

**Modeling and Simulation of Ultra-Large Networks:  
A Framework For New Research Directions**

Prepared under the NSF Grant ANI-0135530

by

Bernard P. Zeigler  
Director, Arizona Center for Integrative Modeling and Simulation  
University of Arizona  
[www.acims.arizona.edu](http://www.acims.arizona.edu)

and

Saurabh Mittal  
ECE Dept. University of Arizona,  
Tucson, AZ 85721

**List of Participants and Observers:**

<b>Participants</b>	<b>Organization</b>
Agarwal, S.	Lucent Technologies
Claffy, K.	San Diego Super Computing Center, SCSD
Colbourn, C.	Arizona State University
Cox, P.	Lockheed Martin
Floyd, S.	Lawrence Berkeley Nat'l Lab/ACIRI/UCB
Fujimoto, R.	Georgia Tech
Gelenbe, E.	Univ. of Central Florida
Ghosh, S.	Stevens Institute of Technology
Hariri, S.	University of Arizona
Hui, J.	Arizona State University
Joslyn, C.	Los Alamos National Laboratory
Kim, T.G.	Korea Advanced Inst. of Science & Tech.
Martinez, R.	University of Arizona
Ng, H.	Naval Research Laboratory
Meyer, D.	Sprint
Nicol, D.	Dartmouth College
Orman, H.	Nortel
Raghavendra, C.	University of Southern California
Reily, G.	Georgia Tech
Sanders, W.	University of Illinois, Urbana Champaign
Sarjoughian, H.	Arizona State University
Wainer, G.	Carleton University
Wilsey, P.	University of Cincinnati
Woodcock, B.	Zocal
Zeigler, B.	University of Arizona
Znati, T.	NSF/University of Pittsburgh
Cam, H.	Arizona State University
Gupta, S.	Arizona State University
Louri, A.	University of Arizona
Reisslein, M.	Arizona State University
Rozenblit, J.	University of Arizona
Sen, A.	Arizona State University
Xue, G.	Arizona State University

# INDEX

<i>Introduction</i>	4
<b>1. Some Critical Issues and Problems</b>	<b>6</b>
<b>2. Some Design Approaches to Solve The Problems</b>	<b>11</b>
<b>3. The Need for Modeling and Simulation</b>	<b>15</b>
<b>4. Research needed to enable M&amp;S to meet the challenge</b>	<b>24</b>
<i>References</i>	29
<i>Links</i>	34

## Introduction

Sponsored by The National Science Foundation Advanced Networking Infrastructure Research Program (NSF/ANIR), a workshop on “Modeling and Simulation of Ultra Large Networks: Challenges and New Research Directions” was held on Nov. 19-20, 2001 in Tucson, AZ, USA. It was organized by The Arizona Center for Integrative Modeling and Simulation (<http://www.acims.arizona.edu>) and The Society for Modeling and Simulation International (<http://www.scs.org>). This report is one of two prepared under the NSF Grant ANI-0135530 that supported the workshop. The first report “*Modeling and Simulation of Ultra-Large Networks: 13 Recommendations for Needed Research*” documents the detailed recommendations that were formulated during the workshop. This second report was prepared after the workshop and is based in part on the workshop discussions and recommendations and in part on subsequent study of the literature and exchanges with workshop participants and others. In contrast to the first report, the present report attempts an attempt to provide a framework in which the recommendations can be understood. An earlier report by the Division on Engineering and Physical Sciences, National Research Council titled “*Looking Over the Fence at Networks – A Neighbor’s view of Networking Research*” dealt with measurement issues, modeling and theory of networking, the need for such new methodologies for design of infrastructure elements. It proposes an approach to migrating new design elements into the Internet called disruptive prototypes. The present report can be viewed as picking up where the NRC report left off – our objectives are to look more deeply into the crucial role of modeling and simulation in networking research and propose some directions for research in modeling and simulation that will enable it to better play this role

In this report, we examine the Internet from the point of view of its future as a ULN. To frame the need for modeling and simulation and associated research directions, we first consider some representative issues and problems that arise in this context. Then we enumerate some possible designs and interventions that have some initial plausibility but need modeling and simulation for their evaluation. On this basis, we place modeling and simulation activities within an iterative cycle of ULN development. The latter enables us to explicate the current capabilities of modeling and simulation and the research needed to meet the challenges of ULN applications.

---

### Case in point: Flat versus Hierarchical Growth

At the end of NSFNet in 1995, the NSF established five Network Access Points (NAPs) in the continental US. Today the number has increased to ten or twelve. This forms the

core<sup>1</sup> of the present Internet. Level 2 consists of router nodes that connect different Autonomous Systems<sup>2</sup> that peer with Level 1 routers. In a recent study<sup>3</sup> it was found that the present Internet is growing in a flat manner with the new nodes mainly being added at Level 2. These Level 2 routers, numbering between 4000 and 6000<sup>4</sup>, communicate to each other through Border Gateway Protocol (BGP) in a peer-to-peer fashion. This large number causes even the robust BGP protocol to be brought to its knees in the advent of accidental switching off of a BGP router, misconfiguration or a router crash due to heavy traffic<sup>5</sup>.

It has been proposed that the Internet should grow in a hierarchical manner to be better managed. To achieve this, overlay networks<sup>6</sup> could be created over Level 2 that deal with the configuration of these BGP routers. The basic tasks of a BGP router would then be to maintain the updated routing tables for the lower tier and to share the traffic in case of bottlenecks in its sub-network.

Clearly, a proposal of this magnitude cannot be introduced directly into the operational Internet. The deployment of the overlay network in the commercial Internet is not appealing to the ISPs since it would bypass their installed infrastructure. Moreover, analytic methods are incapable of predicting the impact of an intervention on such a large scale. Therefore, extensive modeling is needed to characterize the current Internet and the effect of the combined overlay/hierarchical growth strategy. Once developed, tested and proven in a modeling and simulation testbed, with the risks known and managed, the strategy has a better chance of acceptance by all stakeholders in the Internet and its future growth as an Ultra-large scale network (ULN).

---

<sup>1</sup> Labovitz, C., Ahuja, A., and Jahanian, F., "Experimental Study of Internet Stability and Wide-area Backbone Failures", Proc. International Symposium on Fault-Tolerant Computing, June 1999

<sup>2</sup> The Internet is divided into a large number of different regions of administrative control commonly called Autonomous Systems (AS). These ASes usually have distinct routing policies and connect to one or more ASes at private or public exchange points. At the boundary of each AS, peer border routers exchange reachability information to destination IP address blocks or prefixes, for both transit networks and networks originating in their routing domain.

<sup>3</sup> Laboviotz, C., Malan, R., and Jahanian, F., "Internet Routing Instability", Univ. of Michigan., Proceedings of ACM SIGCOMM, September 1997.

<sup>4</sup> These figures correspond to findings in 1997. The extrapolated figures in the present Internet have alarming proportions

<sup>5</sup> Govindan, R., and Reddy, A., "An Analysis of Inter-Domain Topology and Route Stability", IEEE INFOCOMM'97

<sup>6</sup> Overlay networks sets up routers that can share traffic with existing ones to alleviate overflow traffic around autonomous networks. The NRC report suggests that such networks could be employed parallel to the core to test and experiment with new disruptive technologies. Strategies could be developed to migrate over to complete dependence on new technologies once they had been tested in this fashion.

"Looking Over the Fence at Networks-A Neighbor's view of Networking Research", National Research Council

# 1. Some Critical Issues and Problems

As indicated, our later discussion of modeling and simulation will be facilitated by considering how it can help to address some critical issues in Internet development as it increases in size. No claim is made that these issues are the most urgent ones facing Internet development today. Nor are they listed in any particular priority order. They are meant only to be illustrative of current critical issues and to lay the basis for the consideration of research directions for modeling and simulation.

## **1. Scalability of the current Internet to the Ultra-large network of the future may require fundamental changes in its architecture, for example, new routing policies and address mechanisms.**

As the Internet grows in size, house-keeping activities, such as maintaining proper routing information, threaten to consume more and more of its resources. For example, current routing tables currently hold approximately 45,000 Internet Protocol (IP) prefixes<sup>7</sup>. Approximately 6 million routing prefix updates are exchanged at the core every day with as many as 30 million on particular occasions. However, the vast majority (99%) of the routing information is redundant and does not really reflect the actual updates. Moreover, it is becoming increasingly difficult to institute more efficient addressing schemes due to multi-homing,<sup>8</sup> increasing complexity of connectivity, and the legacy of earlier address allocation schemes.

Although redundant updating information is benign in itself, the large volume of prefix updates generates bursty traffic and unnecessary use of network resources. Moreover, bursts of such traffic can cause failures – on the order of 300 updates per second are enough to crash a router due to congestion (see issue 3. below). As the topological complexity grows, the potential for instabilities associated with propagation of router crashes also increases.

Apart from the inefficiency of updating and potential for instability this situation causes, it also impacts scalability toward ULN in that inefficient address allocation limits the number of nodes that can be added within the bounds of a finite address space.

A recent article in the popular press<sup>9</sup> asserts that video, audio, and giant data files are likely to represent the bulk of Internet traffic in the near future. Today's Internet "chokes" on such traffic and the author ascribes the cause to the basic architecture of the Internet. First of all, the main goal of the original (arpanet) design was not speed but

---

<sup>7</sup> Figures borrowed from findings in 1997. Extrapolation of these values to 2002 give a much higher number

<sup>8</sup> Multihoming refers to the existence of multiple connections between an organization (often a business corporation) to the Internet. Multiple links can arise for example, when the organization obtains service from two or more ISPs. Alternatively, Multihoming occurs when there are several externally accessible POPs (points of presence, i.e., visible in global routing table) linked internally within the organization's intranet.

<sup>9</sup> Lagesse, D., "Accelerating the Internet", US News and World Report Magazine, April 22, 2002

flexibility. Second, growth and innovations are happening at the edges where senders and receivers are found, rather than in the core (innermost routers<sup>10</sup>), which was not designed for the new traffic environment that is arising.

Current routing protocols employ distance vector estimates to plan packet routes.<sup>11</sup> Such estimates are based on a picture of the topology of the network as viewed by the packet during its transit to the destination. This topology might not reflect the true state of the network and therefore might well not provide the best route towards the destination. New, more effective, routing policies such as those based on the link states, (to be discussed) will work only on the basis of better addressing schemes.

## **2. The current transmission protocol, TCP is inadequate to characterize and control current traffic flows, let alone the diversity and magnitudes anticipated in the future.**

The current TCP protocol, used as a carrier by BGP, does not allow observing the “day-in-the-life” of a packet (its history of traversing the routers on its path of travel) since the packet header does not include the extra fields need for this purpose. Although alternative approaches to measurement have been developed, such as traceroute, to trace packet paths, they are costly to implement and limited in their data gathering capability. The critical importance of “day-in-the-life” packet information lies in that fact that nearly every network performance parameter depends upon the end-to-end design and diurnal usage patterns in various time zones that create traffic swings and effectively different time varying topologies. Information on this constantly shifting routing of packets, which is hard to trace with the current TCP protocol, is needed to adequately characterize traffic flows.

Apart from its limitations in tracing packet routes, TCP employs an approach to congestion estimation whose credibility may be questioned. The routers use this fundamental scheme as their underlying transmission protocol. The protocol works on a sliding window mechanism and responds to acknowledgment of packet arrival, or lack thereof, by modifying the size of the window (e.g., reducing it by half when acknowledgements fail to arrive). In so doing, the protocol employs packet loss as a proxy measurement for congestion in the round-trip loop. Unfortunately, packet loss is not necessarily reflective of congestion and its use as feedback signal for congestion control may be suspect. For example, corruption of the acknowledgment packet may cause it not to be received by the sender, a situation more likely to occur in wireless communications with their higher error bit rates.

---

<sup>10</sup> Surprisingly there are only ten to twelve core routers managed by independent ISPs, such as ATT, Sprint, etc. (Labovitz, C., and Ahuja, A., “*Experimental Study of Internet Stability and Wide Area Backbone Failure*”,). The numbers correspond to the year 1997.

<sup>11</sup> These routers run BGP protocol to communicate with the huge number of Level2 peers. The BGP is a path distance Vector protocol with a time complexity of  $O(n!)$  in the worst case scenario. Despite its robustness, with this complexity, the BGP algorithm, taking huge times to converge even if the  $n$  is not that large.

Currently there exist approximately 11 versions<sup>12</sup> of TCP used by various groups for their individual networks. The number of such variants suggests that the current version is far from optimum. For example, it is very easy to corrupt a TCP client by making changes in the protocol at the receiver side Operating System. The resulting modified version of TCP can fool the sender and can expand its window size to infinity (as the worst case), eventually leading to the congestion collapse. TCP's vulnerabilities arise from a combination of unstated assumptions, casual specifications and pragmatic need to develop congestion control mechanisms that are backward compatible with previous TCP implementations.<sup>13</sup>

### **3. The Internet has evolved in recent years in such a way that its core structure is vulnerable not only to local router crashes but also to system-wide propagating events they may initiate.**

There are very few algorithms and strategies to reconfigure the system after a system wide failure. The most viable solution still appears to be shutting down the impacted devices, letting the situation stabilize and then restarting the devices again. For example, router-crash is defined as the inability of a router to respond to requests within a time window so that time-out occurs and data packets are lost due to heavy queuing. Each router has approximately 45,000 IP prefixes and shares a minimum of 3 million updates each day. One or two routers can bring the entire Internet down within half an hour of deliberate misconfiguration or accidental switching. By deliberately switching a router on and off every 4 minutes for 15 minutes one can initiate a "route-flap" storm that propagates by initiating a chain reaction. Once triggered it is almost impossible to find its origin or the type of fault that initially caused it (e.g., router misconfiguration, accidental switching, or overload on router due to congestion. Further there is no known remedy to restore normal operation other than switching the routers down for a finite time and then restarting again. At least two days are needed to restore operation to normal. Such an incident occurred on July 19, 2001 and came to be known as the Code Red Virus.<sup>14</sup>

### **4. There are intrinsic features of the current Internet that limit the services it can offer and the Quality of Service (QoS) it can provide.**

Internet users can be classified into four broad categories:

- Researcher

---

<sup>12</sup> Floyd, S., and Paxson, V., "Difficulties in Simulating the Internet", IEEE/ACM Transactions on Networking, February 2001.

<sup>13</sup> Savage, S., Cardwell, N., Wetherall, D., and Anderson, T., "TCP Congestion Control with a Misbehaving Receiver",. An IETF document<sup>13</sup> addresses the problems with TCP6 and its improved version TCPng using a *transport address*. This request for comment document, dated 1994, envisions the limited capability of CIDR addressing and eventually its exhaustion. However, it accepts CIDR addressing as the only the means to buy time to migrate from TCP6 to TCPng or from IPv4 to IPng.

<sup>14</sup> The effect went down to /24 level (Classless Inter-Domain Routing addressing) indicating the high number of HTTP requests on port 80 as the routes became unavailable.

- Professional
- Business
- Leisure

The primary requirement of the researcher is to get the computing resources, usually the grid resources to carry out the heavy computations in minimal amount of time<sup>15</sup>. With the research in fields like astronomy, genetics, microbiology, environmental study and geoscience requiring very high memory and CPU resources, the current Internet fails to provide the cumulative distributed processing power at various locations. Technologies that integrate various computational tools and different Operating Systems (e.g., Globus, AdVice) have been developed to provide the required services. However, the latency and the performance of the Internet renders these designs infeasible.

The professional requires Internet to run applications that involve video-conferencing and time-bound analysis of data. The present infrastructure does not provide reliable means to execute such applications.

Business is a major growth category, including business-to-consumer (B2C) and business-to-business (B2B) applications. Web-enabled business proffers enormous advantages through online catalogs and sales, but incurs the risk of rapidly spread consumer complaints mediated by email or on-the-fly websites. Businesses can greatly increase inventory and production efficiency via distributed and collaborative supply chain information sharing, management and control.

Finally, leisure use includes social interaction through email, games, music downloads, etc. As users grow more skillful, they demand increased high load applications such as video, multimedia, and 3D games.

The current Internet offers each category some of the services that it can offer at that particular instant. Due to lack of predictability, there is no guarantee of repeating the same reliability on another occasion. Moreover, current Internet tools and infrastructure fail to provide the resource usage and work-load measurements needed to develop predictive models on which to base services.

**5. The current infrastructure is inflexible; it has minimal ability to reconfigure itself in response to failures and to accommodate new policies, mechanisms and technologies**

Internet research and development has yielded a number of potential enhancements to the current infrastructure. One can name, for example, developments such as:

- Support for novel network layers
- High speed IPV6

---

<sup>15</sup> Anderson D. P., Kubiawicz, J., "The Worldwide Computer", Scientific American. March 2002.

- Dense wave division multiplexing
- Novel optical network interfaces
- Applications that can model traffic
- Support for high bursty traffic
- Multicasting, and
- Guaranteed QoS.

Such research products are available in the literature and some have been tested and proved in small experimental test-beds by their designers. Although some of these technologies have been known for a decade, no significant progress has been made towards deploying them in the current Internet. One reason appears to be that each of these specific implementation require a minimum apriori guarantee of the Internet which it cannot provide because its behavior is non-deterministic. Perhaps the most significant factor that hampers their deployment is a rigid network infrastructure that does not allow changes to be easily migrated into operation.<sup>16</sup>

## **6. Internet Service providers, ISPs control the current structure of the Internet and experimental access to it**

ISPs are reluctant to share their infrastructure topology (how nodes are connected together) and their customer base. Moreover, ISPs' reluctance to cooperate makes it difficult to trace packets as they move from subnetwork to the next. It would be desirable to be able to release data from various ISPs to non-profit research organizations. This data could be encrypted to restrict access to it by competing businesses.

---

<sup>16</sup> One of the significant examples is the migration from IPV4 to IPV6The IPV6 provides some of the solutions to the current Internet's problems but the transition from IPV4 is a long process and will continue till 2030-2040 from today. During this transition period both the versions of the IP protocol have to coexist. "The Transition to IPV6", ISOC Member Briefing #6, Eric Carmes, Internet Society, January 2002

## 2. Some Design Approaches to Solve the Problems

### 1. Implement new addressing scheme and reassign old addresses

When routers are visible to each other, they must update one another on routing information. However, routers that are subordinate to higher-level routers need not be visible to each other since once a packet arrives at a parent router it can then be routed down to the children. Aggregation of prefixes, also known as supernetting, combines a number of smaller IP prefixes into a single, less specific route announcement reducing their visibility to other routers. Such aggregation is capable of significantly reducing the amount of routing information that is periodically shared among BGP routers. However, its success requires cooperation among ISPs and a well-planned addressing scheme. Such a scheme should ideally approximate a hierarchical addressing structure, where successive levels of a prefix tree denote successively encapsulated nodes. Unfortunately, Multi-homing, increasing complexity of connectivity and the legacy of earlier address allocation schemes make it difficult to institute more efficient addressing schemes.<sup>17</sup> Thus, the problem, under such constraints, is how to reduce, as much as possible, the visible addresses and thereby the volume of required updates. Aggregation through clustering and super-clustering at the core level also will provide better abstraction of the structural Internet topology<sup>18</sup>. One approach is to connect Level 2 routers in an overlay fashion and create tiers between Level 1 and Level 2 so that better visibility is maintained and peer communication is reduced. As part of the solution, one can consider revising

---

<sup>17</sup> As mentioned, there are only 10-12 routers at the core maintained by different ISPs. This is Level 1 of Internet. At the Level 2 there are about 4000-6000 routers connecting a total of 1300 Autonomous Systems (AS) that maintain the “default free” routing table for the visible, reachable IP addresses. Although there are only 1300 ASes, there are 1500 unique AS paths that result from multi-homing as described in the next paragraph. The data corresponds to findings in 1997.

The problem of inefficient addressing is becoming more pronounced with the rapidly increasing number of end-sites that are multi-homed, i.e., they obtain redundant connectivity to the Internet via multiple ISPs. Multi-homing requires that the Area Border routers, near the core of the Internet, maintain the more specific, or longer, prefixes in addition to less specific aggregate block prefixes containing the multi-homed sites. Of all the addresses in the routing tables maintained by these routers, 25% of the addresses are multi-homed. Also the Internet is growing less hierarchical with rapid addition of more and more Internet Exchange Points (BGP routers) developing peering relationships to the already existing base of 4000-6000 routers. Prior to introduction of today’s CIDR addressing, the majority of customer sites obtained their addresses directly from the InternNIC instead of from their ISPs CIDR block. This has resulted today into a large number of globally visible, non-aggregated, IP addresses. (Acronyms: CIDR stands for Classless Inter-Domain Routing, a new addressing scheme for the Internet which allows for more efficient allocation of IP addresses than the old Class A, B, and C address scheme; InternNIC is a registered service mark of the U.S. Department of Commerce. It is licensed to the Internet Corporation for Assigned Names and Numbers (www.internic.net)).

<sup>18</sup> AS level topology, though simple and easy to obtain, is too coarse to model the network proximity of IP addresses. Router level topology is too fine-grained to abstract the required information. The cluster graphs provide more stability than router-level graph, are as easy to obtain as AS graphs and provide the appropriate Internet topology. (Krishnamurthy, B., and Wang, J., “Topology Modeling via Cluster Graphs”, ACM SIGCOMM Internet Measurement Workshop, October 2001)

existing addresses given to clients with new ones from the same address bank and building the logical topology.

As the pressure on hierarchical communication is mitigated, the routing updates can be more easily abstracted and advantages of cluster topology can be exploited. This will bring about the hierarchical growth of the Internet that is the fundamental property required for scalability.

## **2. Increase the packet header size from 32 bit to 48 or 64 bit to accommodate 1 billion or more nodes along with the additional packet information**

A basic consequence of employing a fixed header size is that the number of possible addresses is bounded, albeit exponentially, by the number of bits in the frame. The current 32-bit header is not sufficient to provide for 1 billion IP addresses<sup>19</sup>. The header needs to be increased to 48 or 64 bits to accommodate the addresses needed for future growth. Further, since addresses at the Medium Access Control (MAC) layer are already 48 bits, increasing the IP address size will eliminate the address translation that is currently required<sup>20</sup>. An important secondary benefit of this change in the packet structure is that it can now record more information as a packet traverses the network. As indicated earlier, this “day-life of a packet” information may be useful, among other things, for improved measurement of congestion.<sup>21</sup>

## **3. Revise routing policies and algorithms to benefit from increased address aggregation**

Improved addressing schemes based on cluster topology<sup>22</sup> may reduce the computational burden on routing policies, or possibly eliminate the need for such policies entirely.<sup>23</sup> The reduced computational load may in turn open up possibilities to revise routing policies to exploit the new addressing environment. For example, being based on veridical knowledge of the congestion state of the network, link state protocols can provide

---

<sup>19</sup> It should be noted that the IPV6 contains a 128 bit header and solves the addressing space problem but has its own limitations when deployment is considered in the current scenario.

<sup>20</sup> One such solution is available since 1994 with improved version of TCP implementation called as TCPng. This version defines a TCP header of 48 bits and bringing about the change has negligible impact on communications. Moreover, for speeds above T1 (1.544 Mbps), the increased time is all under 20 microseconds. This modified version of TCP is independent of the Network Layer protocol underneath and can support transition from IPV4 to IPV6 or IPng temporarily by decoupling the Transport Layer and the Network Layer. “Six Virtual Inches to the Left: the problem with IPng”, rfc 1705

<sup>21</sup> “Looking Over the Fence at Networks-A Neighbor’s view of Networking Research”, National Research Council

<sup>22</sup> Perkins, C.E., “Ad-Hoc Networking”, Addison Wesley, Chapter 4

<sup>23</sup> A recent study found that the shortest AS paths represent the routing that would have resulted from a pure (policy-free) hierarchical routing in the Internet. (Tangmunarunkit, H., Govindan, R., Estrin, D., and Shenker, S., “The Impact of Routing Policy on Internet Paths”, In *Proceedings of 20th IEEE INFOCOM*, April 2001).

superior routing to current distance vector protocols. However, their past applicability has been limited to relatively small networks (e.g., networks comprising of few ASes and within a single AS) where the updating they require to maintain the correct routing tables is not overwhelming. With the advent of increased addressing aggregation, the scalability of link-state protocols will be positively impacted and they may emerge as attractive alternatives to current vector-based protocols (as in inter-AS routing protocols).

Routing policies can exploit the abstraction of network topology provided by hierarchical addressing to gain more global views of the network. Greater determinism in path planning can in turn allow better guarantees of the QoS over pre-defined paths with acceptable variation (jitter).

Furthermore, if both geographical and network topology are in conformance with each other then load balancing can be implemented on a global scale thus benefiting from the “diurnal invariant” (time zones that are on opposite sides of the globe alternate in their work/sleep cycles).

#### **4. Adopt new ways of measuring traffic congestion and control policies**

As indicated above, the current transport protocol, TCP, employs packet loss as a proxy for measurement of local congestion, an approach that is open to question in view of the multitude of causes for packet loss. Even the routers at the core Level (BGP) use this as their transmission protocol. New approaches to more direct estimation of congestion might be developed and compared with the prevailing approach. Concomitantly, new control policies can be developed to exploit improved congestion estimation.

Companies, like Network Physics,<sup>24</sup> have emerged in the last two years to deal with the complex issues of network measurement and identifying congestion in the network. They offer commercial software that is quite expensive (of the order of \$50,000 and above). Network Physics claim that they can control the complete path of a packet until it reaches its end user and that the company can optimize this path. They also claim that they can effectively balance the load among the routers and can extract information from the BGP routers.

But solutions like this can prove shortsighted as these end-applications running at the periphery of the Internet will act as traffic boosters for a few entities and will shift the focus towards them. This may result in more delay for the applications that have smaller bandwidth requirements and the network traffic pattern will be dominated by large enterprises who can afford the solution. Moreover, this application doesn't solve the multi-homing issues but rather encourage enterprises to get multi-homed and then optimize the traffic, making the future of the Internet even more complex system. There

---

<sup>24</sup> [www.networkphysics.com](http://www.networkphysics.com)  
[www.arbor.net](http://www.arbor.net)

seems to be no point in developing intelligent applications that operate at the periphery without paying attention to the structure of the Internet.

### **5. Devise new strategies to recover from crashes**

There are not enough recovery mechanisms that deal with router blocking or system-wide network crashes. The vulnerability at Level 2 requires the development of strategies to counter threats such as the Code Red Virus.

### **6. Institute economic policies to incentivize change**

The ISPs play a major role in today's Internet and hold valuable information that can unearth many aspects of the current Internet. Change in the economic policies and considering them as a separate class will benefit both the growth of Internet and economy.

### 3. The Need for Modeling and Simulation

Before the Internet grew to its current size and complexity, it was possible to design and test routing algorithms in small homogeneous networks. Unfortunately, these same algorithms now form the basic protocols used in today's highly heterogeneous Internet. Rooted in this earlier tradition, current tools are not powerful enough to solve today's most pressing problems, such as understanding system wide network behavior and preventing catastrophic breakdowns due to cascading failures. Currently, it is not possible to experiment with the Internet to identify critical points of vulnerability, design and test more robust routing algorithms, and gain overall control of end-to-end behavior. On the other hand, if data and analysis were available to support models of Internet behavior, we could isolate critical failure modes and structural weaknesses and take preventive measure to forestall their occurrence. Better algorithms are required that are resilient to real-world limitations on the available information needed to make optimal decisions. To be able to react quickly to changes in traffic swings and find good transmission paths, these algorithms should be network aware, i.e., that is be able to sense the magnitude and direction of incident and outgoing traffic flows. Theory and models are needed to design and test such improved algorithms. Routing schemes must take into account the composition of the Internet (no longer just nodes and links, but also firewalls, proxies, underlying transport infrastructures, etc.). Real world data, theory, and modeling are needed to determine which elements are relevant how they should be abstracted to design and test such improved routing schemes.

The following matrix summarizes the problems that have been identified and the impact of potential solutions on them. A mark in a row-column slot indicates that row-indexed problem can benefit significantly from the column-indexed solution.

<b>Solutions→</b>	<b>Addressing Scheme</b>	<b>32bit to 48/64 bit</b>	<b>Routing Policies</b>	<b>Congestion Measures</b>	<b>Recovery Strategies</b>	<b>Economic Policies</b>
<b>PROBLEMS</b>						
Scalability	1	2	3			
TCP Inadequacies		4		5		
Core Vulnerability	6		7		8	
Services & QoS	9		10	11		
Structural Inflexibility	12		13		14	
ISP Control						15

To suggest the need for modeling and simulation we list the following “what if” design questions linked to the numbers in the table:

1. How would introducing greater address aggregation (specifically a discrete set of alternative cluster-based schemes) impact scalability, as measured for example, by the number of nodes that could be accommodated at the quality of service characteristic of the current Internet?
2. Clearly increasing the header size exponentially increases the address space and number of nodes that can be addressed. However, which size (48 or 64) would be sufficient to meet the growth of the next 20 years?
3. Are routing policies needed with any of the improved addressing schemes examined in 1? If so which alternative routing algorithms (such as vector-based versus link-based) are most effective in reducing hops and ensuring reliability?
4. How can the extra information in larger headers be exploited to track “a day in the life” of packets? What kinds of metrics and statistics can be collected? Is there a minimum number of bits required for the information to be useful (placing a second constraint on header size)? What algorithms are needed to process the information and are they implementable under infrastructure hardware constraints?
5. Which of a family of congestion measurement schemes are most effective in this regard? Which of these are practically implementable? Which of a family of sliding window control schemes works well together with the superior measurement schemes?

6,7, 8. Will some of the combinations of addressing schemes and routing policies just mentioned also reduce core vulnerability (as determined by a set of metrics in response to set of scenarios)? Which of a family of recovery policies (for example, based on mobile agent technology) works best with such combinations to reduce core vulnerability?

9,10,11. Likewise, will some of the combinations of addressing schemes and routing policies work together with congestion measurement/control algorithms to support a family of Quality of Service approaches for particular user requirements (such as video conferencing, grid-based computing, etc.)? How do these compare with alternative routing policies that are currently under consideration such as MPLS?

12,13,14. Which of the combinations of addressing schemes, routing policies, and recovery strategies can support reconfigurability and ability to accommodate new technologies. Encapsulation of addresses in subnetwork allows experimentation with new technologies to be confined within it. How does a new algorithm work in a subnet and how does it interact with others (e.g. current technology) in other subnets.

15. Which of a family of economic incentive schemes, consistent with minimal regulation, is able to increase willingness of ISPs to share data for the common good? Implement new globally optimal policies?

In the table below, we sketch characteristics of the models needed to address these questions. We note that such models are quite varied in nature, some are traditional network representations, some are not (e.g., complex adaptive systems to study economic incentives). Various levels of abstraction may be employed. Attempts have been made in the recent years to use abstractions to scale the order of simulations<sup>25</sup>.

We illustrate this by calling out the packet abstractions – for many questions not all the information in the packet needs to be employed. Choices of components and the detail in which they are modeled are also dependent on the questions being addressed. Component abstractions are likely to be correlated with packet abstractions, e.g., routers need only be able read destination addresses if packets only carry this information. Although models are quite varied and tuned to the questions they address, they should form a self-consistent family with ability to cross-calibrate their parameters and assumptions.

Model Attributes → Question(s) Addressed	Alternative Design or Study Choices	Input Generator/ Initial Configuration	Output Metric	Components	Packet Abstraction
1	Current flat vs various hierarchical cluster-based addressing schemes	Inter-BGP traffic	Utilization at BGP routers	BGP nodes router	Alive messages
2	Linear vs exponential growth	Demographics/ Technological Advancement/ Applications Growth	Utilization of address space	Nodes	N/A
3	Distance vector-based vs link-based	Source generator of packets	Number of Hops	BGP nodes + lower level nodes	Destination Address
4	48 vs 64 bit header	Source Generator + Load at Node	Amount of information about “day-in-life” of packet that can be gleaned	BGP + lower level nodes	Header substructure allocated to collecting “day-in-life” information
5	Alternative congestion measurement control schemes	Source Generator + Load at Node + hot-spot injection	Response time to clear congestion due to hot-spot injection	OSPF + Nodes	Latency Time, Query Type

<sup>25</sup> Huang, P., Estrin, D., Heidemann, D., “Enabling Large-Scale Simulations: Selective Abstraction Approach to the Study of Multicast Protocols”, International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, July 1998

6,7,8	Alternative addressing schemes and routing policies	Router Failure Injection	Convergence time to consistent routing tables	BGP routers nodes	Alive/Dead Messages
9,10,11	Alternative addressing schemes and routing policies	Source Generator + time varying stochastic loading	Probability of router predicting the actual path taken	Nodes and all hierarchy	Destination Address
12,13,14	Set of technology injection experiments	Source Generator + time varying stochastic loading + Alteration of subnet to represent technology injected	Restoration time to stable operation	Hierarchy and address prefixes	Multiple levels of abstraction
15	Alternative economic incentive schemes	Specification of complex adaptive system model	ISP sharing of data	Economic agents	N/A

We note that such questions are best addressed in a *holistic* manner, i.e., in a manner that will recognize that they must live together in the ULN space. By considering combinations, albeit with the concomitant growth in complexity, we recognize that individual optimizations of schemes, protocols, and algorithms are not likely to carry over to the shared space of operation. Put another way, the design of discrete, identifiable elements of ULNs (such as schemes, protocols, and algorithm) must take into account the interactions that will arise when embedded as parts of the overall whole.

As a consequence of the holistic manner in which ULN design must be addressed, a systems approach to design and concomitantly, the support of modeling and simulation, are critical to success. In the next section, we survey the range of existing network simulators with the ultimate goal of assessing their capability to support the holistic system design requirements just mentioned.

## Current Network Simulators: a Survey

In the last several years, several network simulators have come to a level of maturity to attract considerable use. More recently, simulators are being developed to address issues relevant to ultra-large networks. We survey a number of these simulators as representative of the current state of the art. A number of attributes are considered in order to provide a comprehensive view of the strengths and limitations of current simulator state of the art.

Simulator →	Opnet	Ns-2	Pdns	Glomosim	TeD
Attribute					
Focus	Commercially deployed user-network systems	Network systems under research study	Network systems under research study	Library-based sequential and parallel simulator for Wireless Ad-hoc networks	Tele-communication Networks (Logic of ns simulator being transformed – from event-driven to process-driven)
Description	Commercial simulator used by private users, network designers and architects to predict the feasibility of any deployable network. Works on the underlying process-models that work on finite state machine architecture. Provides extensive documentation	Both educational and research software. One of the first kinds on the league and designed to study the network dynamics of a TCP/IP network system. Many versions of ns are available and now being extended to parallel implementation to achieve scalability.	Extension to the ns simulator. Works in conjunction with Georgia's RTIKIT to process events in a correct timestamp (LBTS) order in the case of distributed simulation. The topology is modeled using (GT-ITM)	GloMoSim is a library written in Parsec for parallel simulation of wireless networks. Communication protocol stack is divided into a set of layers, each with its own API. Models of protocols at one particular layer interact with other layer through these APIs. They are in a process of developing hierarchical routing protocol to handle scaled network routing.	TeD (Telecommunications Descriptive Language) is a tool that brings automated Parallelization of network simulation by transforming its models into functionally equivalent GTW (Georgia Time-Warp). C++ classes in ns become entities in TeD but ns Packet class is implemented as an event. On the negative side, It acknowledges the cost of "state-saving" as the most serious among Time Warp overheads.
Object Orientation	C/C++	C++	C++	Extension to C language	C++
Network Devices	X (drag and drop)	Tcl -C++ scripting files	Tcl -C++ scripting files	Text based configuration files	Text files similar to ns
Network Topology	Static + node mobility	Static	Static	Static (Parsec library)	Static
Hierarchical	X			X and scalable	
Modular	X	X	X	X	X
Support for Protocol (TCP) Programming	X	X	X		
Parallel execution	Research is going on		X	Symmetric multiprocessor	Automated Parallelization
Runtime environment	Event loop	Event loop	RTIKIT (Process-driven)	Parsec Simulation Engine (DEVS approach)	Georgia Time-Warp Simulation engine
Memory Management	X	Dynamic allocation	Dynamic allocation	Tightly coupled Shared and Distributed	Dynamic allocation not required unlike ns
Simulation Resolution (Abstraction)	Packet-level	Packet-level	Packet-level	Packet level	Packet level
Scale of operation (nodes)	Few Hundred	Few Thousand	Hundred of thousands	Few Thousand	Tens of Thousand
Intended Audience	User, Network Designers	Research: protocol design and TCP variants	Network Research	Network research	Network research

Simulator →	SSF	USSF	Guts	MECA	DEVS-DOC
Attribute					
Focus	Large-scale networks described in Domain Modeling Language (DML)	To Simulate complex models with over 1 million components	To simulate Back-bone area networks (WAN)	Multi-computer simulator towards large-scale simulations	DEVS approach to Distributed object Computing
Description	SSF is a discrete event modeling API designed for very large networks and can execute a million or more concurrent TCP/IP flows. SSFNET models are self-configuring and configuring data is hierarchically structured. SSF architecture has just five generic primary classes.	USSF is a framework that runs as an application on an underlying parallel kernel and utilizing its services. The kernels included the WARPED based on optimistic PDES and NOTIME (an unsynchronized PDES kernel). RTEL was developed to reduce the static size of the application modules and in turn its static memory requirements. Its an ideal candidate for simulating large applications that contain LPs of common description.	A high-level wide-area network simulator designed to enable simulation of Internet-scale topologies under a range of realistic work-loads. No network-queues are used in the model and the model allocates the bandwidth to a particular transfer only once and it remains fixed during the transmission of that transfer-block. The simulator tries to simulate by using bulk-transfers as traffic composed in the backbone is of bursty nature.	MECA (multi-path E-cube Algorithms) is a multi-computer simulator with good reconfigurability. The command line switches allow changes without recompiling. The kernel is just 250 lines of code. The negative point being tight coupling of algorithms and data-objects.	This work demonstrates that a high-level representation of networking and computing technologies and components can support distributed hardware architectures.
Object Orientation	Java and C++	C++	X	C++	Java
Network Devices	Object classes in DML	Abstraction into a network component (LP)	Nodes and links	Nodes and Links	Abstraction to various network components
Network Topology	DML (similar to java packages)	Topology Specification language (Runtime Elaboration Language RTEL)	Directed graph	k-ary n-cube wormhole network	Component modeling, static topology
Hierarchical		X			X
Modular	X	X	X	X	X
Support for Protocol (TCP) Programming					
Parallel execution	Symmetric multiprocessor	Network of Dual-Pentiums in and Ethernet network		X	X
Runtime environment	Periodic time-stepped approach	WARPED and NOTIME Simulation Engine	Event loop	Event loop	Event loop
Memory Management	Tightly-coupled shared memory and distributed-memory cluster	Distributed synchronization	Better than ns	X	Java Virtual Machine
Simulation Resolution (Abstraction)	IP Packet	LP (message passing)	No network Queues-transfer level burst messages	Constant length messages	DEVS messages/jobs
Scale of operation (nodes)	Few Hundred Thousand	Few Hundred Thousand	Few hundred nodes	Few Hundred nodes	Within hundred nodes but can scale upto ten thousand nodes
Intended Audience	Network research towards routing protocol simulation Network	Network research including VLSI domain	Research towards coarser grained network models	Network Research towards adaptive routing and broadcasting	Designers of Distributed Collaborative Systems

## Legend:

Attribute	Description
Focus	This describes the basic objective of the simulator designed
Description	General description of the simulator
Object-Orientation	Describes the architecture of the simulator as object-oriented. Languages are usually c/c++/Java
Network Devices	This includes the routers, nodes, links and other entities in a typical network
Network Topology	Describes the overall topology of the network system under study. It may be a directed graph, a static topology, a dynamically created topology or a worm-hole network
Hierarchical	This describes the underlying architecture of the model is component-based and various component levels are defined and model-under-study is constructed using these components. These components at different hierarchies communicate to other components using closure-under-coupling.
Modular	This defines the architecture to be modular in nature and be able to communicate through the interfaces. This defines the boundaries of various entities and their scope of communication.
Validation	This describes whether the results and experiments conducted on the simulator are being validated or not.
Support for Protocol (TCP) programming	This attribute specifies the functionality of the simulator in terms of testing variations in the TCP protocol and improvising it.
Parallel Execution	This explains if the simulator execution can be done in parallel distributive environment
Run-time Environment	Describes the run-time environment of the simulator. It can be either event-driven or process-driven. Its also depicts the simulation engine involved.
Memory Management	Describes the memory management of the simulator. It may be shared, distributed or tightly coupled.
Simulation Level (abstraction)	Defines the lowest abstraction level of simulation. It may be packet-level, transfer(message) level or nodal.
Scale of operation (No. of nodes)	Indicates the maximum number of nodes that can be simulated by the simulator
Intended Audience	To whom the simulator is designed for.

## General Terms used in the table above:

GT-ITM	Georgia Tech – Internet Topology Modeler
RTIKIT	Real Time I.... KIT developed by Georgia Tech to facilitate correct time stamped event execution
LBTS	Lower Bound Time Stamp
Symmetric Multiprocessor	Topology or graph partitioned symmetrically with same number of nodes to each processor
Automated Parallelization	Ability to do the partitioning and assigning the processors is already there in the Georgia Time-Warp Simulation Engine and nothing has to be programmed regarding Parallelization
DML	Domain Modeling Language. A set of java class packages built as a part of APIs for SSF simulator framework. The entire topology is constructed using the specification language.
LP	Logical processes implemented as components representing the different physical processes being modeled. The LPs exchange event information by exchanging virtual time stamped event messages.
WARPED	It is a kernel that presents an interface (API) to build LPs with unique definitions of state. LPs are placed in groups called “clusters”. LPs on same cluster communicate with each other with the intervention of Message Passing System (MPI).
NOTIME	It is unsynchronized PDES (Parallel Discrete Event Simulation) kernel. It mirrors the APIs used by WARPED. LPs are grouped into clusters. Processor level parallelism occurs at the cluster level and each cluster is responsible for communications management and scheduling of the LPs contained in that cluster. NOTIME utilizes MPI for communication.
RTEL	Runtime Elaboration Library. Runtime elaboration provides a tradeoff between the size of the generated code, the compile time of the generated code, and overall simulation time.

As is clear from the table, existing simulators address a range of problems for various intended users. OPNET, the most widely employed commercial simulation environment, is aimed at the problems involved in configuring networks using off-the-shelf components so as to achieve desired performance goals. The focus is on how a particular set of requirements can best be realized within the given structure of the Internet. Thus, OPNET does not support holistic (re)design of a ULN as such, although some of its capabilities can support some of the work needed to do so. For

example, OPNET provides a large library of models of components such as routers and processing “stack” protocols. However, it offers only limited support for expressing and integrating novel models – for components and more critically for architectures – into the library, as would be appropriate for investigating new concepts for ULN design.

NS, is an academic software, developed to enable investigation of TCP/IP protocol variants within realistic network settings. Thus, NS’s underlying objectives are more congruent with ULN needs than is OPNET, but NS also suffers some major limitations in this regard – it’s focus on protocol design problems is not necessarily supportive of other ULN areas of concern such as those raised in the last section. Similar comments apply to simulators that are intended to address particular kinds of networks (such as TeD) or particular problems, of which there are numerous examples in the literature. Some of the recent simulators that are still under research, such as USSF and Guts, address issues of scalability and aim to handle a large number of nodes. The architectural design of OPNET, USSF and DEVS-DOC support scalability through their hierarchical structure. However, only DEVS-DOC is built on the concept of hierarchical, modular model construction supported by the DEVS formalism.

None of the current simulators is complete in addressing the central issues in understanding ULN and supporting new architectural designs. The Internet is an always-on system. As a result, any insertion of new technology or change in operating rules should first be tested in a modeling and simulation environment that can emulate and correctly predict how the new technology will behave after being deployed. To adequately support ULN intervention and design studies, the simulation model should be able to address three primary issues: support for well-defined abstractions, variable-structure, and scalability using hierarchical control structure concepts:

- Since we don’t have a complete theory regarding the behavior of the Internet, the models that underlie current simulators are usually built on purely analytical studies and the resolution is taken down to the lowest relevant level, i.e., the packet level. But to address the ULN issues, the resolution level has to be adjustable to the questions being addressed and there should be a set of clearly defined abstraction levels to choose from.
- A common trait of current simulators is that they deal with a static network topology which remains fixed during the course of the simulation. They are not equipped to study variable-node-network systems such as the BGP systems subjected to malicious manipulation (periodic turning ON-OFF) of one node can prove disastrous.
- Scalability is major concern in the application of widely used simulators, such as OPNET and NS. To address the scalability problem, the model architecture has to be hierarchical so that abstraction is well developed and cross-coupling

between different levels can be done efficiently. The majority of current simulation models are devoid of the hierarchical architecture need to study the behavior of scalable systems. Although today's simulator developers are trying to address the scalability issue, their flat architectural design limits the kind of scalability that can be achieved.

In the next section, we suggest an iterative system design approach framework that can guide the development of appropriate modeling and simulation tools for ULN. This will provide the basis for us to recommend the research needed for new directions in modeling and simulation research for ULN.

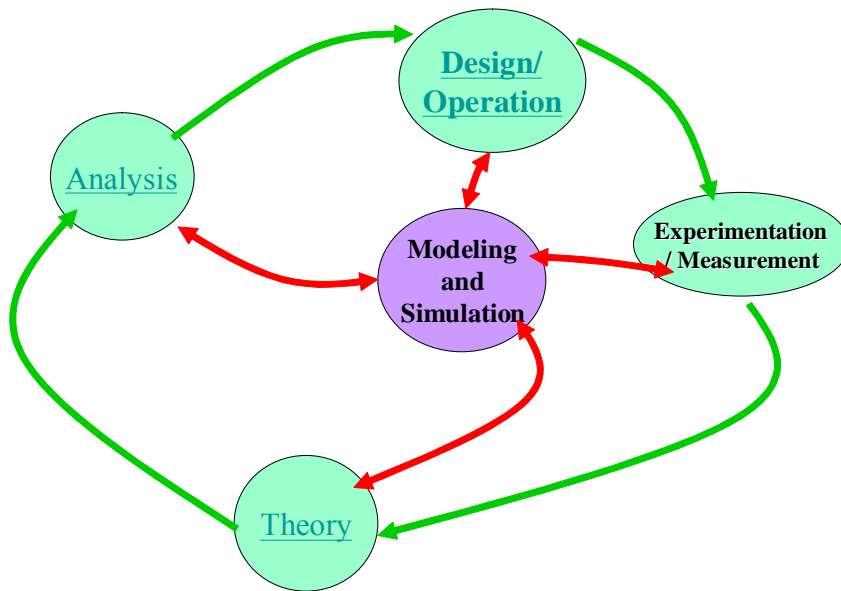
## **4. Research needed to enable M&S to meet the challenge**

As mentioned, the NRC report, *“Looking Over the Fence at Networks – A Neighbor’s view of Networking Research”* advocated new methodologies for design of network infrastructure elements and infusion into practice of such designs. In this report, we expand on this theme by advocating a disciplined design approach to future ULN that will afford such desired systems properties as greater quality of service, reliability, availability, and dependability. As our examples suggest, to meet such goals will require greatly improved design tools and these will necessarily be based on significant advances in the area of modeling and simulation.

System engineering and software development methodologies have moved away from earlier, simplistic “waterfall” concepts, toward more iterative or spiral design approaches. Indeed, the Internet, as we know it today, is the result of an ad-hoc evolutionary process rather than a rational, omniscient systems engineering approach. This was appropriate, given the absence of knowledge about the workings of the Internet, a new domain of human experience. However, with several decades of experience under our belts, it is now appropriate, and critical, to place Internet and future ULN development within an iterative design cycle more in conformance with current systems and software engineering views.

### ***Modeling and simulation at the center an iterative design cycle***

The cycle illustrated in the following figure portrays an iterative software engineering cycle applicable to ULN development. More specifically, it includes not only traditional elements such as analysis and design, but also explicitly recognizes the ongoing need for experimentation and new theory development to support the traditional elements. It also places modeling and simulation in the center of all the other activities with arcs to suggest its interactions with these activities. Indeed, our fundamental thesis is that it crucial that M&S play such a central role in future ULN design. While only M&S can provide the basis for effective coordination of all other activities, its current capability to do so is not adequate for the task. Indeed, the goal of this report to recommend new research directions for M&S to greatly expand its capacity to support the iterative design cycle. Before considering these recommendations in more depth, we will say a few words about the elements in the cycle.



### **Design/Operation**

Software engineering promotes systematic, disciplined, and quantifiable approaches to the development, operation, and maintenance of software-intensive systems. As a software intensive system, ULN development will require this kind of disciplined approach, but as an evolving phenomenon, ULN will require incorporating the following as well:

### **Theory**

Understanding ULN behavior will require theories that suggest new higherlevel explanatory abstractions than traditional packet-level modeling. Such models are unlikely to be analytically tractable, so that the ability to simulate large-scale network models will be critical in developing valid abstraction and predictive capabilities.

### **Experiment/Measurement**

Relevant metrics and measurement tools to acquire them on a continuous and intensive basis are required to inform theory, modeling, analysis, and ultimately design. Experimentation, with some ability to control key variables, is needed to obtain measurements under conditions that can be replicated leading to conclusions that can be generalized. Moreover, aside from learning from such data acquisition in an offline manner, we can employ online measurement to advantage. Specifically, such measurement can be automated, coupled with real-time modeling and embedded into

network operation for improved management and control. Indeed, as the Internet scales up, with more and more high-speed links, overcoming today's measurement limitations is needed to create a situation where we can take control of the bandwidth being created.

### **Analysis**

Traditionally mathematical analysis, such as queueing theory, has helped design network elements. However, traditional approaches do not scale well with the increase in size, heterogeneity, and complexity characteristic of ULN. Still, analysis can provide a global view of system dynamics that simulation can only sample. So what will be needed are new forms of analysis that work in concert with simulation models, with mutual parametric cross calibration and able to provide qualitative hints about their behavior.

### **Modeling and Simulation**

ULN systems and software engineering will require development of models that support theory development, analysis, design and operation and that are calibrated through experiment and measurement. As already mentioned, such models are, in general, going to be intractable to direct solution and will require efficient and correct simulation techniques.

## ***Recommendations for research directions***

Research is needed to enable M&S to play the key-coordinating role within the iterative design cycle. More explicitly, we consider the research needed to enable M&S to support each of the phases in the cycle, as follows:

### **Theory**

Research is needed to enable M&S to provide the means to test theories of ULN behavior and to stimulate their development. Such theories may require introduction of new mathematical concepts as well as borrowing of existing concepts from existing complex systems.

### **Experiment**

Research is needed to link development of models with experiments that can be done on the Internet itself or within testbeds developed for such purposes. Models should help translate general theories of ULN phenomena into concretely testable hypotheses enabling simulations to generate behavior comparable with experiment.

## **Measurement**

Research is needed to enable M&S to support the development of measurement techniques -- whether employed online or for research -- that provide veridical and to-the-point information on the state of the network.

## **Analysis**

Research is needed to enable M&S to work seamlessly traditional analysis techniques so that for example, mathematically tractable abstractions of simulation models can be generated automatically.

## **Design/Operation**

Research is needed to enable new design proposals and/or system interventions to be developed and tested in “virtual” testbeds -- realistic characterizations of the real Internet environment of today or the ULN of tomorrow. Such characterizations will require development of new levels of abstraction and new models within such levels to be able to effectively and efficiently address the questions raised by proposed designs and interventions. Research is needed to enable families of models at various levels of abstraction to be constructed, cross-validated against realistic Internet data, and placed in common repository of models accessible to all researchers. Research is also needed to enable efficient simulation of models using single processor, parallel and/or distributed environments and the widespread dissemination of these simulators.

The following diagram relates the foregoing recommendations to more detailed findings and recommendations that emerged from the Tucson Workshop. The latter are presented in detail in the Appendix.

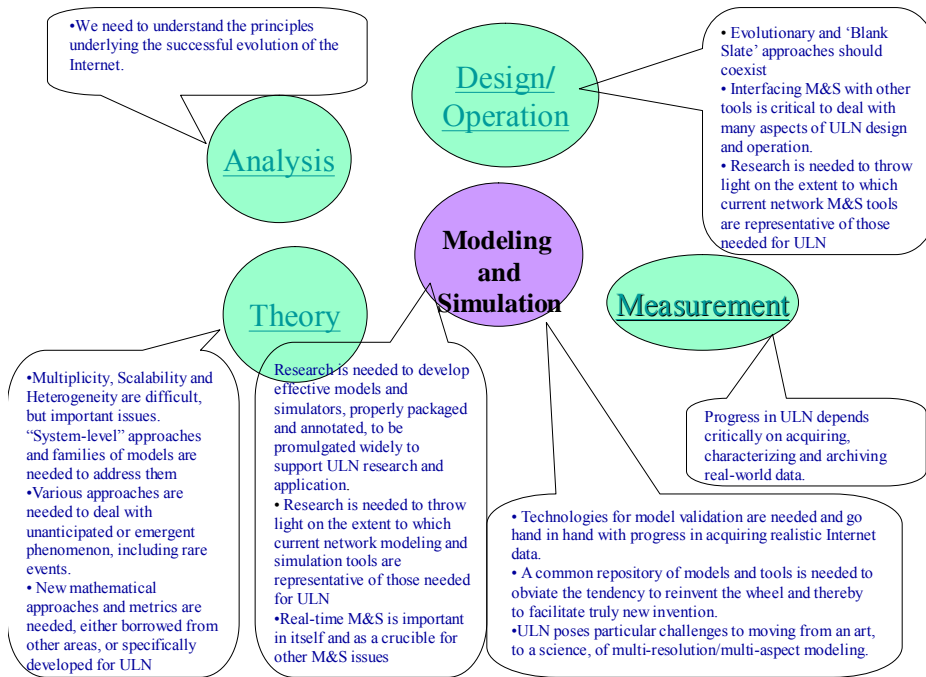


Figure: ULN Recommendations as in ULN Workshop linking to Iterative Design Cycle

## References:

### ***Some things we can change***

#### **1. Implement new addressing scheme**

Krishnamurthy, B. and Wang J., "Topology Modeling via Cluster Graphs", ACM SIGCOMM Internet Measurement Workshop, October 2001

Zhao, X., Pei, D., Wang, L., Massey, D., Mankin, A., Wu, S. F., and Zhang, L., "An Analysis of BGP Multiple Origin AS (MOAS) Conflicts", ACM SIGCOMM Internet Measurement Workshop 2001, November 1-2, 2001

#### **2. Increase the packet header size from 32 bit to 48 or 64 bit**

Hierarchical Routing in IPV6

[http://www.luv.asn.au/overheads/ipv6/ipv6\\_route.html](http://www.luv.asn.au/overheads/ipv6/ipv6_route.html)

IPV6 adds reliability features to the Net

<http://www.ipv6.ru/docs/ebsco/fulltext16.htm>

#### **3. Revise routing policies and algorithms**

Perkins, C. E., "Ad-Hoc Networking", Addison Wesley, Chapter 4

Tangmunarunkit, H., Govindan, R., Estrin, D., and Shenker, S., "The Impact of Routing Policy on Internet Paths", In *Proceedings of 20th IEEE INFOCOM*, April 2001.

#### **4. Adopt new ways of measuring traffic congestion and control policies**

Cao, J., Cleveland, W. S., Lin, D., and Sun, D. X., "The Effect of Statistical Multiplexing on the Long Range Dependence of Internet Packet Traffic", Bell Labs Tech Report, 2002

Krishnamurthy, B., Wang, J. and Xie, Y., "Early Measurements of a Cluster-based Architecture for P2P Systems", ACM SIGCOMM Internet Measurement Workshop 2001, November 1-2, 2001

#### **5. Devise new strategies to recover from crashes**

"Overview of Strengthening Homeland Cyber-Defense", October 18, 2001

[www.ita.org/infosec/summary.pdf](http://www.ita.org/infosec/summary.pdf)

## **Theory**

Aggarwal, S., Courcoubetis, C., and Wolper, P., "Adding Liveness Properties to Coupled Finite State Machines," *ACM Trans. on Programming Languages and Systems*, Vol. 12-2, April 1990, pp. 303-339.

Aggarwal, S., Paul, S., Massey, D. and Calderaru, D., "A Flexible Multicast Routing Protocol for Group Communication," *Computer Networks*, Volume 32, Issue 1, January 2000, pp. 35-60.

Davis, P.C., Fishwick, P.A., Overstreet, C.M., Pegden, C.D., "Model Composability as a Research Investment: Responses to The Featured Paper", *Winter Simulation Conference*, pp. 1585-1591, 2000

Morihama, L., Pasuello, V., Wainer, W., "Validating the behavior of DEVS models", Accepted for publication in *SISO Spring Interoperability Workshop*, Orlando, FL. U.S.A. 2002.

Glinsky, E., Wainer, G., "Analyzing overhead of Real-time DEVS models", accepted for publication in *AIS'2002*. Lisbon, Portugal. 2002.

Joo, S.S., Chung, Y. S., Razouqi, Q., Ghosh, S., "Directions in IP Router Architecture Design And its Validation," 2001, *IEEE International Conference on Software Telecommunications and Computer Networks (SOFTCOM)*, Split, Dubrovnik (Croatia), October 09-12, 2001, pp. 495-503.

Kasputis, S., Ng, N., "Composable Simulations", *Winter Simulation Conference*, pp. 1577-1584, 2000

Riley, G. F. and Fujimoto, R. M., and Ammar, M.H., "A Generic Framework for Parallelization of Network Simulations", *Proceedings of Seventh International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS'99)*, October, 1999

Riley, G. F. and Fujimoto, R. M., and Ammar, M.H., "Stateless Routing in Network Simulations", *Proceedings of the Eighth International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, August, 2000

Zeigler, B.P., Praehofer, H., Kim, T. G., *Theory of Modeling and Simulation: Integrating Discrete event and Continuous Complex Dynamic Systems*, Second edition, Academic Press, 2000.

## **Design/Operation**

Goguen, J., "An Introduction to Algebraic Semiotics,with Applications to User Interface Design ", in *Computation for Metaphor, Analogy and Agents*, ed.Chrystopher Nehaniv, springer LNCS,v.1562,pp.242-291,1999.

Kampis, G., "Self-Modifying Systems", Pergamon,Oxford 1991

Kleinberg, J. M., "Navigation in a Small World ", *Nature* v. 406, 2000

Lindgren, K., "Evolutionary Phenomena in Simple Dynamics ", In Artificial Life II Langton et al. (Eds), Addison-Wesley, 1991

Razouqi, Q., Ghosh, S., "Incremental, Dynamic, Virtual Circuit Connection (IVCC): A New Paradigm for Routing in Future High-Speed Networks," Accepted in IEEE International Conference on Communications, June 11-15, 2001, Helsinki, Finland, pp. 2578-2582.

Schulz, S., Rozenblit, J.W., Mrva, M. and Buchenrieder, K., "Model-Based Codesign". IEEE Computer, 32(8), 60-68, 1998.

## ***Analysis***

Adamic, L., and Huberman, B., "The Web 's Hidden Order ", CACM 44:9,2001

Chapman, W., Rozenblit, J.W. and Bahill, A. T., "System Design is an NP-Hard Process", Systems Engineering Journal, 4(3), 222-229, 2001.

Floyd, S. and Paxson, V., "Difficulties in Simulating the Internet", February 2001. IEEE/ACM Transactions on Networking

Paxson, V. and Floyd, S., "Wide-Area Traffic: The Failure of Poisson Modeling", IEEE/ACM Transactions on Networking, Vol. 3, No. 3, pp. 226-244, June 1995

Pillalamarri, S., and Ghosh, S., "The Impact of Source Traffic Distribution on Quality of Service (QoS) in ATM Networks." Accepted in IEEE International Conference on Communications, June 11-15, 2001, Helsinki, Finland, pp. 2855-2859.

## ***Experimentation/Modeling***

Callaway, D. S. et al., "Are Randomly Grown Graphs, Really Random?", Phys. Rev. E 64,041902,2001

Clark, G., Courtney, T., Daly, D., Deavours, D., Derisavi, S., Doyle, J. M., Sanders, W. H., and Webster, P., "[The Möbius Modeling Tool](#)", *Proceedings of the 9th International Workshop on Petri Nets and Performance Models*, Aachen, Germany, September 11-14, 2001, pp. 241-250.

Christen, G., Dobniewski, A., Wainer, G., "Defining DEVS models with the CD++ toolkit", In *Proceedings of European Simulation Symposium*. Marseille, France.

Deavours, D., Daly, D., and Sanders, W. H., "[Möbius: Framework and Atomic Models](#)", *Proceedings of the 9th International Workshop on Petri Nets and Performance Models*, Aachen, Germany, September 11-14, 2001, pp. 251-260

Ghose, K., Aggarwal, S., Vasek, P., Chandra, S., Raghav, A., Ghosh, A., and Vogel, D., "ASSERTS: A Toolkit for Real-time Software Design, Development and Evaluation", 9th EuroMicro Workshop on Real-Time Systems, June 1997, pp. 224-232.

Hild, D., Sarjoughian, H.S., Zeigler, B. P., "DEVS-DOC: A Modeling and Simulation Environment Enabling Distributed Codesign," IEEE SMC Transactions, Vol. 32, No.1, pp. 78-92.

Newman, M.E., "Models of the Small World ", J. Stat. Phys.101, 819-841,2000

Rozenblit, J.W., "Systems Design: A Simulation-Based Modeling Framework", in Discrete Event Modeling and Simulation: A Tapestry of Systems and AI-based Theories and Methodologies. (Eds.F. Cellier and H. Sarjoughian), Springer Verlag, pp. 107-127, 2001.

Sanders, W. H., "[Integrated Frameworks for Multi-Level and Multi-Formalism Modeling](#)", *Proceedings of PNPM'99: 8th International Workshop on Petri Nets and Performance Models*, Zaragoza, Spain, September 8-10, 1999, pp. 2-9.

Stewart , T., "Six Degrees of Mohamed Atta ",Business 2.0,  
<http://www.business2.com/articles/mag/print/0,1643,35253,FF.html>

## **Measurement**

Fitzek, H.P., Reisslein, M., "MPEG-4 and H.263 Video Traces for Network Performance Evaluation". IEEE Network, 15(6): 40-54, November/December 2001

Riley, G. F., and Ammar, H. M., "Simulating Large Networks: How Big is Big Enough?", Proceedings of First International Conference on Grand Challenges for Modeling and Simulation, January 2002

## **Current Simulators: A Survey**

Bertolotti, S., and Dunand, L., "Opnet 2.4: An Environment for communication network modeling and simulation," in *Proceedings of the European Simulation Symposium*, October 1993.

Bhatt S., Fujimoto R., Ogleiski A., and Perumella K., "Parallel Simulation Techniques", IEEE Communications Magazine 1998, pp. 42-48

Coller, J., Drapper, J., Gorman, M., "ASNT Network Simulation Tools, Survey and Requirements", November 1995

Cowie, J., Nikol, D. M., and Ogeilski, A., "Modeling the Global Internet", Computing and Science and Engineering, 30-36, January 1999

Cowie, J., Nikol, D. M., and Ogeilski, A., H. Liu, J. Liu, "Towards realistic million node Internet Simulations," in International Conference on Parallel and Distributed Processing Techniques and Applications, June 1999

Gadde, S., Chase, J., Vahdat, A., "Coarse-Grained Network Simulation for Wide-Area Distributed Systems," Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDS 2002)., January 2002.

Hild, D., Sarjoughian, H.S., B.P. Zeigler, "DEVS-DOC: A Modeling and Simulation Environment Enabling Distributed Codesign," IEEE SMC Transactions, Vol. 32, No.1, pp. 78-92.

McCanne, S., and Floyd, S., "The LBNL network Simulator.", Software on-line <http://www.isi.edu/nsnam>, 1997, Lawrence Berkeley Laboratory

Perumalla, K., Fujimoto, R., "Efficient Large Scale Process Oriented Parallel Simulations" in *Proceedings of the 1998 Winter Simulation Conference*, December 1998

Premore, B. J., Nikol, D. M., "Parallel Simulation of TCP/IP using TeD", *Proceedings of the 1997 Winter Simulation Conference*, pages 437-443, Atlanta, December 1997.

Perumalla, K., Fujimoto, R., Ogielski, a., "TeD- A Language for modeling Telecommunication Networks," Performance Evaluation Review, Vol. 25, No. 4, 1998, pp. 4-11

Rao, D. M., and Wilsey, P.A., " An Ultra-large Scale Simulation Framework", *Journal of Parallel and Distributed Computing*, Volume 10, Number 1, 18-38, January 2000

Rao, D. M., and Wilsey, P.A., "Simulation of Ultralarge Communication Networks", Intl. Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, MASCOTS'99

Riley, G. F. and Fujimoto, R. M., and Ammar, M.H.,, "Stateless Routing in Network Simulations", in *Proceedings of the Eighth International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication systems*, August 2000

Riley, G. F. and Fujimoto, R. M., and Ammar, M.H.,, "A Generic framework of parallelizations for Network Simulations, " in *Proceedings of Seventh International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems. (MASCOTS' 99)*, October 1999.

Zeng, X., Bagrodia, R., and Gerla, M., "Glomosim: A library of Parallel Simulation of Large Scale Wireless Networks," *In proceedings of the 12<sup>th</sup> Workshop on Parallel and Distributed Simulations*, May 1998

Wu, H., Fujimoto, R., Riley, G., "Experiences Parallelizing a Commercial Simulator", *In Proceedings of the 2001 Winter Simulation Conference*

## **Links**

<http://ww.acims.arizona.edu>

<http://www.ists.dartmouth.edu/cybersecurity/index.htm>