Inverse Multiplexing Overview

With the ever-increasing importance of bandwidth-intensive communications in their day-to-day business, more and more small-to-medium-sized enterprises are outgrowing their T1 or E1 wide area connections. When this happens, they find themselves facing a "bandwidth gap". The next-higher-rate service typically available from their service provider is T3 or E3, a very large step up in bandwidth (roughly, from 1.5 to 45 Mbps for T1/T3, or from 2 to 34 Mbps for E1/E3) that comes with a correspondingly large increase in price. It should not be surprising then that an enterprise having outgrown a T1 or E1 wide area connection often cannot cost-justify leaping the bandwidth gap all the way to T3 or E3.

A bridge over these troubled waters can be found in **Inverse Multiplexing (IM)** a straightforward and easily implemented solution that bridges the bandwidth gap between T1 and T3 when carriers don't offer other incremental level services, such as Ethernet. IM creates scaleable, economical WAN connections of up to, say, 12 Mbps (T1) or 16 Mbps (E1) by bundling up to, say, eight T1 or E1 circuits into a single high-speed link.

The basic concept of IM, illustrated in Figure 1, is simple. A data stream that is too large for a single transmission path is broken into smaller pieces, and the pieces are transmitted over separate transmission paths to the receiving end, where the pieces are reassembled into the original data stream.

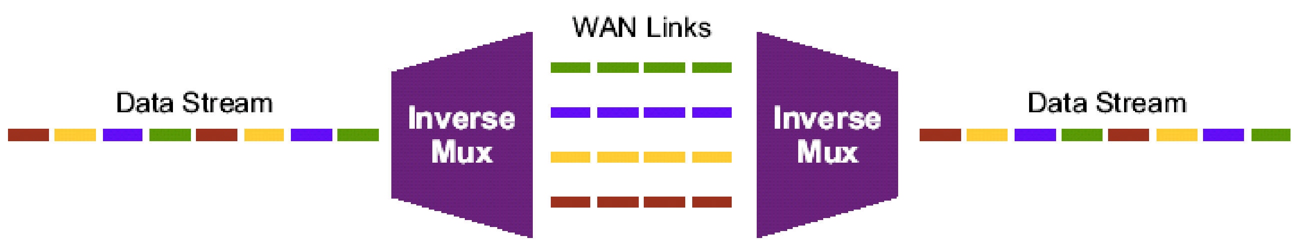

Figure 1 Basic Concept of Inverse Multiplexing

Today a number of network devices, called inverse mu(ltiple)xers, have implemented this concept, with different technology schemes, to create otherwise unavailable multi-megabit WAN links using readily available T1 or E1 circuits and to offer IM solutions that address different performance, availability and interoperability issues.

- **Bit-Based Inverse Multiplexing**

In this method, individual bits are distributed in sequence to the multiple transmission circuits on a round-robin basis. Operation of the inverse multiplexer at the digital signal rate (DSR) layer level of the serial bit stream provides independence from data transmission and network equipment protocols and the ability to be used in virtually any application. Relatively little bandwidth overhead is required to reconstruct received bits into the original sequence. There is no industry standard for bit-based IM; therefore, products implementing this approach must use a proprietary protocol for disassembling and reassembling the data stream. However, since the IM operation is transparent to both the local and wide-area networks, only the inverse multiplexer at each end of the link must recognize and support the proprietary protocol.

The desire for standards-based IM has resulted in several major layer 1 and layer 2 solutions. These approaches have particular benefits and drawbacks related to the protocols on which they are based.

- **Virtual Concatenation (VCAT) with or w/o Link Capacity Adjustment Scheme (LCAS)**

High-order and low-order VCAT has been defined to provide a standards-based protocol that enables client services to be efficiently mapped to SONET/SDH payloads, and LCAS has been defined to make on-demand, hitless bandwidth changes in a VCAT environment a reality. VCAT for SONET/SDH provides the ability to transmit and receive several noncontiguous STS (or VT)/VC fragments as a single flow. This grouping of STSs/VCs is called a virtual concatenation group (VCG). The high-order VCAT notation for SONET is STS-n-Xv, where n is the size of the noncontiguous STS fragments that will be used to transport the entire VCG. The value of X is the total number of n fragments that it takes to make up the total VCG. The VCAT notation for SDH is VC-n-Xv, where the definitions of n and X are the same as used for SONET.

- **Inverse Multiplexing for ATM (IMA) with or w/o Link Addition and Slow Recovery (LASR)**

For ATM-based networks, the ATM Forum has standardized a method of IM in which the data is distributed to multiple T1 or E1 circuits on a cell-by-cell basis and in a cyclic round-robin fashion. The advantages and disadvantages of IMA are essentially the advantages and disadvantages of ATM. The IMA protocol provides advanced Quality of Service (QoS) capabilities, but at the expense of a percentage of bandwidth being used for overhead. IMA is based on the concepts of an **IMA Virtual Link** consisting of multiple physical links, called bidirectional IMA links, and an **IMA Group** referring to the near end or far end of an IMA Virtual Link. The following figure, adopted from Figure 3/I.761, summarizes the IMA managed objects:

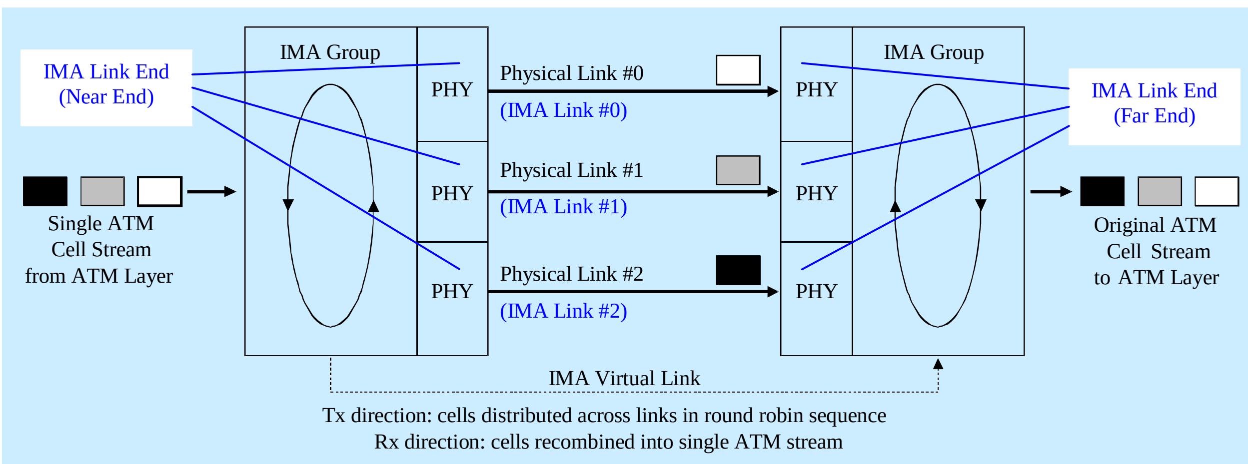

Figure 2 Inverse Multiplexing and De-multiplexing of ATM Cells via IMA Groups

MTNM specifies CTPs at layer rates ATM VC, ATM VP, and ATM NI to model the **ATM transport assembly** consisting of the virtual channel (VC) layer network, the VC to virtual path (VP) adaptation, the VP layer network, and the VP to transmission path (TPath) adaptation by means of transmission convergence (TC) adaptors including TPath termination (ATM cells to/from say SONET/SDH or PDH or DSL payloads), i.e. the transport network interface (NI). For IMA support both the ATM NI CTP and the underlying transport TP need to be enhanced. The ATM NI CTP is replaced by an **IMA group TP** and the respective transport TP is, so to say, supplemented by an **IMA link (end) TP** (see Figure 3 below). The IMA Layer Reference Model (LRM) of I.761 generalizes the B-ISDN Protocol Reference Model (PRM) of I.321 by splitting the transmission convergence (TC) sublayer into an **IMA-specific (upper) TC sublayer** and a **physical interface-specific (lower) TC sublayer** but leaves the physical medium dependent (PMD) sublayer unchanged including transmission media (TMedia) adaptations and terminations. The term "transmission convergence" refers to all functions required to transform a flow of ATM cells into and from a flow of data units (e.g., bits) which can be transmitted and received over the underlying physical medium. The IMA LRM extends the previous VP to TPath adaptation and termination through TC adaptors (ATM cells to/from transport payloads), an essential function of ATM NI CTPs, by an IMA-specific part that adds the capability of IM into multiple interface-specific TC sublayers in a cyclic round-robin fashion and on a cell-by-cell basis.

The ATM IM technique involves inverse multiplexing and de-multiplexing of ATM cells in a cyclical fashion among physical links (bidirectional IMA links) grouped to form a higher bandwidth logical link (IMA Virtual Link) whose rate is approximately the sum of the physical link rates, and each end of an IMA Virtual Link is referred to as an IMA Group (see Figure 2). In the transmit direction, the ATM cell stream received from the ATM layer is distributed, on a cell-by-cell basis, across the multiple links within the IMA link. At the far-end, the receiving IMA unit recombines the cells from each link, on a cell-by-cell basis, recreating the original ATM cell stream. The aggregate cell stream is then passed to the ATM layer. The transmit IMA unit periodically transmits special cells, called IMA Control Protocol (ICP) cells, that contain information permitting reconstruction of the ATM cell stream at the receiving IMA unit after accounting for the IMA link differential delays, smoothing Cell Delay Variation (CDV) introduced by the control cells, and executing other control tasks.

The IMA protocol can preserve service during dynamic bandwidth changes and so IMA allows for the provision of variable bandwidth over time (non-disruptive service modification). This capability of IMA is known as Link Addition and Slow Recovery (LASR) procedure. Whether IMA group and link provisioning tasks can be done in-service (dynamic bandwidth adjustment) or only out-of-service (bandwidth reconfiguration) depends on the support or non-support of the LASR procedure by the IMA equipment.

- **Ethernet Link Aggregation (ELA)**

This IM method employs multiple Ethernet transport links at the Ethernet PHY layer as standardized by the IEEE. Link aggregation allows multiple physical ports to be aggregated together to form a Link Aggregation Group (LAG), such that the Ethernet bridge relay entity can treat the LAG as if it were a single logical port. For bridge functionality (i.e., VLAN, STP, etc), the LAG is considered as a single bridge port (which replaces the physical ports it aggregates). The LAG consists of several full duplex ports operating at the same speed, assumed to be connected via physical links to the same peer system (e.g., Bridge, Host, etc). LAG establishment is initiated by network management and really accomplished by the Link Aggregation Control Protocol (LACP). Via this protocol, the ports on both sides of the same physical link exchange BPDUs and verify that they are part of matching LAGs. Only then they become active as LAG members. ELA provides not only load sharing of traffic between multiple physical links but also high availability and rapid reconfiguration since a physical link failure does not cause failure of the whole LAG (as long as other physical links in the LAG are operational) and traffic from a failed link can be rapidly redirected to other operational links.

- **Voice Signalling Interfaces (V5.2, GR-303)**

Voiceband services include call and service control functions as well as bearer and bearer control functions. Call control is achieved by signalling mechanisms based on a dedicated signalling protocol. The most important voice signalling interfaces between an Access Network (AN) and the Local Exchange (LE) are V-interfaces with Common Channel Signalling (CCS) (see Q.512). There are the E1-based European signalling types V5.1 (see G.964) with a single E1 link and V5.2 (see G.965) with up to 16 E1 links, and the T1-based North American signaling standard GR-303 with up to 28 T1 links. Every E1/T1 link is channelized by TDM into 32/24 voice channels with DS0 capacity that are configured to be either bearer channels or common signalling channels.

A V5 interface is specified by various V5 signalling objects. It contains (logical) c-channels at the AN side and POTS or ISDN user ports at the LE side. The TDM traffic of a V5 interface is transmitted to and received from assigned V5 links. Such a V5 link is an E1 interface that is structured into DS0 channels which either carry signalling information, and are then called physical c-channels, or represent bearer channels. A similar description applies for GR-303. Inverse multiplexing is used to distribute the payload generated from user ports across the V5 links, on a channel-by-channel basis, to guarantee high availability and protection, and to allow overbooking of subscribers by dynamic channel allocation to user ports.

- **Multilink Frame Relay (MFR)**

In this standard developed by the Frame Relay Forum, multiple virtual T1 or E1 circuits are bundled to create a higher-rate channel which is recognized as a single physical interface at the data link layer. The data stream is distributed across the virtual circuits on a packet-by-packet basis. Its advantages include, besides general IM benefits, standardized Service Level Agreements (SLAs), support for variable frame sizes and fragmentation, low latency, and minimal overhead bandwidth.

- **Multilink PPP (MLPPP)**

This packet-based IM method employs multiple Point-to-Point Protocol (PPP) links as standardized by the IETF. It was developed in response to some of the limitations of proprietary load sharing schemes employed on many IP routers. In addition to vendor-independence, its advantages include efficient frame mapping, low overhead bandwidth, connectionless IP environment, and packet fragmentation to prevent monopolization by large data packets. Unlike IMA and MFR, MLPPP is not dependent on switches supporting a particular protocol; rather, it is like bit-based IM in requiring only that the receiving end supports MLPPP.

- **Modelling Inverse Multiplexing with MTNM Objects**

The TMF MTNM NML-EML interface models IM by server layer containment of CTPs and FTPs. Refer to Section 3.1.3 "Inverse Multiplexing (IM) Principles" and Section 3.11.5 "TPs for Inverse Multiplexing (IM)" of the supporting document SD1-18 for general IM modelling considerations.

The current version of the interface supports the IM technologies **IMA/LASR** and **VCAT/LCAS**.

Section 4.5 "ATM ports with Inverse Multiplexing in ports in SDH, SONET, DWDM and ATM NEs", Section 4.6 "Unspecified signals with Inverse Multiplexing in ports in SDH, SONET and DWDM NEs", and Section 4.8 "Ethernet port modelling" of SD1-18 apply MTNM's general IM concepts to ATM and SONET/SDH.

Figure 76 on page 78 of SD1-18 models the managed object "IMA Group" of Figure 2 as a non-connectable, two-layer FTP with upper layer `LR_ATM_NI` and lower layer `LR_Fragment`. The figure also shows the "IMA Link End" managed objects of Figure 2 which are the server CTPs of the **IMA group FTP** and so have `LR_Fragment` as upper layer and the bandwidth of the physical IMA links, e.g. `LR_E1_2M`, as lower layer. The **IMA link CTPs** are named relatively to the containing FTP with respect to their common connectable layer by use of an index (e.g., `/atmnetworkinterface=1/e1=<e>` in case of `LR_E1_2M`). The CTPs at the other end of the SNCs shown in Figure 76/layers.pdf are termed **supporting CTPs** since they support the transport of the splitted ATM cell stream. The physical links that constitute an IMA virtual link can be structured and so the supporting CTPs can be contained in transport PTPs (**unstructured IMA links**) or transport CTPs (**structured IMA links**). The following figure shows the TP hierarchy of a structured IMA virtual link above an STM-4 port:

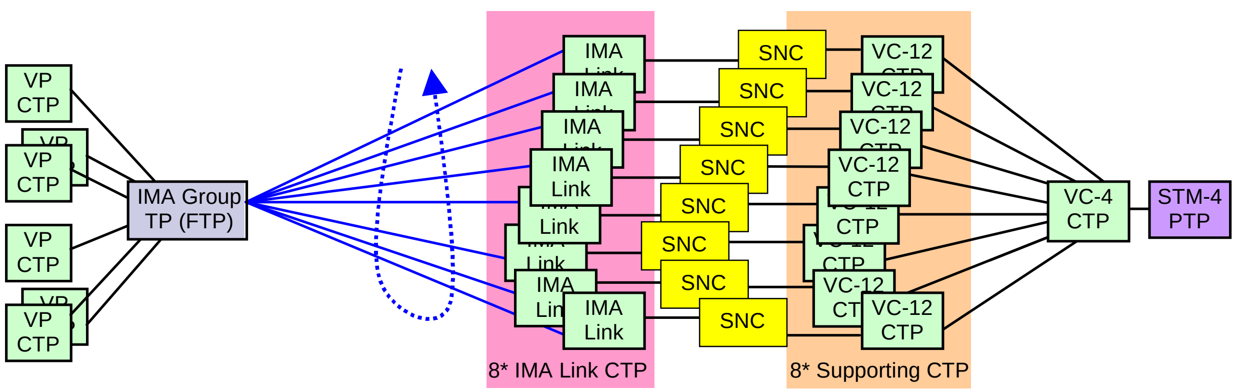

Figure 3 Example of a Structured IMA Virtual Link

Note that the connectable layer rate of an IMA link TP needs not be fixed but may be configured dynamically by the EMS during server CTP generation. However, in many cases the rate is fixed. For example there are IMA cards which have say two IMA groups each above say eight E1 or T1 or SHDSL ports that are permanently assigned. An IMA group TP is called **fixed** if all SNCs shown in Figure 76/layers.pdf are fixed, i.e. all SNCs between its server TPs and their associated supporting CTPs. An IMA group TP is called **flexible** if all SNCs shown in Figure 76/layers.pdf are flexible. A mixture of fixed and flexible SNCs involving IMA link CTPs of the same IMA group FTP is not allowed at the MTNM interface.

The following figure shows all MTNM objects of an IMA plug-in unit with two fixed IMA groups:

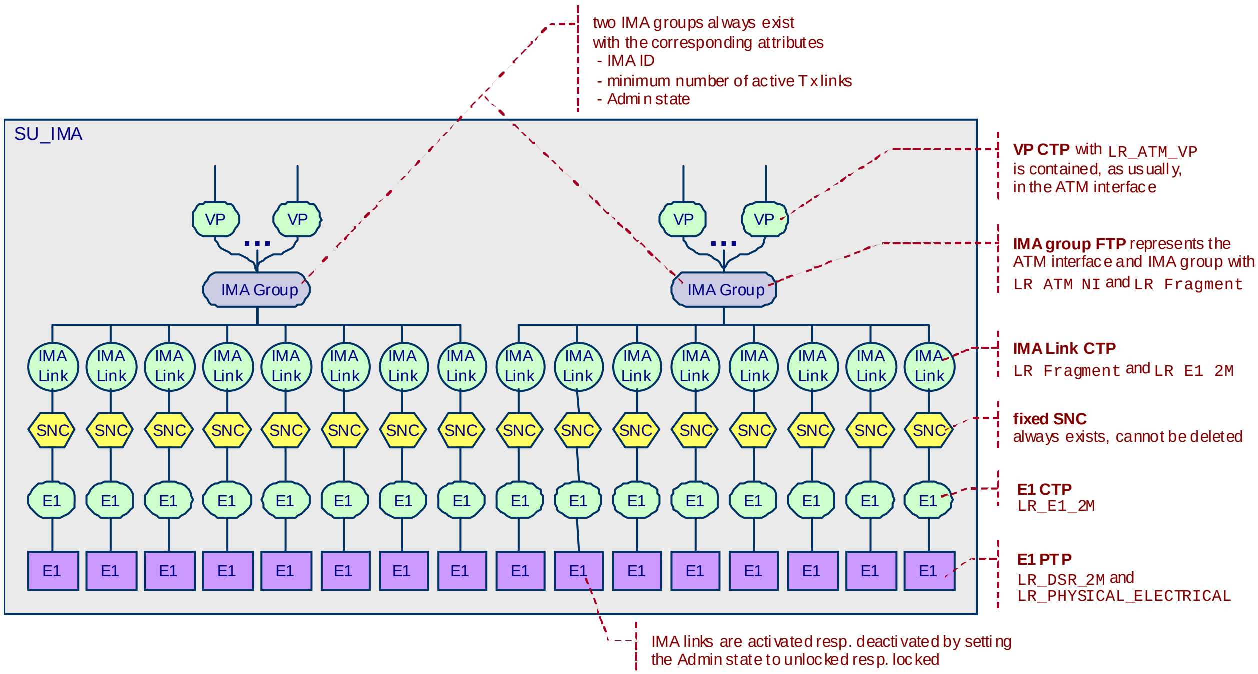

Figure 4 Example of an IMA Equipment with Two Fixed IMA Groups

- **Transmission Parameters for Inverse Multiplexing**

MTNM's transmission parameters for all layers (according to SD1-17) are specified in the supporting document SD1-16 in tabular form. The general parameters for the fragmentation layer rate `LR_Fragment` are listed in the section "General parameters for inverse multiplexing". They apply to all IM technologies, in particular to IMA/LASR and VCAT/LCAS, and refer to `LR_Fragment`'s adaptation function (e.g., the IMA group FTP in case of ATM and the encapsulation CTP in case of SONET/SDH). Technology specific IM parameters are listed in the sections "IMA specific parameters" and "VCAT specific parameters". They include in particular the parameters for `LR_Fragment`'s termination function which is always encoded in the fragments (i.e., the server CTPs of the fragmentation FTP or CTP).

- **Performance Parameters and Probable Causes for Inverse Multiplexing**

Refer to the supporting document SD1-28 for IMA-specific performance parameters, and to the supporting document SD1-33 for IMA-specific and VCAT-specific probable causes of alarms (at both the fragmentation layer and the respective connectable layer).

- **Operational Procedures and Use Cases for IMA**

The following **operational procedures** can be applied to IMA groups and IMA link ends:

1) read IMA group via `getTP()`
2) lock or unlock IMA group: either set "X.721::AdministrativeState" via `setAdditionalInfo()` (see SD1-8) or set "ServiceState" via `setTPData()`, depending on EMS support
3) increase or decrease the required minimum number of active Tx links: set "MinNumTxLinks" via `setTPData()`
4) modify the Tx IMA ID: set "TxImaId" via `setTPData()`
5) modify other transmission parameters of IMA groups, that are writeable by the NMS according to SD1-16, via `setTPData()`
6) read IMA link end via `getTP()`

7) for a fixed IMA group, lock (in a forced resp. deferred/graceful fashion) or unlock a number of IMA link ends in order to modify the transport capacity of the corresponding IMA group: either set "X.721::AdministrativeState" via `setAdditionalInfo()` or (possible for forced locking only) set "ServiceState" via `setTPData()`, depending on EMS support
8) create (resp. load) a flexible IMA group: create and activate SNCs between server CTPs of the IMA group and appropriate supporting CTPs
9) delete (resp. unload) a flexible IMA group: deactivate and delete all SNCs between server CTPs of the IMA group and supporting CTPs
10) modify the assignment of flexible SNCs to a flexible IMA group in order to modify the transport capacity or the routing targets of the corresponding IMA group: create and activate resp. deactivate and delete SNCs between server CTPs of the IMA group and supporting CTPs
11) request dynamic provisioning of an IMA group by the EMS subject to a prescribed bandwidth: use `setTPData()` with the parameters "FragmentServerLayer" and "AllocatedNumber"
12) provision the IMA virtual link between two peer IMA groups as a topological link: create resp. delete a TL having the two IMA groups as end points and `LR_ATM_NI` as layer rate

Since they refer to standard operational procedures on MTNM TPs (or SNCs or TLs) there is no too big impact of IMA group and IMA link administration on the MTNM deliverables and the corresponding use cases are more or less straightforward. The following IMA use cases are nevertheless specified in TMF513:

- "NMS locks (in a forced respectively deferred/graceful fashion) or unlocks a number of IMA links to modify the transport capacity of the corresponding fixed IMA group"
- "NMS requests dynamic provisioning of an IMA group by the EMS subject to a prescribed bandwidth that is communicated as number and connectable layer rate of the IMA links"
- "NMS provisions the IMA virtual link between two peer IMA groups as a topological link"
- "NMS unprovisions an IMA virtual link between IMA groups"
- "NMS creates a flexible IMA group"
- "NMS deletes a flexible IMA group"
- "NMS modifies the transport capacity or the routing targets of a flexible IMA group"

- **Use Cases for VCAT**

The following VCAT use cases are specified in TMF513:

- "NMS creates and activates a point-to-point Ethernet Service using fragmentation"
- "NMS modifies a point-to-point Ethernet Service with fragmentation"
- "NMS deletes a point-to-point Ethernet Service with fragmentation"

Note that these use cases are really not specific to Ethernet but apply also to any other service that can be encapsulated and fragmented across SONET/SDH (e.g., Digital Video Broadcast). The use cases are in fact concerned with the provisioning of the infrastructure trail at `LR_Encapsulation`.

Revision History

Version	Date	Description of Change
3.0	April 2005	
3.0	June 2005	References updated
3.1	November 2005	Version in names of referenced supporting documents deleted.
3.2	November 2007	Changed names of cross-references to the supporting documents.

Acknowledgements

<FirstName>	<LastName>	<Company>

How to comment on the document

Comments and requests for information must be in written form and addressed to the contact identified below:

Keith	Dorking	CIENA
Phone:	+1 678 867 5007	
Fax:	+1 678 867 5010	
e-mail:	Kdorking@ciena.com	

Please be specific, since your comments will be dealt with by the team evaluating numerous inputs and trying to produce a single text. Thus we appreciate significant specific input. We are looking for more input than wordsmith" items, however editing and structural help are greatly appreciated where better clarity is the result.

