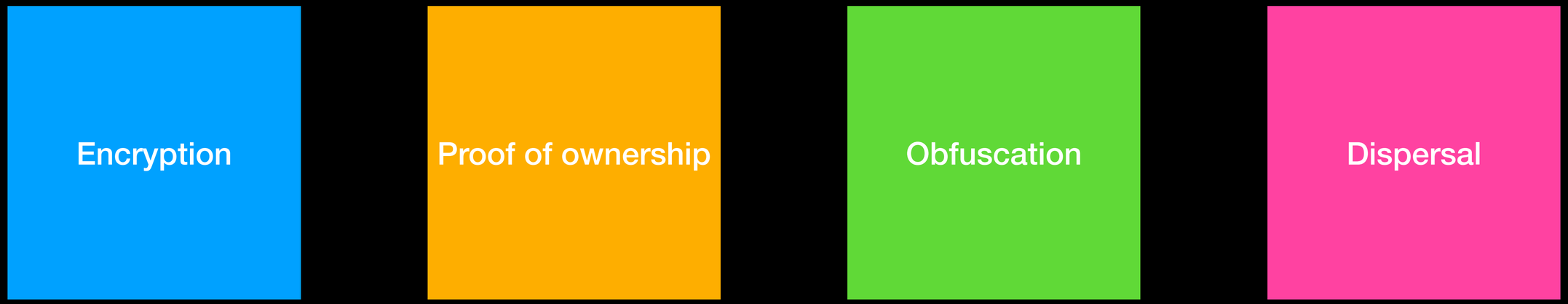
# Data Deduplication Security

## Business Continuity

# Secure Deduplication Solutions

## Shin et al.



# Encryption

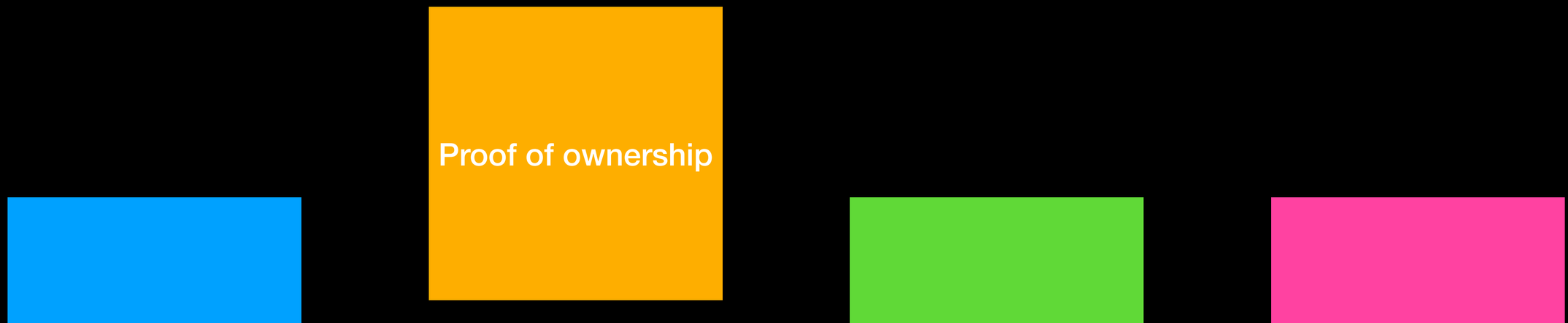
## Secure Deduplication Solutions

- Encryption is an effective control in ensuring the confidentiality, but **traditional encryption approaches can undermine deduplication.**
  - Infrastructure owner will encounter difficulties to identify duplicate data from different cipher texts generated by different users with different keys.
  - Challenging to efficiently store different cipher texts from different users.

# Proof of ownership

## Secure Deduplication Solutions

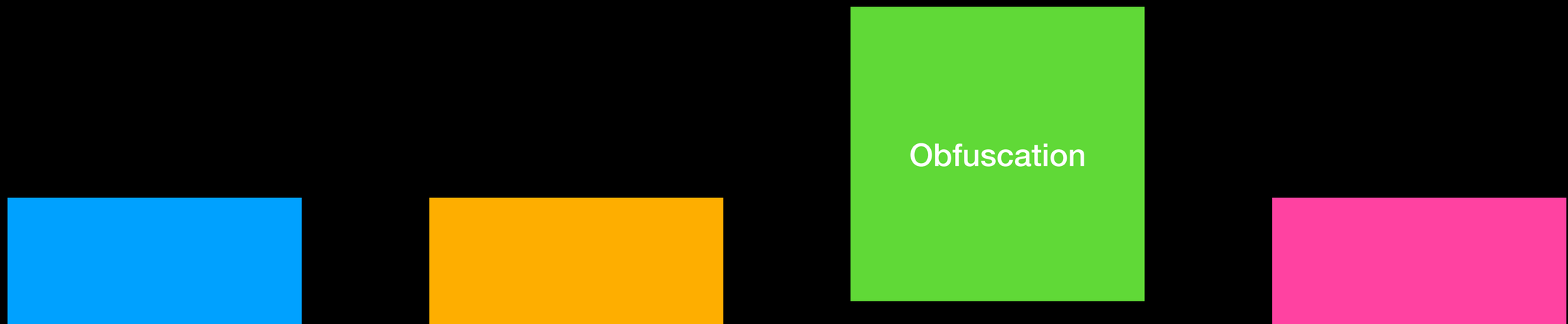
- Proof of ownership (POW) in deduplication can be ownership of the fingerprint of the binary data.
- However, this fingerprint could be easily shared or obtained and so infrastructure providers may want to implement POW protocol.
- POW protocols can be used to confirm that a client is the owner of binary data and it can be retrieved.



# Obfuscation

## Secure Deduplication Solutions

- Deduplication processes can potentially leak information when binary data is not uploaded after binary data is confirmed as already existing on the infrastructure.
- Traffic obfuscation processes can be used to mask and remove the level of information an attacker can infer from data being transferred from clients to infrastructure.



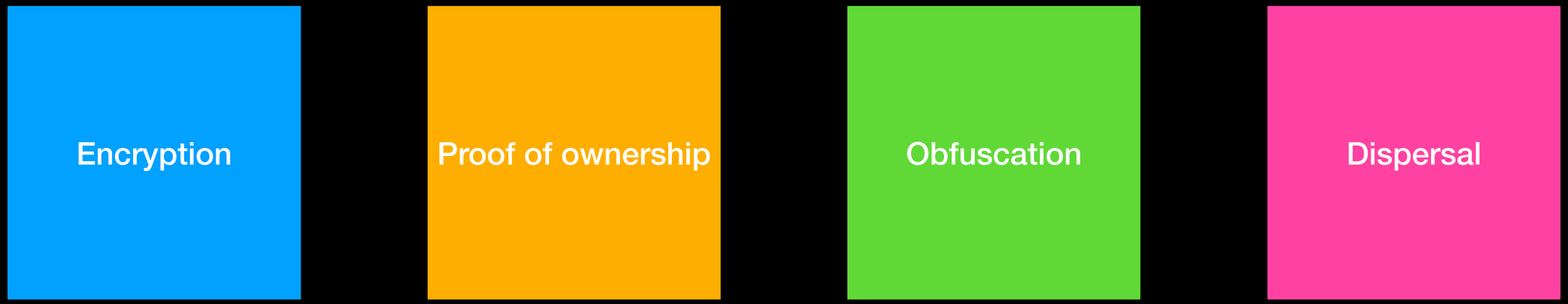
# Dispersal

## Secure Deduplication Solutions

- Deduplication solutions are often discussed in terms of a single infrastructure and single client, but is desirable to spread binary data across multiple deployments. Data deduplication with encryption across multiple infrastructures is challenging.
- An alternative solution is to use secret sharing techniques instead of encryption solutions to disperse information across infrastructures.

# Secure Deduplication Solutions

## Shin et al.



# Data Deduplication Security

## Business Continuity