Internet Engineering Task Force (IETF)
Request for Comments: 7228
Category: Informational
ISSN: 2070-1721

C. Bormann
Universitaet Bremen TZI
M. Ersue
Nokia Solutions and Networks
A. Keranen
Ericsson

May 2014

Terminology for Constrained-Node Networks

Abstract

The Internet Protocol Suite is increasingly used on small devices with severe constraints on power, memory, and processing resources, creating constrained-node networks. This document provides a number of basic terms that have been useful in the standardization work for constrained-node networks.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7228.

Copyright Notice

Table of Contents

1. Introduction

Small devices with limited CPU, memory, and power resources, so-called "constrained devices" (often used as sensors/actuators, smart objects, or smart devices) can form a network, becoming "constrained nodes" in that network. Such a network may itself exhibit constraints, e.g., with unreliable or lossy channels, limited and unpredictable bandwidth, and a highly dynamic topology.

Constrained devices might be in charge of gathering information in diverse settings, including natural ecosystems, buildings, and factories, and sending the information to one or more server stations. They might also act on information, by performing some physical action, including displaying it. Constrained devices may work under severe resource constraints such as limited battery and computing power, little memory, and insufficient wireless bandwidth and ability to communicate; these constraints often exacerbate each other. Other entities on the network, e.g., a base station or controlling server, might have more computational and communication resources and could support the interaction between the constrained devices and applications in more traditional networks.

Today, diverse sizes of constrained devices with different resources and capabilities are becoming connected. Mobile personal gadgets, building-automation devices, cellular phones, machine-to-machine (M2M) devices, and other devices benefit from interacting with other "things" nearby or somewhere in the Internet. With this, the Internet of Things (IoT) becomes a reality, built up out of uniquely identifiable and addressable objects (things). Over the next decade, this could grow to large numbers [FIFTY-BILLION] of Internet-connected constrained devices, greatly increasing the Internet’s size and scope.

The present document provides a number of basic terms that have been useful in the standardization work for constrained environments. The intention is not to exhaustively cover the field but to make sure a few core terms are used consistently between different groups cooperating in this space.

In this document, the term "byte" is used in its now customary sense as a synonym for "octet". Where sizes of semiconductor memory are given, the prefix "kibi" (1024) is combined with "byte" to "kibibyte", abbreviated "KiB", for 1024 bytes [ISQ-13].

In computing, the term "power" is often used for the concept of "computing power" or "processing power", as in CPU performance. In this document, the term stands for electrical power unless explicitly stated otherwise. "Mains-powered" is used as a shorthand for being permanently connected to a stable electrical power grid.

2. Core Terminology

There are two important aspects to _scaling_ within the Internet of Things:

- scaling up Internet technologies to a large number [FIFTY-BILLION] of inexpensive nodes, while
- scaling down the characteristics of each of these nodes and of the networks being built out of them, to make this scaling up economically and physically viable.

The need for scaling down the characteristics of nodes leads to "constrained nodes".

2.1. Constrained Nodes

The term "constrained node" is best defined by contrasting the characteristics of a constrained node with certain widely held expectations on more familiar Internet nodes:

Constrained Node: A node where some of the characteristics that are otherwise pretty much taken for granted for Internet nodes at the time of writing are not attainable, often due to cost constraints and/or physical constraints on characteristics such as size, weight, and available power and energy. The tight limits on power, memory, and processing resources lead to hard upper bounds on state, code space, and processing cycles, making optimization of energy and network bandwidth usage a dominating consideration in all design requirements. Also, some layer-2 services such as full connectivity and broadcast/multicast may be lacking.

While this is not a rigorous definition, it is grounded in the state of the art and clearly sets apart constrained nodes from server systems, desktop or laptop computers, powerful mobile devices such as smartphones, etc. There may be many design considerations that lead to these constraints, including cost, size, weight, and other scaling factors.

(An alternative term, when the properties as a network node are not in focus, is "constrained device".)

There are multiple facets to the constraints on nodes, often applying in combination, for example:

- constraints on the maximum code complexity (ROM/Flash),
- constraints on the size of state and buffers (RAM),
- constraints on the amount of computation feasible in a period of time ("processing power"),
- constraints on the available power, and
- constraints on user interface and accessibility in deployment (ability to set keys, update software, etc.).

Section 3 defines a small number of interesting classes ("class-N" for N = 0, 1, 2) of constrained nodes focusing on relevant combinations of the first two constraints. With respect to available power, [RFC6606] distinguishes "power-affluent" nodes (mains-powered or regularly recharged) from "power-constrained nodes" that draw their power from primary batteries or by using energy harvesting; more detailed power terminology is given in Section 4.

The use of constrained nodes in networks often also leads to constraints on the networks themselves. However, there may also be constraints on networks that are largely independent from those of the nodes. We therefore distinguish "constrained networks" from "constrained-node networks".

2.2. Constrained Networks

We define "constrained network" in a similar way:

Constrained Network: A network where some of the characteristics pretty much taken for granted with link layers in common use in the Internet at the time of writing are not attainable.

Constraints may include:

- low achievable bitrate/throughput (including limits on duty cycle),
- high packet loss and high variability of packet loss (delivery rate),

- highly asymmetric link characteristics,
- severe penalties for using larger packets (e.g., high packet loss due to link-layer fragmentation),
- limits on reachability over time (a substantial number of devices may power off at any point in time but periodically "wake up" and can communicate for brief periods of time), and
- lack of (or severe constraints on) advanced services such as IP multicast.

More generally, we speak of constrained networks whenever at least some of the nodes involved in the network exhibit these characteristics.

Again, there may be several reasons for this:

- cost constraints on the network,
- constraints posed by the nodes (for constrained-node networks),
- physical constraints (e.g., power constraints, environmental constraints, media constraints such as underwater operation, limited spectrum for very high density, electromagnetic compatibility),
- regulatory constraints, such as very limited spectrum availability (including limits on effective radiated power and duty cycle) or explosion safety, and
- technology constraints, such as older and lower-speed technologies that are still operational and may need to stay in use for some more time.

2.2.1. Challenged Networks

A constrained network is not necessarily a "challenged network" [FALL]:

Challenged Network: A network that has serious trouble maintaining what an application would today expect of the end-to-end IP model, e.g., by:

- not being able to offer end-to-end IP connectivity at all,
- exhibiting serious interruptions in end-to-end IP connectivity, or

- exhibiting delay well beyond the Maximum Segment Lifetime (MSL) defined by TCP [RFC0793].

All challenged networks are constrained networks in some sense, but not all constrained networks are challenged networks. There is no well-defined boundary between the two, though. Delay-Tolerant Networking (DTN) has been designed to cope with challenged networks [RFC4838].

2.3. Constrained-Node Networks

Constrained-Node Network: A network whose characteristics are influenced by being composed of a significant portion of constrained nodes.

A constrained-node network always is a constrained network because of the network constraints stemming from the node constraints, but it may also have other constraints that already make it a constrained network.

The rest of this subsection introduces two additional terms that are in active use in the area of constrained-node networks, without an intent to define them: LLN and (6)LoWPAN.

2.3.1. LLN

A related term that has been used to describe the focus of the IETF ROLL working group is "Low-Power and Lossy Network (LLN)". The ROLL (Routing Over Low-Power and Lossy) terminology document [RFC7102] defines LLNs as follows:

> LLN: Low-Power and Lossy Network. Typically composed of many embedded devices with limited power, memory, and processing resources interconnected by a variety of links, such as IEEE 802.15.4 or low-power Wi-Fi. There is a wide scope of application areas for LLNs, including industrial monitoring, building automation (heating, ventilation, and air conditioning (HVAC), lighting, access control, fire), connected home, health care, environmental monitoring, urban sensor networks, energy management, assets tracking, and refrigeration.

Beyond that, LLNs often exhibit considerable loss at the physical layer, with significant variability of the delivery rate, and some short-term unreliability, coupled with some medium-term stability that makes it worthwhile to both construct directed acyclic graphs that are medium-term stable for routing and do measurements on the edges such as Expected Transmission Count (ETX) [RFC6551]. Not all LLNs comprise low-power nodes [RPL-DEPLOYMENT].

LLNs typically are composed of constrained nodes; this leads to the design of operation modes such as the "non-storing mode" defined by RPL (the IPv6 Routing Protocol for Low-Power and Lossy Networks [RFC6550]). So, in the terminology of the present document, an LLN is a constrained-node network with certain network characteristics, which include constraints on the network as well.

2.3.2. LoWPAN, 6LoWPAN

One interesting class of a constrained network often used as a constrained-node network is "LoWPAN" [RFC4919], a term inspired from the name of an IEEE 802.15.4 working group (low-rate wireless personal area networks (LR-WPANs)). The expansion of the LoWPAN acronym, "Low-Power Wireless Personal Area Network", contains a hard-to-justify "Personal" that is due to the history of task group naming in IEEE 802 more than due to an orientation of LoWPANs around a single person. Actually, LoWPANs have been suggested for urban monitoring, control of large buildings, and industrial control applications, so the "Personal" can only be considered a vestige. Occasionally, the term is read as "Low-Power Wireless Area Networks" [WEI]. Originally focused on IEEE 802.15.4, "LoWPAN" (or when used for IPv6, "6LoWPAN") also refers to networks built from similarly constrained link-layer technologies [V6-BTLE] [V6-DECT-ULE] [V6-G9959].

3. Classes of Constrained Devices

Despite the overwhelming variety of Internet-connected devices that can be envisioned, it may be worthwhile to have some succinct terminology for different classes of constrained devices. In this document, the class designations in Table 1 may be used as rough indications of device capabilities:

Name	data size (e.g., RAM)	code size (e.g., Flash)
Class 0, C0	<< 10 KiB	<< 100 KiB
Class 1, C1	˜ 10 KiB	˜ 100 KiB
Class 2, C2	˜ 50 KiB	˜ 250 KiB

Table 1: Classes of Constrained Devices (KiB = 1024 bytes)

As of the writing of this document, these characteristics correspond to distinguishable clusters of commercially available chips and design cores for constrained devices. While it is expected that the boundaries of these classes will move over time, Moore’s law tends to be less effective in the embedded space than in personal computing devices: gains made available by increases in transistor count and density are more likely to be invested in reductions of cost and power requirements than into continual increases in computing power.

Class 0 devices are very constrained sensor-like motes. They are so severely constrained in memory and processing capabilities that most likely they will not have the resources required to communicate directly with the Internet in a secure manner (rare heroic, narrowly targeted implementation efforts notwithstanding). Class 0 devices will participate in Internet communications with the help of larger devices acting as proxies, gateways, or servers. Class 0 devices generally cannot be secured or managed comprehensively in the traditional sense. They will most likely be preconfigured (and will be reconfigured rarely, if at all) with a very small data set. For management purposes, they could answer keepalive signals and send on/off or basic health indications.

Class 1 devices are quite constrained in code space and processing capabilities, such that they cannot easily talk to other Internet nodes employing a full protocol stack such as using HTTP, Transport Layer Security (TLS), and related security protocols and XML-based data representations. However, they are capable enough to use a protocol stack specifically designed for constrained nodes (such as the Constrained Application Protocol (CoAP) over UDP [COAP]) and participate in meaningful conversations without the help of a gateway node. In particular, they can provide support for the security functions required on a large network. Therefore, they can be integrated as fully developed peers into an IP network, but they need to be parsimonious with state memory, code space, and often power expenditure for protocol and application usage.

Class 2 devices are less constrained and fundamentally capable of supporting most of the same protocol stacks as used on notebooks or servers. However, even these devices can benefit from lightweight and energy-efficient protocols and from consuming less bandwidth. Furthermore, using fewer resources for networking leaves more resources available to applications. Thus, using the protocol stacks defined for more constrained devices on Class 2 devices might reduce development costs and increase the interoperability.

Constrained devices with capabilities significantly beyond Class 2 devices exist. They are less demanding from a standards development point of view as they can largely use existing protocols unchanged. The present document therefore does not make any attempt to define classes beyond Class 2. These devices can still be constrained by a limited energy supply.

With respect to examining the capabilities of constrained nodes, particularly for Class 1 devices, it is important to understand what type of applications they are able to run and which protocol mechanisms would be most suitable. Because of memory and other limitations, each specific Class 1 device might be able to support only a few selected functions needed for its intended operation. In other words, the set of functions that can actually be supported is not static per device type: devices with similar constraints might choose to support different functions. Even though Class 2 devices have some more functionality available and may be able to provide a more complete set of functions, they still need to be assessed for the type of applications they will be running and the protocol functions they would need. To be able to derive any requirements, the use cases and the involvement of the devices in the application and the operational scenario need to be analyzed. Use cases may combine constrained devices of multiple classes as well as more traditional Internet nodes.

4. Power Terminology

Devices not only differ in their computing capabilities but also in available power and/or energy. While it is harder to find recognizable clusters in this space, it is still useful to introduce some common terminology.

4.1. Scaling Properties

The power and/or energy available to a device may vastly differ, from kilowatts to microwatts, from essentially unlimited to hundreds of microjoules.

Instead of defining classes or clusters, we simply state, using the International System of Units (SI units), an approximate value for one or both of the quantities listed in Table 2:

Name	Definition	SI Unit
Ps	Sustainable average power available for the device over the time it is functioning	W (Watt)
Et	Total electrical energy available before the energy source is exhausted	J (Joule)

Table 2: Quantities Relevant to Power and Energy

The value of Et may need to be interpreted in conjunction with an indication over which period of time the value is given; see Section 4.2.

Some devices enter a "low-power" mode before the energy available in a period is exhausted or even have multiple such steps on the way to exhaustion. For these devices, Ps would need to be given for each of the modes/steps.

4.2. Classes of Energy Limitation

As discussed above, some devices are limited in available energy as opposed to (or in addition to) being limited in available power. Where no relevant limitations exist with respect to energy, the device is classified as E9. The energy limitation may be in total energy available in the usable lifetime of the device (e.g., a device that is discarded when its non-replaceable primary battery is exhausted), classified as E2. Where the relevant limitation is for a specific period, the device is classified as E1, e.g., a solar-powered device with a limited amount of energy available for the night, a device that is manually connected to a charger and has a period of time between recharges, or a device with a periodic (primary) battery replacement interval. Finally, there may be a limited amount of energy available for a specific event, e.g., for a button press in an energy-harvesting light switch; such devices are classified as E0. Note that, in a sense, many E1 devices are also E2, as the rechargeable battery has a limited number of useful recharging cycles.

Table 3 provides a summary of the classifications described above.

Name	Type of energy limitation	Example Power Source
E0	Event energy-limited	Event-based harvesting
E1	Period energy-limited	Battery that is periodically recharged or replaced
E2	Lifetime energy-limited	Non-replaceable primary battery
E9	No direct quantitative limitations to available energy	Mains-powered

Table 3: Classes of Energy Limitation

4.3. Strategies for Using Power for Communication

Especially when wireless transmission is used, the radio often consumes a big portion of the total energy consumed by the device. Design parameters, such as the available spectrum, the desired range, and the bitrate aimed for, influence the power consumed during transmission and reception; the duration of transmission and reception (including potential reception) influence the total energy consumption.

Different strategies for power usage and network attachment may be used, based on the type of the energy source (e.g., battery or mains-powered) and the frequency with which a device needs to communicate.

The general strategies for power usage can be described as follows:

Always-on: This strategy is most applicable if there is no reason for extreme measures for power saving. The device can stay on in the usual manner all the time. It may be useful to employ power-friendly hardware or limit the number of wireless transmissions, CPU speeds, and other aspects for general power-saving and cooling needs, but the device can be connected to the network all the time.

Normally-off: Under this strategy, the device sleeps such long periods at a time that once it wakes up, it makes sense for it to not pretend that it has been connected to the network during sleep: the device reattaches to the network as it is woken up. The main optimization goal is to minimize the effort during the reattachment process and any resulting application communications.

If the device sleeps for long periods of time and needs to communicate infrequently, the relative increase in energy expenditure during reattachment may be acceptable.

Low-power: This strategy is most applicable to devices that need to operate on a very small amount of power but still need to be able to communicate on a relatively frequent basis. This implies that extremely low-power solutions need to be used for the hardware, chosen link-layer mechanisms, and so on. Typically, given the small amount of time between transmissions, despite their sleep state, these devices retain some form of attachment to the network. Techniques used for minimizing power usage for the network communications include minimizing any work from re-establishing communications after waking up and tuning the frequency of communications (including "duty cycling", where components are switched on and off in a regular cycle) and other parameters appropriately.

Table 4 provides a summary of the strategies described above.

Name	Strategy	Ability to communicate
P0	Normally-off	Reattach when required
P1	Low-power	Appears connected, perhaps with high latency
P9	Always-on	Always connected

Table 4: Strategies of Using Power for Communication

Note that the discussion above is at the device level; similar considerations can apply at the communications-interface level. This document does not define terminology for the latter.

A term often used to describe power-saving approaches is "duty-cycling". This describes all forms of periodically switching off some function, leaving it on only for a certain percentage of time (the "duty cycle").

[RFC7102] only distinguishes two levels, defining a Non-Sleepy Node as a node that always remains in a fully powered-on state (always awake) where it has the capability to perform communication (P9) and a Sleepy Node as a node that may sometimes go into a sleep mode (a low-power state to conserve power) and temporarily suspend protocol communication (P0); there is no explicit mention of P1.

5. Security Considerations

This document introduces common terminology that does not raise any new security issues. Security considerations arising from the constraints discussed in this document need to be discussed in the context of specific protocols. For instance, Section 11.6 of [COAP], "Constrained node considerations", discusses implications of specific constraints on the security mechanisms employed. [ROLL-SEC-THREATS] provides a security threat analysis for the RPL routing protocol. Implementation considerations for security protocols on constrained nodes are discussed in [IKEV2-MINIMAL] and [TLS-MINIMAL]. A wider view of security in constrained-node networks is provided in [IOT-SECURITY].

6. Acknowledgements

Dominique Barthel and Peter van der Stok provided useful comments; Charles Palmer provided a full editorial review.

Peter van der Stok insisted that we should include power terminology, hence Section 4. The text for Section 4.3 is mostly lifted from a previous version of [COAP-CELLULAR] and has been adapted for this document.

7. Informative References

[COAP] Shelby, Z., Hartke, K., and C. Bormann, "Constrained Application Protocol (CoAP)", Work in Progress, June 2013.

[COAP-CELLULAR] Arkko, J., Eriksson, A., and A. Keranen, "Building Power-Efficient CoAP Devices for Cellular Networks", Work in Progress, February 2014.

[FALL] Fall, K., "A Delay-Tolerant Network Architecture for Challenged Internets", SIGCOMM 2003, 2003.

[FIFTY-BILLION]
Ericsson, "More Than 50 Billion Connected Devices", Ericsson White Paper 284 23-3149 Uen, February 2011, <http://www.ericsson.com/res/docs/whitepapers/wp-50-billions.pdf>.

[IKEV2-MINIMAL]
Kivinen, T., "Minimal IKEv2", Work in Progress, October 2013.

[IOT-SECURITY]
Garcia-Morchon, O., Kumar, S., Keoh, S., Hummen, R., and R. Struik, "Security Considerations in the IP-based Internet of Things", Work in Progress, September 2013.

[ISQ-13] International Electrotechnical Commission, "International Standard -- Quantities and units -- Part 13: Information science and technology", IEC 80000-13, March 2008.

[RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

[RFC4838] Cerf, V., Burleigh, S., Hooke, A., Torgerson, L., Durst, R., Scott, K., Fall, K., and H. Weiss, "Delay-Tolerant Networking Architecture", RFC 4838, April 2007.

[RFC4919] Kushalnagar, N., Montenegro, G., and C. Schumacher, "IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals", RFC 4919, August 2007.

[RFC6550] Winter, T., Thubert, P., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, March 2012.

[RFC6551] Vasseur, JP., Kim, M., Pister, K., Dejean, N., and D. Barthel, "Routing Metrics Used for Path Calculation in Low-Power and Lossy Networks", RFC 6551, March 2012.

[RFC6606] Kim, E., Kaspar, D., Gomez, C., and C. Bormann, "Problem Statement and Requirements for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Routing", RFC 6606, May 2012.

[RFC7102] Vasseur, JP., "Terms Used in Routing for Low-Power and Lossy Networks", RFC 7102, January 2014.

[ROLL-SEC-THREATS]
Tsao, T., Alexander, R., Dohler, M., Daza, V., Lozano, A., and M. Richardson, "A Security Threat Analysis for Routing Protocol for Low-power and lossy networks (RPL)", Work in Progress, December 2013.

[RPL-DEPLOYMENT]
Vasseur, J., Ed., Hui, J., Ed., Dasgupta, S., and G. Yoon, "RPL deployment experience in large scale networks", Work in Progress, July 2012.

[TLS-MINIMAL]
Kumar, S., Keoh, S., and H. Tschofenig, "A Hitchhiker's Guide to the (Datagram) Transport Layer Security Protocol for Smart Objects and Constrained Node Networks", Work in Progress, March 2014.

[V6-BTLE] Nieminen, J., Ed., Savolainen, T., Ed., Isomaki, M., Patil, B., Shelby, Z., and C. Gomez, "Transmission of IPv6 Packets over BLUETOOTH Low Energy", Work in Progress, May 2014.

[V6-DECT-ULE]
Mariager, P., Ed., Petersen, J., and Z. Shelby, "Transmission of IPv6 Packets over DECT Ultra Low Energy", Work in Progress, July 2013.

[V6-G9959] Brandt, A. and J. Buron, "Transmission of IPv6 packets over ITU-T G.9959 Networks", Work in Progress, May 2014.

[WEI] Shelby, Z. and C. Bormann, "6LoWPAN: the Wireless Embedded Internet", ISBN 9780470747995, 2009.

Authors' Addresses

Carsten Bormann
Universitaet Bremen TZI
Postfach 330440
D-28359 Bremen
Germany

Phone: +49-421-218-63921
EMail: cabo@tzi.org

Mehmet Ersue
Nokia Solutions and Networks
St.-Martinstrasse 76
81541 Munich
Germany

Phone: +49 172 8432301
EMail: mehmet.ersue@nsn.com

Ari Keranen
Ericsson
Hirsalantie 11
02420 Jorvas
Finland

EMail: ari.keranen@ericsson.com