Collaborative Backup for Nomadic Devices

M.O.Killijian, D.Powell

48th Workshop IFIP WG 10.4 Hakone Japan, 4th July 2005

Context

- The MoSAIC project
  - Mobile Systems Availability Integrity and Confidentiality
- 3 years, 3 partners: LAAS, Eurécom, IRISA
  - Officially started September 2004
  - Funded by French Ministry of Research
- Nomadic device scenario
  - Mostly disconnected operations
  - Opportunistic wireless communication with similar devices
  - Peer-to-peer model of interactions
- Secure Collaborative Backup for Nomadic Devices

MoSAIC Goals

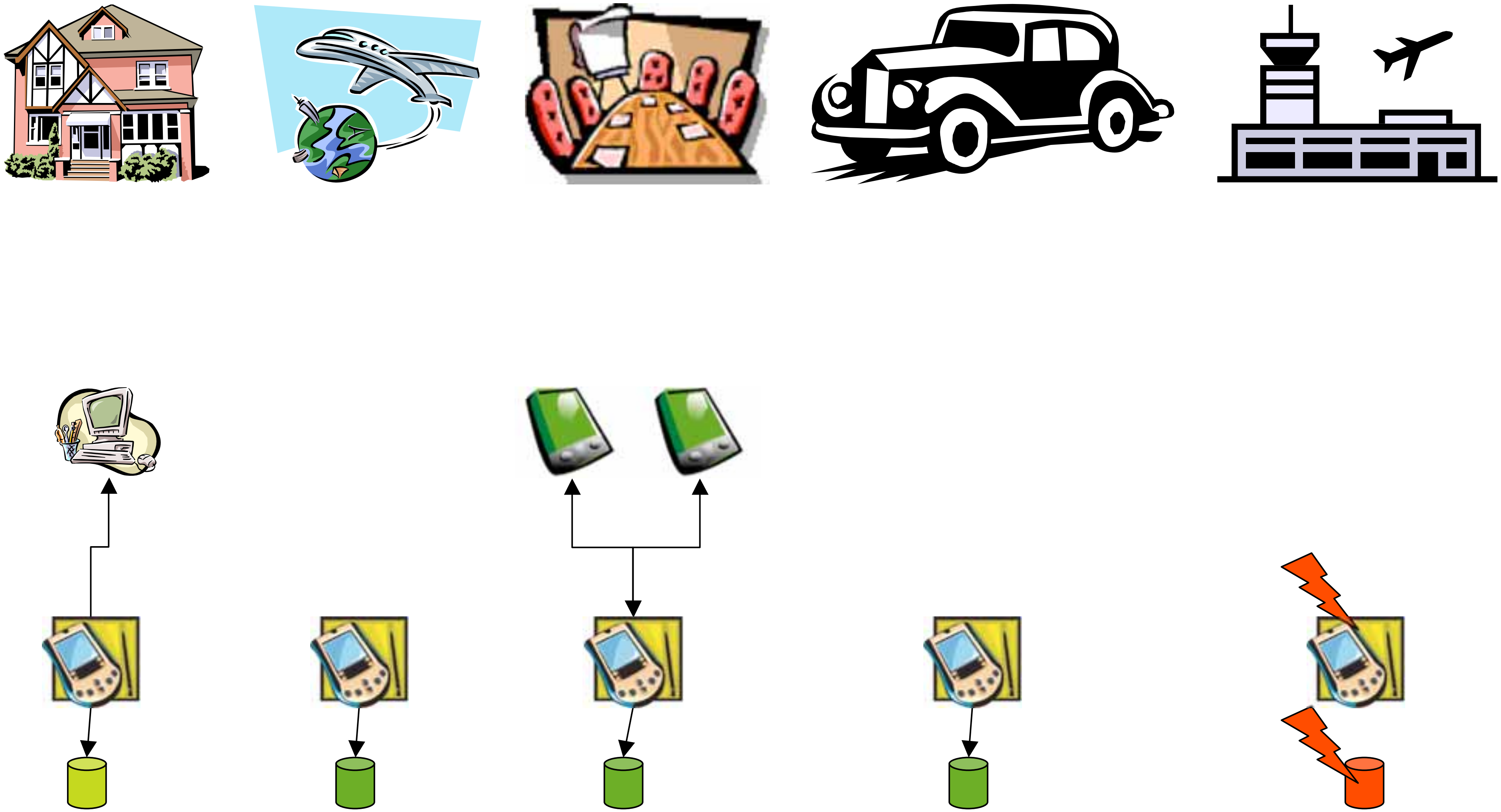
- In this context
  - new distributed algorithms and mechanisms for the tolerance of
    - accidental faults
    - malicious faults
  - without usual strong assumptions
    - synchronous communication
    - global clocks
    - Infrastructure

- New middleware for dependable mobile systems

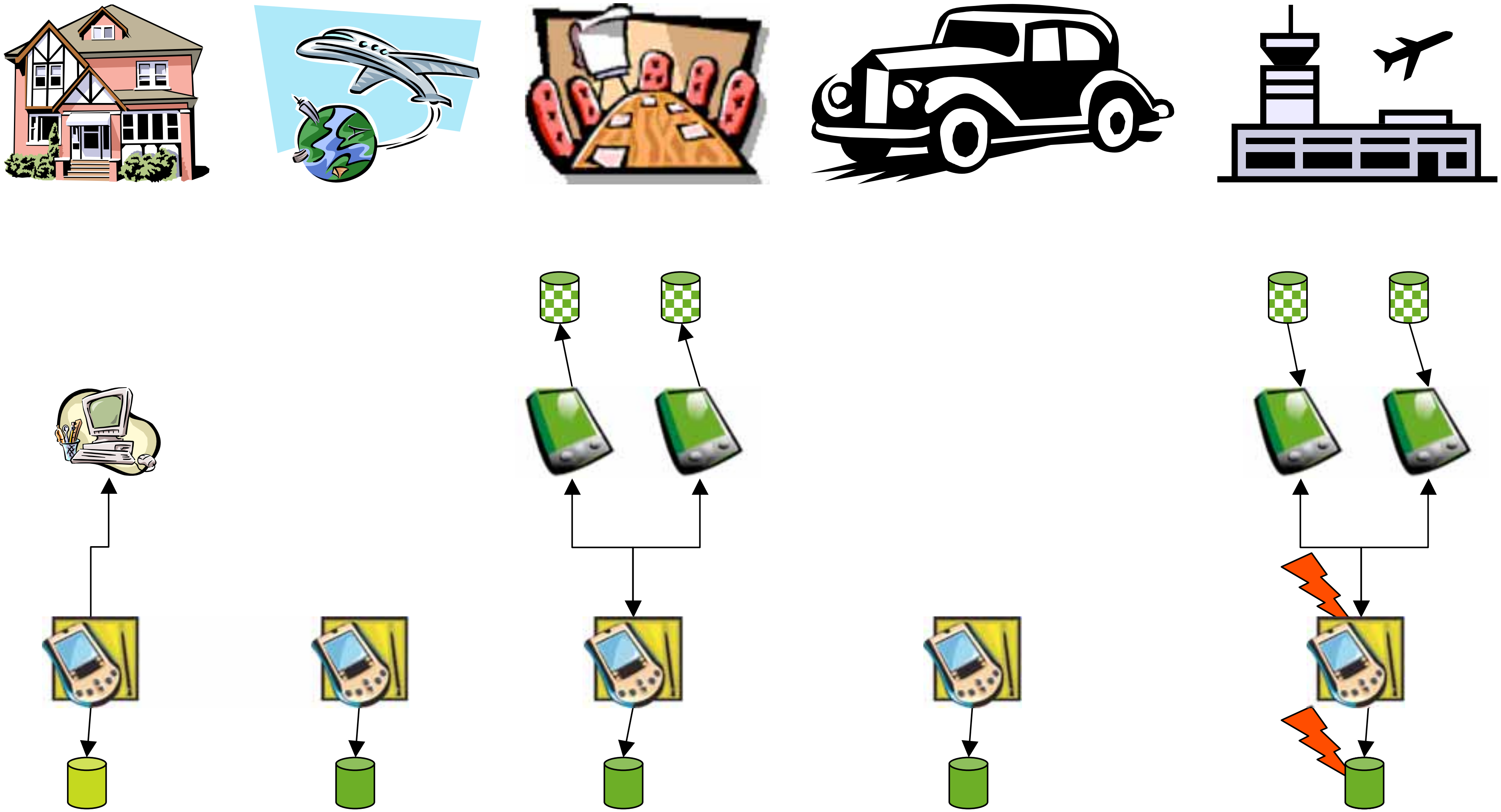
Overview

- **Overview of MoSAIC project**
- Collaborative Backup Systems
- Trust Management
- Current Status

Scenario without MoSAIC



Scenario with MoSAIC



Challenges for Dependability

- Limited energy, computation and storage
- Only intermittent access to a fixed infrastructure
- No prior organization
- Ephemeral interactions
- Critical private data

+ Usual criteria for classic functionalities
  - User transparency
  - Usability
  - etc.

Collaborative Backup

Participants are
- Data owners
- Service contributors

Objectives are
- Integrity and Availability
- Confidentiality and Privacy

Potential faults are
- Permanent and transient faults affecting a data owner
- Theft or loss of a data owner

- Accidental or malicious faults affecting availability of data backups
- Accidental or malicious modification of data backups
- Malicious read access to data backups

- Malicious denial of service (sabotage)
- Selfish denial of service (refusal to cooperate)

Overview

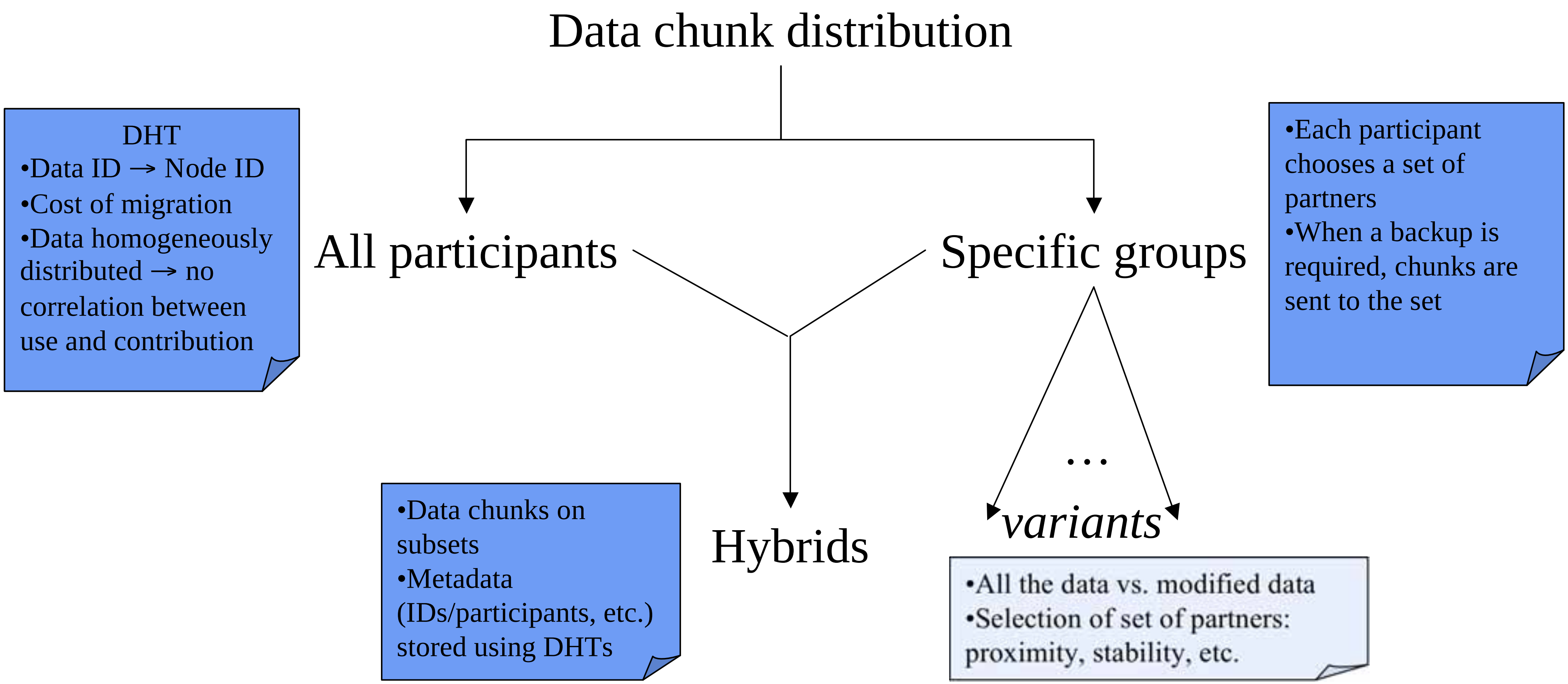
- Overview of MoSAIC project
- **Collaborative Backup Systems**
- Trust Management
- Current Status

P2P Storage Systems

- Peer-to-peer file sharing systems
  - *Overlay networks, DHT, unstructured*
    - GNUnet
    - FreeNet
    - OceanStore

- Peer-to-peer backup systems
  - *Cooperation incentives, trust*
    - Elnikety, Pastiche, PeerStore, pStore for WANs
    - Flashback for PANs

Storage space discovery and allocation



Elnikety *et al.*

- Peer-to-peer backup system on the Internet
  - No unique ID, no certified public keys, no routing
  - Set of partners, point-to-point reciprocal relationships
- Enforces
  - Confidentiality: secret key cryptography (IDEA)
  - Robustness: block redundancy using erasure codes (Reed-Solomon)
  - Integrity: self-checking sub-blocks, crypto hash-keys (HMAC-MD5)
  - Authentication: pairwise shared secret keys (Diffie-Hellman)
- Attacks
  - Selfish DoS: periodic challenges, grace and commitment periods
  - Malicious DoS: protocol against man-in-the-middle attacks

Flashback

- Devices are part of a Personal Area Network (PAN)
  - Same owner: a priori mutual trust
- Permanent fault (or theft) of the data owner
  - Same ID assigned to a new device
  - Reinitialized from backed-up data
- Optimization of the restorable data
  - Limitation of # of copies (function of block priority)
  - Replication rate function of current number of copies
  - Taking into account heterogeneity (energy, storage)
- Backup contracts: notion of lease
  - Duration of lease > expected duration of disconnection
  - Lease renewal at 50% expiry time

P2P vs. MoSAIC

- Fixed and unique IDs: not available
- Bandwidth, duration of connections: not known a priori
- Mobility: partnerships have to change and adapt
- Resource and node discovery: knowing one participant/repository is not enough
- Intermittent connection to fixed infrastructure: mostly disconnected
- Trust mechanisms for disconnected operation: reputation (e.g., using trusted HW)

Overview

- Overview of MoSAIC project
- Collaborative Backup Systems
- **Trust Management**
- Current Status

Tragedy of the Commons

- Why do we need cooperation incentives?
- “Tragedy of the Commons” [Hardin68]
  - Resource sharing
    - Naturally there are disincentives
    - Cooperation implies consumption of ones own resources
  - Selfish users behave as free-riders
    - Consumption without contribution
  - Very common behavior especially in large networks
    - 70% of Gnutella network users do not contribute

Routing in ad hoc networks 1

- Forwarding/routing packets costs
  - Energy, bandwidth, CPU cycles
- Different misbehaving nodes
  - Selfish DoS (passive) - priority is energy
    - Don’t forward packets
  - Malicious DoS (active) - priority is damage
    - Drop packets
    - Send wrong routes
- No a priori trust/confidence
- Enforce cooperation
  - Detection of misbehaving nodes
  - Isolation of misbehaving nodes
  - Stimulate and encourage cooperation

Without excessive resource consumption

Routing in ad hoc networks 2

- Use redundant routes for every packet
  - Increased energy consumption
- Consider false route information as old routes
  - Need a majority of honest nodes
- Use localization information for routing (GPS)
  - Privacy attacks
- **Money** as an incentive
  - Exchange virtual money for routing (e.g., Buttyan’s nuglets)
  - Requires secure kernels/trusted hardware
- Detect misbehavers, give them bad **reputation**
  - Global reputation requires access to servers
  - Local reputation (e.g., Marti’s watchdogs)

Trust Mechanisms

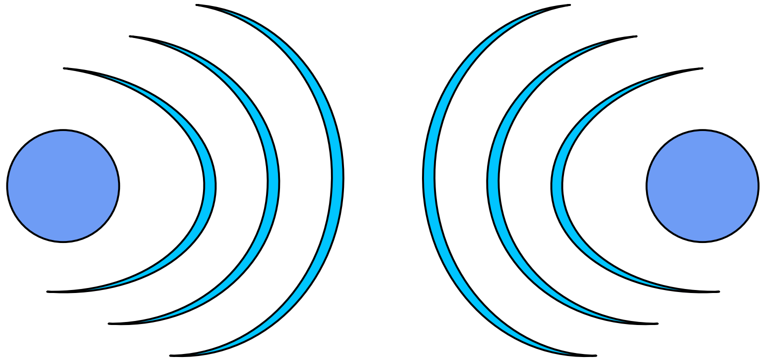
- Traditional key management
  - Public Key Infrastructure (PKI)
  - Trust authority to establish trust between mutually distrusting entities
  - **Centralized trust servers**
- Trust established using long-term accountability
  - Micro-payment against free-riding [Golle]
  - Contributor ratings [eBay, bizrate, etc.]
  - **Centralized rating/bank servers**
- Web of trust
  - Distributed trust model, PGP-like
  - Used primarily for key management
  - Content-centric for reputation-guided searching [Poblano]
  - Peer-centric [Law-Governed Interaction] needs trusted kernels/HW

Overview

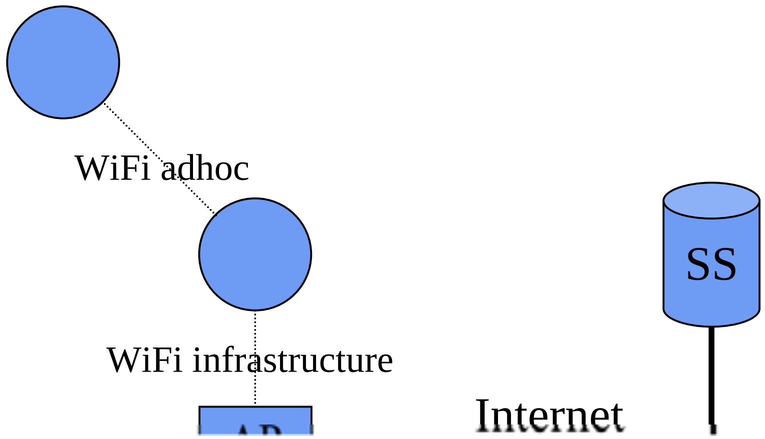
- Overview of MoSAIC project
- Collaborative Backup Systems
- Trust Management
- **Current Status**

Node discovery

- Discovery of MoSAIC nodes
  - Online
  - Creation of ad hoc network
  - Active beaconing:
    low latency vs energy economy

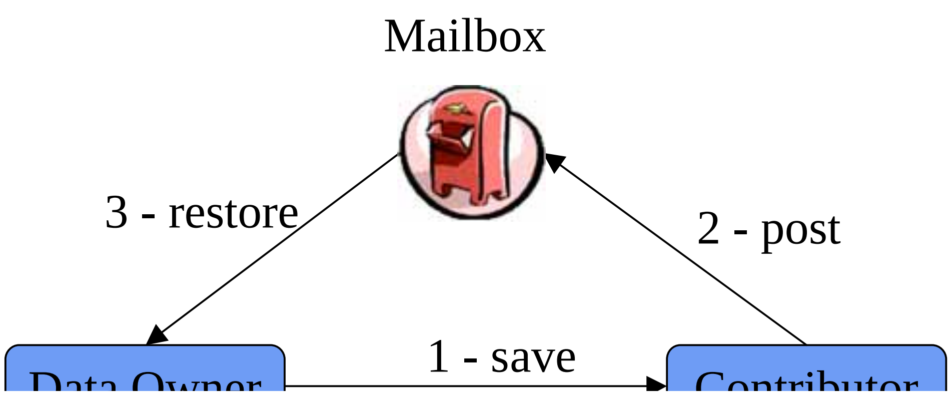


- Discovery of Internet access
  - Be able to backup on reliable storage service
- Ad hoc and infrastructure mode at the same time
  - Cooperation + storage service access



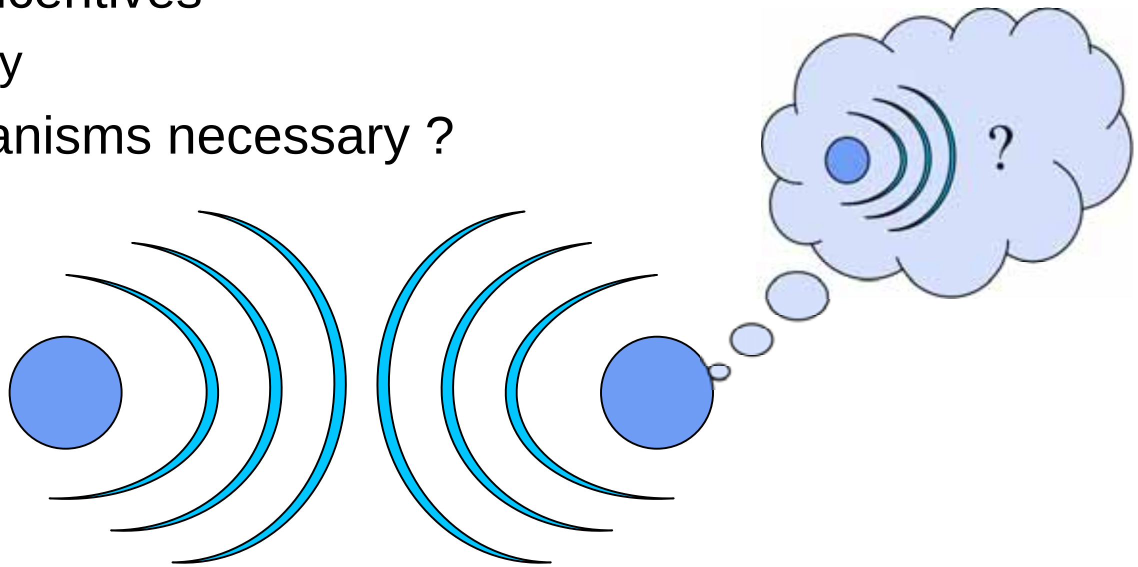
Being Opportunistic

- Opportunistically use connection to Internet
  - "Mailbox" for storing the backup chunks
  - Accommodate several restoration models
    - Push: the contributor sends the chunks back home
      - Internet access, mailbox at the owner's home
    - Pull: the data owner searches for the data when necessary
      - Ad hoc network, mailbox hosted by the contributor
    - Push-pull: storage service as an intermediary
      - Internet access, mailbox hosted by the storage service

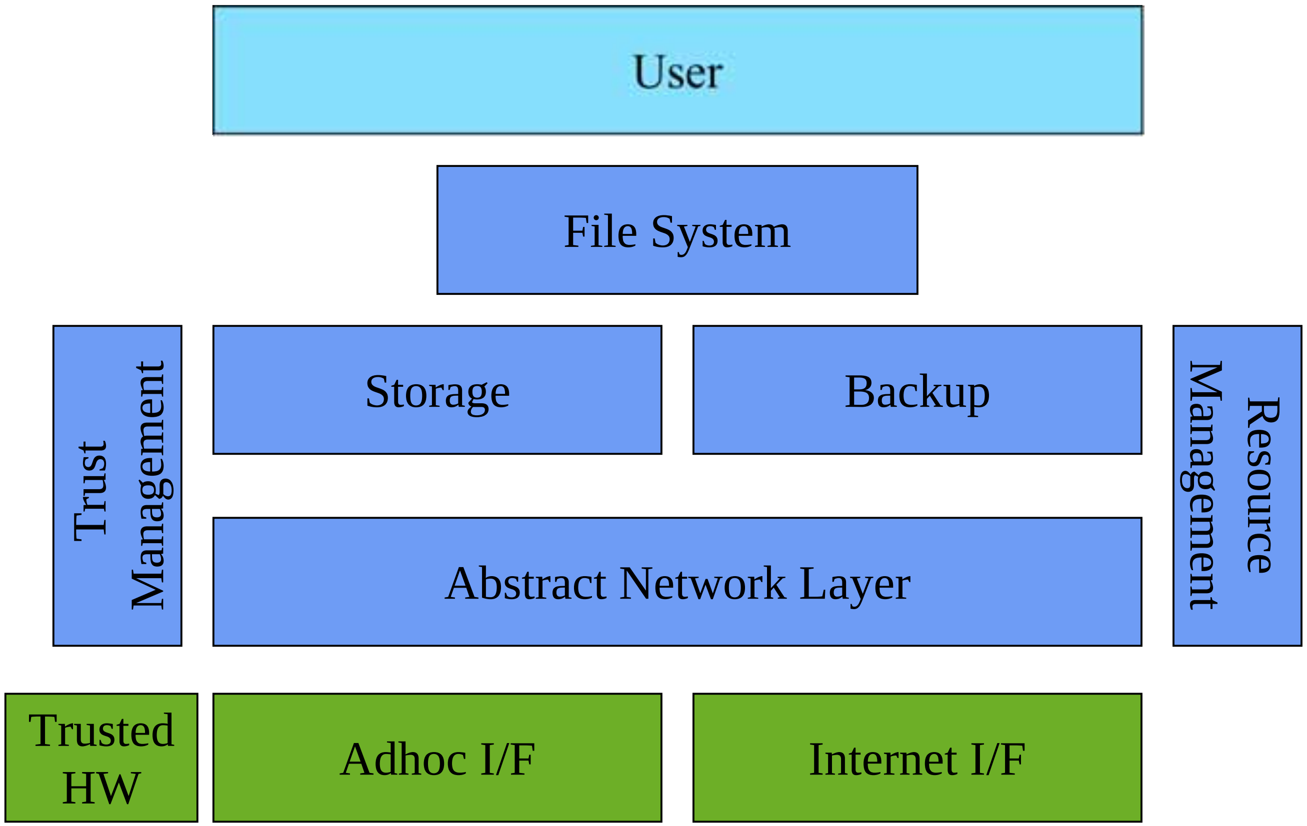


Trust Management

- Classic solutions
  - Participants are almost always connected
- Strong mobility, ephemerous connections, etc.
  - Self-carried reputation (using trusted HW)
    - Checked by other participants
    - Link with the mailbox implementation
  - Collaboration incentives
    - Virtual money
  - Are both mechanisms necessary ?



Architecture



Conclusion

- Scenario for
  - Designing new algorithms
  - Developing new middleware
- Implies fault-tolerance
  - Classic faults
    - Devices: crash of devices (owners and contributors), etc.
    - Data: integrity, confidentiality
  - Interaction faults (selfishness, maliciousness)
- New FT-enabling mechanisms
  - Self-carried reputation, virtual money, etc.
  - Opportunistic Internet backup, P2P interactions
- Project is 10 months old, still a lot to do ….

Collaborative Backup for Nomadic Devices

M.O.Killijian, D.Powell

48th Workshop IFIP WG 10.4 Hakone Japan, 4th July 2005

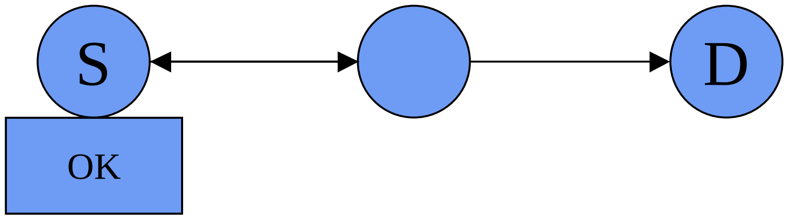
Buttyan's nuglets

- Each node maintains a counter (nuglet)
  - Decreased when sending its own packet
  - Increased when forwarding a packet
  - The counter must remain positive

- The policy must be enforced
  - Use of tamperproof hardware
    - SIMcards, JavaCards, etc.
    - TPM

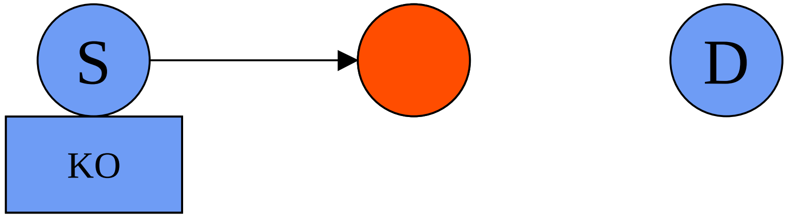
Marti's Watchdogs

- Each node possesses a watchdog
  - When a node sends a packet, the watchdog verifies that the neighbors forward it



Marti's Watchdogs

- Each node possesses a watchdog
  - When a node sends a packet, the watchdog verifies that the neighbors forward it



- Misbehaving nodes are detected: bad reputation
- Limits
  - Collisions
  - Low transmission power attacks
  - False positives
  - Collusion
  - Partial propagation