Documentation

Name Server Predelegation Check

Doc. version:	1.4.1	**Doc. status:**	Final
Doc. date:	01.12.2015	**Doc. name:**	Name Server Predelegation Check-Documentation-DNS Services-V1.4.1-2015-12-01

Imprint

Authors	Department	Phone	E-mail
DNS Services	DNS Services	+49-69-27 235 272	info@denic.de

Document Release

Document version	Released by	Released on
1.4.1	DNS Services	01.12.2015

Distribution List

Name
Members
Public

Inhalt

1. General Information

1.1 About this Document

This document describes the requirements for name servers and zone data to be met so that DENIC can delegate a domain to those name servers. The policy and the related checks are explained in this document.

1.2 Reason

The *Domain Name System* (DNS) is a hierarchically structured, distributed and replicated database for mapping names in addresses and other Internet infrastructure elements. The hierarchical structure and the distribution are implemented by means of *delegation*. The *root* of the namespace is immediately followed by the sub-level of the *Top Level Domains* (TLDs). The rules for the administration of the TLDs are laid down in the RFC 1591 and other documents. The TLD administrator is responsible for reliable maintenance and care of the zone data in accordance with the latest state of the art and for securing operation of the corresponding name servers according to *best practices*.

Due to its distribution and redundancy the DNS is highly fault-tolerant. Nevertheless, network disturbances and configuration errors may cause a situation in which domains are resolved incorrectly or not at all (refer also to RFC 4697[1]). Although the Internet user will notice the failure, they cannot always identify the DNS as its cause. If the constellation is particularly disadvantageous, DNS failures may trigger disturbances in sections of the network that are neither concerned by the cause nor responsible for it, let alone able to eliminate it. Therefore, the TLD administrators verify prior to delegation that certain requirements are fulfilled in order to guarantee proper functioning of the overall system and reliable service – local and remote - for all requestors from the namespace. Delegation is not implemented, unless all requirements are met. There are a number of additional criteria, which are not absolutely necessary, but which provide an improved service quality, if met.

[1] RFC 4697 Observed DNS Resolution Misbehavior (http://www.ietf.org/rfc/rfc4697.txt)

2. Name Server Predelegation Check

2.1 The New Warning and Error Messages Issued in the Course of the Name server Predelegation Checks (Name Server Policy)

Obligatory requirements are described using the word "must" or "must not". Infringements of such rules trigger an ERROR message.

Recommendations are described using the word "should". Infringements of such rules trigger a WARNING message.

2.1.1 General Requirement

All name servers stated in the request must be reachable and authoritative for the requested zone [if not, output of ERROR message].

Explanation: Since the operator of the name server may not be identical with the domain holder nor the adminstrating DENIC member and is thus no contractual partner of DENIC, and since a delegation is to his/her disadvantage, this rule shall ensure that the delegation can be considered approved by the name server, if the zone is served authoritatively. In the remaining aspects, the predelegation check corresponds with the spirit and the text of the RFCs 1034[2] and 1035[3].

2.1.2 Redundant Connection

1. One prerequisite for delegation is the existence of at least two name servers. At least one of them must be connected via IPv4. To carry out the further steps of the check, all IPv4 and IPv6 addresses of each name server are identified or, if feasible, deduced from the request. The request must contain at least one name server with an IP address that differs from all other name servers of the same request. [if not, output of ERROR message].

   Explanation: RFC 1035 expressly requires that all DNS zones are redundantly supported by at least two independent name servers. To avoid negative effects on the TLD servers (see RFC 4697) if one of the name servers of the delegated zone cannot be reached, particular importance is attached to diversity of network topology.

   Example:

   permitted:

   a. *name server 1 with the IP adress 192.0.2.1 and name server 2 with the IP address 198.51.100.1*
   b. *name server 1 with the IP adress 192.0.2.1 and name server 2 with the IP address* 2001:db8:85a3::8a2e:370:7334

   not permitted:

   c. *name server 1 with the IP adress 192.0.2.1 and name server 2 with the IP address 192.0.2.1*
   d. *name server 1 with the IP adress* 2001:db8:85a3::8a2e:370:7334 *and name server 2 with the IP address* 2001:db8:85a3::8a2e:370:7336

[2] RFC 1034 Domain names - concepts and facilities (http://www.ietf.org/rfc/rfc1034.txt)

[3] RFC 1035 Domain names - implementation and specification (http://www.ietf.org/rfc/rfc1035.txt)

2.1.3 Glue Records

1. In principle, the Narrow Glue Policy shall apply: Glue records are entered in the .de / 9.4.164.arpa zone if and only if the name of a name server is definitely located within the delegated domain.
    a. If the name server is located in the domain to be delegated, at least one IPv4 / IPv6 address must be stated (A-/AAAA-RRSet) [if not, output of ERROR message].
    Explanation: The stated constallation requires at least one glue record.
    b. If the name server is not located in the domain to be delegated, but the request still contains at least one IP address (v4 or v6) of that server, this IP address will not be taken into consideration for the request [Output of WARNING message].
    Explanation: DENIC applies the Narrow Glue Policy for both .de and 9.4.e164.arpa, i.e. it accepts glue records only under specific, very restricted conditions, to be precise, if the name server is located within the delegated zone. Possibly stated additional addresses are superfluous and will not be considered. The warning message shall draw the attention to potential input errors.
2. The A- and AAAA-RRSet of a name server must be directly, fully, consistently and authoritatively ascertainable via every IP address (v4 or v6) stated in the request, and the sets must be identical with the data of the request [if not, output of ERROR message].

    Explanation: Since glue data co-exist with the authoritative data, it must be ensured that they are consistent, i.e. that the address data in the glue records are identical with the authoritative address data that have been determined by the "standard way". Moreover, the rule of consistency requires that RRSets (Record Sets, i.e. one or several records of the same type[4]) are always stated in full. That means, for example, that you cannot state only one or two addresses of a name server for the glue records but must include all of them. Last but not least the "Narrow Glue Policy" applies to IPv4 as well as to IPV6 addresses, i.e. if corresponding record sets (A or AAAA) exist in the authotitative data, they must be made availlable in glue records.

[4] RFC 2181 Clarification to the DNS Specification (http://www.ietf.org/rfc/rfc2181.txt)

2.1.4 Requirements Concerning the SOA Data in the Zone

1. Obsolet: The SOA serial number on all name server zones should tally [if not, output of WARNING message].
2. For the following SOA values, guideline values are specified:
    a. "Refresh" should be within the following range: 3600 – 86400 (in seconds) [if not, output of WARNING message].
    Explanation: This value determines the frequency of data reconciliation of the secondary name servers and the primary master. Lower values generate increased DNS traffic and boost the load of the systems involved, higher values may have a negative impact on the up-to-dateness of the data. Since the values must finally be agreed by the operators and the participating name servers, a warning is issued only if the chosen value falls short of or exceeds the "standard" values.
    b. "Retry" should be within the following range: 900 – 28800 (in seconds) [if not, output of WARNING message].
    Explanation: This value replaces the value stated under "Refresh" after the first unsuccessful attempt until either successful reconciliation is made or the "Expire" value is reached. So the value must be below the "Refresh" value. It must be taken into consideration, however, that too small a value may cause load peaks and will trigger a warning message.
    c. The "Retry" value should be 1/8 and 1/3 of the "Refresh" value [if not, output of WARNING message].
    Explanation: This ensures that the relation between the "Refresh" and the "Retry" value is appropriate for the changeover logic to bring about a noteworthy advantage.
    d. "Expire" should be within the following range: 604800 – 3600000 (in seconds) [if not, output of WARNING message].
    Explanation:This value defines the period of unsuccesful reconciliation attemps until a slave stops to further support the zone. Values below one week are very problematic because they may lead to the loss of all authoritative name servers of a zone within a short period, so that the zone remains completely paralyzed due to delegation. 1,000 hours as an upper limit is a frequently used value. If this limit is exceeded, a severe reconciliation problem must be assumed, which should not be ignored.
    e. "negTTL" must/should be within the following range: 180 - 86400 (in seconds) [if not, a WARNING message is issued].
    Explanation: Together with the TTL of the SOA record, this value determines the useful life of negative replies pursuant to RFC 2308. Values that are two high (in this case more than one day) do not noticeably reduce the DNS traffic and/or are curtailed by DNS caches, anyway. So they would be inefficient. Values that are too low (in this case less than three minutes) finally lead to a complete shut down of the "negative caching", which is meant to be avoided.

2.1.5 Requirements Concerning the Further Data in the Zone

1. The NS-RRSet must correspond exactly to the list of name servers stated with the request [if not, output of ERROR message].

   Explanation: RFC 1034 stipulates that the authoritative name server data used in the delegating and the delegated zone must be identical.

2. The existence of a CNAME-RR is not permitted for the requested zone (more precise: at the zone apex) [if not, output of ERROR message].

   Explanation: One and the same node in the DNS tree must not have any additional record types besides the CNAME record. Since a delegated zone requires at least the SOA record and the NS records, the existence of a CNAME record would represent an infringement of the protocol.

3. The referral response (assuming a QNAME of up to 191 Bytes length and counting all required address data, including glue records) must fit into an DNS-UDP package, i.e. must not exceed 512 bytes [if not, output of ERROR message].

   Explanation: The name server of DENIC answers requests for data in delegated zones with a reference to the actually responsible name server on the next hierarchical level. Standard UDP packages allow a useful load of 512 bytes at maximum. To avoid that the answers are curtailed and the question is then resubmitted via TCP, thus leading to a disproportionate workload of the DENIC name server, the above mentioned length restriction is introduced. Since the space consumed is determined by the length of the name server names and their compressibility as well as by the number of the glue records, this calculation method is safer than specifying a maximum limit for the number of name servers.

4. The primary name server stated for the SOA-RR of the requestes zone should be identical on all name servers [if not, output of WARNING message].

   Explanation: This is an additional means to verify the consistency addressed under 2.1.4 (1).

2.1.6 Other Defaults for the Name Servers

1. IPv6 addresses must originate in an address space that is marked as Global Unicast and as ALLOCATED[5] and routable[6]. This applies for all IPv6 addresses of the stated name servers, regardless of the fact whether the address is a glue record or not [if not, output of ERROR message].

   Explanation: IPv6 knows various validity ranges for addresses ("Scoping"). To make the check results unambiguous and comprehensible and to secure uniform accessibility of the name servers throughout the world, only those addresses are accepted that are globally unambigous.
2. Recursive request should not be permitted [if not, output of WARNING message].

   Explanation: For safety reasons and to ensure a correct view on the namespace, authoritative and recursive name servers are strictly separated in the operational praxis.
3. Reachability via TCP should be guaranteed [if not, output of WARNING message].

   Explanation: RFC 1034 and 1035 specify for DNSs the use of UDP as well as of TCP transport, with UDP having priority and being used for the major share of the data traffic. Under certain circumstances (e.g. response size) a resolver may have to switch to TCP, which is expressly supported by RFC 1123.

[5] IANA IPv6 Global Unicast Address Assignments"

http://www.iana.org/assignments/ipv6-unicast-address-assignments/ipv6-unicast-address-assignments.xml

[6] "IANA IANA IPv6 Special Purpose Address Registry"

http://www.iana.org/assignments/iana-ipv6-special-registry/iana-ipv6-special-registry.xml

3. Name Server Predelegation Check for DNSSEC

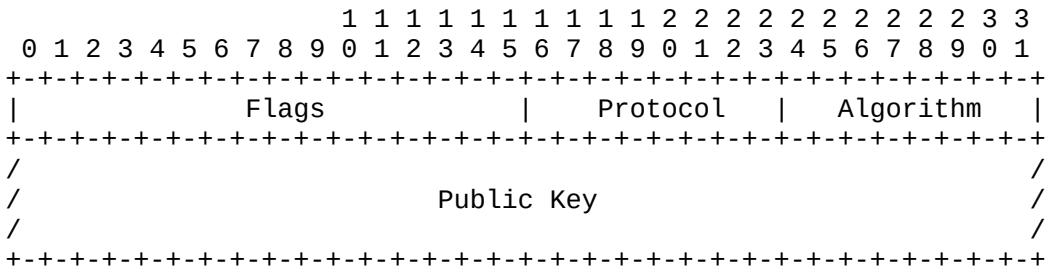
3.1 Technical Background

For validating DNSSEC, a chain of trust must be established. To meet this requirement, the DNSSEC protocol requires that a reference to the key(s) of the delegated zone is entered into the delegating zone. Thus, the chain of trust follows the delegation path. The relevant key is the Key Signing Key of the delegated zone, which is usually flagged as Secure Entry Point (SEP). This information is available in a DNSKEY-RR in the delegated zone and is signed here (at least) by this very DNSKEY. It is not repeated in exactly the same way in the delegating zone for reasons of space. Instead of the actual key, a corresponding finger print is stored in a DS-RR (Delegation Signer).

Provisioning of the key material is effected by means of the DNSKEY-RR.
For each domain, up to five DNSKEY-RRs may be recorded. The corresponding DS records[7] are generated together with the zone (currently precisely one DS-RR per communicated Trust Anchor), signed and distributed with the zone.

3.2 Structure of the Dnskey Record

The format of the Dnskey record is described in chapter 2 of RFC 4034:



Under the flags, you will find one entry respectively for the fields "Zone Key" and "Secure Entry Point". Additional extensions for an automated update of a Trust Anchor are described in RFC 5011 [8]. In the "Algorithm" field the applied public key algorithm is defined. This key determines the internal structure and the size of the actual key data. Additionally, this field is used to signal the use of NSEC3 by means of alias mechanisms.

[7] Cf. RFC 4034 Resource Records for the DNS Security Extensions (http://www.ietf.org/rfc/rfc4034.txt)

[8] Cf. RFC 5011 Automated Updates of DNS Security (DNSSEC) Trust Anchors (http://www.ietf.org/rfc/rfc5011.txt)

3.3 Differentiating between KSK and ZSK

When the DS-RR was initially introduced by the RFC 3658[9], one started to differentiate between the Zone Signing Key (ZSK) and the Key Signing Key (KSK). While the first one sings the actual data (RRSets) in the zone, the latter serves exclusively for authenticating the ZSK. This differentiation allows for the different requirements made to the different keys.

To update the KSK, interaction with the delegating zone is required. Thus, it should not be made too often and needs a key with a longer term of validity (and, in most cases, consequently also more characters). Updating the ZSK is much easier. Thus, this key will usually be shorter. At least with regard to the RSA procedure a shorter key will entail shorter signatures and thus smaller zone files, which in turn will lead to smaller response packets. You will find information about key length and how to update a key in RFC 4641[10].

While the separation makes it easier to select the parameters on the one hand, it makes the protocol more complex on the other. However, separation of the keys is not mandatory. Using only one key instead of a KSK/ZSK key pair - putting up with the above described disadvantages - is consistent with the protocol. Sometimes this option is chosen in the practice. Theoretically, other indirections could be integrated in the key relations. However, since a signature always covers a complete RRSet, any signature on the DNSKEY-RRSet must always authenticate all the keys included in that set. Thus, only the two options ZSK+KSK and ZSK=KSK need to be taken into consideration.

3.4 Utilizing and Publishing Keys

A DNSKEY-RR provisioned for registration is called visible if it is included in the DNSKEY-RRSet of a zone.

3.5 Concept for Proof of Possession

For Certification Authorities (CAs), the concept of the proof-of-possession is of importance. This concept requires that a party requesting a certificate submits a proof showing that it actually has access to the private key associated with the public key for which the certificate is requested. For security reasons, it is recommended to employ this practice also for DNSSEC (cf. chapter 3.6.3).

3.6 Technical Prerequisites and Checks

Before recording the key material in the database, and to ensure correct further processing and proper operation, the key material is subjected to a multi-stage check. First of all, the transferred data itself is checked, subsequently, further tests including the information that can be called in the DNS are carried out.

3.6.1 Parameters of the Provisioned DNSKEY Records

The keys handed over for registration must be unique, i.e. they must differ from one another in at least one field [if not, output of ERROR message]. It may happen that keys (Public Key) are provisioned several times with varying flags. Key tags are not checked on uniqueness [if not, output of ERROR message].

[9] Cf. RFC 3658 Delegation Signer (DS) Resource Record (RR) (http://www.ietf.org/rfc/rfc3658.txt)

[10] Cf. RFC 4641 DNSSEC Operational Practices (http://www.ietf.org/rfc/rfc4641.txt)

3.6.1.1 *DNSKEY: Flags*

In the *Flags* field you must only use bits that are flagged as assigned in the IANA registry [if not, output of ERROR message].

The field is provisioned only as a numerical value (0-65535).

1. Bit 7 (ZONE) must be set. [if not, output of ERROR message].

   Specified in RFC 4034, chapter 2.1.1.

2. Bit 8 (REVOKE) must not be set. [if not, output of ERROR message].

   Follows from chapter 2.1 of RFC 5011. A called-back key cannot be used as Trust Anchor.

3. Bit 15 (SEP) should be set. [if not, output of WARNING message].

   This field is intended to identify the KSK in the DNSKEY-RRSet. It is considered good practice to set this field for KSKs or for Trust Anchor, even if validators shall not evaluate it.

Thus, you can currently use the values 256 (ZONE) and 257 (ZONE, SEP). However, only the latter is accepted without warning.

3.6.1.2 *DNSKEY: Protocol*

The *Protocol* field must have the value "3". Section 2.1.2 of RFC 4034 stipulates this value as mandatory.

3.6.1.3 *DNSKEY: Algorithm*

The *Algorithm* field may contain the values included in the IANA registry and marked there as available values not reserved for private use.

At present, these are the values 3, 5, 6, 7, 8, 10, 12, 13, 14. The only cryptographic algorithms defined are RSA, DSA, ECDSA and GOST in combination with various hash methods.

3.6.1.4 *DNSKEY: Public Key*

The *Public Key* field contains the Base64-encoded version of the public key[11]. The internal structure is determined by the applied algorithm, and so are the corresponding checks:

RSA (relevant for the values 5, 7, 8, 10)

1. The modulus must be 512 to 4096 bit (stated numbers included) long [if not, output of ERROR message].

2. The exponent must not be longer than 128 bit [if not, output of ERROR message].

   The limit values are defined in RFC 3110[12]. As to the exponent, they are also defined by the restrictions to the line length applicable for provisioning.

   The key size is not evaluated.

DSA (relevant for the values 3,6)

[11] Cf. RFC 4648 The Base16, Base32, and Base64 Data Encodings (http://www.ietf.org/rfc/rfc4648.txt)

[12] Cf. RFC 3110 The RSA/SHA-1 SIGs and RSA KEYs in the Domain Name System (DNS) (http://www.ietf.org/rfc/rfc3110.xhtml)

3. The T-parameter may have values between 0 and 8 (both values included) [if not, output of ERROR message].
4. The length must be 213 + T * 24 [if not, output of ERROR message].

   These limit values are defined in RFC 2536[13].

   The key size is not evaluated.

ECDSA (relevant for the values 13, 14)

5. If the algorithm ECDSAP256SHA256 (13) is used, the key length must be 512 bit [if not, output of ERROR message].
6. If the algorithm ECDSAP384SHA384 (14) is used, the key length must be 768 bit [if not, output of ERROR message].

   These values can be deduced from RFC 6605[14], chapter 4.

GOST (relevant for the value 12)

7. The key length must be 512 bit bit [if not, output of ERROR message].

   This value can be deduced from RFC 5933[15], chapter 2.

[13] Cf. RFC2536 DSA KEYs and SIGs in the Domain Name System (DNS) (http://www.ietf.org/rfc/rfc2536.txt)

[14] Cf. RFC 6605 Elliptic Curve Digital Signature Algorithm (DSA) for DNSSEC (https://tools.ietf.org/rfc/rfc6605.txt)

[15] Cf. RFC 6605 Use of GOST Signature Algorithms in DNSKEY and RRSIG Resource Records for DNSSEC (https://tools.ietf.org/rfc/rfc5933.txt)

3.6.2 Visibility and Status of the DNSKEY Records

The DNSKEY-RRSet must be identical for all authoritative servers [if not, output of ERROR message].

At least one of the keys provided for registration must be visible [if not, output of ERROR message]. For every invisible key a WARNING will be issued. Possibly existing additional keys of the DNSKEY-RRSet will not be taken into consideration.

It is generally assumed that the signatures obtained from the various servers are identical. However, this is neither explicitly requested nor verified. In particular for the DSA method, this approach makes online signing possible.

3.6.3 Use of the DNSKEY Records

At least one visible key provisioned for registration must validly sign the DNSKEY-RRSet [if not, output of ERROR message].

This requirement helps to implement the *Proof of Possession*.

3.6.4 Validating with DNSKEY Records

There must be a currently valid validation chain for the SOA-RR of the delegated zone with at least on visible key provisioned for registration [if not, output of ERROR message].

Every zone has a SOA-RR, which is checked by default within the scope of the predelegation check. The validation check is carried out to prevent security lameness.

3.6.5 Comprehensive Rules

Due to the DNSSEC requirements, the authoritative servers and the related infrastructure must satisfy additional needs not covered by the checks of the zone data:

1. All authoritative servers must support DNSSEC, i.e. they must deliver responses consistent with DNSSEC to DO-bit-signed requests [if not, output of ERROR message].
2. All authoritative servers should support TCP and EDNS0 of adequate packet sizes [otherwise a WARNING message will be output].
3. It must be possible to call the signed DNSKEY-RRSet at least via one of these channels (either via TCP or via EDNS0) [if not, output of ERROR message].

The predelegation checks already include a verification of DNS support via TCP.
Here, the EDNS0 variant is added to cover DNSSEC. As described in chapter 2, initially only one of the two check variants must be supported. However, a warning is issued if a second variant is missing. This is to account for the potential consequences with regard to operation and the infringement of chapter 3 of the RFC 4035 (formerly RFC 3226). Item 3 finally stipulates the requirement that must be met anyway with regard to the aforementioned checks: access to the signed DNSKEY-RRSet.