2023-4-1

企业 DNS 建设白皮书

F5 编发

目录

概述

Domain Name System（DNS）自 1987 年被实施以来（RFC1034，1035），已成为网络通信中最重要的核心基础设施。通过将人类难以记忆的通信端点信息转化为易读易记忆的域名，极大的简化了互联网通信。通过映射不变的名称到可变的端点信息，确保了通信双方在端点变化后依然可以快速找到对方。除了作为通信的基本技术核心外，DNS 也是企业对外提供永续数字服务的第一关键，无论企业花费多少资金投入在数据中心或应用的高可用上，如果没有正常 DNS 的解析服务，用户将无法访问到这些服务，企业经营也就无从谈起。可以说没有 DNS，就没有现代网络通信，没有 DNS 就没有企业数字服务。

尽管 DNS 已经存在 30 余年，技术已经非常成熟，但随着技术的变化与演进，我们可以看到 DNS 也在发生着诸多变化，总体来说 DNS 的发展有着以下“四高”趋势：

高安全，DNS 系统需要具备足够的抗攻击性，需要能够抵御放大攻击、畸形报文、水滴攻击，枚举等来保护权威服务器；需要能够做到识别内网 DNS Tunnel，防止攻击者利用 DNS Tunnel 向 C2 服务器发送指令与信息；还需能够抵抗大规模 DDOS，由于 DNS 服务是企业对外提供服务的基础，在遭受 DDOS 时候，不应简单的一刀切的方式采取暴力限流措施，而是能够在 DDOS 防御与提供基础解析服务之间寻求平衡，比如采取更加智能的 DDOS 预测模型来动态抵御 DDOS，部署更高容量的系统来主动承受一定程度上的 DDOS，即便在

DDOS 发生时也能通过降级智能解析服务等措施来保证最基本的解析服务。

高容量，随着 IoT、IPv6、物联网的快速发展，形成了高达百亿终端，万亿解析的解析规模，DNS 系统必须具备足够的解析容量方能应对万物互联下的大规模解析。高容量也是帮助实现高安全的重要方面，在万物互联时代数以亿计的终端在被攻击后可能成为大规模 DNS DDOS 的发起者，这要求 DNS 系统必须具有高容量、高弹性的设计。 此外，高容量还包含 DNS 系统架构自身应可以承载大量解析配置，且具备灵活的拆分与委派能力。

高隐私，DNS 在设计之初采用了明文传输，整个传输过程亦没有认证与保护，导致 DNS 报文易被篡改，缓存系统易被投毒。DNS 系统需要采用相应的技术手段如 DNSSEC，DoHTTPS，DoTLS 等来保证 DNS 内容不被篡改和确保隐私性。

高动态，随着微服务以及 kubernetes 技术架构的普遍应用，DNS 系统正成为整个架构中的关键，大量的服务间通讯需依赖域名解析，部分负载均衡能力也需要依赖 DNS 服务。在微服务技术体系下，一方面服务域名的数量级将会极大提升，另一方域名所映射的端点 IP、服务等信息也呈现高度动态性。因此，DNS 系统需要具备动态的更新能力以随时确保解析最新的端点与服务。

可观测，DNS 是一个网络基础设施，但同时它也是直接面向应用的一个服务，通过洞察 DNS 的解析内容与行为，可以帮助企业了解 DNS 服务运行状态、预警安全威胁，更可以助力企业了解业务运行状态，如多活数据中心运行状况，访客热点区域洞察。在企业强调“数据治理”的理念下，获取并洞察 DNS 运行数据可以帮助企业实现更好的运营。

当前企业对数字化转型进一步深入，现代及边缘应用的部署更加广泛，这对企业 DNS 这一关键数字基础设施的建设提出了更高的要求。为了帮助企业更好的建设 DNS 系统，我们集结了数位 DNS 领域专家编写了此架构白皮书，这些专家都曾是中国大型银行、金融公司 DNS 系统架构的设计者，建设者以及专业服务者。本白皮书将立足通用，可落地以及前瞻性，通过“四大场景八大方案”为您阐述如何构建“高安全，高容量，高隐私，高动态，可观测”的 DNS 系统。

四大场景分别是：

- 外网 DNS
- DoH
- 内网 DNS
- DNS 可观测与运营

八大方案分别是：

- 互联网 DNS 方案
- DMZ DNS 方案
- DoH DNS 方案
- 业务网 DNS 方案
- 办公网 DNS 方案
- 分布式数据库 DNS 方案

- 容器与 PaaS DNS 方案
- DNS 可观测与运营方案

1.外网 DNS 架构

1) 互联网 DNS

需求分析

互联网 DNS 是企业对客服务的最重要基础设施，稳定永续的互联网 DNS 服务是确保业务连续性的首要条件，也是保证企业品牌形象的关键。无论是数地多中心业务多活还是多云策略，即便企业已经对终端类 App 实施了 location service 来减轻对 DNS 服务的依赖，但依然有其他诸多重要服务与应用的业务连续性需要依赖互联网 DNS 服务。互联网 DNS 需要具备对异常服务、异常链路的感知能力，实现快速的业务切换，并基于链路与服务的状态、质量、容量等因素来实现优化的智能解析。因此在架构设计上必须重点加强“安全”、“容量”、“隐私”与“智能”。

互联网 DNS 架构应能够充分保护权威名称服务器，应实现对域名名称过滤、解析类型过滤、畸形报文过滤、反射放大攻击防御等防护工作。不同于其他类型服务的 DDOS 防御，很多情况下 DNS DDOS 攻击发起的报文是完全模拟正常的业务解析，攻击报文的区分识别难度较大，如果采取无差别防御则会产生御敌一千自损八百的结果。因此，DNS 系统对 DDOS 攻击需要具备一定的承受能力，这要求互联网 DNS 架构应在设计容许范围内设计为具有较高解析容量且具备弹性扩容能力的架构。在防护策略上还应具备自主防御行为学习能力，实现动态的抗攻击学习处理，避免人工静态策略导致的防护策略失效。

企业应隐藏提供真实解析服务的权威服务器 IP，在某些局部域名被攻击的情

况下，可灵活操作，针对性的对被攻击域名实施逃逸与降级策略，避免威胁扩散到其他域名。企业还应防止一些Local DNS服务器对权威服务器的不均衡使用，避免解析集中在局部权威服务器上。

互联网 DNS 还应考虑实施 DNSSEC、DoT、DoH 等技术。这些技术可以帮助企业避免 DNS 枚举、缓存投毒、解析污染等问题，保障业务解析的一致性，并避免品牌受损。

互联网链路具有多样性、易故障、不稳定等特点，因此 DNS 系统应能充分感知与处理这些不确定性，无论是线路的故障、应用服务的故障，亦或线路与应用服务叠加的故障等场景，都应始终做出正确且优化的解析。对于一些需要进行报文处理的特定场景，DNS 系统还应具备足够的报文处理能力，并能够实现灵活的 DNS 报文操作。

互联网 DNS 架构自身应采用分布式架构，由于 DNS 的体系结构决定了整体 DNS 系统在节点“可用性”(A) 上已做出了让步 (即可以容许部分 Name Server IP 不可用)，因此互联网 DNS 系统应重点在“一致性 (C)”与“分区容忍性 (P)”上做文章。DNS 系统应保障在任何情形下各个解析单元都能够做出一致性的决策。在某个局部单元无法与其它单元正常通信时，应确保其它可用单元依然能够做出正确且一致的解析决策，不应发生某个局部单元的不可用导致全局性的决策失败或降级。这就要求整个 DNS 解析系统应具备充分的对等的状态同步机制，容忍局部分区的失败并对其隔离，系统不应全局依赖于某个状态控制中心。

对于互联网 DNS 解析的可观测性，应考虑采用易于采集的架构方式，避免数

据系统对接大量解析设备，增加数据归集与缝合的难度。DNS 系统本身应具备对解析日志的分级、格式化与高性能处理，并在必要的时候容许在日志中增加打点信息等。

架构设计

互联网 DNS 架构整体上可设计为“两层四区”结构。

两层分别为：

第一层解析接入控制与分析层

第二层为权威名称解析层

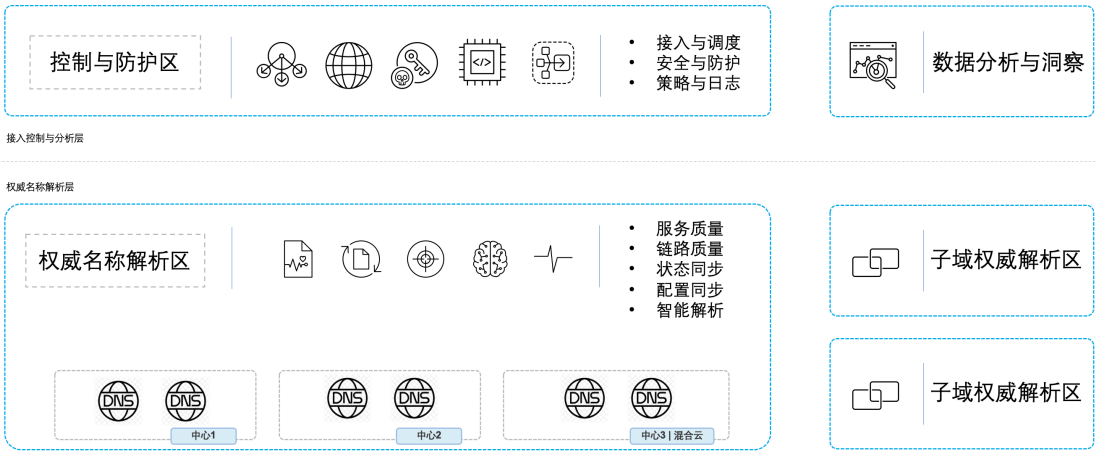
四区则是根据功能分为四个区域：

第一是控制与防护区

第二是数据分析与洞察区

第三是权威名称解析服务区

第四是子域权威名称解析服务区



架构分析

1. 控制与防护区

作为域名解析服务的入口，负责提供互联网 DNS 解析请求的接入。在功能性上该区域主要实现：

接入，提供对外公布的 Name Server IP 的入口，提供标准的 DNS over UDP 53, DNS over TCP 53,以及 DoTLS，DoHTTPs 入口。

调度，该层能够为“名称解析层”提供高级负载均衡服务，包含均匀分配解析请求或根据需要执行基于权重或比率的解析请求分配。该层还需对名称解析层进行高级健康检查以确保及时隔离有问题的名称服务器。可以通过数据报文的分析与控制，实现基于域名、解析类型等的报文调度控制，如将某个子域调度到指定的子域权威层进行解析等。

安全防护，作为直接面向互联网的入口，该层应提供较大的日常抗 DDOS 能力，并在必要时能够将智能解析服务降级为普通解析以提高抗 DDOS 的容忍性，确保业务的基本解析能力。根据安全策略的需要，该层还需提供协议安全、请求

过滤服务，如名称、解析类型、IP 情报、地理位置的过滤，以保证基本的安全基线。基于实际的前置网络结构，反射放大型攻击报文最迟必须在此层终结，以确保权威名称服务器层的安全。

策略，该层可以在网络或应用层面分别提供相应的策略控制，如在网络层面执行限流、对于某些异常请求源可以执行熔断策略、在 DNS 报文里增加 EDNS Client Subnet 信息等策略。

分析，该层具备在不影响解析性能的前提下高速发送解析请求和响应的日志信息到数据分析与洞察区，实现解析的可视化与深度洞察。根据洞察的结果可产生相应的策略规则，通过该层的 API 接口实现自动化调度、策略、安全防护的改变。

2. **数据分析与洞察区**

该区域主要部署相关智能分析系统，该系统接收上文提到的解析日志，根据日志分析并结合实际 DNS 配置元数据来实现可视化的描绘 DNS 解析、系统运行状态的效果。使用智能分析逻辑更好的实现智能化运维。关于该部分的更详细内容请参考本白皮书的第四章内容。

3. **权威名称解析区**

该区域提供域名的最终权威的智能解析，并对业务服务状态执行高级健康检查，根据健康检查结果决定是否将服务作为响应的 RR 记录。根据配置规模和探测逻辑的不同，灵活调配探测作业的性能与效能，保证探测结果的可靠性与一致性。

智能解析，该层除了基本的静态负载均衡解析策略外，还应能够根据业务的

当前服务质量、容量、权重等因素做出符合策略要求的智能解析。在智能决策的规则组合上亦应能够基于多种条件如数据中心、物理位置、逻辑对象等进行组合，且具备层次化的逻辑设定机制，以方便用户可以做出更多的个性化的智能决策规则。

如需要基于运营商链路质量设置智能解析策略，那么还应能够监控运营商链路状态与质量，需要注意的是，对于链路的监控需能够模拟业务的访问，实现从外向内的真实探测，以确保发现链路可能存在的问题，如单栈协议故障、端口被禁等场景，应采取更可靠的探测方式和决策逻辑来避免探测结果的偶然性抖动导致的解析结果波动。

在配置与同步上，权威解析层保存了最终的解析记录配置，应确保解析层所有解析服务实例保持配置的一致性，避免局部或全体实例的配置不一致。所有实例对同一业务的状态表达与决策结论应一致，如果出现不一致情况时候，不一致的实例应能够被隔离，不应对外提供解析服务。这就要求多个实例间应充分解决分布式的一致性，所有实例不应基于某个中心来决策，以避免全局性的失效发生。

4. **子域权威解析区**

在大量的域名存在的时候，企业会希望进行子域拆分以解耦故障域和降低风险，此时可以构建该子域权威解析区域，该区域与权威名称解析区的架构模式完全一致，仅根据实际域名的重要性、解析量、对应服务数量等因素考虑该区域的构建规模。可以根据实际需要构建一套或多套子域权威解析区。在接入控制与防护区实现基于域名的解析拆分，将相关子域解析发送到对应的子域权威区，并对子域权威区实施相同安全防护能力。

技术建议

对于控制与防护区，应重点考虑以下技术建议：

1. 根据规划的线路容量，该区应提供足够的性能以抵御 DDoS。因此该层的设施应在横向与纵向上都能够实现弹性扩展。在横向上，如支持 Anycast 技术，则可以考虑 Anycast。但考虑到大部分用户的实际多线路接入环境难以实施 Anycast，因此应充分考虑在无 Anycast 技术下的水平横向扩展方法，采用例如逻辑拆分调配，资源池与动态组方案。在纵向上应采用能够支持纵向扩容能力的设备。
2. 设备物理接口应支持 100G 接口及接口组捆绑能力。
3. 该区域建议优先考虑透明传递真实解析请求的源 IP 到权威解析区，以方便权威区域基于源 IP 实现基于地理位置的解析。但如果网络结构无法实现透明源地址的透传，则可以考虑 EDNS Client Subnet 技术，但需注意 EDNS Client Subnet 一般会对设备产生更多的性能消耗，因此在容量评估时候应充分考虑该因素。
4. 在 DDoS 防护行为上，建议具备智能学习特性，动态产生防护阀值，以避免人工静态阀值导致的防护缺陷。
5. 提供方便的 API 接口对接分析系统与自动化运维系统，以实现灵活及时的策略调整。
6. 应优先考虑具有 FPGA 能力的设施，借助 FPGA 加速 UDP 处理性能与安全防护效果

7. 与权威解析层自动化联动，实现在大规模攻击下降级智能解析，借助 FPGA 确保最基本的解析服务，以保证业务的可访问性。
8. 具有高性能的日志发送能力，并能自行组织日志格式与内容，可对接多样性的分析系统。
9. 该层在各个物理数据中心分别部署，在各个中心内形成集群组构建自身高可用性。
10. 在该层实施高级负载均衡。
11. 实施报文分析，根据报文内容施加负载均衡策略，或制定相关解析黑白名单能力。

对于权威名称解析区，应重点考虑以下技术建议：

1. 具备多个 40G 以上接口，以保证在单解析实例的通道带宽
2. 基于服务容量、质量、状态以及链路质量、状态的探测，并采用可靠的分布式技术确保探测结果的一致性，应充分考虑对等分布式结构，避免全局性依赖。
3. 灵活的部署额外隐藏探测点，日常可仅用于探测，在必要的时候可以升级为解析节点。
4. 软件实例或硬件实例混合部署，以优化整体方案的性价比。
5. 和控制与防护区自动化联动，实现解析降级、智能调度策略。
6. 在网络上能保证源地址透传下的路径往返一致性。

7. 配置应做到对等的多实例保存，而非集中存在某个控制中心，以避免依赖全局性存储的风险。
8. 容许自行组合服务或链路的健康探测逻辑以实现期望的策略。
9. 智能解析策略能够多层级执行，多维度组合，易于调整和编排逻辑，避免维护大量的规则。
10. 能够自动化发现后端实际业务配置，当后端业务发生实际的上线与下线时候，能够被自动化的发现，避免太多人工操作步骤。
11. 该层在所有数据中心部署，并跨越物理数据中心形成统一的集群共享状态。

2) DMZ DNS

需求分析

DMZ 区作为一个受保护区域，在 DNS 方面主要需要两个方面能力：

1. DMZ 区内的权威解析服务，该权威服务器可以为 DMZ 的具体子域或为 DMZ 的所有需解析的企业域名提供权威解析服务。当然基于现代企业私有云的发展，DMZ 区形态也发生了诸多变化，区内也可能无需权威解析服务。因此在 DMZ 区域内关于权威域名解析服务的需求并不是一定存在。
2. DMZ 区需有 Local DNS 对位于白名单内的外网域名提供转发解析服务。对企业内部域名，将权威解析服务转发到内网 DNS 权威服务上

整体来说，DMZ 区的域名解析服务雷同与内网 DNS 服务，但类如就近性智能解析需求不是主要关注点。鉴于 DMZ 的特殊属性，在 DNS 的安全性上仍然是该区考虑的重要方面。

在容量方面，与互联网及内网 DNS 不同，DMZ 区服务数量或者解析需求量相对来说更加可控，因此一般来说按需部署适度的解析实例。在架构设计中仍应采用可水平扩展的结构设计。

在隐私性方面，DMZ 区服务器需要解析相关外网域名，如果直接将解析请求转发给一般第三方递归服务器，需注意此第三方递归服务器的安全性与可信任度。因此为了提高隐私性，可以考虑在区内部署 DNS over 53 转 DoH 代理，将普通 DNS 解析转化为 DoH 请求发送到外部可靠的 DoH 服务上。在 DMZ 区不宜直接部署区内递归服务器，因为部署递归服务器意味着需要在防火墙上容许任意目的地的 DNS 流量规则，这样的全开规则为 DNS tunnel 类攻击提供了便利性，因此应采用转发器将请求转发或代理到指定可信任的第三方递归服务器上。

DNS 解析请求与日志分析仍然是重要需求，根据实际情况，可将该区日志发送至企业提供的其他数据分析平台，不一定需要在本区内部署独立数据分析与洞察系统。

架构设计

整体上分为“两层五区”设计，部分区域为可选或与企业其他区域能力复用。

两层分别是:

LDNS 接入与防护层

DMZ 权威域名解析与转发器层

五区分别是:

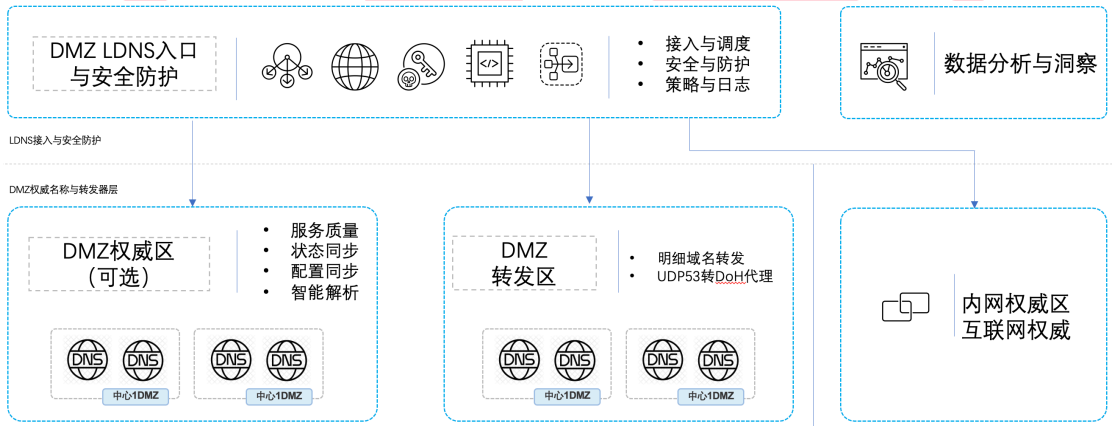
DMZ LDNS 入口与安全防护区

数据分析与洞察区（可复用企业其他区域能力）

DMZ 权威解析区（可选）

DMZ 转发区

内网权威区（企业内网区域，此区域实际不在 DMZ 区）



架构分析

1. **DMZ LDNS 入口与安全防护区**为区内服务器提供集中的 DNS 服务器 IP 入口，该层在各个数据中心的 DMZ 区内部署。提供以下主要能力：

   接入，区内服务器的 DNS 解析请求均发送到该区提供的虚拟 IP 上。

   调度，该区分析判断所解析的域名，如果是 DMZ 区内的特定权威域名，则将解析请求发送到 DMZ 权威区，并执行负载均衡。如果是非 DMZ 区权威域名解析，但仍然是企业权威域名，则发送到内网权威 DNS 或发送到互联网权威 DNS 区进行解析。所有其它非企业权威域名解析，则将其发送到 DMZ 转发区，由转发区负责处理。

   安全与防护，DDoS 防御仍然是一个重要功能考虑，需要防范 DMZ 区内服务器在被攻击后发出大量的 DNS 解析请求，也需要防范不规范的服务器应用程序错误的产生大量的攻击性 DNS 解析。相比互联网区 DNS 的 DDoS 防范能力，DMZ 区抗 DDoS 的规模可以根据实际适当降低，以平衡性价比。对于 DNS 协议特征、域名过滤、IP 情报等基本安全能力与互联网 DNS 区域一致。DMZ 区需额外考虑以下两个能力，一是 DNS Tunnel 攻击，通过部署对 DNS Tunnel 攻击的行为学习能力来阻断借助 DNS Tunnel 向攻击控者的 Command and Control（C2）服务器发送信息数据的行为。二是 DMZ 应用服务可能会连接外部不合法 URL，可以结合有害 URL 库，当所解析域名属于有危害的 URL 的域名范围内时，阻断此类域名解析（注：如果转发区已采用了白名单方式，则可忽略该功能部署）。

   策略，该层可以在网络或应用层面分别提供相应的请求策略，如在网络层面

执行限流、对于某些异常请求源可以执行熔断策略。

分析，该层具备在不影响解析性能的前提下高速发送解析请求和响应的日志信息到数据分析与洞察区，实现解析的可视化与深度洞察。根据洞察的结果可产生相应的策略规则，通过该层的 API 接口实现自动化调度、策略、安全防护的改变。

2. DMZ 权威解析区（可选）

多个数据中心的 DMZ 区形成统一的配置、解析、状态同步能力。配置 DMZ 区专用权威域名。该区在智能解析、健康检查、配置同步、状态同步方面与互联网区 DNS 一致。具体可参考互联网区架构分析。

3. DMZ 转发区

该区可针对具体需要进行转发的明细域名，部署 Forwarder only 型转发器，将请求转发到信任的外部递归服务器。在该区也可以为区内服务器提供 DNS over 53 的普通 DNS 解析请求到 DoH 解析的转发代理，这样可以让区内服务器获得 DoH 的隐私特性，为应用提供更好更安全的外部解析，防止不可靠外部递归服务器的风险。

技术建议

对于 DMZ LDNS 与安全防护区，应重点考虑以下技术建议：

1. 应提供足够的性能以抵御 DDoS。因此该层的设施应在横向与纵向上都能够实现弹性扩展。在横向上，应考虑支持 Anycast 技术以实现在多物理中心实现冗余切换，防止整区不可用故障。在纵向上应采用能够支持

纵向扩容能力的设备。

2. 设备物理接口应支持 100G 接口及接口组捆绑能力。
3. 在 DDoS 防护行为上，建议具备智能学习特性，动态产生防护阀值，以避免人工静态阀值导致的防护缺陷。具备根据报文的特征产生特征码识别。
4. 提供方便的 API 接口对接分析系统与自动化运维系统，以实现灵活及时的策略调整。
5. 应优先考虑具有 FPGA 能力的设施，借助 FPGA 加速 UDP 处理性能与安全防护效果
6. 与权威解析层自动化联动，实现在大规模攻击下降级智能解析，借助 FPGA 确保最基本的解析服务，以保证业务的可访问性。
7. 具有高性能的日志发送能力，并能自行组织日志格式与内容，可对接多样性的分析系统。
8. 具备对 DNS Tunnel 的攻击防御能力。实现基于报文内容的过滤与流量调度。
9. 具有高级负载均衡能力

对于权威名称解析区，应重点考虑以下技术建议：

1. 具备多个 40G 以上接口，以保证在单解析实例的通道带宽
2. 基于服务容量、质量、状态探测，并采用可靠的分布式技术确保探测结

果的一致性，应充分考虑对等分布式结构，避免全局性依赖。

3. 灵活的部署额外隐藏探测点，日常可仅用于探测，在必要的时候可以升级为解析节点。
4. 软件实例或硬件实例混合部署，以优化整体方案的性价比。
5. 和 DMZ LDNS 与安全防护区联动，实现解析降级、智能调度策略。
6. 配置应做到对等的多实例保存，而非集中存在某个控制中心，以避免依赖全局性存储的风险。
7. 具有 DNS over 53 到 DoH 的代理能力。
8. 智能解析策略能够多层级执行，多维度组合，易于调整和编排逻辑，避免维护大量的规则。
9. 能够自动化发现后端实际业务配置，当后端业务发生实际的上线与下线时候，能够被自动化的发现，简化大量人工操作步骤。
10. 该层在所有数据中心部署，并跨越物理数据中心形成统一的集群共享状态。

2.DNS over HTTPS (DOH)架构

需求分析

由于 DNS 技术历史悠久，且在设计的时候并未考虑到安全设计，导致在互联网中存在大量 DNS 问题，主要包括安全问题、隐私问题以及用户访问体验问题。

安全问题

由于 DNS 请求和响应的数据流都是通过明文方式传输，导致在数据流的各个节点中存在潜在的安全隐患，存在广泛的攻击行为。比如 DNS 劫持、DNS 缓存中毒、DNS 数据包伪造攻击等。由于 DNS 请求可以随意发起，并且单个 DNS 请求的性能开销极低，导致在互联网上存在大量的 DNS DDoS 攻击行为。

隐私问题

由于 DNS 查询是未加密的，任何人都可以捕获和查看 DNS 数据包。这意味着不仅 ISP 和 DNS 服务器提供商可以访问这些信息，黑客和其他恶意人员也可以通过监听 DNS 流量来获取用户的隐私信息。例如，一个网络攻击者可以通过监听 DNS 查询来确定用户正在访问哪些网站，并使用这些信息进行钓鱼攻击、广告追踪和其他形式的恶意行为。

此外， 一些 ISP 和 DNS 服务器提供商可能会记录用户的 DNS 查询历史记录，并将其用于广告定位或其他目的。这也会影响用户的隐私和安全。

用户体验问题

由于权威 DNS 域名解析系统以 LDNS 地址为依据判断用户所在地和所属运营商，并以此返回该域名当前最优的解析结果。在实际的互联网场景下，LDNS 并不能完全代表真实的客户端。比如某些用户手工配置了 DNS SaaS 服务，比如 google 的 8.8.8.8，权威 DNS 无法判断这个 IP 的归属属性。还有些场景是运营商内部 LDNS 之间互相转发请求，导致权威 DNS 误判归属属性。一旦归属属性无法判断或者判断错误，则会导致用户无法获取到最优的解析结果，影响用户访问体验。同时由于 LDNS 上存在缓存，解析结果会缓存在 LDNS 上。非最优的解析结果，会影响该 LDNS 所覆盖的所有用户。

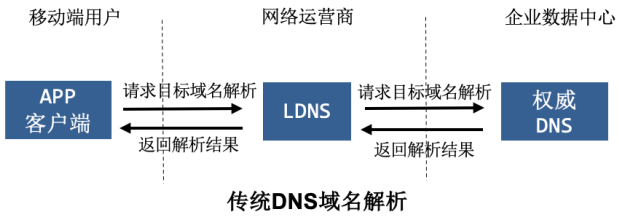
另一方面，在移动端场景下，由于存在网络切换（比如 5G 切换到 WIFI）的情况，叠加智能解析无法获取真实客户端的地址，易导致用户在网络切换前后，对于同一个域名解析到不同的互联网入口，导致应用登出等严重影响用户访问体验的行为发生。除此之外，传统 DNS 域名解析还存在 DNS 解析故障、域名劫持、LDNS 策略错误、LDNS 缓存错误等问题，导致用户访问异常或访问缓慢等问题。

对于传统 DNS 的上述问题， 业界也发展出了两种用于加密 DNS 查询的协议，分别是 DNS-over-TLS（DoT）和 DNS-over-HTTPS（DoH）。它们的目标都是提高 DNS 查询的隐私和安全性。这两种协议在加密方式、协议端口、实现成本、部署方式等方面存在一些差异。DoT 主张维持 DNS 架构和访问方式不变，DNS 各环节通过 TLS 进行加密，主要是运营商在主导。而 DoH 由互联网厂商主导，主张重构传统 DNS 架构，通过通用的 HTTPS 协议来承载传统的 DNS 请求，绕开运营商 LDNS 体系，实现 DNS 流量的端到端可控以及更加灵活的 DNS 调度逻辑。

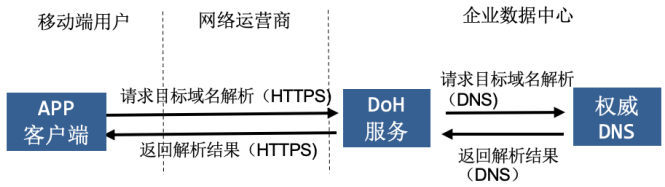
DoH 的前身是 HTTPDNS 技术，由腾讯于 2012 年前后提出通过 HTTPDNS 技术从根源上解决域名解析异常及用户访问跨网缓慢的问题。业界按照自己的理解实现了 HTTPDNS，实现方式参差不齐。随着 HTTPDNS 技术的发展和理念普及，Google 及 Mozilla 基金会也参与推动了该技术的发展进步。在 2018 年该技术正式成为 IETF RFC 规范，RFC 8484，名为 DNS over HTTPS，简称 DoH。

随着 DoH 成为 RFC 规范，经过这几年的发展，DoH 生态已经逐渐完善起来。客户端层面，主流操作系统、浏览器已逐步支持 DoH。如 Google 2019 年 6 月宣布，DNS over HTTPS（DoH）加密 DNS 服务已迈入正式版，使用者现可直接在 dns.google 网域上，以 DoH 来解析网域名称系统（DNS）。Firefox 62 及以上版本中可开启 DNS over HTTPS 配置。 Chrome 是继 Firefox 之后添加 DoH 支持的第二款浏览器。主流浏览器都已逐步支持 DoH。微软在 Windows 10 版本中已增加对 DoH 协议的支持。服务端层面，如 Google DNS 、CloudFlare DNS、阿里云、腾讯云等 DNS 云服务商都已支持 DoH 解析服务。随着时间点的推移，DoH 生态会越来越完善。

架构设计



传统DNS域名解析



DoH域名解析

传统 DNS 域名解析和 DoH 域名解析的数据流如上图所示。客户端通过 HTTPS 协议访问 DoH 服务节点进行域名解析，代替传统方式通过 DNS 协议访问网络运营商 LDNS 服务节点进行域名解析。通过 HTTPS 进行 DNS 解析以及绕开 LDNS，能够获得以下的提升：

1. **解决传统 DNS 的可用性及安全问题。**客户端进行 DoH 技术改造之后，业务域名解析请求可以绕过 LDNS，直接通过 DoH 服务器完成，有效的降低了对传统 DNS 的依赖，从而可规避运营商 DNS 故障、域名劫持、LDNS 策略配置错误、LDNS 域名缓存错误等导致访问异常的问题。同时 DoH 的方式使得第三方无法监测到用户的 DNS 请求，最大程度上规避了用户上网行为等隐私泄露的风险。

2. **提升用户访问体验。**DoH 可以直接获取到用户的真实 IP，可以最大限度地使用智能域名解析的精准调度功能，将用户引导到最优的访问路径上，提升用户访问体验。

3. **实现端到端的控制。**绕开 LDNS 后，DoH 服务直接面对客户端。无论是 TTL 时间的控制，还是流量的精准调度，都能够实现端到端的精细化控制。

在基础架构方面，DoH 相比 DNS 有三个显著的差异，在 DoH 架构设计上需要额外考量。

1. **单个请求处理性能开销增大。**相比传统 DNS，DoH 请求处理性能开销增大，包括 SSL 加解密，HTTP 与 DNS 的协议转化等 CPU 高密型操作。
2. **解析流量成倍增加。**传统 DNS 场景下，DNS 服务器面对的是 LDNS。LDNS 数量有限，且配置有缓存，因此真正由企业 DNS 权威处理的请求量比较小。DoH 场景下，DoH 服务器直接面对海量的客户端，流量预计会有十倍甚至百倍的增长。在这种情况下，如果要满足业务需求，就必须要有**缓存机制。**结合智能解析场景，需要实现基于地址库的缓存机制。
3. **流量波动增加。**传统 DNS 场景下，除非遭受 DNS 攻击，由于 Local DNS 缓存的存在会平抑流量，DNS 流量相对比较平稳。而 DoH 场景下，由于直接面对客户端，每个客户端的行为不一，流量波动会明显增加。为了避免流量暴增影响系统正常访问，需要有**访问控制机制。**

因此，**在进行 DoH 服务架构设计时，重点要考虑弹性可扩展的架构，以灵活应对 DoH 流量波动。**

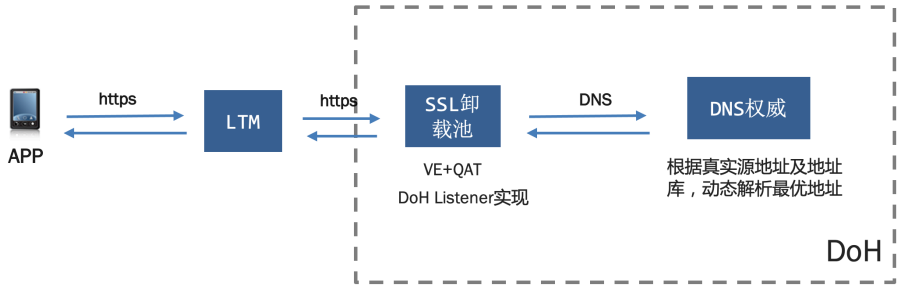
结合互联网应用架构以及业务特点，规划 DoH 服务需具备以下能力：

1. **SSL 卸载。**实现对 DoH 请求的 SSL 卸载，可构建弹性的 SSL 卸载池实现。
2. **源地址透传。**实现将客户端源地址透传到末端的权威 DNS 服务器，实现基于客户端 IP 的智能解析。
3. **DoH 缓存。**实现基于地址库的缓存，即针对不同地域的客户端访问，缓存权威 DNS 不同的解析结果。缓存的时间以权威 DNS 上的 TTL 为准。同时绝大多数 DoH 请求都在这个环节终结，降低对权威 DNS 的压力。
4. **协议转换。**从客户端接收到的 HTTP 请求中提取 DNS 请求，转换成 DNS 协议发送给 DNS 服务器，并将 DNS 响应封装成 HTTP 请求返回客户端。
5. **访问控制。**实现限流、客户端 IP 黑名单、域名白名单等访问控制能力，提升 DoH 服务的运行稳定性，减少安全风险。
6. **流量可视化。**实现 DoH 请求的流量可视化，提升运维智能化能力。
7. **功能灵活扩展。**由于 DoH 技术是个相对较新的技术，随着 DoH 技术和企业业务的发展，具备通过灵活扩展的能力来满足新增的业务需求。在技术选型上需要具备灵活扩展的能力。

架构分析

根据不同的技术路线选择，典型的 DoH 实现架构主要包括以下两种：

1. **原生的 DoH VS 实现。**如下图所示。



由于 DoH 在 2018 年正式成为 RFC 规范，因此各网络厂商逐步开始实现 DoH 能力。以 F5 为例，V16 版本以上实现了原生 DoH 能力，通过创建 DoH Listener 来实现处理 HTTPS 请求，进行协议转换后转发 DNS 数据包到 DNS 权威上。

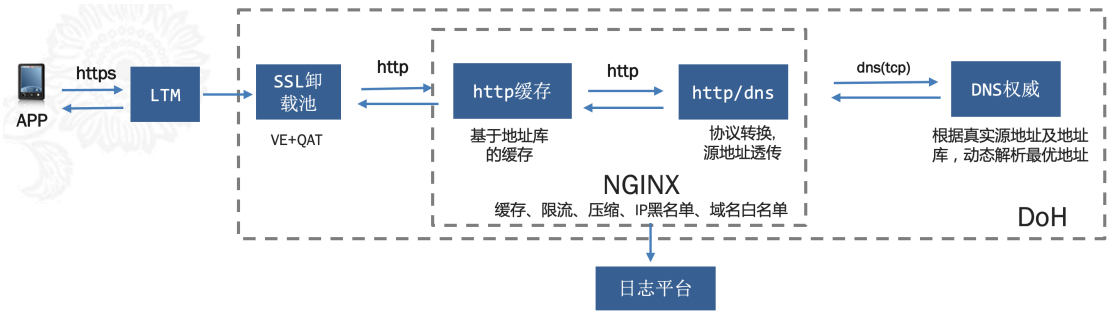
这个 DoH 架构的完整实现主要包含以下两部分：

SSL 卸载+协议转换。功能通过 BIGIP 原生 DoH Listener 实现，以软件资源池的形式部署，实现高弹性。同时部署 INTEL QAT 卡实现高性能 SSL 卸载。

智能 DNS 解析。以纯软的方式部署高性能 F5 DNS 资源池，根据真实源地址及本身地址库，动态解析最优的地址。软件资源池的方式，具备高弹性。但集群的规模受限于 F5 DNS 同步组的规模上限。

该架构的**优势**在于：设备原生 DoH 实现，运维最简单；整体架构高度弹性扩展，适合中小企业客户场景。

2. **NGINX 实现 DoH 服务。**如下图所示：



这个 DoH 架构的完整实现，包含以下三部分：

（1） **SSL 卸载。**实现对 DoH 请求的 SSL 卸载，可构建弹性的 SSL 卸载池实现。

（2） **HTTP 缓存 + 协议转换。**通过部署 NGINX 实现基于地址库的缓存以及协议转换。前者通过 NGINX 原生指令配置实现，后者通过 NJS 脚本开发实现。另外限流、压缩、IP 黑名单、域名白名单等能力通过 NGINX 原生指令实现。

（3） **智能 DNS 解析。**以纯软的方式部署高性能 F5 DNS 资源池，根据真实源地址及本身地址库，动态解析最优的地址。由于 NGINX 上部署了 HTTP 缓存，只有少部分请求会由这个节点进行处理，降低了对该节点的压力。

该架构的**优势**在于：具备完整 DoH 服务所需能力，且可定制化开发扩展；成本最优，无需部署大量的权威 DNS；技术可控性最强。适用于海量用户的大型企业场景。

技术建议

对于 DNS over HTTPS 平台建设，应重点考虑以下技术建议：

1. 应考虑支持 Anycast 技术实现在多物理中心实现 DoH 服务对外发布，且固定 IP 地址，APP 端可以通过固定的 DoH 服务 IP 发起 DoH 请求，减少一次 DoH 服务本身的 DNS 解析，提升用户体验。
2. DoH 作为传统 DNS 服务的衍生，建议在建设 DoH 系统时，综合考虑目前智能解析设备的能力。尽可能复用存量功能节点，降低成本投入。
3. DoH 各功能节点在设计上需考虑解耦，同时基于流量突发的特性，建议尽可能以软件和资源池的方式进行部署，实现快速弹性扩展的能力。
4. DoH 对外暴露的是 HTTPS 443 接口，客户端发起的是 HTTPS 请求。同时由于 DoH 技术相对比较新，可能存在一些潜在安全隐患。建议可专门针对 DoH 请求部署安全防护策略。
5. DoH 系统涉及到 HTTPS 和 DNS 两种协议，相比传统 DNS 系统，日常运维的难度增大。因此在建设 DoH 系统时，需要考虑构建 DoH 流量可视化平台，降低运维复杂度并且实现 DoH 服务运营，提升安全性的同时提升用户访问体验。

3.内网 DNS 架构

1) 业务网 DNS

需求分析

业务网 DNS 主要承担的是业务网生产和测试系统的域名解析功能。其需要满足以下需求：

1. 智能 DNS 解析功能：需要对业务系统的服务状态、服务质量具备探测能力，并根据需要制定适合智能解析策略，实现业务网系统流量的智能调度，或在状态异常时实现自动快速容灾切换。
2. 高可用性：业务网 DNS 架构自身应具备高可用性，保证 24 小时不间断运行，确保业务网系统的能够随时被访问。
3. 高性能和可扩展性：业务网 DNS 架构需要具备高性能的 DNS 解析能力和高配置容量，同时具备良好的可扩展性，以便应对大量 DNS 解析请求，大量域名服务节点配置，大量的健康检查的探测能力，以及各指标高速增长的需求场景，而且能够承受一定量的业务网 DNS DoS/DDoS 的能力
4. 安全性：业务网 DNS 系统需具备 DNS 安全防护能力，实现对域名名称过滤、解析类型过滤、畸形报文过滤、反射放大攻击防御，DNS tunnel 攻击等防护工作。以便应对来自 DMZ 区、办公网或业务网内的 DNS 攻击。

架构设计

业务网 DNS 架构整体上可设计为“三层六区”结构。

三层分别为：

第一层是业务网 DNS 接入层

第二层是业务网 DNS 调度/安全防护与分析层

第三层是业务网 root 根域及权威名称解析层

六区则是根据功能分为六个区域：

第一是业务网 LDNS 区

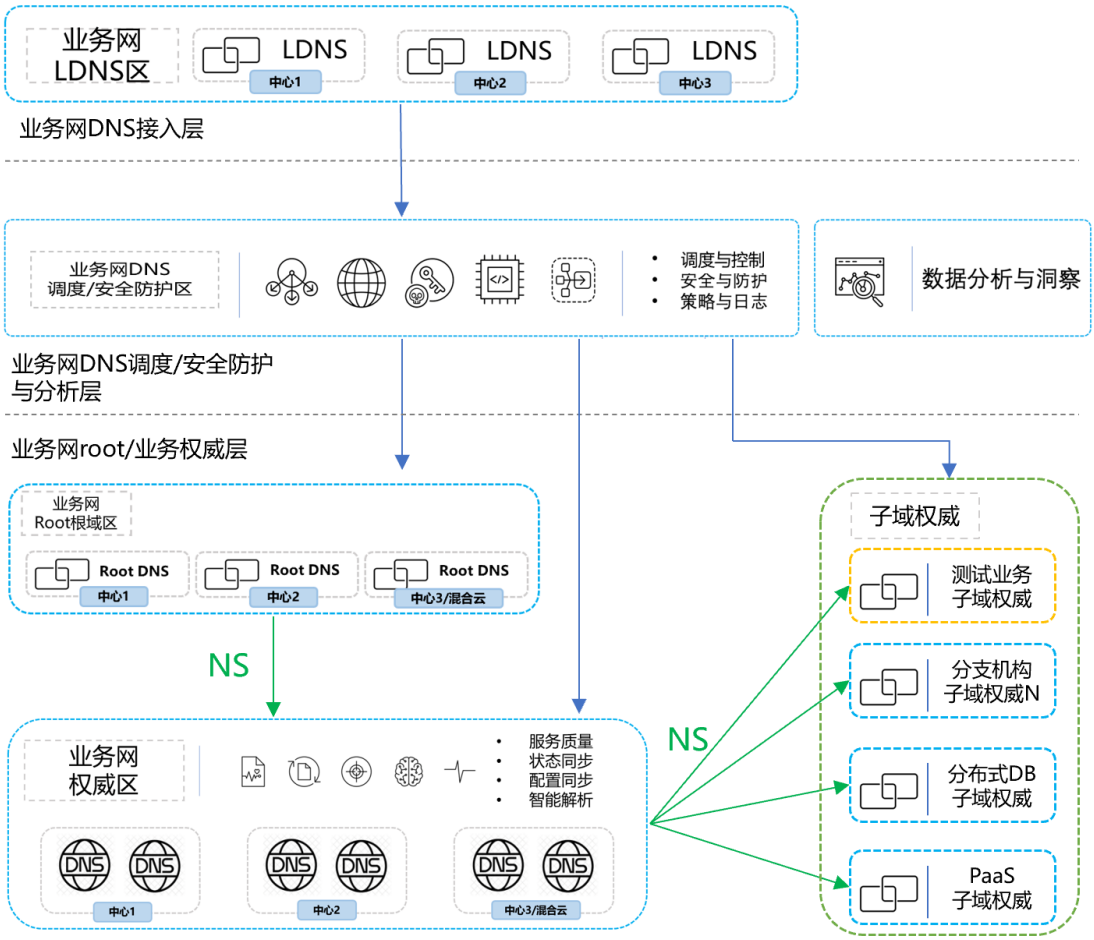
第二是业务网 DNS 调度与安全防护区

第三是业务网 DNS 数据分析与洞察区

第四是业务网 root 根域区

第五是业务网权威名称解析区

第六是业务网子域权威名称解析区



架构分析

业务网 LDNS 区主要功能：

1. 承担业务网系统间互访的域名解析请求的接入
2. 承担各分支机构对业务网各系统的域名解析请求的接入
3. 承担 DMZ 区业务服务对业务网系统访问的域名解析请求的接入
4. 承担办公网对业务网系统访问的域名解析请求的接入

备注：对于一些中小型机构或 LDNS 由总部统一管理的企业或组织机构，业务网 LDNS 层可以与业务网 DNS 调度/安全防护区合并。

业务网 DNS 调度/安全防护区主要功能：

1. 接受业务网 LDNS 转发的业务网系统域名解析请求，并按照预设的策略调度转发到对应的业务网 DNS 系统区域。
2. 分别对各中心业务网 Root 根域区的 DNS 服务器、权威 DNS 服务器、子域权威 DNS 服务器做负载均衡和 DNS 服务的健康监测，为各区域 DNS 系统提供横向的扩展能力，同时提升各区域的高可用性。
3. 执行 DNS 安全防护功能：实现对域名名称过滤、解析类型过滤、畸形报文过滤、反射放大攻击防御等防护工作。
4. 一方面过滤掉 DMZ 或办公网转发过来的或系统间的非业务网系统的 DNS 域名或未配置的 type 类型解析请求；
5. 另一方面对于畸形报文、反射放大、枚举等 DoS/DDoS 攻击进行防御
6. 记录 DNS 解析的日志信息，并转发给数据分析与洞察区，进行大数据分析、审计和溯源等。

业务网数据分析与洞察区功能：

部署相关智能分析系统，该系统接收上文提到的域名解析日志，根据日志分析并结合实际 DNS 配置元数据来实现可视化的描绘 DNS 解析、系统运行状态的效果，更好的实现智能化运维。关于该部分的更详细内容请参考本白皮书的第四章内容。

业务网 Root 根域区功能：

为业务网 DNS 系统提供 Root 根域服务。

备注：业务网 Root 根域可以和权威名称解析区合并

业务网权威名称解析区的主要功能：

1. 为业务网业务系统提供域名的最终权威的智能解析。
2. 该区对业务服务状态执行高级健康检查，根据健康检查结果决定是否将服务作为响应的 RR 记录。根据配置规模和探测逻辑的不同，灵活调配探测作业的性能与效能，保证探测结果的可靠性与一致性。
3. 对于智能解析，除了基本的静态负载均衡解析策略外，权威解析层还需能够根据业务的当前服务质量、容量、权重等因素做出符合策略要求的智能解析。
4. 在配置与同步上，权威解析层保存了最终的解析记录配置，应确保解析层所有解析服务实例保持配置的一致性，避免局部或全体实例的配置不一致。所有实例对同一业务的状态表达与决策结论应一致，如果出现不一致情况时候，不一致的实例应是被自动隔离，不应对外提供解析服务。这就要求多个实例间应充分解决分布式的一致性，所有实例不应基于某个中心来决策，以避免全局性的失效发生。

子域权威名称解析区功能：

当企业或组织机构的存在大量域名的时候，可以对权威 DNS 进行子域拆分构建子域权威 DNS 区域，以实现解耦故障域和降低风险的目的。该区域与权威名称解析区的架构模式完全一致，仅根据实际域名的重要性、解析量、对应服务数量、变更的频次等因素考虑该区域的构建规模。可以根据实际需要构建一套或多套子域权威解析区。接入控制与防护区实现基于域名特征的解析拆分，将相关子域解析发送到对应的子域权威区，并对子域权威区实施相同安全防护能力。

技术建议

对于业务网 DNS 接入层即 LDNS 区，技术建议如下：

1. 对于分支较多的企业或组织机构，建议就近部署 LDNS，一方面利用本地的 DNS Cache 能力可以使客户端获得更快的 DNS 响应，同时避免单 LDNS 区域故障对其他分支的影响；另一方面便于针对不同分支特殊需求制定不同的 LDNS 策略，比如分支机构有自己的子域，则 LDNS 需要将权威域，本地子域或其他子域等不同的 DNS 请求进行区分，并转发到对应的 DNS 解析区域。
2. 为了满足审计溯源，权威区域基于源 IP 进行智能解析实现就近访问的目的等需求，LDNS 需要具备透传真实源 IP 的能力。对于 LDNS 区组网优先考虑透明传递真实解析请求的源 IP 到权威解析区。但如果网络结构无法实现透明源地址的透传，则可以考虑 EDNS Client Subnet 技术（即 ECS），但需注意 ECS 一般会对设备产生更多的性能消耗，因此在容量评估时候应充分考虑该因素。

3. 对于分支机构不多，或需要总部统一纳管的企业或组织机构，可以和业务网 DNS 调度/安全防护层合并，简化架构，减少投资。

对于业务网 DNS 调度 / 安全防护区，技术建议如下：

1. 根据企业或组织机构规模，以及业务的访问量，该区应在提供足够的性能保证正常的 DNS 解析请求的转发的同时，可以抵御内网 DNS DDoS 攻击。因此该层的设备需采用高性能的设备，同时具备横向集群扩展能力。
2. 设备物理接口应支持 100G 接口及接口组捆绑能力。
3. 该区域应确保能够透明传递真实解析请求的源 IP 到权威解析区。
4. 在 DDoS 防护行为上，建议具备智能学习特性，动态产生防护阀值，以避免人工静态阀值导致的防护缺陷。
5. 提供方便的 API 接口对接分析系统与自动化运维系统，以实现灵活及时的策略调整。
6. 应优先考虑具有 FPGA 能力的设施，接触 FPGA 加速 UDP 处理性能与安全防护效果
7. 与权威解析层自动化联动，实现在巨大规模攻击下降级智能解析，借助 FPGA 确保最基本的解析服务，以保证业务的可访问性。
8. 具有高性能的日志发送能力，并能自行组织日志格式与内容，可对接多样性的分析系统。

9. 该层在各个物理数据中心分别部署，在各个中心内形成集群组构建自身高可用性。
10. 具有高级负载均衡能力。

对于业务网权威名称解析区，技术建议如下：

1. 具备多个 40G 以上接口，以保证在单解析实例的通道带宽
2. 基于服务容量、质量、状态的探测，并采用可靠的分布式技术确保探测结果的一致性，应充分考虑对等分布式结构，避免全局性依赖。
3. 灵活的部署额外隐藏探测点，这些探测节点平常仅用于探测，在必要的时候可以升级为解析节点。
4. 软件实例或硬件实例混合部署，以优化整体方案的性价比。
5. 和业务网 DNS 调度/安全防护区自动化联动，实现解析降级、智能调度策略。
6. 在网络上能保证源地址透传下的路径往返一致性。
7. 执行分布式的配置存储，以避免依赖全局性存储的风险。
8. 容许自行组合服务或链路的健康探测逻辑以实现期望的策略。
9. 智能解析策略能够多层级执行，多维度组合，易于调整和编排逻辑，避免维护大量的规则。
10. 能够自动化发现后端实际业务配置，当后端业务发生实际的上线与下线时候，能够被自动化的发现，避免太多人工操作步骤。

11.该层在所有数据中心部署，并跨越物理数据中心形成统一的集群共享状态。

对于业务网子域权威名称解析区，技术建议如下：

1. 为避免测试业务影响生产业务，建议生产和测试业务 DNS 解析系统分离，拆分出测试业务子域权威名称解析区
2. 对于分支机构比较多且各分支希望独立运维自己的 DNS 系统的企业或组织机构，建议为各分支机构拆分独立的子域权威名称解析区，并在各分支机构当地进行部署。
3. 对于规模化部署分布式数据库的场景，对于 DNS 系统的性能、容量和扩展能力要求非常高，则建议拆分独立子域权威名称解析区，具体情况请参见本章第 3）节-分布式数据库 DNS。
4. 对于 PaaS 平台场景，有些 PaaS 架构对于 DNS 的容量要求比较高，且可能变更非常频繁，则建拆分独立子域权威名称解析区，具体情况请参见本章第 4）节-容器与 PaaS DNS。

2) 办公网 DNS

需求分析

办公网 DNS 不同于业务网 DNS，主要承担的是企业或组织机构内部员工的日常办公的域名解析需求，比如：

1. 日常的办公系统的访问：OA/ERP/CRM/财务系统/文件共享系统/内部论坛/内部学习系统或考试系统/邮件系统；
2. 访问外部 Internet 资源：访问外网网站，或者需要通过第三方认证系统进行身份认证才能访问办公系统资源等；
3. 对于一些组织机构，还有对业务网资源访问的需求；
4. 办公 PC 需要进行 AD 域控管理等

对于以上日常办公资源的访问，基本上都是通过域名的方式进行的。鉴于此办公网 DNS 的设计需求也不尽相同，除了要求办公网 DNS 具备高可用性，可扩展性和智能解析功能外，需要着重考虑以下几个方面：

1. DNS 安全性：

（1） 由于办公网环境相对复杂，办公 PC 或办公系统更容易遭受计算机病毒感染或黑客入侵，可能造成 DNS DoS/DDoS 攻击,甚至可能以办公 PC 为跳板攻入 DNS 系统。因此需要办公网 DNS 系统具备 DNS 安全防护能力：实现对域名名称过滤、解析类型过滤、畸形报文过滤、反射放大攻击防御等防护工作。

（2） 为避免波及业务网 DNS 解析和业务系统，办公网 DNS 系统需要与业务网 DNS 系统分离，以达到相对隔离，甚至完全隔离的目的（具体要求视安全或监管要求等情况而定）。

2. DNS 请求转发能力：

（1） 由于办公人员需要访问的资源的分布的不同，办公网 DNS 系统在具备办公系统域名解析能力同时，还需要将访问外部 Internet 资源的域名

解析请求转发到外网的可信 LDNS 上，或将访问业务网生产或测试系统资源的域名解析请求转发到业务网 DNS 系统以便获取对应的 IP 地址。

（2） 如需要对办公网 PC 或 server 进行 AD 域控管理，则需要办公网 DNS 具备将 AD 相关 DNS 解析、注册、反向解析的数据包转发到 AD 服务器上的能力，同时也能够根据 AD 的返回结果再转发给客户端。

3. 域名管控和审计能力：

企业或组织机构可能需要对员工访问的外网 Internet 资源或业务网系统资源进行限制或审计，这要求办公网 DNS 系统具备域名管控能力，通过设置黑白名单或与第三方恶意域名检测或评级系统对接对域名的解析进行管控；同时还可以记录 DNS 解析的日志，用于审计或攻击溯源。

架构设计

办公网 DNS 架构整体上可设计为“三层七区”结构。

三层分别为：

第一层是办公网 DNS 接入层

第二层是办公网 DNS 调度、安全防护与分析层

第三层是办公网权威名称解析及转发层

七区则是根据功能分为七个区域：

第一是办公网 LDNS 区

第二是办公网 DNS 调度与安全防护区

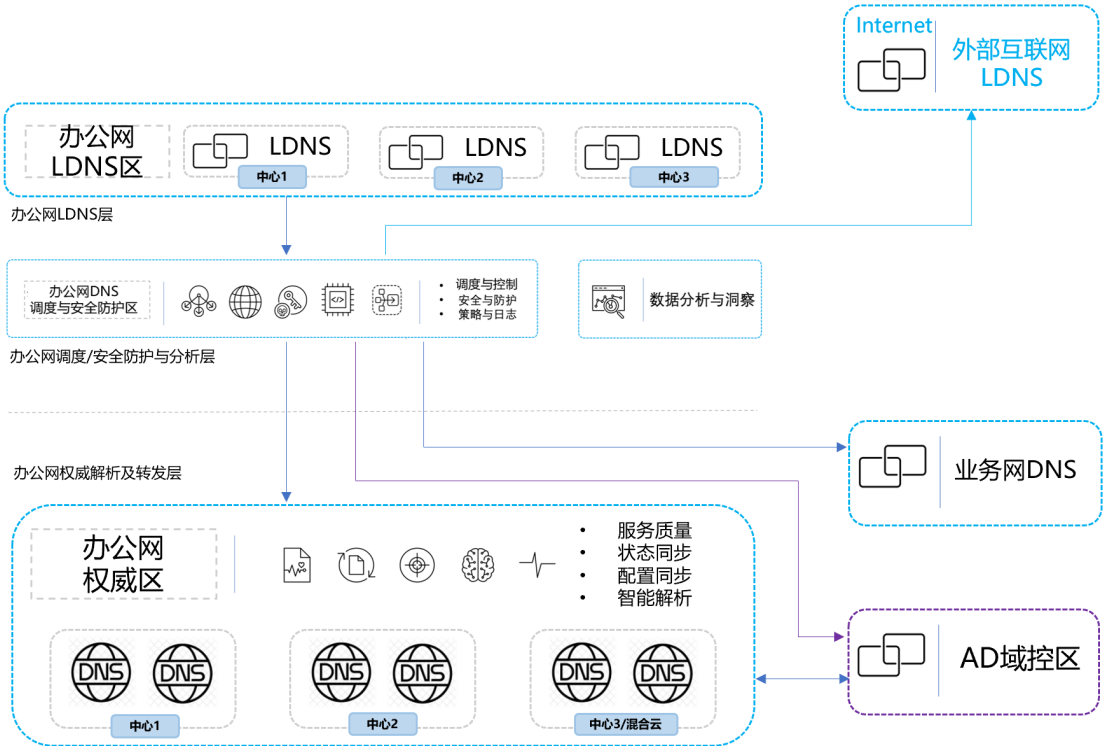
第三是办公网 DNS 数据分析与洞察区

第四是办公网权威名称解析服务区

第五是 AD 域控区

第六是业务网 DNS 区

第七是外部互联网 LDNS 区



架构分析

1. 办公网 LDNS 层的主要功能是将来自办公网的 PC 或 Server 客户端的域名解析请求或 AD 域控请求转发给办公网 DNS 调度与安全防护区，并将获取的解析结果返回给对应的客户端。

   注：如果办公网规模不大，则从性价比考虑该层可与办公网 DNS 调度与安全防护区合并

2. 办公网 DNS 调度和安全防护区，主要功能包括：

   （1） 对办公网 DNS 权威解析系统做负载均衡和服务健康检测监测，并分别为每个中心提供统一的办公网域名权威解析 IP。

   （2） 将办公系统的域名解析请求转发给办公网权威解析区的 DNS 服务上，并将解析结果返回给对应的客户端；

   （3） 将针对 Internet 的域名解析请求转发给可信的外网 LDNS 系统，以便通过其获得 Internet 资源的域名解析结果，并将解析结果返回给客户端。同时还可监测外网 LDNS 的健康状态，如果发现异常，则自动切换到其他可信外网 LDNS 上。

   （4） 将针对业务网的域名解析请求转发给业务网 DNS 系统，并可设置可解析的域或域名的黑白名单，对业务的域名解析进行管控。

   （5） 将针对 AD 域控的请求转发给 AD 域控服务器，帮助完成 AD 相关的 DNS 解析、AD 注册和反向解析等；同时监测 AD 域控的健康状态，提升 AD 域控的高可用。

（6） 通过调整内置的 DNS 防护策略，设置黑白名单，或与第三方恶意域名检测或评级系统对接等实现域名的过滤、解析类型过滤、畸形报文过滤、反射放大攻击防御，域名枚举攻击防御等防护工作。

（7） 记录 DNS 解析的日志信息和 AD 域控相关信息，并转发给数据分析与洞察区，进行大数据分析、审计和溯源等。

技术建议

办公网 DNS 系统与业务网 DNS 系统分离，并对办公网访问业务网系统的域名请求进行安全管控，从而极大程度上保证了各自 DNS 系统的安全性和可用性不受对方的影响。

对于大型或超大型企业或组织机构办公网 DNS 接入层 LDNS 和 DNS 调度/安全防护层的技术建议：

1. 这两层需要具备足够的性能并开启 DNS Cache 功能，在保证正常 DNS 解析请求的同时可以对 DNS DoS/DDoS 攻击具备一定的承受能力。因此该层的设备需采用高性能的设备，同时具备横向集群扩展能力。
2. 设备物理接口应支持 100G 接口及接口组捆绑能力。
3. 在 DDOS 防护行为上，建议具备智能学习特性，动态产生防护阀值，以避免人工静态阀值导致的防护缺陷。具备根据报文的特征产生特征码识别。

4. 提供方便的 API 接口对接分析系统与自动化运维系统，以实现灵活及时的策略调整。
5. 应优先考虑具有 FPGA 能力的设施，接触 FPGA 加速 UDP 处理性能与安全防护效果
6. 与权威解析层自动化联动，实现在大规模攻击下降级智能解析，借助 FPGA 确保最基本的解析服务，以保证业务的可访问性。
7. 具有高性能的日志发送能力，并能自行组织日志格式与内容，可对接多样性的分析系统。
8. 具有可编程能力，实现对 DNS Tunnel 的攻击防御。实现基于报文内容的过滤与流量调度。
9. 具有高级负载均衡能力

对于办公网权威名称解析区，应重点考虑以下技术建议：

1. 具备多个 40G 以上接口，以保证在单解析实例的通道带宽
2. 基于服务容量、质量、状态的探测，采用可靠的分布式技术确保探测结果的一致性，应充分考虑对等分布式结构，避免全局性依赖。
3. 灵活的部署额外隐藏探测点，这些探测节点平常仅用于探测，在必要的时候可以升级为解析节点。
4. 软件实例或硬件实例混合部署，以优化整体方案的性价比。
5. 和调度与安全防护区自动化联动，实现解析降级、智能调度策略。

6. 在网络上能保证源地址透传下的路径往返一致性。
7. 执行分布式的配置存储，以避免依赖全局性存储的风险。
8. 容许自行组合服务健康探测逻辑以实现期望的策略。
9. 智能解析策略能够多层级执行，多维度组合，易于调整和编排逻辑，避免维护大量的规则。
10. 能够自动化发现后端实际业务配置，当后端业务发生实际的上线与下线时候，能够被自动化的发现，避免太多人工操作步骤。
11. 该层在所有数据中心部署，并跨越物理数据中心形成统一的。

3) 分布式数据库 DNS

需求分析

随着微服务或单元化部署的推广，许多用户也开始采用分布式数据库将原来的单体数据库集群服务架构拆分为大量分布式子服务集群，对应不同的微服务或服务单元，如果这些子服务各自拥有自己的域名，此模式下则可能会对 DNS 服务带来的极大的挑战：

1. 分布式数据库的集群数量很多，总体 server 节点的数量必然很大，而且伴随业务的高速增长，数据中心的增加等因素，其集群数量和 server 节点的数量也将面临高速增加，数据库域名的数量同样会面临极大增长的压力。

2. 大量的域名及其快速的增长，带来的海量的域名和海量域名解析请求，而单 DNS 集群架构的容量和解析能力是无法不断持续扩容，这不但要求 DNS 系统架构具备高配置容量、高 DNS 解析能力，还要求 DNS 架构具备持续扩展的弹性扩容能力。

3. 利用智能 DNS 对分布式数据库进行自动的容灾切换，非常重要的一点是对分布式数据库的节点的服务状态探测。大量的分布式数据库节点服务状态的探测，必然会极大消耗 DNS 系统的性能；而且在一些场景下，正常情况副本节点为 down 的状态，且对服务探测无响应，只能等待探测超时，进一步增加了 DNS 对分布式数据库节点探测的性能消耗，这就要求 DNS 系统架构不但需要具备高性能的服务探测能力，还需具备优良的弹性扩容能力。

鉴于以上需求，建议此场景下将分布式数据库的 DNS 进行子域拆分，从业务网权威 DNS 中拆分出来，构建一套可弹性扩容的高性能的分布式数据库子域名分布式架构。

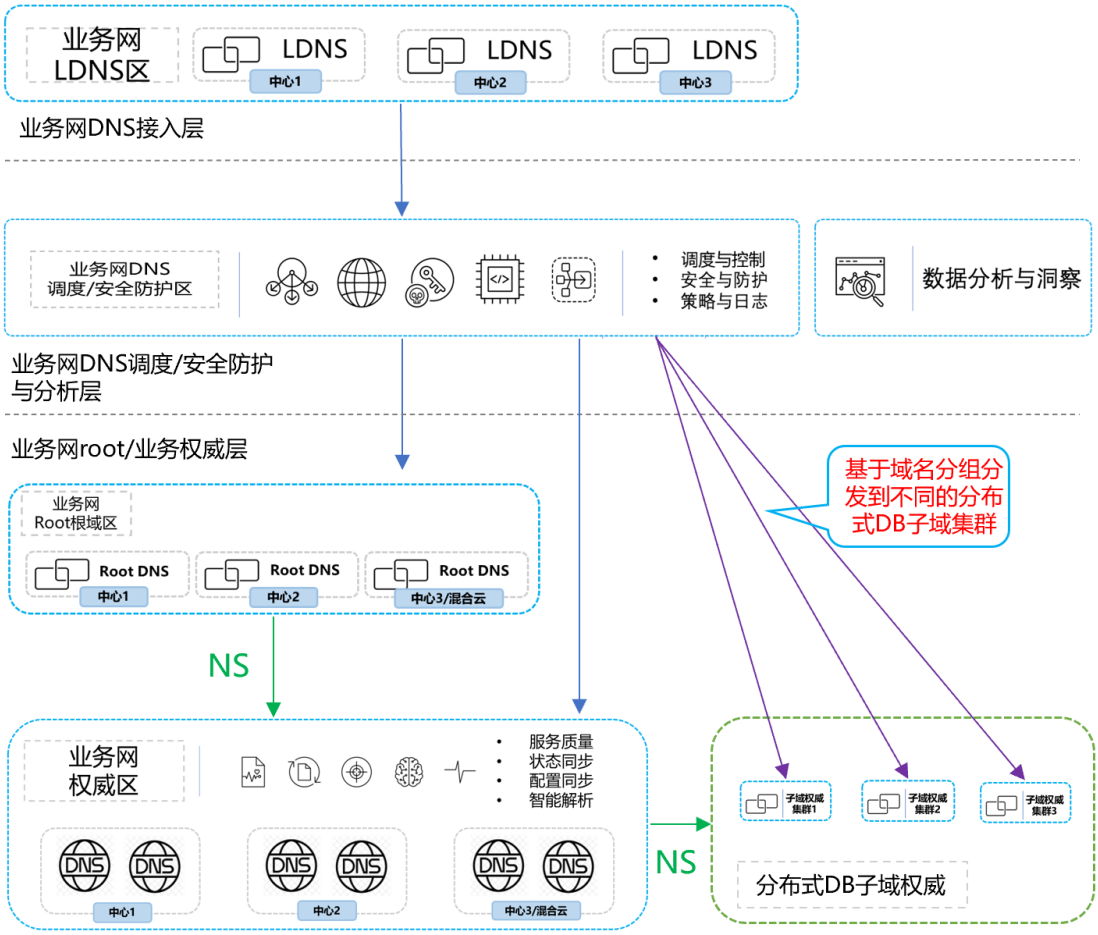
架构设计

分布式数据库 DNS 是业务网 DNS 的一部分，其架构主要涉及：

1. 业务网 DNS 调度/安全防护区

2. 业务网 DNS 权威名称解析区

3. 分布式数据库子域权威区



架构分析

1. 在业务网权威名称解析区通过 NS 方式，将分布式数据库子域拆分出来，并按照每个分布式数据库子域集群承载一定量的分布式数据库节点（即容量阀值），拆分成不同的子域 group 组，每个子域权威集群承担一个子域 group 组的智能 DNS 解析和服务探测。

2. 通过业务网 DNS 调度/安全防护区基于分布式数据库拆分的不同的子域 group 组，将分布式数据库的解析请求分发到对应的子域权威集群上进行处理。

3. 当分布式数据库子域权威集群的承载的数据库节点数量达到容量阀值时，则拆分新的子域权威集群。

技术建议

对于以上分布式数据库使用场景带来的 DNS 系统挑战，构建分布式数据库 DNS 的技术建议如下：

1. 分布式数据库子域权威单台设备的性能要高，以便单子域权威集群可承载更大数量的数据库节点配置、节点探测，及大量的 DNS 解析能力。
2. 单 DNS 集群的能力毕竟有限，可以利用业务网 DNS 调度/安全防护区高性能且灵活的调度能力和负载均衡的能力，构筑容量和性能可弹性扩展的分布式数据库子域权威的分布式架构，从而突破单 DNS 集群的容量和性能瓶颈，承载海量的数据库节点和 DNS 解析请求。

4) PaaS DNS

需求分析

在数字化转型的浪潮中，业务系统为了提高可靠性、可用性和安全性，企业需要采用现代化的云计算技术来开发、部署和管理业务系统。而 PaaS（Platform as a Service）作为一种云计算服务模型，正是为企业提供这样的云计算平台，用于开发、部署和管理业务系统。通过使用 PaaS，企业可以更加高效地进行应用程序开发，通过自动化部署和伸缩功能，实现业务系统的高度可用性和性能，通过可定制化和灵活性，适应不同的业务需求和变化，通过高度可用和安全性，

确保业务系统的连续性和数据安全性。

随着业务系统逐渐转移到 PaaS 系统中，PaaS 集群也逐渐从单集群扩展到多集群，这也要求 PaaS 系统需要专门的 DNS 系统来解决跨 K8s 集群和跨数据中心访问的问题，PaaS 单集群内部的 DNS 系统都已经通过 PaaS 集群实现，而 PaaS 集群外的 DNS 系统则还需要考虑自动化的全生命周期管理、海量域名的支持、服务注册和服务发现等问题。

PaaS DNS 作为多 PaaS 集群系统重要的组成部分需要满足以下要求：

1. 高可靠性和高可用性：需要保证 24 小时不间断运行，具备高可靠性和高可用性，确保应用程序能够随时正常获取解析服务。
2. 域名动态更新能力：具备通过 API 接口实现域名动态更新的能力，实现实时更新 DNS 记录的效果，以确保应用程序和服务的部署和迁移能够及时反映在 DNS 系统中，达到域名自动化的全生命周期管理。
3. 海量域名扩展性：首先需要单台设备支持海量域名的配置，其次整体系统支持高扩展性，以便能够应对不断增长的业务域名数量，最后需要考虑域名系统的分层处理，保证 DNS 系统的高性能和高扩展性。
4. 灵活的负载均衡策略：需要支持灵活的负载均衡策略，以便根据应用程序和服务的负载情况，实现最优的负载均衡策略。
5. 安全性：需要具备安全性能力，包括防止 DNS 欺骗攻击和其他网络安全攻击等。
6. 可视化管理界面：需要具备可视化管理界面，以便管理员能够方便地管理

和维护 DNS 系统，包括对 DNS 记录的管理、对 DNS 设备管理以及对 DNS 设备集群的管理。

7. 兼容性：需要具备与各种云计算平台和容器编排系统的兼容性，以便与其他组件和服务进行集成和协作。

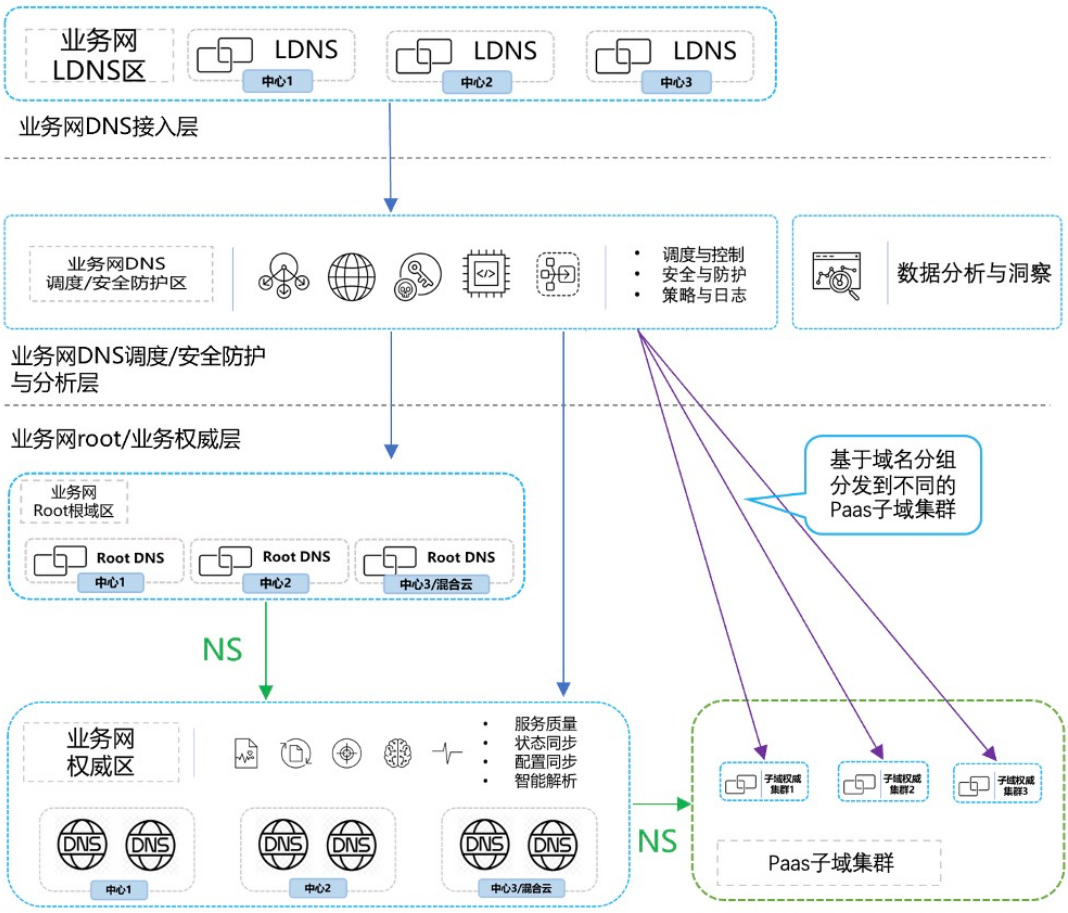
架构设计

容器和 PaaS DNS 是业务网 DNS 的一部分，其架构主要涉及：

业务网 DNS 调度/安全防护区

业务网 DNS 权威名称解析区

PaaS 子域权威区



架构分析

在业务网权威名称解析区通过 NS 方式，对 PaaS 子域进行子域授权，PaaS 子域的解析就由 PaaS DNS 实现。

考虑到所有 PaaS 集群业务量比较大，可以将 PaaS 子域分成多个子域，子域的具体数量根据每个 PaaS 集群承载的业务数量多少来确认。后续当 PaaS 集群数量或 PaaS 集群中业务数量发生变化，可以根据相应情况进行 PaaS 子域数量的调整。

每个 PaaS 子域权威承担响应 PaaS 集群业务的服务探测和智能 DNS 解析，

实现应用客户端获取并访问最优服务节点的效果。

技术建议

PaaS DNS 作为现代应用的重要组成部分，应重点考虑：

1. 高配置量能力：PaaS 业务快速增加，需要配置大量的新域名，需要 PaaS DNS 系统能够承载大量域名的配置，数量级需要达到几万到几十万级别。
2. 高自动化能力：提供方便的 API 接口对接 PaaS 发布系统，以实现灵活及时的域名动态更新。
3. 高可靠性架构：PaaS 平台通过自身的架构实现了平台的高可靠性架构，建议采用多组 DNS 设备配合负载均衡技术实现 DNS 高可靠性架构。
4. 高扩展能力：建议采用多组设备配合负载均衡技术实现高扩展能力。
5. 高安全性：建议部署 DNS DDOS 防护方案，采用具备智能学习特性，动态产生防护阀值，以避免人工静态阀值导致的防护缺陷。

4.DNS 可观测与运营

需求分析

DNS 系统的作为 IT 系统的基础设施之一，在建设完成后，就会转入运营阶段。DNS 的运营包括多个方面：DNS 系统设备管理、域名注册和管理、DNS 安

全管理、DNS 性能管理、DNS 数据管理和备份。

在 DNS 运营中，可观测性就成为最重要的方面，DNS 运营的各个阶段都需要以监控数据为基础，根据数据的不同制定运营方案和优化方案。

DNS 系统设备管理：需要对 DNS 服务器进行管理。DNS 服务器是进行域名解析的关键设备，DNS 运营需要确保部署足够的 DNS 服务器，以便处理来自全球各地的域名查询请求，并确保服务器的稳定性和安全性。DNS 服务器的具体数量就需要根据 DNS 具体的请求和响应数据以及 DNS 服务器的 CPU、内存等设备监控数据进行评估和部署。

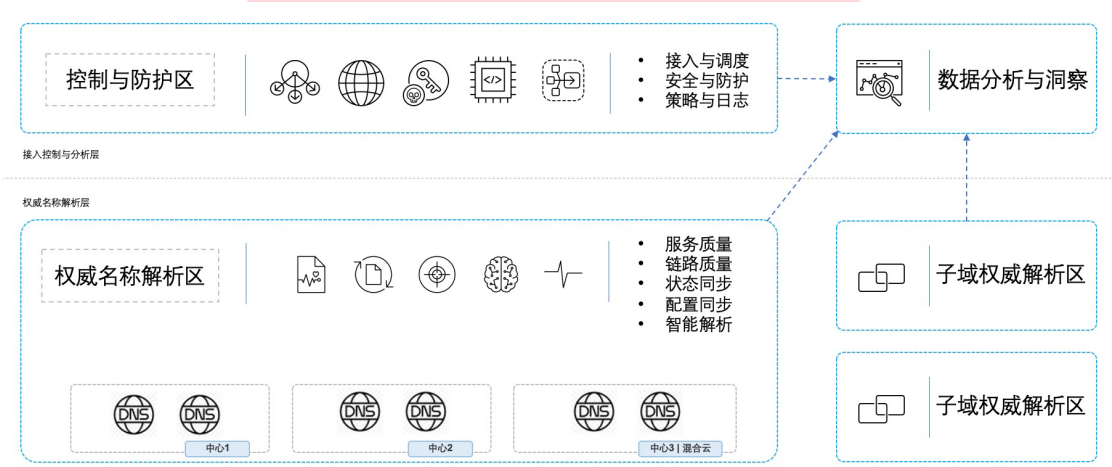
域名注册和管理：DNS 运营需要对域名进行注册和管理，包括验证域名所有者的身份，设置域名服务器，更新域名信息等。在这个过程中，需要确保完成注册和配置过程中各个阶段都是正常的，这就需要通过可观测性，对域名配置的过程进行监控，并根据可观测性的结果数据进行验证，如果出现问题或不合理的地方，则需要进行相应的调整。

DNS 安全管理：DNS 运营需要确保 DNS 系统的安全性，包括防止 DNS 缓存污染、DNS 欺骗等攻击，确保 DNS 服务器和 DNS 客户端的安全，以及确保数据的完整性和保密性。如何发现这些 DNS 安全问题，不能仅仅靠用户投诉或个人简单测试，而是需要通过 DNS 可观测性来实时监控和反馈。

DNS 性能管理：DNS 运营需要不断优化 DNS 系统的性能，包括优化域名解析速度，增加 DNS 缓存，减少网络延迟等。在这个过程中，DNS 可观测性是最重要的数据来源，根据可观测性的数据才能发现性能问题，并对优化后结果进行评估。

DNS 数据管理和备份：DNS 运营需要对 DNS 数据进行管理和备份，确保数据的完整性和可用性，并提供备份和灾难恢复计划。在这个过程中，需要监控数据备份系统的正常运行，以及备份数据的可用性，通过可观测性确保备份灾难恢复计划的正常运行

架构设计



在该架构中，通过 DNS 服务器或 DNS 请求和响应的关键节点（比如 DNS 服务器前部署的负载均衡设备）向数据分析系统发送 DNS 流量的关键指标信息。数据分析系统获取到数据后，需要从以下几个方面进行数据分析，并形成图形化报告：

1. 域名解析时间：DNS 客户端向 DNS 服务器发送请求，然后从服务器接收到响应所需的时间。这个指标是衡量 DNS 性能的重要指标，必须被监测和记录。
2. DNS 解析错误率：由于各种原因（如 DNS 服务器故障、网络故障等）而导致的 DNS 解析错误的比率。DNS 解析错误率的监测可以帮助识别网络和 DNS 服务器故障的根本原因，以便快速发现问题并进行修复和

改进。

3. DNS 流量：DNS 服务器处理的请求和响应的数量。监测 DNS 流量可以帮助评估 DNS 服务器的容量和性能，并预测未来的需求。
4. 域名解析的来源： DNS 请求的来源，例如，是来自哪个地区或哪个网络。了解解析的来源可以帮助识别网络瓶颈或安全问题，并为改进 DNS 性能提供有用的信息。
5. DNS 响应代码：DNS 服务器响应的状态代码。不同的响应代码表示不同的状态，如成功、服务器错误、域名不存在等。了解响应代码可以帮助诊断 DNS 问题，例如，识别是否由于域名拼写错误导致解析失败。
6. DNS 安全指标：DNS 服务器的安全性能指标，如 DNS 缓存污染、DNS 欺骗攻击等。 DNS 安全指标的监测可以帮助识别 DNS 安全问题，并实施相应的防御措施。

定期对观测得到的数据进行分析，在下面几个方面进行运营：

1. DNS 服务器维护：根据 DNS 服务器运行状态，定期进行巡检，及时处理故障。
2. 域名规划：需要根据域名数量的变化，及时对域名层级和子域进行规划和设置，保证业务需要新域名时能够及时申请到，各个子域域名数量控制在合理范围内。
3. 性能及协议优化：根据观测到的数据及时对 DNS 服务器进行扩容、升级，保证满足业务系统对 DNS 解析性能的要求。

4. DNS 安全：利用可观测的数据，合理设置报警参数，及时发现并应对安全问题，确保 DNS 服务器安全性以及解析数据准确性。
5. 数据备份：DNS 运营需要对 DNS 数据进行管理和备份，确保数据的完整性和可用性，并提供备份和灾难恢复计划。

技术建议

针对 DNS 可观测性和运营，建议采用如下方式：

1. 建议数据收集要简洁，只收集必要的准确数据，而不是将所有数据包都保存下来。最好采用日志的方式，便于传输、存储和分析。
2. 在系统关键节点进行全量 DNS 数据收集，建议在负载均衡设备收集数据，保证数据收集的全面性。
3. 分析系统的可视化能力，需要在图形化界面灵活显示各种数据及关联分析。
4. 运营系统的自动化建议，结合观测数据自动生产运营建议，比如达到一定性能要求就需要进行扩容，某子域域名达到一定数量，则需要规划新的子域等。
5. 数据备份的自动化，定期自动备份所有设备的配置数据，并定期对备份数据进行检验，确保灾难时能够及时恢复数据。

5.总结

DNS 在企业的基础架构中，整体上经历了四个阶段：

第一阶段，单中心多链路阶段，在该阶段企业主要关注应用在多条运营商链路下的解析优化。

第二阶段，多中心多活业务阶段，在该阶段企业主要关注在数地多中心模式下 DNS 的能力。

第三阶段，面向移动 App，边缘应用阶段，这也是当前大部分企业所处的阶段，在该阶段中，企业更加重视 DNS 架构的安全性、永续服务能力，以及面向移动与边缘应用的 DNS 技术应用，例如 DoH，k8s externalDNS，混合云 DNS 等。内网 DNS 也从早期的简单解析服务进入了多区、多层、全栈、动态的部署发展模式。

第四阶段，在部分 IT 科技发展较快的企业，DNS 系统也逐步进入了一个更新的发展方向，在该阶段中企业主要关注 DNS 在服务化改造、微服务架构以及 k8s 多集群等方面的能力集成。因此企业的 DNS 架构开始变得更加复杂，并开始呈现新的解耦化特征，例如子域拆分来降低故障域，分层与弹性的架构设计等。

而在 DNS 安全方面，解决 DNS 固有的安全问题也正被企业高度重视，DoH，DoT，DNSSEC 等技术被广泛接受和应用，企业更加重视基于数据洞察 DNS 运行效能和安全预警能力，建设灵活、智能、高安全、可自愈的 DNS 系统已成企业共识。

在本白皮书中，我们分析了内网、外网多个场景，以及可观测方案，通过对这

些不同场景的需求分析给出了相应的建议架构，并围绕架构中的关键技术做出了相关技术建议。这些内容均来自致力于该领域数年的专家们，他们落地了大量实际生产级的 DNS 案例，按章节顺序他们分别是：外网 DNS 章节的作者林静，DoH 章节的作者廖健雄，内网与可观测章节的王文宁、路瑞强。同时在编写过程中我们还邀请了长期为大型客户提供专业支持服务的技术专家：胡兆博、蒋旭、高阳，他们组成了技术专家组，为白皮书提供了诸多专业内容与最佳实践建议。最后还需特别感谢刘旭峰对本白皮书编写的领导与支持。

F5 市场销售热线：400 991 8366
F5 售后支持电话：400 815 5595，010-5643 8123
F5 在线联系：chinainfo@f5.com

F5 官方微信
公司新闻、行业趋势、
F5 社区（活动中心、
资料中心、粉丝论坛）

F5 社区技术群

NGINX
开源社区微信

NGINX 社区微信群

F5 公司北京办公室
地址：北京市朝阳区建国路 81 号华贸中心 1 号
写字楼 21 层 05-09 室
邮编：100025
电话：(+86) 10 5643 8000
传真：(+86) 10 5643 8100
https://www.f5.com.cn

F5 公司上海办公室
地址：上海市黄浦区湖滨路 222 号企业天地
1 号楼 1119 室
邮编：200021
电话：(+86) 21 6113 2588
传真：(+86) 21 6113 2599
https://www.f5.com.cn

F5 公司广州办公室
地址：广州市天河区珠江新城华夏路 10 号
富力中心写字楼 1108 室
邮编：510623
电话：(+86) 20 3892 7557
传真：(+86) 20 3892 7547
https://www.f5.com.cn

DC0923 | OV-GTM-1202116970