Modelling Layer 2 and Layer 3 Device Bandwidths using B-Node Theory

S Cikara, S P Maj and DT Shaw

School of Computer and Information Science
Edith Cowan University
2 Bradford St Mount Lawley, Perth 6050, Western Australia

scikara@gmail.com; p.maj@ecu.edu.au; alidades@iinet.net.au

Abstract

Modern computer networks contain an amalgamation of devices and technologies, with the performance exhibited by each central to digital communications. Varieties of methods exist to measure and/or predict these performance characteristics. "Rule-of-Thumb" is subjective and based on prior experience, typically offering little mathematical rigour. Benchmarks use different scales and units, with comparative results possibly requiring further interpretation. Stochastic modelling uses complex mathematics which can be problematic and difficult to understand and conceptualise to the typical network administrator. As such, the specific technique employed depends on the problem domain and the cost of getting it wrong.

Bandwidth-Nodes (B-Nodes) are a high-level bandwidth-centric abstraction used to de-couple and control the complexity of a particular technology from the underlying implementation. Devices and/or technologies can be modelled as an individual node or as a collection of nodes, describing the overall function and interactions between both the sub-systems and the operating environment.

This paper uses a simple, common measurement method to calculate the theoretical maximum bandwidth of a single and/or collection of B-Nodes. It demonstrates that the efficiency of B-Nodes can be decomposed and shown as a product of all efficiencies contained within that node. Sub-optimal operation and device efficiency and its effect on bandwidth is also introduced. These are empirically validated and incorporated into the B-Node formula, allowing the bandwidth of a network to be calculated to a first approximation for a variety of devices and technologies. Hence, the anticipated performance of a network given a technical specification can be easily and quickly determined.

Keywords: modelling, B-Nodes, bandwidth, performance.

1 Introduction

A wide range of methods, terms, units and metrics are used to describe the performance of a network system. In conjunction with other factors such as price, they are used as an aid to selection (Maj and Veal 2001). In order to be of any practical value, they should be easy to understand and therefore be based on user perception of performance and as such, be simple and use reasonably sized units (Maj, Veal et al. 2000). Many of the results of these methods may require further interpretation and pose additional questions themselves. Others involve the use of complex mathematics and modelling, such as queuing theory, which can be problematic to analyse and difficult to understand and conceptualise to the typical network administrator.

Bandwidth-Nodes (B-Nodes) are a conceptually simple model used to control the detail of a system by the use of abstraction (Maj and Veal 2001). Details of the technical implementation are deliberately hidden as the specific technological execution may change rapidly and vary from device to device.

B-Nodes use a simple formula to determine the anticipated performance of individual components and networks as a whole. Recursive decomposition allows the performance of a node to be assessed by a simple, common measurement- bandwidth. Sub-optimal operation of B-Node efficiencies, including multiple compounded efficiencies, can also be introduced into an existing system, allowing the efficiency of a single or multiple B-Node(s) to be incorporated and evaluated right down to the device, protocol or technology level if so desired.

B-Node experimentation has shown the use of tools such as PING and File Transfer Protocol (FTP) to ascertain the bandwidth of a given configuration (Veal, Kohli et al. 2005). Work to date has not addressed the addition and subtraction of protocols and/or services to a specific device or configuration. This research will focus on empirically validating these variables and modelling each as its own individual sub-B-Node that impacts network performance, either positively or negatively.

Therefore, it is proposed that the anticipated performance of a network given a technical specification can be easily and quickly determined using B-Node modelling.

2 Network Performance

Network Performance is an amalgamation of terms, units and metrics used to characterise and quantify parameters such as delay, packet loss and bandwidth (Coccetti and Percacci 2002). As such, these cannot be simply expressed by a single parameter, and consequently there are numerous metrics and measurement methodologies employed to express such quantities.

As different applications place different requirements on a network, common criteria must be designed to maximise accurate common understanding by end users and service providers of the performance and reliability both of end-to-end paths and of specific 'IP clouds' (Paxson, Almes et al. 1998). For example, Voice over IP (VoIP) is an application that is sensitive to delay but requires relatively small bandwidth, and bulk data transfers that are insensitive to delay but require large bandwidth. As such, different metrics are used to measure the different quantities- delay is typically measured using packet loss and round trip time (Coccetti and Percacci 2002), (Padmanabhan, Qui et al. 2002), (Lai and Baker 2000), and bandwidth is typically measured by capacity, throughput and available bandwidth (Strauss, Katabi et al. 2003), (Lai and Baker 1999), (Prasad, Dovrolis et al. 2003), (Jain and Dovrolis 2002). Benchmarks can be used as an aid to answering these questions, however results may require further interpretation and additional questions may arise (Maj and Veal 2000).

Performance metrics must use concrete and well defined metrics, be repeatable, exhibit no bias for IP clouds using identical technology, exhibit fair and understood bias for IP clouds using non-identical technologies, avoid introducing artificial performance goals and be useful to users and providers in understanding the performance they experience or provide (Paxson, Almes et al. 1998).

Bandwidth, in a network-centric context, quantifies the data rate at which a network link or network path can transfer information (Prasad, Dovrolis et al. 2003). It must address the impact of application data plus overheads required to transport the data, all in a coherent and easily understood manner. Applications that depend on network capacity to transfer significant quantities of data over a single congestion-aware transport connection rely on the Bulk Transfer Capacity (BTC) of the network. BTC is defined as the long term average data rate over the path in question (Mathis and Allman 2001) and is hence defined as:

$$BTC = {}^{data\_sent}\!/_{elapsed\_time}$$

Therefore, the performance as perceived by the user, is constrained by the overall elapsed time an application takes to be executed over the underlying network (Mathis and Allman 2001).

BTC is an active measurement technique that directly probes network properties by generating the traffic required to make the measurement (Claffy and McCreary 1999). This active and direct method of analysis has the undesirable effect of the measurement traffic having a negative impact (saturation) on the performance of other traffic on the link (Coccetti and Percacci 2002), (Claffy and McCreary 1999).

As networks consist of heterogeneous devices and technologies (Maj and Kohli 2002) including computer nodes or hosts, network connection media, protocols, infrastructure and applications, interchanging any of these variables may vary network performance as each of these technologies have differing overheads. Subsequently, there exists a need for unbiased, empirical performance analysis that is simple, easy to use and conceptualise and be based on user perception of performance.

3 B-Nodes

Generally, any performance analysis or benchmark should provide a coherent conceptual model (Maj and Veal 2000). As such, the measurement standard used must be easy to understand, be based on user perception of performance, be simple, and utilize reasonably sized units (Maj, Veal et al. 2000).

Bandwidth Nodes, or B-Nodes, are a bandwidth-centric concept that uses high level abstraction to de-couple and hide the complexity of a particular technology from the underlying implementation (Maj and Veal 2001). They allow B-Nodes to be modelled as individual nodes (*Figure 1*) or as a sequence of nodes linked together (*Figure 2*).

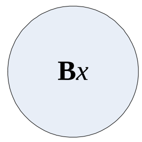

Figure 1 : B-Node

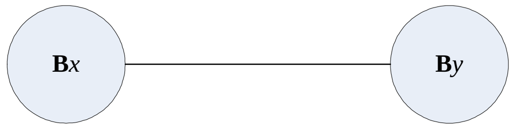

Figure 2: Interconnected B-Nodes

They also allow recursive decomposition to permit a device to be modelled as a collection of B-Nodes (*Figure 3*). A B-Node can also permit full or partial system or device overlap (*Figure 4*) (Maj, Veal et al. 2001).

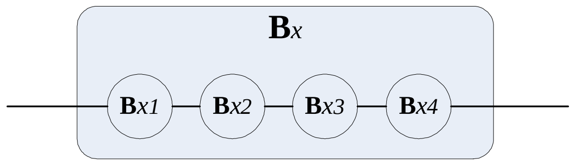

Figure 3: Recursive Decomposition

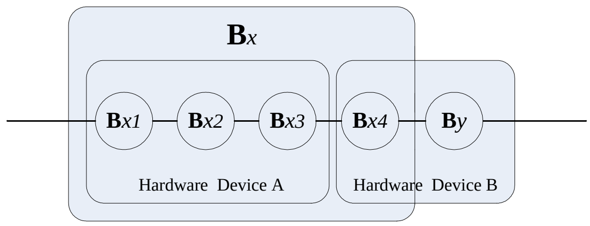

Figure 4: Partial device overlap

Furthermore, "*…each node … can now be now be treated as a quantifiable data source/sink … with associated transfer characteristics (Frames/s or Mbytes/s). This approach allows the performance of every node and data path to be assessed by a simple, common measurement-bandwidth. Where Bandwidth = Clock Speed x Data Path*

Width with the common units of Frames/s (Mbytes/s) … (Maj, Veal et al. 2000).

Operational constraints, including but not limited to processing capacity and interactions between slower nodes, typically influence B-Nodes to perform sub-optimally (Maj, Veal et al. 2001), (Maj and Veal 2001). As such, Maj et. al. has modified the simple bandwidth formula to incorporate sub-optimal operation using an “efficiency” multiplier. Therefore, the bandwidth of a B-Node is defined as:

Bandwidth = Clock Rate x Data Path Width x Efficiency or

$$B = C \times D \times E$$

Equation 1: B-Node formula showing sub-optimal operation

This formula can be applied to the theoretical maximum Bulk Transfer Capacity (and hence bandwidth) for TCP/UDP payloads over 100BASE-TX (100Mbps) Ethernet. All efficiency calculations within this paper are based on this reference protocol. Using the highest level of abstraction, 100BASE-TX has the following transmission characteristics:

Example 1: *Bandwidth* = ?

Clock Speed = 100 MHz

Data Path = $1/8\,B$ (converting bits into bytes), and

Efficiency = 1 (no transport overheads)

Hence:

$$B = 100MHz \times 1/8\,B \times 1$$

$$B = 12.5MB/s$$

Using a lower level of abstraction, 100BASE-TX data encoding uses 4B/5B block coding which means that a 100Mb/s data stream requires 125Mb/s on the media (a 25% speedup resulting in 20% overhead or non-data bits transmitted) (Kaplan and Noseworthy 2000).

Example 2: *Clock Speed* =125 MHz (25% speedup)

Data Path = $1/8\,B$, and

Efficiency = $4/5$ (20% overhead for non-data bits transmitted)

And so:

$$B = 125MHz \times 1/8\,B \times 4/5$$

$$B = 12.5MB/s$$

Alternately, viewing the same problem from an even lower level of abstraction (after MLT-3 coding) the same formula now becomes:

Example 3: *Clock Speed* = $125/4\,MHz$

$= 31.25MHz$ (frequency is reduced to ¼) (Kaplan and Noseworthy 2000)

Data Path = $1/8\,B$, and

Efficiency= $4/5 \times 4 = 3.2$ (20% overhead for non-data bits transmitted)

By reducing the carrier frequency without reducing the data rate, the efficiency is increased by a factor of 4. When it is demodulated at the other end, the efficiency is reduced by the same factor (4 times). So:

$$B = 31.25MHz \times 1/8\,B \times 3.2$$

$$B = 12.5MB/s$$

From this we can see that the regardless of the level of abstraction, the formula still yields the same result- that being the maximum bandwidth of Ethernet is 12.5MB/s. For simplicity, all further calculations and assumptions are based on *Example 1*.

This high level abstraction only deals with 100BASE-TX and its effect on bandwidth. It does not address the subsequent reduction in efficiency additional network protocols and their associated overheads incur, in particular TCP/IP (hence referred to in this document as Ethernet).

4 The Internet Protocol (IP)

Internet Protocol version 4 (IPv4), developed in the 1980’s (Information Sciences Institute 1981), is the most commonly used protocol in today’s networks, and forms the integral basis for what we know as the Internet. As new and powerful applications using the Internet are developed, the underlying protocols operating in the lower layers of the OSI model (the networking protocol stack itself) remain unchanged (Xie 1999).

IPv4 is a network layer protocol that has provision for a 32-bit address space. Modern networks have surpassed IPv4’s capabilities (Tanenbaum 1996). In order to address these and other shortcomings, Internet Protocol Version 6 (IPv6) has been developed (Deering and Hinden 1998) and is slowly being integrated into existing IPv4 infrastructure (Tanenbaum 1996).

IPv6 has a new simplified header format, including a 128-bit address space, which is designed to keep overhead to a minimum. The non-essential and optional fields have moved to extension headers that are placed after the IPv6 header. This reduces the common-case processing cost of packet handling and to limit the bandwidth cost of the new header (Deering and Hinden 1998) allowing for more efficient processing. IPv4 headers are not interoperable with IPv6 headers and hosts must implement both protocols in order to recognize and process both types of headers.

4.1 IP Overhead

By breaking down the various headers, we can analyse and predict the BTC performance degradation incurred between IPv4 and IPv6. By elaborating on Raicu and Zeadally's table (Raicu and Zeadally 2003), we can calculate the *Total Bytes of the Frame on the Wire* for an individual packet, as shown in *Table 1* (grey rows denote new fields introduced by the authors).

Packet Component	IPv4 TCP (B)	IPv6 TCP (B)	IPv4 UDP (B)	IPv6 UDP (B)
Preamble	7	7	7	7
Start of Frame Delimiter	1	1	1	1
Ethernet Header	14	14	14	14
IP Header	20	40	20	40
TCP/UDP Header	20	20	8	8
TCP/UDP Payload	1460	1440	1472	1452
Checksum	4	4	4	4
Interframe Gap	12	12	12	12
Total Overhead	78	98	66	86
Total Bytes of Frame on Wire	1538	1538	1538	1538
Efficiency (%)	94.93	93.63	95.71	94.41

Table 1: IPv4 and IPv6 header overhead showing both TCP and UDP

Using the *Total Bytes of Frame on Wire*, we can calculate the theoretical maximum single packet efficiency using the maximum data payload via *Equation 2*:

$$\text{SinglePacketEfficiency}(\%) = \frac{\text{TCP/UDP Payload}}{\text{Total Bytes of Frame on Wire}}$$

Equation 2: Theoretical maximum efficiency of a single packet

To evaluate the Bulk Transfer Capacity (bandwidth) of 100BASE-TX using these efficiency values, we get the results in *Table 2*:

	No Ethernet Protocol Overhead (Example)	IPv4 TCP	IPv6 TCP	IPv4 UDP	IPv6 UDP
Maximum Line Speed (Mb/s)	100	100	100	100	100
Maximum Line Speed (MB/s)	12.5	12.5	12.5	12.5	12.5
Efficiency of Ethernet (%)	100	94.93	93.63	95.71	94.41
Max Bulk Transfer Capacity (MB/s)	12.50	11.87	11.70	11.96	11.80

Table 2: Bulk Transfer Capacity of IPv4 and IPv6

Using the efficiency percentages from *Table 1*, we obtain the efficiency of a specific protocol ($E_{Ethernet}$) and from this, the computed theoretical maximum BTC for a single protocol (or B-Node) is calculated for 100BASE-TX (*Table 2*). However, the simple B-Node formula (*Equation 1*) does not address multiple B-node efficiencies. It must be extrapolated further to combine the effects of multiple efficiencies and their influence on node bandwidth.

5 B-Node Efficiency Decomposition

By further decomposing *Equation 1*, the efficiency of the B-Node (E) can be shown as a product of all efficiencies (e_i) contained within the B-Node (*Equation 3*).

$$E = \prod_{i=1}^{n} e_i$$

Equation 3: B-Node efficiency product formula

For each B-Node, there is the absolute efficiency, which is the ratio of input to output of each individual B-node, and a relative efficiency which compares the reference value to the output of the B-Node. An example is shown in *Figure 5*.

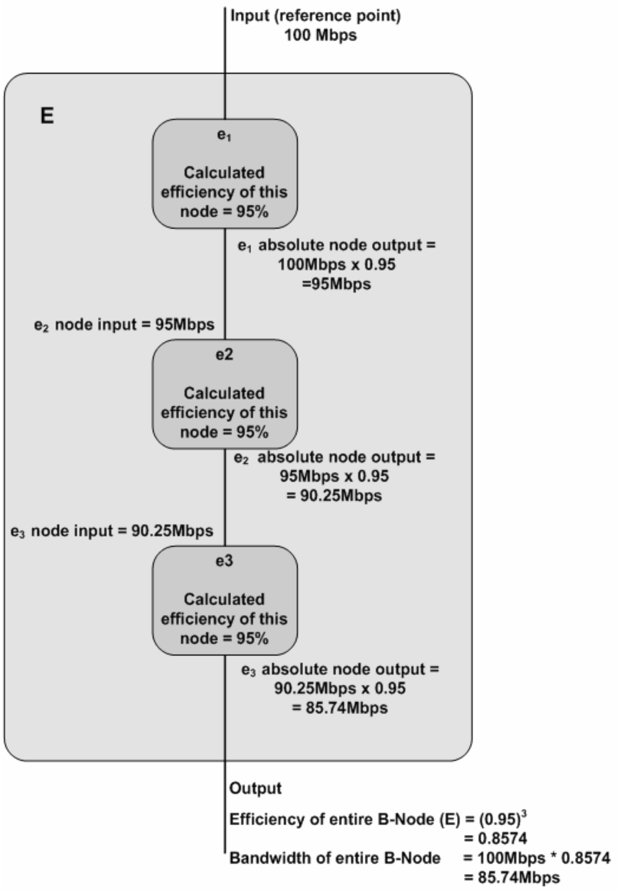

Figure 5: B-Node decomposition example

The component efficiencies (e_i) can further be divided dependent on whether the additional overhead contains Control Packet information or Data Packet overheads.

Data Packet overheads are defined as overheads that are directly added to packets that are transmitting application data. One such example includes Virtual Local Area Network (VLAN) tags. Therefore, e_i, to a first approximation now becomes:

$$E = e_i \times (1 - e_i \times \Delta e_{i+n})$$

$$\text{where} \quad \Delta e_{i+n} = \frac{\text{Additional Data Packet Overhead}}{\text{Data payload}}$$

Equation 4: Data packet efficiency equation

Control Packet information is defined as entirely additional packets used to control link flow. They carry no user data and can typically be viewed as packets that reduce the bandwidth of a link, without transmitting any real application data. Some examples include Spanning-Tree Protocol (STP), Routing Information Protocol (RIP), Enhanced Interior Gateway Routing Protocol (EIGRP) and Open Shortest Path First (OSPF). In this situation, e_{i+n} to a first approximation becomes:

$$e_{i+n} = (1 - \alpha e_{i+n})$$

$$\text{where} \quad \alpha e_{i+n} = \frac{\text{Control packet size per second}}{\text{Link speed per second}}$$

Equation 5: Control packet efficiency equation

The original B-Node formula remains the same, however the e_{i+n} parameter can be interchanged with as many Control Packet or Data Packet efficiencies as required to be added.

To remove an efficiency from a already calculated B-Node, this can simply be achieved by multiplying the B-Node efficiency with the inverse of the efficiency to be removed (*Equation 6*):

$$\frac{1}{e_{i+n}}$$

Equation 6: Efficiency removal equation

This can be applied to both Control and Data Packet efficiencies.

For example, using a B-Node with 5 sub-nodes as defined below:

e_1 is Ethernet Efficiency

e_2 and e_5 are Data Packet efficiencies

e_3 and e_4 are Control Packet Efficiencies

The B-Node formula (*Equation 1*) now becomes:

$$B = C \times D \times (e_1 \times e_2 \times e_3 \times e_4 \times e_5)$$

$$B = C \times D \times (e_1 \times (e_1 \times (1 - (e_1 \times \Delta e_2))) \times (1 - \alpha e_3) \times (1 - \alpha e_4) \times (1 - (e_1 \times \Delta e_2)))$$

As all devices are not created equally, each with their own technological constraints, the B-Node formula does not cater for individual device efficiencies. As such, it must be further expanded to account for these variations in device implementations.

5.1 Device Sub-Optimal operation and its effect on Bandwidth

In an ideal system, an intermediary device such as a switch, router or bridge would have little or no impact on bandwidth. However, this is not always the case. A device itself can introduce latency or processing overheads within a link and hence reduce bandwidth. This may be particularly pronounced in computationally intensive operations such data encryption and decryption.

It is envisaged that there is no one single figure (e_{Di}) for an entire device, rather a figure for each process the device purports to undertake. For example, a router might be particularly fast at switching IPv4 packets, but not very fast at IPv4 encryption using Advanced Encryption Standard (AES) with 256 bit keys. As such, these must be addressed individually. The Efficiency parameter now becomes:

$$E = \prod_{i=1}^{n} e_i e_{Di}$$

Equation 7: B-Node efficiency formula with device sub-optimal operation

The B-Node formula is extrapolated again to take into account this device sub-optimal operation:

$$B = C \times D \times (e_1 e_{D1} \times e_2 e_{D2} \times ..... \times e_{i+n} e_{Di+n})$$

Equation 8: Extrapolated B-Node equation

Using empirically derived results, e_{Di} for an individual process on a particular device can be evaluated.

6 Empirical Validation

6.1 Initial Benchmarking

The initial test bed consisted of two identical 800MHz Celeron dual-stack IBM-compatible PCs with Windows 2003 Enterprise operating system installed. The Intel Pro 100S network interface cards of each machine were directly connected to each other via a crossover cable. This setup (*Figure 6*) forms the benchmark baseline.

To empirically measure the Bulk Transfer Capacity of a link (and hence evaluate B-Nodes), there was a requirement for a single program that could perform IPv4, IPv6, TCP and UDP measurements. In addition to this, it was identified that the performance of a BTC program is often limited by the speed of a disk drive (Spurgeon 2000). Furthermore, the program had to account for this by performing memory-to-memory data transfers. Iperf (NLANR Distributed Application Support Team 2003) was initially evaluated, however erroneous results for IPv6 UDP transfers rendered the program inadequate for the purposes of this experimentation. As such, nuttcp (Fink and Scott 2004) was assessed to meet all the aforementioned requirements. The program's documentation describes "*…its most basic usage is to determine the raw TCP (or UDP) network layer throughput by transferring memory buffers from a source* *system across an interconnecting network to a destination system, either transferring data for a specified time interval, or alternatively transferring a specified number of buffers."*

6.1.1 Initial Benchmark Results (PC to PC)

Using the above method, the efficiency of any introduced or removed system can be calculated and validated. In this case, the efficiency of a BTC test between two identical PCs connected via a cross-over cable was to be assessed. This relative measurement for a minimalistic system was important as it demonstrated the maximum transfer characteristics of an "unloaded" node. All other measurements are calculated relative to these values, either directly or indirectly.

Figure 6 shows the experimental setup consisting of three B-Nodes. The centre node consists of the two identical PCs, the second most inner node is made of Ethernet efficiency (which has already been calculated in *Table 2*), and the outer B-Node, which is the overall efficiency of the node in relation to the input (reference point) and the output (measuring point). This measured value, in conjunction with Ethernet efficiency, allows the empirical calculation of the inner node, and the node efficiency for that specific hardware setup.

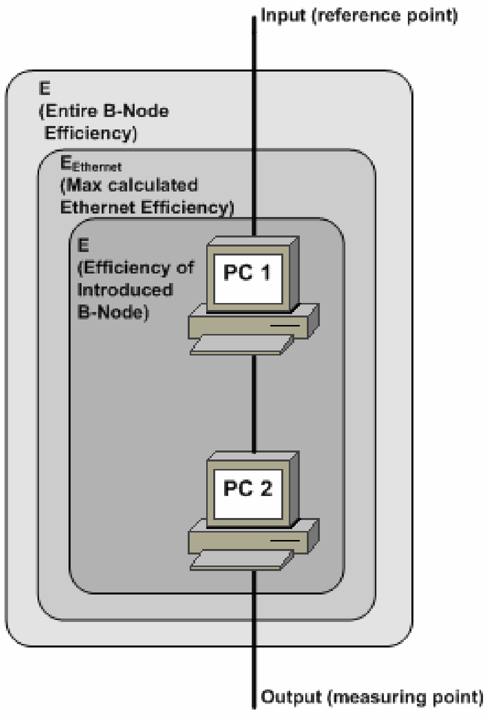

Figure 6: Initial experiment test-bed setup showing B-Node decomposition

From the results (*Table 3*), it can be concluded that both IPv4 and IPv6 TCP have almost optimal (or 100%) efficiencies compared to the calculated value, with both being above 99.78%. Both UDP transfers perform slightly worse than their TCP counterparts (at best almost 1% less) with IPv6 UDP (98.48%) approximately a further 0.5% less than IPv4 UDP (98.93%).

	Theoretical Maximum (MB/s) (from Table 2)	Ouput at measuring point (MB/s)	Difference (MB/s)	Max calculated Ethernet Efficiency (%)	Actual entire B-Node Efficiency (%)	Actual Efficiency (Ed) of Introduced B-Node (%)
IPv4 TCP	11.87	11.84	0.03	94.93	94.75	99.78
IPv4 UDP	11.96	11.83	0.13	95.71	94.66	98.93
IPv6 TCP	11.70	11.68	0.02	93.63	93.43	99.82
IPv6 UDP	11.80	11.62	0.18	94.71	92.96	98.48

Table 3: Bulk Transfer Capacity of IPv4 and IPv6 showing actual efficiency of introduced B-Node.

6.2 Layer 2 Device Measurement

6.2.1 Single Switch Experiments

To calculate specific device efficiencies, the experiment was further elaborated to incorporate both unmanaged (DLink DES1008D) and managed (Cisco 2950 and 3550 series) switches. The equipment was set up as shown in *Figure 7*.

As managed switches have more features available than unmanaged switches, the opportunity to individually test the efficiencies of these was investigated. Initially, a Cisco default switch configuration was tested. In this case, the PCs were in Virtual Local Area Network 1 (VLAN 1), and Spanning-Tree Protocol (STP) was enabled. Various combinations of these were then evaluated including:

1. PCs in VLAN 1 and STP disabled
2. PCs in VLAN 1 and STP enabled (Cisco default configuration)
3. PCs in VLAN 10 and STP disabled
4. PCs in VLAN 10 and STP enabled

Note: On a switch, access ports or non-trunking ports have no VLAN information passed on them. The VLAN tags are not passed through to the PC (and hence do not occupy any time on the wire) and as such, should not impact bandwidth.

It should also be noted that the experiments were conducted using a stable and settled STP network with hello timers set to the default of 2 seconds. Using these parameters, we obtain a calculated maximum efficiency for an STP B-Node to be 99.999663%. This should have negligible impact on BTC.

The experimental setup (*Figure 7*) in this instance consists of four B-Nodes, but with a variable number of ei_{+n} sub-nodes shown in the switch. These variable numbers of sub-nodes in the switch pertain to device specific functionality, such as VLANs and STP. Building up on the methodology introduced in Section 6.1.1, the measured output allows the empirical derivation of the e_{i+n} sub-nodes, and hence, the specific node efficiency for a particular hardware setup, as well as a particular protocol or configuration activated and operating on that device.

From the results obtained, we can see that the introduced B-Node efficiency for both managed and unmanaged switches, regardless of protocols used, was overall fairly constant and close to optimal efficiency (with the minimum being 99.58%, the average being 99.89%). There were instances where the efficiency was greater than 100% (the maximum being 100.6%), but to a first approximation, these can be accounted for in measurement, rounding errors and uncertainties. As such, it can be determined that the addition of a switch within a B-node will have negligible or no effect on Bulk Transfer Capacity.

Analysing the switch sub-node efficiencies also demonstrated that regardless of the VLAN or if STP was enabled or disabled, the result was a negligible impact on bandwidth. The sub-node efficiencies ranged from 99.66% to 100.17%, with an average of 99.97%.

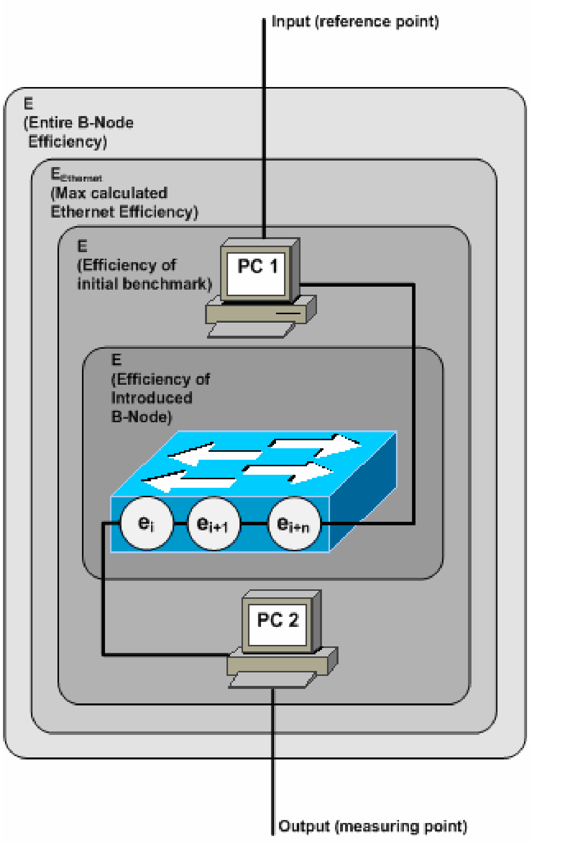

Figure 7: Single switch experiment setup

6.2.2 Dual Switch Experiments

The experiment was then further extended to incorporate two unmanaged (DLink DES1008D) or two managed (Cisco 2950 and 3550 series) switches. The equipment was set up as shown in *Figure 8*.

The additional features that were tested are listed below:

1. PCs in VLAN 1, using 802.1Q encapsulation and STP disabled
2. PCs in VLAN 1, using 802.1Q encapsulation and STP enabled (default configuration)
3. PCs in VLAN 10, using 802.1Q encapsulation and STP disabled
4. PCs in VLAN 10, using 802.1Q encapsulation and STP enabled
5. Same combinations as above, but using Inter-Switch Link (ISL) for encapsulation

The experimental setup (*Figure 8*) again shows four B-Nodes, and a variable number of e_{i+n} sub-nodes. The variable numbers of sub-nodes are setup-specific functionality, such as VLANs and STP and encapsulation type. The specific node efficiency for a particular hardware setup as well as a particular protocol or configuration activated and operating on that device was then evaluated.

From the results, excluding IPv6 TCP on the 3550 (reasons explained further on), we can see that introduced B-Node efficiency is overall fairly constant for both dual managed and unmanaged switches for the features tested (average of 99.92%).

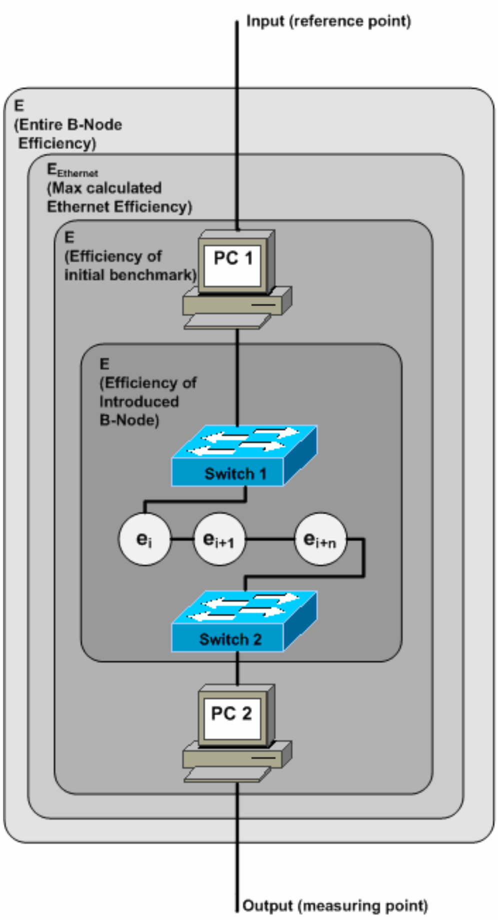

Figure 8: Dual switch experiment setup

On the Cisco 3550 using IPv4 TCP with ISL encapsulation, the efficiency also varies the greatest with respect to the reference value (96.84% and 98.99%). IPv4 and IPv6 UDP with VLAN tagging and ISL encapsulation also had an efficiency that is greater than what can be accounted for in measurement, rounding errors and uncertainties (101.76% to 101.86%). This indicates that the use of these protocols *increases* the efficiency of the B-Node. Possible explanations for this may include the Cisco implementation of these protocols. Further research is required to investigate this phenomena.

IPv6 TCP bandwidth for the Cisco 3550 was significantly lower than for the Cisco 2950 switch (average of 81.62% with approximate 1% deviation from minimum to maximum). Further research is required to explain this, however one possible solution is the software implementation of this particular Internetworking Operating System (IOS) of this switch and its interaction with the congestion control algorithms of IPv6 TCP. This demonstrates that the device efficiency (e_{Di}) for a Cisco 3550 switch using IPv6 TCP is *significantly lower* than for any of the other devices tested.

The sub-node efficiencies (e_{i+n}) showed also that regardless of VLAN, encapsulation or if STP was enabled or disabled, the result was an insignificant effect on bandwidth. The sub-node efficiencies ranged from 99.16% to 101.95%, with an average of 100.22%.

It can be seen from these results that the device (e_{Di}) in conjunction with the protocol used ($e_{Ethernet}$) has the greatest effect on Bulk Transfer Capacity. Ancillary protocols or features (such as STP, encapsulation type and VLANs) have little or no effect on bandwidth. This information would be particularly valuable to a network administrator evaluating and planning network infrastructure.

6.3 Layer 3 Devices

6.3.1 Single Router Experiments

The effect of Layer 3 devices on bandwidth was next to be investigated and empirically evaluated. The device assessed in these experiments was a 2621XM Cisco router, and setup as in *Figure 7* (but with the switch replaced with the router). The results are shown in *Table 4*.

	Theoretical Maximum (MB/s)	Output at measuring point (MB/s)	Difference (MB/s)	Max Ethernet Efficiency (%)	Actual Entire B-Node Efficiency (%)	Efficiency of Introduced B-Node (%)	Efficiency of Introduced sub-B-Node (ei+n)
IPv4 TCP NoCEF	11.87	7.91	3.96	94.93	63.28	66.79	Ref. Val
IPv4 TCP CEF	11.87	7.92	3.95	94.93	63.36	66.87	100.13
IPv4 UDP NoCEF	11.96	7.01	4.95	95.71	56.08	59.24	Ref. Val
IPv4 UDP CEF	11.96	7.04	4.92	95.71	56.32	59.50	100.43
IPv6 TCP NoCEF	11.70	2.5	9.20	93.63	20.00	21.41	Ref. Val
IPv6 TCP CEF	11.70	7.39	4.31	93.63	59.12	63.28	295.60
IPv6 UDP NoCEF	11.80	1.27	10.53	94.71	10.16	10.93	Ref. Val
IPv6 UDP CEF	11.80	7.07	4.73	94.71	56.56	60.84	556.69

Table 4: Single Router Bandwidths

The effect the router has on bandwidth is much more pronounced than a switch. With the exception of IPv6 without using Cisco Express Forwarding (CEF), average TCP bandwidth (65.65%) is consistently higher than UDP (59.86%), with IPv4 TCP (average of 66.83%) having a greater efficiency than IPv6 TCP (63.28%) by approximately 3.5%. Conversely to this, IPv4 UDP (59.37%) is lower than IPv6 UDP (60.84%) by about 1.5%.

Disabling CEF and using IPv6 has the greatest effect on overall router bandwidth. With IPv6 TCP, the efficiency was reduced to 21.41%. IPv6 UDP was approximately 51% less efficient than IPv6 TCP with 10.93%

More pronounced is the effect sub-nodes have on IPv6 router efficiency. By enabling CEF on IPv6 TCP, the efficiency is almost trebled to 295.60%. The result of enabling CEF on IPv6 UDP Bulk Transfer Capacity is even more significant, at 556.69%. Less distinctive is the effect of CEF on IPv4, with the sub-node contributing less than a 0.5% increase in efficiency.

6.3.2 Dual Router Experiments

To quantify the effect of multiple layer 3 devices, 2621XM Cisco routers were paired up and the results noted as follows. In addition, a single Access Control List statement (ACL) was applied to the in and out direction of the ingress interface of *Router 1* and egress interface of *Router 2*. Experimental setup was as in *Figure 8*, with the switches replaced with routers. *Table 5* displays the results.

Average IPv4 TCP performance using dual routers (67.12%) compared favourably with single routers (66.84%), as did IPv4 UDP (dual routers 60.19%) and single routers (59.37%).

Excluding the results obtained from using no CEF, and ACL statements, IPv6 TCP dual routers (DR) were approximately lower by 10.5% than with single routers (SR), to 52.74%.

IPv6 TCP with no CEF was also fairly comparable (DR 20.12% compared to SR 21.41%), as was IPv6 UDP with no CEF (DR 11.10% to SR 10.93%).

Single ACL statements also have significant impact on IPv6 efficiencies. For IPv6 TCP, the statement reduces bandwidth by 13.27% to 39.47%. With IPv6 UDP, this was only reduced by 8% to 52.41%. IPv4 ACL statements improved efficiency by less than 0.6%, which can be accounted for in errors and rounding.

The sub-node efficiencies (e_{i+n}) for IPv4 demonstrated that CEF or an ACL statement does not have an appreciable effect on bandwidth. IPv6 TCP showed that with the introduction of two routers with CEF enabled, efficiency increased to 262.13%, but the addition of an ACL statement reduced this by almost 66% to 197.17%. IPv6 UDP with CEF enabled increased to 544.19% and an ACL statement reduced this by 72.1% to 472.09%

From the results it can be seen that the device sub-nodes (e_{i+n}) in conjunction with the protocol used ($e_{Ethernet}$) has a significant effect on Bulk Transfer Capacity in routers.

	Theoretical Maximum (MB/s)	Measured (MB/s)	Difference (MB/s)	Max Ethernet Efficiency (%)	Actual Entire B-Node Efficiency (%)	Efficiency of Introduced B-Node (%)	Efficiency of Introduced sub-B-Node (ei+n)
IPv4 TCP NoCEF	11.87	7.93	3.94	94.93	63.44	66.95	Ref. Val
IPv4 TCP CEF	11.87	7.92	3.95	94.93	63.36	66.87	99.87
IPv4 TCP CEF ACL	11.87	8.00	3.87	94.93	64.00	67.55	100.88
IPv4 UDP NoCEF	11.96	7.05	4.91	95.71	56.40	59.94	Ref. Val
IPv4 UDP CEF	11.96	7.07	4.89	95.71	56.56	60.11	100.28
IPv4 UDP CEF ACL	11.96	7.12	4.84	95.71	56.96	60.53	100.99
IPv6 TCP NoCEF	11.70	2.35	9.35	93.63	18.80	20.12	Ref. Val
IPv6 TCP CEF	11.70	6.16	5.54	93.63	49.28	52.74	262.13
IPv6 TCP CEF ACL	11.70	4.61	7.09	93.63	36.88	39.47	196.17
IPv6 UDP NoCEF	11.80	1.29	10.51	94.71	10.32	11.10	Ref. Val
IPv6 UDP CEF	11.80	7.02	4.78	94.71	56.16	60.41	544.19
IPv6 UDP CEF ACL	11.80	6.09	5.71	94.71	48.72	52.41	472.09

Table 5: Dual Router Bandwidths

6.4 B-Node Network Performance Analysis

A fictitious network administrator has been given the task to analyse the network shown in *Figure 9*, and to use B-Node methodology to predict the performance of the topology. The technical specification is detailed as below:

1. *PC 1, 2* and *3* are all identical 800MHz PCs
2. *Switch 1* is a DLINK DES1008D switch
3. *Switch 2* and *3* are Cisco 3550 switches
4. *Router 1, 2* and *3* are Cisco 2621XM routers

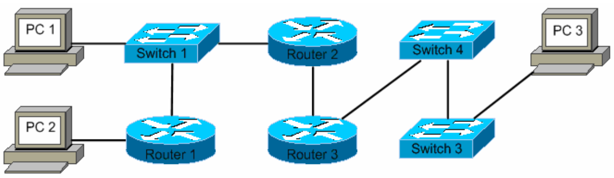

Figure 9: Fictitious network

Assuming no competing transfers, the administrator wants to evaluate the anticipated performance between PC 1 and PC 2 using IPv6 UDP. Router 1 does not use CEF. The B-Nodes for this configuration are:

1. IPv6 UDP Ethernet
2. PC to PC
3. DLink DES 1008D switch, and
4. Cisco 2621XM with no CEF (and hence no sub-nodes)

The B-Node formula hence becomes:

$$B = 100 \times \tfrac{1}{8} \times \begin{pmatrix} e_{IPv6UDP\,Ethernet} \times e_{PC\,to\,PC} \times \\ e_{DLink} \times e_{(Router1\,No\,CEF)} \end{pmatrix}$$

Using the empirically derived results from this research, we get:

$$B = 100 \times \tfrac{1}{8} \times (0.9441 \times 0.9848 \times 100.60 \times 10.16)$$

$$B = 1.19 MB/s$$

The anticipated bandwidth of this configuration is 1.19MB/s. The experimental result obtained was 1.23MB/s which compares favourably with the predicted result.

The network administrator now wants to evaluate the bandwidth between PC 1 and PC 3 using IPv4 TCP transfers, again assuming no competing transfers. Switch 2 and 3 use VLANs, ISL encapsulation but no STP. Router 2 and 3 use CEF. The B-Nodes now are:

1. IPv4 TCP Ethernet
2. PC to PC
3. DLink DES 1008D switch
4. Router to Router (Cisco 2621XM) with a CEF sub-node, and
5. Dual 3550 switches and ISL encapsulation with the sub-node VLANs

The B-Node formula becomes:

$$B = 100 \times \tfrac{1}{8} \times \begin{pmatrix} e_{IPv4TCP\,Ethernet} \times e_{PC\,to\,PC} \times e_{DLink} \times \\ \left(e_{(Router\,to\,Router)} e_{(Router\,to\,Router\,CEF)}\right) \times \\ \left(e_{(Dual\,3550 + ISL\,encapsulation)} e_{(Dual\,3550 + VLANs)}\right) \end{pmatrix}$$

Using the empirically derived results from this research, we get:

$$B = 100 \times \tfrac{1}{8} \times (0.9493 \times 0.9978 \times 0.9988 \times (0.6695 \times 0.9987) \times (0.9701 \times 101.95))$$

$$B = 7.82 MB/s$$

The anticipated calculated bandwidth of this configuration between PC1 and PC3 is 7.82MB/s. Results obtained experimentally compared well to the calculated figure to a first approximation, with a 7.92MB/s bandwidth obtained.

From this, we can see that for a given technical network specification and using B-node analysis, the expected bandwidth for non competing transfers can be calculated to a first approximation.

6.5 Conclusion

B-Nodes may provide a simple, easy to use diagrammatic tool that can be used to hide the complexity of devices and technologies and the performance exhibited by them. Through the use of abstraction, the complexity of a particular technology and its implementation can be decoupled and controlled, allowing them to be modeled as an individual node, or as a collection of nodes showing the overall system structure.

Using the B-Node methodology and its empirical validation, specific technology and device efficiencies have been evaluated and calculated. By decomposing the elements within a configuration and using this information, the simple B-Node formula allows the bandwidth of a network to be calculated to a first approximation, down to the individual components if so desired. Each network communication device, computer nodes or hosts, network connection media and protocols, may be evaluated as required and using this information, the anticipated network performance, given a technical specification, can be easily and quickly determined using the simple B-Node formula, however further investigation and empirical validation of a wider variety of protocols and hardware platforms is required.

7 References

Claffy, K. C. and S. McCreary: (1999): Internet measurement and data analysis: passive and active measurement, http://www.caida.org/outreach/papers/1999/Nae4hansen/Nae4hansen.html

Coccetti, F. and R. Percacci (2002). Bandwidth Measurement and Router Queues. Trieste, Sezione de Trieste.

Deering, S. and R. Hinden: (1998): RFC 2460 Internet Protocol Version 6 (IPv6) Specification, http://www.rfc-editor.org

Fink, B. and R. Scott: nuttcp, v5.1.11 ftp://ftp/lcp.nrl.navy.mil/pub/nuttcp/ 2004

Information Sciences Institute: (1981): RFC 791 Internet Protocol, http://www.rfc-editor.org

Jain, M. and C. Dovrolis (2002). End-to-end Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP Throughput. SIGCOMM, Pittsburgh, Pennsylvania, USA.

Kaplan, H. and B. Noseworthy: (2000): The Ethernets: Evolution from 10 to 10,000 Mbps- How it all Works!, http://www.iol.unh.edu/training/ethernet.html

Lai, K. and M. Baker (1999). Measuring Bandwidth. 18th Annual Joint Conference of the IEEE Computer and Communications Societies.

Lai, K. and M. Baker (2000). Measuring Link Bandwidth Using a Deterministic Model of Packet Delay. Proc. Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, Stockholm, Sweden, ACM Press.

Maj, S. P. and G. Kohli (2002). Modelling Global IT Structures using B-Nodes. 3rd Annual GITM World Conference, New York, USA.

Maj, S. P. and D. Veal (2000). Architecture Abstraction as an Aid to Computer Technology Education. ASEE Computers in Education Division, St Louis, Missouri, USA.

Maj, S. P. and D. Veal (2001). B-Nodes: A proposed new method for modelling information system technology. International Conference on Computing and Information Technologies, Montclair State University, NJ, USA.

Maj, S. P. and D. Veal (2001). Controlling Complexity in Information Technology: Systems and Solutions. IASTED Conference on Computers and Advanced Technology in Education (CATE), Banff, Canada.

Maj, S. P., D. Veal and P. Charlesworth (2000). Is Computer Technology Taught Upside Down? 5th Annual SIGCSE/SIGCUE Conference on Innovation and Technology in Computer Science Education, Helsinki, Finland.

Maj, S. P., D. Veal and R. Duley (2001). A Proposed New High Level Abstraction for Computer Technology. ACM Special Interest Group for Computing Science Education (SIGCSE) 2nd Technical Symposium in Computer Science Education, Charlotte, North Carolina, USA.

Mathis, M. and M. Allman: (2001): RFC: 3148 A Framework for Defining Empirical Bulk Transfer Capacity Metrics, www.rfc-editor.org

NLANR Distributed Application Support Team: Iperf, http://dast.nlanr.net 2003

Padmanabhan, V. N., L. Qui and H. J. Wang (2002). Technical Report MSR-TR-2002-39: Server Based Inference of Internet Performance, Microsoft Research, Microsoft Corporation: Redmond, WA.

Paxson, V., G. Almes, J. Mahdavi and M. Mathis: (1998): RFC 2330 Framework for IP Performance Metrics, www.rfc-editor.org

Prasad, R., C. Dovrolis, M. Murray and K. C. Claffy (2003). Bandwidth Estimation: Metrics, Measurement Techniques, and Tools. IEEE Network. **17:** 27- 35.

Raicu, I. and S. Zeadally (2003). Impact of IPv6 on End-User Applications. IEEE International Conference on Telecommunications (ICT), Tahiti, French Polynesia.

Spurgeon, C. E. (2000). Ethernet: The Definitive Guide, Library of Congress Cataloguing-in-Publication Data.

Strauss, J., D. Katabi and F. Kaashoek (2003). A Measurement Study of Available Bandwidth Estimation Tools. Proc. ACM SIGCOMM Conference on Internet Measurement, Miami Beach, FL.

Tanenbaum, A. S. (1996). Computer Networks. Upper Saddle River, N.J, Prentice Hall.

Veal, D., G. Kohli, S. P. Maj and J. Cooper (2005). A Framework for a Bandwidth Based Network Performance Model for CS Students. 2005 ASEE Annual Conference and Exposition "The Changing Landscape of Engineering and Technology Education in a Global World", Portland, Oregon.

Xie, P. P. (1999). Network Protocol Performance Evaluation of IPv6 for Windows NT. San Luis Obispo, California Polytechnic State University.