Adding Resilience to Message Oriented Middleware (Project Paper)

Jinfu Wang, Peng Jiang, John Bigham

Queen Mary University of London

& MPI-QMUL ISRC

GEMOM

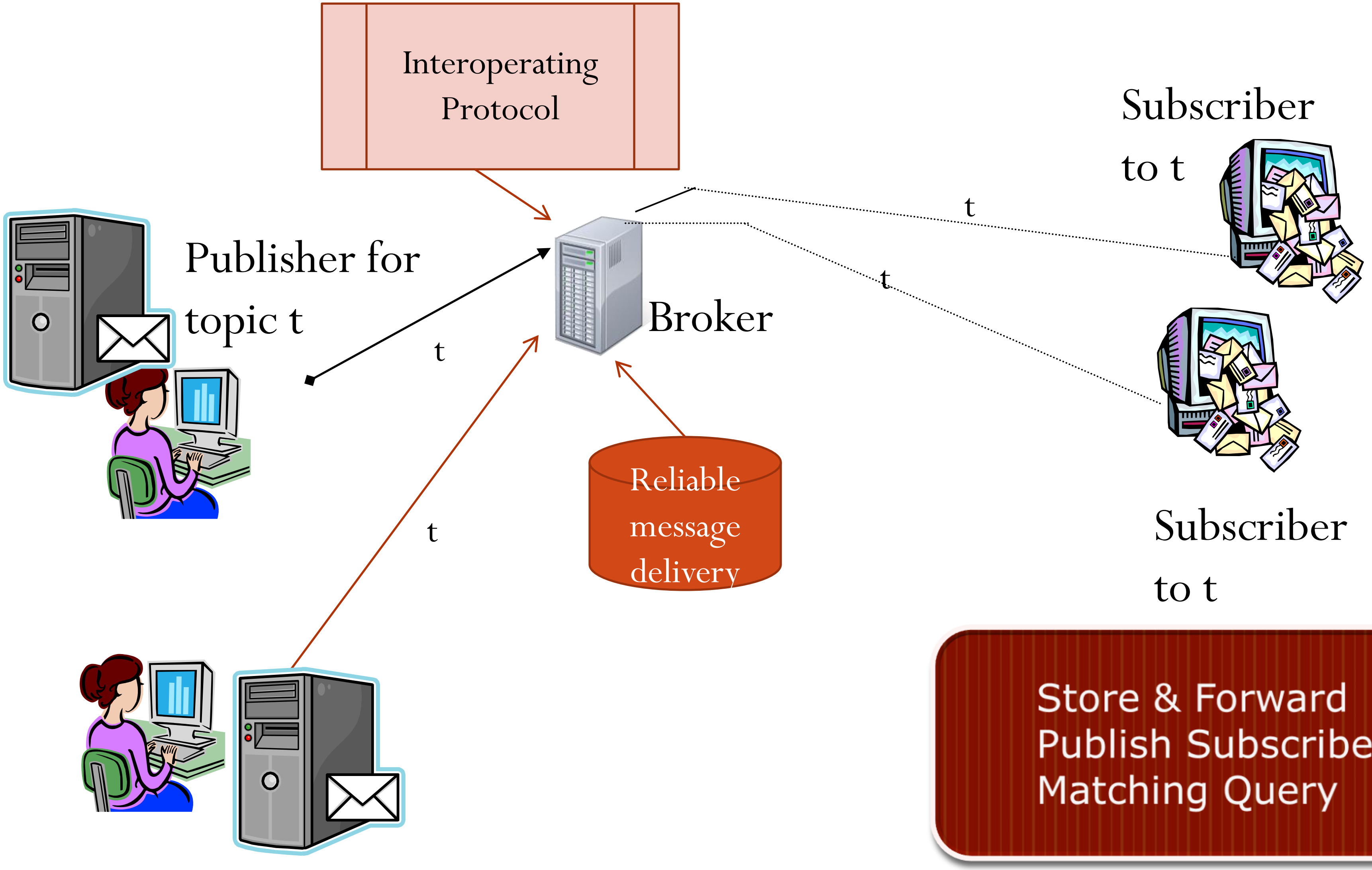
Genetic Message Oriented Secure Middleware

Message Oriented Middleware (MOM)

- Message Oriented Middleware
- Increases the interoperability, portability, and flexibility of an application or set of distributed cooperating applications
  - The "glue" between heterogeneous systems

- Dominant solutions are proprietary - IBM Websphere MQ, Tibco
  - too expensive for everyday use (Cloud-scale)
  - they do not interoperate
- Open standard – Advanced Messge Queueing Protocol (AMQP)
  - Apache QPID, RabbitMQ, Cisco AON
- GEMOM Project
  - Resilience, security, speed for MOM
  - Some of this talk refers to part of this work

What is MOM?



Interoperating Protocol

Publisher for topic t

Broker

Reliable message delivery

Subscriber to t

Subscriber to t

Store & Forward
Publish Subscribe
Matching Query

MOM System

- Multiple Nodes (brokers) are often deployed depending applications
- Both for communications in internal network and across internet
- Clustering and Federation of brokers to improve performance and scalability
- Sometimes MOMs need great speed (e.g. some financial application)
- Sometimes MOMs need great resilience and assured delivery e.g. mission critical application services

GEMOM - Genetic Message Oriented Middleware

- GEMOM – EU FP7 Project
- Goal: Provide resilient messaging service
- Definition of Resilience
  - Ability of the MOM system to provide and maintain an acceptable level of service in the face of both internal and external faults and challenges to normal operation

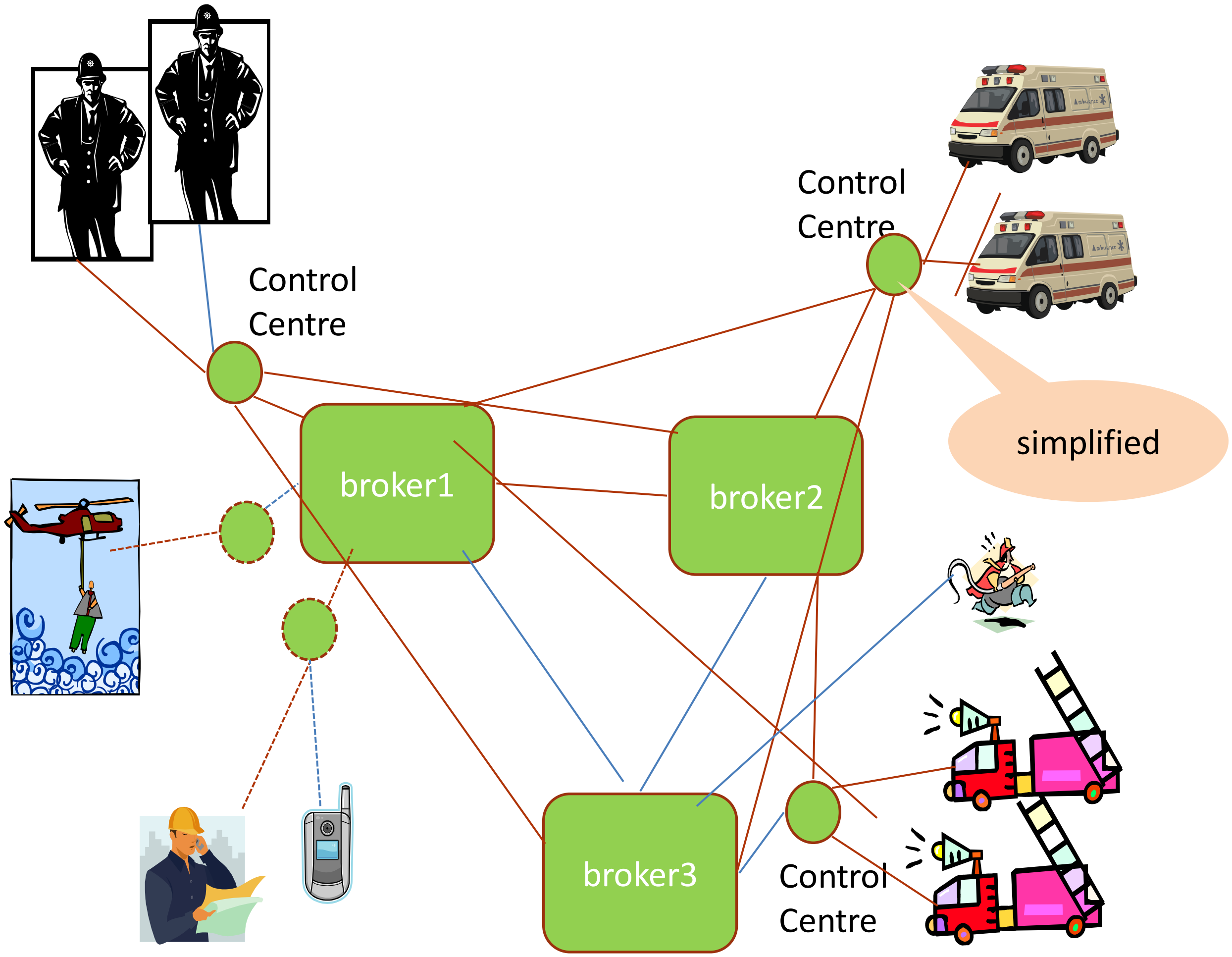
Resilience Challenges

- Internal Challenges,
  - Tolerate burstiness in workload
  - broker faults and failures
- External Challenges:
  - network failures and disruption caused by external reasons
  - DDoS attacks

Approach to achieve resilience

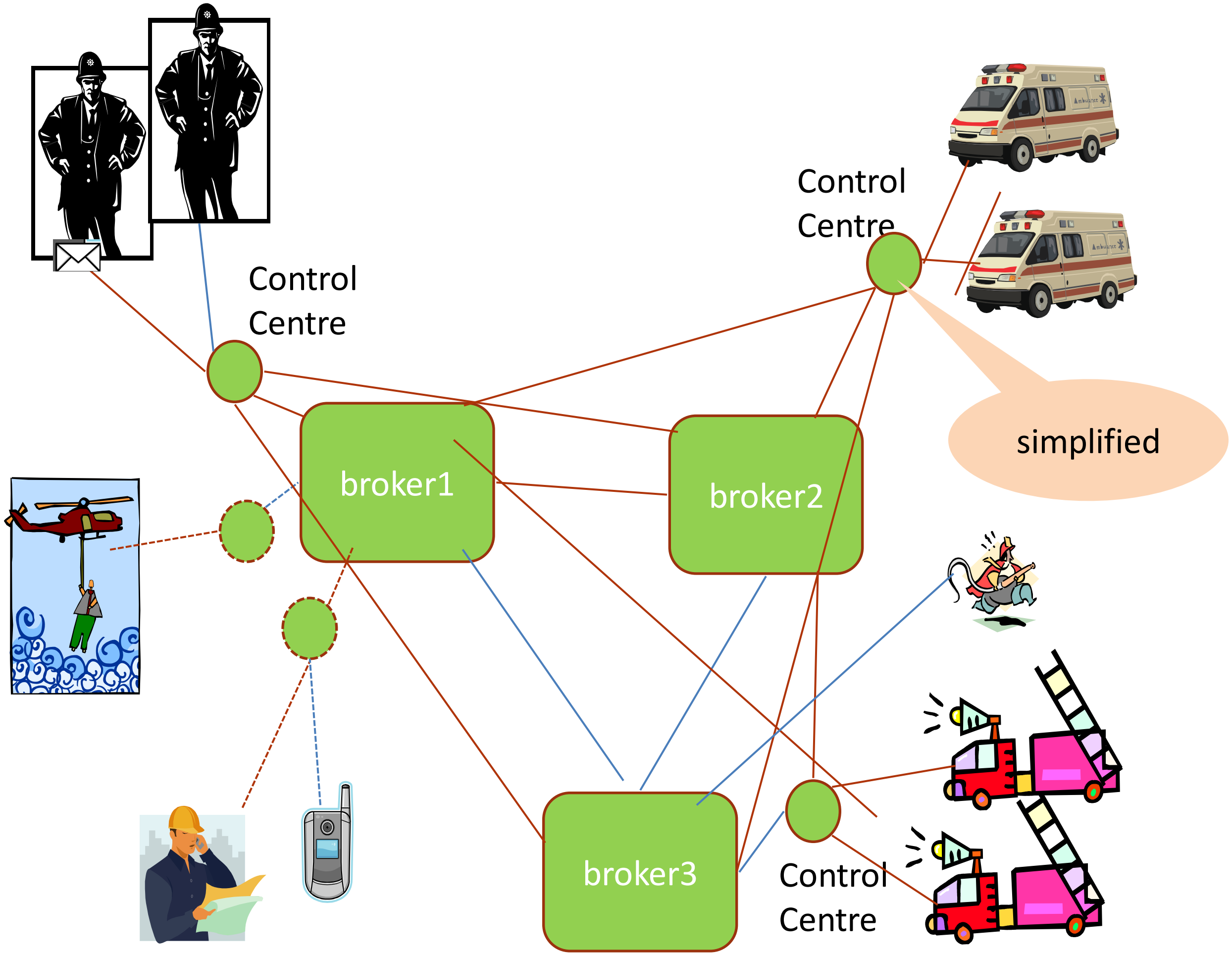
- Tolerating bursty workload and broker faults – Find workload allocation and mirror solutions, while quantify the risk of burtiness
- Use overlay of brokers over internet to response quickly to disruptive link faults via relay brokers
- Defence mechanisms against DDoS attacks, and decentralize the control with P2P Managing Agents

Illustration of issues



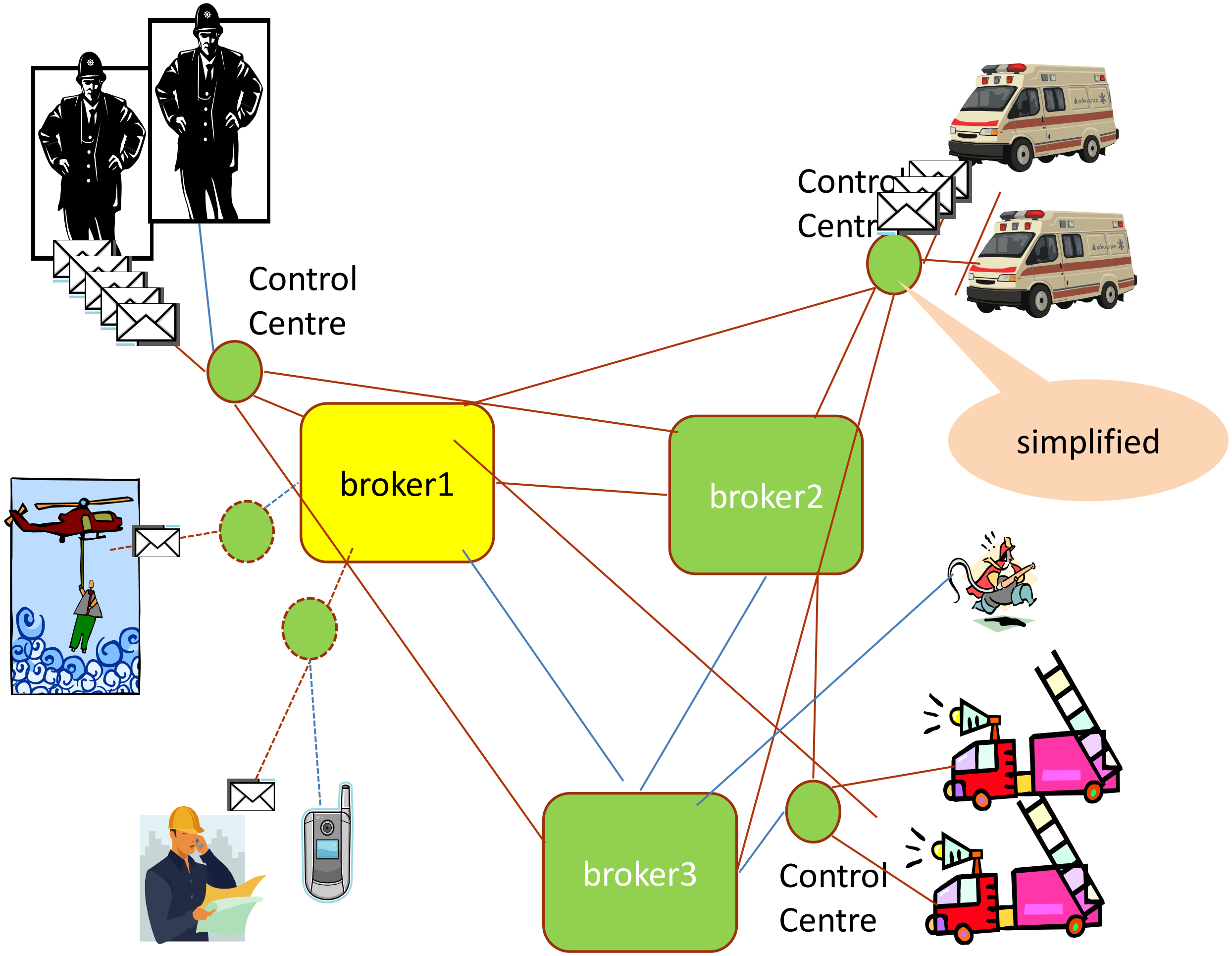
Possible structure of an overlay that can be used to provide resilience

A Simple Message Flow



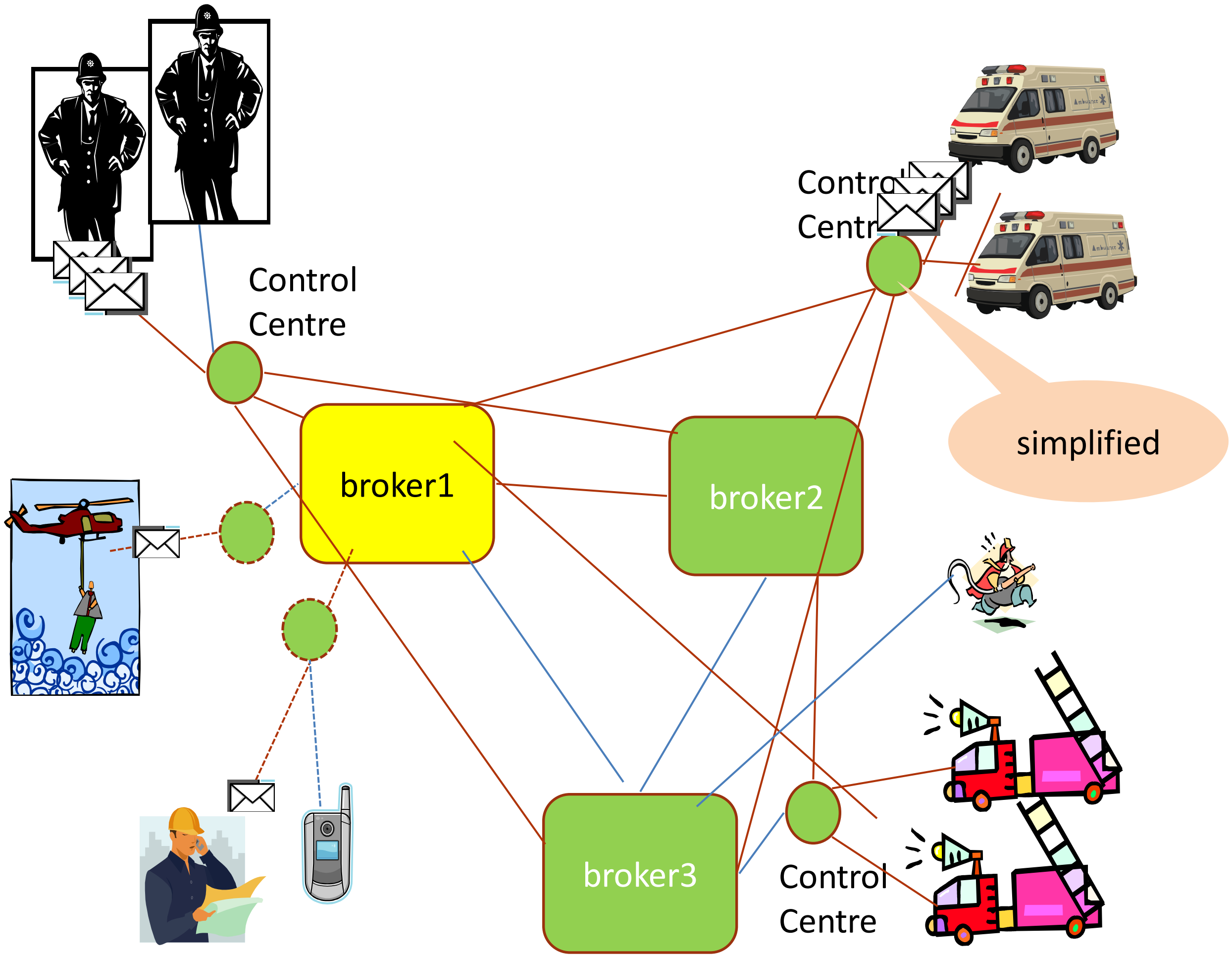
Possible structure of an overlay that can be used to provide resilience

Under stressful load, correlated bursty traffic causes a system overload



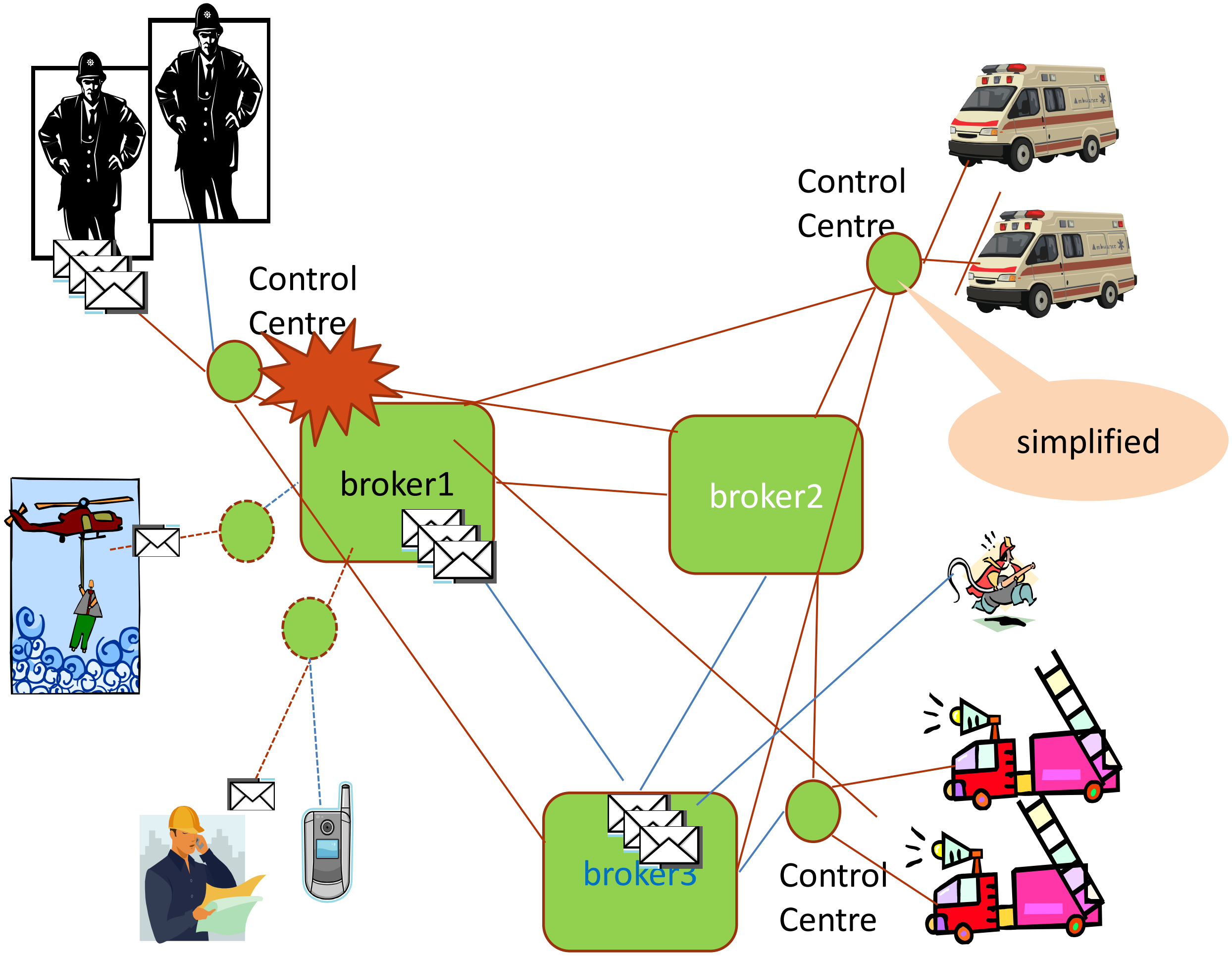
Possible structure of an overlay that can be used to provide resilience

Distribute the correlated bursty traffic to different nodes



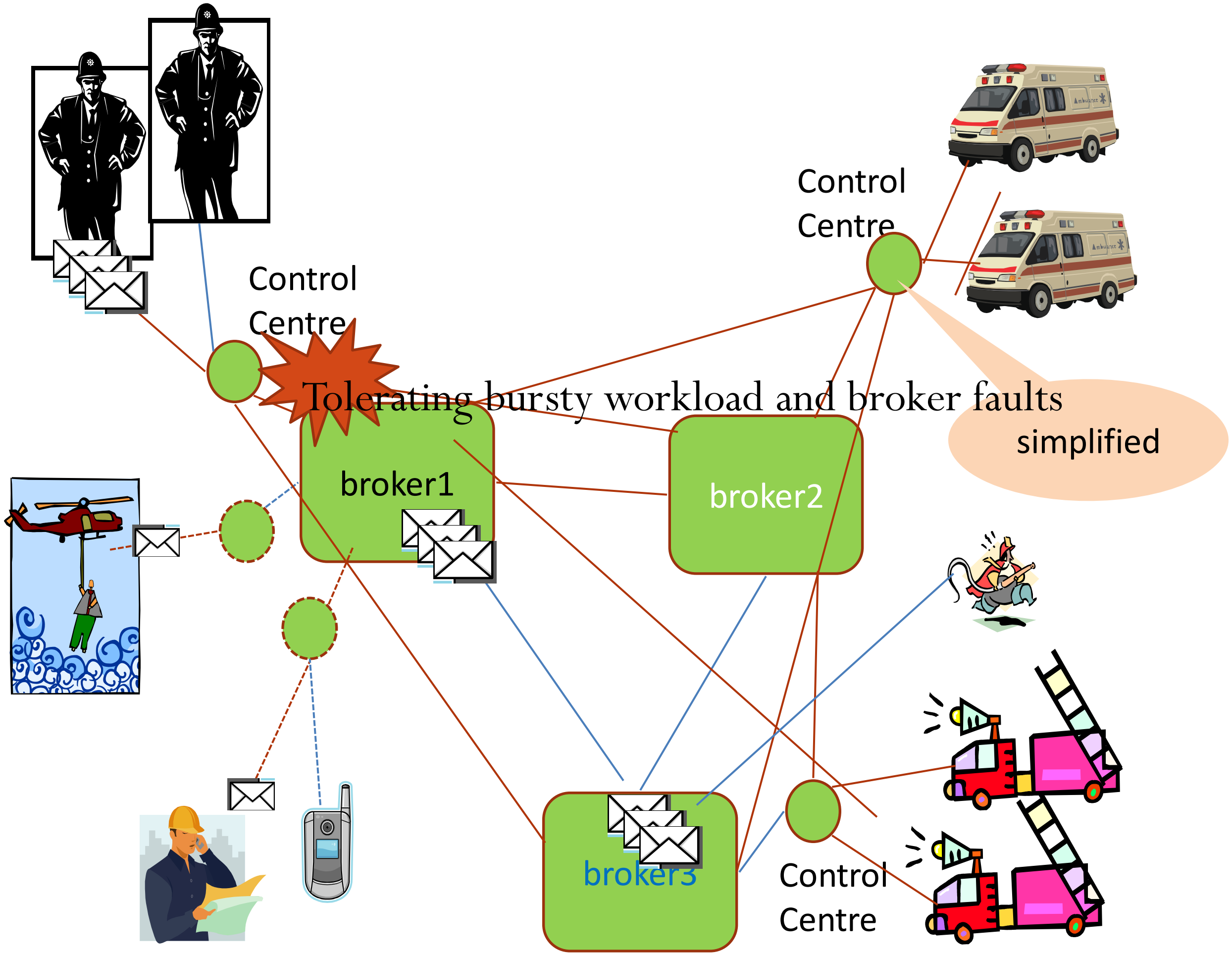
Possible structure of an overlay that can be used to provide resilience

Redundant Mirror is introduced on a disjointed path for important communications



Possible structure of an overlay that can be used to provide resilience

After the recovery, service is switched back from mirror to primary broker



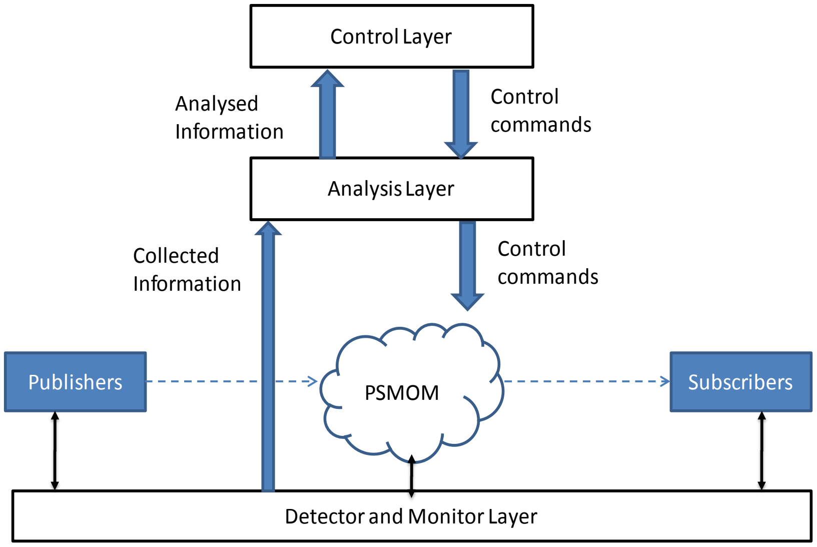
Possible structure of an overlay that can be used to provide resilience

Our Approach towards a resilient MOM system – Current work

- Tolerating bursty workload and broker faults while
  - Quantifying the risk of from super additive effects of correlated workload
  - Finding solutions that distribute workload in the system to minimize the risk of exceeding capacity
  - Finding solutions to introduce redundancy to broker failure by mirroring critical workload to other brokers
- A federated MOM system
  - Some applications require federations rather than just clusters
  - Also increased geographical disjointedness of brokers can result in better resilience compared to just using clusters.

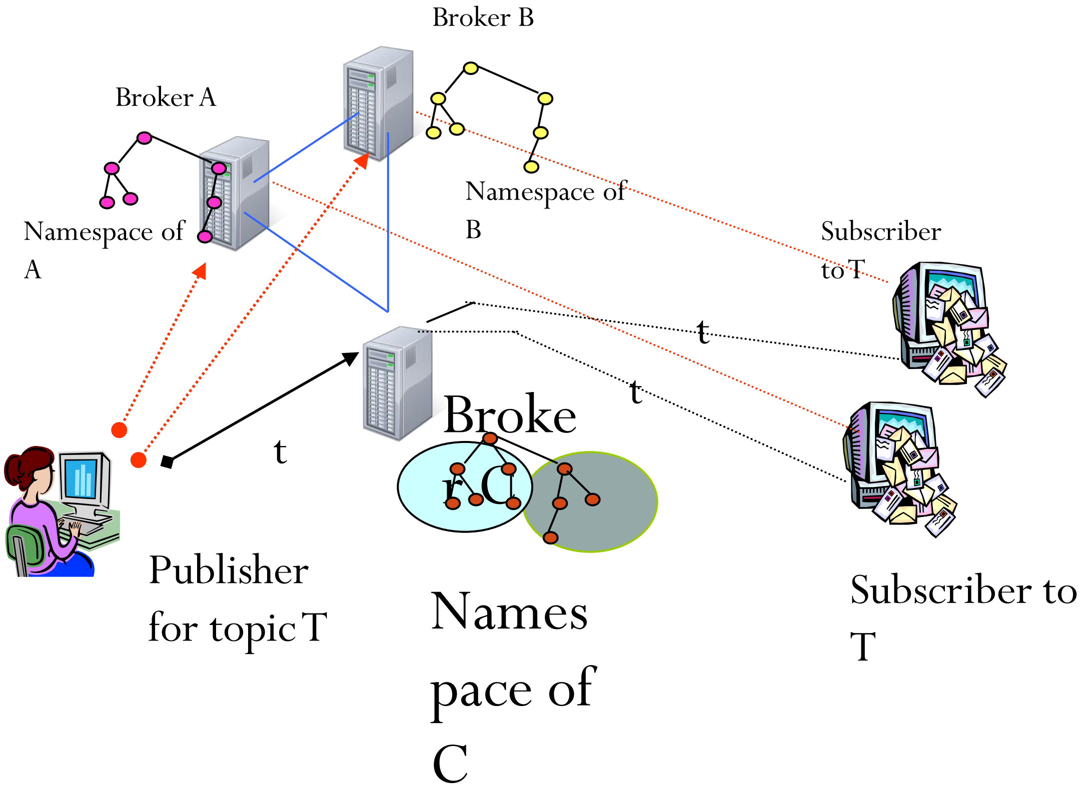
Architecture

- Control – Determines mirror policy by the approach described, augmented by Case Based Reasoning
- Analysis – e.g. context determination such as broker faults or network faults, bottleneck detection using Markov models
- Detector and Monitor – network conditions, broker conditions

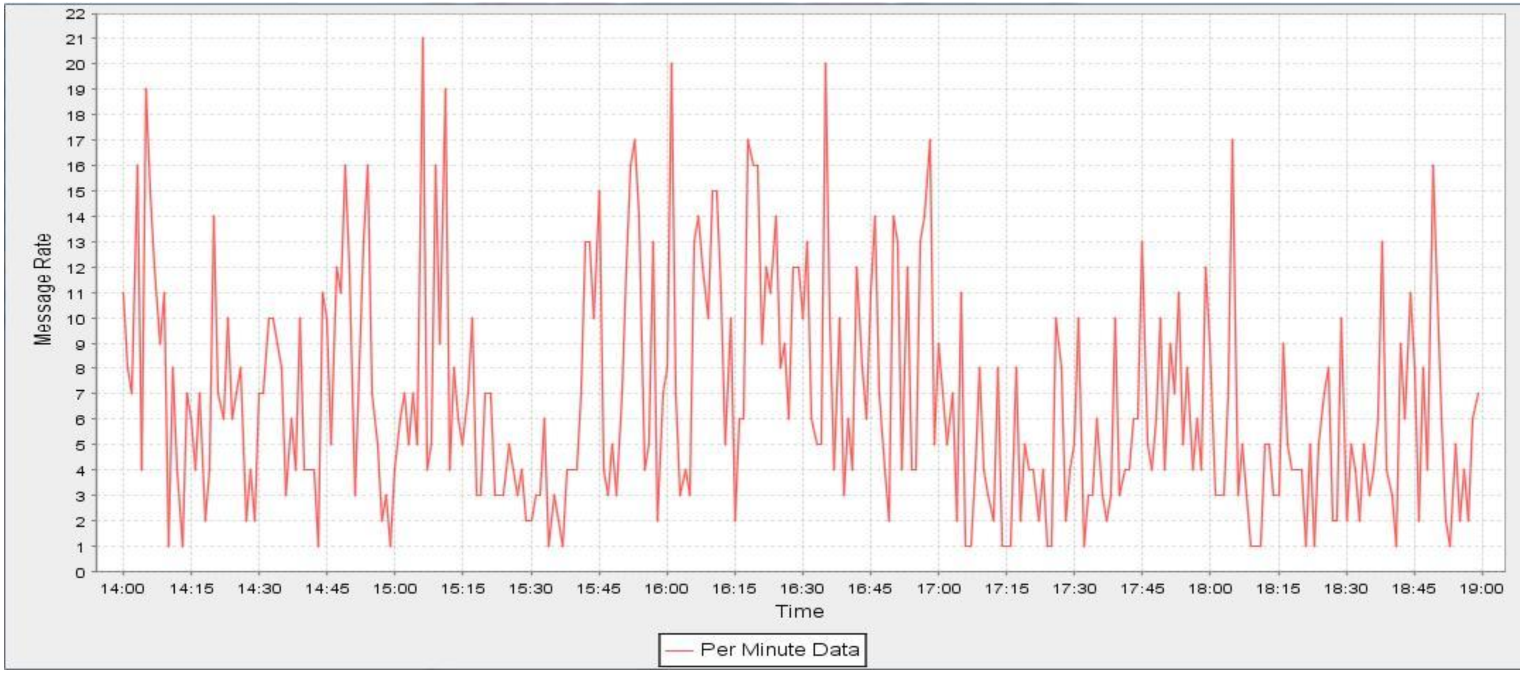


Tolerating bursty workload and broker faults

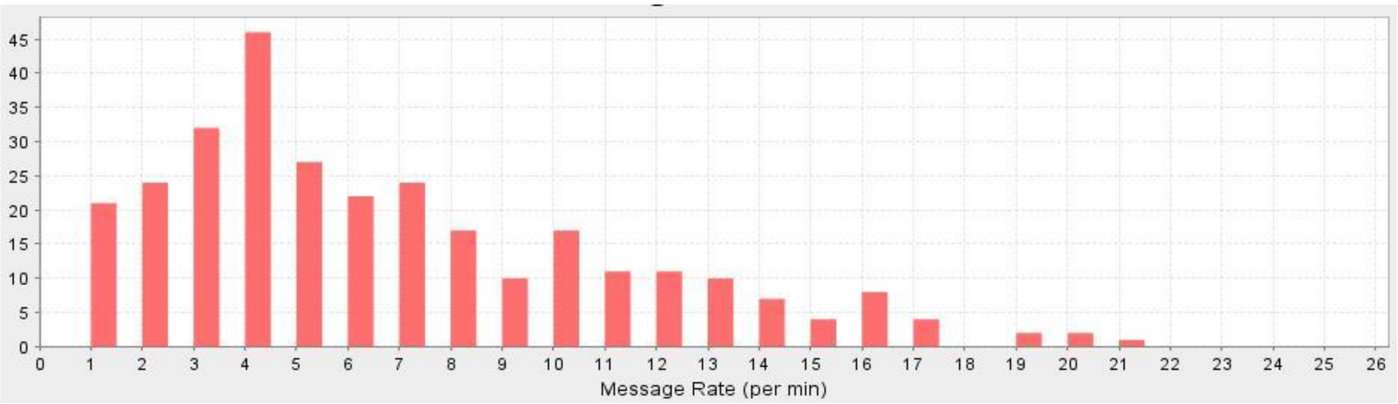
- Problem Formulation
  - Workload from different sources are divided into items.
  - In GEMOM workload is represented by namespace trees



The message rate



The distribution of the message rates



Allocations policies with Quantified Risk

- Allowing for the Super Additive effect of possibly correlated workload (items)
- Method
  - Use an offline analysis to build a Variance Covariance Matrix from the captured the message rate traces of different individual workloads (items)
  - Hence can estimate the variance and mean of combinations of items of workload
  - Hence can compute the probability of exceeding nominal capacity for any combination of items
    - If exceed bounds for broker reject
  - Conduct search through allocations of sets of items to each broker and choose allocation with minimum sum of probabilities
    - Different search algorithms employed
- Allocations form the mirror policy

Planning for Workload Allocation Policies

- Constraints are use in to prune search space in planning solutions
  - Allocate primary workload – constrains, e.g. proximity from source of workload, risk upper bound of workload exceeding the capacity
  - Mirrored workload – constrains e.g. the disjointness of path comparing with mirrored workload, maximum distance to mirror

- and select the solutions to adapt problems
  - E.g. Current CPU and Memory in brokers

- Constrains express concerns of different application

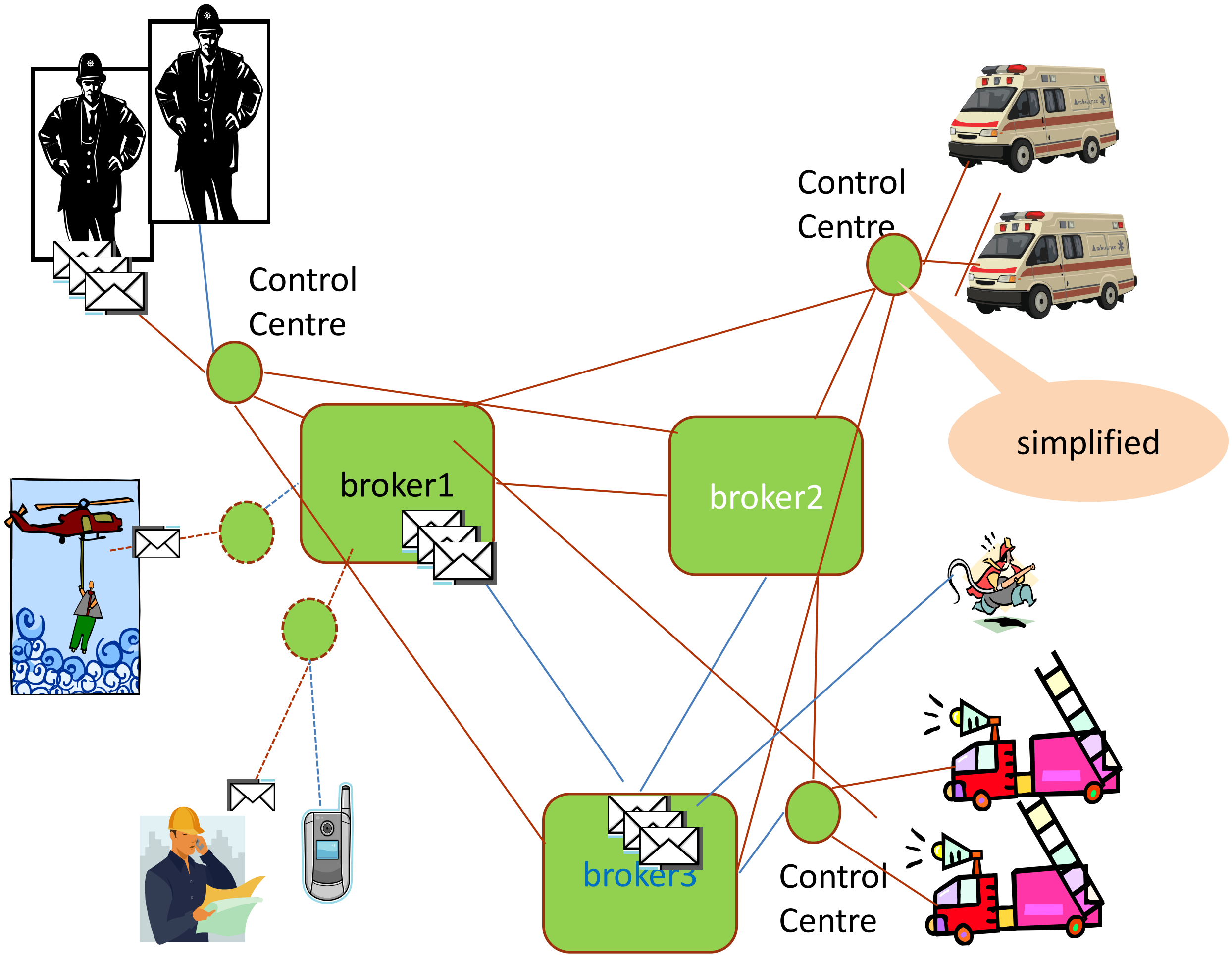
Utility in planning solutions

- Example case of single failure of broker
- Utility of allocating a set of items i to broker1

$$EP = R_1^i - P^i P_r(V_1^i > C)$$

  - C is the capacity of broker 1
  - R is the revenue for carrying the messages (up to the limit C)
  - P is the penalty for exceeding the capacity
  - V is the message rate (or data rate) at broker 1 of the items in set i PLUS existing allocation to broker 1
    - Assume V is normally distributed and var(V) computed from variance covariances of items in item set I and the existing allocated items
- Evaluated at each node of the search through the space of item sets

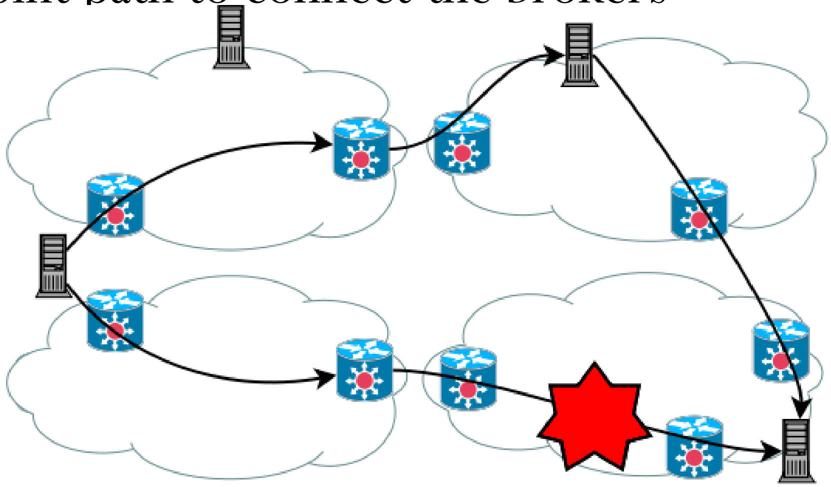
Allocation of primary and mirror workload



Possible structure of an overlay that can be used to provide resilience

Our approach towards a resilient MOM system - Future work

- Use overlay of brokers over internet, to response to disruptive link faults
  - Without assuming we have control over the underlay network
  - E.g. BGP faults takes several minutes or more to re-converge and recover
  - Using a relay broker on a disjoint path to connect the brokers disrupted by link failures



Our approach towards a resilient MOM system - Future work

- Defence against DDoS attacks
  - Diverting traffic from the compromised broker to its mirrors
  - Detect and regulate anomalous traffic
    - Session establishing process
    - Established flows

Validation And Improvement

- Case Studies
  - Financial data analysis
  - Local and international bank money transfer
  - Car toll booth messaging and traffic monitoring
  - Service market place for eGovenment
  - Resilient messaging in emergency context

Relevant Projects

- GEMOM
- RON (Resilient Overlay Network) - MIT
- ResiliNets – KU

Thank you for your attentions

Questions and Suggestions