

# CATS BOF

IETF 53 – Minneapolis, MN

March 2002

`echo subscribe mrcp | mail majordomo@snowshore.com`

`http://flyingfox.snowshore.com/mrcp\_archive/maillist.html`

# Note Well

- All statements related to the activities of the IETF and addressed to the IETF are subject to all provisions of Section 10 of RFC 2026, which grants to the IETF and its participants certain licenses and rights in such statements. Such statements include **verbal statements in IETF meetings, as well as written and electronic communications made at any time or place**, which are addressed to:
  - the IETF plenary session
  - **any IETF working group or portion thereof**
  - the IESG, or any member thereof on behalf of the IESG
  - the IAB or any member thereof on behalf of the IAB
  - **any IETF mailing list, including the IETF list itself, any working group or designated team list, or any other list functioning under IETF auspices**
  - the RFC Editor or the Internet-Drafts function
- Statements made outside of an IETF meeting, mailing list or other function, that are clearly not intended to be input to an IETF activity, group or function, are not subject to these provisions.

# Agenda

- Agenda Bashing & Purpose of BOF 5 min
- Distributed Control of Specialized Voice Services  
Problem Statement 15 min
- Prior Approaches 10 min
- Proposed Charter 15 min
- Work Plan 15 min
- Finish Session Early!
  
- Speaker Verification Tutorial 20 min

# Purpose of BOF

- Consensus on Need for Work to be Done in the IETF
- Consensus on Charter

- Mail List

```
echo subscribe mrcp | \  
mail majordomo@snowshore.com
```

- Archive

```
http://flyingfox.snowshore.com/  
mrcp_archive/maillist.html
```

# Problem Statement

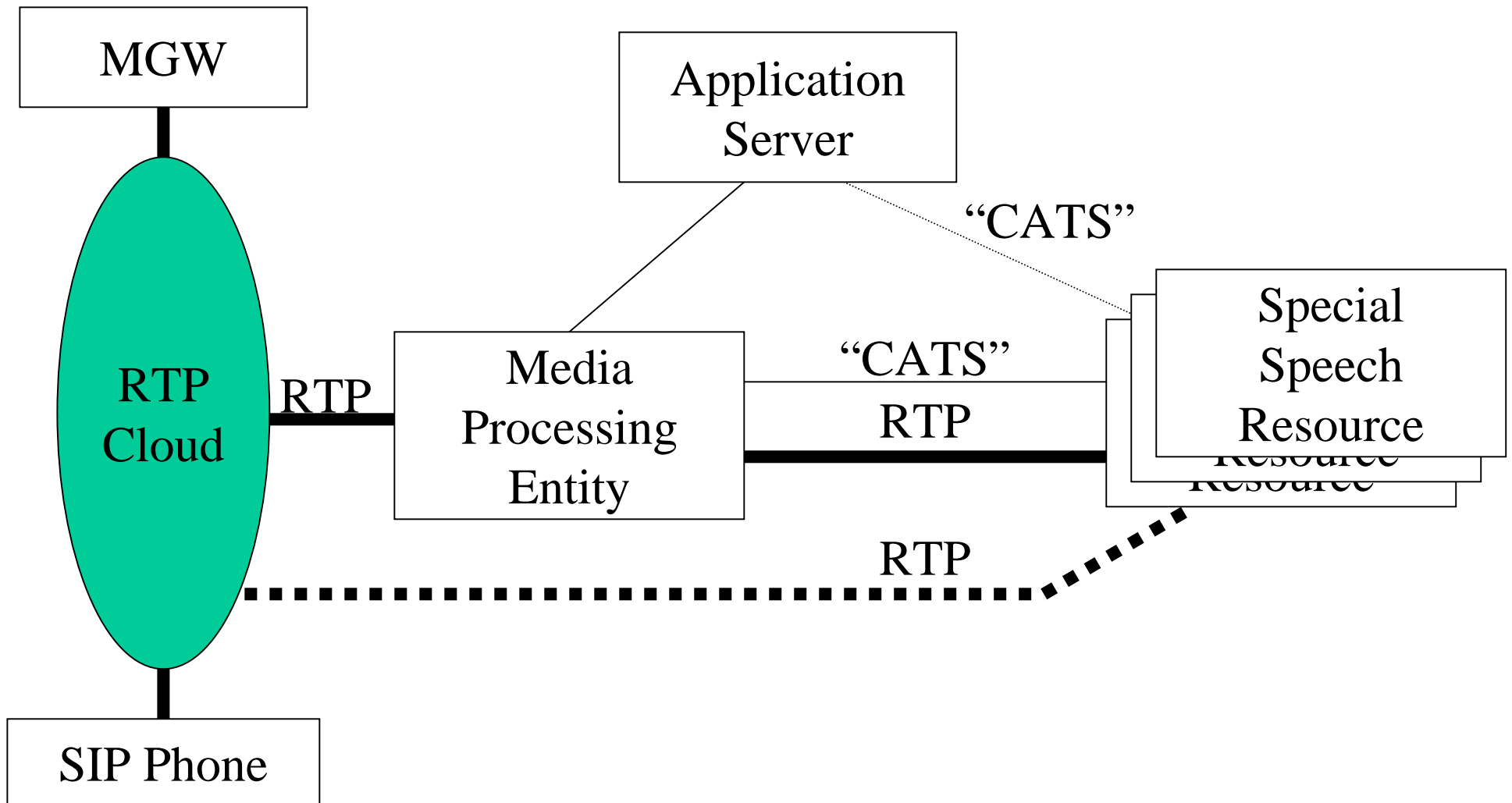
- Distributed, Specialized Voice Processing Services
  - Automatic Speech Recognition (ASR)
  - Text-To-Speech (TTS)
  - Speaker ID/Verification (SV)
- Not ETSI Aurora DSR

# Requirements

`draft-burger-mrcp-reqts-00`

- Sets Out Problem Statement
  - NOT a Protocol, per se
- Framework Called SRCP
  - Before Official Name from AD's
    - MRCP Implied Endorsement
    - Had to Pick Something for Document
  - Name of Framework May Change (CATS?)
  - Alternate Choice: SPEECHSC  
(Speech Services Control)

# Framework (Proposed)



# General Requirements

- Reuse Existing Protocols, Where Sensible
  - Conventions of Use
  - Extensions
  - Something New for Something Different
- Guiding Principle
  - Will Not Jam Something New  
Into Something Old  
If Not Sensible



# TTS Requirements

- Plays Back Text
  - Plain, User Text
    - UTF-8
    - “Human Text Strings”, Per RFC2277
    - Language Identifier, Per RFC3066
  - SSML
  - Others
- Open Issues with List Input:
  - Explicit Text Type
  - Fetch from URL
  - Speech Markers

# Open TTS Issues

- Long-Lived Connections
  - Could they be required?
- VCR Controls
  - Yes; do engines support it?
- Session Parameters (What is a Session)
- Text Over Control Channel? [new]
  - Yes

# ASR Requirements

- Recognizes Speech
- W3C XML Form of the Speech Recognition Grammar Specification
- Static Grammars (ex. Protocol?)
- Utterance Capture

# Open ASR Issues

- ABNF Form
  - Yes
- Session Parameters (What is a Session)
- Requirements for Utterance Capture
  - Simple Indicator for Engine Magic?
  - Protocol Machinery for Streaming Media?
  - Protocol Machinery for File?

# Speaker ID/Verification

- *Dan Burnet*

# Low Latency

- Energy or Speech Detection to Prompt Cut-Off (Barge)
- Critical Human Factors Issue
- “Answer” Has Been Multi-Mode Servers
  - *e.g.*, ASR **and** TTS on Same Server

# Prior Approaches

- Proprietary APIs
- MRCP

# Proposed Charter Specifics

- Develop One or More Protocols Between a Client and a Collection of Specialized Voice Servers, to Serve
  - Speech Recognition (ASR)
  - Text-to-Speech (TTS)
  - Speaker ID/Verification (SV)



# Out-of-Scope Items

- Distributed Speech Recognition (*e.g.*, ETSI Aurora DSR)
- Control of Arbitrary Media Processing Resources (*e.g.*, fax, announcements, recording voice)
- W3C-Domain Activities (*e.g.*, Markup)

# Proposed Charter Goals/Methods

- Satisfy Needs of Distributed Control of ASR, TTS, and SV Servers, As Described in `mrcp-reqts`
- Work With W3C Voice Browser and Multi-Modal Interaction Work Groups on Their Needs and Our Approaches
- Propose Requirements to MMUSIC and Other Groups for Changes to Core Protocols
  - For example, Changes to RTSP
- Create Protocol Extensions or New Protocol Requirements, If Necessary
  - For example, to Satisfy Speaker Verification Requirements

# Work Plan

- Jul 02 Publish Updated Requirements Document
- Dec 02 Publish I-D's Analyzing Existing Protocols for Suitability
- Dec 02 Publish I-D's With Requirements for Core Protocols, and/or
- Dec 02 Publish I-D's for New Protocols
- Mar 03 Drafts to IESG

# Thanks!

Contact: / mail list / archive

Eric Burger <mailto:eburger@snowshore.com>

David Oran <mailto:oran@cisco.com>