MTP-2:

S.E.R.O. Reliable Multicast

Carsten Bormann (cabo@informatik.uni-bremen.de)

Universität Bremen

Jörg Ott, Hans-Christian Gehrcke,
Torsten Kerschat, Nils Seifert

Technische Universität Berlin

MTP-2: S.E.R.O. Reliable Multicast

The Quest for a Reliable Multicast Protocol

– Benefits of ubiquitous multicast transport
  - cf. TCP: may not always be optimal, used anyway
  - must be simple to use, understandable

- ***Target***

– Make use of IP multicast

– Optimize for groupware

- ***Objectives***

– Scalability
  - Can’t use linear algorithms

– Efficiency
  - Will always be compared to TCP

– Robustness
  - Will always be compared to TCP
  - Must be much *more* robust for multicast

– Ordering
  - Make application programmer’s life easier

MTP-2: S.E.R.O. Reliable Multicast

MTP

- Multicast group “web”
- Masters, Producers, Consumers
- ***Global ordering***
- Sequencing by *master* (coordinator)
- Sequence number assigned in *token*
- ***Rate control***
- Negotiate web *rate*
- Master associates tokens with rate
- ***Atomicity***
- *Message acceptance* by master
- Disseminated in web
- ***Scalability***
- *Negative* acknowledgements (unicast)
    - Producer retransmits (cuts into rate budget)
- Divide time into *heartbeats*
- Producers keep packets for *retention* heartbeats

MTP-2: S.E.R.O. Reliable Multicast

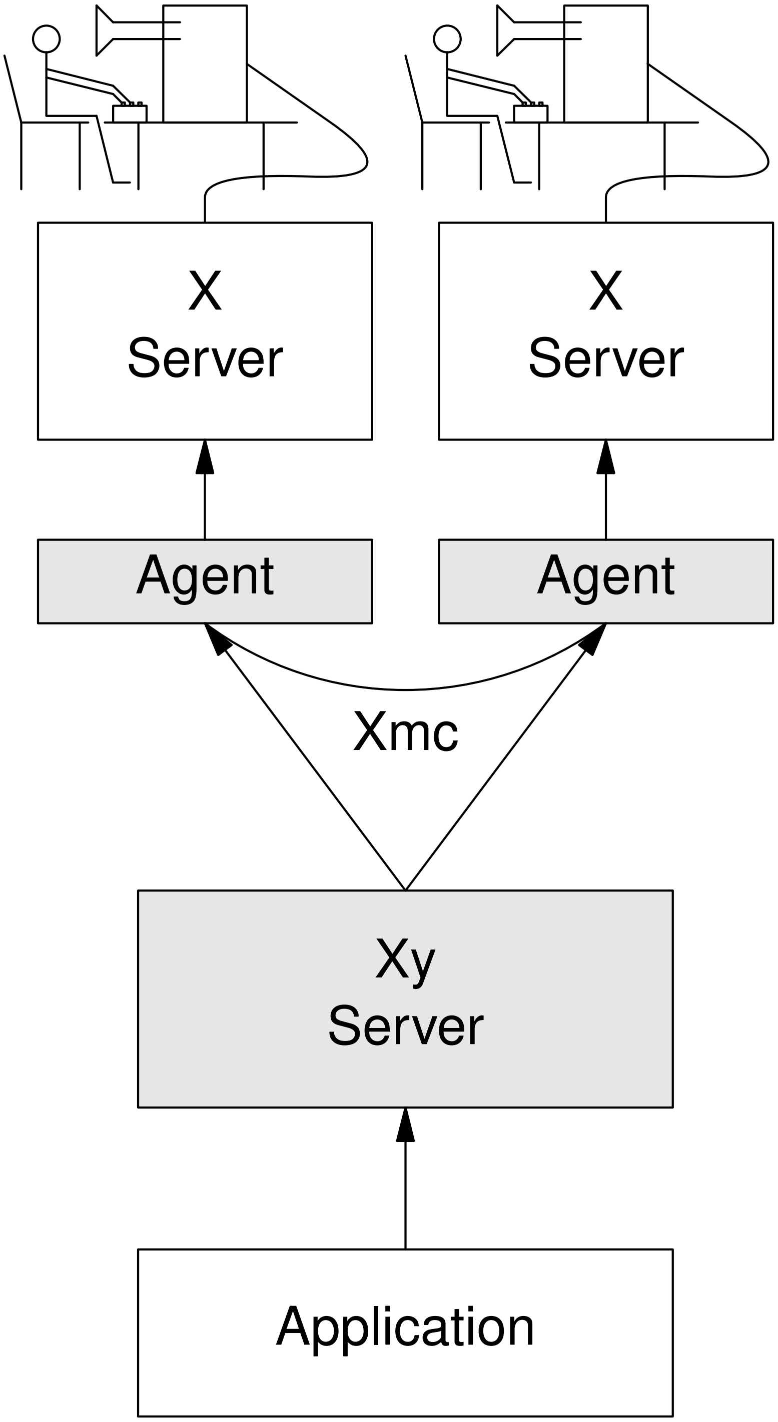
MTP-2

- ***Efficiency***
- – Less reliance on saturation
  - Request retransmission from unknown producer
- – Piggybacking token requests
- ***Delay***
- – Wait for token
  - RTT to master: MTP “feature”
  - Priority for token assignment
- – Ordered Delivery
  - Unsequenced messages
  - Group messages into streams
- ***Master***
- – Single point of failure
  - Master loss recovery
- – Should be able to change voluntarily
  - Master migration
- ***Unicasting***
- – Implicit connection between any pair of members

MTP-2: S.E.R.O. Reliable Multicast

Applications

- ***Xy: X window sharing***



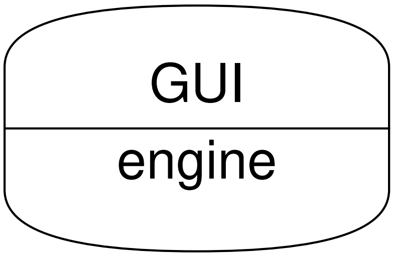
Applications

Fortsetzung

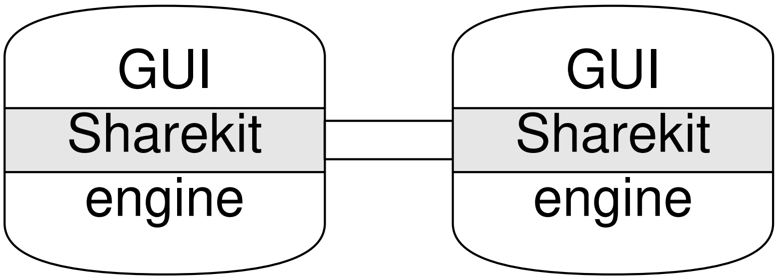
- ***Sharekit: application replication***

  "Hamburger" model of groupware generation

– Before (naïve application)



– After (cooperation-aware application)



  - distribute input events (MTP)
  - provide for initial status exchange
    (triggered by membership control)

MTP-2: S.E.R.O. Reliable Multicast

Status

– MTP implemented & tested (user mode daemon)
  • Part of Xy snapshot release
– MTP-2 implemented (2 versions: user and kernel)
  • Testing under way
  • ~ 10000 lines of code
  • Socket style interface (SunOS 5)
  • Successful use in Sharekit
  – “1 afternoon” conversion time

• ***To do***

– Larger scale testing
  • Validate protocol
  • Validate implementation
  • Validate usability
– Prove recovery algorithms
– Devise method for adaptive parameter changes
– Integrate authentication

