# **#34230 CPE的WAN口配置的多个DNS支持主备切换**

# **1. 版本修订和Review记录**

**修订记录**

|  |  |  |  |
| --- | --- | --- | --- |
| 版本记录 | 版本编号 | 作者 | 说明 |
| 2023-04-23 | V1.0 | @张克尧 | 初稿 |
| 2023-05-11 | V2.0 | @赵昱钦 | 补充4.2.4和遗留问题 |
| 2023-10-24 | V3.0 | @赵昱钦 | 新增4.2.5 （#48748） |

**Review记录**

|  |  |  |  |
| --- | --- | --- | --- |
| **角色** | **Review人** | **Review通过时间** | **备注** |
| 开发 |  |  |  |
| 开发 |  |  |  |
| 安全 |  |  |  |
| 测试 |  |  |  |
| 测试 |  |  |  |

# **2. 术语**

# **3. 需求场景描述**

WAN下虽然添加了多个DNS，但是在目前的实现中，在大部分情况下，实际上在Datapath上只使用第一个DNS。只有在报文的五元组+Transaction ID完全相同时，才会使用第二个DNS。这就导致了当主DNS如果出现故障时，大部分涉及到DNS解析的业务都会出现问题，这其中包含CPE访问Communication Server以及IPdetect的流量，也包含了用户访问互联网的业务流量。

对于公共DNS服务而言，由于这些公共DNS本身足够健壮，一般而言其出现故障的概率不高。但是对于一些运营商自建的DNS，或者是企业自建DNS，其可靠性就远没有那么高。在招商局的项目中，使用了互联网收口，所有Spoke访问互联网的流量都要从Hub走出去，而Hub使用的DNS Server是自建的，并且并没有使用VIP的架构，在这种场景下，一旦主DNS故障，用户所有Spoke访问互联网的流量都会出现问题，这将造成比较大的线上事故。

因此，我们需要在设计上解决主DNS的单点故障问题。由于我们Datapath目前对于DNS报文只做转发，而不会产生新的数据包，我们暂时不考虑在Datapath复制多份数据包的处理模式（这就更类似于完整的Relay功能了，但是这破坏了一对一的原则，变成了一对多，这就不符合现有DATAPATH基于流的基本处理原则）。所以剩下的思路就是通过对WAN上多个DNS Server进行主动探测，并在发现主DNS Server故障，而备DNS Server正常时，能够切换主备，让Datapath使用正常的DNS Server去做解析。

# **4. 功能设计描述**

## **4.1 功能整体逻辑简述和架构设计图**

无

## **4.2 功能细节逻辑描述**

### **4.2.1 Agent新增WAN DNS探测**

在Agent中，对于每个WAN的DNS Server进行探测（这块可以考虑复用探测任务这部分的代码）。使用的域名可以选择IPDetect使用的域名（即Orch的域名）。如果ipdetect是IP，而不是域名，则不进行该探测。这里使用的就是zone2所下的用于ipdetect的Policy，所以不用额外下发Policy。

用户的目标是在10s左右监测出故障，并做切换。但是我们应当考虑单个DNS Request收不到其实并不能完全说明问题，因此一般需要连续若干个DNS Request未收到才应该做主备切换。总体思路考虑：

1. 默认探测周期为2s。即每2s，针对包含了DNS服务器配置的WAN（只要DNS Server不为0，即视为存在DNS Server），向每个DNS Server发送域名为Orch域名的DNS Request，并记录DNS Reply的情况，注意的是这里源地址也应该指定成对应的WAN的地址；

2. 若某个WAN上的主DNS Server连续5次未收到DNS Reply，且备DNS Server5次中收到过DNS Reply，则通知NetIO做主备切换，并向Orch告警；

3. 若某个WAN上主DNS Server在最近100次DNS请求中，有超过10次没有响应，且其未响应数量超过备DNS的2倍以上，则通知NetIO做主备切换，并向Orch告警；

4. 若Orch更新了WAN使得其上的DNS Server发生了变化（不包含仅顺序发生变化，以后我们应该把Orch下发的两个DNS Server看作是对等的），则清空该WAN上关于DNS的现有探测数据。

5. 最好给该功能加一个开关，如果出现问题，可以通过开关关闭

### 4.2.2 NetIO部分

1. ~~NetIO需要为Agent提供对某个WAN进行主备切换的接口。~~NetIO已有为Agent提供WAN主备切换的接口。

2. NetIO在更新WAN配置时，检查DNS Server与之前的配置是否相同（顺序不同也应当视为是相同的），如果DNS Server相同，则不应当更新DNS Server的配置包括顺序

### 4.2.3 Flow生成部分

Flow生成的时候，针对DNS报文，如果其在匹配完一个域名的Policy之后，如果其Outgoing是WAN，且DNS报文的目标地址正好属于该WAN的一个DNS Server，那么就不再对该数据包做DNAT。

### **4.2.4 DnsSwitch功能模块描述**

1. cpeagent.log过滤“Dns switch”查看模块日志
2. DNS包遵循4.2.1中规则1，2s内未收到reply包意为超时；
3. DNS主备切换遵循4.2.1中规则2，规则3
4. 提供cpeagent dnsswitch start/stop开关，后期可增加orch配置，规则5
5. 仅考虑vpnid为0的情况，非0时WAN不发生DNS主备切换

注意：CPE login前，DNS1无效时切换不到DNS2（因WAN未下发，无法注册到DnsSwitch模块）

6、Dns switch日志

**1）使用cmd line方式enable/disable时**

Dns switch module running/stoping

**2) 重置计时器相关日志**

重置计时器时输出相关参数

Dns switch reset count of node, WanId:%d PhyIfName:%s SendCnt:%d MasterDNS:"LW\_HIPQUAD\_FMT"(RecvCnt:%u ConsLossCnt:%u) SlaveDNS:"LW\_HIPQUAD\_FMT"(RecvCnt:%u ConsLossCnt:%u)

发包到达最大值（100）

Dns Switch send %d pkt to reset count(WanId:%d PhyIfName:%s)

主DNS丢包数量大于备DNS

Dns Switch loss more than %d pkt to switch master DNS(WanId:%d PhyIfName:%s)

备DNS连续丢包过多（大于等于5个）

Dns switch loss more than %d consecutive pkt of slave DNS(WanId:%d PhyIfName:%s)

主DNS连续丢包过多（大于等于5个，且备小于5个）

Dns switch loss more than %d consecutive pkt to switch master DNS(WanId:%d PhyIfName:%s)

主备DNS配置相同（忽略顺序改变）

Dns switch clean count and keep same conf with orch(WanId:%d PhyIfName:%s MastrtDNS:"LW\_HIPQUAD\_FMT" SlaveDNS:"LW\_HIPQUAD\_FMT")

wan注册（或更新DNS）

Dns switch wan regist or update(WanId:%d PhyIfName:%s MasterDNS:"LW\_HIPQUAD\_FMT" SlaveDNS:"LW\_HIPQUAD\_FMT"

**3） wan更新netio主dns**

Dns switch update wan dns success, PhyIfName:%s from "LW\_HIPQUAD\_FMT"(%u) to "LW\_HIPQUAD\_FMT"(%u)

**4）dns wtich告警（Type=413, ifStatus=6）**

Send Msg: Type = 413 netId:0 transactionId:46 msgBase{ interfaceReport{ ifStatus:6 wanId:1 ifName:"eth1" dnsList:1920103026 dnsList:16908802 } }

Recv Msg: Type = 414 netId:0 transactionId:46 errorCode:0 msgBase{ interfaceReport{ ifStatus:6 wanId:1 ifName:"eth1" dnsList:1920103026 dnsList:16908802 } }

7. Dns switch Orch告警

例如：Send Msg: Type = 413 netId:0 transactionId:46 msgBase{ interfaceReport{ ifStatus:6 wanId:1 ifName:"eth1" dnsList:1920103026 dnsList:16908802 } }

使用413消息，\_LW\_IF\_STATUS枚举中增加LW\_IF\_STATUS\_DNS\_SWITCH（6），示例中1920103026是发生主备切换后的主DNS。

Orch展示可增加CPE WAN DNS Switch类型告警，表示CPE的DNS发生主备切换，从16908802切换为1920103026

### 4.2.5 DnsSwitch功能拓展（#48748）

#### **经项目反馈，需在原设计基础上拓展，修改内容如下**

* + 支持POP平台
  + 功能默认打开
  + 下发探测域名配置（Orch需要关注）
  + 增加DNS状态上报（Orch需要关注）
  + 增加DNS优先级
  + 增加多DNS场景

#### **DNS校验规则**

##### **连续丢包规则（连续丢包超过五个）**

* 1. 给所有连续丢包超过5个的DNS状态设为不可用。并找出主DNS和第一个可用的备DNS。
  2. 主可用时，则结束。不结束本轮探测。
  3. 主不可用时。如果备DNS可用，则切换到第一个可用的备DNS；如果备也都不可用，则结束。同时完成本轮探测。

##### **2.2 总丢包规则（总丢包大于10为不可用，主比备丢包数量的二倍还多则发生切换）**

* 1. 给所有总丢包数量大于等于10的DNS状态设为不可用，小于10的设为可用。并找出主DNS和第一个可用的备DNS。
  2. 主可用时。如果备可用，则比优先级，优先级高的设置为主DNS；如果备都不可用，则结束。同时完成本轮探测
  3. 主不可用时，如果备可用，则切换；如果备也不可用，则找备中第一个是主丢包数量一半以下的DNS切换，如果没有，则结束。同时完成本轮探测

#### **POP支持DNS主备切换**

需求来源同#34230 CPE支持DNS主备切换，需在POP平台开放功能宏即可（不涉及Orch）

#### **DNS切换告警和状态上报**

* 1. 当发生DNS主备切换时，会发送告警消息到Orch，表示从DNS1切换为DNS2
  2. 当所有DNS状态有变化时，会发送DNS状态消息到Orch
  3. CPE每隔十秒共发送10次给Orch，收到Orch reply后停止发送，超过十次后会丢失消息。

#### **探测域名**

CPE或POP login时，Orch将默认域名下发到设备；当默认域名发生修改，设备重新login时生效。

#### **PB消息设计**

MR：<https://git.appexnetworks.com.cn/LightWAN/sharedoc/-/merge_requests/523>

## **4.3 兼容性考虑**

## **4.4 安全性考虑**

## **4.5 系统资源和性能影响**

## **4.6 CheckList**

|  |  |
| --- | --- |
| **检查项** | **已考虑（是/否/不涉及）** |
| 支持IPv6 | 暂不考虑支持IPv6的DNS |
| 支持Security Zone |  |
| 支持HA |  |
| 支持CPE集群 |  |

# **5. UI交互设计**

# **6. 诊断调试功能**

## **6.1 log日志**

## **6.2 xxx**

# **7. 产品规格**

## **7.1 支持平台**

功能特性1：

|  |  |  |
| --- | --- | --- |
| **设备型号** | **支持**  **(是/否/不涉及)** | **备注** |
| LW1305 |  |  |
| LW2X02 |  |  |
| LW2X08 |  |  |
| LW2X09 |  |  |
| LW300X |  |  |
| LW700X |  |  |
| vCPE |  |  |
| PoP |  |  |
| Windows |  |  |
| Mac |  |  |
| ios |  |  |
| Android |  |  |
| SDK（嵌入式设备） |  |  |

功能特性2：

|  |  |  |
| --- | --- | --- |
| **设备型号** | **支持**  **(是/否/不涉及)** | **备注** |
| LW1305 |  |  |
| LW2X02 |  |  |
| LW2X08 |  |  |
| LW2X09 |  |  |
| LW300X |  |  |
| LW700X |  |  |
| vCPE |  |  |
| PoP |  |  |
| Windows |  |  |
| Mac |  |  |
| ios |  |  |
| Android |  |  |
| SDK（嵌入式设备） |  |  |

## **7.2 规格列表**

|  |  |  |
| --- | --- | --- |
| **设备型号** | **规格** | **备注** |
| LW1305 | 不支持 |  |
| LW2X02 | 4 |  |
| LW2X08 | 4 |  |
| LW2X09 | 4 |  |
| LW300X | 8 |  |
| LW700X | 8 |  |
| vCPE | 8 |  |
| PoP | 128 |  |
| Windows | 不支持 |  |
| Mac | 不支持 |  |
| ios | 不支持 |  |
| Android | 不支持 |  |
| SDK（嵌入式设备） | 不支持 |  |

# 8. 功能缺陷或遗留问题

1、目前没有支持orch告警

2、目前没有支持Orch配置模块功能开关

# **9. 测试建议**

/\* 对qa测试的建议，可以列一些关键的和容易忽视的测试点

# **10. 其它**

/\* 撰写者可按实际需要增加相应章节以丰富文档内容

# 11. 会议纪要

/\* 相关讨论的会议纪要，可直接记录文字，也可附上链接

# **12.** 相关文档

/\* 将和本需求或文档相关的文档链接放在这里，方便查看

/\* 比如开发文档、参考资料等的链接