

## THE PROBLEM ADDRESSED

Reliable and scalable distributed storage systems should handle efficiently where the data should be stored to minimize both the costs and access latency. In addition, they should provide proper mechanisms for data availability in case of node failures. However, traditional data placement and data availability strategies prove to be costly to properly scale up and provide better data availability.

Babel is a set of software solutions for distributed storage systems that aims reducing storage systems cost and improving user quality of experience. It integrates a solution for data placement using machine learning to predict the suitable data storage placement for each type of information, reducing thus the cost of relocating and optimizing response rate. In addition, Babel deploys an innovative erasure coding technique that optimizes repair bandwidth, I/O and required recovery time.

Created in 1982, the LTCI is characterized by its broad scope in the field of information and communication sciences and technologies. The laboratory covers the full theme of networks, signal processing and digital communications from the hardware to the software and applied mathematics.

## TECHNOLOGY

- Flexible **AI-powered** automated data placement for data storage systems. Using **online machine learning** to optimize data placement.
- Efficient erasure coding to improve **data availability** and failure repair.

## COMPETITIVE ADVANTAGES

- On-the-fly and multi-features data placement
- Scalable and reliable data storage
- Reduced infrastructure cost
- Better quality of experience

Babel allows cost savings, via:

- Optimized usage of expensive hard drives
- Reduced network utilization for data relocation
- Reduced data access latency
- Reduced repair bandwidth and repair I/O
- Decreased recovery time

## APPLICATION

- Scalable and reliable data storage system with high data availability requirements and efficient use of infrastructure cost.
- **Targeted Clients:** Cloud and storage providers, Storage software vendors, Large corporates operating their own IT production, Infrastructure providers...

## DEVELOPMENT STATUS

- Automatic data placement: TRL3/4
- Data recovery: TRL4/5

## INTELLECTUAL PROPERTY

2 patents:

- “Une méthode de placement automatique de données pour les systèmes de stockage distribués”, ongoing, priority date 2022/XX/FF
- “Méthodes et dispositifs de codage et de décodage de données”, FR3080196, priority date 2018/04/13

## INVENTORS & CONTACTS

- Nadia Boukhatem, Full Professor at Télécom Paris,  
[nadia.boukhatem@telecom-paris.fr](mailto:nadia.boukhatem@telecom-paris.fr)
- Hana Baccouch, PhD then Postdoc on the project,  
[h.baccouch@bbstechs.com](mailto:h.baccouch@bbstechs.com)
- TTO: [valorisation.transfert@telecom-paris.fr](mailto:valorisation.transfert@telecom-paris.fr)

## LOOKING FOR

- Co-development with industrial partners
- Deployment of the solutions in a client system
- New use cases