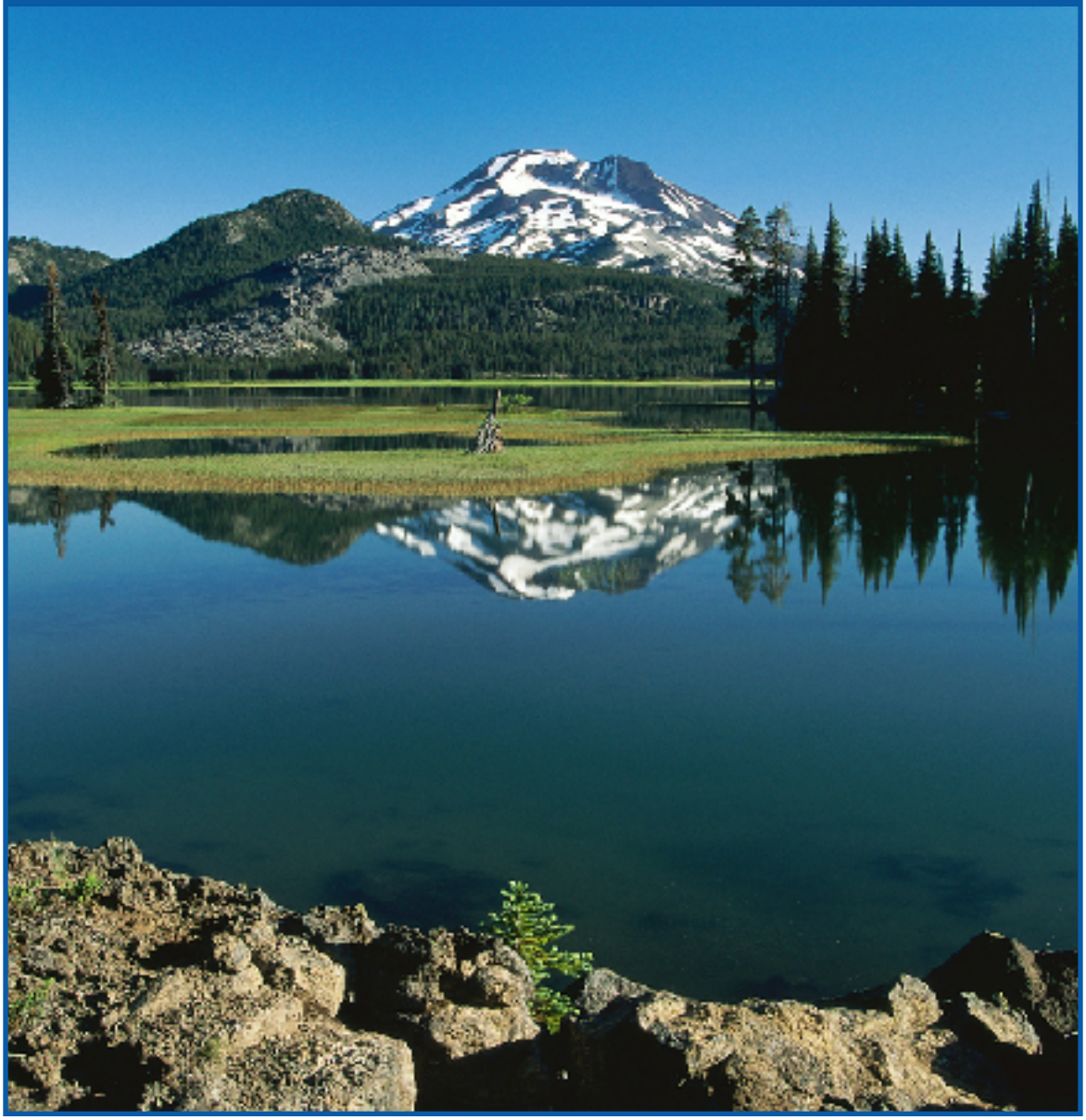
EPA

September 8, 2003

U.S. Environmental Protection Agency

Enterprise Architecture

Status Report 2003



This Page Intentionally Blank

Table of Contents

Foreword iii
Acknowledgments v
Highlights ix
Introduction I-1
Status Report EPA Enterprise Architecture 2003 ..... SR-1
Exhibit A EPA's Enterprise Architecture Version 2.0 ..... A-1
EPA's Enterprise Architecture Version 2.0 ..... A-3
Exhibit B Revisions and Updates to the Target Architecture B-1
Part 1 Strategic and Business Architectures: Alignment of Strategic Architecture with Business and Budget Models B-3
Part 2 Data Architecture: Evolution of the EPA Data Architecture B-15
Part 3 Updates to the Applications Architecture B-21
Exhibit C Enterprise Architecture Sequencing Plan C-1
Introduction to the Enterprise Architecture Sequencing Plan C-3
Part 1 Aligning EPA's Target Enterprise Architecture with E-Government C-5
Part 2 Environmental and Health Protection Architecture Sequencing Plan C-11
Building the Central Services C-15
Schedule C-37
Part 3 Administrative Systems Architecture Sequencing Plan Update C-39
Part 4 Sequencing Applications to the Target C-53
Part 5 Technology Architecture Sequencing Plan C-57
Overview of the Technology Architecture Sequencing Plan C-59
Technology Architecture Sequencing Plan C-63
Technology Architecture Sequencing Plan Diagrams C-67
Addendum - Notes to Technology Architecture Sequencing Plan Diagrams C-77
References C-83
List of Acronyms Acronyms-1

This Page Intentionally Blank

Foreword

As I look back on the brief history of enterprise architecture (EA) at the U.S. Environmental Protection Agency (EPA), I am impressed and proud of the progress we have made. When we submitted our first baseline architecture, prudently titled "Version 0.8," it seemed little more than another formal exercise in "good management." Even when that document won the Office of Management and Budget's (OMB) approval as one of the best submitted anywhere in government, I don't think any of us guessed how important EA would become.

A year later we were struggling to get our arms around the new concepts, complex ideas, and difficult abstractions of what was to become our target architecture. Like traditional architects, we had to take all the requirements, demands, and constraints of our customers and forge them into a single design—if possible, an elegant, buildable, and affordable design. The job was, frankly, considerably more difficult than we'd expected. When our target architecture was published in December 2002, it was due in large part—as I think everyone in the Agency will acknowledge—to the vision, imagination, and energy of our Chief Architect, John Sullivan. Yet it was always a joint effort. John and his team were aided and actively supported by top staff from across EPA. They also had the good fortune to be building on a number of ideas that had been at large in the Agency, in one form or another, for over a decade. Ideas like the Environmental Information Exchange Network, the Model for Integration, and the Central Data Exchange.

During the last nine months we began to see the payoff. We focused, as the President directs, on a limited number of priorities. We identified our most pressing problems—and set out to solve them. The measure of our seriousness is seen in our first action: creation of the new Program Management Office (PMO). Its unassuming title belies its sweeping mission—nothing less than to deconstruct the dozens of mission-critical applications that run EPA and rebuild them around an entirely new infrastructure. To run the PMO we tapped Mike Cullen, one of EPA's most talented and experienced managers. He is building out his team with more top talent from across the Agency. So in addition to an architect, now we have a general contractor.

Our second priority this year was to move beyond the applications and data layers of our architecture and focus on integrating business, strategy, and governance. We aligned our architecture with all the federal reference models issued by OMB, as well as with our new EPA Strategic Plan. We revised our budget structure and human resources planning process so that everything fits together. Our Capital Planning and Investment Control (CPIC) submissions this year are tightly mapped to the architecture. We now have, as OMB describes it, a "clear line of sight" between our services to citizens and the programs and systems that carry them out. Results—as the President's Management Agenda emphasizes—are the bottom line.

This document presents the highlights of this year's work. Version 0.8 raised hardly a ripple in 2001. In 2003, enterprise architecture is the center of attention. I look forward to 2004.

KIMBERLY T. NELSON

Assistant Administrator and
Chief Information Officer

Office of Environmental Information
U.S. Environmental Protection Agency

September 2003

This Page Intentionally Blank

Acknowledgments

The success of an Enterprise Architecture (EA) lies in its ability to be used for business transformation. Thus to acknowledge those who have made the EA successful is broader than just the immediate EA domain.

I would like to thank those who have been involved in further developing the EA. Members of the following workgroups and several significant Agency personnel have made the products of these groups even better:

- The Enterprise Architecture Coordination Committee, for their endurance and commitment to the process, and individual work in exploring programmatic components of the EA with their regional counterparts
- The Administrative Systems Architecture Workgroup, under Lisa Ayala's leadership, for helping align the EA with key E-Government processes
- The Research and Science Architecture Workgroup, under Tom Tracy's leadership, for helping define the research and science components of the EA
- The Technology Architecture Workgroup, under Steve Hufford's leadership, for ensuring the Agency's technical infrastructure planning will support the needs of the target architecture
- The Geospatial Blueprint Team, under Wendy Blake-Coleman's leadership, for leading the direction in setting the geospatial analysis capability of the target architecture
- The ART Users Work Group, under Ethan McMahon's leadership, for their work in developing the agency architecture model
- The Lead Region for Information Management, under leadership from Karen Vasquez and Debra Forman
- The Regional Knowledge Sharing Workgroup, and the work that Debra Forman and Jon Schweiss have led to help define the regional component of the architecture
- The IT Security Staff, under the leadership of George Bonina for their contribution to the Security Architecture

I would like to thank those who have worked with us to integrate the EA into key Agency management processes:

- Mike Feldman and team for working with us to ensure that the EA aligns with the new Strategic Planning and Budget processes of the Agency
- Kirk MacConaughey for working to align the Agency's human capital planning efforts with the EA
- Chuck Cavenaugh and the Capital Planning and Investment Control (CPIC) team for working with us to align the EA with the CPIC process

I would also like to thank those who have worked to develop related components that further the EA. It is the sum of all this work that builds the Agency's EA:

- Sara Hisel-McCoy and the state/EPA action team developing the Core Reference Model
- Chris O'Donnell for working to ensure the Agency's document and records management needs of the Agency are addressed in the EA
- John Harman and Larry Fitzwater for their work on metadata management and the Registry of EPA Applications and Databases (READ)
- Mike Cullen and the new Program Management Office (PMO) for taking the architecture and doing the hard work of putting it into action and starting to build it

Executive sponsorship has been an absolutely essential ingredient to making this successful. Thank you to Mike Ryan and the ASA Executive Steering Board, Henry Longest and the RSA Executive Steering Board, and QIC Technology Subcommittee (QTS) for ensuring that these critical areas are focused on. Thanks to the CTO, Mark Day, our champion and sponsor, for providing us with support and guidance. Last but not least—gratitude for the promotion and championing of EPA's Enterprise Architecture Program that comes from our CIO, Kim Nelson.

While it is this wide participation across the Agency that makes enterprise architecture planning effective, none of it would happen without the devotion of the core Enterprise Architecture Team within OEI's Office of Technology Operations and Planning and our primary Industry partners—SRA International, Inc.—for doing the heavy lifting of EA development. To them I am most grateful.

Thank you.

JOHN J. SULLIVAN

Chief Architect
Office of Environmental Information
U.S. Environmental Protection Agency

EPA Enterprise Architecture Contributors

Joe Anderson, OEI
Lisa Ayala, OCFO
Andy Battin, OW
Len Betchel, OARM
Wendy Blake-Coleman, OEI
George Bonina, OEI
Chuck Cavanaugh, OEI
Sanjib Chaki*, OEI
Barbara Chancey, OEI
Mike Cullen, OEI
Walter Dove*, OEI
Gene Durman, OAR
Mike Feldman, OCFO
Larry Fitzwater, OEI
Laurie Ford, OEI
Debra Forman, Region 3
Bill Grabsch, OEI
Leo Gueriguian, OW
Bob Gunter, OEI
John Harman, OEI
Sara Hisel-McCoy, OEI
Jody Hudson, Region 7
Steve Hufford*, OEI
Lisa Jenkins, OSWER
Joan Karrie, OPPTS
Matt Leopard, OEI
Kirk Maconaughey, OARM
Mark Mahoney, Region 1

* EA Team

Anne Mangiafico*, OEI
Connie Martin, OEI
Joanne Martin, OPPTS
Ethan McMahon*, OEI
David Michals*, OEI
Judy-Ann Mitchell, Region 2
Rebecca Moser, OEI
Cheryl Newton, Region 5
Chris Nugent, OECA
Chris O'Donnell, OEI
Terry Ouverson, OCFO
Dave Patton, OCFO
Kevin Phelps, OEI
John Richards, OPPTS
Caren Rothstein-Robinson, OPPTS
Carolyn Saunders, OECA
Jon Schweiss, Region 10
Tony Studer*, OEI
John Sullivan*, OEI
Ken Tindall, Region 5
Chris Tirpak, OPPTS
Tom Tracy ORD
Karen Vasquez, Region 5
Fred Weeks, Region 1
Nathan Wilkes*, OEI
Dave Wolf, OEI
Jody Zeugner, OEI
Paul Zanter, Region 5

Business Partners

Michael Alford, SRA
Jeff Byron, Byron Consulting
Stephanie Cammarata, SRA
Kerry Casseta, AMS
Darby Chellis, SRA
Jack Deppe, SRA
Ben Didier, SRA
Joshua Fischer, SRA
Chris Galanty, Booz Allen Hamilton
Ted Gonter, SRA
John Horstman, AMS
Tony Ireland, SRA
Peter Lang, SRA
Pat Little, SRA
Megan McCloskey, SRA
John Moriarty, AMS
Jennifer Murphy, SRA
Angela Parks, SRA
Christopher Poillon, SRA
Neal Richards, SRA
Nan Tian, SRA
Paul Triulzi, Perot Systems

* EA Team

This Page Intentionally Blank

Highlights

Aligning EA with the President's Management Agenda

- We have aligned our Enterprise Architecture (EA) to the Agency's planning and budgeting processes. The Office of Management and Budget's (OMB) Federal Enterprise Architecture's (FEA) Business Reference Model (BRM) and EPA's BRM now map to the EPA budget categories to form a common business framework for planning and accountability.
- The 2003 Strategic Plan's performance objectives now map to the federal Performance Reference Model (PRM) for the FY 2005 investment proposals.
- EPA is a Solution Partner in E-Gov and a player in 14 of the 24 E-Gov initiatives under the President's Management Agenda. All of our E-Gov initiatives map with our Capital Planning and Investment Control (CPIC) proposals.
- We have integrated EA and CPIC. EPA is a leader in this arena, sharing its alignment tools with other agencies.
- Our EA governance structures are strengthened. The Chief Financial Officer, Chief Information Officer, Chief Technology Officer, and Chief Architect have worked together to produce the common business framework.

Improving the target architecture to make it actionable

- We have initiated aligning the applications layers of the Environmental and Health Protection Architecture (EHPA) and the Administrative Systems Architecture (ASA). Both will be served by the central services of the target EHPA.
- We are building an enterprise data model to guide construction of the common Framework for Business Warehouses (FBW). This model will be based on a Strategic Information Model (SIM) derived from the Agency's Strategic Plan and goal structure. It will also draw on the Core Reference Model (CRM) of the Information Management Work Group—a common framework for sharing data across the Exchange Network.
- EPA continues its work implementing the target architecture, revising the target to reflect funding and business priorities and extending the baseline ASA architecture to include the regional and program offices.

Starting construction of the central services

- The Office of Environmental Information (OEI) has set up the matrix-managed Program Management Office (PMO) to construct the EPA central services.
- The PMO's workplan calls for all elements of the central services to be operational in 2005, with one pilot system (in the process of being selected) to be running within the central services by that time.
- The PMO is examining new options to construct the FBW using Web services to link Business Area Warehouses and Operational Data Stores.
- Once the pilot system is operational, the Agency's other systems will migrate to the target EA in three "waves." First priority will be given to mission-focused systems, new systems, and "major" CPIC systems.
- The PMO and the EA Team will collaborate on an integrated cost-benefit methodology to be used in the next CPIC cycle to quantify costs and benefits of the target EA.

This Page Intentionally Blank

Introduction

On December 16, 2002, the Environmental Protection Agency (EPA) submitted its *EPA Target Enterprise Architecture* to the Office of Management and Budget (OMB). This document presented the full dimension of EPA's Enterprise Architecture (EA) program, including its use of the Federal Enterprise Architecture Framework, the structure of its system of governance, selection of a Chief Architect, establishment of EA policies and procedures, and creation of an EA Team supported by representatives from across the Agency.

This document adds to that submission—it is not a stand-alone presentation. It outlines the Agency's progress since the December 2002 submission. It updates the EPA Target Architecture and presents version 1.0 of the EPA sequencing plan. It also presents the full current version of the EPA Enterprise Architecture in electronic form. The accompanying CD-ROM contains an executable copy of the EPA Architecture Repository and Tool (ART), which is also available to staff throughout EPA via the Agency's intranet.

The requisite documentation for investment business cases required under EPA's Capital Planning and Investment Control (CPIC) process is included in a companion submission. This documentation includes required risk management plans, security plans, quality assurance plans, contracts, cost-benefit analyses, economic analyses, and alternatives analyses. All CPIC proposals are documented in EPA's I-TIPS system and are mapped to the EA on the accompanying disc.

Over the past nine months, the EPA EA effort has been systematically aligned with the multiple goals of the President's Management Agenda (PMA). EPA has linked its EA and CPIC processes in multiple ways described in this submission. EA is now recognized as a primary authoritative resource of IT planning, from the highest executive levels down. The Agency has strengthened and extended its EA governance processes. It has aligned EA not only to the Federal Enterprise Architecture (FEA) reference models, but also to EPA's budget process, performance tracking system, and soon to the human capital planning process. E-Gov initiatives throughout the Agency are now aligned with the architecture, and several are being positioned to support cross-federal lines of business.

This submission begins with a short Status Report that provides an overview of these and other major developments and how they tie together. It then presents three exhibits that provide more detail. Exhibit A discusses deployment of EPA's Architecture Repository and Tool (ART) and its availability across the Agency to support all levels of EA activity, including program-level architecture work. Exhibit B presents modifications to the target architecture since December 2002, including further definition of the structure of the Enterprise Repository, now referred to as the Framework for Business Warehouses (FBW). Exhibit C presents the EPA sequencing plan.

This Page Intentionally Blank

Status Report
EPA Enterprise Architecture 2003

The Environmental Protection Agency (EPA) is pleased to report many positive developments since it submitted version 1.0 of its target enterprise architecture to the Office of Management and Budget (OMB) nine months ago. We took an ambitious architectural design with far-reaching consequences and turned it into an actionable blueprint. We are now building aggressively from that blueprint.

This year we have repositioned our enterprise architecture (EA) effort to respond to the President's Management Agenda (PMA) on multiple fronts. Much of our work has focused on the highest levels of the EA pyramid: strategic planning, business modeling, budgeting, human capital planning, and performance assessment. We have integrated EA into our budgeting process and our Capital Planning and Investment Control (CPIC) process, including aligning our CPIC proposals with the PMA E-Gov initiatives. We are positioning several of our E-Gov initiatives to become cross-government providers of lines of business. Our EPA is specifically designed to support the environmental information Exchange Network, a major federal/state E-Gov consortium that reduces reporting burdens on states and industry and creates a new paradigm for sharing environmental information among federal and state partners. It offers states and industry a central point of contact for their business transactions with EPA. We have also integrated our architecture with the Federal Enterprise Architecture (FEA) reference models and our new Strategic Plan, allowing us to align our performance goals with our budget and with the Federal EA. We are in the process of extending this integration process to human capital planning.

Our second general focus this year has been to expand our architecture and begin construction. We have built out our EA with more regional and program office content and have started to integrate our separate "business domains" into a single, comprehensive architecture. We have begun construction of our Environmental and Health Protection Architecture (EHPA), setting up a new office to build the EHPA central services infrastructure and moving in key staff to manage it. Our Research and Science Architecture (RSA, formerly the Research and Development Architecture) will complete its target architecture in January 2004 and its sequencing plan in early to mid-2004. It too will align with the EHPA central services.

Our last two submissions, presenting EPA's baseline and target architectures, totaled some 1,800 pages. This submission, including summaries of improvements to the Target Architecture and the presentation of the first iteration of our sequencing plan, is less than 125 pages. Our intent is to report briefly on the highlights of what we have achieved and what we have yet to do. Detail is presented in electronic form on the accompanying CD-ROM, which contains an executable copy of version 2.0 of the EPA Enterprise Architecture (EA). Narrative backup is provided via hypertext links.

As we go forward, EPA continues to dedicate itself to improving its level of services to citizens in protecting and enhancing public health and the environment through results-based planning and management. It will continue its close cooperation with the states, primarily through development

of the Exchange Network to minimize burdens of reporting and data exchange. EPA investments supporting E-Gov initiatives will help us coordinate ever more closely with our federal partners.

We have aligned our architecture with the President's Management Agenda

Over the past nine months we have shifted the EA spotlight to the strategic and planning layers of our architecture. The heart of EPA's version 1.0 architecture submissions (September and December 2002) was application and data integration. Building on ten years of effort, we used EA to address our most pressing IT problem—fragmentation and redundancy in systems and data flows. EA was the lever to advance a sweeping proposal to integrate data flows and consolidate application structures across the major EPA programs.

Now, having gained consensus on a design to resolve system fragmentation and redundancies, we have shifted the spotlight higher up the EA pyramid—to strategic goals, performance measures, business alignment, and capital planning. Four areas stand out: success in aligning the Chief Financial Officer's (CFO's) planning and budgeting processes with the EA at the agency and federal levels, alignment of PMA E-Gov initiatives with our internal project planning and positioning them to play cross-governmental roles, progress in integrating EA with the CPIC process, and improvements to the EA governance structure.

Our planning, budgeting, and EA processes now share a common business framework.

EPA's 2000 Strategic Plan summarized the Agency's mission in 10 major goals. In preparation for the FY 2005 Annual Plan and Budget, the Office of the Chief Financial Officer (OCFO) has been working on a more streamlined, five goal Strategic Plan, backed by solid and quantifiable performance objectives. Details are still evolving, but we expect to provide the FY 2005 Annual Plan and Budget to OMB in this new basic structure.

Meanwhile, the EA Team has been working since December to update EPA's business reference model (BRM) to parallel the evolving structure of the federal BRM. When the federal model separated what an agency does (its services to citizens) from how it does it (its "modes of delivery"), OCFO and the EA Team worked together to align our health and environmental protection services to our new Strategic Plan. They also integrated our administrative business model with our environmental and health protection business model and deepened that combined model to include the details of program office sub-functions. Finally, they completed the coordination of budget categories and the BRM so that the two now map to each other without overlaps or omissions. The result: EPA now has a single business model that maps not only to its new Strategic Plan, but also to its budget structure.

We are in the process of bringing our human capital model into the mix, so that all EPA planning processes can fit together cleanly into a single business modeling framework. Architecture, the budget, and human capital planning will speak the same language by the end of this fiscal year.

Figure SR-1 shows the high level relationship between EPA's BRM and the federal BRM. Detailed cross-walks, along with the full scope and depth of the new EPA BRM and its linkages to budget and human capital, are presented in Exhibit A: Part 1—Alignment of Strategic Architecture with Business and Budget Models.

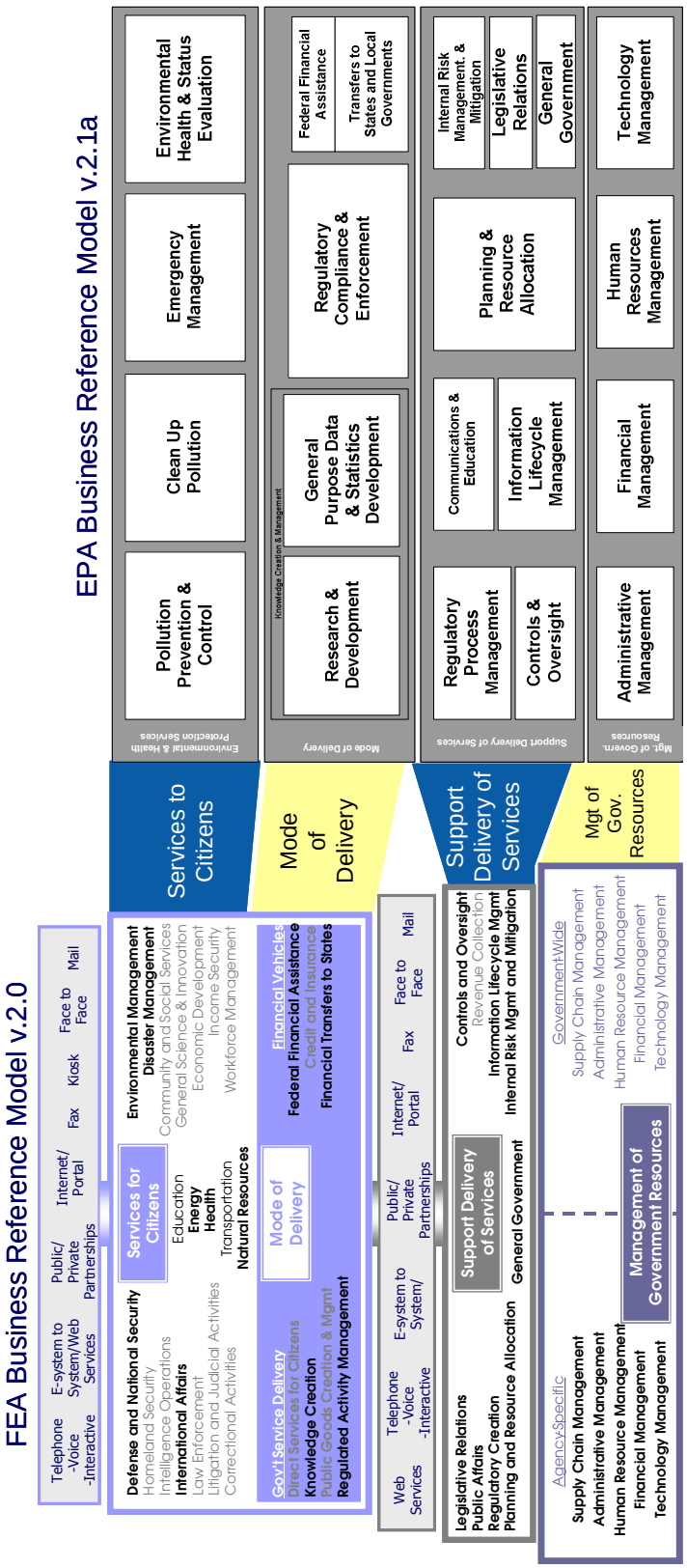


Figure SR-1: Relationship between the FEA BRM and the EPA BRM

This common business framework is the key to integrated performance measurement.

EPA's new five goal Strategic Plan provides specific health and environmental performance goals—defined as objectives and sub-objectives—that commit the Agency to quantifiable results. These performance goals make it possible to create a direct "line of sight" between OMB's high-level view of the Agency's functions, as defined in the FEA BRM and FEA Performance Reference Model (PRM), and EPA's programs and IT investments.

Figure SR-2 below shows an example. OMB lists Environmental Management, and below it, Environmental Remediation, as specific services to citizens. EPA links to Environmental Remediation with its own goal, Preserve and Restore the Land, for which it sets Cleanup and Restore of Contaminated Land as an objective and Make Land Available for Reuse as one of three sub-objectives with specific performance targets.[1]

EPA's Institutional Controls Tracking System (ICTS) keeps track of the status of federal and state sites, providing states and other users with information on safety restrictions that must attach to land parcels when they are made available for reuse (for example, "do not use for residential development"). Thus, EA can demonstrate a clear line of sight between EPA's ICTS investment and our goal of returning contaminated land to productive use.

	FEA Performance Reference Model		2003 EPA Draft Strategic Plan			CPIC
						ICTS
Measure-ment Area	Measurement Category	Generic Measurement Indicator Grouping	EPA Goal	EPA Objective	Measurement Indicator (EPA Sub-Objective)	
Services to Citizens	Environmental Management	Environmental Remediation	Preserve and Restore the Land	Cleanup and Reuse of Contaminated Land (3.3)	Control Risks at Contaminated Sites (3.3.1)	
					Make Land Available for Reuse (3.3.2)	X
					Maximize Potentially Responsible Party Participation at Superfund Sites (3.3.3)	

Figure SR-2: The Line of Sight between Federal Measurement Categories and EPA's Goals and Objectives: Institutional Controls Tracking System (ICTS)

As we complete the final details of our new Strategic Plan, performance measures such as those shown above will become more complete. Our priority is to develop similar performance measures for our administrative systems and research and science business domains.

Other Federal Enterprise Architecture (FEA) Reference Models are part of the complete framework.

Our greatest progress has been in aligning EPA's budget with federal and EPA business categories and performance goals, but we have also made progress in aligning our architecture to the other federal reference models. This year we have begun to align the federal Service Reference Model (SRM) categories to our business functions, systems, programs, and initiatives, but more remains to be done. With the release of the FEA Technical Reference Model (TRM) in June, we also began the process of aligning our own TRM with the federal model.

The FEA TRM differs fundamentally from the EPA model. The federal model includes product-specific content and EPA's does not. EPA does, however, maintain an "IT Roadmap" that lists individual products the Agency has approved for use. Part of the roadmap—the Standards

[1] e.g., "..land will be made available for reuse through cleanup, assessment, stabilization, or other action which indicates that such lands are restored to levels that are protective for the next reasonably anticipated future land use." *2003 Strategic Plan*, DRAFT, March 5, 2003.

Profile—lists our core product inventory. To align properly with the FEA TRM, we will align the Standards Profile with our TRM and then crosswalk the two of them to the FEA TRM. We completed the first draft of this alignment and crosswalk in mid-June 2003 and will continue to validate and extend it over the coming year.

EPA is a strong solution partner of the PMA E-Gov initiatives.

EPA is also an eager solution partner in the PMA cross-federal E-Gov initiatives and is working to ensure that the architecture fully supports them. For the E-Rulemaking Initiative, EPA is leading an interagency team to develop the federal E-Rulemaking docket, which will be hosted for the federal government by EPA at its National Computing Center. On others, we are aligning our business processes, architecture, and investments to leverage several of the developing E-Gov portals. Our Central Data Exchange (CDX) is aligning with the E-Authentication initiative. Our Integrated Grants Management System (IGMS) aligns with the Grants.gov portal and our Integrated Contract Management System (ICMS) with the One Acquisition portal and the Business Partner Network (BPN). PeoplePlus also aligns with the BPN as well as with EZ-Hire. Our geospatial architecture (now integrated with our health and environmental architecture) aligns with Geospatial One-Stop. In all, EPA is a player in 14 of the 24 E-Gov initiatives and has a leadership role in two.

In addition, one of our major E-Gov initiatives is the Exchange Network, a Government-to-Government consortium for sharing environmental information across levels of government—federal, state, and local. As the network continues to grow, all of EPA's major business systems will link to the Exchange Network. States will be able to exchange all their environmental business data through the network, and industry will be able to submit non-confidential data through the network as well. This sets the stage for an unprecedented federal-state-local collaboration in the environmental arena.

This year's CPIC process links closely with the architecture.

Integrating EA with the CPIC process is a major priority of the Clinger-Cohen Act and another area in which EPA has made substantial progress since December 2002.

We fast-tracked the mapping of the new FEA models with EPA's own models to serve the needs of this year's CPIC submissions—and those of the architecture. Within two weeks of the latest FEA releases, we provided CPIC proposal preparers with draft alignments between their own individual proposals and both the EPA and the federal reference models. We used the final, validated alignments they returned to us to update the architecture. EPA is a leader in the federal government in this area. The process was so successful that we are sharing our CPIC/EA alignment tools with other agencies.

CPIC also ensures consistency with E-Gov initiatives.

Following on the efforts undertaken in the FY 2004 CPIC process, we have closely examined the relationships between our major IT investments and the PMA's 24 E-Gov projects. The goal, again, is to ensure consistency and avoid duplication and redundancy.

Beginning with the FY 2004 process, we examined and modified several projects to ensure their consistency with on-going PMA E-Gov efforts in the Government-Citizen, Government-Business, Government-Government, and Internal Effectiveness investment portfolios. Using CPIC as a check on E-Gov alignments, the changes we made contributed to approval of 100% of our investment portfolio. Among these approved investments, for example, is the planned

implementation of data exchange between the EPA grants processing application (IGMS) and Grants.gov in FY 2004.

We recognize that several of the PMA E-Gov projects are still being formulated and that their business cases are evolving. We will continuously monitor our support for existing systems and commit to work with the E-Gov projects as they evolve. This may require modifications or redirection of some EPA projects. At present, however, all E-Gov initiatives in which EPA is involved are reflected in our architecture.

Our EA governance processes have been strengthened and simplified.

With the Agency moving more aggressively to align its strategic and business planning processes with its architecture, change management and governance of the EA have become more important. The Chief Financial Officer (CFO), the Chief Information Officer (CIO), and the Chief Technology Officer (CTO), forged a strong integration of EA, budget, and human resources planning. We also strengthened their vertical lines of authority: the CTO now approves all technical design decisions in the architecture. Management committees assigned to the various components and business domains of the architecture are playing a more active role in setting EA direction, especially as both regions and program offices are setting up their own internal EA programs. We set up a new management committee, the Research and Science Architecture (RSA) Executive Steering Committee, to put our science architecture under the same type of governance as the other business domains.

The structure of the EA governance process, slightly revised from the December 2002 submission, is shown on Figure SR-3 below.

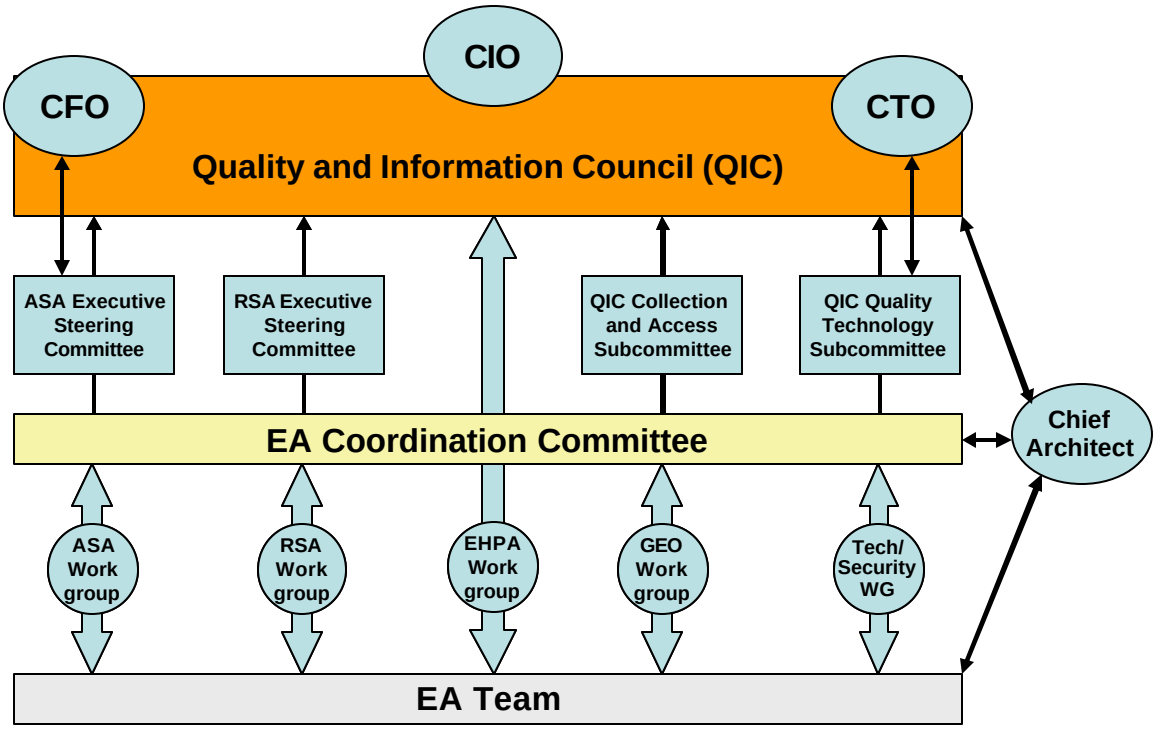


Figure SR-3: Governance Structure for EA Updates

Last year, the QIC created the ASA Executive Steering Committee (ESC) to oversee and coordinate the various elements of the ASA infrastructure. Reporting directly to the Deputy CFO,

the ESC manages and coordinates efforts across multiple administrative organizations to ensure the alignment and integrity of the several initiatives necessary to move rapidly toward the approved target architecture and the future integration of the ASA with the EA. The ESC is responsible for prioritizing and approving all initiatives, as well as ensuring proper sequencing of target implementations. These initiatives include: the Financial Replacement System (FINRS), the New Acquisition System, administrative E-Government initiatives, Enterprise Application Integration (EAI), the administrative portal, PeopleSoft HRMS, various critical upgrades to existing applications, and new business processes to ensure the success of planned initiatives. In FY 2003, the ESC approved a detailed sequencing plan that required coordinated planning among the various administrative capital planning proposals. This year, this plan is being updated to reflect changes in funding priorities, E-Government initiatives, and coordination with the EHPA.

Improvements to the target architecture make it actionable.

In December 2002 we presented an ambitious target architecture—a plan to deconstruct virtually all our major environmental systems and reconstruct them around a shared infrastructure of central services to promote data integration. In the target, all environmental information will enter the Agency through a common Portal to the Central Data Exchange (CDX). Data will be processed by the programs through a series of Operational Data Stores (ODSs) and then transferred to the Framework for Business Warehouses (FBW), a virtual warehouse of shared databases (previously referred to as the Enterprise Repository). All data in the FBW will link to a series of shared databases, called Data Registries, that will provide EPA's authoritative listings of chemicals, facilities, and other common information shared across program areas. The FBW will be the official source of information for users with appropriate access rights—the public, our partners, and our internal staff.

This integrated information management strategy originally applied only to our environmental systems. It did not address our administrative or research data flows because we felt that the data integration needs of environmental flows were the most pressing. We developed separate environmental, research, and administrative "business domains" to reflect this, but we committed to start integrating these domains as soon as possible. This integration is now moving forward.

Not only have we integrated the Agency's business model vertically with our strategic planning and budget operations, we have begun integration *across* the business domains at the application and data layers.

Integrating the applications layers starts with the environmental and administrative business domains.

As we integrated the business processes of our administrative domain (the Administrative Systems Architecture, or ASA) with those of our environmental domain (the Environmental and Health Protection Architecture, or EHPA), it became clear that their applications layers should follow. For example, our ASA systems provide grants to partners to carry out EHPA programs. To measure how our partners use these funds to support our mutual performance goals, our administrative and environmental systems need to communicate with each other. Because the ASA business domain is upgrading its Financial Data Warehouse to serve as the central repository for selected administrative data, integrating this warehouse with the FBW offers the opportunity to link budget data more effectively with business and performance information.

Figure SR-4 shows the evolving form of this integration. (The detailed version of this diagram is shown in Exhibit B: Part 3 –Updates to the Applications Architecture.)

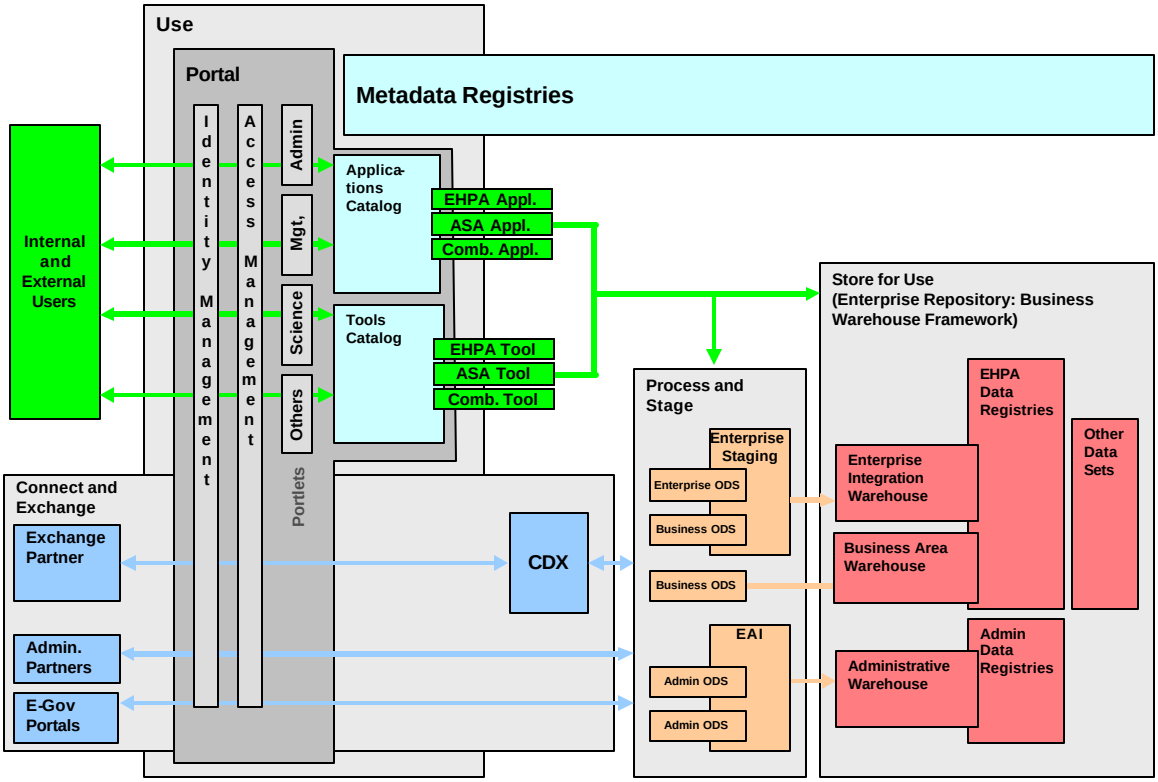


Figure SR-4: Schematic View of the Applications Architecture

In this diagram, the administrative Operational Data Stores (ODSs) are shown in the middle of the bottom border. Sample administrative applications are shown as connected to each other by the Enterprise Application Interface (EAI), described in version 1.0 of the target ASA. This allows transfer of data among administrative applications at the data processing stage. ASA information is then passed to the Administrative Warehouse.

To the upper left, reporting functions within the Portal can now include administrative reporting applications and tools.

Data Storage: Expansion of the Use of Web Services

Of the many issues that have come to light since December 2002, one of the more important is the possibility of creating a new Web services mode for linking business area warehouses. Version 1.0 of the target architecture provided three options for an application to store its data in the FBW.

- It can elect to store it in the **Enterprise Integration Warehouse** (EIW),[2] using the EPA-standard relational database, which will be built to handle the combined data storage needs of multiple mission-critical applications, much as Envirofacts does. The EIW will also contain the common Data Registries.

[2] The EIW represents the original Enterprise Warehouse component of the Enterprise Repository, as originally defined in the December 16, 2002 target architecture. It provides core services, such as the Data Registries, that are leveraged by the Business Area Warehouses, all of which together makes up the FBW.

- It can build its own **Business Area Warehouse**, containing data from one or more systems, such as those run by the Office of Air and Radiation or the Office of Water. Business Area Warehouses must use the same version of the EPA-standard database as the Enterprise Warehouse and must link directly to the common Data Registries.
- If its data is not mission focused, the application can store its data in a different type of database, not necessarily the EPA standard, which can be separate from the Enterprise Integration Warehouse or Business Area Warehouses and not linked directly to the common Data Registries.

For greater flexibility and to take advantage of emerging technology trends, we are extending data storage options to include a fourth mode: linking multiple warehouses with the Data Registries via Web services. This will relax the requirement that all mission-focused data be stored using the EPA-standard database and take advantage of the growing maturity of Web services, which make it more viable than it was when version 1.0 was designed. This fourth mode is basically an extension of the second—it is simply a more flexible way for a Business Area Warehouse to link to the Data Registries and metadata management systems.

No matter which option a mission-focused application chooses, however, most technical requirements will be the same. Data storage must conform to all metadata requirements, including unified data element definitions. The physical data models of individual warehouses, whether distributed or centralized, must map to a unified enterprise data model (discussed below). And all mission-focused data must connect directly to the enterprise Data Registries, which will contain the master copy of information on regulated facilities, substances of interest, organizations, individuals, and other categories of data essential to the regulatory process.

An enterprise data model will integrate the data architectures of the various domains.

Integration of IT systems is of little use unless the data handled by those systems, identified in the Target Data Architecture of December 2002, is also integrated. To do this the Agency needs an enterprise data model—a single model that integrates all repository information into a consistent, normalized schema. At the lowest level, metadata for individual data elements must be reconciled (i.e., we must have common formats for individual fields like telephone numbers, latitude/longitude coordinates, and chemical names). At the highest level, the model's structure and content must mirror the Agency's strategic plan and include all information necessary to track our performance goals.

Our strategy for building the enterprise data model within the context of an overarching data management process is presented in Exhibit B: Part 2: Evolution of the EPA Data Model. It combines bottom-up and top-down approaches as follows:

- We are working closely with the Environmental Council of States' (ECOS) and supporting the state/EPA Information Management Work Group to help develop their **Core Reference Model** (CRM). The CRM is an empirically derived common business framework for sharing environmental information across the Exchange Network. It consolidates data elements that the states use to run their own environmental programs—an invaluable aid for metadata reconciliation at the federal level. We plan to leverage our experience with interagency data architecture development to help the Federal Enterprise Architecture Program Management Office (FEA PMO) develop and administer the FEA Data Reference Model (DRM).

- We are well along in developing a **Strategic Information Model** (SIM), based on the Agency's Strategic Plan, to serve as the general roadmap of the enterprise data model. Using the SIM as a high-level reference assures that the model will fully incorporate (or at minimum reference) all data needed to track our performance measures and deliver on our mission commitments. The SIM will also reference the data necessary to compile EPA's Environmental Indicators, which have already been incorporated, in part, in the new strategic goal structure. The first version of the EPA SIM will be completed in FY 2004.
- We are using **Envirofacts**, EPA's award-winning public access database, as the practical starting point to build the enterprise data model because Envirofacts already contains—and to a degree integrates—data from our 12 most mission-critical data systems.
- Our **Environmental Indicators** project has supported development of performance measures in the new Strategic Plan. The enterprise data model will incorporate necessary data references from the indicators project to ensure that we have access to whatever data is necessary to track our Agency performance measures.
- We are launching an ASA **Data Analysis Initiative** to standardize data management within the administrative domain and provide improved data quality and increased reporting capabilities.

The ASA Data Analysis Initiative is a comprehensive program to reevaluate ASA data from the ground up, including *revising the basic business processes that generate this data* to make them consistent with the needs of relevant E-Gov initiatives, data warehousing, the administrative portal, and Enterprise Application Integration (EAI). This effort will define management processes necessary for identifying, analyzing, prioritizing, reviewing, approving, maintaining, and disseminating administrative data on an ongoing basis. It begins this year with development of models to describe ASA business processes under the target architecture. Once these business processes are defined, we will identify the data necessary to support them, along with the appropriate applications, tools, and procedures necessary for that data's storage and maintenance. Among the expected outputs from this initiative is a data model that will become part of the enterprise data model.

Development of the enterprise data model will begin in the fall and will focus on meeting the near-term needs of the FBW—making sure that the model will accommodate all data elements required by the first system integrated into target architecture in 2005.

We are addressing metadata reconciliation on a parallel track. The Environmental Data Registry (EDR) will become the core of the proposed EHPA Data Element Registry. In FY 2004 we will begin a major multi-year effort to reconcile metadata on the Agency's estimated 15,000 data elements. Envirofacts accounts for about 4,500 elements, of which about 400 have been reconciled. This is a start. We will focus first on the data elements for projects currently under way, as well as for those needed by the pilot system.

We have begun actual construction of the central services.

Our most significant initiative this year is to start actual construction of the central services necessary to support the target architecture. This is the largest IT program in our history. Our strategy is to bring all the components of these central services up concurrently so that one major system can be operational on a pilot basis in 2005.

In January of this year, the Office of Environmental Information (OEI) created the Program Management Office (PMO) as a matrix organization to coordinate the many parts of the target

EHPA central services—now called the Environmental Information Integration and Portal Development (EIIPD) project. Reporting directly to the Principal Deputy Assistant Administrator for OEI, the PMO manages and coordinates up and across the organization by creating a series of construction-oriented "portfolios." These include CDX, the FBW, the System of Registries, the Portal, Analytical Tools, Identity Management, Geospatial Services, and Operational Data Stores.

EPA has transferred one of the Agency's most experienced IT managers from another program office to serve as PMO Director. Other top IT staff have moved to OEI from the program offices to provide necessary skills, knowledge and experience and to enhance working relations between OEI and the programs. Creating the PMO effectively flattens the bureaucratic structure needed to achieve the target, making the operation more efficient and responsive. The PMO is also putting management systems into place to reduce overlapping functions and avoid turf conflicts.

The PMO's first step is an integrated project workplan.

The PMO's first priority has been creation of an integrated EIIPD project workplan that defines in detail all the linkages and dependencies of the target applications architecture. This workplan was developed by first tracing detailed pathways for inputting data through the Portal and CDX, through the ODS systems, and into the FBW. Next, the PMO traced the pathways of queries addressed to the FBW and related data marts via the Portal, with its identity management and access control functions, to the suites of applications and analytic tools that address the business warehouses of the FBW. Finally, it defined all the links and dependencies between these two pathways and the metadata management systems required to ensure data quality and consistency. The resulting flowchart became the map of all necessary system components and projects to be built or coordinated by the PMO.

As part of this integrated project workplan, the PMO is launching a proof of concept project to verify the technical feasibility and implementation costs of using Web services as a means of creating the business warehouse framework, which will permit mission-critical applications to link directly to the Data Registries using multiple database management systems.

The technical steps and project costs of the various options for migrating applications to the FBW will, of course, be different. Next year's CPIC submissions will deal in detail with the options available to each application for transition to the target environment. The PMO and the EA Team will work closely with program offices and provide necessary guidance to ensure that all technical and architectural requirements are met. Details of the development of the basic EIIPD components are presented in Exhibit C: Part 2: Sequencing Plan.

We will base sequencing of applications to the EIIPD on mission-focus, CPIC designation, and system lifecycle phase.

Plans now call for one major program application, yet to be selected, to migrate to the target architecture as a pilot in 2005. Once this system has moved to the target, all the other principal EHPA applications must follow. During transition, applications must continue to operate in their legacy configuration—with inevitable patches and upgrades to keep up with program needs until the new system is ready. To minimize these parallel costs, transition must be as rapid as possible once a system enters the EIIPD pipeline.

Once an application moves into the pipeline, some of the costs of its modernization will fall to the EIIPD and others to its program office. The balance of these costs will depend in large part on how it chooses to link to the FBW. For instance, if it chooses to join it via Web services, the costs of setting up the necessary links will fall to the program office—costs that would not be incurred

at all under other options. Similarly, costs of building a Business Area Warehouse will fall entirely to the program office.

During the proof of concept and pilot phases, OEI will bear the costs of building the central services and integrating the pilot application. Thereafter, we will develop a new funding strategy to appropriately balance costs between OEI and the program offices based on the individual requirements of systems and choices of program offices.

To minimize transition costs, our strategy will be to migrate several "waves" of application systems into the architecture over a period of three to four years. Considerations for setting priorities among the many systems that must migrate to the target include:

Business categories

- **Alignment with strategic priorities:** For strategic reasons, the Agency may wish to give priority to certain classes of systems, such as E-Gov solutions, those that support innovative or voluntary controls, or those related to homeland security. It may wish to emphasize key Agency initiatives, certain goals within the Strategic Plan, or specific priorities within the President's Management Agenda.
- **Major systems first (National Program Systems - CPIC):** Systems considered "major" under the CPIC should receive higher priority than "non-major" or "other" systems. It is logical to focus on the core functions of the enterprise first, and, since data integration is the goal, on the systems that house the core stock of the Agency's data assets, that may support E-Gov initiatives, and whose data flows are directly mandated by Congress. Giving priority to major systems will also tend to ensure participation by all program offices simultaneously.
- **Degree of external coordination and customer impact (E-Gov):** Certain applications hold data of value to external partners, such as homeland security and emergency response, or systems relied upon by the private sector. E-Gov initiatives in general will deserve priority consideration.
- **Resource restrictions:** Low cost projects with high impact and proven benefits may receive higher priority than higher cost projects that offer fewer business improvements.

Technical categories

- **New systems**: No new system should be built "outside the box," so all new systems must be integrated with the EIIPD from the beginning.
- **Current Systems Lifecycle (SLC) state**: We should consider how long it has been since a system's last complete modernization, and how urgent is the need for functional upgrades.
- **Technical complexity:** The technical complexity of particular systems may argue to start transition early, or conversely to wait until critical technologies mature. But either way, technical complexity issues will undoubtedly figure in sequencing decisions.

The Agency is still analyzing details of these migration questions, but the essential result is not difficult to see. New applications and mission-focused systems must get highest priority; all others follow.

The best short list of mission-focused systems includes those that output data to Envirofacts. The core set includes Permit Compliance System, National Compliance Data Base, OECA Docket, NEI, RCRAInfo, AIRS Air Quality Subsystem, AIRS Air Facility Subsystem, the System for Risk Management Plans (SRMP), CERCLIS, SDWIS, TRIS, and STORET. Added to this list might be a few new applications such as OPPIN.

Table SR-1 shows a possible breakdown of three waves of systems migrating to the target.

Table SR-1: Proposed Waves of Applications Migrating to the Target Architecture

	System Group	Comment
Wave 1	All Envirofacts applications	All are mission-critical.
	All new applications	No new application should be built "outside the box."
	All EIIPD components, such as the Data Registries	By definition.
Wave 2	All major applications not included in Wave 1 or Wave 3	The remainder of Tier 1.
	High priority non-major and other applications	Major regional applications and others with significant use across programs and outside the Agency.
Wave 3	All other non-major and other and smaller applications destined for the FBW	To be defined.
N/A	Confidential business information (CBI)	At this time, all CBI is handled in isolated systems for security.
	Legacy systems	Applications that are being retired and/or replaced.

We will conduct an inventory and review of non-major applications to identify those that should also migrate the EIIPD. A number of non-major regional applications may also be priorities for inclusion in Wave 2.

We have started integration of EA with CPIC cost-benefit analysis.

In their review of the target EHPA in December 2002, EPA's Quality Information Council (QIC) asked three questions:

- What will it cost?
- When will it be ready?
- What are its benefits?

OEI and OCFO are still developing complete answers to these questions. OEI has calculated the implementation costs of the major OEI-based central service elements as part of this year's CPIC process, but total costs to the program offices are not yet fully described. OCFO has begun developing cost estimates for the target ASA as part of its current administrative Exhibit 300s. Overall, the Agency must know exactly how its applications will use the central services and what the costs may be of the possible Web services-based warehousing option. In addition, we have yet to fully cost out a number of other important elements in the PMO project workplan, including creating the enterprise logical data model, reconciling program data element definitions

within the enterprise data model, passing program data from CDX to the ODSs via Web services, and creating business area warehouse logical data models.

Once these are known, the costs of all the major pieces will be available. But different options for joining the target involve different mixes of cost elements. Some costs that remain with the program offices with one option move to the enterprise with another option. To track these differences, the EA Team and the PMO are creating an integrated cost-benefit methodology.

This methodology will be available for program office review and training during next year's CPIC cycle. With this as the rule set, we will be able to answer the question of what the target will cost under different scenarios. Many details of this integrated cost-benefit approach remain to be defined, but we believe that this methodology will provide clear benefits in the next CPIC cycle. They will help us fully understand the benefits and costs of our target architecture, further integrate the EA program with the CPIC investment review process, and integrate CPIC proposals (at least those that relate to the EIIPD) with each other.

We have deployed the architecture throughout the Agency using ART.

EPA acquired the Metis architecture modeling tool early in the development of the EA program. By December of last year, we had customized it to model EPA's EA, thereby creating EPA's customized implementation, named ART (Architecture Repository and Tool).

The program offices now recognize ART as a powerful tool for updating the architecture, managing their own business activities, and integrating program-level architecture efforts with the Agency architecture.

- ART contains the full information from the 2002 baseline and target plus all the information developed since then on the Agency target. We have expanded it to include Agency stakeholders, FEA Reference Models, and PMA E-Gov initiatives.
- We have published ART for broad use across the Agency. Staff representing EPA program offices now have on-line access and are trained in how to query the architecture and export information for their own program purposes.
- We have trained selected staff from most program offices in the use of ART. They are now able to model their own processes.
- Most EPA program offices have initiated their own architectures and created detailed program-specific models. They will export information from their own models to the enterprise model.
- We have put a governance process in place to manage updates to the enterprise model. This process is responsible for change management of the enterprise model.

The ART model is contained on the accompanying CD-ROM (launch EA Version 2.0).

Looking forward.

The last nine months have been extraordinarily productive for the EPA EA program. The next year promises to be even more so. Over this time we will:

- **Further tie our CPIC program to the EA** by integrating the cost-benefit analysis of our EIIPD central services with those of the applications that migrate to it. This will let us measure the benefits and costs of the architecture as we develop it.

- **Extend the architecture** to address regional business processes, data, and applications.
- **Integrate EPA operations with evolving E-Gov initiatives.**
- **Build our enterprise data model** and begin to integrate the data definitions of the elements within it. This is a necessary step for building the FBW, but it will also help us measure performance of our programs, link more efficiently to our partners over the state network, and serve the needs of situational analysis.
- **Make the EIIPD operational with a fully integrated pilot application** and complete a test of new Web services options for linking Business Area Warehouses to the central services.
- **Complete an integrated sequencing plan** for migrating the EPA business process and supporting applications to the target architecture.
- **Continue to integrate our architecture domains,** including the RSA target and sequencing plan. The goal of a comprehensive architecture whose services span all programs and functions—environmental, administrative, and scientific—is much closer.
- **Enhance the ART as a cross-Agency management tool,** including new analysis capability to support IT and business managers' strategic decisions.
- **Collaborate with other federal agencies on environmental and human health issues,** furthering the goals of the PMA.
- **Collaborate with states and tribes,** focusing especially on improving the Exchange Network and ensuring compatibility of data definitions across environmental programs at all levels of government.

EPA will continue to be a leader in the EA arena, using its architecture program to improve its environmental and health services and coordinate with its partners and stakeholders.

This Page Intentionally Blank

Exhibit A

EPA's Enterprise Architecture Version 2.0

This Page Intentionally Blank

EPA's Enterprise Architecture Version 2.0

EPA's Enterprise Architecture (EA) has now been widely distributed throughout the Agency via the EA tool—the Architecture Repository and Tool (ART). Presented with this submission is EPA EA version 2.0 in ART. This exhibit provides a brief introduction to version 2.0 and to the methods available in ART to view and query it.

Definition of ART

ART is EPA's authoritative source of information on the EA and a powerful analysis tool for navigating and exploring it. ART's graphical images make EA information understandable to both information technology and business managers. ART helps users visualize the EA, see its structure, and understand how its various models relate to the Agency's goals and business.

ART is based on a commercial off-the-shelf product called Metis, produced by Computas. OEI is managing ART in a collaborative process with program offices and regions.

Deploying ART within EPA

In 2003, the EA team successfully released ART for Agency-wide use. They have:

- Released ART through a web-browser to the Agency.
- Provided access to ART to permit EPA users to edit their offices' activities.
- Established the ART Change Control Board.

Contents and Presentation of the Enterprise Architecture

Since EPA's last EA submission to OMB, the contents of the EA—and the Agency's ability to use it—have expanded considerably. The Agency has added the following content:

- New goals and objectives from the Agency's recently released Strategic Plan.
- Federal PRM, BRM, SRM, and TRM.
- A new EPA Business Reference Model.
- An integrated target applications model.
- 2003 CPIC investment proposals with mappings to EPA and federal reference models.
- EPA stakeholders.

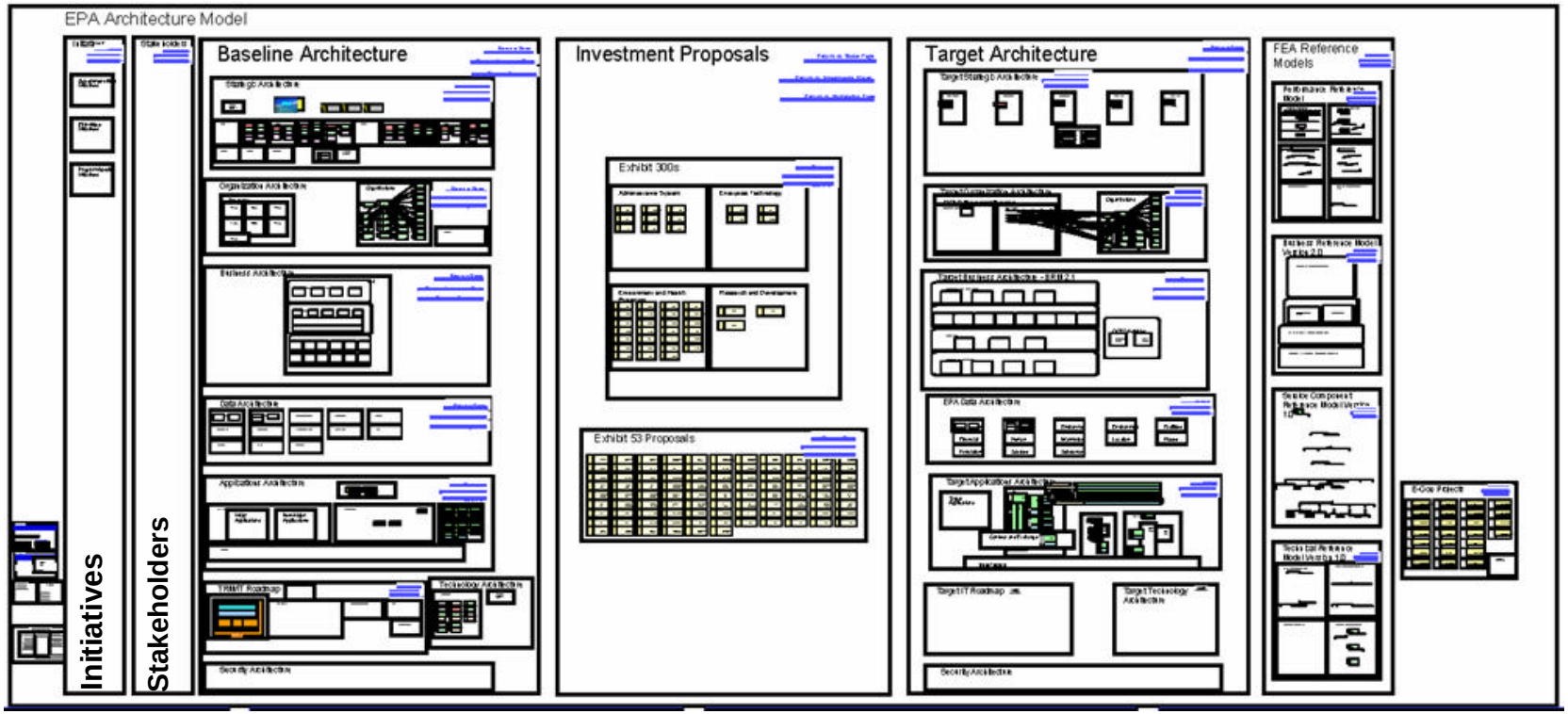


Figure A-1. ART Modeling Areas

The ART user interface has been significantly improved, with a new structure containing EPA baseline and target architectures, and federal reference models (see Figure A-1). The major elements being provided through the ART electronic submission are shown on Table A-1 below.

Table A-1: Contents of ART

1. Individual Models: Entire Model Baseline Architecture Target Architecture Federal Reference Models	**3. Major Relationships** IT Investments mapped to Federal Reference Models and EPA architecture **Federal** — **EPA** PRM — Target goal structure BRM 2.0 — BRM 2.1 SRM 1.0 — Applications TRM 1.0 — TRM EPA BRM mapped to FEA BRM
2. Architectural Layers Strategic Architecture Organizational Layer Business Architecture Data Architecture Applications Architecture Investment Proposals (CPIC) Technology Architecture	**4. Querying Capability**

Figure A-2 shows the ART home page, with links to different views tailored for different audiences.

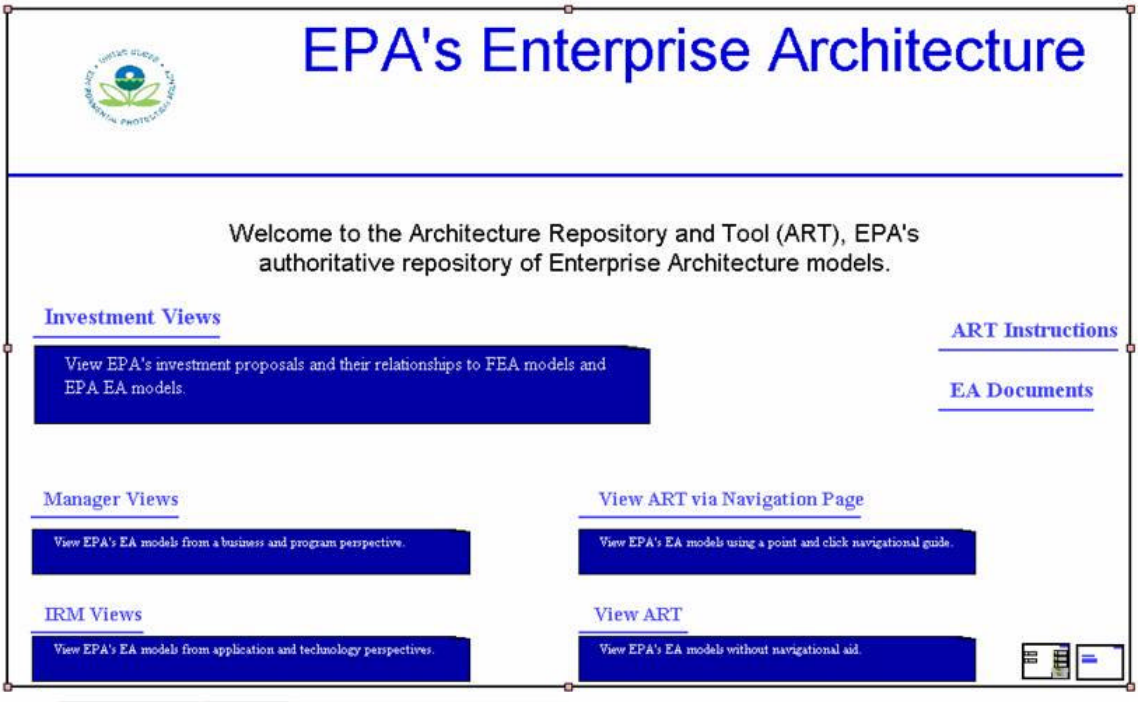


Figure A-2: View of the ART Home Page

Figure A-3 shows a sample query page, where EPA managers see IT investments as they relate to the FEA and EPA EA.

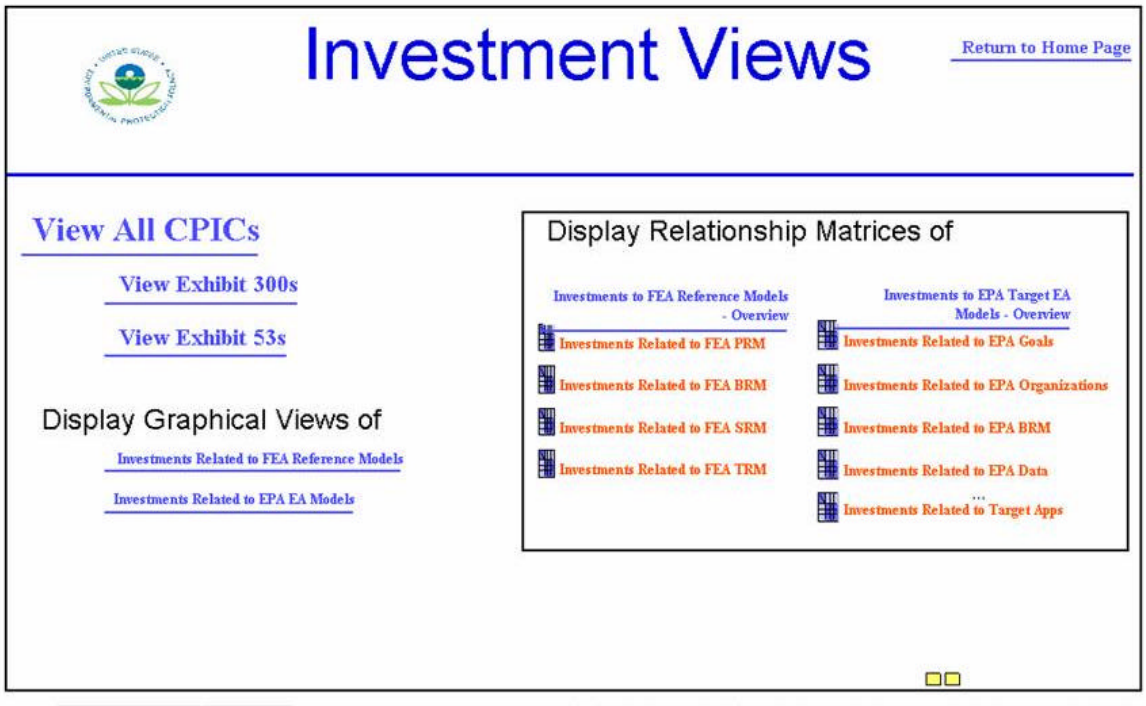


Figure A-3: Investments Views as Displayed in ART

Using ART to View the Enterprise Architecture

To view the EPA's EA, insert the CD-ROM and follow the accompanying self-running instructions.

Exhibit B

Revisions and Updates to the Target Architecture

This Page Intentionally Blank

Part 1
Strategic and Business Architectures: Alignment of Strategic Architecture with Business and Budget Models

EPA faces the same challenges as other agencies in aligning its IT investments, human capital, and financial resources more closely with its mission and goals. Like other agencies, some 40 percent of its workers will be eligible to retire in the next four years, many of them "founding members" who carry with them the memory of EPA's entire institutional history. The need to document Agency processes, and to link these processes clearly with program outputs and environmental outcomes, has never been more acute.

Enterprise architecture allows EPA to relate means and ends in an ever more systematic and scientifically robust way. Since submission of its target architecture to OMB in December 2002, the Agency has aligned its strategic goals, performance measures, budget structure, and human capital planning systems by:

- Revising its Strategic Plan to simplify the Agency's 10 goal structure into a new five goal structure and link them to quantifiable environmental, health, and economic objectives based on the latest scientific analysis.
- Aligning these strategic objectives with the Federal Enterprise Architecture (FEA) Performance Reference Model (PRM) (for CPIC investments only).
- Revising its Business Reference Model (BRM) to parallel the Federal Enterprise Architecture (FEA) BRM structure.
- Aligning the EPA BRM to the Agency's budget structure.
- Aligning the EPA BRM to the Agency's human capital planning structure.
- Providing guidance to the Capital Planning and Investment Control (CPIC) process to align investment proposals to all of the above models.

Alignment of the Strategic Plan to the Federal PRM

The first version of the federal PRM was released on June 13. On June 20, EPA provided guidance to its CPIC reviewers that linked applicable federal PRM measurement categories to the performance objectives of the Agency's draft 2003 Strategic Plan. These performance measures were in turn mapped (provisionally) to each of the 48 individual CPIC proposals under development. The result: a clear line of sight tying high level services to citizens to quantifiable performance measures, and then to the systems holding the data to measure progress.

Figure B-1 below shows an example. OMB lists Environmental Management, and below it, Environmental Remediation, as specific services to citizens. EPA links to Environmental Remediation with its own goal, Preserve and Restore the Land, for which it sets Cleanup and

Restore Contaminated Land as an objective and Make Land Available for Reuse as one of three sub-objectives with specific performance targets.[3]

EPA's Institutional Controls Tracking System keeps track of the status of federal and state sites, providing states and other users with information on safety restrictions that must attach to land parcels when they are made available for reuse (for example, "do not use for residential development"). Thus, EA can demonstrate a clear line of sight between EPA's ICTS investment and our goal of returning contaminated land to productive use.

	FEA Performance Reference Model		2003 EPA Draft Strategic Plan			CPIC
						ICTS
Measure-ment Area	Measurement Category	Generic Measurement Indicator Grouping	EPA Goal	EPA Objective	Measurement Indicator (EPA Sub-Objective)	
Services to Citizens	Environmental Management	Environmental Remediation	Preserve and Restore the Land	Cleanup and Reuse of Contaminated Land (3.3)	Control Risks at Contaminated Sites (3.3.1)	
					Make Land Available for Reuse (3.3.2)	X
					Maximize Potentially Responsible Party Participation at Superfund Sites (3.3.3)	

Figure B-1: Mapping of Federal PRM to EPA Strategic Goals and Individual CPIC Submission

Still to complete are similarly quantifiable goals that create a line of sight tying the federal PRM "Support Delivery of Services" and "Management of Government Resources" measurement categories with administrative CPIC proposals in such areas as Human Resources and Financial Management. The present EPA goal structure for these categories is under development by the Office of the Chief Financial Officer (OCFO).

Alignment of Federal BRM to EPA BRM, Budget Activities, and Human Resources Planning Categories

Over the first and second quarters of 2003, EPA revised its target BRM version 1.0, presented to OMB in December 2002, to serve as its baseline and target BRM version 2.1.

Although there are many opportunities to streamline internal processes and services, EPA's essential lines of business have not changed significantly in recent years and are expected to be roughly the same five years hence. EPA therefore sees no current difference between its baseline and target business reference models. If Congress were to add new responsibilities to the Agency's charter, or if program offices in the future find ways to carry out their goals using new and different business functions, or if the FEA BRM changes, then the target and baseline BRMs may diverge. Until then, EPA BRM 2.1 will support both the "as is" and the "to be" states of its architecture.

The critical changes between EPA BRM versions 1.0 and 2.1 are:

- The Target Administrative Systems Architecture business functions (see Appendix D of the EPA December 16, 2002 submission) have been integrated with the full EPA BRM, substituting for the administrative-level definitions of the previous version.

[3] e.g., "..land will be made available for reuse through cleanup, assessment, stabilization, or other action which indicates that such lands are restored to levels that are protective for the next reasonably anticipated future land use." *2003 Strategic Plan*, DRAFT, March 5, 2003.

- Mode of Delivery has been added as a level of the EPA BRM, paralleling the federal model.
- Mode of Delivery now encompasses both Research and Development and Regulatory Activity Management, which were in different layers of the previous BRM.
- New high-level Environmental and Health Protection Services have been added to version 2.0, paralleling the draft 2003 Strategic Plan.
- These major changes are summarized in Figure B-2 on the next page.

In addition to changing the overall structure of the EPA BRM to match the federal, the Agency has expanded the BRM's coverage to subdivide functions to as many as six layers deep. This allows EPA program offices to make necessary business distinctions at the detail level without losing the overall mapping of the EPA BRM to the federal BRM at higher levels. The overall depth of the EPA model is shown on Figure B-3.

On this figure, the yellow areas denote activities integrated from the ASA December 2002 business model. The orange areas indicate terms and functions adopted directly from the federal BRM version 2.0 to ensure consistency. The horizontally-formatted items are time-sensitive processes that support time-insensitive business functions: e.g., the competition, award and audit steps of contract management are time-sensitive processes underlying the time-insensitive business function of "manage contracts."

Definitions for all these functions and processes have been carefully developed and vetted with EPA programs through the Enterprise Architecture Coordination Committee. (These are available separately on the accompanying CD-ROM: see BRM Definitions.)

Lastly, EPA has cross-walked the BRM against both the federal BRM version 2.0 and the new budget categories developed by the Office of the Chief Financial Officer (OCFO). The first mapping is shown conceptually in Figure B-4. The second is shown in detail on Figure B-5a-d. A similar mapping (not shown) has been made with the human resources planning categories.

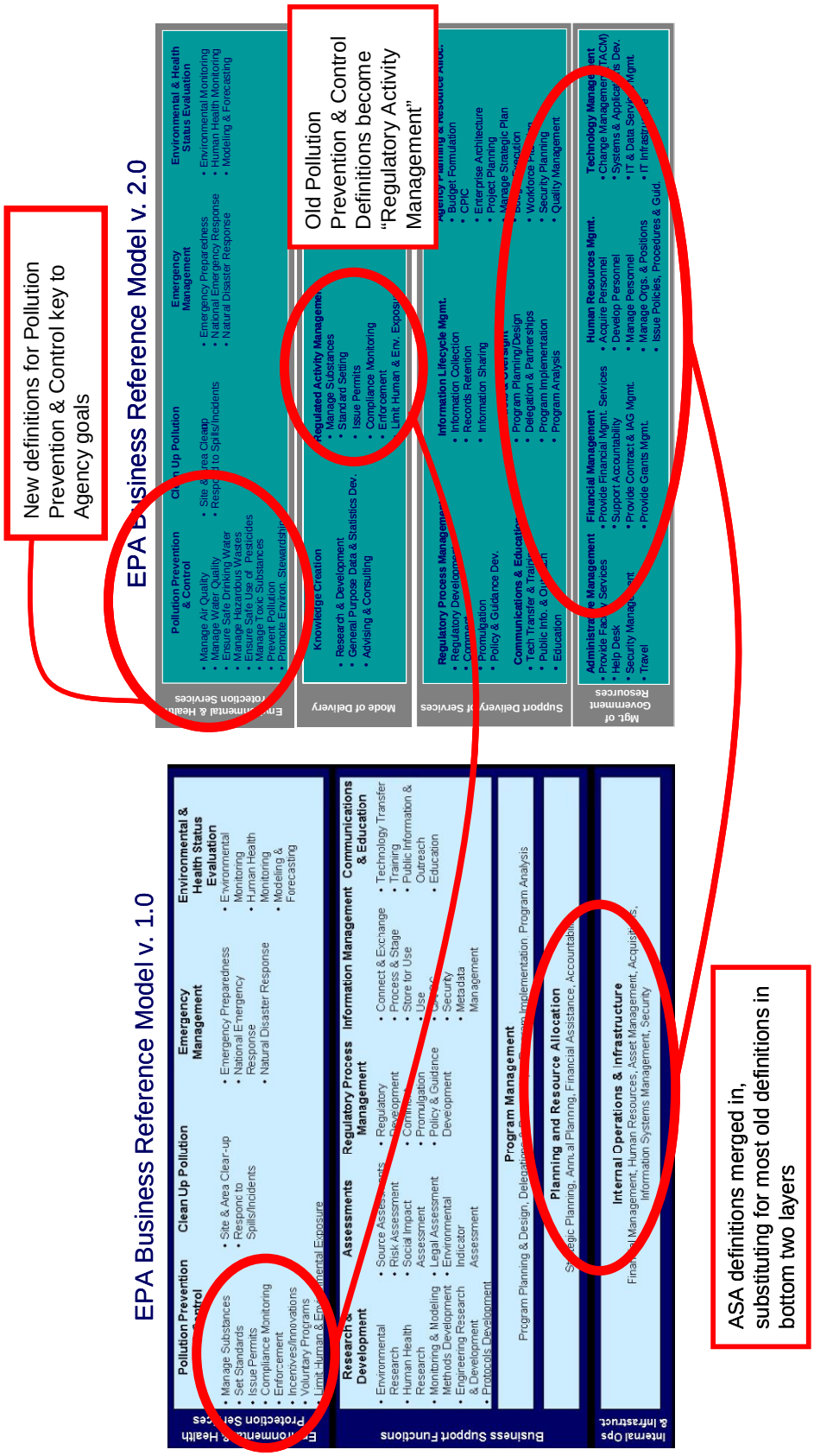


Figure B-2: Principal Changes between EPA BRM Versions 1.0 and 2.1a

Figure B-3: The Detailed Breakout of EPA BRM Version 2.1a

This Page Intentionally Blank

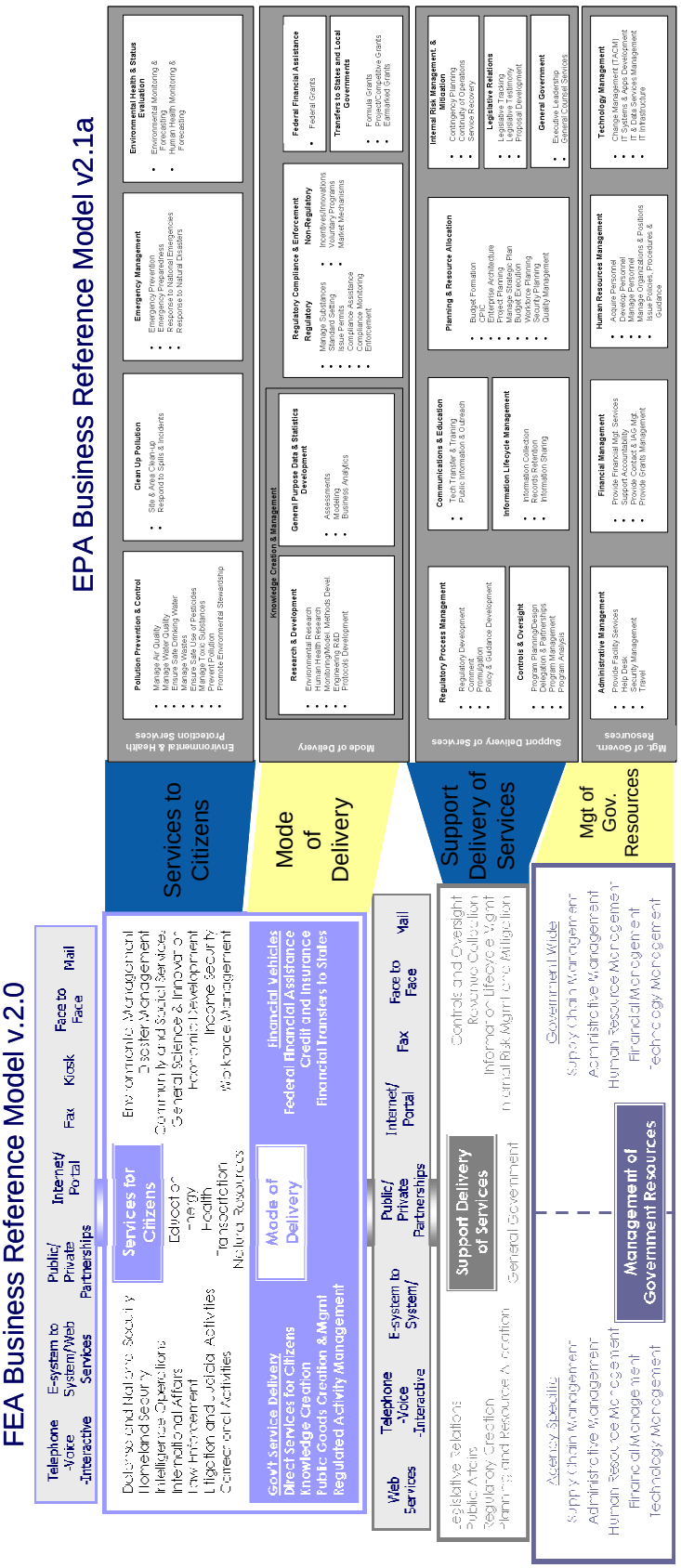


Figure B-4: Conceptual Mapping of FEA to EPA Business Reference Models

Summary

EPA's efforts to revise its business and performance measures to align as clearly as possible with the federal models, as well as with its budget activities and human capital planning process, provide an extremely powerful tool for resource management and planning. For the first time, the Agency is able to draw clear relationships between its general goals, the detailed objectives it sets for meeting these goals, and the human and budget resources available to achieve them.

Next steps in this process include:

- The refinement and expansion of the business model as needed to reflect changes in the target business architecture.
- Expansion of Agency performance measures, particularly in the realm of administrative services (Support Delivery of Services and Management of Government Resources).
- Continued application of these models to the review of capital investments through the CPIC process.

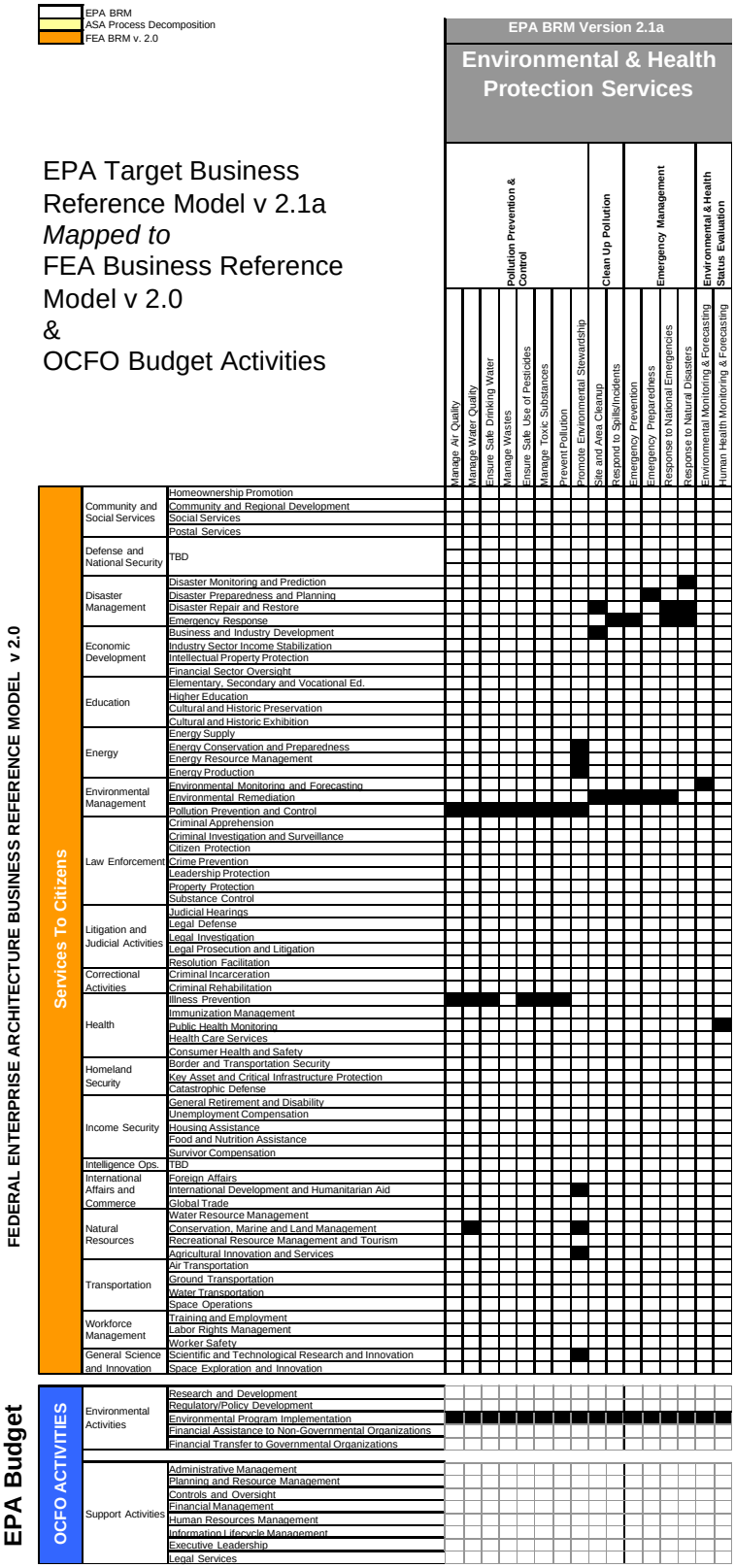


Figure B-5a: Cross-walk of EPA BRM Version 2.1a (x-Axis) to FEA BRM Version 2.0 and OCFO Budget Activities – Services to Citizens

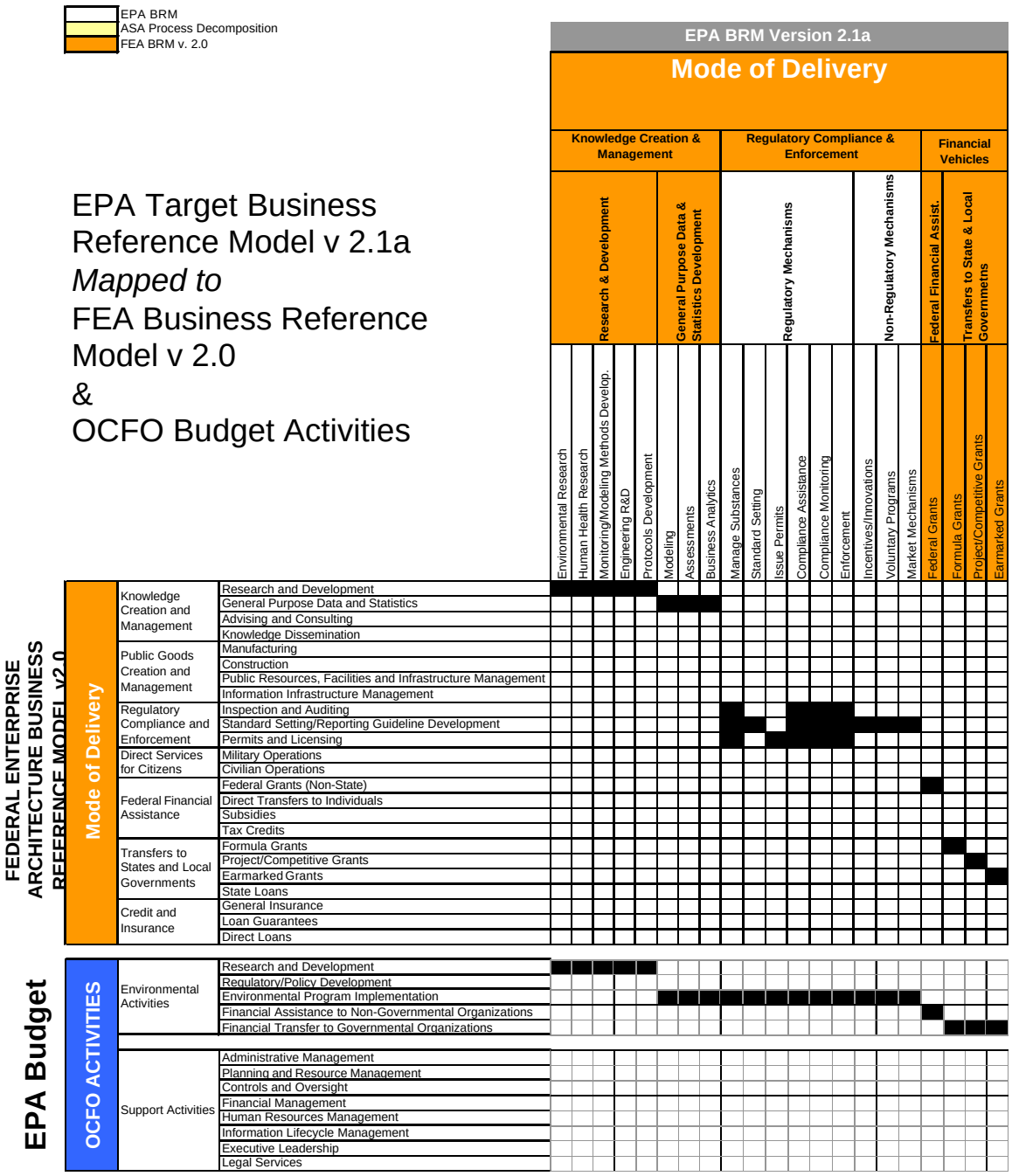


Figure B-5b: Cross-walk of EPA BRM Version 2.1a (x-Axis) to FEA BRM Version 2.0 and OCFO Budget Activities – Mode of Delivery

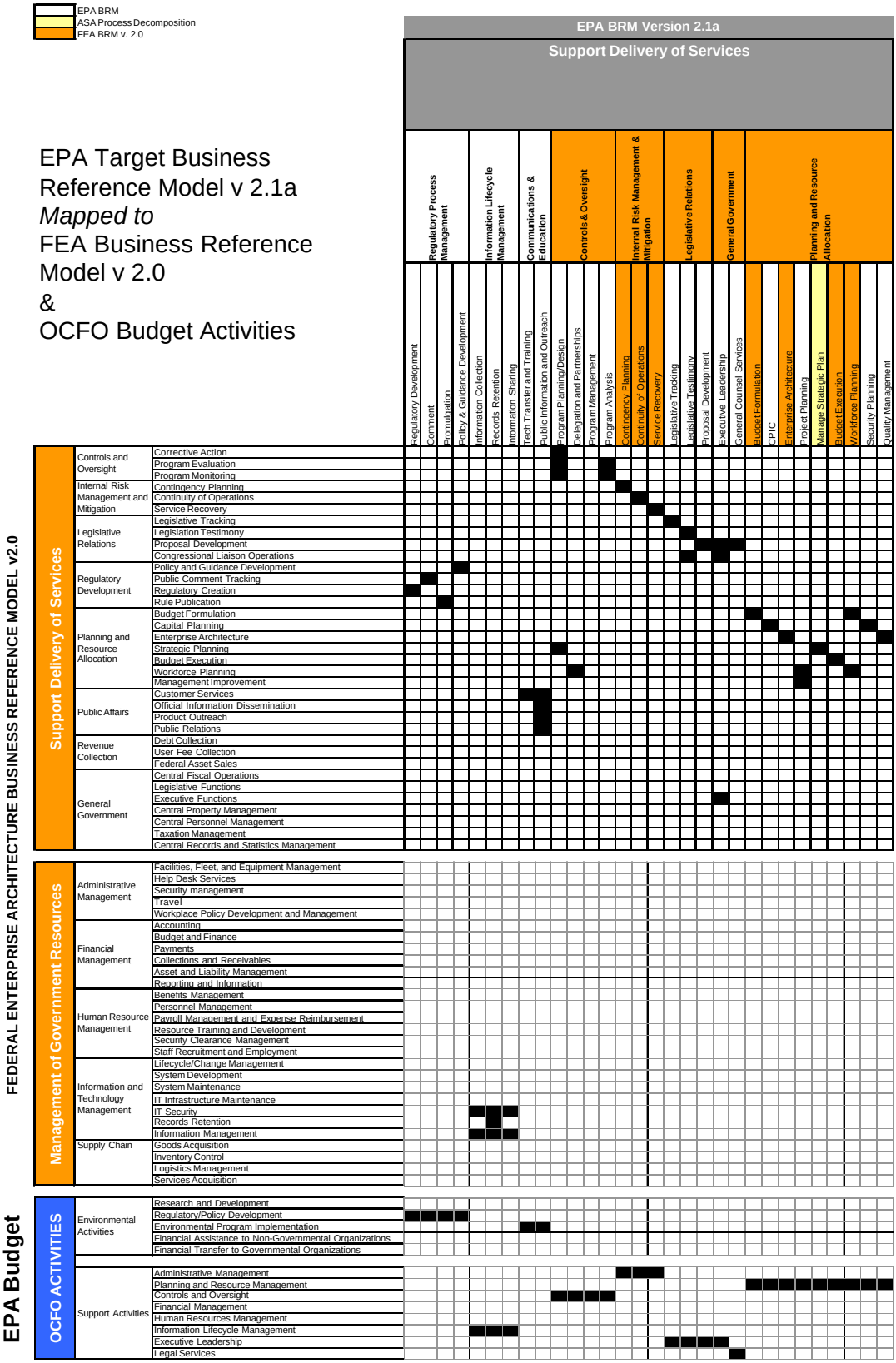


Figure B-5c: Cross-walk of EPA BRM Version 2.1a (x-Axis) to FEA BRM Version 2.0 and OCFO Budget Activities – Support Delivery of Services

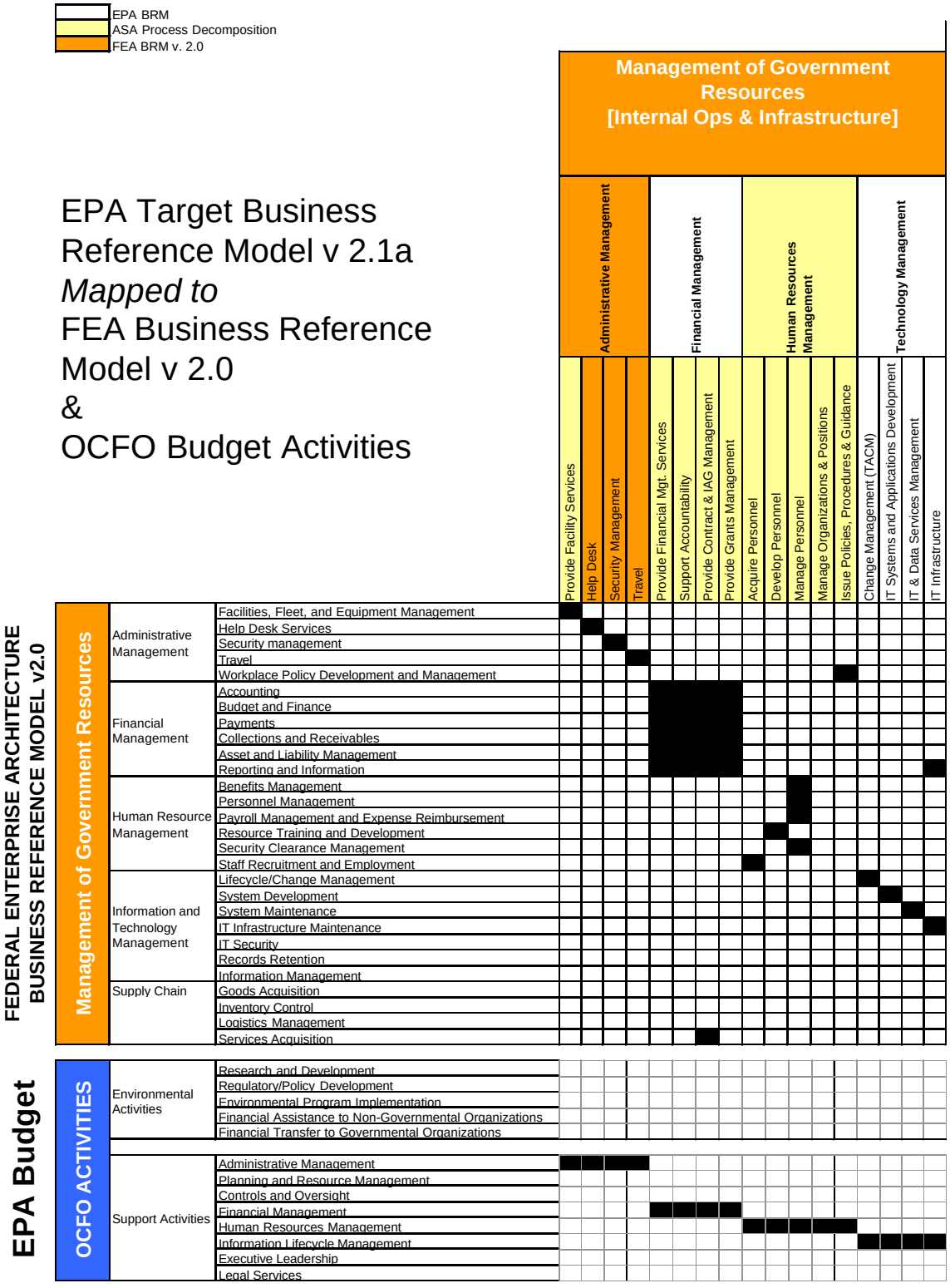


Figure B-5d: Cross-walk of EPA BRM Version 2.1a (x-Axis) to FEA BRM Version 2.0 and OCFO Budget Activities – Management of Government Resources

Part 2
Data Architecture: Evolution of the EPA Data Architecture

The chief goal of EPA's target architecture is to integrate the Agency's data resources. The central element of the target architecture is the Framework for Business Warehouses (FBW) to which all internal and external queries are directed.

Enterprise data storage requires an enterprise data model—a major undertaking. The EHPA Data Architecture and the ASA Target Information Classes, presented in December 2002, created a comprehensive inventory of classes of data used by EPA systems. These now must be integrated and structured into this general model around which the FBW will be built. Even though several options are available for the technical design of the FBW, including using a distributed series of databases linked by Web services, *all options* require the development and deployment of an integrated enterprise data model.

EPA has a candidate model at hand in Envirofacts, the repository of Agency data used for public inquiries and internal analysis. Containing some 4,500 elements, Envirofacts integrates a substantial portion of the program data developed by EPA's 12 most mission-focused applications.[4]

The scope of the FBW will, however, be substantially broader than Envirofacts'. It will hold secure and sensitive information that Envirofacts has never contained, as well as data from a number of program and administrative systems not currently included in Envirofacts. So while Envirofacts is an excellent starting point, the overall task of creating an efficient Agency data model requires more than this single reference.

EPA is approaching the task from five directions. It will build on:

1. The existing **Envirofacts** model, which serves as a point of departure.
2. An original, top-down **Strategic Information Model** (SIM) as a high level blueprint for the enterprise data model.
3. The **Core Reference Model**, created by EPA and the Environmental Council of States (ECOS) as a guide for consistently building and sharing data on the Exchange Network.
4. The Agency's **Environmental Indicators** project, a major reference for the EPA Strategic Plan and performance indicators work.
5. The ASA **Data Analysis Initiative**, described in Exhibit C: Part 3.

Items 2, 3, and 4 are discussed in turn below.

[4] Permit Compliance System, National Compliance Data Base, OECA Docket, RCRAInfo, AIRS Air Quality Subsystem, AIRS Air Facility Subsystem, System of Risk Management Plans (SRMP), CERCLIS, SDWIS, TRIS, and STORET.

The Strategic Information Model

A number of federal agencies have created strategic data models as a central structure for their enterprise data architectures. The General Accounting Office (GAO), for example, has found that having such a model as a guide for systems applications development greatly reduces design time and increases the flexibility and adaptability of systems.

The SIM can help respond to findings and recommendations by providing a clear view of the operative elements and how they relate. For example, working from the EPA Strategic Plan, for instance, it is clear that "designated use" is the common element between what is known about a body of water and the water quality standards EPA expects to be applied in an environmental assessment. The EPA SIM specifies that:

- Water quality standards only make sense in the context of designated use.
- A water quality standard may be appropriate for more than one designated use.
- Every body of water has at least one designated use.
- A body of water may have multiple designated uses.
- Designated use for a body of water may change over time.

Knowing how these elements relate will significantly enhance the quality and completeness of planning to respond to emerging needs and recommendations. The SIM structure for this analysis is shown on Figure B-6.

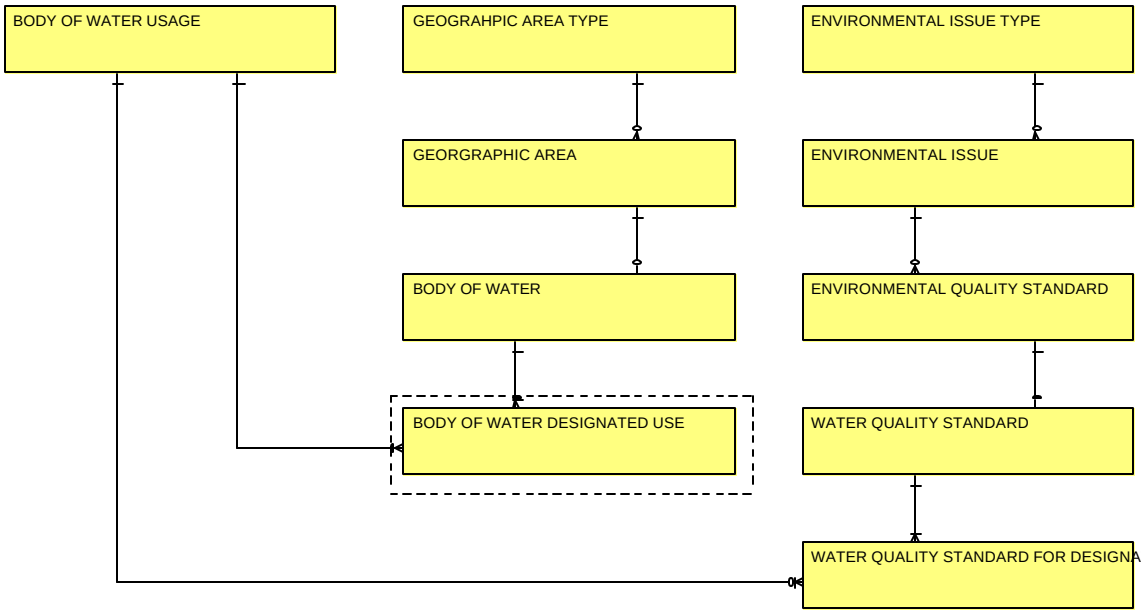


Figure B-6: Example from the EPA Strategic Information Model (SIM)

A SIM is an entity-relationship diagram (ERD) that identifies and defines the information required to implement strategies as described in an organization's strategic plan. A SIM consists of entities, attributes, and relationships created in response to information requirements referenced in the plan.

A SIM is similar to a standard logical data model except for an intentional absence of the sort of detail required for operational-level implementation (e.g., data type, data format, operational control elements, etc.). Keeping the model on the strategic plane makes it possible to develop one model that spans the full extent of the enterprise.

The completed SIM is a fully qualified, fully normalized entity-relationship diagram. This sort of analysis provides insight into underlying key structures for data that may have been viewed as separate and unrelated. In developing the SIM, one models each environmental area separately, then consolidates components that are essentially identical. Preliminary consolidations suggest a common key pattern for observations, standards and performance measures that would expedite development of comprehensive performance indicators and the data exchange network. These findings will be validated by subject matter experts in FY 2004.

In the consolidation phase (the SIM development process is described below), EPA will perform cluster analysis to identify implementable subsets and apply business expertise to group clusters within subject areas. Business-based subject areas provide a rational scheme for assimilating operational-level data models into an enterprise data model. In the assimilation process one of three things will happen:

- There will be a clear home for operational-level data in the higher-level information framework and the element will be assimilated into the enterprise data model, or
- An operational-level element will seem "out of place," which will precipitate a conversation between strategic planners and operational level business experts, or
- Operational-level elements will reveal distinctions not revealed by the higher-level information model and we will update the framework to accommodate the distinction.

EPA's first draft SIM has been completed. Over the first quarter of FY 2004 the EA Team will begin validating the model with the program offices through the EACC.

The Core Reference Model

Building from the opposite direction—bottom up rather than top down—EPA will benefit greatly from work done by the Information Management Work Group (IMWG)[5], which is producing an empirically derived common business "framework" for sharing environmental information across the Exchange Network.

Four states—Michigan, Arizona, Nebraska, and Delaware—assisted in the compilation of the CRM. Its data elements were gathered by examining all environmental reporting forms used by any office in any of the states. The value of this work is that it is fully grounded in documented state needs. In addition, consolidating data elements among the states, which exchange data with EPA, is a valuable aid in metadata reconciliation—one of the highest priorities in implementing the EPA target architecture *and* the environmental information Exchange Network.

The CRM is based on four conceptual components:

- **Data Element:** a single unit of data that cannot be divided and still has useful meaning (e.g., *city name, state name, zip code*).
- **Data Block:** a grouping of related Data Elements that can be used and reused among different information flows (e.g., *address identification*, which includes the component

[5] IMWG was formed by the Environmental Council of States (ECOS) and EPA in 1998.

Data Elements such as *city name, state name, zip code*). There are currently 58 data blocks in the CRM.

- **Compound Data Block:** a grouping of related Data Blocks (e.g., *environmental interest,* which may include Data Blocks such as *contact, address identification,* and *facility activity*.) There are currently 34 compound data blocks in the CRM.
- **Major Data Group:** a logical grouping of related Data Blocks and Compound Data Blocks to fully describe business areas, functions, and entities where EPA and its partners have an environmental interest. There are currently 18 major data groups in the CRM.

Relationships among these elements are shown in schematic form in Figure B-7.

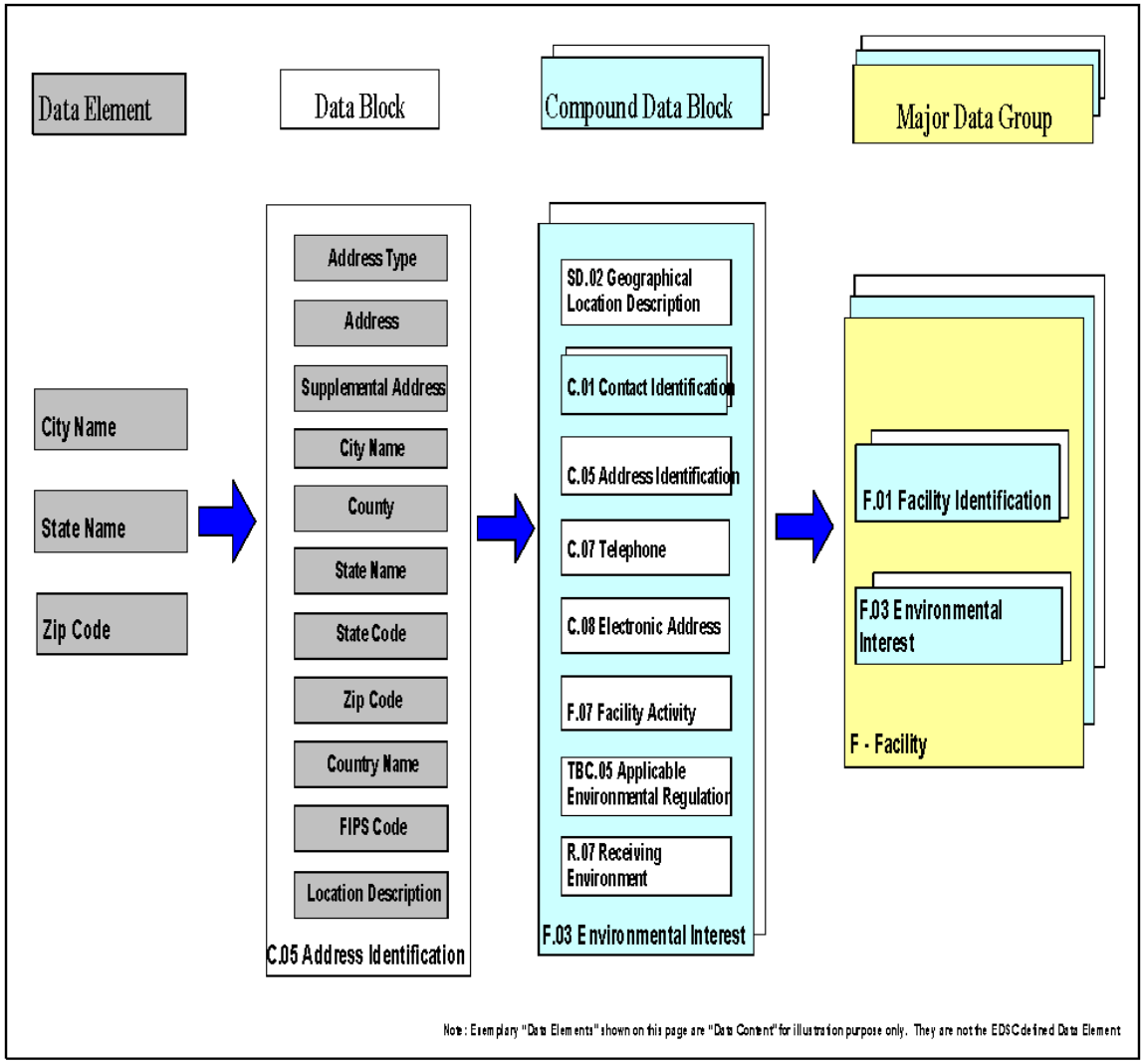


Figure B-7: Relationship of Four Major Conceptual Components of the CRM

The major data groups more fully describe business areas, function, and entities where EPA and its partners have an environmental interest. The enterprise data model will be able to draw from these as a guide to reflect common relationships among data. This information will be particularly important to help align EPA's enterprise data model with the needs of state and tribal partners and

flows of data on the Exchange Network. A complete list of the major data groups is shown in Figure B-8.

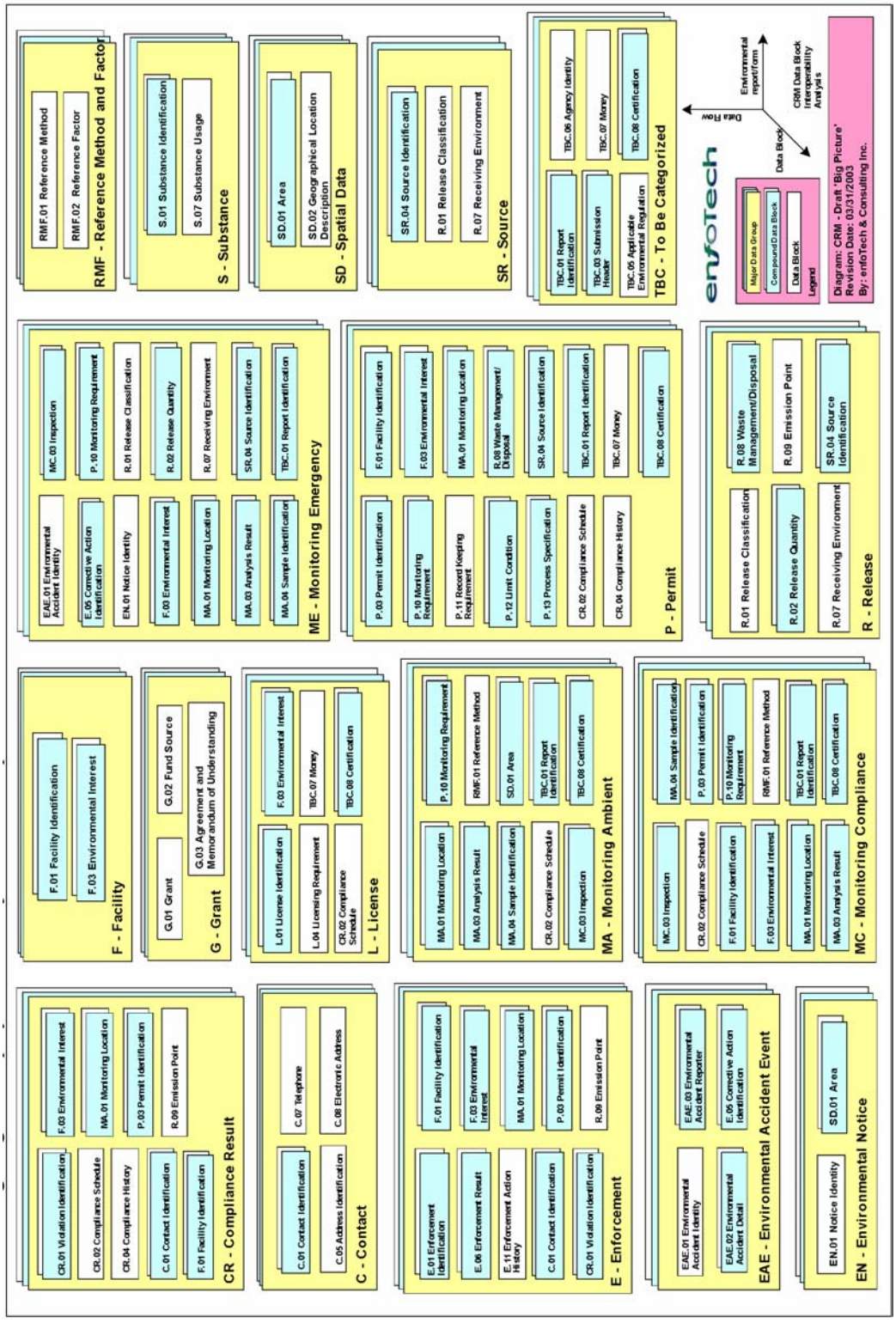


Figure B-8: CRM Major Data Group inventory

Environmental Indicators Project

The third major contributor to the development of the enterprise data model is the EPA Environmental Indicators project, which has contributed substantially to the development of the new 2003 EPA Strategic Plan and the EPA Draft Report on the Environment 2003. This initiative has screened available scientific data to identify the data sources that can best address the questions of highest concern to EPA in the areas of cleaner air, purer water, better protected land, human health, and ecological condition.

The indicators project goes well beyond EPA's internally developed data sets. Much of the data it encompasses will never be directly included in the core FBW, but access to selected data via Web services may well be required to measure compliance with goals and objectives incorporated in the Agency's Government Performance and Results Act (GPRA) goals and the Federal Enterprise Architecture Performance Reference Model (PRM).

Data Sensitivity Considerations

With all internal and external data queries being directed to the FBW, data sensitivity becomes a primary concern. All users will access the FBW through the Agency's Portal, where a central Identity Management System will authenticate their credentials and assign them appropriate access to the Agency's many data sets. Public users will have access only to data that is cleared for public use. Internal users and external partners will have greater access to data, but their access must be tiered according to their differing privileges. For instance, states often mark certain enforcement or inspection-related data to be non-public—perhaps as not to be shared even with EPA staff.

Every data element of the FBW must therefore be classified according to a comprehensive data sensitivity model that determines who can have access, when, and under what circumstances. This data sensitivity model will be developed in parallel with the enterprise data model.

Part 3
Updates to the Applications Architecture

EPA's architectural program was founded in 2000 on several predecessor efforts, among them the "Information Integration Initiative," or I-3, an internal research program whose goal was to integrate the management of data across the Agency. I-3's concept was to decompose major program applications into their functional components: collecting data, processing data, storing data, and using data. Three of these functions (collect, store, and use) would be merged into a common application infrastructure. Only data processing would remain unique to individual programs. These concepts were presented in the *Model for Integration* report (M4I), which was discussed at length in EPA's target architecture submission of December 2002.

I-3's goal was to solve the historical stovepiping problem created (and exacerbated) by EPA's institutional heritage: it was created as a hybrid agency from preexisting programs and authorized under different statutes. The M4I concepts were developed under the current architecture program and became the basis of EPA's target Environmental and Health Protection Architecture (EHPA)—that portion of the architecture that addresses Services to Citizens. The heart of the EHPA is a strategy to provide a shared set of "central services" to major Agency information systems: a shared Agency Portal, a shared Central Data Exchange (CDX), shared reporting tools, and a repository system that couples business area warehouses together. At the same time, EPA convened separate architecture efforts to address administrative systems and research and science systems, whose challenges (and solutions) are fundamentally different from those addressed by M4I.

Integrating the Applications Layers of the Environmental and Administrative Architectures

The Administrative Systems Architecture (ASA) focused on the modernization and consolidation of legacy systems by eliminating redundancies, automating manual operations, identifying systems of record to be the sole source of administrative data (PeoplePlus, Financial Replacement System, etc.), as well as integrating administrative programs with an array of pertinent E-Gov programs under the President's Management Agenda (PMA). Its activities fall under the Support Delivery of Services and Management of Government Resources layers of the FEA BRM.

The Research and Science Architecture[6] (RSA) has focused primarily on improving the management of research projects, and on technology issues of bandwidth and high performance computing.

Although these separate efforts continue today, EPA committed, in its December 2002 submission to OMB, to integrating its three business domains into a single comprehensive architecture. The process is a gradual one. All three architectures have made most progress at the

[6] In the EA Target Architecture Version 1.0, this domain was called the Research and Development Architecture (RDA) because it was centered on the Office of Research and Development . For 2003, the Agency expanded the RDA to encompass all research and science activities and systems in place across all program offices and regions. It therefore changed the name to RSA.

applications layer, because that is where the most pressing problems happen to lie. It makes sense, therefore, that architectural integration should begin with the Target Applications Architecture (TAA), though a great deal of progress has also been made at the business and performance layers (see Exhibit B: Part 1: Strategic and Business Architectures). Here the concept of domain "integration" involves extending the central services to support administrative and science applications as well as environmental applications.

In its target architecture, the ASA solves its application integration problems by applying an Enterprise Application Integration (EAI) interface solution. The administrative warehouse can become a Business Area Warehouse within the general business warehouse framework. Retrieval of administrative information can be handled through the Administrative subportal, that subpart of the Portal that will house administrative data analysis and reporting tools. The result is shown in Figure B-9.

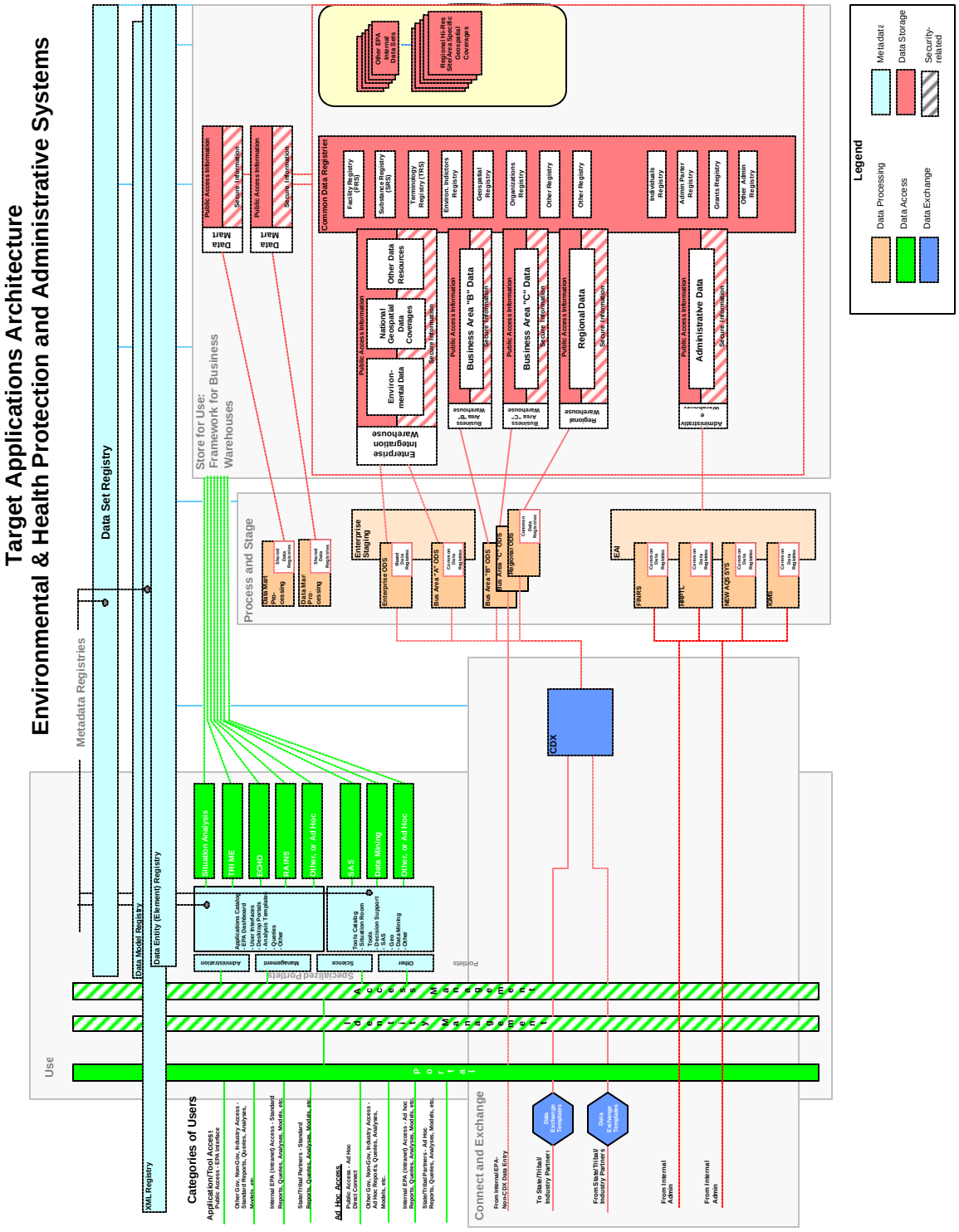


Figure B-9: Schematic View of the Version 2.0 Applications Architecture

This integrated applications architecture is a work in progress. Details remain to be worked out before it is incorporated into the construction schedule of the Enterprise Information Integration and Portal Development (EIIPD) program managed by the Office of Environmental Information

(OEI), which is why it is *not* addressed in EPA's sequencing plan, presented in Exhibit C: Sequencing Plan.

Advantages gained by integrating the applications architecture include:

- **Sharing of data registries:** The EHPA applications architecture relies heavily on a set of shared data registries—segments of the FBW that house commonly used data on such topics as regulated facilities, regulated chemicals, and partner organizations. Extending the common services to support the administrative domain will allow administrative and environmental applications to share data on partner organizations, and perhaps other common data elements. It will be far easier, for instance, to track the environmental performance outcomes of grants given to states or universities.
- **Integration of user access through one Portal:** The construction of a single Agency Portal will integrate identity management and access control. The planned administrative specialty portal will integrate with the general Agency portal, facilitating and simplifying access to EPA information by the public and states as well as by federal workers.
- **Enforcement of data standards through metadata registries:** EPA metadata will be merged with the set of metadata registries that control flows of data within the central services infrastructure. For instance, contact information for grantees and contractors will follow the same metadata standards as contact information for state and tribal partners in environmental programs. This greatly improves the ability of the central services to access information reliably and quickly. It will be invaluable for achieving the goal of improving services to citizens and implementing a wide array of E-Gov initiatives under the President's Management Agenda.

It is still not clear whether data flows to administrative systems should be directed through CDX and the Exchange Network, and whether they should flow directly through the administrative subportal to the ASA operational data stores (ODSs).

EPA will work in coming months with the RSA development team to integrate research and science systems with the enterprise applications architecture. The same logic will apply: the central services should be extended to support, as appropriate, all EPA systems. RSA's administrative management systems should have access to the common Data Registries that address organizational identity and other shared data needs. RSA data storage should be integrated with the FBW. RSA metadata should be actively controlled within the Agency metadata registries (in fact, core aspects of the metadata registry systems, notably the Environmental Information Management System (EIMS), were developed by the Office of Research and Development). Portal functions of identity management and access control will need to be shared by researchers and scientists, and a specialized science subportal will service user access to analytical tools and applications.

Data Storage: Expanding the Use of Web Services

Of the many issues that have come to light since December 2002, one of the more important is the possibility of creating a new Web services mode for linking business area warehouses. Version 1.0 of the target architecture provided three options for an application to store its data.

- It can elect to store it in the **Enterprise Integration Warehouse** (EIW), using the EPA-standard relational database, which will be built to handle the combined data storage needs of multiple mission-critical applications, much as Envirofacts does. The EIW will also contain the common Data Registries.

- It can build its own **Business Area Warehouse** (BAW), containing data from one or more systems, such as those run by the Office of Air and Radiation or the Office of Water. Business Area Warehouses must use the same version of the EPA-standard database as the EIW and must link directly to the common Data Registries.
- If its data is not mission focused, the application can store its data in a different type of database, not necessarily the EPA standard, which can be separate from the Enterprise or Business Area Warehouses and not linked directly to the common Data Registries.

For greater flexibility and to take advantage of emerging technology trends, data storage options include a fourth mode of integration—linking multiple warehouses with the Data Registries via Web services. This will relax the requirement that all mission-focused data be stored using the EPA-standard database.

No matter which option a mission-focused application chooses, most technical requirements will be the same. Data storage must conform to all metadata requirements, including unified data element definitions. The physical data models of individual warehouses, whether distributed or centralized, must map to the unified enterprise data model discussed above. And all mission-focused data must connect directly to the enterprise Data Registries, which will contain the master copy of information on regulated facilities, substances of interest, organizations, and other categories of data essential to the regulatory process.

This Page Intentionally Blank

Exhibit C

Enterprise Architecture Sequencing Plan

This Page Intentionally Blank

Introduction to the Enterprise Architecture Sequencing Plan

EPA's December 2002 submission to the Office of Management and Budget (OMB) presented the Agency's target Enterprise Architecture as broken down into three Business Domains and a number of cross-cutting "component" architectures. Figure C-1 below shows the scheme of the target architecture as it was then.

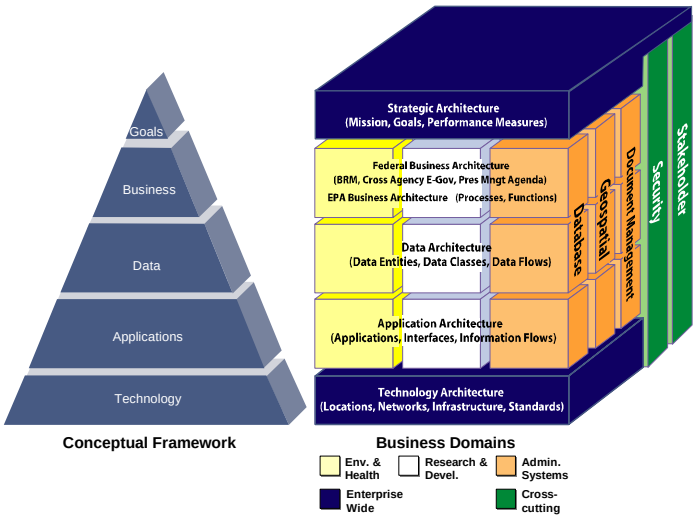


Figure C-1: EPA's Enterprise Architecture Elements

In the December 2002 submission, the Environmental and Health Protection Architecture (EHPA; left, yellow column) included a target architecture but did not include a sequencing plan. This submission provides that sequencing plan, which now includes the geospatial "component" architecture from last year. The Administrative Systems Architecture (ASA; right, orange column) *did* present a sequencing plan—this submission updates it. The Research and Development Architecture (now the Research and Science Architecture, RSA; middle, white column) is still developing its target architecture and sequencing plan, which will be complete by the second quarter of FY 2004.

The centerpiece of the EA sequencing plan is creation of a shared set of central infrastructure services that will require the restructuring of all operational systems. Each system will have to use a central Agency Portal for all data access and for connecting to all applicable external E-Gov systems. Various specialty portals will be constructed within the Agency Portal to service audiences such as administrative personnel and scientists. All systems will connect to a single Identity Management Solution and use common security mechanisms. All will apply commonly available data standards and metadata standards to organize their data, as well as common technology standards and system development practices to build their applications.

Exhibit C is organized in five Parts, as follows:

Part 1: Aligning EPA's Target EA with E-Government discusses the alignment of EPA's target architecture with the E-Government initiatives laid out by the President's Management Agenda (PMA). EPA is an active participant in 14 E-Government initiatives and has a leadership role in two.

Part 2: EHPA Sequencing Plan below presents the Environmental and Health Protection Architecture (EHPA) sequencing plan. The first and largest part of this section is its discussion of the Environmental Information Integration and Portal Development (EIIPD) project. Creation of these IT central services is the main feature of the target EHPA.

Some of the services, such as the Central Data Exchange (CDX), geospatial services, the Facility Registry System (FRS) and the Substances Registry System (SRS) are in service now. In FY 2004 the PMO will build proofs-of-concept for shared data dictionaries, the Portal, Identity Management, the Framework for Business Warehouses (FBW), and Analytical Tools. The central services will be fully functional in FY 2005 with at least one major system fully functional using all components of the central services.

Environmental systems will exchange all their state data and will accept industry submissions through CDX, working with their partners to develop the necessary exchange templates and partnership agreements. They will decide what technical method they will use to connect to the FBW and will redirect their applications to access data from the FBW. They will use shared data registries for commonly used groups of data.

Part 3: Administrative Systems Architecture Sequencing Plan Update section lays out the major ASA IT initiatives necessary to build the ASA target, including steps to install a common technology for integrating system interfaces and to integrate administrative data definitions. ASA applications will use a shared administrative warehouse and administrative data registries. They will use an administrative subportal within the Agency Portal and will integrate their data through the administrative data initiative.

Part 4: Sequencing Applications to the Target discusses how individual program systems will migrate physically to the EIIPD central services. Also discussed within the EHPA sequencing plan are individual program initiatives that will also integrate with, or depend upon, the new central services.

Part 5: Technology Architecture Sequencing Plan presents progress on the Technology Architecture Sequencing Plan (TASP). The Agency's Technology Architecture is a component architecture that supports all the business domains. Its sequencing plan therefore focuses on how the underlying technologies of EPA's Technical Reference Model are being phased in over the next few years.

Part 1
Aligning EPA's Target Enterprise Architecture with E-Government

One of the top priorities in EPA's first target architecture sequencing plan is alignment with federal E-Government initiatives. EPA is an active participant in 14 E-Government initiatives and has a leadership role in two.

EPA is providing federal leadership in E-Rulemaking and E-Records Management:

1. **E-Rulemaking:** (Lead Agency: EPA) EPA is the Managing Partner of the E-Rulemaking Initiative. With the help of other agencies it launched *Regulations.gov* in January 2003. From the Regulations.gov Web site, the public can view and download every federal proposed rule currently open for comment and submit comments online to the federal agencies that have created them. Since its launch, the Web site has received over 1.3 million connections, averaging between 60,000 to 70,000 hits a day and is evidence of the public's involvement in the decisions that affect their lives and their need for a more citizen-centered and results-oriented government. While in the beginning stages of creating the system architecture, the team is taking the opportunity to unify XML tags throughout the rulemaking process. The E-Rulemaking project team is leading a cross-agency collaborative effort with GSA, GPO, NARA, FCC, EPA, OMB, and other agencies to create an XML schema for the entire rulemaking process. The E-Rulemaking Initiative has moved into the crucial second phase of the project—creating a single, government-wide electronic docket system that will give the public access to all rulemakings, supporting regulatory documents, and public comments through one Web site. The E-Rulemaking Initiative has been selected as one of the ten finalists in the Federal Executive Leadership Council Showcase in Excellence Award.

   Potential EPA System Linkage: E-Dockets

2. **E-Records Management:** (Lead Agency: National Archives and Records Administration) EPA has agreed to produce four deliverables under the Enterprise-wide Electronic Records Management (ERM) project of the National Archives and Records Administration's E-Records Management, E-Gov Initiative. These will help other agencies understand the technology and policy issues of procuring and deploying an enterprise-wide ERM system. These products are: 1) guidance for evaluation Capital Planning and Investment Control proposals for ERM systems; 2) guidance on developing agency-specific functional requirements for ERM systems; 3) guidance on developing and launching a ERM proof of concept pilot; 4) a lessons learned paper from EPA's and other federal agency's ERM proof of concept pilots. The four products to be produced under the Enterprise-wide ERM project are based on EPA's Electronic Records and Document Management System Project. This project is an enterprise-wide effort to develop and deploy and electronic records and document management system.

In addition, EPA plans to modify its systems to meet the requirements of 12 other E-Government initiatives by aligning with, and developing robust interfaces to, the following government-wide efforts:

3. **Business Compliance One-Stop (BCOS):** (Lead Agency: Small Business Administration) The Business Compliance One-Stop (BCOS) will help reduce the burden on businesses by helping them find, understand, and comply with governmental laws and regulations through a cross-agency, intergovernmental web portal. To support this effort, EPA is working with the SBA and other Agencies to create a "profiler" that allows businesses to quickly find compliance assistance tools that specifically relate to their business. The profiler will gather certain information about the business activities and use the information to search for and present the business with a list of relevant compliance assistance tools. The profiler will be developed in phases. Initially a database of links to EPA, OSHA, IRS, DOL and DOE compliance assistance tools and resources will be created. Data records will be tagged with appropriate key words to enable the profiler to search the database to identify relevant tools. In the long term, it is expected that the profiler will be capable of seamlessly conversing with Agency databases to identify relevant resources and the maintenance of a separate database will no longer be necessary. SBA anticipates demonstrating the phase I profiler in July 2003.

4. **E-Authentication:** (Lead Agency: General Services Administration (GSA)) The purpose of this project is to deliver a common interoperable authentication solution to match levels of risk and business needs of each E-Government initiative, thereby reducing the number of credentials issued by the federal government. EPA completed a white paper on potential implications of E-Authentication gateway on target security architecture for Central Data Exchange (CDX). The results will be used to help inform CDX and Exchange Network approach to security. EPA also began discussions with NASA and EPA's Region 6 on possible use of gateway in support of New Mexico XL project, but dialogue was disrupted by recent NASA events. EPA circulated recent E-Authentication materials (i.e., policy, risk assessment guide) for internal comment, and also presented at state-EPA forum on security. In addition, EPA developed/coordinated Agency response to a recent GAO survey on E-Authentication/PKI. EPA also conducted an in-house "informal" survey of authentication requirements in early 2003. Over five years, annual requirements include 100,000 new or renewed "registered users" and over 3 million authenticated actions. The results indicate there to be a small net benefit ($350k over five years) for building to gateway. EPA is continuing to engage in the project through GSA-led weekly meetings and related events.

5. **GeoSpatial One-Stop:** (Lead Agency: Department of Interior) This project will provide federal and state agencies with single-point of access to map-related data enabling consolidation of redundant data. OEI has developed a Geospatial Data Index (GDI) which allows EPA employees to search for and link geospatial data and associated metadata from across the Agency (http://intranet.epa.gov/geoindex/). A publicly accessible version of the GDI is being developed, and along with the Geospatial Data Clearinghouse (http://www.epa.gov/nsdi/), it will provide a foundation for developing EPA's portal to the government-wide Geospatial One Stop Initiative.

   Potential EPA System Linkage: Geospatial/GIS

6. **Recruitment One-Stop:** (Lead Agency: Office of Personnel Management (OPM)) The goal of Recruitment One Stop is to position the federal government as the first stop for job seekers by creating an on-line experience that offers all the features sophisticated job

seekers have come to expect from commercial sites. Human resources professionals will have new recruiter tools that will enable them to identify top candidates more easily and to manage the recruiting process more effectively. Upon project completion there will be *one* job search engine for competitive service federal job postings; *one* standard display for vacancy announcements; and one resume builder to create the basic application document. The end result will benefit citizens by providing a more efficient process to locate and apply for jobs, and assist federal agencies in hiring top talent in a competitive marketplace. Ultimately, we hope to see improvement in the quality level of new hires and potentially a reduction in the amount of time required to source candidates.

Potential EPA System Linkage: HR PRO

7. **Grants.Gov:** (Lead Agency: Department of Health and Human Services) The Grants.gov project will simplify the grant application and reporting process across all federal agencies. Grants.gov will produce a simple, unified "storefront" for recipients of federal grants to electronically find opportunities, apply, and report on grants and facilitate the quality and efficiency of operations for grant makers by eliminating the need for them to build stovepipe systems for recipients. EPA has all current grants solicitations posted on-line as required by the "find" phase of the project. EPA is a member of the Electronic Standards Work Group which is defining the apply data elements. EPA is also one of the early participating pilot agencies for the first phase of the project.

   Potential EPA System Linkage: IGMS

8. **Integrated Acquisition:** (Lead Agency: GSA) Through this project, agencies will begin sharing common data elements to enable other agencies to make better-informed procurement, logistical, payment, and performance assessment decisions. This project would automate the interagency ordering of goods and services and the transfer of funds. IGMS and IFMS will be required to exchange data with this system. EPA staff is actively involved in each of the core modules of this project. This work includes the following workgroups: 1) E-Cat Workgroup, which is identifying the difference between E-Catalog, E-Mall, E-storefront and GWACs; 2) Business Partner Network (BPN), in which all business partner information will be maintained by that business partner including the Electronic Funds Transfer (EFT), Tax ID Number (TIN), and Certifications; 3) E-Trans Workgroup, which identified nine common processes that are to be implemented as Phase One of the Integrated Acquisition Environment; 4) FPDS Workgroup, which was formerly known as FAMIS and now the Federal Procurement Data System; and 5) I-GOTs Workgroup.

9. **E-Training (GoLearn):** (Lead Agency: OPM) This project will provide all EPA employees with access to a single point of entry for extensive electronic, web-based training. The intent is to use the new GoLearn site, created in response to the President's Management Agenda, as our Virtual University portal. To fully integrate with GoLearn, the EPA Institute and the Office of Environmental Information have formed an agency-wide E-Learning Workgroup to:

   – Catalog all current and planned uses of E-Training throughout EPA;
   – Prepare and implement an outreach plan to ensure that EPA program managers and employees are aware of GOLEARN opportunities;
   – Develop an Agency-wide strategy for E-Training efforts, including a transition plan to meet OMB's mandatory requirement to move current and future training programs to GOLEARN;

- Establish a policy for granting exemptions for any E-Training effort deemed inappropriate for inclusion in the GOLEARN environment; and
- Develop content, technology, and performance standards for any E-Training program that will continue to reside on EPA servers.

The EPA Institute will introduce E-Learning via GoLearn in a planned approach intended to make usage attractive and desirable to both management and staff. This plan includes designing and implementing a staged agency-wide marketing plan for GoLearn.

10. **Disaster Management Initiative:** (Lead Agency: Federal Emergency Management Agency) The goal of this project is to provide federal, state, and local emergency managers online access to disaster management related information, planning, and response tools. EPA and NOAA are working with FEMA's contractor to adapt EPA's CAMEO to make use of some of the capabilities being developed under the E-Government initiative. EPA is working with FEMA/DHS to address populating the portal for disasterhelp.gov and offer technical assistance to the project's next directions. The Memorandum of Understanding that will transfer funds for this project is being negotiated between FEMA, OMB, and EPA.

    Potential EPA System Linkage: CAMEO

11. **Safecom Wireless:** (Lead Agency: Department of Homeland Security) The goal of this project is provide interoperable wireless solutions for federal, state, local, and tribal public safety organization and ensure they can communicate and share information as they respond to emergency incidents. This project has recently been moved to DHS staff in an attempt to ensure progress on this critical initiative. EPA stands ready to support the new project team when they re-orient the project's direction.

12. **Enterprise HR Integration:** (Lead Agency: OPM) The Enterprise Human Resource Integration (EHRI) initiative will improve the internal efficiency and effectiveness of the federal government by streamlining and automating the exchange of federal employee HR information. EHRI will eliminate the need for a paper employee record, better protect the rights and benefits of the Federal workforce, enable the electronic transfer of HR data throughout the federal sector, and streamline and improve government-wide reporting and data analyses. EPA continues to participate in the EHRI Integrated Product Team component of the project.

    Potential EPA System Linkage: HR PRO

13. **E-Payroll:** (Lead Agency: OPM) E-Payroll is consolidating 22 federal payroll systems to simplify and standardize federal human resources/payroll policies and procedures to better integrate payroll, human resources, and finance. Last year, OMB asked agencies to declare their interest in being considered on of the three to four primary federal payroll providers. While EPA is not a primary payroll provider, it has declared its interest in being a strategic Payroll Service Provider (PSP) partner with one of the final selected providers. Alignment with one of the selected providers will allow EPA to leverage the substantial agency investment made to modernize the human resource and payroll functions. The EPA/PSP partnership would take advantage of EPA's "lessons learned" by having the agency serve as the development lab for a PeopleSoft COTS solution.

    Potential EPA System Linkage: EPAYS, HR PRO, IFMS, FINRS

14. **E-Travel:** (Lead Agency: GSA) The E-Travel vision is to create a unified, simplified service that delivers a cost-effective travel experience, supports excellent management

and results in superior customer satisfaction. GSA announced government-wide Online Booking Engine. The online booking engine, FedTrip™, is a self-service, online booking reservation system that provides federal travelers with convenient 24-hour access to reservations, profiles and itineraries. EPA will use the E-Travel service to unify and simplify travel process for agency employees. EPA has assisted GSA's E-Travel project team develop an end to end solution request for proposals (RFPs).

Potential EPA System Linkage: Travel Manager

This Page Intentionally Blank

Part 2
Environmental and Health Protection Architecture Sequencing Plan

The target EHPA is the most ambitious IT project in EPA's history. It will require the complete reengineering of the Agency's largest and most mission-focused applications.

This work will take place over a period of years under the direction of the newly-formed Program Management Office (PMO) within OEI. The PMO's prime responsibility is to build and run the EIIPD project—the major central services elements defined by the EHPA Target Applications Architecture (TAA). Over the past several months, the PMO has brought together top IT talent from across EPA to form a highly skilled, technically experienced management and project team.

The PMO's overall strategy is to bring up all elements of the EIIPD simultaneously so that one major program system (in the process of being identified) can be fully operational within the system in 2005, resources permitting. Developing the components together will encourage collaborative problem solving and help managers overcome common technical problems as they emerge.

Some of the necessary components—CDX being the most notable—are already operational. Others, such as the Framework for Business Warehouses (FBW), are still at the project development stage. A few, such as the Facility Registry System and the Environmental Data Registry, are operational but must be modified to perform new functions. The PMO's goal for 2004 is to develop proofs of concept for all the EIIPD components, including functionality to support a Portal with single sign on for collection and access to applications. The long term goal for 2005 is to have one major system operational within the system, with all program office data moving through all the EIIPD components.

Management Approach

The mission of the PMO is to coordinate and communicate the design and development of OEI's critical central services components and provide service to the program offices using these components.

The PMO is a matrixed organization within OEI, using staff, knowledge, and expertise from across the organization (see Figure C-2). The PMO includes a program manager, professional support staff, portfolio managers, and contract support. Portfolio managers are responsible for the development and interfaces of the individual components that OEI manages.

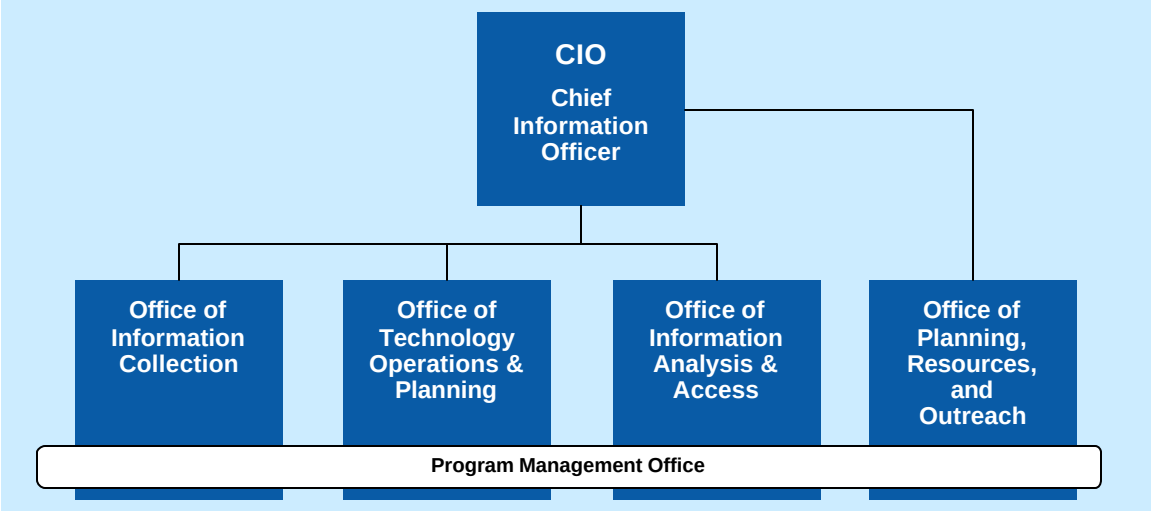


Figure C-2: The Program Management Office in Relation to OEI's Organizational Structure

The EHPA identified six major areas in the target architecture. The PMO has organized this list into eight more detailed project areas and has assigned Portfolio Managers accordingly. The list of original EHPA areas and the corresponding portfolio areas is shown on Table C-1.

Table C-1: Assignment of PMO Portfolio Responsibilities

EHPA Components	Portfolio Managers
Central Data Exchange	Central Data Exchange
Enterprise Repository	Framework for Business Warehouses
Data Registries	System of Registries
Metadata Registries	
System of Access	Portal
	Analytical Tools
	Identity and Access Management
Geospatial Services Program	Geospatial Services
Operational Data Stores	Operational Data Stores

In addition, the EIIPD project will collaborate with the EA Team to ensure a common approach and avoid effort duplication and overlap. The EA Team and the PMO coordinate regularly to ensure consistency and an overall understanding of the status and activities of each project. The PMO will report quarterly to the OEI Board of Directors,[7] which will serve as a governance structure to provide guidance and instruction regarding progress, changes in direction, and opportunities to leverage existing work. In addition, as a part of its management approach, the PMO will hold weekly team meetings with the portfolio managers to review schedule, progress, issues, and possible schedule ramifications.

[7] The OEI Board of Directors includes the Chief Information Officer (CIO)/Assistant Administrator for OEI, the Principal Deputy Assistant Administrator, and the OEI Office Directors.

The EIIPD project will outsource technical work to a diverse, existing contractor base with direct experience in supporting the EIIPD components. It will also benefit from knowledge obtained through the implementation of CDX, the first component of the project in active use. The EIIPD will review documentation, technical analyzes, and strategy pieces developed by the CDX project team, incorporating them into the EIIPD management approach to help avoid duplication and build on existing IT solutions.

Given the impact that EIIPD will have on information management for the Agency and requirements for EPA program office systems, the PMO is implementing a comprehensive communication plan, which it will develop and expand as customer needs change and the project evolves.

EPA's target architecture is likely to evolve as the EIIPD is built. EPA is already exploring integration of the Administrative Systems Architecture (ASA) and Research and Science Architecture (RSA) into the EIIPD structure, which may entail changes to the target EA. The PMO is examining new technical options for constructing the FBW that may also, if adopted, lead to adjustments in the target architecture.

Many, if not most, of the technical and management details of this sequencing plan have yet to be worked out. What follows is a high-level overview of sequencing issues. The Agency will provide OMB with periodic updates.

This Page Intentionally Blank

Building the Central Services

Over the past several months, the PMO and its portfolio managers have been examining the many technical issues that must be resolved to allow construction of the EIIPD. The following pages discuss the current status of these technical issues. As the project continues, more issues may surface and technical direction may change. This discussion reflects thinking and priorities current as of June 2003.

Overall EIIPD Project Sequencing

The EIIPD project includes many interactive components. To clarify how the components fit together, Figure C-3 lays out the interrelationships of all the functions and components of the full EHPA target applications architecture in a simplified fashion. It shows the basic steps and requirements for bringing data into the system (red pathway), querying the system to get information out (green pathway), and managing metadata (blue pathway).

Each object along the pathways represents a functional step that must be completed, or a system component that must be built, before the target architecture can be fully operational.

In 2005, this system will be adequately functional to support the operation of one pilot application, which is in the process of being identified. It will not, however, be fully complete by that time—reconciliation of metadata and other steps may not be fully complete for several years. But the overall structure, logical sequence, and dependencies among systems components and projects is established.

This functional representation of the target architecture embeds Operational Data Stores (ODSs) and Business Area Warehouses (both of which are run by the program offices) into the EIIPD Agency IT central services. The following pages, however, describe the sequencing of EIIPD components only. Integrating ODSs and Business Area Warehouses into the central services is discussed under Application Migration Analysis later in this document. The presentation assumes that a comprehensive funding strategy is in place and that funding is available for these initiatives.

The sequencing plan describes the major activities for each of the major components of the EIIPD project:

- **Central Data Exchange:** A system of Internet-based technologies that facilitate the flow of information from EPA's external partners and stakeholders into the appropriate Agency systems. CDX is an existing system.
- **Framework for Business Warehouses (FBW):** EPA's central data storage mechanism, comprising Agency systems. This is a new component, but it will be initially modeled on the existing Envirofacts system.
- **System of Registries:** A complex set of databases containing Data Registries of use to all systems in the FBW, as well as Metadata Registries to integrate and reconcile the Agency's data element definitions. The EHPA dealt separately with the Data and Metadata Registries, but they are combined into a single PMO portfolio.
- **Analytical Tools:** A set of applications and tools that permit EPA staff, as well as partners outside EPA, to conduct cross-cutting analyses of environmental and health

issues for multiple uses, from managing short term emergencies to examining long term trends. This portfolio performs the analytical and reporting functions the EHPA referred to as the "System of Access" (SoA).

- **Portal:** A gateway through which users of environmental information securely provide data and gain access to Agency information services, analytical products, and data. The Portal was included in the EHPA SoA, but is split out by the PMO as a separate portfolio.
- **Identity and Access Management:** The control of user sign-on, including the assignment of access privileges to all parts of the EIIPD system. This was also included in the SoA. The PMO has separated it out as an independent functionality group. This portfolio is collaborating with the General Services Administration on their E-Authentication solution.
- **Geospatial Services:** A program that will support an internal and external network of shared, distributed geospatial data repositories. This program has been merged into the EHPA; it was previously a separate component of the Agency's EA.
- **Operational Data Stores (ODS):** EPA's individual program office data processing units that receive data through CDX and transmit data to the FBW. This portfolio is still in development because the pilot system has not yet been identified and because it will be staffed outside of OEI. Costs for the ODS components fall to the program offices rather than OEI.

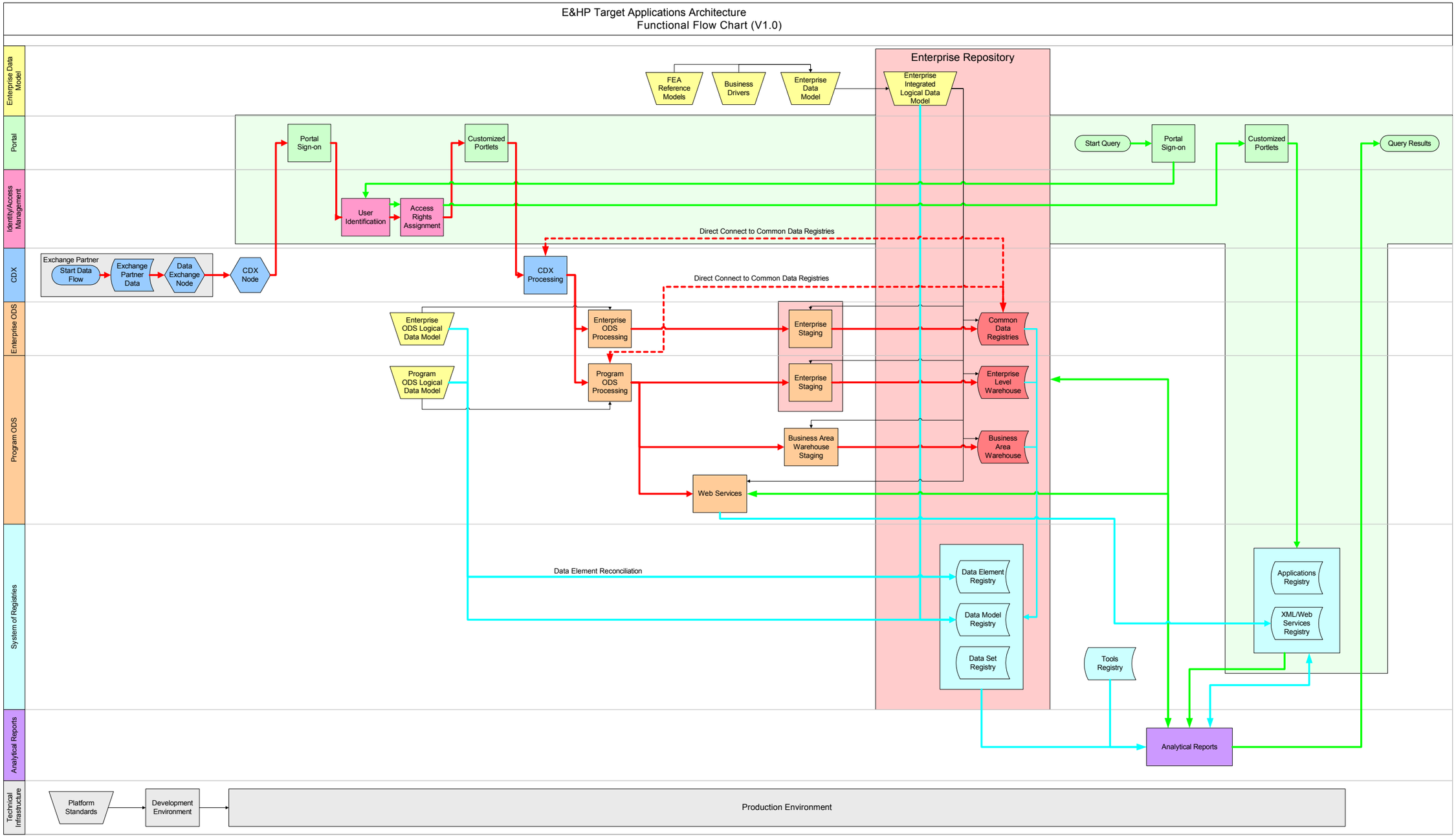


Figure C-3: Functional Workplan: Relationships and Sequencing of the Enterprise Information Integration and Portal Development Project

This Page Intentionally Blank

Central Data Exchange

The Central Data Exchange (CDX) is an existing system of Web services that facilitates the flow of information from external trading partners (states, tribes, and EPA regional offices) into the programs' Operational Data Store (ODS) systems. Seven states currently use CDX to submit and retrieve EPA data. What data is exchanged and how is governed by formal, individual Trading Partner Agreements (TPA).

CDX is EPA's node on the Exchange Network, which has been in existence since 1998. CDX's current priority is on developing additional data flows for existing trading partners and bringing new trading partners into the exchange network. The following data flows are either planned or currently available via CDX (Table C-2):

Table C-2: CDX Data Flows: Current and Planned

Flow	Program Office	Current Status
Facility Registry System (FRS)	OEI	In Operation, but not full production through CDX Node
National Emissions Inventory (NEI)	OAR	In Production
Toxic Substances Control Act Test Submissions (TSCATS)	OPPTS	In Production for: • Electronic Receipt of TSCATS form • Electronic copy of health and Safety Study
Toxics Release Inventory (TRI)	OEI	In Production for: • Electronic Receipt of TRI Reporting Form
Unregulated Contaminant Monitoring Regulation (UCMR)	OW	In Production
Air Quality Sub-system (AQS)	OAR	In Production for: • Ambient Air Monitoring Data
National Emissions Standards for Hazardous Air Pollutants (Radionuclides)	OAR	In Test with Department of Energy
Lead Notification	OPPTS	In Development
Storm water Notification of Intent	OW	In Development
Permit Compliance System (PCS) via Interim Data Exchange Format (IDEF)	OECA	In Production for Interim Data Exchange Format for states that don't require IDEF middleware
Resource Conservation and Recovery Act Information System (RCRAInfo)	OSWER	In Development for: • Subtitle C RCRA Site Identification Form; will be in production by end of FY 2003
Discharge Monitoring Reports (DMRs)	OW	Planned

Flow	Program Office	Current Status
Storage and Retrieval System (STORET)	OW	In Production for Beaches

Impact of Enterprise Data Model

Reviewing the list of current and future flows shows the method by which new data flows are added to CDX. Until an enterprise data model is available, flows are being developed on an ODS-by-ODS basis, with each flow directed to a single ODS. Individualized flows meet the data entry needs of current ODSs, but they do not advance the goal of integrated data management.

EPA has started preliminary work on an integrated data model (see Exhibit B: Part 2 – Evolution of the EPA Data Model). The goal of the enterprise data model is to reorganize data based on its nature and use, rather than on the systems that control it. As the EHPA and the enterprise data modeling efforts move forward, data will no longer be categorized by their controlling systems.

New data flows will be developed based purely on the enterprise data model. Rather than allowing multiple flows to submit the same data, EIIPD will develop a single set of Web services for each category of data. Thus, rather than having a flow specific to the Office of Water's Discharge Monitoring reports and another for the Office of Air and Radiations's ambient air monitoring, a single set of Web services will handle all effluent monitoring data. This set of services will update Business Area Warehouses or the Enterprise Integration Warehouse (EIW) as appropriate, ensuring data integration from the point of data entry.

Unified ODS Data Transfer

The CDX program has concentrated its efforts on providing a single interface for data entry. All trading partner data sent to EPA uses the same set of Web services. Once at EPA, CDX performs a limited amount of processing on the data and passes it to the appropriate ODS using a different method for each ODS. While this solution is the simplest of the ODS systems to implement, it currently creates an extremely complex set of code on the internal data side of the CDX system.

Rather than continue to provide multiple custom methods by which CDX interacts with ODS systems, the EIIPD will adopt a single, or, at most, a dual approach to system integration. The two most likely candidates are Oracle Native Interface (ONI) for ODS systems using the Oracle RDBMS, and Web services implemented on the ODS for non-Oracle architectures. This will simplify and streamline CDX, increasing the number of data flows transacted and decreasing overall development, operations, and maintenance costs.

Common Data Entry Point

CDX is the entry point for data originating from external trading partners, such as for state exchanges and industry submissions, but there is no technical reason for limiting CDX data flows to trading partners. Other external data sources of use to EPA can be accessed via CDX. These flows can be provided as Web services or through some other compatible method. Examples of non-trading partner flows that can be incorporated into CDX include data from the Centers for Disease Control, Dunn & Bradstreet, and the USGS. This would provide a single path for internal users to access external data.

Data Registries Interface

At present, updating of data into the EIIPD Data Registries (currently only the Facility Registry System, FRS, and the Substance Registry System, SRS) is done through the individual systems

using custom routines different for every system. This leads to duplicate code, poor systems connectivity, and increased difficulty in data synchronization and integration.

EIIPD needs a common unified interface for updating these registries. Operational systems should, in the future, never be the input mechanism for updating the Data Registries. Incoming data flows should be routed directly to the registries requiring updates. The logistics of making this change will be substantial, but once a common updating interface is available, data quality and input efficiency will be greatly improved.

To allow the maximum number of systems to interact with this common interface, it must be developed using open standards that are easily implemented on a wide range of technologies. Web services can provide this interface: a layer of Web services placed in front of a registry can expose functions for all systems to use. Streamlining these data flows will ensure that modifications are made in a unified manner.

Framework for Business Warehouses

The FBW is the central data storage mechanism for the EHPA. It is a virtual warehouse against which all users, inside and outside the Agency, are able to perform cross-media, integrated data queries. The architecture of the FBW includes multiple data repositories, implemented in different models, based on use and functionality defined as follows.

Figure C-4 shows the complete EIIPD project in schematic form, including all the options discussed in this section.

Options for Entering into the FBW

The target EHPA architecture presented in December 2002 envisioned three alternative methods of integrating data into the FBW. Since December, the PMO has been working to define the technical details of these alternatives and to explore emerging needs, such as the possibility of linking to the FBW via Web services and the need for data marts to expedite specialized processing. The following sections explore these alternatives.

The EIW

The least complex model for data integration is to transfer ODS to a single EIW, a single instance of Oracle[8] that includes all the Data Registries. Cross-media queries are then addressed to a single database, greatly reducing their complexity.

As new data arrives in a program ODS, it will be made available to the EIW via a staging process. The staging will occur through an Oracle materialized view of the ODS, exposing the data in a schema compatible with the EIW. Data transformation processes (implemented as triggers or stored procedures) will be run to ensure that data is consistent with data already loaded into the FBW. These processes include remapping primary keys, rectifying data inconsistencies, and other functions necessary for analysis and implementation. The resulting data is copied into the EIW.

The EIW is the simplest technical solution for data integration available in the FBW. The Envirofacts system performs almost identical functions for a limited set of data. Implementing the EIW may begin with Envirofacts as a baseline, though the technical specifications of the EIW, staging databases, and data transformation processes have yet to be defined. These specifications will be developed as the initial pilot application is migrated to the EIIPD in 2005.

[8] The EPA standard Relational Database Management System (RDBMS).

EHPA Target Applications Architecture

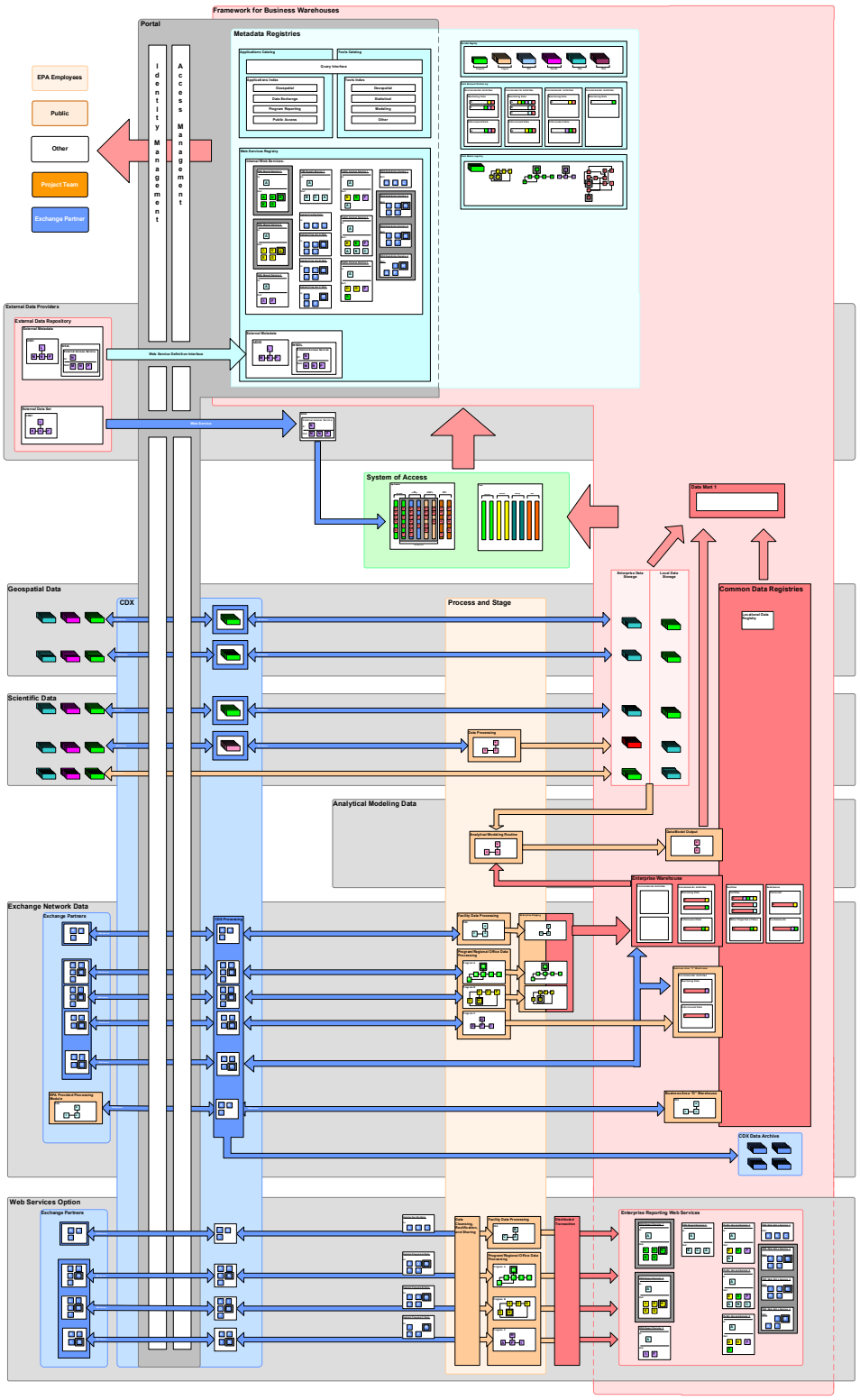


Figure C-4: Complete Schematic of the EIIPD project, Including Web Services Option for Connecting to the FBW

Business Area Warehouses

For programs wishing to maintain their own warehouses rather than store their information in the shared EIW, version 1.0 of the EHPA Target Applications Architecture (TAA) provides for construction of Business Area Warehouses. While this system provides a greater amount of control and flexibility for the program offices, it places several requirements on the Business Area Warehouses that are unnecessary in the EIW. Business Area Warehouses, as specified in the version 1.0 of the TAA, must:

- Never duplicate data stored in other warehouses (as opposed to data marts);
- Use Oracle as their relational database management system (RDBMS);
- Contain their data in a single instance of Oracle, funded, controlled, and maintained by the program office;
- Create database links between the Data Registries instance and their instance;
- Properly format and record their data as part of the Metadata Registries;
- Organize their data in accordance with the enterprise data model.

Provided all conditions are met, program data continues to live in the Business Area Warehouse. Multimedia queries initiated through the Portal, using the Portal's analytical tools and applications, will allow data integration between the Business Area Warehouse, the Data Registries, and EIW.

Methods for integrating data from Business Area Warehouses are well understood. There is little new technology involved in the alignment of such warehouses to meet the needs of the EHPA. The greatest challenge comes in ensuring that the Business Area Warehouses meet the requirements listed above.

Alternative Integration Methods Based on Web Services

In addition to the Business Area Warehouse concept of TAA version 1.0, the EA Team and the PMO are currently validating additional alternatives for warehousing mission-critical data using Web services. Such alternatives can remove the Oracle RDBMS restriction from the Business Area Warehouses. There are two methods by which this can be achieved.

The first is to develop a data connection layer between the ODS and the staging process. Details will depend on the RDBMS in question. For example, if the ODS were developed using Microsoft SQL Server, the connection layer might be a set of functions developed using Object Linking & Embedding (OLE) Database Connectivity. OLE functions would migrate data from the SQL Server ODS into the Oracle Staging Database. The FBW would pull in the staged data as needed.

The second approach involves the development of a Web services library to act as a virtual database. Such a library would be developed jointly by the program offices and the PMO. Both organizations would be required to update this library as the ODS database is updated. Additional metadata requirements would be necessary for these Web services, as outlined in the Metadata Registries section. Queries requiring data in Web service enabled databases would be implemented as Oracle Java applets in the Oracle database. Each applet would make the appropriate Web service calls to the ODS Web services to retrieve necessary data.

While both approaches are technically viable, technical details, costs, and risks will not be fully understood until the proposed pilot study is complete in 2005. The pilot will determine (1) if Web

services can meet the data integration and performance objectives of the FBW and (2) which of the two approaches is superior.

Integrating External and Other Data Sources

The EIW and Business Areas Warehouses (whichever way they are constructed) will contain internal data sets critical to the Agency's mission. Data sets that are not critical to the mission, or that do not exist inside the Agency, would not link directly to the central Data Registries and would not be subject to the technical restrictions that apply to Business Area Warehouses. Such data sets can be broken into three categories: data from non-trading partners (e.g., Dunn & Bradstreet), data used by the Agency but not stored at EPA (e.g., at the Centers for Disease Control), and data from systems that are not appropriate for inclusion in the Warehouses (e.g., the air program's ozone monitoring system, a system that is updated in real time and which does not store records in a warehousing fashion). Data in any of these categories may be available to the FBW via a range of technologies, including Web services, flat files (spreadsheets, text files, etc.), XML documents, and small data systems files (Microsoft Access, FoxPro, etc.).

Specialized Data Marts

In addition to the FBW, which will be normalized through the Enterprise Data Model for maximum flexibility and storage efficiency, program offices will often require custom data models for specialized analysis or to increase reporting performance. They will use data marts to store data useful in analysis, as well as to optimize the schema for analytical processing for certain applications.

Data marts will pull data from the Data Registries, the EIW, and the Business Area Warehouses (via materialized views) and store the information within the FBW. The method of storage and the RDBMS used would be determined on a case-by-case basis to meet the unique needs of each data mart. Each data mart will have a unique set of requirements and specifications. Each will require full lifecycle development, including detailed analysis and design. Systems dependant on the current implementation of these data marts may need to be re-engineered.

Enterprise Data Modeling

Before actual databases and data storage mechanisms can be built, the EA must develop an enterprise data model to ensure that data is stored across all mission-critical data systems in a consistent manner.

The current data model is to store data based on the ODS from which the data originates. As such, each ODS contains data about the same or similar categories in the most convenient fashion for that ODS. A query to obtain all data related to a specific topic (such as permitting data) requires an understanding of each ODS data model.

An integrated data model will simplify the reporting process by reorganizing information based on its type. All warehouses (business area warehouses or the EIW) will use a standard structure for storing each category of data (permits, compliance, etc.). Queries relating to a specific business entity will be written to a single database schema. Such a data model is already employed by the Envirofacts system to provide data integration. This data model can be used as a basis for the remodeling of all enterprise data.

Data Sensitivity Classification

One of the more complex tasks to be addressed by the FBW is data sensitivity. Under the current model, each ODS maintains its own data security by limiting access to data to the users of the

specific ODS. As the FBW consolidates the different ODS data into a unified data source, sensitive data will be stored in the same databases as public access data.

To ensure that data is available only to the appropriate users, the EA will need to create an active data sensitivity model to apply data sensitivity classifications to every data element in the FBW. Because data organization will be based on data types, users will be granted access to only those types of data to which their security classification entitles them. Data query filters and views can be set up to further limit the data to rows and columns available to particular users. OEI will work with programs to define and implement policies and guidance related to data sensitivity classification, as well as authentication, authorization, and validation. These policies and guidance will ensure consistency in approach across EPA applications and interoperability with the federal E-Authentication gateway initiative

The data sensitivity classification system must also integrate with the Portal and Identity Management/Access Control systems. Users will authenticate with the Identity Management system via the Portal and be assigned specific access rights. The FBW must also recognize these access rights to properly limit users' access.

System of Registries

At the core of the FBW is a system of Data Registries. Data contained in the registries is limited to information shared by multiple programs. Registries will contain the fields, keys, and other information necessary for data integration across the ODS systems. Although as many as 23 Data Registries were identified as possibilities in the target EHPA, the following subset are the only ones considered necessary at this time to implement the EHPA and ASA architectures:

- Facility Registry
- Substance Registry
- Environmental Indicators Registry
- Geospatial Registry
- Organizations Registry

As other sets of common data are found necessary, they may be included in the general set of Data Registries.

Facility Registry System

The Facility Registry System (FRS) contains information on the different facilities that EPA has an interest in tracking. It is already implemented as an Oracle database instance. The process of inserting, updating, querying, and integrating data with the FRS is standardized by FRS within FRS tools and interfaces.

To better control FRS data connectivity, the PMO must define and develop a migration path that centralizes the input, update, and querying of registry data. A single data interface will contain all business and technical functions necessary to connect the ODS systems and the FBW to the Data Registries. To preserve functionality, interface wrappers must be developed to encapsulate the functions of the new single interface with the old connectivity methods currently used by the ODS systems. As the individual ODS systems are upgraded, they will be migrated to the new single interface, and the wrapper will be discarded.

In addition to facility information, the FRS system also contains information about organizations external to the EPA. While this information is vital to the needs of the EPA, FRS may not be the best place to house organizational information. No decision has been made as to whether the organizational and facility information will be separated as the enterprise data model takes form. The current organization data in the FRS system could be reorganized into a separate table space as a first step in the creation of the Organizations Registry.

Substance Registry System

The Substance Registry System (SRS) contains several categories of information useful to the ODS systems, including information on chemicals, biological agents, coded waste streams, environmental contaminants, and other substances that affect the environment. While the SRS contains a wide range of data, with over 85,000 substances registered, it does not yet reflect all the substances tracked by the Agency, nor is it complete with respect to the ODS systems to be migrated into the EHPA. A full systems analysis, conducted in conjunction with the development of the FBW, followed by redevelopment work to bring this registry in line with the FBW, will ensure creation of a Data Registry that meets the needs of the architecture. This analysis should also focus on unifying any existing interfaces with other systems into a single data interface, similar to the approach detailed for the FRS.

Other Registries

Of the remaining listed registries, few are actual production systems. In addition, there are most likely other common data elements that would best be organized into their own registries. The process of creating the enterprise data model will better define the final number, type, and scope of data registries.

In addition to the registries listed above, there are several registries originally proposed in other architectures (such as the ASA) should be considered for general use throughout the EA as the ASA and EHPA domains integrate. These include the Individuals Registry, the Grants Registry, and possibly others. When implementing these registries, care will be taken to ensure that they could be reused within the multiple architectural systems under development. The development of the enterprise data model will highlight the proper organization and use of these shared data assets across architectural systems.

Metadata Registries

The most challenging aspect of data integration at EPA is not electronic data exchange between different data sources. This problem has been largely solved over the past decade with the introduction of data connection standards such as Open Database Connectivity (ODBC), Object Linking & Embedding (OLE) Database, Java Database Connectivity (JDBC), and Simple Object Access Protocol (SOAP).

The real challenge for data integration lies in ensuring that data in one database conforms to the same standards as data in others. The Agency estimates that its databases include some 15,000 data elements, of which only a handful (about 400) have been reconciled with each other in the Envirofacts application. The process of standardizing data across Agency programs requires strict enforcement of data standards, for which the Agency has an active and ongoing program. The task of reconciling the full range of data elements will take years, but can be prioritized to address the needs of the first set of mission-critical applications that will migrate to the FBW. To help manage and enforce these standards, EIIPD has developed a series of Metadata Repositories. Database developers, administrators, and reporting system programmers will be required to use the information in these repositories to ensure that their code correctly integrates data.

The EHPA calls for the development of several Metadata Repositories. A few are in development; others are still on the drawing board. The following sections review each metadata repository, its status, and the work remaining.

Data Element Registry

The Data Element Registry (DER) is one of the most critical metadata registries. It will provide a location to record information concerning every field and group of fields in every database incorporated in the FBW, regardless of the method of integration. EPA already has one registry that, with modifications, can serve this function—the Environmental Data Registry (EDR).

The EDR is central to data integration. It will provide a data dictionary for database administrators, developers, and other interested parties to use in implementing new applications and systems. This will ensure that databases are created and upgraded in a consistent fashion. Developers will be able to write integrated reports knowing that they are properly integrating data from multiple systems.

The EDR will require extensive enhancements before it can meet the objectives of the enterprise architecture. Currently, EDR simply lists every field in every database incorporated into it. There has been no effort to reconcile the fields. Thus, if two systems have a primary key field Facility_ID, rather than having a single entry with both systems listed as implementing the field, there are two entries. Such inconsistencies in the metadata for the same field will defeat proper data integration. The PMO must undertake to manage, fund, and strictly govern a multi-year process to reconcile the data elements being merged in the FBW. The DER will properly support data integration only after this work is complete.

The second issue concerning the EDR is that data is not currently organized in a logical format that is easy to reference. Estimates are that there are over 15,000 distinct data elements at EPA. For developers and database administrators to maintain consistency, they must be able to know if the data they are looking for already exists and is defined, or if a new definition is necessary. The architecture calls for the EDR to be organized according to the enterprise data model. This model, once implemented, will allow quick lookup and reference to the contents of the DER.

Applications Registry

Equally important to the EDR is the Registry for EPA Applications and Databases (READ), which will be the Agency's inventory of applications, databases, and other information products (e.g., models). READ meets the Agency's commitment to develop an applications inventory, as called for by OMB Circular A130.

For each application and other information products, there will be a unique record in READ to hold the metadata. Metadata about the applications and databases will include the location of the system, the specific office within EPA that is responsible, and the regulatory basis for the system. Technology metadata, such as programming language and platform also will be included.

Importantly, each READ record will contain links to the data elements in EDR, the substances in SRS, related terms in TRS, and XML tags in the XML Registry, and the actual data in Envirofacts, if applicable.

XML Registry

Extensible Markup Language (XML) is the de facto standard language for transferring complex data over the Internet. The purpose of an XML Registry is to document all the XML "tags" used by EPA partners to transfer their data over the Exchange Network and their internal transfer

within the EIIPD. It will contain complete XML definitions for all data elements and schemas needed by the shared systems (CDX, Web Services, etc.).

Need for an XML Registry is currently satisfied through a partnership between EPA's EDR and Exchange Network's Web site, hosted by the Environmental Council of States (ECOS), which provides the definitions necessary for trading partners to exchange data on the Network. A fully functional XML Registry is planned by EPA and its state partners when issues related to technology and the location of the Registry are worked out.

This registry will most likely be implemented as a library of XML Schema Definition (XSD) documents. These documents will be cross-linked with the EDR so that when developers find a data element of interest, they can locate the XSD document(s) that uses the data element. This is important if developers are designing XML based systems (such as Web services) or if they are modifying the definitions of a data element. The entire XML definition must be kept separate from the EDR because the XSD will typically contain information about the context in which the data is used (the SOAP Message, XML File, etc.), which is neither necessary nor helpful to most users of the DER.

Data Model Registry

Knowledge of how different data elements link to each other within databases is essential when making queries of the FBW. The Data Model Registry (DMR) will contain this information for each database integrated into the FBW, regardless of its integration method. It will complement the XML Registry on the database side, providing the context in which the data elements are organized in their databases.

At the logical level, the DMR will be a collection of Entity Relationship Diagrams (ERDs) that document each database in the FBW. A developer can make use of the DMR when developing queries to determine proper joins and unions and to enforce proper data integration techniques across different databases.

Although the DMR is a vital component to the development of cross-media queries, little work has been accomplished on this system. A suggestion has been made that the DMR be developed using a library of Oracle Designer files, but a complete systems analysis and design is still needed to identify the proper solution. The analysis will include Oracle Designer as a possible solution, as well as other options.

Terminology Registry System

The Terminology Registry System is a Metadata Registry currently being developed as the Terminology Reference System (TRS) and will be used to ensure correct usage and understanding of scientific and environmental terms commonly employed within EPA. This is important for both data and metadata. For data, common terminology and references are necessary to ensure that proper data integration rules are followed and to catalog Web pages, documents, publications, and data so information can be drawn from EPA's data resources. If two data values are integrated based on a misunderstanding of their values, inaccurate results could lead to erroneous decisions, policies, and actions by the Agency. On the metadata side, common terminology is useful to ensure that the meaning of the data is clearly understood and is uniform throughout all data systems using the data element, schema, data set, and so forth.

Data Set Registry

A user looking for information within the EIIPD system should be able to discover what information (data sets) the Agency actually holds. At present there are hundreds, perhaps thousands, of undocumented data sets scattered throughout the program offices, regional offices, and laboratories. Documenting all EPA data sets is the job of the Data Set Registry, which will be the Portal's "roadmap" to Agency data holdings.

The Data Set Registry will provide listings of all the data sources at EPA, or externally used by EPA, with architecture-specific information for those systems that are integrated through the EIIPD. It will provide developers and database administrators with a high-level view of the data available to the FBW. The ability to link data set entries to their corresponding data models in the DMR and, data elements in the DER will be used to provide a "big picture" view of the architecture's data holdings through the Portal.

Currently, the Environmental Information Management System (EIMS) functions as a likely starting point for the Data Set Registry. EIMS provides information on some of the data sets owned by the Agency, but the percentage of data sets cataloged in EIMS is believed to be small. The PMO must complete the EIMS registry or its successor to ensure that all metadata resources are complete and accurate for data integration to proceed.

Portal

The Portal will serve as a gateway for users to provide and access relevant environmental information. It will allow for role-based access with identity management and security solutions, while providing focused information to stakeholders to enable more informed environmental decisions.

Portal Status

Development of a single point of entry to access all components of the central services through a user-friendly interface is critical. The Portal's mission is to provide this single point of entry and to support a seamless process for submitting and retrieving data to and from the FBW.

Phase I of the Portal will focus on existing CDX incoming data flows on the input perspective side, and on the Analytical Tools/Decision Support Services on the output side.

The scope of the Phase I Portal effort will include a selection of initial users—CDX partners, internal EPA users, situation analysis users, and the PMO portfolio managers. Conceptual wireframes have been developed to explore potential users needs. The Portal has been broken into 4 layers of users and usage:

- **Builder Layer**: Serves the EIIPD portfolio managers who will use the Portal to develop a community and facilitate communication between various groups;
- **Participant Layer**: Initially for CDX partners (states, tribes, trusted partners) to get more information/involvement about OEI/CDX and its various information sharing opportunities;
- **Administrator Layer**: Provides internal EPA users with a point of access to what is coming in through CDX;
- **Analytical Tools Users Layer:** Aggregates data from various sources for enhanced decision making capability.

Once data is brought into the FBW, a common method is necessary to control access to the data. Web services provide an almost universal method for accessing data. Most internal systems at EPA can be developed or upgraded to direct their data calls to Web services. These Web services can contain the functions necessary to ensure security, transactional support, and data transfer. The result would be a standard data interface for all applications in the EHPA.

This option should not be confused with the use of Web services to implement alternate integration methods as described in the FBW section. The common integrated data access point simply provides a unified interface for returning data from the FBW. It does not depend on the implementation selected for the FBW.

Interim Technology

Portals can serve multiple purposes to various audiences. EPA's initial Portal implementation will serve the Portfolio Managers, who will use it to support the development of the EIIPD central services—sharing documents and schedules, and updating project plans and milestones. With the multitude of different groups (business units/departments) contributing to the success of this effort, communication across teams and managers will be a critical success factor.

Long-Term Operation

The long-term vision for the Portal includes expanding the audience of users and extending the Portal's functionality. As components of OEI's central services are developed, the Portal will become more robust, leveraging existing applications through the System of Access and building on the identity management and access control efforts. The Portal will provide the ability for Administration to view and manage data coming from external sources and provide agency wide reporting based on this information.

The Portal will provide a single point of access to information currently spread across the Agency. Access to timely, relevant information from a multiple sources through a single screen will enable emergency response and homeland security users to make more informed environmental decisions based on data from potentially wide-spread sources.

Data Security

Agency data resources represent the combined knowledge and information of dozens of ODS systems. This data runs the gamut from publicly available information to data vital to national and homeland security. For this reason, robust security mechanisms are necessary to ensure users access only data they have privileges to obtain.

The functions of an electronic security system are split into two categories. Identity Management provides the mechanisms for the system to know the identity of the current user. Access Control provides methods for a specific user to be granted the rights to perform specific types of functionality and/or access to specific categories of data.

Perhaps the most complex task of the security system is determining how complex a security model is needed. Each ODS currently holds its own security information. These classifications can be used as a basis for the security classification schema. Next, each classification is evaluated against the others to eliminate duplicate classifications (i.e., two classifications from two different ODS systems allow the user the same access in the EHPA). Lastly, a survey of the classifications is completed in conjunction with the EHPA to assure that any new classifications are properly represented.

There are two ways to build classification schemas. The easier is to create a fixed schema in which the classifications are hard-coded into the solution. This is the most straightforward approach, but it is extremely difficult to expand. As such, it is usually only used when the set of classifications is fixed and unlikely to change. The more likely scenario is for EHPA to expand, requiring new roles and sets of access rights. Users can be directly assigned access rights, assigned to groups that are assigned to access rights, or both. This model is significantly more complex to implement and requires greater maintenance, but provides for an extensible solution.

As with the other pieces of security management, the security classification needs of the Agency need intensive evaluation. An in-depth analysis of the security status of each data element must be conducted before a solution is implemented. This can build on the extensive work done by OEI's Office of Information Collection (OIC) on data classifications, and on the work done within the State/EPA Security Challenges Forum in defining security classifications. The Exchange Network is also defining security classifications of Exchange Network exchanges.

Identity and Access Management (IAM)

Enterprise-level identity and access management is essential to the architecture. It used by all other components of the EIIPD project, and, ultimately, by all EPA applications. IAM is used to:

- **Reduce security exposures.** Better management of users and their roles and authorities, including elimination of orphan accounts multiple accounts for single users, systematically reduces security exposures.
- **Improve the end-user's experience.** Users register once and use the same identity and account to access multiple systems. IAM supports single sign-on for applications of like authentication strength requirements. It also supports self-service registration, password changes, and resets.
- **Simplify and improve user administrations.** IAM eliminates redundant processes. It delegates user registration and administration to application owners and stakeholders and aggregates users with similar privileges into common groups and roles.
- **Provide a common framework for application development.** IAM relieves developers of the identity management burden, significantly reducing costs and speeding up application deployment.

The EIIPD Identity Management portfolio will undertake a number of tasks to implement enterprise IAM. The final system will be compatible with government-wide E-Gov E-Authentication principles and infrastructure.

Establish a Single Directory of External Users

The Agency's current directory of internal users must be supplemented by a similar directory of external users. It will be integrated with, but separate from, the current system. Among the information to be stored in this directory is each external user's role and group as these pertain to the applications they access.

Directory data includes demographic information (name, work location, organization, etc.), credential verification information (passwords, security device serial number, fingerprint maps, etc.), and a listing of the functional and data rights that the user has been granted. The most common commercial-off-the-shelf (COTS) database solution is Lightweight Directory Access Protocol (LDAP) database. Nearly all portal and Internet system packages will support this protocol. Alternatives include integrated operating system and/or network security (i.e., Windows

Domain Logins, Novell Logins) and custom user databases implemented from scratch for the particular need.

Implement a Web Access Management Solution

EIIPD will acquire and implement a Web access management solution that centrally manages authentication, access control (role-based, if practicable), and single sign-on across multiple Web applications.

To ensure a strong identity management system, EIIPD will conduct a detailed analysis of EPA's needs. The analysis will include:

- The level of security needed. Multiple types of credentials can be implemented (username and password for non-sensitive data, security devices for classified information, etc.).
- The range of vendors that can provide different solutions.
- The standards to be followed for user databases.
- Whether credentials or database systems need additional physical security. For example, anyone with access to the server generating digital certificates can create their own certificate. This server is usually kept in a secured location with surveillance, guards, etc.
- Governance policies for the system (i.e., how long a password must be, how to obtain. manage. and cancel digital certificates, etc.).
- How the system can be expanded for future increases in security.
- Usability of the system verses the level of security it provides.
- Integration of ID management solution to the E-Authentication gateway, where appropriate.
- Institute Agency policies and guidance to ensure alignment with federal and industry approaches to identity management.

Once users have been identified by the system, they are associated with a list of rights. These rights are interrogated by every system the user interacts with to determine the appropriate controls to enact. A typical scenario may include the following:

- The user logs in and is identified by the Identity Management System.
- The Identity Management System provides the user with a set of access control tokens for each system.
- The Portal interrogates the portal access control token and directs the user to the appropriate page.
- The Metadata interrogates the metadata access control token to determine what applications and tools the user may access.
- Each subportal, application, and tool interrogates the application access control token to determine how to configure the user interface (what data and functions to show, what queries to run, etc.).
- The FBW interrogates the data access control token to determine what fields in the query to return.
- The results appropriate to the user are displayed.

Support Multiple Authentication Techniques

EIIPD will support a variety of authentication techniques to accommodate the full range of data sensitivities that exists among EPA applications and systems.

Three authentication solutions are in widespread use. Because these options are technology-specific, the Technology Architecture and the PMO will work together to define the preferred solution for the EIIPD.

The most common is the Username/Password system. Users are assigned unique usernames and passwords. The theory is that only the users know their passwords. This model is simple to implement, but its usefulness hinges on users keeping their passwords secret and selecting passwords that are difficult to guess.

A more secure solution is a Digital Certificate. Users are issued unique encrypted files containing their information. The user supplies this file to the system, which decrypts it and reads the user information inside. Digital Certificates work well provided that users closely guard the encrypted file.

Another alternative is the User Token, in which a user is provided some physical device (a smart card, a number generator, etc.). The user supplies a username and either the device itself, or the information in the device, to prove their identity. Once again, these systems work so long as the users protect their devices. The simplest to use but the hardest to implement are biometric systems. These systems often require additional hardware at each computer (a fingerprint reader, camera, retinal scanner, etc.), but provide the greatest security because it is nearly impossible to impersonate a biological identity.

EIIPD will also coordinate with GSA as they move toward a federal government E-Authentication solution. To this end, the EIIPD project will analyze requirements and will incorporate the E-Authentication gateway infrastructure as appropriate. In addition, the EIIPD project will engage with EPA, state, and other partners to establish and maintain a governance structure for E-Authentication and to develop a strategy and implementation plan.

Implement Workflow-based Tools

EIIPD will acquire and implement a set of workflow-based tools, integrated with the directory and Web access management solution, for development and execution of user registration and administrative processes.

Modify Applications and Processes

Individual applications and processes will have to be developed or modified to integrate with this new system. EIIPD will provide guidance as necessary for this purpose.

Geospatial Services

EPA's geospatial architecture was one of the predecessor programs of the present Enterprise Architecture. In the Agency's December 2002 submission to OMB, it was a separate, cross-cutting Component Architecture, developed through the Geospatial Blueprint. With the establishment of the Program Management Office, geospatial services become one of the six EIIPD portfolios—a fully integrated part of EPA's target architecture.

The business, data, applications, and technology layers of the geospatial services architecture are now merged with EA. Because of their cross-Agency responsibilities under the Geospatial One-

stop Initiative, however, geospatial services retain special governance functions beyond those of the general architecture. To lead this governance process, EPA plans to appoint a Geospatial Information Officer (GIO).

EPA's Geospatial Blueprint (completed June 2002) now becomes the sequencing plan for integrating geospatial services with the EIIPD. The envisioned program will support an internal and external network of shared, distributed geospatial data repositories. It will provide common application services, based on EPA business needs, which conform to mutually-accepted open standards. Data and applications that can be delivered over the Intranet and Internet via "geoservices" will be key. The program will allow geospatial data and applications to be used in both geospatial and non-geospatial applications, improving overall work processes throughout the Agency.

EPA's geospatial requirements are highly oriented to the needs of the regions, which generate and use the majority of the Agency's geospatial data. This has led to the development of an EPA Geospatial Network that comprises geospatial data, applications, and technologies. The network will significantly improve and simplify the exchange and sharing of GIS databases and applications among regional and program offices, research laboratories, and the National Computing Center (NCC). Figure C-5 below illustrates how the geospatial network and the internal geospatial architecture are integrated with the EIIPD elements.

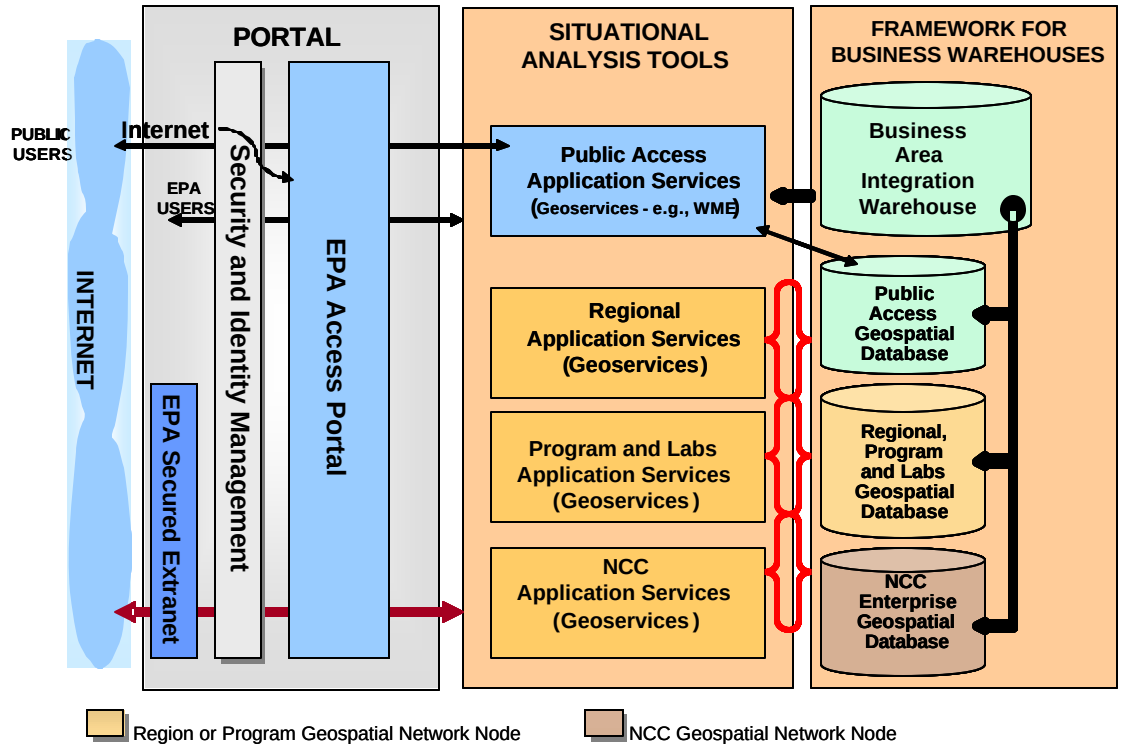


Figure C-5: Internal Geospatial Architecture as Integrated with the EIIPD

Transitioning to the integrated vision of the geospatial services program requires actions at all levels of the architecture: governance, business processes, data, applications, and technology. Chief desired outcomes at each of these levels are as follows:

Governance:	Create a clearly articulated leadership and governance structure that sets geospatial investment priorities; sets guidelines and standards based on interoperability; tracks technical trends; and pursues data investments in partnership with other federal, state, and private entities.
Business Processes:	Enhance business operations by integrating geospatial data, information, applications, and technology to improve decision-making processes.
Data Architecture:	Maintain enterprise-wide indices, registries, and catalogs of data and metadata; create data repositories jointly maintained by headquarters and the regional offices to eliminate duplicate data purchases and unnecessary storage requirements; reflect geospatial data needs in annual budgets and in partnership agreements between the Agency and states or other federal agencies.
Applications Architecture:	Maintain interoperable software based on open standards; make tools available through Web browsers with wireless mobile applications available in the field.
Technology Architecture:	Enable EPA staff to access, process, and manage geospatial data at any time, including sharing data with EPA partners; enable managers to add, delete, and/or replace components of the architecture as needed, based on open standards; allow field sites to provide GIS and other geospatial products via the Web.

Operational Data Stores

The PMO's Workplan calls for one major environmental application to be operational within the EHPA central services on a pilot basis in 2005. Once this application has been chosen, it will become the final portfolio in the PMO management structure. Its manager will come from a program office, not OEI.

As major applications are restructured to use CDX, the FBW, the Portal, Analytical Tools, and the System of Registries, the one function that always remains with program offices is the data processing function, referred to as the Operational Data Store (ODS). Programs are responsible for ensuring that the data they receive through CDX meets data quality standards and is properly processed, as necessary, to serve program functions. Once the data is quality assured and processed, it is stored in the FBW.

Local data storage in the ODS is for review and processing functions only.[9] The degree of processing and quality review done by each ODS will vary. Some will entail extensive data management. In other cases the ODS function may be little more than a direct pass-through from CDX to the FBW.

[9] ASA ODSs may be an exception to this norm. The highly transactional nature of their processing may argue that some day-to-day reporting is more efficiently done at the ODS level, with the Administrative Warehouse storing more summary data and records.

The ODS portfolio will be the test bed to identify issues that will affect systems migrating to the central services after the pilot. No single system can be expected to surface all possible problems, but the ODS portfolio will certainly identify the most important ones. It will deal with at least the following concerns:

- Procedures and standards for receiving data from CDX via Web services.
- Procedures for linking to the Data Registries. These will depend on which FBW storage method the pilot employs: the FBW or a Business Area Warehouse, and if the latter, whether it uses direct database joins or Web services to link to the Data Registries.
- Coordination with CDX to shift the update of Data Registries from the ODS to a common CDX interface.
- Procedures for linking to the Metadata Registries.
- Migration of reporting tools and interfaces to the Portal and to Analytical Tools.
- Integration of Identity and Access Management with the ODS.
- Identification of data elements used by the ODS that may be common to other systems that will migrate to the EHPA later. Such data elements might eventually require their own separate Data Registry.

The ODS portfolio will generate the information needed to provide guidance, technical support to other program offices as they migrate their applications to the target. Its experience will also provide the basis for possible technical modifications to the other EIIPD system components to ensure that the pilot system becomes operational on schedule, with its data flowing through all the EIIPD components in 2005.

Schedule

Below are the high-level milestones for the first year of the EIIPD project. Completion of these critical tasks will lead toward the goal of prototyping one program office system data flow through each of the EIIPD components.

FY 2004, First Quarter

1) Enter into a formal agreement with a program office development partner.
2) Initiate requirements tasks.
3) Convene forum of interested program offices to discuss technical infrastructure issues.
4) Begin collaboration with the soon to be identified OEI data manager on the development of an enterprise integrated data model.
5) Develop Identity Management timing strategy to coordinate with OMB and GSA E-Authentication initiative.

FY 2004, Second Quarter

1) Complete reassessment of requirements for all EIIPD components.
2) Complete a draft of the enterprise integrated data model.
3) Begin presenting the portal operational capabilities demonstration (OCD) to partners for review, comment, and confirmation of scope.
4) Initiate CPIC cost-benefit analysis.
5) Begin EIIPD mid-year assessment and corrections process.

FY 2004, Third Quarter

1) Draft requirements for ODS connections to EIIPD components.
2) Publish interim technology targets.
3) Publish interim guidance on how business areas should transition to the EIIPD components.
4) Begin third party Independent Validation & Verification (IV&V) assessment of EIIPD components.

FY 2004, Fourth Quarter

1) Complete all registries required for metadata management so they are ready for populating.
2) Publish final policy and guidance on how business areas should transition to the EIIPD components.
3) Complete all EIIPD component pilot development and publish IV&V results.
4) Submit integrated CPIC for FY 2006 to OMB.
5) Implement data management processes by OEI data manager.
    a) Registration
    b) Normalization

c) Sensitivity determination

6) Begin EIIPD mid-project assessment and corrections.

Part 3
Administrative Systems Architecture Sequencing Plan Update

The ASA sequencing plan is based on the ASA initiatives (i.e., projects required to migrate from the baseline state to the target state) and is grounded in the ASA Applications Architecture. The sequencing plan was developed by the inter-departmental representatives that comprise the ASA Work Group. Other significant inputs to the development of the ASA sequencing plan included detailed review and analysis of the current CPIC Form 300 for those initiatives/projects that form the core of the effort required to accomplish the target ASA.

The goal of the ASA sequencing plan is to identify and plot the major tasks necessary to move from the Baseline to the Target Architecture along a time line. The sequencing plan provides managers with a comprehensive view of the interdependencies between process improvement and application implementation efforts, and eliminates, as much as possible, potential rework. The implications of project dependencies as described in the sequencing plan drove decisions to either accelerate, or postpone, planned investment in order to mitigate impact on operations and other concurrent initiatives.

The ASA sequencing plan proposes a project schedule based on industry best practices and the Agency's feedback regarding its past experience and expectations. The following paragraphs provide summaries of the functional and technical changes, an overview of the plan to transition from the Baseline Architecture to the Target Architecture based on the ASA initiatives/projects, and a description of the project schedule and task dependencies.

Summary of Changes

Following is a summary of the functional and technical changes necessary for EPA to transition the ASA from its current state to the target state.

Functional Changes

Dozens of the activities supporting the four main business functions that comprise the ASA Business Architecture are expected to change as a result of implementing this target architecture. The target ASA initiatives will provide EPA an opportunity to take a closer look at its business activities as they relate to best business practices supported by the various new and enhanced systems. This sequencing plan proposes business process re-engineering (BPR) efforts to delineate supporting business activities, identify process gaps against potential solutions, and recommend changes to processes/activities as appropriate for each ASA initiative. In addition, functional issues affecting the performance of EPA's administrative functions in the baseline architecture will be addressed as part of the BPR efforts required to achieve the target ASA. Examples of these functional issues include:

- Implementation of managerial cost accounting
- Defining and tracking performance measures

- Regional involvement and communication in the EPA Formulate Budget process
- Managing the Working Capital Fund
- Implementing the matching principle (revenue to expenses) for Superfund Cost Recovery
- Improving the timeliness of financial statement closing process
- Complying with mandatory federal requirements for simplified acquisition procedures
- Streamlining the requisition approval, request for proposal, and purchase order closing processes
- Improving the timeliness of contract evaluation and award
- Elimination of redundant maintenance of financial and human resources data
- Re-engineering of business processes associated with integration with E-Government initiatives
- Implementing the tracking of personnel skills and competencies

These functional issues are examples of the subjects of BPR tasking under the initiatives that are sequenced in this plan to achieve the envisioned Target ASA.

Technical Changes

Changes between the current and target architectures are most evident at the application architecture layer. At the application layer, the 53 current administrative applications are migrated to 37 target administrative applications through replacement, integration, and re-engineering.

In the area of software engineering products, one specific technology plays a major role in the target application and technology layers.

Enterprise Application Integration (EAI) technologies will provide for the development of a common strategy for the integration of ASA applications. EAI will address standardization of interface protocols, use of common interface functional capabilities, and management of communications with internal or external systems. In addition, EAI will provide EPA with the flexibility to support business process changes without the overt need to modify an application's source code (e.g., inter-application workflow or business rule validation through EAI), thus effectively extending the life-cycle of its applications. EPA currently uses this technology for a single interface to IFMS. EPA plans to deploy this technology at the enterprise level to take full advantage of its capabilities.

Target Architecture Planning

The ASA sequencing plan addresses not only how each target application will be acquired, implemented, and rolled out, but also how EAI technologies will be leveraged to provide near-term gains. To provide for an effective transition of EPA's current administrative systems architecture to the target administrative systems architecture, the following nine major initiatives/projects, some of which represent portfolios of systems, are included as part of the overall ASA sequencing plan:

1. PeoplePlus HRMS – HR Management System and Payroll Replacement System
2. EAI – Application Integration
3. FINRS – Financial COTS

4. FDW – Financial Data Warehouse
5. ADW – Administrative Data Warehouse
6. New Acquisition System
7. Relevant E-Government Initiatives
8. ICMS Enhancements
9. IGMS Enhancements
10. Administrative Data Initiative
11. Administrative Portal

These projects support the overall plan to improve EPA's administrative processes and architecture while maintaining the necessary functionality to continue the Agency's mission throughout the transition. The following subsections first describe the scope of the 12 projects. In a later subsection, the proposed schedule and rationale for the time line is discussed. Numerous examples of scheduling and planning integration, and the benefits of integrated planning, are provided.

The revised ASA sequencing plan depicts the general time frames for the major tasks, and is useful as it highlights the many concurrent tasks planned over the next several years, and serves to focus the attention of integrated planning on the challenges and opportunities presented by these concurrent initiatives.

Integrated planning, the determination of task dependencies and appropriate task sequencing, however, requires significantly more detail. The details of the ASA sequencing plan reflect the efforts of an integrated planning approach. This plan depicts the major time frames for the following high-level tasking as appropriate to the specific project:

- Acquisition and Pre-implementation – includes tasking for BPR, acquisition planning, requirements analysis and other pre-implementation analysis activities.
- Implementation – includes tasking for design, programming or COTS configuration, data migration, interface development, testing, and system rollout.
- Training – includes tasking for development of training materials user training.

While system acquisition and pre-implementation analysis activities are critical to the successful deployment of new systems, and are shown in the ASA sequencing plan, there are no dependencies that exist between the major initiatives during this phase of the life-cycle of the various projects. Apart from ensuring that analysis artifacts (both data and functional) are standard, reusable, and shared so that redundant analysis is avoided, these activities have little impact on the sequencing of detailed tasks between initiatives and are therefore not further decomposed.

More significant to the level of detail captured in this target architecture is the implementation phase of each of the projects described in the sequencing plan. The baseline ASA applications architecture shows over 63 custom interfaces between 53 separate applications, revealing the high dependency of current applications on data captured by related applications. While the target architecture substantially simplifies and rationalizes data interchange requirements by

consolidating sources of data and leveraging the technology represented by EAI tools, no new system can be deemed implemented until all downstream data dependent systems are provided with required data. The documentation of target application interfaces is essential to integrated planning as it provides the basis for interface development planning across the administrative segment of EA. Required interfaces are therefore included under the implementation phase for each target initiative, and are based on the information captured to document application interfaces.

Current schedules for several planned E-Government initiatives are also included in the sequencing plan as several major interfaces to these E-Government initiatives are contained in the target ASA. These ASA interface development tasks are dependent on the timely completion of these E-Government initiatives.

Sequencing Plan Dependencies and Planning Integration Highlights

Schedules for relevant E-Government initiatives are included as interface development efforts for ICMS is mandated prior to the time when the new Acquisition System will be available. Additionally, interface development to support the various E-Government initiatives is a major component of target IGMS enhancements.

The sequencing of the target ASA is as follows:

- Early implementation of EAI supports the rollout of the payroll replacement system currently underway. The schedule for EAI interface development focuses early efforts on developing interfaces for several of the major financial and acquisition systems with which the payroll system must interface. These efforts will be leveraged to ensure that unnecessary rework is avoided as payroll system implementation continues apace.
- Implementation of the Enterprise Application initiative as soon as possible allows EPA to realize EAI benefits (e.g., standardization of interface protocols, use of common interface functional capabilities, and management of communications with internal or external systems) early in the transition process. While the focus of EAI implementation in the target ASA is on financial systems, this addresses well over half (60 percent) of the total number interfaces documented in baseline ASA.
- Early implementation of EAI reduces the complexity and risks associated with the implementation of the FINRS on several accounts:
  - The FINRS implementation team is no longer concerned with establishing or determining how interfaces to and from the system will be accomplished (i.e., separation of concerns).
  - The FINRS team needs only to develop a single interface to the enterprise "hub" through a standard method and process. It does not need to create multiple interfaces to multiple systems using methods and processes that could prove to be incompatible with each other.
  - Should the implementation of the FINRS take longer than expected, EPA would have already achieved a higher level of integration and ease of interface maintainability among its legacy systems, thus effectively extending the life cycle of its legacy systems.
- Early implementation of FDW enhancements addresses current and short-term financial reporting needs.

- Later implementation of the Cost Recovery and Imaging project allows EPA ample time to analyze those EPA unique and critical cost recovery requirements not supported by the FINRS. Results of this analysis will result in streamlined cost recovery processes jointly supported by the FINRS and SCORPIOS applications.
- Extended pre-implementation time frames for the FINRS provide a greater degree of certainty that the selected JFMIP certified financial package will more closely meet EPA's requirements.
- Extended implementation time frames for the ADW provide a greater degree of certainty that ADW components will not require re-engineering, and that reporting requirement are sufficiently and exhaustively captured. ADW components are implemented only after several critical source applications (e.g., HR-PTL, FINRS, SCORPIOS, Acquisition System) are fully operational.
- The Payroll application will be the first to provide data to the ADW as its implementation will be completed first, followed by Planning, IGMS, Acquisition, FINRS, and SCORPIOS in that order.
- Implementation of Planning is scheduled to provide sufficient time to have the application in place and fully implemented for use during the budget formulation process.
- IGMS (Grants) integration with IFMS (Core Financial) is currently planned for FY 2003. As EAI implementation plans currently call for development of interfaces to both IGMS and IFMS, these efforts will be leveraged to effect the integration between IGMS and IFMS.
- IGMS (Grants) integration with PeoplePlus (Human Resources) is currently planned for early FY 2004. As EAI implementation plans currently call for development of interfaces to both IGMS and PeoplePlus, these efforts will be leveraged to effect the integration between IGMS and PeoplePlus.

This year, we plan to update the ASA sequencing plan. Changes based on funding priorities, evolving schedules for E-Government initiatives, refinements to current initiatives, and the potential for new initiatives drive updates to the ASA sequencing plan. Each of these items is being researched, and the results of this analysis will be reflected in an update to the ASA sequencing plan that is scheduled to be available by the end of the first quarter of FY 2004.

ASA Initiative Descriptions

Following are more detailed descriptions of the various initiatives included in the update to the ASA sequencing plan, including planned milestone dates. They highlight, as appropriate, the major changes and updates that have resulted from changes in funding priorities, a more thorough review of project dependencies and available resources, changes to the names and estimated completion dates for relevant E-Government initiatives, and refined inter-departmental planning.

FINRS

The Financial Replacement System (FINRS) project has evolved to meet the OCFO's Information Resources Management (IRM) Strategic Vision and includes all functions of the financial business architecture for EPA. The components of this architecture, while in varying stages of their application's life cycle, collectively constitute a single Agency-wide source of official financial information efficiently accessible by authorized customers.

Context and Vision: Comprehensiveness – FINRS will provide an enterprise level financial systems solution for EPA. The target is an integrated suite of applications that may well provide

the modernized systems support required by other government agencies. EPA will migrate to the Department of Agriculture's National Finance Center (NFC) for payroll processing starting in FY 2005 as part of the E-Payroll initiative. In the same vein, EPA will continue to pursue partnerships with other agencies to collaboratively develop a readily adaptable system.

Mission Requirements, Business Practices and Technology – As FINRS progresses there will be ongoing analysis of how to best fulfill EPA's baseline requirements as well as those of potential partners in this effort. Because FINRS is an enterprise solution, COTS ERP systems will continue to be considered along with other COTS based alternatives.

EPA structured FINRS to deliver both short-term and long-term results. For example, in FY 2004 EPA will realize substantial on-going cost savings by implementing the Payroll Time & Labor system. In addition, in early FY 2004 EPA will begin to realize benefits from standardizing interfaces through use of an Enterprise Application Integration tool and improve ad-hoc financial reporting by re-engineering the Financial Data Warehouse and implementing Business Objects, a modern web-based reporting tool. In FY 2004 EPA will more tightly integrate budget and planning by merging the Performance and Environmental Results System into the Budget Automation System. In FY 2004 and FY 2005 EPA will reengineer our existing data warehouse to include enhanced business activity monitoring capabilities including expanded integration with environmental indicators and other administrative areas. In FY 2006, EPA will complete the implementation of a new JFMIP-certified core financial system that complies with recently established federal financial management system requirements.

EPA has conducted alternative evaluations from the following perspective. FINRS will employ COTS whenever practicable (including the modifications of business practices). FINRS will only employ other than COTS solutions when mission critical/legislatively mandated requirements dictate. FINRS will continue to revisit COTS solutions (along with joint business case partners) as commercial products continue to evolve—both to maximize COTS usage and to enhance FINRS attractiveness to a broader user community. EPA will also leverage best practices and E-Gov initiatives in support of the President's Management Agenda and the Federal Enterprise Architecture Business Reference Model. FINRS consists of a portfolio of systems. See Table C-3 for details and associated planned actions and milestones.

Table C-3: EPA Financial Business Architecture Components and Life Cycle Stages

Planned OCFO System	Business Function Supported	Planned System Actions
Financial COTS Application	Provide Annual Planning & Budgeting Provide Financial Management & Services Support Accountability	Retire IFMS, SPITS, ARTS, IDOTS, CPS, Cost Allocation, WCF in 2006; replace with JFMIP-certified COTS software
Planning Application	Manage Strategic Plan Provide Annual Planning & Budgeting Provide Financial Management & Services Support Accountability	PERS and BAS are integrated and BAS is re-engineered in 2005
Cost Recovery & Imaging Application	Provide Financial Management & Services	SCORPIOS is re-engineered by 2007

Planned OCFO System	Business Function Supported	Planned System Actions
Administrative Data Warehouse (ADW)	Provide Annual Planning & Budgeting Provide Financial Management & Services Support Accountability	FDW re-engineered in 2004 into an Operational Data Store
Payroll Time & Labor (PTL)	Provide Financial Management & Services	Implement (replaces EPAYS and CPARS) by 2003
Financial Data Warehouse (FDW)	Provide Annual Planning & Budgeting Provide Financial Management & Services Support Accountability	Replace MARS in 2003; Re-engineer in 2004 into an Operational Data Store
Management Audit Tracking System (MATS)*	Support Accountability	Continue maintenance
Travel Manager (TM)	Provide Financial Management & Services Support Accountability	Replace in FY 2006 with the E-Travel Electronic Travel System (ETS)
* OCFO included this system in its overall systems assessment; however, this system does not require review under EPA's CPIC process given its size and cost.		

In addition to the existing and planned systems noted above, the Financial Replacement System (FINRS) Project will include the selection and implementation of an Enterprise Application Integration (EAI) tool. The EAI tool will support each of the four business functions: manage strategic plan, provide annual planning and budgeting, provide financial management and services, and support accountability.

Finally, EPA plans to modernize the Agency's acquisition systems concurrent with the FINRS Project. Acquisition systems are considered one of the principal systems interfacing with financial systems and constitute one of the major functions within the ASA. As Joint Financial Management Improvement Program (JFMIP) notes, "The principal system that shares information requirements and creates two-way dependencies with the acquisition system is the financial system. These dependencies help to ensure integrity and control in the areas of budget, program management and delivery, external reporting, and data integrity." (Acquisition/Financial Systems Interface Requirements, JFMIP, JFMIP-SR-02-02, June 2002.) Additionally, acquisition systems modernization is one of the objectives of the E-Gov initiatives, particularly as it relates to supply chain management.

Changes and Updates

- Payroll Time and Labor (PTL), operational by FY 2004, will replace EPAYS and CPARS. In FY 2005, EPA will begin receiving payroll services from the National Finance Center (NFC).
- Travel Manager™ implementation shown from the previous ASA sequencing plan was to be completed in August 2004 and is now scheduled for full implementations by FY 2003. It will be replaced with E-Travel's Electronic Travel System (ETS) in FY 2006.

- The FINRS initiative is updated to reflect support for the following E-Gov initiatives:
  - Grants.gov (formerly called E-Grants)
  - Enterprise HR Integration (EHRI)
  - E-Payroll
  - E-Travel
  - Integrated Acquisition Environment (IAE)
  - E-Authentication

HRMS – HR Pro

PeoplePlus includes two significant components: Human Resources Management System (HRMS), and a Payroll Time and Labor system. These are Mixed Life Cycle systems that are in the Operations and Maintenance phase of its life cycle but planned improvements are also being implemented, and tested. Initial planning efforts for this project began in FY 1997. Full implementation will take place in FY 2006. HRMS is in the "control" phase of the CPIC process.

By providing a state-of-the-art, web based, relational database HR information system, EPA is positioned to support Enterprise HR, E-Payroll, Recruitment One-Stop, and E-Training. Each of these initiatives is predicated on the use of government-wide data standards and system interoperability. The HR component of HRMS addresses both requirements.

The HR component of HRMS is a core component of the Agency's ASA. As such, the HR component will serve as the primary source for updating the NetWare Services (NDS) Directory, Notes Domino Directory, Central Mail Directory, etc. It will also consolidate a number of national and local stand-alone locators and databases, thus eliminating redundant data entry and out-of-sync information. The HR component will also interface with the government-wide Employee Express System currently maintained by the Office of Personnel Management.

Changes and Updates

- EPA selected PeopleSoft HRMS for Federal Government (FG) because it offers the most complete set of HR functions with excellent customer/product support. It also has the functions most specifically suited for the federal government and the largest number of federal government clients. EPA's use of PeopleSoft's Human Resources Management System is consistent with the direction of the government-wide move toward a single HR system.
- The HRMS initiative has been updated to reflect support for the following E-Gov initiatives:
  - E-Payroll
  - Enterprise HR Integration (EHRI)
  - Recruitment One-Stop
  - E-Training
  - E-Authentication

IGMS

The Integrated Grants Management System (IGMS) automates the grant and Interagency Agreement (IAG) award and management processes within EPA. IGMS provides on-line development and review of grant guidance, grant work plan negotiation, Agency grant review and funding, and the generation of grant, interagency agreement and fellowship awards. The system supports post award management monitoring, tracking grant performance milestones and closeout activities, and the tracking of utilization of grant funds. IGMS provides managers, grant

specialists, and project officers the ability to monitor grant status and access to current and historical grant, fellowship and interagency agreement data. It can accept both electronic application and reporting data. IGMS is designated an EPA mission critical system. It is a mixed life cycle system in the control phase of EPA's CPIC process. Development is to be complete by the end of FY 2006.

IGMS is aligned with two E-Gov initiatives: Grants.gov and the Integrated Acquisition Environment (IAE), specifically the Inter-Governmental Transaction System (IGTS) and the Business Partner Network (BPN.gov). This investment primarily funds enhancements, which will make it possible for IGMS to exchange data with these systems and to implement grant streamlining requirements emerging from the interagency workgroups established in response to the Federal Financial Assistance Management Improvement Act. Through its connection to Grants.gov, IGMS will receive electronic application and reporting data from recipients. Through its connection with the Intergovernmental On-line Registry, IGMS will electronically place and receive orders for goods and services with other federal agencies. IGMS will use BPN.gov to maintain current business card information for grant applicants/recipients and federal agencies engaged in interagency agreements. In each case, IGMS performs "back office" functions within EPA using these three central federal data sources.

IGMS directly supports the "Expanding Electronic Government Initiative" of the President's Management Agenda in three areas. The first is Grants.gov (formerly called Grants.gov**)**. Grants.gov is the E-Gov initiative that will provide grant applicants and recipients an on-line site, (i.e., Grants.gov) where they can find grant opportunities, apply for grants, and report on grants. While this Grants.gov portal automates key portions of the grant process for recipients, it does not automate the back office grant operations for the grant-making agencies. Agencies will need to download application and reporting data from Grants.gov and process this data in their own grant management systems. IGMS performs the back office functions, providing EPA an electronic means to review, approve, award, and manage the grants after they have been submitted through the Grants.gov portal. EPA has already implemented the E-Find portion of Grants.gov. We are now loading 100 percent of our grant solicitations to the Grants.gov site. IGMS funding in FY 2004 supports the implementation of the interface that will provide Grants.gov application data to IGMS. Funding in FY 2005 supports changes to IGMS providing for the receipt of electronic reporting data from Grants.gov.

The development of Grants.gov is part of a much larger federal-wide effort initiated in response to the Federal Financial Assistance Management Improvement Act of 1999. The Act requires the 26 grant-making agencies to collaborate on a thorough analysis of the entire federal grant process, identifying standard data elements for application and reporting and streamlined regulations and processes. EPA has taken an active role in this streamlining process, participating in the Pre-award, Post Award, Audit, Financial and Electronic Workgroups and their many subgroups, as well as in pilot tests at every phase of Grants.gov development.

The other two projects are both part of the Integrated Acquisition Environment (IAE) E-Gov initiative: the Intergovernmental Transaction System (IGTS) and the Business Partners Network (BPN.gov). The IGTS will provide a common electronic interface for ordering goods and services from other federal agencies and transferring funds between these agencies for intergovernmental transactions. EPA is collaborating with other agencies in the workgroup that is defining the standard data elements for ordering and funds transfer and developing the electronic interface. IGMS will provide ordering information to IGTS electronically for interagency agreements. The FY 2005 investment contains funds for IGMS modifications to provide electronic ordering data to IGTS and to automate the internal approval and funding of interagency agreements within EPA.

BPN.gov is a central electronic registry in which grant recipients, federal agencies engaged in interagency agreements, and contractors can maintain current organization and contact information for the use of the entire federal government. Using BPN.gov eliminates the need for federal contracts, grants, finance, and intergovernmental transaction systems to maintain public address books for external organizations. IGMS will use BPN.gov for grant recipient and federal agency information. EPA will import BPN.gov data and use it to replace external system public address books for IGMS, the Integrated Contracts Management System (ICMS), and the Integrated Financial Management System (IFMS). The FY 2004 investment includes funds for integrating BPN.gov data into IGMS.

Changes and Updates

- References to E-Gov's Grants.gov initiative (shown on the CPIC) has been changed to Grants.gov, its new name.
- The IGMS initiative has been updated to reflect support for the following E-Gov initiatives:
  - Grants.gov
  - Integrated Acquisition Environment (IAE)
  - BPN.gov

ICMS

The Office of Acquisition Management's (OAM) Integrated Contracts Management System (ICMS) suite of applications automates federal acquisition and contract management processes through the generation of solicitations, contract documents and purchase orders, contract modifications, and tasking documents. The suite is comprised of the ICMS application, the Small Purchase Electronic Data Interchange (SPEDI), and Program Office Interface (POI). ICMS is the agency system of record for contract information. ICMS and SPEDI, designated EPA mission critical applications, submit mandated data about EPA's contracting actions to the Federal Procurement Data System (FPDS). All three are nationally deployed systems used by Agency acquisition personnel. Program Office personnel use POI, as well. The Contract Delivery Order Tracking System (CDOTS) was previously reported as part of the ICMS suite, but it is no longer being used in the production environment. The CDOTS database is retained for historical and reporting purposes only.

The ICMS applications are legacy systems in the operations and maintenance phase of the system life cycle, and are in the control phase of the CPIC process. ICMS and SPEDI became production systems in 1995 – 1996. POI was a pilot project until 2002, when it became a production system. OAM plans to retire ICMS, SPEDI and, possibly, POI when different modules of the proposed acquisition system come into production, planned for the years 2005 through 2008. The current applications will be retained until 2010 as a risk mitigation factor for the Acquisition System project. They will be retired as early as practicable. Until such time as these applications are replaced, they will be operated and maintained as production systems, with minimal maintenance modifications required by mandates such as the Integrated Acquisition Environment (IAE) E-Gov initiative.

OAM posts synopses of all solicitations over $25,000 to the Federal Business Opportunities (FedBizOpps.gov) website, with links to each complete solicitation on EPA's website. Past performance information is submitted to the National Institute of Health (NIH) past performance system, and research on past performance is conducted through the Past Performance Information Retrieval System (PPIRS). Planned maintenance modifications to the ICMS suite of applications will increase EPA's use of IAE tools, such as the Central Contractor Registry (CCR). OAM will

also adapt ICMS and SPEDI to report contracting actions as required by the Federal Procurement Data System – Next Generation (FPDS-NG). All federal agencies will be required to report contract actions to FPDS-NG, rather than the current FPDS system, by the end of FY 2004. OAM will eliminate unnecessary enhancements of its legacy systems as the Acquisition System project moves forward.

Changes and Updates

- OAM plans to retire ICMS, SPEDI and, possibly, POI when different modules of the proposed Acquisition System come into production, planned for the years 2005 through 2008. The current applications will be retained until 2010 as a risk mitigation factor for the Acquisition System project.
- The ICMS initiative has been updated to reflect support for the E-Gov's Integrated Acquisition Environment (IAE) initiative.
- EPA will purchase and implement the COTS solution to provide only core acquisition functionality supporting Simplified Acquisition Processing (SAP) and Large Contract Management. This project will be implemented in three stages: SAP, Large Contract, and Purchase Card Web View. In FY 2005, EPA will pilot the SAP module and become operational in FY 2006. The Large Contract Management module will begin implementation in FY 2006 and become operational in FY 2007. The web view of purchase card activity will be available and become operational in FY 2007.
- The Acquisition System initiative has been updated to reflect support for the E-Gov's Integrated Acquisition Environment (IAE). This includes: FedBizOpps.gov, BPN.gov, Federal Technical Data Solution (FedTeDS.gov), and Past Performance Information Retrieval System (PPIRS.gov).

Acquisition System

The Acquisition System investment will provide an intranet-based commercial off-the-shelf (COTS) solution that permits the EPA acquisition community to perform acquisition and business functions in a streamlined, modern response to the mission needs of the program offices. The Acquisition System project will provide acquisition related financial and management information to program offices in real time and allow program managers throughout EPA to aggressively manage mission critical programs. The system will also support performance based budgeting by providing accurate and timely procurement related financial information. The Acquisition System project will be integrated with the financial system and provide end-to-end functionality from program office functions to simplified acquisition and large contracts processing.

EPA will purchase and implement the COTS solution to provide only core acquisition functionality in Simplified Acquisition Processing (SAP) and Large Contract Management. The core capabilities address the minimum requirements of the acquisition process at EPA. These processes touch many users including those responsible for requisitions, program office users, EPA management, acquisition community, finance and property. Functionality that will be provided includes:

- Requisition and work assignment processing
- Simplified acquisition processing
- Contract development and management

- Financial System Interfaces – complete integration with EPA financial systems through EAI tools
- Web access to IAE and other external acquisition tools – linkage to E-Gov tools such as FedBizOpps.gov and BPN.gov

EPA will obtain extended functionality such as on-line catalogs, contractor performance and other information and federal data reporting through the IAE rather than buying or building in house. The Acquisition System project will enable EPA to work with the IAE and will not duplicate or create any process available in IAE.

The Acquisition System project will implement the COTS solution in three stages: SAP, Large Contract, and Purchase Card Web View. In FY 2005, EPA will pilot the SAP module, and in FY 2006 that module will be operational. The Large Contract module will begin implementation in FY 2006, and the module will be operational in FY 2007. In FY 2007 the web view of purchase card activity will be available and the Large Contract Module will be operational.

Results-Based Management Systems

The OCFO included the following Legacy Financial Systems (LFS) in its overall systems assessment:

- Integrated Financial Management System (IFMS)
- Asbestos Receivable Tracking System (ARTS)
- Budget Automation System (BAS)
- Combined Payroll Redistribution and Reporting System (CPARS)
- Contract Payment System (CPS)
- EPA Payroll and Personnel System (EPAYS)
- Inter-Agency Document Online Tracking System (IDOTS)
- Management Accounting Reporting System (MARS)
- Performance and Environmental Results System (PERS)
- Small Purchase Information Tracking System (SPITS)
- Superfund Cost Recovery Package and Image On-Line System (SCORPIOS)

Upgrades to these systems are consistent with Operations and Maintenance phase requirements, and there is no change to the sequencing plan occasioned by these systems. They are included, however, as they remain significant components of the current and future administrative architecture.

Administrative Data Initiative

The ASA Data Initiative comes from the EPA vision of using a single source of data no matter the status or method of collection. The effort serves as the guide for what data will be collected, how it will be formatted, where it will be stored, and how it will be shared with the EPA community while maintaining security standards.

Employee Activity data, Organization Data, and Location Data were identified in the findings of the baseline ASA, and represent mission critical information classes that are created via multiple business processes and stored in multiple applications. This list includes information classes

necessary to meet stated business objectives that are missing or lacking in important ASA applications. These information classes will serve as the focus for data analytics resulting in improved data quality and increased reporting capabilities.

This initiative is in the early planning stages. Work this year includes defining a process to manage administrative data.

Administrative Portal

One of the cornerstones of the ASA target architecture is the EAI tools. Although these tools standardize interface protocols, use common interface functional capabilities, manage communications with internal or external systems, and manage interfaces between administrative applications, EAI does not provide the functional capabilities necessary for external users of the ASA to access administrative systems and information.

An administrative portal will provide secure, personalized, customized, and integrated access to data and applications from disparate sources and locations. The administrative portal will be used primarily to provide and control access to administrative systems and information, and will be able to support new uses for administrative information that is governed by a unique set of business requirements, processes, workflow, collaboration needs, legacy applications, and technologies. By identifying the information requirements specific to administrative lines of business, information solutions will be designed into the portal architecture to meet the specific needs of users external to the administrative domain. These solutions will leverage existing applications and infrastructure to create value and improve business results.

The target ASA currently includes applications owned and maintained by several different offices. Supporting separate portal technology for each of the applications contained in the target architecture would be costly and duplicative. Purchasing portal technology that could be used for all administrative applications would allow offices to share costs, reduce redundancy, and reuse common services.

This initiative is in the early planning stages. Work this year includes a technology review and high-level requirements gathering performed by an inter-departmental work group.

This Page Intentionally Blank

Part 4
Sequencing Applications to the Target

Once the first (pilot) application has moved to the target, all the other principal EHPA applications must follow. During transition, applications must continue to operate in their legacy configuration—with inevitable patches and upgrades to keep up with program needs until the new system is ready. To minimize these parallel costs, transition must be as rapid as possible once a system enters the modernization pipeline.

If the target state were achieved within the usual five year time horizon of enterprise architecture, all conversions would be complete by the end of 2007, five years after the target was designed. Because the pilot application will not be operational in the new state until the end of 2004 (resources permitting), this would leave only three years to transition the rest. "The rest" includes all 20 of the major EHPA applications. In addition, many non-major applications will be important to integrate into the target system. They too often provide data of great value to the public, EPA and its partners, and other federal and state agencies.

In 2004, emphasis will be on the PMO's proof-of-concept projects that will finalize the technical details of the architecture. In 2005, the PMO will bring up the first major system, with all its data flows operating within the EIIPD components. Thereafter, the Agency will migrate other applications into the architecture in a series of waves, based on objective priorities.

Considerations for setting these application migration priorities include:

Business categories

- **Alignment with strategic priorities:** For strategic reasons, the Agency may wish to give priority to certain classes of systems, such as E-Gov solutions, those that support innovative or voluntary controls, or those related to homeland security. It may wish to emphasize key Agency initiatives, certain goals within the Strategic Plan, or specific priorities within the President's Management Agenda.
- **Major systems first (National Program Systems - CPIC):** Systems considered "major" under the CPIC should receive higher priority than "non-major" or "other" systems. It is logical to focus on the core functions of the enterprise first, and, since data integration is the goal, on the systems that house the core stock of the Agency's data assets, that may support E-Gov initiatives, and whose data flows are directly mandated by Congress. Giving priority to major systems will also tend to ensure participation by all program offices simultaneously.
- **Degree of external coordination and customer impact (E-Gov):** Certain applications hold data of value to external partners, such as homeland security and emergency response, or systems relied upon by the private sector. E-Gov initiatives in general will deserve priority consideration.
- **Resource restrictions:** Low cost projects with high impact and proven benefits may receive higher priority than higher cost projects that offer fewer business improvements.

Technical categories

- **New systems**: No new system should be built "outside the box," so all new systems must be integrated with the EIIPD from the beginning.
- **Current Systems Lifecycle (SLC) state**: We should consider how long it has been since a system's last complete modernization, and how urgent is the need for functional upgrades.
- **Technical complexity:** The technical complexity of particular systems may argue to start transition early, or conversely to wait until critical technologies mature. But either way, technical complexity issues will undoubtedly figure in sequencing decisions.

Each of these considerations is discussed briefly below. The Agency has made no decisions on these migration questions, but the outcome is not difficult to see.

Business Categories

Alignment with Strategic Priorities

There will always be strategic and business priorities to consider as specific migration priorities are set. The Agency may wish to emphasize applications that are critical to voluntary compliance or innovative environmental protection, that support major Agency initiatives such as Clear Skies, that support certain goals of the Strategic Plan, or that link to high priority initiatives or systems across the federal government, especially those identified within the President's Management Agenda. It may wish to move all applications within a given business area warehouse at the same time. It may wish to move certain classes of applications, such as administrative support systems or research and science systems, together.

Major Systems First (National Program Systems – CPIC)

Systems identified by the CPIC process as "major" (i.e., with budgets in excess of $1 million per year) tend to include mission-focused national program applications used widely inside and outside the Agency.

Each of the major programs has a relatively small number of applications that are central to its day to day operations and account for a major portion of its IT expenditures for maintenance and operations. Most of these have heavy state usage as well. The best short list of these applications includes those that output data to Envirofacts on a regular basis. The core set includes the Permit Compliance System, the National Compliance Data Base, OECA Docket, RCRAInfo, the AIRS Air Quality Subsystem, the AIRS Air Facility Subsystem, the System for Risk Management Plans (SRMP), CERCLIS, SDWIS, TRIS, and STORET. Added to this list might be new applications such as OPPIN.

These applications also show a good cross-section of program office representation, and, since the Envirofacts data model can be a starting point for creating the enterprise data model of the FBW, this set makes excellent sense for early migration to the target.

Degree of external coordination and customer impact (E-Gov)

A number of systems are critical to partners outside the Agency, including partners in the private sector. Certain systems are critical to E-Gov initiatives. EPA plays the lead role, for instance, in E-Regulation. E-DOCKETS is central to the E-Regulation effort, and for that reason might well be given priority as it migrates to the target.

Resource Restrictions

Cost-benefit considerations and other resource-related restrictions may play a role in setting priorities for migration. Certain non-major applications that offer important benefits may be less expensive to migrate to the target than larger and more complex applications. In the current resource environment, the overall timetable for achieving the target architecture will be resource driven.

Technical Categories

New Systems

New systems must link from the start with the central services. Systems that start development before the system is ready in 2005 will work closely with the EIIPD to integrate their structures with the design of the central services as they evolve.

Current Systems Lifecycle (SLC) State

Most of the major systems are currently in operations and maintenance mode, though many are being marginally upgraded (in "mixed" mode) to integrate additional needed features and functionality. The most important system life cycle factor to consider is the length of time since an application's last major modernization. Several of EPA's systems have been in service for several years without a major upgrade. Some are still mainframe-based. As the Agency moves to a more formalized approach to systems lifecycle management, this may help set priorities by which certain classes of applications are queued for inclusion in the central services.

Technical Complexity

The EIIPD is piloting a number of technical approaches for implementing the central services, including a Web services approach for creating a FBW of loosely connected databases, implementation of new portal technology, and centralized management of metadata. Program offices retain the option of using a variety of options for attaching to the central services, some more technically complex than others. And certain systems, such as STORET and SDWIS, involve the deployment of external counterparts to state and local partners that may require technical linkages through Web services or other means to maintain synchronization. All of these technical complexity factors may weigh in the decision of when and how to move these systems to the target. It may be best to tackle complex issues immediately to ensure adequate time to resolve them. In other cases it may be best to wait until the commercial sector evolves technologies to their next iteration. Security concerns are another technical area that could influence timing of migration to the target architecture. It is essential to maintain security of operations even if that delays implementation of certain systems or system components.

This Page Intentionally Blank

Part 5
Technology Architecture Sequencing Plan

The Technology Architecture sequencing plan (TASP) presents a five-year vision for transitioning from the Agency's baseline Technology Architecture to the target Technology Architecture. The TASP is a strategic plan for attaining the target Technology Architecture (TA) through the sequencing of enterprise technology projects. The CIO Council defines a sequencing plan as "[a] document that defines the strategy for changing the enterprise from the current baseline to the target architecture. It schedules multiple, concurrent, interdependent activities, and incremental builds that will evolve the enterprise." Development of EPA's TASP has been guided by this definition.

EPA's Enterprise Architecture (EA) program includes a focus on EPA's information technology to support EPA's mission and business domains described in the Environmental and Health Protection Architecture (EHPA), Administrative Systems Architecture (ASA), and Research and Science Architecture (RSA) architectural components. The TA is the foundation layer of EPA's Enterprise Architecture pyramid. It represents the technology and infrastructure that enables and supports the Data and Application layers of the enterprise.

In 2002, EPA updated its baseline Technology Architecture and generated its first-ever target TA, Technical Reference Model (TRM), and Standards Profile (SP). In 2003, the focus has been on developing this first sequencing plan to span the five-year timeframe between the current infrastructure and the target architecture. The TASP therefore addresses the Agency's technology plans and implementations from FY 2003 through FY 2007.

This Page Intentionally Blank

Overview of the Technology Architecture Sequencing Plan

Methodology

The primary inputs to the TASP are the requirements and technology directions identified in the target TA. Other key inputs to the TASP include the target architectures of the business domains, in particular the target EHPA and target ASA. Additional inputs include information technology (IT) development programs currently underway, anticipated technology changes from industry, changes in business goals and operational priorities, budget priorities, and federal requirements. Because the target architecture for the Agency's third business domain—the Research and Science Architecture (RSA)—is currently under development, RSA technology requirements were collected as made available by ORD IT managers. Accordingly, RSA requirements for collaboration technologies are probably under-represented in the TASP.

Because EPA has a wide-ranging set of information technology operational plans, the first step in developing a TA sequencing plan focused on collecting and analyzing draft sequencing plans from other EA components, the Agency's Research Agenda, and Capital Planning and Investment Control (CPIC) proposals. The requirements of federated initiatives, such as E-Government and the new Federal Enterprise Architecture (FEA) model, were also considered, including the FEA Technical Reference Model (TRM) and Service Reference Model (SRM).

To garner consensus on the scope, design, and content of the TASP, the Technology Architecture Work Group (TAWG) held several teleconferences in April and May of 2003 and convened on June 9–10, 2003, in Arlington, Virginia. The teleconferences and meeting served as mechanisms to obtain consensus on how best to represent the sequencing of technology implementations and to provide a conduit for regional, laboratory, and program office representation.

The June 9–10 meeting of the TAWG included presentations on the FEA, the EHPA, the ASA, the RSA, the Office of Environmental Information (OEI) Architecture, the OEI Program Management Office (PMO) program, and the Technology Architecture Change Management (TACM) Research Agenda. The PMO program provided input on the coordination of seven specific technology development initiatives related to the EHPA.

During this meeting, the concept and theory of the sequencing plan was presented and discussed, and a draft outline for the TASP was developed. A survey was administered to the TAWG in which participants identified the most important technology directions outlined in the target TA in terms of priority to EPA's business and the representative business segments of those voting. This survey resulted in a priority order for target technologies upon which to focus the sequencing plan.

A straw TASP was developed as a means for discussing the timelines and schedules of TASP technology projects. It was presented to Headquarters Desktop Services Division (HDSD) and National Technology Services Division (NTSD) managers and staff, and led to incorporation of new content, specification of additional detail, and clarification of scheduled and actual timelines for technology implementations. The TASP described in this document is the final product of this iterative process.

Benefits of the Technology Architecture Sequencing Plan

The TASP provides tangible benefits to Agency senior management and technology developers, as described below.

Senior Management: The TASP describes the Agency-wide timeframe for major technology implementations and thus can assist senior management in scheduling project resources, defining annual budgets, and procuring technology. Senior management must coordinate the Agency's technology development to meet customer requirements and to ensure that limited resources are allocated appropriately to meet those needs. The TASP provides a tool that senior management can use to direct and monitor resources used for technology investment and to develop implementation schedules. To further assist management with oversight of schedules, the TASP also highlights areas of "clumping" where too many projects overlap or where there may be periods of sparse activity, indicating that some implementations may need to be rescheduled.

Senior managers also have responsibility for determining whether a proposed investment is consistent with the sequence and priorities in the Enterprise Architecture to ensure progress toward the target architecture. The TASP provides a direct means for managers to monitor and make adjustments to proposed investments to ensure alignment with the Agency's Target Architecture.

Technology Developers: The TASP is a tool that can help EPA technology developers avoid technology conflicts, account for dependencies, and avoid periods of "technology limbo." Developers and implementers must be aware of the technology prerequisites that will directly affect their ability to initiate or complete their work. The TASP provides insight that will help developers anticipate these prerequisites and plan for them accordingly. Conversely, some technologies can be implemented before their time and languish while being of little use to the enterprise. Premature implementation results in resources being unusable and also runs the risk of technology changing and becoming obsolete before it is actually used. For example, implementing a regional wireless network has no benefits to the enterprise if no project is able to retrofit or update the User Environment hardware and software such that it can receive and transmit over a wireless network.

Another benefit of the TASP is that it can provide the foundation for developing a comprehensive risk analysis for implementation projects. A risk analysis would identify adverse effects and suggest mitigation strategies that support realization of the Target Technology Architecture.

The TASP also benefits contractors engaged in developing and supporting the Agency's information technology infrastructure. Since EPA is supported by a range of IT development contractors, the TASP provides valuable information to help foster coordination across contracts and among distinct EPA customer communities.

Scope of the Technology Architecture Sequencing Plan

The TASP presents information on timelines for technology implementations rather than on specific products, processes, or functional steps. The focus is at a high level. It captures current knowledge and priorities relative to the Agency's IT directions.

The TASP is not a management plan, nor does it establish the EA governance processes necessary to promote attainment of the target TA. It is focused on technology implementation, not on the sequence of business and management processes that influence or affect realization of the target TA. While the TASP diagrams reflect technology implementation schedules for

sequencing, they are not project plans or project schedules. Neither project deliverables, the costs of implementation, nor project resource allocations are reflected in the TASP. Each technology implementation typically has its own project plan, with detailed milestones and governance. The pace of implementation is dependent on resource availability.

There are two primary types of projects in the TASP. The first includes those scheduled for the earlier years of the TASP timeframe, which derive from approved and in-progress efforts. The second includes projects derived from longer-term strategic goals and plans that have not yet gone through a review and approval process, but which are considered to be of high priority and vital importance to transitioning to the target TA. Projects in the latter category are typically scheduled in the outer years of the TASP.

Because of the heterogeneity of projects represented in the TASP, the timeline associated with a particular implementation can represent a variety of steps and processes common to a technology project, including feasibility studies, technology assessments, testing, vendor selection, piloting, and deployment. Most of the timelines in the TASP are time sequenced to include the many steps necessary to implement a technology. The end of a timeline for a particular project signifies the point at which the technology is ready for enterprise-wide use and should be included as part of the baseline TA.

Finally, the relationships among timelines do not explicitly show linked dependencies between technology projects. The temporal proximity of the timelines is based mostly on planned implementation schedules. The TASP presents a current view of an IT environment that is continuously evolving with the addition of new technologies, replacement of obsolete technologies, and incremental upgrades to existing technologies. As such, the TASP will require periodic updates.

Next steps for developing the TASP will include further analysis to identify technology implementation gaps, define and diagram dependencies, and evaluate risk. Risk assessment would include sequencing effects and mitigation strategies. In addition, broader, Agency-wide communication of the TASP will serve to further validate the TASP.

This Page Intentionally Blank

Technology Architecture Sequencing Plan

The following text and diagrams present EPA's TASP. Not all the technology directions from the target TA are identified in this sequencing plan. The focus is on known implementation plans and on selected technology directions identified as having a high business priority (see the Methodology section). Because this sequencing plan reflects the organization of EPA's target TA, the TASP addresses each of the seven TRM service areas as a five-year timeframe that extends from FY 2003 (October 1, 2002) through FY 2007 (September 30, 2007). The TASP contains eight separate diagrams, one per TRM service area and one consolidated view. The diagrams are presented in the order described below. An addendum provided at the end of the diagrams contains additional descriptive information related to each of the planned technology implementations.

User Environment (UE)

The UE service includes all aspects of user devices, operating and filing systems for user devices, and office automation, groupware, and utilities. The technology projects and target directions for this area are sequenced in Figure C-7, TASP—User Environment.

Target TA directions for the User Environment include virtual meetings, better collaboration, online training; full offsite connectivity to enterprise IT resources; smaller, faster, more mobile, more integrated, and more usable hardware; automated management and configuration; and reduction of the desktop suite to include only two main operating systems.

The current sequencing includes implementations for increasing desktop compatibility, reducing the number of operating systems, and providing more mobile and better integrated end-user systems.

Applications

The Applications service includes environmental, business, scientific, geospatial, modeling, portal, collection, analysis, statistical, and enterprise applications used by the Agency to carry out its mission. The technology projects and target directions for this area are sequenced in Figure C-8, TASP—Applications.

Target TA directions for Applications include using a web interface for most users' access that will work on smaller, portable devices in addition to PCs, and using Web Services and middleware for real-time connections to other applications, data stores, and common processes. The target TA for Applications also recommends transitioning toward a more unified applications architecture and linking applications into federal E-Gov initiatives.

The current sequencing accommodates implementations for integrating administrative applications using Web services and middleware tools, linking to E-Gov initiatives, providing Web services for the GIS community, and increasing the use of Web interfaces to access data stores and scientific applications and data.

Hosting

The Hosting service includes mainframe, high-performance, Unix, and Windows servers; storage systems; Internet, intranet, and extranet services; and backup and disaster recovery. The Hosting

service includes centralized and centrally-managed distributed resources. The technology projects and target directions for this area are sequenced in Figure C-9, TASP—Hosting.

Target TA directions for Hosting include greatly increased computing and storage capacity for common servers; open systems standards where appropriate; clustered and managed virtual server farms capable of flexibly providing computing and storage resources to applications upon demand; central hosting services (computing, storage, and disaster recovery) for a wide variety of government customers outside the Agency; and cross-servicing with other agencies' IT resources.

The current sequencing includes implementations for unifying user/server directories; enhancing clustering and management of servers; increasing the capacity and flexibility of central and distributed storage systems; evaluating open system standards for mainframe and high-performance computing and for Web services; enhancing and distributing disaster recovery services; enhancing centralized support for geospatial services; and increasing computational speeds on high-performance systems.

Communications

The Communications service includes the transfer of data, voice, and video; transfer protocols; the physical infrastructure for electronic transfer; and the external facilities leased or purchased by the Agency to provide it with wide and local area networks. The technology projects and target directions for this area are sequenced in Figure C-10, TASP—Communications.

Target TA directions for Communications include greatly expanding EPA network bandwidth to support geospatial applications and connecting to other government networks at high speed; enabling a mobile, communications-intensive user environment; standardizing on the TCP/IP protocol; and providing flexi-place employees with broadband connections that link them fully to EPA corporate assets and services.

The current sequencing contains implementations for enhancing employee and partner remote access to EPA services; converting to pure TCP/IP; piloting of wireless networks; redesigning EPA networks for higher bandwidth, security, and failover; piloting high-bandwidth links to the desktop; and upgrading local area networks for higher bandwidth.

Data

The Data service includes database management systems; data and metadata; the integration, migration, and interchange of data; data quality assurance; and data marts and warehouses. These are primarily the logical mechanisms employed to manage and distribute data. The technology projects and target directions for this area are sequenced in Figure C-11, TASP—Data.

Target TA directions for Data include a central data exchange, an enterprise repository, selected data marts and registries, a geospatial data network, uniform approaches for managing metadata, and new tools for data quality, access, transformation, and analysis.

The current sequencing includes implementations for the development and deployment of the Central Data Exchange (CDX), a Framework for Business Warehouses (FBW) for environmental data, a System of Registries for metadata, an integrated geospatial database, and Web services tools for data transformation and access.

Technology Management

The Technology Management service includes systems management, developer support, user support, and records and content management. The technology projects and target directions for this area are sequenced in Figure C-12, TASP—Technology Management.

Target TA directions for Technology Management include more automated techniques for managing IT services, the user environment, distributing software, and licensing; the use of commodity IT services to manage technology via "managed services" arrangements; tools to closely monitor performance and service levels; and a centralized call center using integrated customer relationship management and asset management tools.

The current sequencing plan comprises implementations for centralized software distribution to Netware users; patch management; automated distribution of anti-virus patterns; support for E-Gov initiatives; piloting of desktop managed services; and call center consolidation and enhancements.

Security

The Security service includes all technologies for identity management, perimeterization, data confidentiality and integrity assurance, data availability, surveillance, audit, and forensics. The technology projects and target directions for this area are sequenced in Figure C-13, TASP—Security.

Target TA directions for Security include dynamic, two-factor authentication as the minimum requirement for all data communications with employees and trusted partners; encryption for all sensitive messages and files; security policies and practices to ensure data confidentiality, integrity, and availability; management of user identities, classes, roles, and authorities; defined security perimeters, both within and outside EPA; and monitoring, auditing, and incident forensics.

The current sequencing contains implementations for comprehensive Identity Management with simplified and reduced sign-on; advanced auditing services; file and message encryption; firewall centralization; perimeterization of Agency IT resources; and piloting of more secure remote access methods.

Consolidated TASP—Major Technology Initiatives

A consolidated view of the major technology initiatives is shown in Figure C-14, EPA Consolidated TASP. This figure presents a summary-level diagram of the major projects and technology directions taken from the individual service area timelines. The consolidated view repeats items from each of the individual service diagrams but does not include all the projects from the service area diagrams. Rather, it includes, at a higher level of abstraction, those technologies identified by the TAWG as having the highest business priorities or potential effects on the transition to the TA. In the consolidated diagram, some items may be merged under a higher-level name to simplify the presentation and items are presented in the order of the previous service area diagrams. This integrated diagram enables simpler visualization of sequencing relationships and the enterprise progression of EPA's technology development.

This Page Intentionally Blank

Technology Architecture Sequencing Plan Diagrams

This Page Intentionally Blank

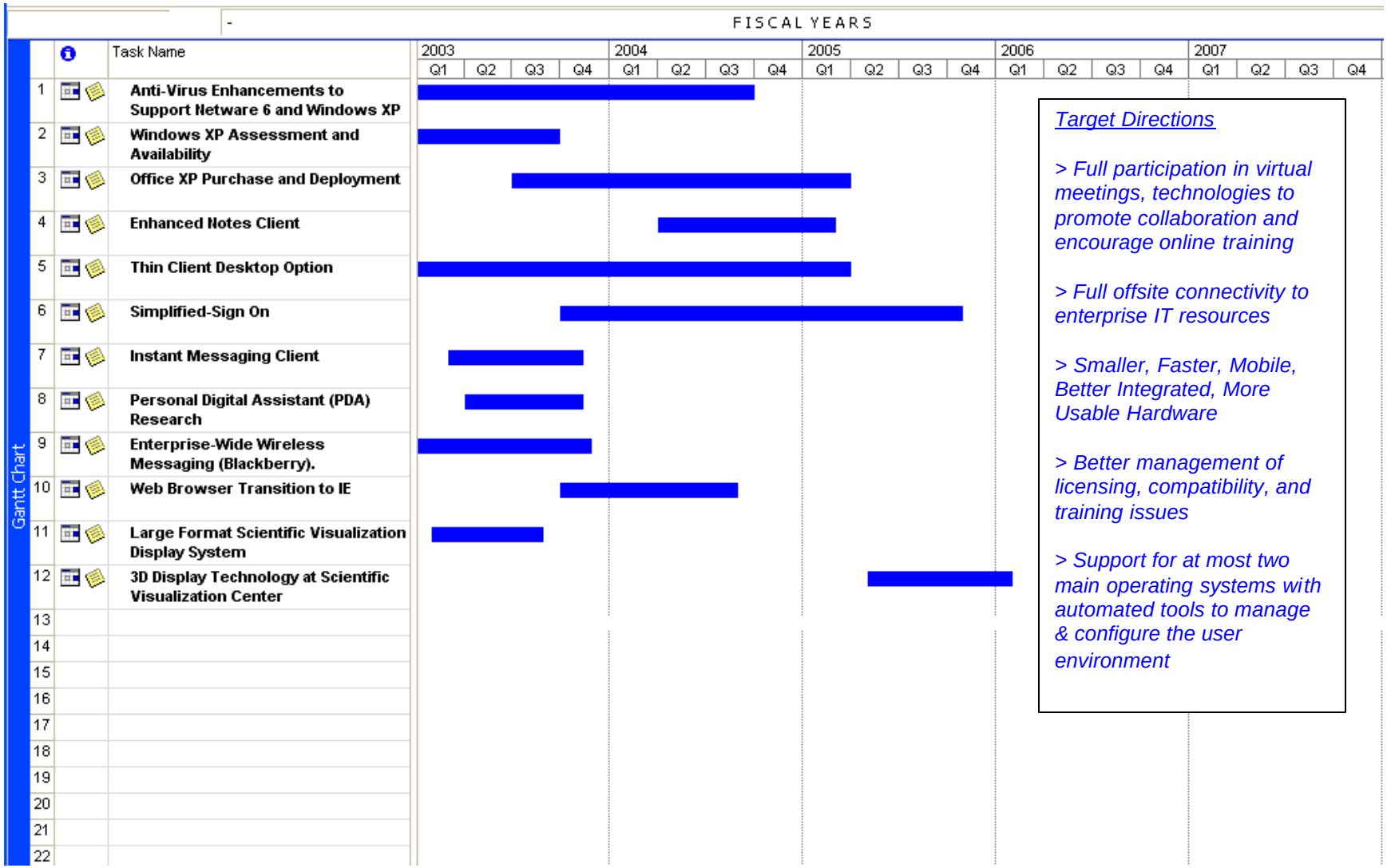


Figure C-7. TASP - User Environment

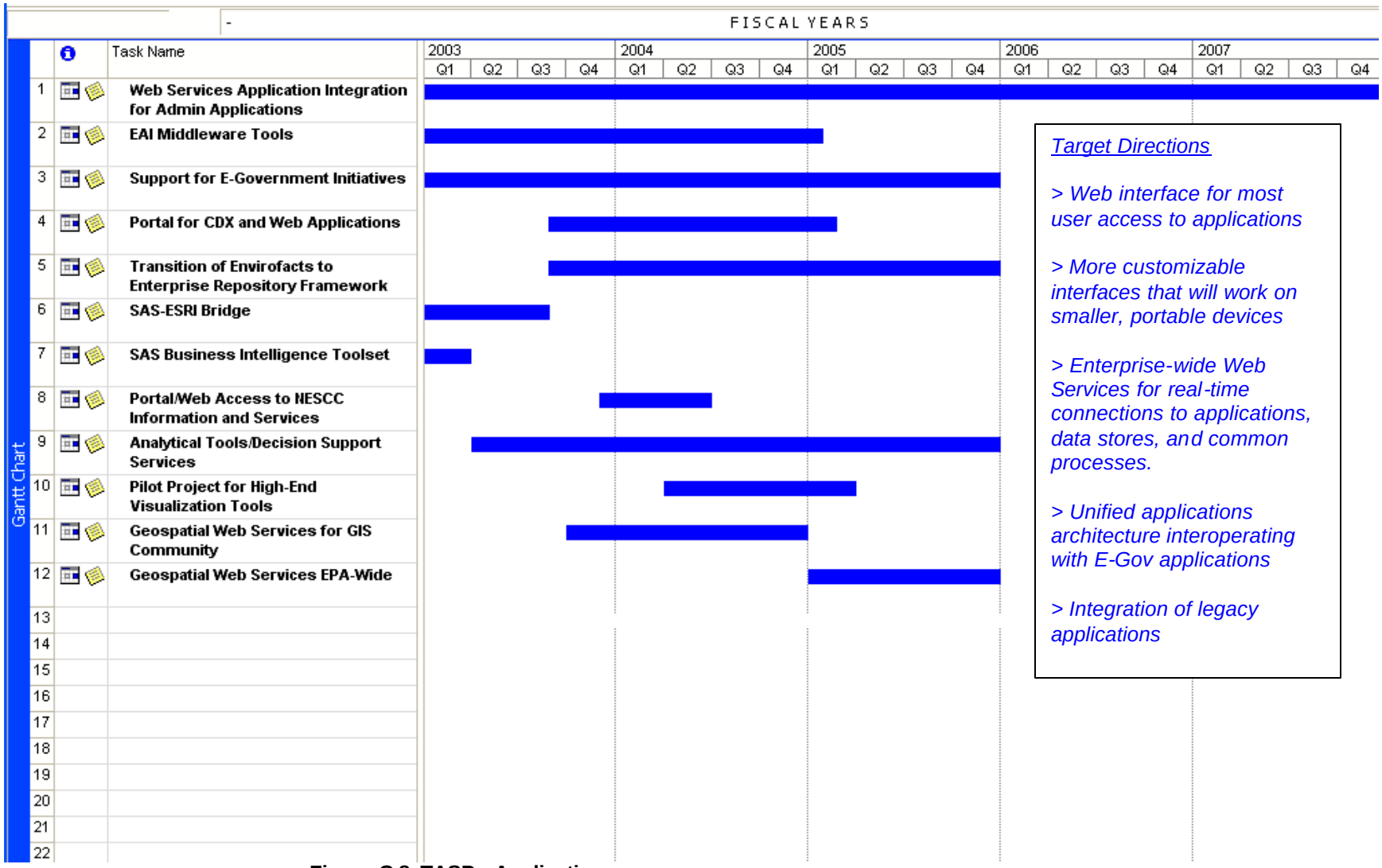


Figure C-8. TASP - Applications

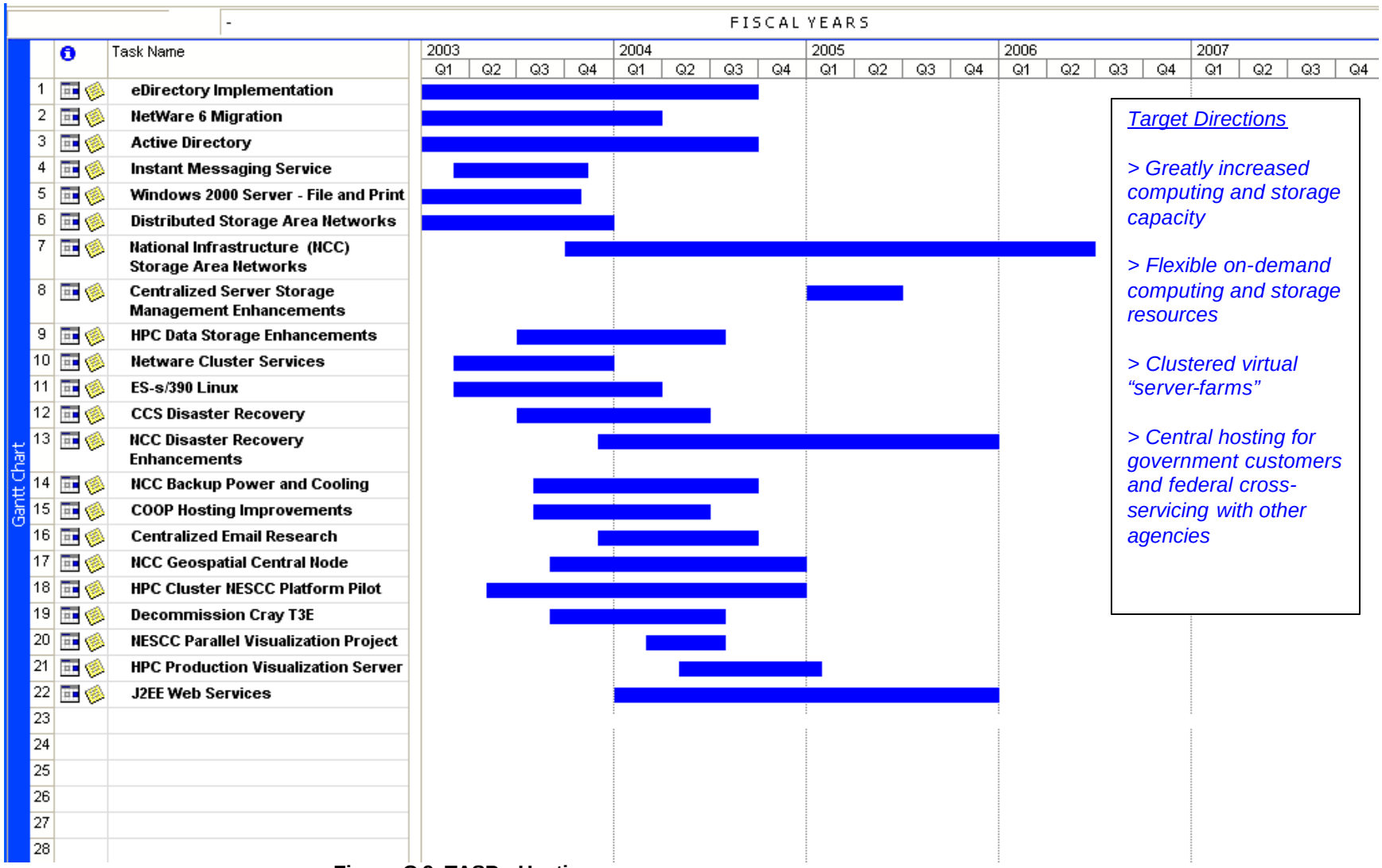


Figure C-9. TASP - Hosting

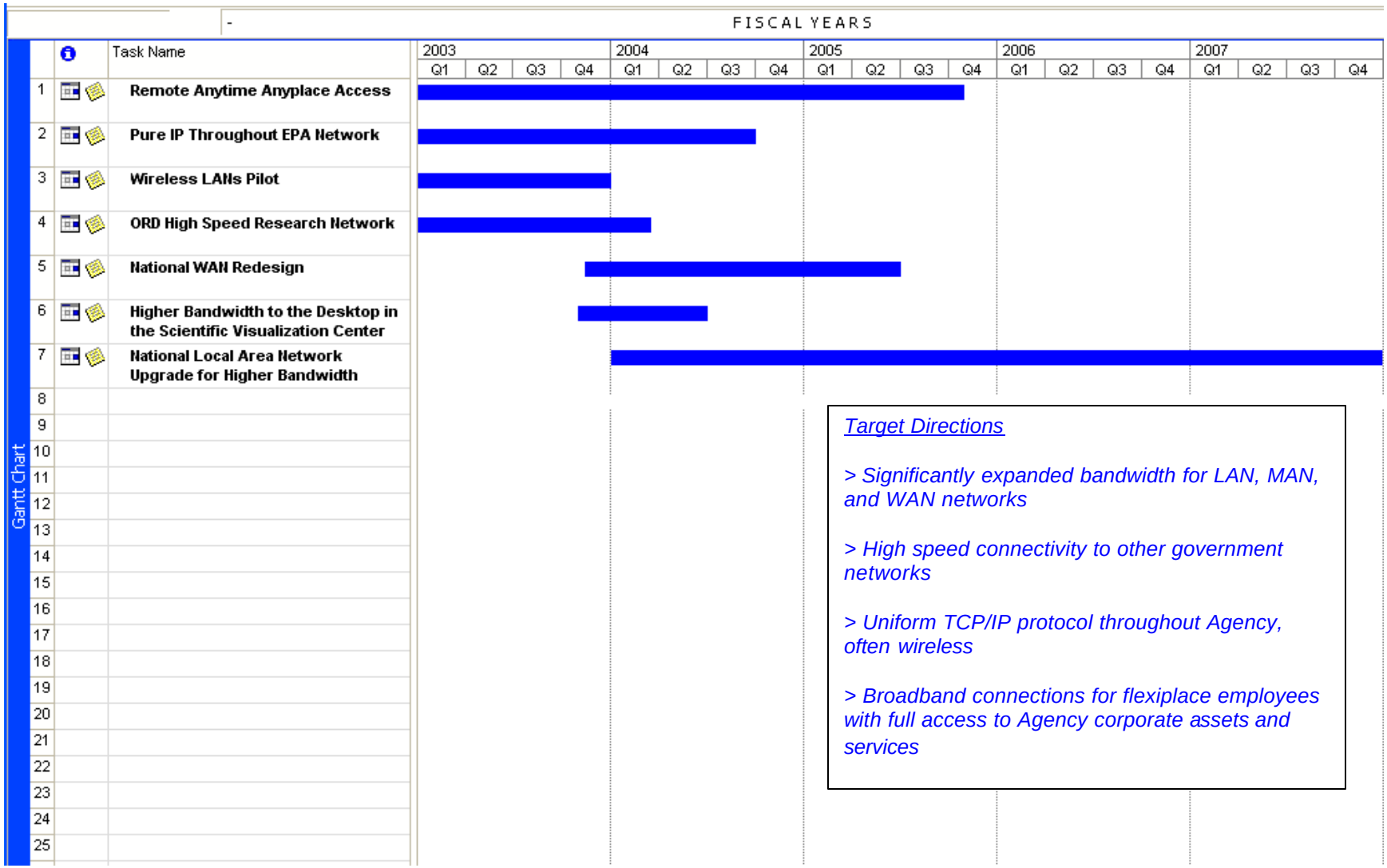


Figure C-10. TASP - Communications

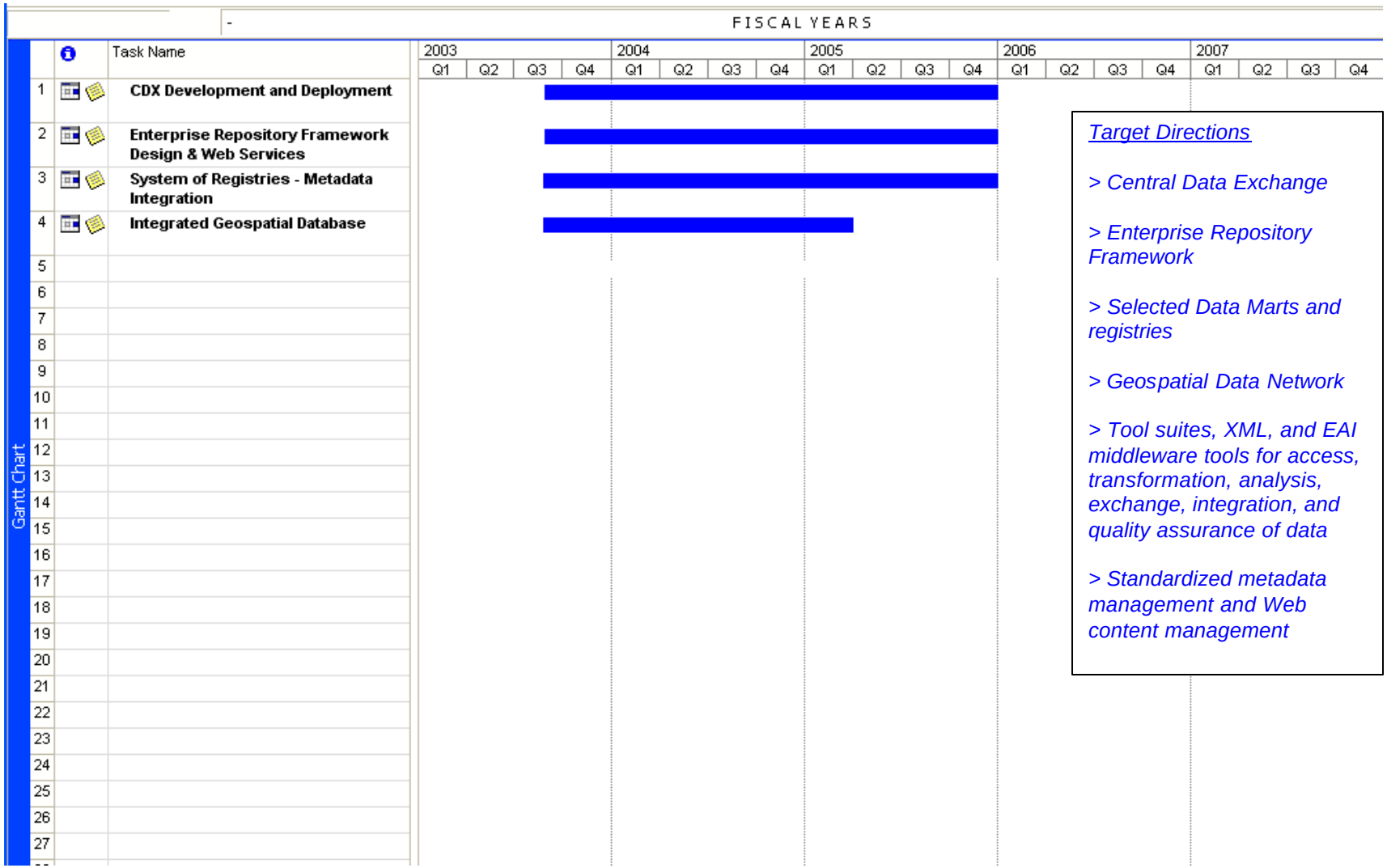


Figure C-11. TASP - Data

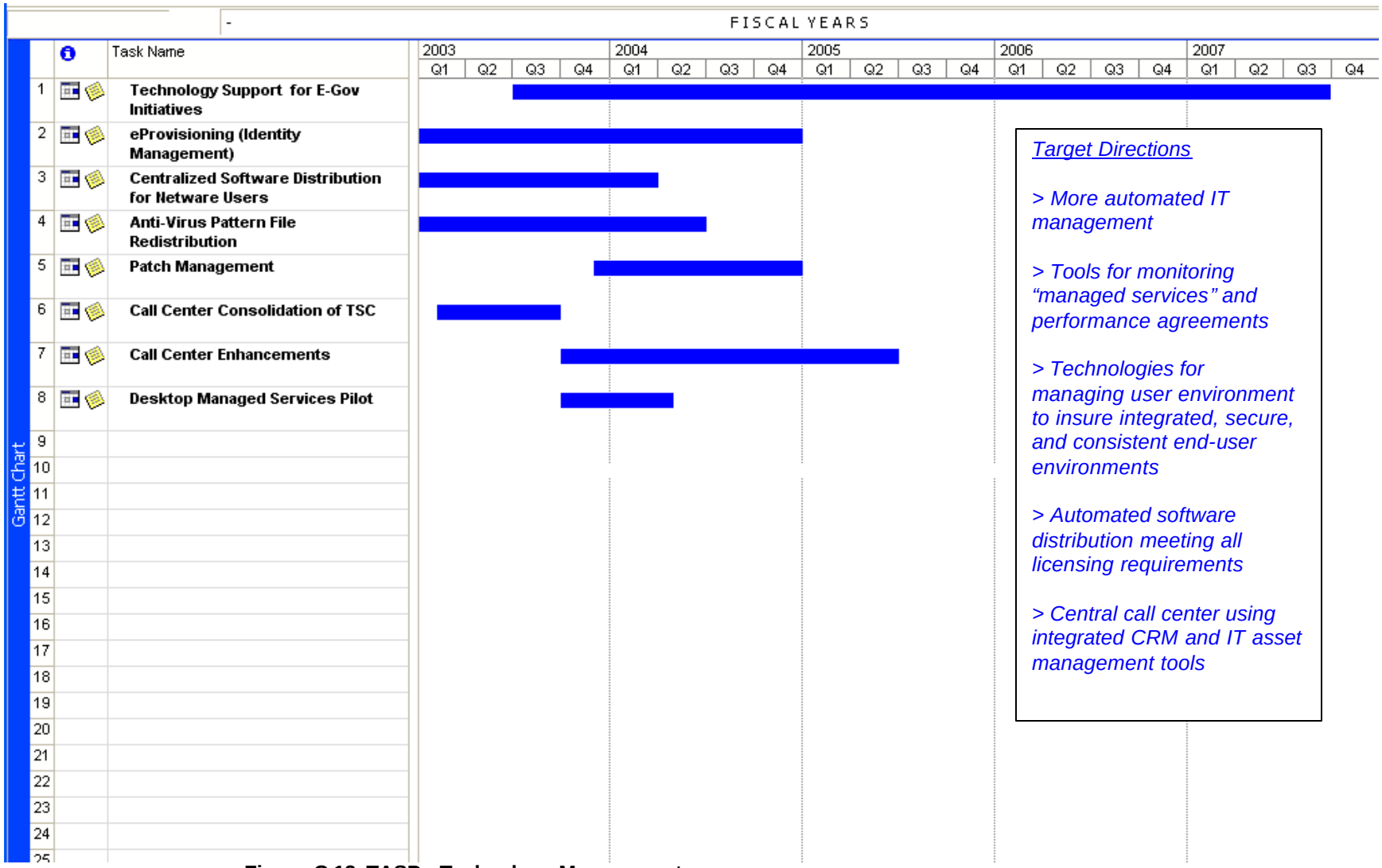


Figure C-12. TASP - Technology Management

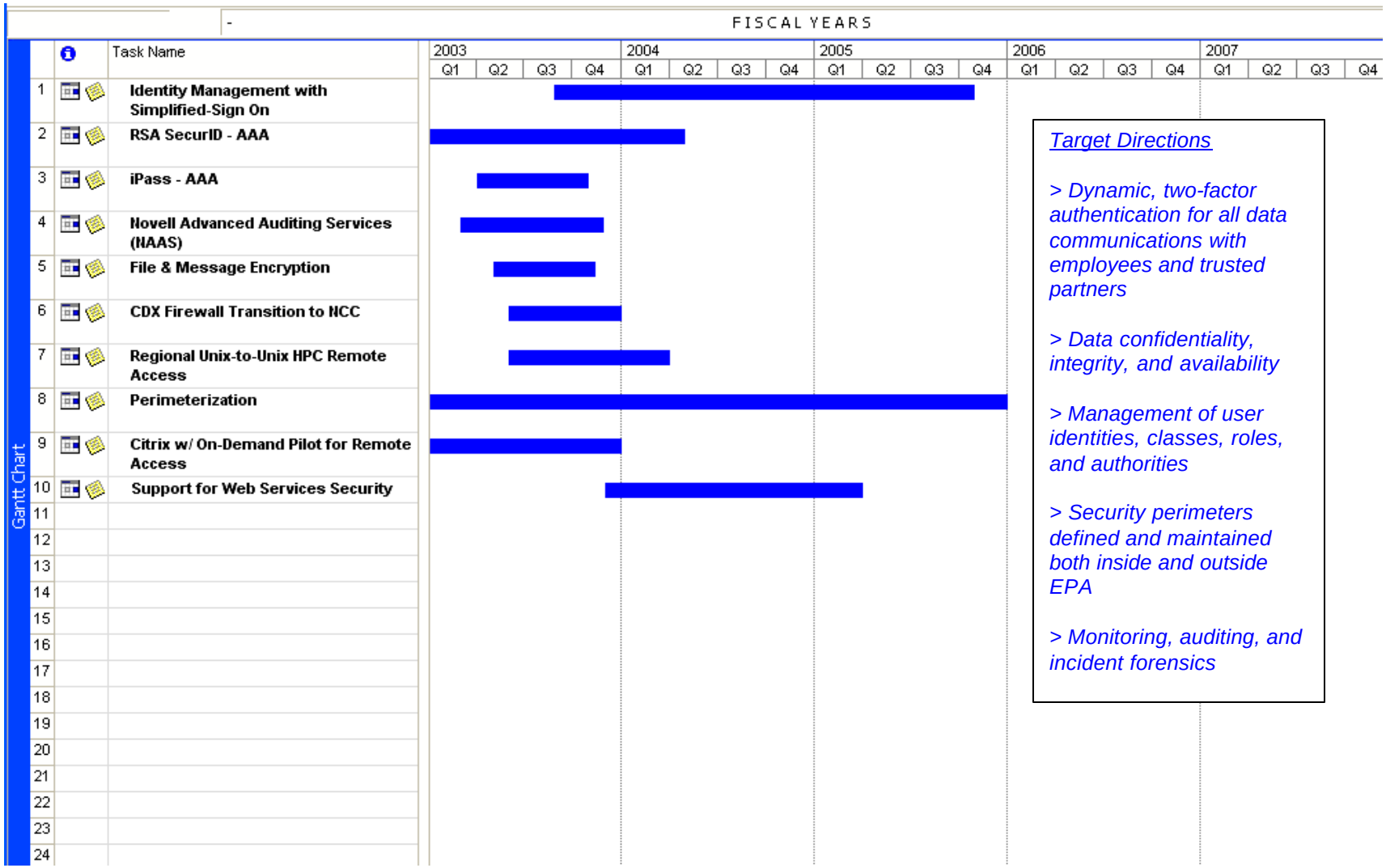


Figure C-13. TASP - Security

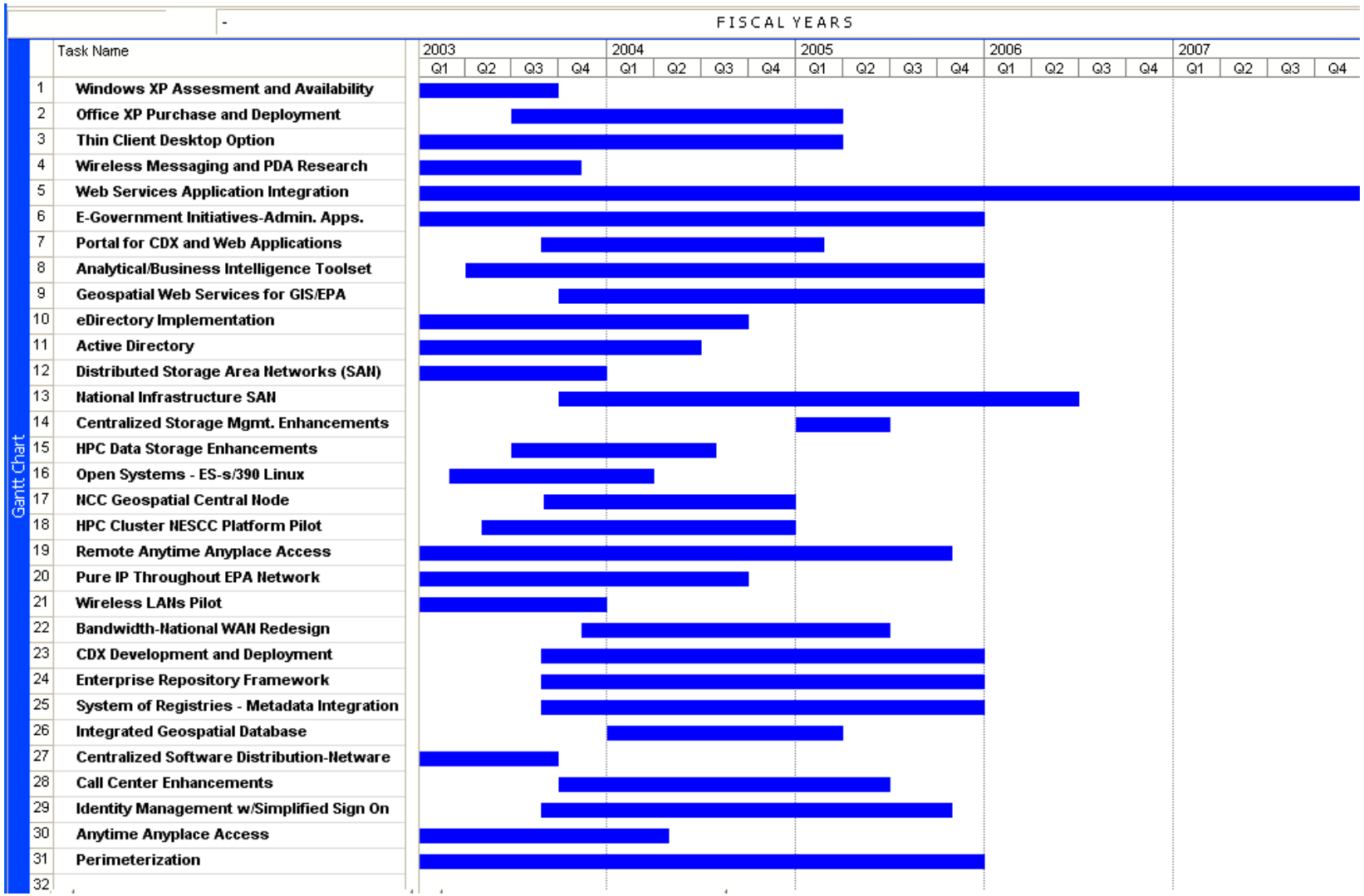


Figure C-14. TASP - Consolidated Technology Architecture Sequencing Plan

This Page Intentionally Blank

Addendum - Notes to Technology Architecture Sequencing Plan Diagrams

This section contains annotated notes for the TASP figures. The annotations are numbered to correspond with the item numbers for tasks as shown in Figures C7-C14.

Figure C-7. TASP – User Environment

Item 1. Purpose: Next version of Norton Anti-virus, provides support for Netware 6 and Windows XP. Dependencies: Completion and approval of white and deployment papers for Windows XP.

Item 2. Purpose: Next version of Windows desktop OS. Dependencies: SCD completion and security review.

Item 3. Purpose: Third level support agreement implementation to Agency. Dependencies: Migration of all desktops to Windows 98, 2000 or XP. Delayed due to the new SmartBuy program.

Item 4. Deployment of latest version of Notes Client to provide enhanced collaboration and better remote access interface.

Item 5. Design, development, piloting, and deployment of Thin Client desktops to selected EPA Regions and Offices that have requested this capability.

Item 6. Purpose: Sign on only once for a particular risk level. Dependencies: E-Directory, Policy Review by TISS.

Item 7. Support available for AOL Instant Messenger (IM) and Microsoft IM.

Item 8. An Agency PDA standard has not been selected but some research is in progress.

Item 9. BlackBerry Messaging Pilot completed in FY 2002. Capability is being made available to the rest of the Agency.

Item 10. Effective July 1, 2003, Internet Explorer will become a co-standard with Netscape as the Agency's standard Web browser. OEI will support both browsers during a transition period. At the end of the transition period, OEI will no longer support Netscape.

Item 11. SVC will complete the large format display system by adding the two remaining display cubes. A briefing of the results of the implementation, deliverable to NESC2 management, HPCEC, and HPCWG.

Item 12. In collaboration with customer input, the SVC will obtain and deploy 3D display technology, such as an ImmersaDesk, CAVE, or visualization wall with 3D projectors and glasses.

Figure C-8. TASP – Applications

Item 1. Use of Enterprise Application Integration middleware and XML technologies to integrate and reduce the number of administrative applications from 53 to 37.

Item 2. Enterprise Application Interfaces to: HR-PTL, FDW, IFMS, IGMS, CPS, ICMS, OMIS, SPEDI, TM, SCORPIOS, IVR, IGORMATS, SPITS, WCF, ARTS.

Item 3. Includes FBO, BPN, FPDS-NG, E-Grants, IGOTS, E-Catalog.

Item 4. Portal with single sign-on for all CDX applications and Web-enabled applications; Phase I - FY 2003, Phase II - FY 2004, Phase III - FY 2005, incorporate all CDX functionality into portal, integrate with E-Gov initiatives, operate and maintain 21 data flows (CDX); expand capabilities to additional nine flows (CDX); 24 states using Node (CDX); make available integrated data tool.

Item 5. Integration of Envirofacts into Enterprise Repository Framework.

Item 6. Use of SAS Bridge software to link to ESRI Geospatial databases.

Item 7. Statistical toolset for data mining, intelligence and reporting; database connections; and other functions.

Item 8. Develop and implement an interface through which customers and stakeholders may launch HPIT jobs, monitor their account usage, gain information about NESC^2 resource availability and services, obtain information about any number of NESC^2 topics, and provide input and feedback to NESC^2 staff and management.

Item 9. Applications for Geo-Enabled Dashboard and Point Enterprise Level Public Access Analytical Tools to the ER.

Item 10. In collaboration with customers, SVC will execute a pilot project as proof-of-concept for high-end visualization tools and applications running on a desktop workstation.

Item 11. First phase scheduled for core GIS user community.

Item 12. Availability of technology to EPA-wide users followed by availability to EPA Partners.

Figure C-9. TASP – Hosting

Item 1. Purpose: Next version of NDS; allows us to do identity management, directory integration, and remote access. Dependencies: Remove all NetWare 4.11 servers from the network; upgrade to a single DS version.

Item 2. Purpose: Next version of Netware; allows us to do remote access and browser based management of many components of NetWare; client-less login. Dependencies: Symantec Anti-virus (SAV) 8.0, completion and approval of white and deployment papers; E-Directory; Novell Advanced Auditing Service (NAAS).

Item 3. Purpose: Directory for Microsoft Windows 2000 and Windows XP - waiver required. Dependencies: Design for Agency; SCD completion and security review, Region 10

Pilot. Parallel ORD Active Directory Project starts 12/2/02 and ends 3/31/04 to transition from Windows NT DNS to Active Directory under Windows 2000/XP.

Item 4. Server support for AOL and Microsoft Instant Messaging.

Item 5. Purpose: Next version of Windows NT which uses Active Directory. Includes Windows 2000 Cluster/Load Balancing. Dependencies: Active Directory Design for Agency; SCD completion and security review, MS Support agreement.

Item 6. Purpose: Improved LAN backups and storage management. SAN deployment will also provide improved data security. Dependencies: Training, deployment guide, funding to complete remaining sites.

Item 7. Integration of Regional SANs at the NCC.

Item 8. Central Client Server managed system to create dynamic and expandable storage with comprehensive backup and failover.

Item 9. Implement the procedures (defined in objective 03.06) for backing up and restoring NESC^2 customer data. The SVC will obtain and install a near-line DVD jukebox with five TB of capacity. Acquire new or additional DMF file management storage unit.

Item 10. Purpose: Provide improved data accessibility. Dependencies: NetWare 6, NAAS.

Item 11. Support for Open Systems operating systems (namely Linux) on Enterprise Server.

Item 12. Disaster Recovery capability for Central Client Server environment.

Item 13. Improvements in networking, firewalls, intrusion detection, backup, hosting, web browsing, and recovery time and coverage for Email. Capability for Continuity of Operations Plan.

Item 14. Implementation planned in 2004 to provide backup power and cooling capability for the NCC. Design phase begins in 2003.

Item 15. Includes Email, Firewalls, Intrusion Detection, and Web Browsing.

Item 16. Evaluate managed service for centralizing Agency's Email servers, consolidation of distributed Lotus Notes Email servers at the NCC.

Item 17. Installation of new Windows and Unix platform servers to provide centrally supported Geospatial services.

Item 18. Use of clustered computers for High Performance Computing (HPC) systems. NESC^2 will test its current collection of applications (such as CMAQ, Fluent, Gaussian), along with several types of tools and supporting software, on a Beowulf Cluster to verify the suitability of using Beowulf Clusters as NESC^2 compute platforms.

Item 19. Phase out of Cray T3E Supercomputer.

Item 20. In collaboration with a customer NESC2 will develop and execute a pilot project employing parallel visualization technology.

Item 21. SVC will obtain and install a server to function solely in a production mode, for both SVC staff and customer use.

Item 22. Includes support for IRUN, Weblogic and Websphere. Target direction for Central Client Server Web services.

Figure C-10. TASP – Communications

Item 1. Purpose: Provide secure, affordable, always on remote access to the Agency. Dependencies: eDirectory, NetWare 6, SAV 8.0, RSA, VPN, ISP.

Item 2. Purpose: Remove IPX traffic from LAN's; improved method of communication. Dependencies: eDirectory, identify all systems that require IPX, upgrade DHCP software.

Item 3. Test wireless local area network access.

Item 4. Project to provide higher bandwidth services within ORD organization for large file transfer, video, and GIS. Pilot scheduled to connect four regional labs. Report of findings due in December 2003.

Item 5. Re-design, configuration and implementation of the Agency's Wide Area Network to improve redundancy, performance, and encryption security.

Item 6. NESC2 will deploy gigabit-per-second telecommunications capability between the SVC servers and the SVC workstations.

Item 7. National infrastructure project to upgrade wiring in buildings for higher bandwidth Ethernet.

Figure C-11. TASP – Data

Item 1. Design and implementation of long term strategy, CDX outreach to Regions and EPA Partners, NTSD production support including ID management, security development, and expansion of capacity.

Item 2. Develop the ER Data Model and Framework, create ER Prototype in development and production environments, integrate with standard program of registries, pilot prototype web services solution, implementation of web services, and Geospatial data improvements.

Item 3. Data and XML validation, metadata registries development, metadata policy development, data normalization, development of Data Model Registry.

Item 4. Will contain the Geospatial portion of the data currently in Envirofacts when this transitions to the Enterprise Repository Framework.

Figure C-12. TASP – Technology Management

Item 1. Management of technology transition to support Electronic-Government initiatives such as E-Dockets, including extra-Agency processing of Agency data, such as Payroll application.

Item 2. Purpose: Directory integration, improved account management across various systems driven by a single source (Peoplesoft). Dependencies: eDirectory, Active Directory, Peoplesoft Version8.0, DirXML, NetWare 6, Notes Directory, Locator Directory.

Item 3. Implementation with ZenWorks for automatic distribution of software to desktop users on NetWare.

Item 4. Purpose: Begin distribution of pattern files all the way to the desktop. Dependencies: Need to know all servers running NAV, Read and File scan access to pattern file subdirectory on servers, additional Bindview licenses, Zenworks for Servers 3.

Item 5. Management of software patching and automatic distribution of software patches. Dependencies: Relies on Zenworks deployment and 100 MB bandwidth to the desktop.

Item 6. Move of RTP Technical Support Center to EPA's Call Center in Chantilly, Virginia.

Item 7. Additional Call Centers in the Agency have been identified (up to 18) and could be consolidated into the EPA Call Center. Transitioning Security Incident Database to the Remedy Service Management Suite and Security Incident call handling. Transition of Remedy to Web-based Remedy 5.1. Transitioning of TSR handling, Change Management Tracking, and Service Level Agreement Tracking to the Call Center.

Item 8. Pilot program for managed desktop services at OEI through computer vendor using three-year equipment refresh cycle.

Figure C-13. TASP – Security

Item 1. Includes Enterprise Identity Management, Software and Hardware Acquisition, Customization, and Implementation. Simplified Sign-On includes sign on once for a particular risk level. Dependencies: E-Directory, Policy Review by TISS.

Item 2. Purpose: Two-factor authentication; part of secure remote access. Dependencies: Licenses, SCD completion and security review, E-Gov Authentication Project registration process.

Item 3. Purpose: Global ISP; part of secure remote access. Dependencies: Licenses, SCD completion and security review, registration process. Currently available for piloting at the Program Office Level.

Item 4. Purpose: Auditing program for NetWare 6, replaces Auditcon. Dependencies: SCD completion, SQL server to read new database, Bindview Version 7.x.

Item 5. Data Integrity Assurance.

Item 6. Transition of firewall to National Computer Center.

Item 7. Regional access to High Performance Computing systems and workstations.

Item 8. Enhancements to securing Agency's networks.

Item 9. Purpose: Thin client solution, part of secure remote access for the agency. Dependencies: eDirectory and NetWare 6; completion and approval of white and deployment papers.

Item 10. Mechanisms for XML Encryption and XML Signature.

Figure C-14. TASP - EPA Consolidated Technology Architecture Sequencing Plan

Since Figure C-14, EPA Consolidated Technology Architecture Sequencing Plan, is a consolidation of the previous seven diagrams, the corresponding notes for Figure C-14 are not repeated in this section.

References

Baseline Enterprise Technology Architecture Update, EPA Enterprise Architecture, August 30, 2002.

Baseline Technology Architecture Version 1.0, EPA Enterprise Architecture, September 9, 2002.

EPA Baseline Administrative Systems Architecture, EPA Enterprise Architecture, July 12, 2002.

EPA Enterprise Architecture Program, Presentation by John Sullivan, Technology Architecture Work Group Meeting, June 9, 2003.

EPA Target Administrative Systems Architecture, Version 1.0, EPA Enterprise Architecture, December 16, 2002.

EPA Target Enterprise Architecture, Version 1.0, EPA Enterprise Architecture, October 7, 2002.

EPA Technical Reference Model, Enterprise Architecture, September 9, 2002.

Federal Enterprise Architecture Framework, Version 1.1, The Chief Information Officers Council, September, 1999.

HDSD Network Projects: Purpose, Prerequisites, and Timelines, U.S. EPA, 2003.

National Environmental Scientific Computing Center Strategic Plan, U.S. EPA, FY 2003.

OEI Blueprint Retreat Lessons Learned, OEI Program Management Office, Presentation by Mike Cullen, Technology Architecture Work Group Meeting, June 9, 2003.

OEI Program Management Office, Status Update, U.S. EPA, 2003.

One-VA Enterprise Architecture Program Management Plan: FY 2003,Version 1.0, June 11, 2003.

The Open Group Architecture Framework (TOGAF), Version 8, "Enterprise Edition," December 2002.

A Practical Guide to Federal Enterprise Architecture, Chief Information Officer Council, Version 1.0, February 2001.

The Service Component Reference Model Version 1.0, A Foundation for Government-wide Improvement, Federal Enterprise Architecture Program Management Office, June 12, 2003.

The Technical Reference Model Version 1.0, A Foundation for Government-wide Improvement, Federal Enterprise Architecture Program Management Office, June 12, 2003.

Technology Architecture Change Management, FY 2003 Technology Research Agenda, U.S. EPA, November 26, 2002.

Treasury Enterprise Architecture Framework, Version 1, Department of the Treasury, Chief Information Officer Council, July 2000.

This Page Intentionally Blank

List of Acronyms

ADA	Administrative Data Analysis
ADW	Administrative Data Warehouse
AIRS	Air Quality Subsystem
AQS	Air Quality Subsystem
ART	Architecture Repository and Tool
ASA	Administrative Systems Architecture
BIMS	Brownfields Information Management System
BPN	Business Partner Network
BPR	Business Process Reengineering
BRM	Business Reference Model
CBITS	Confidential Business Information Tracking System
CDOTS	Contracts Delivery Order Tracking System
CDX	Central Data Exchange
CERCLIS	Comprehensive Environmental Response, Compensation, and Liability Information System
CERES	Coalition for Environmentally Responsible Economies
COTS	Commercial-off-the-shelf
CPIC	Capital Planning and Investment Control
CRM	Core Reference Model
CTO	Chief Technology Officer
DER	Data Element Registry
DMR	Discharge Monitoring Reports
EA	Enterprise Architecture
EAI	Enterprise Application Integration
E-CATALOG	Electronic Catalogs
ECOS	Environmental Council of States
E-GRANTS	Electronic Grants
EHPA	Environmental and Health Protection Architecture
EIIPD	Environmental Information Integration and Portal Development
EIMS	Environmental Information Management System
EIW	Enterprise Integration Warehouse
EPA	Environmental Protection Agency
E-PAYROLL	Electronic Payroll
ERD	Entity-Relationship Diagram
E-TRAVEL	Electronic Travel
FBW	Framework for Business Area Warehouses
FBO	Federal Business Opportunities
FDW	Financial Data Warehouse
FEA	Federal Enterprise Architecture
FISMA	Federal Information Security Management Act
FPDS-NG	FPDS Replacement System
FRS	Facility Registry System
GAO	General Accounting Office
GIO	Geospatial Information Officer
GPRA	Government Performance and Results Act
IAM	Identity and Access Management
ICIS	Integrated Compliance Information System
ICMS	Integrated Contract Management System
ICTS	Institutional Controls Tracking System
IDEF	Interim Data Exchange Format
I-GOTS	Intergovernmental Transactions System
JDBC	Java Database Connectivity
JFMIP	Joint Financial Management Improvement Program
LDAP	Lightweight Directory Access Protocol
M4I	Model for Integration
NCC	National Computing Center
NEI	National Emissions Inventory
OA	Office of Administrator
OAR	Office of Air and Radiation
OARM	Office of Administration and Resources Management
OCFO	Office of the Chief Financial Officer
ODBC	Open Database Connectivity
ODS	Operational Data Stores
OECA	Office of Enforcement and Compliance Assurance
OEI	Office of Environmental Information
OIG	Office of Inspector General
OLE	Object Linking & Embedding
OMB	Office of Management and Budget
ONI	Oracle Native Interface
OPPIN	Office of Pesticide Programs Information Network
OPPTS	Office of Prevention, Pesticides, and Toxic Substances
ORD	Office of Research and Development
OSWER	Office of Solid Waste and Emergency Response

OW	Office of Water
PCS	Permit Compliance System
PMA	President's Management Agenda
PMO	Program Management Office
POI	Program Office Interface
PRM	Performance Reference Model
PRS	Payroll Replacement System
QIC	Quality Information Council
RCRAInfo	Resource Conservation and Recovery Act Information
RDBMS	Relational Database Management System
RSA	Research and Science Architecture
SCORPIOS	Superfund Cost Recovery Package and Image On-Line System
SDWIS	Safe Drinking Water Information System
SIM	Strategic Information Model
SoA	System of Access
SOAP	Simple Object Access Protocol
SP	Standards Profile
SPEDI	Small Purchase Electronic Data Interchange
SRA	SRA International, Inc.
SRM	Service Reference Model
SRMP	System for Risk Management Plans
SRS	Substance Registry System
STORET	Storage & Retrieval System
TA	Technology Architecture
TAA	Target Applications Architecture
TACM	Technology Architecture Change Management
TASP	Technology Architecture Sequencing Plan
TAWG	Technology Architecture Work Group
TOC	Table of Contents
TRI	Toxics Release Inventory
TRIS	Toxics Release Inventory System
TRM	Technical Reference Model
TSCATS	Toxic Substances Control Act Test Submissions
TTA	Target Technology Architecture
UCMR	Unregulated Contaminant Monitoring Regulation
UE	User Environment
UNEP	United Nations Environment Program
XML	eXtensible Markup Language
XSD	XML Schema Definition