Future Architectures: Agency Work Group Computing

Implications for Future SAFE

Summary

A concept for a new architecture for the Agency's work environment is proposed as an evolutionary replacement for the existing mix of clustered word processing (Wang) and centrally provided interactive services, including SAFE. The role of the current systems and the existing SAFE will continue while the transition is taking place. The new architecture migrates as much as possible of the work down to the work group level. By migrating in this direction in a timely manner, growth in size and degradation in performance of the mainframe systems can be capped at a critical point and stand alone processing can be more effectively joined with central processing.

The new architecture consists of personal computers as workstations with a small local area network for each physically co-located work group. The small LAN would include a file server, communications server, print server and be managed by customer personnel. Restricting the size and scope of the LAN to a physically secure area is essential to eliminating otherwise intractable security problems. The small LANs are connected as necessary either to mainframe systems for central service or via bridges between LANs.

The interface between the worker and any other systems is mediated by the PCs. For example, the current VM interface for the analyst to SAFE is entirely replaced for those analysts who have PCs. Those customers who do not have PCs will continue to see the current services, which will interface to the new architecture at regular points so as to allow a mixed population of old versus new environments. The central VM systems remain as general computing resources to which the PC users have 3270 terminal access and continue to serve as the "personal computer" for customers who do not have PCs or do not have the LAN connectivity needed to participate in the new architecture. Some number of such customers are to be expected in perpetuity.

Concurrent with the local computing environment, other changes to SAFE and central services can continue independently. For example, the back end data base technology for SAFE can be targeted for modernization. Eventual transition to a system employing a standardized "structured query language" (SQL) is desirable. The efforts to modernize the front end and back end are fundamentally independent, however. Modernizing the back end is largely an "under-the-covers" effort, while the front end change represents an enormous impact on the system as the customer perceives it. The bulk of this paper will address the front end architecture changes for this reason.

Background

To examine the new architecture, it is most instructive to see how it applies to SAFE. While not limited to "replacing" SAFE, seeing how such a replacement would look under the new architecture represents the best opportunity to prototype and flesh out the architecture with products. In this discussion, both the DI and DO are included as "SAFE" customers and the term "analyst" is used generically for all customers using the services described.

The current SAFE architecture and plans are for four VM front end "user interface" machines, with one back end MVS data base server. More specifically, the MVS back end for CIA handles cable receipt and dissemination, retrospective search of cables, and updates and queries of private index files, which initially will be cables that the analysts "save" with keywords. The VM front end provides the man-machine interface, electronic mail, composition, coordination and printing. The VM system provides basically the AIM system, plus interfaces to the cables handling components unique to SAFE. The system was designed so that the "personal" computing of the analysts was done on VM, with an intelligent, but non-programmable terminal on the desk. The design maximized use of available products and technology when it was chosen, but has not been kept up to date with new technologies, specifically personal computers and work group computing. Similarly, the back end data base system has been eclipsed in technology by newer relational models.

More important than the technology changes is the growth in analysts' requirements. SAFE does not meet significant requirements for an analyst's work environment, specifically in the areas of document production and user-friendly interfaces. Long standing "workstation" oriented requirements are not really part of the design. Rather than just try to "re-implement" SAFE under the new architecture, it is necessary to include current DI and DO ADP requirements as they have already been documented to OIT and see how the future computing environment should look.

By the time the SAFE architecture is fully deployed, many analysts will have personal computers on the desk, and the idea of a "secure analysts' file environment" must conceptually expand to include files and work done on the PC. Moreover, the fundamental overlap between the "personal computing" done on VM and that done of the PC has to be reconciled. The PCs also allow for new functionality and methods of presentation that can be exploited.

Specific Design

The following specific design is proposed as a strawman. Actual products will be called out and an attempt made to design the system by interfacing these products together with existing SAFE components and concepts. Actual product selection and integration would be a task to be performed if the concept is successfully prototyped and accepted. The goal in this discussion is to exhibit the feasibility and identify minimally qualifying solutions, not necessarily the "best".

Equipment:

On each analyst's desk, an IBM PC/AT compatible machine (generic or PS/2 mod 50) with as much screen resolution as can be afforded (1280 pixels wide is available on generics, 1024 on PS/2's so far). Each individual's PC would have a mouse, LAN adapter card and a floppy disk drive modified to be read-only. In vaulted areas, the PCs could also have internal hard disks for software and working storage. However, these hard disks would not be used for storage of anything the analyst wanted to share with anyone and would be strictly up to the individual analyst to backup onto the LAN file server's disks himself. Policy should preclude data bases or anything of value as an Agency "corporate" resource on such private hard disks, and even "private"

information such as PARs and personnel data should be excluded by policy from such disks since they are unprotected from access by anyone with physical access to the PC. For at least a transitional period, each PC should also have a 3270 communications card and attach to the PBX for interactive host access. Theoretically, with the right LAN products, host access can also be achieved via a communications server on the LAN, but the PBX phone access has some advantages for the moment, especially for LANs other than IBM's Token Ring. An alternate link via a LAN based communications server should be provided in any case.

Two specific LAN configurations and products should be supported if at all possible. One is an Ethernet LAN, built from 3COM Etherlink Plus cards, but running Novell's Advanced Netware. Novell has by far the best network and can include excellent security (actually "privacy" in the sense that it is not certified as secure, nor do we care). The other is IBM's Token Ring. For offices with 3270 PC/AT or PS/2's, this is the only real option currently. A hybrid network, running Novell software on the IBM token ring network hardware should also work. This is the best choice for "hedging" one's bet and would be what a pilot office should be wired for unless specific Ethernet requirements are identified.

The wiring would then consist of data grade (equivalent to IBM Type 1) copper wire run inside easy to install Panduit-type covered cable runs along the inside walls of a vaulted area to a chosen point where the actual network is formed by joining the individual wires together via an IBM Token Ring junction box (MAU). Either at this point or in a separate "computer equipment room" (CER) such as where the Wang Alliances sit, the LAN server is placed and wired to the common point. This wire run, and only this one, might be through the building's secure grid of either copper or fiber, depending on distance. It is desirable to physically isolate the LAN server from the office and only allow physical access to it by the system's administrator, as is done for the Wangs.

The LAN file server has enough disks to support the customer population, with sizes of 100MB being typical. The file server would best be either a 386-based or specific Novell 68B server optimized for that task. The 386 has the advantage of doubling as a print server, or even also as a communications server, but putting the print server in the customer office area is preferred anyway, so the 68B would be fine. The Novell LAN server can be configured to be fault tolerant, including automatically double-writing all disks and automatic back-up to tape. It has the best "security" access logs and permission control by the systems administrator, and actually has enough security access control to consider putting the server in the customer's vaulted area instead of a separate CER.

The print server has a high quality laser printer such as the IBM 3812 or newer Pageprinter, and can probably double best as the communications server. It needs to be only a regular PC/AT or PS/2 machine just like the general desk-top model. As the communications server, it provides alternate path host access via an SNA link. For a TRN network, this "alternate" link can be via a TRN "spur" all the way to a mainframe channel-attached 3174 controller at 10 Mbps without even involving the communications server. Thus TRN network with this kind of spur may supplant the need for PBX connections unless the customer has needs to access various systems (like CAMS or DESIST) that can be

reached via the PBX but not via the SNA network that includes the internal Agency systems. The interoperability of such a TRN connection and Novell LAN software is currently unknown, however, which is why parallel investigation of operating over a fully IBM network (IBM LAN software based on Microsoft LAN technology) is needed.

The main function of the communications server is NOT mainframe interactive access, however. The communications server is primarily tasked to provide store and forward exchange of data between the LAN and other LANs or mainframe systems. For this purpose, it needs one or more communications adapters. It is desirable to have direct LAN-to-LAN exchange of items without any mainframe dependencies at all, the communications servers can form a network of their own. In simplest implementations, dial up (secure) phones such as a DIU interface to the headquarters secure PBX would suffice, or there are LAN "Gateway" products that can be investigated. If these gateways can be physically secured, they may offer adequate compartmentation security to link two local work group LANs without allowing customers on one to access data on the other. Research in this area is required before one can safely say the mainframes don't provide the best answer, however. The most important interface for SAFE will be the link between the communications server and the mainframe. This will be a "program-to-program" link between an electronic mail system on the LAN and a host mail server. One method for this is a synchronous communications adapter that can support SDLC/LU6.2 protocols.

Software:

The user interface should be based on a graphical presentation. The migration path for achieving this has been identified by Microsoft and it is to use Windows now and later Windows 2 for DOS 3.x. When and if OS/2 1.1 with the presentation manager becomes desirable, code can migrate to it. The LAN enables a mix DOS and OS/2 machines so this choice can be deferred. The LAN has to provide the office environment for the analyst -- his files, word processing, desk-top publishing, shared authorship/coordination, calendaring, printing, etc. The actual production files should be on to LAN file server, which provides virtual DOS disks with the advantage of being able to share them between customers. By the time this environment goes into "production", I would expect to have a Windows version of MS Word as the answer to both word processing and desktop publishing sufficient for the DI's need. A few offices may also want Aldus Pagemaker, but only for the IA or CPAS careerists, not every analyst. Windows EXCEL should also be available to provide spread sheets compatible with the rest of the Windows suite, although some use of Lotus 1-2-3 will remain, I'm sure.

The less obvious software is the "mail" type interfaces to other analysts, other Agency components, and to the cable systems, including the SAFE dissemination and retrospective search capabilities. The basis for these services in the new architecture has to be an electronic mail and filing environment on the LAN. It turns out that this requirement can largely met by just using DOS files and subdirectory structures along with conventions and a bit of software glue. A prototype of this sort of system is the "Coordinet"
STAT project implemented for CPAS by [redacted] OIT/SAD&E. With relatively little effort (compared to most SAFE development at least), this could be expanded into a usable environment. However, another alternative is to go far beyond the "electronic mail" metaphor and try a commercial LAN-based product

called "The Coordinator" which combines electronic mail with a sophisticated model of paperwork flow and enforces such concepts as action items, suspense dates, and so on. Evaluating such a radically different product will take effort, but is recommended. The status of a graphical (Windows) interface to it is doubtful, however, so it may still take a combination of the "Coordinet" approach as an interface and use of the LAN's mail system under the covers, for example. Fleshing out this area is the primary work involved in prototyping the new system.

On the back end mainframe side, the software effort consists of recasting the services represented by SAFE into a "transactional" model that can be used to service the LAN-based customer in a background fashion. This work would actually proceed in two phases. While prototyping, the services would actually come from the VM front end, where such prototyping is easiest. The analysts would be prompted when necessary for VM passwords so that the PC could, behind the scenes, request services and have things brought down to the LAN in the background. Later, the direct VM link would not be necessary, as the communications server would request that the information be brought down continuously and via a back-channel link, probably to the MVS mainframe.

Concept of Operation

In the full-up architecture, the following is a discussion of the concept of operation and the division of labor between the mainframes and PCs.

The MVS server would continue to do the cable dissemination via profiling as is done today. No apparent need to re-host this process is evident, although something like a Connection Machine could do the process is it ever became a performance bottleneck. The retrospective data base of cables could be kept in INQUIRE on the MVS processor as it is today, or could migrate to a back end data base machine (such as a Teradata) as the processing requirements continue beyond the cost effectiveness of the IBM architecture. The MVS processor would also be a natural host for a store and forward electronic mail "hub", and efforts are underway to use it for this purpose.

The VM processor would continue to be the "PCs" for various people and continue to run SAFE Delivery 3.x code for them. It would also be the host of the AIM system for such people, and communicate with the other AIM systems in the Agency. The bridge between the AIM systems and the LANs and all other "departmental" sorts of systems would be via the planned AIM Gateway to the MVS server, an LU6.2 SNADS link.

PC customers under the future SAFE architecture would receive both AIM mail (from non-yet-converted analysts or other remaining AIM customers) and cables disseminated from the MVS SAFE profiling systems on the PC LAN. This mail and cable traffic will have been forwarded to the LAN and reside on the LAN file server in the PC user's mailbox on a continuous, background basis. The distinction between AIM mail and cables should be largely eliminated, although they should have different "categories" or action types associated with them so the analyst can choose which he wants to review at any one time. Once the mail or AIM file has arrived on the LAN servers, it is deleted from the host side. The decision to retain things and provide the storage for them becomes the responsibility of the customer office entirely. Only the retrospective cable file and the AIM documents actively in use by AIM customers remain on

the hosts, capping the disk storage growth.

"SAVE" files, including PIFs, should be only for the Delivery 3.x customers, not provided at the host for LAN-based customers. The idea of back end data base services for the LAN-based customer should be considered as part of the long term requirements, but not essential for SAFE per se. An indexed collection of corporate information, documents, "very interesting cables", and so on should clearly be maintained centrally, but private index files belong on the LAN. The larger issue of a central store of corporate information is a subject for even farther future consideration. Eventually, out of the SQL standards and the OS/2 "Extended edition" concepts of data base services across networks, one should be able to give the LAN-based analyst access to central data bases and LAN-based databases in a transparent way. Research to see how relational data bases fit into the analyst's environment on the the LAN and on the hosts should be part of the overall effort, but not to be expected in the early versions.

In actual implementation, the MVS dissemination process would still produce mail files of hits to be kept for 30 days. A new process would recognize that certain mailfiles are being followed by analysts who are served on LANs rather than as direct interactive customers. The new cables added to these mail files would be "mailed" down to the individuals via their LAN communications server, along with any AIM mail addressed to them. Some optimization for multiple people on the same LAN getting a hit on the same cable is possible, but needn't be implemented for prototyping and may not even be worth the effort unless the LAN mail system already has provision for this sort of multiple copy addressing. The process of updating profiles is already electronic mail-based in AIM and can be migrated over to PC mail in a fairly obvious fashion. The LAN-based copies of the mail can be kept for 30 days locally if fast searching of them is desired, or the customer can just keep the ones that look interesting on his own file structures as long as desired. The need for a 30-day mail file on the server is an open question that the prototype phase should answer.

Retrospective searching of the combined document file (all the cables ever received) is an interesting design issue. In delivery 1 SAFE, it didn't exist -- only 30 day files were kept. In Delivery 3.x, retrospective search is almost immediate via INQUIRE. In Delivery 2 SAFE, AIM SEARCH is used for searches of "SAVE" files, which is an "electronic mail" interface, with response coming back significantly later via a return message. In the new architecture, a retrospective search that produced a lot of hits would be a big problem if the hits had to come down the communications channel to be reviewed. One potential architecture would be to send the search request as a mail message, just like in Delivery 2 for AIM search, and get all the hits mailed back into a LAN-based mailbox. Limiting the number of "hits" would be mandatory. An alternative is to establish a more interactive connection to a searching task and review the hits without "mailing" them down to the LAN. For example, one could have the analyst "log on" to INQUIRE directly on MVS or to an application like Delivery 1's "TEXT" that actually called INQUIRE under the covers to do retrospective search. The analyst could then request any given "hit" or list of "hits" to be be brought down to the PC for use in composing a report, for example. Choosing between these models, or finding others, is a area to be worked.

Outgoing cables are another interesting area that the LAN solution needs to address, even though SAFE today does not. Presumably, all cable composition and branch-level coordination occurs on the LAN. Given the need for a laser printer on the LAN, one of the requirements for equipment selection of the actual printer hardware should be the ability to produce hard copy cables that can be accepted by the OIT cable systems. One should never have to assume that a mainframe interface is up in order to produce and get a cable out. However, if the mainframe interface is available, by the same token, one should never have to print a hard copy. The software that produces the cable locally should also be able to address it out to the mainframe mail hub and out to the cable network. However, the process of coordinating, authenticating and releasing a cable is probably not commercially available on the LAN-based mail system in a "trusted" way. Unlike AIM, sending a document through a train of people and being able to guarantee that it has gone through them is not a feature of any distributed mail system known. It may be possible to recognize safely that a single individual on the LAN is the one who actually forwarded the cable up to the host, but that is about all one can expect. In instances where that individual has releasing authority, cables could be originated all electronically and actually sent out, but this needs to be investigated a great deal more before electronic cable origination from LANs, or any other departmental computer schemes, can be considered feasible. Note that this is a significant step backwards from AIM's level of capability for authenticated routing and coordination which was designed to be secure enough for cable origination, but was never entirely accepted for that purpose even so.

Stand-alone or Remote Use

One of the main advantages of the new architecture is that, with the possible exception of retrospective search interfaces, it does not require that the mainframe systems be constantly available. They only have to be up long enough, often enough, to get the "mail" through in a timely fashion. If the LAN mail software has alternate routing capabilities for LAN-to-LAN transfer, the availability of the mainframes is further de-emphasized. Note that for the LAN-based customer, no VM availability at all is required in the long run. The link from the LAN to the mainframe mail server can be of relatively low bandwidth if it is available most of the 24 hour day. It need only keep up with the arrival rate of mail and cables (again excluding the retrospective search burst arrival possibility). Statistics on how many analysts are in a physically co-located area and how many cable "hits" they generate in total are needed to see how much bandwidth is required, but it would be surprising if more than a 9600 baud link per LAN were needed to provide "same hour" delivery as required by SAFE for cables. Thus the architecture is not limited to headquarters environments with high bandwidths to the hosts.

In fact, it is quite possible to envision the same architecture where there are no host links at all. The LAN, the PCs, the cable "authoring" software, the office "coordinator" functions all function independently of the SAFE cable dissemination stuff. The same system, subject to EMI security criteria, could just as easily be in the field. The goal of skills portability and symmetry of systems as embodied in the DO's DOLPHIN concepts is met by this architecture. The capacity of this architecture to replace CRAFT seems obvious. And one can take the architecture even farther along this line by hooking the communications server to an interface to the narrative message

network in the field in lieu of the MVS mail server in headquarters, but making it look basically the same to the end customer -- a "mail" interface. The problem of cable releasing authority still exists, but in the field environment may be more tractable where a single releasing authority is more the rule.

Over time, the "mail" metaphor may aid the eventual migration of the narrative networks and the data networks into integrated systems so that the headquarters and field customers really do have the same interfaces. Looking at "mailed" transactions for retrospective cable searches leads to thinking about such transactions for other data base queries such as name traces where the query might be sent not to a data base but to a person who could access data and rapidly transform the results back into a "mailed" response without compromising the security of the data base itself. Likewise, the objects being "mailed" should not be restricted to just narrative messages. Current PC technology allows mailing of images, for example, so the capacity to use electronic mail in lieu of "fax" is already evident.

Similarly, whether for headquarters or remote customers, there is no reason to assume that there is just one central server for cable searches or anything else. The distributed model for new SAFE makes multiple "back ends" more feasible and allows for easier incorporation of new data base engines or alternate computer sites.

Two Tiers vs. Three

The proposed architecture can be characterized as "two tier" in that the processing resides in the PCs and on the mainframe. The file server on the LAN is primarily just another PC providing shared disk access and maintenance. The electronic mail system doesn't really "run" on the server in the sense that customers "log on" and use it. It is the communications task that is doing the mail work -- the rest is done by software in each PC accessing shared files on the server. This is in contrast to the "three tier" departmental computing model, where there are significant applications running on the office-level system and customers "log on" to the departmental machine to use it as a computer. On a LAN, customers enter a password to be able to get to their disks, but that's the extent of the "log on".

The architecture would certainly allow for a departmental computer as part of the LAN, but there appears to be no requirement for it. If the level of "office automation" provided by the combination of PCs and a mail system is not sufficient for some reason, then use of a "departmental" system as a file server plus local "log on" computing can be employed where necessary. However, the burden on the customer office in running a departmental system as opposed to a LAN alone is sufficiently high to design "future" SAFE to consist of only two tiers as the norm. One of the main features of the new architecture is that customer offices operate their local computing, including the LAN and its servers. A systems administrator, similar to the administrators now used for Wang Alliances, must be provided by the customer office. In the two tier model, the LAN and its servers are just more PCs and represent the minimum impact on the customer, whereas departmental computers are generally associated with higher skill levels and effort in systems administration. Another advantage of the two tier approach is that the PCs and the LAN represent relatively small chunks of investment and allow for a

great deal of heterogeneity among brands, even on the file servers. Once departmental computing is introduced, standardization across large segments of the Agency and investment in particular make and models of computers becomes a significant problem.

The proposed architecture does have the problem that IBM compatible workstations are presumed. Other workstations that could coexist on the same LAN and use the same file server and file structures and the same LAN-based mail package could be envisioned, but they do not exist in practice. Developing any code at all on the workstation (such a creating cables) or in the communications server will mean a lock in not to a particular piece of hardware, but to a particular operating system environment, namely DOS with an option to migrate to OS/2. Implementing the architecture for multiple workstation types is not a trivial, and perhaps even an impossible, task and is not envisioned. If the Agency cannot choose a single operating system environment for workstations that are to participate in future systems like SAFE, then a substantial delay in implementation is required, and a reduction in function to even more minimal "mail" interfaces is essentially unavoidable. Although cooperative processing models and transactions all may become standard commodities, once applications code is written at the workstation level to use these standardized interfaces, then the issue is not "standards" but code portability. And the portability of full scale, complex applications including communications, graphic user interfaces, mail systems and data bases from IBM (Windows) to Apple (Macintosh), for example, is essentially a matter starting over at the design phase. There can be no half measures at this point; making a choice and living with it are required.

Actions

The basic components for the new architecture are already available. Work should proceed to assemble and test the necessary components as identified above.

An effort to prototype an analyst's work environment using the new architecture in place of the existing SAFE interactive interfaces should begin.

EG must acquire the systems and personnel to put the tools and interfaces in place, to be followed by a joint prototype development project with customer offices to make the new architecture apply to a system like SAFE, but including current customer ADP requirements as well.

CSPO should cooperate in this prototyping effort and in providing development resources to add interfaces to the existing SAFE systems for dissemination and retrospective searching that can be exploited by the new systems without disrupting the old.

IMS should seriously consider this architecture as a candidate for the base on which to build DOLPHIN and should participate heavily in the cable creation requirements for the PCs.

An evolutionary migration of customers from the old systems to the new should then be planned, with investment in LANs and workstations planned as early as 1989.

OIT should make all components aware of this direction as soon as possible for inclusion into other, non-SAFE systems, including administrative systems and compartmented projects that would otherwise be implemented on word processors or departmental computers.