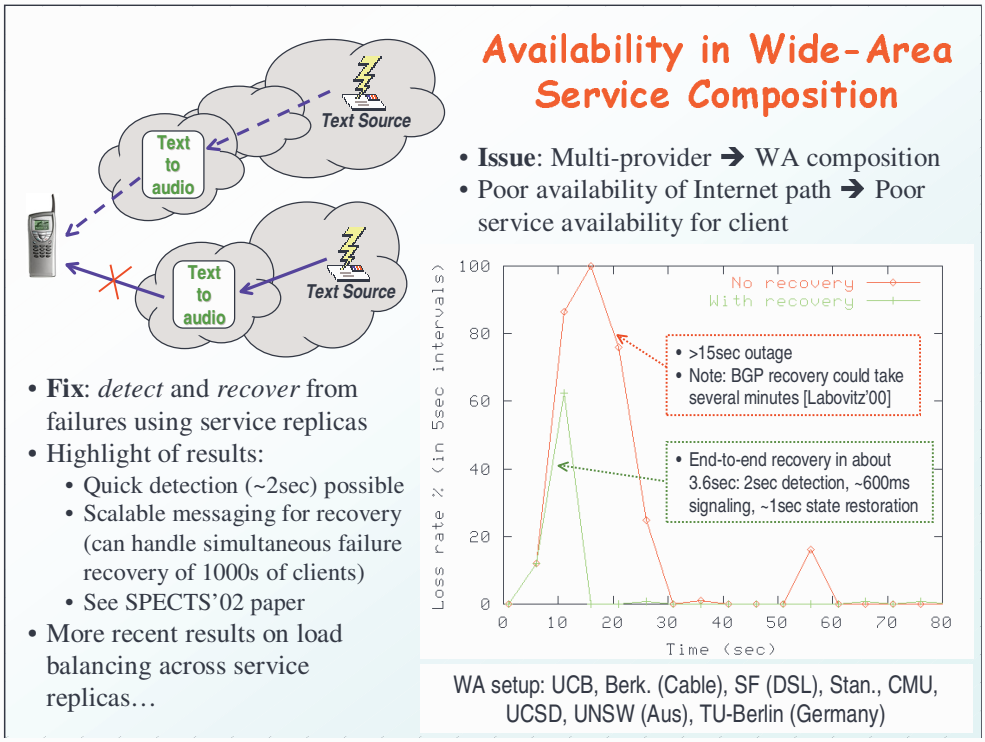

-Composition across providers implies path could stretch across the wide-area

-For instance, the picture shows a service involving a text-source such as email, and a text-to-speech engine

-Wide-area Internet path availability is not great (studies by Labovitz, et.al.)

-This means poor availability for the composed service

-Make use of service replicas to *dynamically* switch from one service instance to another

-We have shown two things:

-Quick failure detection makes sense (within about 2sec), using aggressive heart-beats

-Scalable messaging – when 1000s of client sessions have to restored simultaneously, system does not break down due to message flood

-More details in SPECTS'02 paper

-The graph shows an experiment we ran across the wide-area, across 8 hosts

-These hosts represent university hosts in US, commercial end-points, as well as trans-continental links

-There are two client sessions of the composed text-to-speech application: one with recovery mechanism enabled, one without

-X-axis shows time, as the sessions proceed

-Y-axis shows the loss-percentage of audio packets received at the end-client, computed over 5sec intervals

-The session without any recovery mechanism sees an outage of over 15sec

-Due to recovery, the green line recovers in about 3.6sec (within bounds of end-client buffering)

-We have also studied algorithms for load-balancing across service replicas, in this context of dynamic session recovery to improve availability