IBM Research Report

Technology White Paper on

E-Governance Solution for Government of Maharashtra

Vipul Bansal, Jaijit Bhattacharya

IBM Research Division
India Research Laboratory
Block I, I.I.T. Campus, Hauz Khas
New Delhi - 110016, India.

Table of Contents

1. Current Situation

1.1 Information Technology Revolution and E-governance

Technology is revolutionizing models of information dissemination and commerce all over the globe. As a first step, the use of IT cuts costs by automating manual processes. As a further step, the altered cost structures facilitate new and more effective ways of doing things (by redesigning of the processes themselves). As a still further step, the use of IT makes possible new activities which now become economically viable.

Governments all over the world are trying to utilize IT for various purposes. The initial motivation usually comes from the need to improve efficiency of processes in the government. Examples of such processes may be procurement, approvals, tax collections, registration of various kinds, issuance of certificates, record-keeping of public databases etc. As mentioned above, the first step then is to automate some of the processes. This step may be concurred or followed by the second step comprising re-engineering of the processes. Another set of motivation may come from the need to provide various social services to citizens for improving the quality of life of the citizens. This may manifest in various kinds of information, education, health care and other services offered by governments to citizens over the Internet. A third set of motivation may be to strengthen the democratic foundations of governance, and this may be achieved by information dissemination, open communication and enabling citizen participation (opinion polls, voting etc.). These social services and democratic enablements correspond to the new activities that become economically viable due to the altered cost structures due to the use of information technology.

This white paper will seek to analyze the e-governance opportunities and challenges from the point of view of the Maharashtra government and present an e-governance strategy and a technology cum application model for the state.

1.2 Maharashtra: Current Status, Issues and Challenges

Most of the state government functions are carried out by physical /manual processes involving movement of paper and personnel. On one hand, this calls for a major upfront investment and administrative effort. On the other hand, this represents an opportunity to make use of the latest technologies for maximum benefit and to learn from the experiences of the early adopters of technology.

Government Initiatives

The government of Maharashtra has taken a number of initiatives towards e-governance aimed at enhancing efficiency and effectiveness in government functioning and improving the quality of life of its citizens. In particular,

1. The Maharashtra government plans to computerize the administration and governance of the state, with the following focus :
   a. The deployment should happen in a short time and at a low cost.

b. It should be implementable in a phased manner such that the investments made in each phase can be meaningfully leveraged upon in the subsequent stages.
c. The phasing of the implementation should be such that the applications that provide the maximum benefit are rolled out first.

2. The government (and its various departments and agencies etc.) are trying to use the Internet for providing information to citizens. Noteworthy amongst these are:
   a. Maharashtra state government site (http://www.maharashtra.gov.in/) providing information on the government, the state, community, information related to investments in the state, tourism information and other miscellaneous information. The site is available in both English and Marathi.
   b. Marathi language site of the Nagpur District Collectorate (http://www.nagpuronline.com/nagpurcollectorate/index.htm).
   c. Site of Maharashtra State Electricity Regulatory Commission (http://216.121.216.41/chapter5p1.htm).
   d. Site of the Customs and Central Excise Commissionerate, Nagpur (http://www.cenexcisenagpur.net/).
   e. Maharashtra government's agriculture portal (http://agroadv.gov.in/).
   f. Information Site of Pune Police (http://www.punepolice.com/).

Issues and Challenges

Some of the challenges (relevant to use of IT for e-governance) specific to states in India and particularly so for Maharashtra are as follows:

1. **Geographically dispersed offices:** Maharashtra has around 30 districts. The districts administer about 330 taluks which are located all over the state. Each taluk has around 30 departmental offices through which the state government operates. The taluks are geographically dispersed and many may be located in a difficult terrain. This makes maintenance and support services to user sites extremely difficult and costly and calls for rugged and simple systems, which do not require much maintenance and support.

2. **Absence of reliable and high bandwidth communication infrastructure at taluks:** All district headquarters and some other important cities / towns are connected using VSAT networks operated by the National Informatics Center (NIC) or the state government. This provides reliable connectivity between them with bandwidth ranging from 64 Kbps upto a few Mbps. However, at the taluk (or still further at the village level), the communication infrastructure mainly comprises of telephone lines, which does not have a high bandwidth and is unreliable. While there is interest in laying optical fiber for connecting important cities and towns (which would provide high bandwidth between them within a period of a few years), the story is very different for the taluk and village level, where laying optical fiber would be economically infeasible at least for some time to come.

3. **Information not available in electronic form:** Nearly all the information, like various databases, records, files etc. are currently in physical media (paper). Any kind of use of IT for government operations would require that the underlying data be first converted to

electronic form. Although the transformation can be speeded up with the help of technologies like Optical Character Reader (OCR), this requires considerable will power, time and money.

4. **Lack of coordinated approach:** The use of IT is being driven independently by individual departments with the result that it is leading to the emergence of isolated systems which may not be consistent or compatible. There may also be considerable duplication of effort and wastage of resources in the process. Further, when there is a need to define workflow between such independent closed systems, it may require a significant integration effort. This can be illustrated with two examples:
   a. The government's Internet presence is characterized by a proliferation of individual departmental Web sites each with a separate URL, offering its own department information and online services. This would confuse the end users, who would not know where to look for specific services. Also, it would make offering inter-department services a difficult proposition. Globally, governments are rethinking their Web strategies from their citizens' perspective. Instead of launching online services on a department-by-department basis, they are aggregating services across departments, accessible through a common portal.
   b. Different departments may be developing their own applications, creating own databases and setting up the required connectivity and investing in hardware. On one hand, these databases may be inconsistent. Also, the investments in common applications, connectivity and hardware can be shared across multiple departments. There is a need for an integrated and coordinated approach.

5. **Limited financial resources:** The entire solution, including connectivity infrastructure, hardware, applications, implementation, training and maintenance calls for huge investments which have to come from within the limited resources of the state. This calls for innovative partnering as well as a phased implementation, where the solution can be implemented in a modular way.

6. **Likely Resistance from work force and unions:** There may be a conflict between the interests of the work force currently manning the state government offices and the objectives of the e-governance solutions. One, the use of IT makes information easily available and increases transparency, thereby tilting the balance of power in favour of the end users and citizens. Two, efficiency improvements may render part of the work force surplus, which may call for their retraining and redeployment. This may also face resistance from the employees.

2. E-governance Objectives

The e-governance solution is envisaged to provide an end to end integrated IT support to the government of Maharashtra and to enable the government to efficiently perform its tasks. It seeks to use the emerging state of the art technologies to solve the problems that are specific to the working of the Maharashtra government.

The solution would seek to achieve three classes of objectives as described below:

2.1 Enhanced Government Processes

1. Improve efficiency in governance and administration
   a. Reduce response time in making decisions by streamlining of the governance and making information easily available.
   b. Reduce cost of governance and administration by automating / re-engineering manual processes and making them online for internal as well as external users[1].
   c. Improve revenue generation by facilitating efficiency and transparency in collection of taxes and other revenues.
   d. Improve internal administration of government's own activities like work force administration, monitoring of projects etc.

2. Enhance the effectiveness of governance and administration
   a. Provide consistent and accurate information to decision makers on demand.
   b. Integrate islands of information and allowing the State government to have a holistic picture of the day to day functioning of the government.
   c. Make use of personal and transaction data to find ways of designing better social and administrative programs and their periodic monitoring and evaluation on a continuing basis.
   d. Bring about increased accountability and better auditability in governance.

2.2 Improved Quality of Life for Citizens

1. Make social programs like education (school, college, professional, teacher-training), health care (diagnosis, health education and training, family planning and others), providing information on agricultural products, agricultural practices and prices etc. more effective by:
   a. Improving the reach and accessibility of these programs.
   b. Covering a larger target population base with the same resources due to lower costs.

---

[1] Depending upon the service, the population required to use that service, and other variables, early studies indicate governments across the globe are saving up to 70% by moving services online compared to the cost of providing the same services over the counter. And that figure does not include public costs -- taking a day off work, driving, traffic congestion, parking, and waiting in line.

c. Integrating the programs with the government systems to improve the planning, monitoring and evaluation functions.

2. Provide online information to citizens about government activities, procedures, status tracking etc. including the ability to transact (and pay) online.

2.3 Digital Democracy

1. Improve citizen participation in governance by
   a. Bi-directional communication with government representatives and officials.
   b. Public opinion polling on important matters.

2. Strengthen democratic processes
   a. Provide to the citizens, easier access to the administration and information related to administration and governance.
   b. Facilitate transparency in the administrative processes and decision making.

3. E-governance Solution Framework

A comprehensive e-governance solution which meets the objectives defined in the earlier section will have to address many different requirements which may be present due to different reasons. These reasons may be economic, political, technical and cultural amongst others. The requirements are classified into two categories, (a) generic infrastructural requirements which apply to the overall solution and (b) application requirements, which relate to individual applications that form a part of the overall solution.

3.1 E-governance Solution: Generic Infrastructural Requirements

1. **Phased implementation:** It should be possible to have a unified approach but still implement the solution in phases. This helps to spread the resource requirements over a longer period (reducing the amount of upfront investment required) and learn from early mistakes for the subsequent phases. The earliest applications should be the ones which provide maximum visible benefit.

2. **Easy to maintain and support:** Since the government offices number in thousands and are geographically distributed, it may be economically and technically infeasible to provide a high level of maintenance and technical support to the end users. Therefore, the systems should be rugged, maintenance-free and with little or routine support and upgradation requirements. Also, systems should be simple so that the manpower training and support costs are low.

3. **Integration of inter-department workflows:** It should be possible to support processes involving multi-department and multi-agency workflows. For this purpose, it is necessary that the different departmental offices (and also external agencies) are all interconnected and share the same underlying back-end databases and applications.

4. **Integration of isolated legacy systems:** Some departments or agencies may have already created isolated systems for their own purposes which would need to be integrated into the rest of the solution for creating seamless and interconnected applications.

5. **Scaleable and reliable**: The solution should be capable of scaling with time in terms of number and complexity of applications, number of locations and number of users / usage. It should be reliable, that is, provide assured levels of service for uptime, availability and performance. The solution should also incorporate back-up capabilities and the ability to handle contingencies and recover from failures.

3.2 E-governance Solution: Application Requirements

1. **Support for key verticals:** The solution should support interfaces for government to business (G2B), government to employees (G2E), government to citizens (G2C) and government to government (G2G) interactions as shown in figure 3.1 below.

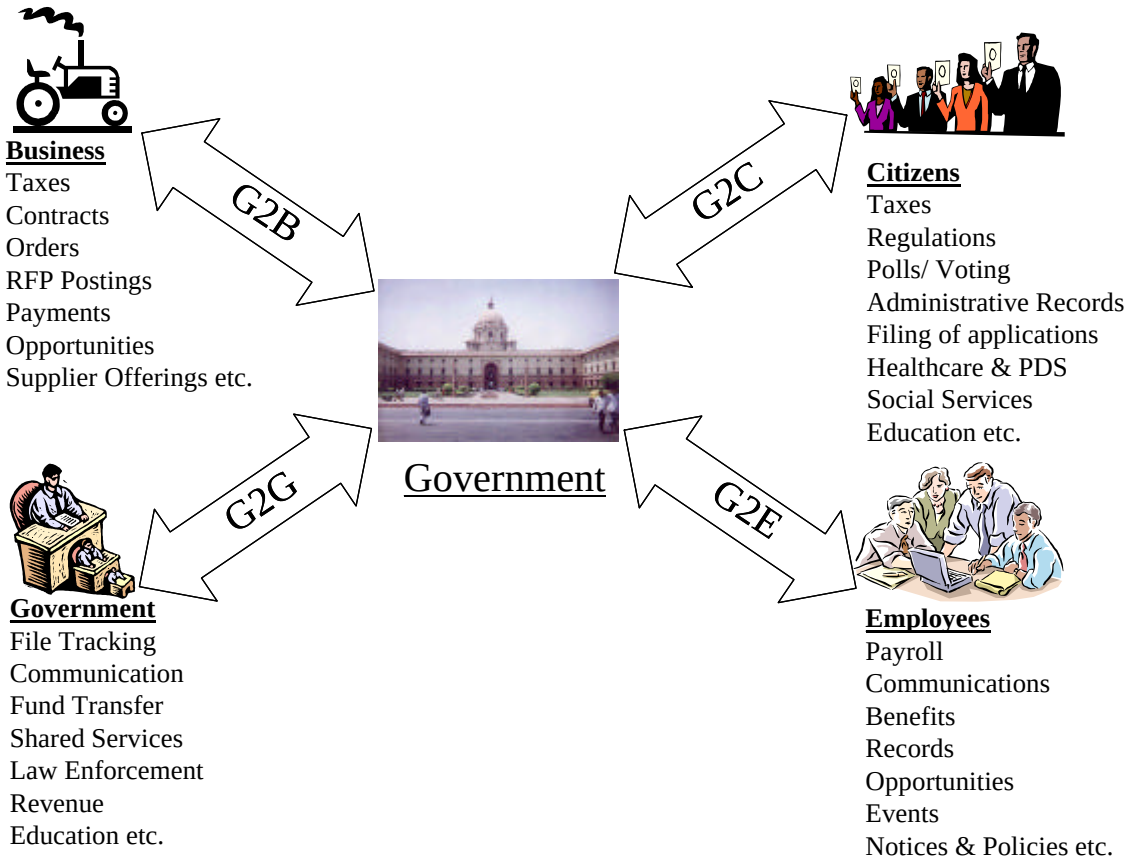

Figure 3.1: Interfaces for Different Verticals

2. **Security features:** Various security services like authentication, multiple levels of access control, confidentiality, privacy, data integrity (prevention of forgery) and non-repudiation should be provided. In addition, ability to generate access logs , especially for sensitive data should be present. The solution should integrate use of public key certificates and digital signatures as applicable in India after the newly passed IT regulations to enable the transactions to have legal validity.

3. **Record retention requirements:** The applications should satisfy the record retention requirements implied by the legal and accounting norms and practices.

4. **Simple human interface:** Human interface should be simple to facilitate greater usage and to keep training costs low. For citizen interfaces, this is all the more important where people may not be literate. Intuitive graphical user interface, use of speech and video technologies etc. may be useful.

5. **Multilingual support:** Native language interfaces would be required both for employee and citizen applications.

3.3 E-governance Solution: Components

A comprehensive e-governance solution would have many components and several activities may need to be accomplished for its successful implementation. Presented below is a simple framework to capture the important components (Figure 3.1 below):

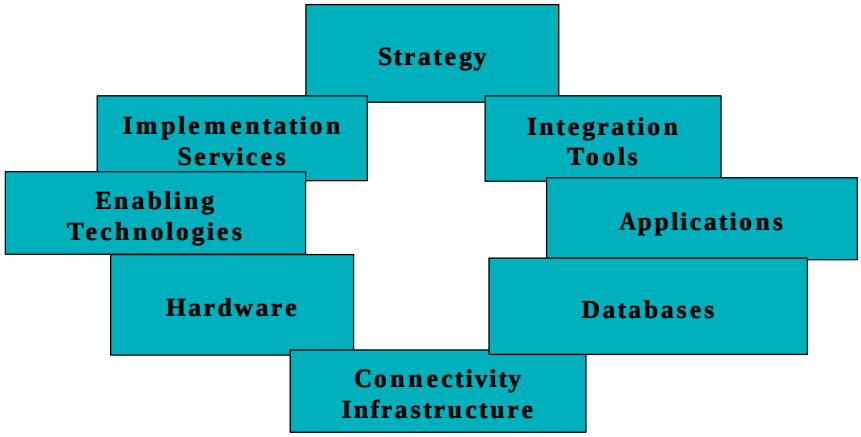

Figure 3.1: Components of an e-governance solution

1. **Strategy:** The strategy defines the scope and the overall approach and answers questions like - which services will be made online, which processes will be automated or re-engineered, how will different departments and different levels of government be integrated, how will solution implementation be phased, what are the milestones etc. Defining a clear strategy and implementation plan is important because it can help avoid duplication and wastage of effort, minimize chances of mistakes and lead to the creation of an integrated system.

2. **Connectivity Infrastructure:** This specifies how various government offices would be interconnected with one another and with external entities (including public). The technologies and scale (bandwidth) of connectivity at various levels should be chosen carefully based on technical feasibility, economic considerations and criticality of requirements.

3. **Hardware:** The hardware description primarily contains specifications for (a) end-user terminals and (b) servers at various levels (taluk, district and capital)**.**

4. **Data Management:** Databases are the starting point of most applications. Since most of the information with the government currently resides on paper media, a giant effort may be required for converting it to electronic databases.

5. **Enabling Technologies:** These comprise technologies such as those for pervasive / mobile access, speech interface technologies, security technologies and solutions,

technologies for conducting electronic transactions and multi-lingual support technologies amongst others.

6. **Applications:** Applications comprise the software used by the end users for their activities. They may be standard applications with or without customization, or custom applications developed specifically for solving the problems at hand. Applications typically make use of one or more databases and may have enabling technologies embedded into them.

7. **Integration Tools:** The integration tools comprise middleware and workflow tools which are used to provide seamless functioning across multiple applications or databases. They are required because (a) some applications / databases may have originated independently and may not be compatible in format or data model with other applications / databases, and (b) for defining complex applications that span across multiple simple applications or multiple levels of users or multiple departments.

8. **Implementation Services:** The implementation services would be needed for development of applications, their integration and deployment.

Besides these components, there may be need for ongoing maintenance and support as well as upgradation of the hardware / software. Training of employees to enable them to use the applications would also be required.

4. E-governance Model for Maharashtra

An e-utility based e-governance model is proposed below for meeting the objectives set forth earlier. In this model, the user-level applications can be viewed as services provided by an e-utility to the end users and they get delivered over a computer network. The chief characteristics of the model are discussed below.

4.1 E-Utility Model: An Overview

The e-utility model of e-governance:

1. Would define the framework (an abstract view) for setting up the entire e-governance solution comprising connectivity infrastructure, databases and applications in such a manner that the users using the system do not have to be bothered about the details of how the connectivity, databases and applications are implemented. The users would only see the services provided to them and their relationships with the other users in the system. Examples of such services may be:
   a. Read and modify access to data of various kinds.
   b. Security services like authentication, access control.
   c. Communication services like e-mail and messaging.
   d. Application services for information, workflow and collaboration.

2. Would provide interfaces to the various stakeholders in governance and ensure their online collaboration through a seamless workflow. In other words, the interfaces with the stakeholders would comprise points of data entry and information release of an otherwise integrated system. At the back-end, there would be a single integrated database system and a unified view of the data (defined by data masters) so that the consistency of data is ensured.

   The various stakeholders are:
   a. **Internal users** - for carrying out government's internal operations, for example, officers in various departments using e-utility services for file movement, project monitoring, revenue monitoring etc.
   b. **Governments (G2G)** - for intra-government coordination amongst different departments and agencies.
   c. **Employees (G2E)** - for better management of government employee-related affairs.
   d. **Businesses (G2B)** - for interaction with business organizations for functions such as procurement, contracts, tendering, tax collections etc.
   e. **Citizens (G2C)** - for providing various citizen services like tax payments, registration services, opinion polling, issuance of certificates and for providing social services like healthcare and education.

   We recommend that the integration of government's internal functions, G2G and G2E should be over an *intranet*. The integration with the outside world, namely the G2B and G2C interfaces should be over *extranet* which can also be the Internet itself. There should be a unified data model for both internal and external users and the internal and

external components of workflow should be seamlessly integrated. For example, tax payments may happen through extranet by citizens and businesses and the aggregate reports on tax collections may then flow to officers in the tax department on the intranet from the same databases which get updated during the payments over the extranet. This would ensure consistency of information and timely and accurate response. It also would also avoid duplication of effort.

3. Would provide numerous benefits in terms of a system which is state of the art, scaleable, easy to grow (in terms of applications, databases as well as users), easy and cost-effective to manage, maintain and upgrade and simple to use.

4.2 Intranet and Extranet

An *intranet*, which would be a computer network private to the state government, would interconnect all the departmental offices upto the taluk level. The intranet would use the same web technologies to link employees across departments for work-flow processing, e-mail, collaboration tools, and government applications. The intranet is thus a network for the government to carry out its own internal activities and for G2G integration. The intranet would be formed by utilizing various networking technologies at different levels (districts, taluks etc.). The employees (end users) would use computer terminals which would all be connected to the intranet. All the applications and data used by the end users would reside on servers and databases connected to the intranet and would be delivered to the end users over the network.

The model also envisages an *extranet* capability, which would link external stakeholders (such as businesses, citizens etc.) to the government systems replacing paper-based and over the counter interactions. The extranet would be used for collaborating with businesses, suppliers etc. and for offering various services and social programs to the citizens. The preferred way of organizing the information on an extranet would be a *portal approach*, which makes information available from a one-stop shop over the Internet. The extranet would allow for information exchange as well as transactions (including payments) and complex workflow processes. The extranet applications also make use of the same databases as the intranet applications. Access control mechanisms would ensure that the external entities have access to only specified portions of the data and with specified permissions and the rest of the data and systems are secured from external users using appropriate security mechanisms.

4.3 Hierarchical Web Server Farm Model for E-Utility

The e-governance solution would employ the web server farm model. A typical web server farm model works as follows: All applications and data are resident in a large centralized center, called the web server farm, which may house a large number of powerful computers. The applications are designed using web technologies so that they can be served on web browsers to client terminals. The end users have low-end computer terminals and use a web browser to access the applications from the web server farm.

A more complex web server farm model may have multiple such farms and the various applications and databases may be distributed across these multiple farms. When a user uses an application, it sees one single application, though in reality, it may actually be navigating back and forth between these farms. Thus, applications may be created by integrating contents from various farms.

A hierarchical web server farm model is recommended for the state government. The model makes use of the natural hierarchy in state administration for distributing the data and applications across various locations. The rationale for the choice of model is discussed later in the section.

The model consists of:

1. **Small to medium servers at taluk levels, one in each taluk:** These house some of the data and applications (or application components) which are used only within a given taluk.

2. **A medium server at each of the district headquarters:** These house some of the data and applications (or application components) which have applicability only within a given district, but are used by more than one taluk within the district.

3. **A large central web server farm at the state capital**: This houses the data and applications (or application components) which have applicability across multiple districts. This may also contain most of the important statewide databases and applications that are of a centralized nature. Most of the extranet data and applications would also be typically hosted on this farm.

The hierarchical schema is depicted in figure 4.1 below.

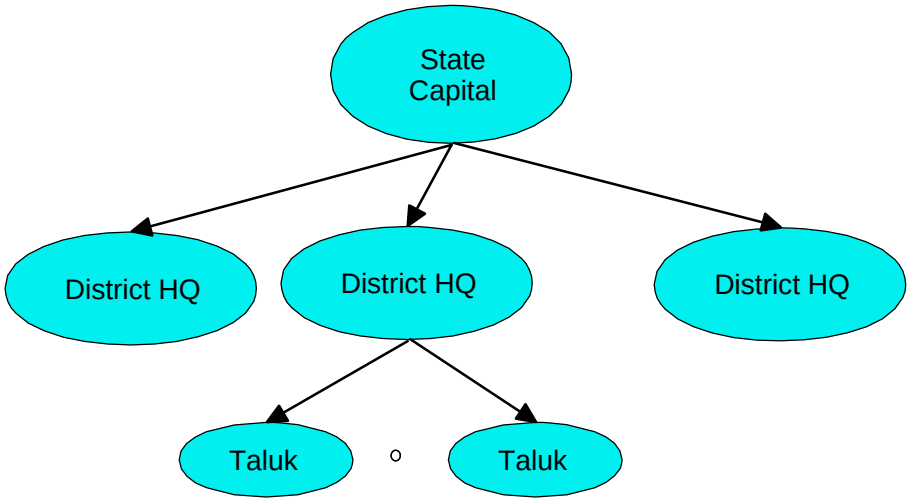

Figure 4.1: Hierarchical Organization of Web Server Farm

A useful way to look at the model is to first think of it as a simple web server farm model which has only one centralized web server farm (located at the state capital) and which houses all the data and applications. In this model, even an application within the same taluk (for example, an e-mail from an employee in a taluk to another employee in the same taluk) would require use of communication bandwidth from the end user in a taluk right upto the state capital. Therefore, to utilize network bandwidth efficiently, it makes sense to locate some applications and data at taluk and district levels so that the local tasks can be accomplished there itself without recourse to higher levels. This also goes well with the hierarchy in the administrative set up.

The implementation of this model would require that the government applications are made web enabled and the dependencies carefully resolved for deciding whether the data should reside in a taluk, a districts or the state capital. The use of distributed web server farms improves network efficiency (and thus cuts down on investment required in setting up connectivity infrastructure as well as recurring network costs) but makes the applications more complex to design and implement and therefore, a right balance should be struck.

'Utility' Model for Applications: An Illustration

An illustration will make it clear how applications would run over the web server farm model. The illustration will also show how internal and external workflow can be seamlessly built based on the requirements of the application. This mode of hosting and delivering applications is often referred to as a *utility model* due to its similarity with public utilities like electricity and water which are delivered on wire (or tap) as and when required.

Consider a simple file application. To create a certain file, an authorized person logs in into the system and uses his web browser to fetch the ‘file create’ application. A message is sent to appropriate web server farm over the network and the returned page is loaded on his browser. Here, the person creates the file by filling in the necessary details including the list of designated authorities to whom the file is to be sent for approval. Once the operation is completed, the file particulars now reside on the web server farm database. The 'next designated authority' specified in the file application now gets a message that a file is waiting his approval. He can view the file by sending a 'view file' request in his browser to web server farm. The file comes to his computer, he appends his notes and approves the file.
At any time, the status of a file can be tracked by an authorized person (which may even be an applicant citizen) by querying the file application on his browser from the web server farm. The application will typically show who have approved it and with what comments, on whose 'desk' is the file currently pending approval and so on. An advantage of the utility model is that the file creator or a file user does not have to bother about whether the file is at a particular taluk or district. The system presents the file as an abstract object and also takes care that only the authorized personnel can see or modify the file.Thus, the utility model utilizes the power of networking to create and run applications in an effective manner. It minimizes both hardware and software requirements on end user PCs and ensures consistency of information.

5. Maharashtra Solution: Details

This section discusses the technology options with the Maharashtra government for some important components of the e-governance solution and provides broad recommendations.

5.1 Connectivity Infrastructure

The goal here is to have connectivity right upto each department at the taluk level. The problem is handled at 3 levels in the following way:

1. Intra-taluk connectivity - linking various departments within the same taluk into a single Local Area Network (taluk-LAN) - where a combination of Ethernet and wireless technologies is recommended.
2. Taluk-to-district connectivity - linking the taluk-LAN for each taluk within a district with the district headquarters - where two options, namely leased lines and VSATs are spelt out along with implications of the use of either technology.
3. Inter-district connectivity - linking all the district headquarters together and with the state capital - where use of the existing VSAT infrastructure is recommended.

The technology options available for each level are discussed later in the section.

Bandwidth Requirements

A first cut estimation of bandwidth required at various levels is as follows:

1. We assume that most of the local requirements in a taluk (like local e-mail, editing, simple number crunching etc.) are met out of the taluk server itself over a LAN. This LAN needs to have adequate bandwidth to allow terminals in a taluk to work without any visible networking delays. Such a LAN with a few dozen terminals typically needs at least 1 Mbps of shared bandwidth, providing on an average 30-40 kbps per terminal.

2. The taluk-district traffic is likely to originate mainly from transaction type of data (which includes e-mail). Assume that about one-third of the terminals (about 10 out of 30) at any time are doing such transactions (the rest are doing general local work over the LAN as mentioned above). Assuming about 12 transactions per hour and about 5 page downloads per transaction (each page being on an average 10 kilobytes), this translates to about 10*12*5*10 = 6000 kilobytes per taluk per hour (or about 13.33 kilo bits per sec - kbps). If we assume that the in the complete workflow, an equal volume of the transactions are also happening through the extranet (citizens & businesses), then that adds another 13.33 kbps of traffic. Due to the hierarchical organization of databases and applications, assume that one-third of it can be served from the taluk itself without resource to the district or higher levels. This gives a receive bandwidth requirement for a taluk-district link to be (13.33+13.33)*0.67 = 17.8 kbps. If we provision approximately 2-4 times for the peak (as compared to average), then a taluk-district link should provide a bandwidth of approximately 32 - 64 kbps per taluk.

3. The capital to district traffic will comprise the data required at taluks which needs to be fetched from the state capital and the data which is required due to usage at the district offices themselves. Assume only half of the data supplied by a district to a taluk comes from the capital (rest comes from other taluks and the district server itself). Also assume that the data required from the capital at the district offices themselves is approximately the same volume as that required by all the taluks in the districts. For 10 taluks in a district, this gives an average receive bandwidth requirement from the capital to a district of about 10* (17.8 * 0.5 *2) = 178 kbps. Again provisioning higher for peak (provisioning can be lesser due to more number of users being multiplexed), we need about 256 - 512 kbps of receive bandwidth per district-capital link.

The final determination of bandwidth should consider more details on the number of end user terminals, the number and type of applications and the location of various application and data components (whether in taluk, district or capital). A more decentralized application model would require lower bandwidth for taluk-to-district and inter-district connectivity.

Intra-Taluk Connectivity (taluk-LAN)

The problem at hand is to interconnect about 30 departments located within a taluk into a LAN. The departments could be located within a few hundred meters, but may also be separated by a few kilometers.

Logical Networking Model

The logical networking model for a taluk-LAN is shown in figure 5.1 below. The end user terminals connect to the taluk-level server through a local area network. The citizens' terminals may also connect to the same taluk-level server through an extranet (Internet). The taluk-level server is connected to the district HQ through some communication technology.

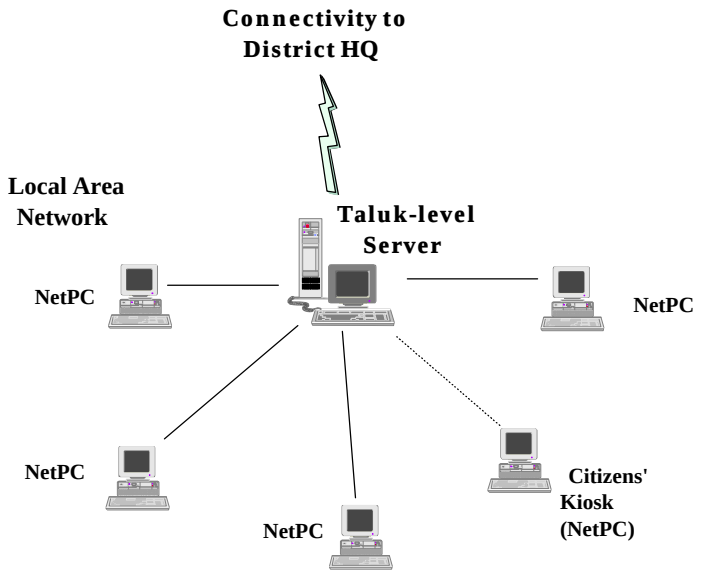

Figure 5.1: Logical Networking Model for taluk-LAN

The NetPCs use applications over the taluk-LAN through the taluk-level server. The applications and data may be coming directly from the local taluk-level server on the LAN or may actually be fetched from servers located at higher levels (district HQ or capital).

Choice of Networking Technologies

Keeping in mind that the NetPCs would have little or no resident data or applications and all of them would be served through the taluk-LAN, it would be desirable to have a reliable shared (shared across all terminals in the LAN) bandwidth of about 1 - 5 Mbps for a taluk-LAN. The following technology options are available for creating a taluk-LAN:

Technology	Remarks
1. Ethernet	Ethernet is an ideal and inexpensive technology deployed worldwide for LANs and can provide 10-100 Mbps of shared bandwidth. However, a single ethernet LAN can extend only upto about 1 km. It may be desirable to create multiple small ethernet LAN clusters and interconnect them using some other technology into a single logical LAN if the terminals in a taluk are located more than a kilometer apart.
2. Optical Fiber	High cost of laying fiber and no such high bandwidth requirements within a taluk imply that putting fiber right upto each of the 30 terminals in a taluk would not be justified, though use of fiber for connecting small segments can be considered in some cases.
3. Wireless	NIC has recently started using hub based wireless data network technology[2] (using Radio Frequency - RF) which provides shared 2 Mbps bandwidth to connecting users - users may be located upto about 12 kms. A small wireless equipment is needed at the user / LAN side to connect to the hub. This (or other similar wireless technologies) may be very useful for interconnecting either the individual NetPCs themselves or clusters of NetPCs to a hub. The hub can be located next to the taluk-level server.
4. Dial-up	Dial-up gives a bandwidth of only upto 28 kbps (max.) and the connectivity is very unreliable. This may not be suitable for a LAN type of setting, especially where the NetPCs do not have much resident data or applications and depend entirely on the connectivity to the taluk-level server.

We would recommend that for creating an intra-taluk LAN in each of the 330-odd taluks, the terminals within the same taluk and located within a small distance can be interconnected with 10/100 Mpbs Ethernet LAN which is a worldwide standard for local area networks. Where some departments are located at distances more than permissible in an ethernet LAN, clusters of such departments be formed such that all PCs within a given cluster are interconnected on an Ethernet LAN. These clusters (independent ethernet LANs) can be integrated (interconnected) into a single logical LAN using either wireless access technology as described above or using segments of optical fiber (depending on the distances involved,

[2] More information is available on website: http://www.nic.in/Informatics/archive/inf98oct/

number of such clusters to be connected and terrain). This logical LAN thus interconnects all the terminals in the taluk and the taluk-level server and also connects to the district HQ.

Taluk-to-District Connectivity

Each district has about 10 taluks. The taluk-LAN of each taluk needs to have a connectivity with the district HQ.

Logical Networking Model

The logical model of the network is similar to that for the taluk-LAN. As shown in figure 5.2 below, the district-level server is connected by communication links to each of the taluk-LANs in the district. (There can be a variant in which different taluk-LANs may be directly connected to one another and with district HQs and state capital).

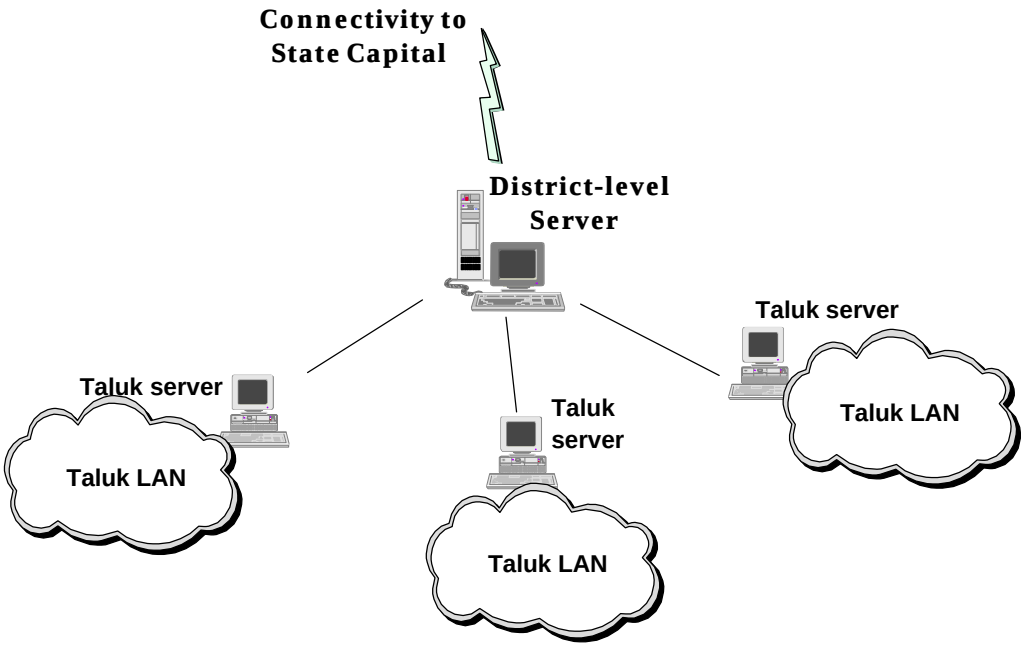

Figure 5.2: Logical Networking Model for Taluk-to-District Connectivity

Choice of Networking Technologies

The link from a taluk LAN (or equivalently taluk server) to the corresponding district HQ is a relatively low speed link (32-64 kbps per taluk would be fine) which carries only the data which needs to go out of the taluk. The following technology options are available for the purpose:

Technology	Remarks
1. Leased Line	A low capacity (32-64 kbps) leased line from the taluk to the district may be a good option, given the low bandwidth requirement, small distances to be connected (10-100 km) and high initial costs of other technologies.
2. VSAT	VSATs using many different technology variations are available and can be rapidly deployed[3]. VSATs should be considered if the government wants to use a highly centralized model for applications, where the taluks would need to directly connect to the state capital and to other taluks and the bandwidth requirements would also be large. For more decentralized application models, leased lines may be more cost-effective since the geographical distances as well as bandwidth requirements would be small in a decentralized model. A mix of leased lines and VSATs (or even VSATs alone) can also be considered if leased lines are not available.
3. Satellite + dialup hybrid	This is a hybrid technology which uses satellite broadcast for receiving data and uses a dialup / ISDN / leased line for the reverse transmission. They are much less expensive compared to VSATs but the reliability of the reverse channel can be a question mark, especially if it uses a dial-up link.
4. Optical Fiber	Fiber for all taluk-to-district links may be an expensive and time consuming proposition at the current juncture. At present, fiber may be considered only in the absence of other alternatives. At a later stage, when IT and Internet would have reached the masses, it may be economically viable, and maybe even necessary, to network all taluks using fiber technology.
5. Dial-up	Dial-up connectivity may not be appropriate due to its unreliable nature, since the connectivity of an entire taluk to the rest of the state would depend on this link.

We would recommend leased lines if the government wants to pursue a hierarchical application model as proposed earlier in the document, where applications and data reside on taluk-level, district-level and capital-level servers. If however, the leased lines are not easily available, or if the government wishes to pursue a highly centralized model, with only a single web server farm at the state capital then VSATs would represent a better option. Amongst VSAT technologies, broad band VSATs may be more suited given the web-based application model. The use of C-band or extended C-band would be better than Ku-band due to the wet climatic conditions of some parts of the state.

Inter-District Connectivity

[3] A careful review of existing VSAT infrastructure in the state should be done to identify the VSAT technologies that would integrate well with the existing network. Also, other parameters like nature of applications, bandwidth requirements, rain fade (especially for Ku band) etc. would need to be considered.

The district HQs in the state are already connected using VSATs to each other and to the state capital. The same network can be utilized for the present purpose by appropriately scaling the bandwidth as may be required. If this is not possible, deployment of more VSATs should be considered.

5.2 Web Server Farm

The logical model for a web server farm has already been presented in the earlier sections. The government would need to decide on the extent of centralization / decentralization of data and applications. As indicated earlier, the centralized approach results in simpler applications and data models but require high bandwidth right from the NetPCs in the taluks upto the state capital. On the other hand, a partly decentralized approach would place some data and applications at servers located in districts and taluks. This would significantly cut down the bandwidth requirements (and thus cut down investment required in setting up connectivity infrastructure as well as recurring costs), but would require somewhat greater effort in the application development step.

Taluk Level Server

At the taluk level, a high end PC or a low-end IBM Netfinity server should be sufficient to take care of workload of about 30 NetPCs. It would contain the applications (or application components) local to a taluk and databases (or database portions) that are applicable only in a single taluk. The NetPCs would connect to the taluk server for all their needs. The taluk server would serve the request locally if possible or else forward the request to the district HQ server if required.

District Level Server

At the district level, an IBM Netfinity server or an equivalent machine should be sufficient to handle requests from the taluks in an average district. This server contains only a few applications (or application components) and databases (or database portions) that are local to a district but are applicable across multiple taluks within the district. The main function of the district server would be to route the requests from taluks to the state capital or to other taluks within the same district.

Central Web Server Farm (State Capital)

This houses a large number of centralized applications and databases and can receive requests from taluks and districts. The alternatives for server set up for the farm comprises (a) a cluster of IBM Netfinity servers which may consist of specific servers dedicated for specific classes of applications, or (b) an IBM RS/6000 system, which could also be a super computer of the right computing power which can run all the applications.

The sizing of the computing power and the storage required for this farm would need to take into account the number and type of applications, the expected number of end users and the extent of usage by them, the extent of centralization /decentralization of the applications and databases and the levels of service required from the web server farm. The farm would be

scaleable in terms of both computing power and storage so that new applications and end users can be added in the future.

The central web server farm being the chief storehouse of data and applications, the farm solution would also contain means of providing the following services:

1. Redundancy - so that in case of failure of some hardware / software, the tasks are immediately taken over by an equivalent component.
2. Security - using firewall, intrusion detection and other optional features, especially to protect the system from people accessing the contents from extranet (Internet).
3. Back-up features - either a periodic back-up facility on a storage device or a real-time mirroring of the data on an alternate site.

The web server farm can be implemented by the government on its own or its services may be outsourced using a service model where IBM builds and maintains the server farm and charges the state government on an ongoing basis. The applications would however need to be owned by the government.

5.3 End User Hardware Equipment

The end user terminals in the taluks comprise low cost NetPCs which have the following characteristics:

1. They can connect to a network (taluk-LAN) and run applications (primarily web-based applications) over the network.
2. They have adequate compute power and RAM (memory) but little or no secondary storage (hard disk).
3. There are no stand-alone applications stored on these NetPCs. All applications must come over the network. Also, there is normally no facility for input and output of data (using floppy, CD etc.) at the user's end on the NetPC. Such facility may be made available on the taluk and district server systems on a per-use permission basis.

The use of linux-based NetPCs is recommended due to various reasons like (a) low deployment cost since linux is freeware, (b) rugged operating system requiring much less help desk support, thereby implying much lower maintenance and support costs, and (c) superior natural ability to work in a networked computing environment as compared to Windows-based systems which are more suited for stand-alone PCs.

5.4 Data Management

The objectives of data management in the present context are to ensure consistency of data (and ensure that data is not duplicated or double-entered for different applications) while simultaneously ensuring high availability and achieving high performance in access (read as well as modify).

An important component of the e-utility is the data management sub-utility which provides reliable data to a host of government applications. The data management organization defines the external interfaces to access and update data, the access control policies, policies for record retention and archiving as well as the backup and recency policies.

The major issues to be handled in the database organization are:

1. **Creation of data masters by normalization of the government data:** The inter-linkages and overlaps amongst databases for various applications would need to be documented and the resulting data mesh would have to be passed through a large normalization exercise to evolve a set of data masters (data organization having a unified data model) which would prevent duplication of data and allow a large number of applications to reuse the same databases.

2. **Directory organization:** Since many databases would be very large, they would be best organized as directory services to facilitate high performance access and scaleability. Although it would be premature to recommend a specific directory organization at this stage, the LDAP (Lightweight Directory Access Protocol) is a popular directory protocol supported by most software and hardware vendors. It allows a hierarchical name space capable of supporting large amounts of information and specifies the communication between the directory client and directory server using the directory access protocol. Having a directory organization would allow faster searches, retrieval, update, insertion and deletion in databases.

3. **Distributed versus centralized:** A centralized database organization is much simpler and easy to implement and manage. However, it may not result in efficient utilization of networking bandwidth and therefore a completely centralized model would call for significantly higher investments in the communication infrastructure. Therefore, a partially hierarchical database organization was recommended in the earlier section. The idea is to keep only that portion of data locally (at a taluk or a district) which is needed only by the local users and has no applicability outside the local domain.

4. **Handling data consistency in hierarchical web server farm model**: In a hierarchical web server farm model, data may reside at multiple levels (capital, district and taluk). Consider a database, say citizens' database, and suppose that parts of this database reside at respective taluks and the taluk data is then replicated at district and state levels. A person who *reads* the data should get the same (consistent) answer, irrespective of where the person accesses the data from. For such databases, which are distributed at more than one level (taluk / district / capital), the following model of coherence is recommended for ensuring data consistency:
   a. Any given data item is stored on at most one location within a single level. Thus, two different taluk servers cannot be the home for the same data item. A data item which needs to be used at more than one taluk must be stored at a district level server (or state level server). Similarly, a data item which needs to be used in more than one districts must only be stored at the state level server.
   b. A *read* access (access which does not involve any modification to the data) can be made from the nearest available copy of the data. Thus, a terminal in a taluk may

perform the read on the data stored on its taluk server, or on the district server or the state level server, whichever is the nearest storage location for the data item.

c. A *write* access (access which alters the stored data) is deemed to be complete only when the alteration has been successfully performed at all levels where the data is stored. The update of multiple levels can be real-time or batched, depending on the networking technologies deployed and the nature of the application.

This model achieves the task of providing a consistent view of the government databases. At the same time, it leads to efficient utilization of network bandwidth by saving on bandwidth usage for the *read* operations on data.

5. **Record Retention and Data Backup**: The legal and accounting requirements for record retention, recency and archiving should be adhered to in designing and configuring the data management system. The backup for data should be systematized, so that there is an automatic backup (either on same location) or an online mirroring on a remote site. More on this is discussed under system management section.

6. **Data Security**: The security issues are important in data management and are discussed later under system management section.

5.5 Applications and Integration Tools

The applications would be built around the various services provided by the e-utility and would provide interfaces to the various categories of end users.

It is a good practice to standardize on various aspects of application development, especially when developing a large solution which would be implemented in stages. These aspects include:

1. **Underlying programming model:** In this case, the programming model is that of an e-utility where various data and compute services and tools for building and integrating applications are offered as abstract services by the e-utility in a transparent manner over a computer network.

2. **Language and Platform:** Standardizing on implementation platform and language can be a tricky issue in large systems. The applications should be web-based (they can be accessed over a web browser requiring no other software to be installed on user terminals). Java is fast emerging as the language of choice for web application development and so is XML for messaging and inter-application data exchange. The applications should make use of open standards as far as possible to allow inter-operability.

3. **Documentation**: Standardized documentation formats are necessary for maintaining and enhancing the applications. Documentation includes (a) technical documentation for system developers, (b) configuration and administration documentation for system administrators, and (c) user manuals for end-user training.

Need for Integration Tools

Large and complex systems such as a statewide e-government solution can over a period of time have a large number of applications which may need to talk each other, or may need to be integrated into a single application. This requires the use of integration tools which perform two main tasks:

1. **Integration of heterogeneous systems** (which may have different data definitions, different platforms etc. and may have evolved in different directions) - The use of a *middleware* software[4] should be seriously considered keeping in view the large size of the system. The middleware resides between any two applications that need to talk to each other and provides for interconversion of data and message formats and also provides many value added services for integration of applications. This is particularly useful in integrating legacy systems into the mainstream or for integrating systems that may have evolved independent of one another.

2. **Creation of workflow between interfaces for complex processes:** A typical process may require collaboration between citizens, government officials and business organizations. To build applications that integrate various interfaces into a seamless workflow, a useful approach is to use application development tools[5] which facilitate modeling of complex processes and the implementation of these processes into workflow applications.

5.6 Project Implementation and Management

The implementation of the solution should be phased according to resource availability and the priorities of the state government. The implementation should start with high-benefit and low-risk services and processes.

Project Life Cycle

To manage the enormous complexity involved in developing an e-governance system, the project life cycle adopted should satisfy the following criteria:

1. Supports phased development and implementation.
2. Has mechanism to verify and validate each phase.
3. Supports iterative modifications to the system.
4. Provide flexibility to develop a phase while the previous phase is being implemented.

We propose the following model for the development planning of the system as shown in figure 5.3 below :

---

[4] For example, *MQ Series Integrator* is a popular middleware product from IBM. For more information, visit http://www.software.ibm.com.

[5] For example, *MQ Series Workflow* is a popular workflow integration product from IBM. For more information, visit http://www.software.ibm.com.

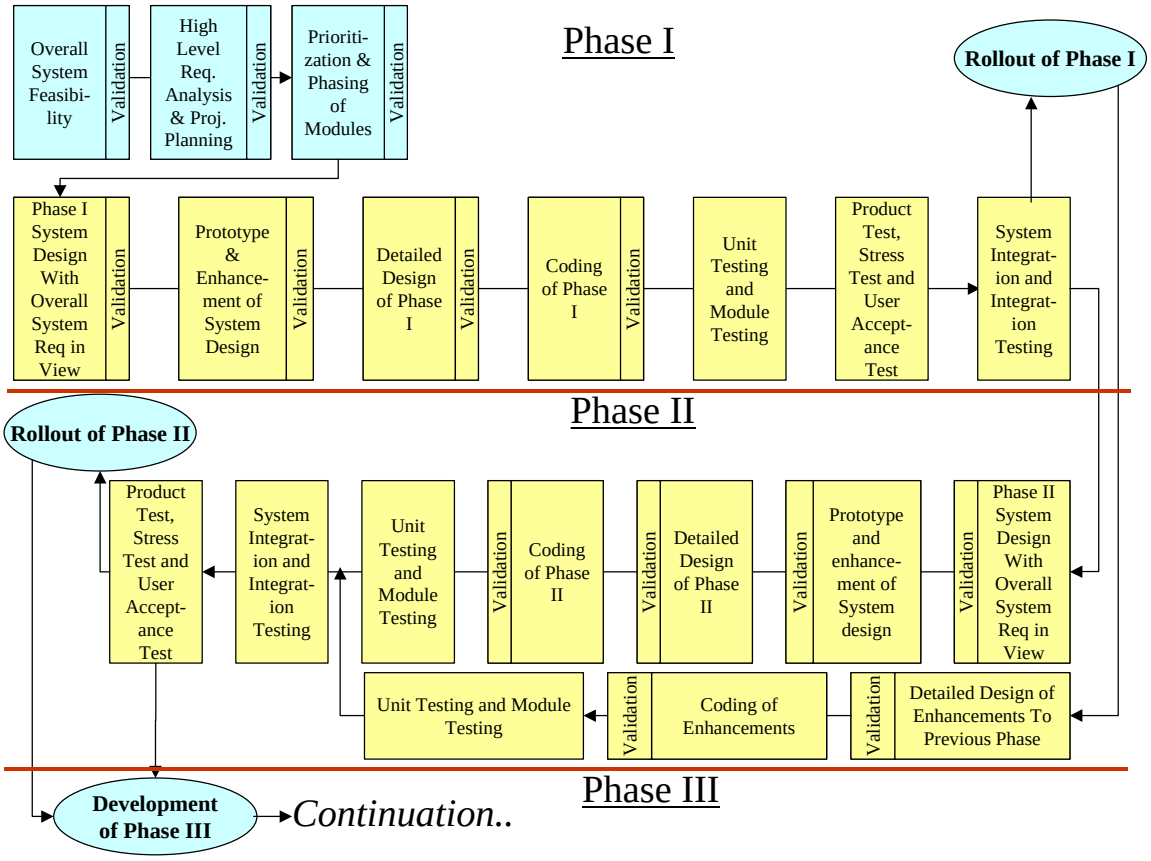

Figure 5.3 Project Development Planning

Delivery Approach

The following approach to delivery is recommended:

1. **Incremental Delivery:** The delivery is phased into *pilot and mass roll-out periods*. During the pilot phase the processes and systems are verified and, if necessary, modified. This is done in a selected few taluks. In addition, the roll-out team resources are trained for the mass roll-out period. The incremental delivery approach is aimed for minimizing the risk of the delivery and maximizing the efficiency in roll-out period.

2. **Total Delivery:** Total delivery refers to delivering the processes, systems and related training (as opposed to plain installation of the system). The delivery of process which the system supports maximizes the benefits of the delivery. The training effort ensures that both the processes and the system are adapted in location smoothly and effectively.

3. **Industrialization of Roll-out:** In order to maximize the efficiency and minimize the required calendar time, the roll-out effort needs to be industrialized through the usage of standardized working methods (roll-out methodology) and development of pre-developed roll-out deliverables.

4. **Early Benefit Delivery:** The goal of this approach is to deliver a bulk of the benefits as early in the project life cycle as possible. This can be achieved by prioritizing of functionality / applications, prioritizing of roll-out locations and roll-out organization arrangements enabling 'overlapping' roll-outs. Also, the components that have significant positive externalities (like communication infrastructure) for the other components should be developed earlier.

5. **Local Resource Utilization:** The project also should involve and exploit local resources as much as possible. The process delivery should adhere to the concept of local workshops where roll-out effort is 'kicked-off'. The processes and the work related to their adaptation are introduced in these workshops. This facilitates the local administration to agree and adopt to the changes. The local resource utilization is vital for performing the total delivery.

Early Tasks

The early phases should aim to accomplish the tasks which lead to significant positive externalities being generated for the rest of the system implementation and also the tasks which provide some immediate visible benefits. Thus, the following is recommended in the early phases:

1. Devise a complete strategy with respect to the overall approach:
   a. We recommend a comprehensive intranet (for G2G and G2E) + extranet (portal for G2C and G2B interfaces) strategy which uses a common web server farm model and a network-based utility model for applications.
   b. We also recommend a partially decentralized hierarchical approach to organizing data and applications on server farm / servers at various levels.

2. Establish connectivity infrastructure between the taluks, districts and the state capital. The basic equipment of the network should be deployed and the network should be configured to run on a low initial bandwidth for the various levels of connections. The basic equipment should be capable of scaling to higher bandwidth later on.

3. Identify of the key initial applications and implement them. The key applications should be the ones which provide maximum visible benefits and have the highest chances of success. Applications should be selected in all spheres - government's internal operations, citizen services and quality of life - related applications.

4. Provide training to the employees for using the applications.

The solution can then be enhanced on an ongoing basis in terms of (a) implementation of new applications and extending and improving earlier applications and (b) scaling the bandwidth (and also communication equipment and hardware if required) to take care of increased usage.

Crucial Success Factors

The crucial success factors for successful project implementation would be:

1. **Clarity of vision and objectives and a clear strategy:** The objectives and business case behind the solution must be clearly defined, so that a consistent message for why the new system is needed can be delivered.

2. **Commitment of senior people:** It is vital that a few key persons behind this change are identified as sponsors of the effort, and that they are committed to driving the change. These people should be known across the state-wide governance machinery and be respected senior administrators (bureaucrats and politicians) within the areas covered by the system, and they must be ready to stand up and speak for the system at all times.

3. **Participation of impacted users:** All impacted users should be involved in the development and implementation of the system. This factors in all the requirements as well as creates a sense of ownership of the system among the users. Participation helps to transform resistance to positive commitment and co-operation to the change as shown in the figure 5.4 below.

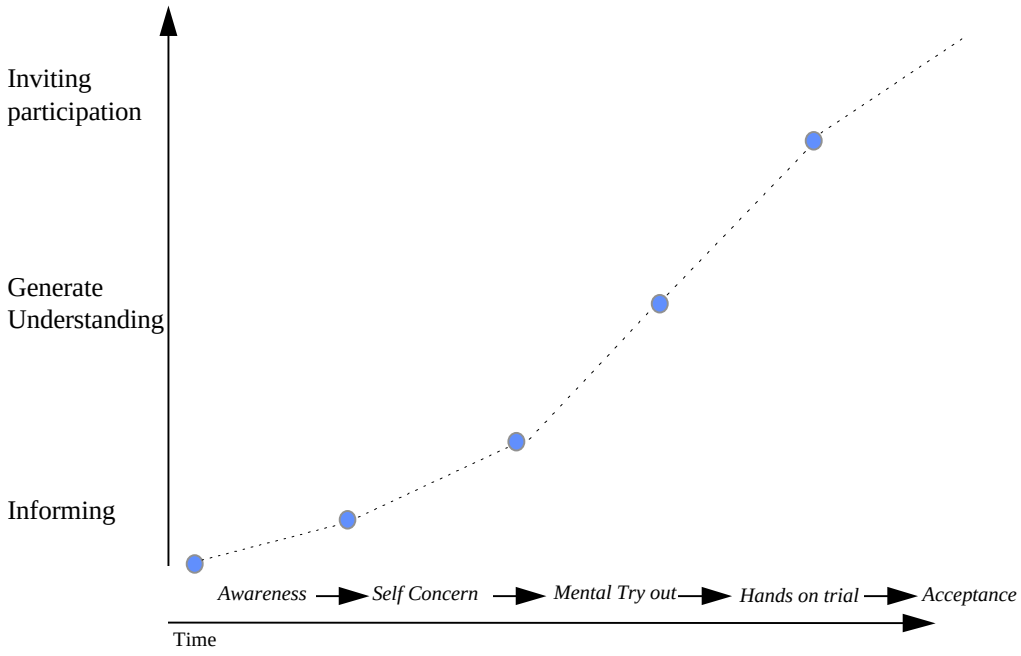

Figure 5.4 Participation and Acceptance in Change Management

6. System Management

System management includes system administration, network administration, database administration, system support and help desk and training. The hierarchical web server farm solution should have a strong team at the state capital with a single personnel at each district. The personnel have to be mobile in order to provide support to the taluks, as and when required.

1. **System Administration :** System administration would consist of user accounts management, security management, application performance monitoring and support, transaction control and workflow management and software distribution and control.
   a. **User accounts and certificate management :** A strong certificate and user accounts management system needs to be implemented. Such a system will lay down procedures and processes to manage and monitor user accounts and associated capability based certificates. It will support multiple levels of user categories and control access rights to various records based on the user category, using capability based certificates. It will monitor and report attempts of unauthorized access of any record using any login. It will also modify the access rights/ capability certificate of existing users based on their changed responsibilities. Finally, it would manage the online audit trail based on the login, password and the capability based certificate.
   b. **Security management :** Security has two aspects: application and data security and physical security.
      i. Application and data security - The solution will provide means for authentication, access control and access logs, confidentiality, intrusion detection and intrusion logs, data integrity and non-repudiation. These are provided through (a) a strong user accounts and certificate management, and (b) application level support including encryption and use of digital signatures.
      ii. Physical security involves physical monitoring and protection of the equipment and software / data. This may be achieved by housing the equipment in secure constructions, deployment of personnel etc. We propose the use of rugged hardware / software and use of remote backing up of the data which would minimize the need for deploying personnel.
   c. **Application performance monitoring and support :** System performance issues will come up from time to time due to addition of new applications and users, increased usage, accumulation of unwanted data, archiving etc. To preempt any detrimental effect on system performance, the system will be monitored through a set of parameters. This can also be automated to a certain extent. When a deterioration of the system performance is noticed, or when such a complaint is notified to the system administrator, it would be possible to take corrective measures immediately in order to prevent disruption in the functioning of the system.
   d. **Transaction control and workflow management :** Over a period of time, the process rules defining the system will change. Such change would require managing the workflow applications in the e-governance solution. Also, under exceptional circumstances, it may be required to reverse entries made in the system, which a normal user cannot reverse due to data integrity reasons.
   e. **Software distribution and control :** As the solution is rolled out and additional functionality added, it becomes critical to manage the version in different parts of the

system and in different offices. The web server farm model solves this problem to a great extent as most of the data and applications would reside at central web server farm (or at district & taluk servers) and none on user terminals.

2. **Network Administration :** It should define the acceptable service levels for uptime and performance and ensure adherence to them. It should also continuously appraise the bandwidth requirement based on added network traffic and the system performance and also prevent and monitor security breaches in the network. The use of rugged networking equipment and plug-and play technologies have been proposed in the networking solution to keep the personnel requirements low and due to the difficulties in managing such a large network spread over the entire state.

3. **Database Administration :** It will be responsible for data security and database management. Database administration will include defining and maintaining standardized data and access models and also routine functions like record management, data archiving and purging so that the size of the database remains manageable.

4. **System Support and Help desk :** The system support and help desk will provide the first level of support to the users. The web server farm model and the use of NetPCs would help to greatly reduce the help desk requirements. We would recommend that the solutions and applications be so designed (and with simple user interfaces) that the support required at taluk level be minimized (served over telephone from the district).

5. **Training :** In order to run the system smoothly, a separate organization for imparting training round the year is required. Personnel from the support and help desk can double as trainers, keeping the support manpower for the system low. Again, simple design of application interfaces are needed to ensure that the training requirements are small.

7. Advantages of Proposed Solution

1. Integrated and coordinated approach which would:
   a. Provide maximum benefit due to statewide coordination.
   b. Ensure consistency of data.
   c. Avoid duplication / wastage of effort.
   d. Make complex and interdependent applications possible.

2. Investment creates an asset which can be leveraged upon in the future:
   a. The taluk-LANs and the state wide connectivity infrastructure would be a permanent asset of the state government independent of the choice of applications.
   b. Subsequent applications can be added with minimal incremental costs.
   c. The model integrates with the Internet in a natural way.

3. Allows phased implementation with rapid initial deployment:
   a. The networking technologies proposed can be rapidly deployed.
   b. It takes less time to build central web based applications than to build and deploy independent-PC based resident software applications.
   c. The model allows the solution to start functioning with only a few initial applications and the government can keep adding new applications to the integrated system with time.

4. Low initial investment requirements to start operating:
   a. Single copy of applications and databases required (on the web server farm) as compared to multiple copies on all PCs. This cuts down software costs.
   b. Since departmental PCs need not store any softwares, they can be inexpensive NetPCs with a very low configuration. This is a significant saving since the number of such PCs is about 10,000.
   c. The web server farm can be very conveniently outsourced using a service model, freeing the government of the initial investment required on that front.

5. Low recurring costs:
   a. The web server farm being a centralized entity, outsourcing costs of maintenance, support and upgradation are much less.
      i. No requirements of maintaining and supporting / upgrading multiple copies of software on users' PCs.
      ii. NetPCs without hard disk and without external input/output devices are virtually maintenance free and do not need much technical support.
   b. Government need not deploy a large workforce for internal management and support of the system and can concentrate on its core activities.
   c. The model uses the natural hierarchy in government to optimize on the costs of network usage. It ensures fast access to most locally used applications over the high speed taluk LAN and uses wide area bandwidth only when required.