A Model for Estimating Installed First Costs of NGPN Translation in PCS Systems

Ravi Jain[*]
Applied Research
Bellcore

Abstract. Future wired networks with an ATM backbone will need to support PCS and wireless subscriber services. One of the key functions required to support PCS and wireless subscribers with non-geographic phone numbers (NGPN) will be the ability to efficiently identify which Home Location Register (HLR) database serves the subscriber. (Note that the same functionality is also needed to serve subscribers with portable phone numbers.) In a previous paper we have presented a scheme for NGPN translation based upon distributed, dynamic hashing. The scheme uses a hash function in the Visitor Location Registers (VLR) and a set of distributed Translation Servers (TS) which store the NGPN-to-HLR mapping.

In this paper we develop a simplified model for estimating the capital investment cost of deploying the NGPN translation scheme we have proposed. The model focuses on the incremental cost of providing NGPN service by estimating the Installed First Cost (IFC) of deploying TSs. The cost model assumes that PCS subscribers adopt NGPNs following the classic Fisher-Pry model of technological adoption, and takes into account user mobility, caching, the projected growth in the performance of computing servers, and projected decline in their cost, as well as inflation, the time value of money, and population growth. The cost model does not take into account operational (OA&M) or recurrent capital expenditures, or tax effects. The cost model is fully parameterized, allowing network planners to substitute appropriate values to evaluate "what-if" scenarios. We illustrate its use for an example scenario. We also point out that the cost model itself is quite generic and may be applicable to other network costing and operations planning tasks.

Subject areas: Network planning and economics, Personal communications, Network deployment and management.

1. Introduction

One of the key functions required to support PCS and wireless subscribers with non-geographic phone numbers (NGPN) in PCS systems with ATM backbones will be the ability to efficiently identify which Home Location Register (HLR) database serves the subscriber. In a previous paper [1] we have presented a scheme for NGPN translation which uses a set of distributed Translation Servers (TS). In this paper we develop a simplified model for estimating the capital investment cost of deploying the TSs. The cost model is quite generic and may be applicable to other network costing and operations planning tasks.

We begin by describing, in this section, the motivating application for this cost model and summarize the solution we have proposed. In sec. 2 we derive the cost model in detail, and in sec. 3 we present an example of its application for a hypothetical scenario. Finally, we end with some conclusions.

1.1 Background on NGPN Translation

The problem of NGPN translation is to determine the identity of the signaling network database which serves a Personal Communications Services (PCS) subscriber, when the only relevant information available is a non-geographic phone number (NGPN).

Currently, fixed telephone subscribers are assigned a geographic phone number, which contains enough information to determine how the signaling messages required to set up a call to the subscriber are to be routed through the signaling network. For proposed PCS systems, subscribers will be assigned NGPNs (e.g., 1-500-XXX-XXXX), which do not contain this information; a process called Global Title Translation (GTT) has been designed for this purpose. GTT is executed at signaling switches called Signaling Transfer Points (STPs), and essentially translates a subscriber's NGPN to the identity of the Home Location Register (HLR) database which serves that subscriber. For future PCS systems in which the wired backbone is an ATM network, however, in-band signaling will be used to set up calls, so STPs may not be used for signaling and GTT cannot be performed.

NGPN translation is required in three situations: (1) when a PCS subscriber with a NGPN crosses Registration Areas (RAs) served by different VLRs, (2) when a PCS subscriber with a NGPN is called (by a fixed or PCS subscriber), (3) in some implementations, when a PCS subscriber originates a call.

In order to keep this document self-contained, the scheme presented in [1] is briefly described as follows; see Fig. 1.

1. When any of the situations requiring NGPN translation occur, the non-geographical PN is presented to a switch; depending upon the architecture deployed, this may be a Mobile Switching Center (MSC) or Wireless Switching Center (WSC), or a Service Switching Point (SSP).
2. The switch forwards the NGPN to the VLR serving that switch. The VLR performs a *hash function* upon the binary representation of the NGPN, to obtain a value *f(NGPN)*, where *f* is the hash function. This specifies the ID of an entity called a Translation Server (TS).

---

[*] Address correspondence to: Ravi Jain, Applied Research, Bellcore, 331 Newman Springs Rd, Red Bank, NJ 07701. Phone: (908) 758-2844. Fax: (908) 758-4371. Email: rjain@bellcore.com

0-7803-3925-8/97 $10.00 ©1997 IEEE

Translation servers are entities not present in current and proposed PCS architectures; they were introduced for the purpose of NGPN translation. Note that TSs are logical entities; physically they may be databases co-located with HLRs.

3. The VLR launches a query to the TS specified by *f(NGPN)*, passing it the value *NGPN*. The TS contains a table mapping the NGPN to the ID of the HLR serving that NGPN.
4. The TS responds to the VLR by returning the HLR ID.
5. The VLR uses the HLR ID to continue with the registration, call delivery, or call origination signaling operations as usual.

The VLR can maintain a cache of NGPN translations to avoid querying the TS. Thus when presented with a NGPN for the first time, the VLR performs a hash and queries the indicated TS to obtain the ID of the serving HLR. It then stores the NGPN-to-HLR mapping for that NGPN in its cache. If presented with the same NGPN a second time, the VLR can search its cache first. If the mapping is found (a cache hit), a hash and a query to the TS is avoided; otherwise, the TS performs a hash and queries the TS as usual.

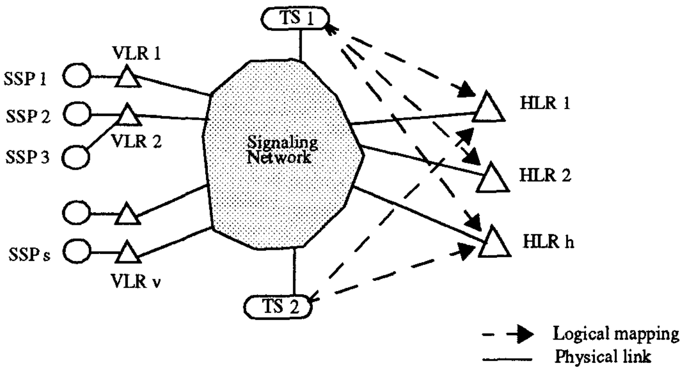

Figure 1: Architecture for NGPN Translation Scheme

2. The Cost Model

We have developed a model for estimating the capital investment costs of a phased deployment of the architecture proposed above for NGPN translation, focusing on the costs represented by the TS databases required for storing the NGPN-to-HLR mapping. The model estimates non-recurring, one-time costs, called *Installed First Costs* (IFC) [2], which typically consist of material (equipment, freight, sales taxes, etc.), installation (labor, vehicles, etc.), engineering (engineering personnel's time and related expenses) etc. The model does not consider recurring capital costs associated with long-lived equipment (e.g., capital repayment, depreciation, return on capital and income taxes), or operational, or OA&M costs.

IFC expressed as Present Worth of Expenditures (PWE). It is clear that as the number of users with NGPN increases, so will the number of TSs required to provide NGPN translation. However, typically the TSs will be deployed in a phased manner, over time, rather than all at once when the service is introduced. We will express the capital investment costs in terms of the present worth of these future expenditures, i.e., we will estimate how much capital investment these future expenditures represent today. This allows future investment costs to be accounted for in terms of today's dollars, and is called the *Present Worth of Expenditures* (PWE).

We present a brief step-by-step description of the IFC cost model.

2.1 Total translation load

We have previously [1] developed a simple analytical model of the total rate of NGPN translations in a geographical region assuming that translations are required for call origination as well as delivery and for registrations. The total rate of NGPN translations is expressed as:

$$T = 2pn\left(c(1-h) + \frac{2mv}{\pi s}\right)$$

where Total number of users = n, Penetration, i.e., fraction of users which are NGPN users = p, Rate of calls originated by users = Rate of calls terminated at users = c calls/sec., Fraction of NGPN users which are mobile at any given time = m, Speed of mobile users = v, Each registration area is a square of size = s, Cache hit ratio, i.e., fraction of translation requests for which correct information is cached = h.

For the purposes of this derivation, we will assume that many of the variables above are fixed. Thus we write T in terms of a function, α, of the cache hit ratio h and fraction of mobile NGPN users, m.

$$T = pn\alpha(h, m)$$

2.2 Service adoption model

To estimate how the workload increases with time, we employ a commonly-used model of technology substitution, called the Fisher-Pry model (or Pearl model) [3]:

$$p(t) = \frac{pmax}{1 + e^{-B(t-A)}}$$

where $pmax$ is the maximum penetration of this service, $p(t)$ is the penetration at time t, and A and B are the Fisher-Pry parameters. The parameter A specifies the number of years after which 50% penetration will be reached, while the parameter B

specifies the rapidity of the penetration. Thus the variation of the total workload with time is:

$$T(t) = \frac{n\alpha(m,h)\,pmax}{1+e^{-B(t-A)}}$$

It is assumed that after the penetration *pmax* is reached, further increases in the total workload will be primarily due to an increase in the population, *n*; these increases are considered to be part of the cost of operations and (e.g. upgrades), rather than capital IFC, and so are not considered further in the model.

2.3 Machine performance and deployment model

It is assumed that TS database machines will be deployed in a phased manner as the total transaction workload grows. However, from the trends in the computer industry, it is clear that performance of any given class of machines increases from year to year, so that machines deployed in future will have a greater relative capacity to meet the workload. For the period 1966-1990, for example, the performance of minicomputers has grown at a rate of 20% per annum, while for the period 1979-1990, the performance of microprocessors has grown at 35% per annum [4]. These performance increases are primarily due to performance increases in the underlying technology used to implement the machines.

We assume the performance $q(t)$ of a machine of a certain class at time t is assumed to be:

$$q(t) = q(1+g)^t$$

where performance of a machine initially, i.e., at the time deployment is started = q, and Rate of performance growth = g.

We need to model how system engineers will chose to add machines in the future. The assumptions we make are:

- A certain class of machines will be chosen to provide the service (e.g., minicomputers). To avoid problems of interoperability and maintenance of heterogeneous machines, all future machines will be of the same class. Thus, for example, if the workload increases, the new machines chosen would not be of an entirely new class (e.g., supercomputers).
- Initially, a single machine will be deployed, and will be used as long as it continues to meet the performance objectives (delay, throughput etc.) of the workload. At that point, a new machine, of the same class, will be added to handle the increased workload. Both machines will continue to be used until they collectively can no longer handle the increased workload. Thus new machines will be added only as required, and old machines will continue to be kept in service.

Thus at time $t = 0$, a machine of performance q will be deployed. This machine will suffice until some time t_1, at which point the workload $T(t_1)$ will exceed q. A machine of a similar class will now be deployed; however, due to the performance growth of machines, the new machine will have performance

$$q(1) = q(1+g)^{t_1}$$

This process will be continued until the maximum transaction workload is reached, which occurs when penetration reaches *pmax*. Thus it is possible to calculate the time instances ($t = 0, t_1, t_2, \ldots, t_k$) at which new machines must be deployed.

In a sense, this phased deployment model is a conservative model, in that the risks of converting from one class of machines to another are avoided, and immediate cash flows at each point of time are minimized. Obviously, in any real situation deployment will be driven by a host of technical and non-technical factors.

2.4 Machine costs

We assume that a machine bought in the future will have all three factors effecting its PWE: the time value of money at some interest rate r, inflation at some rate f, and a quoted price decline d. The third factor reflects the fact that the quoted prices of machines often decline with time; note that these are declines in the *quoted* price, distinct from the effects of inflation or time value of money. The reasons for these price declines include the effect of high-volume production (sometimes referred to as the "learning curve" effect; e.g. see [5]), competition, etc.

Thus the PWE of a machine which costs w today will be, if deployed at time t,

$$w(t) = \frac{w(1-d)^t(1+f)^t}{(1+r)^t} = \frac{w}{(1+R)^t}$$

where R is a "composite discount rate" which accounts for all three component factors, and can be calculated once they are given.

2.5 Population growth

For some services, population growth may simply be equal to the rate of the growth of the general population; however, the service may be targeted at certain regions of the country, or certain age groups, or certain groups (e.g., businesses) which may have a significantly greater growth rate.If the population at the start of the deployment is n, the population at time t, given a growth rate of *ngrow*, will be

$$n(t) = n(1+ngrow)^t$$

3. Example Calculations with the Model

The model has been implemented as a computer program. We now demonstrate its use for a hypothetical scenario; we will calculate the PWE, and illustrate how it varies with various parameters. We stress that this example is intended *for* *illustration purposes only* and does *not* necessarily indicate the deployment costs of TSs for any real situation; obviously, for any real or considered deployment numerous cost factors and variables pertaining to the situation at hand must be considered, some of which are not captured in this model.

3.1 Scenario

Consider a hypothetical geographical region with the following statistics. The total population of this area is 19.55 million, of which 77%, i.e., $n = 15$ million people, are over 18 years of age. We assume that users are moving at pedestrian speeds, i.e., $v = 5$ km/hr. During the busy hour, the rate of call originations and deliveries is 1.4 calls/hour, i.e., $c = 1.4/3600$ calls/sec.We assume that each registration area is a square of side $s = 5$ km. We also assume a maximum penetration of $pmax = 25\%$ and inflation rate, $f = 3\%$ For any given example, these parameters can be modified and the model calculation performed again.

We have assumed a baseline set of parameter values, and then varied each individually (See Table 1 for baseline and varied parameters).

Variables and Parameters	Symbol	Baseline value
1. Fixed parameters:		
Current population	n	15 million
Mobile (pedestrian) velocity	v	5 km/hr
Busy hour call origination and termination rate	c	1.4 calls/ hour
Length of the side of a Registration Area	s	5 km
Maximum service penetration	pmax	25%
Inflation rate	f	3%
3. Baseline parameters (varied as needed):		
Fisher-Pry: years to reach 50% penetration	A	2 years
Fisher-Pry: rapidity of service adoption	B	3.0
Fraction of mobile NGPN users	m	10%
Cache hit ratio	h	0.0
Current performance of chosen machine class	q	500 trans/sec
Rate of machine performance growth	g	10% per yr
Current quote for installed cost of machine	w	$200,000
Rate of decline of quoted price	d	10% per yr
Interest rate	r	10% per yr

Table 1: Summary of assumed fixed and baseline parameters for model in the example scenario

3.2 Impact of technology adoption rates

We first examine the effect of modifying the Fisher-Pry parameters on the PWE. In Fig. 5 we plot the PWE, assuming all the baseline parameters given in Table 1, except that A and B are varied as shown. It can be seen that as service adoption becomes slower (i.e., the time to penetrate 50% increases) the PWE of the phased deployment decreases. This is because machine deployments are pushed out further into the future, and since the composite discount rate, $R > 0$, these result in a decrease in the amount of capital that must be set aside for future IFC. The curves are slightly concave.

Similarly, as B, the rate of service adoption, increases, PWE decreases. However, note that for $B > 1$, the PWE is relatively insensitive to B, for any given choice of A in this range. This is to be expected, since in this hypothetical example, the 50% penetration point is assumed to be reached very soon (2 - 4 years); thus there is limited impact fast adoption can have.

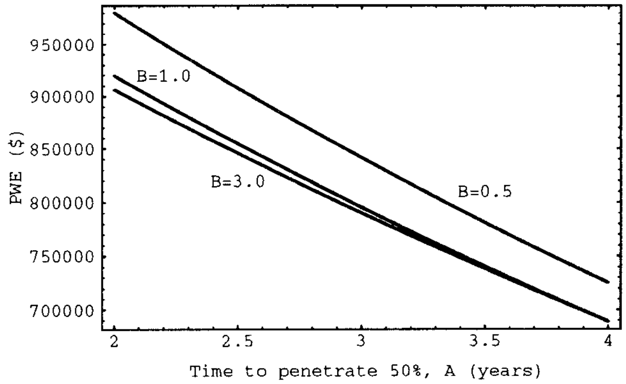

Figure 5: Impact of rate of service adoption on IFC

3.3 Impact of machine performance

In Fig. 6 we fix all the parameters to the values shown in Table 1, except for the q, the current performance of the chosen machine class, and its growth rate, g.

It can be seen that as machine performance increases at a faster rate, the PWE of the cost decreases; this is because in future, fewer machines are required to satisfy the workload. This can be seen also in terms of the "waterfall" nature of these curves. As g increases, PWE decreases only slightly until a threshold value, at which a sharp drop takes place: this represents the point at which an entire machine becomes unnecessary.

As q, the baseline machine performance, decreases, several interesting effects can be seen. Firstly, since we have held the cost of the baseline machine constant (at $w = \$200{,}000$), obviously machines of lower performance will result in higher cost, and higher PWE. In addition, as the machine growth rate is increased, the PWE curve for the lower-performance machine shows more "falls"; this is because machines are deployed in

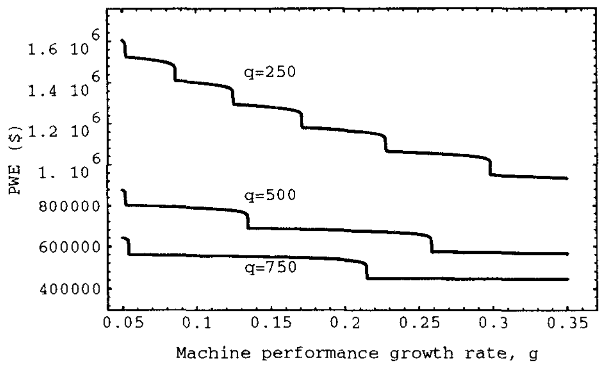

Figure 6. Impact of machine characteristics on IFC

smaller performance increments, so there are more opportunities where deploying an entire machine becomes unnecessary.

3.4 Impact of user behavior

In Fig. 7 we evaluate the model for the situation in which all baseline parameters specified in Table 1 are used, except for the cache hit ratio, *h*, and the percentage of NGPN users which are mobile, *m*. These two parameters represent calling and mobility characteristics of the users.

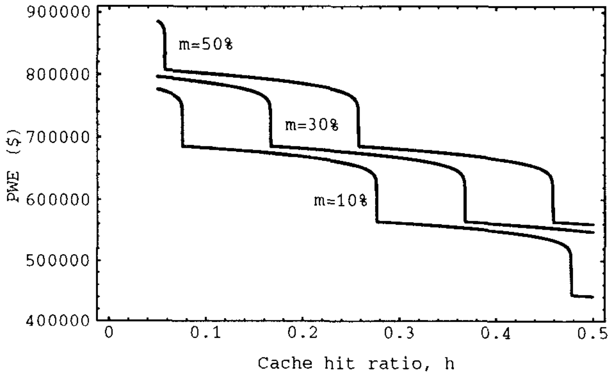

Figure 7: Impact of user characteristics on IFC

In particular, the cache hit ratio is a measure of how frequently the NGPN-to-HLR mapping information stored in the VLR's cache is present and correct. The information stored in the VLR will either not be present or be incorrect in several situations, e.g. (1) the NGPN user has not registered at a Registration Area served by that VLR (at least, recently), or (2) no call has been placed to the NGPN user from that VLR (at least, recently), or (3) the information is out of date because the NGPN user has changed service providers. Clearly, it is difficult to predict cache hit ratios accurately, since it is highly dependent not only upon user calling and mobility patterns but the size of the cache used at each VLR. We have previously studied caching in detail [6], and found that it can reduce the number of HLR accesses (and hence NGPN translations) by 30-50% under certain assumptions [7].

Figure 7 shows that as the hit ratio increases, the PWE decreases, once again in a waterfall fashion. As the hit ratio increases, more and more translation requests are satisfied at the VLRs, and the total translation load falls linearly; at some point, it falls enough that an entire machine is not needed. Figure 7 also shows the effect of subscriber mobility.

4. Conclusions

We have developed a model for estimating the Installed First Costs (IFC) of capital investment required for a phased deployment of NGPN services. The model takes into account numerous factors, including subscriber calling and mobility patterns, cache hit ratio, cost and performance of machines as well the growth rates of these values, inflation, the interest rate, and the decline in the quoted price of machines with time. (Due to lack of space we have not presented the results on the effect of varying interest rate and machine cost decline.) The model is implemented as a computer program which allows the network planner to examine "what-if" scenarios.

Acknowledgments. Many thanks to Martin Eiger for useful discussions on the development of the cost model, and Li Fung Chang for her support and encouragement of this work.

References

[1] R. Jain, S. Rajagopalan, and L. F. Chang, "Phone Number Portability for PCS Systems with ATM Backbones Using Distributed Dynamic Hashing," *IEEE Journal on Selected Areas of Comm.*, Feb. 1997.

[2] AT&T (Construction Plans Dept.), *Engineering Economy*, 3rd ed., McGraw-Hill, 1977.

[3] J. Martino, *Technological Forecasting for Decision Making*, 3rd ed., McGraw-Hill, 1993.

[4] D. Patterson and J. Hennessy, *Computer Architecture: A Quantitative Approach*, Morgan Kaufmann, 1990.

[5] M. Eiger and C. Harrington, "Present Value of Expenditures for Phased Fiber in the Loop Deployment", pp. 1834-1839, *IEEE Globecom*, 1991.

[6] R. Jain, Y.-B. Lin, C. Lo and S. Mohan, "A caching strategy to reduce network impacts of PCS", *IEEE Journal of Selected Areas in Comm.*, Oct. 1994.

[7] H. Harjono, R. Jain and S. Mohan, "Analysis and simulation of a cache-based auxiliary location strategy for PCS", *IEEE Conf. Networks for Pers. Comm. (NPC '94)*, 1994.