QoS in Forwarding Strategies for ICN: New Algorithm and Evaluation Through Simulation

Ivanes Araujo, Andrey Silva, Aldebaro Klautau, Neiva Linder

Abstract—Several forwarding strategies (FS) were already proposed to leverage packet delivery in information-centric networking (ICN). However, less attention has been given to FS potential to improve quality of service (QoS) in ICN. This work addresses the impact of FS on the performance of voice application, video streaming and file transfer protocol (FTP). Particularly, we investigate the performance gain when distinct FS are tailored to these traffic categories using named data networking (NDN), a prominent ICN architecture. This work also proposes a new forwarding strategy, called inverse pending interests (IPI), which can be used for low-priority traffic. Given that distinct FS are simultaneously applied, we also evaluate their interoperability using three scenarios and observing QoS metrics for the applications mentioned above with different priorities for each traffic. The work was conducted within a multiobjective optimization framework, which allows defining dominant sets of solutions and enriches the analysis when compared to aiming at a single optimal solution. The simulations were performed using the ndnSIM simulator and the results indicate, for example, that multicast and best-route are the most suitable FS for voice application, while the request forwarding algorithm could better serve video streaming. The proposed IPI algorithm is beneficial when low-priority (FTP) traffic is present, and its adoption does not significantly impair other FS.

Index Terms—ICN, NDN, Forwarding strategies, QoS.

I. INTRODUCTION

In information-centric networking (ICN), the communication is based on content requests and not on endpoints, as legacy Internet protocol (IP) networks [1]. ICN targets several IP issues such as bottleneck and latency reductions [2]. Additionally, ICN can raise network performance with forwarding strategies (FS), that can benefit from its structure in forwarding decisions.

In IP, before the emergence of multipath routing, forwarding actions were mostly limited to a single next-hop while more complex network path decisions were made in the routing. With multipath routing, however, different paths could be accounted for forwarding, increasing the relevance of this decision [3]. In spite of that, the forwarding decisions on IP networks can still face limitations due to multipath routing restrictions such as establishing only loop-free routes [4] and maintaining a per-flow state to forward packets from the same flow by the same path, to avoid variable latency that might affect packet ordering and congestion detection. For instance, as conceived in [5], multipath TCP accounts for multiple paths by the use of subflows, which essentially represent a flow operating over an individual path, similar to a regular TCP connection.

Although these constraints might reduce the possibilities of FS proposals in IP, in NDN, loop detection is already present with the use of a nonce mechanism in interest packets, and there is no need to limit the forwarding of packets from a specific flow to only one path since NDN favors content retrieval instead of end-to-end connections. In this manner, a plethora of FS has been proposed considering different metrics for forwarding decisions such as delay, content popularity and cache hit ratio [6], [7], [8].

Nevertheless, the use of FS for service differentiation still needs further investigation. Since different applications have distinct requirements regarding the quality of service (QoS), the use of an appropriate forwarding strategy for each application might increase user experience [9]. However, in current state-of-art, a considerable number of FS are evaluated without taking into consideration specific traffic with its particular demands. Furthermore, FS are usually tested without taking into account their performance regarding coexistence, i.e., how it would perform in an environment where other FS are used in the same node for different applications. To address this issue, we evaluate several FS in NDN, with ndnSIM simulations, considering three different applications: Voice over NDN (VoN), video streaming and file transfer protocol (FTP). To better analyze which FS could be more effective concerning QoS improvement, in combination with other FS in use, the set of quality values obtained are approached as a multiobjective optimization problem (MOP), used in a weighted sum method. The results are also analyzed regarding the dominant set concept in MOP, to address the impact of one forwarding strategy in another one. Moreover, aiming to empower the use of FS for QoS, we propose a new forwarding strategy adapted for low-priority services, called inverse pending interests (IPI). Our contributions, then, are threefold:

- Evaluation of QoS potential of FS for specific services, in a MOP context.
- Impact analysis of forwarding strategy coexistence.
- Proposal of a new forwarding strategy for low-priority services.

Ivanes Araujo, Andrey Silva and Aldebaro Klautau are with Federal University of Pará, Belém-PA, Brazil. emails: ivanes@ufpa.br, aldebaro@ufpa.br, andrey.silva@ufpa.br.
Neiva Linder is with Ericsson Research, Kista, Sweden. email: neiva.linder@ericsson.com.
Digital Object Identifier: 10.14209/jcis.2019.28
The rest of this paper is organized as follows: Section II presents the related work regarding QoS and forwarding strategy in ICN networks. Section III describes the features of evaluated NDN FS, followed by details of our proposed forwarding, IPI, in Section IV. Section V covers the methods used for the analysis of FS performance. Section VI provides information about the parameters used for simulation. In Section VII, simulation results are evaluated regarding FS efficiency in congested scenarios and multiple FS utilization. Finally, Section VIII presents the conclusions of this work.

II. RELATED WORK

QoS for ICN has been explored mainly concerning architecture and FS. Here, some representative works in QoS are summarized, regarding advances in architecture, followed by forwarding strategy proposals, which are more aligned with this work.

In [10], a QoS model is designed for ICN with multiple QoS levels. For example, if no other constraint is provided, a default QoS is adopted according to the type of traffic, but it can be overwritten by QoS parameters either from the network provider or users and/or service providers, allowing content priority. Another architecture is proposed in [11], based on differentiated services (DiffServ), where packet marking is defined, as well as changes in interest aggregation and cache policies, adapting the use of QoS for ICN.

Regarding FS, a new forwarding strategy is developed in [6], accounting for QoS in the network by the measurement of bandwidth, round-trip time (RTT), and the use of modified interests that possess QoS information. This forwarding strategy is tested against best-route (BR) forwarding strategy, achieving better results than the latter one, concerning mean delivery time and hop count of data packets. Another forwarding strategy is proposed in [7], where weights are defined for each kind of traffic to forward interest packets. Depending on the content request that arrives in a router, it will have a different priority regarding which interface should be chosen, derived from its weights. The goal, in [7], is to explore the capability of using information about various services in a single forwarding strategy, increasing QoS. The work in [8] defines the forwarding priority of an interest to an interface based on its class. Interests from high priority classes are forwarded by interfaces that can provide smaller delivery time. Each class has a delay threshold, and if the experienced average delay is lower than this limit, interests from this class will be queued, allowing classes over their threshold to forward packets by these interfaces.

Although the efforts in the literature represent significant advances for service differentiation in ICN, none of them treat the influence of individual FS in distinct applications. FS mentioned above ([6], [7] and [8]), are focused in adaptable strategies that can cope with distinct demands. Namely, one could phrase the goal of such contributions as following: if only one forwarding strategy could be used for all traffics, how could it be leveraged for QoS purposes.

Our goal, instead, is to evaluate approaches that can be suitable for a specific application. Instead of analyzing exclusively how a single forwarding strategy can perform for a variety of traffics, we aim to establish which forwarding strategy could better satisfy a single service, in a context where decisions of one forwarding strategy can influence another forwarding strategy performance. The principle that each service requires a proper forwarding strategy is also outlined in [9], which details FS functionalities for NDN. Furthermore, our IPI proposition differs from previous work since it intends to be used for low-priority applications, giving the possibility to use FS for service prioritization in ICN. To the best of our knowledge, this is the first work in the literature to deeply explore the mentioned contributions.

III. FORWARDING STRATEGIES IN NDN

In this section, FS are detailed considering the NDN forwarding pipeline, where the request of a content is made by the transmission of an interest packet, which is then responded with a data packet, i.e., with the corresponding content. When receiving an interest, an NDN node will follow a forwarding process as defined in the upper part of Figure 1. A node, by the arrival of an interest packet, will first check if it has the requested content in the Content Store (CS). If CS has the content, the corresponding data packet is sent downstream. Otherwise, the incoming interface of the interest will be recorded in the pending interest table (PIT). If this is the first entry in the PIT for this interest, it will be forwarded to one of the upstream interfaces defined in the Forwarding Information Base (FIB). The decision of which specific interface will be used in forwarding is made by the forwarding strategy. The bottom part of Figure 1 describes when a data packet arrives in an NDN node, where a check is made to identify whether a corresponding interest packet was recorded in the PIT. If this check succeeds, the data will be stored in CS and it is transmitted to all the corresponding interfaces recorded in the PIT.

Fig. 1: Forwarding Process in NDN, adapted from [12].

From the NDN procedure above, FS are responsible for choosing the proper interface for interest transmission, which is the focus of this work. The creation of new FS is encouraged in NDN architecture [9], making space for several proposals in the literature [13], [14], [15], [16], [17].

For this work, to evaluate exclusively forwarding without cache influence, no FS related to cache or content popularity, such as [13], are considered. Similarly, FS including package drop functionalities, such as [14], are also excluded from this
evaluation. This exclusion is motivated by the fact that, in QoS context, package drop policies can be seen as an overlay feature, i.e., policing and shaping can be applied after any forwarding decision, resulting in package drop/mark if needed. In this manner, investigated FS always choose valid interfaces to forward interests. Thus, we adopt FS present in ndnSIM simulator multicast (MCS), BR [15] and also relevant ones from the research community (on-demand multipath interest forwarding (OMP-IF) [16] and request forwarding algorithm (RFA) [17]). A brief description of their forwarding actions is made below:

- MCS forwards the interest to all available interfaces, e.g., all possible interfaces listed in the FIB, except for the one from which the interest arrived [9].
- BR chooses the interface with the lowest routing cost to transmit the interest. Re-transmissions are forwarded to the lowest cost interface, not previously used to forward this interest [9].
- In RFA, each interface has a weight inversely proportional to its average number of pending interests. An interface is chosen probabilistically, according to the defined weights [17], i.e., interfaces with less pending interests tend to be chosen for forwarding, creating a load balancing scheme.
- OMP-IF has different forwarding probabilities for each interface, per name prefix, guided by the goal of having the lowest RTT. Besides, for the selection of the possible interfaces to transmit the interest, a path discovery is made, where each interface used for packet transmission follows a disjoint path from other interfaces [16], if possible.

In the next Section, the functionality of IPI is described.

IV. IPI FORWARDING STRATEGY

FS are usually designed with the aim of providing a better overall QoS for end users. Therefore, less attention is given to low priority applications, that would not necessarily need to have its interests forwarded by the shortest path or closer cache-hit node, for instance. To address this issue, a new forwarding strategy is proposed, called inverse pending interests (IPI), with the focus on sending interests through the most congested interface, up to a certain limit. To get further insight into IPI operation, Figure 2 characterize an example of IPI aimed behavior. In this example, the depicted router is traversed by two different services A and B, the former with higher priority than the latter. Service A is forwarded with a strategy that can provide, e.g., the lowest delay path. Service B, however, needs to be forwarded in a way that A can continue to flow by the best path without higher competition. If service B used the same forwarding strategy of A, then both would flow by the best path, creating higher competition for resources while waisting the possibility to forward packets by the higher delay interface. Hence, IPI could be used for service B, since its principle is to forward packets by the most congested or higher delay path. As later detailed in this section, to benefit from NDN built-in capability to record pending interests (PIT), IPI forwards most of the packets by the interface with the higher number of pending interests.

More details of IPI operation are described in the Algorithm 1, which is executed after the forwarding process from Figure 1, i.e., when cache, interest loop and PIT were already checked, and the interest is considered ready to be forwarded upstream. Minor functions are described in Subsection IV-A, and a summary of the terms used henceforth are provided in Table I.

![fig2.png](image)

**Fig. 2:** Example of IPI operation in conjunction with another forwarding strategy.

Regarding main behavior, IPI will start by an exploration phase to define the interface $i$ with the most average pending interests $P_i$ (adopted in this work as an exponential moving average, as performed in [17]), which can be considered the slower available interface. For a fair $P_i$ comparison, during the exploration phase, forwarding is performed aiming an even distribution of interests among interfaces. Thus, considering $\mathcal{E}$ the set of all available interfaces $i$ through which interests can be sent (interfaces present in the FIB, excluding the one from which the interest came), forwarding behavior in exploration phase can be modeled as:

$$
\min(s_i(t)), \forall i \in \mathcal{E}
$$

(1)

where $s_i$ is the number of interests sent until a given time $t$, through an interface $i$. The be performed until a specific event occurs, such as a limit number of sent packets reached or an interest times out, which will be discussed in Subsection IV-A. After this phase, an arrival of a new interest will trigger the determination of the interface $j$ that satisfies:

$$
\max(P_i(t)), \forall i \in \mathcal{E}.
$$

(2)
In case of a tie between interfaces, \( j \) is chosen randomly among tied ones. Interface \( j \) will be the main forwarding choice, considering that IPI will be used for services with less strict requirements. When in congestion, however, interests will be forwarded according to:

\[
\min(p_i(t)), \forall i \in E
\]

where \( p_i \) is the number of pending interests in a given time \( t \). This condition is made to not over penalize interests under IPI since even in the case from Expression 3, when interests follow the less loaded interfaces, other service differentiation features, such as scheduling, might still be applied to these packets. Hence, this design aspect of IPI can avoid bottlenecks and unnecessary losses on traffic forwarded with this strategy.

To determine congestion in the network, an indirect approach similar to TCP congestion control is adopted, where the condition \( \mathcal{E}_j = \emptyset \) must be attended to consider that the network is congested, with the subset \( \mathcal{E}_j (\mathcal{E}_j \subset \mathcal{E}) \) defined as:

\[
\mathcal{E}_j = \{ i : p_i(t) < W_i(t) \}
\]

where \( W_i \) is a window defined in number of pending interests in a given time \( t \). In this case, as well as for Expression 3, no average is taken into consideration since, for congestion detection, \( p_i \) must represent the actual number of pending interests.

If no congestion is detected the interest will be forwarded by the interface \( j \) if \( j \in \mathcal{E}_j \), otherwise it will be forwarded by the interface \( i \) that attends:

\[
\max(p_i(t)), \forall i \in \mathcal{E}_j.
\]

This decision aims to prevent high priority services, possibly forwarded by other FS, from resource competition with contents processed by IPI. Moreover, in case of a tie between interfaces, the one by which the interest will be forwarded is chosen randomly among the tied interfaces.

**Algorithm 1 IPI Algorithm**

```plaintext
if exploration_phase = true then
    Forward interest according to Expression (1)
else
    if main_interface_not_chosen = true then
        Choose main interface \( j \) following Expression (2)
    end if
    if \( \mathcal{E}_j \neq \emptyset \) then
        if \( j \in \mathcal{E}_j \) then
            Forward interest by interface \( j \)
        else
            Forward interest according to Expression (5)
        end if
    else
        Forward interest according to Expression (3)
    end if
end if
```

**A. Minor Functions**

For IPI operation, exploration phase stop/re-start and \( W_i(t) \) must be defined. However, several approaches could be adopted for those cases, and it should be highlighted that their definitions hereafter are an embodiment of IPI behavior, but do not preclude other approaches.

The exploration phase will stop, i.e., will be set to false, when an interest times out. This decision is based on the fact that if there is no packet loss, thus no congestion is detected, there is no need to penalize low priority transmissions, while also expending more network resources. Therefore, IPI will keep its forwarding decision based on Expression 1, which does not penalize the traffic, and further service differentiation can be sufficiently performed with proper content scheduling. Concerning the re-start of the exploration phase, it should be triggered when the node is not loaded, in a way that measurements from average pending interests in each face can be correctly obtained. Hence, when the number of pending interests from Expression 5 is \( p_i(t) = 0 \), exploration phase is re-started.

With respect to window \( W_i(t) \), that must detect congestion in the interface \( i \), a classical behavior of adaptive increase multiplicative decrease (AIMD) congestion control mechanism is adopted [18]. Initially with value \( 1 \), \( W_i(t) \) will be increased by 1 for each data packet received by interface \( i \). The evolution of \( W_i(t) \) over time can be defined from [18] as:

\[
\frac{d}{dt} W_i(t) = \alpha \frac{R_i(t)}{R_i(t)} - \beta W_i(t) W_i(t - R_i(t)) T_{io}(t - R_i(t)) \quad (6)
\]

where \( R_i(t) \) is the RTT in time \( t \), \( T_{io}(t) \) is the probability of an interest time out, \( \alpha \) and \( \beta \) are the values of increase and decrease of \( W_i(t) \), respectively. An increase will happen when a data packet arrives by interface \( i \), while a decrease will happen when a packet times out. Values for \( \alpha \) and \( \beta \) in AIMD congestion control are typically adopted as \( \alpha = 1 \) and \( \beta = 1/2 \) [18].

**B. Pending interests per face with IPI**

The main IPI objective can be understood as forwarding interests primarily by interface \( j \). Therefore, the number of pending interests per face must be analyzed. To understand \( p_i(t) \) when using IPI, \( W_i(t) \) will be discussed, since it represents an upper bound for \( p_i(t) \) values when not in congestion. To this end, it is considered that a node has \( A(t) \) interests per second to forward by any interface \( i \in E \), which will determine \( W_i(t) \) evolution.

When \( A(t) < W_j(t)/R_j(t) \) (case 1), \( W_j(t) \) will be the only interface growing since IPI always send interests through interface \( j \) if \( j \in \mathcal{E}_j \). When \( W_j(t)/R_j(t) < A(t) \leq \sum_{i \in \mathcal{E}_j} W_i(t)/R_i(t) \) (case 2), other interfaces will also have their window increased. A priori, interfaces with small RTT values tend to have a bigger window growth than slow RTT ones. This would actually differ from desired IPI behavior, as the slowest interface would not be the one forwarding more interests. But considering that without packet losses, every \( W_i(t) \) can grow in cycles of RTT \( R_i(t) \), interfaces in fact would have greater periods for their cycle. Interface \( j \)
can be inferred from Expression 2 to be the one with the greater RTT value, and it would have its window increased every $R_i(t)$. Any other interface would have to wait for interests to be forwarded by $j$ until $p_j(t) = W_j(t)$. Thus, the interface $i$ with the smaller RTT, other than $j$, must wait $W_j(t) - p_j(t)$ packets to be forwarded by $j$ to forward the first interest. Similarly, when starting to transmit by an interface $i$, given the forwarding condition from Expression 5, it will continue to forward by the same interface until $p_i(t) = W_i(t)$.

Hence, considering that a new interest must be forwarded every $1/A(t)$ period, an interface $n$ would have its cycle period defined as $R_n(t) + L_n$, where $L_n$ is:

$$L_n = \sum_{k \in \mathcal{E} : p_k(t) > p_n(t)} \frac{W_k(t) - p_k(t)}{A(t)}. \quad (7)$$

The value of $L_n$ would be greater for interfaces with small $R_n(t)$, which could reduce the difference of window growth rates. For instance, with $R_n(t) + L_n > R_j(t)$, $W_j(t)$ would have a higher growth rate than $W_n(t)$. When $A(t) > \sum_{i \in \mathcal{E}} W_i(t)/R_i(t)$ (case 3), the network is in congestion, leading to packet transmissions by the interface with lower pending interests, thus a possible balance between $p_i(t)$ while in congestion. In this case, however, interfaces with lower RTT can have a bigger window growth than $W_j(t)$, which should be treated at other prioritization tools such as shaping and scheduling. Additionally, it is possible to have $p_i(t)$ higher than $W_i(t)$, since it represents a congestion scenario. A situation with $p_i(t) \leq W_i(t)$ can be achieved again when $\mathcal{E} \neq \emptyset$.

C. Loop-free constraint

Usually, FS rely on loop-free routing to properly forward packets in the network [4]. Nonetheless, even when loop routes exist, most FS might still satisfy a considerable amount of interests. This can be illustrated in Figure 3, which depicts a possible forwarding behavior.

![Fig. 3: Possible forwarding decision for RFA.](image)

In Figure 3, router $A_1$ must forward an interest either to router $A_2$ or $A_3$ to arrive at the given content source. When using, e.g., RFA, even if the interest follows the path described by dashed lines, the number of pending interests in router $A_1$ would increase for the interface with router $A_2$, leading to forward interests mostly by the interface with router $A_3$. Similarly, after some increase in the number of pending interests in the interface from router $A_3$ to $A_1$, router $A_3$ would mostly send interests through its interface with the content source. Hence, RFA can work under not loop-free routes, giving the possibility to exploit new paths that usually cannot be considered in loop-free routes. In contrast, IPI in router $A_1$ would probably choose the interface with router $A_2$ as the main face for transmission, given its probability to have more average pending interests due to loop routes. Additionally, router $A_3$ would likely choose its interface with router $A_1$ as its main interface for the same reason. This configuration leads to main interface choices resulting in loop transmissions which, as detailed in Subsection IV-B, would represent most of the transmitted interests. In the case of $W_j(t)/R_j(t) < A(t) \leq \sum_{i \in \mathcal{E}} W_i(t)/R_i(t)$, even if $W_j(t)$ will be prevented from growing due to the lack of satisfied interests, this will push interest forwarding to faster interfaces that should be loaded by applications carried with other FS.

Hence, although this work is not focused on routing protocols, it should be emphasized that IPI must operate under the constraint to use loop-free routes to have its proper behavior, i.e., $\mathcal{E}$ set defined in IPI must consider only interfaces that will result in loop-free paths.

V. EVALUATION METHODS

In this section, the methodology to evaluate the FS in Sections III and IV is described. Firstly, the performance of each forwarding strategy is assessed according to the QoS obtained for each application used in the scenarios described in Section VI, as detailed in Subsection V-A. Secondly, as each forwarding strategy might improve the quality of one specific service while reducing other service qualities, the set of quality values obtained from FS are further analyzed in a MOP context, discussed in Subsection V-B.

A. QoS of each service

As mentioned in Section I, the applications used in the simulated scenarios are VoN, video streaming, and FTP. Their details are given in Section VI, while this Subsection will cover QoS metrics related to each application.

Considering the referred traffics, VoN was analyzed with the use of conversational quality estimation of mean opinion score (MOS-CQE) [19]. To obtain MOS-CQE values, the mapping described in [20] was used, converting the rating factor $R$ in MOS-CQE $1 - 4.5$ scale. $R$ is calculated with a simplified E-model, also described in [20], that accounts for packet loss and delay impairments, with a maximum value of 4.5 when mapped to MOS-CQE. This value can be transduced to user satisfaction according to Table II [20].

<table>
<thead>
<tr>
<th>MOS-CQE</th>
<th>User Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.34</td>
<td>Very satisfied</td>
</tr>
<tr>
<td>4.03</td>
<td>Satisfied</td>
</tr>
<tr>
<td>3.60</td>
<td>Some users dissatisfied</td>
</tr>
<tr>
<td>3.10</td>
<td>Many users dissatisfied</td>
</tr>
<tr>
<td>2.58</td>
<td>Nearly all users dissatisfied</td>
</tr>
</tbody>
</table>

To evaluate video streaming, the model proposed in [21] for dynamic adaptive streaming (DAS) was used, which is
based on the video quality sum of all segments received, minus variations between consecutive segments quality, and re-buffering time events. The model permits to assign different weights for the impact of quality variations and re-buffering time called $\lambda$ and $\mu$ respectively. The values adopted for those parameters are extracted from [21] ($\lambda = 1$ and $\mu = 3000$), representing that 1 second of re-buffering time has the same impact of a 3000 kbps reduction in a video segment bit rate. Moreover, to give further knowledge about what the video quality represents, the final obtained value is divided by the number of received video segments to have an average quality per segment in bits per second. Table III shows those average segment qualities and what can be inferred about the video streaming experienced by the user. The video streaming application is assumed to use scalable video coding (SVC), having three layers with $L_0$ representing the lower layer and $L_2$ the highest one (see Subsection VI-A for more details about video application). Hence, if the average segment quality is at least 1.382 Mbps, which is the maximum attainable value, it means that users could watch the video always at the highest quality ($L_0+L_1+L_2$). If the average segment quality value is at least 975 kbps, it is equivalent to users watching the video with at least $L_1$ quality ($L_0+L_1$). For example, a value of 1.1 kbps is equivalent to users having, besides always $L_1$, also $L_2$ for some segments of the video. This same logic is applied to the row of 640 kbps value, while average segment qualities below this value are equivalent to have only $L_0$ video quality and also re-buffering times. It is said to be an equivalent behavior because since quality variation between segments and re-buffering time are accounted in the video quality, a value of, e.g., 1.1 kbps can mean that users were always watching the video with $L_2$ quality, but with a lot of stalling periods due to re-buffering. However, to facilitate the interpretation of the results in Section VII, we will use this equivalent behavior as a guide for video evaluation.

<table>
<thead>
<tr>
<th>Average Segment Quality</th>
<th>Equivalent video streaming behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.382 Mbps</td>
<td>Users have the highest video quality $L_2$</td>
</tr>
<tr>
<td>0.975 Mbps</td>
<td>Users have at least $L_1$ video quality</td>
</tr>
<tr>
<td>0.640 Mbps</td>
<td>Users have at least $L_0$ video quality</td>
</tr>
</tbody>
</table>

FTP traffic is evaluated in terms of the ratio of requested interests that could be satisfied (0-1). As not having a strict delay or jitter constraint, no intermediate ratio values are mapped to a specific user satisfaction. Hence, the FTP quality is considered to be excellent as the ratio value gets close to 1 and poorer as it gets closer to 0.

### B. QoS in MOP

Quality for VoN, video streaming and FTP can be seen as competing objectives, e.g., the enhancement of one quality might affect another one. Thus, for the performance evaluation of FS, the selected three application qualities are analyzed as a MOP. In this context, an objective function vector can be defined as $\overrightarrow{f}(\overrightarrow{x}) = [f_1(\overrightarrow{x}) \ f_2(\overrightarrow{x}) \ \ldots \ f_N(\overrightarrow{x})]$ with $\overrightarrow{x} = [x_1 \ \ldots \ x_k] \in \mathcal{X}$, which is the feasible set of decision vectors. $\overrightarrow{F}(\overrightarrow{x})$ expresses the set of individual objective functions $f_N$. Each $f_N(\overrightarrow{x})$ maps a specific event in $\overrightarrow{x}$ for a value that can represent the performance of this event. For a system considering a single $f_N(\overrightarrow{x})$, the optimal solution is the $\overrightarrow{x}_0$ that can maximize or minimize $f_N$. However, MOP usually consists of conflicting goals, e.g., the solution $\overrightarrow{x}_0$ that maximizes $f_1(\overrightarrow{x}_0)$ does not always maximizes $f_2(\overrightarrow{x}_0)$. Therefore, a MOP usually has a set of optimal solutions instead of just one, where trade-offs can be evaluated regarding the increase in a given $f_N(\overrightarrow{x})$ value at the cost of the reduction of another objective function value.

In MOP, one can say that a solution $\overrightarrow{x}_1$ dominates another $\overrightarrow{x}_2$, if the following conditions are attended [22] (hereafter, maximization is assumed):

$$\forall i \in \{1, 2, \ldots, N\} \quad f_i(\overrightarrow{x}_1) \geq f_i(\overrightarrow{x}_2),$$
$$\exists j \in \{1, 2, \ldots, N\} \quad f_j(\overrightarrow{x}_1) > f_j(\overrightarrow{x}_2)$$

(8)

The solutions that are not dominated by any other, w.r.t. $\mathcal{X}$ set, are called Pareto front, and represent the Pareto-optimal solutions set $\mathcal{P}$ [23], called interchangeably as optimal solutions henceforth. For instance, Figure 4 depicts a MOP with two objective functions ($f_1(\overrightarrow{x})$ and $f_2(\overrightarrow{x})$). For each $\overrightarrow{x}$, $f_1(\overrightarrow{x})$ and $f_2(\overrightarrow{x})$ values are obtained (circles). Values that are not dominated by any other are highlighted (dotted line), representing the Pareto front.

![Fig. 4. Pareto front example for a MOP with two objective functions.](image)

Adopting MOP and the Pareto front concept, we can define our objective function vector as:

$$\overrightarrow{F}(x) = [f_1(x) \ f_2(x) \ f_3(x)]$$

(9)

where $f_1(x)$, $f_2(x)$ and $f_3(x)$ represent the quality values for VoN, video streaming and FTP respectively, called henceforth voice, video and data quality, for simplicity, and normalized by the highest attainable value for each application (4.5 for voice, 1.382 Mbps for video and 1 for data quality). Since $X$ should contain all the possible triples of FS, a solution $x$ is defined as $x \in X : x = [ \text{forwarding strategy for VoN} / \text{forwarding strategy for video streaming} / \text{forwarding strategy for FTP} ]$. For instance, $x_0 = [\text{MCS/BR/RFA}]$ represents the solution where MCS was used for VoN, BR for video streaming and RFA...
for FTP, while \( \vec{F}(\vec{x}_0) \) is the vector expressing the application qualities for VoN \((f_1(x_0))\), video streaming \((f_2(x_0))\) and FTP \((f_3(x_0))\). With 5 possibilities of FS for each application (MCS, BR, RFA, OMP-IF and IPI), e.g. \([MCS/BR/RFA]\) is a different solution from \([MCS/RFA/BR]\), we have a set of 125 solutions in \(X\).

Apart from deriving the Pareto front, it is also possible to define a single solution to a MOP. This can be performed by attributing weights to each \(F_i(\vec{x})\) to compose one objective function \(F_i(\vec{x}) = w_1 f_1(\vec{x}) + w_2 f_2(\vec{x}) + ... + w_N f_N(\vec{x})\), where \(w_N\) is the weight attributed to the objective function \(N\). Hence, a weighted sum can be used to collapse a MOP into a single objective function that can be optimized. For convenience and to facilitate weight attribution, \(f_N(\vec{x})\) values are usually normalized such that \(\sum_{i=1}^{N} w_i = 1\). Furthermore, if the weights are positive for every \(f_N(\vec{x})\), the optimal solution determined with \(F_i(\vec{x})\) is also part of the optimal Pareto set \(P\), which is guided by the concept of a dominant solution [24].

Both solutions from single objective function derived from weight attribution (in Subsection VII-A) and Pareto front ones (in Subsection VII-B) are used in this work.

VI. Scenario

In this section, the framework used to test FS in Sections III and IV is described. Multiple simulation runs were made with sufficient time to enable the tested FS to perform packet forwarding decisions for several packets. The simulations were conducted in a customized ndnSIM, provided by [7], with the support to estimated voice, video and data quality, as well as the respective applications needed and the FS RFA and OMP-IF implemented (MCS and BR are default strategies in ndnSIM). IPI was implemented by the authors and the scenarios defined next were designed in this customized ndnSIM.

Three scenarios were adopted, as depicted in Figure 5. To exploit FS use in the network, each scenario expresses real topologies from [25], Italian academic and research networks (GARRB), European optical network (EON) and Germany backbone network (GBN), where each one was chosen w.r.t. values of average nodal degree \(A_d\) and average path length \(A_l\).

As representing the average number of links connected to each node, \(A_d\) must be 3 or more. The motivation for this relies on the fact that when making a forwarding decision, a node will already exclude the incoming interest interface from its options. Consequently, with less than 3 interfaces, nodes would have no interface options and would just forward the interests through the only possible interface.

Likewise, \(A_l\) value should be 2 or more, which can be observed from average path length equation \(A_l = \frac{1}{N(N-1)} \sum_{a,b} d(n_a, n_b)\), where \(N\) is the number of nodes in the network, and \(d(n_a, n_b)\) denotes the shortest distance (in number of hops) between node \(n_a\) and \(n_b\). According to this \(A_l\) definition, \(A_l < 2\) implies that usually the nodes can reach final destination (e.g., nodes linked to servers from Figure 5) in the next-hop. In this manner, little can be exploited from FS since most of the times only one node along the path from the user to the server might be used for forwarding decisions. Moreover, to not limit FS, routes between nodes and servers were defined considering at least two paths for most of the cases, to have more than one interface option for forwarding decisions.

![Fig. 5: Scenarios used for FS evaluation with \(A_d\) and \(A_l\) indicated.](image-url)
Regarding the users in the network (triangles), each one represents a bundle of the three considered applications: VoN, video streaming, and FTP. Each application in each user starts according to a uniform distribution between 0 and 3 seconds of simulation. To facilitate visualization, link rates in Figure 5 are represented merely by an M, indicating link capacity in Mbps, while dotted lines represent ideal links. Link capacities are designed with the intention to stress the network, with GARRB as the less congested network and GBN as the most congested one. For all the scenarios, propagation delays are, in general, distinct for each link.

In all scenarios from Figure 5, for the sake of FS evaluation, the NDN cache feature is disabled, and each user requests a unique content, excluding the possibility of traffic aggregation in the network. More details about traffic considerations are given in Subsection VI-A.

A. Applications Details

Current applications that use TCP/IP architecture may need adaptations to work in ICN architecture. Push-based functionality, concerning video over IP (VoIP) behavior, is not a native feature in NDN [26]. Hence, previously proposed techniques to address VoIP issues in ICN will be considered in this work. More specifically, voice push-based actions are modeled as the work in [27]. Furthermore, voice application considers a fixed rate of 64 kbps, with G.711 voice codec. Since the modeled voice application works under NDN, although VoIP inspired, it is called VoN. For video application, we assume SVC as DAS application, each video segment with average layer quality \( L_n = \{640, 335, 407\} \) kbps, buffer-based adaptation logic [28], and dataset from [29]. FTP traffic is modeled as a uniform distribution in a range of 0 up to 6 Mbps. Interest lifetime, i.e., the remaining time before an interest times out, is chosen accordingly for each application as 0.05, 1 and 2 seconds for VoN, video streaming, and FTP, respectively.

It should be noted that VoN has a traffic rate much lower than the video and FTP traffic, which could lead to the analysis of how much each application impacts on the overall amount of traffic in the network. For instance, when voice quality is enhanced to the detriment of video and data quality, the overall throughput obtained from applications can be actually lower than other configurations (bad voice quality but good quality for the other applications), since video and FTP applications represent heavier traffic. Nevertheless, this work aims to explore QoS analysis, where a higher throughput cannot necessarily be translated to a higher QoS, as detailed in Section V. Therefore, the results presented in Section VII are based on voice, video and data quality, regardless of the total throughput achieved in the network, or demand percentage of one traffic compared to other applications in the network.

Moreover, to be able to use a forwarding strategy for each kind of service, the network should map interest requests into applications, and group together equal ones under the same forwarding strategy. Nevertheless, it is out of the scope of this paper the exact procedure of classifying services in ICN and how to apply a specific forwarding strategy to specific traffic. There are different manners pointed out in NDN architecture to classify services, but the topic is still an open issue [30]. Thus a perfect map between traffics used in this work and chosen FS is considered. Similarly, it should be noted that as ICN is evolving, there are still several open issues under discussion, such as the feasibility of deployment [31], forwarding processes in heterogeneous environments [32], and routing schemes [33] [34] [35], which are out of the scope of this paper.

VII. RESULTS DISCUSSION

FS performance is addressed in this section, where it must be highlighted that the \( X \) set contains all the possible triples of FS. An optimal solution \( x \) (as in multiobjective description in Section V) is, then, defined as \( x \in X : x = [ \text{ forwarding strategy for VoN / forwarding strategy for video streaming / forwarding strategy for FTP } ] \).

In the case of the weighted sum method, attributing weights \( w_1, w_2 \) and \( w_3 \) for the objective functions \( f_1(x), f_2(x) \) and \( f_3(x) \) respectively. Considering VoN as the application with higher priority, followed by video streaming and lastly FTP, weights are defined such that \( w_1 > w_2 > w_3 \). For improved readability, results for a particular choice of weights \((w_1 = 0.45, w_2 = 0.35 \text{ and } w_3 = 0.2)\) are presented. While other combination of weights could be picked as an example, the particular choice above has a reasonably fair distribution of values for \( w_1, w_2 \) and \( w_3 \), while still attending the aforementioned condition.

In Subsection VII-B, the analysis is performed as a MOP using Pareto fronts \( P \in X \). Even though this analysis is not dependent on weights as defined for the weighted sum method in Subsection VII-A, the weights are mentioned also in this section for the sake of comparison with MOP analysis using Pareto fronts.

A. FS evaluation for QoS

Tables IV-VI show, for each scenario, the quality values of voice, video and data, with \( f_1', f_2' \) and \( f_3' \) representing the truncated objective function values without normalization \( (f_2' \text{ in Mbps}) \). Those values are used to give further insight about what are the acceptable values for each quality. But note that \( f_1' = f_3 \) since data quality is already computed in percentage of satisfied interests. Moreover, \( F_i^{\text{max}} \) represents the optimal solution for the weighted sum method considering the weights \( w_1 = 0.45, w_2 = 0.35 \text{ and } w_3 = 0.2 \). Similarly, \( F_N^{\text{max}} \) represents the solution that maximizes \( f_N \) (which can be obtained by \( w_N = 1 \) and 0 for the other weights). In Table IV, equal values are present since the difference between them is smaller than the number of digits used in the Table.

### TABLE IV: Quality values without normalization for GARRB.

<table>
<thead>
<tr>
<th>FS Triple</th>
<th>( f_1' )</th>
<th>( f_2' )</th>
<th>( f_3' )</th>
<th>( F_1^{\text{max}} )</th>
<th>( F_2^{\text{max}} )</th>
<th>( F_3^{\text{max}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR/RF/IF</td>
<td>4.225</td>
<td>1.285</td>
<td>0.990</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR/RF/RA</td>
<td>4.225</td>
<td>1.289</td>
<td>0.987</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OMP/IF/RF</td>
<td>4.184</td>
<td>1.293</td>
<td>0.994</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OMP/IF/IF</td>
<td>4.155</td>
<td>1.148</td>
<td>0.994</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As it can be seen from the optimal solutions \( F_i^{\text{max}} \) in Tables IV-VI, IPI is always present as the forwarding strategy for FTP.
with reasonable performance for all the objective functions considering the maximum ones achieved from $f_1^{max}$ in each scenario. Nevertheless, different weights could be adopted to derive $F^{max}s$. Thus, we discuss how other weights could affect the investigated applications and if IPI would still be present as the forwarding strategy for FTP.

A higher $w_1$ would push the optimality to $f_1^{max}$, but with considerable degradation in video quality regarding GBN scenario. Comparing $f_2$ in Table VI for $F_s^{max}$ and $F_s^{max}$, we can see that video quality in the former case represents a video played with at least the base quality level of the video layers ($L_0$). But in the latter case, a video is watched with at maximum $L_0$ and possibly lots of stalling times since it is around 27% smaller than the minimum value defined in Table III. For EON scenario, the $f_1^{max}$ solution does not have any great decrease in quality values and also has IPI forwarding FTP traffic. In GARRB scenario, gains from $f_1^{max}$ in $f_2$ are roughly negligible, while some small degradation occurs in $f_2$ and $f_3$ in this case.

A higher $w_2$ would push the optimality to $f_2^{max}$, but voice quality for EON and GBN scenario would be less than 3.10, which is considered, according to Table II, that nearly all users are dissatisfied with VoN performance. In GARRB scenario, $f_1$ has a small degradation, but also a little increase in $f_2$ and $f_3$, thus it could still be a reasonable FS triple choice and which also has IPI forwarding FTP traffic.

Considering $f_3^{max}$ our optimal solution, derived from a higher $w_3$, we would have a considerable decrease of video quality for GBN scenario, degradation of video and voice quality for EON scenario, and also some reduction for voice and video quality in GARRB scenario. These service quality decreases do not favor the adoption of $f_3^{max}$ as a suitable solution, since $F_s^{max}$ can have a better performance.

Therefore, regarding $F_s^{max}$ and attained $f_3$ values, it can be verified that IPI is the most suitable forwarding strategy for low priority services, such as FTP, with the possibility to increase the quality of this service without great impact in high priority traffic.

Concerning VoN application, FS used in $F_s^{max}s$ and $f_1^{max}$ are basically BR and MCS. BR is suitable for VoN since it forwards packets always by the same interface if it is not an interest retransmission, reducing jitter. By multicasting interests, MCS can also achieve great voice quality, but it can significantly impact other FS operation, as discussed in Subsection VII-B. In fact, this MCS feature prevents its presence as an optimal solution for video or data quality, given that those applications have greater packet size, and forwarding them with MCS would severly impact other application performances. RFA is not present as the forwarding strategy for VoN application, which is already expected, since RFA can dynamically change forwarding interface for every arriving interest, causing significant jitter in VoN packets.

Video streaming application can be evaluated regarding FS used in $F_s^{max}s$ and $f_2^{max}$. RFA is mostly used since, besides dynamically change its forwarding interface, it tends to use the less loaded (or faster) interface. Thus, even if this feature might increase jitter, that would affect its use for applications such as VoN, it is a suitable feature for video streaming, since packet bursts may arrive, requiring to be satisfied in the lower possible time. OMP-IF sends interests through the lowest delay interface, which could also favor video streaming, but does not change the interface as fast as RFA, i.e., after it had discovered proper forwarding paths they will be used until this forwarding information times out, and discovery process is triggered again. BR, due to the frequent use of only one interface for transmission, does not cope with traffic bursts present in video streaming. MCS, as previously said, can strongly impact other used FS, which does not make it suitable for video streaming application.

### B. Impact of FS coexistence

Figure 6 depicts the normalized values for voice, video and data quality for the FS triples. Regarding axis limits in Figure 6, data quality, for all the FS combinations, is always above a minimum value of around 50%. In contrast with these results, voice and video minimum values can reach 20%. These lower bounds are predictable results, since VoN and video streaming are applications with more constraints, specially VoN, due to low latency and jitter requirements, while it is considerably affected by the FS triple choice and also affects depicted Pareto surfaces. Since these surfaces represent the optimal values for the triple voice, video and data quality, they tend to have at least one high-quality value (e.g., close to 100%). But since FTP is the less constrained application, most of the Pareto front values are concentrated in data quality near 100%, while the other two applications have a bigger span of values in the surface.

In Figure 7, the normalized Pareto front values of FS triples are depicted. As it is already expected, weighted sum solutions are also Pareto-optimal solutions. But despite $F_s^{max}s$, other solutions from $P$ have considerable performance. Taking 3.6 and 0.640 Mbps from Tables II-III as a reference of acceptable values, which can be normalized to 0.8 and 0.463 respectively, we can analyze which optimal solutions were able to attend all application qualities reasonably (i.e. equal or above these normalized values) and if IPI is used on them.

In this manner, for GBN scenario, acceptable solutions are roughly MCS/RFA/RFA, MCS/IPI/RFA and MCS/IPI/IPI, which is the $F_s^{max}$. Although, the last two solutions differ only in the forwarding strategy used for FTP traffic, when using IPI for FTP traffic in the latter case (MCS/IPI/IPI), even if resulting in a slightly decrease in video quality of 1.8%, data

### TABLE V: Quality values without normalization for EON

<table>
<thead>
<tr>
<th>FS Triple</th>
<th>$f_1$</th>
<th>$f_2'$</th>
<th>$f_3'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1^{max}$</td>
<td>3.535</td>
<td>1.095</td>
<td>0.934</td>
</tr>
<tr>
<td>$F_1^{max}$</td>
<td>3.622</td>
<td>1.031</td>
<td>0.951</td>
</tr>
<tr>
<td>$F_2^{max}$</td>
<td>2.82</td>
<td>1.128</td>
<td>0.99</td>
</tr>
<tr>
<td>$F_3^{max}$</td>
<td>3.096</td>
<td>0.589</td>
<td>0.972</td>
</tr>
</tbody>
</table>

### TABLE VI: Quality values without normalization for GBN

<table>
<thead>
<tr>
<th>FS Triple</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>$f_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1^{max}$</td>
<td>3.538</td>
<td>0.659</td>
<td>0.904</td>
</tr>
<tr>
<td>$F_1^{max}$</td>
<td>3.84</td>
<td>0.464</td>
<td>0.886</td>
</tr>
<tr>
<td>$F_2^{max}$</td>
<td>2.774</td>
<td>0.794</td>
<td>0.709</td>
</tr>
<tr>
<td>$F_3^{max}$</td>
<td>3.605</td>
<td>0.485</td>
<td>0.932</td>
</tr>
</tbody>
</table>
quality is increased by 4.3%. When MCS/IPI/IPI is compared to MCS/RFA/RFA, the former has the data quality 4.7% higher but video quality 8.6% lower. Despite lower video quality, both could be considered $F^{\text{max}}_S$ depending on chosen weights in Subsection VII-A - a shift in the weights for $w_1 = 0.45$, $w_2 = 0.4$, $w_3 = 0.15$ would shift $F^{\text{max}}_S$ to MCS/RFA/RFA, while $w_1 = 0.6$, $w_2 = 0.3$, $w_3 = 0.1$ would shift $F^{\text{max}}_S$ back to MCS/IPI/IPI - which makes both acceptable solutions.

For EON scenario, acceptable solutions are roughly MCS/RFA/IPI, MCS/IPI/IPI, OMP-IF/RFA/IPI, OMP-IF/OMP-IF/IPI, OMP-IF/IPI/IPI and OMP-IF/RFA/RFA. These solutions have similar results for all application qualities, and all have IPI as the forwarding strategy for FTP, except for OMP-IF/RFA/RFA. This solution, however, has the lowest voice quality among the considered ones. Although having a high video quality, the fact that $w_1 > w_2$ does not favor OMP-IF/RFA/RFA. Nevertheless, by having $w_2 >> w_1$, the optimality is actually pushed to BR/RFA/IPI solution.

For GARRB scenario, the Pareto-optimal solutions attend the application quality constraints and were already discussed in Subsection VII-A, except for BR/IPI/IPI, which has mostly lower quality values compared to the other solutions.

Besides IPI, another relevant aspect of the $P$ set in Figure 7 is the overall interaction between FS. As it can be remarked, none of the solutions include the use of a single FS for all traffics, reinforcing that although FS are usually evaluated separately, their use in conjunction can result in a higher quality for the investigated traffics.

MCS, for example, has generally a performance below the thresholds from Tables II-III, as MCS/MCS/MCS for GBN scenario with voice/video/data quality as 1.876/0.474 Mbps/0.511. But used with other FS, as MCS/IPI/RFA for the same scenario, it can achieve 3.584/0.672 Mbps/0.863 which can reasonably attend application quality constraints.

Moreover, among optimal solutions, one forwarding strategy is typically used only for one service. As said before, MCS is only present for VoN traffic, due to its high impact on other FS. In cases where BR or OMP-IF are used for more than one service, it results in poor service quality, e.g., for video application with BR/BR/RFA and OMP-IF/OMP-IF/IPI in EON scenario. When RFA is used for more than one traffic, it does not strongly decrease service qualities, i.e., below thresholds from Tables II-III. This is due to the load balancing behavior of RFA, while also dynamically choosing the forwarding interface for each incoming interest. The same can be said of IPI due to Expression 3 in the algorithm, which makes it perform as RFA for high congestion levels.

VIII. CONCLUSION

This work addressed the operation of FS for QoS with three different applications - VoN, video streaming and FTP, by the use of simulations performed in ndnSIM. These applications were evaluated concerning the adopted forwarding strategy for each traffic in a MOP context, where the improvement in a given service quality could degrade others. Furthermore, this work described the development of a new forwarding strategy, called IPI, which is adapted for low priority applications.

Regarding the evaluation of FS MCS, BR, RFA and OMP-IF, it was possible to attest that MCS and BR were the most suitable forwarding strategy for VoN. Video streaming was better served with RFA, while FTP could be better served mostly with IPI. From which we can also deduce that IPI strategy can be efficiently used for low priority traffics, to not increase resource competition for higher priority applications.

Another conclusion from this work concerns the use of different FS for each application for performance improvement. MCS has proven to considerably interfere with the operation of other FS, indicating it must be carefully used, e.g., only...
is concluded that the proper coupling of FS is a relevant feature to generate higher results concerning the achieved quality for each investigated application, making FS a suitable prioritization feature in ICN networks.

**References**


