Automated allocation and configuration of dual stack IP networks

Wilfried Daniels, Bart Vanbrabant, Danny Hughes, Wouter Joosen

iMinds-DistriNet, KU Leuven, Heverlee, B-3001, Belgium
Email: {firstname.lastname}@cs.kuleuven.be

Abstract—The manual configuration and management of a modern network infrastructure is an increasingly complex task. This complexity is caused by factors including heterogeneity, a high degree of change and dependencies between configuration parameters. Due to increasing complexity, manual configuration has become time consuming and error prone. This paper proposes an automatic configuration tool for dual stack IP networks that addresses these issues by using high level abstractions to model the network topology and key parameters. From this high level configuration model, low level configuration files can be generated and deployed. A key parameter specified in the high level model is the network prefix of the entire network. When translating a configuration model to configuration files, IPv4 and IPv6 subnets need to be allocated. We provide an allocation algorithm to do this allocation in the most efficient way. Evaluation of our approach shows that there is a significant increase in efficiency compared to manual configuration.

I. INTRODUCTION

The manual configuration of modern network infrastructures is a very time consuming task. To make matters worse, human mistakes occur frequently. Studies have shown that up to 36% of the errors that cause downtime for ISPs are operator errors [1]. In more than 50% of cases, operator errors took the form of configuration errors. Finding and correcting configuration errors takes a lot of time and resources. In fact, 75% of Time To Repair (TTR) hours are spent searching for and correcting operator configuration errors. The resulting outages can be very costly [2], [3]. In summary, the manual configuration of networks is inefficient due to the complexity of the configuration problem. This complexity arises from a number of sources:

The first source of complexity is dependencies between configuration parameters. For example, the configuration of a new primary DNS server requires a chain of reconfigurations as the DHCP server has to be reconfigured to supply hosts in the network with the IP of this new DNS server and the secondary DNS servers will have to be reconfigured to do zone transfers from the new DNS server. The failure to adjust one of these parameters can result in an interruption of the network service.

A second point is the increasing heterogeneity of networks. Each device or operating system has its own configuration system and syntax. The network administrators have to become familiar with all of these different systems to configure the entire network. Learning to use these systems is time consuming, and the risk of configuration errors for administrators working with new systems is significantly higher.

Finally, networks are becoming very dynamic. The topology and requirements frequently change, causing reconfiguration. These reconfigurations are difficult to support manually. The effort to manually reconfigure a network is approximately linear to the size of the network. For this reason, manual reconfiguration is very inefficient for large networks with a high degree of change.

This paper addresses the aforementioned problems by using a configuration model to specify high level configuration parameters. This model-based approach is implemented in a tool that focuses on configuring and managing a dual stack network infrastructure. In the high level configuration model abstractions of the network components and relations between them are specified. This way dependencies between low level parameters are resolved automatically, and a uniform interface is offered that encapsulates heterogeneity. The proposed model provides a uniform centralized interface to the configuration of the network infrastructure, which increases the efficiency of configuration and management.

The main contributions of this paper are twofold. The first contribution is a model-based approach to configuring a dual stack network infrastructure. By introducing a model containing a high level configuration specification, the complexity of the configuration problem is reduced. The second contribution is the design and implementation of a prefix allocation algorithm. Using this algorithm, a single prefix specified in the high level model is split and allocated to all the subnets in an optimal way.

The remainder of this paper is structured as follows. In section II we identify the requirements of the proposed tool. Section III further explains the design of both the tool and the configuration model. Section IV describes the prefix allocation algorithm. In section V we evaluate the proposed tool. Section VI discusses related work. Finally, we conclude in section VII.

II. REQUIREMENTS

The proposed automatic configuration tool has to satisfy the following requirements:

- Efficiency: The main aim of the tool is to improve the efficiency of configuration by decreasing complexity. This way configuration can happen with less effort and
the likelihood of configuration errors is lower. The tool achieves this by using a high level configuration model to specify the configuration.

- **Support for dual stack IP networks**: The IPv4 address space is depleting, despite measures taken in the 90’s like CIDR [4] and NAT [5]. The demand for IPv4 address space has only increased the last few years, with the advent of internet enabled mobile devices like smartphones and the many new internet users in developing countries. The only worldwide accepted solution for this problem is the migration to IPv6. A first step in this migration is running networks in a dual stack configuration. Special attention is paid to make sure dual stack networks are supported by the tool as well.

- **Support for critical configuration aspects**: The aim of the proposed tool is to set-up basic network functionality in a network infrastructure, facilitating deployment of more advanced services like mail servers or remote file systems. A minimal set of services is selected which are necessary for a functional network, including configuration of the following aspects:

  - Host configuration protocol: The configuration of DHCPv4 [6], DHCPv6 [7] or SLAAC [8] in all the subnets of the network.
  - DNS Server: The configuration of DNS servers in a master-slave configuration, including the generation of forward and reverse zones containing all hosts.
  - Generic servers: A hosts in a subnet can be designated as a server, which will be automatically configured with a static IP and are included in the DNS records. After this basic configuration other services can be configured on top of these generic servers.

III. DESIGN

In order to meet the proposed requirements, several preliminary design decisions were made.

A. IMP framework

The proposed configuration tool is built on the IMP framework [13]. The Integrated Management Platform is a generic framework for system configuration tools that manage an entire infrastructure. In most cases, the configuration parameters of different network components are closely related. These relations are mostly determined by the physical network topology. When a parameter changes, its related configurations should be altered as well. IMP facilitates the resolution and updating of these parameters by providing a framework to develop an object oriented declarative modelling language, in which relations between entities can be modelled at a high abstraction level. This high level configuration model can then be translated to the various low level configuration files on every machine.

Both the definition of the configuration model, containing the definitions of the various entities and their relations, and the translation mechanisms can be isolated in modules. This way tools developed with IMP are modular and extendible.

B. High level configuration model

An example of a high level model is shown in Listing 1. This high level model serves as input for the tool. Based on this model, the tool will generate and deploy the low level configuration files of the various machines in the network.

```plaintext
Listing 1. Example of a high level configuration model
```

The advantage of specifying the configuration in a high level configuration model is that the complexity of the configuration problem is reduced, consequently reducing the effort and time required to enact configurations. Dependencies between low level parameters can be derived from the model and are automatically kept consistent. Heterogeneity is encapsulated in the high level abstractions, and the configuration model uses a uniform platform independent syntax. Because the configuration is centralized in the configuration model and deployment is automated, the reconfiguration of the network infrastructure can happen efficiently. The system administrators can think on a higher level without being preoccupied by the low level configuration parameters.

IV. PREFIX ALLOCATION ALGORITHM

To further increase the efficiency and achieve a higher level of abstraction, the allocation of prefixes to subnets is automated. In the configuration model, a prefix is specified for the entire network. When translating the configuration model to the various low level configuration files, this prefix has to
be divided over the subnets of the network. This has to be done in the most optimal way, without wasting address space and possibly allowing aggregated routing. The subnet prefixes are allocated based on the network topology and the input network prefix. The network topology consists of a number of interconnected routers and subnets, with one designated border router connecting the network to the Internet. Prefix allocation happens differently for IPv4 and IPv6. The prefix allocation algorithms proposed are roughly inspired on an algorithm to allocate an IPv6 prefix discussed in [14]. The algorithm proposed in [14] is not fit for the prefix allocation required, mainly due to the strict need of a tree topology. Another problem is the assumption that routers should be connected via point-to-point links, and subnets with hosts are only allowed as leaves in the tree topology. Due to these inflexibilities, some of the basic ideas like the use of pools and metrics are reused to design new algorithms for both IPv4 and IPv6 prefix allocation.

A. IPv4

In the case of IPv4, address space is limited. The given network prefix has to be divided as efficient as possible over the subnets. Due to this, no special care is taken to make routing aggregated along the topology. The IPv4 allocation algorithm is displayed in Listing 2. The allocation algorithm uses a global pool which contains all the prefixes of varying length available. All the subnets in the network are assigned a prefix derived from this pool. The needed prefix length of each subnet is determined by their size, which is specified in the input configuration model. This size should always be rounded up to a power of 2. When the pool does not contain a prefix matching the needed size, an attempt is made to split a prefix of a smaller length until the desired length is reached. The residual prefixes are stored in the pool for reuse. For example, when splitting a /8 to get a /11 prefix, a /9, /10 and an extra /11 are generated. These prefixes are put back into the pool. This ensures an efficient allocation which minimizes the use of the address space.

Used variables and functions:
• \( P \): Pool of all available prefixes of varying length.
• \( M_s \): Desired subnet size for subnet \( s \), rounded up to a power of 2.
• \( \text{prefix.split}() \): Splits a prefix of length \( l \) in 2 prefixes of length \( l - 1 \).
• \( \text{neededPrefixLength}(\text{size}) = 32 - \log_2(\text{size}) \). Calculates the minimum prefix length needed by a subnet based on its size.

B. IPv6

Contrary to the IPv4 prefix allocation, the IPv6 allocation algorithm doesn’t have to be as economical with the address space. Also, the prefix length of an IPv6 subnet is always /64 per specification [15]. Because of this the prefix length doesn’t have to be calculated based on the size of the subnet. The proposed IPv6 allocation algorithm will try to allocate the subnet prefixes in a way that allows aggregated routing, but still tries to minimize address space usage. The algorithm operates in 3 steps.

1) Step 1 - Dijkstra: In the first step the network graph is reduced to a tree with the border router as root. Because deeper trees will cause a less ideal prefix allocation, a shortest path tree is preferred. Dijkstra’s algorithm is used to generate a shortest path tree from the network graph, minimizing the distance between each router and the border router. The output of this step will consist of the connected child routers of each router per interface.

2) Step 2 - Metric propagation: In the second step each router and each interface is assigned a metric denoting the amount of /64 subnets that are below the router or interface in the tree. The metric propagation algorithm is displayed in Listing 3. The algorithm is written recursively and the network tree is traversed depth-first. The output of this step are the metrics of each interface and router and also a set of of router interfaces that are the default gateway interface on their subnet.

Listing 2. Pseudocode prefix allocation IPv4 subnets

```java
void allocateSubnets() {
    for (s in Subnets) {
        length = neededPrefixLength(M_s);
        s.prefix = getPrefix(length);
    }
    return;
}

// Returns a prefix of the given length, splitting prefixes in the pool when necessary
Prefix getPrefix(length) {
    pref = getAvailablePrefix(length);
    while (pref.length < length) {
        pref, pref2 = pref.split();
        P.add(pref2);
    }
    return pref;
}

// Returns a prefix from the pool smaller or equal to the given length
Prefix getAvailablePrefix(length) {
    if (length < 0 || length > 32) {
        printError(“No available prefixes!”);
    } else if (!P.containsPrefixWithLength(length)) {
        return getAvailablePrefix(length - 1);
    } else {
        pref = P.getPrefixWithLength(length);
        P.remove(pref);
        return pref;
    }
}
```

Listing 3. Pseudocode metric propagation IPv6 subnets

```java
void metricPropagation() {
    for (r in Routers) {
        M_r = Dijkstra(r);
        for (i in Interfaces) {
            M_i = metricCell(int metric);
            if (i isTransitInterface()) {
               路由 is a transit subnet of a router. A subnet is a transit
```
3) Step 3 - Prefix allocation: In the final step the prefixes are allocated to the subnets top to bottom in the reduced tree. The algorithm is shown in Listing 4. The same pool system is used as in the IPv4 allocation algorithm, except that in this case there is a pool for each router and interface. Each pool contains the prefixes that can be allocated to the subnets below the interface or router it belongs to. Initially, the network prefix is added to the border router pool. Subsequently, in the assignPrefixRec() function, the router interfaces of the border router are allocated a prefix from this pool based on their metric. The allocated prefixes are added to the corresponding interface prefix pools. From these pools, the connected subnets are allocated a /64 prefix when the connecting interface is a gateway interface. Also the child routers connected through these interfaces are allocated prefixes and the assignPrefixRec() function is called recursively for these child routers. This continues until the leaf subnets are reached and all the subnets have a /64 prefix assigned.

Used variables and functions:

- $C_i$: Output from step 1.
- $M_i$: Output from step 2.
- $G$: Output from step 2.
- $P_r$: Prefix pool of router $r$. Contains all prefixes available for this router.
- $P_i$: Prefix pool of interface $i$. Contains all prefixes available for this interface.
- $\text{neededPrefixLength}(\text{metric})$: $64 - \log_2(\text{metric})$. Calculates the minimum prefix length required according to the given metric.
- $\text{prefix.split}()$: Splits a prefix of length $l$ into $2$ prefixes of length $l - 1$. 

//Entry point
void propagateMetric() {
    announceMetric(border);
}

//Recursive function, calculates the metric for the given router
int announceMetric(Router r) {
    $M_i = 0$;
    for (i in r.Interfaces) {
        $M_i += \text{announceMetric}(\text{childRouter})$;
    }
    $G.add(i)$;
    if ($\text{subnet}.\text{isTransitSubnet}()$ || $\text{subnet}.\text{hasGateway}()$ && !$\text{subnet}.\text{isTransitSubnet}()$) {
        $G.addGateway(i)$;
        $\text{subnet}.\text{addGateway}(i)$;
    }
    $M_f = \text{metricCeil}(M_i)$;
    $M_r = M_f + M_i$;
    $G.add(M_f)$;
    $\text{subnet}.\text{addGateway}(i)$;
    }
    $M_r = \text{metricCeil}(M_r)$;
    return $M_r$;
}

Listing 3. Pseudocode metric propagation IPv6

//Entry point
void assignPrefix() {
    $P_{\text{border}}.\text{add(Network.getIPv6Prefix())}$;
    assignPrefixRec(border);
}

//For a given router, assign prefixes to connected subnets and child routers and recurse over children
void assignPrefixRec(Router r) {
    for (i in r.Interfaces) {
        length = neededPrefixLength($M_i$);
        $P_r.add(\text{prefix})$;
        if ($G.contains(i)$) {
            $\text{subnet}.\text{isTransitSubnet}()$ = getPrefix($P_r$, 64);
        }
        for (c in $C_i$) {
            length = neededPrefixLength($M_c$);
            $P_r.add(\text{prefix})$;
            assignPrefixRec(c);
        }
    }
}

//Returns a prefix from the given pool of the given length, splitting prefixes when necessary
Prefix getPrefix(length) {
    if (length < 0 || length > 64) {
        printError("No available prefixes!");
    } else if (prefix.length < length) {
        return getAvailablePrefix(prefix.length - 1);
    } else {
        return getAvailablePrefixWithLength(length);
    }
}

//Returns a prefix from the given pool smaller or equal to the given length
Prefix getAvailablePrefix(length) {
    if (length < 0 || length > 64) {
        printError("No available prefixes!");
    } else if (prefix.length < length) {
        return getAvailablePrefix(length - 1);
    } else {
        prefix = prefix.split();
        return prefix;
    }
}

Listing 4. Pseudocode prefix allocation IPv6

V. Evaluation

The primary requirement of the configuration tool is to increase the efficiency of the configuration problem. In the evaluation of the tool an attempt is made to quantify this increase of efficiency. The tool is evaluated by comparing an automated configuration with the tool to a manual configuration. The configuration is carried out for an existing network topology currently in use at the Department of Computer Science of the KU Leuven shown in Figure 1.

A good measure for the increased efficiency is the comparison of the amount of effort the network administrators have to do in both cases. This effort translates to the amount of lines of configuration that have to be written. The amount of configuration lines contained in the input configuration model is compared to the number of lines in the various output configuration files.
To give a more fine grained view of the increase in efficiency, these lines of codes are categorized based on their complexity. An overview of the categories:

- **Addresses**: This category consists of all the lines with an IP address or a netmask. These parameters are very error prone to configure manually. Often dependencies exist between addresses, increasing the complexity.

- **Regular parameters**: This is a residual group containing all the lines that still contribute to the configuration but do not fall under the **Addresses** category. Frequently these lines are switches activating or deactivating a function.

- **Configuration metadata**: This category contains all the lines which do not contribute any useful configuration, but still are required to understand the structure of the configuration file. This category consists of brackets ( { }, ( ), [ , ] ) and keywords.

- **Comments and whitespace**: These are the lines in a configuration file that are completely meaningless configuration wise. The effect of the configuration would be identical without these lines.

The result of the evaluation is shown in Table I. The input configuration model is compared to the output configuration files. When using the configuration model, network administrators have to produce approximately 6.7 times less lines of configuration. This ratio further increases when only the most complex category **Addresses** is compared. In this category, the ratio is roughly 12.8. This shows the positive effects of using a high level configuration model. The amount of complex parameters that have to be produced by the administrators decreases. This is due to the abstractions of the various network elements and their mutual relations in the configuration model. Dependencies between low level parameters are derived from the configuration model, avoiding repetition of these parameters. This way less configuration parameters are needed when expressing the configuration in a high level model.

The compilation time the tool needs to translate the high level model is negligible. Our largest configuration model with 70 managed hosts takes less than 5 seconds to compile.

### Table I

<table>
<thead>
<tr>
<th></th>
<th>Configuration model</th>
<th>Configuration files</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addresses</td>
<td>46</td>
<td>590</td>
</tr>
<tr>
<td>Parameters</td>
<td>75</td>
<td>328</td>
</tr>
<tr>
<td>Meta</td>
<td>29</td>
<td>81</td>
</tr>
<tr>
<td>Comments</td>
<td>37</td>
<td>538</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>187</td>
<td>1537</td>
</tr>
<tr>
<td><strong>Total without comments</strong></td>
<td>150</td>
<td>999</td>
</tr>
</tbody>
</table>

**VI. RELATED WORK**

The related work can be divided into automatic configuration of network infrastructures and prefix allocation algorithms.

#### A. Automatic configuration tools

There are several automatic configuration tools commercially available that partially solve the problems described in the introduction. Examples of state of the art configuration management tools often used in production systems are Cfengine [16], Chef [17] and Puppet [18]. These configuration management tools are used by large dot-com companies like Facebook, eBay, Twitter and Oracle [19], [20], [21].

In contrast to the proposed configuration tool, these tools do not work at a high abstraction level and only facilitate the distribution and management of configuration. Administrators will still have to specify all the low level parameters of each configuration file, resulting in a costly inefficient configuration.

In addition to commercially available tools, some work has been done in research regarding model based configuration particularly for heterogeneous systems. Narain et al. [22] propose a solution for the network configuration problem by using an object oriented configuration model based on first order logic. A solver is used to generate the low level parameters using model finding. This solution differs from our approach by specifying configuration in first order logic. This complicates configuration specification; the administrator has to be familiar with first order logic concepts like inference, quantification and implication. Furthermore, when predicates are written in an inefficient but correct way, a solution may not be found.

Elbadawi et al. [23] present a framework that translates a high level configuration specification to low level parameters of each network component. The tool builds on top of NETCONF, an IETF standardized XML based RPC protocol for network management. The tool coordinates the configuration of interdependent network devices it manages via NETCONF by supplying a global NETCONF interface itself.
B. Prefix allocation algorithms

Atallah et al. [24] propose an algorithm for IPv4 prefix allocation and prove the optimality of the algorithm mathematically. The key difference with the proposed algorithms is the lack of support for aggregated routing.

Bhatia et al. [25] propose a distributed algorithm allocating IPv6 prefixes to MANET (Mobile Ad-hoc Network) nodes using ICMPv6. While a lot of effort is done to recuperate prefixes from dropped out or leaving nodes, no special care is taken to allocate prefixes in an aggregated way. Also, our allocation algorithm guarantees prefixes and routing happens along the shortest path tree with the border router as root. Whereas, in this approach interconnections between nodes are made randomly and a shortest path route to the border router is not guaranteed.

A more advanced dynamic prefix allocation algorithm for MANETs is proposed by Jelger et al. [26]. Introducing the concept of prefix continuity, they can guarantee prefixes are allocated in an aggregated way. In addition, path length is taken into account.

Kong et al. [27] developed a dynamic IPv6 prefix allocation algorithm inspired by memory management techniques using a paging mechanism. Bookkeeping guarantees subnet prefixes are split optimally and registers when subnets are added or removed. Route aggregation is not taken into account.

VII. CONCLUSION

In this paper, an automatic configuration tool for dual stack IP networks is proposed. The primary requirement of this tool is to improve the efficiency of network configuration.

The proposed tool reduces this complexity by using a high level configuration model. From the evaluation, it is clear that the tool significantly increases efficiency. In particular, the amount of complex error prone parameters is reduced.

In the future, the tool will be expanded to include other configuration aspects. As the tool is built on top of the IMP framework [13], support for aspects can be isolated in modules. These can be network related (e.g. firewall configuration), or application oriented (e.g. distributed file systems, web servers). By including these aspects in the high level configuration model, the same increase in efficiency is achieved for these aspects. Furthermore, dependencies between aspects are automatically derived and configured. For example, when adding a web server to the configuration model, the corresponding firewall rules will be automatically adjusted to allow traffic to and from the web server. By adding support for more aspects, the configuration tool can evolve towards a modular integrated configuration platform.

ACKNOWLEDGEMENTS

This research is partially funded by the Research Fund KU Leuven.

REFERENCES


