

# **Programm- und Projektplanung per Videokonferenz in Max-Planck-Gesellschaft und Helmholtz-Gemeinschaft – Eine Frage der Qualität**

Bunne, S.<sup>1</sup>; Pfeiffenberger, Dr. H.<sup>1</sup>; Schwenn, Dr. U.<sup>2</sup>; Stöckigt, K.<sup>2</sup>

<sup>1</sup> Alfred Wegener Institut (AWI), Bremerhaven, HGF;

<sup>2</sup> Max-Planck-Institut für Plasmaphysik (IPP), Garching, MPG

sbunne@awi-bremerhaven.de  
hpfeiffenberger@awi-bremerhaven.de  
schwenn@rzg.mpg.de  
kfs@rzg.mpg.de

**Abstract:** Intensive Tests und Betriebserfahrungen an AWI und IPP erlauben die Aufstellung von Qualitätsanforderungen für Videokonferenzen als Arbeitsmittel in der Programm- und Projektplanung in HGF und MPG. Diese werden gegliedert nach Anforderungen, medientechnischer und netzwerktechnischer Realisierung detailliert beschrieben. Empfehlungen für weitere Aktivitäten werden gegeben.

## **1 Einleitung**

Nach extensiven Tests im Jahr 2003 haben sich die beiden Wissenschaftsorganisationen MPG und HGF entschlossen, Rahmenverträge mit dem DFN abzuschließen, um dessen Mehrwertdienst „DFNVC“ zu nutzen. Die Hauptanwendung der Videokonferenz wird in beiden Organisationen in der Reduzierung der Reisetätigkeit leitender Wissenschaftler bei der Programm- und Projektplanung gesehen.

Die Anforderungen der sehr kritischen Zielgruppe von Wissenschaftlern einerseits und deren Abgleich mit den extrem vielfältigen Varianten und Optionen der Videokonferenztechnik andererseits bedürfte zur sorgfältigen Analyse im Grunde eines Buches. Davon haben sich die Autoren beim Aufstellen einer detaillierten Gliederung zu diesem Artikel überzeugt.

Im gegebenen Rahmen bleibt uns daher nur, die Notwendigkeit einiger Schlüsselforderungen und unserer Schlussfolgerungen bzw. Lösungen durch den Nachweis zu belegen, dass das Gegenteil bzw. andere plausible Optionen nicht funktionieren bzw. hinreichen ( wohl aber die von uns gewählten Lösungen ). Die zum Teil anekdotische Evidenz gerade im Hinblick auf Akzeptanz können wir in vielen Fällen schon deshalb nicht statistisch durch größere Stichproben absichern, da die vorgesehene Zielgruppe sich für solche Untersuchungen im Allgemeinen nicht erwärmen kann und die unfreiwillige Teilnahme an einer qualitativ unzureichenden Konferenz mit hoher Wahrscheinlichkeit direkte Auswirkungen auf die Finanzierung von Folgeversuchen hat.

Die angesprochenen Tests – insbesondere zur Qualität von Mehrpunktkonferenzen unter Nutzung des DFNVC-Dienstes – konnten nur mit Hilfe und Engagement einer großen Zahl von Kollegen aus den Infrastruktureinrichtungen unseren Forschungsorganisationen und der DFNVC – Administration durchgeführt werden, denen wir hiermit danken. Auch Ihre Zeit ist wertvoll und also begrenzt.

Letzten Endes also wird die „Frage der Qualität“ nur etwa so beantwortet werden können: Wenn in zwei bis drei Jahren die Nutzung der geschaffenen Infrastruktur nicht den Erwartungen entspricht, war entweder das Szenario falsch beschrieben oder die Qualität unzureichend. Dies ist sicher keine wissenschaftlich/methodisch zufrieden stellende Art, Systemoptimierung anzugehen. Es sind allerdings schon wichtigere Systeme (z.B. das IP-basierte Internet) entstanden, ohne dass a priori wissenschaftlich präzise untersucht worden war, ob diese für ihren letztendlichen Einsatz hinreichend wären.

Daher soll unser Artikel gerne so verstanden werden, dass wir versuchen, einen System- und Qualitätsstandard für einen bestimmten Auslegungspunkt zu „setzen“ und erhoffen, dass sich konkurrierende Systeme und Auslegungen daran messen und Diskussionen einen Ausgangspunkt erhalten.

Bei diesem Auslegungspunkt grenzen wir uns bewusst durch einen höheren Qualitätsanspruch gegenüber früheren "Best Practise" - Papieren ab. [1,2]

Dazu erfolgt im Kapitel 2 eine Beschreibung der Nutzungsszenarien und daraus abgeleitet, in Kapitel 3 eine Anforderungsanalyse mit übergeordneten Gesichtspunkten wie Benutzerprofil, Effizienz, Audio-, Video- und Präsentationsqualität. Die medientechnische Realisierung in den in dieser Arbeit hauptsächlich analysierten mittelgroßen Seminarräumen erfolgt in Kapitel 4. Hier werden Zuverlässigkeit, Ton-, Bild- und Präsentationsübertragungen, sowie Benutzeroberflächen durch Einsatz von Mediensteuerungen und Personalanforderungen definiert. Die Anforderungen an die Transportdienste - IP und für Sonderzwecke ISDN -werden in Kapitel 5 diskutiert. Auch Adressierung, Sicherheitsaspekte, Einsatz von Mehrpunktkonferenzen. und Techniken der Präsentationsübertragung werden hier beschrieben. In Kapitel 6 geben wir eine Übersicht der von uns als notwendig erachteten Weiterentwicklungen.

## 2 Nutzungsszenarien

An den 80 Instituten der Max-Planck-Gesellschaft und den 15 Forschungszentren der Helmholtzgemeinschaft mit ihren ca. 130 Standorten, die fast entsprechend der Bevölkerung über Deutschland verteilt sind, s.Abb. 1, arbeiten ca. 36.000 Mitarbeiter. Mit Etats von insgesamt 3,4 Milliarden € / Jahr bearbeiten sie Fragestellungen aus Grundlagen- und angewandter Forschung, häufig unter Einsatz von sehr kostspieligen wissenschaftlichen Infrastruktureinrichtungen, wie Beschleunigern, Fusions-Experimenten, Flugzeugen oder Schiffen. [3,4]

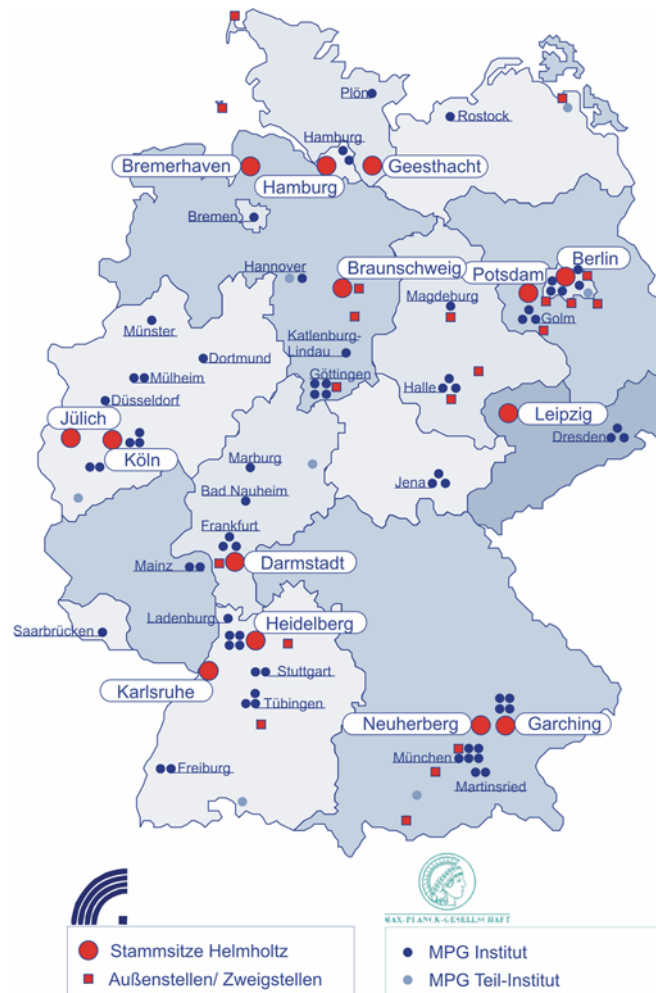


Abbildung 1 Standorte der Max Planck Gesellschaft und der Helmholtz Gemeinschaft

Um diese großen Mittel möglichst produktiv zu nutzen, haben Geldgeber (BMBF) und die Organisationen selber Strukturen innerhalb der Organisationen geschaffen, die intensive, flächendeckende Koordination und Kooperationsstrukturen voraussetzen.

Als Beispiel möge die Gliederung der Helmholtz-Gemeinschaft, nach ihrer noch voran schreitenden neuen Förderstruktur, dienen, s. Abb. 2. Unterhalb der Ebene der 6 Fachbereiche befinden sich ca. 30 Programme, die eine Laufzeit von 5 Jahren haben. Spezielle Themen von gemeinsamem Interesse werden darüber hinaus in etwa ebenso vielen Special Interest Groups bearbeitet.

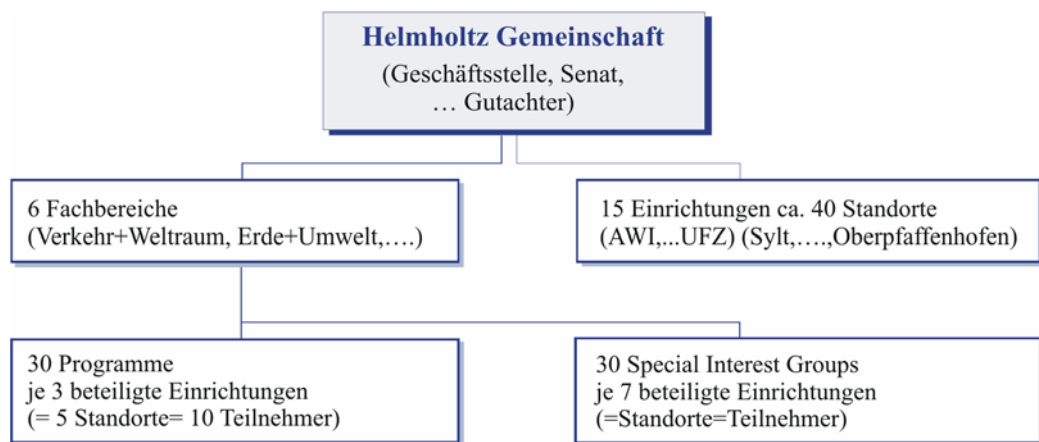


Abbildung 2 Gliederung der HGF

An Besprechungen zur (Detail-)Planung zukünftiger oder zur organisatorischen Abwicklung laufender Programme werden demnach je ein bis zwei Wissenschaftler von 3 bis 7 Standorten teilnehmen. Diese Gruppen werden sich ein- bis mehrmals pro Jahr treffen müssen.

Auf die Frage, ob Videokonferenzen den absehbar damit verbundenen großen Reise- und damit Zeitaufwand ersetzen können, kann man nur antworten mit einem deutlichen: „Nein – allenfalls deutlich reduzieren!“ Jedem, der an derartigen Treffen schon teilgenommen hat, ist klar, dass die Teilnehmer (oder die Fraktionen!) ein gewisses Mindestmaß an Smalltalk und anderen sozialen Interaktionen pflegen müssen, die sich in diesen Szenarien<sup>1</sup> nicht durch Videokonferenzen darstellen lassen. Jedoch kann die erhebliche Anzahl der physischen Treffen unter geeigneten organisatorischen Randbedingungen erheblich reduziert werden.

Nehmen wir an, dass die Reduktion der Treffen mit extensiver sozialer Interaktion bis zu 80% betragen könnte, ein beträchtliches Potential also gegeben wäre. Es verbleibt also, die notwendigen technischen Randbedingungen für akzeptable und produktive Kommunikation bei mehr formalen Treffen zu definieren. Diese Bedingungen konnten innerhalb der Institute der Autoren zum Beispiel bei Sitzungen der wissenschaftlichen Leitungsgremien – verteilt auf Garching und Greifswald im Falle des IPP und Bremerhaven, Potsdam, Helgoland und Sylt im Falle des AWI – exploriert werden. Dabei wurde seitens der Teilnehmer immer wieder die positive Erfahrung einer grundsätzlich neuen Kommunikationsform betont („Irgendwo zwischen Telefon und persönlichem Treffen halt“). Die Gesamteffektivität der Entscheidungsvorgänge sei erheblich gewachsen.

Als Beleg für eine bereits eingespielte, stabile und in den beschriebenen Szenarien erfolgreich genutzte Infrastruktur mag das IPP mit seinem Hauptstandort Garching und dem Teilinstitut Greifswald dienen [5]: Die Videogruppe des IPP betreibt derzeit in Garching einen Großen Hörsaal, einen großen Seminarraum beim Experiment ASDEX Upgrade, zwei Besprechungsräume für das Direktorium, die wissenschaftliche Leitung und die Verwaltung, drei mittelgroße Seminarräume, und etwa 15 Desktopsysteme (Polycom ViaVideo). In Greifswald werden ein großer Seminarraum, ein Besprechungsraum für die Leitung und Verwaltung, drei mittelgroße Seminarräume und etwa 10 ViaVideos betrieben. Die Einrichtung erfolgte durch Ausschreibungen und Outsourcing (Hörsaal, große Seminarräume, teilweise auch Direktoriumsräume), und eigene Installationen (aus Finanzierungsgründen teilweise provisorisch) für die kleineren Räume und die Desktop-Systeme. Alle Settop- und Raumsysteme sind über ISDN und IP erreichbar, dabei werden die eingebauten MCUs häufig für Zuschaltungen von ISDN Video- aber auch von reinen Telefonteilnehmern verwendet. Alle H.323 Systeme, derzeit etwa 40, sind am RZG Gatekeeper registriert, der alle H.323 Datenströme durch einen Proxy leitet und damit die RZG Firewall ergänzt. Der OpenSource GNU-Gatekeeper und -Proxy läuft seit Ende 2002 äußerst stabil.

---

<sup>1</sup> In anderen Szenarien, z.B. qualitativ hochwertigen eins-zu-eins Konferenzen zwischen den Büroarbeitsplätzen der Autoren, kann durchaus eine Situation ähnlich der einem Büro mit zwei Arbeitsplätzen entstehen.

### 3 Detaillierte Anforderungsanalyse

Im Einzelnen haben wir folgende Faktoren – etwa in Reihenfolge ihrer Relevanz – als notwendig für den erfolgreichen Einsatz ermittelt:

- Die Zuverlässigkeit des Systems muss so hoch sein, dass ein Komplettausfall, der die Teilnahme verhindert, extrem selten ist. Gerade in der Einführungsphase können Ausfälle die Akzeptanz auf lange Zeit vernichten. (Der Referenzwert hierzu mag etwa der Ausfall des öffentlichen Verkehrs oder entsprechend unerwartete Autobahnstaus sein.)
- Die Verständlichkeit des übermittelten Tons muss – bei einem Minimum an Kooperation der Beteiligten, etwa in Form deutlicher Aussprache – so gut sein, dass nicht allein schon der Aufwand des akustischen Verstehens zur Ermüdung führt und die eigene Beteiligungsfähigkeit mindert. Bereits die Notwendigkeit, einen korrekten Abstand zu einem Konferenzmikrofon einzuhalten, kann die Teilnehmer auf nicht akzeptierte Art ablenken (oder gar zu taktischen Spielen verführen). Es ergibt sich nebenbei, dass die Tontechnik dergestalt optimiert sein muss, dass ein entfernter Sprecher in eine Diskussion eingreifen können muss, ohne dass er dazu mehr unternehmen muss als ein lokal Anwesender.
- Der (potentielle) Wert des Bildes wird ad absurdum geführt, wenn die Ausleuchtung oder der Kamerablickwinkel kein deutliches Bild der Sprechenden ergibt oder die Wiedergabe auf einem zu kleinen oder überstrahlten Monitor („in der Ecke“) dargestellt wird.
- Wenn der Eindruck von Präsentationen an einem der Standorte eine Rolle spielt, müssen diese an allen Standorten in gleicher Qualität wiedergegeben werden. Dies betrifft sowohl die Auflösung und den Kontrast – was im Falle der Wiedergabe mit Fernsehauflösung nicht gegeben ist – als auch die Synchronität mit dem Vortragenden – die im Falle der Übermittlung des vollen Bildes mit niedriger Bandbreite leidet.
- Schließlich muss der Aufwand, die Technik in dieser Qualität in Bereitschaft zu setzen und die Verbindung aufzubauen, so gering sein, dass die ersparte Zeit nicht an anderer Stelle wieder verbraucht wird. Insbesondere sollte es möglich sein, dass einer der Beteiligten, eine Sekretärin oder eine andere, nicht medientechnisch vorbelastete Person die Konferenz in wenigen Minuten in Gang setzt und weiter steuert, ohne dadurch übermäßig abgelenkt zu werden.
- Alle diese Kriterien führen naturgemäß zu hohen Investitions-, Installations- und Wartungskosten. Diese dürfen aber den ökonomischen Wert des Systems nicht in Frage stellen.

Zuletzt in dieser Analyse sei nochmals ein Wort zur Bedeutung der Fokussierung auf Planungsvorgänge gesagt: Dergestalt gut funktionierende Kommunikationssysteme haben das Potential, Interaktionen und Zusammenhalt in einer geografisch verteilten Organisation über die Wissenschaftsplanung hinaus deutlich zu verbessern. Dieses mehr qualitative – jedenfalls in Geld schwer auszudrückende - Argument kann aber nur in Ausnahmefällen zur Begründung einer großen Investition genutzt werden. Das Faktum aber, wenn es sichtbar wird, kann sehr wohl helfen, den weiter laufenden Aufwand, weitere Verbesserungen der Qualität, oder weitere entsprechende Infrastrukturinstallationen zu rechtfertigen.

#### **4 Realisierung durch Medientechnik**

Die möglichen und nützlichen Anwendungsszenarien der Videokonferenztechnik – gerade auf dem beschriebenen Qualitätsniveau - sind vielfältig. Im Hinblick auf die Realisierung kann man sie am ehesten anhand der Anzahl der Teilnehmer je Standort und damit der Größe der genutzten Räume kategorisieren. Im Falle der beschriebenen Planungsszenarien könnten wir uns also auf Büros oder kleine Besprechungsräume beschränken. Dies wird bei der Realisierung auf Direktorebene in der Max-Planck-Gesellschaft auch so gehandhabt.

In der etwas ausgeweiteten Ebene der leitenden Wissenschaftler jedoch überwiegt die Nutzung kleiner bis mittlerer Seminar- oder Besprechungsräume. Auf diese werden wir uns im Folgenden konzentrieren – nicht ohne zu bemerken, dass die entsprechenden Realisierungshinweise auch für die zwei anderen Szenarien gelten – Videokonferenzen vom Arbeitsplatz aus (Desktop), oder Übertragungen zwischen Hörsälen oder großen Seminarräumen sowie für Konferenzen unter Einbeziehung aller 3 Szenarien.

Um optimale Konferenzbedingungen zu erreichen, müssen einige, wenn nicht alle der in Abb. 3 im Überblick gezeigten Maßnahmen in jedem Falle getroffen werden. Im folgenden gehen wir auf die wichtigsten näher ein.



Abbildung 3: Beispiel eines optimal ausgestatteten Videokonferenzraums im Europäischen Patentamt (© EPO, DCTV, VIPLA) für bis zu 10 Teilnehmer.

- (1) Blendfreie Beleuchtung von Teilnehmern und Tisch
- (2) Regelbare Hintergrundbeleuchtung, neutrale Hintergrundfarbe
- (3) Mehrere Kameras (hier 2 zentral zwischen den Displays, zwei an den Wänden, eine an Rückwand)
- (4) Zentrales (Kardioid-Ebenen-)Mikrofon
- (5) Lautsprecher
- (6a) Display für Video, (6b) Display mit Smartboard für Präsentationen
- (7) Berührungsempfindlicher Monitor für Mediensteuerung
- (8) Monitor für Präsentations-PC im Raum
- (9) Klimaanlage
- (10) Verdeckte Anschlussfelder für Laptops oder Presenter (s.u.)

#### 4.1 Zuverlässigkeit und Reproduzierbarkeit

Insgesamt beschreiben wir technische Anforderungen, die manchen als hoch erscheinen werden. Selbst wenn der Leser geneigt wäre, bei der praktischen Umsetzung hier oder dort Kompromisse einzugehen: Auch in diesen Fällen empfehlen wir zur Realisierung des ersten und wichtigsten Kriteriums, der Zuverlässigkeit und Reproduzierbarkeit der Qualität, dringend eine feste Installation der Komponenten in einem bestimmten Raum.

Unsere Erfahrung zeigt, dass das Herumfahren der Komponenten auf einem Wagen, so ökonomisch attraktiv es erscheinen mag, inakzeptable Risiken in sich trägt:

- Der „harmloseste“ Fall ist der Anschluss des Ethernet an „irgendeinen“ Switchport. Er birgt eine hohe Wahrscheinlichkeit vermeidbarer Paketverluste – mit Folgen bis hin zu Aussetzern im Ton, s.u.
- Die falsche Platzierung des Mikrofons hat eine schlechte Tonqualität - bis hin zu starken Rückkopplungen und dementsprechendem Pfeifen – zur Folge, die eine effektive Verständigung unterbinden kann.
- Der „vergessene“ oder falsch gesteckte Ethernetanschluss (oder das fehlerhaft gesteckte Audiokabel), erst beim Versuch des Verbindungsaufbaus bemerkt, führt zu Zeitverlusten beim Beginn der Sitzung, bis die Korrektur erfolgt ist. Eine Reihe wichtiger Leute müssen warten und mindestens ein Wissenschaftler, der falsch gestöpselt hat, muss sich von einem Techniker helfen lassen.

Solche Peinlichkeiten werden der Technik oder der betreuenden Infrastruktureinrichtung angelastet werden und diesem Projekt, geplanten Verbesserungen oder nachfolgenden Projekten möglicherweise ein Ende setzen - zumindest, wenn sie sich wiederholen.

#### **4.2 Tonqualität**

Die Anforderungen an die Tonqualität gehen bei den hier betrachteten Szenarien über die von Verbindungen zwischen zwei Personen deutlich hinaus:

Die Trennung des Nutzsignals Sprache vom Hintergrund – zum Beispiel Rascheln, Tassenklappern, Murmeln, vor allem aber auch das Echo im Raum – wird nicht, wie beim Gespräch an einem Tisch, durch die Möglichkeit der räumlichen Ortung erleichtert. Gleich laut Sprechende werden – je nach Mikrofon – u.U. drastisch verschieden wiedergegeben, wenn sie verschieden weit vom Mikrofon entfernt sind.

Diesen Effekten kann und muss man, um die oben angedeuteten Ermüdungseffekte zu vermeiden, auf mehrere Arten entgegen treten, insbesondere

- am leichtesten durch eine größere analoge Bandbreite, in Tonfrequenzen ausgedrückt. Hier setzen wir die Ton-Kodierung mit G.722 [6] ein, die eine Bandbreite von 7 kHz erbringt, das Doppelte der ISDN-Bandbreite (mit dem Codec G.711)
- durch den Einsatz eines oder mehrerer Mikrofone sehr hoher Qualität und geeigneter Richtcharakteristik, z.B. die Grenzflächenmikrofone von Tandberg oder das Kardioid-Ebenen-Mikrofon von Microtech Gefell.

- wenn dies notwendig und finanziell und architektonisch machbar ist, durch Dämpfungsmaßnahmen im Raum, z.B. durch selektives Anbringen von geeigneten Platten oder Paneelen („Wilhelmi-Platten“, stoffbespannte Akustikpaneele), aber auch durch die sorgfältige Wahl besonders geräuscharmer Beamer oder der Montage der Beamer in schalldämmenden Gehäusen. Teppiche erbringen zwar keine wirkliche Schalldämpfung, reduzieren aber Geräusche von Schritten und Stuhlrücken und sind deshalb Parkett- oder Linoleumböden in jedem Fall vorzuziehen.

- Es ist von großem Vorteil, die Mikrofone körperschallisoliert, oder von den Besprechungstischen getrennt zu befestigen bzw. aufzustellen. Sie sollten unbedingt so angebracht bzw. schlecht zugänglich sein, dass zumindest einem unabsichtlichen Verändern der Mikrofonplatzierung und -ausrichtung vorgebeugt ist. Die Kabel sollten fest in einem Boden-/Kabelkanal verlegt sein.

- Es versteht sich von selbst, dass diese Maßnahmen nur dann sinnvoll sind, wenn auf beiden Seiten ein entsprechend gutes Audiosystem mit hochwertigen Lautsprechern zur Verfügung steht – was i.a. aber deutlich einfacher und billiger zu erreichen ist, als auf der Mikrofonseite.

- Die Qualität der elektronischen Echosperrung (Echo-Cancellation)<sup>2</sup> ist bei neuen Anlagen hoher Qualität völlig ausreichend. Auf den Einsatz externer Echosperrungen („Gentner“) kann in der Regel verzichtet werden. Bei älteren oder „billigeren“ Videokonferenzsystemen ist aber auch dieser Faktor genau zu untersuchen. (Wir setzen zurzeit noch durchgehend Tandberg-Systeme neuerer Bauart ein. Einige Systeme (z.B. ältere VCON oder Sony) müssen nach einer Änderung akustischer Verhältnisse – zum Beispiel Bewegung von Personen, die sich nahe beim Mikrofon befinden – einige Zehntelsekunden „lernen“. Während dieser Zeit können bereits irritierende Echos auftreten, die in Punkt-zu-Punkt-Verbindungen noch tolerabel, bei Mehrpunktkonferenzen jedoch äußerst störend sind.

### 4.3 Bildqualität

Die Qualitätswahrnehmung des nicht vorbelasteten Teilnehmers ist unserer Erfahrung nach durch das Fernsehen mit der dort vorgegebenen Auflösung (704 x 576 Pixel) und vor allem der gerade bei Talkshows perfekten Ausleuchtung und Bildregie geprägt. Wenn ein Teilnehmer deshalb dem neuen Medium Videokonferenz nicht einfach ablehnend gegenüber steht, wird er oder sie jedenfalls deutlich schlechtere Ergebnisse bemerken und möglicherweise zunächst nicht tolerieren wollen.

Wir streben daher folgende Charakteristiken und Einstellungen an:

---

<sup>2</sup> Echocancellation dient dem Entfernen des durch die entfernte Lautsprecher/Mikrofonkombination zurück übertragenen eigenen Signals. Diese Aufgabe wird durch die Echozeitverzögerung erschwert, die durch die Komprimierung (En- und De-Coding) von Ton und Bild bei Videokonferenzen zusätzlich zu weiteren Delays durch die IP-Übertragung, hervorgerufen wird. Einige Mehrpunktkonferenzanlagen unterstützen Transcoding, d.h. individuelles Um-Kodieren des Signals für jeden Teilnehmer. Ältere Anlagen fügen dabei bis zu einer Sekunde weitere Verzögerung hinzu.

- Eine Auflösung wie beim Fernsehen, genannt 4CIF, entsprechend dem PAL Standard, bei mindestens 10 bis 15 Bildern pro Sekunde, die wir mit dem Codec H.263+ bei 384 bis 768 kBit/s Gesamtbandbreite erreichen (wenn der Bildhintergrund nicht zu unruhig ist! Bücherregale, feingerasterte Wandverkleidungen oder schwankende Spitzlichter, etwa durch die sonnenbeschienenen Führungslöcher in Jalousien können diese Unruhe bewirken.). Tabelle 1 führt die gebräuchlichen Bildauflösungen auf.

Bildformat <sup>3</sup>	Auflösung	Videocodec
SQCIF	128 x 96	H.263 [7]
QCIF	176 x 144	H.261 [8], H.263
CIF	352 x 288	H.263, H.264 <sup>4</sup>
4CIF	704 x 576	H.263(+), H.264
XGA	1024 x 768	H.239 <sup>5</sup> [9]
16CIF	1408 x 1152	H.263(+) <sup>6</sup>

Tabelle 1: Übersicht der in Videokonferenzen üblichen Bildformate inkl. Auflösung und verwendeter Videocodecs

- Die Gesichter der Sprechenden müssen gut ausgeleuchtet sein – insbesondere nicht von oben oder nur von einer Seite. Darüber hinaus sollte der Hintergrund farblich so gestaltet und/oder separat ausgeleuchtet sein, dass die automatische Belichtungssteuerung die Gesichter auch bei Totalen richtig belichtet. (Die Effekte unterschiedlicher Ausleuchtung kann man bei der Betrachtung „mit eigenen Augen“ kaum abschätzen, da der visuelle Apparat diese meist unmerklich korrigiert. In fixierten Büroumgebungen für eine Person kann man die gewünschte Ausleuchtung noch mit einer geschickten Ausrichtung zwischen Person und Kamera sowie einer entsprechenden Wahl von Vorhangstellung und angeschalteten Beleuchtungsmitteln erreichen. In Besprechungsräumen, zum Beispiel mit einer U-förmigen Anordnung der Sitzplätze, ist dies nur selten ohne spezielle Hintergrund- und Beleuchtungsmaßnahmen möglich (Hinweise in Abb. 3).

---

<sup>3</sup> CIF ist eine Abkürzung für Common Interchange Format

<sup>4</sup> Der neue Codec-Standard H.264 erlaubt dem H.263 gleichwertige Bildqualität mit etwa der halben Bandbreite. Leider wird dieser Standard derzeit noch nicht von allen Herstellern so implementiert, dass er zwischen verschiedenen Geräten interoperabel ist, auch die relevanten MCUs unterstützen H.264 noch nicht durchgängig.

<sup>5</sup> H.239 erlaubt Präsentationen bis XGA und ca. 5 fps. Zur Interoperabilität gilt leider auch hier das zu H.264 gesagte.

<sup>6</sup> 16CIF hat sich nicht durchgesetzt

- Die gute Ausleuchtung der Teilnehmer muss aber gewährleisten, dass die Displays bzw. Projektionen zur Darstellung der Präsentation und der entfernten Teilnehmer nicht direkt beleuchtet werden. Dies bedarf also auch eines besonderen Augenmerks bei der Ausrichtung der Beleuchtung und der Montage der Leinwände oder Monitore. Hinreichend starke Beamer, bzw. Helligkeit der Displays sind vorzusehen.

- Bei bis zu drei Teilnehmern kann eine Kamera mit einer Einstellung ausreichen, um die Teilnehmer hinreichend darzustellen. Leitende Wissenschaftler, auch solche, die in der Lage und willens sind, die Konferenz selbst aufzubauen, werden aber nicht immer daran denken, die Kameraeinstellung zu optimieren, da verständlicherweise das Hauptinteresse sofort dem Bild der Gegenseite gilt. Feste Sitz- und Kamerapositionen und entsprechende Standardeinstellungen für alle Standorte können das Problem in diesen Fällen entschärfen. Sollen mehrere Sprecher (etwa an einer U-förmigen Tischanordnung) angemessen wiedergegeben werden, sind zusätzliche Kameras notwendig und eine angemessen einfach gestaltete Bedienoberfläche der dann notwendigen Mediensteuerung (s. Kap. 4.5) muss die Bildregie über eine begrenzte Zahl von Presets – die Kamerawahl, -winkel und Zoom umfassen – ermöglichen.

#### 4.4 Präsentationen

Angesichts der Bandbreitenproblematik (s.u.) und der Forderung, dass alle Beteiligten in den Genuss gleicher Auflösung kommen sollen, haben wir uns als gemeinsamen Nenner, der noch ausreichende Qualität erbringt, schon vor einiger Zeit auf eine Auflösung von 1024 x 768 (XGA<sup>7</sup>) als Standard festgelegt. Da derzeit noch kein Hersteller den H.239 ITU Standard [9] in der Art und Weise umgesetzt hat, so dass „jeder mit jedem“ kann (siehe Tabelle 2 im Kapitel 5.5), wird sowohl in der HGF als auch in der MPG derzeit VNC [10] zur Übertragung von Präsentationen verwendet. Bei einzelnen Veranstaltungen innerhalb der MPG (IPP Garching – IPP Greifswald) wird auch noch NetMeeting verwendet<sup>8</sup>. Weitere Erläuterungen folgen in Abschnitt 5.5.

Es ist hier zu bemerken, dass trotz aller Bemühung, hochwertige Verfahren einzusetzen, Animationen – z.B. effektvolle Übergänge zwischen Folien oder gar schrittweise eingeblendete Elemente auf einer Folie – weitgehend zu vermeiden sind oder aber einen deutlich verlangsamten Ablauf der Präsentation bedingen. Computeranimationen in voller Auflösung und eingebettete Filme können derzeit nicht ohne Einbussen übertragen werden und sollten deshalb vorläufig nicht für Präsentationen verwendet werden. Alternativ kann ein separater Movieclip (mit Fernsehcharakteristik) unter Umständen auf einem zweiten Kanal oder anstelle des Bildes des Sprechers übermittelt werden.

---

<sup>7</sup> eXtended Graphic Array

<sup>8</sup> Da NetMeeting über kurz oder lang dem Microsoft Messenger weichen wird, wird der Einsatz von NetMeeting bereits stark reduziert.

Für Vortragende, die noch Folien mitbringen oder zur wünschenswerten Aufnahme von physischen Objekten, deren lokaler Präsentation und Übertragung, bietet sich das Verwenden eines sog. Presenters an. Dieser sollte ebenfalls eine Auflösung von 1024 x 768 an einem PC-kompatiblen Monitoranschluss unterstützen und ein gut korrigiertes Zoomobjektiv ohne Verzeichnung und Bildfehler haben. Dieses Signal kann dann über einen Multiplexer oder eine Kreuzschiene anstelle des Bildes einer PC-Präsentation gezeigt werden.

#### **4.5 Mediensteuerung**

Zu einer optimal reproduzierbaren Durchführung von Konferenzen ist neben der Beleuchtung und Bildregie, der festen Anordnung aller Komponenten – Leuchten, Kameras, Leinwände, Tische, Personen, ... – auch eine Mediensteuerung unabdingbar, die alle beweglichen (Vorhänge) und veränderbaren oder schaltbaren Komponenten (Lichtstärke, Kameras) zu Beginn der Konferenz in Standardeinstellungen bringt und danach nur sinnvolle und koordinierte Änderungen zulässt. Über die Mediensteuerung können fest programmierte Szenarios abgerufen werden, die neben den lokalen Einstellungen auch die Anwahl der gewünschten Teilnehmer und sogar das Starten und Steuern der entsprechenden Technik an den angerufenen Standorten erlauben. Dies erlaubt eine ganz erhebliche Reduzierung des Hilfspersonals, setzt aber natürlich beträchtliche, aber einmalige, Investitionen voraus.

#### **4.6 Personal- und Investitionsaufwand**

In unseren Instituten ist es nahezu unmöglich, zusätzliche Planstellen für die Betreuung der Videokonferenzräume zu schaffen. Vielleicht im Gegensatz zu Universitäten besteht meist auch nicht die Möglichkeit, die Betreuung individueller Konferenzen an studentische Hilfskräfte oder ähnliche Personen zu delegieren. Unseres Erachtens beruht diese Vorstellung ohnehin auf einer Illusion - sie reduziert die Verfügbarkeit des Systems und verlagert den Personalaufwand auf das Anlernen der häufig wechselnden Hilfskräfte.

Kommerzielle Settop- oder Raumsysteme (H.320 und/oder H.323) zusammen mit einer minimalen Displayausstattung, etwa einem Fernseher, können im Allgemeinen von „jedermann“ eingeschaltet werden und der Aufbau einer Verbindung ist nicht schwieriger als die Programmwahl im Fernsehen verbunden mit dem Wählen einer Telefonnummer. Komplexere Einstellungen, wie die unter Bildqualität geschilderten, können nur mit Hilfe einer kommerziellen Mediensteuerung realisiert werden.

Wir müssen hier aber festhalten, dass ein komplettes Outsourcen auch nicht möglich ist. Eine gut geplante und realisierte Mediensteuerung erlaubt in den meisten Szenarien die Bedienung durch die Teilnehmer selbst. Selbstverständlich muss dennoch ein Assistent / eine Assistentin zur Verfügung stehen, etwa wenn neue Teilnehmer die Räume erstmals benutzen wollen, und/oder besondere Funktionen abgerufen werden sollen, die spezielle Kenntnisse erfordern. Alle komplexeren Arbeiten an der Technik sollten im Rahmen von Wartungsverträgen mit den ausführenden Firmen geregelt sein. Hier ist insbesondere auf eine schnelle Einsatzbereitschaft zu achten.

Gegen die MBone-basierte Videokonferenztechnik haben wir uns daher nicht nur aus Qualitätsgründen entschieden – sie erbringt die von uns bisher geschilderten Qualitäten nicht oder nur mit unverhältnismäßig größerem technischen Aufwand. Vielmehr hat uns schon zu Beginn der Aktivitäten speziell für das AWI und das IPP beeinflusst, dass die über viele Jahre regelmäßig betriebene und vom Szenario durchaus vergleichbare Telekonferenz Bayrischer Rechenzentrumsleiter (s. DFN-Projekt TKBRZL[11]) an jedem Standort bei jeder Konferenz eine Vorbereitungszeit von einer halben Stunde einer qualifizierten IT-Person erforderte. Vergleichbare Erfahrungen gibt es am IPP aus den Jahren 1998-1999 und im Rahmen der Europäischen Fusionsforschung mit dem aus der Hochenergiecommunity stammenden VRVS System [12], das auf den MBone Tools basiert. Auch hier wird derzeit noch in erheblichem Umfang hochqualifiziertes Personal gebunden, das anderweitig sinnvoller einsetzbar wäre.

Die angesprochene Arbeit eines Assistenten / einer Assistentin umfasst auch die allgemeine Betreuung und Systempflege unabhängig von den Zeiten, zu denen Konferenzen stattfinden. Erfahrungsgemäß beträgt der Bedarf für ein Settop-System in einem kleinen Besprechungsraum ohne Präsentationsmöglichkeit und besondere Beleuchtungs- und andere Vorrichtungen ca. 2-4 Personenwochen/Jahr und der für ein Raumsystem in einem mittleren Besprechungsraum mit allen geschilderten Einrichtungen 4-8 Personenwochen/Jahr. Voraussetzung hierfür ist jedoch eine technisch weitgehend homogene Herkunft und Ausstattung der von einer Person zu betreuenden Systeme und Räume. Weiter reduziert ein entsprechender Wartungsvertrag die notwendige Qualifikation des Assistenzpersonals auf technischer Ebene erheblich. Die Manpowerangaben beziehen sich auf die derzeit weitgehend ausgelasteten Seminarräume des IPP in Garching und Greifswald (ca. 5-15 VC / Woche / Raum à 1-4 Stunden).

Die Investitionskosten betragen für die erwähnte Minimallösung etwa 10.000 €, für die zweite Lösung, mit allen Einbauten (Raumdämpfung, ... bis hin zu fast geräuschlosen Beamern) bis zu 100.000 €. Bei der Beurteilung dieser Aufwendungen ist zu beachten, dass ein für medial aufwendige, lokale Präsentationen geeigneter Raum für 20 bis 25 Personen bei jeder Gruppenbesprechung mit jeweils mehreren Teilnehmern zwischen zwei Standorten Personal- und Reisekosten vermeidet, die ohne weiteres fünfstellig werden können.

## 5 Realisierung durch Netzdienste

Bis zu diesem Punkt haben wir versucht, die Fragen zum Transport durch Netzwerke weitgehend zu vermeiden. Andeutungsweise ist jedoch bereits klar, dass die Bandbreitenanforderung eher über normalerweise genutzte und verfügbare ISDN-Bandbreiten hinausgeht:

Nach unserer Erfahrung reicht die Video- und Tonqualität von Punkt-zu-Punkt-Verbindungen im ISDN-Netz bei 384 kBit/s, also 6 B-Kanälen bzw. 3 S0-Anschlüssen durchaus für die von uns beschriebenen Ansprüche. Zwei Gründe sprechen aber gegen eine ausschließliche Verwendung von ISDN:

- Wenn über den gleichen Kommunikationskanal auch Präsentationen übermittelt werden, werden für diese typisch bis 128 kBit/s verwendet. Dies ist für viele Präsentationen zu wenig, kann die Bildqualität des eigentlichen Videos unter Umständen aber deutlich mindern.
- Die von Dienstleistern betriebenen Systeme, die Mehrpunktkonferenzen ermöglichen, sind bei diesem Qualitätsniveau nur zu drastischen Verbindungspreisen zu „mieten“.

Darüber hinaus legt die zunehmende „Konvergenz“ – z.B. der Trend zu VoIP – nahe, eine Videokonferenzinfrastruktur von Anfang an auf IP-Transport auszulegen.

Die mit diesen Faktoren (noch) verbundenen Probleme und deren Lösungen bzw. Ansätze dazu werden wir im Folgenden diskutieren.

### 5.1 Transportoptionen

Wie summarisch festgestellt, wollen wir grundsätzlich auf den Transport über ISDN verzichten. Jedoch sehen wir zwei Gründe, ISDN als Fallback-Lösung an jedem Standort und beim Dienstleister für Mehrpunkt-Konferenzen mit vorzusehen: Einerseits pflegen fast alle unserer Institute Kontakte zur Wirtschaft, bei der - falls überhaupt vorhanden - Videokonferenzsysteme auf ISDN und H.320 beruhen. Andererseits wollen wir für Fälle vorbeugen, wenn entweder der lokale Internetanschluss gestört oder ein entfernter Teilnehmer nur über unzulängliche Internetverbindungen erreichbar ist.

An dieser Stelle soll – zur Ehrenrettung des ISDN – noch bemerkt werden, dass ISDN nicht nur einen synchronen und damit sehr gleichmäßigen Transport bietet, sondern auch einen geringeren Overhead hat. Die Qualität, die man bei 384 kBit/s IP-Bandbreite wahrnimmt, ist – subjektiv beurteilt – im Fall von ISDN wohl schon bei 256 kBit/s erreicht.

Das Internet Protokoll (IP) bietet darüber hinaus nur einen „unzuverlässigen“ Transport, d.h.: Pakete können verloren gehen. Die Herstellung einer zuverlässigen Verbindung mittels TCP kommt hier nicht in Frage, da das Nachfordern verlorener Pakete in der Realtime-Übertragung zu größeren Aussetzern bzw. Verzögerungen beim Vorhalten entsprechender Puffer<sup>9</sup> führen würde, als tragbar ist.

Das Paketverlustniveau, das im Allgemeinen noch ohne deutlichen Qualitätsverlust tragbar ist, liegt bei etwa 0,5%. Zu beachten ist allerdings, dass bei den entsprechenden Anzeigen der Videokonferenzsysteme offenbar auch übermäßig verzögerte Pakete als verloren gezählt werden. Jedenfalls lassen sich diese Anzeigen nur selten in klare Übereinstimmung mit Verlustraten bei ping-Serien bringen. Dies kann zu einem wesentlichen Problem bei der Kommunikation des Systembetreuers mit dem Netzwerkbetreuer werden: Letzterer sieht bei allen Standardtests – ping, traceroute, ftp-Durchsatz,... – keine Probleme, also liegt „es“ nicht an „seinem“ Netzwerk!

Wir haben uns auch an Versuchen mit aufwendigeren Messverfahren [13,14] beteiligt. Auch hier kamen wir nicht zu Ergebnissen, die klare Hinweise zu Verbesserungen im Netzbereich oder auch nur zu Ursachen für von den Systemen angezeigte Paketverluste gegeben hätten.

Unter diesen Umständen muss man also versuchen, die üblichen Probleme nach Möglichkeit auszuschalten. Es gibt jedoch keinen Grund, mutlos an die Aufgabe heranzutreten: Schon seit 1999 veranstaltet eine Gruppe an der Ohio State University jährliche „Megakonferenzen“ [15] mit bis zu 150 Teilnehmern aus aller Welt [16]. Das dabei erreichte Qualitätsniveau ist, angesichts der internationalen Verbindungen und der Größe des Mehrpunktnetzes, beachtlich. Schon bei diesen Konferenzen zeichnete sich ab, dass meist die lokalen Einrichtungen und lokale „Kleinigkeiten“ Ursache mangelnder Qualität sind.

In den Kernnetzen des G-WiN, GEANT oder von Abilene und sogar bei Verbindungen zwischen diesen kann man generell von geringen „echten“ Paketverlusten und geringen Latenzschwankungen ausgehen. Es verbleiben die lokalen Netze und der Zugang zum Wissenschaftsnetz:

Unsere Faustregel zum lokalen Netz besagt, dass die Systeme nur an einem bandbreitenmäßig sehr überdimensionierten, geschwichten lokalen Netz angeschlossen werden sollten (Overprovisioning). Insbesondere sollte der eigentliche Anschluss<sup>10</sup> und die gesamte Strecke full duplex betrieben werden.

---

<sup>9</sup> Selbst wenn man die lokale Bearbeitungszeit mit 0 annimmt, addieren sich die Hin- und Rücklaufzeiten in relevanten Verbindungen leicht auf wahrnehmbare Verzögerungen, die man schon aus Gründen der ungestörten Kommunikation vermeiden möchte. Darüber hinaus erschweren lange Laufzeiten aber noch die Aufgabe der Echosperren.

<sup>10</sup> Ein besonders Problem bereiten bestimmte Switches, die bei einer Autosensingeinstellung regelmäßig neu verhandeln und dabei für Augenblicke den Transport aufhalten. Dies geschieht bei Datenübertragungen unmerklich; bei Videokonferenzen führt es zu Aussetzern. Daher sind die Einstellungen zu Bandbreite und Duplex fest vorzunehmen

Die „Last Mile“ vom Kernnetz zum eigenen Router der Einrichtung wird häufig aus finanziellen Gründen mit hoher prozentualer Auslastung betrieben. Dies führt im Allgemeinen<sup>11</sup> zwar nicht zu Paketverlusten, jedoch aufgrund der Warteschlange, die sich am Interface bildet, zu Laufzeitschwankungen, die zu einer Degradation der Videoverbindung führen kann.

An dieser Stelle sei bemerkt, dass die Strategien und Techniken der Endgerätehersteller, um mit solchen Verlusten umzugehen, sehr verschieden sind und sich gerade in den letzten Jahren deutlich verbessert haben: Offensichtlich haben die arrivierten Hersteller mittlerweile durchgängig Strategien eingebaut, bei dauerhaft großen Verlusten die genutzte Bandbreite zu verringern. Die eingesetzten Techniken sind aber zwischen den Herstellern offenbar nicht ganz kompatibel und sind offenbar beim Einsatz in Mehrpunktkonferenzen deutlich unwirksamer als in Zweipunktkonferenzen mit Geräten vom gleichen Hersteller.

Verlorene Pakete können bis zu einem gewissen – herstellerabhängigen – Niveau „überspielt“ werden. Längere Tonaussetzer oder gar Knacklaute sind nur in geringstem Masse zu tolerieren. Bildfehler sind eher zu tolerieren, sie können jedoch sehr ablenkend sein (wenn etwa ein Hersteller früher grüne Balken zeigte oder bei vielen Implementierungen Artefakte wie z.B. Bildüberlagerungen auftauchen.)

Die Lösung durch eine zugesicherte, globale Ende-zu-Ende-Qualität (etwa unter dem Stichwort Resource Reservation Protocol) ist, wenn überhaupt je möglich, noch sehr weit hinter dem Horizont. (In von einem Betreiber Ende-zu-Ende beherrschtem Netz kann dies offenbar mit proprietären Implementierungen erreicht werden, jedoch nicht „im Internet“).

Kurzfristig hilfreich können dagegen auch punktuelle Maßnahmen, wie die gezielte Priorisierung („Low Latency Queue“)<sup>12</sup> von Videokonferenzpaketen auf stark belasteten Strecken oder solchen mit geringer Bandbreite sein.

Eine derzeit am IPP in Kooperation mit dem AWI durchgeführte Simulationsstudie soll näheren Aufschluss über das Verhältnis „verfügbare (schwankende) Bandbreite – prozentueller Paketverlust“ geben.

---

<sup>11</sup> Hierbei sei jedoch anzumerken, dass in Zeiten hohen Datenaufkommens die Überlastkontrolle (congestion management) des lokalen Routers „zuschlagen“ kann. Dabei werden je nach implementiertem Algorithmus, z.B. RED (Random Early Detection), Pakete "frühzeitig" verworfen.

<sup>12</sup> Dieses - wohl proprietäre - Verfahren führt alle Pakete einer bestimmten Klasse - typischerweise VoIP und Videokonferenzpakete - gemeinsam über eine Queue, die sie bevorzugt gegenüber allen anderen Queues mit "gerechter" Aufteilung der Bandbreite auf die Nutzer der Leitung.

## 5.2 Adressierung

Die Adressierung von H.323-Systemen erfolgt meist auf zwei Arten: Durch IP-Adressen oder so genannte E.164 Nummern [17], die wiederum das Aussehen von Telefonnummern haben. Letztere eröffnet tatsächlich die Möglichkeit der Verbindung zu Voice-over-IP (VoIP) Telefonie, aber auch das Hinzunehmen eines Teilnehmers zu einer Konferenz nur per Sprache. Dieses, aber auch eine ganze Reihe anderer Argumente und vor allem Erfahrungen führen uns dazu, nur mittels E.164 zu arbeiten.

Die Übersetzung von E.164 Nummern zu IP-Adressen wird von einem (H.323-) Gatekeeper wahrgenommen, der z.B. auch die Funktion der Zugangskontrolle übernehmen kann. (s. auch Proxy-Funktion, im nächsten Kapitel)

Durch die internationale Koordinierung der Nummernvergabe ist ein Global Dialing Schema, GDS [18], entstanden. In diesem vermittelt ein nationaler Gatekeeper, der hierzulande vom DFN betrieben wird, zwischen den Gatekeeper einzelner Institutionen und anderen nationalen Gatekeepern. (Es muss hier angemerkt werden, dass nicht alle Gatekeeper-Implementierungen problemlos in dieser Weise zusammen arbeiten. Wir empfehlen daher jedermann strikt den GNU-Gatekeeper und -Proxy, s.u.).

Die nationalen Gatekeeper „verwalten“ somit auch die nationalen Vorwahlen (im Fall des DFN die 49) und erleichtern damit auch die internationale Adressierung und Anwahl beträchtlich. Voraussetzung für diese ist naturgemäß, dass die Gatekeeper von Institutionen mit ähnlicher Sorgfalt verwaltet werden wie deren DNS-Nameserver. (Beispielsweise verwaltet der Gatekeeper des IPP den Nummernbereich (hinter der Vorwahl) (+49) 89 3299 6001...6999 und der des AWI einen Bereich von (+49) 471 4831 2801...2899. Es empfiehlt sich dabei, die verwendeten Durchwahlnummern mit denen des Telefonsystems abzustimmen und diese dort nicht zu verwenden.) Mit der Ausnahme des Videokonferenzservice des ESNNet, der bis vor kurzem ausschließlich auf ISDN basierte, stützen sich alle Services im Umfeld der Nationalen Wissenschaftsnetze auf diese Technik und das GDS.

Der wesentliche Grund, diesen Punkt unter dem Titel „Qualität“ zu erwähnen, ist die Einfachheit und Zuverlässigkeit des Verbindungsaufbaus: Insbesondere Sicherheits- und Priorisierungsmaßnahmen werden noch für einige Zeit dazu führen, dass IP-Adressen häufig geändert werden, falls die Videokonferenzsysteme überhaupt direkt, d.h. ohne zwischengeschalteten Proxy, erreichbar sind. In naher Zukunft wird auch die Verwendung von ENUM [19] die Qualität von H.323 Videokonferenzen merklich erhöhen, da der Anwender nun nur noch eine Nummer wählen muss, und sich dabei „keine Gedanken“ über den zu verwendenden Service machen muss.

Die Verwendung einer Gatekeeper/Proxy Kombination bringt noch weitere Qualitätsvorteile sowohl für den Benutzer als auch den Administrator mit sich. Benutzer können, wie bereits oben erwähnt, E.164 Nummern zur Anwahl verwenden. Deren Desktopsysteme können durch Firewalls gesichert werden. Der Administrator kann mittels des Gatekeepers Anzahl und Dauer von Konferenzen in ermitteln<sup>13</sup> und diese Statistiken beispielsweise in Berichten an die Leitung und/oder in Jahresberichten der Institution verwenden. Weiterhin kann man mittels externer Tools beispielsweise das monatliche Datenaufkommen auf dem Gatekeeper messen, wenn die Proxyfunktionalität für alle Verbindungen (intern ↔ intern, extern ↔ intern, interne ↔ extern, extern ↔ extern) verwendet wird. Mit Hilfe solcher Statistikfunktionen konnte im IPP gezeigt werden, dass im Monatsmittel ca. 1000 Konferenzen durchgeführt werden, die einen Datentransfer von etwa 150GB bedingen. Weiterhin konnte festgestellt werden, dass derzeit ca. 42% aller Verbindungen externe<sup>14</sup> Teilnehmer einschließen.

### 5.3 Sicherheit

Die Protokollfamilie H.323 bedingt einen sehr komplexen Verbindungsaufbau und nutzt dabei mehrere IP-Verbindungen für Signalisierung, Ton und Bild parallel zu einander, auf verschiedenen, teils dynamisch ausgehandelten IP-Ports.

Als Firewallstrategien kommen – grob zusammengefasst – nur solche in Frage, die entweder

1. Systeme, die nicht durch eine Firewall geschützt werden, wenigstens so schützen, dass verschiedenen Services, beispielsweise telnet und http nicht von außen zu erreichen sind und vor allem die Konfigurationsservices, die bei vielen Videokonferenzsystemen darüber erreichbar sind, mit einem Passwort schützen,
2. einen großen Bereich von Ports für alle Quell- und Zieladressen zum Verbindungsaufbau in beiden Richtungen öffnen,<sup>15</sup> oder
3. eine Firewall-Lösung verwendet, die unter anderem (!) H.323 Pakete auf Signalisierungsports genau inspizieren, die ausgehandelten dynamischen Ports freischalten usw., oder
4. einen dedizierten Proxy verwenden, mit dem sowohl interne als auch externe Systeme wie unter 2) kommunizieren dürfen und auch die Audio- und Video-Datenströme durch diesen Proxy gelenkt werden.

---

<sup>13</sup> Solche Statistiken sind immer ein gutes Mittel um die generelle Akzeptanz und Funktionalität des Systems und der verwendeten Technik aufzuzeigen.

<sup>14</sup> Externe Teilnehmer in diesem Fall sind sowohl eigenständige Videokonferenzsysteme, aber auch MCUs.

<sup>15</sup> Während des Setups jeder Verbindung handeln die Videokonferenzsysteme einen zu verwendenden Portbereich aus. Der dazu verwendete Portbereich beinhaltet alle Ports größer 1024 und kleiner 65535. Im Regelfall wird dies durch das rufende System induziert. Bei MCU Konferenzen gibt die MCU den zu verwenden Portbereich vor. Daher empfehlen wir keine Portbeschränkungen auf Proxies, selbst wenn dies möglich ist. Mehr Informationen dazu findet man auch in [20,21]

Für die Strategien 1 und 2 wird man – zu Recht – kaum Akzeptanz beim Verwalter einer Firewall finden können – zumal, wenn man in anderen Szenarien als den bisher geschilderten an den Einsatz von PC-basierten Einzelplatz-Systemen denkt.

Die Strategie 3 scheitert auch aktueller Erfahrung nach daran, dass kommerzielle Firewall-Systeme (z.B. Cisco PIX) zwar als H.323-fähig vermarktet werden, sie aber immer wieder an Upgrades, z.B. zur H.323 Version 4 scheitern. Auch wenn dies „nur“ einige Wochen dauert: Diese Service-Unterbrechung ist nicht tragbar.

Aus unserer Sicht ist die Kombination GNU-Gatekeeper und –Proxy die derzeit einfachste und empfehlenswerteste Lösung des Problems [20,21]. Sie müsste aus Sicht der Sicherheit ausreichend sein und induziert, auf Standardhardware unter Linux implementiert, keine erkennbaren Qualitätsprobleme bei Durchsatz und Latenz-Schwankungen.

#### **5.4 Mehrpunktkonferenzen**

H.323 sieht zur Implementierung von Mehrpunkt-Konferenzen eine Mehrpunkt-Konferenzeinheit (MCU) vor. Diese realisiert Verbindungen zwischen den einzelnen Teilnehmern und ihr selbst. Sie empfängt die eingehenden Ton- und Bilddatenströme, errechnet daraus ein den Anforderungen entsprechend aufbereitetes Mischsignal und sendet dieses Mischsignal an die einzelnen Teilnehmer. Sind alle teilnehmenden Endgeräte vom selben Typ und arbeiten alle mit denselben Verbindungsparametern, also z.B. dem Videostandard H.323, einer Bandbreite von 768 Kbps, einer 4CIF Bildauflösung und einem H.263 Standard, ist eine robuste Mehrpunkt-Konferenz mit den uns heute zur Verfügung stehenden Geräten möglich.

Derzeit stehen uns folgende Varianten zur Verfügung:

- Eingebaute MCUs in den Endgeräten (z.B. Polycom, SONY und Tandberg) für bis zu 4 Standorte, die sowohl über IP (H.323) oder über ISDN (H.320) teilnehmen können. Damit stehen automatisch Gateways zwischen IP und ISDN zur Verfügung. In diese MCUs können sich auch ein oder mehrere Telefonteilnehmer einwählen.
- Die vom DFNVC betriebenen Radvision MCUs, darunter eine für die HGF und die MPG speziell beschaffte, erlauben das Aufbauen von Mehrpunkt-Konferenzen mit erheblich mehr Teilnehmern. Auch beim DFNVC sind Gateways für ISDN, IP, und auch schon für SIP verfügbar.

Mit den erwähnten neueren Soft- und Hardwareversionen der MCUs sind prinzipiell interessante Erweiterungen möglich. Die MCU kann auch so betrieben werden, dass einzelne Teilnehmer mit verschiedenen Bandbreiten und Codecs arbeiten, und jeder mit dem für ihn optimalen Ergebnis beliefert wird (Transcoding). Praktisch empfehlen können wir dies zurzeit nur eingeschränkt, da wir in unseren umfangreichen Tests und realen Konferenzen noch keinen wirklich stabilen Betrieb dieser Art ohne sporadische bis häufige Qualitätseinbußen für alle Teilnehmer erreichen konnten. Die Ursache hierfür dürfte vor allem in den unterschiedlichen Implementierungen der entsprechenden Standards durch verschiedene Hersteller von MCU und Endgeräten liegen. Wir denken dabei vor allem an die Algorithmen, mit denen aus vorgewählten oder zur automatischen Auswahl freigegebenen Bandbreiten und Codecs diejenigen ausgewählt werden, mit denen dann Verbindungen aufgebaut werden.

So sind derzeit stabile optimale Ergebnisse bei mittleren und großen Teilnehmerzahlen kaum anders zu erzielen als unter Einsatz von Geräten des gleichen Herstellers (auf gleichem Softwarestand!). Im Falle von MPG und HGF haben wir daher erfolgreich für eine homogene Ausstattung mit jeweils mindestens einem Tandberg-System geworben und durchgesetzt [5].

### **5.5 Präsentationsübertragung**

Die internationale Standardbehörde ITU hat im Juni 2003 den diesbezüglichen Umbrellastandard H.239 verabschiedet (s. Abschnitt 3.4). Er behandelt das Senden von Präsentationen mit maximal XGA (1024x768 Pixel) Auflösung durch einen Teilnehmer und die Wiedergabe durch einen oder mehrere (MCU-) Teilnehmer. Hierbei ist die mögliche Bildrate beschränkt durch die Bandbreite, da die Bandbreite proportional zu dem Produkt aus Bildrate und Bildauflösung ist. Höhere Auflösung bei gleich bleibender Bandbreite erzwingt also eine Reduzierung der Bildrate.

H.239 erlaubt den Einsatz verschiedener Videocodecs, dabei wäre natürlich H.264 von zusätzlichem Interesse, da dieser Kompressionsstandard effektiv Bandbreite spart.. Zurzeit bieten lediglich VCON und Polycom in ihren PC-basierten Systemen diese Kombination an. Hier sind erheblich höhere Bildraten von bis zu 15 fps bei 3 Mbps Bandbreite erreichbar - leider wieder nur in proprietären Lösungen. Die meisten hardwarebasierten Geräte sind noch auf maximal 384 Kbps für H.264 beschränkt und setzen die Kombination H.239 und H.264 für Präsentationen überhaupt noch nicht ein.

Der Stand der Interoperabilität zwischen Geräten verschiedener Hersteller wird laufend von der Videogruppe des RZG getestet. Die Ergebnisse werden unter [22] veröffentlicht und regelmäßig aktualisiert. Die kürzlich angekündigte engere Zusammenarbeit zwischen Sony und Radvision, sowie die Ankündigungen von Tandberg und Polycom im Sommer 2004 die H.239 konforme Softwareversionen E.4 bzw 7.0 frei zu geben [23], lassen hoffen, dass diese in höchstem Maße benutzerfreundlichen Taktiken der Hersteller wenigstens in diesem Punkt in den nächsten Monaten der Vergangenheit angehören werden. Dann können auch die Radvision MCUs beim DFNVC mit der neuen Software für H.239 aufgerüstet werden.

Eine seit langem bewährte, robuste und kostenfreie Alternative sind die verschiedenen VNC-Versionen [10]. VNC steht für „Virtual Network Computing“ und erlaubt das Austauschen und wechselseitige Bedienen einzelner Fenster oder des gesamten Desktops des sendenden PCs (des „Servers“) durch einen oder mehrere Teilnehmer („Viewer“). Die Zugriffsrechte können dabei flexibel eingestellt werden. VNC ist für alle Plattformen (Windows, UNIX, LINUX, Mac-OS) verfügbar und erlaubt auf der Viewerseite eine Skalierung des empfangenen Bildes auf die gewünschte Auflösung. Bemerkenswert ist, dass Tandberg VNC benutzt, um Präsentationen über H.323 und H.320 zu ermöglichen, auch wenn das Endgerät (vor allem die kleineren Modelle) kein DuoVideo erlaubt, aber am selben LAN wie der Präsentations-PC hängt. Diese Lösung bringt jedoch weitaus schwieriger zu beantwortende Fragen zur Sicherheit mit sich.

## **6 Laufende und notwendige Entwicklungen**

Die Ausführungen in dieser Arbeit haben an vielen Punkten deutlich gemacht, dass von allen beteiligten Partnern auf allen Gebieten Weiter- bzw. Neuentwicklungen dringend erforderlich sind. Die Partner sind dabei:

- Einrichtungen und Institute der HGF und der MPG
- Hersteller der Endgeräte und MCUs
- Firmen zur technischen Realisierung der Technik in den VC-Räumen
- DFN VideoConferencing Service

Die Gebiete auf denen Entwicklungen laufen, bzw. notwendig sind, lassen sich so zusammenfassen:

- Ausstattung und Einrichtung der Räume nach Nutzeranforderungen
- Funktionalität (Audio, Video, Präsentationen)
- Konferenztypen
- Endgerätestandards
- ITU Standards
- Netzwerküberwachung: Sicherheit, QoS

Die Einrichtungen bzw. Institute müssen sich darüber klar sein, dass Videoconferencing nicht zum Nulltarif erhältlich ist. Videoconferencing ist hier als zusammenfassender Oberbegriff für alle behandelten Komponenten definiert. Die Qualität, der Leistungsumfang und die Akzeptanz von Videoconferencing (Q) ist proportional dem Produkt aus Manpower (M) für den Betrieb, Geld (€) für Investitionen, Outsourcing (O) und Zeit (t):

$$Q \sim M \epsilon O t$$

Wichtig an dieser Faustformel ist, dass kein Faktor Null sein darf. Die Mbone Befürworter glaubten, dass Videoconferencing zum Nulltarif ( $\epsilon = 0$ ) möglich sei. Leider geht dann auch die Qualität und die Akzeptanz auf Null, auch wenn man die Netzinfrastruktur von der VRVS Gruppe bezieht ( $O > 0$ ) und (un-)endlich Zeit  $t$  investiert. Aber auch die Outsourcingbefürworter müssen einräumen, dass ohne Manpower in den Einrichtungen, die für endliche Zeit verfügbar ist, keine noch so hohen Investitionen akzeptable Qualität für die Teilnehmer sicherstellen können.

Wir befürworten intensive Entwicklungsarbeit zu folgenden Probleme:

- Die Gerätehersteller müssen gezwungen werden, die von ihnen selbst verabschiedeten ITU Standards so zu implementieren, dass alle wesentlichen Funktionen zwischen Systemen aller Hersteller interoperabel zur Verfügung stehen. Dies gilt insbesondere für den H.264 Videostandard zur Übertragung hochwertiger Videobilder und den H.239 Standard zur Übertragung von Präsentationen bis zu XGA Auflösung mit möglichst hoher Bildrate.
- Die Palette an Diensten, die der DFN im Rahmen des DFNVC anbietet, muss weiter erprobt, zuverlässig angeboten und gründlich dokumentiert werden.
- Der „Premium Service“ des DFNVC für HGF und MPG muss die vertraglich geregelten Anforderungen erbringen. Darunter fallen Unterstützung beim Aufbau von Firewall- und Gatekeeperlösungen, sowohl personell als auch mit technischen Hinweisen und Hilfestellungen, bis hin zum (kostenpflichtigen) Betrieb von Gatekeepern und Proxies bei den Einrichtungen durch Personal des DFNVC.
- Zusammen mit dem DFN müssen Verfahren zur Netzwerküberwachung und vor allem zur Garantie hochwertiger Verbindungen zwischen den Endgeräten und den zentralen MCUs entwickelt und bereitgestellt werden.
- Die erwähnten QoS Maßnahmen sind vom DFN zentral zu koordinieren und allen Nutzern des DFNVC zur Verfügung zu stellen. Hierbei ist vor allem die Betreuung der HGF und MPG im Rahmen des Premium Service als Vorreiter zu gewährleisten.
- Die Mitarbeit möglichst vieler Einrichtungen (bisher vor allem AWI und IPP) zur Erreichung dieses Ziels ist auf breiter Basis anzustreben. Das heißt Mitarbeit aller Rechenzentren und der involvierten Mitarbeiter.

## **7 Zusammenfassung und Ausblick**

Die in dieser Arbeit beschriebenen Eigenschaften einer funktionierenden und von allen Benutzern akzeptierten Videoconferencing Infrastruktur sind nicht nur in die Zukunft gerichtete Forderungen an Anbieter und Dienstleister oder Absichtserklärungen sondern basieren auf mehrjährigen Entwicklungen und Erfahrungen in einigen HGF und MPG Einrichtungen. Dazu gehören die Aktivitäten bei DESY, die bisher überwiegend auf ISDN basierten, die Einsätze beim AWI, wo neben ISDN- auch Satellitenverbindungen genutzt und in Megaconference II und III eingebunden wurden, und die praktischen Erfahrungen aus nunmehr mehr als 5 Jahren ISDN und IP Konferenzen zwischen dem IPP in Garching und dem Teilinstitut in Greifswald.

Wir sind also überzeugt, dass qualitativ hochwertige, zuverlässige und mit überschaubarem finanziellen und vor allem vertretbarem Personalaufwand zu betreibende Videokonferenzen zunehmend zu einem akzeptierten und im weiteren unverzichtbaren Mittel in der wissenschaftlichen Zusammenarbeit und Projekt- und Programmplanung in HGF und MPG werden. Dies ist bereits jetzt in der Hochenergiephysik und Energieforschung mit ihren wenigen großen Beschleunigern und Großexperimenten, die von Teams aus aller Welt benutzt werden, sichtbar. Als nächstes werden wohl die Disziplinen folgen, in denen vom Thema her multinationale Zusammenarbeit in Projekten notwendig ist – wie der Erdsystemforschung. Schließlich werden sich wohl auch zum Beispiel EU-Projekte dieses Mittels bedienen, wenn erkannt wird, dass diese Technik Zusammenarbeit und Produktivität fördert.

Somit erscheint uns die Weiterentwicklung dieser Kommunikationsform als wesentliches Element beim Aufbau einer eScience-Infrastruktur. Eine Zusammenstellung der unabdingbaren Aktivitäten und der dazu notwendigen Partner sind im letzten Kapitel dieser Arbeit gegeben.

## Literaturverzeichnis

- [1] Video Development Initiative : *Videoconferencing Cookbook*;  
<http://www.videnet.gatech.edu/cookbook/>
- [2] Kompetenzzentrum für Videokonferenzdienste, Dresden: *Videokonferenz-Handbuch*;  
<http://vcc.urz.tu-dresden.de/vc-handbuch/>
- [3] Max-Planck-Gesellschaft zur Förderung der Wissenschaften: *Zahlen und Fakten*;  
<http://www.mpg.de/ueberDieGesellschaft/profil/zahlenUndFakten/index.html>
- [4] Helmholtz-Gemeinschaft Deutscher Forschungszentren: *Daten, Zahlen, Fakten*;  
[http://www.helmholtz.de/de/Wir\\_ueber\\_uns/Profil/Daten\\_\\_Zahlen\\_\\_Fakten.html](http://www.helmholtz.de/de/Wir_ueber_uns/Profil/Daten__Zahlen__Fakten.html)
- [5] Schwenn, U.: *Video over IP - Establishing an H.323 Videoconferencing Infrastructure for the German Max-Planck-Gesellschaft*, IMTC/Wainhouse Research European Forum, May 2003, Geneva, Switzerland; [http://www.rzg.mpg.de/vc/docs/IMTC-WR\\_D1USchwennSes6\\_2.pdf](http://www.rzg.mpg.de/vc/docs/IMTC-WR_D1USchwennSes6_2.pdf)
- [6] ITU-T Recommendation G.722 (1988), 7kHz audio-coding within 64 kbit/s
- [7] ITU-T Recommendation H.263 (1998), Video coding for low bit rate communication
- [8] ITU-T Recommendation H.261 (1993), Video codec for audiovisual services at p x 64 kbit/s
- [9] ITU-T Recommendation H.239 (2003), Role management and additional media channels for H.300-series terminals
- [10] TightVNC; <http://www.tightvnc.com/docs.html> ;  
RealVNC; <http://www.realvnc.com/documentation.html>
- [11] Abschlussbericht des DFN-Projekts "Telekonferenz der Bayerischen Rechenzentrumsleiter", <http://www.tkbrzl.rze.uni-erlangen.de/doc/TKBRZL.pdf>
- [12] *VRVS Frequently Asked Questions*  
<http://www.vrvs.org/Documentation/faq.html#Basics01>
- [13] ITEC-Ohio; *H.323 Beacon Troubleshoot Before, During, and After Videoconferences*,  
[http://www.itecoho.org/beacon/h323beacon\\_overview.pdf](http://www.itecoho.org/beacon/h323beacon_overview.pdf)
- [14] WiN-Labor Erlangen; *Verfahren der Leistungsbeschreibung für den DFN-Internet-Dienst*; [http://www-win.rze.uni-erlangen.de/ippm/dfn-internet-qos\\_de1.html](http://www-win.rze.uni-erlangen.de/ippm/dfn-internet-qos_de1.html)
- [15] OARNet; <http://www.megaconference.org>,
- [16] Pfeiffenberger, H.: *K-12 Students Explore Work and Life in an Antarctic Station via H.323 (live)*, 17th Annual Conference on Distance Teaching & Learning, Madison, Wisconsin . 2001; <http://www.awi-bremerhaven.de/Publications/Pfe2001b.pdf>

- [17] ITU-T Recommendation E.164 (1997), The international public telecommunication numbering plan
- [18] Verharen, E.: *The Global Dialing Scheme Explained*, <http://www.vide.net/help/gdsintro.shtml>
- [19] Huston, G.; *ENUM – Mapping the E.164 Number Space into the DNS*, [http://www.cisco.com/warp/public/759/ipj\\_5-2/ipj\\_5-2\\_ENUM.html](http://www.cisco.com/warp/public/759/ipj_5-2/ipj_5-2_ENUM.html),
- [20] Stöckigt, K.; *Videokonferenzen aus sicheren Netzen*, Bachelor Thesis, GWDG Göttingen, 2002; [http://www.rzg.mpg.de/vc/docs/kfs\\_bsc.pdf](http://www.rzg.mpg.de/vc/docs/kfs_bsc.pdf)
- [21] Stöckigt, K.; *Videoconferencing & Firewalls, GnuGK – A cheap solution for everybody*, 2. Workshop des VCC, Dresden, 2002; [http://www.rzg.mpg.de/vc/docs/kfs\\_gnugk.pdf](http://www.rzg.mpg.de/vc/docs/kfs_gnugk.pdf)
- [22] RZG-Videoconferencing Projects; <http://www.rzg.mpg.de/vc/projects.php>
- [23] Schwenn, U.; Private Communication by Wainhouse Research, Sony, Radvision, Tandberg, Polycom