Workshop on

Future Heterogeneous Networks

Mountain View, California, March 24-25, 2011

Organizing Committee

Vincent Chan, Massachusetts Institute of Technology, Chair

John Chapin, Massachusetts Institute of Technology

Pat Elson, NASA Ames Research Center

Darleen Fisher, National Science Foundation

Victor Frost, National Science Foundation

Kevin Jones, NASA Ames Research Center

Grant Miller, National Coordinating Office for Networking and Information Technology Research and Development (NITRD)

Table of Contents

Figures, Tables and Lists

1 Executive Summary

The workshop **FutureHetNets 2011** on "Highly Controllable Dynamic Heterogeneous Networking" was held on March 24-25, 2011 at NASA Ames Research Center in Mountain View, California. It was sponsored by the Large Scale Networking Coordinating Group (LSN) of the Networking and Information Technology Research and Development (NITRD) interagency community, and supported by the National Science Foundation (NSF) and NASA. The workshop brought together 74 leading researchers from US institutions to discuss the research and development activities needed to enable the end-to-end, scalable, highly controllable, secure heterogeneous networks of the future.

The workshop was designed to uncover tough networking problems, especially those driven by the properties of new physical layer communication systems, and explore architecture constructs that may provide realistic solutions for the realization of an integrated heterogeneous network. Workshop participants were encouraged to consider fundamental architectural changes, covering the entire stack from the Physical to the Application Layers that hold the potential for major breakthroughs in heterogeneous network performance.

Computing power will increase dramatically in the near future with the development of advanced multi-core processors and cloud computing and storage. The limiting factor on how fast new applications will develop is the availability of high network speeds and much better quality of service at reasonable costs. Device technologies and hardware subsystems are mature enough to provide at least two orders of magnitude increase in network speeds. However, left to incremental developments, the current Internet architecture will not be able to support heterogeneous applications over heterogeneous networks at affordable costs. The principal reason is that the current architectural partitioning of networks into layers has run its course and is approaching the saturation point for further major improvements.

The current layer partitioning may not be the best for investigating radically different and better network architectures. Indeed, recent developments in the physical layer of communication systems lead to functions and behaviors that do not fit current network architecture assumptions regarding physical layer properties. Some examples are:

- wireless nodes with multiple antennae supporting higher-layer control of bandwidth, antenna beam patterning and power allocations;

- optical networks capable of dynamically establishing and modifying high-bandwidth end-to-end transparent light-paths; and
- communication satellites incorporating packet switching in which the combination of high latency and inherent multicast lead to routing behaviors fundamentally different from terrestrial wireline routers.

Few upper layer network architectures in recent years have taken advantage of the new space of possibilities at the physical layer. Substantial network performance improvements can be achieved if future network research emphasizes and exploits the interplay of the physical layer and the higher network layers including applications. Hitherto network management and control has not demanded focused investigation. Sessions were short and have small to moderate sizes, enabling fluctuations to be smoothed by data plane statistical multiplexing. This limited the required response rate of the control plane. In the future this approach will no longer be sufficient, due to new behaviors such as “elephant sessions” that occupy an entire optical network wavelength for durations as short as seconds, and broadband wireless sessions bursting at multi-gigabits per second. Effects like these will reduce the required speed of control plane adaptation from minutes to seconds and sub-seconds. With changes of such magnitude, network management scalability and network state sensing and propagation are big concerns. This is more than an isolated problem in the control plane. The physical, routing and transport layer architectures can and must be designed and tuned to relieve the pressure on the control plane. Thus, a key addition to networking research is a scientific treatment of network management and control and its interactions with network layer architectures. With a new architecture the network may finally have the responsiveness and cost structure to match and utilize the capability of today’s emerging device and physical layer technologies. This report summarizes a number of suggested research directions that the participants thought would be fruitful to pursue. We feel we are at the dawn of a new network research era where disruptive rather than incremental improvements are there for the taking. We have consciously made the decision to include all reasonable and rational suggestions for research directions in the report, without yielding to the urge to make value judgment on which subsets of ideas would be the best. Indeed it is too early and presumptuous to tell which ideas would be the winners without further deeper investigations and understandings!

Vincent Chan, November 20, 2011.

2 The Challenge for Heterogeneous Networks

The overall goal for research on heterogeneous networks is: **to enable users to obtain and share necessary and timely information in the right form over an integrated heterogeneous dynamic network that is scalable, evolvable and secure**. The performance metric most important to end users is the ability to get the information they need in a timely and secure fashion. The data must be in useful form within the constraints of the network such as bandwidth, delay, display capabilities and resolution. The network must be an integrated network so users can communicate with each other despite using different communication modalities that have different characteristics and capacity constraints (e.g. fiber, wireless, satellite communications; Figure 2.1). In addition to heterogeneity in communication modalities, the network must support multiple other forms of heterogeneity, including significant variation in user session requirements, protocol stacks, and rate of change of capacity and connectivity. The network must be able to grow to global scale and incrementally evolve as user requirements and the operating environment evolve over time.

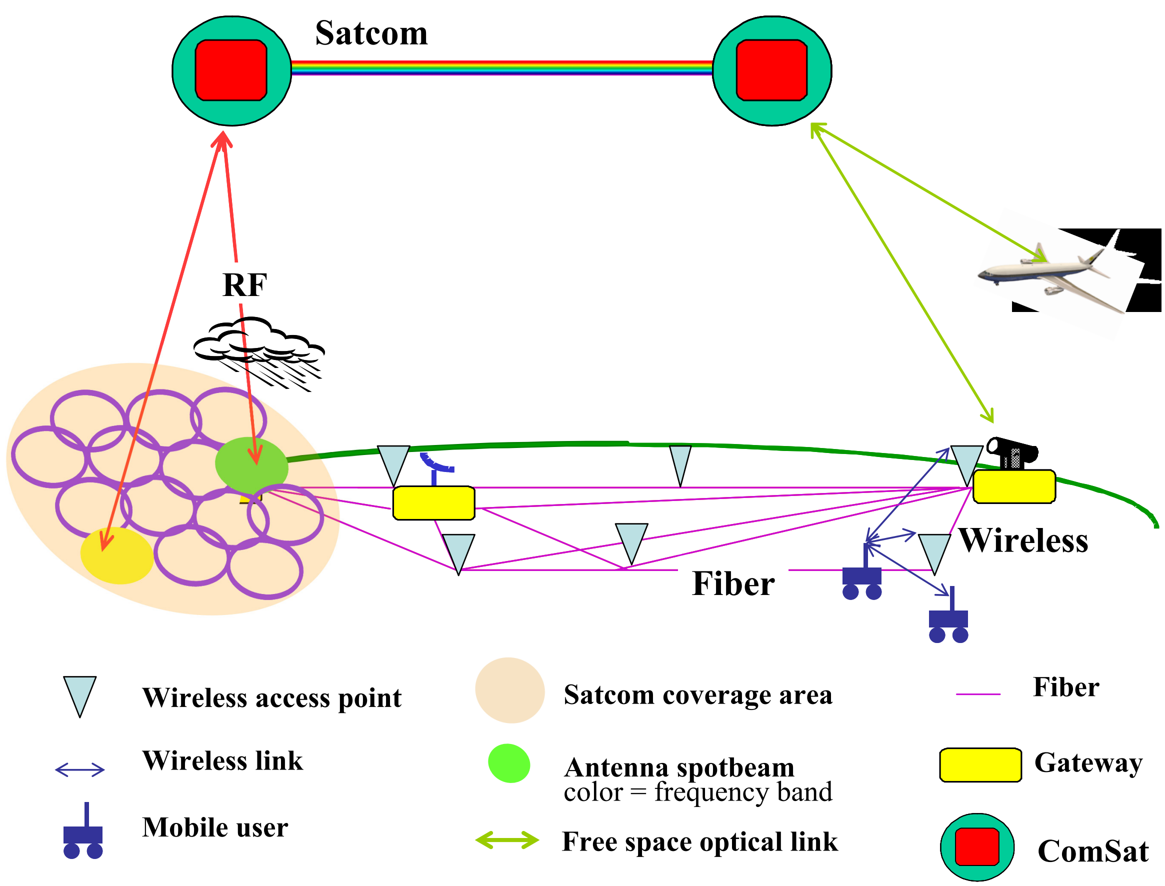

Figure 2.1 Heterogeneous network with interconnected modalities

2.1 Types of heterogeneity

Research on heterogeneous networks includes work focused on multiple different types of heterogeneity. Each is important to address efficiently and effectively in a future integrated network design.

1. Heterogeneous communication modalities: terrestrial wireless, optical fiber, satellite communications, free-space optical links, and new technologies not yet widely deployed.
2. Heterogeneous channel types: “channel type” refers to the rate of change and predictability of change of channel characteristics including bandwidth, delay, and error rate. Potential channel types include long-term stable channels, slowly varying channels, quickly varying but predictable channels, and stochastic channels.
3. Heterogeneous technology generations, integrated into a single interoperable network via gateways or other interconnection techniques.
4. Heterogeneous services, some packet-oriented and some circuit-oriented, including voice, video, file transfer, short message.
5. Heterogeneous transaction sizes and transfer rates, ranging from short messages sent by low-end mobile devices or sensors to multi-terabyte files sent by high-end systems.
6. Heterogeneous application quality of service requirements:
    a. From strict time deadline to best effort only
    b. From no error acceptable to high error rates tolerable
    c. From no downtime acceptable to random disconnections tolerable
7. Heterogeneous network management and control systems and administrative domains.
8. Heterogeneous protocol types beyond TCP/IP.

The imperative is to develop an integrated network architecture that supports all the above types of heterogeneity simultaneously while using all available resources efficiently and without requiring excessive operator intervention.

2.2 Examples

The following examples illustrate the range of requirements and characteristics encompassed by the above types of heterogeneity. This set of examples is not meant to be complete. It is intended to help the reader to appreciate the challenge faced by future heterogeneous network designs.

1. User characteristics
    a. Streams and bursty files
    b. Time deadline and indifference

    c. Guaranteed delivery and best effort
    d. Low and high rates (8 orders of magnitude differences: low kbps to 100Gbps)
2. Communication modality
    a. Fiber with high bandwidth: relatively stable (becoming less so in future dynamic fiber networks) but potentially with a high delay-bandwidth product
    b. Ethernet with copper and fiber both unmanaged and managed with multiple connectivity
    c. Wireless with stable infrastructure; two different cases are fixed users and mobile users.
    d. Mobile ad hoc networks ("MANET"): dynamic connectivity and capacity, frequent data drop-outs
    e. Satellite networks with bent-pipes, circuits, or packet switch in the sky: dynamic capacity and drop-outs when the communication links are tuned for maximum capacity rather than lower capacity with high stability as is common today.
3. Network types
    a. Packet switching
    b. Circuit switching: carrier class Ethernet (≥ 100Gbps)
    c. Hybrids: Multi-Protocol Label Switching (MPLS), Generalized Multi-Protocol Label Switching (GMPLS)
    d. Dynamic very fast circuit switching (the new problem is dynamic circuit setup and scheduling with high efficiency)
4. End-to-end networking protocols
    a. IP
    b. Heterogeneous protocol translated at the gateways between different modalities
5. Services (there are probably more modes in between these two extremes):
    a. A robust minimum "hard-core" network service providing the communications essential for time critical and high reliability-demand applications. The high resource cost of this service means it can only be used for a fraction of carried traffic and will offer only low to moderate data rate.
    b. A much higher rate best effort "soft-shell" network service appropriate for applications without extreme reliability demands.

It is evident that there is a large range of requirements and technology characteristics. There is every likelihood that more types of heterogeneity will surface in the future. It is vital that the network architecture incorporates the capability to adapt to new network properties and requirements over time. This perhaps is the biggest challenge in the architectural design of future networks.

2.3 Need for fundamental research

The current Internet does not support the types of heterogeneity listed above with good efficiency and with acceptable end user performance, delay and quality of service. There are many reasons that lead to these problems, which can be traced to the fundamental ways in which modern communications modalities and requirements differ from the environment in which the Internet and its protocols were developed.

1. Wireless modalities – which include radio frequency, free-space optical, and less widely-used modalities such as terahertz and acoustic – are much more dynamic than the wired networks for which the Internet and its protocols have been optimized. The time scale of change of key parameters such as capacity, delay and connectivity can be as fast as fractions of a second.
2. Many of the newer modalities are expensive enough – sometimes in financial terms, more often in terms of power or spectrum consumption – that it is essential to achieve high utilization to effectively support user communications needs. This rules out the over-provisioning strategy used in the Internet and its protocols to provide acceptable quality of service for voice, video and mission-critical applications.
3. The Internet design assumes that the boundary between communications modalities corresponds to the boundary between independent (often competing) administrative domains, so the management of each constituent network is handled independently with minimal information exchange across the boundary. In future heterogeneous networks, a single enterprise will often support a common mission by combining multiple modalities, each with an independent management system.
4. In networks where an end-to-end data path may cross subnetworks of heterogeneous channel types, functions of the transport layer – such as deciding when to transmit, how much transmit at each point in time and when to retransmit – are best handled differently for each subnetwork. In the development of current Internet protocols, the relative homogeneity of channel types across wired networks enabled an architecture where transport layer functions are performed by end devices with minimal network information. Thus transport layer functions cannot be specialized to subnetwork needs without gross layer violations (e.g. TCP performance enhancement proxies).

5. The Internet was originally designed for a relatively narrow but generic range of services. The generic best-effort packet service provided by IP functions well for the services traditionally supported by the Internet. As the Internet has evolved into the primary global communications network, new heterogeneous service requirements have been added such as hard time deadline and very large file transfers that are not effectively supported by the current Internet strategy of a generic best-effort packet service plus overprovisioning.

These environmental changes are well-known and many proposals have been made to address each one through linear extensions or minor changes to the existing Internet protocols. Thus far the piecemeal proposals have not coalesced into a broad solution with a clearly articulated pathway for stable long-term growth. A fundamental reconsideration of network architecture and protocols – considering all layers jointly without merging them into one – may be needed to provide a broad solution.

There have been many debates regarding whether fundamental changes to the Internet's architecture are necessary. At this time, the debates cannot be resolved. The research necessary to discover fundamentally new approaches and properly assess them has not been done.

2.4 General research areas

With the high degree of heterogeneity expected in future large-scale systems, a research approach driven by specific modality or service needs is unlikely to lead to coherent architectural solutions for the overall problem of heterogeneous networks. Research investigation must be structured around broad questions that apply to wide areas of the design space. The following list of research areas indicates the type of research questions likely to prove productive.

1. How can the attributes of emerging communication modalities and the physical layers that exploit them be characterized in a way that permits analysis and optimization of higher layers?
2. Can one protocol suite adapt to all modalities or is it necessary to use multiple protocols with application layer gateways?
3. How do algorithms in heterogeneous parts of a network cooperate to accomplish a common goal?
4. Does the protocol stack need to be repartitioned for proper support of the various types of heterogeneity?
5. What are the network level abstractions that can cover the space of future

communication modalities?

6. How should networks be layered, organized, controlled, and used in the presence of rapid changes in underlying attributes and user needs? The rate of change may be several orders of magnitude faster than in the current Internet.
7. Are there new analytical models that provide insight into design, evaluation, optimization and correctness of heterogeneous networks?
8. What transport mechanisms can coordinate the widely different data rates of different subnets?
9. How can network management and control and user resource allocation be made much more responsive even though integrated over heterogeneous networks?
10. How can network management and control respond to requirements in a wide range of time scales from quasi-static provisioning to per session controls for large transactions?
11. How should resources be discovered and addressed across a complex and rapidly changing network?
12. What new security vulnerabilities are created by the architectural changes required to support various forms of heterogeneity, and how should these be addressed?

These are types of questions that the network community has tackled before. However, the increased range of heterogeneity to be tightly integrated into an efficient network and the appearance of new network properties in some modalities create the need for new investigation and fresh solutions. In addition, the need to increase data rates by orders of magnitude may require the development of new constructs for future network and communication architecture.

3 Requirements and drivers

The FutureHetNets 2011 workshop opened with presentations on the requirements that will drive future heterogeneous networks and on possible other developments that will shape the context for any research performed. This section summarizes the views expressed from several perspectives.

3.1 Commercial Service Providers: Meeting the Needs of Growing Traffic[1]

Network service providers face multiple challenging issues as they seek to simultaneously address the technical challenges of operating large-scale networks under growing traffic demands and the business challenges of providing an affordable and competitive service to customers. While every technical issue is also a business challenge, and vice versa, the issues can be segregated into those that are driven primarily from the technical side and those driven primarily from the business side.

Table 3.1 Issues facing commercial service providers

Technical Issues	Business Issues
Total capacity; Higher link speeds	Network Upgrades
Network congestion	Network Neutrality
Traffic Growth	Usage based Charging; Congestion Pricing
Lower cost/bit	Lower cost of operations
Video and IPTv; Cloud Services; New Services	Go up the "food chain"
Information Centric Networking	Delivering Content
Machine-to-machine (M2M) communication	
Multicast	
Reliability and Restoration	

[1] Information in this section summarizes a presentation by K. K. Ramakrishnan of AT&T Laboratories Research, Florham Park, NJ.

3.1.1 Wireline networks

The next figure shows the breakdown of traffic flowing across the inter-city network core of a major US Internet Service Provider (average traffic, summer 2010).

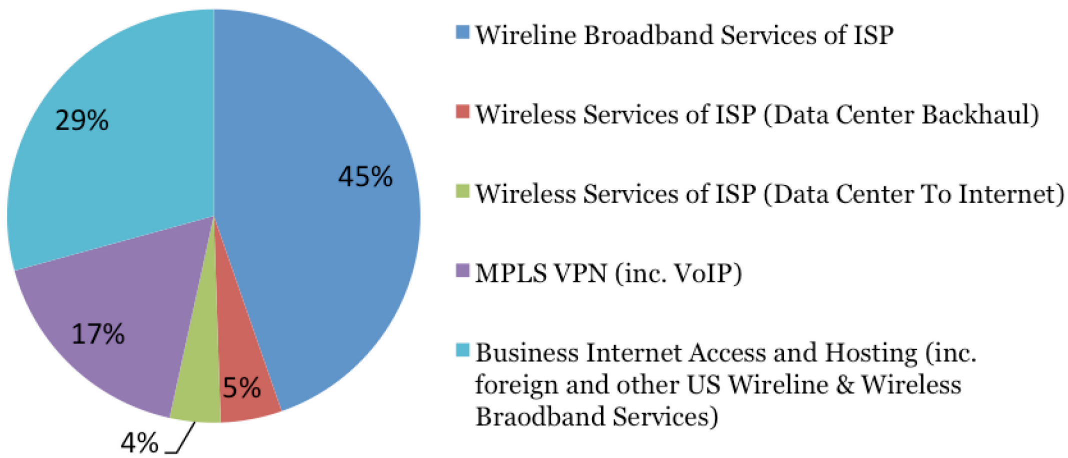

Figure 3.2 US network core average traffic, 2010

The chart shows that for core (inter-city) networks, wireline broadband services are the largest source of traffic. Business traffic accounts for a large share as well. Wireless is not yet significant.

Traffic has been growing rapidly, due to two causes. The first is that each broadband subscriber has generated and consumed an increasing amount of data. The increase has followed a stable exponential growth curve over 32% CAGR (Source: US DSL Downstream traffic per subscriber).

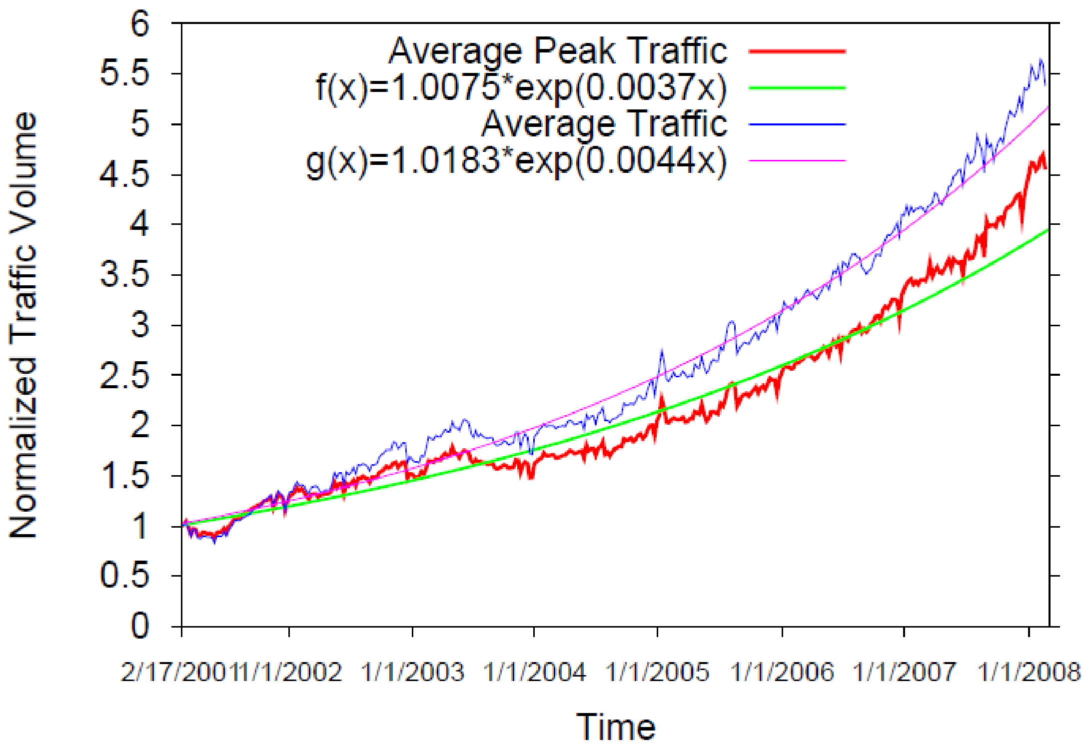

Figure 3.3 US DSL downstream traffic per subscriber

The second primary contributor to increased traffic volume is the increased penetration of broadband to households. The average growth rate was 31%/year from 2000 to 2009 in the US. Annual growth decreased from 48% to 4% as the US market approached saturation (Source: OECD reports from 34 countries).

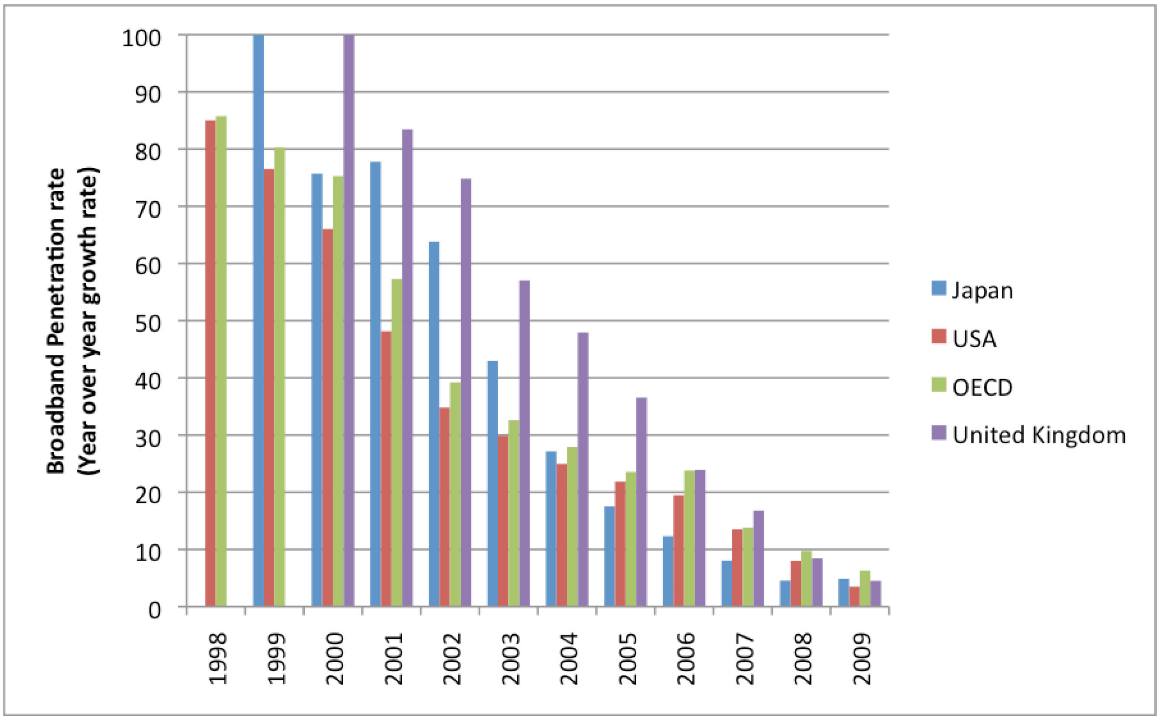

Figure 3.4 Penetration of broadband to households

The traffic flowing over the inter-city core network can be categorized into various services based on the protocol used. The chart shows that HTTP has become the work-horse protocol for many applications, accounting for 60% of the traffic in recent years. (Source: US ISP backbone; Netflow).

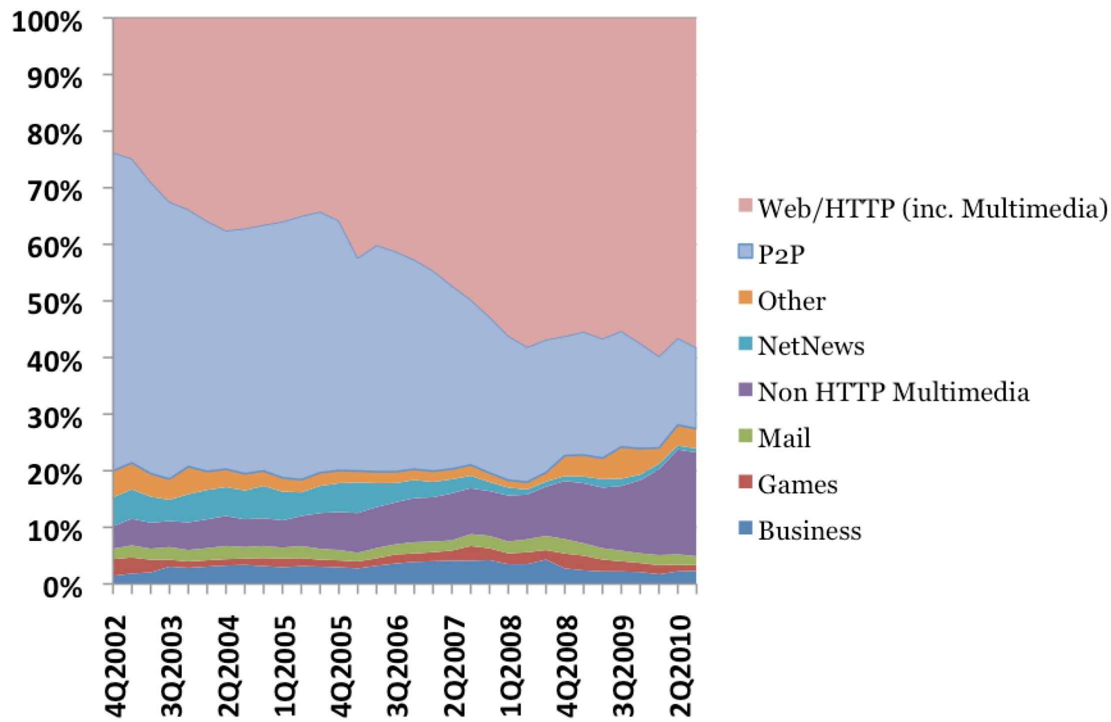

Figure 3.5 US core traffic by protocol

Examining the per-user DSL downlink data shows that the preeminent driver for the growth of HTTP protocol traffic has been the rise of video over the Internet. Lately it has been growing at an annualized growth rate of 83%. This is a major driver for service provider network evolution.

3.1.2 Wireless Networks

Service providers have found that their customers intensely desire mobile services. Despite the rapid penetration of smartphones in the consumer market, tablets may be an even bigger success: the adoption rate for the iPad was the fastest ever for a new product category, surpassing the DVD player (CNBC.com, October 4, 2010). 61% of AT&T postpaid customers now use integrated devices.

The overall growth in mobile data traffic over the last 4 years was 800%. Mobile data bandwidth usage worldwide rose 73 percent during the second half of 2010. Video streaming is the fastest growing application, accounting for 37% of mobile bandwidth. (Source: Allot Communications, Feb. 2011.)

Service providers are aggressively increasing wireless capacity to meet these challenges. However the near term approaches being used (smaller cells, exploiting WiFi) will likely need to be enhanced to sustain growth over the longer term. Moreover, the rapid capacity growth has led to pressure to upgrade the cellular technology even before completing rollout of the previous technology across the network, let alone having earned a return on investment from the current technology.

3.1.3 Comparison of wireline and mobile services growth curves

It appears that mobile data services are following the same pattern as wireline data services, roughly 8 years later. The Y axis in the following chart of monthly average traffic for a US ISP is a log scale; the backbone traffic (blue) grew almost 4 orders of magnitude while mobile traffic (red) is continues to grow. (Note that the rate of growth has been debated.)

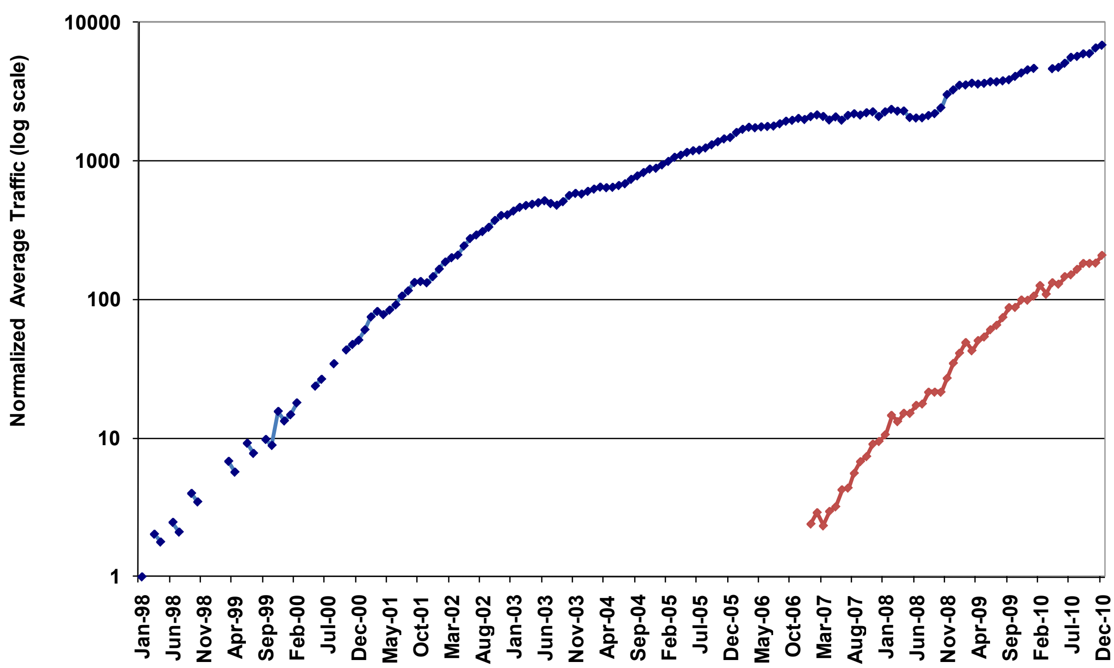

Figure 3.6 US monthly average traffic, wireline vs. mobile services

3.2 Military: Creating an Assured Joint DOD and Interagency Interoperable Net-Centric Enterprise[2]

The primary goal of a defense network should be: **To enable defense users to reliably obtain and share necessary and timely information in the right form over an**

[2] This section is a summary of the report of the same title published by the Defense Science Board, March 2009.

integrated heterogeneous dynamic network which is scalable and evolvable, Figure 3.7. The performance metric most important to the end users is the ability to *reliably* get the end products they need in a *timely* fashion with *integrity*.

In defense applications, timeliness is important due to the nature of warning messages and dynamic command and control in military operations. Though in the civilian sector there are time deadline requirements as well, e.g. electronic stock trading, applications in the defense sector with timely message delivery requirements are much more prevalent and important, and may occupy a larger fraction of the traffic. The timeliness and reliability requirements together likely will drive the future defense network architecture to a state that is different from the Internet; at least leads its development in these dimensions for some number of years[3].

The data must be in usable form within the constraints of the network such as bandwidth, delays and display capabilities and resolution. In the defense sector, there will be a very large dynamic range of the capability of end user devices. More importantly, the users at the network edge can have very different network service variability and in many cases severely challenged. Providing reliable and timely service to these challenged users will be a significant network architectural hurdle.

The network must be a single integrated and also heterogeneous network so users connected to different communication modalities can communicate with each other. The degree of heterogeneity is much larger in the defense sector due to its long lifetime legacy systems and lack of an unified architecture in the past four decades. For the new defense network to run efficiently and economically the architecture may have to select the most current and important modalities for integration and let some legacy and less important systems be served by "converters" appropriately placed in the network, sacrificing some responsiveness.

More so than the civilian sector, the network modalities in the defense sector are extremely heterogeneous in channel properties, data rates and delay performance. Thus, integration from the Physical Layer to all higher Layers of the protocol stack of the network is imperative and must at conception accommodate the reality of the heterogeneity of communication modalities and capacity constraints. Many of the tough network problems at or near the tactical edge are yet to be fully solved and it should be a first priority to address these problems immediately before they impact large acquisition programs.

---

[3] John M. Chapin and Vincent W.S. Chan, "Ultra high connectivity military networks", IEEE Military Communications Conference 2010.

A new architecture study is necessary to uncover tough networking problems, especially those tied to physical layer communication system properties, and explore architecture constructs that may provide realistic solutions for the realization of an integrated heterogeneous defense network.

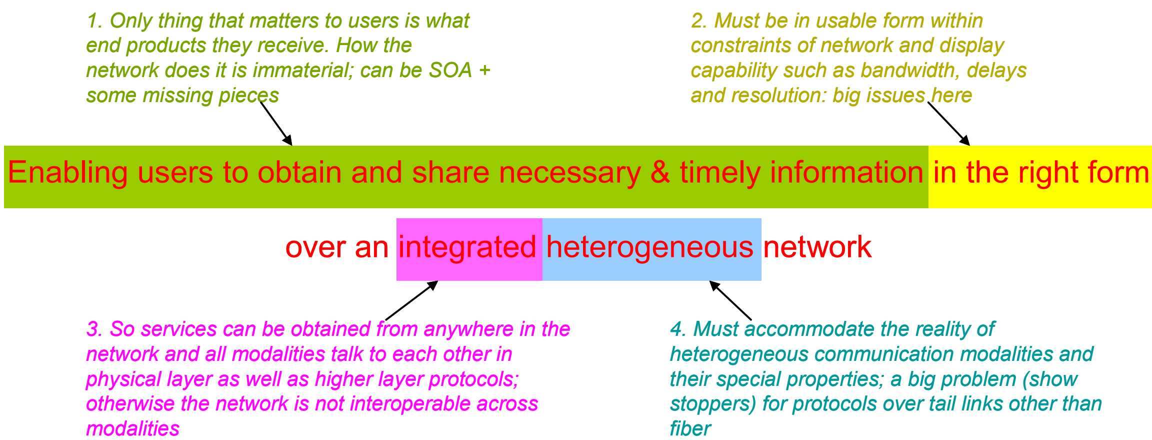

Figure 3.7 Integrated heterogeneous defense network goals

The following is a set of high level architecture goals of an integrated heterogeneous defense network. These high level goals should be used to guide research:

1. The Defense Network should be heterogeneous and accommodate multiple types of networks of different generations, with integration into a single interoperable network via gateways between disparate subnets and a Master Control Plane for subnet integration.

2. The Defense Network should have the following two-tiered network services, Figure 3.8:

   - A robust minimum “hard-core” network service providing the communications essential for successful defense operations in applications such as network management, warning messages and command and control of military operations. This service will have modest rates but solid connectivity at all times with low delays. It should be designed and tested to be both physically survivable and highly secure.

   - A much higher rate “soft-shell” network, which is less robust but offers higher rates and supports commercial-like services such as web-based Service Oriented Architectures.

3. The integrated network should support voice, video, data, hybrid IP as well as circuit based services.

4. There are critical outstanding architecture and technology issues in infrastructure-less wireless networks and some SATCOM networks, Figure 3, that need to be addressed immediately and the deliberations used for the construct of the defense network architecture.

5. It is important to examine the new concept of a Master Control Plane (MCP) to facilitate internetworking and Information Assurance. It is critical for the integration across subnets/domains, interoperability, priority/policy enforcement, access control and security, Figure 3. The MCP can monitor network services in each component networks and sense attacks and compromised subnets as well as reroute around affected assets.

6. The proliferation of information and data services to the tactical edge will generate and consume vastly more bandwidth in the future and will saturate the backhaul network. Network-aware Distributed Information Service (NDIS) technology should be employed near the network edges so that application behavior adjusts as the network changes to maintain critical services on the network, Figure 3. NDIS can be distributedly implemented and works with the MCP to discover resources in the network and broker the best available service for the user application. When satisfactory service cannot be provided with current resources, it can also work with the MCP to reconfigure network resources to satisfy such needs on a priority based basis.

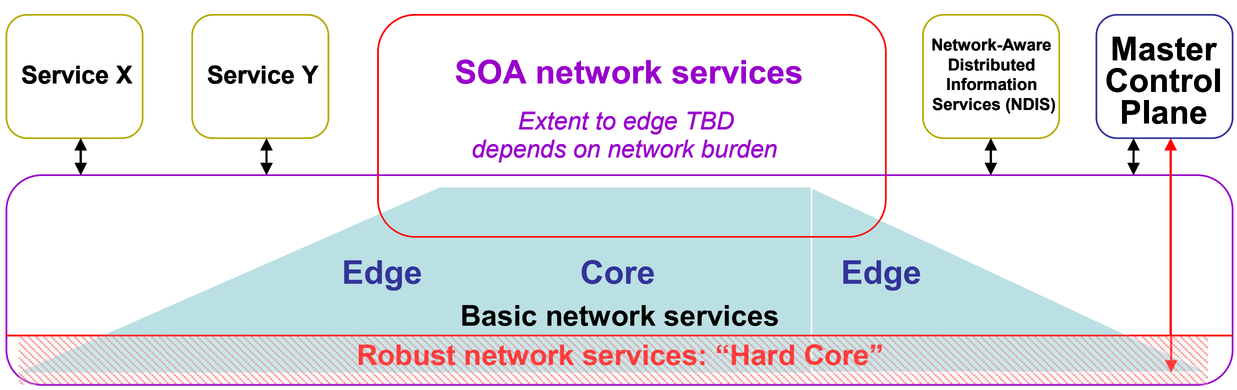

Figure 3.8 Defense network architecture with two-tiered services

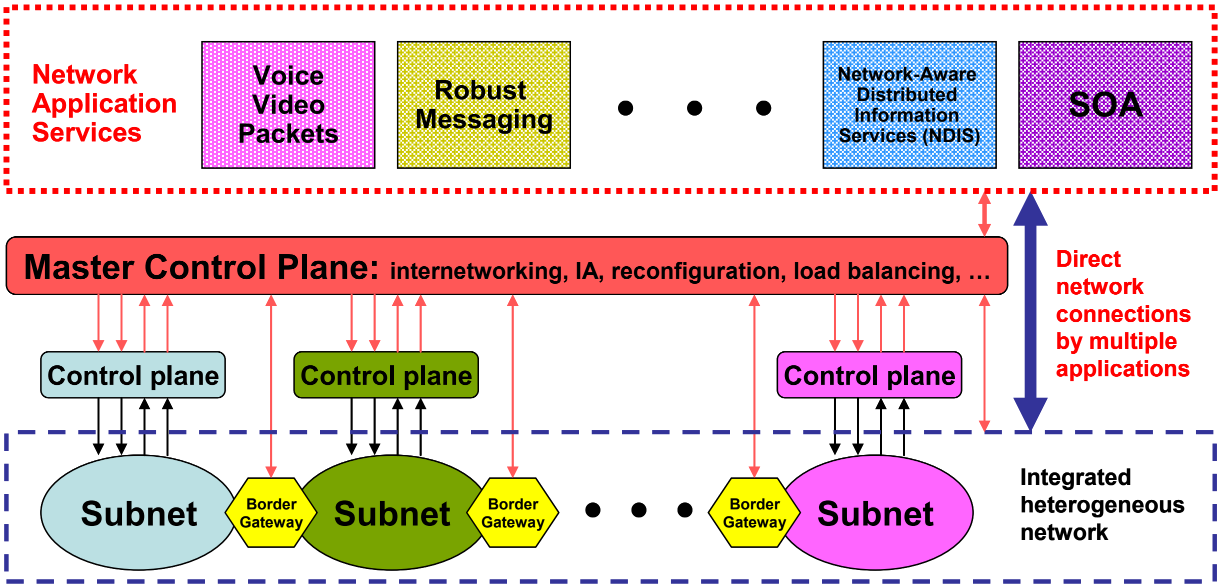

Figure 3.9 Integrated heterogeneous defense network with voice, video, data, hybrid IP and circuit based services

3.2.1 Major tasks ahead

1. Assessment of defense network status and preliminary identification of problem areas and preliminary suggestions for R&D areas.
2. Articulation and justification of comprehensive list of problem areas to be addressed and quantify consequence of failure to provide solutions.
3. Create preliminary architecture concept and identify and initiate the necessary critical research areas.
4. Generate if possible a "Strawman Architecture" and approach to solution of critical network problems as defined. This will support flushing out issues that need more research.

3.3 Basic research: Future Internet architecture projects

The National Science Foundation has sponsored projects to conceive, design and evaluate trustworthy Future Internet architectures. The concepts developed in the FIA projects are exploring fundamental changes in the Internet design that if adopted will interact strongly with efforts to make fundamental changes in support of heterogeneous networking. Four projects are currently in progress: Named Data Networking, MobilityFirst, NEBULA, and XIA (eXpressive Internet Architecture).

3.3.1 Named Data Networking

The NDN FIA project is based on the observation that content is what users and applications care about. The fundamental shift in network design concerns what is named in the network. Current networks name locations, for example with IP addresses. In NDN, data is named rather than location. Data objects (such as photos) become first-class entities with the network responsible for locating them when requested.

The underlying architectural principles of the Named Data Networking FIA project are: Packets indicate what (content) not who (IP address); Packet contains <name, data, signature>; and user-centric trust based on securing named data. Design goals include: retain the hourglass in the architecture; expand the end-to-end principle for robustness in the face of failure; design flow balancing into the thin waist for self regulation; separate routing and forwarding; and facilitate user choice and competition.

3.3.2 MobilityFirst

The MobilityFirst FIA project is based on the observation that mobility of users and nodes is now the norm, which is fundamentally different from the static wireline topology in which today's Internet evolved. Cellular, wireless sensors, machine-to-machine, smart grid & vehicular nets integrate physical world awareness and control into Internet applications. There are frequent disconnections, energy constraints and real-time control applications. All of this results in changes in service, trustworthiness and management.

The architecture uses generalized delay-tolerant networking to provide robustness even in presence of link/network disconnections. GDTN integrated with the use of self-certifying public key addresses provides an inherently trustworthy network. The resulting architecture does not rely on gateways or overlay accommodations.

3.3.3 NEBULA

The NEBULA FIA project is based on the observation that the rise of cloud computing has fundamentally shifted data communication patterns and hence the requirements on the Internet. The network becomes an always-on utility where cloud computing data centers are the primary repositories of data and the primary locus of computation. The data centers are connected by a high-speed, extremely reliable and secure backbone network, with parallel paths between data center and core. Secure access and transit, policy-based path selection and authentication during connection establishment become key design features.

3.3.4 eXpressive Internet Architecture (XIA)

The XIA FIA project is based on the observation that the only constant in the Internet is the high rate of change. Thus a future Internet architecture must be designed to efficiently adapt to accommodate unknown future entities – including hosts, content, and services – while supporting ongoing communications among current entities.

For each type of entity, XIA defines a narrow waist that dictates the application programming interface (API) for communication and the network communication mechanisms. XIA enables flexible context-dependent mechanisms for establishing trust between the communicating principals. Design goals include:

- an evolvable set of first-order principles for communication;
- intrinsic security not dependent on the correctness of external configurations, actions or data bases;
- a narrow waist enabling flexible mechanisms for trust management, bridging human-readable to machine readable, intrinsically secure identifier; and,
- all other network functions supported as service out of or atop the architecture.

3.3.5 Summary

All the FIA projects are expected to address a range of challenges facing the Internet architecture. They must provide solutions for: sophisticated multi-layer trustworthiness; network and configuration management; scalability and control of system complexity; predictable performance; integrating the Future Internet with the optical network; and performance evaluation and comparison of different architectures. The goal is for each project to weave together its various approaches and mechanisms into a coherent, overarching candidate design for a future Internet.

4 Research issues at each layer

Sections 4 through 6 report the detailed results of focus groups that investigated specific issues during FutureHetNets 2011. Because the groups met and drafted their reports independently, the topics covered in each section are varied and occasionally overlapping. Readers seeking a integrative summary of the workshop results may wish to skip to Section 7.

This section reports the results of the focus groups on per-layer research issues that met on the first day of FutureHetNets 2011. Three parallel working sessions considered the physical layer, the network layer, and combined transport and applications layer issues.

4.1 Physical layer

Network growth along multiple dimensions—traffic, servers, subscribers, applications—is the key trend that is generating the need for innovation in the physical layer. In particular, networks today have matured to the point that sustained growth will cause communication technologies to approach scaling and capacity limits. Optical systems are rapidly approaching the spectral efficiency limits for single fiber transmission. RF communications are already constrained by capacity limits. The consequence of these trends is a focus on efficiency and the use of alternative degrees of freedom such as space, cell size, and parallel systems or modalities including novel transmission bands or media. Indeed, these physical constraints are a major driver for increased heterogeneity in future networks. Network scaling becomes a major challenge when moving in these directions because the energy and footprint tends to grow linearly as parallel systems are deployed, resulting in near exponential increase. The maturation of both communication technologies and networks also force new directions. For example, communication systems can no longer rely on microelectronics for energy reductions. In fact, device limitations on chip power densities have motivated the semiconductor community turns to system level, 'more than Moore' solutions.

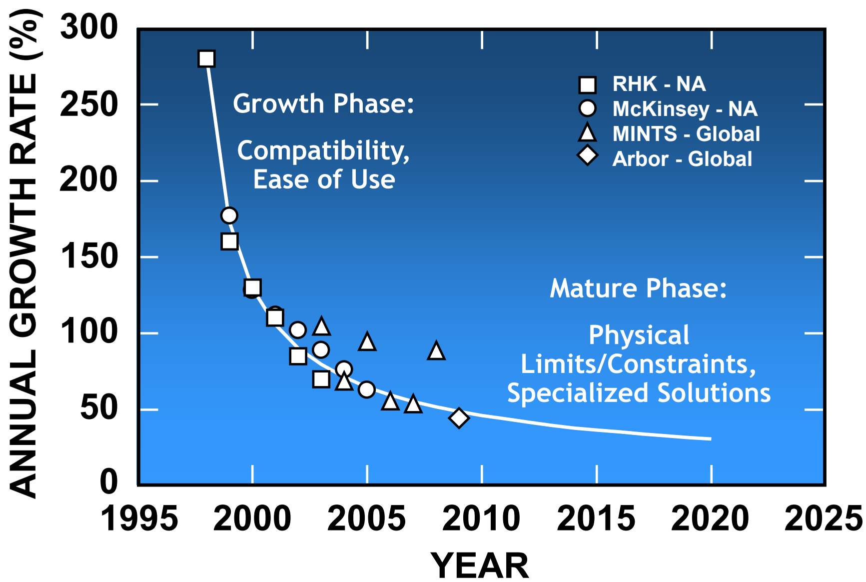

Figure 4.1 Annual growth rate of North American Internet backbone traffic[4]

Data networks have transitioned from rapid, early growth phase to steady, mature growth phase with different technology requirements.

Physical integration of technologies comprising heterogeneous networks, enabling scalable architectures, is an overarching research challenge. From this perspective, research in all areas should address the relevant physical constraints of energy, footprint (including materials and media), and spectrum. Traditional performance requirements, including latency, security, quality of service, mobile coverage, and wireline access, will need to be revisited in light of these new physical constraints and concurrent design requirements. These challenges are identified in List 4.2.

List 4.2 Network Growth Driven Research Challenges

- Physical
    - Energy
    - Footprint, Materials, Media
    - Spectrum
- Performance

[4] Adapted from D. C. Kilper, G. Atkinson, S. K. Korotky, S. Goyal, P. Vetter, D. Suvakovic, and O. Blume "Power Trends in Communication Networks," IEEE J. Sel. Top. Quantum Electron., vol. 17, no. 2, pp. 275-285, 2011.

    - Latency
    - Security
    - Quality of Service
- Accessibility
    - Mobile coverage
    - Wireline access

Unique research challenges will also arise that are specific to the heterogeneous nature of future networks and are shown in List 4.3.

List 4.3 Heterogeneous Physical Layer Network Issues

1. Layering and virtualization, where best to draw the line: network layers and virtualization provide a level of abstraction that simplifies the network operation and management, and can lead to efficiency improvements. These come at a cost of additional overhead, interfaces, and hardware. Finding the right level of virtualization and layering will be essential to scalable heterogeneous solutions

    a. Diversity between applications: ranging from sensor networks to scientific computing, from text messages to immersive telepresence, from centralized content delivery to peer to peer content sharing, how can future physical layer networks support such a wide range of applications?
    b. Unified source management and monitoring: can we find a platform to enable management across communication modalities? Find a unified control protocol so that technologies and networks can be used with seamless integration?
    c. End to end performance measures: in order to invent new end to end, cross-modality, heterogeneous solutions we need methods to predict, measure, and evaluate performance.
    d. Mobile agility and impact on wireline: as mobile traffic increases, what are the consequences on wireline networks? Will greater agility be required in fixed access networks?
    e. Network evolution: standards development and semiconductor cycles can be a decade or more, federated networks create complex technology adoption patterns, how best can research impact networks?

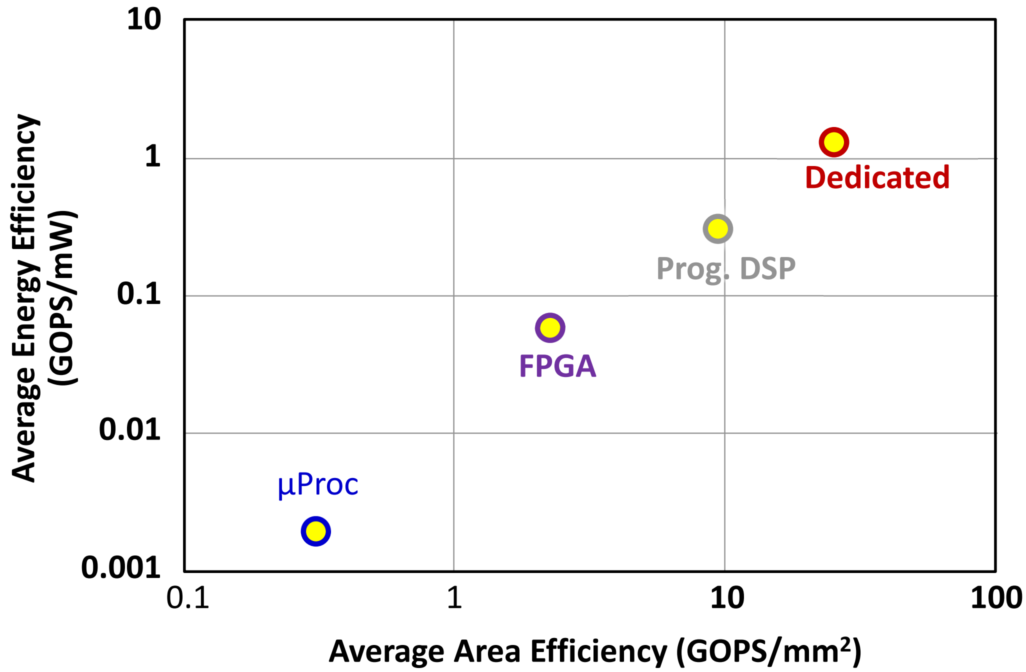

Figure 4.4 Area and energy efficiency of processing technologies used in networks[5]

Dedicated technology solutions, such as ASIC processors, achieve efficiencies needed for future network growth. New methods to achieve intelligent, hybrid inter-working of dedicated hardware will be required for future heterogeneous networks.

High bandwidth services and on-demand networks are requiring new wired and wireless solutions. Examples include private networks, both wavelength and sub-wavelength, vehicular high speed communication, Internet of things, and mobile video services. In many cases it's not just about more bandwidth, but the characteristics of that bandwidth—quality of service, security, and flexibility. These performance and flexibility requirements will often work against network efficiency. Less network partitioning (fewer fences) can allow adaptability, although some partitioning may be necessary to support the limitations of the technology. In general physical layer support at higher layers has two sides:

- Complicates upper layer design to support decisions previously confined to lower layers
- Optimization can be performed at higher layers that can't be accomplished otherwise

Adding intelligence to the physical layer such as monitoring or control plane functionality can assist higher level optimizations. Today, the higher network layers often look to the physical

[5] ISSCC & VLSI 1999-2011, averaged.

layer for stable, low latency, and infinite bandwidth. Relative to packets or historical end user access rates, wavelength connections largely have these characteristics. Going forward, however, this will no longer be the case. Even understanding to what extent the physical layer can be made dynamic—measurement and feedback time scales, control loop delays, stability criteria—is not well understood, particularly under multi-cast, heterogeneous service conditions.

4.1.1 Physical Layer Technologies

Technological challenges for heterogeneous networks are broken down by network segment: core, access, and mobility. Many networks today have far more segments such as interoffice, metropolitan, regional, and back-haul. In this report, core refers to all such network segments primarily involved in the distribution of data between network nodes and access refers to segments that aggregate and distribute data, usually in large tree networks to the end users. Where the division occurs and how many segments are used is an open question for future networks. Each network segment furthermore may include multiple communication modalities such as wireless, fiber optics, or satellite.

4.1.2 Core Networks

Today core networks are almost entirely based upon fiber optics due to the high data capacities associated with traffic that has been aggregated. In fact as traffic has increased, optical technologies have moved steadily toward the edge and even into the access and home. Exceptions include technologies such as microwave backhaul used to reach remote locations. Increasing capacity needs will move core networks toward parallel architectures with multiple fibers per link (even 100's-1000's of fibers) in large mesh networks. Today's core networks are heavily congested with routers acting like numerous toll booths providing operator handover and service processing. An efficient, streamlined superhighway will facilitate future networks. The technology base for such networks does not exist today, including integrated multi-fiber/multi-spatial mode switches, amplifiers, and transceivers as well as the algorithms to optimize and manage these technologies in mesh networks of the scale needed. Networks today that are composed of order 100 nodes, 100 fibers, and 100 wavelengths per fiber will be replaced by networks with of order 1000 nodes, 10^5 fibers with 100-1000 wavelengths per fiber or other network configurations that provide similar levels of capacity and connectivity.

High performance computing and network services such as cloud computing are expected to increase traffic heterogeneity in terms of flow size and service requirements—new services and higher bandwidth services will continue to emerge, while the low end—voice, text messages, etc.

is not going away. Large flows requiring capacities best handled by optical signals will need to be managed along with small packet-size flows, perhaps increasing due to sensor networks and "smart" energy aware paradigms such as smart grid. Several research focus areas were identified to address this challenge: integrated hybrid circuit and packet hardware, physical layer aware and programmable control planes, and sub-wavelength bandwidth resource allocation. Focus areas specific to optical technologies include: fast and reconfigurable photonic switch fabrics, flexible bandwidth/grid optical networks (moving away from a rigid wavelength grid), and physical layer/optical network control algorithms. Physical layer control planes may include the transfer of wavelength connection properties to higher layers—physical plant details, modulation formats, and physical performance parameters such as accumulated dispersion or signal to noise ratio. While many of these general topics have been in the literature for some time, new solutions are needed for the current technology base and from the perspective of heterogeneous networking requirements.

Figure 4.5 illustrates the role of heterogeneity in understanding core network evolution. Today, large transparent networks have been commercially deployed, however, these are static networks in which wavelength switches are used for flexible provisioning and resource management. Dynamic wavelength capability ("dynamic transparency") is not possible in core networks today largely because of the complexity of designing high capacity transmission systems coupled with the high reliability requirements in core networks. In recent years much work has gone into optical packet and burst switching technologies, anticipating the need for very high capacity network capability. Under specific conditions that allow for small buffers and short reach transmission, some optical packet switching may become advantageous. Highly heterogeneous traffic and network requirements, however, prefer opaque solutions that make best use of the agility and processing power of electronics. These opaque or 'digital' solutions do not scale well to high capacities. Future core networks will likely be composed of hybrid technologies spanning circuits to packets and flexibly utilizing electronic and optical technologies to best advantage.

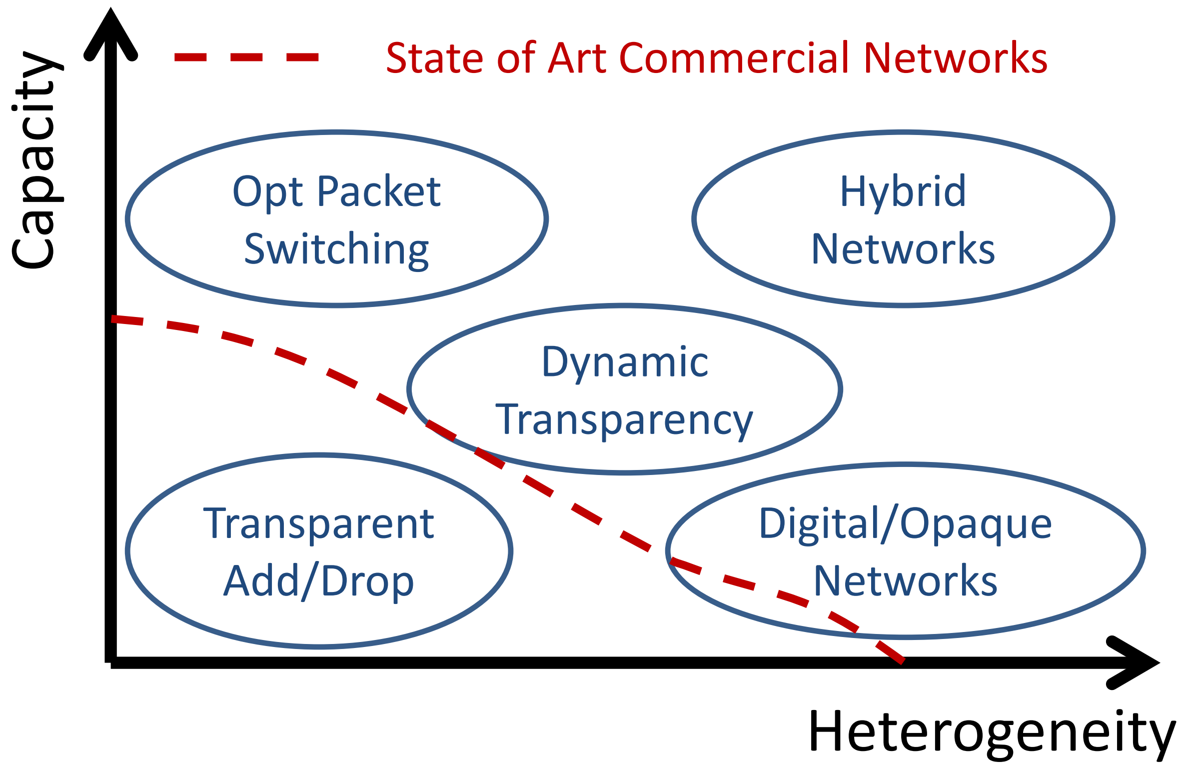

Figure 4.5 Efficient network paradigms for different capacity and service heterogeneity requirements

Hybrid networks support cross-layer end to end traffic management spanning circuit to packet capabilities using combinations of optical and electronic technologies.

Network performance and quality of service management will be especially challenging in future core networks. These functions typically require deep packet inspection or bit level processing, which does not scale well at the high capacities of core networks. Many of these functions are being pushed to the edge. Providing a level of security or quality of service in the physical layer may help to facilitate this trend, effectively clearing the toll booths from the super-highway of the Internet.

Data centers, cloud services, and content delivery are expected to be integral to future networks. The transfer of data between data centers, preparing and backing up content for example, will need to be considered along with the distribution to the end users. Designing networks specifically to support these data center applications may lead to improved efficiency and performance.

4.1.3 Access Networks

Future access networks from in-building and home to the core are expected to be shaped by the convergence of fixed and mobile networks. Identifying future proof solutions that can smoothly evolve through this convergence will facilitate progress in access. Areas for convergence include the use of optical access networks to support distributed or multiple input multiple output (MIMO) antenna systems and combined wireline and wireless using radio over fiber techniques.

Access networks will need to detect and adapt to multi-user environments and multiple access modalities. Today connections between devices in the same room often involve communication deep into the core of the network. Keeping this data, in particular high bandwidth traffic such as video or file transport, in the edge will help to reduce network traffic and improve performance and latency. Opportunistic networking in terms of spectrum use, modality (optical, rf, or satellite), and ad-hoc connections can lead to similar improvements. High frequency microwave or free space optical (through the room lights) are examples of new modalities that might have application for in-room communications and would take advantage of opportunistic functionality. Developing optical access systems that are more programmable or 'digital' in nature would help to facilitate this flexibility. Maintaining consistent levels of security across these modalities will be a challenge, but will be essential for adoption particularly in the case of shared access media.

4.1.4 Mobility, RF, and Satellite

Radio systems present a unique set of technology challenges and opportunities. The trend toward cognitive or adaptive capabilities will continue, with focus on both capacity and energy efficiency. Fixed spectrum allocation is fundamentally flawed. Figure 4 shows the current spectrum allocation in North America. Even though the spectrum is allocated it is almost unused at any given location. This is even more true as we go to higher frequencies. Furthermore, the applications that will use wireless connectivity are continually changing and future radio systems will exploit capabilities that we cannot anticipate today. Frequency is only one of the three dimensions to exploit, time and space give new opportunities. This flexibility, however, can lead to increased interference and may require new dynamic interference coordination techniques.

In addition to more efficiently using existing bands, using new spectra regions has potential to unlock new bandwidth. For example, high frequency millimeter radios operate in the 30-300 GHz range (1-10 mm wavelength) with large amounts of unused and unlicensed spectra (7 GHz at 60 GHz). Misconceptions about path loss and propagation at 60 GHz have hindered past use of these bands and the technology to process signals at 60 GHz is now fully integrated on mainstream CMOS.

New RF technologies can facilitate these trends. Active, multi-element, and large scale antennas can make better use of spatial diversity. New materials such as meta-materials can enhance antenna and RF amplifier performance. Antenna design can also enable beam forming, providing greater energy efficiency.

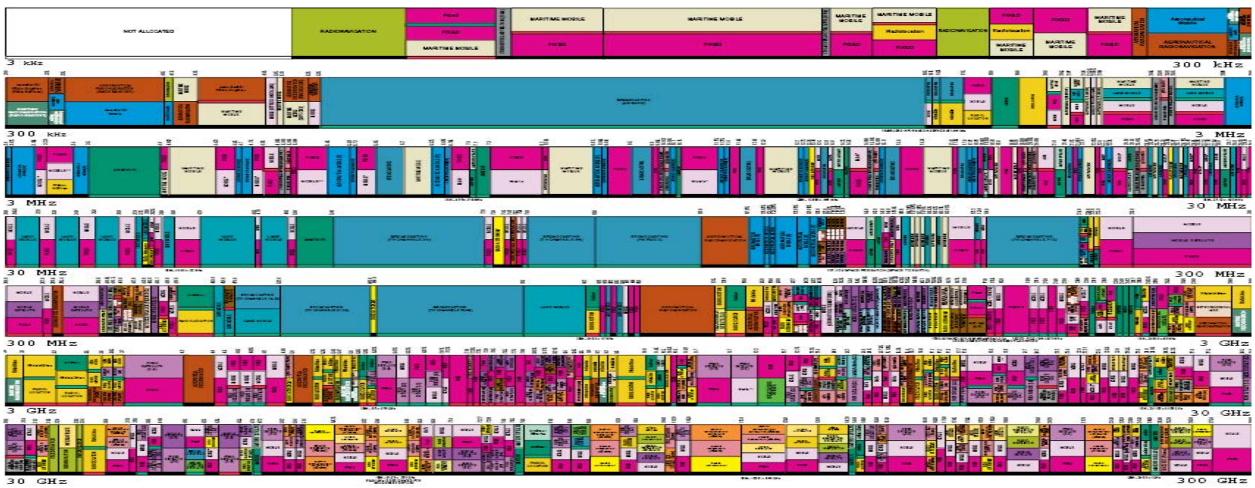

Figure 4.6 Spectrum allocation in North America[6]

4.1.5 Metrics

In a heterogeneous network environment, new end-to-end performance metrics are needed to capture the inter-working across network modalities and heterogeneity-enabling technologies. For example energy/bit is a common efficiency metric that can be extended from the device level to an end to end network. However, in a network scenario the energy must include both the data path as well as the control path and the bits need to be extended to the network Goodput or end-to-end application data. Other common metrics such as bit/second/Hertz, cost/bit/second, and subscribers/Hertz might be similarly extended. Technology specific metrics such as expressing quantities on a per kilometer basis for wireline or per kilometer squared for wireless will still be needed and balanced against the larger system or network metrics. This end to end perspective also brings focus to metrics such as fairness (among technologies and users), availability and coverage.

---

[6] US National Telecommunications and Information Administration

While care must always be taken in setting up the test cases to evaluate performance against a metric, this will be especially true of future complex, distributed communication systems. Examples include variability in traffic granularity and statistics, peer to peer vs data center to user, modality combinations, network hardware configurations, and subscriber distributions (e.g. geographic, content popularity, and QoS). A well-defined metric is only useful if it is consistently applied under conditions that are meaningful to the intended application. Understanding these conditions will be an important enabling tool for progress in heterogeneous networks.

4.1.6 Industry Partnership

Future networks will rely on large complex systems with substantial infrastructure requirements. Similar to the silicon electronics industry, emerging elements of the base hardware are beyond the resources of academic research. For example, coherent optical transceivers at 100+ Gb/s require increasingly specialized high performance devices. Recent research on high speed transmission is almost entirely driven by industrial labs. Close collaboration with industry is a necessity. However, the focus of this collaboration should not be to solve today's problems, but instead to understand where the future of the industry is heading, to identify the major impediments to that future, and to gain access to the technologies and techniques that will facilitate research into that future. Obtaining input from industry is not enough. Research labs today are international and industrial facilities will expand in regions where support is available. Similar to Europe and Asia, US industry needs to be incentivized to engage in long term research in collaboration with academia through cost-sharing and direct funding of industrial research.

4.2 Network layer

4.2.1 Virtualization, Community, Programmable

Virtualization and cloud computing are key for future heterogeneous networks. Virtualization involves the abstraction of communication and computing resources for networking. Cloud Computing involves making these abstracted network resources to facilitate computing applications, data storage, and network communications.

Cloud computing requires high quality services such as high bandwidth, low delay, massive storage, and highly distributed processing capacities. The glue for these network resources is virtualization technologies that provide flexible and on demand allocation of resources in highly abstract representations, relieving the end users of the cloud the need to physically manage these resources and to maintain irrelevant details. Virtualized resources can be represented as files which are easy to migrate, replicate, and store.

The cloud effectively forms a community, a network. Programmability provides the abstraction and ease of control over the network resources to form infrastructure, platforms, software, and application as services.

This vision for future computing raises many fundamental networking issues:

- How to create virtualization techniques for different layers: physical layer, mac layer, network layer, transport layer, and application layer?
- How to virtualize basic network element, like IO, router, network, switch, data content? What is the new addressing method for them?
- In a virtualized network, what information should be transparent? What are the security issues? Where data should be encrypted, at the lower or the higher layers or in storage?
- How to better organize the future network to utilize social network structure and interests of users?
- How to create a sustainable virtualized network that is energy efficient and reliable?
- How to migrate data in a virtualized network, for example among the cloud service providers?
- What is the fundamental difference between the physical connectivity and virtual connectivity and who should be responsible for the QoS managed physically or virtually?
- What are the new programming abstractions that are necessary to facilitate virtualization and cloud computing?

Other issues are also important:

- Is it feasible to build up a virtualized network over existing infrastructure? What are the conflicting policies and regulations?
- What are the business models for the new virtualized network?

4.2.2 Security and Transparency

Security remains an important issue in designing the network layer. User identification and privacy are fundamental to secure operation of the network layer.

On the other hand, transparency and network data collection are also important for efficient network operations. Here we want to jointly consider how these two antagonistic attributes of the network layer could be harmonized. A few of the issues are identified:

- Tension among anonymity/privacy, transparency/manageability/security
    - E.g., wireless jamming for specific packet types
- Relax the service model exposed to the higher layer by the network layer
- An anonymization service to help eliminate cover channels
- Degree of security support at the network layer to help enable security

4.2.3 Application Driven

The applications currently experience the network based on very coarse granularities – availability or unavailability of timely information passed through the communication layers consisting of the physical layers to the transport layers. As various new applications start using the Internet as the prime communication infrastructure, the range of requirements vary a lot – from hard real-time guarantees to interactive applications to massive data flows from large instruments.

One of the key capabilities that would make the application and network interaction more responsive is network awareness of application requirements and vice-versa, application awareness of network issues. It is clear that full feedback of all application requirements to the network or detailed statistics of the network to the application is not practical or desirable. Tuning the interaction to the right level of abstraction and frequency so that it enhances the quality of experience and quality of service delivery is one of the key research areas to investigate.

It is important to consider this feedback loop between the applications and the network to have a programmable "dial" so that each side can directly adjust awareness at each end. This programmable dial helps the application to be responsive to its user needs based on changing network conditions or "soft" failures. For example, a network

One suggested approach is for the network layer to expose a set of abstracted, service primitives to the applications. The applications can then combine these service capabilities to tune the network services to meet their desired needs. This lego-block approach allows the applications flexibility to demand custom services from the network. The main research area is techniques create an abstract model of a network. The abstraction should take into account both static and dynamic states. The challenge then is how to create, manage and offer a service over an abstracted network topology that consists of various heterogeneous elements of the network.

4.2.4 Scalable Mobility, Ubiquity

Future networks are evidently characterized by the mobility of nodes and sub-networks and the need for ubiquitous connectivity. This sub-group focused on the concerns revolving around scalable support for mobility and ubiquity. These concerns are driven both by the demands of applications for seamless network connectivity but also by the RF spectrum limitations that are pushing the connectivity providers to deploy smaller range base-stations (in order to increase the network capacity) therefore resulting in high network topology dynamics. The following are some of the key issues to be addressed in order to enable heterogeneous networks with seamless access that links-agnostic for the applications but still scalable and resource-aware from the system perspective.

Ubiquitous connectivity: Mobile devices should be able to maintain seamless connectivity across heterogeneous networks relying on various network interfaces including multi-radio, optical, and wired. These network interfaces cooperate (instead of co-existing as is the case today) efficiently reconfiguring to adapt to the links characteristics, network dynamic, and application needs.

Applications connectivity: Applications should be provided end-to-end connectivity without having to be concerned with the specific network interface(s) or address(es) used by the mobile node as these devices move across networks. Such seamless connectivity requires some level of identification to maintain consistency but at the same time raises concerns about security such as the capability of malicious entities to locate and track users. Seamless connectivity from/to heterogeneous network interfaces with multi-source-interface, multi-destination-interface, and multi-paths communication implies that new routing paradigms and algorithms have to be developed.

Infrastructure mobility: In future networks, part of the infrastructure might itself be mobile such as is the case in the disaster recovery, military networks, but also cooperative networks of mobile devices.

End-to-end routing over heterogeneous network to meet application needs: While applications are concerned with end-to-end connectivity independently of the intermediary heterogeneous networks, a mobile device will have identities (e.g., addresses) and will utilize resources across multiple domains. This calls for new concepts of domains, and the need for new mechanisms for information exchange within and between domains. New intra and inter domain routing mechanisms that are application specific with multi-path capabilities will be key to the efficiency, scalability, and robustness of future heterogeneous networks.

Up/down scalability: One of the key obstacles to many innovative networking paradigms has always been scalability. For future networks to be successfully deployed, they have to scale both to a large number of players with high bandwidth demands but also down to devices with limited resources and capabilities. The scalability should encompass and account for the heterogeneity in terms of the number of users, domains, speed, delay, and bandwidth.

Secure and robust geo-location: Many applications require access to geo-location information independently from the device peripherals and existing infrastructure. Future networks should provide for a geo-location service but should overcome the security problems that result from existing technology such as location spoofing and users tracking.

Critical control-plane: Most network designs achieve resource efficiency through complexity of the control-plane. This makes the control-plane the most critical part in terms of resiliency to adversarial attacks. Mobility and ubiquitous access provides adversaries with many opportunities for attacks that are difficult to detect, locate, and track. Designing a control plane that is robust to faults, attackers, and insider adversaries is a key issue that needs to be addressed in future heterogeneous networks.

4.2.5 End-to-End Quality of Service: market driven with service guarantees

Today service level agreements (SLAs) for communication services include standard performance metrics such as average and peak bandwidth, minimum availability (e.g., 99.999% availability), maximum bit error rates, and maximum packet drop rates. These metrics are sufficient for a wide variety of applications, but for applications in the future they may be too coarse and inflexible. For example, in the case of say 99.9% availability over a period of a month (31 days), it is only required that the communication service has total accumulated outage of at most 0.01% of the time or 45 minutes in a month. Thus, a single outage can have length over 40 minutes. For some applications, availability of 99.9% may be acceptable as long as no outage can be longer than say 5 minutes. Hence, it is important for SLAs to be formatted so that it can be more flexible to accommodate new and finer grained services and constraints. The goal of e2e reliability and quality of service in general becomes even more challenging when underlying networking substrates within each domain is more diverse and heterogeneous across multiple domains. Inter-domain interfaces allowing swift and effective expression of performance requirements are crucial to glue together services and performance guarantees offered by such diverse set of domains.

Economic Granularity: Also in networks today, the amount of time to set up a communication connection through a service provider domain can take weeks and even months. Future heterogeneous networks with highly scarce (wireless) resources will need a much faster and granular inter-domain economics so that it is possible to express availability of opportunities to other domain owners. To facilitate such speedier connection set up across multiple domain owners, an automated process should be developed that matches clients and their service needs to network capabilities. Most of the time, even though the underlying network resources may be vast, the end-to-end paths composed by the inter-domain protocols cannot leverage such vast underlying resources effectively due to uninformed policies of domain owners and too rigid contracting mechanisms or coarse economic tools unable to swiftly and effectively establish reliable end-to-end paths.

Technologies: Some of the technologies that may have impact include but not limited to OpenFlow and Cloud Computing technologies. In the case of Cloud Computing, new interconnection architectures for data centers and optical technology on a chip have the potential for significant impact. A specially difficult and critical research problem is the reduction of control complexity between packet and optical networks.

4.2.6 Performance and Energy Efficiency

Bandwidth requirement, per application as well as aggregated, is expected to continue to grow exponentially. As transmission bandwidth continues to grow rapidly with the advent of gigabit Ethernet to terabit Ethernet, the demand on switching a large number of such high capacity links exceeds the capability of current designs. This creates new requirements.

- Terabit switching on a chip, using massive on chip IO connectivity. Of equal importance to high throughput is low latency of nanoseconds.
- Switching at various layers, e.g. physical and IO such as PCIe, MAC and LAN such as Ethernet, routers and Internet, and application routing.
- Network addressing may have to be re-examined. Is IP address all there is for the purpose of routing? As the computer is being de-constructed and virtualized on the network, are other forms of addresses, e.g. byte-orient memory addresses, block-oriented storage addresses, various kinds of LAN addresses becoming more important in the network layer? How can we

perform address resolutions and reversed address resolution more flexibility and efficiently at the network layer?

- Do we need new routing algorithms beyond Bellman and Dijkstra type of shortest path algorithms? Can network virtualization provide new concepts for network routing?
- Reliability issues need to be incorporated into network routing and provisioning
    - Requirement for five 9s reliability
    - Redundant network elements/connectivity
    - Most traffic doesn't really need 5 9s
- Energy efficiency metric may be crucial in transmission, routing, processing and storage of information. We may need to consider various metrics such as
    - Energy used/bit
    - Total energy consumed
- Smart apps optimize use of network to satisfy user experience, i.e. more application routing and security management
    - Take advantage of network APIs
- Smart networks optimize transmission of app traffic to satisfy user experience
    - Take advantage of app APIs and user-behavior/interests
- User controls how data is sent over networks - opportunistic
    - Multi-homing; Dynamic, Rapid address look-up
    - Source routing
    - Make routers dumber, push smarts to the edge

There is a tradeoff between reliability and energy efficiency. Reliability is fundamentally built on redundancy, and the higher the level of redundancy, the greater the amount of energy used per actual unit of work done. However, there is an opportunity in this. Historically most networks have provided the same level of reliability (and hence used the same level of redundancy) for all traffic. The reliability demands of the most stringent application end up driving the design and hence the energy efficiency of the network. Most traffic, however, has less stringent demands. If the level of redundancy provided for each component of the overall traffic flow were adjusted to match the actual reliability requirements of that traffic flow, substantial energy savings could be achieved. Several architectural innovations must work together to achieve this vision.

- Smart apps optimize use of network to satisfy user experience. They do this by taking advantage of new network APIs that expose appropriate and useful information.
- Smart Networks optimize transmission of app traffic to satisfy user experience. They take advantage of new network APIs that provide information about user behavior and requirements.
- End applications opportunistically control how data is sent over the network. Supporting this requires features including multi-homed nodes, rapid changes in the DNS database (or its future equivalent), and some level of source routing, in addition to the new network APIs that provide the information needed to make the right decision based on current conditions.

4.2.7 Conclusions

This section has presented a wish list for Future Heterogeneous Networks. Due to lack of time and the complexity of the matter, few solutions were explored during the workshop. However, a number of technologies deemed important for Future Heterogeneous Networks at the network layer were identified:

- IO, LAN, and Internet virtualization
- Terabit switch on a chip
- Application and security aware routing
- New methods for mobile ad-hoc network that are scalable and robust beyond MANET.

4.3 Transport and application layers

4.3.1 Transport layer issues

A key challenge is a transport layer that supports composition of heterogeneous networks more effectively than current transport layer designs. The following issues need to be considered:

- What is the end-to-end service available on a given composition of networks
- An application needs to negotiate with the network service on capabilities
- There is a need for an interface for application to query the network for info' as well as the ability to negotiate with the network for required capabilities.

- Also need an interface for application to tell the network about it's traffic profile. While further study may be needed on whether this 'bi-directional negotiation' across the application and network layer interfaces are desired, there appears no reason to exclude them at this time.
- What would then be a minimal, but sufficient, interface look like?
- We see the need to exploit multi-path opportunities, especially with heterogeneous capabilities. A transport capable of exploiting the diversity across paths is desirable.
- As energy concerns grow, there is a desire for an option for a "greenest path first"
- Role of DNS to help with these issues needs to be explored?
- Security and privacy issue make this much harder
- End-to-end could mean layer 4 to layer 4 communication capability.

The current Transport protocol, TCP is not working well for high loss wireless environments, nor at speeds of 10Gbps or higher. It appears highly likely that future heterogeneous networks will present applications with a choice of multiple transport layer protocols/systems. One area for research is how best to determine which protocol or system to use in a given instance in a system with heterogeneous transport layers. A key component of this is the congestion control protocol.

- Can protocol/ Congestion Control (CC) algorithm selection be done automatically based on loss rates, packet arrival rates, etc?
- Can it be changed dynamically after the session is set up?
- What information needs to be passed to the application to help it make better decisions?
- How to prevent applications from always requesting the most aggressive algorithms? This comes from a need to be 'optimal' across all the sessions that share the network resources.
- What protocol/ CC algorithm should be used on dedicated circuits?
- What support is needed from the NIC to allow for better CC algorithms (e.g.: precision packet arrival timers).
- What is the role of multi-path TCP? How well can it exploit path diversity?
- What about increasing redundancy to reduce need for retransmissions in high loss environments? Wireless is becoming ubiquitous, so tolerance for lossy links needs to be a core feature of the transport layer, rather than an afterthought. This is viewed as a key enhancement needed when considering heterogeneous wireless networks.

4.3.2 Application view of the network

There is a clear need for more feedback from the network that supports end-to-end analysis of performance and reliability issues.

- Need better ways to monitor the end-to-end data path. This includes application and disks, as often the problem is not the network
- Need to understand which layer is losing packets/congested and why.
- Monitoring needs to work across administrative domains, which creates authentication and authorization challenges. There can also be privacy challenges, particularly in monitoring of vehicular and other mobile networks where node location information may disclose the location and motion of individuals.

The application's view of the network is established by the session layer, which today primarily presents the abstraction of a Socket. A Socket is a 1-to-1 or 1-to-many data pipe established by specifying the addresses of the nodes that are its endpoints. A possible radical change in the application view of the network is a session layer that generalizes the services provided by today's Content Distribution Networks (e.g. Akamai). This is called data or content centric networking. Here are some key characteristics of content centric networks:

- Change from point A to point B model (IP model) to a "just get the data" model
- The traditional End-to-end model is not efficient for data dissemination. (need to use best protocol for a given segment of the path – this seems misplaced)
- How to "name" the piece of data to deliver, making it possible to route the data based on the name.
- Role of multi-source TCP? (reword: content centric networks could use an enhanced transport protocol that supports true multi-point to multi-point communication (as evidenced by a multiple sender, multiple receiver multicast transport protocol).
- Is there a role for storage/caching in a content-centric network. This has the possibility of enhancing efficiency of dissemination (throughput, latency and more judicious overall utilization of network resources).
- Privacy issues for CCNs: being addressed by some current proposals
- Fairness issues may be more complicated (What does 'fairness' mean?)

The key architectural question is whether content centricity is something viewed at the application interface/session layer (analogous to a socket) or whether it is a fundamental construct in the core of the network (which we view as the combination of the current transport and network layers).

5 Cross-layer research issues

This section reports the results of the focus groups on cross-layer research issues that met on the first day of FutureHetNets 2011. Two groups met in parallel. One investigated issues in the high-speed fiber core and the other investigated issues in wireless and satellite communications.

5.1 *Fiber core*

There are several architectural options can enable significant increases in the communications capabilities and performance delivered to the user. Realizing these performance gains particularly within the emerging highly heterogeneous networking environment will require transformative changes to the network resources can be dynamically controlled and allocated. At the physical fiber layer, future research is driving increased bandwidth density and transmission capacities with emerging technologies, modulation schemes, and coherent detection. To fully exploit these vast increases in transmission capacity at the fiber core, future network architecture must embrace high flexibility in providing for and adjusting the granularity of the data. In the ideal case this flexibility and control of the data at fine granularity creates the notion of a continuous optical physical layer: where any size data of any packet length, data rate, or number of wavelengths, could be dynamically allocated, accessed, employed, released and redeployed by the higher network layers. Toward realizing this architectural goal, care must be taken to ensure that increased flexibility does not lead to network instability.

An example of emerging physical layer technology which supports the notion of the architectural options with dynamically flexible fine granularity data control is the super-channel orthogonal frequency division multiplexing (OFDM). This recently proposed technology can support bandwidths at high transmission rates of terabits per second or more with greater flexibility in sizing the bandwidth of communication links by allowing them to have a flexible number of subcarriers. Therefore a link can be "right sized" to the amount of traffic that it carries, more precisely matching requirements to resources.

These technological breakthroughs are enabling a rethinking of the current wavelength structure of a standardized wavelength grid, leading to a possible gridless or a finer grained grid. Building upon these approaches in future architectural designs will reduce the need for guard bands and lead to more efficient usage of the network resources.

The development and implementation of components under these future architectural options need to embrace interoperability and flexibility. For example, for component insertion the need for frequency flexibility and possibly programmability is a critical consideration. Important challenges to these classes of flexibility driven architectures will be the overall physical layer performance and quality of transmission.

Though the optical layer can provide extremely high communication rates at the physical layer, in many cases it is underutilized at higher layers in the network. Architectural options should be explored to address this issue. One option is to allow larger packets or even unlimited packet size to improve communication performance for a certain applications. Another option is to avoid bottlenecks in the network by using end-to-end L1/L2 paths and L3 tunneling. Consideration must also be given to the design of heterogeneous networks that allow paths to be constructed from multiple layers (e.g., L1/L2,L3). Extremely high capacity links will likely carry a large aggregation of traffic streams, so efficient methods of flow aggregation are needed to optimize the use of the capacity. Also, networks today carry a lot of duplicate information, so a critical approach to improve the bandwidth usability is to reduce or eliminate this redundancy in data.

Another architectural option is to improve the control and management of networks at L1, L2, and L3 with a uniform control plane architecture. Today each layer has its own set of protocols including IP, Ethernet, MPLS, OTN, and GMPLS, to note just a few, and the overall complexity can be a barrier to fully utilizing the network bandwidth. To overcome this barrier, a scalable uniform control plane for dynamic L1/L2 switching should be developed. The current protocols at these layers have similar functions, and this observation leads to the consideration of, as well as a justification for, a uniform control plane architecture. Such architectural concepts could also be extended to other layers.

Another architectural option is to have two (or more) co-existing network infrastructures on the same physical infrastructure using different fibers or lightpaths. For example, one infrastructure would be the current Internet, while another infrastructure could be dedicated to support high speed switching. This approach enhances the potential for evolving networks because it enables entirely new network infrastructures without having to change existing network infrastructures.

The current Internet has been extremely successful in providing connectivity to a wide variety of applications and over all types of physical communication media. However, its architecture of one-size-fits-all can lead to inefficiencies. The next step in improving network performance is to design more heterogeneity into the network, allowing it to more efficiently provide service to different applications and to support "purpose-built" networks.

5.1.1 Cross layer interaction and management

In general, cross layer interaction can potentially improve network operation and performance. A critical challenge is in identifying and defining the interactions between the layers. In addition, the amount of information must not overwhelm the network causing other problems. The amount of information must be minimal. And all designs must have inherent potential for future evolution.

Exposing lower layer information and functionality (especially at the physical layer) to higher layer processes and functions could improve network operation and performance. This is certainly true for new technologies at higher data rates (100G and higher), where more intelligence is required at the physical layer. Lower layer information can be available to even network applications allowing applications to utilize the network by being network aware.

Another approach is "application aware networks" or "application driven networking" where the network can optimize its service depending on the application. This requires careful design to ensure solutions do not cause instabilities in the network.

Research on crosslayer interactions should explore new abstractions, APIs and may eventually lead to standardization of interfaces. It may even lead to designing protocols that can cross layers. But it should preserve simplicity and not diminish the ability to separate subsystems to evolve separately and enable innovation at individual layers.

To improve network service, network architectures are needed that provide better coordination among requirements, services, and resources. For example, coordination can be dependent on time scales, where over short time scales, distributed coordination is used to react quickly to events but over longer time scales, central coordination is used to optimize the network.

Architectural design should anticipate resources that are owned by federated domains. Currently almost all large scale infrastructure is owned and managed by single entities. In the future, multiple domains will have to coordinate requirement fulfillment and resource requests.

5.1.2 Metrics

New metrics should be considered and new instrumentation will be needed. When addressing individual network issues, care is needed to avoid being too narrow in scope since improving one aspect of a network can often lead to unintended and unwanted consequences in other areas.

Metrics are typically from the user or network perspective. A generally accepted user-oriented metric is end-to-end throughput or "goodput". But this may not completely capture the utility of a network since some services may have a great impact for users but with lower communication bandwidth. New metrics are needed to capture network utility. For example, a measure of usability of a network is the amount of information that is being shared. A related issue is what types of metrics can be integrated with user control. For example, additional metrics can lead to more knobs and options so that the network becomes more usable.

An accepted network-oriented metric are actual capital and operational expenditures (CAPEX and OPEX), and energy and power costs have become another key considerations. Metrics are also needed to measure the complexity to configure them and their devices since networks are difficult and complicated to manage.

Metrics are needed to measure network efficiency because improving user services are often at the expense of increasing network costs. An example is the ratio of total amount of network service provided to the users divided by the total network cost.

Finally, heterogeneous networks will undoubtedly present new services, systems, and architectures. New measurement tools are required to address the new capabilities of these networks.

5.2 Wireless and Satellite Communications

The core issue in wireless and SATCOM cross-layer optimization is the variability of link and path properties. In this environment, resource management decisions such as transmission scheduling and bandwidth allocation need to be handled in a fundamentally different way from the mechanisms in the current Internet that evolved in more stable wireline networks.

A cross-layer design that more effectively manages the variability of wireless networks will affect all levels of the system (physical layer, resource management, transport and applications). Information and control need to flow both "up" and "down" the protocol stack. Application requirements should drive choice of physical topology, scheduling of data transfers, and decisions about which packets to prioritize, buffer, or drop when capacity is limited. Information from the physical layer should flow into the network layer to optimize packet forwarding decisions.

Information also needs to flow horizontally, from link to link across a network, and from network to network across a heterogeneous Internetworking system. Supporting horizontal information flow more effectively than in current systems requires innovation in the control plane itself.

5.2.1 Different modalities

When considering different wireless communications modalities, the primary challenges are different.

- Vehicular: The system must identify and use “good” links. The set of good links is constantly changing as vehicles move with respect to each other and encounter shadowing and fading due to environmental obstacles. Key areas for research are appropriate metrics, ways to share information, and how to effectively exploit a combination of heterogeneous link types to improve the user experience.
- Satellite: The system must optimize throughput, latency and sharing given high capacity and power variability in the physical satellite links. Key areas for research are to better understand the cause of losses and predict them, and to enable random channel access without prior reservation for more efficient short transfers.
- Cellular: The system must optimize handoff and transmission scheduling while managing dynamically varying load and reallocating radio resources between neighboring cells. Much of the information required to do this effectively is known only to client devices (application requirements, channel measurements) and ways are needed to share this information with the network core and across networks at minimal cost and while respecting administrative and privacy concerns.

5.2.2 Service predictability

A network must provide a predictable level of service to enable application designers to assure user needs are met. A key challenge created by wireless and SATCOM systems is how to mask or exploit the high variability of wireless links to support a negotiated or implicit service level agreement with each application. Service level is not the same as a promised metric value (for example, rate or delay). Instead, the service level describes what fraction of offered transactions (message sends, video packets, etc.) are completed at the promised metric value. A low service level reduces network implementation challenges, because the variability of the underlying wireless links is simply exposed to applications. However, a low service level increases application complexity, and may prevent some applications from being implemented.

Achieving a high service level simultaneously with efficient use of the underlying communications resources is a major cross-layer challenge for wireless networks. The simple way to achieve a high service level is to reduce the promised metric value (for example, communication rate) to approximately the worst-case instantaneous performance of the underlying physical link or network. This hides variability from the applications at the cost of wasting capacity. In highly variable links, such as satellite links through the turbulent atmosphere, the amount of capacity wasted can be significant.

One promising approach to achieve high service level combined with efficient resource use is to combine diversity and collaboration.

- Diverse frequency, technology, relays for independence, congestion mgmt
- Appears to require multihomed endpoints on independent networks
- Use collaborative communications to reduce impact of PHY issues
- Diversity receive to overcome shadowing and multipath
- Distributed MIMO, collaborative decoding, beamforming, etc

A cross-layer architecture is needed to implement these techniques.

- Topology discovery by MAC layer, used by physical layer collaboration
- Applications negotiate explicit service level with network or select from menu of services that include implicit service level.
- Export information about dynamically changing network/link characteristics to support route selection and data transfer decisions across heterogeneous networks

6 Heterogeneous network research issues

This section reports the results of the focus groups on heterogeneous network research issues that met on the second day of FutureHetNets 2011. Two groups met in parallel to investigate the same issues; their results are combined in this section.

6.1 Challenges

There are a number of challenges that currently impair adoption, utility, or performance of heterogeneous networks. Each of these challenges corresponds to a topic area to be considered when developing a broad architectural solution.

Resource integration and seamlessness: Heterogeneous networks are currently less efficient and harder to utilize than homogenous networks because of a lack of resource integration, and because users/applications must work around problems associated with the boundaries and the differences between the modalities. Network and resource management is normally organized per modality, with minimal support provided for enterprise-level cross-modality resource management. This results in misallocation of resources and slow response to changing load or component failures.

User/Application Control and Flexibility: Heterogeneous networks are currently more rigid and provide a lower level of user/application control than homogenous networks. Users experience the network as inflexible because poor resource management and lack of seamlessness leads to static pre-configuration (of capacity allocations, prioritizations, routes, and so on) that is complex and difficult to change quickly. The low level of integration between the separate modalities significantly constrains the choices available to application designers for Quality of Service and bandwidth guarantees.

Robustness: Heterogeneous networks are currently fragile and fail more often than homogenous networks, due to the same poor resource management and static pre-configuration issues just described.

Abstraction of network details: There are currently significant challenges to overcome to deploy high performance applications on a heterogeneous network. A primary challenge is to achieve high performance without requiring the application to be customized for the networks and technologies in use. Users and applications who don't want to deal with the heterogeneity of the network are often forced to cope with it.

Greenness and Battery Life: The poor resource management and low flexibility of heterogeneous networks causes them to currently consume more power than is necessary. It is difficult to preserve the battery life of resource-constrained components because of the distributed network control and inefficient design. There is no way to optimize the use of end-to-end network capacity for power efficiency.

Security, privacy and trustworthiness: Heterogeneous networks are currently less secure than homogenous networks. When there are multiple heterogeneous services or modalities, the increased system complexity makes it harder to differentiate normal behavior from attacks. The greater complexity makes it harder to control the network to assure security policies are followed and to restore service after an attack. In a large or heterogeneous network, the "attack surface" is larger. This is the set of component types (e.g. software loads, waveform designs) an attacker could potentially target and which thus must be secured. Finally, denial of service attacks can exploit the diverse communications paths of heterogeneous networks to route around detection and defense mechanisms.

While privacy issues are not fundamentally different in heterogeneous networks from current homogenous networks, the inherent heterogeneity can create subtle channels for information leakage that are not easily predicted.

Heterogeneous networks should be more trustworthy than homogenous networks, but without appropriate architectural design, they can end up being only as trustworthy as the least robust subnetwork.

6.2 Fundamental solutions

We identified a set of architectural options that provide fundamental solutions valuable in addressing many of the problems identified in the problems section.

Integrated control across modalities: New control mechanisms should be developed that provide integrated resource management and control across the modalities making up the heterogeneous network.

- An architecture is needed that supports enterprise-level control of resource allocation and other management decisions made in each subnetwork.
- A key component of that architecture is cross-layer and cross-modality information sharing – necessary information, but not too much as to bog down the system.

- Congestion control could benefit from distributed computation and storage as part of the network infrastructure.
- Routing can interact with integrated control to achieve path diversity based simultaneously on estimation of path characteristics and network-level needs for load balancing and defragmentation.

Integrated control faces challenging questions.

- There is a control loop for adaptive applications and a control loop for network performance, these are currently decoupled. How should these loops be coupled? What are the relevant time scales? How to insure stability of the QoS?
- Should there be a human in the loop? How to present the network information to the user? How to gather the network information that you need? How to trust the information received from the probe?
- What are the trade-offs between hiding vs exposing network variability and/or state? When to hide and when to expose network variability and/or state based on magnitude of changes and time scale.
- Accountability, beyond SLA; trade off between transparent access and internal details.

New forms of cross-administrative-domain information exchange: Cross-domain is distinct from cross-modality information exchange. There are cases when two communications modalities (technologies) are controlled by a single administrative domain, and vice versa, cases where two administrative domains connect networks that exploit the same communications modality. Cross-domain information exchange in the current Internet is provided by BGP. New mechanisms are needed to support tighter integration in large-scale heterogeneous networks. One of the key design elements is making just the right amount of information available across domains.

Make gateways an integral part of the network data & control plane: Gateways are the devices—often containing routers—that connect two or more communications modalities. Today gateways are engineered as an afterthought when the modalities need to be connected. Normally they are built ad-hoc and independently for each pair or group of modalities. Since they are connected to multiple networks, they do not belong fully to any one network, and their internal resources and behavior is not fully controlled by the control plane of any of the networks they are connected to. From the data plane perspective, the special characteristics of a gateway are not considered; it appears either as a normal router or as a normal link.

- The network should be rearchitected with the gateways as first-class entities, designed into and exploited by the data and control planes.
- In the data plane, it is appropriate in some circumstances for the transport layer to run from gateway to gateway rather than end-to-end. A gateway that performs transport layer functions can be called a "transformer" (by analogy to transformers in power distribution networks) that matches heterogeneous segments in an end-to-end path. The goal of an architecture based on transformers is to prevent the end-to-end path from seeing a combination of weakest properties of all segments. For examples, the transformers may make space / time tradeoffs between segments, and may duplicate and diversity route packets over unreliable segments.
- In the control plane, gateways can serve a critical monitoring and enforcement mechanism for enterprise-level resource management and control of the collection of heterogeneous modalities.

Add storage into the network: This was proposed as a general architectural change but few specific applications were mentioned. One discussed by the group was to cache copies of packets at the far end of a high-delay or expensive modality (e.g. SATCOM) enabling retransmission from there of packets lost downstream without resending across the expensive modality.

Network-level identity for services and data: Today the only network-visible names are node addresses and low-level protocol names. A number of valuable mechanisms (better prioritization, automatic dissemination, others) could be built if high-level services and data objects are given network-visible names.

Multiple user level services: One driver for architectural innovation is that there are multiple design centers: low price + mostly good enough versus high price + guaranteed performance. User wants the best service possible at an agreed-upon price point. The application should have the choice of which service to use. Techniques to support this more effectively than in current networks include the following.

- Allocating networking resources, transport functions, etc. to satisfy the multiple needs (explore virtualization as potential solution)
- Where is the best place to make those distinctions?
- Identify future modalities that make up the heterogeneous network
- Physical reach of segments (match specialized requirements to reach)

Admission control / scheduling for traffic that needs special handling:

- Advanced reservation and negotiation
- Tradeoffs between price and time of scheduling
- Pre-emption

6.3 Security and robustness

The discussion of security and robustness identified a number of approaches that might contribute to a coherent heterogeneous network architecture.

- Cross-layer link diversity and cooperative anti-jam strategies exploiting game theory
- Positive identification of endpoints via a mixed technical/regulatory solution
- Combine cryptographic and incentive-based mechanism design to preserve privacy in the context of cross-layer information exchanges
- Data-centric security (rather than link-level security)
- Better visualization tools for network forensics
- Composable trust frameworks (trusted systems if built from trusted components)
- Robust information transfer via control plane/gateway infrastructure
- Common authentication / accounting mechanisms over heterogeneous networks (e.g. WiFi and cellular)

6.4 Network Management

Large heterogeneous networks cannot be designed and managed as single homogeneous entities. The current BGP does not work and needs to be replaced. To solve this problem, several proposals were discussed:

- Think of the network as an aggregation of organizations, with heterogeneous structures and needs. Larger network management can be organized around standardized rules rather than strict control. A framework of rules can be established so all sub-networks cooperate properly. Some of the rules relate to information sharing to achieve QoS, security, etc. The network overhead could be kept low this way. New rules could be imposed as market needs change, including times of emergency. The user's experience needs to remain seamless.

- A hierarchical management structure could be used, with levels of control that span either geographical or domain/timescale based. The network becomes locally managed yet globally aware. The top plane analogous to NATO. Could have centralized cost/utility model.
- Content-aware network management: the management structure could be data centric. Data could be labeled, including priorities. The names can automatically become part of security.
- Management could be hierarchical, peer-to-peer, and federated.

The following were ideas that resonated with panelists but did not present holistic solutions:

- Send control signals on a separate channel. This allows the channel to become universal, with the required information for control and management available across modalities (technologies) and across domains (administrations).
- Make network resources "sliceable" at all levels. Investment is needed to increase the granularity and agility of the resources. Physical layer needs to slice by capacity and temporally. Each layer can be made more flexible so that it interfaces easier with other layers. Multiple-layer implementation otherwise difficult.
- Application can define the network parameters so that bandwidth is reserved and QoS is guaranteed.

Additional issues that were brought up that could facilitate management structures:

- Innovative metrics for heterogeneous networks
- Better visualization tools for network management
- Better knowledge management of information about networks
- Benefits of satellite communications should be integrated and exploited within management structure.
- Effective decision support.

6.5 Quality of Service

These are architectural options to improve the assurance that a promised QoS will be met, and to enable providing more flexible QoS options to developers/applications/users. More specifically, these architectural techniques are oriented to enable:

- end-to-end situational awareness to determine whether SLAs can be met

- network controls topology to meet application needs
  - (e.g. assign wavelength)
- application-appropriate caching, retransmission, replication, delay tolerance, resource management of network storage, provided as an integral network service
- opportunistic exploitation of diverse networks by smart applications

Redefine/extend notion of QoS to be at application level: Today QoS is expressed in low-level network metrics like latency and jitter. It would be much more useful and flexible to enable applications to describe their desired QoS at a much higher abstraction level. One example is a computing cluster that needs some guaranteed level of connectivity among its nodes.

Use dynamic SLAs: An SLA is a Service Level Agreement defining a certain quality of service. Normally they are static. As an architectural option, use dynamic SLA's where the quality of service offered is conditional on external factors. For example, bandwidth will be X MB/sec when network is in normal operating mode, Y MB/sec when network is declared in emergency mode.

Architectural notion of independent islands: We coined the term "island" to describe a region of the network that is relatively homogenous internally and hence capable of tight management and strong QoS guarantees. Architecturally the network should be considered as a collection of islands. Each island exports its service capability – expressed as a dynamic SLA – rather than low-level info like BGP describes such as link up/down. Each island has the option of doing its own L4 transport on data flows if necessary to meet its SLA guarantees.

Make gateways between islands a 1st class entity in network: The boundaries between the islands are gateways. The gateways are connected by an enterprise-level control plane that manages their resources – including storage – to guarantee end-to-end dynamic SLAs.

New interfaces between applications and control plane: The goal is application-aware networking and network-aware applications.

Virtualization of communication resources to support multiple networks making different resource use/QoS tradeoffs: Rather than attempting to support all application needs in a single network (single routing layer, transport layer, naming mechanism, etc.), this architectural approach proposes sharing the available physical resources (channels, transmit power, fibers, etc.) among multiple networks. Each of the sharing networks can then be specialized and optimized to support a class of applications whose needs cannot be met using the techniques selected for other networks. Virtualization at the appropriate level enables overall resource management and isolation between the separate networks. One important area for research is automatic conflict resolution between different networks that compete for the same (physical) resources.

6.6 Where to invest in testbeds

One sorely needed experimental testbed is a GENI like facility for lower layer / cross layer work. This is especially true for optical physical layer innovation, which is so costly to implement.

A large virtual distributed testbed could be designed, with GENI as one component.

Heterogeneous network testbeds require built-in heterogeneity. One way to achieve this heterogeneity is by integrating existing testbeds. A superstructure could be designed to accomplish this, including optical, dynamic spectrum access, satcom, and other wireless networks. These wireless networks would form an important part, yet are currently difficult to develop because of spectrum licensing problems. A spectrum license would have to be obtained to permit open experimentation

A broader impact might be made if testbed opportunities were combined with a simultaneous goal of bringing broadband to underserved (rural) areas. Outside-the-box innovation might result from setting such a goal.

6.7 Role of Commercial R&D

The panel considered the role of commercial organizations and what portion of the advancement in heterogeneous networking might be more efficiently undertaken by these entities. Commercial organizations were thought to be investing heavily in

- Single-domain security: the protection of their own resources is paramount , with little concern to what happens to the data outside the boundaries of the administrative domain.
- Cloud, content-based networking: this is seen as a great potential source of new revenue.
- Cross-modality integration: larger companies that own/manage networks composed of multiple modalities (optical and wireless, access and backbone, etc.) continue to research more cost and bandwidth efficient interfaces between them.

On the other hand, there are specific areas that single commercial entities are unlikely to support:

- Inter-domain information sharing: each administrative domain is protective of its information, and hesitant to share with other domains data that would be necessary to facilitate end-to-end security, QoS, and access. Some data is seem as proprietary, and other (such as current network state) can be seem as exposing weaknesses in their infrastructure.
- Large optical testbeds: these have not been developed, largely because of cost.
- Fundamental architectural changes and disruptive technologies: some panelists felt that commercial organizations could not stray too far from their existing networking vision.
- Infrastructureless wireless: little commercial investment is being made in multi-hop ad-hoc networks for mobile users.

6.8 High-payoff application areas

Some application areas that require inherently heterogeneous networks, provide some structure, yet yield a rich platform of research problems, were identified. These are also seen as having a potentially broad societal impact:

- Vehicular networks
- Public safety
- Education
- Health care – personal area networks

Of these four, public safety was investigated in more detail. Heterogeneity plays a key role in emergency networks for public safety by improving robustness through diversity. There are three primary use cases.

Urban: When there is an urban disaster such as 9/11 in New York, networks get overloaded. A heterogeneous network provides critical capability for emergency responders to work around an overload in one of the networks they depend on.

Rural and remote: Disasters of interest include earthquakes, tsunamis etc., which can disable the only installed network in a region. A backup network is essential to maintain connectivity. Requirements for a rural/remote backup network is that it be instantly deployable and self organizing (for example, by using satellite backhaul). It must support incremental deployment by layers.

Critical infrastructure backup: Current critical infrastructures (power grid, cellular network) each depend on an internal data communications network. Making the infrastructure robust against such attacks depends on having a backup network. It should be possible to share a common backup backbone among several infrastructure systems.

For a heterogeneous network to offer benefits in these use cases, it must satisfy several requirements.

Incremental deployment: The design must enable applications running over the old network to be transferred step by step onto the new network without a "big bang" changeover.

Ongoing testing: The emergency network must be designed and provisioned in such a way that it can be regularly exercised in non emergency periods.

Heterogeneity with interoperability: Multiple technologies must be used to achieve diversity, but this cannot compromise interoperability among agencies that are required to collaborate to meet mission goals.

Extended operation without relying on the power grid: Perhaps using energy scavenging or other innovative power sources.

Efficient support for multicast: Multicast is a predominant communications mode among first responders.

Self organizing: The requirement for rapid deployment and robust operation under changing conditions means that the network cannot require manual configuration or reconfiguration for correct functioning, and it must adapt to the environment in which it is deployed. Example techniques include spectrum scavenging and DWDM dynamic overlays.

Key performance measures for the backup network – a network whose purpose is to maintain emergency services when the primary emergency network goes down – include the following.

- Robustness (including resistance to DDOS attacks)
- Speed of deployment
- Ability to interoperate among different (unpredictable) technologies
- Level of service delivery; functional priorities (citizens protection before news reports..); TSP (telecom service priority)
- Cost (must be lower per node/user than the operational net)
- Geographic flexibility (e.g. satellites)

7 Findings

FutureHetNets 2011 was sponsored by NITRD, NSF and NASA to provide guidance on future research investments in heterogeneous networks. This section presents specific, actionable recommendations for research investment extracted from the diverse discussions of the workshop focus groups.

This section describes the takeaway from the workshop as seen by the organizers. Readers interested in the diversity of viewpoints and insights presented at the workshop are encouraged to review the detailed reports in sections 4-6.

Section 7.1 describes an overall vision for research in future heterogeneous networks. Sections 7.2 through 7.3 present specific architectural innovations ("nuggets") that have significant value for addressing the challenges of heterogeneous networks.

7.1 Overall vision

There is no doubt that computing power will increase in the near future with the development of advanced multi-core processors and cloud computing and storage. Coupled with processing power increase, the I/O speeds of new generations of processors will also increase substantially.[7] The limiting factor on how fast new applications will develop is the availability of high network speeds and much better quality of service at reasonable costs.

Device technologies and hardware subsystems are mature enough to provide at least two orders of magnitude increase in network speeds. However, left to incremental developments, the current Internet architecture will not be able to support heterogeneous network applications over heterogeneous networks at affordable costs. The principal reason is that the current architectural partitioning of networks into layers has run its course and is approaching the saturation point for further major improvements.

However, optimization that entirely breaks down the layer structure is not the correct path for future network research. If structured as an amorphous system with no boundaries, the network becomes a single large problem that is too complex for global optimization. Some form of partitioning is important for modeling and subsequently developing the right architecture for each network function.

[7] The new generation of 128 core Intel processors will have 2x10 Gbps I/O speeds using "Lightpeak" an optical I/O interface.

The current layer partitioning may or may not be the best one for investigating radically different and better network architectures. Indeed, recent developments in the physical layer of communication systems lead to functions and behaviors that do not fit current network architecture assumptions regarding physical layer properties. Some examples are:

- wireless nodes with multiple antennae supporting higher-layer control of relative bandwidth to partner nodes;
- optical networks capable of dynamically establishing and modifying high-bandwidth end-to-end transparent paths; and
- communication satellites incorporating packet switching in which the combination of high latency and inherent multicast lead to routing behaviors fundamentally different from terrestrial wireline routers.

Few if any upper layer network architectures in recent years have taken advantage of the new space of possibilities at the physical layer. It is the joint vision of the workshop participants that substantial network performance improvements can be achieved if future network research emphasizes and exploits the interplay of the physical layer and the higher network layers including applications. Due to the heterogeneity of subnets and applications, a fundamental question to ask is: whether a single suite of protocols can perform in these networks at high efficiency, or if an adaptive multiprotocol hybrid architecture is needed.

With current traditionally focused network research running into performance/cost barriers, it is time to re-examine network architectures from a fresh viewpoint. In particular, the network control plane needs much more attention. The current highly suboptimum control architecture designed in an ad-hoc fashion needs to be turned into a science with better modeling and quantitative analysis.

Hitherto network management and control has not demanded focused investigation. Sessions were short, enabling most short duration fluctuations to be smoothed by data plane statistical multiplexing. This limited the required response rate of the control plane. In the future this approach will no longer be sufficient, due to new behaviors such as "elephant sessions" that occupy an entire optical network wavelength and broadband wireless session bursting at gigabits per second. Effects like these will reduce the required speed of control plane adaptation from minutes to seconds and sub-seconds. With changes of such magnitude, scalability and network state sensing and propagation are big concerns. How we should deal with such dynamics and still ensure affordability is a wide open problem.

This is more than an isolated problem in the control plane. The physical, routing and transport layer architectures can be designed and tuned to relieve the pressure on the control plane. Thus, we believe a key addition to networking research is a scientific treatment of network management and control and its interaction with network layer architectures. With a new architecture the network may finally have the responsiveness and cost structure to match and utilize the capability of today's emerging device and physical layer technologies.

7.2 Network structured as collection of islands

This and the following sections describe specific architectural innovations ("nuggets") worthy of focused investigation.

The workshop participants coined the term "island" to refer to a region of the network that is relatively homogenous and hence capable of tight management and strong QoS guarantees. Current network design assumes that each integrated network is an island, where "integrated" refers to a network whose elements are deployed and managed to support a common mission. The necessary architectural innovation for heterogeneous networks is to develop technologies that permit an integrated network to contain a collection of islands rather than just one island.

Transformers: Current network design uses ad-hoc, unmanaged, single-purpose designed gateways to transfer data between islands. In an integrated network structured as a collection of islands, the gateways become a first-class part of the architecture that the workshop participants called a "transformer."

A transformer is an active component that interfaces and couples the heterogeneous segments in an end-to-end data path. It is active at layer 4 (the transport layer) and may terminate or initiate the transport protocol, duplicate packets, initiate resends, cache data, and perform other functions traditionally not implemented inside the network. The transformers are coupled to a enterprise-level control plane that uses them to observe network behavior, manages their resources, and exploits their capabilities for the benefit of network operators and distributed applications.

Composable Service Level Agreements (SLAs): End-to-end performance predictability and SLAs cannot be provided in current heterogeneous networks, where the only information exported from each subnetwork is the simple up/down information represented in BGP. In a network structured as a collection of islands, techniques are needed by which each island exports its service capabilities as a dynamic SLA (see section 7.6) that can be composed to provide end-to-end SLAs.

Integration across islands: New approaches are needed in multiple areas of the network architecture to realize the vision of an integrated network structured as a collection of islands.

- Common authentication/accounting mechanisms
- Information sharing behavior controlled by rules rather than manual control
- Use islands with different modalities to provide redundancy for seamless service after failure, especially in field situations where one end node can be connected to multiple islands (for example, a software-defined radio that can be reconfigured between terrestrial wireless and satellite communications modalities).

7.3 Integrated cross-layer control plane

Currently, resource management is strictly partitioned. Channels are managed by a central Network Operations Center with strong guarantees (for example, wavelengths in optical networks, or RF frequencies in wireless). Capacity within a channel is managed by switches and routers on a packet-by-packet basis with poor guarantees. For future heterogeneous networks, it will be necessary to develop a control plane that blurs this distinction and integrates the two management domains.

- Enable providing dynamic sub-channel bandwidth resource allocation to individual flows.
- Overcome complexity barrier to using full network capacity.
- Needed in order to exploit tunneling and on-demand network resources (see section 7.4).

Two architectural approaches appear promising in order to implement this vision.

- Resource management can be partitioned by time scale rather than by layer (fast-response management is fully distributed while centralized mechanisms respond to slower shifts).
- Resource management can be structured as a hierarchy, with a coordination plane interfacing to and synchronizing the management planes implemented within each island, administrative domain, or modality.

These various management approaches should be integrated into an architecturally coherent design able to optimize resource allocation at all layers of the network.

7.4 On-demand networks

Today heterogeneous networks normally offer just a few services, each intended to meet the needs of a broad class of users. It will be possible to improve scalability, reduce resource use, and more efficiently support a heterogeneous application mix by developing ways to specialize network services more precisely to application needs. Workshop participants called this concept "on-demand networks." Specific areas for innovation include:

- Improve the speed of allocation and response to change at all layers, flexibly allocating resources where needed.
- Develop control techniques to assure safe and efficient behavior of a much more dynamic network.
- Enable the underlying communications resources to be shared by multiple networks each specialized for different application or service needs.

Several architectural approaches appear promising in order to implement this vision.

- Exploit virtualization much more aggressively and fundamentally.
- Exploit tunneling as a core part of a coherent heterogeneous architecture, for example RF over fiber and end-to-end L2/L3 tunnels bypassing intermediate switches and routers.
- Exploit new physical layer technologies in fiber core (super-channel OFDM) to provide variable capacity allocation, rather than multiples of a fixed bandwidth.
- Deploy an integrated cross-layer control plane (section 7.3).

7.5 New interface between applications and control plane

In the current Internet architecture there is extremely limited information flow between applications and the network control plane: applications must infer network state solely from latency and packet loss, while the network must infer application requirements and future behavior solely from packet statistics. This approach though valuable for scalability becomes highly performance limiting in heterogeneous networks.

The necessary architectural innovation is a new interface between applications and the network control plane supporting the exchange of appropriately abstracted information to enable application-aware network behavior and network-aware application behavior. Key challenges are to preserve network scalability, modularity and security in the presence of this new information pathway.

Promising approaches to implement this vision include the following.

- Applications can express their requirements as a new form of SLA (section 7.6).
- Network permits applications to select traffic paths at a high abstraction level (e.g. which modalities are used, which transformers will cache the data) but hides routing details within a modality.

There are key unknowns that must be investigated through basic research.

- How does one create an abstract dynamic model of a network. The abstraction should be able to describe both static and dynamic state.
- How to assure stability when you couple the control loops of adaptive applications and adaptive networks

If successful this approach will have other benefits beyond improving the performance of individual applications.

- There will be less network resource impact per application, since smart applications will be able to select the timing and destination of data transfers to match resource availability in network
- The network may control its topology to meet application needs, for example by reconfiguring multiple-element antennas on wireless nodes to create a high-speed channel for a planned large data transfer.
- The energy consumption penalty caused by high reliability guarantees can be reduced. High reliability requires redundancy; redundant equipment wastes power. However, most traffic doesn't need the high reliability provided by default. The number of hot backup components/links can be reduced if applications describe their actual reliability needs.

7.6 New forms of service level agreements

Service level agreements (SLAs) are the language for describing quality of service requirements and performance guarantees. Today SLAs are contractual items, written in English and normally restricted to static specifications of a few parameters such as packet loss rate and percentile latency guarantees. The result of these coarse inflexible SLAs is required overprovisioning and sluggish service delivery changes. The innovation opportunity is to develop a machine-readable

SLA representation that is capable of expressing much richer information. Such SLAs could be used in multiple ways in a future coherent heterogeneous network architecture:

- New SLAs can allow sub-networks (“islands”) to express their capabilities to the enterprise-level control plane of an integrated network (sections 7.2 and 7.3).
- New SLAs automate the process of matching new client requests with network capabilities. This supports admission control incorporating pricing and pre-emption, enables integration of data-centric networking model, and enables on-demand networks (section 7.4).
- New SLAs provide an appropriate level of abstraction for the interface between smart applications and the heterogeneous network control plane (section 7.5).

The new service level agreement technology should satisfy multiple requirements. The SLAs should be not only machine-readable but also machine-negotiable, in real time. The SLAs should be *dynamic:* able to express that requirements or performance will vary over time based on specified conditions. The SLAs should be *composable:* enabling the end-to-end SLA to be computed based on the SLA offered by each segment in a composite path. The SLAs should be *application-level:* expressed in terms of the concepts visible to and useful for application developers rather than network implementers (this reduces developer burden and also, critically, preserves modularity enabling network developers to innovate).

7.7 Summary

The preceding discussion is intended as the starting point for a wider investigation of heterogeneous networks. The research vision and architectural nuggets described in this section capture only a portion of the rich set of insights and conceptual frameworks discussed at the workshop. An even broader range of opportunities will arise as a larger community engages with the critical problems of these networks. However, the concrete targets identified here show that the time is ripe for substantial research investments. The problems and opportunities have come together enabling significant progress in the near future. From this perspective, the FutureHetNets 2011 workshop provided a valuable milestone in the development of the heterogeneous networks field.

8 Appendices

8.1 Scope of workshop

The following language was approved by the workshop's sponsors.

Future heterogeneous networks will have multiple modalities (wired, wireless, satellite) and range from low to ultra-high user data rates (100 Gbps+). Major challenges facing the deployment of such networks are:

1. coordination of the widely different data rates offered by different subnets to the users;
2. efficient support of heterogeneous voice, video and data services with several orders spread in transaction sizes and rates and with heterogeneous priority;
3. development of a common set of protocols that work across disparate networks with very different physical layer attributes; and,
4. development of a network management and control architecture that is capable of efficient and secure operation over heterogeneous networks.

To provide the basis for research programs that will address these challenges, the workshop participants were asked to explore the concept of "Highly controllable dynamic heterogeneous networking." Free exploration was encouraged. The initial agenda included:

1. understanding the flexibility and characterizing the attributes of future physical layer architectures;
2. dynamic networking in network management and control and in user resource allocation over heterogeneous networks;
3. controls over a wide range of time scales from quasi-static provisioning to per session controls for large transactions;
4. re-examination and reconstruction of the entire network protocol stack and repartition of the stack if necessary; and,
5. cross-domain and cross-modality management and scheduling.

8.2 Workshop structure and schedule

Thursday, March 24, 2011		Friday, March 25, 2011	
8:00	Welcome and introduction	8:00	Cross-layer breakout outbriefs
8:20	Objective of study	9:00	Heterogeneous networks breakouts
9:00	Requirement presentations	10:30	Break
	Future Internet Architecture	11:00	Heterogeneous networks continue
	Darleen Fisher, NSF	12:00	Lunch
	NASA applications that impact future network design	13:00	Heterogeneous networks continue
	Ray O'Brien, NASA Ames	14:30	Break
	Meeting the needs of growing traffic: challenge for wireless	15:00	Heterogeneous networks outbriefs
	K. K. Ramakrishnan, AT&T labs research	16:15	Discussion and next steps
	Defense Network Needs	17:00	Workshop ends
	Vincent Chan, MIT		
10:20	Break		
10:45	Layer breakout sessions		
12:00	Lunch		
13:00	Layer breakouts continue		
14:00	Layer breakout outbriefs		
15:00	Break		
15:30	Cross-layer breakout sessions		
17:00	Adjourn		
18:00	Banquet		

8.3 Agendas for each breakout session

To kick off discussion, each breakout session described in section 8.2 was given a set of questions.

8.3.1 Per-layer sessions

The list below refers to an "multiple instances" of a network layer. This refers to the range of system designs at that layer. For example, at the physical layer it refers to multiple modalities such as fiber, wireless, and SATCOM. At the network layer it refers to the network design comprising topology management, routing, and related functions. At the transport layer it refers to different protocols that solve transport requirements including end-to-end reliability and congestion management.

1. What new technologies, new designs at this layer, and new applications are potentially coming?
2. What would be key drivers that would lead to their deployment?
3. How can the behavior of current and future instances of this layer be statistically characterized and abstractly represented to other layers, particularly for networks with intermittent links and rapidly changing rates.
4. What new properties that can substantially diminish performance will future instances of this layer manifest to other layers? Characterize the space of possibilities, not just a point design.
5. (Higher layers) What problems will this layer encounter with the new physical modalities with dynamically changing properties?
6. What are appropriate quantitative performance metrics?
7. What are the major open problems when multiple disparate instances of this layer are composed into an heterogeneous network?
8. What are the appropriate metrics to assess progress in this area?

8.3.2 Cross-layer sessions

1. Describe the key architectural options that need to be explored to achieve a significant jump (~10x) in communications capability in this area. What are the major barriers to realizing each one?
2. How could the open problems identified in the previous breakout session, and the barriers described in the previous bullet point, be simplified or become more tractable with better cross-layer interaction and/or new forms of support from the network management system?
3. What abstractions, API innovations, or new architectural approaches can be used to support better cross-layer interaction while preserving modularity, simplicity, and the ability of the separate subsystems to evolve independently.
4. What are the appropriate metrics to use to assess progress in communications capability in this area?

Important issues for wireless and SATCOM cross-layer: intermittent and dynamically varying rate channels, rapidly changing link state and effect on routing, transport layer misinterpretation of observables, application layer performance degradations such as excessive delays.

Important issues for high speed fiber cross-layer: Very high speed (>100Gbps), long reach and dynamically switched channels, rapidly changing link state and effect on routing especially complexity of session scheduling, transport layer inefficiencies at high link speeds, application layer performance degradations such as excessive delays.

8.3.3 Heterogeneous network architecture

1. List the key end-user visible problems or limitations that arise in current large-scale heterogeneous networks. Select the top 4 or 5 that need to be addressed through research investment.
2. Identify key architectural concepts or options that create the possibility of significant improvement in these problems or limitations.
3. Focus on the areas outside/crossing the traditional layers: security, management, quality of service, robustness. What progress is needed, what new concepts hold the potential for breakthrough improvements, how can innovative architecture enable new solutions?
4. Describe appropriate large-scale experiments and/or testbeds that are vital for progress in this area. Identify the ways in which current assets/projects (GENII, PlanetLab, etc.) are useful or do not fully meet the requirement.
5. In what areas will ongoing commercial investment trajectories lead to ongoing progress; in what areas will commercial R&D fail to solve critical problems, and why won't commercial R&D address those problems?
6. What topics are appropriate for high-risk high-payoff basic research investment in this area?

8.4 Participants

There were 74 participants: 51 from academic institutions, 10 from private industry and 13 from government agencies and research laboratories.

Table 8.1 FutureHetNets 2011 participants

Keren Bergman, Columbia University Andrey Bobyshev, Fermilab Maite Brandt-Pearce, University of Virginia Bob Brodersen, Berkeley Wireless Research Center Richard R Brooks, Clemson University Vincent Chan, Massachusetts Institute of Technology Gee-Kung Chang, Georgia Tech John Chapin, Massachusetts Institute of Technology Hao Che, University of Texas at Arlington Angela Chiu, AT&T Labs Research Loren Clare, Jet Propulsion Laboratory Bharat Doshi, Johns Hopkins University Tarek S. El-Bawab, Jackson State University Pat Elson, NASA Ames Research Center Darleen Fisher, National Science Foundation Victor S. Frost, University of Kansas Mario Gerla, UCLA Paul Grams, NASA Ames Research T. Russell Hsing, Telcordia Technologies, Inc. Rose Qingyang Hu, Utah State University Joseph Hui, Arizona State University Anura Jayasumana, Colorado State University Chuanyi Ji, Georgia Tech Kevin L. Jones, NASA Mohsen Kavehrad, Pennsylvania State University Leonid Kazovsky, Stanford University James Kempf, Ericsson Research Daniel C. Kilper, Bell Labs, Alcatel-Lucent Alan Kirby, Massachusetts Institute of Technology Sitaram Kowtha, Johns Hopkins University Tony Li, Cisco Systems Jin Li, Microsoft Research Shan Lin, Temple University Malathi Veeraraghavan, University of Virginia Joe Mambretti, Northwestern University Z. Morley Mao, University of Michigan Jim Martin, Clemson University Nicholas F. Maxemchuk, Columbia Univesity Lee W. McKnight, Syracuse University	Muriel Medard, Massachusetts Institute of Technology Grant Miller, National Coordination Office Inder Monga, Energy Sciences Network Sayandev Mukherjee, DOCOMO Communications Labs USA Inc. Guevara Noubir, Northeastern University Tim Owen, DoD HPCMP/DREN Chunming Qiao, SUNY Buffalo K. K. Ramakrishnan, AT&T Labs. Byrav Ramamurthy, University of Nebraska-Lincoln Srinivasan Ramasubramanian, University of Arizona Sundeep Rangan, NYU-Poly Brian Rathgeb, Department of Defense Glenn Ricart, M. Y. Medy Sanadidi, UCLA Aileen Y. Sansone, Chief of Naval Operations Strategic Studies Group Galen Sasaki, University of Hawaii Srinivasan Seshan, Carnegie Mellon University Kang G. Shin, University of Michigan Allen Shum, SPAWAR Systems Center Pacific Alex Sprintson, Texas A&M University Peter Steenkiste, Carnegie Mellon University Ravi Sundaram, Northeastern University Violet R. Syrotiuk, Arizona State University Brian Tierney, Joe Touch, USC/ISI Pramode Verma, Univ. of Oklahoma-Tulsa Vinod Vokkarane, University of Massachusetts Dartmouth Tilman Wolf, University of Massachusetts Amherst Andrew Worthen, MIT Lincoln Laboratory Jin Yang, Verizon Communications Inc. S. J. Ben Yoo, University of California, Davis Murat Yuksel, University of Nevada - Reno Lixia Zhang, UCLA Hongwei Zhang, Wayne State University Tao Zhang, Telcordia Technoloiges Inc.

8.5 Acknowledgements

The workshop organizers and participants deeply appreciate the financial support of NSF for this workshop. We learned a lot about each others views on network research and the interactions stimulated many new ideas at and after the workshop. Thanks to Taieb Znati, former Division Director, Division of Computer and Network Systems, NSF, for initiating the idea for the workshop. Thanks also to Darleen Fisher and Victor Frost of NSF for their help in carrying this idea to fruition.

The National Coordinating Office for Network and Information Technology Research and Development (NITRD) helped the coordination of this workshop among government agencies: primarily NSF and NASA. Grant Miller of NITRD did a lot of behind the scene work to get interested agencies together for the workshop.

The participation and generous support of NASA is also greatly appreciated. Kevin Jones in addition to representing NASA and being a participant was the gracious host for the workshop at NASA AMES Research Laboratory. Ms. Pat Elson was responsible for all the logistics at the workshop. The workshop ran smoothly without a hitch. Most of all, she made our workshop during two chilly and rainy days at AMES, warm and welcoming.

We want to thank Dave Foss, the Assistant Director for Information Technology Services of the Research Laboratory of Electronics, MIT, for setting up the Website of the Workshop.

Dr. John Chapin took the lead in the organization and writing of the report. His many hours of investment on this workshop have made the workshop fruitful and the report coherent.

Finally, my assistant, Ms. Donna Beaudry, spent hours with travel, finance, keeping the workshop website current and on report productions. We would not have pulled this off without her hard work.