

Section 1.0 Introduction and Project Scope of Work

1.1 The University of Victoria

The University of Victoria, one of Canada's leading universities, has earned a reputation for innovation and engagement in the critical issues of our times. UVic serves more than 17,500 students and is widely recognized for its interdisciplinary research initiatives, the non-traditional approach of its professional schools, its extensive co-operative education programs, its pioneering work in distance education, and its support for innovative teaching. The University is located on Canada's spectacular west coast, in the capital of British Columbia. The 160-hectare campus is a 15 minute drive from downtown Victoria.

1.2 Background and Nature of the Project

TAPoR (Text-Analysis Portal for Research) will build a unique human and computing infrastructure to support research in the humanities with advanced text analysis tools. This CFI funded project will establish six Canadian regional centres, that will collaborate to provide scholars with access to electronic texts and text analysis software, through a shared web portal. Each centre (node) will establish a computing environment to support this work, including production and development servers and a small research lab with workstations. The six centres will share expertise and run services which will share data and computing cycles with each other.

The TAPoR project is based at McMaster University, and consists of a network of six of the leading Humanities computing centres in Canada: McMaster, University of Victoria (in collaboration with Malaspina UC), University of Alberta, University of Toronto, Université de Montreal (law) and University of New Brunswick.

1.3 Project Description

At the University of Victoria, our computing environment will be composed of servers which will authenticate users, provide them with advanced searching facilities for electronic texts, connect the texts with a variety of commercial and not-for-profit software tools, and create a secure environment in which researchers can explore texts, save and store results, and share their work with students and colleagues.

For the servers, we envisage a three-tier architecture. The web server will provide user access, user accounts, and connectivity services (through secure protocols such as https,

ssh). Behind this server are the applications server, where the applications logic resides and the applications execute; and the database server, which provides database services and access to large text repositories. We will combine the web server and the applications server in a single computer. There will also be a server to deliver streaming media content.

The production environment is kept separate from the development environment. In the development environment we do testing and development; only selected users have any access here. The web and applications development server and the database development server are intended to substitute for the production servers should one of the production servers fail. There is no development server to pair with the streaming media server.

The equipment will be housed in the UVic Computing Centre computer room, a secure, environmentally controlled room with continuous power supported by UPS (uninterrupted power supply) and generator backup. The UVic Computing Centre operates a number of computers in this location, and provides support for systems programming, operations, maintenance, back-up, and recovery processes for programs and data, and secure off-site storage for disaster-recovery situations. The back-up system is implemented using the Tivoli Storage Manager using an IBM 3494 robotic tape library. The computer Centre has a staff of 11 professionals and 8 technical support staff.

1.3 Project Requirements

A detailed proposal for pricing is required for the initial purchase together with a description and costing a potential upgrade for the third year. We recognize that it is difficult to prepare cost estimates for years 3, 4, and 5. We request that proponents please describe the anticipated evolution of their equipment over this period.

The Project is proposing two separate options on which proponents can bid (see Proponent Options 1 & 2 in the Technical Requirements section). The selection will be based on the overall qualities of the bids, the criteria for which are detailed on the Evaluation Schema attached to the RFQ. Please refer to the Evaluation Schema when preparing your bids.

A brief description of the design and the technical requirements are given in Appendix A.

1.5 Selection of Proponent

The selection will be based on the criteria described in this document. It will depend on the four key areas that are described in Appendix A. These include: Technical Requirements, Cost of the System, Operational Costs, and Proponent Assessment.

The RFP will be used to select a Proponent for system in the first year of the project with the option to purchase the proposed solution for the third, fourth and fifth year, subject to review.

1.6 Canada Foundation for Innovation

The CFI and the British Columbia Knowledge Development Fund (BCKDF) will provide up to 80% of the cost of the TAPoR IT Systems. Universities are expected to find funding partners to share the remaining 20% of the project costs. One of the ways this can be done is through negotiating "special discounts" with Proponents. A Proponent discount becomes an eligible contribution to the project when it represents an additional discount from the price normally charged to an educational institution or most favored customer.

The proponent MUST include a CFI in-kind contribution to be competitive in the bidding process.

In addition, the final award will take into consideration value-added contributions beyond the matching funds. Benefits may take various forms such as cash, products or services in-kind, promotional items and others, in exchange for marketing value back to the external organization.

Appendix A: Detailed Description and Requirements

A.1 Introduction

The TAPoR IT Systems will serve UVic Researchers in the Humanities, the UVic Humanities Computing and Media Centre programming and technical staff, and a variety of users accessing the systems from outside the UVic campus, including researchers and staff from others nodes in the TAPoR Project. The users of these systems and the lab stations linked to them will have a wide variation in their requirements: textual analysis; database search and retrieval; web-based services via a "web portal"; media serving, storing and retrieving; access to many tools to perform a wide variety of humanities-related computing functions; and storage, retrieval and back-up via an existing Network Storage Facility.

The types of functions will vary. The media server will serve large amounts of often culturally sensitive media-based materials for research purposes; the database will process requests for research materials including large media and text files. The Web and Applications servers will provide access to the TAPoR Portal and a variety of tools. Some users will work exclusively with large amounts of small text files while others will

work with large, high resolution media files including images, sound and video--some of it compressible.

Linux is the preferred operating system of the servers. Other operating systems solutions will be considered, but the issue of compatibility with the other network nodes (which are running Linux) must be clearly addressed. The applications software needs to be Unicode-aware.

The system must satisfy this diverse user community.

A.2 Evaluation

The evaluation of the proposals will be based on four main criteria: the technical requirements, the cost of the equipment, the operating costs and the qualifications of the Proponent.

Proponents are free to present a solution of their choice. However, proponents must demonstrate that their proposed solution will integrate and function with the existing infrastructure.

The equipment will be used to service the Humanities research community. Reliability is an important feature; however, it is not necessary that the facility maintain 24x7x365 availability. We are willing to consider alternatives.

A.3 Technical Requirements

The technical requirements are subdivided into four areas: the general design, the hardware configuration / requirements, software requirements / environment, and maintenance and support.

A.3.1 General Design

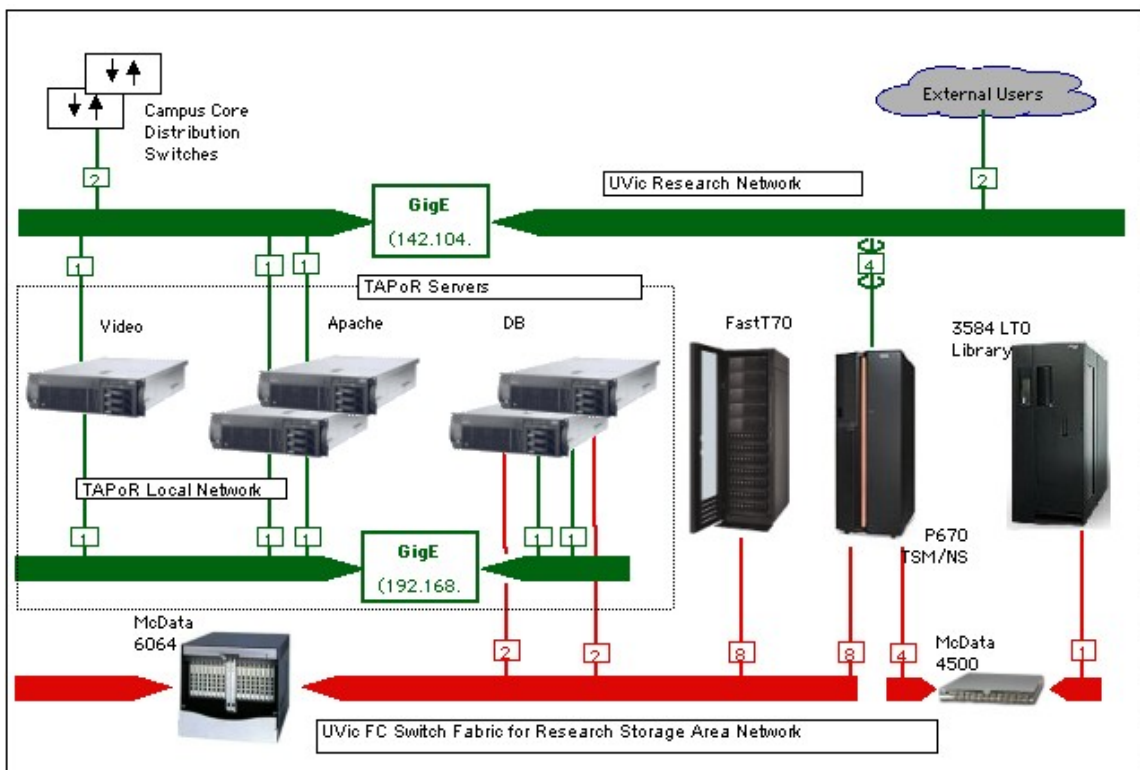
Figure A.3 shows the preferred technical solution. The intent of this figure is to provide an outline of our requirements. The servers are contained within the outlined area listed as TAPoR Servers.

The servers will connect to each other through Gigabit Ethernet, and both the media server and the two web / applications servers will connect to the UVic Research Network through Gigabit Ethernet. The database servers will connect to the storage facility (disk array, controller, and tape library) via a 2 GB/s fibre channel (FC) SAN. The figure

indicates the estimated number of both Ethernet and SAN connections necessary for this design. The media server and the two web / applications servers will require two each of the 1000baseTX gigabit Ethernet connections. The two database servers require one each of the 1000baseTX gigabit Ethernet connections, plus two each of the 2 GB/s dual fibre channel connections to the SAN. Connections of the two databases to the SAN and the Mcdata 6054 switch ARE part of the bid specifications.

External users will connect directly to the media and web servers only, via the UVic Research Network, and a dedicated gigabit Ethernet line will be placed directly into the TAPoR lab area in the Humanities Computing and Media Centre, where the majority of the research will take place.

Figure A.3 The Preferred Technical Solution



A.3.2 Hardware Requirements / Configuration

Servers

Proposals offering either multiple independent dedicated servers or a single multi-processor, partitionable server are acceptable; however, the former is preferred.

The following is a generic description of the desired hardware configuration. These are suggested minimum configurations. Proponents may respond with other comparable configurations, providing documentation to explain and justify their responses.

a) Two identical Web and Applications Servers for PRODUCTION and DEVELOPMENT. Each server requires a dual processor, minimum 2GB RAM, mirrored local disk storage for Operating System, and two 1000baseTX gigabit Ethernet connections.

b) Two identical Database Servers for PRODUCTION and DEVELOPMENT. Each database server requires a dual or quad processor, minimum 8GB RAM, mirrored local disk storage for Operating System, one 1000baseTX gigabit Ethernet connection, and two dual fibre channel connections to the existing Network Storage System.

c) One Streaming Media Server for PRODUCTION. This server requires a dual processor, minimum 2GB RAM, mirrored local storage for Operating System, and two 1000baseTX gigabit Ethernet connections.

d) Data storage of 2TB for PRODUCTION. The disks will fit directly into an existing disk subsystem providing RAID 5 for protection against hardware failure. The existing disk subsystem is attached to the SAN via multiple fibre channel cables for redundancy, performance and load balancing. It is capable of sustaining read/write performance of 100-200 MB/s within the SAN fabric for each TB of disk storage.

e) Standard equipment rack(s) to house the 5 servers and 2TB data storage system.

An Uninterrupted Power Source (UPS) and a Keyboard, Video, Mouse (KVM) switch, are NOT REQUIRED as this equipment is already in place in the University of Victoria Computing Centre, where the servers will be housed. The Computing Centre will provide UPS services and backup power via a diesel generator.

Existing UVic Research Infrastructure - Storage Area Network (SAN)

The University of Victoria has existing information processing infrastructure for research computing which includes a storage area network (SAN). The SAN (see figure A.3 above) is comprised of:

IBM FastT700
IBM P670 TSM/NSF
IBM 3584 LTO Library
Fibre channel switch fabric using a McData 6064 and a McData 4500

Proponent Options

Proponent Option 1 – Bid On Complete Solution

Proponents can bid on providing a complete solution which includes the 5 servers, 2TB of data storage, 40 LTO2 tapes for the data storage system and high speed interconnect between the 2TB of data storage and the 2 data base servers. Bids are anticipated to be in the range of \$150,000 to \$200,000 before applicable taxes.

Proponent Option 2 – Bid On Solution Which Interfaces With Existing UVic Storage Area Network

Proponents can bid on providing just the 5 servers and accompanying gigabit Ethernet connections. For this option, the 2TB of data storage would be provided by an upgrade to the existing UVic research storage area network. The upgrade to the existing UVic storage area network would be purchased from IBM and consist of:

One (1)	IBM Netbay 42U Rack with power distribution units
One (1)	IBM FAStT7 – Storage Server (director) ($\frac{1}{4}$ of the capacity of this device would be assigned to this project and the remaining $\frac{3}{4}$ capacity would be available to other research user(s) of the UVic storage area network.)
One (1)	IBM FAStT EXP700 (drawers)
Fourteen (14)	2Gbps FC 146GB 10K Hot Swap HDD (drives)
Two (2)	McData 6064 2Gb Fibre 4 port mod w/4 sw
Forty (40)	LTO2 Labeled Cartridges

Bids for just the 5 servers are anticipated to be in the range of \$100,000 to \$125,000 before applicable taxes.

Server Environment

Proponents should describe their proposed server environments in these terms:

- i.. reliability, scalability
- ii.. processing power (emphasis on non-floating point performance and input/output speed)
- iii.. user authentication via Kerberos and Grid Canada (www.gridcanada.ca/ca)
- iv.. ability to collaborate with other TAPoR nodes not at UVic; which will include interoperability with other varieties of Unix and Linux
- v.. interoperability with other platforms
- vi.. availability of systems management software and its ease of use
- vii.. interoperability with common programming languages (Java, Perl, PHP, Ruby)

A.3.3 Software Requirements / Environment

Systems Management Software

Systems management software is required on all the servers. Proponents are expected to describe the nature of systems logs which are created, and whether reporting software is being provided. Sample reports or a demonstration of reporting would be helpful.

Operating Systems

Linux is the preferred operating system. Other operating systems solutions will be considered, but the issue of compatibility with the other network nodes (which are running Linux) must be clearly addressed.

Proponents must address any issues of compatibility with the other network nodes, which will also run Linux.

The application software must be Unicode aware. Proponents must address any issues with providing full Unicode support via the Systems Management Software or support for applications packages (see list of applications software below) that provide full Unicode support.

Software Environment

These are typical and anticipated software solutions which we expect to run in this environment. They are provided here for information purposes only. **Proponents are not expected to supply or install them.**

- a) Web server: Java; Apache; Samba; ssh
- b) Applications server: Java; text analysis software (see below for details)

- c) Database server: Oracle; MySQL; PostGres; text management software (see below for details)
- d) Media server: RealMedia, open source streaming video package

The applications server will be primarily delivering computing cycles for text analysis and text processing. Commercial packages which will be implemented include:

Package	Environment	More Information
TIBCO Transform Engine – XSLT transformations of data	Java	www.tibco.com
Ontopia applications – XML topic map processing and manipulation	Java	www.ontopia.net
XTeXT– searching and browsing of text collections	Java	www.isagn.com/#software
Omnimark – batch transformation of textual data	binaries for various flavors of Unix	www.stilo.com/index.html
HyperPo –textual search and display [written in Perl]	Perl	http://huco.ualberta.ca/HyperPo/
Exgenix – automatic document conversions	Linux	www.exegenix.com/technology/technology_faq.php

The development version of the applications server will be used by developers. It will support a J2EE development environment, with an emphasis on text technologies extensively using XML, which will have services such as the following installed on it (these are typical examples; not all of them will be necessarily implemented):

Package	More Information
Orion: J2EE-based server	www.orionserver.com
BEA Systems Weblogic: J2EE development environment	www.bea.com
JBoss: J2EE development environment	www.jboss.org
Tomcat	jakarta.apache.org/tomcat
XP: XML parser in Java	jclark.com/xml/xp
Xerces & Xalan: XML parser, XSLT processor	xml.apache.org/xalan-j and xml.apache.org/xerces-j
Perl: programming language	www.perl.com
Ruby: programming language	www.ruby-lang.org/en
PHP: web database scripting	www.php.net

The database server will be running SQL database systems, and will also deliver textual data from text collections and text stores. The primary database system will be: Oracle (version 9i). There will also be the need to run other database systems: we will implement: MySQL PostGres, as well as repository packages such as TIBCO Canon (XML DTD and Schema repository)

A.3.4 Hardware Warranty, Maintenance, Support and Upgrade

Proponents are encouraged to provide a full 365 days first year warranty from the date of installation on all parts and labour. Travel, shipping and other expenses incurred due to warranty repair shall be included at no additional cost. The period will begin to run at the time the system is placed in complete and full-time operation. The proponent agrees to repair the system or replace any necessary parts at no charge during this period.

Proponents should provide a maintenance schedule that describes all required licensing, renewal and on-going costs for a period of five years from the time of initial purchase. The maintenance agreement should provide for coverage for business days Monday to Friday 9:00 to 17:00, with the option to pay for the 5 year warranty up front.

Proponents are asked to provide information on the following items:

- i.. Details and options for telephone support.
- ii.. Mechanisms for reporting hardware and software problems.
- iii.. Identify Victoria or Vancouver resident technical support staff who would be assigned to the project.
- iv.. Describe your normal office hours and mechanisms for obtaining access to technical support for emergency response.

Proponents should describe a strategy, with estimated costs, for upgrade or renewal of the hardware at the 3 year mark of the project's life-cycle.

A.3.5 Network Connectivity

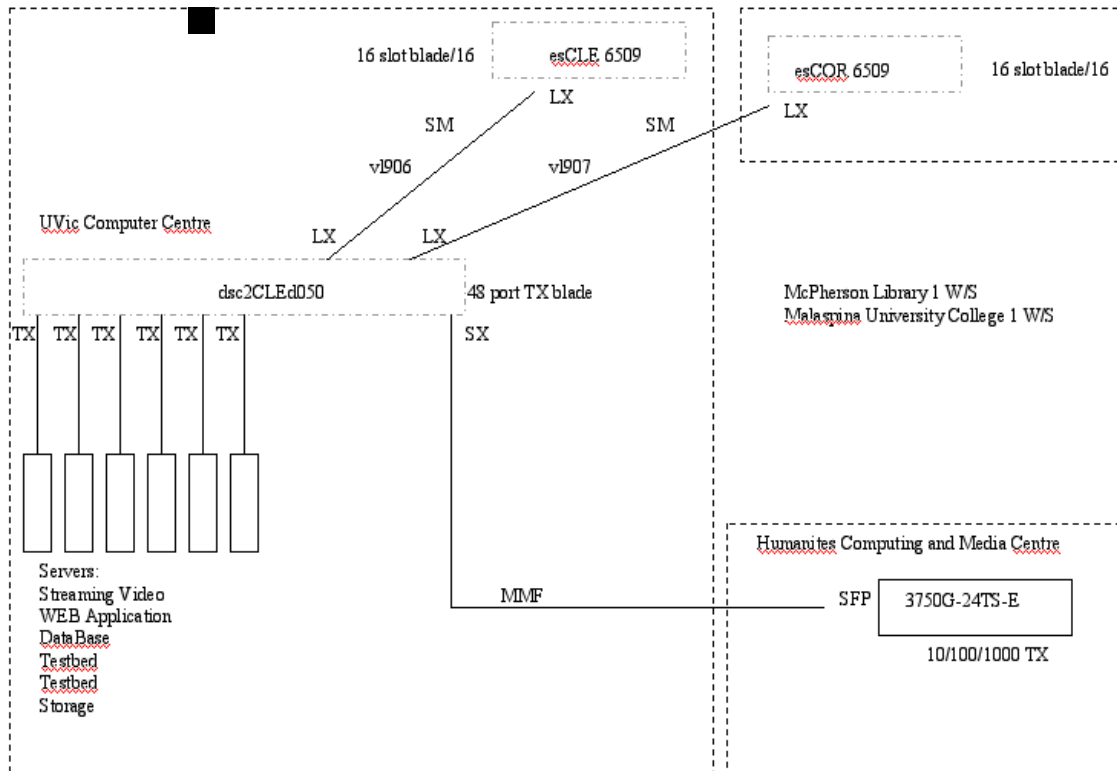
See figure A.3 above for a block diagram representation of the connection of the servers to the UVic research data network and of the servers to the fibre channel switch fabric connection to the existing UVic research storage area network.

Server / Client Connection

The servers are to connect to each other and to the UVic campus research data network via 1000baseTX gigabit Ethernet connections. The servers will have a dedicated line to the Humanities Computing and Media Centre, where 8-10 project workstations will be

housed. The line will be 1000baseTX capable and route to the HCMC via a 24 10/100/1000TX switch to be located in a switching closet near the HCMC. Other Ethernet lines are to run to switching closets in other locations via the UVic Research Network. See Figure A.3.5 for a diagram of the Server / Client connection.

Figure A.3.5 Networking Diagram of TAPoR Servers to Clients



A.3.5 Cost of Equipment

Proponents are required to provide the total costs for the equipment in year 1. This should include the cost of each hardware item, the cost of the tape media, the software costs and the cost of the warranty and maintenance. The total costs should also include any expenses for shipping, handling, installation and commissioning. The total should include all taxes.

The pricing MUST include the cost after all exceptional discounts. In addition, Proponents are REQUIRED to provide the normal discounted price to UVic. The difference between the normal university discount and the quoted price is considered by CFI to be the industrial matching funds.

The total cost should be based on the technical requirements specified in this document.

Proponents are requested to provide the cost per terabyte of additional online disk capacity.

Innovative solutions that have a higher total cost will be evaluated fairly.

Proponents are required to give an estimate of the cost to upgrade the equipment in year 3. Proponents should identify the hardware and software required to upgrade the equipment and, if possible, proponents should indicate potential technology and performance changes that may occur in year 3.

Proponents must commit to providing the same (or better) discounts in year 3 as provided in year 1.

Proponents are required to describe clearly the base pricing and the amount of the qualifying CFI discount they are able to provide on each item.

A.5 Installation and Commission

Proponents are required to specify the delivery time for all pieces of hardware and software.

Proponents must provide a detailed description of the installation and commissioning procedures during Year 1. An estimate of the time to install and commission the facility is required. Proponents should indicate the manpower and assistance that they will provide during this phase.

Proponent must provide a description of the installation and commissioning procedures for the upgrade of the equipment in Year 3.

A.6 Mechanical and Electrical Requirements

Proponents must demonstrate that all electrical hardware meets the requirements of the University of Victoria.

Proponents are required to provide the electrical requirements for each hardware unit. This should include the voltage, current and power requirements, and the number and type of electrical connections.

Proponents are required to provide the cooling requirements for each hardware unit. If

cooling fans are provided, then proponents should indicate the floor space required for cooling.

A.7 Operational Costs

Manpower

Proponents must describe the manpower required to manage the proposed server system. The should include the manpower to manage the hardware components (servers, storage disks), and to address the user requirements.

Match to Staff Skills

The existing manpower at the UVic Computing Centre has a variety of skills and a wide level of experience. The group has in-depth knowledge of the AIX and Windows NT operating systems and the Tivoli Storage Management (TSM) suite, as well as limited knowledge of the Solaris and Linux operating systems.

The existing manpower in the Humanities Computing and Media Centre has in depth knowledge of a variety of programming languages, and a wide level of experience with the creation and use of Humanities related software and websites, in addition to extensive knowledge of Windows 2000 and Macintosh operating systems. The HCMC staff have some knowledge of Linux operating systems.

Training

Proponents should describe the training opportunities that they provide.

Proponents should describe any local or on-site training opportunities.

A.8 Proponent Assessment

In an earlier section, proponents have been asked to provide information. It is required that Proponents establish their credentials and that their proposed solution meets our technical criteria.

Credentials

Proponents should highlight their reputation, experience and qualifications in the server marketplace.

Proponents should describe their resources in British Columbia that would be of value to

the proposed TAPoR IT system.

Proponents should describe their existing relationship with UVic and their expectations for the future.

Proposed Solution

Proponents are expected to provide a clear and complete proposal.

Proponents should highlight any exceptional discounts or value-added benefits.

4.0 CFI In-Kind Contributions

The Canada Foundation for Innovation (CFI) is an independent corporation established by the Government of Canada in 1997. The Foundation's goal is to strengthen the Canadian capacity to carry out world-class research. CFI generally funds projects with total projected costs in excess of \$100,000. On average, the CFI contributes 40% of the eligible costs of the projects funded. The BCKDF generally matches the CFI funding with a 40% contribution. Universities are expected to find funding partners to share the remaining 20% of the project costs. One of the ways this can be done is through negotiating "special discounts" with vendors. A vendor discount becomes an eligible contribution to the project when it represents an additional discount from the price normally charged to an educational institution or most favored customer.

Pricing Format

Proposal pricing must be in the following format to adhere to CFI requirements:

List Price \$ _____

Educational Discount \$ _____

In-Kind Contribution \$ _____

Net Price \$ _____