ICN 2022
The Twenty-First International Conference on Networks

Optimization of the Virtual Network Function Reconfiguration Plan in 5G Network Slicing

Hanane BIALLACH[1,2]
Mustapha BOUHTOU [1]
Dritan NACE[2]

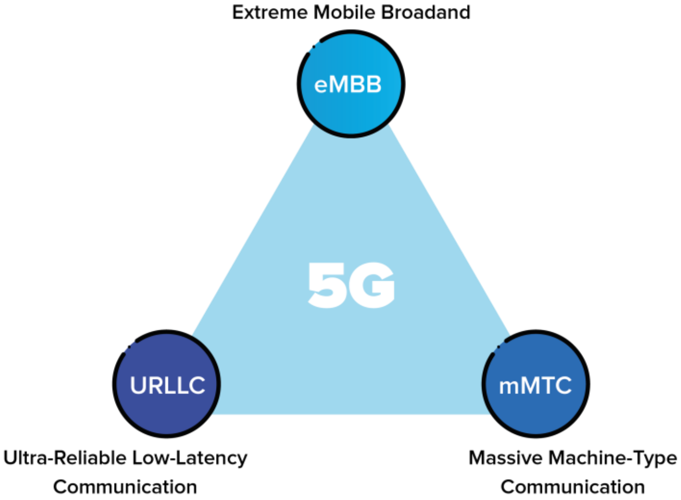
1 Orange Innovation, Châtillon, France.
[2] Heudiasyc Lab, Compiègne, France.

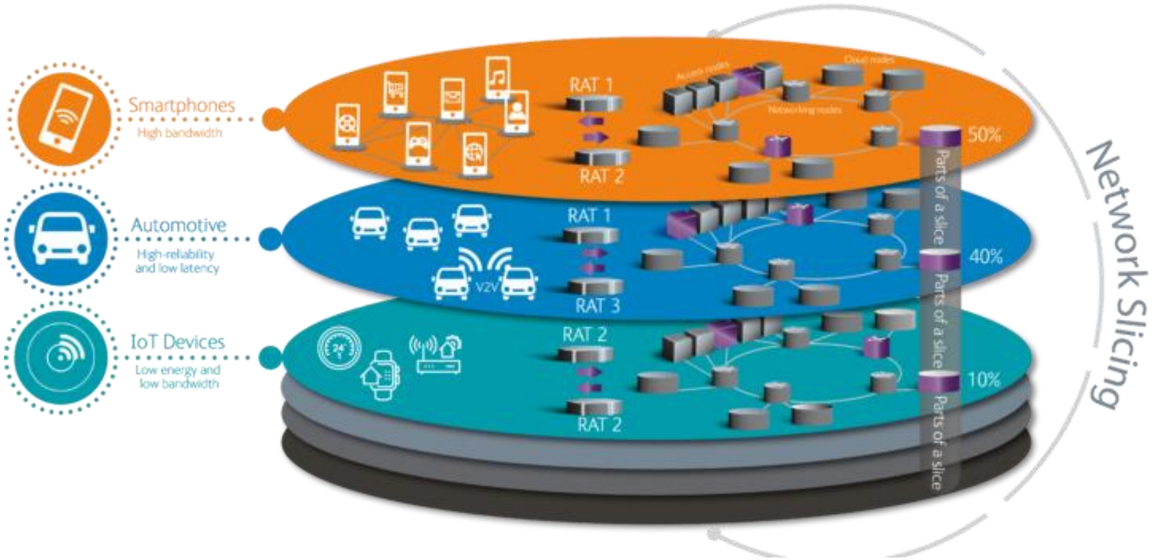
Contact email: hanane.biallach@hds.utc.fr

Speaker : Hanane BIALLACH

Hanane BIALLACH received the engineering degree in Information Technology Management from the National Institute of Posts and Telecommunications of Rabat, Morocco. The master's degree in computer science from MINES of Saint Etienne, France. She is currently pursuing the PhD degree with Orange Innovation, châtillon, France, in collaboration with Heudiasyc lab of Compiègne, France. Her research interest include 5G network slicing, VNF reconfiguration and migration, optimization.

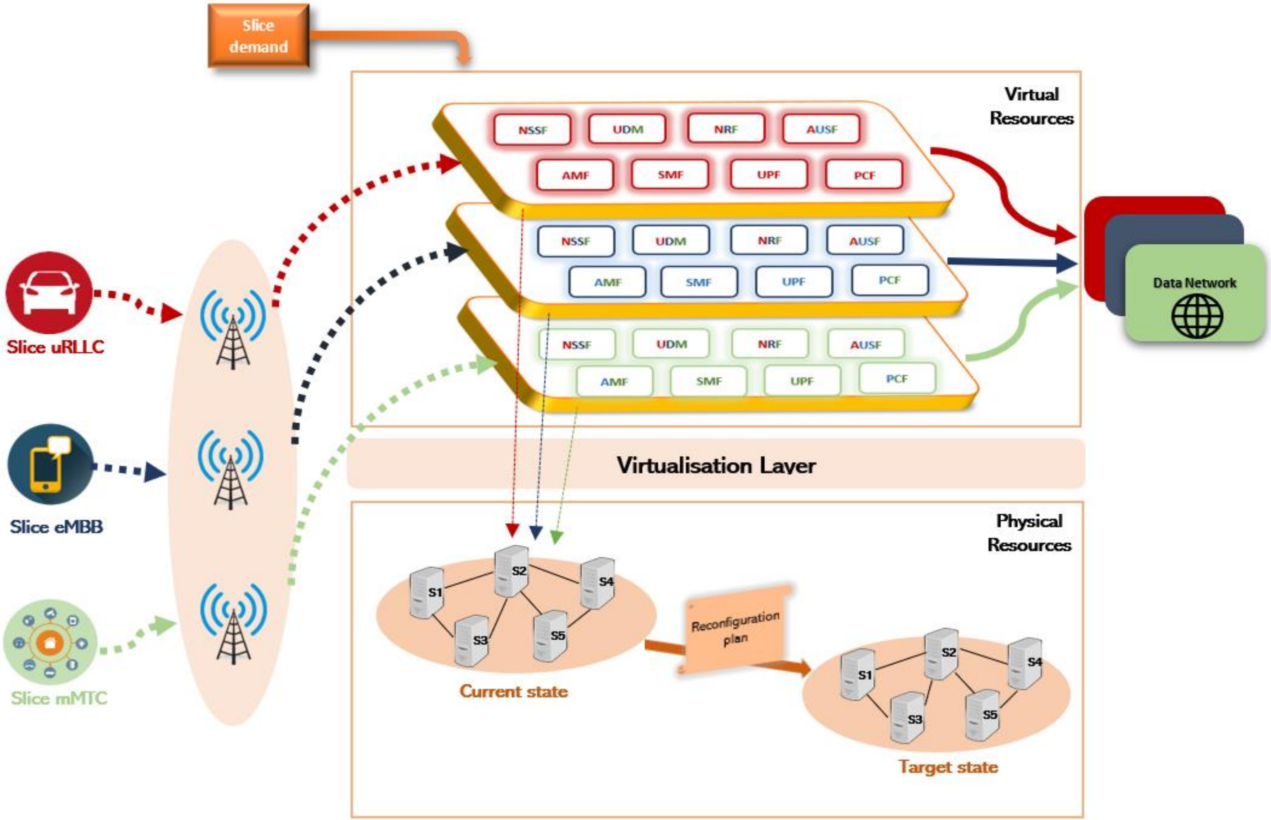
5G network slicing





❖ Future networks will support a multitude of services thanks to slicing. [1][2][3]

2. Problematic



NSSF: *Network Slice Selection Function*, **UDM**: *Unified Data Management*, **NRF**: *Network Repository Function*, **AUSF**: *Authentication Server Function*, **AMF**: *Access and Mobility Management Function*, **SMF**: *Session Management Function*, **UPF**: *User Plane Function*, **PCF**: *Policy Control Function*

- The reconfiguration of 5G network slices is a reallocation of the Network Function Virtualization (NFV) to adapt the utilization of network resources to the occurred changes.

- The VNF migrations are performed by two types of migrations:
  - **Hot migration** the running VNFs are moved between the source and target servers without disconnecting the service or application. [4]
  - **Cold migration** the VNFs are moved between servers by powering off the VNF on the source server, moving it to the target server then powering it back up on the target server.

Objective function / Metrics

- Service interruption (interruption duration)
- Duration of total migration process
- Slice Service Level Agreement (SLA) penalities (availability..)

Constraints

- Server Capacity
- VNF migration duration
- Dedicated and shared VNF

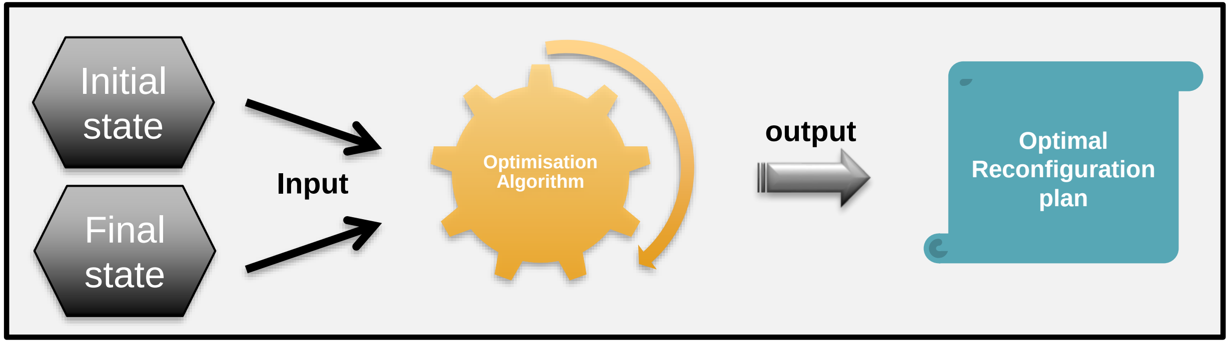
VNF Migration mode:

- Hot migration
- Cold migration

VNF Reconfiguration problem

- ❖ A lot of studies tackle the NP-Hard problem of VNF/VM placement [5][6][7][8][9] .. But, how to execute the migrations to reach a given placement is even less tackeled?
- ❖ Few research studies are dealing with the VNF reconfiguration problem
- ❖ The problem of dynamic management and optimization of reconfiguration plans for slicing has been less treated !

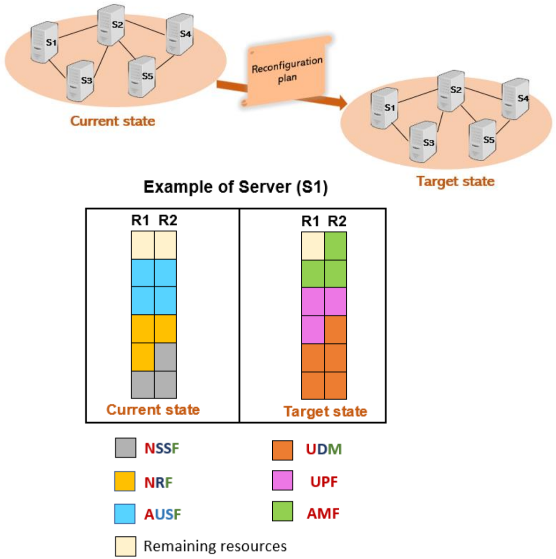
Main scope



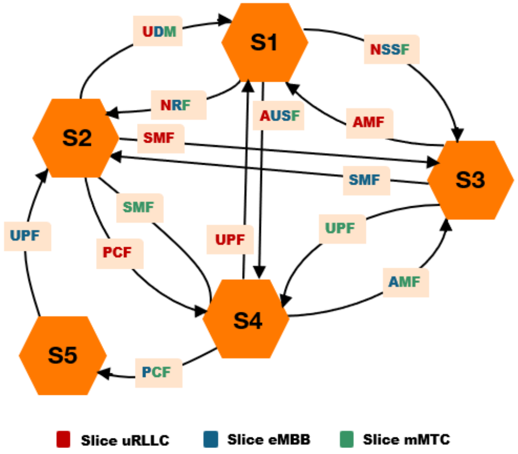
❖ The problem of VNF reconfiguration can be seen as :

a) A multi-dimension multiknapsack reconfiguration problem

(The multidimensionality stands on the number of resources concerned(i.e., CPU and memory), while the knapsack aspect is captured by the servers, each of them representing a knapsack.)



b) Ordering edges of a graph



❖ If the graph is acyclic, there is a feasible solution without interruption [10] and can be found in polynomial time [11] using Topological Sorting (TS) algorithm.

❖ If the graph is cyclic, the solution is not necessarily feasible without interruptions.

The idea : Modelize the state dynamics of the server during stages of VNFs migrations

The problem is NP-Hard, it is modeled by integer linear programming:

- **The objective function**: Minimize the migration and interruption duration.
- **Decision variables**: are binary variables, indicating the stage where VNF_i is migrated or interrupted.
- **Capacity constraint**: we calculate the capacity of each server s at each stage k according to the interruption / the migration of VNF.
- **Integrity constraint**: insures that each VNF can be migrated only in the destination server, and it can be interrupted only in the source server.
- **VNF Migration duration constraint**: finds the maximum migration duration that should be minimized.
- **Interruption duration constraint**: The interruption duration is considered as the number of stages between the VNF interruption and VNF migration. In live migration the interruption duration is negligible. In cold migration, the VNF interruption is performed at least in one stage. Both of live and cold migration are taken into consideration.

ILP model

TABLE I

TABLE OF NOTATIONS

Notation	Description
N	number of stages
N_v	number of VNFs
N_s	number of servers
k	order / stage of the reconfiguration
x_{ik}	a binary variable indicating that VNF_i is migrated in order k
y_{ik}	a binary variable indicating the stage where VNF_i is interrupted in source host
$O_{(s)}$	set of VNFs originating from server s
$D_{(s)}$	set of VNFs targeting server s
C_s^k	represents the residual CPU capacity of server s in stage k
R_s^k	represents the residual RAM capacity of server s in stage k
cap_i^{cpu}	represents the occupied CPU capacity of VNF_i
cap_i^{ram}	represents the occupied RAM capacity of VNF_i
δ_i	represents the interruption duration of VNF_i
T	represents the migration duration of a given VNF
β_i	represents the cost of service interruption, which is the SLA availability of each VNF_i
α	represents the migration cost of all VNFs

$$x_{ik} = \begin{cases} 1, & \text{if the } VNF_i \text{ is migrated in stage } k\,; \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

$$y_{ik} = \begin{cases} 1, & \text{if the } VNF_i \text{ is interrupted in stage } k\,; \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

$$\sum_{k=1}^{N_v} x_{ik} = 1 \,;\, \forall i \in \{1,..,N\} \quad (3)$$

$$\sum_{k=1}^{N_v} y_{ik} = 1 \,;\, \forall i \in \{1,..,N\} \quad (4)$$

$\forall s \in \{1,..,N_s\} \,;\, \forall k \in \{1,..,N\}\,;$

$$C_s^k - \sum_{i \epsilon D(s)} x_{ik} cap_i^{cpu} + \sum_{i \epsilon O(s)} y_{ik} cap_i^{cpu} = C_s^{k+1} \quad (5)$$

$\forall s \in \{1,..,N_s\} \,;\, \forall k \in \{1,..,N\}\,;$

$$R_s^k - \sum_{i \epsilon D(s)} x_{ik} cap_i^{ram} + \sum_{i \epsilon O(s)} y_{ik} cap_i^{ram} = R_s^{k+1} \quad (6)$$

$$C_s^k \geq 0 \,;\, \forall s \in \{1,..,N_s\} \,;\, \forall k \in \{1,..,N\} \quad (7)$$

$$R_s^k \geq 0 \,;\, \forall s \in \{1,..,N_s\} \,;\, \forall k \in \{1,..,N\} \quad (8)$$

$$\sum_{k=1}^{N_v} k y_{ik} = \sum_{k=1}^{N_v} k x_{ik} \,;\, \forall i \in \{1,..,N\} \quad (9a)$$

$$\sum_{k=1}^{N_v} k y_{ik} \leq \sum_{k=1}^{N_v} k x_{ik} \,;\, \forall i \in \{1,..,N\} \quad (9b)$$

$$\sum_{k=1}^{N_v} k y_{ik} = \sum_{k=1}^{N_v} k x_{ik} + 1 \,;\, \forall i \in \{1,..,N\} \quad (9c)$$

$$\sum_{k=1}^{N_v} k y_{ik} \leq \sum_{k=1}^{N_v} k x_{ik} + 1 \,;\, \forall i \in \{1,..,N\} \quad (9d)$$

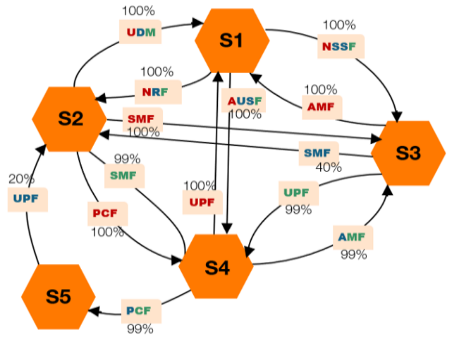
$$\delta_i = (\sum_{k=1}^{N_v} k x_{ik} + 1) - (\sum_{k=1}^{N} k y_{ik}) \quad (10)$$

$$\sum_{k=1}^{N_v} k x_{ik} \leq T \,;\, \forall i \in \{1,..,N\} \quad (11)$$

$$min(\sum_{i=1}^{N_v} \beta_i \delta_i + \alpha T) \quad (12)$$

5. Experimental Results

Evaluation according to the nature of slices



Start Step	Actions	Is interrupted ?	Interruption duration
1	Migrate **NSSF** to S3	NO	--
1	Migrate **AUSF** to S4	NO	--
1	Migrate **SMF** to S3	NO	--
1	Migrate **UPF** to S1	NO	--
1	Migrate **PCF** to S5	NO	--
2	Migrate **NRF** to S2	NO	--
2	Migrate **UPF** to S4	NO	--
2	Migrate **UDM** to S1	NO	--
2	Migrate **PCF** to S4	NO	--
3	Migrate **AMF** to S1	NO	--
3	Migrate **SMF** to S2	YES	3
3	Migrate **AMF** to S3	YES	1
3	Migrate **UPF** to S2	YES	3

Fig.1 **(a).** The VNFs migration with the SLA availability according to each service slice.

Fig.1 **(b).** The reconfiguration plan for all slices.

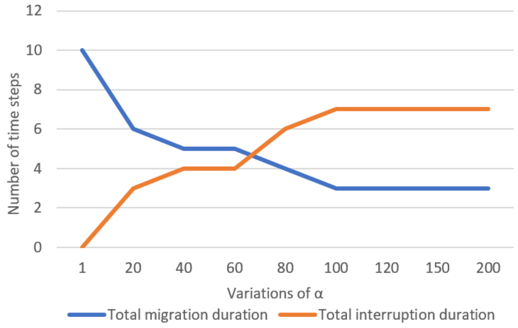


Fig.2 (a). The total migration and interruption according to α variations.

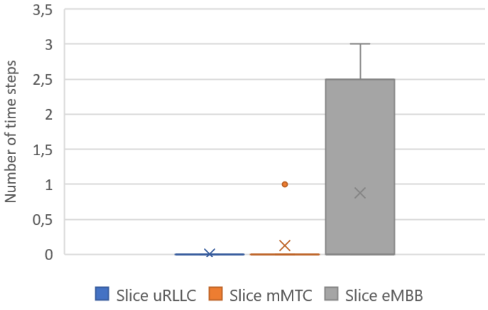


Fig.2 **(b).** The interruption duration of each slice.

- ❖ All the experiments were conducted on a machine with Core i7-6600U CPU and 16 Go of RAM
- ❖ Datasets are randomly generated with different sizes of graphs using NetworkX of python
- ❖ The ILP model is solved using the CPLEX Optimization Studio V12.8

$$min(\sum_{i=1}^{N_v} \beta_i \delta_i + \alpha T) \qquad \beta_i \begin{cases} = 100\% & \text{High availability} \\ \geq 99\% & \text{Average availability} \\ < 99\% & \text{Low availability} \end{cases}$$

- ○ The ILP model finds the optimal solution that minimizes both of migration and interruption duration.
- ○ The ILP model takes into consideration the availability of each slice while minimizing the interruption duration.

Evaluation according to the nature of datasets – **Acyclic graph**

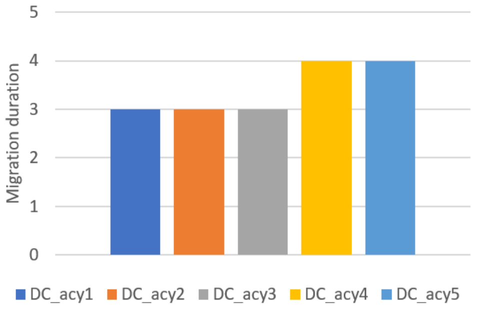


Fig.3 (a). Total migration duration.

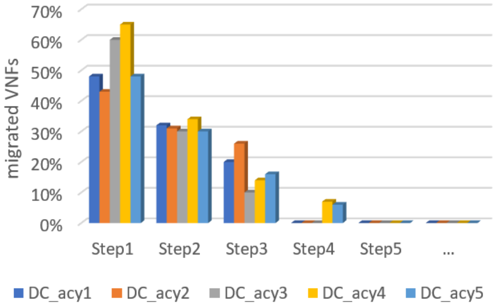


Fig.3 (b). Percentage of migrated VNFs per steps of migrations.

TABLE II
DATASETS OF ACYCLIC GRAPHS

Instances	Servers	VNFs	Slices
DC-acy1	10	25	6
DC-acy2	20	35	11
DC-acy3	40	60	12
DC-acy4	50	120	24
DC-acy5	80	150	35

TABLE IV
COMPARISON BETWEEN THE ILP MODEL AND TS ALGORITHM FOR ACYCLIC GRAPH

Instances	**ILP:** Execution time (s)	**ILP:** Best objective	**TS:** Execution time (s)	**TS:** Best objective
DC-acy1	0.33	3	0.000532	25
DC-acy2	1.06	3	0.000324	35
DC-acy3	2.08	3	0.000949	60
DC-acy4	17.76	4	0.001346	80
DC-acy5	36.19	4	0.001257	150

- The ILP model solves the VNF reconfiguration problem without interruption and with 0% of optimality gap.
- The ILP finds the optimal solution that minimizes the migration duration while the Topological Sorting (TS) algorithm [12][13] finds a feasible solution without taking into consideration the migration duration.
- The ILP gives a solution where the VNFs are migrated from the early steps and in parallel as long as possible, and with minimum migration duration.

Evaluation according to the nature of datasets – **Cyclic graph**

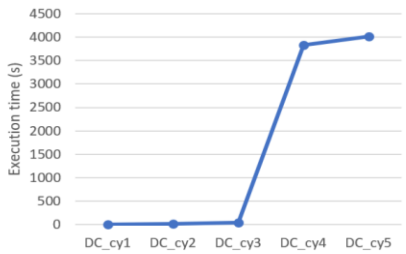


Fig.4 (a). Execution time variations.

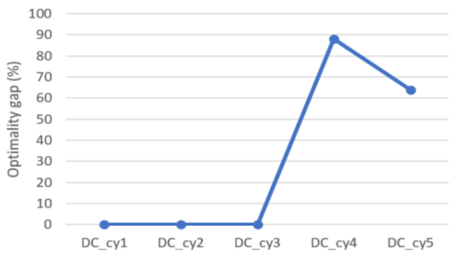


Fig.4 (a). Optimality gap variations.

TABLE III			
Instances	Servers	VNFs	Slices
DC-cy1	10	30	8
DC-cy2	20	45	12
DC-cy3	40	70	15
DC-cy4	50	120	25
DC-cy5	80	146	32

TABLE III
DATASETS OF CYCLIC GRAPHS

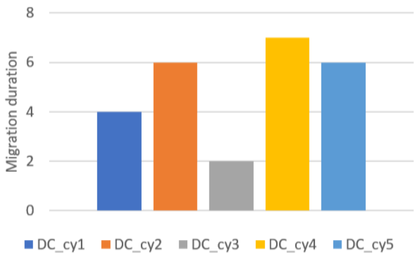


Fig.5 (a). Total migration duration.

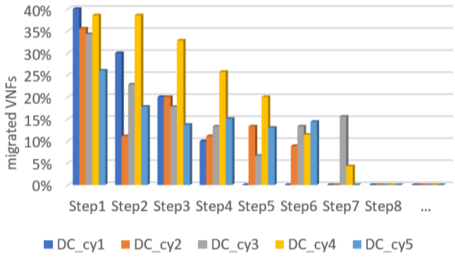


Fig.5 (b). Percentage of migrated VNFs per steps of migrations.

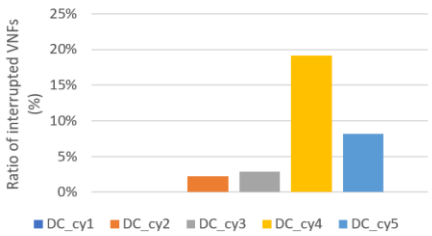


Fig.6 (a). Ratio of interrupted VNFs.

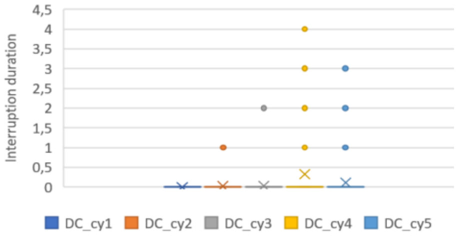


Fig.6 (b). Interruption duration for each instance.

- The execution time increases for DC-cy4 and Dc-cy5 instances due to the np-hardness of the problem.
- Optimality gap are often at 0% except in case of DC-cy4 and Dc-cy5 instances which need more time to find an optimal solution.
- The ILP provides an interesting solution in terms of migration duration and VNF interruption.
- The ILP model complexity depends strongly on the presence of cycles.

Conclusion

- We have proposed an ILP-based solution for the problem of slice reconfiguration in the context of 5G networks.
- We evaluate the proposed model according to the service importance taking into consideration the SLA availability metric, and according to the nature of datasets whether it is an acyclic or cyclic graph.
- The evaluation results show that the ILP model yields good solutions, in terms of minimizing the total migration and VNF interruption duration, and respecting the Slice SLA availability.

Future work

- We plan to propose a heuristic based on topological sorting algorithm in order to improve the convergence time and allow dealing with larger instances.

References

[1] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, “Network slicing in 5g: Survey and challenges,” IEEE Communications Magazine, vol. 55, no. 5, pp. 94–100, 2017.

[2] I. Afolabi, T. Taleb, K. Samdanis, A. Ksentini, and H. Flinck, “Network slicing and softwarization: A survey on principles, enabling technologies, and solutions,” IEEE Communications Surveys Tutorials, vol. 20, no. 3, pp. 2429–2453, 2018.

[3] “5g; procedures for the 5g system (3gpp ts23.502 version 16.5.0 release 16),” (2020-07). [Online]. Available: https://www.etsi.org/deliver/etsi ts/123500 123599/123502/16.05.00 60/ts 123502v160500p.pdf

[4] C. Clark, K. Fraser, S. Hand, and al., “Live migration of virtual machines,” in Proceedings of the 2nd NSDI. USENIX Assoc., 2005.

[5] A. Laghrissi and T. Taleb, “A survey on the placement of virtual resources and virtual network functions,” IEEE Communications Surveys Tutorials, vol. 21, no. 2, pp. 1409–1434, 2019.

[6] X. Li and C. Qian, “A survey of network function placement,” in 2016 13th IEEE Annual Consumer Communications Networking Conference (CCNC), 2016, pp. 948–953.

[7] P. T. A. Quang, A. Bradai, K. D. Singh, G. Picard, and R. Riggio, “Single and multi-domain adaptive allocation algorithms for vnf forwarding graph embedding,” IEEE Transactions on Network and Service Management, vol. 16, no. 1, pp. 98–112, 2019.

[8] O. Soualah, M. Mechtri, C. Ghribi, and D. Zeghlache, “A green vnf-fg embedding algorithm,” in 2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft), 2018, pp. 141–149.

[9] J. J. A. Esteves, A. Boubendir, F. Guillemin, and P. Sens, “Location-based data model for optimized network slice placement,” in 2020 6th IEEE Conference on Network Softwarization (NetSoft), 2020, pp. 404–412

References

[10] Renaud Sirdey, Jacques Carlier, Hervé Kerivin, Dritan Nace. On a resource-constrained scheduling problem with application to distributed systems reconfiguration. European Journal of Operational Research, Elsevier, 2007, 183, pp.546-563

[11] J. Bang-Jensen and G. Gutin. Digraphs. Theory, Algorithms and Applications. Springer-Verlag, 2002.

[12] D. E. Knuth and J. L. Szwarcfiter, “A structured program to generate all topological sorting arrangements,” Information Processing Letters, vol. 2, no. 6, pp. 153–157, 1974.

[13] D. Ajwani, A. Cosgaya-Lozano, and N. Zeh, “A topological sorting algorithm for large graphs,” ACM J. Exp. Algorithmics, vol. 17, sep 2012. [Online]. Available: https://doi.org/10.1145/2133803.2330083