-going group discussion.

Broadband Data: Much of this will be IP traffic from a public safety device to a server, possibly in the cloud. Although this capability can be handled by existing LTE equipment, it is important that 5G consider the following public safety use cases:

- High-resolution security cameras monitoring public spaces and property with the captured images/video analyzed to alert authorities when incidents occur or persons or interest are detected.
- Drone- or robot-based surveillance systems to monitor remote areas.
- Wireless sensors and tracking devices used for intrusion detection, bio and chemical hazard detection and emergency personnel tracking.

The data generated by these and many other modalities will significantly strain 4G radio link and networks.

Besides these needs specifically for public safety officials, 5G systems will need to support legacy public safety features such as Public Warning Systems (PWSs), emergency calling, Multimedia Emergency Services (MMES) and lawful intercept. To support all such use cases, future wireless networks must provide a robust, highly reliable, resilient and low-latency communication infrastructure.

2.5 PSTN SUNSET

The Public Switched Telephone Network (PSTN) sunset in North America is scheduled prior to 2020. With the general industry trend of migrating towards wireless communication, it is expected that in the 2020 timeframe and beyond, wireless broadband networks will be commonly used to replace the PSTN. Therefore, the 5G ecosystem must also serve today’s landline needs. For 5G networks to be considered as a viable replacement to PSTN, they must exhibit the same levels of reliability and robustness. In addition, PSTN services primarily serve stationary customers that do not require support for mobility. The concept of mobility-on-demand that can simplify the packet core and make it scalable should be explored.

2.6 CONTEXT-AWARE SERVICES

The past decade has seen a tremendous rise in the use of always-on, Internet-connected devices. The users of such devices are consistently bombarded with information, most of which may not be relevant or actionable for them. For the most part, existing service models require users of such devices to reach out to the Internet to get the useful information and/or service that they desire.

In such a service model, amongst other things, users first have to figure out the best match for their request and then find out how to get to it. With the ever-increasing amount of available information, it is quite evident that this service model is not scalable. A desired approach is for a service to be context aware and be able to provide a seamless delivery of the right set of information at the right time using the right means. This approach can also be described as instead of the user going to the Internet and figuring out a way to fulfill its needs, the Internet comes to the user with the right information.[3]

3 CO-EXISTENCE OF LTE END-TO-END ECOSYSTEM WITH 5G

3.1 RELATION BETWEEN 5G RADIO ACCESS AND LTE

There seems to be a general agreement within the industry that the beyond-2020 radio access will consist of two tracks:

There will be a continued evolution of LTE. In this context, "evolution of LTE" implies that the evolution is constrained by backwards compatibility, implying that earlier-release LTE devices should still be able to access the carrier. Thus, evolution of LTE is specifically relevant for frequencies below 6 GHz being deployed by LTE prior to 2020.

In parallel, there will be a new radio-access technology not constrained by backwards compatibility. This new radio-access technology will, at least initially, target new spectrum that is not already deployed with LTE. Thus, it will include spectrum above 6 GHz, but may also include new spectrum below 6 GHz, if such spectrum would be made available. In a longer-term perspective, new radio-access technology may also migrate into spectrum used by LTE.

In the foreseeable future, LTE and new radio-access technology will exist in the same network with, at least in many cases, the new radio-access technology operating on higher frequencies in dense local deployments and LTE providing wide-area coverage on lower frequencies. It is also possible that with new spectrum allocation in the lower frequencies before 2020, the new radio-access technology is deployed as the coverage layer to carry control signaling and IoT traffic more efficiently than LTE. Some of the new mission critical use cases requiring low latency and high reliability, and other use cases requiring massive scalability, will need wide area coverage and may not be efficiently supportable using evolved LTE with the backward compatibility restriction. Furthermore, the new radio interface could be designed so that there is tight coupling between low and high band deployments of the new interface; for example, to support beam tracking. Thus, this white paper foresees at least the following two deployment models emerging over time.

---

[3] "5G: Personal Mobile Internet beyond What Cellular Did to Telephony", *Gerhard Fettweis, Siavash Alamouti,* IEEE Communication Magazine, February 2014.

3.1.1 NEW RADIO ACCESS TECHNOLOGY TIGHT INTERWORKING WITH LTE

To maximize efficiency and service level of such joint deployments, there should be possibility for tight interworking between LTE and the new radio-access technology. This includes, for example, dual-connectivity where the new technology provides high-speed user plane connectivity when available while LTE provides continuous coverage and control-plane connectivity. As another example, a downlink user plane could be provided via a new radio-access technology on higher frequencies, while an uplink user plane, due to reduced coverage on higher frequencies, is provided by LTE on lower frequencies.

This kind of interworking between LTE and the new 5G radio-access technology will imply some constraints on the later. This will primarily be valid for higher layers. However, it may also impact the basic timing structure which would need to be aligned with that of LTE, at least on some level. For example, on could envision a timing structure with “sub-frames” being sub-multiples of 1 ms, allowing for lower latency but still being aligned with the LTE 1 ms structure. In this sense, LTE may condition the actual choice of basic timing structure for the new radio access technology, and evolved LTE should also respect it for the sake of backwards compatibility.

In previous generations, tight interworking between radio access technologies was not necessarily sought as systems from future generations did not depend on a previous generation for their operation. Thus, 2G, 3G and 4G systems traditionally work independently (even if some kind of loose interoperation may be in place in the form of inter-RAT handovers, reselections, etc.). In this tight interworking approach for 5G, the situation may be different because of the broad scope in foreseeable frequencies (from cmWaves to mmWaves), requirements and use cases (from massive IoT deployments to ultra-reliable low-latency access). Deployments addressing more than one such frequencies and/or use cases may require tight interworking between evolved LTE and the new radio-access technology, as in the above mentioned examples of dual connectivity or uplink/downlink split.

Notwithstanding such backwards compatibility, evolved LTE may incorporate some of the technical solutions currently under study for 5G. In fact, some of the hottest topics for 5G are not exclusive of it, but are already being discussed for current 4G (like network virtualization, IoT or massive Multi-Input Multi-Output (MIMO), among others). This means that probably some (if not all) of these techniques will be developed for evolved LTE to a certain extent. It is therefore natural to question how the evolution of LTE will take place, and whether that evolution will ultimately merge within 5G or not.

Since LTE may provide control plane connectivity and wide coverage in this tight coupling approach, the evolution of LTE is an important question that needs to be addressed. The robustness of LTE control mechanisms (such as mobility, resources management, scheduling, etc.) at lower frequencies makes it very attractive as an anchor technology for the different access techniques foreseen in 5G.

However, this option has a price to pay which is twofold:

- The need to adapt 5G control mechanisms to the specifics of LTE radio access; and
- The need to evolve LTE towards the direction set by 5G.

The first implication means that, for example, low-frequency control of 5G radio access must comply with LTE numerology and basic constraints (e.g., the poor spectral confinement of the Orthogonal Frequency Division Multiplexing (OFDM) signals).

The second implication, about evolving LTE towards the direction set by 5G, is part of a more profound discussion about the alternatives for the evolution of LTE which are discussed below in Section 3.2.

3.1.2 NEW RADIO ACCESS TECHNOLOGY LOOSELY INTERWORKING WITH LTE

In this approach, the new radio access technology in low frequencies serves as the primary carrier carrying control signals or uplink traffic for new radio access technology in the high frequencies deployed in dense areas for higher capacity. The low and high frequency deployments of the new radio access technology may operate in carrier aggregation mode possibly with a common medium access control layer or in dual connectivity mode. In this approach, the new radio access technology, free from backward compatibility restrictions, may be designed to operate as the control layer for high frequency deployments with appropriate frame structure and provision for carrying control signals for high frequency user plane. Even in this model, the new technology will be coupled with already existing LTE deployments in dual connectivity mode where necessary to increase wide area network throughputs.

3.2 ALTERNATIVES FOR THE EVOLUTION PATH OF LTE

LTE ecosystem (comprising of LTE and LTE-Advanced standards) has expanded stunningly in the last years, due in part to the great performance that 4G networks can achieve in terms of spectral efficiency and number of connected devices. Given the huge amount of Capital Expense (CAPEX) and Operating Expense (OPEX) already spent on 4G networks over the last years, it is not surprising that operators want them to be alive for a long time to come.

Under this perspective, there are three factors in the current 5G debate that can give rise to significant uncertainties with respect to 4G evolution:

- The different waveform candidates that are being researched as candidates for 5G (such as Filter-Bank Multi Carrier (FBMC), Universal Filtered Multi Carrier (UFMC) or Generalized Frequency Division Multiplexing (GFDM[4]), to name a few).
- The alternative multiple access schemes that are being researched and for which intra-cell orthogonality is not guaranteed.
- The use of millimeter wave and upper centimeter wave frequencies[5] which may motivate the use of either of the two above techniques (and other ones).

Waveform candidates proposed as evolutions of OFDM aim at improving some of the known weaknesses of OFDM, such as the bad spectral confinement of the signals or the need for tight time/frequency synchronization.

The implications of the evolution path of LTE can be very significant if LTE is considered as the primary anchor radio access technique for 5G technology. There are basically three different approaches for this evolution path:

---

[4] Filter Bank Multicarrier (FBMC), Universal Filtered Multi Carrier (UFMC) and Generalized Frequency Division Multiplexing (GFDM) represent evolutions of OFDM waveforms with improved spectral and time-domain characteristics.

[5] Millimeter wave frequencies are those between 30 and 300 GHz, while centimeter waves extend from 3 to 30 GHz. Upper centimeter wave frequencies include those frequencies above 10 GHz not traditionally used for cellular access.

- Gradually converge the evolution of LTE towards the specifics of 5G radio access, particularly in terms of multiple access techniques (if different than in current LTE) and/or involved numerology, under the limits imposed by the backwards compatibility with previous releases.

- Keep the evolution path of LTE basically unchanged in terms of multiple access schemes and numerology, and rather adapt 5G control mechanisms to the specifics of LTE.

- Keep LTE and 5G as separate systems, thereby progressing at different paces without any actual compatibility between them (as in previous cellular generations).

The first option (where the evolution of LTE would adapt to the specifics of 5G) seems unlikely given the long run of current 3GPP LTE Releases, which precludes any drastic change in its basic foundation (like numerology or waveform design) for the sake of backwards compatibility.

The second option (where 5G control mechanisms would be based on LTE since its very inception) seems more reasonable, but somehow constrains the foreseeable changes in radio access (at least for 5G control tasks). By way of example, for the control plane, no waveform changes would be allowed and intra-cell orthogonality would be assumed.

The third option (where LTE and 5G run different and parallel paths) would make sense if evolved LTE is finally agreed not to be an anchor radio access technique for 5G technologies. In this case, LTE evolution may be decoupled from 5G, thus motivating the need to provide the necessary signaling for interoperation between both systems[6].

Both the first and second options point towards having sufficient commonality between 5G radio access and the evolution of LTE. The third option would however yield separate (non-compatible) systems, as traditionally happened in previous 3G and 4G cellular systems. In this case, significant work would have to be conducted in 3GPP for efficient interworking between 4G and 5G, which will be unavoidable in practical systems.

It is expected that some or all of these unknowns will be clarified when the first 3GPP Study Item on 5G RAN technical requirements starts in December 2015.

3.3 CO-EXISTENCE OF LTE WITH NEW 5G RADIO ACCESS AND CORE

3.3.1 5G NETWORK AND INTERWORKING OBJECTIVES

The Next Generation Mobile Networks (NGMN) Alliance and 4G Americas have both advocated certain principles for the design of the 5G core network and access technology interfacing architecture. NGMN's white paper lists several key requirements relevant to 5G System and Packet Core, including the following:

- A functional decoupling of core and RAN network domains.

- A RAN technology agnostic architecture where introduction and connection of new radio technology will be possible in a plug and play manner.

[6] Although it is not unlikely that 3G systems (and even some 2G systems) have to coexist with 5G by the 2020 timeframe, interoperation with 3G/2G is not usually considered.

- Independent evolution of core and RAN network domains should be possible.
- C- and U-plane functions should be clearly separated, with open interfaces defined between them, in accordance with Software Defined Networking (SDN) principles.
- A decoupling of hardware and software functions of network elements in all network domains.
- A cost and effort efficient upgrade path so that operators can leverage significant investments into existing operational infrastructure and maximize its utilization.
- Real-time and on-demand network configuration and automated optimization.
- Flexible and cost efficient network operation.
- Maximize utilization efficiency of available network resources.
- Dynamic relocation of network resources, fully controlled by the operator.

4G Americas also lists several 5G requirements that are relevant to development of a next generation of packet core:

- The 5G network architecture should be such that the air interface and core network can evolve and scale independently of each other.
    - Changes/enhancements to one domain should not mandate changes/enhancements to the other.
- 5G networks must also support multi-Radio Access Technology (RAT) connectivity efficiently and effectively.
    - Provide access to agnostic packet core across multiple radio technologies (e.g., cellular, Wi-Fi, etc.) to support uniform authentication, session continuity, and security across radio technologies.
    - Provide plug and play capability where a new access technology may be attached to the packet core without any modifications.

These principles will lead to a new architecture that:

- Minimizes access specific nodes.
- Has separation of user and control plane.
- Provides mobility on-demand.
- Implements operator policies via SDNs.
- Implements Quality of Service (QoS) in an access agnostic manner.
- Implements Authentication, Authorization, and Accounting (AAA) in an access agnostic manner.
- Supports use of service specific core via network virtualization and service specific network slicing.

While some of the features like NFV are independent of network architecture and hence can and will be implemented for existing Evolved Packet Core (EPC), other aspects like access agnostic QoS and AAA, mobility on-demand, access agnostic mobility tunnels and use of SDNs for policies would result in a significant change to the current packet core. Thus, the 5G packet core may not be an evolution of EPC, but rather a brand new solution.

During the development of LTE in 3GPP and for its coexistence with legacy 3GPP technologies, a tight coupling approach was adopted that allowed seamless service continuity for legacy services. While the differences between the General Packet Radio Service (GPRS) (2G/3G packet core) and EPC were not as significant, a tight coupling resulted in an increased packet core complexity involving interworking between 2G/3G and EPC packet core entities. In the case of the 5G packet core, tight coupling with legacy can be reduced for further simplification. For instance, legacy interworking towards circuit switched domain in the 2G and 3G networks can be reduced.

Like the 2G/3G packet core, EPC is not access agnostic and has dependencies on the LTE RAN. For example, EPS bearers and QoS principles are access specific QoS concepts that would need to be made access agnostic in 5G architecture. Similarly, other aspects like mobility protocols should be revisited. The 5G packet core should also consider introducing concepts such as mobility on-demand, common AAA mechanisms and the need for a mobility anchor point which is access agnostic. This is to ensure that the new packet core can easily be upgraded independent of the access technology and also interwork with a variety of access technologies. With significant changes expected in the packet core, a tight coupling with legacy LTE RAN that has not been decoupled with EPC would result in an introduction of a similar complexity as was seen in LTE and 2G/3G interworking. The following strategies could be considered as possible options for integrating LTE and 5G:

- Option 1: 5G packet core interfacing with multiple access technologies.
- Option 2: Introducing an interworking function to enable interworking.

3.3.1.1 OPTION 1 – 5G PACKET CORE INTERFACING WITH MULTIPLE ACCESS TECHNOLOGIES

The 5G packet core is developed in such a way that it interfaces with multiple access technologies, including LTE and 5G. LTE evolves in a manner such that the evolution of LTE supports two protocol stacks. One protocol stack allows LTE RAN to connect to existing EPC while the second allows LTE RAN to connect to the new 5G packet core. In other words, this option requires the evolved LTE RAN to support the 5G control and user-plane stacks to interface with the 5G core. In addition, the evolved LTE RAN may continue to support the LTE legacy control and user-plane stack required to interface with EPC. Operators who have upgraded their networks to support 5G would benefit from being able to use the same protocol stack regardless of whether or not a supporting User Equipment (UE) is in the LTE evolution or 5G coverage. Other operators who do not deploy 5G can continue to use the evolution of LTE with EPC using the old protocol stack.

Internet Protocol (IP) session continuity between LTE evolution and 5G can be provided in this option as part of 5G architecture and is not discussed in this section any further.

Service continuity between legacy RATs and 5G can be provided through higher layer means (e.g., Multipath Transmission Control Protocol (MPTCP)). Common 5G user-plane functions (e.g., Transmission Control Protocol (TCP) optimization) as depicted in Figure 1 can also be accessed from legacy RATs. However, there is no control-plane interface between the legacy cores and the common user-plane functions. This implies that the legacy cores cannot “control” those u-plane functions (e.g., control the

TCP optimization function or other functions like the Deep Packet Inspection (DPI) function) that traffic from the legacy cores may be routed to.

It should be noted that the LTE only UE(s) cannot register with the 5G core network and thus need to be registered with the EPC core. It should also be noted that legacy LTE RAN and 2G/3G RAN are expected to connect to EPC and GPRS core respectively; thus, operators with legacy LTE RAN and 2G/3G RAN deployments are expected to maintain EPC (for legacy LTE RAN), GPRS (for 2G/3G RAN) and 5G packet core (for the evolution of LTE and 5G RAN). This option is illustrated in Figure 1 below:

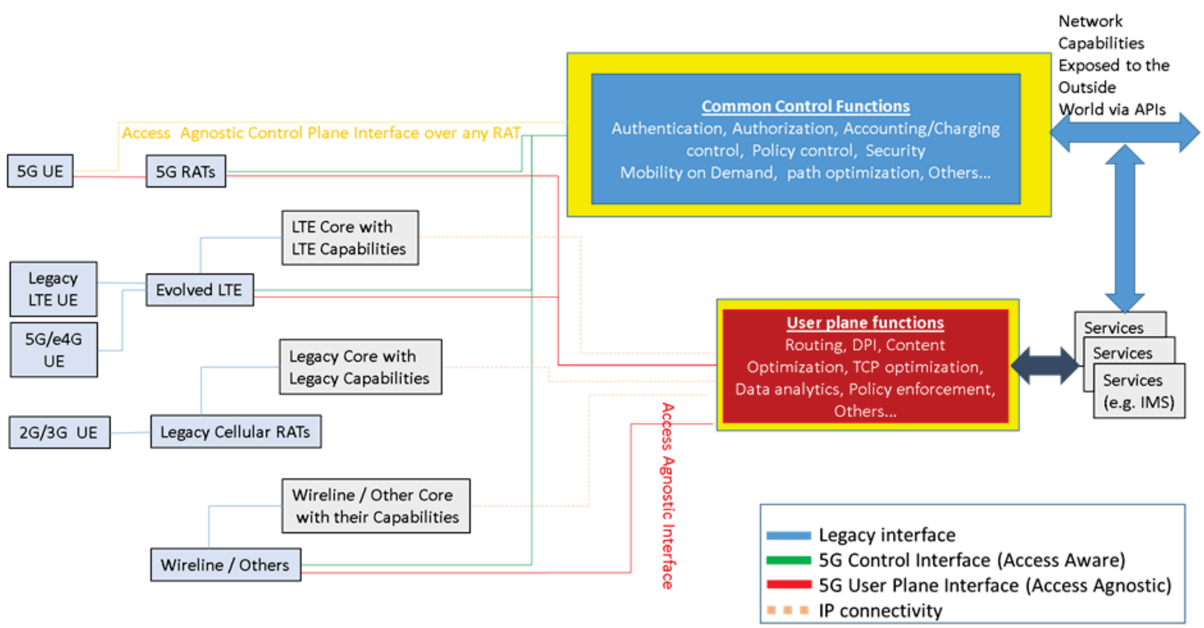

Figure 1. Architecture for Access Technology Interfacing.

3.3.1.2 OPTION 2 – INTRODUCING AN INTERWORKING FUNCTION TO ENABLE INTERWORKING

This option assumes that operators will independently maintain and evolve their LTE and 5G networks. To achieve this, a special interworking function between 4G RAN and the 5G packet core network will enable interworking between legacy LTE and 5G for both control and user-plane. Thus, there is no impact to LTE radio access network and 5G packet core network.

IP session continuity between 4G and 5G networks can be provided in this option. Service continuity between legacy RATs and 5G can be provided through higher layer means (e.g., MPTCP).

Similar to Option 1, it should be noted that the LTE only UE(s) cannot register with the 5G core network directly; thus, they need to register with the EPC core or interworking function that supports the necessary functionality. This option is illustrated in Figure 2 below:

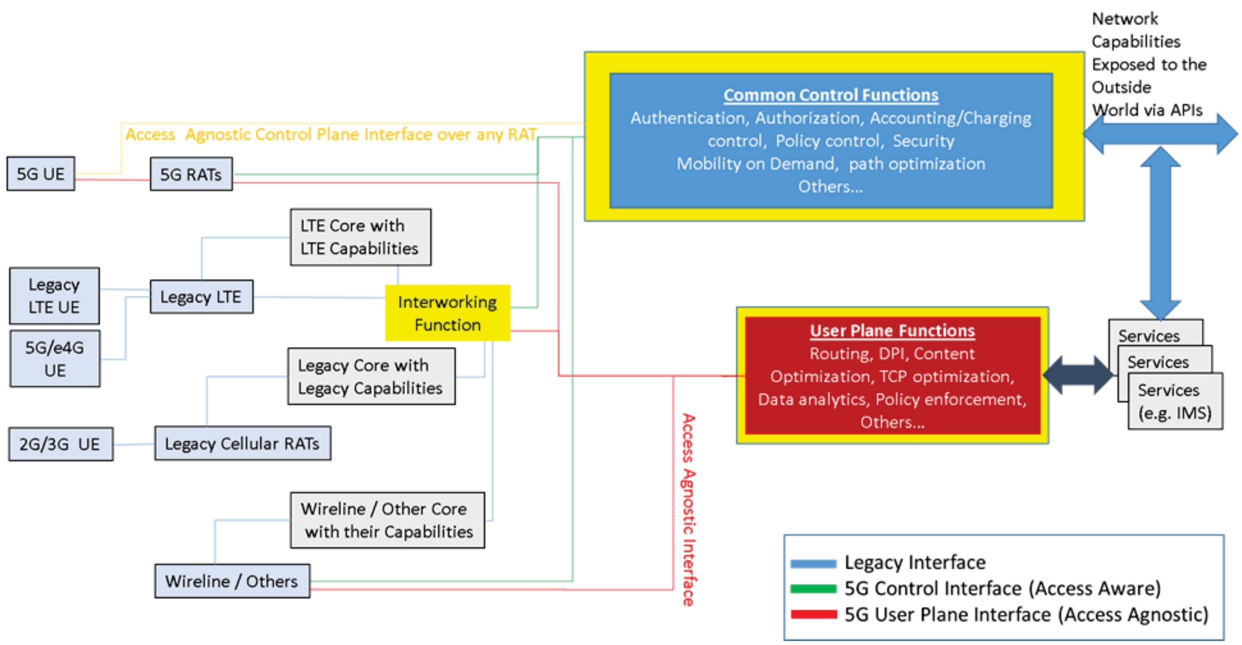

Figure 2. Architecture for Access Technology Interfacing with Interworking Function.

3.4 4G AMERICAS 5G PERFORMANCE REQUIREMENTS

Table 1 below is 4G Americas' view of the 5G performance requirements for various situations and environments. The legend for the values contained in Table 1 is defined in Table 2.

Table 1. 4G Americas 5G Performance Requirements.

Scenario	End-to-End Latency	Mobility	Data Rate		Battery Life	Reliability
			DL	UL		
Mobile Broadband						
Dense Urban	Medium	Low	High	Medium	Short	Medium
Urban	Medium	Low	High	Medium	Short	Medium
Suburban	Medium	High	Medium	Low	Short	Medium
Rural	Medium	High	Medium	Low	Short	Medium
Remote[1]	Low	Low	High	Medium	Short	Medium
Automotive (section 2.1.4)	Low	High	Medium	Medium	Medium	High
Extreme Video,	Low	Low	High	Medium	Short	Medium

[1] While it may not be feasible or practical for commercial deployment of 5G networks in remote locations, the availability of 5G networks in remote locations for Public Safety could be critical. Establishment of 5G networks in remote environments may be performed on an as-needed basis using deployable network equipment. Such deployable network equipment may be vehicular based, aircraft based, air-dropped containers, etc.

Scenario	End-to-End Latency	Mobility	Data Rate		Battery Life	Reliability
			DL	UL		
Virtual Reality and Gaming Applications (section 2.2)						
Public Safety (section 2.4)	Medium	High	High	Medium	Medium	High
PSTN Sunset (section 2.5)	Medium	Low	Low	Low	Medium	Medium
M-Health and Telemedicine (section 2.1.3)	Medium	Low	High	Medium	Medium	High
Smart Cities (section 2.1.2)	Medium	Low	Medium	Medium	Long	Medium
Sports and Fitness (section 2.1.5)	Medium	Low	Low	Low	Short	Low
Increased Density of Data Usage (section 2.3)	Medium	No	High	Medium	N/A	Medium

Table 2. Legend for 4G Americas 5G Performance Requirements.

End-to-End Latency	High (> 100 ms)	Medium (10 - 100 ms)	Low (< 10 ms)
Mobility	No: 0 m/s (static / nomadic)	Low: ~1 m/s (pedestrian)	High: ~>10 m/s (mobile)
DL/UL Data Rate	Low: < 25 Mbps	Medium: 25 – 200 Mbps	High: > 200 Mbps
Battery Life	Long: Years	Medium: Weeks	Short: Days
Reliability	High: > Five 9's	Medium: Four 9's	Low: < Four 9's

4 REQUIREMENTS FOR 5G

Requirements can be subdivided into two categories:

- User-driven requirements in terms of quality of experience, user satisfaction, reliability and speed of the connection.
- Network-driven requirements in terms of network operation and management.

4.1 USER-DRIVEN REQUIREMENTS

4.1.1 BATTERY LIFE

Several of the IoT applications involve battery operated sensor networks that are out in the field and transmit data only occasionally. Wide-scale deployment of 5G-based sensor networks would be possible only if much longer battery life and/or reduced energy consumption by such devices guarantees their unattended operation over a duration spanning years.

4.1.2 PER-USER DATA RATE AND LATENCY

Per-user data rate and latency attributes for a network defines the typical data rate and round-trip delay (respectively) that users experience. Ultimately, the values for these attributes determine the types of applications that can be supported on a network. It is estimated that by 2020, there will be a new class of data rate hungry services with low latency requirements. Previous work has shown that applications in the future such as augmented reality, 3D gaming and “tactile Internet” will require a 100x increase in achievable data rate compared to today and a corresponding 5x to 10x reduction in latency. 5G networks must therefore be designed to meet these data-rate and latency requirements.

4.1.3 ROBUSTNESS AND RESILIENCY

5G networks will increasingly be used as the primary source of communication as a replacement network for PSTN after its sunset. They also will support emergency communications and public safety, including during and after disasters. A key requirement for such use cases is for the network to be robust, reliable and resilient. This requirement would also need to ensure the ability to defend against security attacks such as Denial of Service (DoS) for mission-critical applications such as public safety, smart grids and natural gas and water distribution networks.

4.1.4 MOBILITY

5G systems are expected to support both very-high-mobility scenarios (e.g., high-speed trains and planes), as well as scenarios with low to no mobility for end devices. The technology therefore should be able to cope efficiently with such extreme situations by providing mobility on demand based on each device’s and service’s unique needs and capabilities.

At the same time, machine-type communicating devices can require nomadic access to the network with the purpose of sending reduced amounts of data in mostly static locations. For these nomadic devices, reliability and resilience could be the most important network features than mobility support. Finally, the need for extreme data rates (or extremely low latencies) at specific situations can usually be satisfied with very stable channel conditions in stationary devices. Future 5G systems will have to cover such extreme cases, from no mobility to future high-speed trains or even possibly aircraft.

4.1.5 SEAMLESS USER EXPERIENCE

Current cellular systems provide very high peak quality, such as very high peak data rates. However, the quality often varies substantially over the coverage area of interest. For example, the achievable data rates can be substantially lower for devices far from the base station site or in indoor locations. 5G wireless access should deliver a much more consistent user experience, irrespective of the user’s location. Thus, the achievable quality (e.g., achievable data rate and latency) should be the key quality indicator, where achievable quality should be defined as the quality experienced with perhaps 95 percent probability, rather than peak data rates.

5G will likely comprise a collection of layers, technologies and frequency bands that should seamlessly interwork when moving across networks, layers and/or frequencies. Interruption times of the order of a few milliseconds for both inter-RAT and intra-RAT handovers can be expected in this sense. Services such as ultra-high definition video or tactile Internet will require end-to-end latencies in the order of 1 millisecond. Interruption times well beyond that could destroy the attractiveness of these services.

4.1.6 CONTEXT-AWARE NETWORK

With MTC and greater diversity in human communication devices, it becomes increasingly important for the network to provide the correct resources to meet the unique needs of each application and device. This is possible only if the network is context-aware and hence can dynamically adapt to meet those needs. For example, this means that: 1) full mobility need not be provided to MTC devices that are stationary; 2) 3GPP mobility management and paging need not be provided for services that require only device-initiated communication; and 3) resources that are configured to support long battery life, high reliability, low latency, low cost, secure communications and global roaming is truly needed. Optimizing resource allocation in this manner makes possible simpler, lower cost, application-tailored devices and lower network costs because only the necessary resources are used. It also enables a better end-user/device experience. Mobile operators gain additional abilities for creating service plans customized for individual customers, groups of customers or market segments. An example is creating one set of plans for MTC customers that use a limited amount of network resources, and a second set of plans for smartphones that use a large amount of network resources.

Context includes network awareness, such as the availability of alternative multi-RAT, small cell and macro networks of varying capabilities, application and device awareness with associated service requirements, subscription context such as operator preferences for providing service and subscriber analytics. Awareness of these attributes makes it possible for the network to dynamically adapt to the needs of devices and applications rather than have applications adapt to today’s one-size-fits-all set of access characteristics.

4.2 NETWORK-DRIVEN REQUIREMENTS

4.2.1 SCALABILITY

Support for IoT use cases will be key to the success of 5G networks. An expected 10X-100X increase in the number of devices, primarily because of M2M services, requires network elements that can scale up gracefully to handle this growth. This requirement is true both for the user plane and the control plane. For example, the 5G network should be able to scale well to handle signaling traffic, such as for authentication/authorization for large numbers of IoT devices. Another example is that the user plane must be able to scale well to handle (in)frequent and small data transmissions from large number of devices.

When it comes to supporting a mix of traffic from IoT applications and more traditional services such as voice and video, the term “scalability” has an additional dimension. 5G networks should be able to support both high-data-rate/low-latency conventional services alongside M2M applications that require much lower bandwidths. Each M2M vertical will likely have its own unique traffic pattern. Both frequent and infrequent (bursty) data transmission, for example, will have to be supported in an efficient manner. For example, traffic pattern and transmission requirements (data rate, latency) from earthquake/tsunami warning sensors will be quite different than for traffic from a vending machine. Similarly, many devices in 5G networks will be stationary or nomadic and require no mobility support or only occasional mobility support. 5G network designs therefore should not assume mobility support for all devices and services but rather provide mobility on demand only to those devices and services that need it.

4.2.2 NETWORK CAPACITY

Experience suggests the expectation of a 1000x – 5000x[7] [8] rise in traffic over the next decade. To handle this explosive increase, a key requirement for 5G networks will be to increase traffic-handling capacity – which is defined as the total traffic that the network can handle while still maintaining QoS.

4.2.3 COST EFFICIENCY

With an expected increase in the total network traffic, and the need to stay competitive, the next generation of mobile networks should provide a significant cost benefit over the current generation. The cost improvement should be at least as good as or possibly much better than what has been experienced in going from 3G to 4G. In this context cost refers to both the OPEX and CAPEX of delivering a byte of data to the subscriber. Network Function Virtualization (NFV) will play a key part in achieving cost reductions.

4.2.4 AUTOMATED SYSTEM MANAGEMENT AND CONFIGURATION

In 5G deployments, the network density is expected to significantly increase for a number of reasons including higher data volume density and the use of higher frequency spectrum. To better manage the CAPEX and OPEX of running a network with a much higher number of network nodes, a key requirement is that 5G networks will be able to self-configure as much as possible.

4.2.4.1 STATUS OF SON TECHNOLOGY IN 4G

One of aspects of Self-Organizing Networks (SON) functionality is to provide for efficient automatic network optimization to boost network quality and cut OPEX and CAPEX.

In 4G, the SON technology is represented mostly by Radio Access Network (RAN) SON functions such as Mobility Load Balancing (MLB), Coverage and Capacity Optimization (CCO), etc., running at two different network layers; some belong to distributed SON and some to centralized SON, potentially with coordination between the layers.

There are also network self-organization functions running in Core Network (CN) with certain means of coordination with RAN SON. User Plane Congestion (UPCON) management is an example of such function.

4.2.4.2 GENERAL DIRECTION OF SON EVOLUTION TOWARDS 5G

In 5G, SON is expected to be a coherent functionality that integrates all SON functions across network (i.e., RAN and CN), network/management layers and Radio Access Technologies (RATs) with efficient coordination between the centralized and distributed components. Such solution should be adapted to virtualized network.

[7] "Evolutionary & Disruptive Visions Towards Ultra High Capacity Networks" IWPC's MoGIG White Paper Proposal, 2014.
[8] "Scenarios, requirements and KPIs for 5G mobile and wireless system", METIS Document number: ICT-317669-METIS/D1.1.

4.2.4.3 HOW EVOLUTION TO 5G AFFECTS SON

As noted in Section 2.1, there will be a continued evolution of LTE resulting in Evolved LTE.

It is important to provide a view on SON evolution from 4G to Evolved LTE as required by integration of the Evolved LTE component into the 5G network. Two routes of the SON evolution to e4G can be identified:

- SON Evolution to Evolved LTE SON in non-virtualized network.
- SON Evolution to Evolved LTE SON as driven by virtualization.

4.2.4.4 SON EVOLUTION TO EVOLVED LTE SON IN NON-VIRTUALIZED NETWORK

Desired development in this direction for SON (resulting in "Evolved LTE SON"):

- Provide for sufficient coordination capabilities between centralized SON and distributed SON with the goal to improve efficiency of combined (hybrid) SON.
- Provide for joint operations of RAN SON and self-organization functions in core network such as UPCON.
- Provide for efficient coordination between RATs: Evolved LTE and 5G.
- Provide for efficient collection of comprehensive information of the state of the network including RAN, CN and UE.

4.2.4.5 SON EVOLUTION TO EVOLVED LTE SON AS DRIVEN BY VIRTUALIZATION

In a virtualized network, Virtual Network Functions (VNFs) may be running in different Network Virtual Function Infrastructure (NFVI) domains (e.g., for different levels of RAN centralization). For example, CN VNFs may be running in the central data center of the network operator while RAN VNFs may be running in regional or local data centers to provide for lower latency. It is expected that the virtualized network will include VNFs supplied by different vendors.

In view of this, evolution of SON driven by virtualization is expected to address the following aspects:

- Allocation of SON functions to different NFVI domains to follow different levels of RAN centralization.
- SON components hosted by the corresponding NFVI domains will interface to RAN VNFs via standardized Application Program Interfaces (APIs) to support multiple RAN VNF vendors.[9]

[9] ETSI's Mobile Edge Computing (MEC) project provides an example of similar approach. The White Paper "Mobile-Edge Computing" (https://portal.etsi.org/Portals/0/TBpages/MEC/Docs/Mobile-edge_Computing_-_Introductory_Technical_White_Paper_V1 18-09-14.pdf) outlines Radio Network Information Services (RNIS) interface between the

- There will be intra-SON APIs to provide coordination between SON components located in same NFVI domain or different domains and/or in different network layers (e.g., SON API of Small Cell Forum).

4.2.5 NETWORK FLEXIBILITY

The 5G network architecture should allow the RAN and the core network to evolve and scale independently of each other. Changes/enhancements to one should not mandate changes/enhancements to the other. In order to achieve this goal, the RAN and the packet core should avoid mutual dependencies. Additionally, 5G networks must also support multi-RAT connectivity efficiently and effectively. This includes the ability to:

- Provide an access-agnostic packet core across multiple radio technologies (e.g., cellular, Wi-Fi) to support uniform authentication, session continuity and security.

- Provide plug-and-play capability where a new access technology may be attached to the packet core without any modifications.

Decoupling the packet core from the RAN will also help efforts to further flatten out and simplify the network. Flattening improves network scalability by enabling some of the functions to be pushed closer towards the user/edge.

Revisiting the current architecture for mobility and operator policy management can help minimize interdependency between the RAN and packet core. As part of this effort, new mobility protocols, the concept of mobility on-demand and use of Software Defined Networking (SDN) for providing and enforcing operator policies should be explored.

Recent trends in virtualization envision a separation of the control and data planes, as well as a decoupling of hardware and software so that network functions are mainly driven by software (with hardware equipment being as generic as possible). These principles would enable greater flexibility in deploying network functions on demand.

One final recommendation is for the 5G network architecture for small cells to be more neutral host capable. 3G and 4G have generally required multiple small cells (one per operator) to be deployed in an area or building in order to provide coverage for all 3G/4G devices that enter the area or building. This has been one of the main disadvantages of 3G/4G small cells compared to Wi-Fi. Trends towards wideband Radio Frequency (RF) and Centralized RAN (CRAN) should help evolve small cells to be more neutral host capable, but challenges such as fronthaul requirements should be addressed as the industry moves towards 5G small cells.

4.2.6 ENERGY EFFICIENCY

Rather than only maximizing spectral efficiency, there is an ever growing concern about the energy consumption per bit (expressed in Joules/bit) that represents a measure of the energy efficiency. Network functions should not convey excessive energy (both radiated and consumed by the network

---

3rd party software running at the Base Station MEC platform and the BS software to collect information on the status of the cell. In this approach the role of the MEC platform is similar to NFVI.

infrastructure). More interestingly, energy consumption could be adapted to the current traffic conditions to achieve significant energy savings in off-peak situations.

4.2.7 COVERAGE

Although coverage is ultimately limited by the band in use, 5G will take special care in improving coverage for IoT-related applications to make 5G viable for this emerging market. While it is clear that coverage largely depends on the frequency of operation and density of deployment sites, special actions can be taken so as to ensure optimal coverage for specific services such as IoT, public safety and other critical systems.

4.2.8 SECURITY

Mission-critical applications such as smart grids, telemedicine, industrial control, public safety and automotive, have strict security requirements to defend against intrusions and to ensure uninterrupted operations. 5G should address the following security objectives:

- **Integrity:** Ensure information is not tampered with either accidentally or deliberately during transit. This includes the ability to authenticate the source of the received information and the ability to authenticate the recipient.
- **Confidentiality:** Keep sensitive information away from unauthorized users. This includes proper user authentication, data protection through encryption, etc.

4.2.9 DIVERSE SPECTRUM OPERATION

To serve the use cases outlined in the previous section, 5G is expected to operate in a diverse set of spectrum bands. These include traditional sub-6 GHz cellular bands for coverage and low-power operation, to above-6 GHz bands including millimeter spectrum for ultra-high data rates. As the propagation characteristics and hardware implications of these bands are expected to be substantially different, 5G systems would need to accommodate these requirements in radio access, network architecture, protocol and modem design considerations.

4.2.10 UNIFIED SYSTEM FRAMEWORK

Use cases described in the previous section have very diverse, and sometimes conflicting, requirements. For example, sensor applications generally require low data rates and can tolerate high latencies. Meanwhile, evolving applications such as telemedicine require high data rates and low latencies. 5G should be as flexible and extensible as possible to support existing and future use case requirements, thus avoiding the need to introduce a dedicated system for each emerging use case.

4.3 ASPECTS OF 4G NETWORK ARCHITECTURE THAT CAN BE ENHANCED BY 5G

The basic principles of 4G network architecture were conceived several years ago prior to the explosion in mobile broadband usage. Many of the requirements that have emerged since then have been handled by incremental modifications to the basic architecture. Although capable of satisfying the requirements, 4G architecture does so by increasing the complexity of existing functional components and by adding new functional components. This paper describes some of the limitations in 4G architecture that can potentially be improved in 5G architecture.

4.3.1 ENHANCEMENT OF NETWORKING FLEXIBILITY

With the advent of small cells in indoor environments such as offices, there is a need for some traffic to be routed locally while other traffic needs to access Mobile Network Operators (MNOs) or third-party services. For example, as a result of enterprise Bring-Your-Own-Device (BYOD) policies, devices increasingly have multiple “personalities,” with some applications communicating within the private office environment while other applications communicate with Internet-based consumer services. For local offloading, 4G requires a separate mobile Packet Gateway (PGW) deployed locally largely because mobile-network-specific tunneling is employed for all traffic.

Advances such as Content Distribution Network (CDN) virtualization are being made that enable content caching closer to the device at various locations in the transport network between the base station and the core network. Intelligent-content-request routing mechanisms are being proposed in the context of content-centric networking, some of which may be applicable to traditional CDNs as well. In order for applications on the mobile device to leverage such innovations, a local mobile gateway has to be deployed and devices need to support multiple Access Point Name (APN) connectivity or the network has to initiate an APN switch, which may result in some disruption of ongoing services.

In addition to deploying local gateways, new APNs have to be provisioned in the network and devices, which is sometimes a complex process for MNOs. A solution that eliminates the need for such a specialized mobile specific local gateway, new APN provisioning and associated signaling is preferable.

4.3.2 ADDITIONAL SUPPORT FOR ESSENTIAL FUNCTIONS AS FUNDAMENTAL ATTRIBUTES OF NETWORKING LAYER

Like mobile architecture, Internet architecture continues to evolve to support new use cases. One challenge is that IP was designed at a time when the network’s fundamental objective was to transport data packets between fixed communication hosts quickly and efficiently. With packet headers naming the communication hosts via the IP addressing scheme, the network task has been simply to forward the packets hop-by-hop from one host to the other. Significant elegance and operational efficiency derived from routing protocols have enabled automatic topology mapping and pathology-free routing without significant operator intervention.

In response to the emergence of more elaborate use cases and usage demands (e.g., mobility, content distribution, security), the networking community has incrementally added new functions as either overlays on the existing network or as specialized elements in the network to address the need. The key point is that the importance of capabilities considered essential today was not fully appreciated at the time the original Internet design was conceived. As a result, these capabilities were not incorporated as fundamental elements in the original design. So over the years, the cost of managing and operating the network has progressively increased, largely due to the added complexity introduced by the persistent stream of functionality patches and overlays (support for mobility amongst them).

5G architecture should re-examine the mobile network architecture from the perspective of how it stands to benefit from research in the Future Internet Architecture, and in particular network architectures and protocols that implicitly support mobility, security and content caching (storage) as fundamental components of the network design.

4.3.3 PROVIDING MORE FLEXIBLE MOBILITY SOLUTIONS

It is well known that even in a mobile network, many devices are stationary. For example, video constitutes 55 percent of the traffic, and people are generally static when watching video. Many M2M devices, such as utility meters, are stationary. Furthermore, even if the device is moving, maintaining the same IP address is not required for proper functioning of many applications.

An example is HTTP Adaptive Streaming (HAS). This application works by downloading twp seconds worth of chunks that are buffered in the client with a deep buffer of tens of seconds. If downloading of a chunk is interrupted by a handoff, a new TCP connection with a new IP address can be set up to download the same chunk again. Another example is tracking devices on pets/people where the device is periodically updating the network about the wearer's location. If an active session is interrupted a new session can be set up to perform the location update.

The functions that handle device mobility in 4G are oblivious to the specific requirements of the applications and devices. The same seamless mobility is always provided, incurring additional signaling, processing, memory and bandwidth overheads. More specifically, inefficiencies in 4G include the following:

1. Establishing and modifying the tunnels when they are not needed incurs a number of signaling messages between various network elements. The RAN signaling protocol used in LTE to track both idle and active UEs is an adaptation of legacy 2G circuit voice systems designed for relatively low volumes. Voice calls are typically of longer duration and happen much less frequently than data interactions, especially when data transactions are associated with large volumes of M2M devices. As a result, today's 3GPP handling of mobility entails significant RAN and core network signaling overhead that is unnecessary for devices and applications that are primarily static or nomadic.

2. There is overhead associated with additional headers that are added to every packet, and this consumes additional bandwidth on the backhaul links. The overhead can be significant when packets are small. Processing resources are unnecessarily consumed at the mobile gateway nodes and the base station to encapsulate and de-encapsulate packets in tunnels.

3. Because the tunneling encapsulation and de-encapsulation can occur only at special router nodes designed to handle the associated signaling messages, packets cannot always be routed using the shortest path. In particular, if the destination is closer to the base station to which the device is attached compared to the SGW/PGW nodes, then a specialized solution (SIPTO/LIPA) with a local SGW/PGW is required to avoid triangular routing. Additional cost is incurred to deploy this local SGW/PGW. And if the device has two applications, one that is local and one that is not, then the device needs to support two PDN connections and needs to attach to two PGWs. This essentially doubles the signaling involved.

4.3.4 EXPANDED FORM OF MULTI-RAT INTEGRATION AND MANAGEMENT

Wi-Fi/4G interworking is becoming increasingly prevalent, resulting in devices communicating over multiple radio access technologies. This interworking also has produced network topologies, such as "trusted non-3GPP access," that support seamless access selection, authentication, bearer plane inter-operation and in some cases seamless mobility. This trend is expected to continue with devices communicating over multiple air-interface types.

However, because the supporting network architectures for the different air-interface types were defined independently by different standards bodies, today there is little commonality in network functions and procedures. This fragmentation leads to a coarse level of interworking based on redirecting the mobile to an alternative access technology, where it exercises a different set of network procedures. For example, signaling on different access links is largely separate resulting in duplication of mobility, authentication and policy signaling for each access technology. Moreover, the selection and routing of traffic over the different technologies is left to the device, with some policy-based guidance functions such as the ANDSF. However, this limits the potential to steer traffic across different technologies based on dynamic criteria such as network status. For example, the network cannot alter whether a users' video should be sent over Wi-Fi versus over cellular as a function of network conditions and steer traffic accordingly.

4.3.5 ENHANCED EFFICIENCY FOR SHORT-BURST OR SMALL-DATA COMMUNICATION

Smartphone applications and many M2M devices frequently exchange short bursts of data with their network-side application. When there are no other communication needs, the devices have only a small amount of data to send but nevertheless have to go through a full signaling procedure to transmit the data. This wastes battery life, spectrum and network capacity.

To handle this type of transaction more efficiently, the network needs to support a truly connectionless mode of operation, where devices can simply wake up and send a short burst of data. Upon reception of the short burst, device and application-related state information can be retrieved from a controller function and resources to handle the packet allocated accordingly. Some attempts have been made to address this in 4G through some tweaks. 5G will offer an opportunity to include the requirement upfront in the design, thereby leading to superior solutions.

4.3.6 EXPANDING CONTEXT INFORMATION KNOWN TO THE NETWORK

Today 3G/4G access and the mobile core network have limited knowledge of the device and even less knowledge of applications that are requesting access. Static device and subscription information is available in the 3GPP Home Subscriber System (HSS) data base. This typically includes information about device type and subscribed services (e.g., SMS, voice). Standardized use of this information has been largely limited to determining the authorized PDN connections for the UE and authorized QoS. The network knows even less about the applications being used on devices. Almost all smartphone applications are provided "over the top" by third parties. Unless there are add-on deep packet inspection or analytic tools, the network has no visibility into these applications or their needs.

This contrasts sharply with knowledge obtained by applications and devices that produce big-data-based subscriber analytics, which enable targeted advertising and context-relevant subscriber offers. 5G needs a richer and more flexible method for the network to obtain and utilize information relevant to deciding how network resources should be allocated in the context of operator policy.

4.3.7 SELF-ORGANIZATION

For 5G, SON functionality will continue the evolution path started at 4G, so observations in Section 4.2.4.5 will be applicable. On top of that, the 5G SON concept will provide for the following, as required by 5G technology advances:

- 5G needs a coherent SON solution that integrates all SON functions across network (i.e., RAN and CN), across network / management layers and across RATs. In particular new technology elements (e.g., Section 6) should be addressed.
- The SON architecture should be adapted to virtualized network; above considerations for 4G virtualization are applicable.
- 5G SON should provide for efficient collection of comprehensive information of the state of the network including RAN, CN and UE.

- The SON architecture should allow for easy integration with management (OAM) and service orchestration.
- The SON architecture should allow for easy integration between RAN, CN and UE SON functions, between centralized and distributed SON components.
- The SON architecture should allow for application aware optimization.
- The SON architecture should be able to address interoperation of multiple technologies in 5G, such as multiple Radio Access Technologies (RATs).
- The SON architecture should address network slicing.
- The SON architecture should provide for scalable solution in view of expected raise in the number of Network Elements in HetNet and the number of subscribers in combined Broadband Access & Internet of Things networks.

5 REGULATORY CONSIDERATIONS

The regulatory environment should not be overlooked as 5G technologies are being developed. It is anticipated that existing regulations will be applicable to the 5G environment. The following challenges need to be addressed to support regulatory requirements in 5G:

- **Location Accuracy:** Location accuracy requirements for emergency calls (e.g., 911) continue to evolve, with tighter and stricter regulations being imposed by the regulators. The 5G network must comply with the existing location accuracy requirements, as well as the anticipated stricter location accuracy requirements to be imposed within the next few years. The stricter location accuracy requirements could include both indoor and outdoor environments and an altitude component, and they could be applicable to dense urban, suburban, rural and remote environments. Specifically for indoor locations, where GPS cannot fulfill even current requirements, access subsystem in 5G networks will have to assist in achieving the needed accuracy requirement. In addition to location accuracy requirements for emergency calls, location accuracy may also be applicable to the public safety first responders, such as firefighters inside a burning structure.
- **Lawful Intercept:** It is anticipated that existing requirements for lawful intercept capabilities will continue and may be expanded as communication options and choices evolve. The existing lawful intercept architecture is based upon the principles that communication paths traverse centralized network elements, which can be monitored for lawful intercept purposes. However, the 5G network has the potential for communications paths that do not traverse centralized network elements (e.g., direct device-to-device communications, mesh network communications). So one technical challenge is to develop a 5G architecture that enables communications without transversal of centralized network elements while complying with lawful intercept regulations.
- **Tower Sharing:** 5G networks will have to support multiple radio technologies. Today's cell sites typically have one set of antennas – increasingly MIMO – for each RAT. This design becomes an issue at sites shared by multiple operators and for tower companies. As this may lead to the inability of deploying 5G services, with ripple effects on other licensed services, the 5G architecture should support solutions to minimize the number of antennas in shared multi-RAT environments. Additionally, regulators should be encouraged to limit the scope of tower sharing because this undermines operators' ability to innovate and shield those innovations from rivals.

- **Flexible Spectrum Use:** Even when fiber is available near a cell site, wireless backhaul allows faster deployment until the fiber can be installed. In many cases, both rural and urban, wireless backhaul often becomes the only technologically or economically feasible alternative. On the other hand, it is expected that 5G will use spectrum above 6 GHz. Therefore, ideally any spectrum used for 5G access should be flexible enough to also be used as backhaul. Spectrum licenses should be flexible enough to allow operators to meet the rollout demand while being capable of using 5G spectrum for backhaul when appropriate.

- **Mandated Digital Roaming:** There are two different types of domestic roaming. In the first, customers can roam to get service when they're outside their operator's coverage area. In the second, customers can roam to get service when they're inside their operator's coverage area but their operator has weak or no signals in that particular spot, what's known as a "black hole." The second type is much more demanding on both networks' resources. It is important to note that some countries mandate domestic roaming on digital technologies. Therefore, 5G networks should be capable of coping with such demand, while regulators should be aware about the more stringent requirement imposed on the latter scenario.

- **Critical Infrastructure:** Most telecom networks are classified as critical infrastructure because mobile broadband services are becoming an essential part of daily life. That responsibility will only increase while supporting IoT services. In order to fulfill this mandate, 5G networks must be robust, reliable, resilient and secure, as specified in Section 4.1.3. Therefore, self-healing functionality, such as domestic roaming and network sharing, should be considered. Specifically in the case of network sharing, multiple carriers support is an essential requirement, one that's not really addressed today. This would ensure coverage and backhaul backup, which are often the weakest link of a wireless network. Those improvements should also be recognized and supported by adequate actions from regulators.

- **Emergency Telecommunication Service (ETS):** This service gives government users priority access to the next available channel in crisis/disaster situation, when networks often get congested. For example, some countries (e.g., the U.S., Canada) have set up Wireless Priority Service (WPS) in partnership with mobile operators. Therefore, 5G networks should be capable of supporting these essential services. In addition, roaming with public safety mobile broadband networks, such as FirstNet in the U.S., may be required.

- **Public Warning System (PWS):** 3GPP TS 22.268 already provides requirements for implementation of emergency alert systems around the world. Examples include Earthquake and Tsunami Warning System (ETWS), Commercial Mobile Alert System (CMAS) in the U.S., EU-ALERT in Europe and Korean Public Alert System (KPAS). It is important that such functionality be maintained while deploying 5G networks. It will also be important for 5G networks to be capable to fulfill another important PWS aspect not fully covered currently, support for multiple languages.

- **Accessibility:** Mobile broadband services are part of daily life, so 5G services must be accessible to people with disabilities, as is the case with 3G and 4G.

- **Use of SIM, E164 and TAC:** From a regulatory perspective, as 5G networks are going to all-IP, it may be appropriate to consider alternatives to the use of SIM, E164 (international public telecommunication numbering plan) and Type Allocation Code (TAC), which uniquely identifies each mobile device's model number and version. For example, is IMEI still the way to capture the right device for legal interception, or should a new approach be taken? TAC (a subset of the IMEI), was developed to uniquely identify a wireless device. However, it is now found that it may not be appropriate anymore for an operator to identify a specific version of a model number. In the case if SIMs, should it be further improved or a new approach taken, although use of IP certificate may not be an improvement?

6 POTENTIAL TECHNOLOGIES FOR 5G

This section describes potential technologies for 5G to address the market drivers and use cases described in Section 2, the LTE and 5G co-existence requirements in Section 3, the requirements for 5G identified in Section 4 and the regulatory considerations defined in Section 5. Specifically, this section will discuss the following potential technologies:

- Massive MIMO
- RAN Transmission at Centimeter and Millimeter Waves
- New Waveforms
- Shared Spectrum Access
- Advanced Inter-node Coordination
- Simultaneous Transmission Reception
- Multi-RAT Integration and Management
- Device-to-Device Communications
- Efficient Small Data Transmission
- Wireless Backhaul/Access Integration
- Flexible Networks
- Flexible Mobility
- Context Aware Networking
- Information Centric Networking (ICN)
- Moving Networks

6.1 MASSIVE MIMO

MIMO employs multiple antennas at the transmitter and receiver, and is a well-known technique to increase the spectral efficiency of a wireless link. When devices have only a few antennas, multiple base station antennas can be used to simultaneously serve multiple users using the same time frequency resource. This requires knowledge of the channel between the base station antennas and the receiver antennas so that appropriate pre-coding can be employed to eliminate interference between signals transmitted to different users. Single-user MIMO and multi-user MIMO are both part of the 4G standards.

Massive MIMO extends the multi-user MIMO concept by dramatically increasing the number of antennas employed at the base station to be significantly larger than the number of users being served simultaneously in the same time-frequency block. With hundreds of antennas serving tens of users simultaneously, spectral efficiency can increase 5x to 10x, while users on a cell's fringes can maintain high throughput.[10] Furthermore, the pre-coding required for each user's signal reduces to simple conjugate beamforming. The major challenge of acquiring channel information at the transmitter is solved by employing time-division duplexing, where the same spectrum is used in both the DL and the UL,

[10] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," IEEE Transactions on Wireless Communications, vol. 9, no. 1, pp 3590-3600, Nov 2010.

ensuring that the DL channel is nearly the same as the UL channel. Users transmit orthogonal pilots on the UL from which the UL channel is estimated and then used for the conjugate beamforming on the DL.[11] UL pilot transmissions may not necessarily be orthogonal across cells because only limited resources can be devoted to pilot transmissions. This results in what is called “pilot contamination,” producing channel estimation errors. Pilot contamination is mitigated using pilot reuse, where the pilot sequences are reused only in cells outside of the immediate neighborhood of the cell where it is used.

Advances in radio and antenna technology are required to cost-effectively deploy a large number of antennas at the base station. While channel propagation is reciprocal, the receive and transmit paths may not be. Thus antenna calibration may be required to account for any substantial difference that may arise between DL and UL. It should be noted that Massive MIMO does not significantly increase the peak rate to a single user as it inherently needs multiple users to be served simultaneously to achieve the high spectral efficiency.

6.2 RAN TRANSMISSION AT CENTIMETER AND MILLIMETER WAVES

Mobile-communication networks have, until now, almost exclusively operated on frequencies below 3 GHz. However, extension into higher frequency bands, including frequencies above 10 GHz, is being considered for 5G. Frequencies of 3 GHz to 30 GHz are in the centimeter wavelength band, and frequencies of 30 GHz to 300 GHz are in the millimeter wavelength band.

The main benefit of frequencies above 10 GHz is the potential availability of large amount of spectrum and, perhaps even more important, large continuous spectrum chunks. The latter is needed to enable the very wide transmission bandwidths, such as several hundred MHz. Such transmission bandwidths are needed for efficient support of multi-Gbps data rates.

The main drawback of using higher frequencies is higher path loss. This can be partly compensated for by the use of more advanced antenna configurations, making use of the reduced size of the basic antenna elements at higher frequencies. Indeed, by keeping the size of the overall antenna configuration the same on both the transmitter and the receiver side, in combination with beam-forming, the overall path loss may actually be reduced as the frequency increases. This is one of the reasons for the current use of higher frequencies, in combination with highly directional antennas, for wireless backhaul.

However, this is true only for line-of-sight conditions. In non-line-of-sight conditions, which are the typical propagation situation in mobile communication, there are additional path-loss-degrading factors such as:

- Reduced diffraction, leading to higher path loss due to shadowed locations.
- Higher attenuation when propagating through walls, for example, leading to higher path loss in indoor locations covered by outdoor base stations.

Despite this, recent studies have shown that higher frequencies, up to at least 30 GHz, can be used for wireless access in non-line-of-sight conditions, assuming relatively short-range (100-200 meters) links. Even higher frequencies are being considered for ultra-dense deployments with even shorter access-node inter-site distance.

---

[11] T. Marzetta, “How much training is required for multiuser MIMO,” Proceedings of Asilomar, CA, Oct. 2006.

It seems likely that the use of higher frequencies will be one important component of 5G wireless access. However, higher frequencies can serve only as a complement to lower frequencies, providing high capacity and high data rates in dense urban environments. Lower frequencies should remain the backbone providing full wide-area coverage.

World Radio Conference (WRC) 2015 will focus on new spectrum mobile communication below roughly 6.5 GHz. Identification and assignment of spectrum above 10 GHz for mobile communication is expected to be on the agenda for WRC in 2018/2019. Thus, higher frequencies above 10 GHz may be available at the time of 5G initial deployments, which are expected around 2020.

6.3 NEW WAVEFORMS

The new waveforms in 5G include advanced multi-carrier transmission and non-orthogonal transmission.

6.3.1 ADVANCED MULTI-CARRIER TRANSMISSION

LTE radio access is based on OFDM transmission in both DL and UL: conventional OFDM for the DL and DFT-precoded OFDM for UL. OFDM transmission, which is a kind of multi-carrier transmission scheme, is also a candidate for 5G radio access. However, several other/modified multi-carrier transmission schemes are also under consideration for 5G radio access. These include (see e.g.,[12] for more details):

- Filter-Bank Multi-Carrier (FBMC) transmission
- Universal Filtered Multi-Carrier (UFMC) transmission
- Generalized Frequency-Division Multiplexing (GFDM)

Common for these transmission schemes is that they, at least in principle, can provide a more confined spectrum compared to conventional OFDM. This is relevant for spectrum-sharing scenarios. It should be noted that the confined spectrum is a property of the fundamental waveform and that transmitter nonlinearities may cause additional spectrum spreading that may reduce these waveforms' benefits.

The more confined spectrum is also assumed to make the above transmission schemes less reliant on-time synchronization to retain orthogonality between different transmissions. This may be valuable especially for UL transmission with requirements on very low access latency as the need for time-consuming synchronization procedures may be relaxed or even avoided.

6.3.2 NON-ORTHOGONAL TRANSMISSION

4G radio access is based on orthogonal transmission for both DL and UL. Orthogonal transmission avoids interference and leads to high system capacity. However, for rapid access of small payloads, the procedure to assign orthogonal resources to different users may require extensive signaling and lead to additional latency. Thus, support for non-orthogonal access, as a complement to orthogonal access, is

[12] "5G Waveform Candidate Selection", 5GNOW deliverable D3.1, http://www.5gnow.eu/node/52.

being considered for 5G. Examples include Non-Orthogonal Multiple Access (NOMA)[13] and Sparse-Code Multiple Access (SCMA).[14]

6.4 SHARED SPECTRUM ACCESS

The Federal Communications Commission (FCC) has been aggressive in its efforts to make new spectrum available for mobile communications. The FCC's 2010 National Broadband Plan concluded that 1.2 to 1.7 GHz of new spectrum is required to sustain capacity expansion required to meet anticipated growth in wireless data traffic. The FCC is pursuing shared spectrum access as a way to free up spectrum. Secondary users would be allowed to use the spectrum when the primary or incumbent is not using the spectrum in a given geography at a given time. The FCC has published a Notice of Proposed Rule Making (NPRM) for three-tier spectrum access in the 3.5 GHz band: incumbent, protected and general authorized access. A spectrum access server (SAS) manages the allocation of spectrum between these tiers.

Some new technologies are needed for mobile networks to use the shared spectrum:

- The RAN must be capable of interfacing with the SAS to request and receive spectrum allocations, and to provide the SAS with spectrum-sensing information from the base stations.

- The base stations must be spectrum agile and capable of spectrum sensing. Spectrum sensing means monitoring the frequency channel for use by the primary. With spectrum agility, users served in one channel are migrated to a new channel and then, to protect the incumbent, stop transmitting in the first channel when the SAS says to. This can be achieved with wideband adaptive radios. The migration should be designed in such a way that the users do not see an interruption when the switch to the new channel happens. That seamlessness can be accomplished through channel aggregation schemes so that when one of the channels is temporarily blocked, communication can continue on the remaining channels.

6.5 ADVANCED INTER-NODE COORDINATION

Interference between radio access nodes is the limiting factor of current wireless networks. In addition, the extreme densification of 5G radio nodes necessitates specific interference-avoidance solutions, both from the network and the device sides. One such solution is based on exchanging information between schedulers at the network side in order to avoid interference. This technique leverages current inter-cell interference coordination schemes (LTE Rel. 8) as well as Coordinated Multi-Point (CoMP) (LTE Rel. 11 and 12). However these solutions are currently not so effective because of a number of drawbacks. One drawback is the upgrades in backhaul transport that are required to support the necessary exchange of information with very small delays. Another significant drawback of CoMP-based solutions is inter-cluster interference: Cells are grouped into clusters for resources coordination, but the optimal cluster configuration that minimizes inter-cluster interference at a reasonable coordination complexity is still an open question.

---

[13] "Non-orthogonal Multiple Access (NOMA) for Cellular Future Radio Access", Vehicular Technology Conference (VTC) Spring), 2013.

[14] "Sparse-code multiple access", PIMRC, 2013.

Research in 5G is being conducted to overcome the aforementioned issues and facilitate inter-node coordination even with legacy backhaul networks. One such research area, originated in LTE Rel-12 and anticipated for Rel-13, deals with pre-compensating any foreseeable delay and jitter impairments of the backhaul. Other approaches try to relax inter-node coordination burden by allowing some degree of interference between nodes that must be handled at the receiver.

Centralization of radio processing functions reduces the signaling burden and would therefore be a driver for efficient inter-node coordination. However, the required capacity for the front haul network (transport network between the central baseband unit and the remote radio heads) would be one of the challenges when aggregating large numbers of base stations.

6.6 SIMULTANEOUS TRANSMISSION RECEPTION

Today's wireless systems dedicate spectral or temporal resources to UL and DL channels. Simultaneous transmission reception would enable sharing of the available resources between both directions of communication. Clearly, this would be a more efficient use of the available spectrum and in theory could double the current link capacity. This may seem a small gain compared to the 5G capacity needs and compared to what can be achieved through techniques like massive MIMO. However, the value of simultaneous transmission reception for 5G may not necessarily be in its capacity gains but in possible improvements in signaling and control layers. Removing the fundamental assumption of having separate UL and DL could allow 5G systems to be designed with a new approach and possibly facilitate the achievement of 5G goals in ways that may not be obvious now.[15]

It is generally not possible to only use the same channel for simultaneous transmission of UL and DL signals. The key challenge is the large power differential between the strong self-interference due to the device's own transmissions and the weak signal of interest coming from the distant transmitter. However, this has not deterred researchers from pursuing the goal of simultaneous transmission reception. Experimental demonstrations of simultaneous transmission reception for wireless communication systems have been reported since 1998[16] [17] [18] [19] [20] [21] [22] [23]. The work so far has been successful in reducing this interference by up to 85 dB[24], which has generally been achieved through a combination of analog cancellation, hardware cancellation and digital cancellation techniques. The interference-reduction levels achieved so far are sufficient for Wi-Fi-type systems over very short distances, where the received signal is strong enough to make the difference between the device's transmit power and received signal. These levels of interference reduction are clearly not sufficient for current cellular systems, which operate with

---

[15] Jung Il Choi, et al, "Beyond full duplex wireless", Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on, 4-7 Nov. 2012,pp 40 – 44.
[16] S. Chen, M. Beach, and J. McGeehan, "Division-free duplex for wireless applications," In IEEE Electronics Letters, 34, 147–148 (1998).
[17] D. W. Bliss, P. Parker, and A. R. Margetts, "Simultaneous Transmission and Reception for Improved Wireless Network Performance," In IEEE/SP 14th Workshop on Statistical Signal Processing, pp. 478–482 (2007).
[18] B. Radunovic, D. Gunawardena, P. Key, A. P. N. Singh, V. Balan, and G. Dejean, "Rethinking Indoor Wireless: Low power, Low Frequency, Full-duplex," Technical report, Microsoft Research (2009).
[19] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," In Proceedings of ACM Mobicom, pp. 1–12 (2010).
[20] M. Jain, J. Choi, T. M. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time full duplex wireless," In Proceedings of ACM Mobicom, pp. 301–312 (2011).
[21] M. A. Khojastepour, K. Sundaresan, S. Rangarajan, X. Zhang, and S. Barghi, "The case for antenna cancellation for scalable full-duplex wireless communications," In Proceedings of Hotnets, pp. 1–17 (2011).
[22] A. Sahai and G. Patel and A. Sabharwal, "Pushing the limits of Full-duplex: Design and Real-time implementation," In arXiv.org:1107.0607, (2011).
[23] A. Sahai, G. Patel, and A. Sabharwal, "Asynchronous Full-duplex Wireless," In Proceedings of COMSNETS, pp. 1–9 (2012).
[24] Melissa Duarte, "Full-duplex Wireless: Design, Implementation and Characterization," Ph.D. Thesis, Rice University, April 2012.

much higher transmit powers and in greater path-loss environments. However, the necessary interference reduction levels for a 5G system operating in the millimeter wave range may be more realizable due to lower transmit powers.

The current experimental success in achieving simultaneous transmission reception in Wi-Fi-like scenarios is encouraging. However, it is by no means close to practical implementation in a cellular environment. Further refinement of the interference cancellation techniques and possibly completely new approaches to interference cancellation are necessary to make simultaneous transmission reception a reality in practical systems. The interest in simultaneous transmission reception as a possible feature of 5G systems could drive further work in this area that would bring the current experimental achievements closer to practical implementation.

6.7 MULTI-RAT INTEGRATION AND MANAGEMENT

The ever-increasing number of RATs to be supported in a given deployment makes it crucial to consider multi-RAT integration and management issues. The objective is to facilitate uniform multi-RAT management and convergence among disparate technologies, both 3GPP and non-3GPP, such as Wi-Fi. Operation efficiency and user experience would be dramatically improved by automatically steering devices to the most suitable RAT in a seamless way. While multi-RAT management has been an important aspect in previous mobile generations, 5G's user-driven requirements foresee a seamless user experience when moving across networks. That makes multi-RAT integration more critical, particularly for services such as ultra-high-definition video or tactile Internet.

Given the likelihood of having multiple, heterogeneous wireless access points available in ultra-dense scenarios (e.g., 5G, LTE, 3G and Wi-Fi), some kind of decoupling between the user and control planes should be provided in order to separate the user payload from the necessary signaling. Multi-RAT integration may also consider simultaneous connection to multiple RATs in an opportunistic manner.

The expected impact on the network of such schemes is the introduction of a logical entity that coordinates resources among multiple RATs. To this end, the introduction of virtualization techniques may facilitate this point by enabling the instantiation of network functions upon demand, without having to change the network topology and/or architecture. Software-defined network functions can cope with different RATs by instantiating the necessary network functions upon demand, without the need to physically deploy additional network nodes for multi-RAT management.

6.8 DEVICE-TO-DEVICE COMMUNICATION

The introduction of direct Device-to-Device (D2D) communication for LTE began during 3GPP Release 12. At this stage, the considered D2D functionality was relatively limited, mainly focusing on D2D communication for public safety communication and D2D proximity detection for more general commercial applications.

However, as part of overall 5G discussions where technology components such as highly integrated backhaul/access (see Section 6.10) and more general multi-hop communication is being considered, D2D communication should definitely also be included as a possible technology component. One should then consider direct D2D communication as a more general tool that is a well-integrated part of the overall wireless-access solution. Besides what's depicted in Figure 1, this should include:

- The use of direct peer-to-peer D2D communication as an overall more efficient mode of transmission when nearby devices have end-user data to convey between each other.

- The use of direct D2D communication as a means to extend coverage beyond the reach of the conventional infra-structure (device-based relaying).

- Cooperative devices where high-speed inter-device communication provides means for "joint" transmission and/or reception between multiple devices, thus opening up for more efficient communication with the network-device communication. Note that this can be seen as a kind of coordinated transmission/reception ("CoMP") but on the device, rather than the network side.

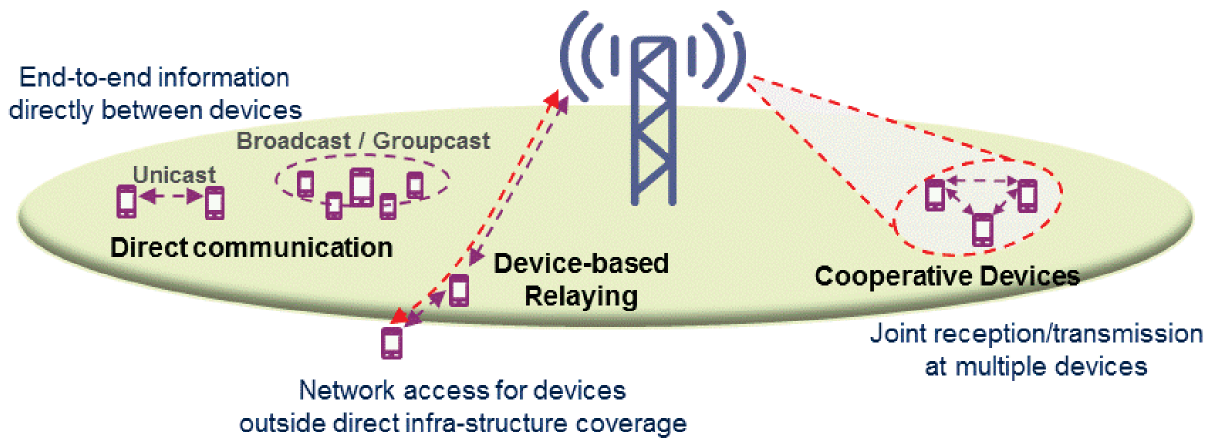

Figure 3. Device-to-Device Wireless Access Solution.

The key thing should be to see D2D communication as well-integrated part of the overall wireless access solution that is being taken into account already from the start of the 5G definition, rather than being a later-introduced "add-on."

D2D communication can be carried out in licensed spectrum under network control. D2D communication can also be carried out in, for example, unlicensed spectrum. Also in that case, an overlaid network, operating in the same or other spectrum, can be used to control/assist the D2D communication for enhanced efficiency and performance. However, D2D communication should also be possible in scenarios where there is no network coverage available, in which case the D2D link has to be possible to establish without network control/assistance.

6.9 EFFICIENT SMALL DATA TRANSMISSION

Small data bursts are most efficiently handled by connectionless access whereby a device wakes up and sends a short user plane data burst on a common 5G carrier using contention based access, eliminating dedicated radio resource control and higher layer signaling between the mobile and the network. Connectionless access reduces device power consumption and saves network resources when the size of the data burst is small compared to the overhead needed to set up bearers and establish security and device context for scheduled access. For 5G, the radio system must support connectionless contention based access multiplexed with scheduled access on a common carrier, with flexible allocation of resources that may vary along with traffic demand.

Smartphones and other devices capable of both scheduled and connectionless access may use connectionless access to send short data bursts and scheduled access for longer transactions. Example applications for connectionless access include the polling of servers (e.g., for email), sending instant messages (IM) and sending TCP keep-alives by applications that require long-term connectivity through Network Address Translations (NATs). In contrast, MTC sensor devices may only support connectionless user plane access; saving complexity and expense in the device in addition to further extend battery life.

The savings in network resource usage by these connectionless-only devices may enable a lower cost tier of service, expanding the range of applications for which mobile wireless access makes economic sense.

To support connectionless access, a device must attach or otherwise establish a context in the network that allows the receiving base station to validate the packet as belonging to a legitimate, authenticated device. A security context must also be established whereby the network can ensure user plane integrity and privacy for the transmitted burst. Public/private key infrastructure, or an extension of the 3GPP security architecture, with the context information retrieval keyed from the user plane packet header may be used for this purpose.

Packet processing for connectionless access may be modeled on OpenFlow Software Defined Networking (SDN) procedures that handle the processing of unidentified packets arriving at a switch. For example, the following steps may be taken after the device attaches to the network, is authenticated, and establishes with a controller a security context and a device context indicating support for connectionless access.

1. The device (at a future time) selects a base station and sends a user plane data burst.
2. The base station sends the packet or packet header to the controller for validation, much as an OpenFlow switch would send a packet for which it had no flow table entry to an OpenFlow controller.
3. The controller authorizes access and sends security context information to the base station along with forwarding instructions, much as an OpenFlow controller would provide a flow table entry to a switch.
4. The base station deciphers the packet contents and forwards the packet according to the received instructions.

After a data burst has been sent, the device may wait for a response from the network and transition to a connected state as warranted by further transmissions.

6.10 WIRELESS BACKHAUL/ACCESS INTEGRATION

In many places around the world, wireless backhaul constitutes a major part of cellular's overall backhaul portfolio, especially in less populated areas, where other backhaul solutions are not economically and/or technically feasible. Wireless backhaul is typically based on a proprietary radio-link technology operating under line-of-sight propagation conditions. In the early days of mobile communication, wireless backhaul often used frequencies in the 2 GHz range and even lower. Today, backhaul is being concentrated in spectrum above 6 GHz, including the millimeter band (above 30 GHz). A few years ago, wireless backhaul started to be deployed in the 70-80 GHz spectrum (E-band), where the backhaul can transport up to 5 Gbps. With the increased interest in small cells, including those deployed indoors, there is an increased interest in wireless backhaul solutions that can operate under non-line-of-sight conditions. There is still an assumption of a backhaul-specific radio-access technology operating in spectrum separate from the spectrum used for the access (Base Station (BS)-UE) link. The backhaul link may, in these cases, operate in millimeter bands. Lower frequencies also are being discussed as possibilities for small cell backhaul.

There is no fundamental reason why the basic radio-access technology needs to be different between the (wireless) backhaul link and the access link. This is especially the case taking into account that:

- The use of higher frequencies is being considered also for the access link (see Section 6.2).
- Lower frequency bands are being considered for wireless backhaul, as discussed above.
- Flexibility to use spectrum for both mobile and backhaul.

The radio conditions for a backhaul link to a small “under-roof-top” base station are more similar to a BS-UE link than to a conventional macro-cell backhaul link.

In the case of *integrated backhaul/access*, one sees the access and backhaul link as just two links in an overall unified wireless access solution:

- The same (set of) radio-access technology(ies) can be used on the access (BS-UE) link and the backhaul link.
- The backhaul and access link relies on the same spectrum pool. Note that this does not necessarily mean that the same carrier frequency is used on the two links at the same time and in the same location.
- Resource management, QoS optimization and other tasks are carried out jointly for the backhaul and access links.

The aim of access/backhaul integration is to achieve more efficient use of technology and spectrum and to improve the overall performance/quality of the overall end-to-end link. It should be noted that, from a radio-link point-of-view, integrated backhaul/access can be seen as a kind of multi-hop communication.

6.11 FLEXIBLE NETWORKS

Elastic computing and storage in data centers has ushered in a new need for highly dynamic networking that is being addressed by SDN. The benefits of SDN in data center networking, both within and between data centers, have been firmly established. Extending beyond the data center, creation of on-demand virtual private networks with dynamically configurable capacity in the optical links between enterprise premises and the data center is also being explored.

It should be noted that the narrow view of SDN includes only OpenFlow switches and OpenFlow controllers that connect to the OpenFlow switches to centrally manage packet forwarding, with management support from an OpenFlow Config. However, the broader view of SDN is one where the emphasis is on network programmability through open northbound interfaces, a configurable policy framework, resource discovery and optimization and an SDN control framework that is separate from the forwarding plane.[25] The SDN controller may exercise control over the network elements not only through OpenFlow but also through other interfaces such network management interfaces. This report takes this broader view of SDN when applying it to the domain of wireless networks.

The application of SDN in 5G may extend beyond control of transport resources to include the wireless network’s policy framework. Examples include how functions such as QoS management and application of forwarding plane functions via service chaining are applied in an SDN-enabled packet core. More

[25] http://www.opendaylight.org/project/faq#1.

broadly, an SDN framework can also be applied to wireless control functions where, for example, mobility management, security, charging and optimization become applications subtended by an SDN controller.

NFV virtualizes network applications and network functions such as those provided today by dedicated core network components, media servers and management functions. Virtualization enables these functions to run on hypervisors using shared, off-the-shelf data center computing infrastructure. NFV is already employed in 4G implementations and will be commonplace well before 5G. Going further, NFV can be exploited as technology fundamental to the design of 5G, where it is expected that the majority of the wireless network functions in the RAN and the core to be virtualized, thus increasing flexibility to meet varied demands on control and data processing from a large variety of applications. Interconnection amongst these functions and programming of the forwarding plane to support mobility is most flexibly achieved using SDN techniques.

Closely related and complementary to NFV is orchestration technology that enables automated deployment and management of network functions and network services. The orchestration framework is key to realizing the cost savings of moving to virtualized implementations. Latency and networking requirements may necessitate placement of virtualized functions in close proximity to each other or at local data centers close to the subscriber. In some cases, network functions key to achieving very low latency may be better implemented without NFV. The MANO work group within ETSI NFV forum has defined an orchestration framework that can be applied to 4G and more deeply integrated into the 5G network architecture.

NFV and the major principles of SDN, such as the separation of control plane and data plane, network abstraction and programmability of the network by external applications all should be broadly applicable to wireless networks. This potentially enables new services, reducing the time and effort needed to implement those services and reducing network costs.

The ability to rapidly establish network resources based on application demands and operator policy has many benefits. Rather than manually configuring fixed resources on dedicated hardware, which may not match dynamically changing application demand, operators can shift bearer plane assets as needed and allocate to users only those resources needed to satisfy device and application requirements. This enables network operators to better monetize their network assets and quickly and more easily roll out new end-user services. In addition, CAPEX savings comes from more efficient use of network assets and the use of simpler, lower cost bearer plane networking infrastructure. OPEX savings is obtained from auto-configuration and simpler network configurations without wireless-specialized forwarding plane elements.

As distributed network functions running on dedicated hardware are phased out in favor of NFV, it is expected SDN for wireless networking to facilitate meeting data center networking requirements for the virtualized wireless RAN, core network and applications. In the data center, SDN is needed to provide dynamic configuration and scalability of network resources as wireless network functions are instantiated, VMs are reassigned, traffic grows and new radio infrastructure is deployed outside of the data center. For wireless access, SDN programmability can be used to simplify access selection and packet routing, with forwarding rules directing packet flows both in the mobile and base stations. This real-time programmability realizes the goals of context-aware networking, so intelligence information is actually applied in the selection of packet routes, particularly when multiple air interface options are available.

6.12 FLEXIBLE MOBILITY

Flexible mobility for 5G embraces a selection of options, which may be dynamically assigned to a device or application according to the device and application context, or statically configured for specialized devices and applications. Flexible mobility consists of two components: one for managing mobility of active devices and a second for tracking and reaching devices that support a power-saving idle mode. Assigned mobility may range at one extreme, beginning with no "active mode" mobility, with no support for idle mode as is typical of Wi-Fi access today. The other extreme is full support for active and idle mode mobility similar to mobile 3G/4G. Gradations of flexible mobility bridge the gap between these extremes, allowing for independent assignment of idle-mode mobility on a per-device basis, and active mode mobility on a per-application basis as shown in the examples in Table 3.

Table 3. Examples of Flexible Assignable Mobility for 5G.

Idle Mode Mobility	Active Mode Mobility	Application	Current Tech
Yes	Yes	Smartphones, Tables for mobile communication	3G/4G
None	None	Always-on, battery in-sensitive	Wi-Fi
None with sleep/ coma mode in device	None	When the device need not be reachable from the network after idle transition and only the device initiates transactions (e.g., MTC long battery life)	--
Yes	None	For nomadic, long-battery life devices that must be reachable from the network	--
None	Yes	Always-on, battery insensitive with seamless continuity when active	--
None with sleep/ coma mode in device	Yes	When the device need not be reachable from the network after idle transition with seamless continuity when active	--

Here idle mobility involves the tracking of the device location while in "sleep mode." Active mobility refers to establishment of IP anchors, where tunnel endpoints are updated when the UE moves between base stations so that seamless IP session continuity may be maintained.

Offering a range of mobility options not available today in 4G enables a better match between the needs of the device and application and network resources. This has advantages for both the network and user/device. Specifically:

1. "**No idle mode mobility**" means that context and state information for tracking the device need not be stored in the network, saving resources when applied to a plethora of MTC, sensor-type devices. There are two cases for no idle mode mobility: The device may not support idle mode and be always-on, similar to Wi-Fi, or it may support a sleep/coma mode whereby the device de-allocates Tx and Rx resources, hibernates until it wishes to initiated a transaction and is not reachable from the network until that point. The latter case is particularly useful for a sensor-like device where battery life is of paramount importance and network-initiated communication is not required.

2. "**No active mode mobility**" means that the network need not establish and maintain user plane tunnels and store related state information. IP addresses are allocated local to the base station, allowing for more efficient routing of locally available content. The user/device benefits from lower latency due to more direct routing when compared to active mobility with a centralized anchor. This mode is most appropriate for stationary and nomadic devices.

With active mode mobility, the network may flexibly assign an IP anchor to one of several network elements, such as a base station router, aggregation router, edge router or, in the case of "no active mobility," a local prefix is assigned with no anchor. A device such as a smartphone may have multiple active mode mobility contexts for different applications with, for example, a central anchor point assigned for a mobile communication application, and no anchor assigned for HTTP Internet access, where IP session continuity is not required and content can be provided locally.

Flexible mobility is enabled by two critical technologies. The first is context awareness. The network needs context information about the application's needs and the device's capabilities to determine the appropriate level of active and idle mode mobility to assign. The second technology is SDN control of the transport path. Flexible active mode mobility is possible only if mobile anchors can be dynamically assigned at a central point, close to the device, in between or not at all at the time an application is invoked. This requires a context-aware controller that can program transport elements to establish tunnels where needed and forward traffic accordingly.

6.13 CONTEXT-AWARE NETWORKING

The network cannot provide resources tailored to serve a wide range of devices and applications without context information that goes significantly beyond that available in 4G. Context awareness allows the network to adapt to the needs of applications within the framework of network constraints and operator policy. This is preferable to the alternative, where applications adapt to the constraints of a one-size-fits-all set of service characteristics on a default bearer as is typical in 4G.

In addition, the ability of both the network and device to use context awareness (e.g., location, historical usage pattern, subscriber preferences) can help further enhance the user experience. This ability also enables the concept of the Internet coming to the user with the most relevant and timely information rather than the user having to go to the Internet to retrieve information and then to filter out the irrelevant pieces of information. Context awareness includes awareness of the following:

1. **Network Analytics**, including alternative RATs, network layers (macro cell, mmWave, small cell, Wi-Fi) and corresponding congestion levels, capabilities and performance characteristics.

2. **Subscriber Analytics**, including subscription attributes, wireless activity level, loyalty management status, experience analytics, historical subscriber activity, location history, current location, subscriber contacts, location context (e.g., work, home, mall) and application usage.

3. **Device Attributes and Capabilities**, including information on single function vs. multi-function devices, device support for specialized applications, MTC vs. subscriber devices and radio and network optimization capabilities.

4. **Application Requirements**, including QoS requirements (e.g., delay, throughput, latency), connection reliability, access price, power consumption and security level.

5. **Subscriber Preferences**, including preferred access options, power savings vs. performance and access cost.

6. **Operator Policies and Subscription Context**, including allowed services, service attributes and QoS.

Context information may be gathered from the device, network monitors, network elements, network data bases and analytics platforms. It is processed by the network when a device attaches or an application is invoked and results in a determination of service attributes that govern how the device and application will be treated by the network. The service attributes for access may for example include cost, reliability, power consumption, security level, QoS and mobility. The service attributes for access may be mapped to configurable 5G features, which are then assigned by the network.

For example, context information may determine that low-cost access, with no support for active mobility and long battery life, is best for providing service to a nomadic sensor device that attaches to the network. The network as a result configures connectionless access with low priority, simple IP networking with no tunneling and an idle-mode wake-up period of one day.

6.14 INFORMATION CENTRIC NETWORKING (ICN)

As discussed in Section 4.3.2, 5G should be based on new network architectures and protocols designed specifically with support for mobility, security and content caching as fundamental design criteria. Information Centric Networking (ICN) – for example, as realized in the Named-Data Networking (NDN)[26] and Content-Centric Networking (CCNx)[27] programs – is emerging as a leading architecture that can meet such design criteria. ICN approaches are focused on the support of future Internet evolution and, in particular, support of new communication models that focus on the distribution of information rather than the communication of data packets between endpoints. This section provides details of ICN including architecture aspects as well as a discussion of potential use cases and benefits of ICN.

6.14.1 INTRODUCTION

A major objective of 5G network is to reshape the economics of mobile networking to efficiently serve new emerging use cases and markets (e.g., IoT) as well as drive support for current usage models (especially video-related) to lower cost points. This challenge extends beyond the mobile network and the following observations can be made regarding the contemporary network:

1. **There is a mismatch between Internet usage and architecture:** Usage has significantly evolved over the past decade and today is mostly centered on information dissemination and retrieval. Users search for information over Google, watch videos on YouTube and share information on Facebook. The information-centric nature of current Internet usage has turned the network into a medium to connect people with information or content, a distribution network conceptually different from the communication network defined by the host-to-host architecture principles which led to the design of IP. Some Use Cases for sensors and machines are expected to continue to drive this model for information sharing. There has been exponential growth of digital information diffused over the Internet, eased by cheaper storage and bandwidth supports, and driven by the increasing popularity of highly demanding services, such as cloud computing or video delivery. Nevertheless, the Internet architecture (and by extension, that of the mobile network) is still based on the end-to-end, connection-oriented model that can be improved in the future to more economically address emerging trends.

2. **Overlay Solution Model:** The approach to address the design-usage mismatch has been to overlay purpose-directed solutions on the existing infrastructure. A large range of on-top solutions, such as Content Delivery Networks (CDNs) and peer-to-peer networks, have been designed and widely deployed to overcome this mismatch at application layer and today carry a large fraction of Internet traffic. The reliance on overlay solutions has progressively increased

[26] http://named-data.net/project/archoverview/
[27] http://ccnx.org/what-is-ccn/

complexity in the network ecosystem created new players (e.g., CDN providers, ISPs and content providers) and rendered overall network management a difficult task from both the operational and business points of view. Technical inefficiencies also emerge as consequences of the overlay model, for example dynamic content-to-location binding, mobility management, especially in presence of low-latency services, multicast, multi-homing, etc. In principle, overlay solutions are most effective when the service provided by the overlay is narrowly directed to a specific usage model or its decoupling is required to facilitate parallel evolution. The cost derives from the need to manage a separate and distinct layer (which is one reason why transport over wired portion of the mobile network infrastructure is more costly than fixed networks).

3. **A sustainable solution:** The expected traffic growth in the mobile segment, fostered by broadband access penetration and powerful mobile terminals, motivates innovation at the network layer to effectively support this increased traffic demand. Information Centric Networking (ICN) and its implementations, Named-Data Networking (NDN) and Content-Centric Networking (CCN), identify a new networking paradigm that leverages content awareness at the network layer to simplify the network architecture. Network operations are driven by content names, rather than location identifiers (IP addresses) to gracefully enable user-to-content communication.

   Embedding the content name into network primitives allows for a more agile connectionless transport model driven by the end-user that is not bound to a network addressable interface. This simplifies the management of multi-homed communication (mobile and fixed) and provides autonomous control of forwarding based on factors such as congestion and local cache states. Due to content-awareness carried by content names, the network is able to route requests toward nearest content replicas, exploiting temporary in-network caches and adaptive request routing for a more efficient and cost-effective data delivery. Finally, the symmetric-routing attribute endows the networking layer with natural mobility support. The inclusion of support for mobility, security and storage in the networking layer leads to a simpler, cost-efficient architecture that intrinsically supports modern communication usage patterns.

The more prominent characteristics of the ICN architecture are discussed below. Note that in the context of this section, ICN refers to the related architectural approaches defined by the NDN and CCNx implementations.

6.14.2 ICN ARCHITECTURE

6.14.2.1 NAMING DATA

The basic idea of ICN is to enrich network-layer functions with content awareness so that routing, forwarding, caching and data transfer operations are performed on topology-independent content names rather than on IP addresses. Data are divided into a sequence of chunks uniquely identified by a name and permanently stored in one or more servers. Naming data chunks allows ICN network to directly interpret and treat content according to its semantic with no need for DPI (Deep Packet Inspection) or delegation to the application layer.

The naming convention does not need to be specified and can be application-specific: only a hierarchical structure, similar to that already adopted by HTTP, is required for entries aggregation in name-based routing tables.

The hierarchical naming scheme uses hierarchically ordered labels (e.g., URIs). More precisely, a name is composed by a variable number of components (not necessary human-readable as URIs), organized in a hierarchical structure.

Authenticity: One of the most important differences between name-centric and traditional host-centric networking is that data are retrieved by name rather than location. Hence, in ICN architectures the data authenticity verification (i.e., the verification of the publisher of a named data object) is an important challenge. Data authenticity is achieved by applying a digital signature (hash of name plus data-object through publisher's key) to named data object with hierarchical naming scheme.

Update and versioning: In ICN, routable object names are globally unique. Hence, updating an object or creating a new version of an object corresponds to the creation of a new object. With the hierarchical naming schema, a component of the data-object's name can be considered as its version.

Name Encoding: As previously mentioned, ICN names can be potentially unbounded. An important challenge is to define an efficient name-encoding scheme in order to achieve: i) fast name parsing; ii) reduction of the space needed for carrying the name in ICN packets; and, iii) flexibility.

Encoding proposals satisfying these requirements exist and leverage type/length/value encoding with component offset encoding that is highly-flexible, compact and faster to parse.[28]

6.14.2.2 MOBILITY SUPPORT

In ICN mobility is managed in a very different way as interfaces do not have network addresses so a change in physical location does not imply a change of address in the data plane. One sees support for mobility emerging naturally from the architecture, especially in the case of consumer mobility. The transport model is connectionless and pull-based, hence mobility does not affect the communication. The consumer sends Interest packets which are routed in the network over one or more paths and the Data flows back towards the client following the paths built by the Interests. In a case of a 'move-before-get' the consumer simply re-expresses the Interests for the Data not yet received and the network is able to fetch it from local caches. Various strategies exist to support producer mobility and real-time group communication leveraging the multi-path forwarding capability intrinsic to ICN and dynamic updating of local forwarding tables without requiring explicit signaling.[29] The distributed in-network caching of ICN facilitates smooth handoffs and prevents service quality degradation during mobility-driven connectivity transitions.

6.14.2.3 DATA-CENTRIC SECURITY

Current Internet security is made available by means of ad-hoc protocol extensions such as DNSsec, IPsec and TLS. TLS provides web security by encrypting a layer 4 connection between two hosts. Authenticity is provided by the web of trust (certification authorities and a public key infrastructure) to authenticate the web server and symmetric cypher on the two end points based on a negotiated key. In presence of TLS many networking operations become unfeasible (e.g., filtering, caching, and acceleration).

The security model of ICN is radically different. Instead of securing by encrypting simply connections, the ICN object-security model allows the separation of security actions regarding privacy, data integrity and data confidentiality all of which leverage an existing web of trust based on certification authorities and a

[28] M. Mosko, I. Solis, *CCNx Messages in TLV Format, IETF Internet Draft (Status Experimental),* https://tools.ietf.org/html/draft-irtf-icnrg-ccnxmessages-00

[29] J Auge, G Carofiglio, G Grassi, L Muscariello, G Pau, X Zeng, *Anchorless Producer Mobility in ICN*, 2nd ACM Conference on ICN (ICN 2015), Poster Session http://conferences2.sigcomm.org/acm-icn/2015/program.php

public key infrastructure. The security actions performed by the producer and consumer are directly at network layer with content identification provided in data names. All data is integrity protected while confidentiality (via data encryption) is optional. Integrity protection guarantees the authenticity of the data bound to the name by including the producer signature of the data plus its name.

The atomic security service provided by ICN guarantees that the producer has published a piece of data with the name available in the packet. Naming semantics are determined by the application and may (or may not) be designed to reveal the identity of the payload. This service enables location-independent secured content access. Denial of service attacks based on cache poisoning can be blocked using signature verification techniques. However, the cost is not negligible and recent work has started to build network layer trust management that does not require in-network signature verification by using the concept of interest-key binding.

6.14.2.4 NAME-BASED ROUTING AND FORWARDING

ICN network routers process user requests (Interests) by name in a hop-by-hop fashion towards a permanent copy of the content. To this goal, every router has a name-based routing table storing one or more potential next hops towards a set of content items. Dynamic forwarding algorithms select a given next hop according to specific metrics (e.g., time-monitored delay), and with objective of achieving optimal throughput while minimizing network cost.[30] ICN routers also keep track of received Interests in order to return content chunks to the user following the reverse request path (symmetric routing).

The content delivery process (illustrated in Figure 4) is driven by three basic communication mechanisms: name-based request routing, pull-based connectionless transport and in-network caching. The ICN forwarding process is described below.

6.14.2.4.1 ICN ROUTING OPERATIONS

Upon reception of a Request packet from an input interface, intermediate nodes perform the following operations:

1. Cache lookup, to check if the requested Data chunk is locally stored. In case of cache hit, the Data is sent through the interface the request is coming from. Otherwise,

2. Pending Interest Table (PIT) lookup, to verify the existence of a pending Request for the same Data chunk. If yes, the Interest is discarded since a pending query is already outstanding. If not, a new PIT entry is created and a

3. Forwarding Information Base (FIB) lookup via Longest Prefix Matching returns the interface where to forward the Interest (selected among the possible ones).

FIB entries are associated to name prefixes. Data may come from a server, or from any intermediate cache along the path with a temporary copy of the Data packet. Forwarding operations are illustrated in the following figure[31]:

---

[30] G Carofiglio, M Gallo, L Muscariello, *Joint Hop-by-hop and Receiver-Driven Interest Control Protocol for Content-Centric Networks*, ACM SIGCOMM Workshop on information Centric Networking (ICN), Helsinki, Finland, August 2012.
[31] *Named Data Networking (NDN) Project 2011-2012 Annual Report*

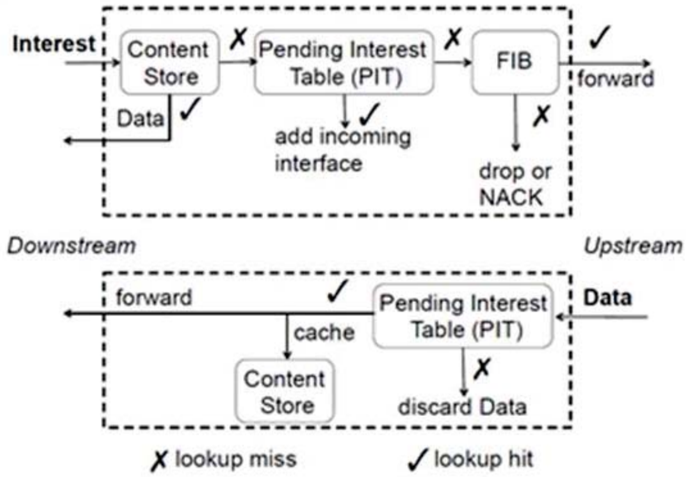

Figure 4. ICN Forwarding Engine.

6.14.2.5 PULL-BASED CONNECTIONLESS TRANSPORT

Differently from current sender-based TCP/IP model, ICN data transfer process is triggered by user requests addressed to chunks of the requested content item (i.e., pull based model). Rate and congestion control is performed at the end user by means of a connectionless, yet stateful transport protocol with the following characteristics:

- No connection instantiation, support for user/content mobility;
- Support for retrieval from multiple sources, a priori unknown at the user (e.g., intermediate caches);
- Support for multipath communication (to improve user performance and traffic load balancing).

6.14.2.6 IN-NETWORK STORAGE

Network nodes temporarily store content items in order to serve future requests for the same content.

Whenever an Interest is received at an ICN node, it first checks if the requested chunk is present in the local cache. If this is the case, the content is returned back to the user. Otherwise, the request is forwarded to the next hop by the ICN request routing.

It is worth noting that the labeling of packets by name in ICN enables routers to temporarily store and locally retrieve data packets for two new purposes:

- Reuse: subsequent requests for the same Data can be served locally with no need to fetch data from the original server/repository.
- Repair: packet losses can be recovered in the network, with no need for the sender to identify and retransmit the lost packet.

---

http://named-data.net/wp-content/uploads/2013/08/ndn-proj-pub.pdf.

- Simple cache management policies and coordination techniques allow an efficient allocation of distributed in-network storage resources at very low computational overhead and without requiring the complex management of today's CDN.

As illustrated in the following figure, the presence of distributed in-network storage and of name-based lookup permits to automatically move copies of popular content close to the users.

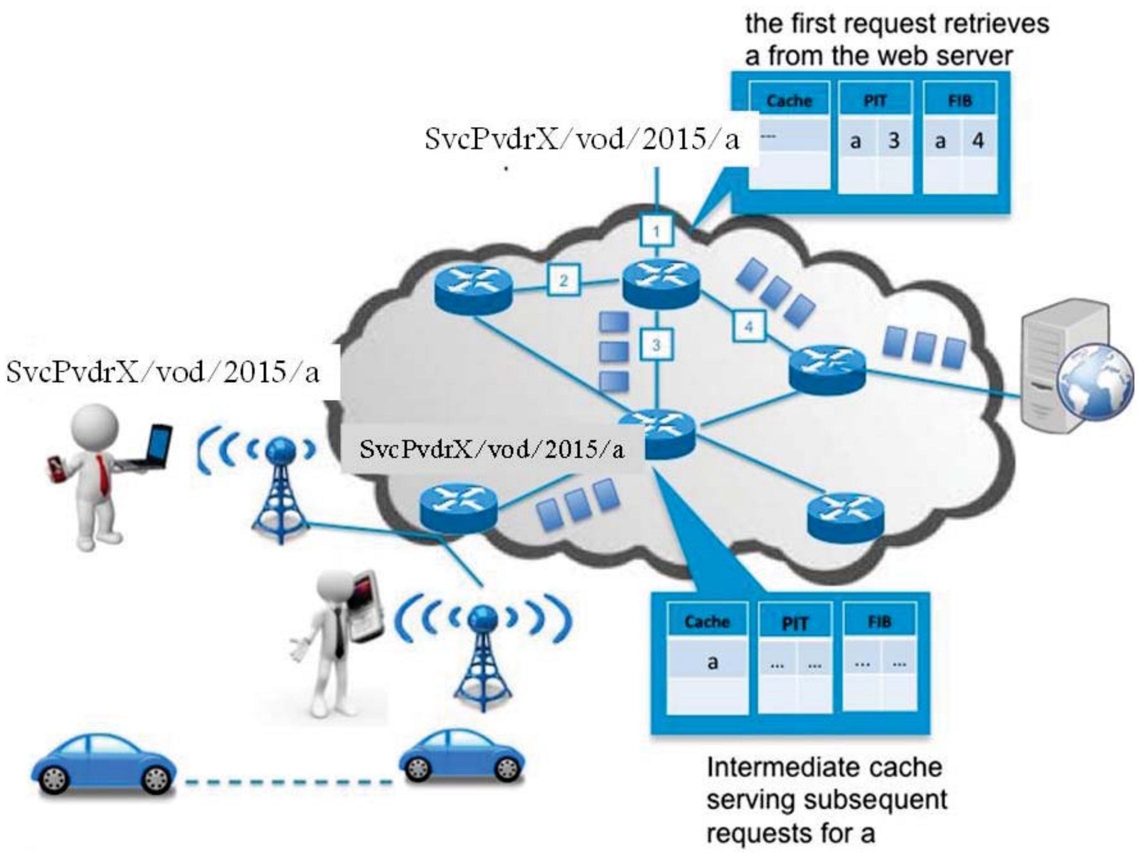

Figure 5. ICN Data Retrieval.

6.14.3 EXAMPLES OF ICN BENEFITS AND USE-CASES

6.14.3.1 SIMPLIFIED AND ENHANCED NETWORK SUBSTRATE

Moving the barycenter of the communication from location to the information itself brings significant advantages beyond the enhanced delivery performance.

First, it leads to a network architecture that is greatly simplified: application naming directly translates into name-based network operations, so it eliminates the name-to-location indirection that currently needs to be resolved by DNS. As a consequence, all recognized inefficiencies due to misuse of DNS for content routing would be overcome (dynamic content-to-location binding, coarse granularity in location mapping, poor reactivity due to entries' time-to-live).

Second, content awareness provides ISPs with a fine-grained knowledge of the carried traffic and of its requirements, which in turn may permit a more optimization of resources allocation than in the overlay model.

In addition, the presence of a unique caching infrastructure that all services can leverage, coupled with an agile connectionless transport model, leads to a reduction of transport cost (bandwidth/energy savings).

Finally, the data-centric security model disentangles integrity, confidentiality and trust issues from the transport model and enforces object security independent from the underlying network layer.

6.14.3.2 EXAMPLES OF APPLICATION AREAS

ICN architectural principles are designed to build an agile network substrate that can accommodate future innovation either at the network and application layers in a scalable manner. While ICN design addresses fundamental aspects of the network architecture, the benefits of its use can be illustrated by the following use cases:

Content delivery: The natural application of ICN is content dissemination and delivery, especially mobile content delivery, where network edge caching can considerably enhance delivery performance. According to user requests and hence content popularity over time, the content moves in time and in space due to the symmetric routing and in-network caching principles. Unlike traditional CDNs where request routing requires centralized orchestration for load-balancing purposes and where content placement requires complex optimization tools, the distribution of content in caches close to the edge is automatically triggered by user requests in ICN. Request load balancing and multipath forwarding can be achieved by concurrent use of multiple network interfaces in parallel at the receiver (e.g., in the context of multi-homing) and by hop-by-hop forwarding performed at ICN routers. This results in important latency and bandwidth reduction in addition to simplifying network management.

Sensor Networks/IoT: The current host-centric TCP/IP model appears not suitable for large distributed and low-power devices: a richer naming, for data, applications and devices is needed to simplify the communication model. Also, the communication is naturally one-to-many or many-to-one device over a largely distributed network. ICN is considered a promising network-layer candidate for its namings, but additionally for its light transport model and highly distributed forwarding.

Disaster Scenarios and Network Resilience: Unpredictable natural disaster or human-caused disasters like a terrorist attacks have shown that the current network infrastructure is vulnerable and mobile communications play a fundamental role in recovering such failures in the infrastructure.

The challenges faced in the aftermath of a disaster are the automatic reconfiguration of fragmented networks, an accrued congestion of the available network resources and an authentication to be managed in a decentralized manner. ICN has recently appeared as a promising candidate due to its connectionless and receiver-driven transport which does not require a connection setup, nor an end-to-end routing information. Interest can be routed by adopting an hop-by-hop forwarding strategy based upon incomplete network knowledge and satisfied by distributed in-network storage capabilities. The security model associated to objects rather than channels also result very appealing in such scenarios.

Smart Cities (healthcare, energy distribution, vehicular networks, etc.): The capability of ICN to cope with a dynamic, mobile and often infrastructure-less network make it suitable for supporting the development of new applications meant to facilitate our daily life: e-health, intelligent transportation system, smart energy distribution systems are emerging in the broader context of smart cities. The challenges ICN may effectively relieve are here those of scalable distribution of information in a mobile and failure-prone environment, again the data-centric security model and the service-aware Quality of Experience that can be tuned and enforced in the network by leveraging content names. The latter aspect would be critical for accommodating services with widely different requirements on top of the same distribution and caching network infrastructure.

6.15 MOVING NETWORKS

Mobility support in 5G will likely extend to very high speeds such as 350 km/h and beyond and even aircraft communications, although this possibility still remains unclear. However, both the device and cell can be moving like in D2D, V2V or V2I scenarios. Ultimately, the concept of a cell becomes blurred in favor of a more general concept of connectivity, where the network follows the movement of the user rather than the opposite (as usually conceived in previous mobile generations).

The management of nomadic and moving cells presents a number of issues – such as activation/deactivation of cells, trajectory prediction and handover optimization – because users will rapidly traverse multiple cells in a very short time. Additionally, the Doppler shift caused by very high relative movement between transmitter and receiver can challenge the use of millimeter waves.

7 SPECTRUM

Traffic growth continues to increase at exponential rates. The old adage that “if you build it, they will come”, does not hold true. Users are coming faster than the infrastructure can be built. The trend that started in the 1990s of merging IT and cellular systems continues. Data services are expected anywhere, at any time and at an ever-faster rate. Low latency is not a luxury, but a basic requirement.

The industry has estimated capacity resources to start becoming saturated around the 2018 timeframe. These estimates are based on existing technology and do not take into account incremental steps of air interface capacity improvements. When improvements such as beamforming, load balancing, channel efficiency and CA are taken into account, estimates show that perhaps it will be 2020 when new technologies will make their commercial debuts.

Besides just technology advances and system architecture evolution, it is clear that additional spectrum allocated for mobile broadband will be required to meet the projected demand[32]. The industry recognizes that new spectrum below 6 GHz will be difficult to obtain and that spectrum above 6 GHz may not have the desired propagation characteristics for wide-area coverage, although in many cases, it will be suitable for high-density system deployments. The simple truth is that future networks must have the ability to utilize the entire range of spectrum ranging from below 1 GHz up to 6 GHz and well into the millimeter wave ranges efficiently and seamlessly.

Mobile operators reacting to the demand of the marketplace, advancement of enabling technologies and convergence of communication, information and entertainment are offering an increasingly vast range of personal, business, public service, wide area data, location and MTC across a wide variety of environments.

In order to meet the challenges, both the application of advanced technology and use of appropriate higher spectrum bands with larger channel bandwidth are anticipated from a data rate perspective to provide support for the large factor increase projected for 5G system traffic. From a capacity viewpoint, the use of higher spectrum bands may also support new architecture models promoting increased effectiveness of the systems.

---

[32] See Report ITU-R M.2290-0 (12/2013) “Future spectrum requirements estimate for terrestrial IMT”

As noted previously, spectrum above 6 GHz (principally in the multi-GHz ranges) is already being investigated for its technical suitability for mobile broadband deployments to support 5G. There is also consideration of use of these bands at national levels and/or at the World Radio Conference level (targeting WRC-19).

One-size-fits-all air interfaces, which have been the typical generational solutions of the past twenty plus years, may no longer be the total solution. Although subsequent generations of wireless may indeed use new air interfaces, the industry fully expects that future technology will depend more on heterogeneous network deployments utilizing multiple air interfaces in order to meet throughput, coverage and capacity needs. Additionally, the concepts of spectrum sharing and unlicensed operations must be part of any 5G vision. Wi-Fi will continue to grow in importance.

Spectrum is already highly fragmented and is not likely to change in the future. Roaming, a challenge once thought solved, is becoming a resurging issue with all of these new spectrum bands. The concept of a 50-band mobile device is needed from a global ecosystem viewpoint. But in practical terms, it is not possible. Hence, the continued need to attain globally harmonized frequency bands in future spectrum allocations is still an urgent and pressing quest.

Furthermore, the ability to deliver the best QoS and utilize the best radio band available consistent with the QoS requirement is a key feature of mobile equipment of the future. It may be required in future equipment (and hence, supported by system specifications) to operate in reasonable and harmonized ways not only on dedicated spectrum bands, but also to incorporate operation under licensed assisted access schemes, shared spectrum arrangements or other concepts.

8 CONCLUSIONS AND RECOMMENDATIONS

An end-to-end 5G system has to be architected to meet the expected demand in 2020 and beyond. Figure 2 illustrates a comprehensive view that must be considered in the initial planning process for 5G.

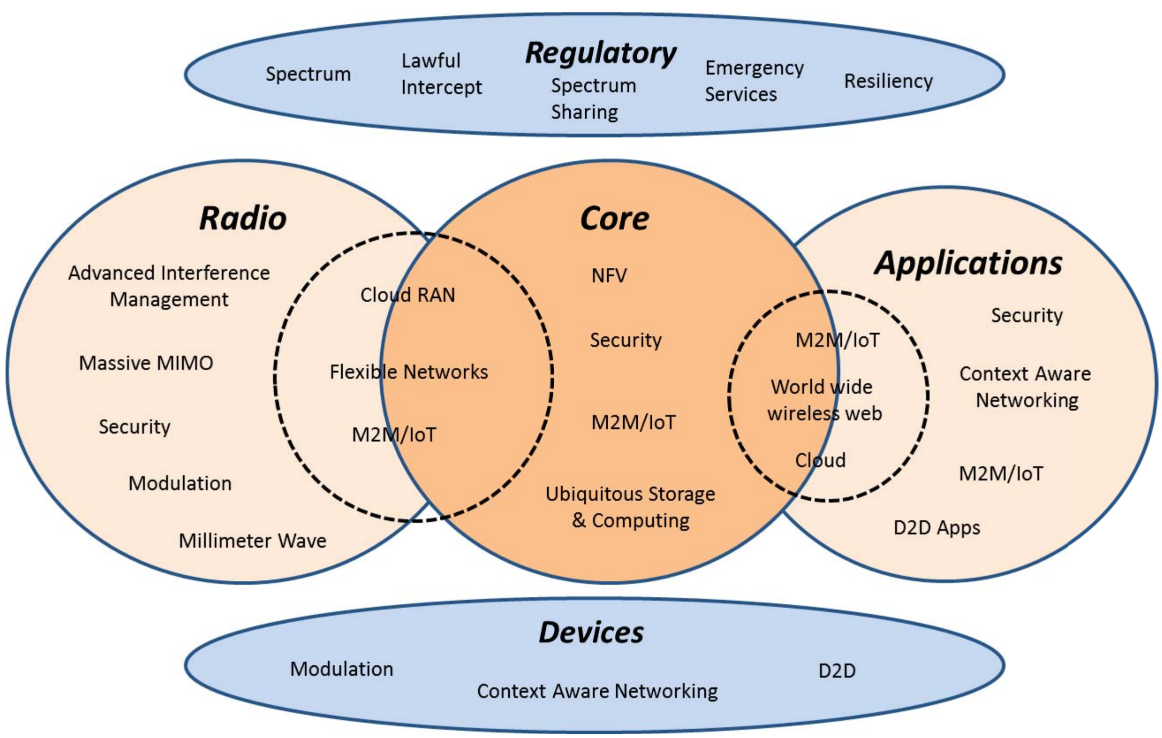

Figure 6. A Preliminary View of an End-to-End 5G Ecosystem.

The key 5G principles highlighted in this white paper are:

- As 5G is defined and requirements developed, it must include the entire 5G ecosystem (e.g., air interface, devices, transport and packet core).
- 5G development should provide global harmonization under a single framework and allow time for true advances of technology, feasibility studies, standardization and product development.
- It is critical that the countries of the Americas invest in 5G research.
- Avoid debate (at least initially) on what 5G is. 5G does not (yet) describe any particular specification in any official document published by any standardization body.
- 5G planning should consider all major technology advances on the road to 5G.
- Wherever feasible, features being discussed as 5G requirements should be implemented as LTE-Advanced extensions before the full 5G is available. This will also give time to recoup the investment in 4G.
- There are ongoing enhancements in LTE-Advanced that will continue through 2018. 5G is envisioned to have initial deployments around 2020. It must be recognized that significant breakthroughs in new radio transmission interfaces may be accompanied by a break in backward compatibility.

In conclusion, it is recommended that the market drivers, use cases, requirements, regulatory considerations and technology elements described in Sections 2-7 all be taken into consideration in the further development of the end-to-end 5G system.

ACRONYM LIST

3D	Three Dimensional
3G	Third Generation
3GPP	Third Generation Partnership Project
4D	Four Dimensional
4G	Fourth Generation
5G	Fifth Generation
AAA	Authentication, Authorization and Accounting
ADAS	Advanced Driver Assistance Systems
ANDSF	Access Network Discovery and Selection Function
API	Application Program Interface
APN	Access Point Name
BS	Base Station
BYOD	Bring Your Own Device
CAPEX	Capital Expense
CA	Carrier Aggregation
CAT	Computerized Tomography
CCN	Content-Centric Networking
CCO	Coverage and Capacity Optimization
CDN	Content Distribution Network
CMAS	Commercial Mobile Alert System
CN	Core Network
CoMP	Coordinated Multi-Point
CRAN	Centralized Radio Access Network
D2D	Device-to-Device
DFT	Discrete Fourier Transform
DL	Downlink
DNS	Domain Name Service
DNSsec	Domain Name System Security Extensions
DoS	Denial of Service
DPI	Deep Packet Inspection
EPC	Evolved Packet Core
ETS	Emergency Telecommunication Service
ETWS	Earthquake and Tsunami Warning System
FBMC	Filter-Bank Multi Carrier
FCC	Federal Communications Commission
FIB	Forwarding Information Base
Gbps	Gigabits Per Second
GFDM	Generalized Frequency Division Multiplexing
GHz	Gigahertz
GPRS	General Packet Radio Service
HAS	HTTP Adaptive Streaming
HSPA+	High-Speed Packet Access Plus
HSS	Home Subscriber System
HTTP	Hypertext Transfer Protocol
ICN	Information Centric Networking
IM	Instant Message
IMEI	International Mobile Equipment Identity
IMT	International Mobile Telecommunications
IoT	Internet of Things
IP	Internet Protocol
IT	Information Technology
ITU	International Telecommunication Union
Km/h	Kilometers per hour
KPAS	Korean Public Alert System
LIPO	Local IP Access
LTE	Long Term Evolution
M2M	Machine-to-Machine
MANO	Management and. Orchestration
MHz	Megahertz
MIMO	Multiple Input, Multiple Output
MLB	Mobility Load Balancing
MMES	Multimedia Emergency Services
mmw	Millimeter wave
mmWave	Millimeter wave
MNO	Mobile Network Operator
MPTCP	Multipath Transmission Control Protocol
MRI	Magnetic Resonance Imaging
MTC	Machine Type Communication
NAT	Network Address Translation
NDN	Named-Data Networking
NFV	Network Function Virtualization
NFVI	Network Virtual Function Infrastructure
NGMN	Next Generation Mobile Networks Alliance
NOMA	Non-Orthogonal Multiple Access
NPRM	Notice of Proposed Rule Making
OFDM	Orthogonal Frequency Division Multiplexing
OPEX	Operating Expense
PDN	Packet Data Network
PGW	Packet Gateway
PIT	Pending Interest Table
PSTN	Public Switched Telephone Network
PWS	Public Warning System
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
RF	Radio Frequency
Rx	Receive
SAS	Spectrum Access System
SCMA	Sparse-Code Multiple Access
SDN	Software Defined Networking
SGW	Signaling Gateway
SIM	Subscr ber Identity Module
SIPTO	Selected Internet IP Traffic Offload
SMS	Short Message Service
SON	Self-Organizing Network
TAC	Type Allocation Code
TCP	Transmission Control Protocol
TLS	Transport Layer Security
TS	Technical Specification
Tx	Transmit
UE	User Equipment
UFMC	Universal Filtered Multi Carrier
UL	Uplink
UPCON	User Plane Congestion
URI	Uniform Resource Identifier
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
VNF	Virtual Network Function
Wi-Fi	Wireless Fidelity
WPS	Wireless Priority Service
WRC	World Radio Conference

ACKNOWLEDGEMENTS

The mission of 4G Americas is to advocate for and foster the advancement and full capabilities of the 3GPP family of mobile broadband technologies, including LTE-Advanced and beyond to 5G, throughout the ecosystem's networks, services, applications and wirelessly connected devices in the Americas. 4G Americas' Board of Governors members include Alcatel-Lucent, América Móvil, AT&T, Cable & Wireless, Cisco, CommScope, Entel, Ericsson, HP, Intel, Mitel, Nokia, Qualcomm, Sprint, T-Mobile US, Inc. and Telefónica.

4G Americas would like to recognize the significant project leadership and important contributions of Brian K. Daly and DeWayne Sennett of AT&T and Anders Svensson of Ericsson, as well as representatives from member companies on 4G Americas' Board of Governors who participated in the development of this white paper.