ETSI GS E4P 008 V1.1.1 (2021-05)

Europe for Privacy-Preserving Pandemic Protection (E4P); Back-End mechanisms for pandemic contact tracing systems

Disclaimer

The present document has been produced and approved by the Europe for Privacy-Preserving Pandemic Protection ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference
DGS/E4P-008

Keywords
covid, eHealth, emergency services, identity,
mobility, pandemic, privacy, security, smartphone

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 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

Users of the present document should be aware that the document may be subject to revision or change of status.
Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Notice of disclaimer & limitation of liability

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.
No recommendation as to products and services or vendors is made or should be implied.
No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.
In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

Copyright Notification

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are 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 Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, 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.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

BLUETOOTH® is a trademark registered and owned by Bluetooth SIG, Inc.

Foreword

This Group Specification (GS) has been produced by ETSI Industry Specification Group (ISG) Europe for Privacy-Preserving Pandemic Protection (E4P).

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**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.

Introduction

The COVID-19 pandemic has generated significant challenge for many countries and their citizens and showed that digital technologies could play an important role in addressing this pandemic and future pandemics. Various applications, services and systems for contact tracing (identification and notification of those who come in contact with a carrier) have been developed in different regions.

Despite the similar goal of automated detection of COVID-19 exposure as a complementary solution to manual tracing (interviews with people diagnosed with COVID-19 to track down their recent contacts), their functionality, technology, scale, required data and limitations are different and may not interoperate.

These systems are currently being deployed in different countries and many more are expected in the near future. In particular, mobile devices with their contact tracing applications can support public health authorities in controlling and containing the pandemic. In that purpose, E4P has been created to provide a technical answer to pandemic crisis not limited to COVID-19 by specifying interoperable contact tracing systems.

1 Scope

The present document specifies back-end mechanisms for Pandemic contact Tracing Systems, including:

- architecture;
- information flow;
- protocols for sharing proximity data of contacts;
- the requisite APIs (Application Programming Interfaces); and
- privacy, data and system security.

Sufficient technical detail will be included to also facilitate means of interoperability in a later Work Item (ETSI GS E4P 007 [3]. Each Digital Contact Tracing System is characterized by its degree of compatibility with the E4P requirements work item (ETSI GS E4P 003 [1]).

2 References

2.1 Normative 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.

Referenced documents which are not found to be publicly available in the expected location might be found at https://docbox.etsi.org/Reference/.

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 necessary for the application of the present document.

[1] ETSI GS E4P 003 (V1.1.1): "Europe for Privacy-Preserving Pandemic Protection (E4P); High level requirements for pandemic contact tracing systems using mobile devices".

[2] ETSI GS E4P 006 (V1.1.1): "Europe for Privacy-Preserving Pandemic Protection (E4P); Device-Based Mechanisms for pandemic contact tracing systems".

[3] ETSI GS E4P 007 (V1.1.1): "Europe for Privacy-Preserving Pandemic Protection (E4P); Pandemic proximity tracing systems: Interoperability framework".

[4] IETF RFC 5246: "The Transport Layer Security (TLS) Protocol Version 1.2".

[5] IETF RFC 2616: "Hypertext Transfer Protocol -- HTTP/1.1".

[6] IETF RFC 3629: "UTF-8, a transformation format of ISO 10646".

[7] IETF RFC 5905: "Network Time Protocol Version 4: Protocol and Algorithms Specification".

[8] ZIP File Format Specification, Version: 6.3.9.

NOTE: Available at https://pkware.cachefly.net/webdocs/casestudies/APPNOTE.TXT.

[9] Annex D.1 of NIST FIPS186-4.

[10] FIPS PUB 180-4: "Secure Hash Standard (SHS)".

NOTE: Available at https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.180-4.pdf.

[11] "Exposure Notification Bluetooth® Specification", v1.2 April 2020, Google Apple.

NOTE: Available at https://developers.google.com/android/exposure-notifications/exposure-key-file-format.

[12] PRIVATICS team, Inria, France, Fraunhofer AISEC, Germany: "ROBERT: ROBust and privacy-presERving proximity Tracing, version 1.1", May 31st, 2020.

NOTE: Available at https://github.com/ROBERT-proximity-tracing/documents and https://hal.inria.fr/hal-02611265/en/.

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] Frequently Asked Questions about the Corona-Warn-App.

NOTE: Available at https://www.coronawarn.app/en/faq/#anonymous.

[i.2] Criteria for the Evaluation of Contact Tracing Apps.

NOTE: Available at https://github.com/corona-warn-app/cwa-documentation/blob/master/pruefsteine.md#unlinkability.

[i.3] PRIVATICS team, Inria, France, Fraunhofer AISEC, Germany: "ROBERT: ROBust and privacy-presERving proximity Tracing, version 1.1", May 31st, 2020. .

NOTE: Available at https://github.com/ROBERT-proximity-tracing/documents, https://hal.inria.fr/hal-02611265/en/.

[i.4] G. Kessibi, M. Cunche, A. Boutet, C. Castelluccia, C. Lauradoux, D. Le Metayer, V. Roca: "Analysis of Diagnosis Key distribution mechanism in contact tracing applications based on Google-Apple Exposure Notification (GAEN) framework (version 1.2)", September 2020. .

NOTE: Available at https://github.com/3rd-ways-for-EU-exposure-notification/project-DESIRE, https://hal.inria.fr/hal-02899412/en/.

[i.5] Swiss Confederation: "Replay Attacks", June 14th, 2020, section "Unmasking users by eavesdropping EphIDs".

[i.6] Tijmen Schep: "Corona Detective".

NOTE: Available at https://www.coronadetective.eu/.

[i.7] O. Seiskari: "BLE contact tracing sniffer PoC".

NOTE: Available at https://github.com/oseiskar/corona-sniffer.

[i.8] S. Vaudenay: "Centralized or Decentralized? The Contact Tracing Dilemma".

NOTE: Available at https://infoscience.epfl.ch/record/277809.

[i.9] Ghazaleh Beigi, Huan Liu: "A Survey on Privacy in Social Media: Identification, Mitigation, and Applications". ACM/IMS Transactions on Data Science, March 2020, Article No. 7.

NOTE: Available at https://doi.org/10.1145/3343038.

[i.10] IETF RFC 8446: "The Transport Layer Security (TLS) Protocol Version 1.3".

3 Definition of terms, symbols and abbreviations

3.1 Terms

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

diagnosis key: secret code from which ephemeral identifiers for a given period of time can be derived with the help of a cryptographic function

ephemeral identifier: unique device identifier exchanged with another device during a proximity event

proximity event: event recorded by the software on a mobile device corresponding to proximity to other device with an active interoperable applications, and which meets the predefined criteria for an event to be recorded (e.g. duration)

3.2 Symbols

Void.

3.3 Abbreviations

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

ACK	Acknowledgement
BCD	Requirement "Back-end Centralized Diagnosed"
BCE	Requirement "Back-end Centralized EBID"
BCP	Requirement "Back-end Centralized Privacy"
BCR	Requirement "Back-end Centralized Registration"
BCS	Requirement "Back-end Centralized Setup"
BCX	Requirement "Back-end Centralized Exposure"
BDB	Requirement "Back-end Decentralized BaseURL"
BDD	Requirement "Back-end Decentralized Data-Structures"
BDK	Requirement "Back-end Decentralized Key Download"
BDO	Requirement "Back-end Decentralized One-Time-Token"
BDP	Requirement "Back-end Decentralized Parameters"
BDR	Requirement "Back-end Decentralized Retrieve Test Results"
BDT	Requirement "Back-end Decentralized TLS"
BDU	Requirement "Back-end Decentralized Key Uplead"
BE	Reference Point "Back-end - External System"
BF	Reference Point "Back-end - Federation"
Bluetooth® LE	Bluetooth Low Energy
CDN	Content Distribution Network
DB	Reference Point "Device - Back-end"
DCTS	Digital Contact Tracing System
DP3T	Decentralized Privacy-Preserving Proximity Tracing
EBID	Ephemeral Bluetooth® IDentifier
ESR	Exposure Status Request
GAEN	Google Apple Exposure Notification
GDPR	General Data Protection Regulation
GUID	Globally Unique Identifier
HSM	Hardware Security Module
HTTP	Hypertext Transfer Protocol
IP	Internet Protocol
JSON	JavaScript Object Notation
MB	Mbits/s
NTP	Network Time Protocol
PKI	Public Key Infrastructure
QR	Quick Response
TAN	Transaction Authentication Number

TEE	Trusted Execution Environment
TLS	Transport Layer Security
URL	Uniform Resource Locator
UTF	Unicode Transformation Format

4 Definition and Extent of Back-End Mechanisms

4.1 Mapping to Reference Model from Requirements Document

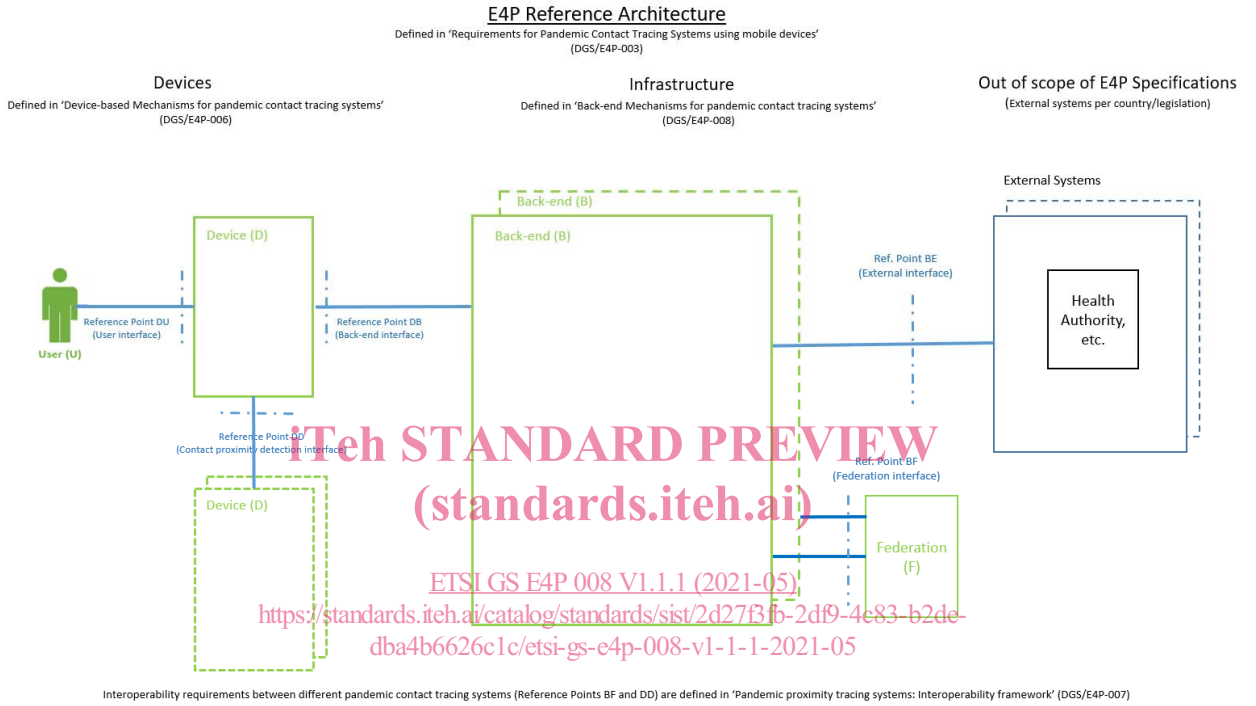

Figure 1: E4P Reference Architecture

The present document covers all back-end mechanisms necessary to make a digital contact tracing system functional. Depending on the nature of the system, certain functions reside in the central infrastructure or in the device.

Functions west-bound from the reference point "DB" are covered in ETSI GS E4P 006 [2].

Mechanisms that facilitate interoperability between digital contact tracing systems, east-bound from the reference point "BF", like a federation function between systems using the same design approach, or between systems with different design approaches, are addressed in ETSI GS E4P 007 [3].

Digital contact tracing systems need to be integrated into a regional or national regime of handling a pandemic. Functions in this context can be e.g. testing or setting of parameters to calibrate or adapt the system. These functions are eastbound of the reference point "BE" and out of scope of the ISG E4P.

4.2 Systems and components out of scope

Deployment of a Digital Context Tracing System (DCTS) requires a suitable technical and organizational integration with regional or national authorities and health entities, part of or working under authorization of the public health authorities (in the scope of overall contact tracing initiated by public health entity/ies).

DCTS provides the technical means to securely measure, collect, calculate, and communicate data according to and in the context of an overall scheme defined by epidemiological and legal experts that have the respective scientific and formal authorizations. The definition of the parameters to tune a DCTS, e.g. distance and time of proximity to be measured, risk calculation algorithms, legal requirements on data handling, etc. are defined and decided on outside of the DCTS. The E4P Reference Architecture refers to this complex as "External Systems" and defines the Reference Point BE (Back-End - External Systems) for the respective information flows.

Figure 2 shows some examples of information flow types initiated by External Systems that might or might not be handled by the Infrastructure as defined in the present document. If the information flow handled by the Infrastructure the information has to make use of the Reference Point BE. An example from the present document would be clause 6.1.6.

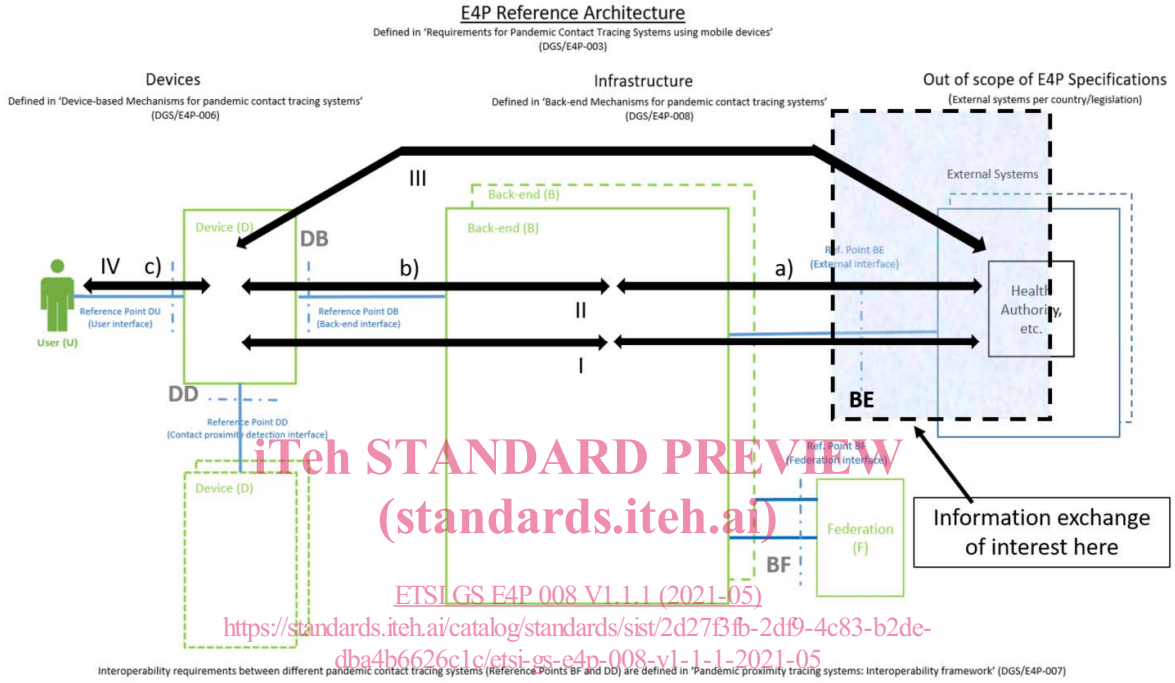

Figure 2: E4P High-level Reference Architecture indicating different types of information flows with External Systems

"Receive Test-Result" is informally represented in Figure 2 by the arrows II indicating the steps:

a) test result received by the Back-end;

b) test result retrieved by the Device;

c) test result displayed and acted on by the user.

In case of a positive test result, the release of 'diagnosis keys' is initiated in one of two ways: by an action by the user or by an action by the App software on the device; in the latter case the DCTS App informs the user.

The arrow I represent an information flow of the type "Configuration-Parameter Download", where the parameters affect the functions performed by the device, but are not necessarily displayed to the user.

The arrow III represents an "out-of-band" flow, where information to the user is bypassing the DCTS infrastructure.

The arrow IV represents all interactions between the users and their smartphones in relation with DCTS; this includes interactions with the DCTS App, but may include interactions with other Apps, e.g. a browser or e-mail client used to obtain useful information related to the use of DCTS.

In case of the decentralized approach the downloaded Diagnosis Keys come together with a digital signature to allow the Mobile Device to verify, that the Diagnosis Keys are genuine and were not manipulated during transport. The concrete way how the Mobile App gets the signature test key is out of scope of the present document (see clause 6.1.3).

In clause 6.1.6, the reception of test results from the responsible authority shall be authenticated. How this authentication is done is out of scope of the present document.

In clause 6.1.7, it is out of scope of the present document how the health authority decides to issue a teleTAN in the case a laboratory does not support the procedure described in this clause.

In clause 6.1.8, the details regarding the configuration parameters to alter/tune the risk calculation algorithm are out of scope of the present document.

5 Back-end mechanisms for pandemic contact tracing systems

5.1 Decentralized approach

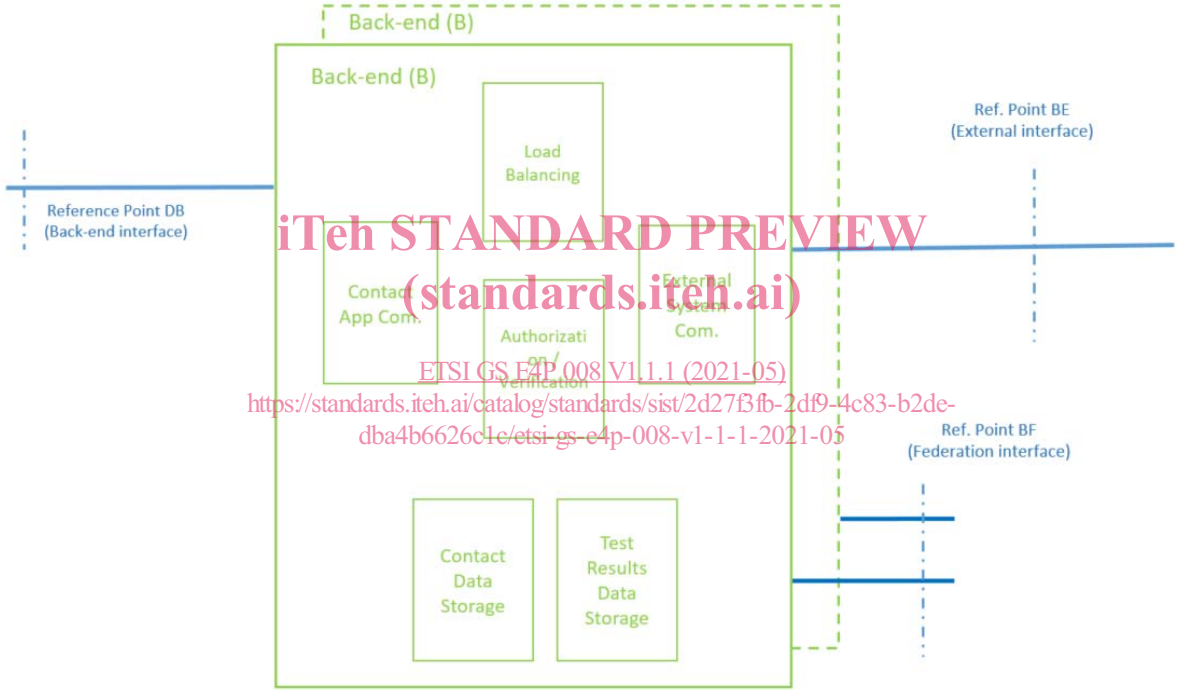

Figure 3: Back-end mechanisms overview (decentralized)

External System Communication

The External System Communication component interacts with external systems (e.g. Health Authorities, Test Laboratories) in order to acquire or provide necessary information (e.g. test results, parameters to tune the overall system or the exposure calculation formula, etc.) to make the overall Contact Tracing System functional.

NOTE 1: The E4P Reference Architecture also shows a Reference Point BF towards functions or modules realizing federation between two or more Digital Contact Tracing Systems. However, no federation specific functions inside a Back-end are described in the present document. For details regarding federation see ETSI GS E4P 007 [3].

Contact App Communication

The Contact App Communication component interacts with the Mobile Device to Upload Diagnosis Keys (e.g. if User is diagnosed) and to download periodically the Diagnosis Keys (needed to calculate the Users exposure to other diagnosed Users) and parameters to tune the risk calculation. In the latter case a Load Balancing Function may be used to cope with the number of mobile devices. The Contact App Communication component is also used to perform some "dummy" communication between Mobile Devices and the Back-end to make the Upload unobservable.

Authorization/Verification

The Authorization/Verification component authorizes specific Users for specific actions (e.g. Uploads) by appropriate means (e.g. one-time PINs) and verifies the correct authorizations in the respective steps.

Contact Data Storage

The Contact Data Storage component stores the sets of Diagnosis Keys the DCTS has to handle and provides access to them according to the specified protocols and rules.

Test Result Data Storage

The Test Result Data Storage component stores the test results necessary to forward to Users and provides access only to verified authorizations.

Load Balancing (e.g. CDN)

The Load Balancing component is an optional function that replicates needed resources in order to guaranty maximum availability of the DCTS. Below are some illustrative calculations regarding the expected amount of data which needs to be transmitted to each Mobile Device in case of a decentralised DCTS.

NOTE 2: The following calculation is based on GAEN [i.1].

In the decentralised system design each Mobile Device locally checks if the device was in close proximity to other mobile devices of diagnosed users. Therefore, each Mobile Device needs to know all the relevant diagnosis keys from diagnosed users. In the current system design "relevant keys" usually refers to the diagnosis keys used by an diagnosed user during 14 days (depending on the epidemiological necessities, e.g. at least 14 days for COVID-19) prior to the point in time at which the user uploads his/her diagnosis keys to the back-end.

Moreover, the diagnosis keys change on a daily base, i.e. there is one diagnosis key per mobile device per day. Taking this as input for the calculations and assuming that each user uses only one mobile device, each newly diagnosed user will upload 14 diagnosis keys to the back-end, which need to be distributed to all other mobile devices.

A single diagnosis key consists of 16 random bytes. Therefore, each newly diagnosed user contributes at least $14 \times 16 = 224$ random bytes to the amount of data which need to be distributed to each other mobile device per day. Note that because of the random nature of these bytes they cannot be compressed. Assuming 10 000, 100 000 or 1 000 000 newly diagnosed users this will lead to an amount of data of 2,24 MB, 22,4 MB or 224 MB respectively.

Encoding the diagnosis keys using appropriated data structures will require some extra bytes. Given that the protocol buffer data structure `TemporaryExposureKey` as describe in clause 6.1.1 is used for encoding diagnosis keys, the additional overhead will be roughly 11 bytes per diagnosis key.

These bytes can be compressed making an exact calculation of the amount of data traffic necessary to transmit the diagnosis keys difficult. In a worst case scenario it would need $14 \times (16 + 11) = 378$ bytes per newly diagnosed user. So, for 10 000, 100 000 or 1 000 000 newly diagnosed users this will lead to an amount of data of 3,78 MB, 37,8 MB or 378 MB respectively.