An agent-based model of IPv6 adoption

Ioana Livadariu∗, Ahmed Elmokashfi∗ and Amogh Dhamdhere†

∗Simula Metropolitan, 0167 Oslo, Norway
Email: {ioana,ahmed}@simula.no
† CAIDA/UCSD, La Jolla, CA 92093, U.S.A
Email: amogh@caida.org

Abstract—Despite having been proposed more than 20 years ago, IPv6 deployment has been very slow. The imminent depletion of IPv4 address space has recently motivated stakeholders to actively promote IPv6. These efforts, however, have only led to a relatively modest increase in the overall uptake. This outcome is expected given the complexity of the adoption landscape and the involved economic intricacies. Aiming to offer better insights into this process, we present the first data driven computational cost centric model of IPv6 adoption. Our model is grounded in empirical data yet parsimonious that is it focuses only on factors that are key to the modelled transition. We validate our model using historical snapshots of addresses allocations and then use it to explore a set of what if scenarios. Our findings paint a bleak picture for IPv6 adoption, predicting it to be decades away.

I. INTRODUCTION

Following its commercialisation in 1994, the Internet has evolved in unforeseen directions in terms of both usage and reach. Yet, the core part of the protocol suite, TCP/IP, that underpins the Internet has been fairly static. At the network layer, the IPv4 protocol has largely remained the only protocol used for addressing and routing [1]. IPv4 addresses are 32-bit identifiers that provide a maximum of approx. 4.3 billion addresses. However, the unexpected expansion of the Internet prompted, as early as in 1992, efforts to optimise the IPv4 address space usage [2]. These efforts eventually lead to introducing and standardising the IPv6 protocol in 1998 as a permanent alternative to IPv4 [3]. IPv6 addresses are 128-bit identifiers expanding the address space from $2^{32}$ to $2^{128}$ addresses. IPv6 also comes with other improvements like simplified header, or address auto-configuration mechanism. Thus, IPv4 to IPv6 transition was bound to face both technological and economic challenges.

Two decades after, IPv6 deployment remains scant and the question of whether IPv6 will ever take over remains elusive. There was arguably no compelling reason for deploying IPv6 in the first decade of the 2000s. IPv4 addresses were abundant, IPv6 support in end systems was largely missing and existing implementations were immature resulting in performance penalties. In 2011, IPv4 depletion became a reality when the Internet Assigned Numbers Authority (IANA) allocated the last /8s available at its disposal to regional internet registries (RIRs) [4]. Subsequently, all RIRs except the African Network Information Centre (AFRINIC) have either depleted their /8 or have started rationing it [5]–[8]. Innovative solutions have emerged to deal with this shortage that includes trading of IPv4 addresses and deployment of large scale network address translation devices known as Carrier Grade NATs (CGNs) which allow a single IPv4 address to be shared among many users. IPv6 deployment has accelerated. Today, the core of the Internet fully supports IPv6, 25% of users access Google over IPv6 [9] and almost all major content providers are dual-stacked [10], i.e., support IPv4 and IPv6. In total, there are half a billion end users with IPv6 support [11]. While these numbers are encouraging, IPv6 deployment remains slow and uneven. 80% of the current half a billion users come from India, the US and Brazil propelled by the decisions of few ISPs, mainly mobile operators, to embrace IPv6. Populous countries like China and Russia as well as the majority of the developing world are by large IPv4 countries, which hints at potential splintering and digital divide. There is currently no acceptable hypothesis that explains what has been driving IPv6 adoption post 2011. Neither the depletion of IPv4 space nor competition nor demand for addresses to meet customers growth seems to solely explain it [12].

The depletion of IPv4 address pool has created a new reality, where addresses are commodified and CGNs are breaking the Internet end-to-end principle. Furthermore, it is envisioned that the demand on addresses will rise markedly, if the expected surge in Internet of Things (IoT) materialises [13]. The way this new reality is going to influence IPv6 adoption is far from well understood. Hence, there is a need for a representative model that captures key factors of the adoption process, yet flexible enough to entertain different what-if scenarios. For instance, operators can choose to satisfy their address demand by buying addresses, deploying CGNs, becoming dual stacked, using some transition technologies or any combinations of these. Also operators’ demand depends on their business type (access vs content), growth in customer base and their geography. Furthermore, a successful model must account for various costs e.g. IP address prices, cost for running a CGN, etc.

This paper proposes a novel cost-centric agent-based model of IPv6 adoption that is representative yet parsimonious. It ensures representativeness by modeling all relevant stakeholders and by operating at the granularity that is equivalent to the way IP addresses are grouped in the global routing system. The lack of backward compatibility and killer features make the decision on opting for IPv6 largely cost driven, hence the focus of our model. Expressing these costs as well as other model parameters (e.g. how many customers a network has) can be speculative though! Thus, we follow a data driven approach that draws on existing measurements to parametrize the model. Further, we ground the costs in plausible economic
principles. We choose agent-based modeling to capture the individuality and heterogeneity of network operators. Our model is parsimonious including a mere 16 parameters. We validate the model by starting it in 2006 and 2012 and comparing its outcome in terms of deployment decisions to what was measured in reality. Subsequently, we employ the model to explore the impact of two scenarios on IPv6 adoption—an accelerating demand on IP addresses and the IPv4 transfer market impact. Our findings paint a bleak picture for IPv6 adoption. They show that currently there is no pressing need from an economic point of view for adopting IPv6. A full adoption remains an inevitable outcome, yet maybe decades away. Even a fast growth in demand for addresses will not significantly speed up the transition process.

II. RELATED WORK

Modeling the diffusion of new technologies or products has been well studied in the marketing literature. Bass proposed a model to forecast first purchase of long-term sales patterns for new products and technologies based on expected utility gain [14]. Fout and Woodlock also used utility gain to analyze diffusion of innovation on the food product market [15]. Many other studies have built on this work modeling the decision process at the user level, impact of network externalities and the role of convertors [16]–[19]. Intuitively, the diffusion of new technology depends on the gain that adopters get from deploying the new technology. However, studies have shown that adoption of a new technology is influenced by the network externalities [20], [21], the existence of a related technology infrastructure [22], the communication channels as well as general knowledge about the new technology [23], [24].

Hovav et al. [25] applied a diffusion-based model for the IPv4 to IPv6 transition and found that success of a new technology depends on its compatibility with the existing standard, but also on the external influences like regulation and investments. Joseph et al. [19] investigated the impact of convertors on the IPv6 adoption and found that this process can be influenced by the efficiency of the convertors. The work published by Sen et al. [26] converged to the same conclusion regarding the role of convertors. Nikkhah et al. [27] build a two-side market model where service providers connect their customers using different types of connectivity (public IPv4/IPv6 addresses, private IPv4 addresses), and found that co-ordination among service providers can contribute to the IPv6 adoption. In their follow-up work [28] they included an empirical quantification of the IPv6 migration, and proposed a simple model of IPv6 adoption for different stakeholders: service providers, content providers and end-users.

Our proposed modeling work differs in four ways from previous research: it is cost-focused and does not assume that deploying IPv6 affords any additional utility, it is faithful to the real-world challenges of gradual transition by organizations, it is data-driven, and we use a computational (rather than analytical) approach to solve the model and investigate what-if scenarios relating to IPv6 deployment.

III. DATASETS

World Bank data: We use World Bank’s Internet penetration rate [29] and total population per country [30] datasets to compute the number of Internet users (subscribers) per country. We further distribute these users to the access provider networks within each country. We also employ the Internet user growth rate to estimate the Internet user growth within each country.

RIR delegation files: The RIRs maintain and publish files that contain details of the allocated and assigned Internet resources, i.e., IP address blocks and Autonomous System Numbers (ASNs) [31]. Using these information we identify and count the networks (ASes) registered within each country.

BGP data: We employ routing data from Routeviews [32] and RIPE [33] to count the number of advertised IPv4 and IPv6 addresses. We employ the number of routed IP blocks to initialize the IP addresses used by the networks in our model.

AS classification dataset: We use CAIDA’s AS classification classification scheme that divides ASes into “Enterprise” networks, ”Content” and ”Transit/Access” providers [34] to initialize the business type of the networks in our model.

List of transferred address blocks: Trading of IPv4 addresses blocks has emerged as a mechanism for prolonging the usability of IPv4 address space. RIRs periodically publish the approved intra-RIR and inter-RIR IP transactions. We leverage these lists in our model validation described in Section V.

IV. MODEL PARAMETERIZATION AND EXECUTION

We employ agent-based modeling an approach that is used for simulating complex systems comprised of numerous individual components, for modeling IPv6 adoption process. Accordingly, we model each network as a selfish agent that seeks to minimize the overall cost associated with its current IP addresses assignment configuration. The model executes iteratively, running in rounds until it reaches an equilibrium. Our model is cost-focused and the core component of our model is the network’s decision process. Building a computational model of the IPv6 adoption decision process faces at least four modeling challenges: faithfully representing players (agents) and their range of possible actions, capturing the costs of operating IPv4 and/or IPv6 infrastructure, formalizing and efficiently executing the decision process among agents and identifying an equilibrium. We further describe the main components of the model.

Model Execution. We use the information extracted from the four datasets above to initialize the model parameters (see Sec III). Figure 1 shows an overview of the model. In each iteration, we randomly select a network, and execute its decision process (shown as grey before completing the process). As outcome of this process, the network can optimize or not optimize its overall cost (shown with green and red after completing the process). This continues until all nodes are visited. At the end of an iteration, we can be left with nodes that managed to reduce their cost and those that did not. We then compare the fraction of networks that have reduced their overall cost with a chosen threshold. If the fraction is smaller than the threshold, the model execution continues to the next iteration. Otherwise, we consider that the system has reached equilibrium.

Active Players. Several stakeholders contribute to determining the fate of IPv6 adoption, i.e., content and access (including mobile) providers, transit providers, regulators, end users and end systems developer. Most transit providers have already deployed IPv6 [35], while all major OSes have long supported IPv6. Hence, the content and access providers are the active...
(decision-making) players in our model. RIRs and end users are passive players. We rely on CAIDA’s AS classification dataset [34] to identify access or content provider networks. Our analysis shows that only a third of access provider were deploying IPv6 at the end of 2017. We further detail how we model the number of customers each access and content provider serves.

Access Providers (AP): APs offer connectivity to end-users and earn revenue from their customers. Several sources provide aggregated values for the number of Internet Users, for e.g., [11], [29]. However, individual APs do not usually publish these numbers. To estimate the number of customers an AP serves, we use the Internet penetration rate and the overall population per country from the World Bank records [29], [30] to obtain the number of Internet users per country. We then identify which ASes operate in each country by looking at the RIR delegation files. The next step is to assign users to the APs that operate in a particular country. To this end, we need to pick a suitable probability distribution. The available statistics indicate that the market within a country is dominated by a few internet service providers (ISPs) [36]. Thus, we model the distribution of users per AP as log-normally distributed. This is a reasonable assumption since the log normal distribution is known to arise from multiplicative processes, i.e., the number of a customers an AP serves is a percentage of its current customer base [37]. Given the lack of empirical data, we use the standard log normal distribution, i.e., the location and scale parameters are set to 0 and 1, respectively. Finally, we use BGP routing data to map users to APs in each country. Specifically, we assume that the AP with the largest user base is the one that advertises the largest IPv4 space and continue ascendingly. The number of users per AP then grows at rate $g$ as the model progresses. This rate is also extracted from the World Bank projections for Internet penetration growth [29].

Validating users to APs mapping. Figure 3 shows our distribution of Internet users per RIR for 2012, while figure 2 shows the same number in 2016 when using data from by APNIC Labs [38]. The latter data covers 13183 small and large size access providers networks across the globe. The modeled distribution closely matches APNIC Labs’ measurements both quantitatively and qualitatively (see the order of RIRs).

Content Provider (CP): CPs revenue is a function of the number of Internet users that access their content which in turn reflects the popularity of the content providers. Several studies indicate that content popularity follows a skewed distribution [39], [40]. We therefore rely on the Zipf distribution [41] to assign popularity to CPs. To capture the presence of a few main CPs we set the distribution shape parameter to 2. To account for the fact some CPs offer content of global relevance while others have a more local appeal, we assume that a low fraction of CPs are globally accessible, e.g., Google, Netflix, YouTube and Amazon [42], [43]. Accordingly, we consider 10% of the CPs to be globally accessed. The remaining CPs offer content to users within their local region. To capture how Internet users access CPs, we assume that only part of the AP customers access the content offered by a particular CP. To this end, we randomly assign fractions of AP’s subscribers that access the content offered by CPs. Let $n_a$ be the number of APs. Also let $U_i$ be the total number of customers that AP $i$ serves and $U_j$ the number of end-users that access the content offered by CP $j$. If $f_{ij}$ the fraction of AP $i$’s end-users that access content offered by CP $j$, we can express:

$$U_j = \sum_{i=1}^{n_a} f_{ij} \times U_i, \quad f_{ij} = 0, 1$$

Address Space. We use from BGP data [32], [33] to determine the number of addresses used by the networks in our model. To this end, we extract the IPv4 and IPv6 address blocks advertised by APs and CPs. Next, to avoid double counting IPs, we use Cittadini et al. [44] classification technique to filter out the provider agreeable (PA) space from the selected IP blocks. Note that the classification algorithm is described for the routed IPv4 address blocks. We apply the same methodology for the IPv6 address blocks. Thus, we retain IPv4/IPv6 blocks that are not covered by any other IP blocks to initialize the number of IPv4 and IPv6 addresses for the networks in our model.

Networks can acquire IP addresses during each iteration from their local RIR, which in turn acquires IP blocks from IANA. The distribution of address blocks is regulated though the RIRs’ allocation policies). Moreover, networks can sell (buy) IPv4 addresses from (to) other networks given that the RIRs of the seller and buyer networks have implemented transfer policies. Our model captures the allocation and transfer policies along with the main restrictions imposed by these regulations – limiting the number of IP blocks that networks can acquire and regional IPv4 transfer markets.

Possible actions. APs can assign to customers using different types of IP addresses - public IPv6 and/or public/private IPv4.
addresses. We model five device classes (address buckets) that capture the addresses assigned to the subscribers:

1) IPv4-Only (IPv4): only IPv4 public addresses;
2) Dual-Stack (DS): IPv4 and IPv6 public addresses;
3) IPv6-Only (IPv6): only IPv6 public addresses;
4) Carrier Grade NAT (CGN): private and public IPv4 addresses;
5) Dual-Stack with private IPv4 addresses (CGN-DS): IPv4 and IPv6 public address, and private IPv4 addresses;

CPs can choose whether they would like to offer content over IPv4 and/or IPv6. For such networks, we consider only the first three device classes from the above list.

Cost component. We model the costs networks incur by deploying IPv4, IPv6 and/or satisfying their addresses space needs with IPv4-based technologies. Based on these costs and their numbering configuration networks incur an overall cost that we express as a percentage of their revenue.

Revenue: APs use a subscription-based model to offer connectivity to their users. We compute the profit from each customer as a percentage of the subscription prices and refer to this value as the profit margin. Instead of assigning a particular subscription price for each AP, we use a uniform price. While this is a simplification, broadband prices around the world generally vary in a limited range within one order of magnitude [45]. We set the price to 10$ per month, matching the lower end of the market today and assuming that prices in the rest of the world will continue to decrease. CPs use an indirect revenue model. Internet end-users consume freely content and this content generates revenue through strategies like selling advertising space or collecting and using user information. We assume that CPs earn from each user that access their content a profit that directly depends on the CP popularity. For each CP we model this profit as the inverse of the network’s popularity.

Cost of acquiring IPv4 addresses: Networks can satisfy their address space needs by acquiring IPv4 addresses on the IPv4 transfer markets. However, the monetary details of IP transaction are not public. The closest estimation of IPv4 address prices is offered by a few IPv4 brokers that publish average prices values for the IPv4 transferred blocks, e.g. [46]. An analysis of these values indicates that IPv4 prices remain within the same range of values over time. Specifically, these values vary between $11 and $18 per IP. In absence of a particular model of IPv4 price growth, we model the price as increasing with a factor $x$ between iterations. In our analysis, we use a linear price growth.

IPv4/IPv6 capital overhead: IPv4 (IPv6) deployment costs are modeled as the costs associated with the capital overhead (e.g., hardware, support) for deploying IPv4 (IPv6). We use the economics of scale concept [47] to capture these costs, i.e., the average cost of production decreases as the production volume increases. When the number of networks deploying IPv4 (IPv6) increases, the cost associated with the deployment decreases. Let $n^a(4)$ and $n^c(4)$ be the number of APs and CPs that deploy IPv4 (IPv6). Let $n^a$ and $n^c$ be the number of APs and CPs, respectively. The IPv4 (IPv6) deployment cost is:

$$C^a_{4(6)} = \left(1 - \frac{n^a(4) + n^c(4)}{n^a + n^c}\right) \times 100 \quad (%)$$

In our simulations, all the networks in our model deploy IPv4 and only a fraction of these networks deploy IPv6. Based on Eq. 2, IPv4 deployment cost ($C^a_{4}$) is set to zero, while the cost of deploying IPv6 ($C^a_{6}$) is high.

CGN capital overhead and breakage cost: CGN deployment costs comprise the costs of purchasing and operating CGN devices. For such networks, we also model a breakage probability that captures the poor performance and application breakage when accessing certain services behind the CGN equipment. Recent studies [48], [49] report that a significant percentage of networks deploy CGN. Thus, we consider the capex to be close to zero. However, we model the opex and breakage cost to rapidly increase as the number of end-users behind the CGN increases, i.e., the CGN compression factor ($c$). To this end, we use the quadratic function to capture the loss incurred by the networks deploying CGN: $c^2$. We choose this function form since it captures an increasing rate of change that will eventually plateau, i.e., stacking more users behind a single public IP will quickly increase breakage probability, which will plateau once we hit the maximum possible number of users. Note that the number of users behind a CGN has a lower (an upper) limit: $c_{min}(c_{max})$. The maximum value is determined by the number of available ephemeral ports [50]. We set this number to a max of 1000, since many ports are already reserved. We set the lower value to 10, which is the NAT compression factor reported by Grover et al. [51]. We express the breakage cost as a percentage of the revenue an AP receives from a user as follows.

$$C_{br}^{o/br} (c) = \frac{e^2 - e_{min}^2}{e_{max}^2 - e_{min}^2} \times 100 \quad (%)$$

IPv4-IP6 translation capital overhead and breakage cost: Customers with IPv6 addresses that access content over IPv4 require a translation mechanism. We model the translation cost as capital overhead and breakage cost ($C^{o/br}_{tr} (c)$). To determine these costs, we rely on existing measurements studies [52], [53] that report failure rates for transition technologies. Published in 2010, both studies reported high values for these rates – between 9% and 20% by Aben [52], and 13% by Huston [53]. Five years later, Huston’s study reported a small improvement in the failure rates – between 8% and 10%. We set this cost to 8% for the scenarios described in section VI.

Total cost: We compute for each network the capital overhead and breakage probability costs associated to each of their addresses buckets. The network’s overall cost is the sum of the address bucket costs. The capital overhead of an address bucket is given by the IP address type assigned to the bucket. We compute each bucket cost as a percentage of the profit value per user that a network gains from an end-user assigned to that bucket. Hence, the overall cost is the percentage lost of the network’s revenue. Let $d_{a(c)}$ be the number of device classes for the access (content) providers in our model, i.e., $d_{a(c)} = 5$ (3).

The overall cost ($T$) is given by:

$$T = \sum_{i=1}^{d_{a(c)}} C^o_i + \sum_{i=1}^{d_{a(c)}} C^{br}_i$$

For AP $x$, the overall cost ($T_x$) depends on the number of users per address bucket ($n^x$), the capital overhead and breakage probability of each bucket. The total cost for any CP $y$ depends on whether the network offers content over IPv4 and/or IPv6 ($a^y$), and any breakage costs due to AP subscribers accessing the offered content. We express the overall cost as follows:

$$T_x = \sum_{i=1}^{d_{a(c)}} u^x_i \times \left( C^o_i \sum_{j=1}^{n_x} f_{xj} \times \left( \sum_{i=1}^{d_{a(c)}} C^{br}_i \right) \right)$$

364
\[ T_y = \sum_{j=1}^{d_c} C_j \times a_j^y + \sum_{i=1}^{d_a} C_i^{br} \times \left( \sum_{k=1}^{n_a} f_{ky} \times U_k \right) \]  

**Decision process.** During each iteration of our model, both APs and CPs go through their own decision process. This process is comprised of two sequential steps: *addition* and *optimization*. In the first step, the network updates the number of its subscribers, while in the second step the network seeks to find an address space configuration that optimizes its current overall cost, i.e., reduces the cost with at most \( \theta \). We discuss in section V the model convergence for different values of \( \theta \). Hence, the network can alter its number of subscribers and IP addresses, as well as the incurred costs. We next present the decision process for APs and CPs: let \( U \) and \( A_{4/6} \) be the overall number of subscribers and IP addresses, and \( u \) and \( a_{4/6} \) be the number of subscribers and IP addresses per bucket and \( T \) be the overall cost. \( u_{4/6} \) is the number of IP addresses that the network acquires from the RIR/IPv4 transfer market.

**AP decision process:** APs assign end-users to different device classes to minimize the total cost of the address space. The steps of the AP decision process are listed in Algorithm 1. A provider first updates its number of overall subscribers, and distributes the new users its buckets. Next, the AP network gradually moves \( x \) subscribers from the highest-cost buckets to the lowest-cost buckets until it reaches a deployment that minimizes its overall cost.

**Algorithm 1 Access provider decision process**

**Input:** \( T \): initial configuration \((U, A_{4}, A_{6}, T)\)  
**Output:** \( O \): output configuration \((U', A'_{4}, A'_{6}, T')\)  
1. \( U' \leftarrow U + g \times U \)  
2. \( A'_{4} \leftarrow A_{4} + w_{4}^{a} + w_{6}^{a}, A'_{6} \leftarrow A_{6} + w_{6}^{a} \)  
3. \( s \leftarrow \text{sort ascending cost buckets} \)  
4. \( i_{\text{max}}, i_{\text{min}} \leftarrow \text{highest-cost, lowest-cost bucket index} \)  
5. while \( i_{\text{min}} < i_{\text{max}} \) do  
6. while \( u[4][i_{\text{max}}] \geq x \) do  
7. \( u[4][i_{\text{max}}] \leftarrow u[4][i_{\text{max}}] - x, \)  
8. \( u[4][i_{\text{min}}] \leftarrow u[4][i_{\text{min}}] + x \)  
9. \( a_{4}, a_{6} \leftarrow \text{update number of addresses per bucket} \)  
10. \( A_{4} \leftarrow \text{sum}(a_{4}), A_{6} \leftarrow \text{sum}(a_{6}) \)  
11. if \((T \geq T') \text{ or } (A_{4} > A'_{4}) \text{ or } (A_{6} > A'_{6})\) then  
12. \( u[4][i_{\text{max}}] \leftarrow u[4][i_{\text{max}}] + x \)  
13. \( u[4][i_{\text{min}}] \leftarrow u[4][i_{\text{min}}] - x \)  
14. \( a_{4}, a_{6} \leftarrow \text{update number of addresses per bucket} \)  
15. \( i_{\text{min}} \leftarrow i_{\text{min}} + 1 \)  
16. break  
17. end if  
18. end while  
19. \( i_{\text{max}} \leftarrow i_{\text{max}} - 1 \)  
20. end while

**CP decision process:** CPs have to decide whether to offer their content over IPv4 and/or IPv6 based on their capital overhead and breakage costs. During each iteration, each CP updates the number of users that access its content. Based on its address space needs, the network then tries to acquire IP addresses from its local RIR and/or from the transfer market. Finally, the CP computes the overall costs for possible configurations and selects the lowest-cost configuration.

**Output.** Our model can allow us to understand the network deployment over time for access and content providers. Altering different parameters can provide insights into whether networks are going to adopt IPv6 as well as the adoption rate.

V. **Model Validation**

In absence of detailed information on the economics of the transition technologies, the impact of social factors and the incentives for participating in the IPv4 transfer market, validating our model should be considered *best-effort*. We evaluate the model *correctness, convergence and robustness*. **Model correctness.** We investigate whether the model’s predictions align with trends extracted from empirical data. To this end, we run simulations with the staring points in 2006 and 2012, i.e., we parametrize the model according to the number of APs and CPs and their corresponding address configuration. **Starting point 2006.** We initialize our model with 7583 APs and 914 CPs. At that time IPv4 address space depletion was not regarded as an immediate issue and RIRs did not implement strict allocation policies and transfer policies. Hence, in our simulation we do not restrict the number of IP addresses that networks can acquire from their local registry and we do not activate the IPv4 transfer market. For the CGN deployment, we rely on the value reported by Grover et al. [51] to initialize the compression factor to 10. IPv4-only is the lowest-cost bucket for APs. Since we do not impose any allocation thresholds, most networks acquire in the first iteration the number of IP blocks that allows them to move their subscribers to the lowest-cost bucket. The simulated networks manage to minimize their overall cost within one iteration which correspond to one year (i.e., 2007) and the model reaches equilibrium. Our model estimates that in 2007 94% of the networks rely on IPv4-only deployment and only 5.03% of these networks deploy IPv6. Half of these networks come from RIPE, 26.5% from APNIC and 16.3% from ARIN. From 2006 to 2007 we estimate that IPv6 deployment increases with 0.1%. Analyzing the routing data for 2007 indicates a similar IPv4/IPv6 deployment; specifically we find that 94.42% of the providers deploy IPv4. 5.57% of the networks were deploying IPv6. Breaking down this number per RIR we find that 45% of these networks come from RIPE, 25% from APNIC and 22% from ARIN. **Starting point 2012.** Following the IPv4 address space depletion [4], RIRs imposed restrictions on the IPv4 allocations. During the same period, networks from three RIRs were already acquiring IP blocks on the IPv4 transfer markets. Moreover, studies indicate that APs were already deploying CGN to prolong the usage of their IPv4 address space [48], [54]. To capture these conditions we activate the IPv4 transfer markets and initialize the CGN compression factor with 100, the value reported in CAIDA’s IPv6 survey [55]. We initialize the number of APs and CPs with 15095 and 1714, respectively.

We evaluate the model’s predicted CGN and IPv6 deployment evolution during the first six iterations of our simulations which correspond to the period of time from 2013 to 2018. The black and blue line in fig. 4 show the CGN and IPv6 deployment, respectively. APs appear to deploy CGN. Moreover, CGN deployment is increasing over time. More than 97% of these networks use deploy CGN in mixed configurations. Half of the networks deploying CGN come from RIPE, 22% from ARIN and 18% from APNIC.

We compare our findings with results published in a recent study on CGN deployment in the wild [49]. The authors found 4100 ASes deploying CGN in the period from July 2014 to July 2016. We cross-check our model predictions only for the results inferred in the second half of their measurement period.
(from July 2015 to July 2016), since their raw data significant increased during this period. Similar to our predictions, for the considered period of time, the number of CGN inferred networks is increasing linearly and approx. 66% of the networks come from RIPE. We also predict that the number of networks that deploy IPv6 to increase linearly. Analyzing the number of network that advertise IPv6 prefixes in the global routing system reveals also a linear trend.

We activate both the inter-RIR and intra-RIR IPv4 transfer markets across all regions allowing networks to exchange IPv4 blocks with any other network on the markets. As of 2012, inter-RIR and intra-RIR were implemented only by two and three RIRs, respectively. These policies varied across different regions; for e.g., APNIC policies limited the number of IPv4 blocks network exchanged in the IP transactions whereas RIPE policies did not limit the number of such IP blocks. Taking this into account, we expect to observe difference between the reported data and our model outcome. Most buyers in our model come from RIPE, followed by ARIN and APNIC. We use the lists of transfers, that are reported by RIRs, to extract the number of buyers [56]–[59]. Starting from 2013, the highest fraction of buyers was coming from RIPE. From 2015 on, ARIN comes second and APNIC comes third. Summary. Starting our model at two points in time that correspond to different phases in the run up to IPv4 depletion show that our model largely capture the major trends. Note that the model also reproduces emerging trends like identifying the RIR with the highest CGN deployment and that with the largest number of IPv4 buyers.

Model convergence. The termination condition of our simulations depends on the networks minimization cost threshold (θ) and the number of networks that reduce their overall cost at the end of each iteration. Hence, we investigate the model converge by varying θ from 0 to 0.1% and computing the percentage of networks that reduce their overall costs at the end of each round. Figure 5 shows this percentage when starting our simulations from 2006. For θ ≥ 0.00001%, all simulated networks reduce their overall cost. Only 89% of the networks reduce their costs when θ = 0. At least 90% of the APs manage to completely reduce their address space cost as these network incur loss as long as Internet users located behind CGN boxes access their content. Our analysis reveals that the system appears to stabilize in the last third of iterations across all simulations. Specifically, the networks distribute their users, on average, in a similar manner regardless of the cost threshold. When starting our model with the 2012 configuration we find that 90% of the networks reduce their address space cost within 83 iterations of the model. Similar to the previous simulations, the model stabilizes in the last part of the simulation, i.e., the percentage of networks that reduce their cost increases with

Fig. 5. Model convergence for different cost thresholds (θ).

2%. Running our model over 100 iterations is problematic due to double floating-point round-off error that may occur in operating large numbers in C++. However, given that our model stabilizes long before reaching this stage, we believe that relaxing our two convergence constrains does not tangibly affect the simulations outcome. Hence, for the scenarios in sec. VI we set θ = 0.001% and stop the simulations when at least 90% of the networks reduce their overall cost.

Model robustness. In our model, we randomize two parameters – the order in which networks play during each iteration and the fractions of AP subscribers that access the content offered by CPs. We evaluate how robust is the model to the randomness due to these parameters by running the second scenario above, i.e., starting the simulation with 2012 configuration, 30 times and analyzing whether these converge to the same network deployment. Our model takes 4.45 hours on average to converge and reaches equilibrium in approx. 80 iterations. Figure 6 shows the range for the percentage of APs that assign their end-user to different address buckets at equilibrium. These results indicate a low variability across the deployments. Most networks distributed their users across the IPv4, IPv6 and DS buckets. For this address bucket combination we register the highest variability, i.e., inter-quartile range is 1.37% and the range is 11%.

Fig. 6. Variability of the percentage of APs at equilibrium that distribute their subscribers to different address buckets.

VI. WHAT-IF SCENARIOS

Having shown that our model is able, qualitatively and to some extent quantitatively, to capture the process of IPv6 adoption, we further investigate the role of possible drivers and inhibitors of this process. Thus, we investigate whether an explosion in demand for addresses via a step rise in the number of internet of things devices and the IPv4 transfer markets could potentially make IPv6 a viable alternative cost-wise. We set the simulations to start in 2015 and 2012.

A. Rise of the Internet of Things (IoT)

The Internet has undergone an unforeseen growth. Over the years, the overall number of Internet users has grown exponentially - from 2000 to 2018, this number has increased with approx. 970%. Moreover, half of the world population are already Internet users [11]. Available statistics report 17.68 billion connected devices in 2015, which account for an approx. 15% increase than that the previous year [60].

Global uptake. We use our model to investigate how the increase in the number of connected devices impacts IPv6
deployment. To this end, we increase the number of devices per subscriber \((D)\) globally. To better understand the model outcome, we disable the IPv4 transfer market for this scenario. **Impact on adoption.** Figure 7 shows the average fraction of IPv6 deploying networks over time for different values of \(D\). As expected a large demand for addresses will accelerate IPv6 adoption. The IPv6 adoption rate directly depends on \(D\). For \(D < 200\) we observe a clear difference in the number of rounds necessary to deploy IPv6, whereas for \(D \geq 200\) this number remains at approx. the same value. Under all growth rates, networks respond initially to the increasing demand by deploying CGN. Surprisingly, even at a sustained 400% growth, a full adoption of IPv6 is predicted to take a decade. Currently the IoT market growth is estimated between 25% to 30% [61], which appears too slow for driving IPv6 adoption. **Cost evolution.** IPv6 deployment rate appears to follow the S-curve across the different growth values; slow adoption for the first 10% of the Internet users, followed by a rapid increase for approximately half of the users and then again a slow increase in the deployment rate for the remaining users. Examining the evolution of APs address assignments in this scenario. Initially, APs satisfies their address demand using existing IPv4 addresses and CGNs. Then, a small fraction of APs will take up IPv6 as the only assignment option once the CGN cost becomes prohibitive. This corresponds to the slow growth phase in the S-curve. Finally, once a small fraction of networks, around 10%, chooses IPv6 only, the cost of IPv6 drops below the CGN cost ushering the start of the exponential growth phase. This indicates that having a sustainable yet small fraction of networks that offer IPv6 only connectivity is crucial for propelling the adoption process. A necessary condition for this is that most content should be available on IPv6. **Differences between RIRs and networks.** The explosion in number of devices results in a massive adoption in the three largest RIRs in contrast with the two smallest, i.e., LACNIC and AFRNIC. All ASes are predicted to deploy IPv6, however the final address assignment is going to be mixed with some users remaining behind IPv4\(^1\). For the three largest RIRs, the ratio between the average number of users with IPv6 to those with another configuration that is not IPv6 is \(\approx 10^{12}\). The same ratio is \(\approx 10^4\) for AFRNIC and LACNIC. This correspond to \(\approx 10^{10}\) users without IPv6 in each RIR. Figure 8 shows this spread of this ratio per network over time when the IoT deployment growth is 100% across all regions. Note that the absolute numbers are not the key goal of our analysis and are likely to be off by some margin since we are not controlling for population growth and technological innovations. Nevertheless, these numbers indicate that a need driven uptake of IPv6 will likely result in clear differences between industrialized countries and developed countries. **Regional uptake.** We have also experimented with a regional increase in the demand of addresses instead of the global case above, i.e., the demand is limited to a single RIR. This is more plausible given today's digital divide and differences in the Internet penetration rates [29]. The lower panel in Figure 7 shows the overall fraction of APs deploying IPv6 following a massive increase in IoT devices within just one region, i.e., APNIC, ARIN and RIPE. Neither of these alternatives would propel IPv6 deployment beyond 40% of all APs. The numbers of adopting APs in each case largely correspond to the number of APs coming from the region with a massive increase in demands for addresses. This is mainly because of the strict regional nature of the management of Internet resources, which limits the scope of changes in one region to itself. While the order that RIRs bring have proved useful in terms of a better manageability of resources, they counterintuitively seem to stifle IPv6 deployment. RIRs keep demands for IPs local thus preventing it from affecting networks in other regions. **B. Mature IPv4 Transfer markets** IPv4 transfer markets have emerged as a mechanism for prolonging the IPv4 address space usability. In the last few years, the market size has increased significantly [56]–[59]. Recent work reported that in the last years the number of IPv4 addresses exchanged on the transfer markets is comparable with the number of IPv4 allocated addresses from the RIRs [62]. Given these observations, we study whether the IPv4 transfer markets can substitute the RIR pool of IPv4 available addresses and the market impact on the IPv6 adoption. To this end, we propose a scenario in which the IP transfer markets are the only source of IPv4 addresses, i.e., networks can obtain IPv4 addresses only from the IPv4 transfer markets. We activate both the inter-RIR and intra-RIR IPv4 transfer markets across all regions allowing thus networks to exchange IPv4 blocks with any other network on the markets. Moreover, we run our simulation when networks can exchange unlimited and limited number of IP addresses during one transaction on the market; we further refer to these two cases as UINTER-UINTRA and LINTER-LINTRA. We allow the immediate return of the networks on the market and model the IPv4 address prices to follow a linear increase. We estimate price growth value from reports published by IPv4 brokers, i.e., 0.02, 0.025 and 0.015 for APNIC, ARIN and RIPE, respectively [46]. Given these small values and the lack of available data for AFRINIC and LACNIC, we choose to keep the same price

---

\(^1\)Assuming that all content is dual-stacked. A deviation from that may result in a complete eradication of IPv4.
for the two RIRs. We show in figure 9 the evolution of the five addresses buckets. Initially, the cheapest solution is to assign user to IPv4 or CGN buckets. However, after approx. 20 iterations assigning users to the IPv6 bucket becomes cheaper.

**IPv4 transfer market transient role.** The evolution of the APs deployment follows the same pattern for both configurations. We plot in figure 10 the fraction of users connected over public IPv4 and IPv6 addresses, and private IPv4 addresses. Initially, networks distribute the majority of their end-users to IPv4-enabled buckets by assigning them either IPv4 public addresses or by placing them behind CGN boxes. During this period of time assigning users to these two buckets is cheaper than any other bucket. After 20 iterations, due to the growth in the connected users, none of these two solutions continues to be viable and networks start connecting their subscribers over IPv6. With more networks deploying IPv6, the cost of this bucket decreases and ultimately becomes cheaper to assign end-users to the IPv6 bucket than to the CGN one.

**IPv6 uptake.** Our simulations reveal that the market size is decreasing over time, indicating that the markets do not satisfy the demand for IP addresses as the number of users connected to the Internet grows. Most of space on the market is acquired by networks in RIPE, ARIN and APNIC. IPv4 transfers markets do not appear to substitute the adoption of IPv6. Depending on the implemented transfer policies, the markets appear to have a short-term mild impact on the distribution of subscribers to the IPv4-enabled buckets. Connecting users over IPv6 becomes a viable solution regardless of the implemented transfer policies. Similar to the other two scenarios, the main driver of the IPv6 uptake is the user growth.

**VII. Discussion**

**Limitations.** Modeling a complex transition like the IPv6 adoption requires employing several assumptions. We keep our model static, i.e., the number of networks remain fixed. This decision is motivated by the fact that the numbers APs and CPs grow at a relatively slow pace as well as the absence of a reasonable model that captures the evolution of the Internet from an economic point of view. Our model is cost driven and thus we are not capturing effects like peer pressure and voluntary uptake. This should have a little impact on our findings since both effects thus appear to have limited impact on IPv6 adoption [12]. In absence of empirical data on CGN cost, we have modeled it as quadratic. We believe this choice is on the conservative end of possibilities. Note that mobile operators are known to use CGN, which seems to function reasonably well. Picking a slower growing cost, e.g. linear, will only delay the predicted adoption rate, which will not affect our results qualitatively. Furthermore, one can argue that we make several assumptions, that are not empirically grounded, about the IPv4 transfer market with respect to who sells, how much an IP cost and how the does cost vary as a function of various factors. While this can be partially true, we have used the best available data. We can draw some insights from the IoT scenario with low/modest increases (e.g. 20% or 50%) in the number of devices, (see Sec. VI-A), where the market is not existent and there is demand for addresses. The IoT case shows that a full adoption is decades away but faster than the case where we have active markets. Hence, a conservative more expensive market will lie some where in between.

**Results and implications.** Our results paint a bleak picture of the IPv6 adoption process. Barring a very fast rising demand for addresses, a full IPv6 adoption is likely several decades away. This is because the cost of using CGN is predicted to remain low in the foreseeable future. Furthermore, the needed demand rate should be several multiples of today’s demand. We hypothesise that the anticipated IoT explosion is likely not to speed up the adoption process since many of these devices will be operated in a client-server fashion (i.e. they initiate communications) and can thus function very well behind CGNs. The deregulated nature of the Internet access and content markets will make this an intricate endeavour. As expected, content availability over IPv6 seems to be the only key requirement for the adoption process to start. Another finding is that the uptake will quickly accelerate once a small fraction of access providers become IPv6-only speakers. Hence, the development in this fraction can be considered as a relevant metric for gauging the uptake process. Different efforts for tracking IPv6 adoption have recently shown that we are still lacking a clear idea on what drives IPv6 adoption [12], [63]. Furthermore, after several years of growth, IPv6 adoption appears to have slowed down in 2018. These observations agree with our findings, which indicate that currently there is no pressing need, cost-wise, for deploying IPv6. The current progress seem to belong in the initial adoption phase, where networks remain unable to depend only on IPv6.

**VIII. Conclusion**

We have presented a computational agent-based model that can be used to study the implication of different factors on the process of IPv6 adoption. This model is cost-centric, empirically grounded and represents networks as selfish agents that continuously attempt to reduce the cost associated with their numbering configurations. We have leveraged this model to explore the impact of massive demand of addresses, regulatory intervention and IP addresses trading on the adoption process. Our findings paint a bleak picture showing that, in absence of extraordinary rise in demands for addresses, IPv6 adoption is several decades away.

**REFERENCES**


