



ELECTRONIC PAYMENT SYSTEMS 101

Amitabh Saxena, Managing Director, Digital Disruptions



USAID
FROM THE AMERICAN PEOPLE

fhi360
THE SCIENCE OF IMPROVING LIVES

Definitions and History

In its simplest form, a payment is any exchange of value between two parties, where usually Party A offers a form of currency in exchange for a good or service provided by Party B. The advent of nation-states issuing fiat currency, such as paper bills and bronze coins, which unlike gold or silver coins, did not have an intrinsic value, helped take what we now call “cash” payments to a global scale.

While the traditional definition of a payment has the receiving party providing a service, certain definitions are more broad to include transfer of value between two parties without an explicit service being provided (e.g., donating money to a religious institution, providing funds to a family member in need). In this brief we will consider the case when the end party is an individual and the sender is either another individual or a government—i.e., Person-to-Person or Government-to-Person payments—though technically it could also include Business-to-Business (e.g., transfer between one financial institution and another).

In either of these cases, in all countries **cash** is still the overwhelming choice as a consumer payment mechanism, particularly for transactions of small volumes (less than \$100 USD). Centuries-old, paper **checks** are still prominent in developing countries—and indeed in some developed ones, such as the United States—and represent a relatively straightforward way to initiate a payment. Electronic systems developed initially to process checks between one bank and another, and then later used to process various types of electronic transactions between consumers and merchants (for example, a user instructs a bank to automatically deduct their checking account to pay a monthly bill of a utility company) or among institutions themselves. This so-called “Automated

DEFINE

BANKING VS PAYMENTS

At the individual (rather than institutional) level of the broader financial services industry, we distinguish between consumer payments and retail banking. The latter refers to the provision of a full suite of core financial products—chiefly savings, loans, and insurance—to individuals. While many banks also provide some payment services (e.g., debit card issuance, online money transfer), these are complementary offerings and are usually separated from the retail arm of the bank.

Clearing House” (**ACH**) is still used today in the US and has similar forms in other countries. For real-time, larger volume processing, many banks use a **wire transfer** system to send funds between them.

Owing to the impact electronic payments can have to financial inclusion (see sidebar on next page), this Payment Systems 101 brief will focus specifically on **Card Networks**, including basic principles, main actors, and card network economics. Payment Systems 201 will delve more deeply into emerging systems such as mobile money and the potential of enhancing the current bank card payment network — or creating a new, interoperable system entirely.

Payments and Financial Inclusion: What's the Link?

Enabling a safer, faster, cheaper, and more convenient way to make and accept payments is a key component of financial inclusion. Among these characteristics, compared to more traditional forms such as cash and check, electronic payments can offer a superior option for both consumer and merchant. This takes the form of quickly transferring funds to a far-away family member by text message, enabling a small store to pay its multiple suppliers using a prepaid card, or receiving government-funded social benefit transfers directly in a recipient's bank account. This topic will be covered in greater detail in Electronic Payment Systems 201.

Card Network Fundamentals



While various large merchants issued their own credit cards in the 1940s and 50s in the US, the history of general purpose payment cards (not limited to use at a single merchant) begins in the US state of California. In 1958, Bank of America, which had a dominant market position in the state, mailed unsolicited credit cards, consisting of a paper card, with a revolving, unsecured \$300 credit line, to several thousand people in a small city outside of San Francisco. It franchised the operations, dubbed BankAmericard, to other banks in the 1960s, and formalized itself as a separate organization called Visa in 1976. A rival group of California banks created a similar card-based network in 1979: MasterCard.

It is important to note that up until their IPOs (Initial Public Offerings) in the late 2000s that made them public corporations, both Visa and MasterCard were owned and operated by their member banks effectively as a cost center, and even today they offer their global networks as a common, shared platform. (For this reason they were often referred to as “card associations” rather than independent companies.) American Express and Discover, two other dominant card payment companies, evolved differently: they operate a global network — which until recently was not available to banks—as well as issue credit to consumers and manage relationships with merchants. Such “3-party” models are discussed in more detail later in this brief.

These payment companies perform three main function related to operating a network:

1. Define membership rules. These include creating policies and operating rules governing a franchise model, such as standardizing technical implementation, calculating overall fee structures, defining merchant acceptance standards, and drafting policies for resolving disputes or fraud. Crucially, they also set guidelines on which actors can take part as members,

which is usually limited to banks. (These “franchise rules” remain a core responsibility of what global payment networks such as Visa and MasterCard do today.)

2. Build brand and trust. Payment companies seek to build trust for consumers to use its network, and to raise awareness for merchants to accept it. Brand building and reinforcement is a crucial part of what payment networks do, as witnessed by the hundreds of millions of dollars in marketing that is spent annually by each of the global brands.

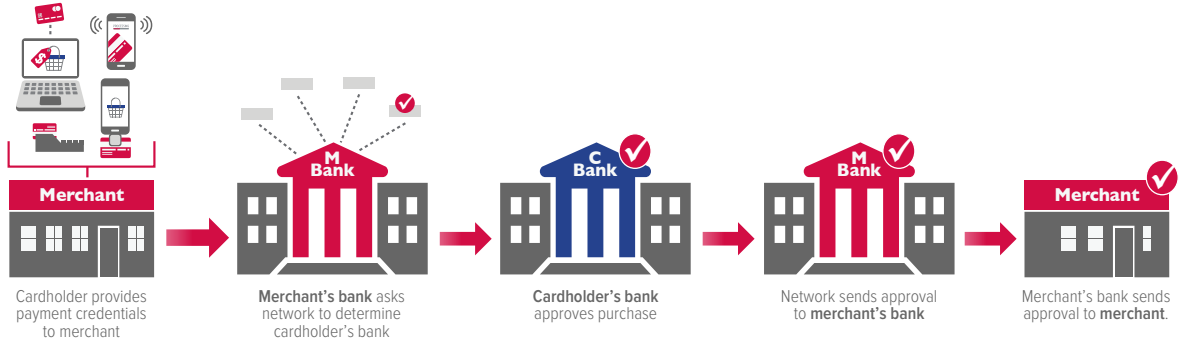
3. “Process” transactions. All payment networks perform some degree of data “processing”, although the specific activities and mechanisms to do so vary. In addition to what is referred to as “switching”, most, if not all, facilitate three main functions: authorization, clearing, and settlement.

DEFINE

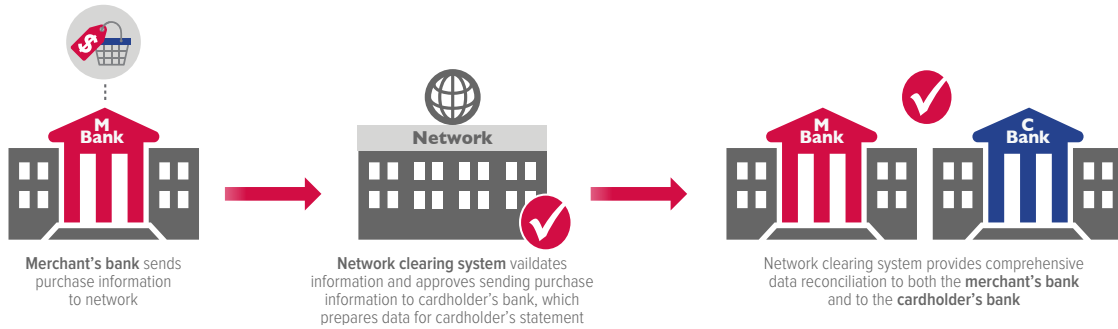
Switching refers to correctly routing a transaction message from one party to another in the network. This, by design, is a core responsibility of a network with multiple “nodes”, here represented by financial institution members.

TRANSACTION PROCESSING IN-DEPTH

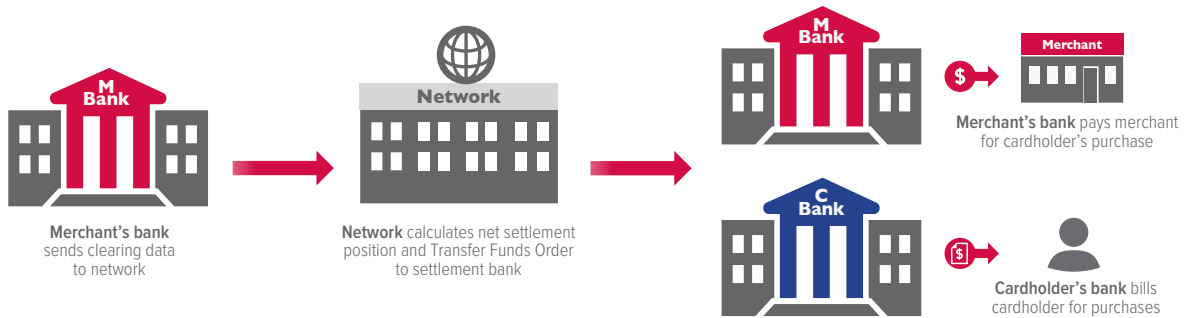
a. **Authorization** basically means permitting a transaction to go through. In most cases it is the cardholder's bank that is responsible for authorization (for example, to verify that its customer has sufficient funds in their debit card account), though the network may perform some pre-validation (checking for fraud, or that correct information has been entered) before passing it through. This is the first message that is "switched" by the network to the cardholder's bank.



b. Once the transaction is authorized, the network is responsible to help **clear** the transaction: in other words, reconciling the amounts, in addition to any fees, sending information to both parties detailing the amount to be credited or debited, and referencing the consumer or merchant account in question.



c. **Settlement** occurs when funds are transferred from one bank to another. In card systems this process usually occurs at the end of every day (not in real-time) and represents the net position of all participating banks for all transactions that occurred during that period. It is important to highlight that payment networks do not handle the actual fund transfer themselves, but provide the instructions to do so; payment networks usually designate a separate "settlement" bank for members to move funds in and out of.



Note: Payment networks do not wholly perform authentication (validating the cardholder's identity), but may facilitate it by routing information to the responsible party. For example, for cards that require a consumer to enter their PIN, the network will capture this information, along with information on the actual payment and card account, and route it to the cardholder's bank, which will then both authenticate the user and authorize the transaction.

Domestic Switches

While the above global payment networks process most card transactions in the US, in developing countries the situation is often quite different. In many, though not all, cases, so-called domestic switches—companies usually owned by local banks—perform many of the functions described above. (Their structure and function resemble what MasterCard and Visa were before they went public: owned and operated by banks for the benefit of its members.) Because these domestic switches operate exclusively in the country in question, transactions that are cross-border—where the consumer and merchant banks are in different countries—must eventually pass through a global network that connects the two.

Moreover, these switches can be used to make a single payment system operate (such as debit systems, which are the most common), and often run at a lower overall cost to the banks to process domestic transactions than international schemes such as Visa and MasterCard. Finally, as its remit is limited by its borders, domestic switches are often able to customize products to the local market, and do so much quicker than its global counterparts.

Payment Participants

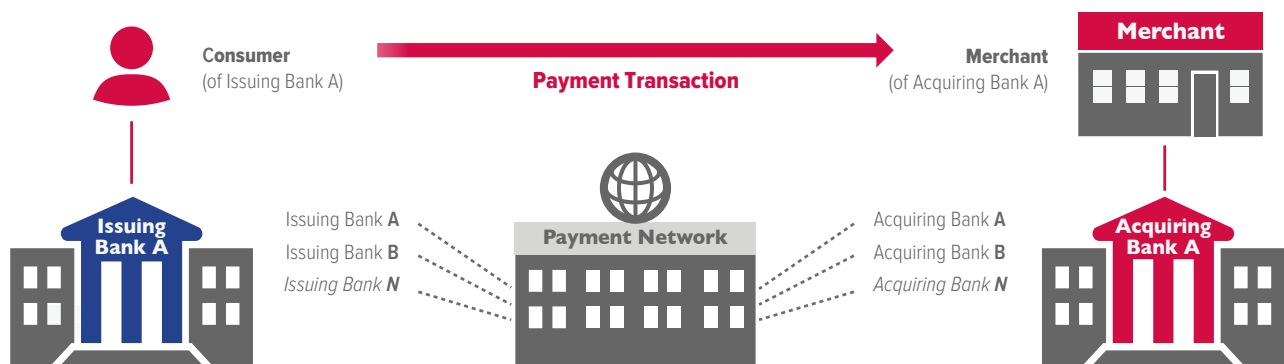
The main participants in card payments are described in the figure below. In addition to a payment network, there are four main participants: two types of financial institutions (Issuing Bank and Acquiring Bank) and two types of end-users (Consumer and Merchant).

The Consumer and Merchant are the most straightforward: the former, sometimes called the cardholder, has a debit or credit card and initiates a purchase transaction by providing payment credentials (card number and in some cases, a secret PIN or a verification code on the back of the card) to the merchant. The Merchant is any regulated business that accepts the payment card, holds an account at a bank where transactions are paid into, and can either be physical or online.

The Issuing Bank is a regulated financial institution which owns the account of the customer (either credit or debit), and is responsible for marketing, issuing the actual plastic card, authorizing and settling transactions, and providing ongoing customer service. Similarly, the Acquiring Bank manages the relationship with the merchant: it has overall responsibility for marketing (i.e., acquiring) the merchant, installing the necessary hardware (such as a Point-of-Sale terminal), software, and branding at the merchant site, processing and settling transactions, and providing ongoing merchant service. Many acquiring banks outsource some of the functions to other parties who provide services "on-behalf" of the bank, but in all cases the acquiring bank has ultimate responsibility for the merchant account. Moreover, many, though not all, retail

banks have both issuing and acquiring units; in the event that a payment transaction is processed by a consumer and merchant bank that belong to the same financial institution, the so-called "on-us" transaction is in almost all cases handled within the bank and does not route through the payment network. (Note that in the previous schematic on transaction processing, the Acquiring Bank is depicted as "Merchant Bank" and Issuing Bank as "Cardholder Bank").

The figure below represents an **open-loop platform**, whereby the payment network operates as a hub-and-spoke model: multiple acquiring banks and issuing banks can join, each bringing their own set of merchants and consumers. The strength of the network is proportional to the number of acquiring and issuing banks (and corresponding end-users) in its system, which explains the current dominance of global networks such as MasterCard and Visa, which connect thousands of banks in nearly every country in the world. It is also important to note that, as previously mentioned, while an open-loop network may make efforts to raise brand awareness to both merchants and consumers, it technically does not "own" the relationship with either end-user.



In contrast, **closed-loop networks**, as the name implies, restrict membership to other bank parties. The most prominent example is American Express, which until recently, acted as a single entity comprising three roles: payment “network” (from a transaction processing standpoint), issuing bank, and acquiring bank. This is referred to as a Three-Party Model. In this case American Express does directly manage relationships with

consumer and merchants alike. It has since begun licensing its network as stand-alone service to other financial institutions, which is why some banks (such as Banco Bradesco in Brazil and ICICI in India) now issue American Express cards. Other common closed-loop platforms include many gift card and transit cards, where the only point of acceptance is the retailer or public transport provider in question.

Card Network Economics

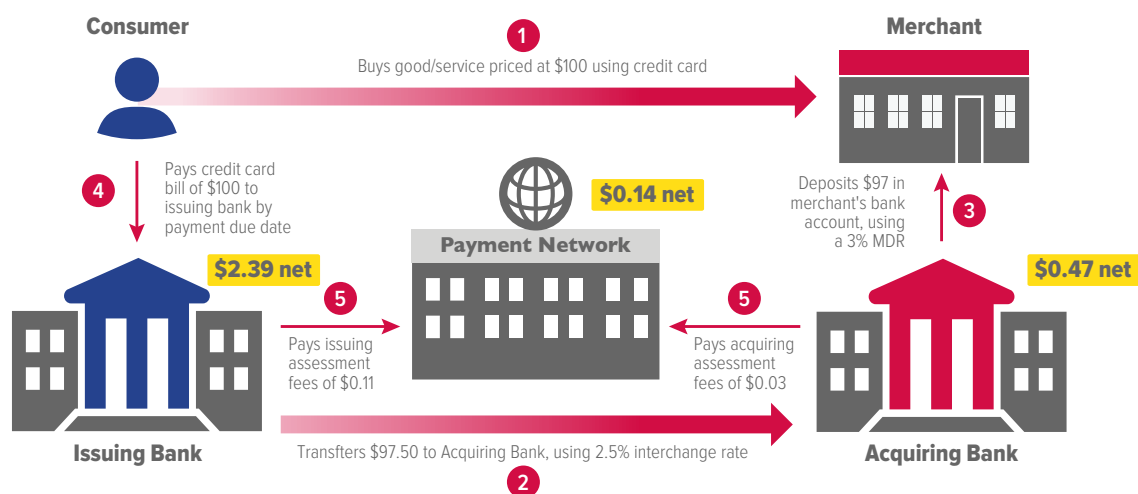
The two most important fees in card networks are **interchange** and **merchant discount rate (MDR)**. As its name implies, the MDR is the fee that a merchant pays to accept a card payment transaction. Often this is a percentage based on the purchase price of the good or service being sold (e.g., 3%) and sometimes includes a fixed fee for each transaction (e.g., \$0.20).

The MDR varies on a number of factors—whether it’s a debit or credit card, the type of loyalty programs associated with the card, if it’s an online or offline purchase—and varies by the type of merchant (supermarkets have lower rates than small stores, for example) and total monthly payment volume. Note that the MDR is one of several fees that merchants pay (for example, registration fees and POS terminal rental fees,) but it is usually the merchant’s largest cost component for accepting cards.

Interchange is the fee that an acquiring bank pays an issuing bank, and thus is a concept limited to open-loop networks. The rationale is that the issuing bank, particularly for credit card transactions, takes the risk of not getting repaid by the cardholder in case of loan default; since debit card transactions do not have this risk, interchange rates are substantially lower. Because interchange is a cost for the acquiring bank while MDR is a revenue item, the latter is usually priced as “cost-plus” pricing—i.e., the interchange rate plus a small margin that

represents gross profit for the acquiring bank. It is important to highlight that the payment networks do not receive revenue from interchange, but rather from a variety of “assessments”—either based on transaction number, volume, or both—on each transaction, in addition to monthly fees. The step-by-step payment funds flow are depicted in the example below, using a \$100 purchase price and 3% MDR as an example (all values are hypothetical), as well as the net revenue gained by each participant from the transaction in question.

While in all cases the acquiring bank sets the MDR, the entity setting interchange depends on each market. In some cases, such as in the United States, it is set by MasterCard and Visa and available to the public¹ (in most countries it is not published). In others, such as in Peru, it is set by a consortium of banks. Still in other countries, such as India and Colombia (and in the United States, for debit cards), it has been regulated by governments.



1. See <http://usa.visa.com/download/merchants/visa-usa-interchange-reimbursement-fees-april2013.pdf> and http://www.mastercard.com/us/merchant/pdf/MasterCard_Interchange_Rates_and_Criteria.pdf for more information.