# Chameleon: Keeping data safe for the naïve and thrifty

Ansley Post and Peter Druschel

MPI-SWS

NOTE: Slides contain notes which include additional information for readers of the slides

# Application Scenario

- Home Users / Small Businesses

# Application Scenario

- Home Users / Small Businesses

# Application Scenario

- Home Users / Small Businesses

What is RAID?

Is my data safe?

# Application Scenario

- Home Users / Small Businesses

What is RAID?

Is my data safe?

How much is this going to cost?

# Application Scenario

- Home Users / Small Businesses

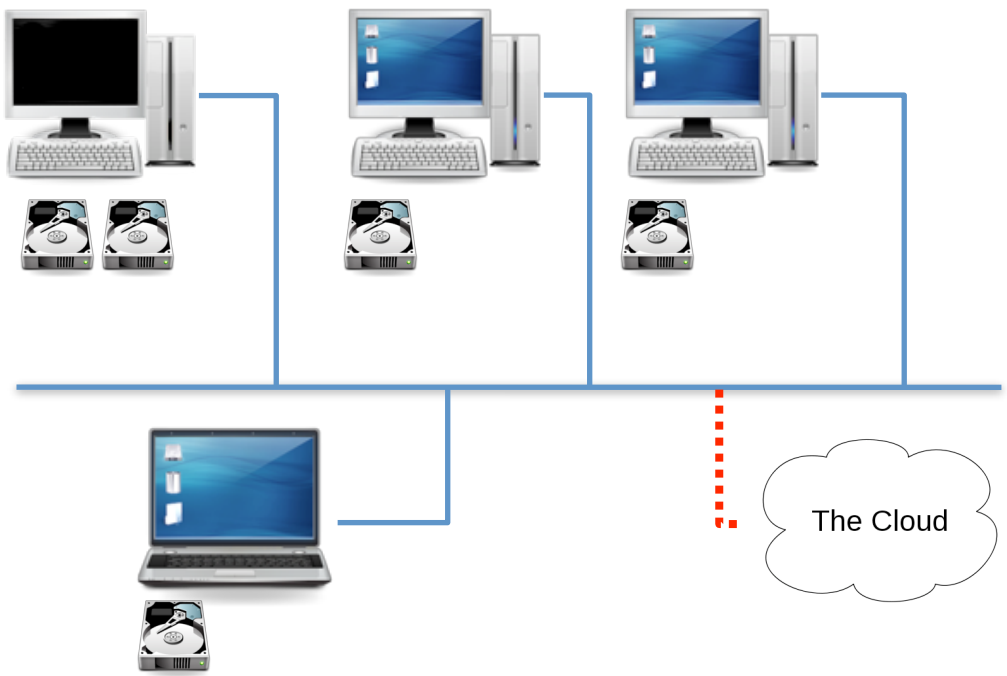
What is RAID?

Is my data safe?

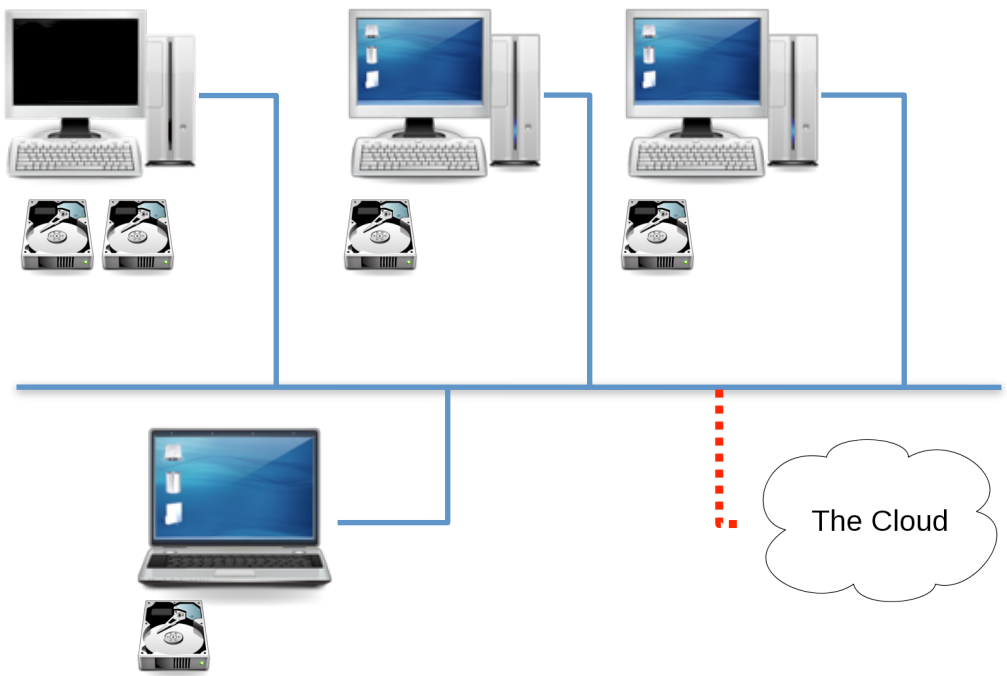
Am I capable of managing this?

How much is this going to cost?

# Example

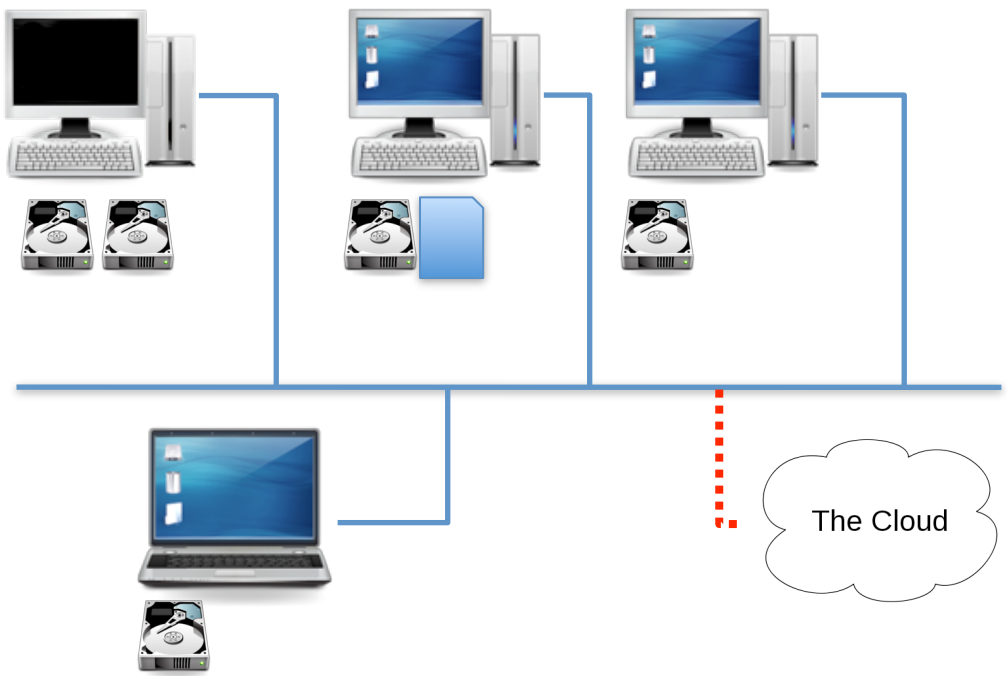


# Example



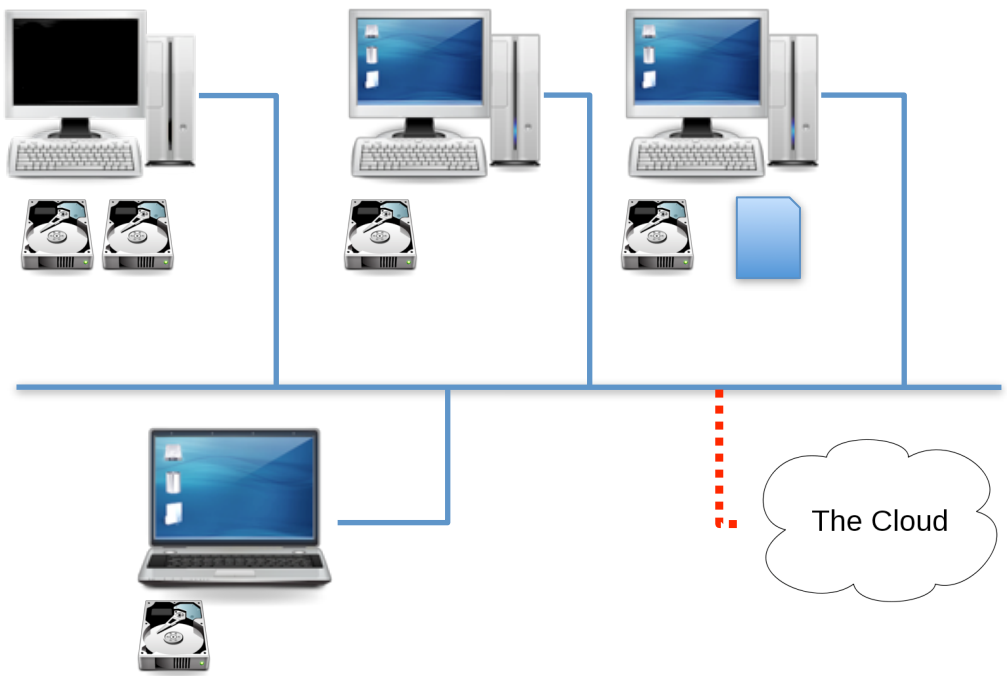

Many possible ways to replicate, place and encode data

# Example



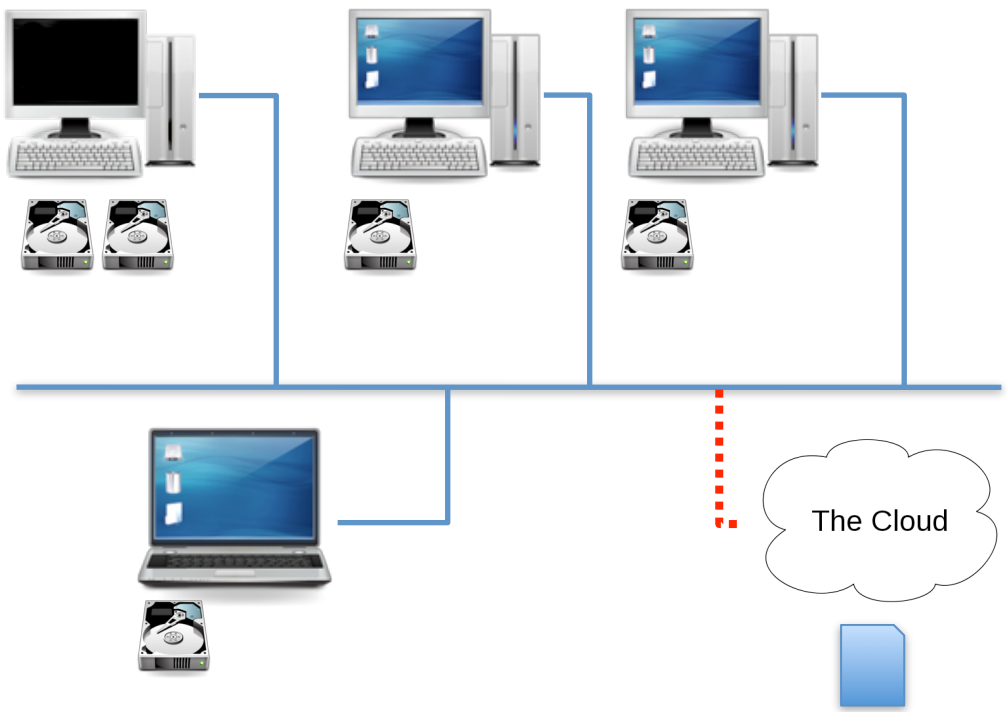

Many possible ways to replicate, place and encode data

# Example



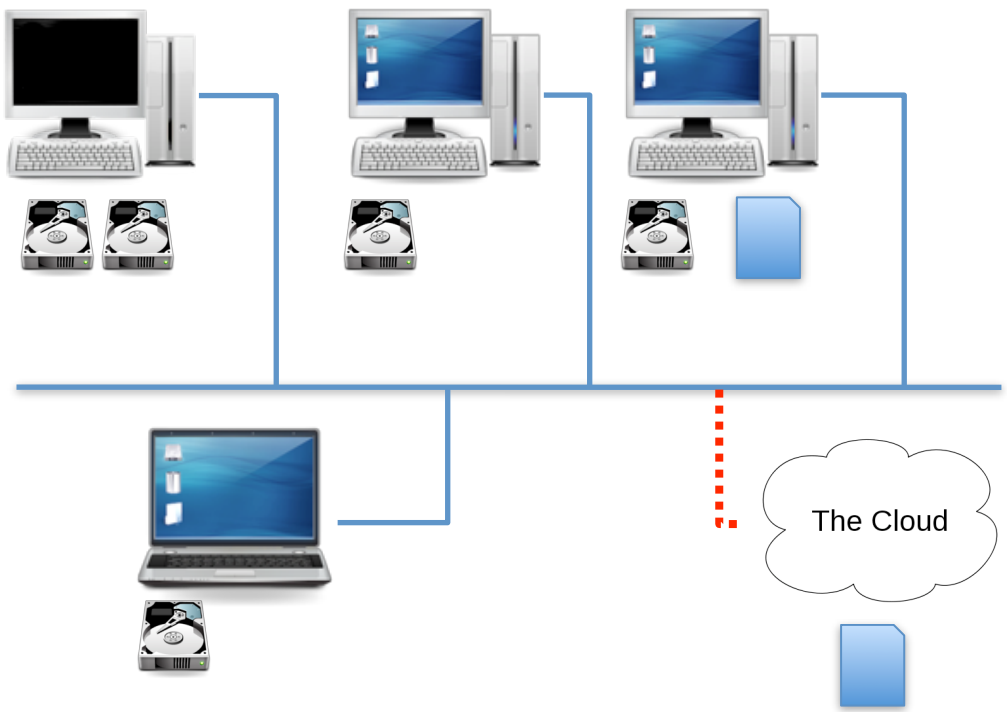

Many possible ways to replicate, place and encode data

# Example



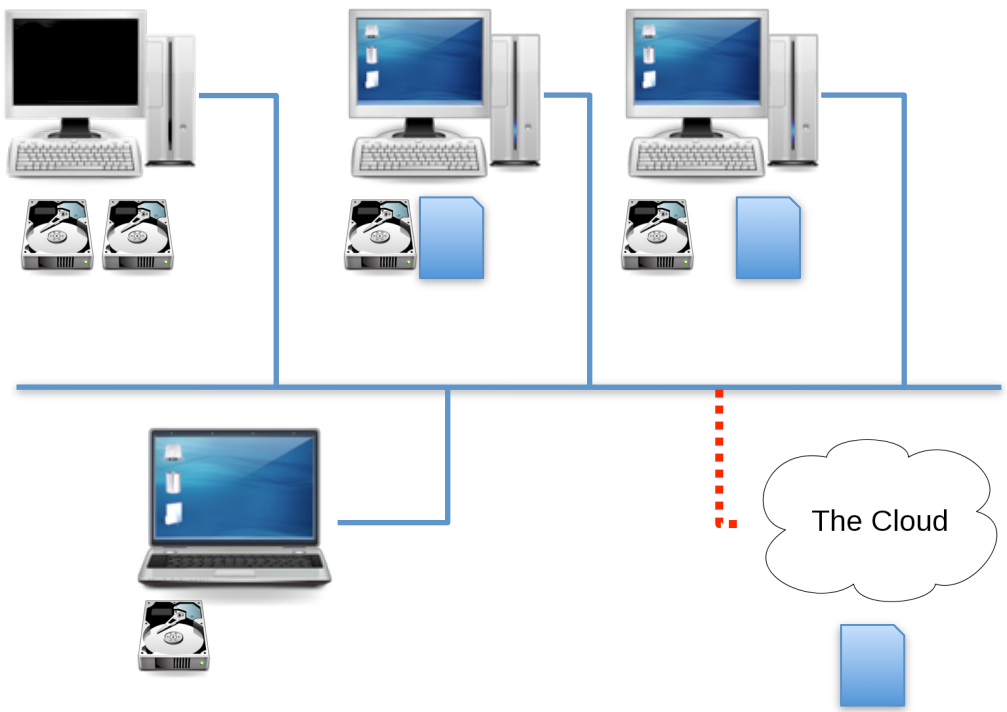

Many possible ways to replicate, place and encode data

# Example



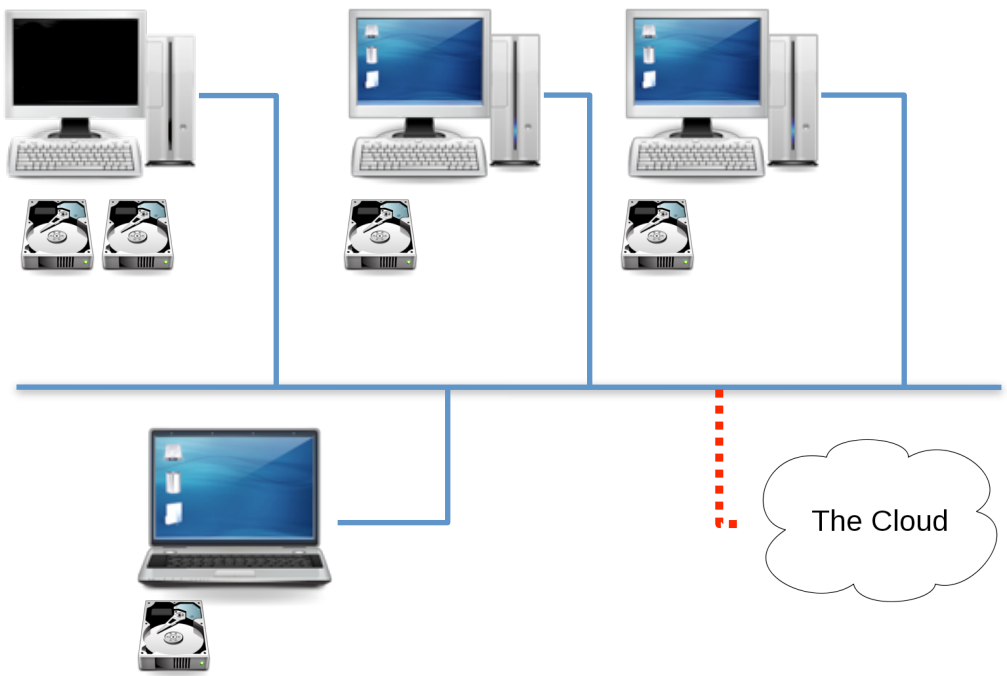

Many possible ways to replicate, place and encode data

# Example



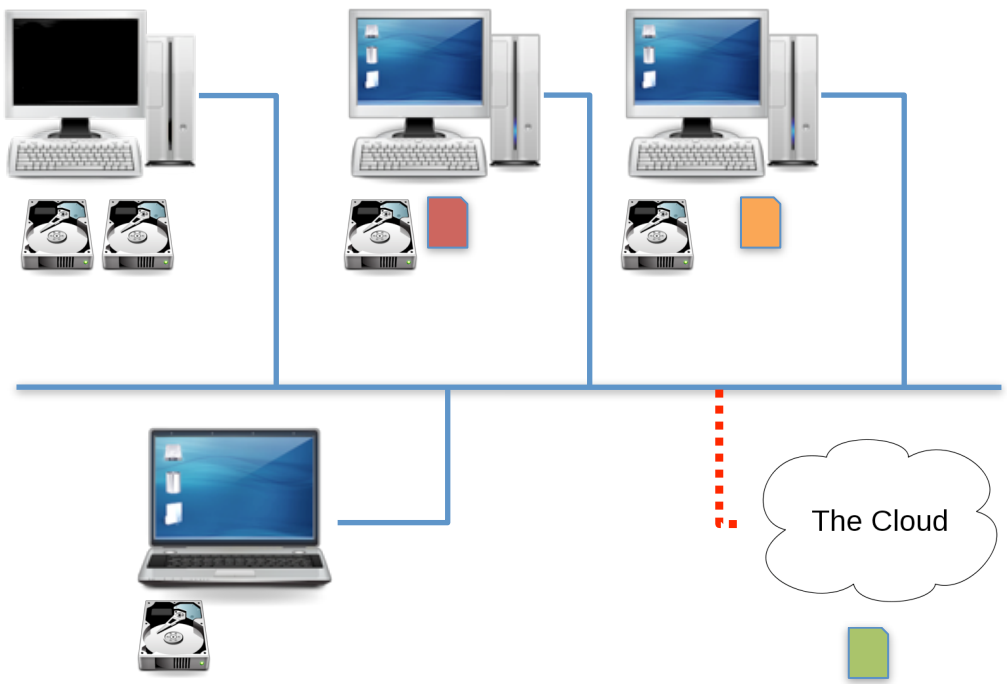

Many possible ways to replicate, place and encode data

# Example




Many possible ways to replicate, place and encode data

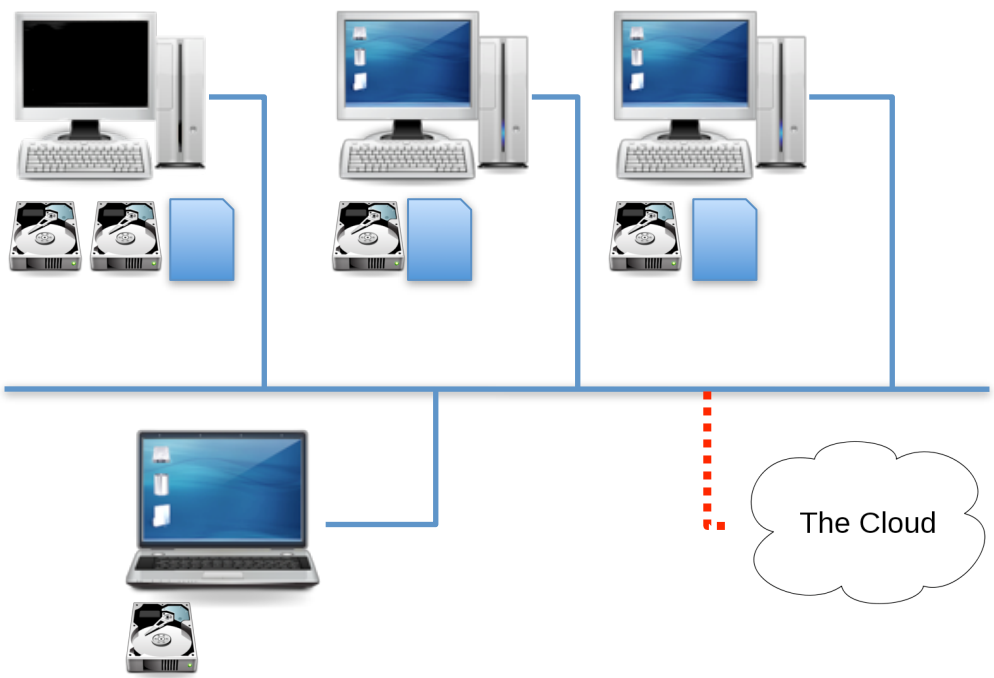
# Example




Many possible ways to replicate, place and encode data

# Offline storage

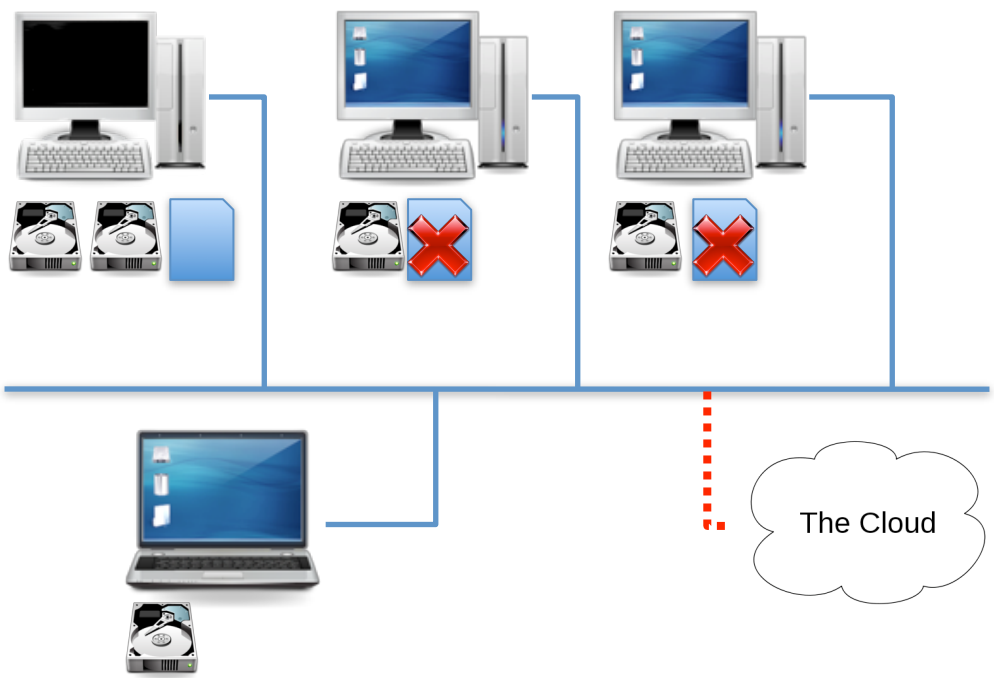
- Ideally, data is stored on at least one offline device
  - Enables recovery from catastrophic failure




NOTE: All online copies may become corrupted due to virus, operator error, software bug. Offline copies provide insurance against catastrophic failure.

# Offline storage

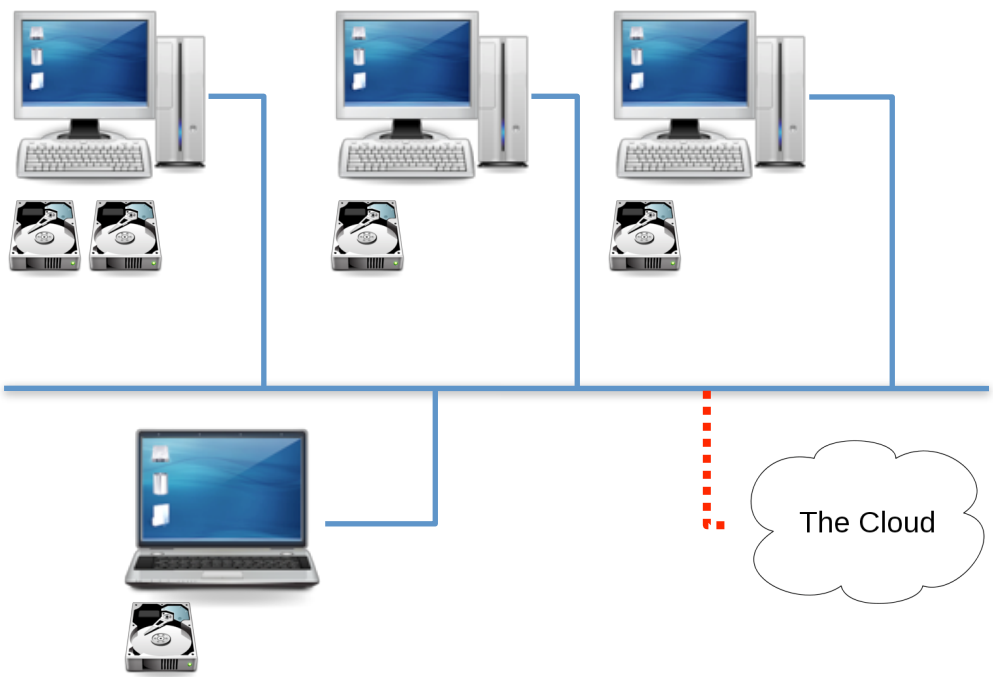
- Ideally, data is stored on at least one offline device
  - Enables recovery from catastrophic failure




NOTE: All online copies may become corrupted due to virus, operator error, software bug. Offline copies provide insurance against catastrophic failure.

# Anti-availability protocol

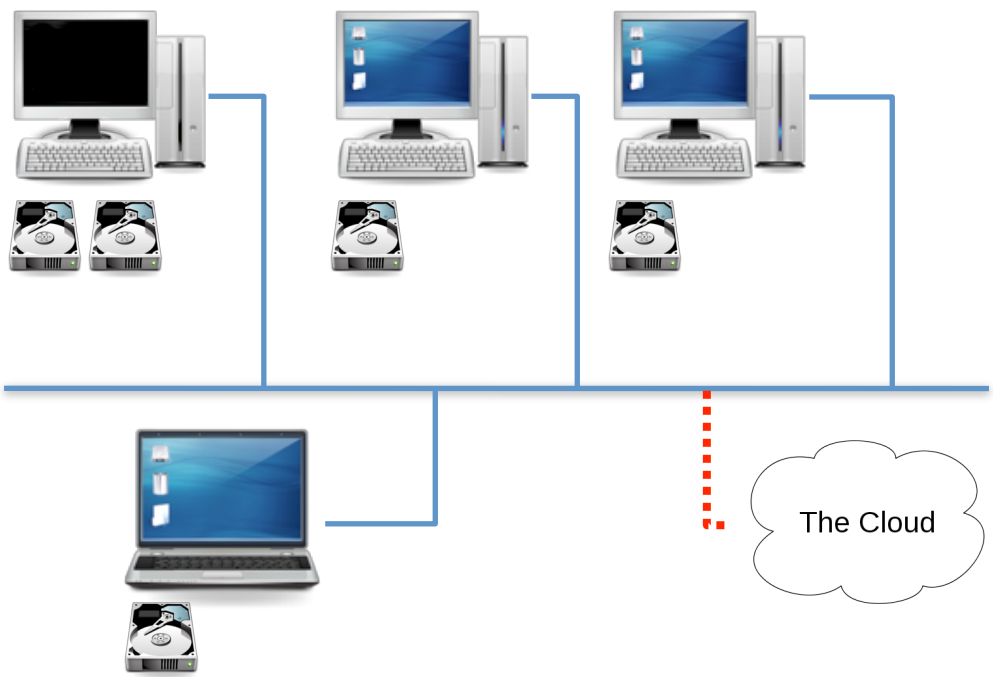
- Writes are only allowed if a certain number of devices are offline
  - Enforced by a small trusted kernel




NOTE: For clarity, a machine is shown as offline, but in practice only a storage device must be offline.

# Anti-availability protocol

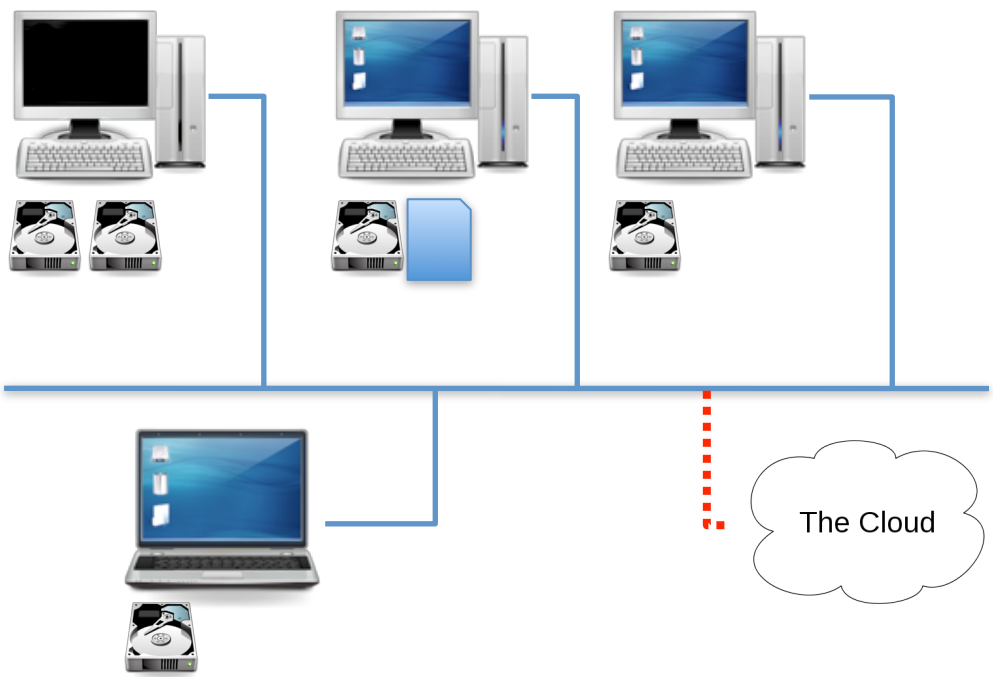
- Writes are only allowed if a certain number of devices are offline
  - Enforced by a small trusted kernel




NOTE: For clarity, a machine is shown as offline, but in practice only a storage device must be offline.

# Anti-availability protocol

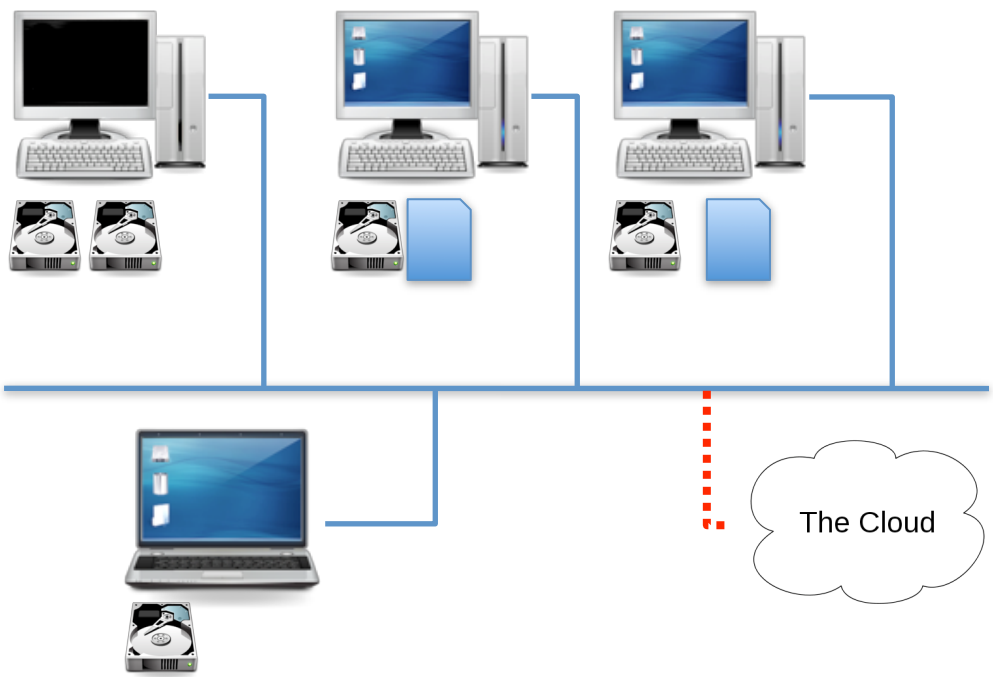
- Writes are only allowed if a certain number of devices are offline
  - Enforced by a small trusted kernel




NOTE: For clarity, a machine is shown as offline, but in practice only a storage device must be offline.

# Anti-availability protocol

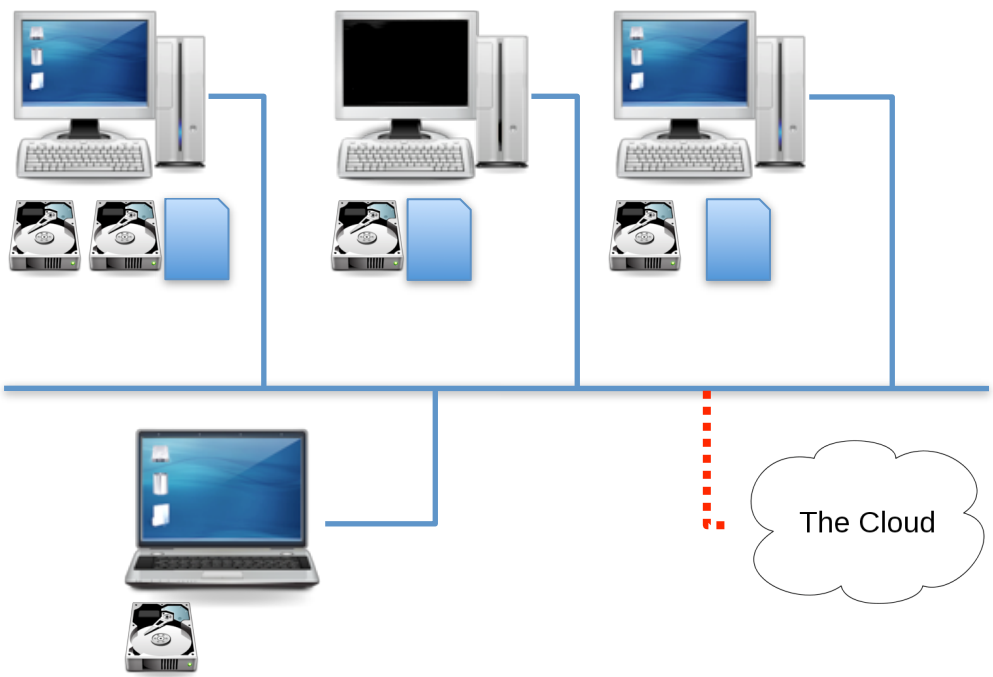
- Writes are only allowed if a certain number of devices are offline
  - Enforced by a small trusted kernel




NOTE: For clarity, a machine is shown as offline, but in practice only a storage device must be offline.

# Anti-availability protocol

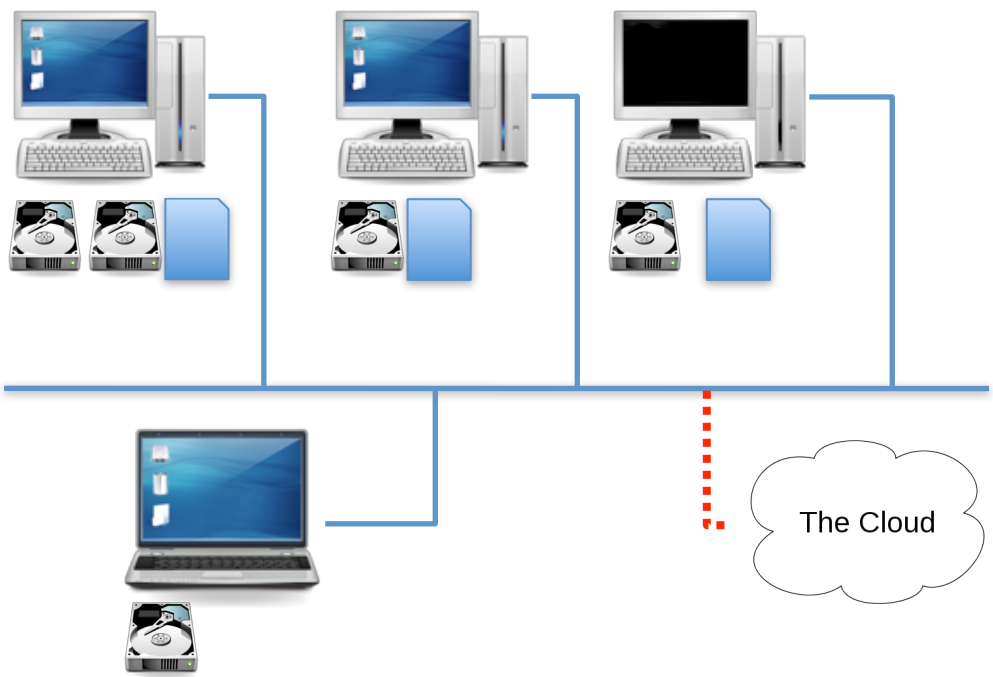
- Writes are only allowed if a certain number of devices are offline
  - Enforced by a small trusted kernel




NOTE: For clarity, a machine is shown as offline, but in practice only a storage device must be offline.

# Anti-availability protocol

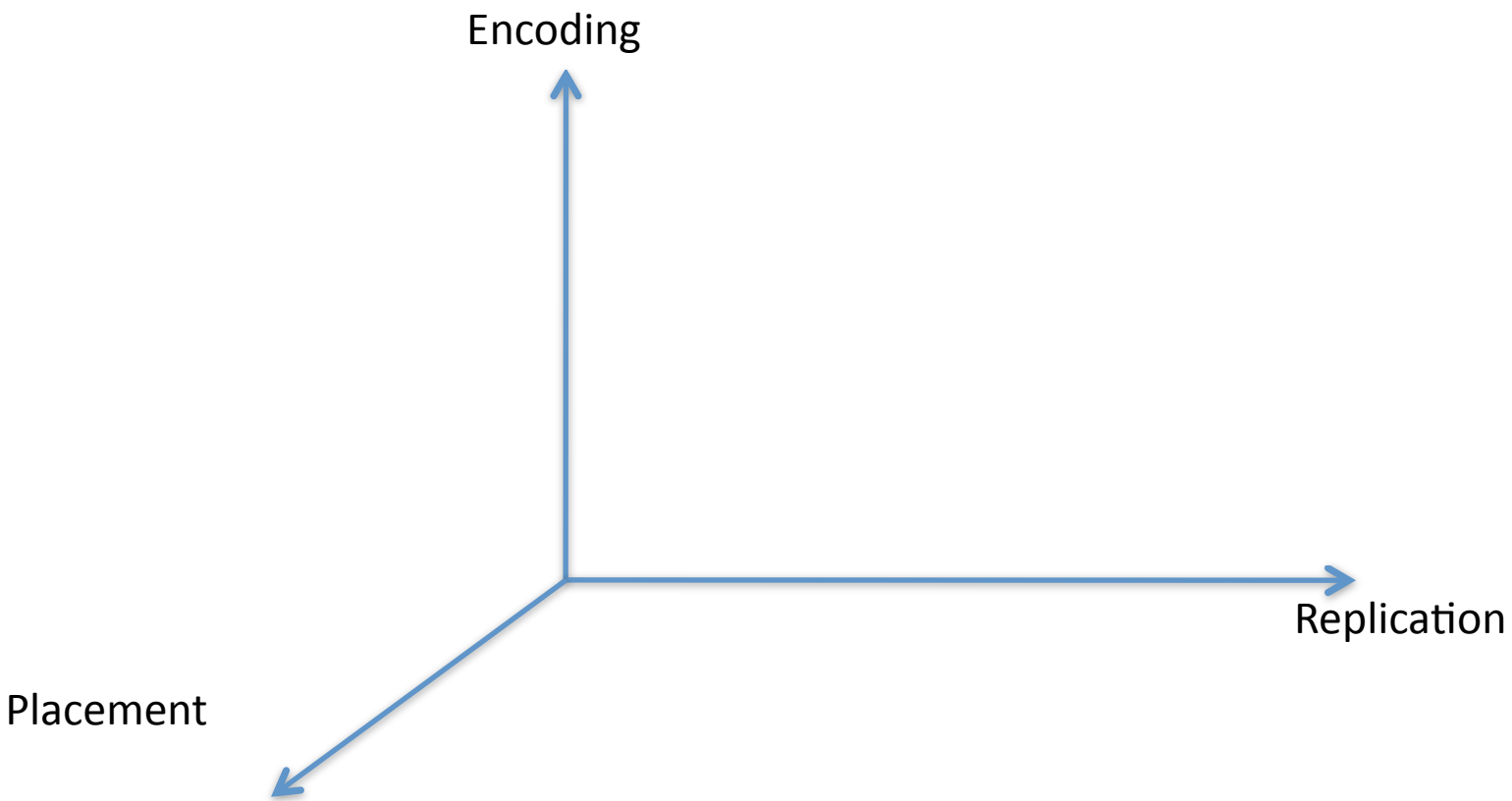
- Writes are only allowed if a certain number of devices are offline
  - Enforced by a small trusted kernel




NOTE: For clarity, a machine is shown as offline, but in practice only a storage device must be offline.

# Adaptive data storage

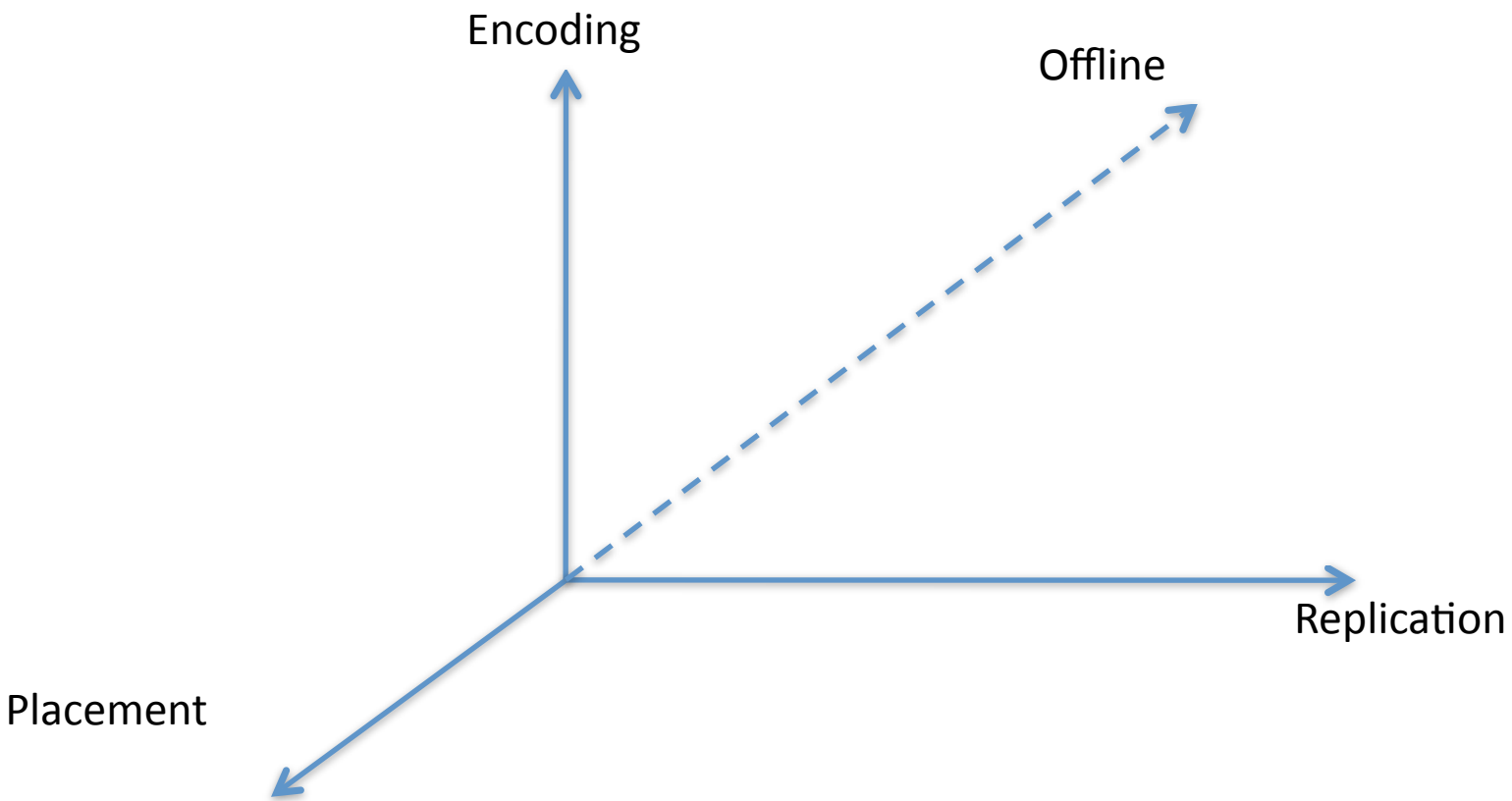
- Use linear programming to select and adapt storage configuration




NOTE: Huge number of possible tradeoffs. Well defined objective allows automatic selection of best configuration.

# Adaptive data storage

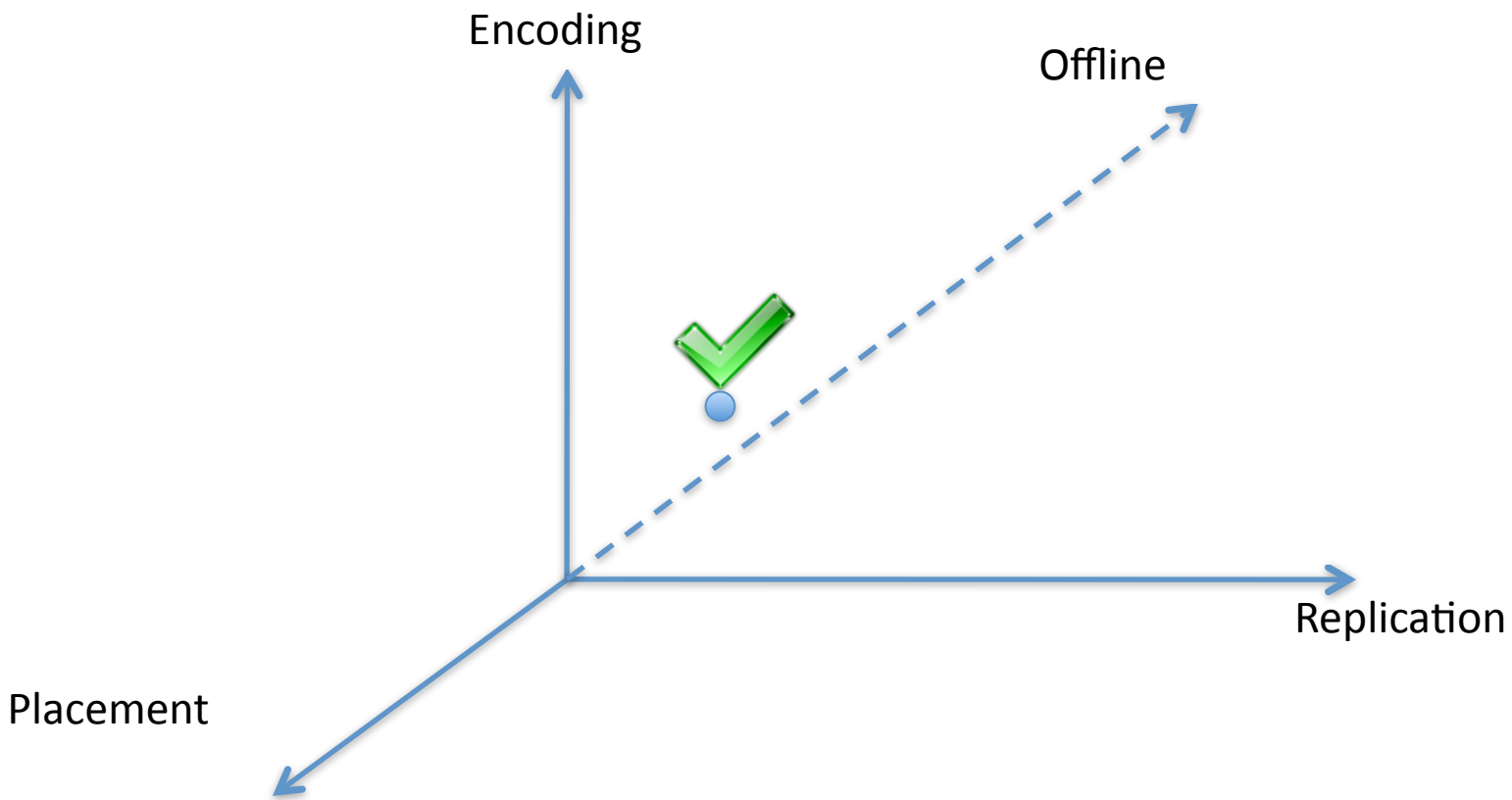
- Use linear programming to select and adapt storage configuration




NOTE: Huge number of possible tradeoffs. Well defined objective allows automatic selection of best configuration.

# Adaptive data storage

- Use linear programming to select and adapt storage configuration




NOTE: Huge number of possible tradeoffs. Well defined objective allows automatic selection of best configuration.

# Conclusion

- Currently finalizing design
- Prototype, named *Chameleon*, is under development

Contact: abpost@mpi-sws.org for more information