In this paper we propose a new technique to reduce the program footprint and the instruction fetch latency in high performance architectures adopting long instructions in the memory. Our technique is based on an algorithm that factors long instructions into instruction patterns and encoded instructions, which contains no redundant data and it is stored into an I-cache. The instruction patterns look like a map to the decode logic to prepare the instruction to be executed in the execution stages. These patterns are stored into a new cache (P-cache). We evaluated this technique in a high performance architecture called 2D-VLIW through trace-driven experiments with MediaBench and SPEC programs. We compared the 2D-VLIW execution time performance before and after the encoding, and also with other encoding techniques implemented in computer architectures. Experimental results reveal that our encoding strategy provides a program execution time that is up to 69% better than EPIC.