Comm_002

A Proposed Architecture and Access Protocol Outline for the IEEE 802.11 Radio LAN Standards Part II

Jonathon Y.C. Cheah
Hughes Network Systems
10790 Roselle St.
San Diego CA92121
March 1991

Abstract:

This paper deals with the issues confronting the radio LAN connectivity and proposes a network architecture and access protocol that will satisfy most of the user requirements. The architecture and the protocol attempt to address the need of a wide cross section of applications which require wireless connections. The goal is to accommodate as completely as possible of the common usage of this medium, without penalizing other specialized applications. The solution should provide high end products to cater for applications areas where propagation environment is hostile and the data integrity requirements are stringent as well as low cost, low power devices to service portable computers. All products compliance to the standards should be inter-operable. In addition, the architecture and the access protocol are also designed to side step other more troublesome aspects of the radio communications. Problems such as site survey for propagation characteristics modeling or coordinated redundancy protection schemes. To be commercially viable, these skilled labour intensive installation procedure will have to be avoided.

I Introduction

The use of radio LAN products operating under FCC Part 15 rules has been wide spread for a number years. The advent of the FCC Part 15 rules on spread spectrum has undoubtedly prompted a further increase in interest at the market place for these products. However, irrespective of the desirability of the radio LAN concept, the standards making process is hampered by the needs of various applications which have diverse requirements and interests. The requirements of a radio LAN connected shopping mall cash register has a different set of priorities compare with that of a crane driver's computer in a steel mill over a blast furnace. Likewise, a lap-top computer would have yet a different set of constraints with respect to its connectivity requirements.

The well accustomed performance niceties of LAN standards in IEEE Project 802 and that of others in the past ten years has presented a pre-conceived expectation for the wireless LAN standards in its functional and performance requirements. It is difficult to change this expectation although wireless LAN has the inherent handicap of a shared universal "ether" medium in comparison with a benign individual cable medium. Nevertheless, the standards will have to fulfil what the expectation dictates.

An unsuccessful attempt was made by the IEEE802.4L group to implement the Token Bus protocol within the confine of the indoor radio communications properties. Much work has been expended by the IEEE802.4L group before its demise. It is clear that the challenges viewed through IEEE 802.4 standards will differ to some extent with respect to IEEE 802.11 standards. Nevertheless, challenges in adapting a shared medium with the expectations of the accustomed cable LAN functionalities remain the same.

II The Standards' Requirements and Problem Issues

There are basically two sets of requirements IEEE 802.11 will need to meet. The first is the basic functional requirements that are dictated by the IEEE 802 standards performance guide lines. The second, a set of perceived user requirements that would determine the usefulness of the standards in meeting market demands in the wireless applications environment.

The performance and the functional requirements as declared in the IEEE 802.11 standards Project Authorization Request (PAR) stated an exception to the IEEE P802 functional requirements:

> The proposed standard will meet all of the 802 Functional Requirements except that the probability that a MAC Service Data Unit (MSDU) reported at the MAC service interface contains an undetected error, due to operation of the conveying MAC and Physical Layer entities, shall be less than 5×10^{-14} per octet of MSDU length and the MSDU loss rate will be less than 4×10^{-5} for MSDU of 512 octets, in a minimally conformant network.
>
> A minimally conformant IEEE 802.11 network will meet these requirements over a minimally conformant radio service area. IEEE 802.11 will define standard approaches

to allow minimally conformant system to be enhanced to achieve full 802 functional requirements over the radio service area

Table 1. provides a list [2, 3, 4, 5] of the necessary user needs that the standards should meet in order to assure its usefulness. The list stated in Table 1. was summarized from a large number of user submissions to the IEEE 802.4L and the current IEEE 802.11 groups. It is easily to realize that the viable wireless LAN architecture and protocol will need to satisfy as much of the these needs and requirements as possible.

Moreover, it is also important to support transparent interoperability among all conformant systems without penalizing the any specific group of users in unnecessary cost or complexity or functionality overheads than they required.

The problem of propagation characteristics and interference profile of various frequency bands and environment scenario has also been examined extensively[3]. Given the disparity in these parameters, it would point to the solution that extensive site survey procedure may be necessary prior to a network installation. This is certainly an unacceptable proposition in a great number of applications. Although the choice of a suitable modulation/demodulation technique has a great deal of potential in minimizing this problem, cost and the uncertainty in frequency/bandwidth assignments would make sole reliance on this solution less attractive. Thus, the burden of mitigating this problem will have to be shared by the architecture and protocol of the network.

The cost issue would also dictate a minimum transmission hardware overhead. This consideration points towards a hardware platform where transmitter and receiver can share major part of the RF circuitry. The use of a single-frequency half-duplex design would appear to be optimum. Based on this hardware platform concept, the following sections will attempt to introduce a suitable architecture and protocol for use in wireless LAN.

III Synchronous Network with Slotted Aloha DAMA

For ease of discussion, it is tentatively assumed that Direct Sequence Spread Spectrum (DSSS) modulation technique is used. This modulation technique is one of the two that are prescribed by FCC Part 15 with much increased emission power level and signal bandwidth in a few specific ISM bands. It will be clear that this modulation technique is not absolutely necessary as far as the multiple access protocol scheme presented here is concerned. Similarly, it is also convenient to assume that a set of suitable orthogonal codes is available for code division isolation purposes. This isolation mechanism also has its equivalent counterparts in other modulation techniques.

(a) Basic Service Area

The basic Synchronous Network with Slotted Aloha Demand Assignment Multiple Access scheme (SN-SADAMA) begins with a Basic Service Area (BSA). A BSA is defined as an area that can be serviced by a Head-End Controller (HEC) only. The area is analogous to a "cell" in the cellular telephony context. However, the "cell" concept differs from BSA by the desire and practicality to keep the BSA as small as possible, and thereby simplify the radio propagation problems. This aspect will be discussed later.

A HEC is an autonomous controller with only two distinctive interfaces; Forward Communications Interface (FCI) and the Reverse Communications Interface (RCI). FCI interfaces with the physical layer of the radio propagation medium, and RCI interfaces with a point to multi-point back-bone connection medium, either wired or wireless. For this discussion, the RCI connectivity will be assumed to be provided by a suitable high speed back-bone LAN or LAN/WAN network.

In the FCI direction, when a HEC is powered up initially, it listens for transmissions using one code at a time within its code table. In the event it receives a legible transmissions, it would acquire synchronization with the received transmission source, otherwise it proceeds to operate with its own clock frequency. It would select a code from its code table that is not used by the transmissions in its surroundings. The fault condition when synchronization errors exist among transmissions received will be dealt with in a later section.

Figure 1 shows a flow chart that depicts the basic HEC operation. The HEC transmits in predetermined time slots. The time slot boundary is synchronous among all HECs within the reception distance. The time slot is marked with code marker with excellent autocorrelation property such as a Barker code.

After bit synchronization operation with its surrounding is achieved, a HEC begins with transmission in all time slots except three. They are the ALOHA slot [ALOHA][1], the Remote unit Acknowledgement slot [RACK] and the Synchronization slot [SYNC]. The absolute slot timings of [ALOHA] and [RACK] are the same among all the HECs within the reception distance. The [SYNC] slot is randomly assigned by a HEC within each HEC Transmission Frame (HTF). The structure of a HTF is shown in Figure 2. Apart from the three overhead slots mentioned above in which a HEC listens to, there is only one overhead slot HEC transmits in independently. This is the Poll and Assign SloT [PAST]. The [PAST] slot can be viewed as the beginning of a HTF. Like [ALOHA] and [RACK], its timing slot location is also fixed universally among all HECs. A HEC uses this slot to inform all the Remote Units (RUN) the slot assignments of the oncoming HTF and polling the successful ALOHA sender and receiver(s) pairs. The structure of [PAST] slot is arranged so that ongoing time slot assignment is placed in the beginning of the time slot, and the successful ALOHA access polling will be placed at the end of the time slot. This is to allow more time for the RUNs accessing [ALOHA] slot more time to settle down in the receiving mode.

To increase the chances of successful ALOHA access, the length of the [ALOHA] slot can be increased and divided into sub [ALOHA] slots, so that each sub-slot is a legitimate access opportunity. However, the bandwidth efficiency is affected directly by the [ALOHA] slot overhead.

The [SYNC] slot is a time slot to enable the LAN coverage area to extend beyond a single BSA. This time slot provides a means that allows the coverage area to remain small. Thus,

[1] Throughout this paper, parentheses are reserved for introducing acronyms and square braces are used for time slot identifications

harsh propagation environment can be mitigated by another degree of freedom, that is, by limiting antenna illumination area. In this way, contiguous coverage in such an environment remains possible without the necessity of site survey preparation. Unlike the [ALOHA], [RACK] and [PAST] slots which are universally fixed in all BSAs through synchronization, [SYNC] slot timing position is selected in each HTF randomly and the timing information is broadcast in [PAST] for every frame. HEC will select the timing slot m randomly , where $\{m : m \in (G \cap \overline{P})\}$. $P = \{[RACK], [ALOHA], [PAST]\}$ and G is all the time slots available.

A HEC maintains synchronization with all other neighboring HEC transmissions through the [SYNC] slot. All transmission ceases during the [SYNC] slot within a BSA, this is to allow the HEC and all the RUNs within a BSA to listen to all other neighboring BSAs' transmissions. Each [SYNC] slot will be listened to by a single code within the HECs' and the RUNs' code tables. Synchronization of the code table is not necessary or needed. The code is selected in a cyclic manner through the code table one at a time except the code that is being used in the BSA. The [SYNC] slot serves a number of purposes. Both HEC and RUNs maintained bit and time slot synchronization with their counterparts in all other BSAs. A HEC will glean transmission power information from its neighbours, and thus allow it to assess its contribution to the spectrum noise level. A RUN would assess the power received from all the codes to determine which BSA it should be in by comparing the strongest signals with that from its current BSA. This operation will be discussed in the next section.

A RUN has an identical layered architecture as a HEC. This is absolutely necessary to meet the peer to peer communications and the no-single-point-of-failure requirements, as shown in Figure 3. When a RUN has a filled data buffer, it transmits a network access request during the [ALOHA] slot. A CSMA/CA or /CD protocol may apply here for efficiency but this is not absolutely necessary. It then listens during the [PAST] slot for its request to be broadcast. If it is successful in accessing the network, it would expect to get time slot assignments in later [PAST]s within a preset time out number of frames. If its network access attempt is not acknowledged during the [PAST] it would select a random number n where $\{n : n \in S\}$ and re-try at n^{th} [ALOHA] slot. The size of S is constrained by the maximum network access delay time specification. The transmission operation timing chart of a RUN is depicted in Figure 5.

After a connection assignment is allotted by the HEC between a requesting RUN and a receiving RUN (or RUNs) the HEC simply re-broadcasts the RUN's transmissions according to the HEC's assignment during [PAST] as shown in Figure 2. In this way, a virtual bit pipe is provided, and the various ARQ[1] strategies are now possible. A RUN Transmission Frame (RTF) is shown in Figure 4 and its operation flow chart is shown Figure 4.

(b) Extended Basic Service Area

It can be seen that all BSAs that is in the suitable reception distance of one another will be synchronized in bit timing to within a tolerance dictated by the propagation delay. This contiguous coverage, albeit separated by code isolation, is called the Extended Basic Area

(EBSA).

There are two basic EBSA operations that involve crossing a BSA boundary. They are:

- A RUN trying to get in touch with another RUN in different BSA (analogous to "roaming" in cellular telephony).
- A RUN crossing the BSA coverage while in communication (analogous to "hand-off" in cellular telephony).

When a RUN $\hbar_1$ places a request to contact another RUN $\hbar_2$ during the [ALOHA] slot, the [PAST] slot in the serving area of $\hbar_1$ polls $\hbar_2$ as usual. If $\hbar_2$ does not respond during the assigned [RACK], HEC will relay the request through the RCI back-bone connection to all the member HECs linked by the RCI net. This cable LAN that RCIs share is assumed to be broadband and has bandwidth capacity that will support the desirable response time requirement. In receiving a request from its RCI, a HEC will respond to the request in the [PAST] slot in the same manner as it would with a successful [ALOHA] in its FCI operation. If $\hbar_2$ responds in BSA_2 [RACK], its response will be relayed back through RCI network back to BSA_1. The EBSA operation HTF and RTF are shown in Figure 6. HEC_1 will assign a time slot to $\hbar_1$ without repeating the transmission of $\hbar_1$, while HEC_1 repeats the data packets of $\hbar_1$ through its RCI port. It also monitors the $\hbar_2$ reception ACK and relay that back to HEC_1 for its HTF slots. In this way, the spectrum for the forward and reverse communications is equitably shared between BSA_1 and BSA_2 via the RCI network.

It is important to note that RCI operation time delay can easily be accommodated within the HTF structure if indeed the delay would require the ACK to appear in later than desired time slot.

By following the operation above, it is now easy to examine the problem of a situation where $\hbar_1$ is in motion. Assuming $\hbar_1$ is moving into BSA_2 while it is in communication with $\hbar_2$. $\hbar_1$ has during [SYNC] slot determined that the signal quality of BSA_2 has greatly improved over its serving BSA signal. When a preset decision threshold is reached, in its transmit signal, $\hbar_1$ presents a request to freeze slot assignment. The decision threshold should have suitable inherent hysteresis. After HEC acknowledges this request, HEC will not alter the slot assignment for $\hbar_1$ in the subsequent frame. $\hbar_1$ will proceed to change code and place a connection request in BSA_2 [ALOHA] and monitors the respective [RACK]. With a successful access into BSA_2, $\hbar_1$ will then terminate the time slot it is allocated in its serving BSA. Figure 6a depicts such an operation.

(C) Extended Service Area

An Extended Service Area (ESA) is defined as a service area which does not share a contiguous coverage with EBSA either by geographical separation or by other physical differences such as modulation or bit rate. However, the service area is connected to all the BSAs through the RCI back-bone LAN/WAN net. This definition also implies that no RUN can move through ESA without loss of service or physical reconfiguration. Apart from this definition and the physical incompatibility, the communication connectivity concept is essentially identical to EBSA. In the situation where there is a bit rate difference. The burden of adjusting for transmission message package sizes will be borne by the back-bone RCI connection. It is easy to see that such bit rate differences is transparent to SD-SADAMA.

The concept of ESA also allows a single node to access multiple PHYs. So long as each individual PHY is connected to the same RCI back-bone LAN. It is easy to envisage that the same RUN accessing both Infra-Red (IR) and radio PHYs.

(D) Peer to Peer Communications

As shown in Figure 4, when a RUN is powered up, it would attempt to synchronized to a HEC. In the absence of a HEC, a RUN has all necessary capability to act as a pseudo-HEC except without the RCI communication support. In place of the RCI, the resident computer will substitute its own data traffic. This action is designed to simplify the complexity in the standards that would otherwise necessary to support independent peer to peer connectivity. The peer to peer communication in the absence of a HEC is perceived as an important capability in the portable market. In this market, the power drain is a vital factor. Thus it is essential that a RUN can be inhabited from volunteering as a pseudo-HEC, and when it serves as a pseudo-HEC and thus creating its own BSA, it will cease its operation as a pseudo-HEC when there is no traffic within a preset number of frames. A RUN serving as a pseudo-HEC will also identify itself as such in its [PAST] slot, and all its transmission time slots by setting a pseudo-HEC bit in the control bit field. This is to ensure that it will have an automatic migration of traffic out of its BSA, as soon as a regular HEC becomes operational. There is one exception. The RUN that has data traffic with a pseudo-HEC will be prevented from migration by being deny the "Assignment Freeze" request. This is to ascertain the integrity of the ongoing data traffic with the pseudo-HEC user port and to avoid another added protocol complexity of a pseudo-HEC having to migrate to a regular HEC while maintaining its own BSA.

Thus, in the event a regular HEC becomes available, all the active RUNs in a BSA serviced by a volunteer RUN, will recognize this fact in the [SYNC] operation, and will then migrate to the new BSA as soon as possible through the new BSA's ALOHA process.

All the functions of the peer to peer communications can be easily implemented through a configuration management table strategy so that a particular peer to peer communication session can be tailored to user's requirements.

Figure 7 shows the operation involving peer to peer communications.

(E) Automatic Redundancy Support

After a HEC powers up, it listens for all other HECs in existence before it determines its own operation code, and output power. However, if there is a very strong signal emanating from another HEC exceeding a predetermined threshold, the HEC will not begin operation and remains in standby mode. This is shown in Figure 1. There is no reason for a HEC to set up its own BSA if the coverage area is being served. A HEC in standby mode will monitor all the traffic assignments of the signal that violates the preset threshold. If the strong HEC signal ceases, the standby HEC will begin operate in exactly the same manner as the HEC signal did beginning at the next [PAST] slot, and maintain all previous assignments. The failure detection by the standby HEC can be assumed to be secure at all time, since the standby HEC has been receiving the operating HEC signal at over threshold level. The redundancy standby will be able to accept all its BSA traffic without interruption, but will terminate all EBSA and ESA traffic as it has no RCI connection information. all the terminated traffic will have to undergo a new connection request cycle through [ALOHA].

The trade off of the termination of services brings forth a important implementation flexibility. Any HEC can connected at any point within a EBSA, and the HEC will automatically serve as a backup if it deems itself not useful within operating environment at the time, i.e. there is already a better signal serving the area. The HEC is autonomous and thus lessens the human factor in the installation process. Moreover, this scheme eliminates the complexity of redundancy configuration, which in most cases required skilled human intervention.

Redundancy protection for RUNs can be achieved by one or more standby RUNs in parallel connection with the operating RUN. There are a number of redundancy protection strategies possible within this architecture and protocol scheme.

- A standby RUN is paralleled with the operating RUN. Its shadows every function of the operating RUN except it does not activate its transmitter. A mutually exclusive transmitter switch shared by the main and the standby RUNs determines the operating status of the RUNs.
- A standby RUN maintains time slot assignments through the next BSA with the second strongest signal. It keeps maintenance of a secondary connection at all times. When the main RUN fails, or in this case, the operating BSA fails, the standby RUN will the bear the data traffic in the next subsequent frame. However, without adding additional complexity to the protocol, this redundancy protection is only possible if both ends of the data traffic have standby RUNs to maintain secondary connections.
- Tiers of redundancy protection are possible that are transparent to the protocol. An ESA of different modulation technique or transmission PHYs can be deployed to superimpose on the BSA. In this way, combination of redundancy protection schemes can be used simultaneously.

In terms of redundancy protection, the troublesome single shared "ether" medium can be a blessing.

IV Time-out Fault Management Strategy.

It is important to restrain from excessively complex fault control mechanism, where in some instances, this aspect of the protocol dominates large percentage of the protocol overheads. In this proposed architecture and protocol, the structure allows a simple, direct time-out strategy. In all request-reply exchange, a suitable preset time-out termination of the connection can be implemented. It prevents complex connection bookings.

One exception may appear to exist to this general fault management concept. This is a condition when a HEC receives legitimate but disparaging synchronization timings from the surrounding HECs. This scenario is possible if a large number of BSAs is connected in a long loop. In this case, the fault condition can be corrected at the modem level. The greatly erred synchronization timings will not be integrated in the synchronization loop, although the existing of such signal will slow down the synchronization acquisition time. In SN-SADAMA, however, this start-up transient delay is transparent to the protocol because unless a RUN or a HEC becomes fully synchronous, the protocol will not be operational.

V Closing Discussion

As it is clear that the DAMA nature of the protocol allows a highly efficient and flexible way of utilizing the available bandwidth. The frame by frame DAMA through [PAST] slot enable real time adaptation to the traffic demands. There is an added advantage in EBSA and ESA operations. Half of the bandwidth necessary in these operations are now carried by the back-bane LAN, conserving the valuable and scarce "ether" medium resource for wireless use. Within the BSA, the maximum data bandwidth available to one user is slightly less than 0.5 of the information bit rate of the system, whereas for EBSA and ESA operation, the maximum bandwidth available will approach 100% of the bit rate. The efficiency limitation will be dependent on the size of the four overhead time slots, the addresses and the control overhead bits needed in the data transaction.

The outline of architecture and protocol for IEEE 802.11 described in this paper has omitted finer details of implementation. The recommendation on the actual data transfer protocol stacks are intentionally omitted. This is an option that may be best excluded from the standards as it will be specific to the users' convenience. The recommendation to the choice of the back-bone LAN will directly impact the choice of key operation parameters within the wireless standards, however, it is still advantageous to exclude as much as possible specific choices and accommodate these problems under " Recommended Guide Lines"

The choice of imbedded Forward Error Correction (FEC), and the question of its relevance to be included in the standards can be a serious decision, as it impacts the interoperability requirements and the problem in satisfying the conditions set out within the context of the IEEE 802.11 PAR. So far, this problem can be easily dealt with within the structure of the architecture and the protocol described here and a decision can be easily made.

There are a number of other low level issues outstanding, but none has been noted that will pose a direct violation problem that is un-containable. Nevertheless, there are significant analytical and simulation works remain to verify the performance of this architecture and protocol for the wireless medium.

VI Conclusion

In this paper, an outline of a proposed architecture and protocol suitable for IEEE 802.11 standards has been described. The suitability of its applications lies in the fact that it attempts to provide a frame work where a diverse applications requirements and entrenched LAN performance expectation can be satisfied. The actual performance merits such as traffic efficiency and others require further theoretical analyses or by simulation means.

References

[1] D. Bertsekas and R. Gallager, *Data Networks* New Jersey:Prentice-Hall, 1987.

[2] S. Y. Seidel, K. Takamizawa and T. S. Rappaport, *Radio channel modeling in manufacturing environments Parts II and III* IEEE Project 802.4L document No. P802.4L/89-14a

[3] T. S. Rappaport and C. D. Mcgillem, *UHF characterization in factory* IEEE Project 802.4L document No. P802.4L/89-14b

[4] D. C. Johnson, *Retail measurement results* IEEE Project 802.4L document No. P802.4L/89-21

[5] V. Hayes, *Wireless MAC and PHY specification archive* IEEE Project 802.11 document No. P802.11/90-15

Table 1: A summary of attributes deemed necessary in wireless LAN standards.

Applications	Propagation Environment			Functional Requirements
	Standard Deviation (dB)	Prop. Exp.	RMS delay Spread (ns)	
Heavy Industries, Factories.	4-7	3.3-4.2	80-140	Single node uses multiple PHYs. No single point of failure. Flexible redundancy protection. Operability with varied data rates. Nodes moving at vehicular speeds. Simple distribution system. Isochronous data. Interference protection
Shopping Mall, retail stores.	5-10	1.8-2.4	100-140	Flexibility in coverage design. Low cost Stationary nodes. Easy deployment of nodes. low installation cost. low/high traffic segregation. Wide coverage area. Large node population
Office(desk and portable computers)	2-7	3.3-4.0	≤ 50	Extremely low cost. Voice is desired. Volume market product. Ease of use. Low power consumption. Peer to peer communications. Connectivity to LAN/WAN. Low network start up cost. Low expansion cost. Nodes moving at pedestrian speeds. Automatic peer to network connection. Interoperability within the standards.

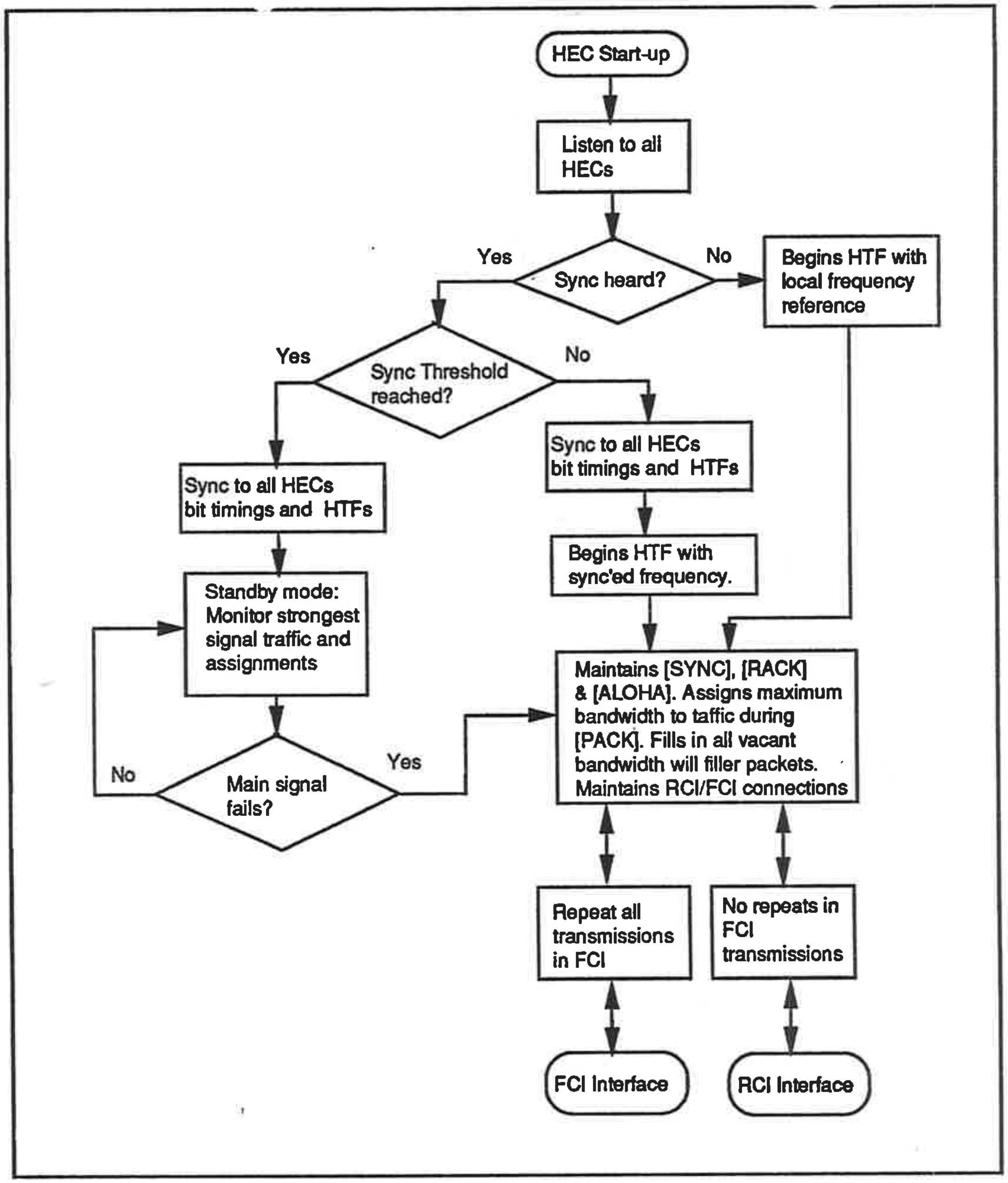

Figure 1: The HEC flow-chart showing the the operation of a HEC

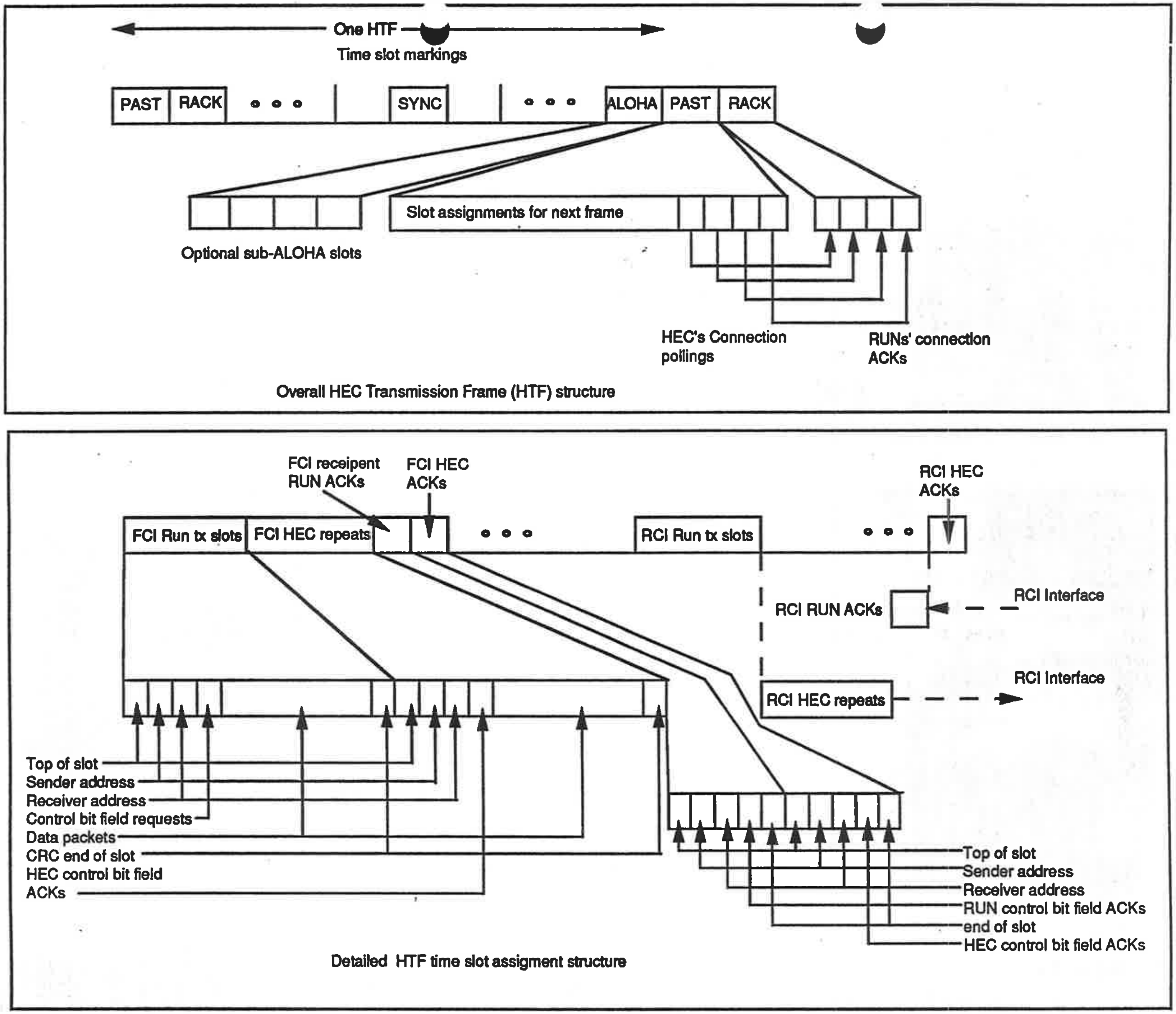

Figure 2: The proposed structure for the HEC Transmission Frame (HTF)

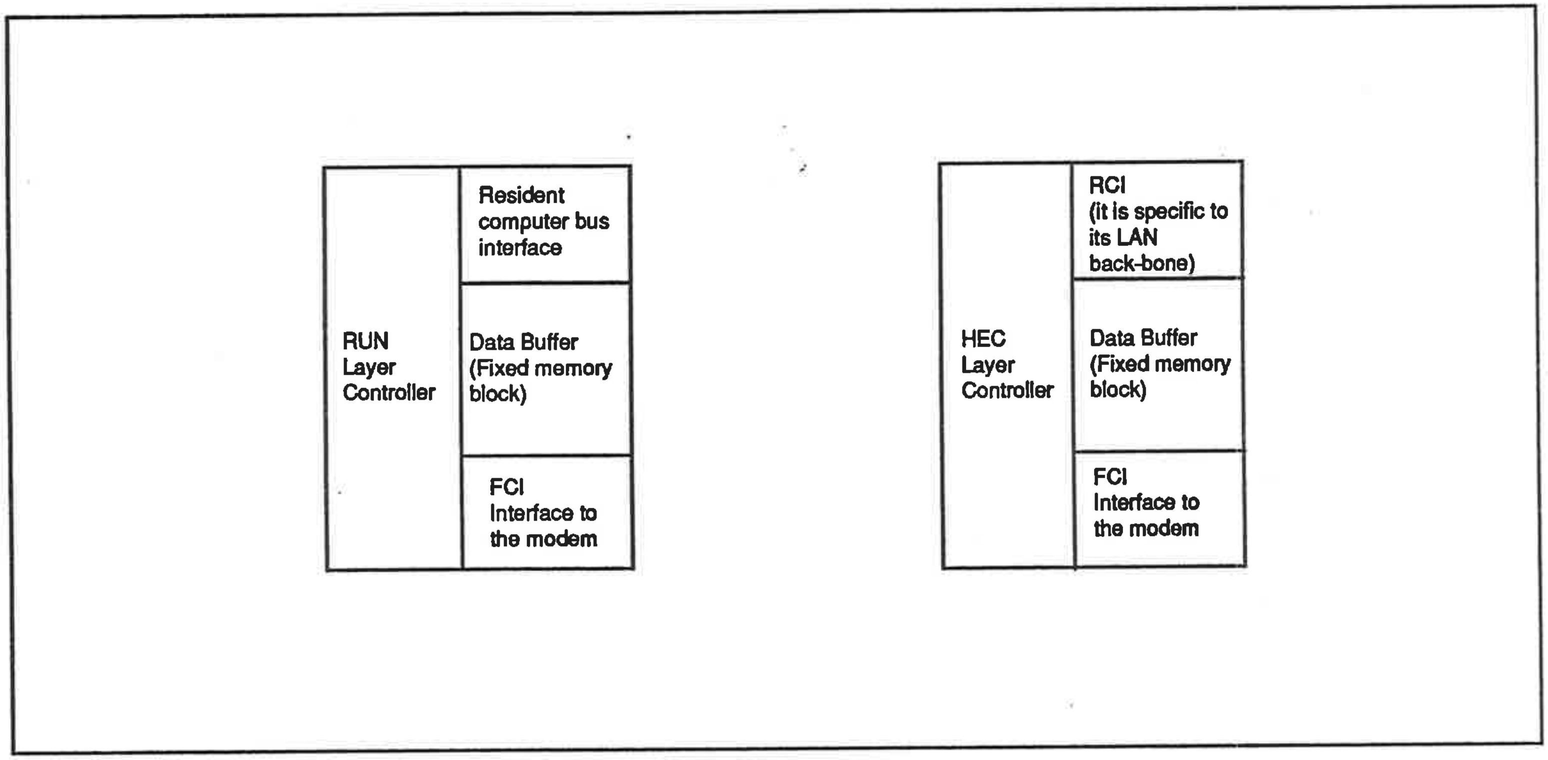

Figure 3: A proposed RUN/HEC Architectural structure. A specific assignment of the communications layers is intentionally avoided.

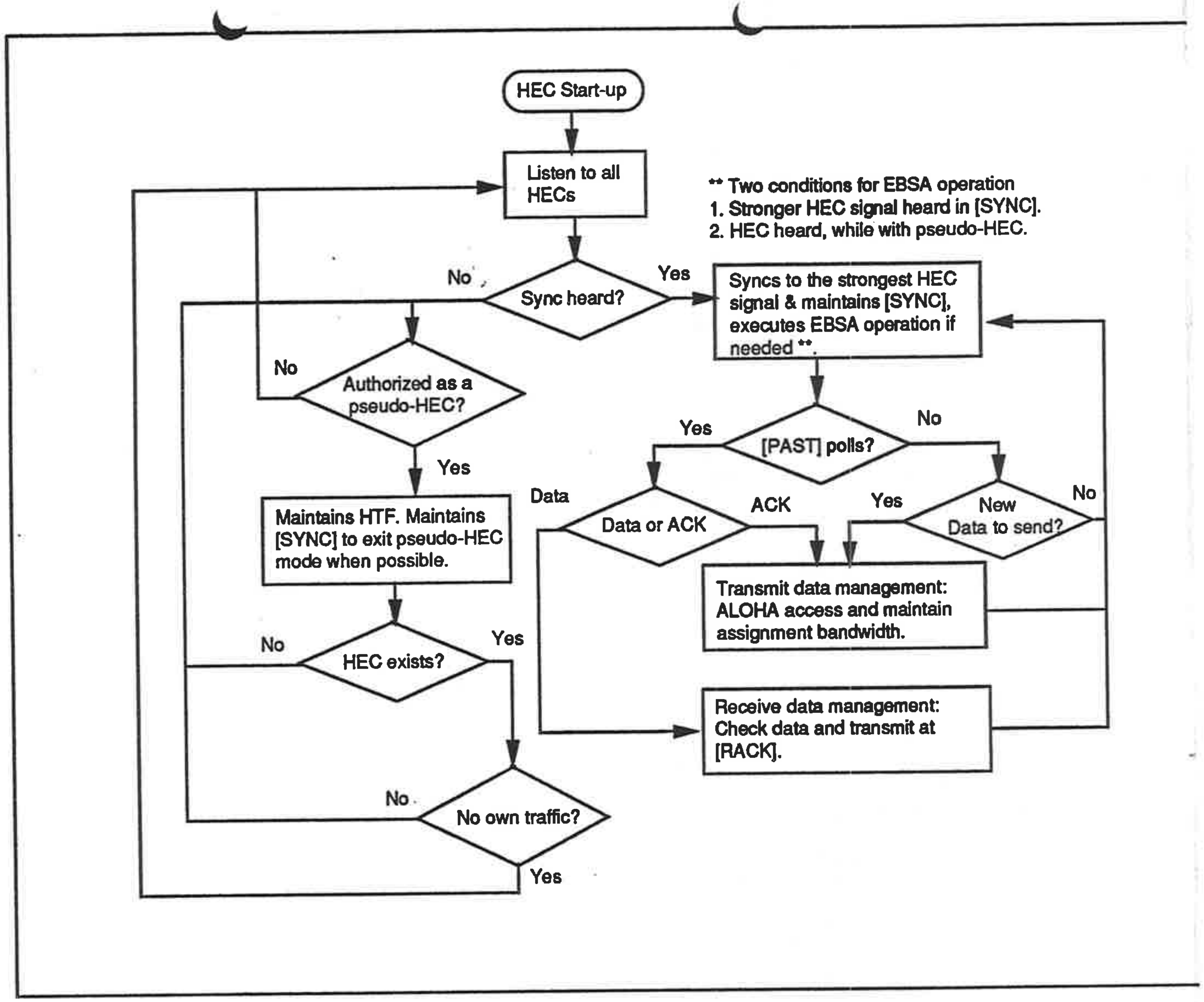

Figure 4: The RUN flow-chart showing the the operation of a RUN

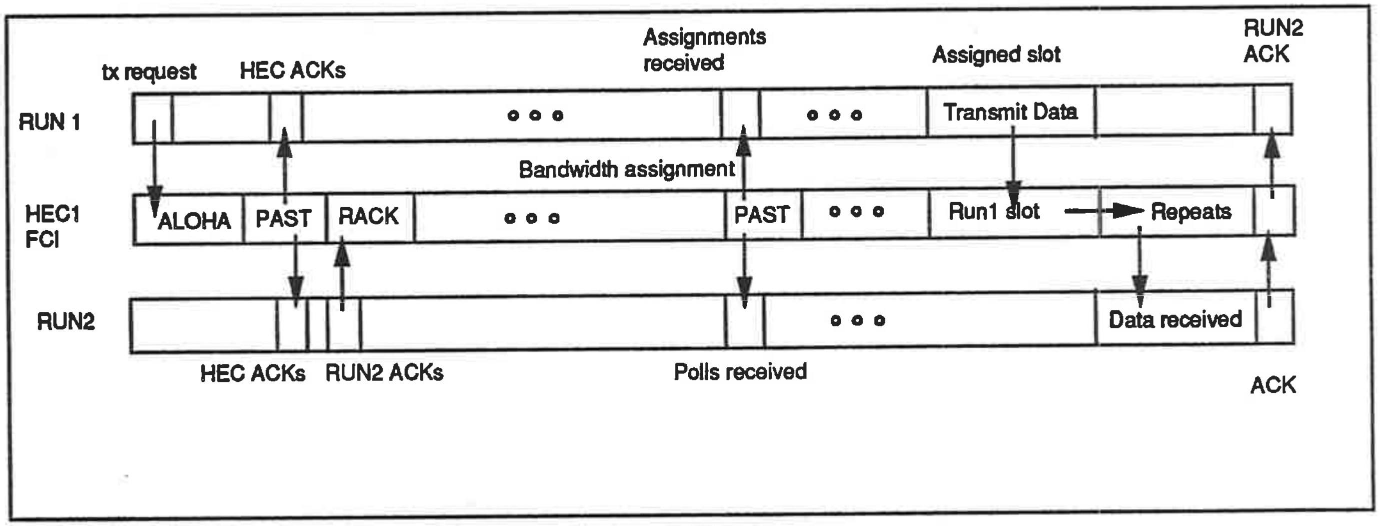

Figure 5: The detail BSA transmission operation timing chart.

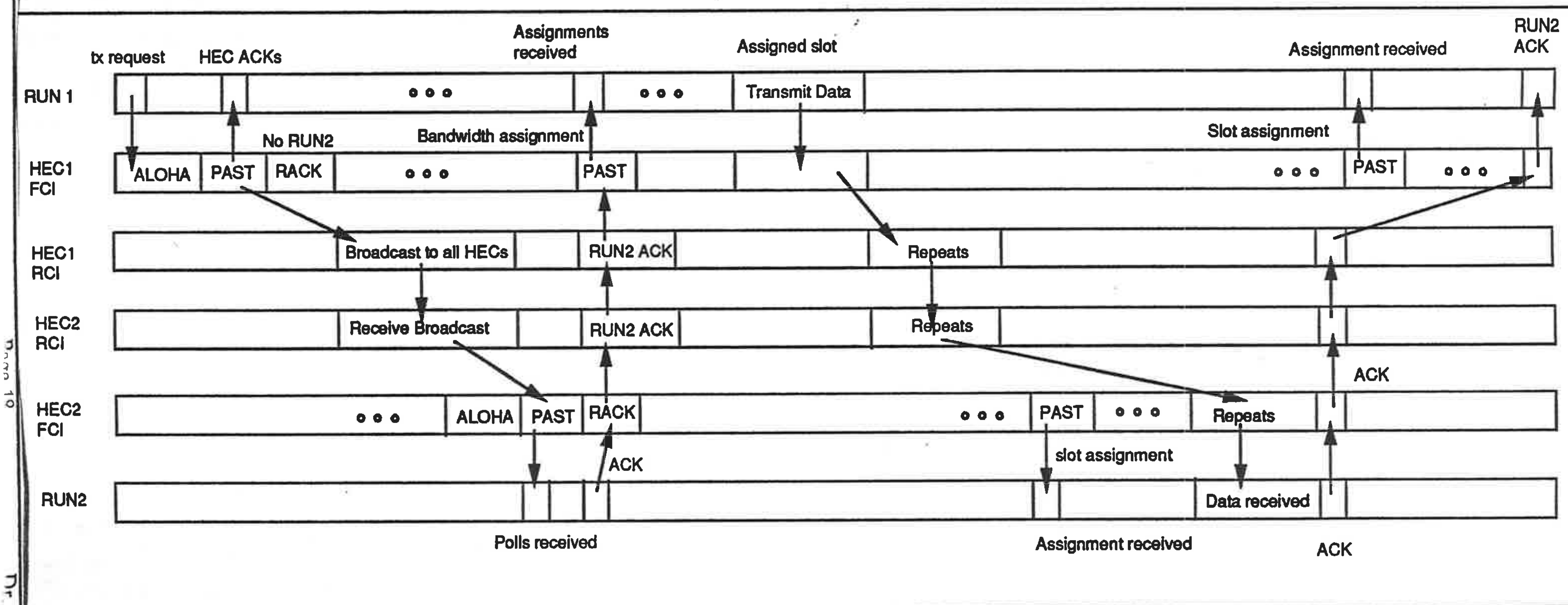

Figure 6a: The detail EBSA transmission operation timing chart between 2 fixed RUNs

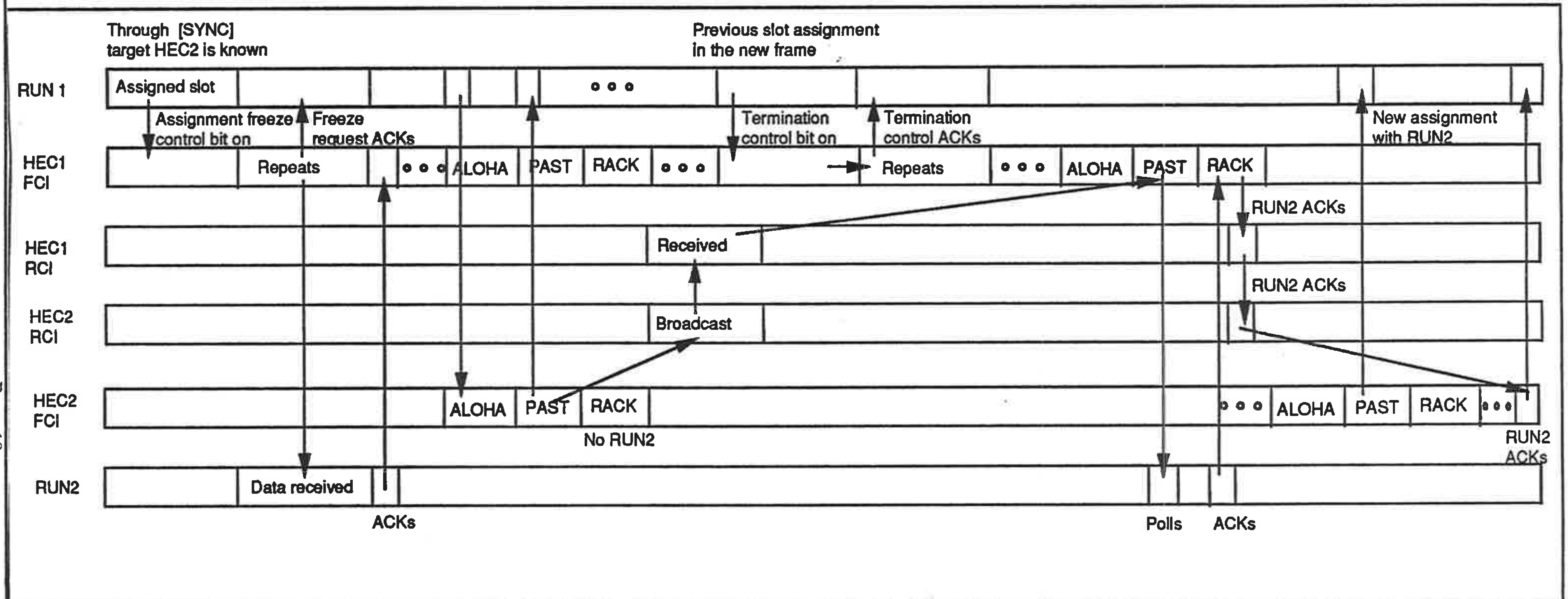

Figure 6b: The detail EBSA transmission operation timing chart between a fixed RUN and a mobile RUN in a typical way.

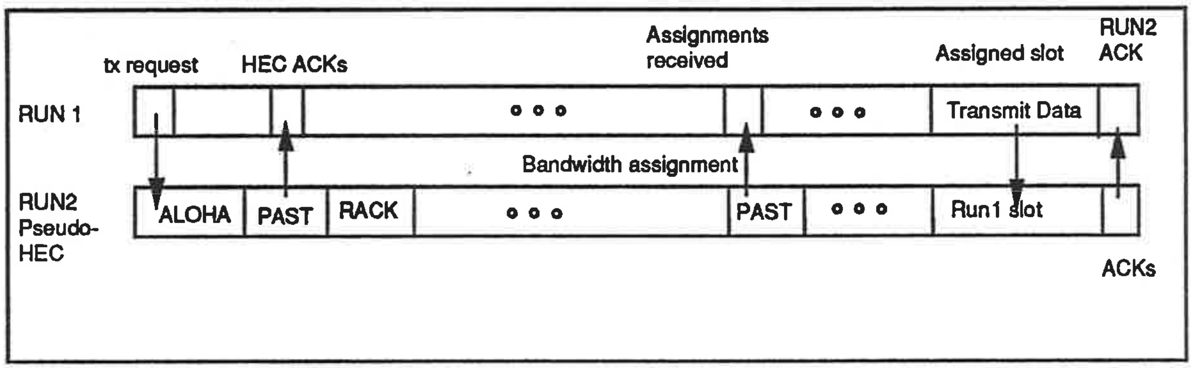

Figure 7: The detail peer to peer transmission operation timing chart.