ETSI GR IP6 009 V1.1.1 (2017-03)

IPv6-based Industrial Internet leveraging 6TiSCH technology

Reference
DGR/IP6-0009

Keywords
6TiSCH, IPv6, network

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Group Report (GR) has been produced by ETSI Industry Specification Group (ISG) IPv6 Integration (IP6).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Executive summary

The Industrial Internet will enable deep process optimization in multiple industries by introducing Information Technology (IT) capabilities, such as Big Data and virtualization, to improve Operational Technology (OT) processes while reducing the OPEX, with the convergence of the IT and OT network. At the core of this revolution, a new breed of Deterministic Networks will provide enhancements that are required to fully emulate the traditional serial links and field buses that are widely deployed in that space over IPv6.

Deterministic Networking is a new (to IT networks) level of guarantee for network-based services, based on time, resource reservation, and enforcement. Deterministic Networking provides the capability to carry specified unicast or multicast data streams for real-time applications with extremely low data loss rates and bounded latency. Deterministic Networking technology allows for guarantees of 'worst-case' delivery. More precisely, the worst-case data loss and latency are guaranteed in a consistent fashion as multiple services are deployed on a common converged network infrastructure.

Deterministic Networking adds key capabilities to the Internet (wired, wireless, Layer 2 and Layer 3) to support time-sensitive mission-critical applications on a converged enterprise infrastructure. These capabilities are required to drive the connection of billions of things, and make available the vast amounts of data that IoE applications generate.

Deterministic Networking is a quantum step beyond existing QoS mechanisms. It implies time synchronization on all the nodes, often including source and destination, the centralized computation of the deterministic paths from a global perspective for a better optimization, new traffic shapers and schedulers within and at the edge to protect the network, and new hardware for time-triggered access to the media.

6TiSCH [i.107] enables Deterministic Networking over Low-Power Radios, controlled by a central intelligence called a PCE. At the same time, 6TiSCH allows traditional best effort IPv6 flows routed with the RPL routing protocol to utilize the portions of the bandwidth that are not allocated to deterministic flow. This way, the collection over IPv6 of traditionally unmeasured data can scale to vast numbers without interfering with the more critical flows for which all the necessary resources are reserved.

Introduction

It all started with point-to-point copper wires, transporting analogue signals for short messages, then telephone and television, industrial measurements and commands, anything though initially not data. Digital data networks, and then packet networks, came last; but with the advent of determinism, the late comers now show the potential to federate all original forms of wired and wireless communication and lead to the final convergence of all communication networks.

A generic and cheap replacement to serial cables to provide connectivity to all sorts of devices, coupled with resource-sharing meshed networks, are now required to simplify the cabling and drive the costs down in many industries, from transportation to manufacturing. Simple as it may seem to emulate the legacy forms of serial communications, reproducing the various aspects of a point-to-point electric cable over a multi-hop packet network is actually the hardest thing to do. Yet, the need is becoming more and more pressing, as:

1) managing all the existing sorts of cables and buses has become an increasingly costly complexity in many aspects of our lives; and

2) point-to-point wires will not scale to serve the exploding needs of the Internet of Things.

A paper on "Integrating an Industrial Wireless Sensor Network with your Switched Ethernet and IP Network [i.67]" was presented at the Emerson Exchange 2008 conference in Washington. The paper discussed how Wireless Sensor Networks (WSNs), which are in essence cheaper and faster to deploy than traditional wired field-buses, could leverage the entire network to connect the sensors to a centralized controlling application located afar on the carpeted floor, for Industrial supervisory control or logging. At the same time, the paper stressed issues that are raised when integrating a classical, often proprietary industrial automation network, with tight response time and availability constraints, into a wider IP network based on packet-switched and Internet technologies. With this and a collection of other papers [i.69], [i.70], [i.71] and [i.72], the realization is now coming that with techniques such as flow isolation, high availability and a new generation of Quality of Service (QoS), the times of the convergence of these networks are finally approaching.

1 Scope

The present document outlines a general architecture for an Industrial Internet, providing motivation for the deployment, and some technical guidelines with a focus on deterministic and low power technologies, for a prospective IPv6-based Industrial Internet leveraging deterministic wireless technology. The present document elaborates on deterministic networking, wired and wireless, for application in the Industrial Internet.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] IETF RFC 6371: "Operations, Administration, and Maintenance Framework for MPLS-Based Transport Networks".

[i.2] J. Araujo et al.: "High availability automation networks: PRP and HSR ring implementations" in: Industrial Electronics (ISIE), 2012 IEEE International Symposium on, IEEE, 2012, pp. 1197-1202.

[i.3] IETF RFC 7471: "OSPF Traffic Engineering (TE) Metric Extensions".

[i.4] D. Beller and R. Sperber: "MPLS-TP-The New Technology for Packet Transport Networks" in: DFN-Forum Kommunikationstechnologien, vol. 149, 2009, pp. 81-92.

[i.5] IETF RFC 6119: "IPv6 Traffic Engineering in IS-IS".

[i.6] M. S. Borella et al.: "Methods for determining sendable information content based on a determined network latency", US Patent 6,182,125, Jan. 2001.

[i.7] IETF RFC 7490: "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)".

[i.8] A. Colvin: "CSMA with collision avoidance" in: Computer Communications 6.5 (1983), pp. 227-235.

[i.9] IEC 62439-3:2009: "Industrial communication networks - High availability automation networks - Part 3: Parallel Redundancy Protocol (PRP) and High-availability Seamless Redundancy (HSR)".

[i.10] IEC 62591:2016: "Industrial networks - Wireless communication network and communication profiles - WirelessHART™".

[i.11] IEC 62734:2014: "Industrial networks - Wireless communication network and communication profiles - ISA 100.11a".

[i.12] IEC 62601:2015: "Industrial networks - Wireless communication network and communication profiles - WIA-PA".

[i.13] IEC 61850: 2016 SER: "Communication networks and systems for power utility automation - ALL PARTS".

[i.14] S. S. Craciunas and R. S. Oliver: "SMT-based task-and network-level static schedule generation for time-triggered networked systems" in: Proceedings of the 22nd International Conference on Real-Time Networks and Systems, ACM, 2014, p. 45.

[i.15] IETF draft-ietf-6tisch-6top-sf0-02D: "6TiSCH 6top Scheduling Function Zero (SF0)".

[i.16] Y. Fang and Y. Zhang: "Call admission control schemes and performance analysis in wireless mobile networks" in: IEEE Transactions on vehicular technology 51.2 (2002), pp. 371-382.

[i.17] IETF RFC 4655: "A Path Computation Element (PCE)-Based Architecture".

[i.18] IETF RFC 5960: "MPLS Transport Profile Data Plane Architecture".

[i.19] M. Goraj and R. Harada: "Migration paths for IEC 61850 substation communication networks towards superb redundancy based on hybrid PRP and HSR topologies" in: Developments in Power Systems Protection, 2012. DPSP 2012. 11th International Conference on, IET, 2012, pp. 1-6.

[i.20] IETF draft-ietf-detnet-use-cases-11: "Deterministic Networking Use Cases".

[i.21] T. Hasegawa et al.: "Industrial wireless standardization - Scope and implementation of ISA SP100 standard" in: SICE Annual Conference (SICE), 2011 Proceedings of, IEEE, 2011, pp. 2059-2064.

[i.22] K.-i. Hwang: "Energy efficient channel agility utilizing dynamic multi-channel CCA for ZigBee RF4CE" in: IEEE Transactions on Consumer Electronics 57.1 (2011), pp. 113-119.

[i.23] D. M. Ingram, P. Schaub, and D. A. Campbell: "Use of precision time protocol to synchronize sampled-value process buses" in: IEEE Transactions on Instrumentation and Measurement 61.5 (2012), pp. 1173-1180.

[i.24] H. Kirrmann et al.: "HSR: Zero recovery time and low-cost redundancy for Industrial Ethernet (High availability seamless redundancy, IEC 62439-3)" in: Proceedings of the 14th IEEE international conference on Emerging technologies & factory automation, IEEE Press, 2009, pp. 203-206.

[i.25] H. Kopetz et al.: "The time-triggered ethernet (TTE) design" in: 8th IEEE International Symposium on Object-Oriented Real-Time Distributed Computing (ISORC'05), IEEE, 2005, pp. 22-33.

[i.26] W. Liang et al.: "Survey and experiments of WIA-PA specification of industrial wireless network" in: Wireless Communications and Mobile Computing 11.8 (2011), pp. 1197-1212, issn: 1530-8677, doi: 10.1002/wcm.976.

NOTE: Available at http://dx.doi.org/10.1002/wcm.976.

[i.27] Z. Lin and S. Pearson: "An inside look at industrial Ethernet communication protocols" in: White Paper Texas Instruments (2013).

[i.28] J. D. MacKay: "Applications and Opportunities for the IEEE 1588 Standard in Military Applications", tech. rep., DTIC Document, Jan. 2007.

NOTE: Available at www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA474269.

[i.29] S. Mangold et al.: "Analysis of IEEE Std 802.11 e for QoS support in wireless LANs" in: IEEE wireless communications 10.6 (2003), pp. 40-50.

[i.30] E. Mannie: "Generalized multi-protocol label switching (GMPLS) architecture" in: Interface 501 (2004), p. 19.

[i.31] N. McKeown et al.: "OpenFlow: enabling innovation in campus networks" in: ACM SIGCOMM Computer Communication Review 38.2 (2008), pp. 69-74.

[i.32] S. Meier and H. Weibel: "IEEE 1588 applied in the environment of high availability LANs" in: 2007 IEEE International Symposium on Precision Clock Synchronization for Measurement, Control and Communication, IEEE, 2007, pp. 100-104.

[i.33] A. Morell et al.: "Label switching over IEEE Std 802.15.4e networks" in: Transactions on Emerging Telecommunications Technologies 24.5 (2013), pp. 458-475.

[i.34] T. Neagoe, V. Cristea, and L. Banica: "NTP versus PTP in com puter networks clock synchronization" in: 2006 IEEE International Symposium on Industrial Electronics, vol. 1, IEEE, 2006, pp. 317-362.

[i.35] M. Nixon and T. Round Rock: "A Comparison of WirelessHART™ and ISA100. 11a" in: Whitepaper, Emerson Process Management (2012), pp. 1-36.

[i.36] B. A. A. Nunes et al.: "A survey of software-defined networking: Past, present, and future of programmable networks" in: IEEE Communications Surveys & Tutorials 16.3 (2014), pp. 1617-1634.

[i.37] A. de la Oliva et al.: "An overview of the CPRI specification and its application to C-RAN-based LTE scenarios" in: IEEE Communications Magazine 54.2 (2016), pp. 152-159.

[i.38] R. S. Oliver, S. S. Craciunas, and G. Stöger: "Analysis of deterministic ethernet scheduling for the industrial internet of things" in: 2014 IEEE 19th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), IEEE, 2014, pp. 320--324.

[i.39] IETF RFC 6378: "MPLS Transport Profile (MPLS-TP) Linear Protection".

[i.40] IETF RFC 5543: "BGP Traffic Engineering Attribute".

[i.41] M. R. Palattella et al.: "Standardized protocol stack for the internet of (important) things" in: IEEE Communications Surveys & Tutorials 15.3 (2013), pp. 1389-1406.

[i.42] S. Petersen and S. Carlsen: "WirelessHART versus ISA100. 11a: the format war hits the factory floor" in: IEEE Industrial Electronics Magazine 5.4 (2011), pp. 23-34.

[i.43] IETF RFC 5673: "Industrial Routing Requirements in LowPower and Lossy Networks".

[i.44] K. Pister and L. Doherty: "TSMP: Time synchronized mesh protocol" in: IASTED Distributed Sensor Networks (2008), pp. 391-398.

[i.45] G. Pujolle: "SDN (Software-Defined Networking)" in: Software Networks, pp. 15- 48.

[i.46] IETF draft-ietf-6tisch-dtsecuritysecure-join-00: "6tisch Secure Join protocol".

[i.47] IETF RFC 5714: "IP Fast Reroute Framework".

[i.48] IETF RFC 2212: "Specification of Guaranteed Quality of Service".

[i.49] F. Shu et al.: "Packet loss analysis of the IEEE Std 802.15. 4 MAC without acknowledgements" in: IEEE communications letters 11.1 (2007), pp. 79-81.

[i.50] K. Srinivasan et al.: "The β-factor: measuring wireless link burstiness" in: Proceedings of the 6th ACM conference on Embedded network sensor systems, ACM, 2008, pp. 29-42.

[i.51] B. Sundararaman, U. Buy and A. D. Kshemkalyani: "Clock synchronization for wireless sensor networks: a survey" in: Ad hoc networks 3.3 (2005), pp. 281-323.

[i.52] J.-C. Tan and W. Luan: "IEC 61850 based substation automation system architecture design" in: 2011 IEEE Power and Energy Society General Meeting, 2011.

[i.53] R. Teixeira et al.: "Dynamics of hot-potato routing in IP networks" in: ACM SIGMETRICS Performance Evaluation Review, vol. 32, 1, ACM, 2004, pp. 307-319.

[i.54] IETF draft-thubert-6tisch-4detnet-01: "6TiSCH requirements for DetNet".

[i.55] IETF draft-ietf-6tisch-architecture-10: "An Architecture for IPv6 over the TSCH mode of IEEE Std 802.15.4".

[i.56] IETF draft-ietf-6lo-backbone-router-02: "IPv6 Backbone Router".

[i.57] IETF draft-ietf-detnet-architecture-00: "Deterministic Networking Architecture".

[i.58] IETF draft-thubert-6lo-forwarding-fragments-03: "LLN Fragment Forwarding and Recovery".

[i.59] IETF draft-ietf-6tisch-minimal-21: "Minimal 6TiSCH Configuration".

[i.60] IETF draft-ietf-6tisch-minimal-security-00: "Minimal Security Framework for 6TiSCH".

[i.61] IETF draft-ietf-6tisch-6top-protocol-03: "6top Protocol (6P)".

[i.62] T. Watteyne, A. Mehta, and K. Pister: "Reliability through frequency diversity: why channel hopping makes sense" in: Proceedings of the 6th ACM symposium on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks, ACM, 2009, pp. 116-123.

[i.63] T. Watteyne et al.: "Industrial IEEE Std 802.15.4e networks: Performance and trade-offs" in: 2015 IEEE International Conference on Communications (ICC), IEEE, 2015, pp. 604-609.

[i.64] T. Watteyne et al.: "Mitigating Multipath Fading Through Channel Hopping in Wireless Sensor Networks" in: International Conference on Communications (ICC), IEEE, Cape Town, South Africa, May 2010.

[i.65] IETF RFC 3630: "Traffic Engineering (TE) Extensions to OSPF Version 2".

[i.66] IETF RFC 5286: "Basic Specification for IP Fast Reroute: Loop-Free Alternates".

[i.67] E. Ziouva and T. Antonakopoulos: "CSMA/CA performance under high traffic conditions: throughput and delay analysis" in: Computer communications 25.3 (2002), pp. 313-321.

[i.68] Cisco, Emerson: Integrating an Industrial Wireless Sensor Network with Your Plant's Switched Ethernet and IP Network".

NOTE: Available at http://www.controlglobal.com/assets/14WPpdf/140303-Cisco-Emerson-WirelessNetworks.pdf.

[i.69] Atos White Paper: "The convergence of IT and Operational Technology".

NOTE: Available at https://atos.net/content/dam/global/ascent-whitepapers/ascent-whitepaper-the-convergence-of-it-and-operational-technology.pdf.

[i.70] Cisco White Paper.

NOTE: Available at http://www.cisco.com/c/en/us/products/collateral/se/internet-of-everything/white-paper-c11-735380.pdf.

[i.71] Statseeker White Paper: "The IT/OT Convergence - Bridging the Gap".

NOTE: Available at http://www.isssource.com/wp-content/uploads/2015/11/111115statseeker-IT-OT-Convergence-White-Paper.pdf.

[i.72] NexDefense White Paper: "IT/OT Convergence - Bridging the Divide".

NOTE: Available at http://ics.sans.org/media/IT-OT-Convergence-NexDefense-Whitepaper.pdf.

[i.73] IEEE/IETF Coordination meeting: "Deterministic Networking".

NOTE: Available at https://www.iab.org/wp-content/IAB-uploads/2013/01/tsn-nfinn-Deterministic-Networking-BOF-0914-v1.pdf.

[i.74] Schneider Electric: "How the Convergence of IT and OT Enables Smart Grid Development".

NOTE: Available at http://cdn.iotwf.com/resources/10/How-the-Convergence-of-IT-and-OT-Enables-Smart-Grid-Development_2013.pdf.

[i.75] http://www.computerweekly.com/opinion/Big-data-to-unlock-value-from-the-Industrial-Internet-of-Things.

[i.76] http://sloanreview.mit.edu/case-study/ge-big-bet-on-data-and-analytics/.

[i.77] http://www.ge.com/digital/blog/everything-you-need-know-about-industrial-Internet-things.

[i.78] http://www.investors.com/news/ge-courts-silicon-valley-investors-for-digital-industrial-push/.

[i.79] ODVA Industry Conference: "Time Sensitive Network (TSN) Protocols and use in EtherNet/IP Systems".

NOTE: Available at https://www.odva.org/Portals/0/Library/Conference/2015_ODVA_Conference_Ditzel-Didier_TSN.pdf.

[i.80] IEEE 802.15™: "Wireless Personal Area Networks (PANs)".

NOTE: Available at http://ieee802.org/15/pub/TG4.html.

[i.81] IEEE 802.15.4e™-2012 :"IEEE Standard for Local and metropolitan area networks-- Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs) Amendment 1: MAC sublayer".

NOTE: Available at http://standards.ieee.org/findstds/standard/802.15.4e-2012.html.

[i.82] https://fieldcommgroup.org/technologies/hart/hart-technology.

[i.83] EtherCAT Technology Group: "Industrial Ethernet Technologies: Overview".

NOTE: Available at http://www.ethercat.org/2011/italy/download/04_Industrial_Ethernet_Technologies_IT_1102.pdf.

[i.84] National Instruments: "Ethernet-Based Industrial Communication Protocols - Connect LabVIEW to PLCs".

NOTE: Available at ftp://ftp.ni.com/pub/events/nits/presentations/2008/ethernet_protocols.pdf.

[i.85] Available at http://www.iebmedia.com/index.php?id=8096&parentid=63&themeid=275&hft=65&showdetail=true&bb=1&PHPSES.

[i.86] ABB White Paper: "Introduction to WISA".

NOTE: Available at http://www.millennialnet.com/MillennialNet/media/Resources_Media/WhitePapers/WhitePaper_IntroductiontoWISA_V2.pdf.

[i.87] Harman International: "Understanding IEEE 1722 - AVB Transport Protocol - AVBTP".

NOTE: Available at http://www.ieee802.org/1/files/public/docs2009/avb-rboatright-p1722-explained-0903.pdf.

[i.88] AVnu Alliance™ White Paper: "AVB for Profesional A/V Use".

NOTE: Available at http://avnu.org/wp-content/uploads/2014/05/AVnu-Pro__White-Paper.pdf.

[i.89] http://www.electronicsweekly.com/news/products/analog/low-power-radio-standard-simplifies-sensor-networks-2012-10/.

[i.90] http://www.ieee802.org/1/pages/tsn.html.

[i.91] https://www.opennetworking.org/sdn-resources/sdn-definition.

[i.92] SAE International: "Comparison of CAN, FlexRay, and Ethernet Architectures for the Design of ABS Systems".

NOTE: Available at http://papers.sae.org/2011-01-0453/.

[i.93] http://www.cisco.com/c/en/us/td/docs/net_mgmt/ciscoworks_lan_management_solution/4-2/user/guide/configuration/config/configvlan.html.

[i.94] http://www.cisco.com/c/en/us/td/docs/net_mgmt/ciscoworks_lan_management_solution/4-2/user/guide/configuration/config/vrf.html.

[i.95] http://www.mixedcriticalityforum.org/about/timing-analysis/.

[i.96] http://www.embedded.com/electronics-blogs/cole-bin/4406659/Deterministic-networking--from-niches-to-the-mainstream-.

[i.97] http://www.itential.com/blog/overhyped-sdn/.

[i.98] http://www.cisco.com/c/en/us/products/collateral/cloud-systems-management/network-services-orchestrator/datasheet-c78-734669.html.

[i.99] https://www.arista.com/en/um-eos-4172f/eos-section-45-1-introduction#ww1148008.

[i.100] http://avnu.org/proav/.

[i.101] https://www.sdxcentral.com/sdn/definitions/who-is-open-networking-foundation-onf/.

[i.102] http://www.3gpp.org/.

[i.103] IEEE 802.15™: "WPAN Assigned Numbers Authority (ANA)".

NOTE: Available at http://www.ieee802.org/15/ANA.html.

[i.104] IANA Protocol Parameter Assignments.

NOTE: Available at http://www.ietf.org/iana.html.

[i.105] GE: "Industrial Internet: Pushing the Boundaries of Minds and Machines".

NOTE: Available at http://www.ge.com/docs/chapters/Industrial_Internet.pdf.

[i.106] https://www.geoilandgas.com/our-voice/power-one.

[i.107] IPv6 over the TSCH mode of IEEE 802.15.4e (6tisch).

NOTE: Available at https://datatracker.ietf.org/wg/6tisch/charter/.

[i.108] IEEE 802.11™: "Wireless LANs".

NOTE: Available at http://standards.ieee.org/about/get/802/802.11.html.

[i.109] Deterministic Networking (detnet).

NOTE: Available at https://datatracker.ietf.org/group/detnet/charter/.

[i.110] IEEE 1901.2™-2013: "IEEE Standard for Low-Frequency (less than 500 kHz) Narrowband Power Line Communications for Smart Grid Applications".

NOTE: Available at https://standards.ieee.org/findstds/standard/1901.2-2013.html.

[i.111] IEEE 802.11ah™: "IEEE Approved Draft Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Amendment 2: Sub 1 GHz License Exempt Operation".

NOTE: Available at http://www.wi-fi.org/discover-wi-fi/wi-fi-halow.

[i.112] IEEE 1588™: "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems".

NOTE: Available at http://standards.ieee.org/findstds/standard/1588-2008.html.

[i.113] IEEE 802.1AS™-2011: "IEEE Standard for Local and Metropolitan Area Networks - Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks".

NOTE: Available at http://standards.ieee.org/findstds/standard/802.1AS-2011.html.

[i.114] IEEE 802.1Qbu™-2016: "IEEE Standard for Local and metropolitan area networks -- Bridges and Bridged Networks -- Amendment 26: Frame Preemption".

NOTE: Available at http://standards.ieee.org/findstds/standard/802.1Qbu-2016.html.

[i.115] ETSI EN 300 328: "Wideband transmission systems; Data transmission equipment operating in the 2,4 GHz ISM band and using wide band modulation techniques; Harmonised Standard covering the essential requirements of article 3.2 of Directive 2014/53/EU".

3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

3GPP	3rd Generation Partnership Project
6lo	IETF IPv6 over Networks of Resource-constrained Nodes WG (continues 6LoWPAN)
6LoWPAN	IETF IPv6 over LWPANs WG
6TiSCH	IETF IPv6 over the TSCH mode of IEEE Std 802.15.4 WG [i.55]
AMI	Advanced Metering Infrastructure
AMR	Automatic Meter Reading
ARC	Available Routing Construct
AS	Autonomous System
AVB	Audio/Video Bridging
BBR	BackBone Router
BER	Bit Error Rate
BGP	Border Gateway Protocol
BIER	Bit Indexed Explicit Replication
BTLE	(or BLE) BlueTooth Low Energy
CA	Collision Avoidance
CAC	Call Admission Control
CAN	Controller Area Network
CAPEX	CAPital EXxpenditure
CDU	Channel Distribution/Usage
CIP	Common Industrial Protocol
CLI	Command-Line Interface
CRC	Cyclic Redundancy Check
CSMA	Carrier Sense Multiple Access
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
DAD	Duplicate Address Detection
FCC	US Federal Communications Commission
G-MPLS	Generalized MPLS
GMPLS	Generalized MultiProtocol Label Switching
GPS	Global Positioning System
HMI	Human Machine Interface
HSR	High-availability Seamless Redundancy
ICT	Information and Communication Technologies
IE	Information Element
IESG	Internet Engineering Steering Group
IFS	Inter Frame Space
IGP	Interior Gateway Protocol
IoT	Internet of Things
IP	Internet Protocol
IPv6	Internet Protocol version 6
IRT	Isochronous Real-Time
ISDN	Integrated Services Digital Network
IS-IS	Intermediate System to Intermediate System
ISM	Industrial, Scientific, and Medical
IT	Information Technology
LAN	Local Area Network
LDP	Label Distribution Protocol
LFA	Loop-Free Alternates

LLN	Low-Power and Lossy Network
MAC	Medium Access Control
MPLS	MultiProtocol Label Switching
MPLS-TP	MPLS-Transport Profile
MRT	Maximally Redundant Trees
ND	Neighbor Discovery
NFC	Near Field Communication
NFV	Network Functions Virtualization
NIC	Network Interface Card
NTP	Network Time Protocol
OAM	Operations, Administration and Maintenance
ONF	Open Networking Foundation
OPEX	OPerational EXpenditures
OSPF	Open Shortest Path First
OT	Operational Technology
PCE	Path Computation Element
PCEP	Path Computation Element Protocol
PCI	Peripheral Component Interconnect
PDR	Packet Delivery Ratio
PER	Packet Error Rate
PHY	PHYsical layer
PMN	Professional Media Networking
PRP	Parallel Redundancy Protocol
PSK	Pre-Shared Key
PTP	Precision Time Protocol
QoS	Quality of Service
REP	Resilient Ethernet Protocol
RFC	Request For Comment
RPL	IPv6 Routing Protocol for Low-Power and Lossy Networks
RSVP	Resource Reservation Protocol
SDH	Synchronous Digital Hierarchy
SDN	Software-Defined Networking
SERCOS	SErial Real-time COmmunications System
SONET	Synchronous Optical NETworking
SR	Segment Routing
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TE	Traffic Engineering
TEAS	Traffic Engineering Architecture and Signaling
TSCH	Time-Slotted Channel Hopping
TSMP	Time Synchronized Mesh Protocol
TSN	Time-Sensitive Networking
TSPEC	Traffic SPECification
TTE	Time-Triggered Ethernet
TV	TeleVision
VLAN	Virtual Local Area Network
VRF	Virtual Routing/Forwarding
WG	Working Group
WIA-FA	Wireless Networks for Industrial Automation - Factory Automation
WIA-PA	Wireless Networks for Industrial Automation - Process Automation
WISA	Wireless Interface for Sensors and Actuators
WPAN	Wireless Personal Area Network
WSN	Wireless Sensor Network

4 Converging Networks for the Industrial Internet

4.1 On Operational Technology

"Operational technology (OT) is hardware and software that detects or causes a change through the direct monitoring and/or control of physical devices, processes and events in the enterprise" - Source: Gartner.

In practice, OT refers to industrial networks, which focus on highly reliable, secure and Deterministic Networking. In OT environments, Deterministic Networks are characterized as providing a guaranteed bandwidth with extremely low packet loss rates, bounded latency, and low jitter [i.73]. OT networks are typically used for monitoring systems and supporting control loops, as well as movement detection systems for use in process control (i.e. continuous manufacturing) and factory automation (i.e. discrete manufacturing), and protection systems in the SmartGrid [i.74].

Due to its different goals, OT has evolved in parallel but in a manner that is radically different from Information Technology/Information and Communications Technology (IT/ICT), which until now relied on selective queuing and discarding of IP packets to achieve end-to-end flow control over the Internet, and provided limited guarantees in terms of delivery and latency. For that lack of determinism and an associated lack of trust, IT and OT networks have been maintained segregated.

4.2 Enabling the IT/OT convergence

The motivation behind the original vision of an Industrial Internet [i.105] was that a single percentile point of operational optimization [i.106] would enable massive savings across many vertical industries. But reaching additional levels of optimization is not an easy task; it would require collecting and processing huge amounts of live measurements by widely distributed OT sensing and IT analytics capabilities [i.75], [i.76], [i.77] and [i.78].

These currently missing measurements comprise all sorts of diagnostics and sensor data that in many cases today are captured but not reported, either because the cost of wiring the devices would be prohibitive, or, if a device has connectivity to a control network, because the control network is kept isolated from the IT network and the wider Internet for security reasons.

In order to avoid skyrocketing operational costs in maintaining highly trained teams for multiple solutions, the missing measurements have to share the same infrastructure (network and management) as the deterministic OT flows. It results that the Industrial Internet vision can only be achieved by the convergence of IT and OT, whereby the network becomes capable of emulating the properties of deterministic OT circuits in the same fabric that serves traditional best effort IP applications.

In response to needs from different vertical OT industries, new efforts at the IETF and the IEEE [i.79] to enable traffic that requires bounded latency in a worst case scenario and is generally sensitive to packet loss and/or jitter, for application in a large variety of use cases [i.20] with a high degree of operational criticality.

The upcoming protocols will support a mix of deterministic and classical best-effort traffic to be transported across Ethernet bridges and over IP networks, respectively. In particular:

- A first generation of Ethernet-based standards, called Audio/Video Bridging (AVB) [i.87], was developed at the IEEE Std 802.1™ [i.88] for the Professional Media Networks.
- With TimeSlotted Channel Hopping (TSCH), IEEE Std 802.15.4 evolved into a highly predictable, quasi-deterministic Medium Access Control (MAC) [i.89] technology.
- The Time-Sensitive Networking (TSN) [i.90] Task Group (TG) at the IEEE and the Deterministic Networking (DetNet) Working Group (WG) at the IETF are now generalizing those methods to transport deterministic flows across Ethernet bridges and over IP networks, respectively.

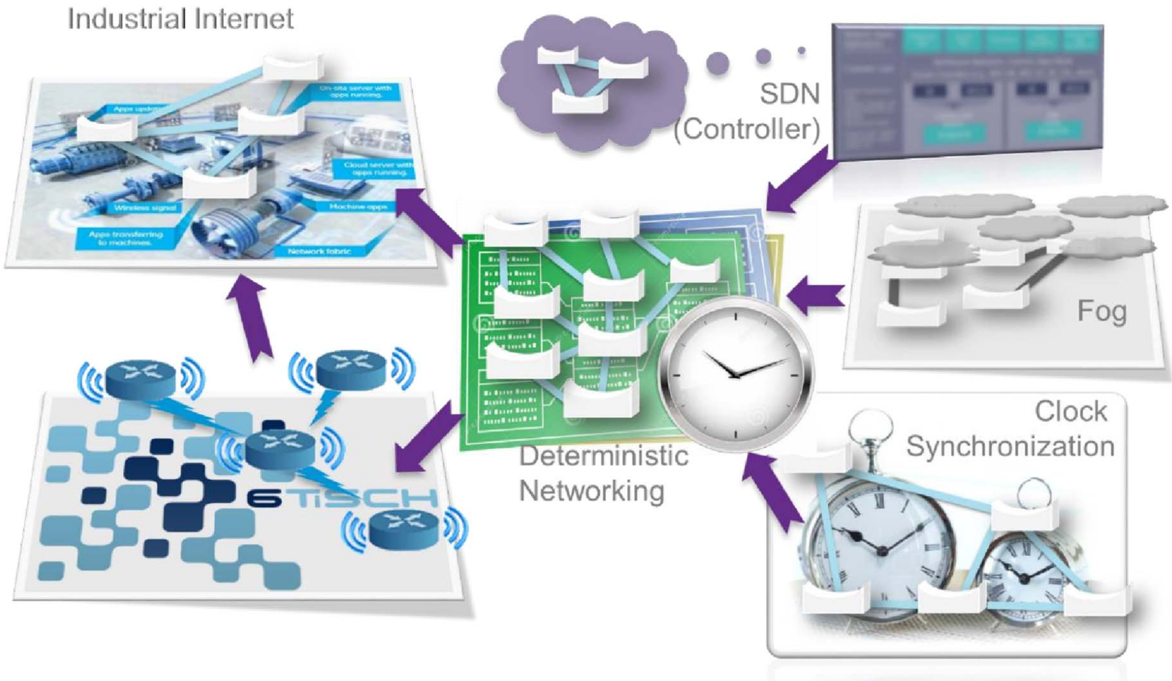


Figure 1: Enabling trends

In figure 1, the trend to bringing determinism in IT protocols relative to other trends that lead to it is positioned, and ones that in turn it enables. *Deterministic Networking* refers to this new trend that is now perceived as the key enabler for the convergence of IT and OT networks.

On the one hand, determinism builds on recent trends of *Clock Synchronization* [i.51] to control time-shared network resources, and on Software Defined Networking (SDN) [i.91] and [i.45] for the centralized visibility on network resources and capability to achieve a global optimization of the network.

On the other hand, Deterministic Networking is a key enabler for a series of new trends that involve the replacement of a serial cable or a multidrop synchronous bus over a switched fabric. Examples of such trends include the migration to IP of Professional Media Networks such as found in production studios and concert arenas, the replacement of dedicated automotive buses such as the Controller Area Network (CAN bus) and FlexRay [i.92] by cheap 100 MBps Ethernet over twisted pairs in the new models of cars, and the Compute Platform Disaggregation whereby the internal bus inside a server such as a Peripheral Component Interconnect (PCI) bus is replaced by an Ethernet switched fabric.

Among those trends, the focus is on the Industrial Internet; in that case, deterministic signals that are used to control industrial processes have to be isolated, and yet share the physical media with other information such as diagnostics and sensory data that are used to monitor and improve the processes.

4.3 The path to the IT/OT Convergence

At first, people realized that an on/off electrical signal could be perceived at very long distances and the telegraph was born. They realized that the voltage could be modulated and the analogue audio signals, which still available today from devices equipped with those legacy jack and RCA connectors, joined the party.

In the last twenty years, voice, data and video networks have converged to digital over IP. Mail delivery has become quasi-immediate and volumes have multiplied; long distance voice is now mostly free and the videophone is finally a reality; TV is available on-demand and games became interactive and massively multi-player.

The convergence of highly heterogeneous networks to IT/ICT resulted in significant drops in price for the end user while adding distinct new value to the related services.

Yet, and even though similar benefits can be envisioned when converging new applications over the Internet, there are still many disjoint branches in the networking family tree; many use cases where mission-specific applications continue to utilize dedicated point-to-point analogue and digital technologies for their operations.

Even as a number of industrial protocols are now migrating to open standards such as Ethernet and IP, the typical OT network is kept isolated from the IT network and operated by a different crew of OT specialists, which yields double operational expenses compared to a converged network infrastructure and management.

Forty years ago, OT people found that information encoded as an analogue modulation of current could be carried virtually instantly and with no loss over the distance; the basis of industrial control was laid out. Even today, most control signals rely on modulated current, typically between 4 mA and 20 mA, to report the variation of an observed phenomenon.

Then came digitization, which enabled to carry more data and control the device, but also introduced latency to industrial processes, time to encode a series of bits on a link and transport them along, which in turn may limit the amount of transported information [i.83], [i.84] and [i.6].

The need to save cable and simplify wiring lead to Time Division Multiplexing (TDM) of signals from multiple devices over shared digital buses, each signal being granted access to the medium at a fixed period for a fixed duration; with TDM, came more latency, waiting for the next reserved access time.

Statistical multiplexing, with Ethernet and IP, was then introduced to achieve higher speeds at lower cost, and with it finally came jitter and congestion loss.

Some OT applications evolved to compensate for the transport degradation with recovery mechanisms and jitter absorption buffers, at the expense of yet some additional latency. However, this did not seem acceptable for all, and in order to avoid those unwanted statistical effects, competing and not interoperable solutions appeared, driven by multiple standard defining organizations (SDOs), consortia and individual vendors.

Notable examples are, in the wired space [i.27], PROFINET Isochronous Real-Time (IRT), POWERLINK, SERCOS III, CC-Link IE, and to some degree Modbus TCP and EtherNet/IP which are now converged in the CIP specification, and in the wireless space, ISA100.11a and WirelessHART [i.35], WIA-PA and WIA-FA, iPCF [i.85] and WISA [i.86].

In the real world today, operational signals are still massively carried as a simple analogue modulation of the electrical current, over costly point-to-point wires. In order to replace those wires with a cheaper Ethernet-based switched fabric federating multiple access links, a limited subset of deployed control networks made all the steps towards digital statistical multiplexing; but even those cannot interoperate with IT technology, and the convergence is stalled.

The main technical reason for not converging those networks derives from the limits of the protocols that sustain the Internet as it is known today. With multi-hop operations and statistical multiplexing, IP technologies lack *"determinism"*.

4.4 The case of Low-power Lossy Networks

The quality of transmission over IEEE 802.15.4 [i.55] radios is affected by multiple elements, such as the relative location of objects in the environment and interferers of all kinds; these elements may be extremely difficult to control (e.g. radars) and may change brutally (e.g. a door opens); it results that a given channel cannot be expected to remain stable over a long period of time, and that some Channel Agility is required to guarantee a degree of service continuity over a long period of time.

When a radio transmission fails, adding diversity to the transmission characteristics improves the chances to avoid the cause of the failure and thus, those of a successful retransmission. A retransmission over the same channel adds only time diversity, and unless the cause of the loss is really transient, it is bound to fail for the same reason the original transmission did. But, as discussed by Srinivasan [i.50] in the case for IEEE Std 802.15.4 [i.55] and IEEE Std 802.11 [i.29] networks, other forms of diversity can help alleviate the issue.

Considering that those radios are highly sensitive to multipath fading [i.64], and that in turn multipath fading is highly sensitive to both location and frequency, it makes sense to add channel diversity to the retransmission. Trying this, it was found that switching channels dynamically based on a variation of the link quality also yields transient periods of instability (more on this in clause 6.4).

So in the years 2003-2007, providing deterministic services over IEEE Wireless Personal Area Networks (WPANs) [i.80] such as IEEE Std 802.15.4 [i.55] Low-Rate WPANs (LoWPAN) appeared to be extremely challenging. Around that time, Kris. Pister developed a novel approach to combine frequency diversity and channel hopping based on time synchronization, and they introduced the quantum leap that brought determinism over low-power wireless, with the Time Synchronized Mesh Protocol (TSMP) [i.44].

TSMP introduced a scheduled combination of frequency and time diversity that defeats most interferences and can reach wire-equivalent reliability on battery-operated devices. TSMP opened the way to Time-Slotted Channel Hopping (TSCH) [i.63], which was published as an amendment [i.81] in 2012, and is now updated and retrofitted in the mainline IEEE Std 802.15.4 [i.55].

TSCH was immediately adopted by the Process Control industry, and two competing industrial WSN standards were developed in the following years, both leveraging this technique; first came WirelessHART, which shipped with HART7 [i.82], and then ISA100.11a [i.21], which brought a limited support of IPv6.

The above standards are optimized for application in industrial Process Control; they are not designed to access the global Internet. In order to realize the Industrial Internet vision for wireless, there was a need to develop a new series of open standards combining best-effort and time-sensitive IPv6 traffic over TSCH. This work has started in 2013 at the IETF 6TiSCH WG and is now delivering its first round of standards.

However, wire-equivalent reliability is still not determinism. For safety and critical control applications, there is a need to guarantee the timely delivery of packets, even in the face of accidental situations such as the failure of a device or an obstacle moving in the way of the transmission. For this, additional forms of diversity, such as spatial diversity, and sometimes implementation and even technology diversities, are required.

5 What is Deterministic Networking?

5.1 Common definitions (from Web encyclopaedia)

In mathematics and physics:

"*A deterministic system is a system in which no randomness is involved in the development of future states of the system. A deterministic model will thus always produce the same output from a given starting condition or initial state.*"

In philosophy:

"*A deterministic system is a conceptual model of the philosophical doctrine of determinism applied to a system for understanding everything that has and will occur in the system, based on the physical outcomes of causality. In a deterministic system, every action, or cause, produces a reaction, or effect, and every reaction, in turn, becomes the cause of subsequent reactions. The totality of these cascading events can theoretically show exactly how the system will exist at any moment in time*."

A sense of what Deterministic Networking is has emerged as the capability to effectively emulate point-to-point wires on switched networks that were initially designed to serve IT/ICT and then Internet of Things (IoT) applications, such as Ethernet and IEEE 802.15.4 [i.55], and the IETF is now extending the Pseudo-Wire emulation of Ethernet over IP to provide Deterministic Networking (DetNet) services over larger networks.

Various flavours of this concept can be found around, for instance with the concept of Network Slicing at the 3rd Generation Partnership Project (3GPP) [i.102]. With these methods, an overlay, which is effectively a logical structure of meshed tunnels, would inherit physical properties such as a portion of the available buffers and bandwidth.

This clause details the characteristics of Deterministic Networking, how the concept can be achieved on wired infrastructures, and the direction taken by the standards bodies to enable deterministic properties on packet networks.

Determinism in a network brings the guarantee that a particular information is transported across the network in a tight window of time, and that a periodic process will be repeated identically every time. Determinism is a required property in the power grid, to ensure that high-tension lines breakers can be activated within milliseconds, in public transportation to make sure that automated vehicles are operated safely for their passengers, and in industrial automation for control loops.

To further delineate the concepts behind Deterministic Networking, let us use some analogies.

5.2 The train analogy (to control loop traffic)

The analogue of a congestion loss in the railway system, that is the collision of two trains using the same rails at the same time, is avoided in the real world by fully scheduled operations that repeat, day after day, the predetermined schema that is a train schedule.

Figure 2 illustrates the sheer complexity of computing an optimal schedule for multiple trains that will share a same infrastructure, with the goal to minimize the end-to-end time and the constraint to avoid collisions by ensuring that at most one train is present on any section of the rail at any point of time.

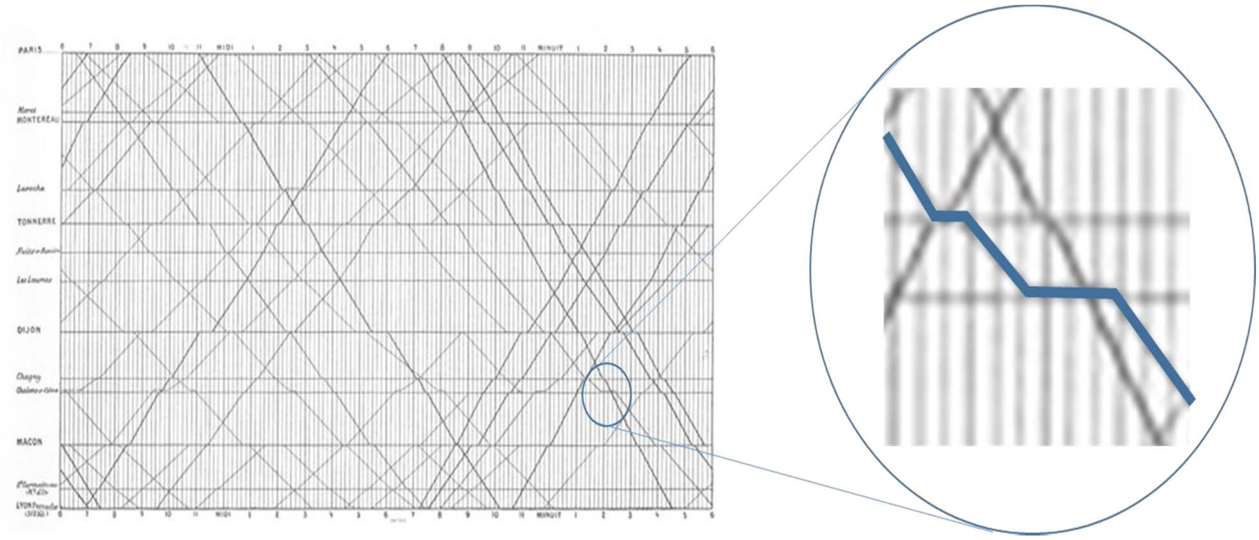


Figure 2: A Railway System, illustrating the complexity of a global schedule

The flat portions of the schedule that are enlarged on the right of figure 2 (coloured in blue) illustrate that, in order to avoid collisions, a train may be temporarily stopped on a garage way, until the pre-determined time when it can safely leave the station. The overall travel experience may not be as fast as a *hot potato* [i.53] model whereby the train is expelled as soon as the way is open, but scheduling avoids queues and results in a much more predictable experience for the passengers.

Likewise, a deterministic flow may experience short buffering in intermediate hops to guarantee that there are never two deterministic packets scheduled for transmission at overlapping times, in other words that the link is always available at the scheduled time of transmission; because of the extra fixed latency that is added at each hop, most deterministic flows will incur a higher latency than with best effort hot potato forwarding.

5.3 The bus analogy (to deterministic circuit switching)

The goal of the reserved bus lanes in a city is to avoid that the bus is delayed by traffic jams and guarantee a good and repeatable experience for the passengers.

Say a bus takes thirty minutes to travel from a given bus stop down the street from a user's home, all the way to his office. If there is a bus every ten minutes, then the transit time for a particular user of the public service will always be somewhere between thirty and forty minutes, depending on how lucky he is on a particular day, and this, regardless of the traffic in the car lanes.

This user will certainly not use every bus that passes by, but he knows that there will be one soon enough, and that, forty minutes later at worst, he will be at his desk.

As it goes, a single bus line may not take our user from his departure point to his final destination, and he may have to hop in another bus midway; the additional transport time for him yields huge benefits in reduction of operational complexity and thus an increased profitability for the bus company.

A bus line is analogue to the reserved circuits that can be scheduled in a packet-switched Deterministic Network, and as illustrated in figure 3, an overall bus transport fabric is quite alike a complex deterministic switching system. Regardless of the load, a Deterministic Network will guarantee a periodic transmission opportunity with a bounded latency for asynchronous commands and alerts.

In that context, the ultimate realization of a Deterministic Network appears to be a perfect emulation of the good old serial cable over a packet-switched network, transporting the exact same application, at a much lower cost than a full mesh of wires.

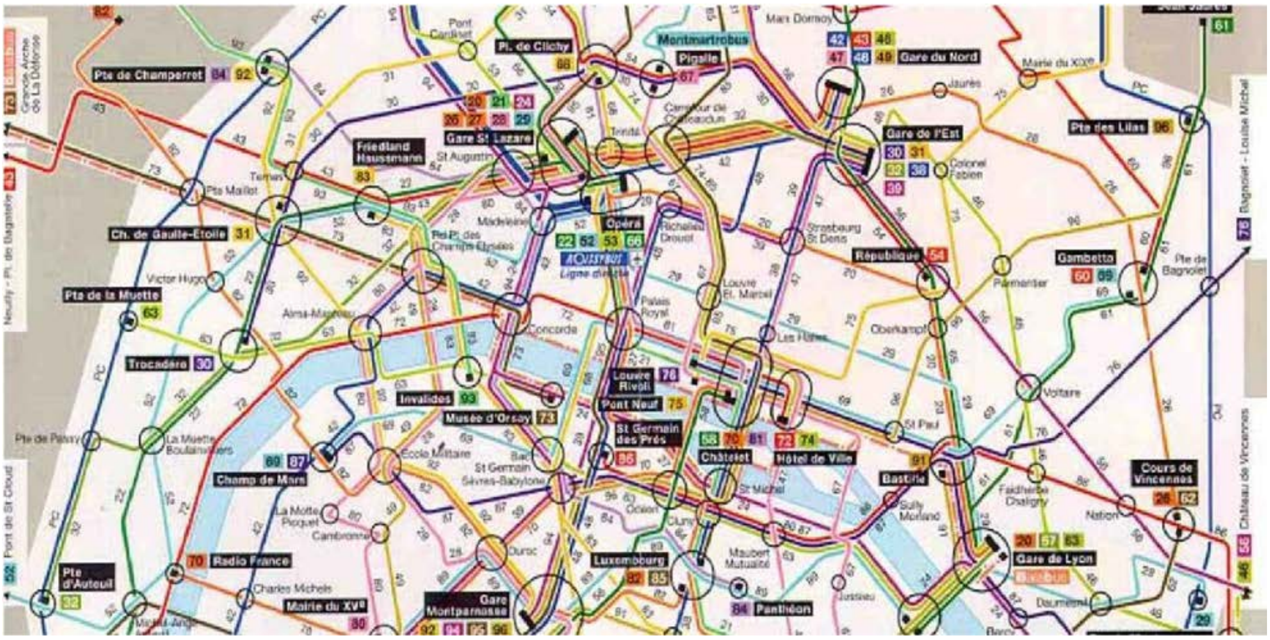


Figure 3: The bus transport fabric

5.4 The vacation place analogy (to time-sharing)

For those of us who do not have 52 weeks of vacation per year, clever marketing has invented this concept of time sharing acquisition whereby an individual owns a flat in a nice vacation resort for just one week a year. That week, the individual will use the flat as if it was always his, will not find unexpected people in his bed at night, and, provided that all owners are decent enough, he will have no clue whatsoever of how many other owners also live in that flat at other times along the year.

As seen in clause 5.2 that determinism requires that packets from a particular flow may need to be held in an intermediate node until the scheduled time when the free access to the transmission medium is guaranteed. This means that enough buffers have to be available during that period to hold the packet; by analogy to the vacation place, there cannot be an unexpected occupant in the master bed.

The object of Deterministic Networking, which is to remove chance from the picture, requires tying physical resources to the protected flows. With scheduled operation, a same resource may be affected to different flows at different times, and the duration of this affectation, thus the number of deterministic flows that can make use of a same resource, is directly affected by the precision of the shared sense of time in the network.

It results that the reservation system that locks physical resource to well-dimensioned and identified flows have to be aware of the device capabilities such as clock precision and forwarding latency, as well as the amount of buffers, timers and queues that are available to control in that particular device.

With appropriate shaping, a deterministic flow is fully isolated from any influence from other traffic, with no leak, no loss, and no latency whatsoever that could be imputed to other flows, whatever the load on the network. This isolation goes beyond that provided by a Virtual LAN (VLAN) [i.93] and its Layer 3 equivalent, the Virtual Routing and Forwarding (VRF) [i.94]; those only prevent leak and eavesdropping, but cannot protect against congestion loss and latency induced by some other traffic that happens to share the same physical resources.

In other words, Deterministic Networking brings a new level of isolation and guarantees that are critical to converge OT control flows onto a shared IT infrastructure spanning the campus or the factory. To realize this to its full extension, a strict policing and shaping have to be performed that filters out misbehaving devices, whether it is an external attack or a failing network node that may, for instance, repeat the last frame forever at line speed. In turn, the isolation brings a new form of security whereby attempts to influence a flow by injecting another in the same physical infrastructure becomes totally inefficient.

5.5 The casino analogy (to statistical effects)

The law of large numbers says that long term, the casino wins. Yet, though it rapidly becomes statistically improbable to keep winning, there is no bound to the potential gains that a player may achieve. By analogy, statistical multiplexing and arguably buffer bloat yields unbounded latency, or loss.

In wireless, CSMA with collision avoidance (CA) operations detect a collision and apply a randomized back off to avoid that it happens again. Though the chances decrease exponentially, there is always a possibility that a frame again for a theoretically unbounded number of times [i.67].

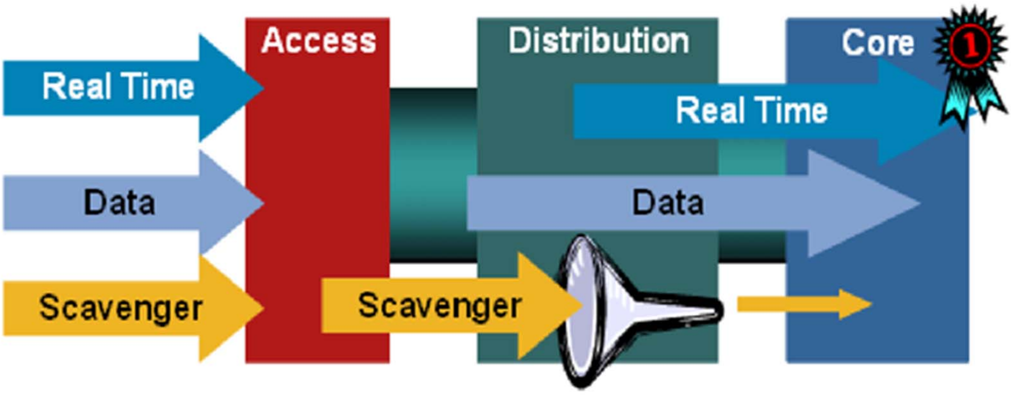


Figure 4: traditional Quality of Service

Similarly, the hot potato [i.53] forwarding operation in a switch or a router entails a statistical chance of too many flows outgoing a certain link, queues forming and eventually packets being dropped. The imperfect experience of voice over IP generally results from this sort of effect, in particular over Wi-Fi where the data rate may vary considerably and Call Admission Control (CAC) [i.16] is very complex to achieve. The only protection in classical routers and switches is to drastically under-utilize the medium, with a very limited amount of critical traffic of high Quality of Service (QoS) [i.95], as illustrated in figure 4.

Ever increasing the bandwidth can make it so that statistical effects such as congestion loss and out-of-bound latency become negligible; and in addition, appropriate QoS settings enable to push the unwanted statistical effects onto vulgar packets, and protect relatively low rates of critical flows. This has been the recurrent answer from packet-switched statistically multiplexed networks against the better-groomed behaviour of transports designed for circuit mode communications, such Synchronous Optical Networking (SONET) and its Synchronous Digital Hierarchy (SDH) variant, or T1 Time-Division Multiplexed (T1-TDM) and its Integrated Services Digital Network (ISDN) variant.

So far, this has been a winning strategy for most use-cases, but it falls short when the needs of the critical flows approach the effective throughput of the wires, e.g. for large time-sensitive flows of uncompressed video, where tens of Gbps have to be delivered at a precise time across a switched fabric in professional studio.

This is why until now, the extreme OT uses cases such as industrial motion control, which operate control loops at several hundred times per second, had to resort to specialized standards or proprietary designs, as detailed in clause 4.3.

In contrast, scheduled operations can utilize the medium for critical flows up to high loads, limited only by the precision of the clocks and the compute capability of the central controller. Voice calls and industrial control loops are like trains, the traffic is predetermined and with adequate dimensioning and scheduling, the network can be programmed to ensure a seamless and repeatable experience. This requires new shapers, a fine sense of time, and a better control of individual physical resources such as buffers in memory.

A deterministic flow has to traverse the network in the same predictable fashion every time, regardless of the load of the network. There can be no observable influence whatsoever from any other flows. A Deterministic Network may be primarily loaded with deterministic flows and still maintain its predictable properties unmodified for each of them.

This goes beyond the capabilities of the current QoS-based Internet technologies, where an increased load in one flow generates delays and losses, which are rapidly observable, on adjacent flows.

5.6 Transporting OT traffic

Transporting OT traffic requires a network that is reliable and jitter-free, and can be trusted to transport periodic and asynchronous commands with a bounded latency. An ever-increasing bandwidth is not always a valid response in the face of high loads of time sensitive flows found in OT applications. This strategy was not accepted by the industrial community, which developed its own adaptions of Ethernet and IP to meet its specific needs. As opposed to QoS-based networks, a Deterministic Network has to always retain its properties, even under high loads of critical flows.

To achieve this, it is necessary to schedule the timely operation of the network with the granularity of numerous tiny physical resources. While scheduling enables to provide the required guarantees, it may also yield an additional latency for the transported flows. A precise synchronization of the network has to be obtained, so the required resources can be reserved for the individual packets within that particular flow at their precise time of arrival. Over the distance, tiny flows have to be aggregated into larger ones that can be processed as a single entity.

With Deterministic Networks, the worst-case data loss and latency can be guaranteed in a consistent fashion when multiple services are deployed on a common converged network infrastructure. A deterministic flow is completely isolated in its own time-shared set of physical resources and cannot be influenced by any other traffic in any observable fashion; from the perspective of the application, a deterministic end-to-end connection appears as a dedicated point-to-point wire.

As Deterministic Networking capabilities are deployed, wired and wireless links that are today limited to provide Internet connectivity can be used to replace any of the cables that are found in day-to-day applications, RCA connectors on the stereo gear, RS232 serial cables, fancy bus connectors in cars, 2 and 4-wires cables in industrial control networks, all of them. A shared sense of time enables to transport -in fact to recreate- a clock signal that is used to synchronize both ends of an emulated serial cable. Moreover, the perfect isolation between flows that is obtained to guarantee the required latency is also an improvement on the security side, and a factor of trust in the convergence for OT people.

Awareness of Deterministic Networking is now spreading [i.96], and the new technology is generally perceived as an evolution of quality of service to bring a new level of guarantees for network-based services. True as it is, what Deterministic also and mostly brings to Networking is more revolutionary; it is the capability, for the first time, to carry any signal that was ever transported across a point-to-point wire, for any form of application that men ever devised, over multi-hop packet networks.

6 Enabling Determinism in a Network

6.1 The precursors

6.1.1 On Fast Reroute

Use-cases where two non-congruent paths are set up to ensure either a full redundancy, or at least a rapid fail-over, by selecting at the ingress between two pre-computed paths, abound in the art of networking. An example of such, the Linear Protection [i.39] of the MPLS Transport Profile [i.18] (MPLS-TP) [i.4] uses specific OAM frames [i.1] to monitor the liveliness of the routes and make rapid fail-over decisions.

To avoid the complexity of setting up non-congruent paths, and in order to react even quicker from where the problem actually happens, the IETF has devised the concept of fast-reroute [i.47], which means rerouting around a failure from anywhere inside the network.

The IETF has proposed two approaches for IP fast-reroute, the original IP LoopFree Alternates (LFA) [i.66] and [i.7], which attempts to find a path around a failing node or a failing link, with a variable coverage, and a new technique for IP Label Distribution Protocol (LDP) using Maximally Redundant Trees (MRT) [i.14], which draws a pair of trees in any bi-connected topology, each tree connecting every node and both rooted at the destination, in such a fashion that a breakage only blocks one of the trees, so that the other tree offers a path to destination.

One key benefit that is found in both techniques is to be compatible with, and actually leverage, an existing OSPF-based infrastructure; and one limit to both approaches is the lack of control on the generated alternate paths, in particular with respect to their relative costs. The detour may be long and expensive, in particular with MRT, which computes global trees.

6.1.2 On SDN and Traffic Engineering

Software Defined Networking (SDN) [i.36] promotes a model that simplifies the network operation by automating the deployment of network resources from a centralized controller; at the extreme, the networking gear becomes a basic subservient to the master controller in practice, though, a distributed routing plane may still be associated with the SDN operation so as to reach the controller and handle the general purpose bulk of the traffic. Though still highly overhyped [i.97], the momentum started with Openflow [i.31] is getting traction in campuses and cloud data centres through various vendor incarnations such as the Network Element Drivers [i.98] of Cisco's Network Services Orchestrator and Arista's Directflow [i.99].

The art of Traffic Engineering (TE) at the IETF leverages routing protocol extensions to link-state Interior Gateway Protocols (IGP) such as OSPF [i.65] and [i.3] and IS-IS [i.5], as well as extensions to BGP [i.40], to report the topological information to a central routing component, which is implemented by a Path Computation Element (PCE) [i.17]. Relying on that topological information learned, the PCE computes diverse paths and assign flows to those paths. But this technique yields a lot of human intervention and does not yet support deterministic properties.

More closely related to OT applications of IOT and low-power radios, Industrial Routing Requirements in Low-Power and Lossy Networks [i.43] discusses the need for the LLN routing protocol to compute multiple forwarding solutions. The centralized approach in TE is echoed by the best practice found in the art of industrial and vehicular networks, operating Time-Triggered Ethernet [i.25], IEC 62591 [i.10] and IEC 62734 [i.11]. The limit of the model is probably the scalability of the controller, considering that the optimization of a Time-Triggered schedule is an NP-complete [i.9] and [i.38] problem with a complexity that grows rapidly with the number of engineered flows.

6.2 Expected benefits in wired networks

A perfectly Deterministic Network would ensure that every packet would reach its destination, and would always do so in due time. In an imperfect world, Deterministic Networking nearly eliminates packet loss, with the associated goal to guarantee a worst case latency for a packet, and all this whatever the overall network conditions are, effectively emulating the case of a point to point serial cable.

The main cause of data loss in a wired switch fabric is a statistical effect called congestion loss, whereby, at a particular moment, multiple flows entering a switch converge to a same outgoing port, in volumes that exceed the capacity of that port to output the traffic. It results that some of the traffic has to be dropped, and one of the desired effects of Quality of Service (QoS) is to intelligently select the frames to be discarded.

Making networks more deterministic eliminates this statistical effect by maintaining at all time the amount of critical packets within the physical capabilities of the hardware. This can be achieved by the use of time-shared resources (bandwidth and buffers) per circuit, and/or by shaping and/or scheduling the packets at every hop.

Equipment failure, such as a switch rebooting, a broken interface adapter, or an unplugged physical wire, is a secondary source of data loss. When a breakage occurs, multiple packets are lost in a row before the flows are rerouted or the system may recover. This is not acceptable for critical applications such as related to safety. A typical process control loop will tolerate an occasional packet loss, but a loss of several packets in a row will cause an emergency stop (that is, typically after 4 packets lost, within a period of 1 second).

Making networks more deterministic improves the resiliency against breakages and statistical transmission loss such as due to cosmic particles, typically by adding redundancy in the network path.

Finally, since the operation of a Deterministic Network rely on precisely applying a tight schedule, and the worst-case time of delivery has to be guaranteed, a shared sense of time is propagated throughout the network, which can be exposed to and leveraged by other applications.

6.3 Making Ethernet deterministic?

On precise time: A synchronization of clocks in the order of the microsecond can be achieved by a software implementation using the Precision Time Protocol (PTP) [i.32] IEEE 1588 [i.112], or one of its derived profiles, such as IEEE 802.1AS [i.113]. With hardware assistance, this can be brought down to tens of nanoseconds for use in Smartgrid [i.23] and even down to nanoseconds for military [i.28] and 4/5G so-called fronthaul and crosshaul [i.37] applications.

With a precise shared sense of time, the switches can guarantee the exact forwarding time at each hop for TSN applications, and in-time delivery at the egress for AVB applications. Precise time is leveraged by Time-Triggered Ethernet (TTE) [i.25] to control the exact time of transmission; with TTE, the sender and the listener can agree that a certain packet belongs to a certain deterministic flow just because it is transmitted at a certain precise time, without the need to tag the packet. Keeping the Ethernet frames unmodified enables inter-working between TTE and classical Ethernet switches.

On timely transmission: Another key aspect to provide deterministic latency guarantees is to ensure that the medium is free and can be accessed with no wait when the time comes for sending a deterministic packet.

In the case of full-duplex switched Ethernet, a sender owns the transmit medium, so there is no need to defend against an interfering transmission from other parties; a collision may be avoided either by preventing transmission ahead of the scheduled time by a guard time that is more than the transmission time of the largest frame, or by suspending a frame being transmitted to free the medium for the deterministic packet, which is the solution that the IEEE as selected with 802.1Qbu [i.114] Frame Pre-emption.

By construction of the schedule, only one frame may be programmed during any particular window of time sized for one frame of maximum size, and there can never be a contention between deterministic frames.

If a non-deterministic frame is being transmitted at the precise time scheduled for a deterministic frame, its transmission is interrupted and a CRC is attached to validate the partial transmission. The deterministic packet can then be transmitted in time and in full, and then the non-deterministic transmission can resume starting at the offset where the transmission was interrupted, yielding no access latency for the deterministic frame.

On redundancy: A Deterministic Network brings resiliency against physical and logical failures and guarantee the continuity of operations in all conditions; this requires the physical redundancy of each involved piece of networking equipment, and the capability to compute non-congruent paths between source and destination and leverage Frame Replication and Elimination techniques.

High-availability Seamless Redundancy [i.24] (HSR) for that purpose. The International Electrotechnical Commission (IEC) further standardized those methods [i.9] for both the industrial [i.2] and Smart grid [i.52] applications, and is now evolving the IEC 61850 [i.13] standard for substation communication networks to include them [i.19].

On reliability: Finally, critical applications such as professional Audio/Video Bridging (AVB) [i.100] demand a packet delivery ratio (PDR) that orders of magnitude better than the capabilities of a simple chain of switches along a path. Additional diversity, such as replication and elimination of a packet over non-congruent paths, enables to reach the required figures. The math is simple: If the probability of loss along a path is P, then, if total diversity and independence can be achieved between two parallel non-congruent paths, then the probability of loss with Replication and Elimination over those two paths becomes P^2. With a loss ratio of an Ethernet fabric that is typically in the order of 10^{-5}, sending redundant copies over two fully diverse paths yields a 10^{-10} loss ratio, which enough to burn a Blu-ray disk without a scratch in an AVB studio. When applied to highly lossy multi-hop media such as a wireless link with, a worse than 10 % packet error rate (PER) without retries is not uncommon [i.49]. A simple redundancy, whether it is based on rings or parallel paths, still yields at best a PER of 1 % or worse, which is far from the 10^{-5} wire equivalence that industrial networks are after.

6.4 Making wireless deterministic?

On transmission reliability: The radio medium, in particular in the industrial, scientific, and medical (ISM) bands that are shared not only between data networks but also with all sorts of interferers such as microwave ovens and radars, is orders of magnitude less reliable than classical wired networks such as Ethernet over copper or fibre; providing deterministic services over wireless appears extremely challenging.

Co-channel interferences are not the only possible causes for a frame loss; for instance, physical obstacles may happen to move in the way of the transmission and block the communication. Multi-Path Fading, which is due to multiple reflections that may reinforce or cancel one another out a few centimetres away, is a major cause of transmission errors; it does not affect all channels in a same fashion, and which channels are impacted is highly sensitive to the relative position of the sender and the receiver, and to their environment.

In short, the quality of a given channel is affected by multiple parameters that may vary brutally, and IEEE 802.15.4 transmissions over a fixed channel cannot be expected to remain stable over a long period of time. It results that basic IEEE 802.15.4 [i.55] implementations that operate on a fixed channel will often suffer from intermittent delivery issues, and can only apply to low-end applications for which a consistent reliability is not a concern.

Channel Agility was added to improve the availability of the radio links by permanently sensing the channel, looking for an increase in Bit Error Rate (BER), and switching channel when loss becomes too high. This technique represents a clear improvement from the art of fixed channel, but it only fixes an error condition reactively, and a period of lossy transmissions is experienced before a new channel is selected. This is well suited for applications such as AMI/AMR metering, which do not require a wire equivalent reliability, but, at the same time, this does not provide the deterministic guarantees that industrial applications require.

A predictive technique that would enable to switch to a better channel before the problem even occurs would be ideal but early attempts [i.22] did not make it to mainstream. Effectively, one can leverage transmission statistics, observe activity on other channels and remember channels with a bad transmission record, so as to black list them. This may be efficient to protect against a stable Wi-Fi co-channel interferer, but there is no way to know if this is effectively the case, and if the problem will last and for how long. The physical phenomenon that is the most common cause of channel degradation, moving an object or starting a process, cannot be fully predicted by the radio device in most practical situations, and making Channel Agility proactive has appeared so far to be a red herring, yielding more complexity than actual benefits.

Time-Slotted Channel Hopping (TSCH) [i.63] is the best of breed with its simple per-packet channel rotation called channel hopping. TSCH brings Time (by repeating failed transmissions) and Frequency (by switching channel) diversities.

But unless there is an infinite time to attempt an infinite number of retries, there is always a limit to the reliability of a system. In short, deterministic delivery within a constrained time cannot be guaranteed over a wireless medium. A realistic goal is to optimize the delivery ratio, and this is best achieved by combining all possible forms of diversity.

On deterministic channel access: The other major difficulty with wireless links is that the medium is shared, meaning that not only this node, but other nodes, may be transmitting at the time intended for a deterministic packet, which would have to wait a variable time till the end of that current transmission. Then, with CSMA/CA [i.8] mechanisms, it would have to introduce a random delay to obtain access to the channel, and then, in case of a collision, wait for an additional exponential back off time.

In wireless, a technique like the IEEE 802.1Qbu [i.114] - Frame Pre-emption - is not really workable, because apart from new specific attempts for full duplex-radios, there is no way to interrupt a remote talker that is blinded by its own signal.

Depending on the specific radio technology, packets may be blocked at the MAC or PHY layer and then sent together in a single transmission opportunity; data rates may vary dramatically with the distance and the environment; it results that the exact duration of the current transmission can be very hard to predict, and that a guard time that would guarantee that the medium is empty at scheduled time (see clause 6.3) on a wireless medium would necessarily be very long and highly wasteful in terms of bandwidth to cope with highly variable transmission duration.

The only way to approach determinism in wireless transmissions -most people will prefer to use terms such as "highly predictable" instead of deterministic- is to schedule all the transmissions, and that is what TSCH and industrial WSNs do.

A schedule such as illustrated with coloured codes in figure 5 controls at which time and on which channel a frame is forwarded between which pair of nodes.

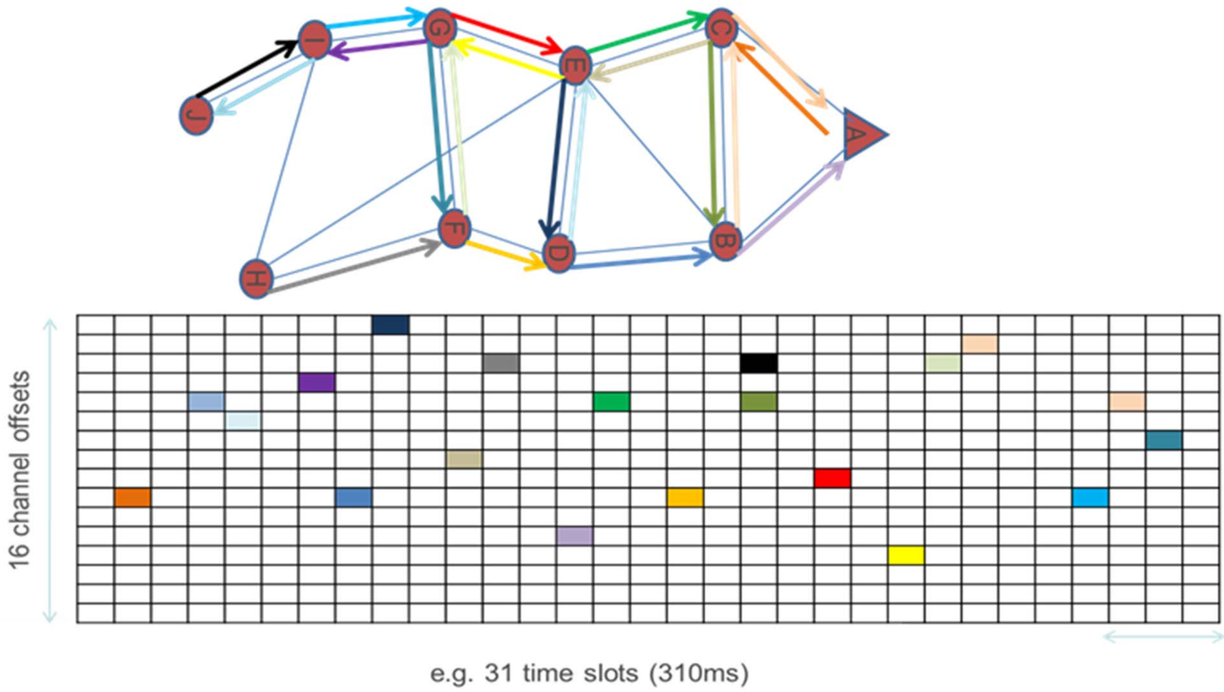


Figure 5: A TSCH schedule

Scheduling implies that all nodes control the precise time of emission, which in turn requires a shared and precise sense of time. It is, for instance, possible to achieve in the order of tens of microseconds clock synchronization over an IEEE 802.15.4 link [i.55].

With fully scheduled operations, it is now possible to guarantee the time of delivery for those packets that make it to destination, in a deterministic fashion.

Additional Benefits from scheduling in wireless: In addition to the benefits listed in clause 6.2, scheduling provides specific value to the wireless medium. On the one hand, scheduling reduces transmission losses and with TSCH and its industrial derivatives, a wire-equivalent loss ratio of 10^{-5} can be obtained: to achieve this, routes are computed so as to enable at least two forwarding solutions for every node, which ensures spatial diversity, whereas retries over the TSCH MAC provide both time and frequency diversity and effectively combats co-channel interference as well as multipath fading.

On the other hand, scheduling optimizes the bandwidth usage: compared to CSMA/CA operations, there is no blank related to IFS and exponential back off in scheduled operations, though some minimal Clear Channel Assessment may be needed to comply with the local regulations such as ETSI EN 300 328 [i.115]. And because TSCH time slots provide a full time sharing operation, there is no limit to the ratio of guaranteed critical traffic.

Finally, scheduling plays a critical role to save energy; in IoT, energy is the foremost concern, and synchronizing sender and listener enables to maintain them in deep sleep at all times when there is no scheduled transmission; this optimizes sleeping periods by avoiding idle listening and long preambles; TSCH enables battery operated nodes that actually forward packets in a mesh topology for multiple years [i.63].

In a nutshell: While scheduling transmissions can guarantee the time of delivery, it is impossible, in the ISM band, to keep all possible interferers at bay. Co-channel interference, as well as the self-inflicted Multi-Path fading, which is due to echoes of the transmission, are unavoidable. In other words, there is no way to guarantee the delivery of all frames.

It takes different mitigation techniques to avoid the different issues that affect wireless transmissions. To combat them all, all possible forms of diversity should be leveraged, in the spatial domain by routing over multipath, in the temporal domain by retrying transmissions or sending copies over parallel paths at distinct times, and in the frequency domain with frequency hopping (within frames) or channel hopping (between frames).

Several times along the way to the destination, the work presented in the present document replicates and then eliminates the copies of a packet that are forwarded along parallel disjoint paths, hopping between frequencies with each transmission so that a copy along one path does not interfere with a copy along the other.

7 The IETF DetNet architecture

7.1 Positioning of work

Deterministic Networking refers to the allocation of pre-determined physical resources in the network (queues, buffers, transmission medium) for well-characterized flows that are known a-priori, in order to avoid the statistical effects that lead to a poor bandwidth utilization, uncontrolled jitter, and congestion loss. Bandwidth that is not actually used by deterministic flows is available for use by other traffic.

Figure 6: DetNet Components

In more details, a path is nailed down with a particular set of resources at particular times, and the forwarding behaviour ensures that the right packets are forwarded at the right time to make use of these resources.

As illustrated in figure 6, this requires capabilities for:

- Hop-by-hop synchronization, which enables to restore an apparent end-to-end clock signal that can be associated with the timely distribution of data streams, which in turn can be decoded to reproduce precisely phased analogue signals;
- Time-based resource reservation coupled with enforcement methods, which eliminates collision loss and provides the capability to transport unicast or multicast data streams of predefined characteristics such as data rate and bounded latency.

Work has started at the IETF DetNet WG [i.109] to enable the establishment and maintenance of deterministic paths over Layer 2 Bridged and Layer 3 routed segments, by defining a common representation (data models) of the physical resources and network topology, and standardizing the protocol flows and interface mappings that are used to set up the flows.

Work will also be needed for management, with in particular network control frames that are carried along a deterministic path to assess its health and performance - aka Operations, Administration, and Maintenance (OAM) frames, but also in-band means to trace a packet from the last replication point from which it may have strayed into the network.

7.2 The architecture in a nutshell

The Deterministic Networking Architecture [i.57] applies a centralized approach for a limited amount of deterministic flows, which share the network with a more classical distributed path computation and statistical multiplexing operation for traffic of lesser criticality and requirements.

In that approach, diverse applications can push their requirements over a Northbound Interface to a centralized controller, which translates these requirements in term of complex routes with replication and elimination capabilities, and pushes the result onto the network over a Southbound Interface, including precise operating schedules and time-based resource reservations.

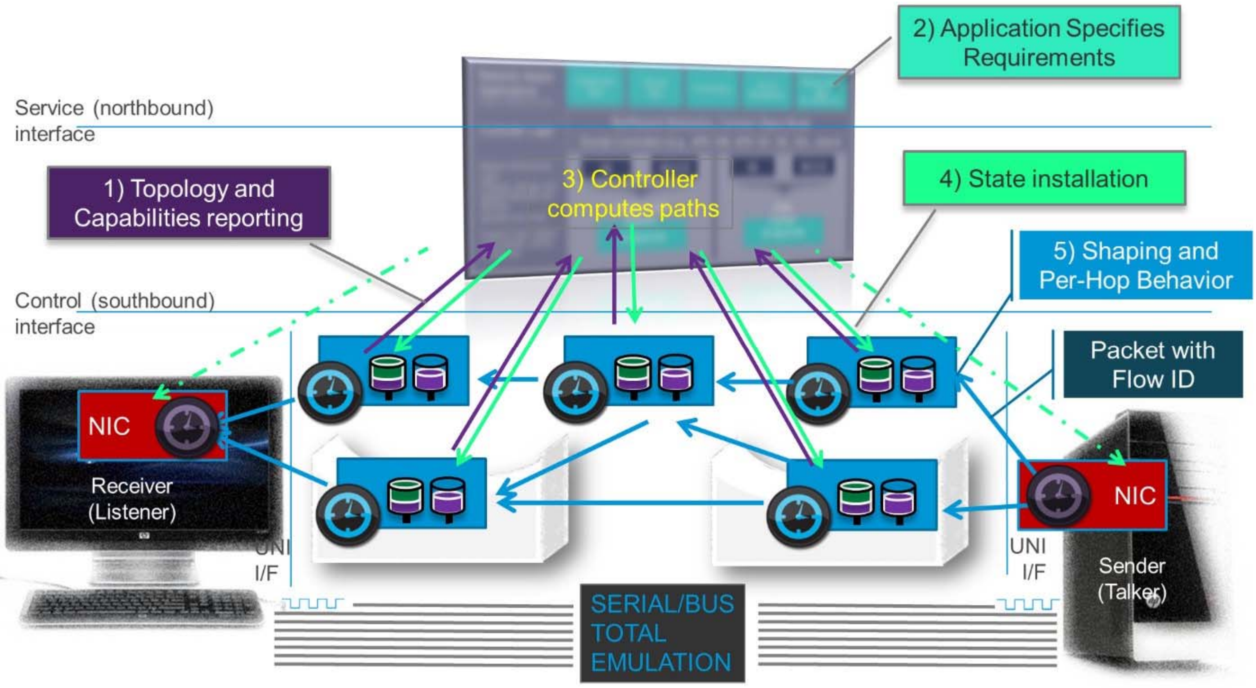


Figure 7: The DetNet Architecture

Figure 7 illustrates the steps that take place in order to setup a deterministic path (in blue), and the interfaces (in purple and green) and data models that DetNet should standardize. Following the numbers on that figure, the main steps are:

1) the definition of data models to report the topology and the devices capabilities to the controller which is aware of the application requirements and can perform;

2) the computation of a path that matches those needs;

3) the protocol elements to request a path set up for a given flow and configure;

4) the Network Interface Card (NIC) in the end nodes, and the time-shared reservation of physical resources in the network nodes along the end-to-end path; and

5) the forwarding behaviour for each flow.

In short, the SDN model comprises three layers, the application, which is not covered cover in the present document, the network, which loses some of its distributed features, and the controller that seats in between and implements the needs of the former using the capabilities of the latter.

Clauses 7.3 and 7.4 explore the DetNet abstractions for the network and the controller.

7.3 Networking in DetNet

The art of L2 with AVB, or of L3 with IntServ (RSVP), is already capable to establish a serial path in a distributed fashion as depicted in figure 8, with end-to-end latency guarantees [i.48] at the expense of wasted bandwidth.

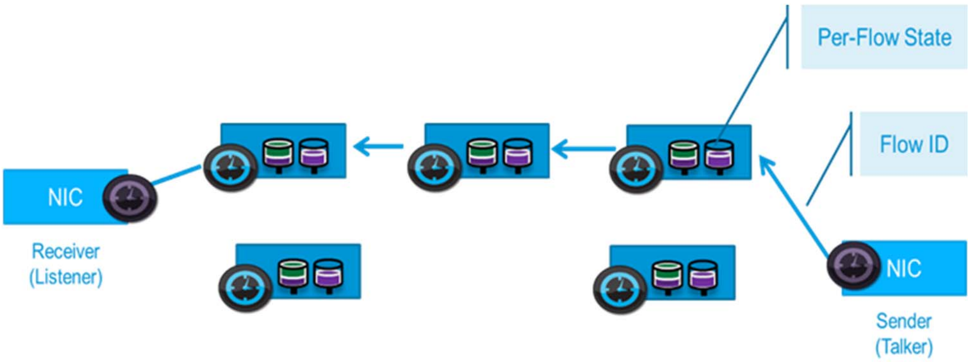


Figure 8: Single Path

So arguably, DetNet could use a distributed architecture as well.

However, to reach a higher reliability, DetNet adds Replication, Retry (wireless) and Elimination of individual packets to the problem, making the path more complex and more global, as illustrated in figure 9.

The complexity of the DetNet path makes solving the problem of minimizing the overall latency and the amount of resources even more difficult. It was determined from the art of Traffic Engineering that optimally distributing a number of known flows across a shared network is an NP-Complete problem and that the most advanced optimization will be obtained by leveraging a central PCE with a global view of the network.

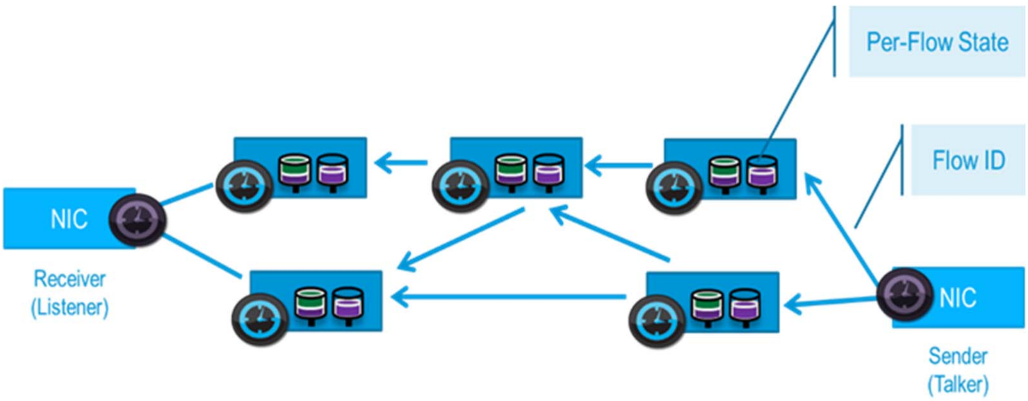


Figure 9: Complex Track

This is why, as opposed to IntServ, the DetNet architecture relies on a controller with a total view of the system capabilities and application requirements to optimize globally the allocation of resources and install the related state in a transactional fashion.

Figure 10 represents a classical view of a network as perceived by a networking engineer, with switches that forward frames at Layer 2, routers that route packets at Layer 3, over both wired and wireless connectivity, with all sort of medium between routers. From the perspective of that engineer, the elements of concern are means of logical connectivity, e.g. Virtual LANs, and communication layers, L3 routing vs. L2 switching.

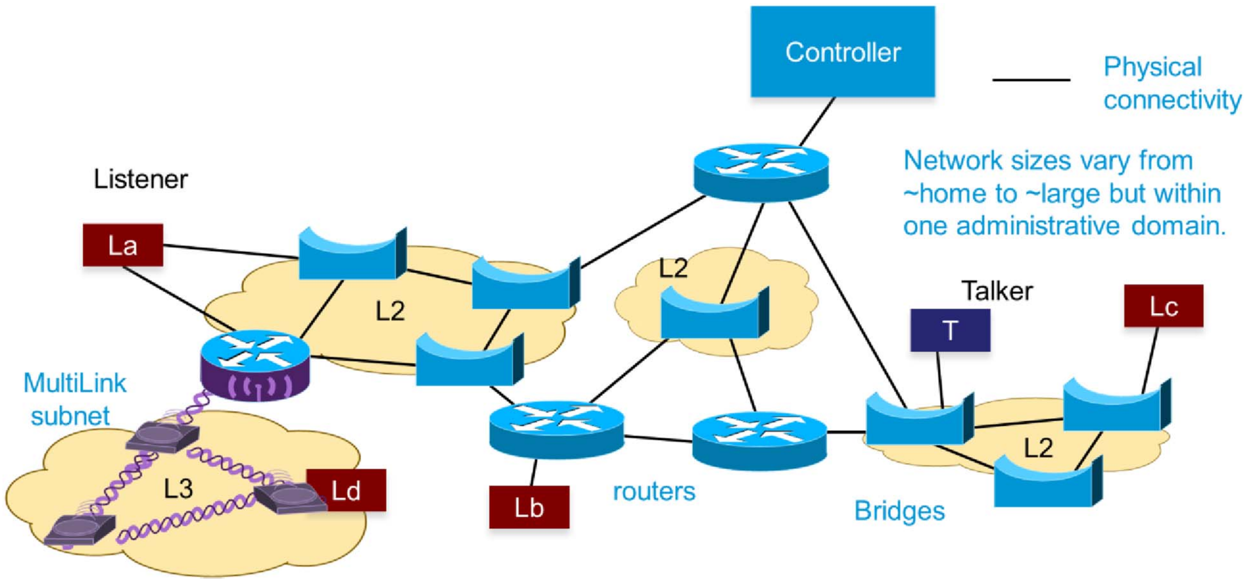


Figure 10: Perspective of a networking engineer (Source N. Finn)

In contrast, figure 11 represents the view of a Deterministic Networking Controller.

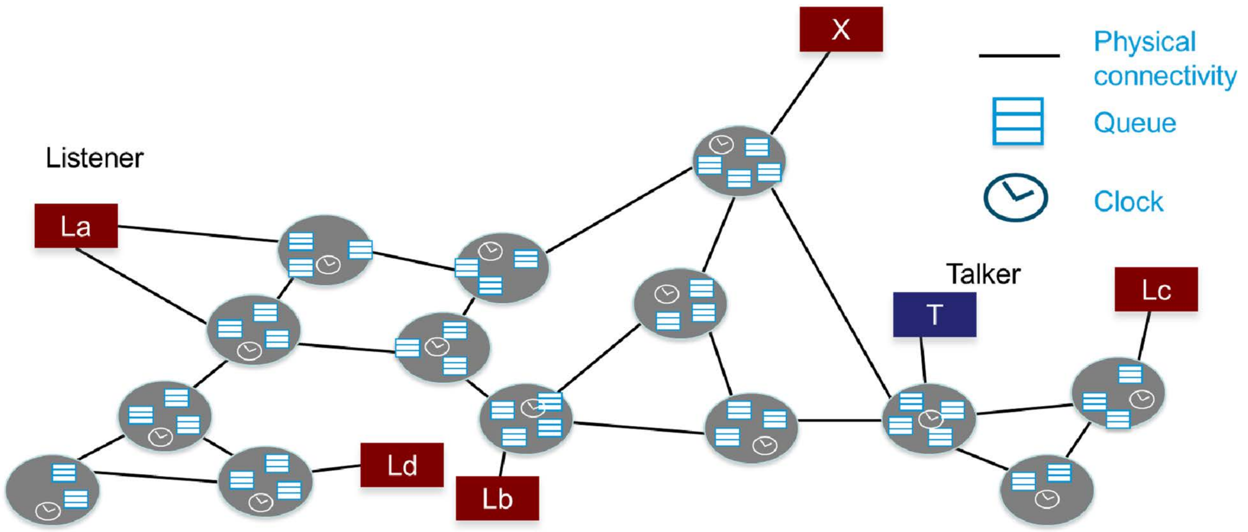


Figure 11: Perspective of the controller (Source N. Finn)

From the Deterministic Networking perspective, the network is a collection of interconnected boxes and wires, regardless of whether they are implementing as Layer 2 or Layer 3 boxes.

What counts is the physical capabilities in those devices to store reliably some amount of packets and resend them at a precise time along a scheduled path.

7.4 Controlling a Deterministic Network

7.4.1 Reporting the topology to the controller

Clause 7.3 presented the simplified view of the network from the DetNet controller perspective. This clause elaborates on the steps that are presented in figure 7, of reporting the topology to the controller, implementing the needs of the application in the network, and automating the network operation.

In order to compute a deterministic path, the controller needs to learn not only the connectivity between the networking pieces of equipment, but also their capabilities down to amounts of buffers and timers that can be reserved, which types of shapers are available on the device, and with which precision the device is synchronized to the rest of the network.

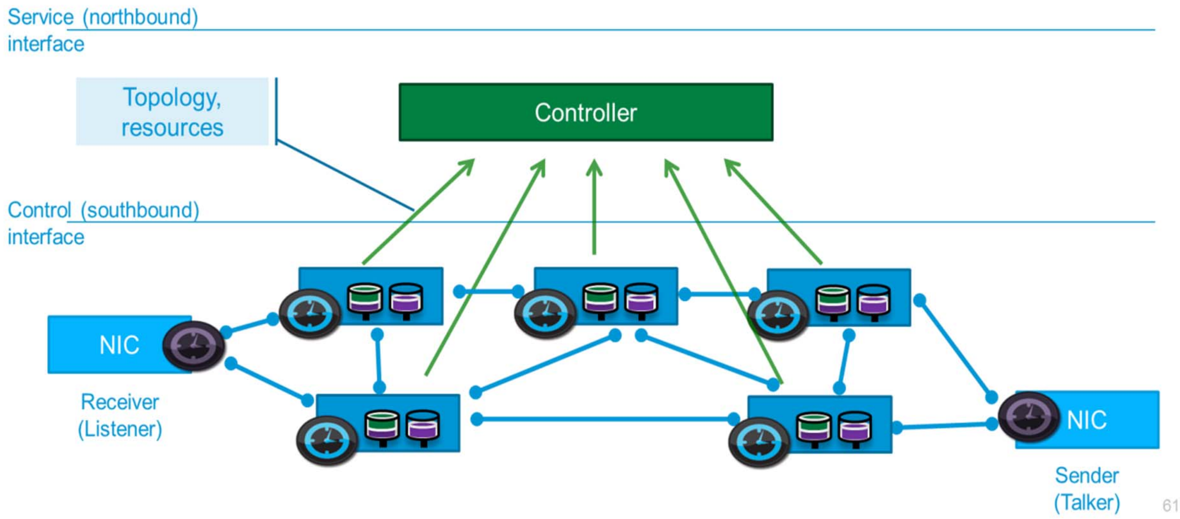


Figure 12: Reporting capabilities and topology to the Controller

The most probable outcome is that DetNet will inherit from work done in the art of TE at the Traffic Engineering Architecture and Signaling (TEAS) working group to report the topology and the device capabilities to the controller over a control interface called the Southbound Interface.

This means that DetNet needs to extend the data models that exist in TE to report, in particular, the physical characteristic that are relevant to the deterministic path computation, which in turn may differ from one medium to another; for instance, IEEE 802.15.4 [i.55] TSCH has a concept of time slots associated to frequency hopping that does not exist in Ethernet.

7.4.2 Implementing the needs of the application

The creation or the modification of a flow is triggered when an application, residing on an end-device, a management console like an Industrial Human Machine Interface (HMI), or a Broadcast Control System in Professional Media Networking (PMN), needs a deterministic communication channel to, say, transport the content of a movie to a Blu-ray disk burner for mastering.

A time sensitive application residing on the end-device (a mobile phone) or some third party hardware (a central HMI) provides a Traffic Specification (TSPEC) that reflects the transport requirements from its perspective, expressed as volumes of data, sensitivity to jitter, loss, and latency. The application makes its request to the controller over a service interface called the Northbound Interface, which is defined by the Open Networking Foundation (ONF) [i.101], providing the TSPEC as parameter.

The controller translates those requirements in terms that are actionable by the network devices. In many cases, the controller will schedule a complex path across the network, which can be seen as an evolved form of a circuit.

The relevant state is pushed in the end and intermediate systems, indicating how to identify a particular flow and the operation to be executed on that flow, like the precise time of forwarding and the behaviour of the shaper, e.g. credit-based vs. time-triggered.

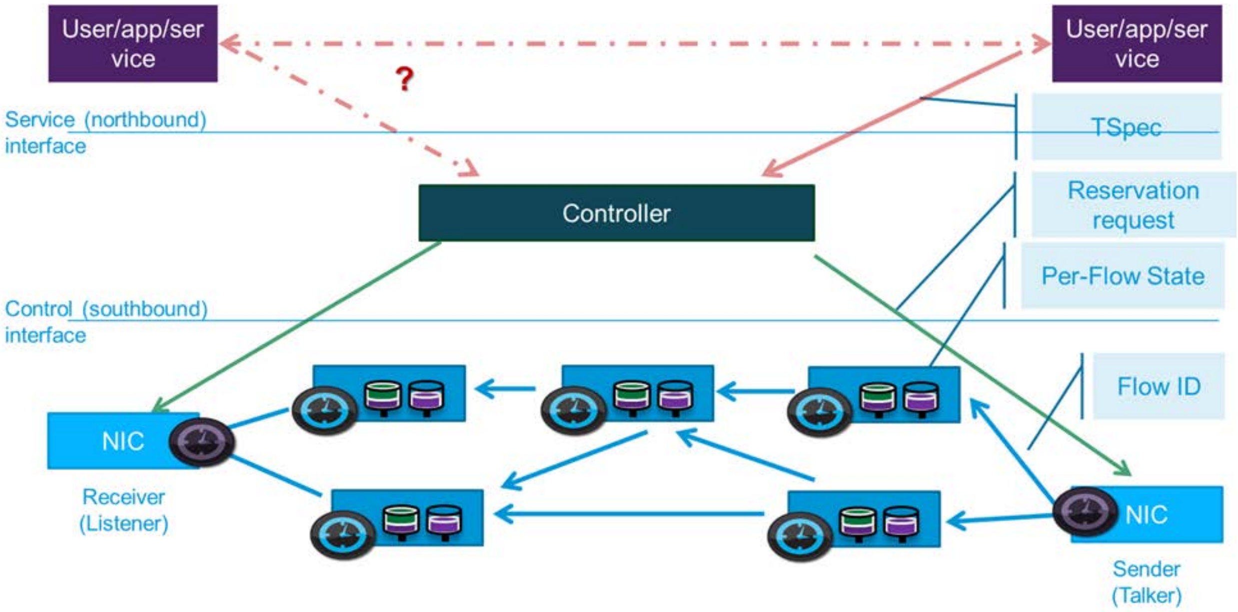


Figure 13: Setting up a new flow

As shown in figure 13, DetNet needs to define the new data models that are required to set up a path that supports new features such as packet Replication, Retry and Elimination, and the definition of tagging elements (i.e. Flow ID, and packet marking) to be used to identify the flow as it is being forwarded along the complex path.

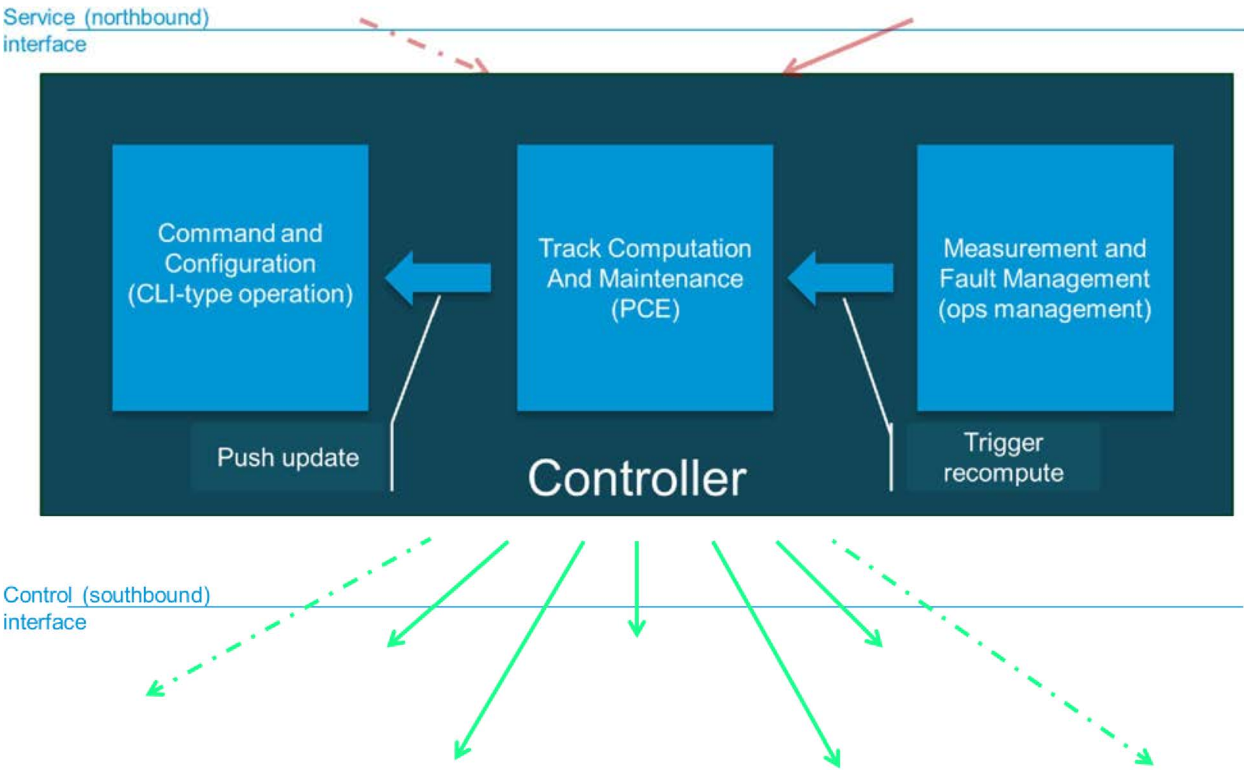


Figure 14: Inside the controller

7.4.3 Automating the network operation

The SDN model used in DetNet enables a degree of automation that help reduce the operational expenses associated to the deterministic network.

To make that happen, the controller participates to a permanent control loop that observes the network to validate whether it still matches user needs, triggers actions when that is not the case, performs an update computation (that is the PCE's piece), and then reprograms the network, which is done by the combination of a driver for the particular network element and a protocol such as the PCE Protocol (PCEP), a REST or a Command Line Interface.

Figure 14 represents the innards of a controller, which effectively is composed of a classical Path Computation Element (PCE) that figures out the optimized layout for the complex paths, a management component that supports the life-cycle of the paths, and Command and Configuration drivers that trigger the required operation over the interfaces that the particular networking hardware support.

This process may be triggered in an open loop, for instance by a user who configures a new device on some HMI, or in a close loop as an automated reaction to an event that is processed in the measurement and fault management entity.

7.5 Limits and perspectives

Deterministic Networking is not a matter of standard bodies and networking layers, (L2 vs. L3); nor is it about virtualization. What it is really about is a tight control of physical operations and scheduling of real buffers and queues, receiving, shaping and resending real packets at precise times with dedicated Hardware.

Such operation requires a network that is precisely synchronized; DetNet will inherit the precise clock synchronization from work done at other bodies, such as but not limited to, GPS, IEEE 802.1AS [i.113], the Precision Time Protocol (PTP) IEEE 1588 [i.112], or the Network Time Protocol (NTP) [i.34]. The precision of the synchronization limits the capability to tightly control the resources and thus the amount of deterministic traffic that can be applied onto the network.

Other limits of the centralized approach are on the one hand the complexity of the computation that hinders the scalability of the solution, and on the other hand the control plane overhead and the delays that are required to install a new path, or modify an existing path, from the controller across the network to all the intermediate nodes.

More of a constraint than a limit, the DetNet operation at Layer 3 will require services from lower layers to achieve the required properties end-to-end; the DetNet group will collaborate with IEEE 802.1TSN, which is responsible for Layer 2 operations, to define a common design to support deterministic applications. A number of abstractions such as the end-host operation should be defined in a fashion that is agnostic to the choice of network used for the connectivity. This common design should ensure that the definition of the southbound and northbound interfaces is kept homogeneous between Layer 2 and Layer 3 so as to enable various inter-working models, such as DetNet transporting TSN on a deterministic pseudo-wire, or a mapping interface between a TSN network and a DetNet network at the edge of a deterministic L2 fabric.

The weakest link along a path will limit the whole chain, and an imperfect mapping at the interconnection between two networks with slightly different operations may induce a reduction of end-to-end results. The interconnection between a wired and a wireless network, in particular, can be expected to yield complex mapping issues.

The following chapters explore how deterministic can be applied on wired and wireless media, and which particular benefits are expected in either case. Determinism can be practically achieved, but through different methods and with different capabilities. The interconnection and the enablement of end-to-end deterministic capabilities are still a complete green field, open for further research.

8 The art of low-power wireless sensor network

8.1 A highly predictable wireless

With TimeSlotted Channel Hopping, scheduling transmissions minimizes the chances of collision and associated loss. Scheduling is achieved through TDM, by slicing time and affecting time slots to particular transmissions. This scheduled mode of operation is particularly adapted to well known, periodic flows for which a schedule can be computed in advance.

TSCH combines TDM with channel agility in order to defeat interferences, in particular Multi-Path Fading, which generally affects 2 to 4 channels out of the 16 available with IEEE Std 802.15.4 [i.80] in the 2,4 GHz band, and co-channel interference, in particular when it is located in a limited number of adjacent channels.

TSCH recognizes that smarts in channel selection do not pay off in practice, and all channels that are not black-listed are equally used; they are tried in a pseudo-random order, hopping between non-adjacent channels at each transmission; and though typically, at any given point of time, several channels present a high BER for a particular pair of devices, a series of retransmissions over a sequence of alternate channels eventually bypasses the issue after a few attempts, and an industrial-class reliability can effectively be achieved.

The complexity in TSCH is elsewhere; it comes from the need to ensure that the receiver is tuned to the same channel as the sender at the precise time of the transmission; this requires additional protocol elements to synchronize the network and schedule the transmissions. A TSCH schedule can be viewed as the program of a mechanical piano that would play the same tune repeatedly, whereby the channels used would be the music. But the analogy stops here, since at the next iteration of a schedule, all transmissions are rotated by a pseudo-random number.

An example is given in figure 15, with 3 iterations of a schedule, and 4 possible transmissions (A → T), (B → A), (D → A) and (C → E) , the first 3 being best effort, to be prioritized in A to avoid scheduling collisions, while the last is protected.

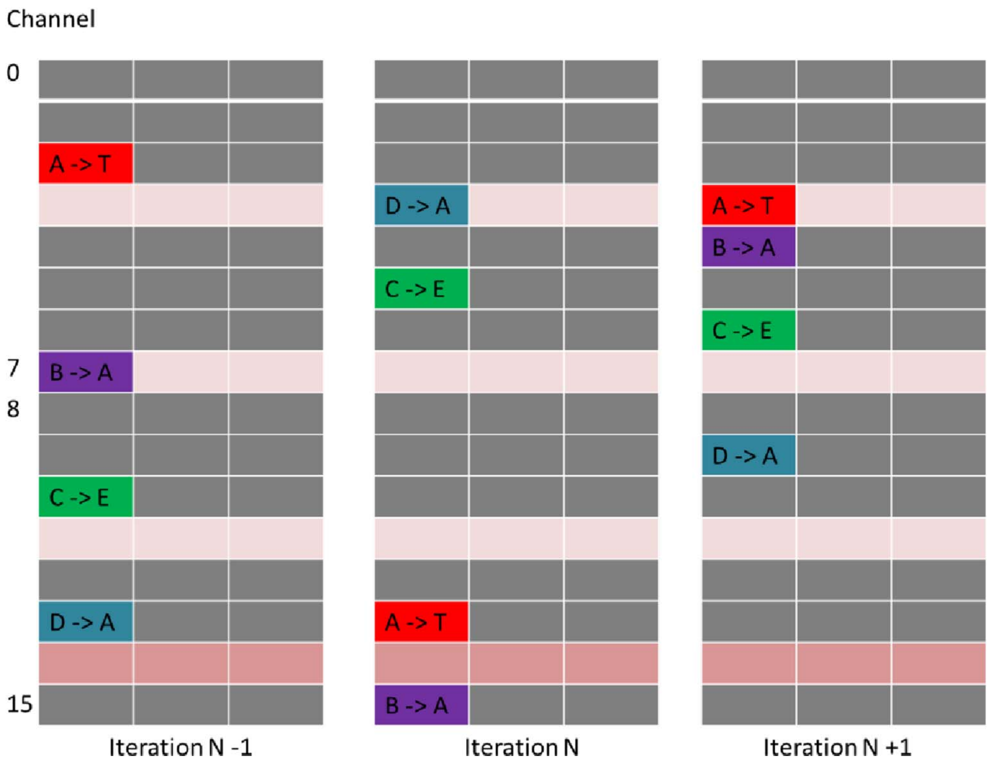


Figure 15: TSCH iterations

Table 1 shows the example of a pseudo-random sequence for 16 channels that is used in this example.

Table 1: A pseudo-random channel hopping sequence

5	6	12	7	15	4	14	11	8	0	1	2	13	3	9	10

Say that a strong interferer causes loss on channel 14 (marked brick red), then any transmission on that channel is mostly doomed. But then, say that the particular transmission from node D to node A (D *7*→ A) experiences multipath fading on channels 3, 7 and 10, which represented in pink in figure 15.

At iteration N, the transmission will fail, but an retransmission at iterations N+1 of the schedule will succeed, because it is on a different channel.

The strategy is effective in combating both co-channel interference and multipath fading, and reach wire-equivalent reliability [i.62] and [i.64], given enough time to perform all the required retransmissions.

But it falls short in the face of physical and logical failures such as equipment breakage, reboot, or network desynchronization, and the latency incurred in fast reroute may be too high to guarantee in-time delivery.

In contrast, as discussed in clause 9, 6TiSCH Tracks provide multi-path redundancy, which addresses these failure cases, and reduces the jitter incurred in retries and rerouting for a better determinism in end-to-end transmissions, arguably at the expense of additional energy spending for which some remediation are proposed.

8.2 WSNs in Industrial Process Control

IEC 62591 [i.10] and IEC 62734 (ISA100.11a) [i.11] are the major industrial WSN standards in use today in Process Control networks [i.35] and [i.42]; IEC 62601 (WIA-PA) [i.12] was developed in parallel in China, also for Process Automation applications [i.26]. Interestingly, WIA offers a faster FA version for Factory Automation, trading IEEE 802.15.4 [i.55] for an IEEE 802.11 [i.29] Physical (PHY) layer.

In order to avoid collisions and ensure the transmission of a packet at an exact time, deterministic radio operations require a fully scheduled MAC such as TSCH and LTE/5G.

Both IEC 62591 [i.10] and IEC 62734 [i.11] use a variation of the IEEE Std 802.15.4 [i.55] TSCH MAC, which is optimized for ultra-low power activities and is a natural match for low-frequency periodic flows, such as control loops, and both rely on the centralized routing model promoted by SDN.

A Controller called System Manager, or Network Manager, respectively, computes all the routes in the mesh network. Those routes are generally multipath, to augment the spatial diversity that is offered to the transported flows and to fast-route around the interference and breakages dynamically. Due to the complexity of solving the NP-complete problem of multi-path route optimization, those networks are not designed to scale beyond roughly one hundred nodes, and are generally too costly to efficiently address large scale monitoring applications such as required for the Industrial Internet.

In spite of their common roots and design points, each of these standards is defined from the PHY layer all the way to the application as a monolithic silo, with no desire to interoperate with one another beyond the regulatory capabilities (ETSI and FCC) to share the spectrum with other technologies.

Because a needful device may only exist in one of those standards, practical use cases often need to deploy more than one of those standards, which means that different hardware is installed and maintained by different OT specialists, which multiplies the operational expenses (OPEX), and appears to be a major limitation that to their wider deployment.

This contrasts with the end-to-end principle that guides the Internet designs, with a network that is agnostic to the applications and can be shared between multiple existing and any upcoming ones.

It is usual that a green field starts with highly specialized proprietary or semi-proprietary solutions; this tends to maintain the prices high and limits the adoption of the new technology. Soon enough, open standards based on the Internet Protocol (IP) -that would be IPv6 for the IoT- and the end-to-end principle eventually take over.

Bringing IP eliminates the need for gateways to provide connectivity to a wider network such as the Internet, and enables a common network infrastructure and a shared management for all applications, which drives the OPEX down and results in a larger acceptance. This transition can be expected for industrial WSNs as well.

But the Industrial Internet is also - and a lot - about reporting non-critical data such as diagnostics and for which the incumbent protocols are not a cost-efficient solution. The next problem for industrial wireless is thus to extend highly predictable WSN technologies to support IPv6, and in the process to share bandwidth and other physical resources with non-deterministic traffic, reaching higher scales at lower costs.

8.3 6TiSCH and best effort IPv6

The application of wireless technology in the operational space has enabled a variety of new devices to get interconnected, at a very low marginal cost per device, at any distance ranging from Near Field to interplanetary, and in circumstances where wiring may not be practical, for instance on fast-moving or rotating devices.

While critical monitoring was initially the main application, large scale/best effort capture of missing measurements for analytics purposes is now the fastest growing application of wireless technology in OT; a study by ABI Research shown in figure 16 indicates that it should represent more than the half of the deployed devices within the next 3 years.

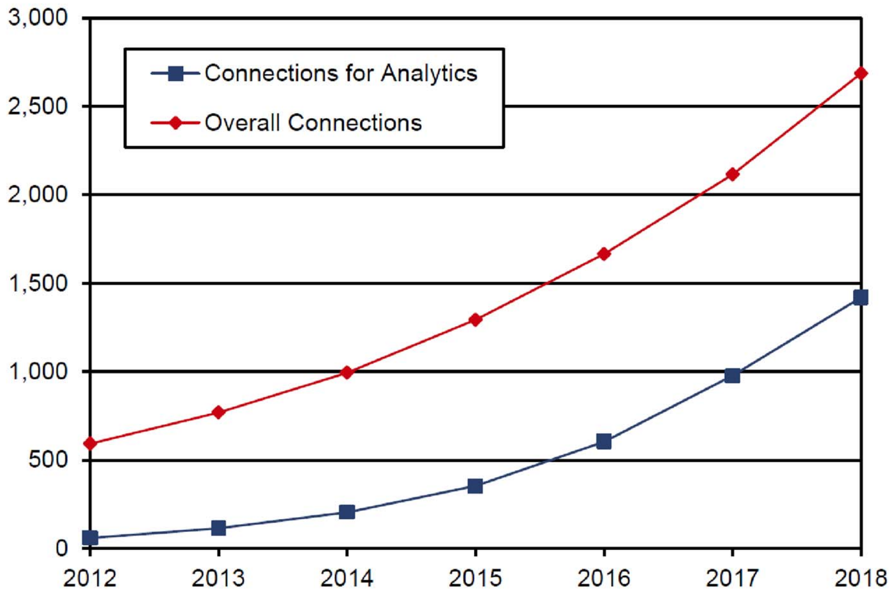


Figure 16: Growth in analytics-related measurements. Source: ABI Research

To support this growing class of traffic, standardization work has started at the IETF 6TiSCH WG that enables IPv6 over the TSCH MAC technology for best effort statistically multiplexed traffic. This way, a large number of rare transmissions can share TSCH resources that are left unused by deterministic control traffic, without interfering with it. The challenge is to adapt to the dynamics of these transmissions and preserve energy.

The 6TiSCH Architecture discusses techniques for allocating resources in the form of chunks of time slots that allow the physical separation of the two types of traffic, enabling an effective coexistence whereby the highly predictable properties of time-sensitive flows are maintained in the presence of stochastic traffic.

6TiSCH addresses this additional challenge and allows for a mix of stochastic (best effort) IPv6 flows with such well-known deterministic flows while preserving the deterministic properties regardless of the load imposed by other flows. While the work on a protocol stack [i.41] for best effort is well-advanced at the IETF, and though the vision is clearly to apply the methods defined at the IETF DetNet WG, it remains to be ensured that the way a path is signalled in wired networks is fit for the wireless medium as well.

The pillar of this stack, 6LoWPAN, has emerged as the suite of protocols that enables IPv6 on any low-power personal-area device, however small. Multiple Internet-Drafts are available from the IETF that specify 6LoWPAN over such Low-power Lossy Networks (LLN)s as power-line (IEEE 1901.2-2013 [i.110]), BACnet, NFC, ZWave, Bluetooth Low Energy (BTLE) and even IEEE 802.11ah [i.111].

To clarify a common misunderstanding, 6LoWPAN does not compare with a full solution standard such as WirelessHART and ISA100.11a; 6LoWPAN is the standard way to support IPv6 over LLNs, meant to be included in a full solution such as the above. 6LoWPAN provides an Adaptation Layer and a novel registration-based IPv6 Neighbor Discovery (ND) operation for duplicate address detection (DAD) that saves inefficient multicast operations.

The 6TiSCH architecture [i.55] combines 6LoWPAN for IPv6, RPL for routing between constrained devices, and a suite of IETF protocols, for application over IEEE 802.15.4 TSCH [i.55]. In a fashion, 6TiSCH generalizes ISA100.11a and WirelessHART to enable Industrial Internet use case over IEEE 802.15.4 TSCH [i.55], though it has to be noted that for the most part, the 6TiSCH architecture is NOT specific to that MAC.

The extent of the problem space for 6TiSCH encompasses one or more LLNs, which are reachable over one or more LLN Border Routers (6LBRs). The LLNs may be federated through a common high-speed backbone link such as an Ethernet switched fabric operating classical IPv6 ND.

As illustrated in figure 17, 6TiSCH introduces the concept of a Backbone Router (6BBR) [i.56] functionality that performs routing and proxy operations to aggregate the LLNs and the backbone link into a single Multi-Link Subnet.

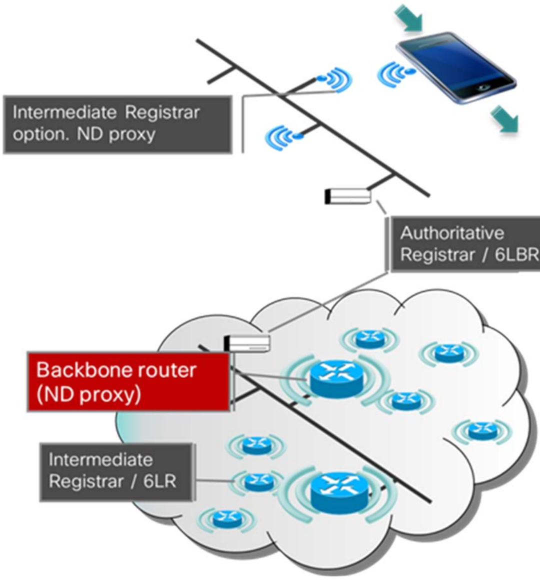


Figure 17: 6TiSCH Backbone Router

With the currently defined operation, the backbone router proxies on IPv6 Neighbor Discovery (ND) operation over the backbone on behalf of the LLN devices, enabling them to remain asleep as they are looked up from backbone devices. As a result, the BBR attracts the packets from the backbone and routes them to the destination LLN devices through the operation of RPL.

The first deliverable of the WG, the 6TiSCH Minimal Support [i.59], describes a slotted-aloha operation over a static schedule composed of shared timeslots, which means that any node can transmit or listen on these slots; the Minimal Support specification is complete at the time of this writing, going through the IESG review process as a best practice document.

A companion document 6TiSCH Minimal Security [i.60] is on the works to enable the initial settings of Layer 2 keys and a secure connectivity for the configuration of the device. Since the Minimal Security assumes the one-touch manual setting of Pre-Shared Keys (PSK) on each individual device, there is an additional desire to enable fully autonomic operations and avoids the manual intervention.

Based on certificates installed by the vendor on the device, and a backend collaboration between the vendor and the target domain, the 6TiSCH Secure Join Protocol [i.46] will cover the phase that precedes the Minimal Security and enable the initial trust and key exchange between the device and the domain, eliminating the need for a PSK.

The group is now addressing the challenge to make the schedule dynamic, which involves a Scheduling Function [i.15] that allocates and releases time slots dynamically for parent/child unicast communication along the RPL graph, based on the observed needs of the current flows, and a peer-to-peer protocol [i.61] between adjacent nodes to negotiate the time slots.

When this work is complete, 6TiSCH may re-charter to work on adapting DetNet to wireless and ultimately enable the IT/OT convergence for WSNs.

9 The vision of 6TiSCH centralized scheduling

9.1 A converged wireless network

In order to fully realize its architecture, the 6TiSCH WG has to make work together:

- A Distributed Routing for large scale monitoring (RPL) to enable the co-existence of low-criticality flows of IPv6-based Industrial Internet, ensuring the separation of resources between deterministic and stochastic and leveraging IEEE/IETF standards (IPv6, IEEE 802.15.4 TSCH [i.55], 6LoWPAN, etc.), and this work is well underway.

- A Centralized Routing for Time-Sensitive flows for mission-critical data streams such as monitoring, control loops, diagnostics and alerts; a deterministic reach back to Fog or Cloud based application is provided for virtualized loops and measurement files, following the work initiated at DetNet for the abstractions that are common to all networks.

9.2 PCE vs. 6TiSCH

With 6TiSCH, a PCE controls the network via Command and Configuration interface that implements specific device drivers, e.g. CLI, NETCONF or CoMi. This PCE inherits from the overall DetNet design but requires some specific awareness for such notions as channels, which has to be added to the common DetNet design.

A similar concept arises when implementing DetNet on fibre optics, which requires the awareness of the light wavelength -the *lambda*- that is used for multiplexing. In that context, Multi-Protocol Label Switching (MPLS) was already generalized as G-MPLS [i.30] by adding the implicit context of the lambda to make a switching decision.

In other words, as opposed to switching based only on octets in a frame header, GMPLS uses the physical context of the packet, here the lambda at which it was received, to make the switching decision.

In "Label switching over IEEE 802.15.4e networks" [i.33], it is observed that with scheduled IEEE 802.15.4 TSCH [i.55], a frame can be switched based on the particular time and the particular channel at which it was received. This can effectively be considered as a form of G-MPLS, and enables to save octets when transmitting on a constrained medium.

9.3 6TiSCH base elements (time slots, schedule, chunks and bundles)

Based on IEEE 802.15.4 TSCH [i.55], 6TiSCH defines a new concept that is global to the network, called a *Channel distribution/usage* (CDU) matrix, as illustrated in figure 18; a CDU matrix is composed of so-called *cells*, each of a duration of one network *timeslot*.

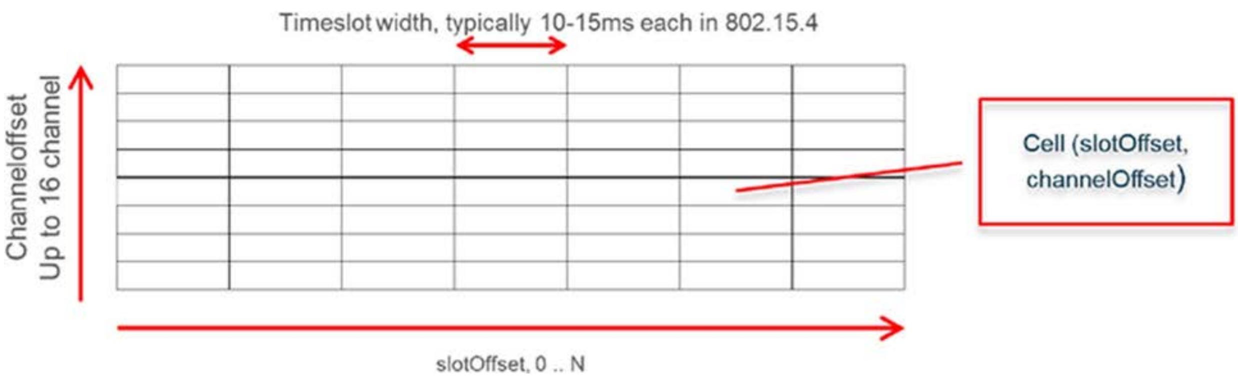


Figure 18: 6TiSCH CDU matrix

The CDU matrix provides the global characteristics of the network, such as the physical properties of the cells, and the channels used. The height of the matrix equals the number of available channels (indexed by *ChannelOffset*), and its width is the period of the iterative network scheduling operation (indexed by *slotOffset*).

The width of the CDU matrix is to be in phase with the period of the application. If a node needs to support different applications with incompatible periods, then multiple CDU matrices can be defined to accommodate those different periods, and the transmissions associated to different CDU matrices may periodically overlap.

The CDU matrix also describes how cells are grouped in *Chunks* of similar properties. A Chunk is a well-known list of cells, well distributed in time and frequency, within the CDU matrix. In other words, a chunk represents some unit of bandwidth and can be seen as the generalization of a transmission channel in the time/frequency domain.

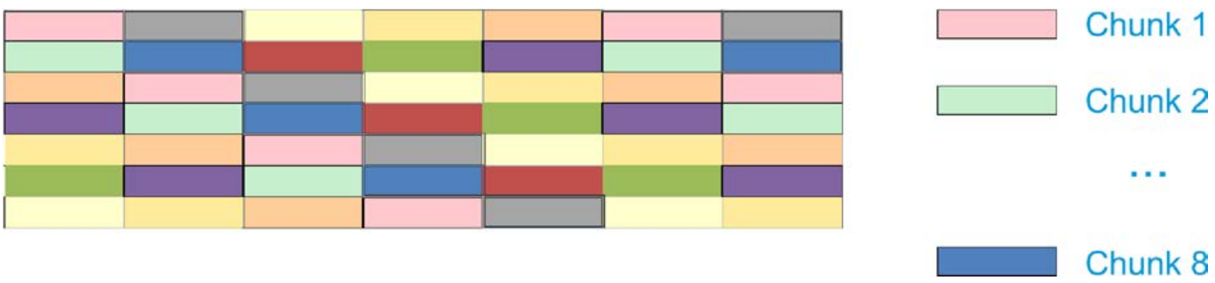


Figure 19: Example CDU matrix partitioned in Chunks

The partition of the CDU in chunks, as shown in figure 19, is globally known by all the nodes in the network to support an appropriation process that is left to be fully defined, by which a node gets ownership of the cells in the Chunk. A node that appropriates a Chunk gets to decide which transmissions will occur over the cells in the chunk within its interference domain.

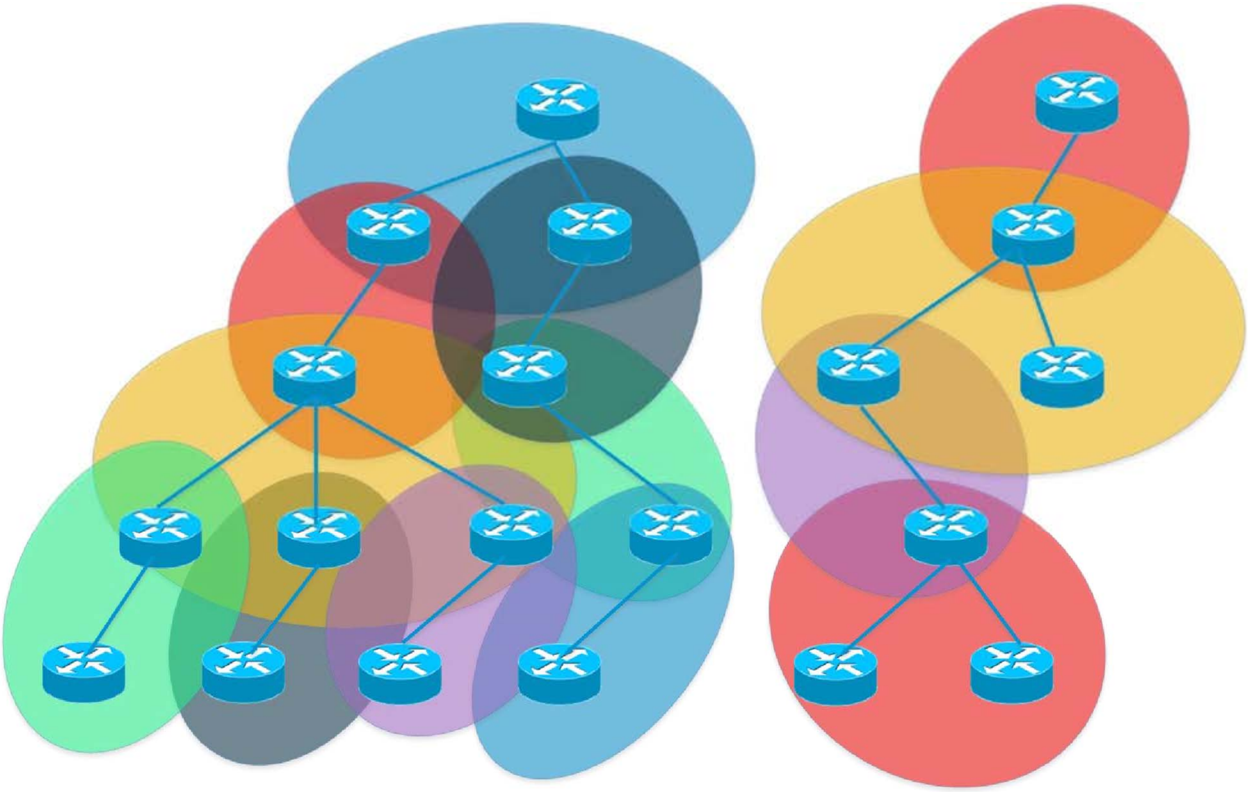


Figure 20: Spatial application of Chunks

The idea is that a node, typically a RPL parent, appropriates a chunk, and then uses it to communicate with its children. The appropriation process will enable a parent to grab a whole chunk and validate that this chunk is not used within its interference domain.

A *schedule*, on the other hand, is a MAC-level abstraction that is local to each node; a node's schedule represents the minimal knowledge that this node needs to participate to the network. To express a schedule, TSCH defines the concept of *slotframes*, which represent the timely layout of *timeslots*, and *bundles*, which represent their logical relationship (e.g. they belong to a same IPv6 Link).

A *slotframe* comprises a series of timeslots of equal length and priority, indexed by *slotOffset*; a timeslot represents the activity of this node at that slotOffset, e.g. which cell is being used, and whether it is a reception or a transmission. The duration of a slotframe is aligned with that of a CDU matrix as illustrated in figure 21.

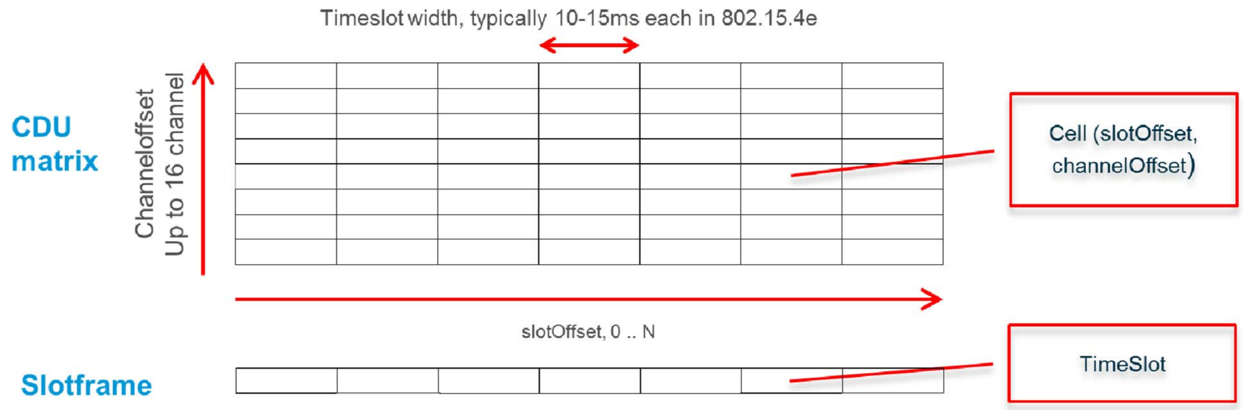


Figure 21: 6TiSCH slotframe

A node may need to support concurrent applications, and these applications may need to transmit or receive at a same time. To resolve such collisions, the application flows with different priorities are scheduled on different slotframes, ordered by application priority.

Decision to transmit/receive is made at the beginning of a timeslot based on existing frames in queues and slotframe priorities, as illustrated in figure 22.

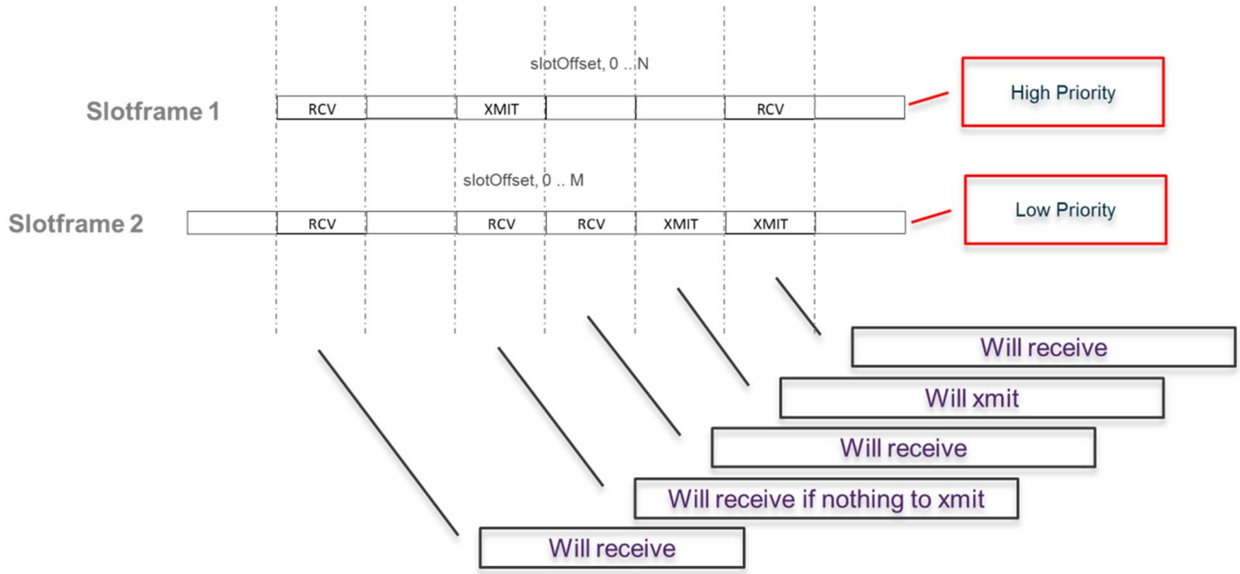


Figure 22: 6TiSCH Slotframe prioritization and schedule execution

With TSCH, one node wakes up to receive a frame when another wakes up to transmit, on a timeslots of their respective schedules that are associated with a same cell of the CDU. Both nodes wake up at the exact same time -modulo a small guard time to cover for the clock drift- and on the exact channel, which it rotating with the same pseudo-random algorithm on both ends.

This is how power-constrained devices can stay in low-energy deep sleep mode when not involved in a communication.

With 6TiSCH, the bandwidth that is allocated for a particular purpose is represented as a *bundle* of timeslots as illustrated in figure 23. A node that owns of a chunk of cells will place them in matching bundles on both ends.

When applied to Layer 3, a bundle participates to an IP Link between adjacent nodes, whereas at Layer 2, a bundle participates to a path between a source node and a destination node and this node may only be an intermediate forwarder.

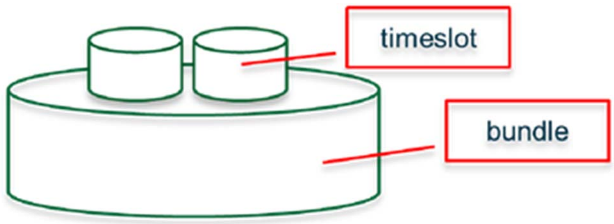


Figure 23: A 6TiSCH Bundle

Figure 23 illustrates that the establishment of a Layer 3 Link abstraction for IPv6 communication between a node A and a node B requires setting up a pair of bundles, one in each direction.

If the traffic is asymmetrical, it may be that the number of cells in the respective bundles is not the same in both directions.

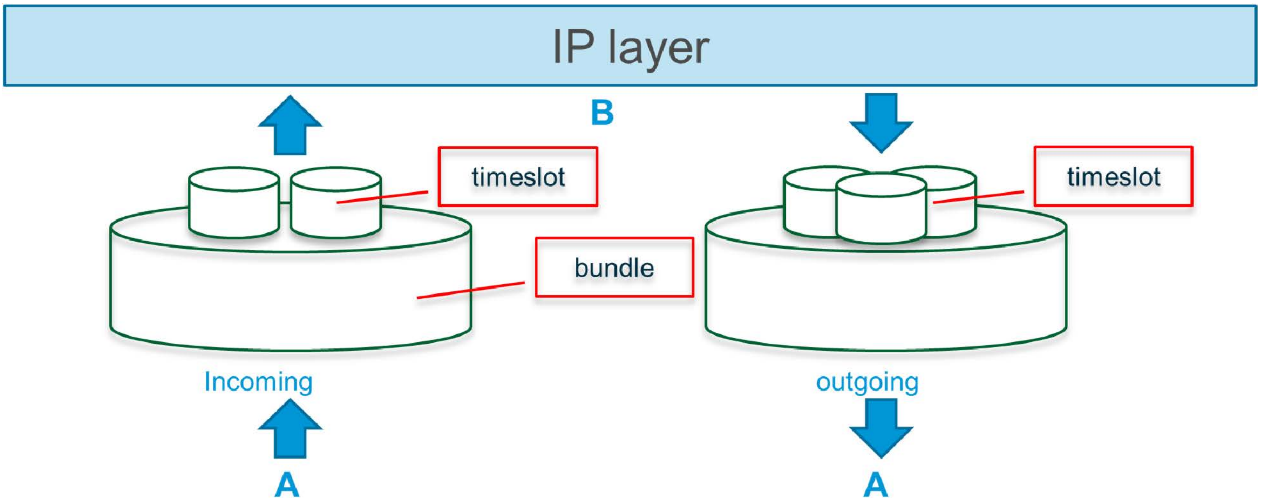
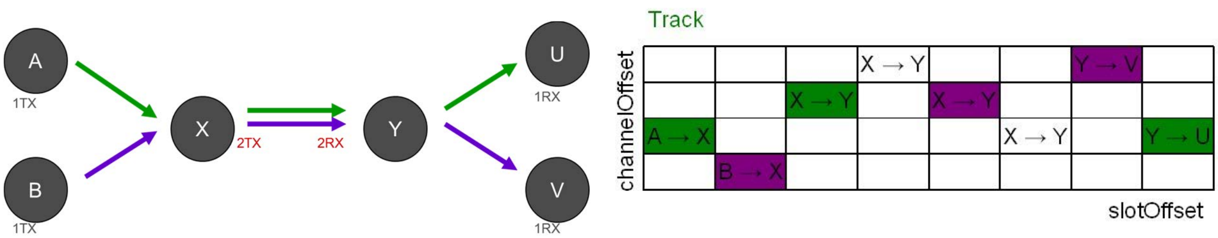


Figure 24: A 6TiSCH Layer 3 Bundle

For Layer 2 communications, 6TiSCH introduces the concept of a *Track*, and that of *Track switching*. The simplest form of Track is a serial switched path such as shown in figure 25, with a purple serial Track from a node A to a node U, and a green one from a node B to a node V.



(a) Layer 2 operation at the 6top sublayer (b) Corresponding cells in the CDU matrix

Figure 25: Forwarding along 2 simple serial Tracks, A to U and B to V

In that model, as illustrated in figure 26, the bundle that corresponds to receive slots in a node B for transmission from a node A is now paired in the schedule of node B with a bundle of transmit slots from B to C. This sequence of paired bundles from node to node determines a circuit with a particular amount of bandwidth between a source node and a destination node along a serial path.

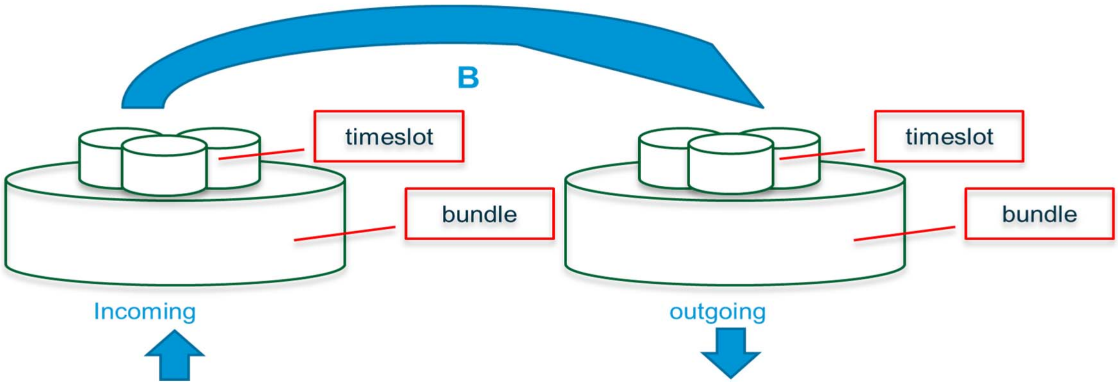


Figure 26: A 6TiSCH Layer 2 Bundle

This simple mapping applies to serial switched paths such as shown in figure 25, but with 6TiSCH Tracks, more complex constructs are required to express the full semantics of Frame Replication, Retries and Elimination, as further discussed in clause 9.5.

It is possible, for instance, to use bundles of equivalent timeslots, so the next timeslot in a transmit bundle could be used for a retransmission of a same frame, and a different transmit bundle is then used to represent the Replication. A single receive bundle can then be used to represent the Elimination of multiple copies of a same frame, regardless of the previous hop.

In both Layer 2 and Layer 3 cases, the timeslots in the receive bundle in node A use the same cell as the matching timeslot used for reception in node B, but the direction, transmit vs. receive, is opposite.

This can be ensured in a distributed fashion by the 6top Protocol [i.61], which leverages a new IEEE 802.15.4 Information Element (IE) in IEEE 802.15.4 [i.55] frames to update the cells that are allocated between adjacent 6TiSCH nodes; that new IE was delegated by the IEEE Assigned Numbers Authority (ANA) [i.103] to the IETF and allocations will be handled by the ICANN Internet Assigned Numbers Authority (IANA) [i.104].

The schedule is enforced by the 6top sublayer; 6top resides at the upper Layer 2, above the IEE802.15.4 Medium Access Control (MAC) but below the 6LoWPAN sublayers for compression and fragmentation, and the IP(v6) Layer above it.

As illustrated in figure 27, a classical IOT device will operate all layers, and may either switch or route a packet. Operating in 6top is akin to MPLS switching, which is sometimes seen as an intermediate Layer, like a Layer 2.5.

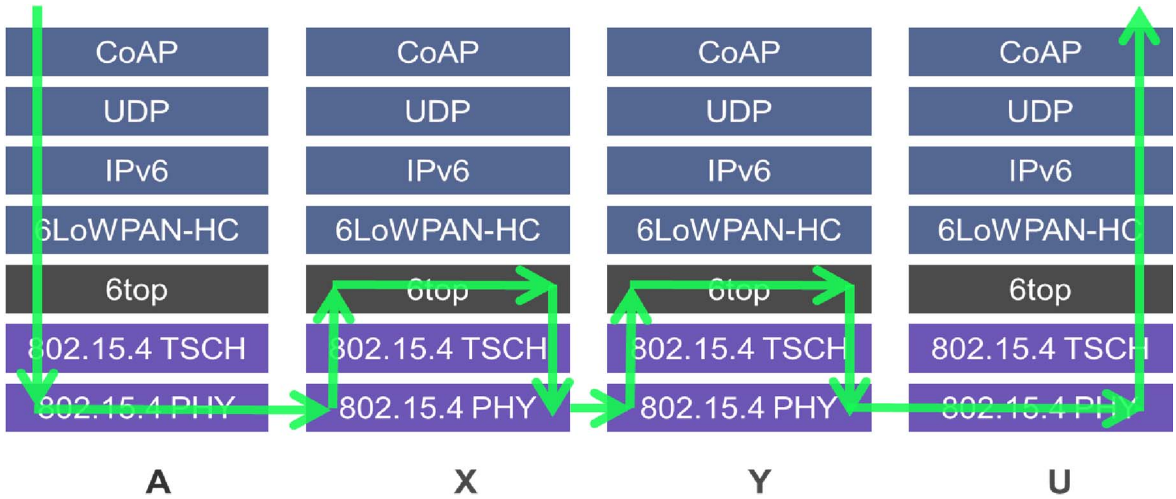


Figure 27: G-MPLS switching along Tracks

It is worth noting that, because the timeslot information is sufficient for the 6top sublayer to make a switching decision, and this, regardless of the payload of the frame, 6LoWPAN fragments can be forwarded along the G-MPLS-switched path without the need of reassembly in intermediate nodes.

This was true for any frames in the art of IEEE 802.15.4SCH and remains true for 6TiSCH Tracks.

This contrasts with normal 6LoWPAN fragment operation in Layer 3, aka *route-over*, mesh networks. In a traditional route-over 6LoWPAN network, fragments have to be reassembled at each hop in order to enable the IP routing operation based on the IP header that is only present in the first fragment.

Work on LLN Fragment Forwarding and Recovery [i.58] is ongoing to improve that situation and forward individual fragments all the way. To achieve this, the draft also leverages a switching operation on the fragments that is akin to MPLS, this time based on locally significant fragment IDs.

9.4 Applying DetNet to 6TiSCH

This clause shows how a variation of the Available Routing Construct (ARC) technology can be leveraged to form the ladder, and then how timeslots can be scheduled and ACKs optionally suppressed to save additional energy.

Briefly, the major steps for 6TiSCH centralized operations as controlled by a PCE are as follows:

- Discovering and exposing the 6TiSCH topology to the PCE/controller
- Programming a full schedule per node vs. TE-like path setup
- Retransmission vs. elimination and replication
- Detecting and rerouting around interferences

The WG has documented which kind of operations may be required from a central computer to establish a deterministic path over the 6TiSCH network. Some of this text was integrated in the Deterministic Networking Use Cases [i.20] Work Group document, in clause 5 Wireless for Industrial.

A more complete description of the 6TiSCH requirements for DetNet and the concept of a Track can be found in the 6TiSCH requirements for DetNet [i.54] and [i.20]. That document details the dependencies on DetNet and PCE controller to express topologies and capabilities, as well as abstract state that the controller has to program into the network devices to enable deterministic forwarding operations.

In order to cope with the high loss ratio on the wireless medium, there is a need to support the concept of Replication and Elimination, as combined with the concept of Retransmission. It results that a Track may be a lot more complex than a serial sequence of hops, and involve correlating multiple transmissions and reception as relating to a same packet.

In other words, a node may need to:

- Expect more than one copies of the same frame at some subsequent slotOffset/channelOffset and eliminate the duplicate.
- Forward replicated copies of that frame at multiple slotOffset/channelOffset.
- Perform retries when the copies are not received, which is a concept that is not present in wired Deterministic Networks.

9.5 Forwarding along 6TiSCH Tracks

6TiSCH Tracks extend the model of G-MPLS so that the physical properties of a transmission indicate not simply a next hop but a full context for Replication, Retries and Elimination, all done without the need to add information inside the frame itself, and leverage that concept to form complex Tracks.

A deterministic flow is qualified with information such as:

- *bandwidth requirements*, which translates in a number of cells in a bundle;
- *precise time of packet transmission*, which translates in the slotOffset in the schedule for those cells; and
- *maximum latency*, which translates in the alignment of schedules along a path, which is computed by a PCE so as to meet the end-to-end deadline.

In order to closely indicate the deterministic scheduling of every individual packet, this work suggests a number of operational rules for 6TiSCH Tracks:

- The IEEE 802.15.4 [i.55] destination MAC address in the frames is always set to multicast (0xFFFF), meaning that all nodes that have a Track programmed to listen to that particular timeslot are expected to accept the frame and handle it.
- one (or more than one) bundle are assigned uniquely to each frame in the flow; in other words, all the cells in the bundle carry a copy of a same frame.
- Multiple cells in a bundle indicate that a packet is to be retried in case of a loss; an acknowledgement is the indication that the other cells in the bundle will not be used. This denotes an OR operation between the cells in the bundle.
- Multiple transmit bundles denote a Replication. Regardless of the success or failure of the transmissions over bundle A, bundle B will be tried. This denotes an AND operation between the cells in the bundle.
- Multiple receive bundles indicate an Elimination. Unless it is desired to capture statistics on successful transmissions, there is no point in listening to the next bundles once the frame is received.

9.6 Enabling the convergence

In the use case of Industrial Internet, industries are after the next percentile point of operational efficiency to reduce down-times and operational expenditures (OPEX), and save natural resources and energy. A new form of optimization is emerging, which requires collecting and processing of live "big data", that is huge amounts of missing measurements, by widely distributed sensing and analytics (Fog) capabilities.

For reasons of cost and complexity of operation, as well as available spectrum, there is a need of convergence of deterministic industrial networks acting as silos onto IP and to share bandwidth with non-deterministic traffic, reaching higher scales at lower costs.

In the art of OT, industrial networks are deployed to transport deterministic flows (e.g. control loops) over Deterministic Networks (serial, TDM buses). In the IT art, best effort networks (Ethernet, Wi-Fi) transport statistically multiplexed traffic (IP). In both cases, there is a natural fit and multiple technologies have been successfully deployed over the last 40 years.

Converging deterministic and stochastic flows on a same network requires breaking that natural coupling. Either the network is deterministic and the challenge is to transport best effort traffic, or the network is deterministic and the challenge is the reverse. As it goes, both have been tried. Applying QoS (IEEE 802.11e [i.108] or derivatives) and with a widely underutilized physical medium, it is possible to make a stochastic packet-switched network transport Deterministic Flows. But under load from best effort traffic, this solution fails rapidly, faster on Wi-Fi than on Ethernet since IEEE 802.11e is a statistical QoS operation [i.29].

Table 2: Matching Layer capabilities

Type of MAC / Type of traffic	Deterministic (e.g. Control Loops)	Stochastic (e.g. classical IP)
Deterministic (e.g. 802.15.4 TSCH)	Good fit Adapted to centralized routing and fully scheduled operation All industrial protocols are here	Difficult but achievable: requires dynamic allocation of transmission resources (6TiSCH)
Stochastic (e.g. Zigbee, Wi-Fi)	Problems with channel access (guard time) Lead to gross over-provisioning CSMA alone cannot provide hard guarantees	Good fit Adapted for IP traffic, distributed routing and statistical multiplexing with RED

The 6TiSCH architecture defines the other way around, with the challenge to enable stochastic IP traffic over a deterministic TSCH MAC. The idea is that a PCE reserves *hard cells* from the time/frequency matrix CDU matrix for deterministic flows; those cells cannot be reused of displaced; they are allocated when a flow is established, along complex paths called *Tracks*. The reservation ensures that a flow that is placed on a Track cannot be influenced whatsoever by stochastic IP flows, which can only use the unreserved cells.

What is free in the CDU matrix is partitioned in chunks of *soft cells* that are made available for best effort traffic. This Time and Frequency Division Multiplexing technique is how 6TiSCH ensure the co-existence of deterministic and best effort traffic on a same medium.

10 Conclusion

Over the last forty years, most of the communication technologies that people use on a daily basis, mail, books, music, voice and video, have converged onto digital networks, and, by and large, the Internet, not only bringing costs down but also adding unprecedented new value such as immediate service and improved interactivity.

In a similar fashion, converging IT and OT technology on a shared network yields vastly unrealized benefits in multiple vertical industries, including but not limited to manufacturing, commercial sector, building automation, vehicles, and the power grid. Despite the huge potential, the convergence is not happening; OT networks are still typically purpose-built, proprietary, using serial point-to-point wires, and operated as physically separate networks, which multiplies the complexity of the physical layout and the operational (OPEX) and capital (CAPEX) expenditures, while preventing the agile reuse of the compute and network resources. In some cases, the operational resistance comes from a lack of trust in the technology and between different professional groups; only time, probably decades, will change that. In other cases, the limitation is technical and the lack of determinism in IT networks now appears as the gating factor.

Bringing determinism in Information Technology (IT) networks will enable the emulation of those legacy serial wires over IT fabrics and the convergence of mission-specific OT networks onto IP. The IT/OT convergence onto Deterministic Networks will in turn enable new process optimization by introducing IT capabilities, such as the Big Data and the network functions virtualization (NFV), improving OT processes while further reducing the associated OPEX. There are several existing and emerging use-cases for Deterministic Networking. Audio/Video Bridging, (AVB) for the entertainment industry, Professional Media Networking, (PMN) targeted at the broadcast industry and many applications of the Internet of Things (IOT) in the context of the Industrial Internet are three significant categories. In one example, Industrial Internet potentially yields tens of billions of savings in various industries by optimizing industrial processes.

Deterministic Networking technology allows new Quality of Service (QoS) guarantees of 'worst-case' delivery. More precisely, the worst-case data loss and latency can be provided in a consistent fashion as multiple services are deployed, augmenting the load of the network with no measurable impact on existing flows whatsoever. Based on time, resource reservation, and policy enforcement by distributed shapers, Deterministic Networking provides the capability to carry specified unicast or multicast data streams for real-time applications with extremely low data loss rates and bounded latency, so as to support time-sensitive and mission-critical applications on a converged enterprise infrastructure.

Both wired and wireless networks are evolving towards more determinism, in particular with work done at the IEEE 802.1 for bridged Ethernet networks, and at IEEE 802.15 [i.80] for Low-power Wireless PANs, but the techniques used in wired and wireless environments are largely different; the DetNet group at the IETF is now considering the establishment of end-to-end paths with Deterministic properties from the perspective of Layer 3, hopefully to be applied at 6TiSCH for the particular case of LWPANs.

New capabilities are required to drive the connection of billions of things, and make available the vast amounts of data that are generated by IoT applications and do so in accordance with specific application performance requirements beyond our traditional Internet network technologies.

Deterministic Networking Solutions and application use-cases require capabilities of the converged network that is beyond existing QOS mechanisms. Key attributes of Deterministic Networking are:

- Time synchronization on all the nodes, often including source and destination.
- The centralized computation of network-wide deterministic paths.
- New traffic shapers within and at the edge to protect the network.
- Hardware for scheduled access to the media.

The applicability of the various techniques in the art, and of those proposed in this manuscript, really depends on the use case:

- If an industrial application has no degree of liberty in terms of acceptable jitter and latency whatsoever, then the traditional technique of modulating current over a point to point wire is probably still the best, if not the only option. But the cost and operational complexity of deploying new cables in an existing production facility hinders the addition of new devices for upgrades and enhancements.
- If a high rate and an ultra-precise determinism is required, but some limited latency is acceptable, then a solution based on high speed deterministic Ethernet could be considered; but even if that does not require end-to-end wiring but only from device to the switched fabric, deploying Ethernet comes at the incremental cost for deploying wires the devices, which may or may not be doable in a particular environment, or may be hardly affordable, depending on the use case. When applicable, for instance for low speed Process Control applications, huge savings in incremental deployment time and cost of goods may be achieved by relying of wireless solutions as opposed to wires.
- If energy is critical and some rare interruptions of service are acceptable, for instance for non-critical monitoring applications, then a hot potato forwarding along a serial TSCH path will be optimal, providing both the best average delivery time and the lowest energy consumption.
- But if the prominent goal is to get every packet to the destination and energy is only secondary to it, then a full Replication and Elimination offers the best chances of success. If the latency budget permits, one might then balance an even more complex Track and more spatial diversity with a mix replication and retransmission.

Enabling determinism on wired and wireless networks separately is certainly not the end of the journey. It can be foreseen that maintaining deterministic properties at the interconnection of wired and wireless networks will be problematic; and this is a green field for future research, but certainly not barring from already deploying the technology in homogeneous wired or wireless control networks.

Annex A: Authors & contributors

The following people have contributed to the present document:

Rapporteur:
Mr Pascal Thubert, Cisco Systems, Inc.

History

Document history		
V1.1.1	March 2017	Publication