A Framework for a Bandwidth Based Network Performance Model for CS Students

D. Veal, G. Kohli, S. P. Maj

Edith Cowan University
Western Australia

J. Cooper

Curtin University of Technology
Western Australia

d.veal@ecu.edu.au

Abstract

There are currently various methods by which network and internetwork performance can be addressed. Examples include simulation modeling and analytical modeling which often results in models that are highly complex and often mathematically based (e.g. queuing theory). The authors have developed a new model which is based upon simple formulae derived from an investigation of computer networks. This model provides a conceptual framework for performance analysis based upon bandwidth considerations from a constructivist perspective. The bandwidth centric model uses high level abstraction decoupled from the implementation details of the underlying technology. This paper represents an initial attempt to develop this theory further in the field of networking education and presents a description of some of our work to date. This paper also includes details of experiments undertaken to measure bandwidth and formulae derived and applied to the investigation of converging data streams.

Introduction

There are many ways of approaching computer networking education. Davies notes that: *"Network courses are often based on one or more of the following areas: The OSI model; Performance analysis; and Network simulation"* [1]. The OSI model is a popular approach that is used extensively in the Cisco Networking Academy Program (CNAP) [2] and in other Cisco learning materials. With respect to simulation Davis describes the Optimized Network Engineering Tools (OPNET) system that that can model networks and sub-networks, individual nodes and stations and state transition models that defines a node [1]. However, Davies gives no indication as to the accuracy of this simulation or of the limits of its application. The development and testing of networking simulations may depend upon student's possessing knowledge and having suitable experience in computer programming languages that many networking students might not possess.

Performance analysis can use analytical based models that are often specialized in their area of application and may involve the use of complex mathematics which could be problematic for many computer networking students. Its advantages include the use of powerful mathematical tools. However, these models may not be based upon real systems. Queuing is a form of analytical modeling [3]. Networks of queues can also be analyzed [4, 5]. However, queuing analysis often assumes a Poisson distribution, which is not the case with most networks [6,7]. Performance analysis in computer networking can be based upon various models; and bandwidth in MB/s is a common performance indicator. Students can perceive differences in performance with respect to bandwidth which can be used as a starting point for a constructivist based performance model.

Constructivism

The dominant theory of conceptual understanding in education is constructivism. This has been extensively tested in the field of science and mathematics education [8-11]. Although Constructivism is a foundation of many modern teaching practices it has not been influential within computer education [12]. The knowledge the learner has already constructed will affect how new knowledge is interpreted [13], *"Put in the simplest way, to understand what someone has said or written means no less but also no more than to have built up a conceptual structure that, in a given context, appears to be 'compatible' with the structure the speaker had in mind. – and this compatibility, as a rule, manifests itself in no other way than that the receiver says and does nothing that contravenes the speaker's expectations"*[14]. The importance of conceptual understanding in the field of computer science education has been highlighted by the ACM who note that: *"Concepts are the raw material for understanding new information technology as it evolves"* [15]. Hence there is a need to develop a performance based model in computer networking curricula that has its foundation based upon students' common conceptual understanding. The B-Node model provides such a foundation based upon the students' own perceptions of computer network performance [16].

A Bandwidth Model

B-Node models are bandwidth centric high level abstractions which are independent of the underlying implementation details of a particular technology [16]. The rapid rate of technological change within the field of Computer and Networking Technology (CNT) has given rise to the need to future-proof student learning. This is achieved by omitting details of the underlying technological implementation, which may change rapidly as the technology progresses [17,18]. Furthermore additional advantages of a bandwidth centric approach are that bandwidths can be readily measured in computer networks and bandwidths are often included in network equipment specifications. As B-Nodes are high level abstract models they allow the possibility of recursive decomposition into their lower level component parts [16].

Potential Problems with a Bandwidth Centric Approach

A bandwidth centric approach needs to address the problem of latency. Hennessy and Patterson, under a heading of *"Fallacies and Pitfalls"* mention the: *"... pitfall of using bandwidth as the only measure of network performance. ... this may be true for some applications such as video, where there is little interaction between the sender and the receiver, but for many applications such as NFS, are of a request-response* *nature, and so for every large massage there must be one or more small messages … latency is as important as bandwidth"* [19].

In this paper latency is addressed by subsuming its effects under a heading of bandwidth. Bandwidth is defined as being equal to the size of the file transmitted from source to destination, divided by the time taken to send that file. Latency is simply included in the time to transfer the file. Networks and computers may produce more bandwidth and yet result in a lower performance when compared to another system that produces better performance with less bandwidth [20]. Yet any measure of performance has its drawbacks and may not yield meaningful results in all circumstances.

Deterministic B-Node Models

The authors chose to concentrate upon an initial deterministic model rather than a stochastic model. Lai and Baker note that: *"Deterministic models are typically easier to work with mathematically than stochastic models, enabling us to find an analytical solution rather than a numerical one. Unfortunately, a deterministic model implies modeling with absolute certainty and many things cannot be known with enough certainty to make this practical"* [21]. However, a deterministic model can allow the checking of the basic parameters upon which to base a more complex stochastically based model. These initial parameters are to be determined by experimental results.

The Experiments

As an approach to these difficulties of modeling network performance, the authors have designed a solution that has analytical components and yet is both relatively simple to use and quantitatively based. These investigations commenced with experiments using PING to determine a baseline for subsequent bandwidth measurements. Table 1 shows the shorthand used for describing a given chain of networked of devices.

Device chain	Abbreviations
PC to fiber link to PC	PFP
PC to PC	PP
PC to switch to PC	PSP
PC to Router to PC	PRP
PC to Switch to two Routers to Switch to PC	PS2RSP
PC to Switch to three Routers to Switch to PC	PS3RSP

Table 1: Device chain abbreviations

Extended PING is a program used to check connectivity at the first three layers of the OSI model. This was used under Windows XP to measure the latency between a source and a destination. The version of PING used under Windows XP only gives a resolution of 1 ms and is of unknown accuracy. To address this problem we used a baseline of PCs connected by a single crossover cable. The results are shown in Figure 1 which shows the size of the data transferred against the transfer times. This also provides a measure of the bandwidth as 5.48 MB/s. It should be noted that these results are based upon binary Megabytes where as 100 mbps is denary based. The actual bandwidth is twice this value, approximately 11 MB/s, due to Round Trip Time (RTT). Multiple values were taken to obtain a better approximation of bandwidth. The slope for respective devices gives bandwidths which are tabulated in table 2.

Devices	Line plotted on graph	Bandwidth (MB/s) (Slope x 2)
PP	A	11.1
PSP	B	11.4
PSRSP	C	7.9
PS2RSP	D	7.5
PS3RSP	E	7.3

Table 2: Legend for PING transfer time graph

It can be seen from Figure 1 that when a switch replaces a crossover cable as the link between two PCs this introduces a latency of much less then 1 ms and the link operates approximately at line speed line speed (approximately 12 MB/s) as the slopes of lines 'a' and 'b' are approximately equal. The latency introduced by the addition of a switch is negligible at this level, as shown by the horizontal differences of points on the lines of 'a' and 'b'. The latency of adding extra routers and is shown by the time between lines 'c' and 'd' and lines 'd' and 'e' which is approximately 1ms.

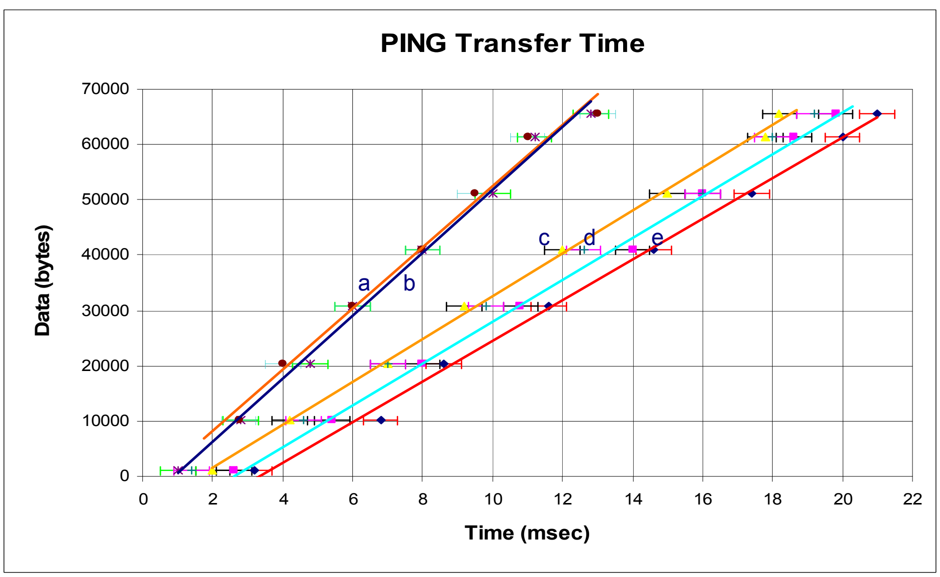

Figure 1: Data verses PING transfer time

The bandwidth drops from 11.4 MB/s to approximately 7.5 MB/s as shown by the difference in twice the slope of appropriate lines. It should be noted that fast switching was enabled by default on all routers and switches [25] .

FTP Transfer Time Experiments

Further experiments were conducted by transmitting across a series of devices using File Transfer Protocol (FTP). This was done to investigate larger file transfers, since PING can only send data up to 64KB. Figure 2 shows the bandwidths obtained using different configurations an example of which shown in Figure 3. This showed that the introduction of a router in a series of (B-Nodes) decreased the bandwidth to 7.5 MB/s.

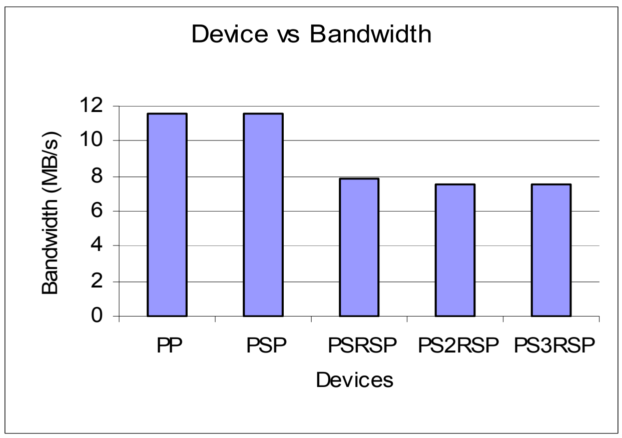

Figure 3: Device verses bandwidth for FTP transfers

Negative Bandwidth Experiments

In order to model the reduction in bandwidth due to the addition processing on devices (e.g. a router) the authors measured the decrease in observed bandwidth when processing is taking place compared to the bandwidth when this processing is not taking place, and they defined this as 'negative bandwidth'.

To investigate further the effect of processing in the router upon the bandwidth of traffic flowing through it, Access Control Lists (ACLs) were used to provide a measure of 'negative bandwidth'. ACLs can adversely affect router performance [22]. Significantly in product documentation the authors could not find where the effects of ACLs were quantified even though such a measure can be useful. To obtain such a measure experimentally, extended ACLs of varying numbers of Access Control Entries (ACEs) were applied to router Fast Ethernet interfaces and effects upon bandwidth noted. The results are shown in Figure 4. It can be concluded that, with the equipment used, the decrease in bandwidth is approximately 32kB/s per additional ACE as indicated by the slopes. (The dashed line is represents a file size of 150MB and the solid line indicates a file size of 15MB).

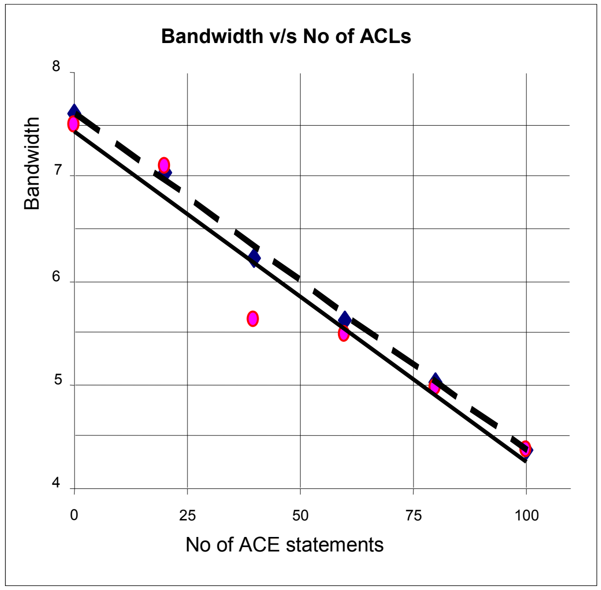

Figure 4: Bandwidth verses number of ACE statements

B-Node Model Requirements

The B-Node model would need to handle both series and parallel configurations. These are represented as shown in Figures 5 and 6. Given a chain of devices in series, the resultant bandwidth is the minimum individual bandwidth of given set of individual bandwidths of each of these devices:

$B(Resultant) = Min \{B_1 , B_2 , B_3 , \ldots\}$.

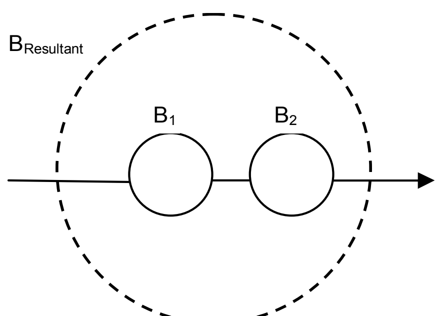

Figure 5: B-Nodes in series

This minimum result was also noted by Jain and Dovrolis [23].

ACL's were applied to a series of routers and the major effect on the overall source to destination bandwidth was dominated by the router with the maximum number of ACLs statements. Furthermore, this principle can also be seen to apply when routers were introduced into the chain the bandwidth dropped from 11.5 MB/s to 7.5 MB/s.

As well as routers in series, data can also flow via parallel routes through an internetwork as show in Figure 6. The division of such flows can depend upon the particular routing protocol implemented [24, 25]. This requirement led to a resultant bandwidth which is the sum of given set of parallel bandwidths:

B (Resultant) = Sum {B1, B2, B3 , …}

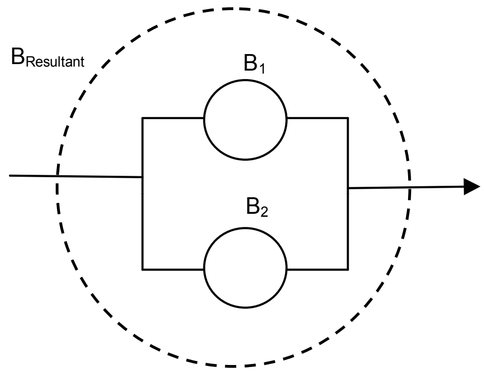

Figure 6: B-Nodes in parallel

Modeling the Flow of Data

To allow for more complex configuration the modeling of networks and internetworks via bandwidth based B-Nodes requires a consideration of converging, diverging, counter-flow and crossing data-flows. These have the representations shown in Figure 7.

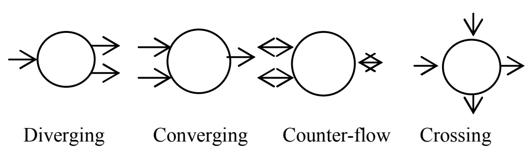

Figure 7: Data flow representations

Converging Data Flows

Files were sent across a network from a source PCs to a destination PCs that is from $P_1$to P_4, P_2 to $P_{5\ \mathrm{and}}$ P_3 to P_6 as shown in Figure 8.

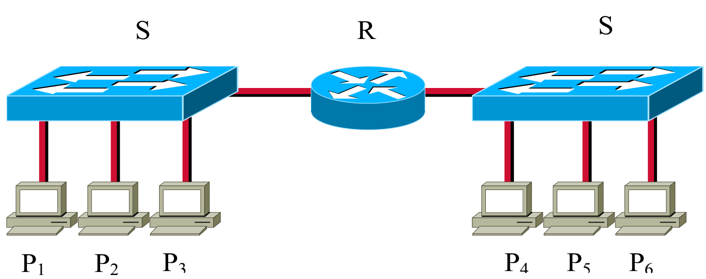

Figure 8: Experimental configuration

Files of different length, such that $L_3 > L_2 > L_1$, were first sent individually between PSRSP configuration. A value of channel capacity C, which is the bandwidth of the source to destination PC using FTP, 7.5 MB/s was used in calculations. This was the average obtained previously in over 100 readings. This meant that the minimum channel capacity 'C' did not occur in the PC but in the routers. Next the effect of multiple files sharing a link of capacity C of was investigated. The version of FTP used only gave transfer times therefore an experimental design technique used to reduce timing errors was used to send a series of files across the network. Whereby each subsequent file sent was smaller than the previous file and also arrived ahead of it, as shown in Figure 9.

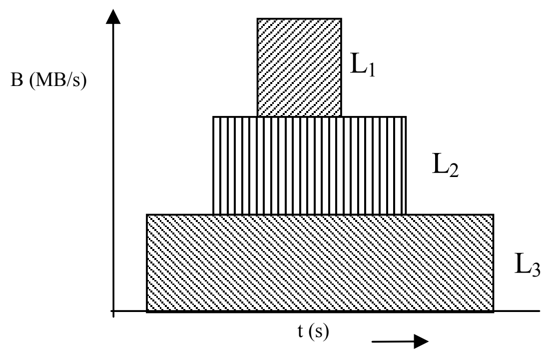

Figure 9: File transmission sequences

As, the initial bandwidths of the file transmission without any other data sharing the channel was 7.5MB/s = C.

The Convergence Formula

A hypothesis was that the files would "squeeze" into the channel and so become elongated with respect to time. This can be seen by considering the diagrams in Figure 10. Lie presents a similar argument for such elongation [21]. This is in keeping with what was expected from equal priority packet switching.

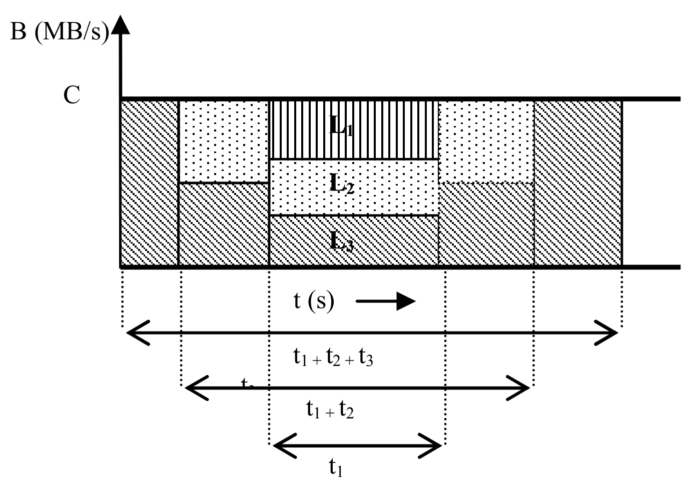

Figure 10: File Bandwidths 'Squeezed' into a Channel

If three files of equal bandwidth attempt to share the channel at the same time then each gets a bandwidth of C/3. The following formula derivation assumes equal bandwidths on each data flow of data sharing a channel, equal priority packet switching and also that the channel capacity $C \leq$ sum of the individual bandwidths of the data flows. This gives the following series for the times to transfer the files and results in the formula shown for determining total transfer time. Assuming nested file transmission times as in Figure 10 and that N represents the number of files to be transmitted.

Where $L_0 = 0$ and $B_0 = 0$.

It can be seen by considering Figure 9 that time $t1 = L_1/(C/3) = 3L_1/C$

Similarly t_2 = (remaining data in L2)/ bandwidth
$= (L_2-L_1)/(C/2) = 2(L_2-L_1)/C$

$$t_3 = (L_3-L_2)/C$$

$$t_{total} = t1 + t2 + t3 \ldots tN$$
$$= N\,L1/C + (N-1)(L2 - L1)/C \ldots + (L_n - L_{n-1})/C$$

This series results in the following formula for the nth term:

$$T_n = \frac{(N - n + 1)\,(L_n - L_{n-1})}{C}$$

File L_1 will arrive with a time of t_1 and file L_2 will arrive with a time $t_1 + t_2$ and so on. These results were confirmed by experiment within an accuracy of 5% or better which confirms how the FTP files converged. These experiments were needed to form a foundation upon which to build more complex models to form networks and internetworks. Such internetworking models can then be made more complex as show in Figure 12.

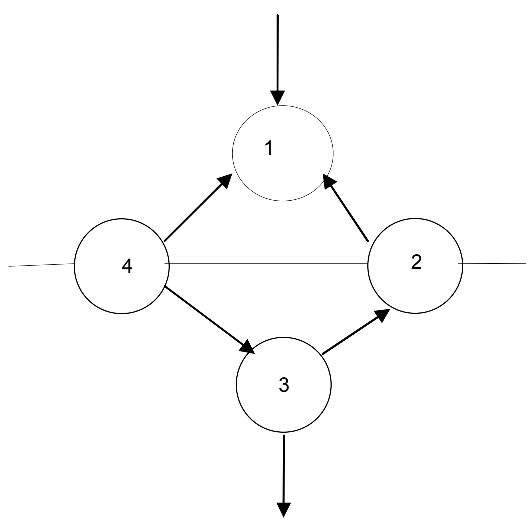

Figure 11: B-Nodes with Recursive Decomposition

The B-Node model needs to undergo further development and testing and its breakdown points need to be determined.

Conclusions

Constructivism offers the possibility of meaningful conceptually-based measures of performance. B-Node models are high level abstractions that are decoupled from the underlying technological detail. From the work to date, under the constraints of the equipment used, it may be concluded that the switches operated at line near to a line speed. Furthermore, introducing routers into the path of the data flow reduced bandwidths.

The authors have quantitatively evaluated the negative bandwidth introduced per extra extended ACE statement. Further investigations are being undertaken to determine the effects of counter-flows and cross-flows. The B-Node model needs further investigation and its breakdown points need to be determined. Further work is planned to repeat these experiments with a range of routers switches and PCs and to use a range of other protocols such as HTTP as the transfer program. Work to date suggests that this simple quantitative model may support student learning.

Bibliography

1. N, Davies, S. Ransbottom, & D. Hamilton, Teaching computer networks through modeling. *ACM SIGada ada Letters,* XVIII(5): 104-110. (1998).
2. Cisco, Cisco networking academy program: First year companion guide. 2 ed. (Indianapolis, IN: Cisco Press. 2001).
3. L. Kleinrock, Queuing systems vol I: Theory. New York, John Wiley and Sons. (1975).
4. T.G. Robertazzi, Computer networks and systems: Queuing theory and performance evaluation. Telecommunication networks and computer systems, ed. M. Gerla, P. Kuehn, and A. Lazar., New York: Springer-Verlag. (1990).
5. D. A. Menasce, & A.F. Alemeida, Capacity planning for web services: Metrics, models, and methods Upper Saddle River, NJ: Prentice Hall. (2002).
6. D.J. Lilja, Measuring Computer Performance: A practitioner's guide Cambridge: Cambridge University Press. (2000).
7. V.Paxton,.& S. Floyd, Wide area traffic: The failure of Poisson modeling. *IEEE/ACM Transactions on Networking*, 3(3): 226-244. (1995).
8. Cobern, W.W. Contextual constructivism: The impact of culture on the learning and teaching of science. In *Annual Meeting of the National Association for Research in Science Teaching*. Lake Geneva, WI: NARCT. (19910.
9. J. Confrey, Learning to listen: A student's understanding of powers of ten., in radical Constructivism in Mathematics Education, E. von Glasersfeld, Editor. Kluwer Academic Publishers: Dordrecht, Holland. 111-138. (1991).
10. R. Driver, & B. Bell, Students' thinking and the learning of science: A constructivist view Leeds, UK: Centre for Studies in Science Education University of Leeds. (1985)

11. G.H. Wheatley, Constructivist perspectives on science and mathematics learning. Science Education, 75(1): 9-21. (1991).
12. M. Ben-Ari, Constructivism in Computer Science Education. In *Twenty Ninth SIGCSE Technical Symposium on Computer Science Educ*. Atlanta, GA: ACM Press. (1998).
13. J. P. Mestre, Cognitive Aspects of Learning and Teaching Science, in Teacher Enhancement for Elementary and Secondary Science and Mathematics: Status, Issues, and Problems, S.J. Fitzsimmons & L.C. Kerpelman, Eds. National Science Foundation: Washington, D.C. 3-54. (2000).
14. von Glasersfeld, E., Cognition, Construction of Knowledge, and Teaching. Synthese, (80), 121-140. (1989).
15. ACM, Computing Curricula 2001 ACM/IEEE Joint Task Force Computer Science Final Report. ACM. (2001).
16. S.P.Maj, & D. Veal. B-Nodes: A Proposed New Method for Modelling information System Technology. In the I*nternational Conf on Computing and Information Technologies ICCIT'2001*. Montclair State University, NJ. (2001).
17. S. P.Maj, & D. Veal, Computer technology curriculum - A new Paradigm for a New Century. *Journal of Research and Practice in Information Technology November (Formerly Australian Computer Journal)*, 32(3/4) 200-214. (2000).
18. S. P. Maj, D. Veal, & R. Duley. A Proposed New High Level Abstraction for Computer Technology. In *SIGCSE ACM 2nd Technical Symposium in Computer Science Educ*. Charlotte, NC. (2001).
19. Hennessy, J.L. & D.A. Patterson, Computer Architecture: A quantitative approach. San Francisco, CA: Morgan Kaufmann. (1996).
20. McComas, B., Can the Pentium 4 recover? 2001, InQuest Market Research.
21. Lai, K & M. Baker. Measuring link bandwidths using a deterministic model of packet delay. In *ACM SIGCOMM 2000 Conf*. Stockholm, Sweden: ACM Press. (2000).
22. Davies, P.T., Securing and Controlling Cisco Routers., Boca Raton, FL: Auerbach Publications, (2002).
23. Jain, M. & C. Dovrolis. End-to-End Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP Throughput. in *ACM SIGCOMM 2002 Conf.* Atlanta, GA: ACM Press. (2002).
24. Lammie, T. & A. Pfund, CCNP Support Study Guide. 2 ed., San Francisco, CA: Sybex. (2002).
25. McGregor, M., CCNP Cisco Networking Academy Program: Advanced Routing. 2 ed., Indianapolis, IN: Cisco Press. (2001).

DAVID VEAL

David received an honors degree in theoretical physics from the University of York in England. After completing a Post Graduate Certificate in Education from the University of Keele after which he lectured in physics at South Devon College UK for 10 years. He now lives in Western Australia where he has taught computing, mathematics and physics at high school level. He now lectures in computing science at ECU in Perth, Western Australia. His areas of research include: Competency-based assessment techniques in computing science, modeling of computers and networks to aid student understanding, and Graphical User interfaces for the partially sighted.

GUPREET KOHLI

Gurpreet is a PhD student at Edith Cowan University with three years of experience in Lecturing and Developing Network and Data Communication units at Edith Cowan University. Gurpreet is currently investigating web services and capacity planning of e-business sites as part of his research at ECU.

PAUL MAJ

Associate Professor S. P. MAJ is a recognized authority in the field of industrial and scientific information systems integration and management. He is the author of a text book, 'The Use of Computers in Laboratory Automation', which was commissioned by the Royal Society of Chemistry (UK). His first book, 'Language Independent Design Methodology - an introduction', was commissioned by the National Computing Centre (NCC). Dr Maj has organized, chaired and been invited to speak at many international conferences at the highest level. He has also served on many national and international committees and was on the editorial board of two international journals concerned with the advancement of science and technology. As Deputy Chairman and Treasurer of the Institute of Instrumentation and Control Australia (IICA) educational sub-committee he was responsible for successfully designing, in less than two years a new, practical degree in Instrumentation and Control to meet the needs of the process industries. This is the first degree of its kind in Australia with the first intake in 1996. It should be recognized that this was a major industry driven initiative.

JAMES COOPER

James Cooper obtained a PhD from the University of Western Australia where his research interest was dynamic machine vision. James now lecturers in Computer Science at Curtin University of Technology in Perth Western Australia where he has undertaken research into network performance analysis and automatic vision systems.