ReOpt: Near-Optimal Region Division for Low-Latency Regional Anycast

Yimeng Xu[†], Minyuan Zhou[†], Congyin Wang[†], Jiaqi Zheng[†*], Guihai Chen[†*], Jie Wu[‡*]
[†]The State Key Laboratory for Novel Software Technology, Nanjing University, China
[‡]China Telecom Cloud Computing Research Institute, China & Temple University, USA

***Abstract*—Regional IP anycast enhances traditional global anycast by organizing infrastructure into geographically-defined regions, each advertising distinct IP prefixes to attract local client traffic, though this approach introduces two key challenges: suboptimal intra-region routing from rigid geographic boundaries and cross-region path inflation due to excessive prefix multi-announcement. To address these, we propose ReOpt, an optimized anycast framework that dynamically minimizes latency through three key mechanisms: (1) real-time RTT measurements between client-site pairs and assess pairwise site preferences for each client, (2) intelligent multi-announcement strategies that enhance routing flexibility while maintaining stability, and (3) country-level region partitioning that simplifies DNS management while preserving geographic optimization. We formulate this as the Latency-Minimized Anycast Region Partition (LMARP) problem, prove its NP-hardness, and develop polynomial-time approximation algorithms with guaranteed approximation ratios. Our experimental evaluation on the PEERING testbed demonstrates that ReOpt-optimized regional anycast achieves significant latency improvements, reducing 90th percentile client latency by 4.6-9.3% across diverse partitions compared to conventional regional anycast, demonstrating its effectiveness in adapting to real network conditions beyond static geographic constraints.**

***Index Terms*—Anycast, CDN, Regional anycast, approximation algorithm.**

I. Introduction

IP anycast forms the backbone of critical Internet infrastructure — from DNS to CDNs — by advertising identical IP prefixes across geographically distributed sites [1], [2], [3]. Although BGP-driven traffic distribution inherently balances load, it fails to ensure optimal client-to-site assignments. Clients often connect to distant sites due to policy-driven routing (e.g., prioritizing customer paths over potentially lower-latency alternatives), degrading both latency and user experience [4], [5], [6], [7], [8], [9], [10]. This catchment inefficiency, well-documented in prior research, persists despite anycast's widespread adoption, highlighting a fundamental trade-off between BGP's operational simplicity and performance optimization [11], [12].

To address these limitations, regional IP anycast has emerged as a hybrid solution that leverages DNS-based client mapping while preserving anycast's operational benefits [13]. By partitioning sites into distinct regions (each announcing a unique anycast prefix) and using country-level IP geolocation [14], [15] to map clients during DNS resolution, this approach maintains anycast's simplicity while constraining routing decisions within regional boundaries, thereby reducing BGP-induced path inflation [16], [17].

However, recent measurement studies reveal that while regional anycast improves worst-case latency performance (e.g., reducing North American 90th-percentile latency from 110ms to 38ms in Imperva's network [18]), it introduces two significant operational challenges. First, the fixed geographic boundaries of regions can cause clients near region edges to be assigned to suboptimal sites despite better alternatives being physically closer. Second, the multi-announcement mechanism, designed for redundancy, can unexpectedly recreate path inflation problems similar to those in global anycast. These limitations fundamentally arise from current region partitioning schemes that prioritize administrative and geopolitical considerations over network performance metrics. Addressing these technical challenges is essential not only to validate the transition from global to regional anycast architectures, but also to fully realize the performance benefits promised by regional anycast deployments [18].

We propose ReOpt, a novel optimized anycast framework designed to systematically minimize end-to-end latency through three interconnected technical innovations. First, our framework initiates by establishing a comprehensive latency profile through real-time RTT measurements between all client-site pairs. This measurement process employs active probing techniques conducted at regular intervals [19], while simultaneously collecting passive traffic metrics from existing connections. For each client $u \in \mathcal{U}$, we construct a complete preference relation $\succ_u$ over the set of sites $\mathcal{S}$ based on the measurement results, where $s \succ_u s'$ indicates site s is strictly preferred over s' for client u [20]. Second, ReOpt incorporates intelligent multi-announcement strategies that go beyond conventional BGP advertisement approaches. By carefully orchestrating prefix announcements, we achieve enhanced routing flexibility while rigorously maintaining network stability — preventing common issues like suboptimal path selection. Third, the system introduces the country-level region partitioning scheme that provides dual benefits: it significantly simplifies DNS configuration management through clear administrative boundaries while still preserving the advantages of geographic optimization. This tripartite approach enables ReOpt to outperform traditional regional anycast implementations by dynamically balancing three critical factors: network performance (through latency minimization), opera-

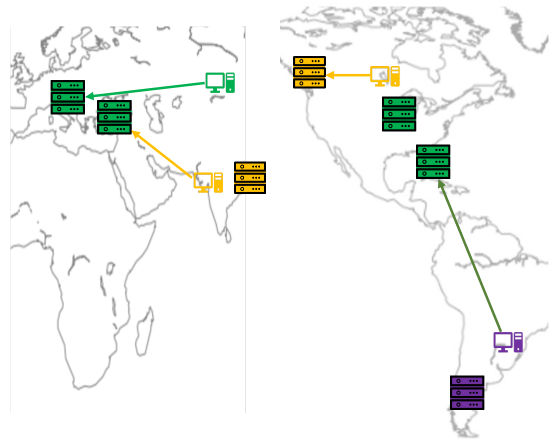
* Corresponding authors.
This work was supported in part by the NSF of China (62422207).

tional efficiency (via automated measurement and adjustment), and administrative practicality (with its hierarchical region management system). The framework's architecture particularly addresses the limitations of conventional geographic anycast by replacing rigid boundary definitions with adaptive, measurement-driven routing policies that respond to actual network conditions rather than theoretical proximity.

Our first contribution is that we not only formulate LMARP as an optimization framework but also derive its NP-hardness. Given a set of clients and geographically distributed sites, we first collect real-time RTT measurements between client-site pairs and assess pairwise site preferences for each client. Using this data, we construct a complete preference ordering for each client while incorporating three critical constraints: (1) BGP multi-announcement mechanisms that govern prefix advertisements, (2) quantifiable preference relationships that reflect latency and stability requirements, and (3) country-aware routing policies that ensure compliance with administrative boundaries. The resulting optimized region partition minimizes path inflation while maximizing the likelihood that clients are routed to their preferred (i.e., lowest-latency) sites, even when such assignments cross traditional geographic or administrative borders.

Our second contribution is the design of an efficient and scalable approximation algorithm ReOpt for solving LMARP. We illustrate our approach in three steps from simple to complex. Algorithm 1 addresses Simplified-LMARP (Section III-B) by initializing each site as an independent region and iteratively merging pairs, achieving an $\mathcal{O}(\log(m-\ell))$ approximation ratio, where m and ℓ denote the number of sites and regions, respectively. Algorithm 2 incorporates multi-announcement constraints, merging regions while dynamically reassigning sites to satisfy these constraints, while maintaining the same $\mathcal{O}(\log(m-\ell))$ ratio. Algorithm 3 solves the original LMARP by integrating both multi-announcement and country-aware constraints: it initializes regions per country to satisfy site constraints and introduces a user-site assignment function to comply with user constraints, yielding an $\mathcal{O}(\log(p-\ell))$ ratio (p is the number of countries). All three algorithms run in polynomial time, making them PTAS (Polynomial-Time Approximation Scheme) algorithms.

Our third contribution comprises a comprehensive real-world evaluation using the PEERING testbed [21], which employs BGP-speaking routers across 14 globally distributed sites with tier-1 ISP peering to emulate production CDN conditions. We benchmark ReOpt against geographically partitioned baselines and demonstrate its superior flexibility in optimizing performance-driven region design. Unlike rigid geographic boundaries, ReOpt intelligently merges and splits regions — e.g., combining Australia with North America or partitioning Europe into finer sub-regions — to minimize latency and improve routing efficiency. Multi-announcement sites (e.g., Madrid serving multiple regions) further enhance coverage and traffic control. These optimizations enable ReOpt to outperform traditional partitioning while maintaining comparable region counts, proving its effectiveness in real-world CDN deployments.



(a) *Edgio's 3-region partition.* (b) *Imperva's 6-region partition.*

Figure 1: *Sites and clients belonging to the same region are visually distinguished by matching color codes. The directional arrows represent empirically observed network paths collected during active client access measurements.*

II. Background and Motivation

In this section, we first introduce the background of global and regional anycast before analyzing the limitations of deployed regional IP anycast, focusing on Edgio and Imperva — two leading CDNs ranked among the top 15 worldwide.

A. Global anycast and regional anycast

IP anycast is a widely used routing technique in which multiple geographically distributed sites advertise the same IP prefix, enabling client traffic to be routed to the "nearest" site based on BGP path selection. This approach improves latency and load balancing for critical services like DNS and CDNs without requiring explicit client redirection. However, global IP anycast — where all sites announce a single prefix — faces catchment inefficiency: BGP's policy-based routing often prioritizes business relationships (e.g., customer routes over peer routes) rather than performance metrics like latency, leading to suboptimal client-site mappings and increased delays. Regional IP anycast addresses this inefficiency by dividing sites into geographically bounded regions (e.g., continents or large countries), each announcing a distinct IP prefix. By combining DNS resolution with IP geolocation, clients are mapped to their nearest regional prefix, effectively confining traffic within predetermined geographic boundaries. Unlike global anycast, this approach maintains operational simplicity while giving operators precise control over client-to-site mappings — without requiring BGP modifications. As a result, regional anycast significantly reduces catchment inefficiency by strictly limiting maximum client-site distances while preserving anycast's scalability benefits.

B. Limitations of deployed regional IP anycast

Prior research has demonstrated performance limitations in regional anycast implementations [18]. Building on these findings, we examine two major CDNs that employ regional anycast - Edgio and Imperva recently. Their deployments serve as representative case studies that motivate our investigation into optimization opportunities. Edgio uses 3-4 regional partitions while Imperva employs 6 (including separate prefixes for the U.S. and Canada), both aligning regions with geopolitical boundaries. We measure Edgio's 4-region and Imperva's 6-region partition to motivate our work. We use RIPE Atlas's user-defined measurement capability to first resolve regional anycast IP addresses through DNS queries to selected domains, then perform latency measurements by sending ping packets to these resolved IPs. Finally, we find two inefficiencies persist that diminish these gains.

First, rigid geographic boundaries in regional anycast often result in suboptimal routing decisions. For example, our measurements of Edgio's deployment reveal that a client in Russia is routed to Amsterdam despite having a lower-latency path to India (Figure 1a). Similarly, in Imperva's deployment (Figure 1b), Canadian clients near U.S. sites are forced to connect to Canadian-region sites due to strict country-based partitioning, incurring over 30 ms of additional latency. These cases demonstrate how inflexible region definitions can degrade performance even when closer alternatives exist.

Second, cross-region site announcements can introduce latency overhead when BGP routing policies diverge from geographic proximity. As shown in Figure 1a, when a European site announces an Asian prefix, some Indian clients may be routed to this distant location. We observe similar suboptimal routing in Imperva's deployment (Figure 1b), where a Florida-based site advertising prefixes for both North and Latin America causes some Brazilian clients to experience increased latency due to BGP's policy-driven path selection rather than minimal-distance routing.

These findings highlight a critical gap: existing region partitions are static and geography-driven, ignoring latency dynamics and intra-region connectivity. This paper aims to address the limitation of regional anycast by proposing and evaluating a latency-optimized region partition scheme. This study can provide new insights into how to design a regional anycast (combining region partition and strategic cross-region announcement) that achieves low client latency worldwide. Additionally, if this study experimentally validates the performance benefits of improved regional anycast, it could motivate more CDNs to adopt regional anycast.

III. An Optimization Framework

In this section, we introduce our optimization framework.

A. Problem Statement and Formulation

We first give a problem definition of LMARP. The key notations are shown in Table I.

Table I: Key notations in this paper.

Input	$\mathcal{C} = \{c_1, c_2, \dots, c_p\}$	Set of countries
	$\mathcal{R} = \{r_1, r_2, \dots, r_\ell\}$	Set of regions
	$\mathcal{S} = \{s_1, s_2, \dots, s_m\}$	Set of sites
	$\mathcal{U} = \{u_1, u_2, \dots, u_n\}$	Set of clients
	$RTT(u, s)$	RTT between client u and site s
	$RTT(s, s')$	RTT between site s and site s'
	$\succ_u$ and $\succeq_u$	Client u's preference over $\mathcal{S}$
Output	$U2S(u, s)$	zero-one variables
	$S2R(s, r)$	zero-one variables
	$UC2R(c, r)$	zero-one variables
	$SC2R(c, r)$	zero-one variables

Definition 1. *The Latency-Minimized Anycast Region Partition Problem (LMARP) is defined as follows: Given a set $\mathcal{U}$ of n clients, a set $\mathcal{S}$ of m sites, and for each client $u \in \mathcal{U}$, a total order $\langle \mathcal{S}, \succ_u \rangle$ representing site preferences, we incorporate country-aware constraints through country-specific subsets $\mathcal{S}_c \subseteq \mathcal{S}$ (sites) and $\mathcal{U}_c \subseteq \mathcal{U}$ (clients). These constraints require that (1) all sites in $\mathcal{S}_c$ belong to a single region r_1, and (2) all clients in $\mathcal{U}_c$ connect to a single region r_2. The objective is to find an ℓ-region partition of $\mathcal{S}$ (represented by binary variables $S2R(s, r)$), client-to-site assignments $U2S(u, s)$ and country-to-region assignments $SC2R(c, r), UC2R(c, r)$ that minimize total client latency.*

Using the problem definition and notation established above, we formulate Program 1 for LMARP.

$$\text{minimize} \quad \sum_{u \in \mathcal{U}} \sum_{s \in \mathcal{S}} RTT(u, s) \cdot U2S(u_i, s_j) \qquad (1)$$
$$\text{subject to} \quad \text{(1a), (1b), (1c), (1d), (1e), (1f), (1g).}$$

The objective function minimizes the total RTT across all clients $u \in \mathcal{U}$. All decision variables $U2S(\cdot, \cdot)$, $S2R(\cdot, \cdot)$, $UC2R(\cdot, \cdot)$ and $SC2R(\cdot, \cdot)$ are binary indicators: $U2S(u, s) = 1$ if client u connects to site s, and 0 otherwise; $S2R(s, r) = 1$ if and only if site s is assigned to region r (i.e., $s \in r$); $UC2R(c, r) = 1$ if and only if all clients $u \in \mathcal{U}_c$ connect to region r; $SC2R(c, r) = 1$ if and only if all sites in country c are covered by region r (i.e., $\mathcal{S}_c \subseteq r$). We now formally introduce and analyze the constraints (1a), (1b), (1c), (1d), (1e), (1f), (1g) and discuss them in detail.

$$\sum_{s \in \mathcal{S}} U2S(u, s) = 1, \ \forall u \in \mathcal{U} \qquad (1a)$$

Constraint (1a) enforces that each client $u \in \mathcal{U}$ is assigned to exactly one site $s \in \mathcal{S}$.

$$\sum_{r \ \in \mathcal{R}} S2R(s, r) \geqslant 1, \ \forall s \in \mathcal{S} \qquad (1b)$$

Constraint (1b) ensures that each site $s \in \mathcal{S}$ is assigned to at least one region $r \in \mathcal{R}$. We call it "multi-announcement" if site s is assigned to multiple regions.

$$U2S(u, s) \geqslant \sum_{s' : s \succ_u s'} U2S(u, s') \cdot S2R(s, r) \cdot S2R(s', r), \quad \forall u \in \mathcal{U}, s \in \mathcal{S}, r \in \mathcal{R} \qquad (1c)$$

Constraint (1c) ensures that client u selects the site s with the highest preference among all sites in region r. The term $S2R(s, r) \cdot S2R(s', r)$ evaluates to 1 if both s and s' belong to the same region r, and 0 otherwise. The inequality indicates that a site s with higher preference will have a correspondingly larger value of $U2S(u, s)$. Note that the preference relation $\succ_u$ is derived empirically from comprehensive network measurements. Section V details our methodology for measuring pairwise preferences via active RTT probing and passive monitoring of all client-site pairs, and constructing complete preference orderings through multi-criteria optimization incorporating both latency stability and path reliability metrics.

$$\sum_{r\in\mathcal{R}} SC2R(c,r)=1,\ \forall c\in\mathcal{C} \tag{1d}$$

$$\sum_{r\in\mathcal{R}} UC2R(c,r)=1,\ \forall c\in\mathcal{C} \tag{1e}$$

$$(SC2R(c,r)=1)\Rightarrow(S2R(s,r)=1), \tag{1f}$$
$$\forall c\in\mathcal{C}, s\in\mathcal{S}_c, r\in\mathcal{R}$$

$$(UC2R(c,r)=1)\Rightarrow(\bigvee_{s\in\mathcal{S}}[U2S(u,s)\wedge S2R(s,r)]=1), \tag{1g}$$
$$\forall c\in\mathcal{C}, u\in\mathcal{U}_c, r\in\mathcal{R}$$

Constraints (1d) and (1e) ensure that each country is assigned to exactly one region. Constraint (1f) guarantees that every site $s\in\mathcal{S}_c$ is covered by the same region as its associated country c. Constraint (1g) ensures that each client $u\in\mathcal{U}_c$ connects to the region r covering c through a site assigned to r. $a\Rightarrow b$ is true means that a is false or $a\wedge b$ is true. $(\bigvee_{s\in\mathcal{S}}[U2S(u,s)\wedge S2R(s,r)]=1)$ means that there exists site s satisfying $U2S(u,s)=1$ and $S2R(s,r)=1$. Country-aware constraints serve two critical functions in DNS operations. First, they align with the practical realities of IP geolocation, where country-level granularity represents the most reliable and widely-adopted standard for accurately mapping client requests to region-specific DNS responses. This approach accommodates the inherent limitations of geolocation databases while maintaining sufficient precision for most operational needs. Second, by enforcing consistent treatment for all clients within a country's borders, the system eliminates the need for complex per-client customization, reduces configuration errors, and simplifies policy enforcement. This dual benefit of operational practicality and administrative simplicity makes country-level constraints particularly valuable for large-scale, globally-distributed CDN infrastructures.

B. Hardness Analysis

We present a simplified version of LMARP, denoted Simplified-LMARP. Simplified-LMARP relaxes the original formulation by removing all country-aware constraints (1d), (1e), (1f), (1g) and simplifying multi-announcement to single-region assignment ($\sum_{r\in\mathcal{R}} S2R(s,r)=1$, replacing (1b)). We then establish the NP-hardness of Simplified-LMARP through a formal reduction. This result consequently proves that the original LMARP is also NP-hard.

Theorem 1. *Simplified-LMARP is NP-hard.*

Proof. We establish the NP-hardness of Simplified-LMARP through a polynomial-time reduction from the ℓ-median problem.

$$\text{Instance } \mathcal{I}=\begin{cases}\mathcal{R} & =\{r_1,r_2,\ldots,r_\ell,r_{\ell+1}\}\\ \mathcal{S} & =F\cup\{f_\infty\},\ \|F\|=m\\ \mathcal{U} & =X\\ RTT(u_i,s_j) & =\begin{cases} d(x_i,f_j) & ,\ f_j\neq f_\infty\\ \infty & ,\ \text{otherwise}\end{cases}\end{cases}$$

Since the ℓ-median problem is known to be NP-hard [22], this reduction consequently proves that Simplified-LMARP is also NP-hard. Based on an instance of ℓ-median problem (V,d,F,X), we construct a Simplified-LMARP $\mathcal{I}$ as above.

When the RTT between sites $\mathcal{S}=\{f_1^u,\ldots,f_m^u,f_\infty\}$ and client u form a non-increasing sequence $\text{RTT}(u,f_\infty)\geq \text{RTT}(u,f_1^u)\geq\cdots\geq\text{RTT}(u,f_m^u)$, we derive the client's preference ordering as $f_\infty\succeq_u f_1^u\succeq_u\cdots\succeq_u f_m^u$, where $\succeq_u$ represents the weak preference relation of client u over anycast sites.

Consider any $(\ell+1)$-region partition $\mathcal{R}_{\ell+1}=\{r_1,\ldots,r_{\ell+1}\}$ of $\mathcal{S}$ with $f_\infty\in r_{\ell+1}$ (without loss of generality). If any non-f_∞ region $r\in(\mathcal{R}_{\ell+1}\ominus r_{\ell+1})$ satisfies $|r|>1$, we can refine the partition by moving some site $f\in r$ to $r_{\ell+1}$, creating modified regions $r'=r\ominus\{f\}$ and $r'_{\ell+1}=r_{\ell+1}\oplus\{f\}$, where $\ominus$ and $\oplus$ are set minus operation and set union operation. For clients previously preferring f in r, they will now select their highest-preference alternative in r', which by construction cannot increase their RTT (and thus cannot worsen the objective). Crucially, since f_∞ remains the global preference in $r'_{\ell+1}$, this transfer preserves all other clients' optimal site selections and RTT values.

The finite solution space guarantees the existence of an optimal solution by the extreme value theorem. Our analysis reveals an optimal $(\ell+1)$-region partition $\mathcal{R}_{\ell+1}^{\text{OPT}}$ with the structure $r_k^{\text{OPT}}=\{f_{\pi_k}\}$ for $k=1,...,\ell$ and $r_{\ell+1}^{\text{OPT}}=\{f_{\pi_{\ell+1}},...,f_{\pi_m},f_\infty\}$, where $\{\pi_i\}_{i=1}^m$ is a permutation of $1,\ldots,m$. This induces an ℓ-median candidate solution $M=\{f_{\pi_1},...,f_{\pi_\ell}\}$. If M were suboptimal, replacing it with an optimal ℓ-median solution $M^{\text{OPT}}=\{f_{\pi'_1},...,f_{\pi'_\ell}\}$ would yield a new partition $r_k^{\text{OPT}}=\{f_{\pi'_k}\}$ (for $k=1,...,\ell$) and $r_{\ell+1}^{\text{OPT}}=\{f_{\pi'_{\ell+1}},...,f_{\pi'_m},f_\infty\}$ that reduces the total RTT — contradicting $\mathcal{R}_{\ell+1}^{\text{OPT}}$'s optimality. Thus, M must indeed be optimal for the ℓ-median problem.

In conclusion, we have demonstrated that any solution to the Simplified-LMARP would yield a solution to the NP-hard ℓ-median problem through our polynomial-time reduction, thereby establishing the NP-hardness of the Simplified-LMARP as formally stated in Theorem 1. □

IV. Algorithm Design and Analysis

In this section, we present a progressive approach to solving LMARP through three increasingly complex scenarios. We first address Simplified-LMARP by developing Algorithm 1. Building upon this foundation, we extend our solution to multi-announcement scenarios through algorithmic adjustments, yielding Algorithm 2. Finally, we incorporate country-aware constraints to solve the original LMARP, culminating in the design of Algorithm 3 as our comprehensive solution.

A. Solving Simplified-LMARP via Reverse Greedy Algorithm

We present Algorithm 1 for the Simplified-LMARP in Section III.B, employing a bottom-up approach that initializes each site as an independent region and iteratively merges optimal region pairs until reaching the target ℓ regions. Unlike top-down partitioning strategies that divide a single universal region, our method evaluates all possible merges at each step, computes the resulting total client latency, and selects the pair that minimizes this metric to form the new partition. This process continues until the desired number of regions is achieved, ensuring both efficiency and optimality in the merging sequence.

Specially, Algorithm 1 begins by initializing $\mathcal{R}_m$, where each site forms its own region (line 1). The algorithm then iteratively processes the current region partition: at each step, it identifies and merges the pair of regions whose combination yields the minimal performance degradation (lines 3-5). This process repeats until the partition size reduces to the target ℓ regions (lines 2-5). The $\mathcal{O}(m^3)$ time complexity arises from evaluating $\frac{t(t-1)}{2}$ potential merges per iteration for $t = m, m-1, ..., \ell+1$, maintaining polynomial-time efficiency throughout the computation.

Algorithm 1: Reverse Greedy Algorithm (RGA)

Input : $\mathcal{R}, \mathcal{S}, \mathcal{U}, RTT(u,s)$, Preference $\succ_u$
Output: $S2R(s,r)$, $U2S(u,s)$

1 Initiate $R_m = \{r_1 = \{s_1\}, r_2 = \{s_2\}, \ldots, r_m = \{s_m\}\}$.
2 **for** $t = m, m-1, \ldots, \ell+1$ **do**
3   Let $\mathcal{R}_t = \{r_t^1, r_t^2, \ldots, r_t^t\}$ be the current region partition.
4   Let $\mathcal{R}_t^{i,j} = \left[\mathcal{R}_t \ominus \{r_t^i\} \ominus \{r_t^j\}\right] \oplus \{r_t^i, r_t^j\},\ 1 \leqslant i < j \leqslant t$
      /* $\ominus$ and $\oplus$ in line 4 is set operation. */
5   Then $\mathcal{R}_{t-1} = \arg\min_{\mathcal{R}_t^{i,j}} \{Obj(\mathcal{R}_t^{i,j})\},\ 1 \leqslant i < j \leqslant t$
      /* $Obj(\mathcal{R}_t^{i,j})$ represents the minimum objective function value of $\mathcal{R}_t^{i,j}$, subject to constraints (1a) and (1c). */
6 Set $S2R(s,r)$ and $U2S(u,s)$ according to $\mathcal{R}_\ell$ and $Obj(\cdot)$.

Theorem 2. *The approximation ratio of the Reverse Greedy Algorithm (Algorithm 1) is $\mathcal{O}(\log(m-\ell))$, where m is the number of sites and l is the number of regions.*

Proof. We define a projection function $\Pi(r_x, \Phi)$ that maps a region $r_x \in \Phi'$ to an optimal target region $r_y \in \Phi$ by minimizing the total RTT between site $s(u, r_x)$ and $s(u, r_y)$ for all clients u assigned to r_x in the partition Φ': $r_y = \Pi(r_x, \Phi) = \arg\min_{r\in\Phi} \left\{ \sum_{u \mid u \to r_x} RTT\left(s(u,r_x), s(u,r)\right) \right\}$, where $u \to r_x$ indicates that client u connects to region r_x in Φ', and $s(u, r_x)$ (resp. $s(u,r)$) denotes u's serving site in r_x (resp. r).

Let $\mathcal{R}^{\text{OPT}} = \{r_1^{\text{OPT}}, r_2^{\text{OPT}}, \ldots, r_\ell^{\text{OPT}}\}$ denote the optimal region partition. At iteration t, define the projected partition $\Psi_t = \{\psi_1^t, \psi_2^t, \ldots, \psi_{p_t}^t\}$ where each $\psi_j^t = \Pi(r_j^{\text{OPT}}, \mathcal{R}_t)$ for $j = 1, \ldots, p_t$ with $p_t \leq \ell$. The following inequalities hold:

$$
\begin{aligned}
&Obj(\mathcal{R}_{t-1}) - Obj(\mathcal{R}_t)\\
&\leqslant \min_{r_t^i \neq r_t^j \in \mathcal{R}_t \ominus \Psi_t} \{Obj(\mathcal{R}_t^{i,j})\} - Obj(\mathcal{R}_t)\\
&\leqslant \frac{2}{(t-p_t)(t-p_t-1)} \sum_{r_t^i \neq r_t^j \in \mathcal{R}_t \ominus \Psi_t} \left[Obj(\mathcal{R}_t^{i,j}) - Obj(\mathcal{R}_t)\right]\\
&\leqslant \frac{2}{t-p_t}\left[Obj(\Psi_t) - Obj(\mathcal{R}_t)\right]\\
&\leqslant \frac{2}{t-\ell}\left[Obj(\Psi_t) - Obj(\mathcal{R}_t)\right]
\end{aligned}
$$

Given that $\mathcal{R}_t \ominus \Psi_t \subseteq \mathcal{R}_t$ and by the definition of the objective function $Obj(\cdot)$, we have $Obj(\mathcal{R}_{t-1}) \leqslant \min_{r_i^t \neq r_j^t \in \mathcal{R}_t \ominus \Psi_t} \{Obj(\mathcal{R}_t^{i,j})\}$. This establishes the first inequality. The second inequality follows since the minimum is bounded above by the average. For the third inequality, we examine client u's contribution to the left-hand side (LHS), which occurs if and only if u's connected region undergoes merging. The total contribution is bounded by $\frac{2}{(t-p_t)(t-p_t-1)} \cdot (t-p_t-1) \cdot [RTT\left(u, s(u,\Psi_t)\right) - RTT\left(u, s(u,\mathcal{R}_t)\right)]$, where $s(u,\Psi_t)$ and $s(u,\mathcal{R}_t)$ represent u's serving sites under partitions Ψ_t and $\mathcal{R}_t$ respectively. Summing these contributions across all clients $u \in \mathcal{U}$ validates the third inequality. Finally, the fourth inequality holds because $p_t \leq \ell$ by construction.

Prior to analyzing $Obj(\Psi_t) - Obj(\mathcal{R}_t)$, we establish two foundational assumptions: (1) triangle inequality holds for all latency measurements ($RTT(u,s), RTT(u,s')$, and $RTT(s,s')$, $\forall\ u \in \mathcal{U}$ and $\forall\ s \neq s' \in \mathcal{S}$), which is a common network property; and (2) the projection function $\Pi(\cdot,\cdot)$ satisfies the following inequality.

$$
\begin{aligned}
&\sum_{u \mid u \to r_x^{OPT}} RTT\left(s\left(u, r_x^{\text{OPT}}\right), s\left(u, \Pi(r_x, \Psi_t)\right)\right)\\
&\leqslant \sum_{u \mid u \to r_x^{OPT}} RTT\left(u, s\left(u, r\left(u, \mathcal{R}_t\right)\right)\right),
\end{aligned}
$$

where $s(u,r)$ denote the highest-preference site for client u in region r, and $r(u, \mathcal{R}_t)$ represent client u's connected region in partition $\mathcal{R}_t$. Given the premium inter-site links, complex client-site relationships, and the projection function's definition, we maintain that this assumption remains valid.

In addition, we have the following inequalities.

$$
\begin{aligned}
&\sum_{u \mid u \to r_x^{OPT}} RTT\left(u, s\left(u, r\left(u, \Psi_t\right)\right)\right)\\
&\leqslant \sum_{u \mid u \to r_x^{OPT}} RTT\left(u, s\left(u, \Pi\left(r_x^{OPT}, \Psi_t\right)\right)\right)\\
&\leqslant \sum_{u \mid u \to r_x^{OPT}} \Big[RTT\left(u, s\left(u, r_x^{\text{OPT}}\right)\right)\\
&\qquad + RTT\left(s\left(u, r_x^{\text{OPT}}\right), s\left(u, \Pi(r_x^{OPT}, \Psi_t)\right)\right)\Big]\\
&\leqslant \sum_{u \mid u \to r_x^{OPT}} \left[RTT\left(u, s\left(u, r_x^{\text{OPT}}\right)\right) + RTT\left(u, s\left(u, r\left(u, \mathcal{R}_t\right)\right)\right)\right]
\end{aligned}
$$

The first inequality follows from the definition of $r(u, \Psi_t)$, since the sum where clients select their optimal regions cannot exceed the sum using projected regions. The second inequality holds by our triangle inequality assumption, which generates terms of $Obj(\mathcal{R}^{\text{OPT}})$. The third inequality results from our projection function assumption, enabling the transformation of inter-site RTTs into terms of $Obj(\mathcal{R}_t)$.

By aggregating the inequalities across all $r_x^{\text{OPT}} \in \mathcal{R}^{\text{OPT}}$, we obtain the bound $Obj(\Psi_t) - Obj(\mathcal{R}_t) \leqslant Obj(\mathcal{R}^{\text{OPT}})$. This leads to our key result: $Obj(\mathcal{R}_{t-1}) - Obj(\mathcal{R}_t) \leqslant \frac{2}{t-\ell} Obj(\mathcal{R}^{\text{OPT}})$. Summing over iterations $t = m, m-1, \ldots, \ell+1$ yields the final approximation ratio of $\sum_{i=1}^{m-\ell} \frac{2}{i} \approx \mathcal{O}\left(\log(m-\ell)\right)$. ∎

B. Solving Simplified-LMARP with Multi-Announcement constraints

We now extend our approach to incorporate multi-announcement constraints, allowing sites to advertise multiple regional prefixes. During each iteration, alongside the region-merging operation that minimizes total latency increase, we introduce an additional optimization step: for each unassigned site-region pair (s_j, r_k), we compute the potential latency reduction from assigning s_j to r_k, then implement the most beneficial assignment. This process repeats until no further beneficial assignments exist, yielding an optimized partition that satisfies multi-announcement constraints. To maintain solution quality, we prevent premature multi-announcement operations when regions are too small, as early application could lead to region domination and degraded performance. The procedure maintains polynomial-time complexity, as each assignment evaluation requires only polynomial-time computations relative to the problem size.

In Algorithm 1, each iteration performs region merging by selecting candidate regions r_i and r_j to minimize the objective function increase, creating $r'_i \leftarrow r_i \cup r_j$ and updating site-to-region mappings as $S2R(s, r'_i) \leftarrow S2R(s, r_i) \vee S2R(s, r_j)$ for all $s \in \mathcal{S}$. The extended version (Algorithm 2) handles multi-announcement scenarios through a two-phase approach: (1) an initial merge phase (lines 2-3) generates an ℓ-partition without multi-announcement, followed by (2) an optimization phase (lines 4-5) that strategically flips $S2R(s_x, r_y)$ mappings from 0 to 1 to maximize objective improvement. This phased approach maintains polynomial-time complexity while preserving solution quality, with the second phase iterating until convergence by selecting optimal site-region pairs (s_x, r_y) for announcement activation.

Algorithm 2: RGA-MA

Input : $\mathcal{R}, \mathcal{S}, \mathcal{U}, RTT(u, s)$, Preference $\succ_u$
Output: $S2R(s, r)$, $U2S(u, s)$

1 Initiate $\mathcal{R}_m = \{r_1 = \{s_1\}, r_2 = \{s_2\}, \ldots, r_m = \{s_m\}\}$.
2 **for** $t = m, m-1, \ldots, \ell+1$ **do**
3 $\quad \mathcal{R}_{t-1} = \arg\min_{\mathcal{R}_t^{i,j}} \{Obj(\mathcal{R}_t^{i,j})\},\ 1 \leqslant i < j \leqslant t$
4 **while** $\exists Obj(\mathcal{R}'_\ell(x, y)) < Obj(\mathcal{R}_\ell)$ **do**
5 $\quad \mathcal{R}_\ell = \arg\min_{\mathcal{R}'_\ell(x,y)} \{Obj(\mathcal{R}'_\ell(x, y))\}$
$\quad$ /* Ensure that there exist no $r_1, r_2 \in \mathcal{R}'_\ell(x, y)$ such that $r_1 \subseteq r_2$, where $\mathcal{R}'_\ell(x, y)$ denotes the set $\mathcal{R}_\ell$ under the condition $S2R(s_x, r_y) = 1$. */
6 Set $S2R(s, r)$ and $U2S(u, s)$ according to $\mathcal{R}_\ell$ and $Obj(\cdot)$.

Let $\mathcal{R}_t^{\text{merge}}$ and $\mathcal{R}^{\text{ma}}$ denote the solutions obtained after phase (1)'s each iteration (lines 2-3) and phase (2) (lines 4-5), respectively. And $\mathcal{R}^{\text{OPT-ma}}$ and Ψ_t^{ma} represent the optimal multi-announcement-enabled solution and its corresponding projection. Following the proof methodology of Theorem 2, we derive the following key inequalities.

$$Obj(\mathcal{R}_{t-1}^{\text{merge}}) - Obj(\mathcal{R}_t^{\text{merge}}) \leqslant \frac{2}{t-\ell}\left[Obj(\Psi_t^{\text{ma}}) - Obj(\mathcal{R}_t^{\text{merge}})\right]$$
$$Obj(\Psi_t^{\text{ma}}) - Obj(\mathcal{R}_t^{\text{merge}}) \leqslant Obj(\mathcal{R}^{\text{OPT-ma}})$$

Considering lines 4-5 of Algorithm 2, we establish that $Obj(\mathcal{R}^{\text{ma}}) \leqslant Obj(\mathcal{R}_\ell^{\text{merge}})$, which consequently preserves the inequality $Obj(\mathcal{R}^{\text{ma}}) - Obj(\mathcal{R}_m) \leqslant Obj(\mathcal{R}_\ell^{\text{merge}}) - Obj(\mathcal{R}_m)$. While multi-announcement enables potentially better optimal solutions, Algorithm 2 maintains the same $\mathcal{O}(\log(m-\ell))$ approximation ratio as Algorithm 1, demonstrating its theoretical robustness.

C. Solving LMARP

In this section, we incorporate country-aware constraints into our formulation, culminating in the complete optimization framework presented as Program 1. This extended model maintains all the characteristics of the previous problems while enforcing geographical compliance through explicit country-aware restrictions, thereby providing a comprehensive solution to the original problem with real-world deployment considerations.

Compared with Algorithm 2, we make two changes in terms of the initial solution and the total latency calculation. First, taking each site as an independent region in the beginning would violate the site-country constraint. We initialize the solution as the set of sites in each country. So, the corresponding constraint will always be satisfied. Second, the previous total latency calculation function $Obj(\cdot)$ does not consider the client-country constraint. We replace it with a new one $Obj^c(\cdot)$ and the calculation function can be aligned with the corresponding constraint. Note that the algorithm introduces two new output variables: $SC2R(c, r)$ for site-country constraints and $UC2R(c, r)$ for client-country constraints.

Building on the optimization framework from Section III, we incorporate country constraints by defining country-specific site and client sets $\mathcal{S}_c$ and $\mathcal{U}_c$. These enforce that: (1) all sites in $\mathcal{S}_c$ must be contained within a single region, and (2) all clients in $\mathcal{U}_c$ must connect to a single region. To guarantee satisfaction of constraints (1d) and (1f), we initialize the solution with country-level site groupings $\{\mathcal{S}_c : c \in \mathcal{C}\}$ rather than individual sites, preserving these constraints throughout all merge and multi-announcement operations. For client assignments, we formulate a constrained objective function $Obj^c(\cdot)$ that explicitly accounts for (1e) and (1g).

Algorithm 3: RGA-MA-C (ReOpt)

Input : $\mathcal{C}, \mathcal{R}, \mathcal{S}, \mathcal{U}, RTT(u, s)$, Preference $\succ_u$
Output: $S2R(s, r)$, $U2S(u, s)$, $SC2R(c, r)$, $UC2R(c, r)$

1 Initiate $\mathcal{R}_p = \{r_1 = \{\mathcal{S}_{c_1}\}, r_2 = \{\mathcal{S}_{c_2}\}, \ldots, r_p = \{\mathcal{S}_{c_p}\}\}$.
2 **for** $t = m, m-1, \ldots, \ell+1$ **do**
3 $\quad \mathcal{R}_{t-1} = \arg\min_{\mathcal{R}_t^{i,j}} \{Obj^c(\mathcal{R}_t^{i,j})\},\ 1 \leqslant i < j \leqslant t$
4 **while** $\exists Obj^c(\mathcal{R}'_\ell(x, y)) < Obj^c(\mathcal{R}_\ell)$ **do**
5 $\quad \mathcal{R}_\ell = \arg\min_{\mathcal{R}'_\ell(x,y)} \{Obj^c(\mathcal{R}'_\ell(x, y))\}$
6 Set $S2R(s, r)$, $U2S(u, s)$, $SC2R(c, r)$ and $UC2R(c, r)$ according to $\mathcal{R}_\ell$ and $Obj^c(\cdot)$.

Algorithm 3 extends the previous approach with country-aware constraints. It initializes with a solution $\{\mathcal{S}c_h\}_{h=1}^p$ (line 1) to satisfy constraints (1d) and (1f), then performs operations similar to Algorithms 1 and 2 but using the constrained objective function $Obj^c(\cdot)$ (lines 2-5), ensuring compliance with (1e) and (1g). The projection function $\Pi(\cdot,\cdot)$ inherently preserves country constraints without modification. Crucially, Theorem 2's analysis applies to this extension, maintaining both the $\mathcal{O}(\log(p-\ell))$ approximation ratio and all country-aware constraints.

V. Evaluation

In this section, we use real-world experiments on the anycast testbed to evaluate ReOpt.

A. Setup

We leverage two public measurement platforms: RIPE Atlas [23] for global vantage points and the PEERING testbed [21] for controlled anycast experimentation.

RIPE Atlas is a global measurement infrastructure comprising over 10,000 geographically distributed probes, each capable of performing periodic predefined measurements. Each probe's location (latitude, longitude) is publicly documented. We leverage RIPE Atlas's custom measurement capability to conduct DNS queries, ping tests, and traceroutes. Our study utilizes all 5,092 active probes spanning 137 countries. However, RIPE Atlas has a well-documented geographic bias, with most probes concentrated in Europe and North America. This uneven distribution could skew performance evaluations of regional anycast by over-representing these regions. To mitigate this bias, we analyze results by geographic region and present performance metrics separately for each region: (1) **EMEA** (Europe, Middle East, and Africa) has 3,626 unique probes. (2) **NA** (North America, excluding countries in Central America) has 804 unique probes. (3) **LatAm** (South America and countries in Central America) has 125 unique probes. (4) **APAC** (Asia and Pacific) has 537 unique probes.

The PEERING testbed enables researchers to inject BGP announcements into the global Internet routing system. It integrates Verfploeter [19] to measure catchment areas per anycast site and determine client preferences. With 14 operational sites and the ability to announce two /24 IP prefixes, we leverage PEERING to conduct a controlled comparison between traditional anycast and ReOpt-optimized regional anycast performance.

Benchmark Schemes: We evaluate three commercially deployed regional IP anycast services [18] as performance baselines in our study: Edgio-3 (EG3), Edgio-4 (EG4), and Imperva-6 (IM6). EG3 spans America (North and South), Asia-Pacific (APAC), and Europe-Middle East-Africa (EMEA). EG4 covers North America, South America (as separate regions), APAC, and EMEA. IM6 encompasses the United States, Canada, South America, APAC, EMEA, and Russia as distinct operational regions.

Data Imputation: While our methodology enables measurement of round-trip times (RTTs) between any client-site pair and can determine client preferences for any pair of sites, two key limitations persist: (1) incomplete RTT measurements due to testbed constraints, and (2) partial preference data that may not form a complete or consistent total order across all sites for a given client. In this section, we present our data imputation strategy to address these gaps and describe the subsequent processing steps to derive a robust total preference order for analysis. For the first limitation, we employ a geospatial estimation approach based on established Internet latency-distance relationships. Following the methodology of [24], we approximate RTT values between clients and sites using their geographical great-circle distances.

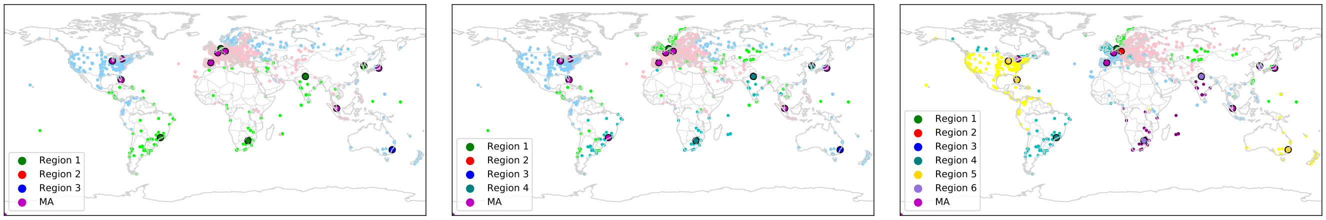
$$RTT \approx \frac{\text{geo-dist}}{2c/3} \times 2 = \frac{\text{geo-dist (km)}}{1 \times 10^5 \text{(km/s)}} \cdot \frac{1000\text{ms}}{1\text{s}},$$

where $c \approx 3 \times 10^5$ (km/s) is the velocity of light. For the second limitation, we begin by modeling each client u's measurement results as a round-robin tournament between all site pairs. For every pair of sites (i, j), we assign binary outcomes: if client u prefers site i over j, we record this as "i defeats j" with corresponding scores of $+1$ for i and -1 for j. When measurement data is unavailable for a pair, we treat it as a draw (0 points for both sites). This process generates a complete scoring matrix for each client-site combination. Next, we rank all sites based on their aggregate scores, which would produce a strict total order if the data were complete and consistent. In cases where sites have identical scores (indicating either missing data or true equivalence), we break ties using the measured RTTs from client u to the sites in question. This two-stage sorting mechanism — first by preference scores, then by latency — ensures we derive a deterministic total preference order for each client.

B. Performance Comparisons

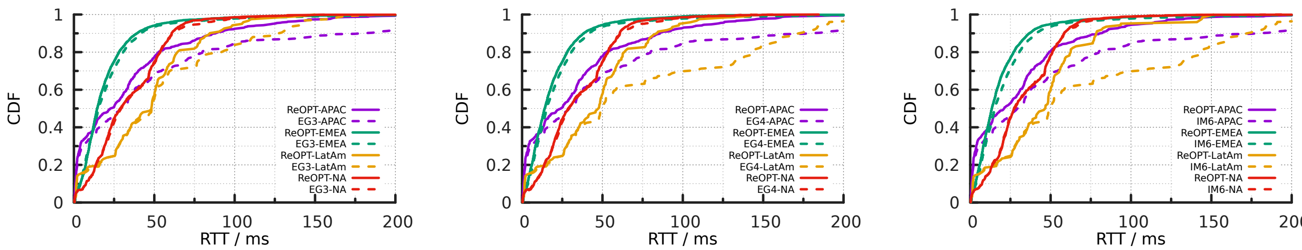
We establish geographically partitioned regions as our baseline using measured configurations from Edgio-3, Edgio-4, and Imperva-6 [18], generating ReOpt-optimized partitions with matching region counts for direct comparison. Unlike the baseline's strict geographical boundaries, ReOpt employs a flexible partition strategy tailored for performance optimization, as illustrated in Figure 2. In the 3-region partition, ReOpt splits America into North and Mid-South, merging the latter with Africa, India, and other discrete blocks into Region 1, while combining Australia (separated from APAC), North-East Europe (from EMEA), and North Asia (from APAC) with North America to form Region 2, and merging Southeast Asia (from APAC) with the remaining EMEA as Region 3. The 4-region partition follows a similar structure but separates part of South America, South Africa, and India from Region 1 into an independent Region 4. For the 6-region partition, Europe splits into West and East as independent regions, with Canada joining South America and Russia merging into East Europe, unlike IM6 where they were standalone regions. The figure also visually highlights sites that adopt a multi-announcement strategy, such as the Madrid site, which announces prefixes for both Region 2 and Region 5 (as shown in Figure 2c), enabling a single site to serve multiple regions simultaneously while enhancing coverage and allowing finer-grained traffic control across region boundaries.

To evaluate ReOpt's effectiveness, we compare its performance-optimized partitioning against the baseline's rigid



(a) *ReOpt's 3-region partition.* (b) *ReOpt's 4-region partition.* (c) *ReOpt's 6-region partition.*

Figure 2: *The visualization represents PEERING sites as large colored dots and client probes as small dots. Sites supporting multi-announcement are labeled with MA.*



(a) *Edgio-3 v.s. ReOpt-3.* (b) *Edgio-4 v.s. ReOpt-4.* (c) *Imperva-6 v.s. ReOpt-6.*

Figure 3: *CDFs of RTT for Edgio-3, Edgio-4, Imperva-6 and ReOpt-optimized regional anycast partitions.*

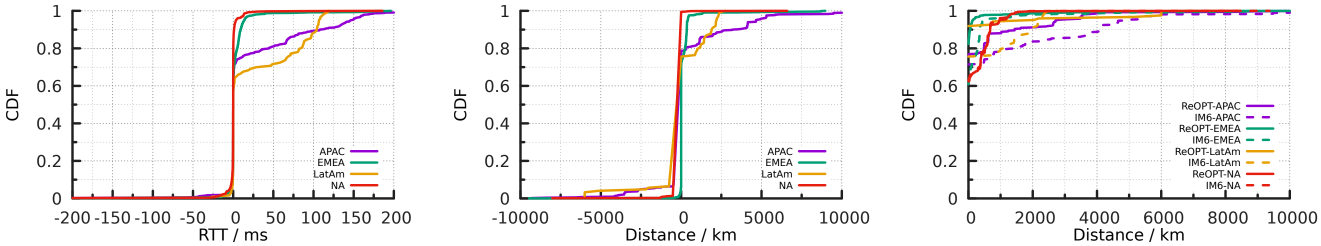
geographical strategy, using RTT as the key metric. Figure 3 shows RTT improvements across ReOpt's 3-, 4-, and 6-region configurations, with the 90th percentile RTT reduced by 4.6 ms (3-region), 5.4 ms (4-region), and 5.2 ms (6-region) compared to their baseline counterparts. These gains highlight ReOpt's ability to mitigate latency regardless of region granularity. While improvements in EMEA and NA are modest, APAC and LatAm see significantly larger reductions: the 90th percentile RTT drops by 33.7% (3-region), 51.4% (4-region), and 54.6% (6-region) in APAC, and by 50.9% (3-region), 52.2% (4-region), and 54.8% (6-region) in LatAm, underscoring the advantages of performance-aware partitioning. These results demonstrate that ReOpt effectively addresses performance inefficiencies in regions where traditional anycast configurations fall short, particularly in geographically dispersed or underserved areas. The findings highlight the critical role of adaptive, performance-driven partitioning in enhancing regional anycast performance, proving that rigid geographical boundaries are often suboptimal for latency optimization.

C. Deep Dive to the Improvements

This section analyzes the performance advantages of ReOpt over traditional partitioning methods through an in-depth examination of the ReOpt-6 configuration. As demonstrated in Figure 4, ReOpt-optimized regional partitioning significantly improves both client-to-site distances and RTTs compared to conventional continent-based Imperva-6 (IM6). Our results reveal a strong positive correlation between clients achieving reduced physical distances and those experiencing latency improvements under ReOpt. Notably, 23.6% of all clients connect to geographically closer sites in the optimized scheme, with particularly significant gains in the APAC and LatAm regions: 21.4% and 24.6% of regional clients respectively benefit from reduced distances, translating to substantial latency improvements of 83.4 ms (APAC) and 93.6 ms (LatAm) on average. These findings demonstrate ReOpt's ability to overcome the limitations of coarse-grained geographical partitioning through its performance-aware optimization approach.

The performance gains achieved by ReOpt originate from its novel partitioning strategy, which: (1) clusters clients with their empirically observed lowest-latency sites, and (2) strategically aggregates countries only when doing so preserves routing quality for all affected clients. Unlike traditional geographic approaches, ReOpt optimizes directly for measured network performance rather than physical proximity, sometimes producing geographically non-intuitive but latency-optimal groupings. For instance, while the grouping of Australia with the United States (shown in Figure 2c) might appear suboptimal due to potential trans-Pacific latency penalties, our measurements confirm zero instances of Australian clients actually being routed to US sites. By contrast, conventional continent-based partitioning (where Australia belongs to APAC) results in 8.0% (7/87) of Australian clients being routed to Delhi-based sites, incurring a substantial 165.9 ms average latency penalty. This demonstrates how ReOpt's performance-driven methodology avoids such pathological routing cases that occur in traditional geographic schemes.

To evaluate path inflation, we measure the additional distance clients traverse under different partitioning schemes. Figure 4c demonstrates that ReOpt achieves significantly more efficient client-to-site routing: 78.7% of clients connect to their geographically closest site — a 16.0 percentage point improvement over traditional partitioning. Notably, ReOpt nearly eliminates extreme path inflation (>4000 km) for APAC clients (1.3% vs. 11.4% under traditional schemes). These results highlight a fundamental architectural difference: while



(a) *RTT difference.* (b) *Distance difference.* (c) *Extra distance.*

Figure 4: *CDFs of (a) RTT, (b) distance difference, and (c) path inflation between imperva-6 and ReOpt-6.*

conventional geographic partitioning minimizes tail latency by restricting clients to regional sites — often forcing suboptimal intra-region connections — ReOpt's design simultaneously reduces both extreme path inflation and tail latency through its intelligent client-site matching strategy.

D. Large-scale Simulations

In this section, we conduct simulations using randomly generated data to evaluate our algorithms' solution quality and runtime performance. We employ Gurobi [25], a state-of-the-art convex optimizer, running on an Intel i5-12400F CPU with 32GB RAM. Our experimental setup consists of 30 countries, a target of 3 regions, and randomly generated RTT and preference matrices. Client and site counts are detailed in Table II, with clients and sites (if $|\mathcal{S}| > |\mathcal{C}|$) randomly distributed across countries. We benchmark ReOpt against the offline optimal solution (OPT) computed by the solver, with results presented in the following.

Table II: SIMULATION RESULTS

		(Number of Sites, Number of Clients)		
		(10, 200)	(30, 1000)	(200, 5000)
Total RTT	ReOpt	73.69 s	446.26 s	37.78 min
	OPT	70.88 s	493.24 s	Out of Memory
Running Time	ReOpt	0.70 s	9.10 min	8.49 h
	OPT	2.43 h	$>$ 24 h	Out of Memory

Table II demonstrates ReOpt's performance in terms of solution quality and computational efficiency. The upper half reveals that our approach achieves excellent RTT results, particularly in the (10, 200) case. The lower half shows ReOpt's superior runtime performance, solving the same (10, 200) instance in under 1 second compared to OPT's 2.5 hours — a difference attributable to OPT's computational complexity from nonlinear 0-1 integer constraints (Eqs. (1c)-(1g)). Scaling analysis reveals an important efficiency characteristic: while problem size grows proportionally (e.g., $\frac{30}{10} < \frac{200}{30}$), runtime increases sub-linearly ($\frac{9.10\text{min}}{0.70\text{s}} \gg \frac{8.49\text{h}}{9.10\text{min}}$). We infer that it occurs because when $|\mathcal{S}| > |\mathcal{C}|$, runtime depends primarily on $|\mathcal{U}|$ rather than $|\mathcal{S}|$, as $|\mathcal{C}|$ bounds site-related computations. Consequently, ReOpt maintains high efficiency even for large-scale problems, making it practical for real-world deployment.

VI. RELATED WORK

We review the prior arts on anycast and partition problem.

Global anycast and regional anycast. Prior work has extensively studied IP-layer anycast deployment and performance, with particular focus on root DNS servers [1], [4], [5], [6], [7], [8], [9], [26], [27], [2], [28] and global CDNs [29], [2], [28]. Recent advances in regional anycast include LinkedIn's reported latency improvements through private CDN migration [13] and Zhou *et al.*'s systematic analysis of deployment strategies and optimization opportunities [18]. While existing approaches like Zhou's rely on geolocation-based partitioning, our work fundamentally differs by constructing regions based on direct latency measurements rather than geographical proximity, enabling more accurate performance optimization.

Partition optimization. Our problem can be framed through the lens of facility location theory by mapping "customer-facility" relationships to "client-site" interactions. Existing work on Facility Location Problems (FLP) offers valuable algorithmic insights: foundational studies have developed constant-factor approximation algorithms using LP rounding [30], dual-fitting [31], factor-revealing LP [32], primal-dual methods [32], and local search [33]. These approaches extend naturally to the k-median problem, whose Lagrangian relaxation reduces to FLP [32], with [32], [33] providing theoretical guarantees for such extensions. Notably, [34] introduces a reverse greedy algorithm achieving bounds between $\Omega(\frac{\log n}{\log \log n})$ and $\mathcal{O}(\log n)$. However, our LMARP introduces unique challenges: the distance metric incorporates multidimensional network characteristics beyond geometric distance, and clients within a single region may connect to multiple sites (unlike traditional FLP/k-median assignments where cluster elements share a single facility). These constraints necessitate novel partitioning strategies that account for heterogeneous client-to-site mappings while preserving theoretical guarantees.

VII. CONCLUSION

This paper presents a novel approximation algorithm designed to minimize total RTT in regional IP anycast systems. We establish three key contributions: (1) the development of a polynomial-time algorithm with provable approximation guarantees, (2) comprehensive theoretical analysis of its performance bounds, and (3) empirical validation through both real-world PEERING testbed experiments and trace-driven simulations. Our evaluations demonstrate significant improvements in total RTT compared to existing approaches while maintaining computational efficiency competitive with optimal solutions.

References

[1] H. Ballani and P. Francis, "Towards a global ip anycast service," in *Proceedings of ACM SIGCOMM*, 2005.

[2] T. Koch, K. Li, C. Ardi, E. Katz-Bassett, M. Calder, and J. Heidemann, "Anycast in context: A tale of two systems," in *Proceedings of ACM SIGCOMM*, 2021.

[3] X. Zhang, S. Lin, T. Huang, B. M. Maggs, K. Schomp, and X. Yang, "Characterizing anycast flipping: Prevalence and impact," in *Proceedings of PAM*, 2025.

[4] H. Ballani, P. Francis, and S. Ratnasamy, "A Measurement-Based Deployment Proposal for IP Anycast," in *Proceedings of ACM IMC*, 2006.

[5] S. Sarat, V. Pappas, and A. Terzis, "On the use of anycast in dns," in *Proceedings of ACM SIGMETRICS*, 2006.

[6] H. U. L. Colitti, E. Romijn and A. Robachevsky, "Evaluating the effects of anycast on dns root name servers," *RIPE document*, 2006.

[7] Z. Liu, B. Huffaker, M. Fomenkov, N. Brownlee, and K. Claffy, "Two days in the life of the dns anycast root servers," in *Proceedings of PAM*, 2007.

[8] J. Kuipers. (2015) Analysing the k-root anycast infrastructure. [Online]. Available: https://labs.ripe.net/Members/jh_kuipers/analyzing-the-k-root-anycast-infrastructure

[9] Z. Li, D. Levin, N. Spring, and B. Bhattacharjee, "Internet anycast: Performance, problems, and potential," in *Proceedings of ACM SIGCOMM*, 2018.

[10] A. S. M. Rizvi, T. Huang, R. Esrefoglu, and J. Heidemann, "Anycast polarization in the wild," in *Proceedings of PAM*, P. Richter, V. Bajpai, and E. Carisimo, Eds., 2024.

[11] C. Miao, Z. Zhong, Y. Xiao, F. Yang, S. Zhang, Y. Jiang, Z. Bai, C. Lu, J. Geng, Z. He, Y. Wang, X. Zou, and C. Yang, "Megate: Extending wan traffic engineering to millions of endpoints in virtualized cloud," in *Proceedings of the ACM SIGCOMM*, 2024.

[12] X. Liu, S. Zhao, Y. Cui, and X. Wang, "Figret: Fine-grained robustness-enhanced traffic engineering," in *Proceedings of the ACM SIGCOMM*, 2024.

[13] R. Maheshwari, "Tcp over ip anycast - pipe dream or reality?" https://engineering.linkedin.com/network-performance/tcp-over-ip-anycast-pipe-dream-or-reality, 2015.

[14] B. Huffaker, M. Fomenkov, and kc claffy, "Geocompare: a comparison of public and commercial geolocation databases," Tech. Rep., 2011. [Online]. Available: https://www.caida.org/publications/papers/2011/geocompare-tr/

[15] F. Chen, R. K. Sitaraman, and M. Torres, "End-user mapping: Next generation request routing for content delivery," in *Proceedings of ACM SIGCOMM*, 2015.

[16] N. Spring, R. Mahajan, and T. Anderson, "The causes of path inflation," in *Proceedings of ACM SIGCOMM*, 2003.

[17] Z. Shmeis, M. Abdullah, P. Nikolopoulos, K. Argyraki, D. Choffnes, and P. Gill, "Localizing traffic differentiation," in *Proceedings of ACM IMC*, 2023.

[18] M. Zhou, X. Zhang, S. Hao, X. Yang, J. Zheng, G. Chen, and W. Dou, "Regional IP Anycast: Deployments, Performance, and Potentials," in *Proceedings of ACM SIGCOMM*, 2023.

[19] W. De Vries, R. de Oliveira Schmidt, W. Hardaker, J. Heidemann, P.-T. de Boer, and A. Pras, "Broad and load-aware anycast mapping with verfploeter," in *Proceedings of ACM IMC*, 2017.

[20] X. Zhang, T. Sen, Z. Zhang, T. April, B. Chandrasekaran, D. Choffnes, B. M. Maggs, H. Shen, R. K. Sitaraman, and X. Yang, "AnyOpt: Predicting and Optimizing IP Anycast Performance," in *Proceedings of ACM SIGCOMM*, 2021.

[21] B. Schlinker, T. Arnold, I. Cunha, and E. Katz-Bassett, "PEERING: Virtualizing BGP at the Edge for Research," in *Proceedings of ACM CoNEXT*, 2019.

[22] C. H. Papadimitriou, "Worst-case and probabilistic analysis of a geometric location problem," *SIAM Journal on Computing*, 1981.

[23] R. N. Staff, "Ripe atlas: A global internet measurement network. internet protocol journal," 2015.

[24] D. Bhattacherjee, W. Aqeel, S. A. Jyothi, I. N. Bozkurt, W. Sentosa, M. Tirmazi, A. Aguirre, B. Chandrasekaran, P. B. Godfrey, G. Laughlin, B. Maggs, and A. Singla, "cISP: A Speed-of-Light internet service provider," in *Proceedings of USENIX NSDI*, 2022.

[25] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," 2023. [Online]. Available: https://www.gurobi.com

[26] X. Fan, J. Heidemann, and R. Govindan, "Evaluating Anycast in the Domain Name System," in *Proceedings of IEEE INFOCOM*, 2013.

[27] L. Wei and J. Heidemann, "Does anycast hang up on you?" in *Proceedings of TMA*, 2017.

[28] D. Cicalese and D. Rossi, "A longitudinal study of ip anycast," *ACM SIGCOMM Computer Communication Review*, 2018.

[29] M. Calder, A. Flavel, E. Katz-Bassett, R. Mahajan, and J. Padhye, "Analyzing the performance of an anycast cdn," in *Proceedings of ACM IMC*, 2015.

[30] F. A. Chudak and D. B. Shmoys, "Improved approximation algorithms for the uncapacitated facility location problem," *SIAM Journal on Computing*, 2003.

[31] K. Jain, M. Mahdian, E. Markakis, A. Saberi, and V. V. Vazirani, "Greedy facility location algorithms analyzed using dual fitting with factor-revealing lp," *Journal of the ACM*, 2003.

[32] K. Jain and V. V. Vazirani, "Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and lagrangian relaxation," *Journal of the ACM*, 2001.

[33] V. Arya, N. Garg, R. Khandekar, A. Meyerson, K. Munagala, and V. Pandit, "Local search heuristic for k-median and facility location problems," in *Proceedings of ACM STOC*, 2001.

[34] M. Chrobak, C. Kenyon, and N. Young, "The reverse greedy algorithm for the metric k-median problem," *Information Processing Letters*, 2006.