iFogStorC: a Heuristic for Managing IoT Data Replication Storage and Consistency in a Fog Infrastructures

Presented by
Mohammed Islam Naas

Thesis Director
- Jalil Boukhobza

Supervisors
- Philippe Raipin Parvedy
- Laurent Lemarchand

Performance and Scalability of Storage Systems (Per3S)
INRIA Bordeaux, Talence, France.
January 25th 2019.

UBO Université de Bretagne Occidentale

Lab-STICC

orange™

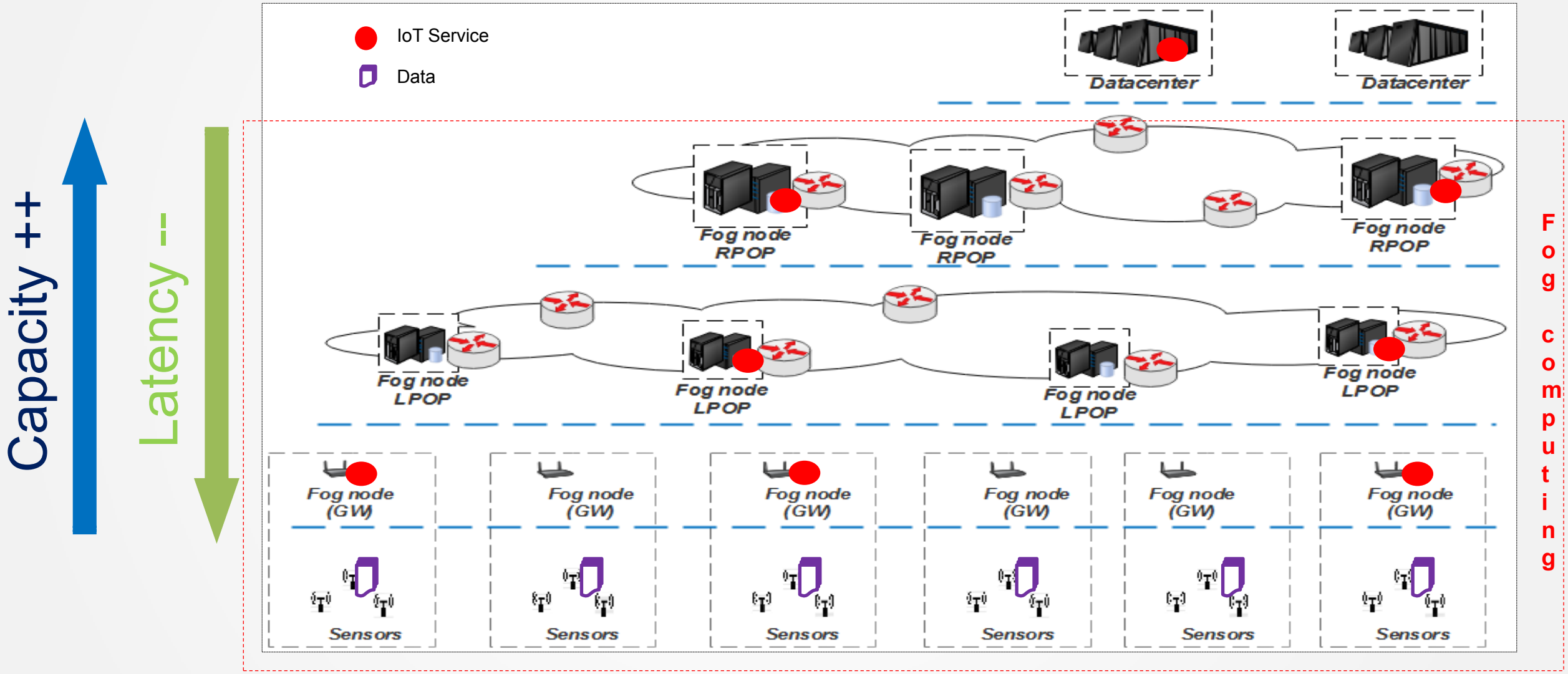
Agenda

- Context
- Problem statement
- Solution
- Results
- Conclusion

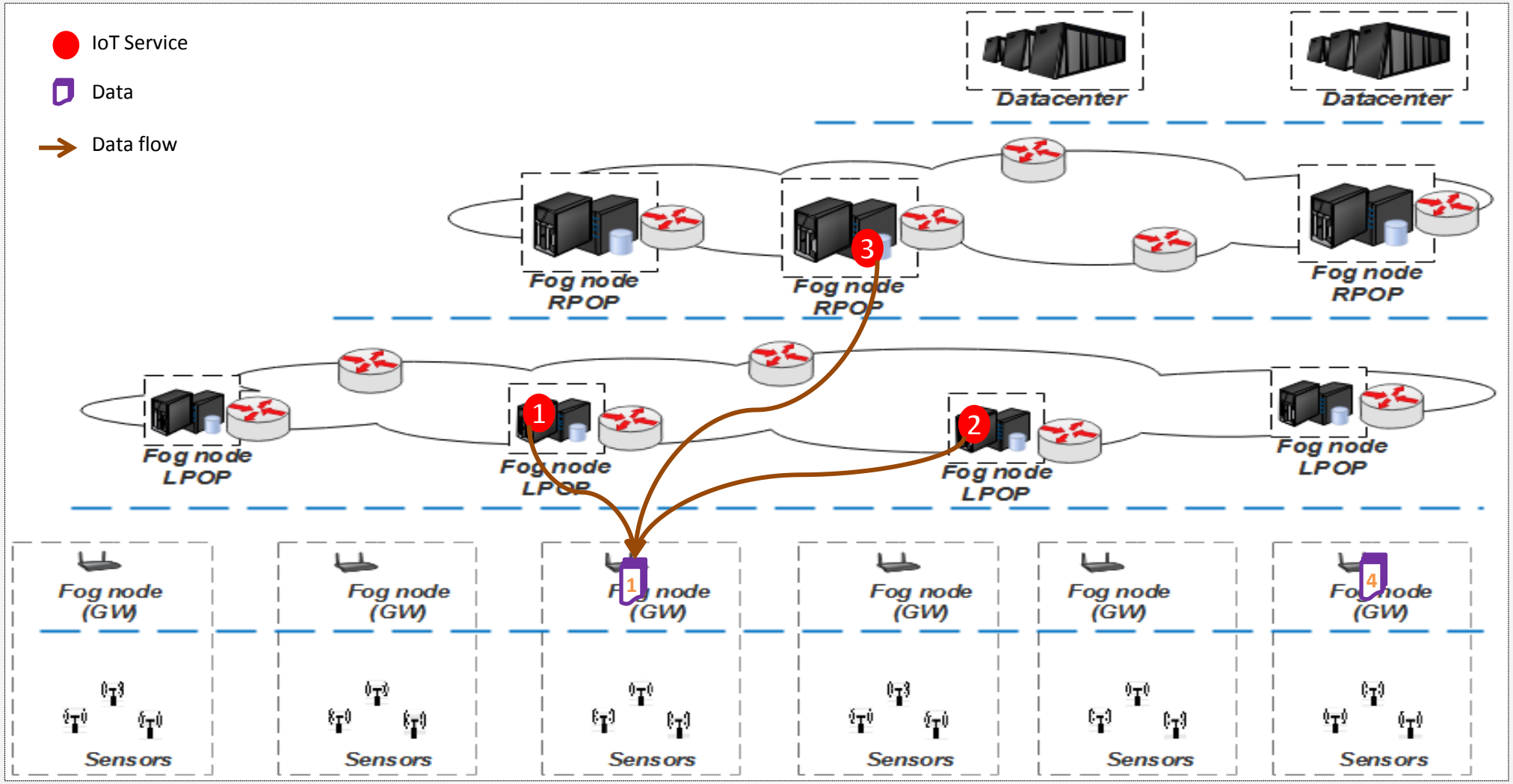
Context: Fog computing

Fog computing: is ”*a highly virtualized platform that provides compute, storage and networking services between end devices and Cloud data-centers, typically, but not exclusively located at the edge of network.*” Bonomi et al.

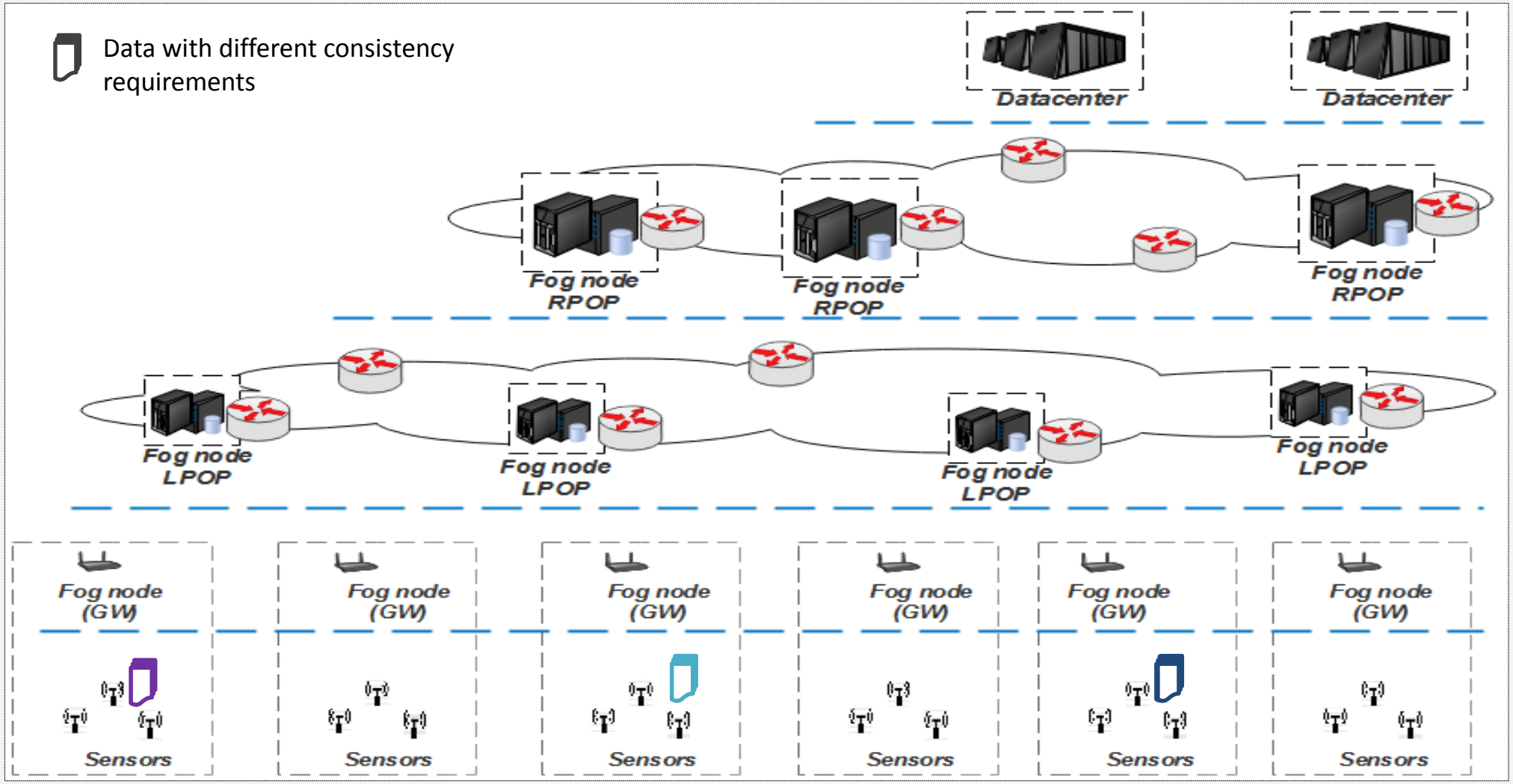
Context: Fog computing



Problem Statement: Data sharing



Problem Statement: Replicas synchronization time

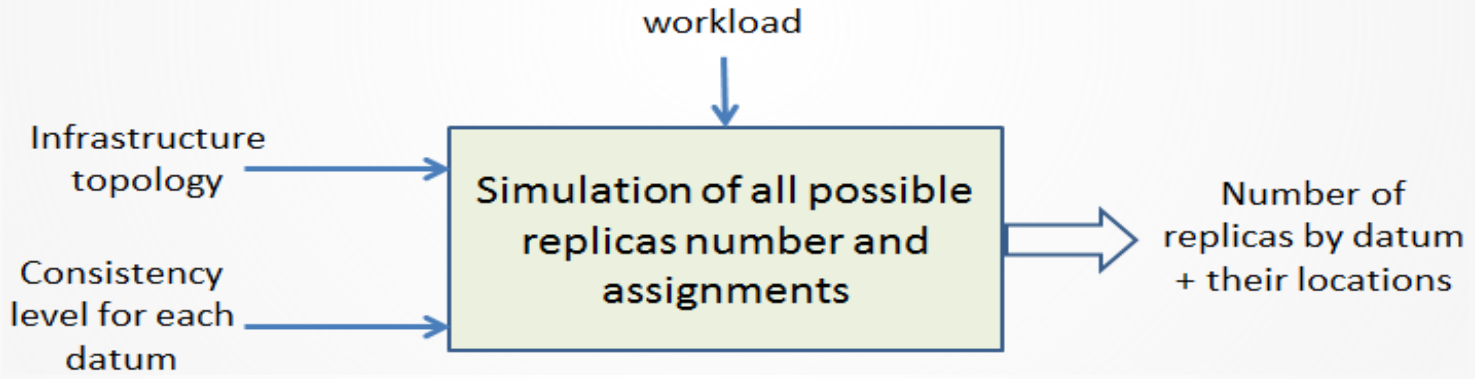


Problem Statement

How can one choose the right number of replicas and place them in order to minimize the overall service latency while respecting a given consistency level ?

Motivation

Exact solution: enumerate all replica placement possibilities.



Complexity: the number of possible assignments= $d \times \left(C_n^{P_{min}} + \cdots + C_n^{P_{max}}\right)$

- d: number of data units
- n: number of Fog nodes
- P_{min}: number of Fog nodes
- P_{max}: number of Fog nodes
- C_n^P: a combination of P from $n, C_n^P = \frac{n!}{P!(n-p)!}$

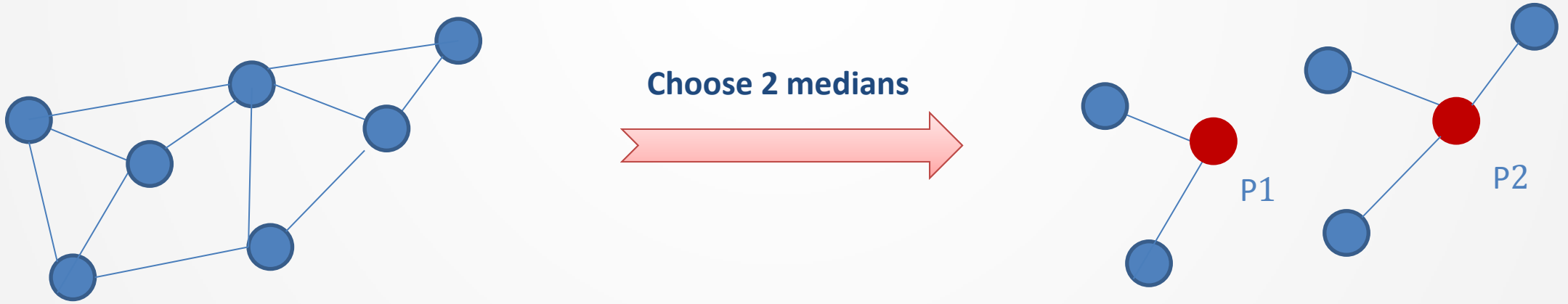
Combinatorial Explosion

- E.g. for $n = 20, P_{min} = 3, P_{max} = 5, d = 20$ ➔ 322,335 possible assignments

Approach: iFogStorC

- Idea: reduce the number of possible assignments/placements
  - ➔ considering only Fog nodes located in the shortest paths between producers and consumers.
  - ➔ In addition, only the P-median nodes are chosen.

- P-median: choose P nodes from a graph so that the sum of all distance between each node of the graph to a median node is minimized.



Approach: iFogStorC

- Algorithm:
- For All Data :
  - Get All shortest paths nodes
  - For $P \in [P_{\min}, P_{\max}]$, number of replicas:
    - Find P-median to place P replicas (using CPLEX)
    - Estimate the latency overhead of this assignment (a micro simulation is done using iFogSim)
  - Choose P with the minimum latency overhead.

- Complexity : the number of possible assignments = $d \times (P_{max} - P_{\min} + 1)$.
- E.g. for $n = 20, P_{\min} = 3, P_{\max} = 5, d = 20$ ➔ 60 possible assignements.

Evaluation: Methodology

- **Metrics of comparison:**
  1. Overall service latency.
  2. Freshness of data.

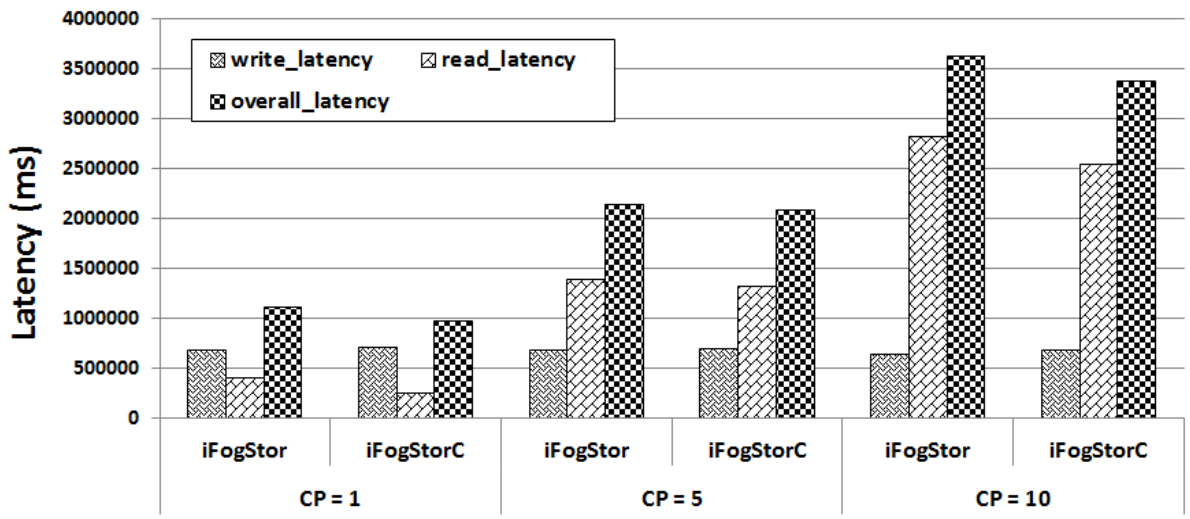
- **Storage strategy:**
  1. **iFogStor**: finds the storage location for data (no replication) minimizing the overall system latency.
  2. **iFogStorC** : our heuristic that finds the number of replicas and their storage location while reducing the system latency and respecting a given consistency level.

- **Data consistency protocol:**
  1. **Strong**: ensures that several replicas have the latest version.
  2. **Eventual**: if there are no writes, the system converts replicas to the last version.

- **Workload:**
  - **Distributed**: data are consumed by nodes from the whole infrastructure.
  - Number of consumers that share the same data by : 1, 5 and 10.

Evaluation: Results

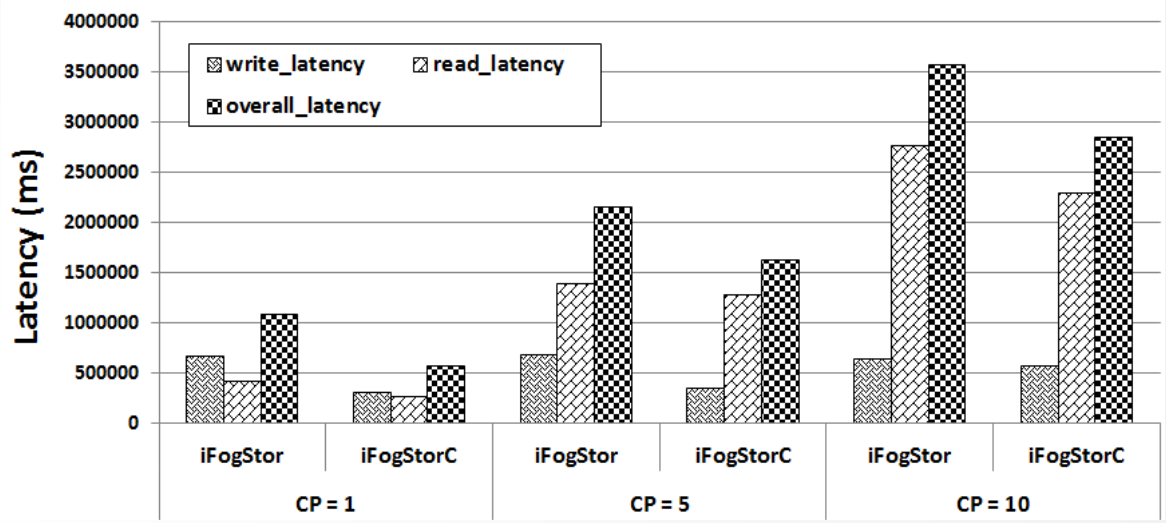
Strong consistency



Read requests stats (CP = 5)

Metric	iFogStor	iFogStorC
Latest version	89.76%	89.96%

Eventual consistency



Read requests stats (CP = 5)

Metric	iFogStor	iFogStorC
Latest version	90.09%	7.34%

Conclusion

Context

- Fog extends Cloud services to the network edge.
- Fog computing has the ability to host IoT applications.

Problem statement

- Data sharing and replica synchronization → degrade overall service latency → replica placement problem.

Contributions

- **iFogStorC**: a heuristic that finds the replicas number and their location to reduce the service latency while respecting a given data consistency level.

Future work:

- Evaluate the execution time of **iFogStorC** for different scenarios.
- Evaluate other workloads.
- Evaluate other consistency protocols.
- Consider other metrics such as nodes availability or energy consumption.
- …

References

[1]: M. I. Naas, P. R. Parvedy, J. Boukhobza, and L. Lemarchand. ifogstor: An iot data placement strategy for fog infrastructure. In 2017 IEEE 1st International Conference on Fog and Edge Computing (ICFEC), volume 00, pages 97–104, 2017.

[2]: M. I. Naas, L. Lemarchand, J. Boukhobza, and P. Raipin. A graph partitioning-based heuristic for runtime iot data placement strategies in a fog infrastructure. In SAC 2018: Symposium on Applied Computing. ACM, 2018.

[3]: M. I. Naas, J. Boukhobza, P. Raipin and L. Lemarchand. An Extension to iFogSim to Enable the Design of Data Placement Strategies. In 2018 IEEE 2nd International Conference on Fog and Edge Computing (ICFEC), Washington DC, USA, 2018.

[4]: Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012, August). Fog computing and its role in the internet of things. In Proceedings of the first edition of the MCC workshop on Mobile cloud computing (pp. 13-16). ACM.

[5]: Vangelis Gazis, Alessandro Leonardi, Kostas Mathioudakis, Konstantinos Sasloglou, Panayotis Kikiras, and Raghuram Sudhaakar. 2015. Components of fog computing in an industrial internet of things context. In Sensing, Communication, and Networking-Workshops (SECON Workshops), 2015 12th Annual IEEE International Conference on. IEEE, 1–6.

[6]: Vernon Turner, John F Gantz, David Reinsel, and Stephen Minton. 2014. The digital universe of opportunities: Rich data and the increasing value of the internet of things. IDC Analyze the Future (2014).

[7]: Ibm ilog cplex optimization toolkit. http://www-03.ibm.com/software/products/en/ibmilogcpleoptistud

[8]: H. Gupta, A. V. Dastjerdi, S. K. Ghosh, and R. Buyya. ifogsim: A toolkit for modeling and simulation of resource management techniques in internet of things, edge and fog computing environments. Technical Report CLOUDS-TR-2016-2, Cloud Computing and Distributed Systems Laboratory, The University of Melbourne, Australia, 2016.

[9]: A. V. Dastjerdi, H. Gupta, R. N. Calheiros, S. K. Ghosh, and R. Buyya. Fog computing: Principals, architectures, and applications. 2016.

[10]: Bermbach, D., & Kuhlenkamp, J. (2013). Consistency in distributed storage systems. In Networked Systems (pp. 175-189). Springer, Berlin, Heidelberg.