CHAPTER 2

FPGA-BASED IP CORE PROTECTION

2.1 LITERATURE REVIEW

The protection of IP has become a serious issue as intercompany subsystem design exchange becomes more commonplace. Over 150 companies including all the major players formed the Virtual Socket Interface Alliance (VSIA) in March 1997 to address the IP protection issue (Ralf 1999). The VSIA IP protection development working group (VSI 2001) has identified three main approaches to secure IPs. First is a deterrent approach, where the owner uses legal means to stop the attempts of illegal distribution. This is done by using patents, copyrights and trade secrets. Second is a protection approach, where the owner tries to prevent the unauthorized usage of the IP physically by license agreements and encryption. Protection techniques are mostly based on model encryption and they fall short in securing the designs or tracking them, in case they are stolen or reused without permission. For such reasons a third detection approach has been introduced, where the owner detects and traces both legal and illegal usages of the designs as in watermarking or fingerprinting techniques.

2.1.1 Deterrent Approach

Whenever a company in the EDA industry invests a significant amount of time, money or effort to develop some form of competitive software or hardware, it is important for them to protect such proprietary
interest legally for the advantage in the EDA marketplace (Wollinger et al 2004). Typically, this is accomplished by acquiring patents, copyrights or trade secrets (David 1987). This protection provides the owner of the property with significant rights and benefits including: - Enhancement of company valuation; Source of future income; Recoupment of corporate investment in research and development; Control of market share; Deterrence against misappropriation or infringement; Recovery of lost profits or damages and negotiation of leverage in defensive cross-licensing situations.

2.1.1.1 Patent Protection

Among the different forms of intellectual property protection, patents may extend the broadest scope of protection. Patents protect innovative, useful, non-obvious ideas and concepts and provide a patent owner with the right to exclude others for 17 years from making, using or selling inventions. Inventions that are patentable generally include devices, compositions of matter and processes. Examples of patentable subject matter include electronic hardware and computer software, particularly unique methods for logic design like partitioning, synthesis, simulation, emulation, testing and verification. If a court determines that one or more claims of the patent are infringed, then the rights can translate into injunctions to prevent further violation of the patent or the award of damages can be granted or both can be considered. These damages are based on either lost profits or reasonable royalty depending on circumstances.

Many companies mistakenly believe that patenting a design is an effective protection against design theft. While a patent certainly discourages reputable companies from even attempting to reverse engineer a product, it does little to discourage design piracy. IP pirates have little regard for the law or court and in fact usually operate in countries with weak or nonexistent IP protection. Even if the pirate can be prosecuted in a country that respects IP
rights, such a proceeding can be both lengthy and costly. However using
patent protection in addition to physical design protection is a good way to
provide additional product security, especially the brand protection.

2.1.1.2 Copyright

Copyright is a set of exclusive rights granted to the author or
creator of an original work. These exclusive rights include the right to copy,
distribute and adaptation of the work. Copyright does not protect ideas but
only their expression. In most jurisdictions copyright arises upon fixation and
does not need to be registered. Copyright owners have the exclusive statutory
right to exercise control over copying and other exploitation of the works for
a specific period of time, after which the work is said to enter the public
domain. Uses covered under limitations and exceptions to copyright such as
fair use do not require permission from the copyright owner. All other uses
require permission and copyright owners can license or permanently transfer
or assign their exclusive rights to others.

Although it is usually difficult to directly prove that an alleged
infringer actually copied a work, we can often infer that copying took place
by showing that the alleged infringer had access to the work. This is
particularly applicable to mass-marketed software, since in these types of
software, the means to reverse engineer is readily available to the public.
Once a court determines that the copyrighted software has been infringed, a
variety of remedies may be available. These include injunctions to prevent
further infringement, impoundment of the infringing articles and even
criminal charges when infringement is shown to be wilful. If the copyright is
properly registered then one can receive either actual damages or statutory
damages.
IP modules designed by one company and sold in a non-physical form (e.g. HDL, netlist, layout) to others; do not have a natural physical manifestation. The IP blocks are modular and designed to be integrated within other systems, usually on the same chip. As a result of the flexible, intangible nature of these modules, IP copyright protection has become a problem. A thief needs to only resell or reuse an IP module without even reverse engineering the design, as a result of which the proof of IP ownership becomes difficult to assert due to its inherently abstract nature.

2.1.1.3 Trade Secret

Trade secrets have long been a popular method for protecting software. However, trade secret protection is usually ineffective for protecting mass marketed software unless one cannot ascertain the secrets from the form of the software distributed. Also unlike the other forms of intellectual property, trade secret law is governed by the individual states rather than by the federal government, hence the applicable law may vary somewhat from state to state. The variety of subject matter that can be given trade secret status is quite extensive and can encompass software and other related subject matter, formulas and programming techniques, compilation programs, device, methods, techniques or processes. No filing or registration is necessary to obtain a trade secret. However, one must be able to say that certain circumstantial factors are present for the subject matter to qualify as a valid trade secret. These include: whether one can derive economic value from the trade secret while the subject matter is a secret; whether the subject matter is really a secret and not in the public domain; and whether one continually take reasonable efforts to ensure that the subject matter remains a secret. Compliance with the third factor can be a complex and often expensive undertaking. As the owner of a valid trade secret, one may be entitled to various remedies if one can show that the trade secret was misappropriated,
procured by improper means like theft or disclosed by one who had a duty not to disclose it. Reverse engineering of rightfully obtained (such as purchased) subject matter is another permissible means of obtaining a trade secret. If a court determines that misappropriation of the secret has indeed occurred then the remedies available typically include injunctions preventing additional harm to the trade-secret owner, compensatory damages for the actual damage resulting from the misappropriation and when misappropriations are shown to be wilful even punitive damages are possible.

However, these deterrent approaches are effective only where appropriate laws exist and are enforced. Attitudes towards design-ownership rights vary significantly worldwide, making this type of deterrent approach not wholly effective in places where it matters the most. Countries that source counterfeit goods tend to be places where design ownership rights laws and enforcement are weak or ambiguous. These measures do not actively prevent fraud or theft but their presence may deter the would-be attackers and hence it can be beneficial to advertise them.

2.1.2 Protection Approach

2.1.2.1 Licensing Agreement

License agreements are excellent tools for extracting an IP owner’s rights in court once a violation is detected and proven. A license agreement when present is a document that a user must agree prior to using an IP and it generally states that the user must use a valid, legally obtained copy of the IP and only use it for a specified period of time. While the concept of licensing ready-made IP cores for FPGAs is particularly suited for small-scale operations, the current pricing strategy works heavily against customers. Since the IP core providers cannot limit the number of deployment instances, they cannot differentiate in pricing between the two classes of customers.
Kean suggests a pay-per-use licensing scheme where IP blocks can be purchased for one-time use only (Kean 2002a). Although this provides a means to profit from FPGA IP, there is nothing in place to physically protect the IP itself. In addition, an IP vendor can only profit once from each client. Similar to software licensing schemes, the licensing of hardware IP often involves a license agreement combined with a separate component used for physical protection of the design. A secure mechanism and licensing model for protecting IP cores against IP over-deployment is discussed in Soudan and Adi (2008) wherein a solution is proposed that allows the IP to be licensed for specific uniquely identifiable devices. Here, it is impossible to deploy the IP into any additional device without the explicit involvement of the IP provider.

However the strength of licensing agreement is mainly built on the assumption of goodwill between the IP owner and the IP customer. The main issue is the lack of a reliable mechanism for IP providers to collect royalties for the IP cores they license out. Enforcing licensing agreements often requires a combination of techniques. Some software license enforcement techniques include watermarking and/or fingerprinting for fraud identification, code partitioning, software/hardware tokens and dynamic code decryption.

2.1.2.2 Encryption Techniques

SRAM-based FPGAs are volatile and require an external configuration memory, so they are the least secure of available FPGA devices. Cloning and reverse engineering are the major threats to SRAM-based FPGA. Obfuscation based protection and use of sophisticated encryption techniques are the most effective and practical countermeasures.

Obfuscation is a technique by which the description or the structure is modified intentionally to conceal its functionality and thereby making it
significantly more difficult to reverse engineer. Obfuscation techniques can be classified into two main categories: passive techniques, which do not directly affect the functionality of the system and active techniques, which directly alter the functionality of the system. Often the active hardware obfuscation techniques (Rajat and Swarup 2009) are key-based such that normal functionality of the obfuscated design can only be enabled by the successful application of a single pre-determined key or a sequence of secret keys at the input. In contrast, the passive techniques modify the circuit description in a soft form (e.g. syntactic changes), such that it becomes difficult for a human reader to understand the functionality of the circuit. In the case of obfuscation of HDL (Hardware Description Language technique) (Brzozowski and Yarmolik 2007), the code is reformatted by changing the internal net names and removing the comments, so that circuit description is no longer intelligible to the human reader.

Encarnacion et al (2007) suggested an obfuscation technique in which the HDL source code is modified by removing the comments and changing the internal net-names followed by a simple string substitution strategy. This is done in such a way that the functionality of the system remains unchanged but the logic description becomes incomprehensible. A major shortcoming of the passive approaches is that it does not affect the functionality of the IP core and thus cannot prevent it from being stolen by a hacker, who may use it as a black box module (Meyer-Base et al 2011). Most widely available solutions to the IP authentication problems are based on bitstream encryption. Software-based configuration bitstream encryption is used to safeguard system the design IP as it is loaded onto the FPGA. Once transferred, the configuration bit stream is decrypted on the FPGA chip.

The SRAM-based FPGA bitstream security problem arises because an attacker can probe the connection between the FPGA and the external non-
volatile memory during the configuration and obtain a copy of the programming bitstream. Many solutions to the problem have been proposed over the last years. One of the earliest suggested mechanisms for providing bitstream security to the FPGA is contained in a US patent assigned to Pilkington Microelectronics (Austin 1995). This approach protects against both piracy and reverse engineering. The primary disadvantage of this approach is that it requires non-volatile memory and hence non-standard processing of FPGA device which increases the cost. A variant of the Pilkington scheme is used by the Xilinx in their Virtex II family wherein instead of providing on-chip EPROM (Erasable Programmable Read Only memory) to store the cryptography key; the key is stored in the key register with its own power supply lines (Ralf 2004). Here an external battery maintains the state of register when the equipment is power off, which adds cost to the system and also reduces the overall reliability, as a momentary loss of power will delete the key.

Kean (2001) has proposed a strategy based on a static secret key already inserted during the manufacturing process. Here, the issue of key transfer is solved by including the cores both for encryption and decryption in the FPGA. Kean (2002b), Simpson and Schaumont (2006) have proposed more complex protocols for a complete solution. But both of these approaches require the implementation of additional security features in the FPGA and also the participation of FPGA manufacturer whenever a bit stream is to be encrypted for a particular FPGA. The problem of key storage is overcome by using Physical Unclonable Functions (Lim et al 2005; Kumar et al 2008) to store the keys in a secure way. Tuyls et al (2006) have described how a combination of bitstream public key encryption and a key extracted from a Physical Unclonable Functions works for the IPP of the FPGA.
Yip and Ng (2000) have introduced a partial encryption scheme in which the configuration bitstream is partially encrypted and then loaded onto a separate RAM (Random Access Memory) built into the FPGA. The security of this technique relies on the fact that the bitstream file is hard to reverse engineer. Another possibility which has been suggested is the anti-piracy scheme analogous to the dongles, where a separate security chip, normally a Complex Programmable Logic Device (CPLD) is provided and connected to user I/O on the FPGA (Kessner 2000). The disadvantage of this scheme is the cost of the external CPLD and the FPGA resources required to implement the security circuitry within the user design. Dynamic IPP techniques that uses both public-key and symmetric cryptography but does not burden the FPGAs with usual overhead of public key cryptography has been proposed by Tim et al (2007) and Adi et al (2006). But these approaches require few modifications to the current FPGA technology. Finally, in practice, most of the encryption based IPP approaches have eventually been broken and it is expected that stronger forms of encryption combined with more thorough system engineering will likely improve the reliability of the encryption techniques.

2.1.3 Detection Approach

Various kinds of IP marks such as watermarks, fingerprints are embedded into the design to ensure IPP through identification of legal owner. Inclusion of robust marks incurring low design overhead is of major concern of IPP technique. On the other hand convincing the verification of marks while maintaining robustness of the process, that is without leaking any relevant information is also a prime issue. Mark verification is necessary during selling a design IP as well as later on, if any claim is raised. During selling, IP vendor’s signature is extracted from the design to convince the buyer that the product is from a bona fide IP owner. The buyer’s signature is
also extracted to ensure buyer’s protection against repudiation attack. However, the effectiveness of detection based IP protection schemes is limited by the fact that these techniques are passive and hence cannot prevent the usage of the stolen IP. However, the identification marks may serve as initial evidence in court or trigger further investigation in the case of IP infringements.

2.1.3.1 Watermarking Techniques

Watermarking is the direct embedding of additional information into the original content to be protected. The functional correctness must be preserved in the case of IP core watermarking. Watermarking (Fan and Tsao 2003) is a mechanism for identification that is designed to (1) identify the author, source, used tools, techniques and also the recipient of the IP (2) be nearly invisible to human and machine inspection (3) be difficult to remove and (4) be permanently embedded as an integral part of the design.

Most methods for watermarking IP cores focus on either introducing additional constraints on certain parts of the solution space of the synthesis and optimization algorithms or by adding redundancies to the design. Lach et al (1998a) introduced a watermarking approach for watermarking the bitfile-cores codes, wherein the signature is embedded into the unused look-up tables (LUTs). Later the approach was improved by using many small watermarks instead of one large one (John et al 1999) and here the size of the watermarks are only of the size of a lookup table. The advantages of small watermarks are that they are easier to search and also for the verification process only a part of the entire watermark positions must be published.

Constraint-based methods were originally introduced by Kahng et al (2001), which restrict the solution space of an optimization algorithm by
setting additional constraints that are used to encode the signature. The different methods for embedding the constraint-based watermarking in FPGAs are by exploiting the scan-chain (Chip-Hong and Aijjiao 2010), preserving nets during logic synthesis (Kirovski et al 1998), placing constraints for CLBs in odd/even rows (Kahng et al 1998a), altering the transistor width (Bai et al 2007) or routing constraints with unusual routing resources (Kahng et al 1998b). An approach to watermark a HDL-core by watermarking the scan chain has been proposed by Lin et al (2004). The method to watermark the netlist-cores by preserving certain nets in the synthesis and mapping step has been pointed out by Darko et al (1998) and Alpert and Kahng (1995). Adarsh et al (2003) introduced a technique for watermarking the bitfile-cores, wherein the delay on selected paths is measured and added to find the new timing constraints. Jiliang et al (2012) proposed a FPGA bitfile watermarking method that can achieve the goals of zero-overhead without affecting the function and performance of design.

A common problem of many watermarking approaches is that of the verification of the presence of the marks. The existence and the characteristic of a watermark must be disclosed for verification, which enables the possible attackers to remove the watermark. To overcome this obstacle, Adelsbach et al (2004), Li and Chang (2006) have presented a so-called zero-knowledge watermark scheme which enables the detection of the watermark without disclosing relevant information.

The potential sources of information that can be used for extracting a watermark are configuration bits, the power consumption, electromagnetic (EM) radiation and the temperature. In the case of verification by extraction of FPGA bits (Ziener et al 2006; Schmid et al 2008), the bits can be analyzed to detect structures that can carry a watermark or that can be used to identify an IP core. Here lookup table contents are used which are suitable for
watermarking and IP core identification. Ziener et al (2010a) verified the signature over the power consumption patterns of the FPGA. The advantages of power watermarking are that: the signature can be easily read out from a given device; only the core voltage of the FPGA must be measured and recorded; No bits is required which needs to be reverse-engineered; and also these methods work for encrypted bits where methods extracting the signature from the bits fail. Moreover, these techniques can sign netlist cores as well as bit cores and many watermarked netlist cores can be integrated into one design. The results are superposition and interferences which complicate or even prohibit the correct decoding of the signatures. To achieve the correct decoding of all signatures, multiplexing method is proposed by Ziener et al (2010b).

It is also possible to extract signatures from the electromagnetic (EM) radiation of the FPGA (Daniel and Jurgen 2005). An advantage of this technique is that a raster scan of the FPGA surface with an EM sensor can also use the location information to extract and verify the watermark. Unfortunately, more and more FPGAs are delivered in a metal chip package which absorbs the EM radiation. Finally, a watermark might be read out by monitoring the temperature radiation (Kean et al 2008). The concept is similar to the power and EM-field watermarking approaches but here the transmission speed is drastically reduced. This is the only watermarking approach which is commercially available and here reading the watermark from the FPGA may take up to 10 minutes.

2.1.3.2 Fingerprinting Techniques

Fingerprinting (Pfitzmann and Waidner 1997; Crouch and Patel 2008) is a special case of watermarking targeting primarily the FPGA designs. Fingerprinting approach makes use of both, the watermarking technique and the design tiling technique to create secure signatures. Any effective
fingerprinting scheme should achieve the following goals: (1) The marks must be difficult to remove. (2) It must be difficult to add a mark to the released design. (3) The mark must be transparent. (4) The mark must have low area and timing overhead and also require little added design effort.

The methodology of the first IP fingerprinting technique in the literature (Lach et al 1998b) is based on solution partitioning. By partitioning an initial solution into a large number of parts and by providing for each part several different realizations, one can realize a fingerprinting scheme with relatively low performance impact for their application. However this technique cannot be applied to design steps that do not have natural geometric structure and that are sensitive to the cost of the solution. More importantly, the technique has relatively low resilience against collusion attacks since it produces solutions with identical global structure. Finally the time overhead associated with creating fingerprinted solutions is relatively high.

A generic fingerprinting methodology is proposed by Caldwell et al (2004) that apply to arbitrary optimization synthesis problems. It combines existing watermarking techniques and iterative approaches for solving the optimization problems. This fingerprinting technique can be applied for partitioning, graph colouring and placement that simultaneously provide high collusion resiliency and low runtimes. However the resilience of the method to various attacks has not been addressed. The fingerprinting approach designed by Lach et al (2001) creates protection for the FPGA based IP by inserting a unique marker identifying both the origin and recipient of the design. The fingerprinting process produces a secure mark and a unique physical layout for each design instance recipient. But there is area and timing overhead associated with this technique. Methods to dramatically increase the stability and robustness of the digital fingerprint identity by the proper choice of input sequences are dealt by Patel et al (2009). Finally it may to be noted
that the key problem in any of the existing fingerprinting techniques for the intellectual property protection is the trade-off between the collusion resiliency and runtime.

(Kindly note that henceforth in this thesis, the term IP cores or reconfigurable cores will denote the SRAM-based FPGA IP cores and the term FPGA will denote SRAM-based FPGA. Further these terms may be used interchangeably).

2.2 MOTIVATION FOR THE RESEARCH

IP security is the primary concern of companies or IP developers whose competitive advantage is derived from their ability to implement complex proprietary designs. It is also the primary concern for manufacturers of mid-to-high volume consumer electronics whose market share and profitability may be eroded by cloned or counterfeit versions of their product.

IP reuse has gained widespread acceptance throughout the semiconductor industry. IP protection is a complex subject. Though there are several legal means to protect one’s IP, the legal procedures are complex and expensive. Often small and medium firms that provide IPs are unaware of these laws and rules. Also most small and medium firms cannot afford expensive lawyers to register their IP with the concerned agencies. Social deterrents are provided by legal systems and they rely on people’s good social conduct and aversion from being prosecuted and incarcerated. The deterrent techniques that protect the IP works by law like the copyrights, trade secrets, patents are effective only if the misconduct can be proven and the appropriate laws exist. Furthermore, the attitudes towards design-ownership rights vary significantly from country to country, making this type of deterrent not
wholly effective in places where it matters the most, which is in countries where most counterfeit goods are manufactured or in which the ownership rights tend not to be enforced.

The detection approaches like watermarking and fingerprinting provide reactive protection that can detect the theft or fraud and deliver evidence of misconduct. But these techniques cannot prevent IP infringements. Reactive deterrents provide detection or evidence of intrusion and fraud that may help in applying the social tools that are available. Sometimes these approaches may deliver only suspicious facts which however, may be enough to trigger further investigations. Though the detection based approaches do not actively prevent fraud or theft, their presence may deter the would-be attackers and it is sometimes beneficial to advertise them.

The protection approaches are active mechanisms that make use of physical or cryptographic mechanisms to prevent the theft or fraudulent usage of the protected work. Active protection like encryption techniques is highly effective if implemented correctly and has the highest deterrent degree. Further, combined with the social deterrents, active deterrents can help to convince a court that the designer has taken the appropriate measures to protect the design and that the perpetrator had actively circumvented them. But all the encryption techniques for IPP provide only the confidentiality of design; it does not say anything about its authenticity. Authentication is very important in IP industry as it gives an assurance on the identity of source of the bitstream and also ensures the authorised use of the IP cores.

Further, all the existing IPP techniques are capable of only detecting the IP infringements and have very limited capability to prevent their illegal use. As reuse of IP is promoted in SoC environment, access
control becomes essential for the IPs. In order to reuse an IP component, company should purchase the IP from its genuine vendor in legitimate way. Further, its reuse in design house and in SoC application environment should be protected. Unauthorized use of an IP by a company or any other adversary renders loss of revenue to the genuine IP owner thus causing economic damage to the IP vendor. Moreover, while the demand for IP in today’s system designs is increasing, the lack of efficient protection technique for IP in reconfigurable designs still makes it a risky business venture. What is needed for IP providers is the ability to authenticate and securely integrate their logic into third-party systems. Also in the system designer side the ability to authenticate and verify that the third-party IP has not been tampered with is important as well. These requirements are addressed in this research.

As SRAM-based FPGA continues to become more powerful and cheaper, it becomes more attractive to IP designers as well as more susceptible to attackers, who will attempt to exploit any security weakness. Unfortunately, the security of reconfigurable hardware has, until recently, largely been ignored. Further, in all the existing encryption based IPP techniques for SRAM-based FPGA IP cores, the decryption key is part of the IP core that is to be protected. It is possible for an unauthorized person who has access to the hardware IP module to decipher the key. So, there is a need for an effective key storage mechanism that is secure and authentic. This issue is also addressed in this research work. The aim of this research is to develop an efficient access control protection mechanism for the reconfigurable IP cores that would combine the advantages of wireless protection, mutual authentication, remote activation, encryption protection and secure decryption key management mechanism, in a single protection scheme, a desirable safety feature for security sensitive applications.
2.3 SIGNIFICANCE OF THE RESEARCH

- This research proposes a new field of IP protection technique namely the wireless preventive approach (as shown in Figure 2.1), for the active protection of SRAM-based FPGA IP cores.

- This research work explores the possibility of application of the significant short range communication technology in the wireless application areas namely the Radio Frequency Identification (RFID) for the protection of reconfigurable IP cores. The RFID is expected to add intelligence and capabilities to IP providing organizations by its identification, tracking and tracing nature.

- This work incorporates an efficient wireless authentication and activation mechanism to provide access control protection for the IP cores. Moreover, the main advantage of reconfigurable devices which is the flexibility given by the reconfiguration capabilities is also exploited in the protection schemes.

- An important feature in encryption based IPP techniques is that of key management. Unlike other IPP techniques, the decryption key is not part of the reconfigurable IP core that is to be protected. Hence a pirate cannot access the bitstream decryption keys used even if he/she has physical access to the reconfigurable IP cores.

- Finally, unlike conventional RFID security system applications which make use of ready to use RFID tags available in the market, here a customized reconfigurable
RFID tag is realised and the same is embedded in the hardware logic of the FPGA to be protected.

Figure 2.1 Framework of FPGA-Based IP Core Protection Techniques