Architectural Considerations for Providing Carrier Class Telephony Services Utilizing SIP...

W. Marshall, K. K. Ramakrishnan, E. Miller, G. Russell, B. Beser, M. Mannette, K. Steinbrenner, D. Oran, J. Pickens, P. Lalwaney, J. Fellows, D. Evans, K. Kelly, F. Andreasen

AT&T, CableLabs, 3Com, Cisco, Com21, General Instrument, Lucent Cable, NetSpeak, Telcordia

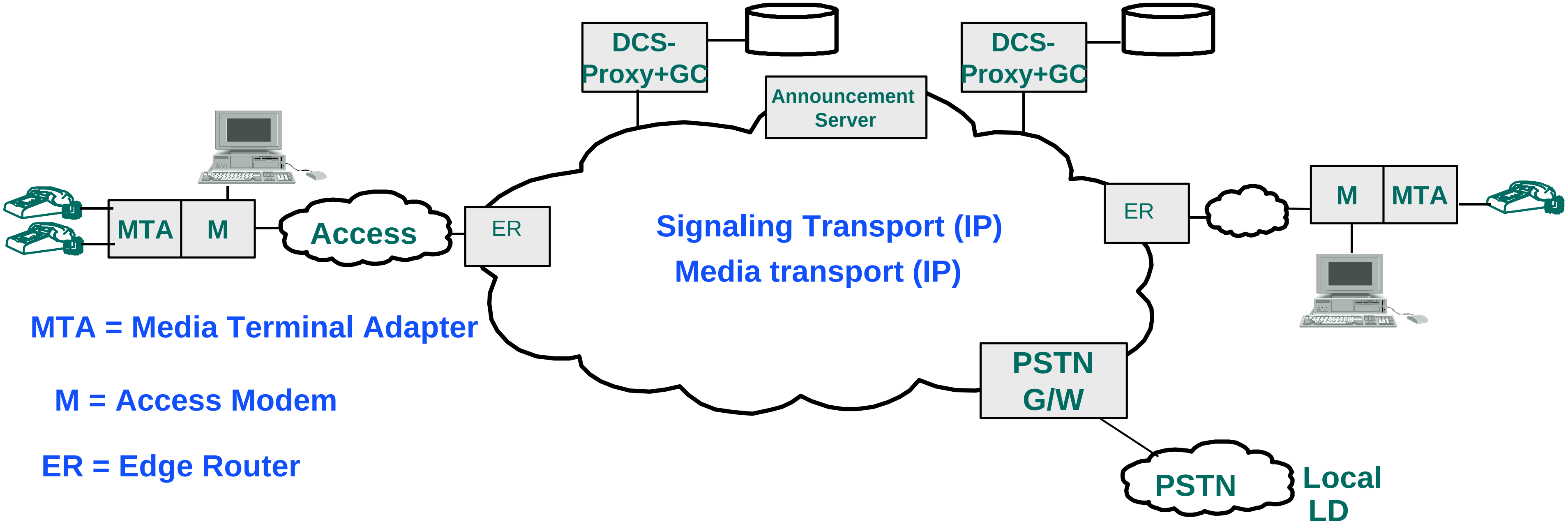
July 1999

IETF Presentation

Requirements from a Service Provider’s Perspective

- Need for differentiated quality-of-service is fundamental
  - must support resource reservation and admission control, where needed
  - hope SIP enables lots of new services; also desire to meet needs of current users
- Allow for authentication and authorization on a call-by-call basis
- **Can’t trust** CPE to transmit accurate information or keep it private
- Need to guarantee privacy and accuracy of feature information
  - e.g., Caller ID, Caller ID-block, Calling Name, Called Party
    - privacy may also imply keeping IP addresses private
- Protect the network from fraud and theft of service
  - critical, given the incentive to bypass network controls
- We must be able to operate in large scale, cost-effectively
  - don’t keep state for stable calls in proxies; end-points can keep state associated with their own calls

Distributed Call Signaling Framework



DCS-Proxy+GC

DCS-Proxy+GC

Announcement Server

MTA

M

Access

ER

Signaling Transport (IP)

Media transport (IP)

ER

M

MTA

PSTN G/W

PSTN

Local LD

MTA = Media Terminal Adapter

M = Access Modem

ER = Edge Router

- Designed as a complete end-to-end signaling architecture for PacketCable
  - Philosophy: encourage features and services in intelligent end-points, wherever technically and economically feasible
  - “DCS-Proxy” designed to be scalable transaction server
  - Resource management protocol provides necessary semantics for telephony
  - “Gates” (packet classifiers) at network edge allow us to avoid theft of service

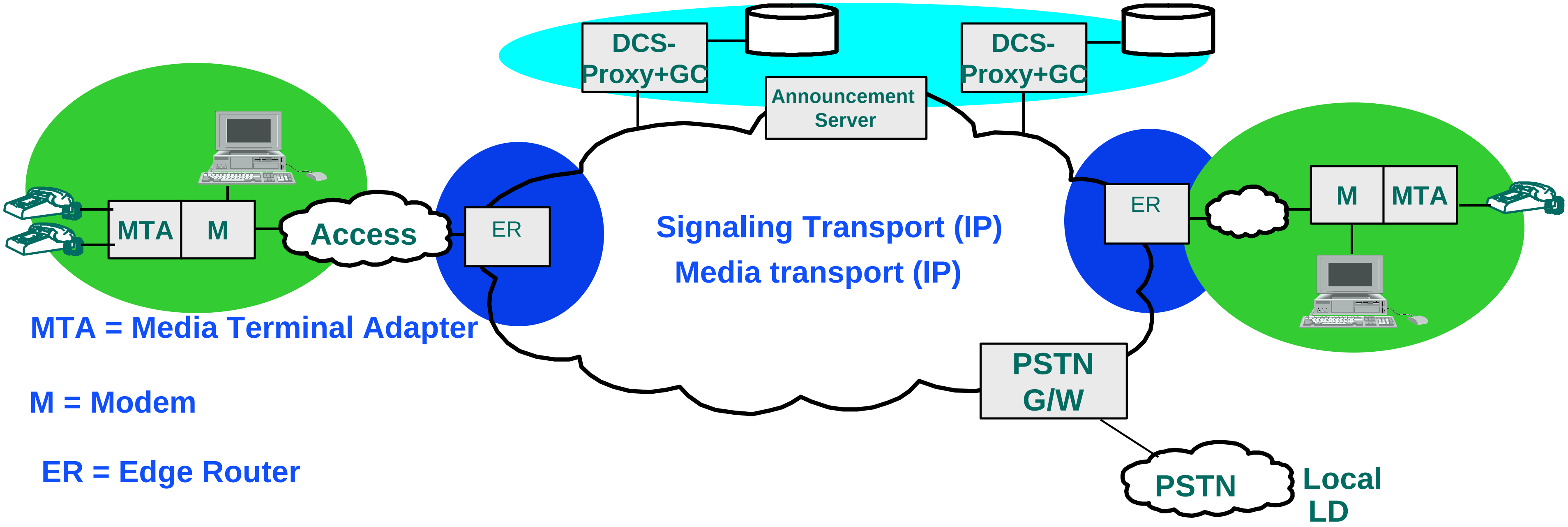
DCS Architecture

- Enhances SIP With Carrier Class Features
  - Resource Management
  - Privacy
  - Authorization and Theft of Service issues
- Tight Coupling Between Call Signaling And QoS Control
  - Prevent Call Defects: don’t ring the phone if resources are unavailable
  - Prevent Theft Of Service: associate usage recording and resource allocation, ensuring non-repudiation
    - » provide the ability to bill for usage, without trusting end-points
    - » ensure quality requirements for service are met (e.g., don’t clip “Hello”)
- Care taken to ensure untrusted end-points behave as desired
- Privacy mechanisms built into architecture

DCS Architecture

- Makes use of end-point intelligence
  - useful from the point of view of new feature creation
- Distribution of state
  - Clients keep Call State
  - Edge Routers keep Connection State
  - DCS-Proxy only keeps Transaction State
- Failure model minimizes service impacts due to component outages

DCS Architecture



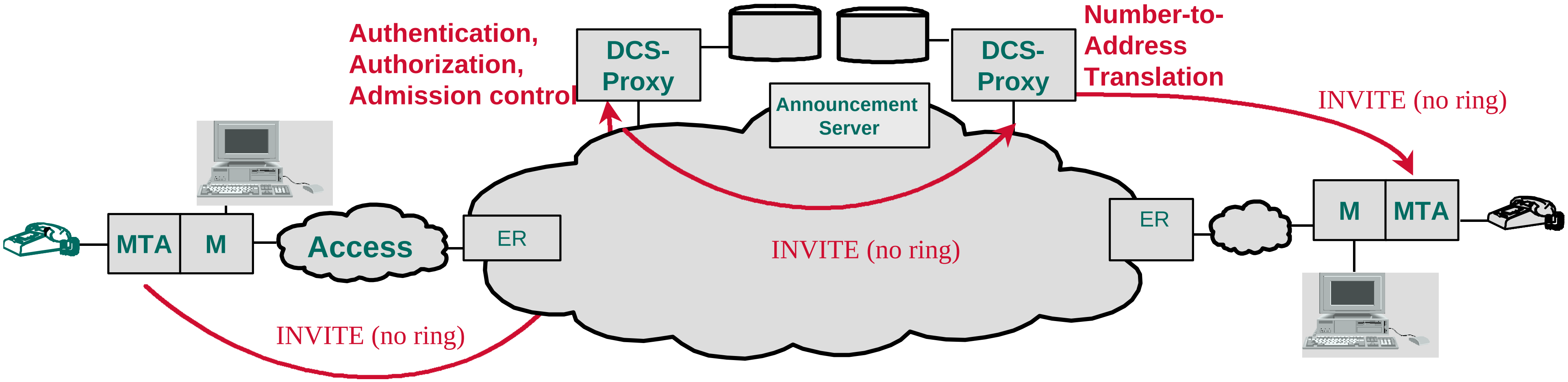
MTA = Media Terminal Adapter

M = Modem

ER = Edge Router

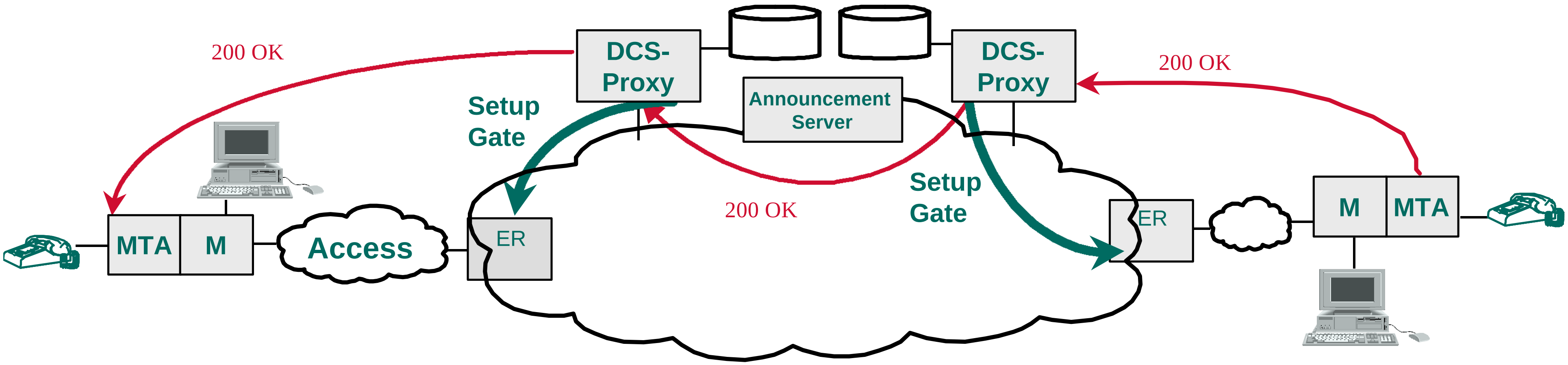
Call State **Connection State** **Transaction State**

Example Call Flow



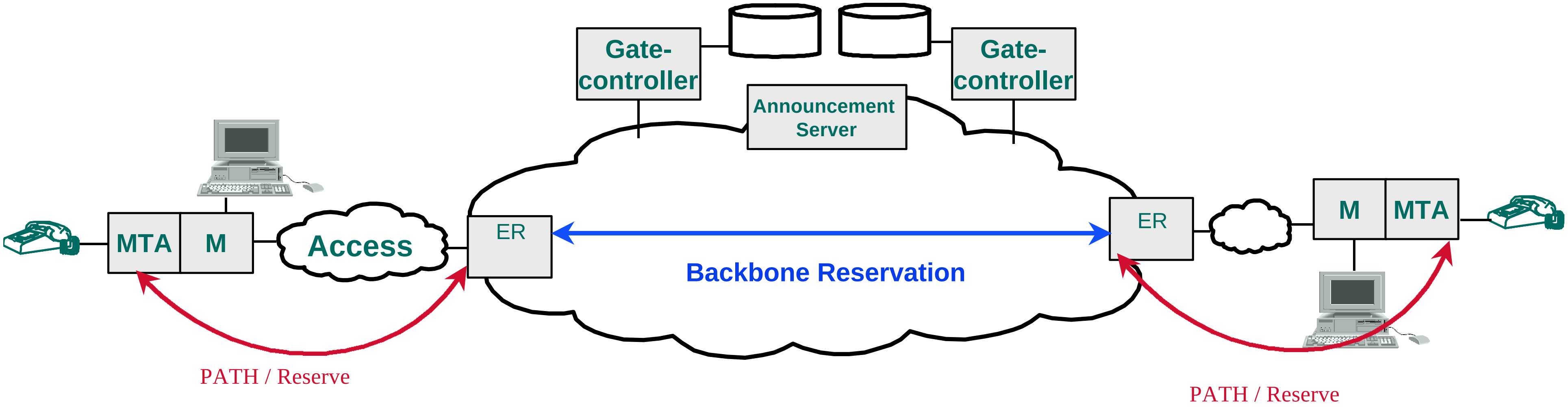
- ◆ MTA issues an INVITE to destination E.164 (or other) address
  - – don’t know yet “what” resources are needed to “where”
  - – provider may choose to block a call if resources are unavailable
    - » but P(blocking) may be ≥ P(call defect)
      - ➪ call defect: when the call fails after the parties are notified
- ◆ Originating DCS-proxy performs authentication and authorization
- ◆ Terminating DCS-proxy translates dest. number to local IP address

Example Call Flow (contd…)



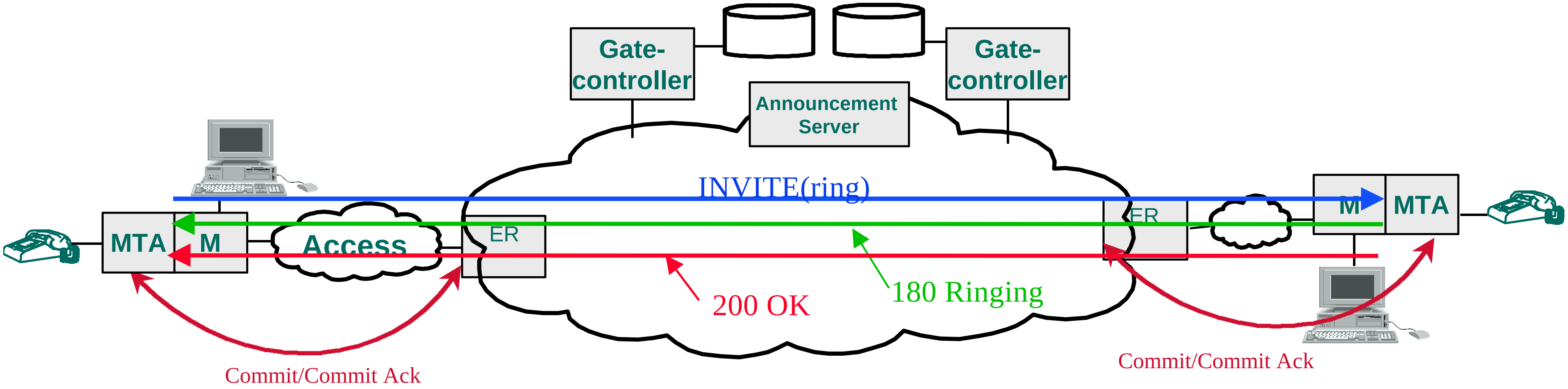
- 200 OK communicates call parameters and gate identity to MTA
- Gate controllers setup “gates” at edge routers as part of call setup
  - gate is described as an “envelope” of possible reservations issued by MTA
  - gate permits reservation for this call to be admitted
- Policy may be exercised either at Gate controller or associated policy server

Resource Management: 1st Phase



- MTA initiates resource reservation
  - access resources are “reserved” after an admission control check
    - » this insures that resources are available when terminating MTA rings
  - backbone resources are “reserved” (e.g., explicit reservation or “packet marking”)
- Originating MTA starts end-to-end handshake with terminating MTA
  - originating MTA sends INVITE(ring), terminating MTA sends 180 RINGING, 200 OK

Resource Management: 2nd Phase



- MTA knows voice path is established when it receives a 200 OK
- MTAs initiate resource "commitment"
  - resources "committed" over access channel
    - CMTS starts sending unsolicited grants; usage recording is started
  - commitment deferred until far end pick up, to prevent theft of service; allow efficient use of constrained resources in access network
- Commit opens the "gate" for this flow

Signaling Performance Requirements

- Short post-dial delay
  - no perceptible difference in post-dial delay compared to circuit-switched network
- Short post-pickup delay
  - delay from when the user picks up a ringing phone and the voice path being cut-through should be small
    - called party's "hello" must not be clipped
    - calling party's response to hearing the "hello" must also not be clipped
- Probability of Blocking: a metric to which provider may engineer net
- Probability of Call Defect (i.e., call that has both parties invited to and then fails) due to lack of resources needs to be much smaller
  - target rates not necessarily under the control of the provider
- Flexibility in deployment of DCS-Proxy: start small.

DCS: Profile With Extensions

- 6 Internet Drafts have been submitted:
- Architectural Overview
  - » draft-dcsgroup-mmusic-arch-00.txt
- Resource Management And Call Authorization
  - » draft-dcsgroup-mmusic-resource-00.txt
  - » draft-dcsgroup-mmusic-call-auth-00.txt
- Privacy
  - » draft-dcsgroup-mmusic-privacy-00.txt
- Proxy-to-Proxy Communications
  - » draft-dcsgroup-mmusic-proxy-proxy-00.txt
- Distributed Call State
  - » draft-dcsgroup-mmusic-state-00.txt